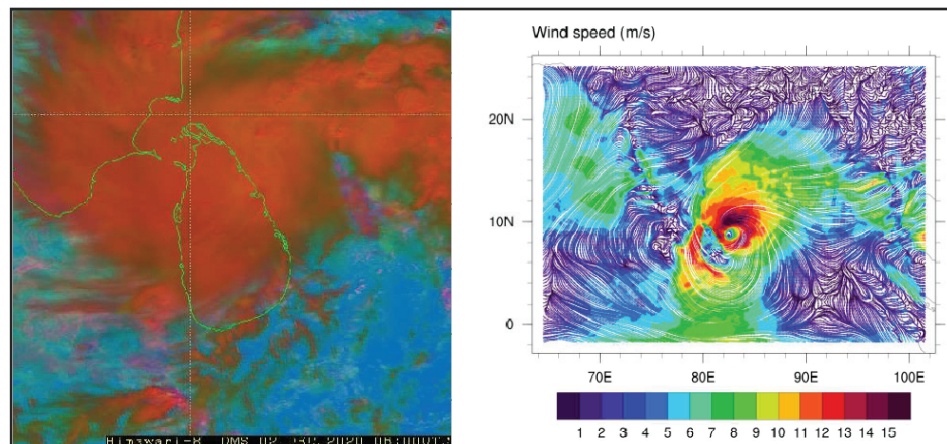


Journal of the National Science Foundation of Sri Lanka





NATIONAL
SCIENCE
FOUNDATION

JOURNAL OF THE NATIONAL SCIENCE FOUNDATION OF SRI LANKA

Editorial Board

Ajit Abeysekera (Editor in Chief)
A. Atputharajah
J.K.D.S. Jayanetti
L.P. Jayatissa
P. Prasad M. Jayaweera
Jagath Manatunge
S.S.N. Perera
Rohini de A. Seneviratne
Saman Seneweera
P. Wijekoon
M.J.S. Wijeyaratne

Language Editor

R.D. Guneratne

Editorial Office

Nadeeja Wickramarachchi (Principal Scientific Officer)
Uthpala T. Karunarathne (Senior Scientific Officer)
Upuli Ratnayake (Scientific Officer)

International Editorial Advisory Board

Chamil Abeykoon, UK
Dilanthi Amaratunga, UK
Dilantha Fernando, Canada
Leslie Gunatilaka, USA
Saman K. Halgamuge, Australia
Kithsiri W. Jayasena, Australia
Vassilios Kapaklis, Sweden
Wah Yun Low, Malaysia
Thomas Mathew, USA
Shanthi Mendis, Switzerland
Javier Francisco Ortega, USA
Malik Peiris, Hong Kong
Kamal Premaratne, USA
Nalin Samarasingha, USA
Ravi Silva, UK
Christopher C. Steel, Australia

Impact Factor : 0.682

Publication : Published quarterly (March, June, September and December) by the National Science Foundation of Sri Lanka.

Subscription : Foreign : US \$ 75 (SAARC countries)
US \$ 150 (other countries) } inclusive
Local : Rs. 2000 } of airmail
postage

Accepted as standing orders or on a calendar year basis.
Single issues - Available on request;
Rs. 600 per copy

Payment must accompany all orders. Remittances in favour of the National Science Foundation of Sri Lanka.

Manuscripts : Research Articles, Research Communications, Reviews and Correspondence in all fields of Science and Technology may be submitted for consideration for publication. A guide to the preparation of manuscripts is provided in each issue. The guidelines may also be obtained by visiting the NSF website.

No responsibility is assumed by the National Science Foundation of Sri Lanka for statements and opinions expressed by contributors to this Journal.

Manuscripts and all correspondence relating to them should be sent to the Editorial Office, National Science Foundation, 47/5, Maitland Place, Colombo 07, Sri Lanka.

Fax: 94-11- 2694754

E-mail: jnsf@nsf.gov.lk

JNSF home page: <http://www.nsf.gov.lk/index.php/nsf-science-magazine>

Publication : A publication fee of US\$ 150 will be levied for each fee manuscript except when the corresponding author is affiliated to a Sri Lankan institution.

Copyright : © National Science Foundation of Sri Lanka

Articles in the Journal of the National Science Foundation of Sri Lanka are Open Access articles published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed anyway.

Indexing : The JNSF is indexed in Science Citation Index Expanded, Journal Citation Reports/Science Edition, BIOSIS Previews, Zoological Record, Biological Abstracts, Chemical Abstracts, Scopus, TEEAL, Ulrich's, AGRICOLA and EBSCOhost

**JOURNAL OF THE
NATIONAL SCIENCE FOUNDATION
OF SRI LANKA**

Volume 50 Number 3

September 2022

C O N T E N T S

EDITORIAL

- 519 **Seeing is believing**
Sumedha Jayanetti
-

REVIEW ARTICLES

- 521 **Rice production in nutrient-limited soils: Strategies for improving crop productivity and land sustainability**
LBD Suriyagoda
- 541 **Heavy metals and food safety in Sri Lanka: A review**
U Samarajeewa
-

RESEARCH ARTICLES

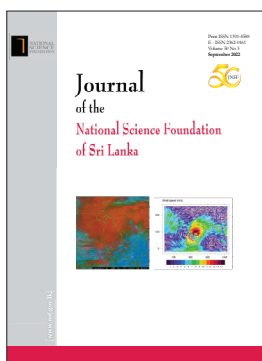
- 563 **Fuzzy based congestion detection and control algorithm for energy efficient wireless sensor network (WSN)**
A Ramasamy-Rajeswari, S Ganapathy, K Kulothungan and A Kannan
- 577 **Evaluating the potential of an open sensor network to support reservoir pre-release decision making**
E Warusavitharana, PKS Mahanama, M Cannata, R Ratnayake, BH Sudantha and D Strigaro
- 589 **High prevalence of glaucoma-associated CYP1B1 mutation (p.G61E) in primary congenital and open angle glaucoma patients in Pakistan**
K Yousaf, R Bashir, K Balqees, S Naz, N Munir and F Aslam
- 595 **Effect of cylindrospermopsin on the hepatotoxicity in wistar rats**
HASN Abeysiri, JKP Wanigasuriya, TS Suresh, DH Beneragama and PM Manage
- 607 **On the construction of unitaries representing minimal inner toral polynomials**
CSB Dissanayake and UD Wijesooriya
- 613 **A coupled system of stochastic differential equations for probabilistic wind speed modelling**
HMDP Bandarathilake, GWRMR Palamakumbura and DHS Maithripala
- 625 **Failure prediction of solid resilient tyres due to kerb impact: A finite element modelling approach**
WAAS Premarathna, JASC Jayasinghe, CD Senanayake, KK Wijesundara and P Gamage

- 639 **Screening of potential aerobic denitrifying bacteria for nitrate removal from water**
A Kirisan, N Gnanavelrajah and RR Ratnayake
- 651 **Physicochemical and sensory attributes of inulin incorporated set yoghurt prepared using cow and buffalo milk**
KADC Gnanarathna, MTLK Jayasumama, KWDP Weragama, DCS Gunasekara, WVVR Weerasingha, A Chandrasekara and PHP Prasanna
- 663 **Fractal characteristics of creeping discharges propagating on nano-epoxy composite insulators immersed in coconut oil**
S Ediriweera, PB Jayarathna, R Samarasinghe, and R Lucas
- 675 **Forecasting the track and the intensity of the cyclone Burevi using WRF**
TD Gamage, U Sonnadara, S Jayasinghe and S Basnayake
- 685 **Anti-inflammatory activity of nonyl 8-acetoxy-6-methyloctanoate, isolated from the cultured marine diatom, *Phaeodactylum tricornutum* : mediated via suppression of inflammatory mediators in LPS-stimulated RAW 264.7 macrophages**
KW Samarakoon, JY Ko, JH Lee and YJ Jeon
- 695 **Implementation of a fuzzy logic approach for a smart production system**
D Dihovicni, M Mišćević, NR Kovačević and D Kreculj
- 705 **Non invasive automated approach for eczema lesions segmentation using colour space normalization**
H Nisar, Y-K Ch'ng and KH Yeap

RESEARCH COMMUNICATION

- 717 **Bioactive cyclo-(S-Pro-R-Leu) from *Aspergillus flavus*, the marine endophytic fungus from brown alga, *Dictyota kunthi***
Haroon M. Haniffa, H. Ranjith W. Dharmaratne and M Yasin Mohammad

Guidelines for Contributors



Cover: Satellite image captured by JMA (left) and forecasted 10-m wind speed (right) of cyclone 'Burevi' at 0600 UTC on 02 December 2020
See *J.Natn.Sci. Foundation Sri Lanka* 2022 50(3): 675 – 684

EDITORIAL

Seeing is believing

The James Webb Space Telescope, the largest and the most powerful space telescope that has ever been constructed by NASA, has begun its science operations orbiting the Sun in line with Earth on a trajectory located at an anti-sunward point approximately 1.5 million kilometers away from Earth. In Webb, revolutionary technologies have been used to study the phases of cosmic history of the universe. The continual expansion of the universe has caused the very first stars and galaxies formed over billions of years ago that emitted ultraviolet and visible light to undergo redshifts to observe that radiation in the infrared range. Unlike the Hubble space telescope which views the universe in visible and ultraviolet light, Webb is equipped with an infrared telescope to focus on infrared radiation capable of peering through gas and dust clouds to see distant objects with an unprecedented resolution and sensitivity. Webb is also capable of studying the nearby universe making it a powerful tool to explore the planets and other bodies in our solar system. Already, to the excitement of scientists and readers, headlines appear in news reports on exoplanets, old star clusters and distant galaxies with their spectacular images taken by Webb with much improved clarity. Scientists, with time will be able to come up with new information about the universe with interpretations, models and even new theories.

The newest technologies NASA has used in the Webb include a foldable sunshield which prevents the heat and light from the Sun, Earth, and the Moon reaching the telescope's heat-sensitive optics, a foldable mirror system in the appropriate dimensions to receive the infrared radiation emitted from celestial objects and maintenance of lasting cryogenic operating temperatures to prevent emission of infrared radiation by this massive observatory itself. Webb is considered an engineering marvel that has become a reality due to meticulous planning over the years and its implementation by an international collaboration between NASA and its

partners in Europe and Canada involving thousands of engineers and hundreds of scientists representing a number of universities, organizations, and companies from U.S. and several other countries.

Thus, Webb serves as a classic example that unifies the scientist, the engineer and experts representing other areas in the quest to finding the origin and evolution of the universe. The team on its way to reach this goal has made groundbreaking inventions, innovations leading to new materials and several new technologies which can be used for the benefit of humanity. Applications are already seen in the fields of medicine and engineering. This project shows some parallels to projects launched several decades ago in exploring the subatomic particles which led to the introduction of global communication pathways like World Wide Web that revolutionized the entire communication industry. Webb can be seen as yet another example for how fundamental science can help shape and impact society.

In the context of Sri Lanka, lessons can be learnt from the approach made by NASA to successfully create this massive, amazingly powerful and precisely controllable space observatory. Webb is an example of teamwork which has been guided by proper leadership knowledgeable in correct scientific approach and methodologies which were put into practice. In Sri Lanka, in a culture that focusses on individual glory at personal and institutional level, at times at any cost, Webb shows that strategies must be in place with proper leadership to encourage the collective efforts of scientists, engineers and relevant experts in the country to address the national scientific and technological challenges it is facing at the moment. When considering the size of Sri Lanka, such an approach seems timely and viable. Webb has provided the impetus to believe in ourselves to use human ingenuity to address issues related to energy, environment, agriculture and earth resources etc. recognizing them as high priorities.

Sumedha Jayanetti

REVIEW

Crop Science

Rice production in nutrient-limited soils: Strategies for improving crop productivity and land sustainability

LBD Suriyagoda

Department of Crop Science, Faculty of Agriculture, University of Peradeniya, Peradeniya.

Submitted: 30 June 2021; Revised: 18 April 2022; Accepted: 22 June 2022

Summary: Productivity improvement of rice cultivation in most parts of Africa and Asia has slowed down, which is largely due to declining soil fertility. Widely reported soil fertility constraints are deficiencies of nitrogen (N), available phosphorus (P), sulphur (S) and zinc (Zn). The amount of N, P and potassium (K) removed with the harvest of 5 t grain ha⁻¹ is estimated as 27–97 kg N ha⁻¹, 4–20 kg P ha⁻¹, and 8–32 kg K ha⁻¹. Straw produced from one ha of land also contains 25–40 kg N, 3.5–10.5 kg P, 70–85 kg K, 2.5–5 kg S, 15–20 kg Ca, 5–15 kg Mg and 200–350 kg Si. Soils do not have an inherent capacity to replenish these amounts of nutrients over the very short period. Most rice farmers also do not have the ability to apply nutrients due to financial limitations and/or lack of awareness. In order to sustain rice cultivation in those fields, soil fertility constraints must be addressed without prioritising unrealistic synthetic fertiliser applications. Improving soil organic carbon and nutrient pools by incorporating organic manure is needed. Crop rotation, including fallow, minimal or zero tillage, site-specific nutrient management, and introduction and promotion of biological nutrient enrichment are other alternative and promising soil-carbon and nutrient-improvement strategies. The combination of these measures is known as an integrated plant nutrient management system (IPNS), and it must be implemented to maintain and/or improve soil fertility. Apart from agronomic interventions, cultivar selection and crop improvement programmes need to particularly focus on low-input soils, as farmers are resource-poor. Due to the complex soil, socio-economic, and institutional structures in these regions, an integrated and participatory strategy incorporating all stakeholders is essential for implementing agronomic interventions and productivity improvement programmes.

Keywords: Organic, productivity, soil fertility, sustainability.

CURRENT CONTEXT

Most rice is grown in Asia (143.4 million ha), Africa (10.5 million ha) and the Americas (7.2 million ha). Moreover, 57% of the global rice cultivation is under irrigated lowland conditions. The remaining rice is grown under rainfed lowland (31%) or upland (9%) conditions, and less than 3% as deep-water or mangrove-associated rice culture (Haefele *et al.*, 2014). The rate of increase in global rice yields has slowed down, limiting the availability of milled rice for consumption (Dobermann *et al.*, 2002; Ladha *et al.*, 2003) (Figure 1). It is also estimated that world rice production must increase by approximately 1% annually to meet the growing demand for food resulting from population growth and economic development (Peng *et al.*, 2004; Mishra *et al.*, 2018; Saito *et al.*, 2018). But increases in rice production have remained below the level required in many countries (Cooper *et al.*, 1999; Ray *et al.*, 2012). Despite this growing demand, the extent used for rice cultivation remains unchanged in the recent past (Figure 1). Therefore, it is essential to revisit the necessary efforts need to be made to ensure sustainability and improve the productivity of rice-cultivating systems.

Input used in rice cultivating systems

Declining soil fertility is one of the main factors for stagnating rice yields in Africa and Asia (Timsina & Connor, 2001; Dobermann *et al.*, 2002; Saito &

* Corresponding author (lalith.suriyagoda@agri.pdn.ac.lk;  <https://orcid.org/0000-0002-5253-2717>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

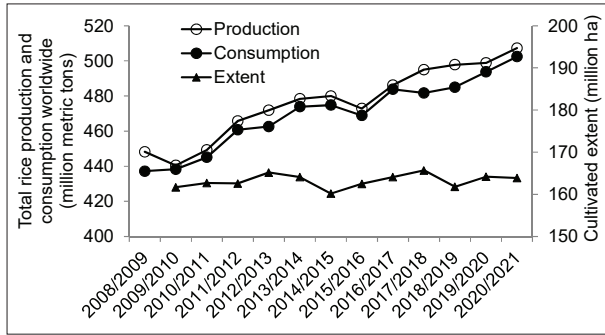


Figure 1: Changes in global rice production, consumption and cultivation extent in the past decade (Note: data were extracted from <https://www.statista.com/statistics/271969/world-rice-acreage-since-2008/> on 10.04.2022)

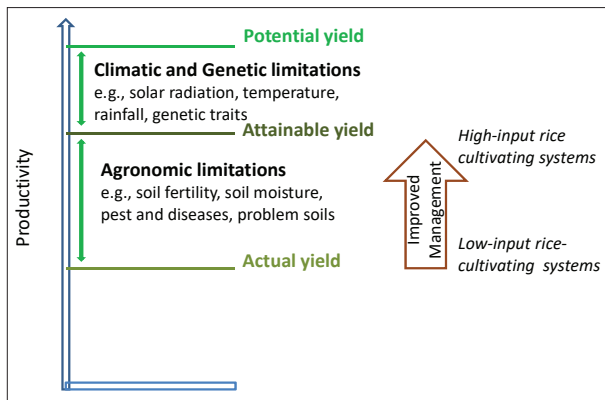


Figure 2: Yield gaps of rice and the factors associated. Note: the gaps between actual and attainable, and attainable and potential are not drawn proportionately.

Futakuchi, 2009; Weerakoon *et al.*, 2011; Rodenburg *et al.*, 2014; Katsura *et al.*, 2016; Mayamulla *et al.*, 2017; Stewart *et al.*, 2020). This has been caused to a considerable extent by the continuous crop and residue removal, negative nutrient balances and soil degradation (Cassman *et al.*, 1998; Dobermann *et al.*, 2002; Nhamo *et al.*, 2014; Kekulandara *et al.*, 2019). Therefore, many research studies have been conducted around the globe to improve the productivity of rice-based cropping systems. One of the critical attempts in most of this research is to enhance soil fertility along with other inputs and

management, such as irrigation water supply, pest and disease control, and mechanisation. Such systems can be classified as intensive, high-input rice-cultivating systems (Figure 2). Due to the application of all required inputs and intensive crop management, farmers in high-input rice-cultivating systems could cover a higher percentage of the yield gap between attainable and actual yields by increasing productivity. Even though improving soil fertility to achieve high productivity is desirable, one of the main problems faced by many rice farmers in Asia and Africa is the difficulty in applying recommended rates of synthetic fertilisers. This is due to their poor economic status, high fertiliser costs, unavailability of inputs, and/or production risks (Weerakoon *et al.*, 2011; Haefele *et al.*, 2003, 2014; Nhamo *et al.*, 2014; Mayamulla *et al.*, 2017; Stewart *et al.*, 2020).

The low-input rice-cultivating systems defined in this study have the nutrient-supplying capacity (either by inherent soil nutrient pools or through the application of fertilisers) less than the amount and rate required by the rice plant to reach attainable grain yields ($t\ ha^{-1}$) in a particular environment, when other inputs or characteristics (*e.g.*, drought, pest and diseases, Fe-toxicity, low-CEC, low-pH) are not yield limiting (Figure 2). One or more nutrients may act as rate-limiting variables in this system. In low-input rice-cultivating systems, farmers have a variety of options for enhancing soil fertility, despite recommendations from extension services to rely primarily on synthetic fertilisers, which are typically supplied at a subsidised rate (Weerahewa, 2004; Asai *et al.*, 2009; Druilhe & Barreiro-Hurlé, 2012; Nhamo *et al.*, 2014; Stewart *et al.*, 2020). Due to the continuous increase in fertiliser prices, more and more rice fields are likely to become nutrient deficient in the future (Shepherd *et al.*, 2016). Despite this, the importance of rice production in Africa and Asia has significantly increased over the last decades (Nhamo *et al.*, 2014), as rice plays a pivotal role in improving household food and nutritional security as well as national economies in these regions (Diallo *et al.*, 2016; Katsura *et al.*, 2016; Stewart *et al.*, 2020). Therefore, to sustain rice cultivation in low-input agricultural systems, soil fertility constraints must be addressed before other technologies and policies can become effective in improving productivity. When doing so, soil management options besides applying synthetic fertilisers, such as the addition of organic matter, crop rotation, and tillage practices, should receive high priority as those are vital to overcoming nutrient limitations.

Existing nutrient limitations

Soil fertility differences exist from continental scale to paddock and between and within farms (Tsujiyamoto *et al.*, 2013; Nhamo *et al.*, 2014; Katsura *et al.*, 2016). Among the three major nutrients, nitrogen (N) is the most widely limiting, while phosphorus (P) ranks second, and potassium (K) deficiency is rarely reported (Dobermann *et al.*, 2003; Ladha *et al.*, 2003; Shen *et al.*, 2004; Rodenburg *et al.*, 2014; Diallo *et al.*, 2016; Somaweera *et al.*, 2016, 2017; Stewart *et al.*, 2020). Moreover, deficiencies in sulphur (S) and zinc (Zn) are common in many regions to variable degrees (Buri *et al.*, 2000; Duxbury *et al.*, 2000; Ladha *et al.*, 2003; Haefele *et al.*, 2014; Rodenburg *et al.*, 2014). Apart from nutrient limitations, constraints related to pH and cation exchange capacity (CEC) also occur in isolated areas or problem soils (Ladha *et al.*, 2003; Haefele *et al.*, 2004, 2014; Shen *et al.*, 2004; Gami *et al.*, 2009; Saito & Futakuchi, 2009; Rodenburg *et al.*, 2014; Diallo *et al.*, 2016; Dossou-Yovo *et al.*, 2016). In general, the occurrence of nutrient deficiency decreases in the sequence of uplands > rainfed lowlands > irrigated lowlands > deepwater/mangrove areas (Haefele & Konboon, 2009; Haefele *et al.*, 2014; Nhamo *et al.*, 2014). This is caused by the transport of nutrients and soil particles from higher to lower parts of the landscape and due to the differences in soil moisture availability. In contrast, soils with problems such as Fe-toxicity are more frequent in the lower parts of the landscape and in locations where seepage water accumulates (Becker & Asch, 2005; Haefele *et al.*, 2014; Rodenburg *et al.*, 2014; Suriyagoda *et al.*, 2017; van Oort, 2018). Based on the type of nutrient limitation, rice-growing soils can broadly be categorised as;

- (i) problem/very poor soils (show multiple constraints such as iron and aluminium toxicity, low CEC and pH, high P-fixation, shallow soils, limited water-holding capacity, and deficiency in macro- and micro-nutrients),
- (ii) poor soils (no constraints as above, but low in many macro and micro nutrients including N and P), or
- (iii) good soils (no constraints, mostly limited by N and/or P only) (Haefele *et al.*, 2014).

All these soils are found both in high and low-input rice cultivating systems (Figure 2). However, in this study, we mainly focus on the second and third categories of rice-growing soils, which are limited by macro- and micronutrients.

Average recovery efficiency (*i.e.* percentage taken up of the amount applied) of fertiliser N by rice plants is

only about 30 to 47% (Dobermann *et al.*, 2002; Ladha *et al.*, 2005). The low recovery efficiency of N is mainly due to high losses, making it unavailable for current or future crops. The recovery efficiency of P is less than 10% in most instances (Sanchez & Salinas, 1981; Dobermann *et al.*, 1998), but values up to 20% have also been reported (Holford, 1997). Unlike N, most excess P is retained in the soil after forming complexes, making it available for future crops after solubilisation at the optimal range of pH (Lambers *et al.*, 2008; Sirisena & Suriyagoda, 2018). In most rice-growing systems in Asia and Africa, K balances are negative (Dobermann *et al.*, 1998; Ladha *et al.*, 2003; Haefele *et al.*, 2004) unless crop residues are incorporated (Wihardjaka *et al.*, 1999; Sahrawat, 2000). However, mass balances of K inputs, K uptake, and exchangeable K in rice soil show that a large part of K uptake is from non-exchangeable pools, and the mobilisation of non-exchangeable K is apparently plant-induced (Wihardjaka *et al.*, 1999). Moreover, high soil K reserves can buffer negative K balances for decades (Haefele *et al.*, 2004; Shen *et al.*, 2004; Diallo *et al.*, 2016; Somaweera *et al.*, 2017). Irrigation water also contains more K than N and P, allowing K to move across rice fields in cascade irrigation networks (Suriyagoda *et al.*, 2017). Therefore, K deficiency in rice is less prevalent than N and P and most common in high-yielding systems.

Rice grain N, P, and K content (expressed on mg g⁻¹ basis) reported from different countries are mainly in the range of 9-15 mg g⁻¹ for N, 1.0-4.3 mg g⁻¹ for P and 0.6-4 mg g⁻¹ for K (Dobermann *et al.*, 1996, 2002; Timsina & Connor, 2001; Xu *et al.*, 2015; Somaweera *et al.*, 2016, 2017). Moreover, internal N-use efficiency (*i.e.* kg grain DW kg⁻¹ N in aboveground biomass) ranges between 31-112 kg grain DW kg⁻¹ N uptake, and for P and K, the values are 192-900 kg grain DW kg⁻¹ P uptake and 32-115 kg grain DW kg⁻¹ K uptake, respectively (Witt *et al.*, 1999; Dobermann *et al.*, 1996, 2002; Dwivedi *et al.*, 2004; Buresh *et al.*, 2010; Das *et al.*, 2009, 2014; Haefele *et al.*, 2003; Singh *et al.*, 2005; Xu *et al.*, 2015). Therefore, the amount of N, P and K stored in aboveground biomass when producing 5 t grain ha⁻¹ are 45-161 kg N ha⁻¹, 5-26 kg P ha⁻¹, and 43-156 kg K ha⁻¹. It is also reported that about 55-65% of N, 65-70% of P, and 15-20% of K absorbed by rice plants are stored in the grains (Dobermann & Fairhurst, 2000; Sahrawat, 2000). Therefore, the amount of N, P and K removed with the harvest of 5 t grain ha⁻¹ can be estimated as 27-97 kg N ha⁻¹, 4-20 kg P ha⁻¹, and 8-32 kg K ha⁻¹. The extensive range of N, P, and K removal is due to the differences in grain N, P, and K contents, mainly affected by the variation in soil fertility.

Most farmers cannot apply sufficiently high rates of N, P, and K to support the uptake of elements to reach expected yield targets. Consequently, as soil nutrient limitations develop, an increase in nutrient-use efficiency (or a decrease in tissue nutrient content) can be anticipated. As a benchmark, straw N, P, and K content less than 10 mg N g⁻¹, 0.6 mg P g⁻¹ and 1.8 mg K g⁻¹ at maturity are generally considered deficiency thresholds for N, P, and K when identifying soils, which have a low capacity to supply N, P and K to rice plants (Dobermann *et al.*, 1998; Dobermann & Fairhurst, 2000; Vandamme *et al.*, 2016). Tillering is the most sensitive stage of rice crop development when P deficiency symptoms arise, although N deficiency symptoms appear throughout the crop life cycle (Fageria, 2003; Somaweera *et al.*, 2017). Therefore, the reported N-, P- and K-use efficiency values, straw N, P, and K contents, and sensitive crop developmental stages can be used when identifying the existence and severity of N, P, and K deficiencies in rice culture. Moreover, due to the discrepancy between the growth stage at which critical P level is identified (*i.e.*, at maturity) and the most sensitive growth stage of rice to P deficiency (*i.e.*, at tillering), the relationship between those two needs to be explored.

Zinc deficiency is one of the most widespread micronutrient deficiencies in rice production in Africa and Asia (Buri *et al.*, 2000; Dhanapala, 2000; Duxbury *et al.*, 2000; Timsina & Connor, 2001; Quijano-Guerta *et al.*, 2002; Ladha *et al.*, 2003; Wissuwa *et al.*, 2006). Availability of Zn is generally decreased after flooding rice soils and is associated with an increase in soil pH after flooding (Fageria *et al.*, 2011). Significant positive correlations between total soil carbon and available P and Zn were also found in West Africa (Buri *et al.*, 2000). Thus, knowledge of rice tolerance to many stresses is useful when looking for Zn-deficiency resistant genotypes because much of the information on other nutrient deficiencies, such as P, is available in most rice germplasms.

Apart from Zn, S deficiency occurs in certain parts of Asia, Africa and America (Blair *et al.*, 1979; Dobermann *et al.*, 1998; Yamaguchi, 1999; Buri *et al.*, 2000; Dhanapala, 2000; Duxbury *et al.*, 2000; Tsujimoto *et al.*, 2013). As a benchmark, straw S content less than 0.5 mg S g⁻¹ at maturity is generally considered the deficiency threshold for S (Dobermann *et al.*, 1998). When N, P, and K were applied without S in Ghana, biomass production was not increased while the tissue N content and N:S ratios increased, indicating S deficiency (Tsujimoto *et al.*, 2013). In another study, Yamaguchi (1999) found that the maximum quantity of S needed to alleviate the deficiency

in rice was about 30% of the N requirement. Moreover, the topsoil sulphate-S content was positively correlated with total soil carbon, available P, and CEC (Buri *et al.*, 2000), indicating the occurrence of multiple nutrient deficiencies and the importance of soil properties in determining the availability of mineral elements. Consequently, S deficiency diagnosis in rice plants with a single variable often leads to an error of judgement. Therefore, the evaluation is suggested to take several variables into account, *e.g.*, total-S, total-N/total-S ratio, sulphate-S, and sulphate-S/total-S ratio (Yamaguchi, 1999), requiring detailed laboratory procedures. This has also become a limitation when detecting S deficiency in most of rice fields at its initial stages.

Silicon (Si) is increasingly considered as a beneficial element for plant growth. Rice is efficient in accumulating Si from the soil, and shoot Si content is about 10% of shoot dry weight (Yamamoto *et al.*, 2012). Silicon in soil may be lost if the Si-cycle is interrupted (Meharg & Meharg, 2015). Most of the Si taken up can be returned to the soil with the incorporation of straw and paddy husk. However, rice cultivation over centuries with considerable straw removal can deplete the soil Si pool, risking the sustainability of rice cultivation in certain regions (Savant *et al.*, 1997). Apart from the positive effects of Si on plant growth, there is evidence that Si fertilisation ameliorates Fe, manganese (Mn), and arsenic (As) toxicities (Fu *et al.*, 2012; Meharg & Meharg, 2015; Suriyagoda *et al.*, 2018), and improves resistance to pest and disease incidences (He *et al.*, 2015). Moreover, Si fertilisation decreases Ca, Fe, Mg, N, and P in rice straw (Ma & Takahashi, 2002), demonstrating a negative relationship between Si and other essential element levels in rice. However, if rice plants can produce the same or higher grain yields with reduced nutrient content (*i.e.*, enhanced nutrient-use efficiency) due to improved Si nutrition, applying Si to Si-deficient fields could be advantageous.

Unlike the occurrence of deficiencies of most of the nutrients in rice fields as discussed above, the presence of a high content of iron (Fe²⁺) in soil has caused toxicity to rice plants, mostly in lowlands in Africa and Asia, causing even complete crop failure (Becker & Ash, 2005; Fairhurst *et al.*, 2007). It is reported that up to 12% of the rice lands in Africa are potentially affected by iron toxicity (van Oort, 2018). The Fe²⁺ concentration (expressed as mg L⁻¹) in soil solution that affects rice yield can vary in the range of 10 - 2000 mg L⁻¹ (Becker & Ash, 2005). Toxicity caused by Fe²⁺ is a complex disorder, as it is associated with soil acidity, low-CEC, and nutrient deficiency (Haefele *et al.*, 2014). Moreover,

due to the presence of a high concentration of Fe^{2+} in lowland soils, availability of Mn, Zn, P, and K in the soil can be low, causing deficiencies of those nutrients to rice plants (*i.e.*, multiple nutrient disorder) (Fageria & Zimmermann, 1988; Suriyagoda *et al.*, 2017).

AGRONOMIC INTERVENTIONS TO IMPROVE FERTILITY IN RICE FIELDS

Improving soil organic carbon (SOC) and nutrient pools

Addition of organic matter

Continuous cultivation on the same land, removal of harvested grains, lack of incorporation of organic matter, including straw produced in preceding seasons, constant and fast decomposition of retained organic matter, and losses due to runoff have led to a decline in organic carbon content in some of the rice-growing soils (Singh *et al.*, 2005; Das *et al.*, 2014). Most of the rice-growing soils in the globe have SOC values in the range of 1-3%, and values less than 0.5% have also been reported (Mukhopadhyay *et al.*, 2013; Xu *et al.*, 2015; Diallo *et al.*, 2016). It is also known that SOC content is lesser in aerobic soils than in anaerobic soil, and in tropical soils than in temperate region. (Sahrawat, 2004). Therefore, tropical soils experiencing aerobic or frequent wetting and drying conditions decompose SOC rapidly and release N (Sahrawat, 2006). Even though the application of external sources of N such as urea is preferred to organic manure low in N and soils deficient in N to enhance the decomposition, such applications are not required in most tropical soils due to the inherently fast rate of organic matter decomposition.

Soil organic carbon can improve nutrient-retention capacity, water-stable aggregates, and porosity, and reduce bulk density, thereby facilitating crop establishment, root growth and nutrient uptake (Lal *et al.*, 1989; Timsina & Connor, 2001; Ladha *et al.*, 2003; Dwivedi *et al.*, 2004; Bi *et al.*, 2009; Suriyagoda *et al.*, 2014) (Table 1). As the C/N ratio of soil organic matter is relatively constant, incorporation of organic matter also improves soil N pools (Cassman *et al.*, 2002). However, the application of synthetic N, P, and K fertilisers does not improve long-term SOC or nutrient pools (Gami *et al.*, 2009). Management practices can significantly affect soil carbon storage through carbon inputs and losses. Still, detection of the changes in SOC is often difficult due to the small magnitude of changes relative to the total stock, except in long-term studies (Conen *et al.*, 2003). Despite

the importance of maintaining adequate SOC pools to ensure a favourable soil environment for rice plants and increasing awareness among farmers, there has been insufficient attention to this aspect (Dossou-Yovo *et al.*, 2016; Stewart *et al.*, 2020). Therefore, an understanding of the benefits of increasing SOC should be raised among rice farmers. Moreover, long-term investigations of the changes in SOC and nutrient balances, particularly in nutrient-limited soils, are required.

Improving SOC pools increases grain yield of rice in China (Pan *et al.*, 2009; Huang *et al.*, 2013; Hu *et al.*, 2016; Zhou *et al.*, 2020), Laos (Roder *et al.*, 1995; Asai *et al.*, 2009), Senegal (Diallo *et al.*, 2016), East and South Africa (Nhamo *et al.*, 2014 and references therein), Nepal (Gami *et al.*, 2009), Bangladesh (Hossain *et al.*, 2016), Thailand (Wonprasaid *et al.*, 1996; Homma *et al.*, 2003), Cambodia (Ly *et al.*, 2016) and Sri Lanka (Sirisena *et al.*, 2016). In China, grain productivity increased by 21% and 24%, and C accumulation by 72% and 103%, respectively, when pig manure (8400 kg ha⁻¹) or straw (4500 kg ha⁻¹) was applied (Pan *et al.*, 2009). Hu *et al.* (2016) reported that the application of a lower amount of organic manure, *i.e.*, 2100 kg ha⁻¹, can be recommended for the implementation of organic rice farming, and higher rates of organic manure application may increase the risks of pests and diseases. Huang *et al.* (2013) also observed a yield increase by 5.2% when crop residues were retained. In Thailand, grain yield increased, on average, by 49% with the application of 6.25 t farmyard manure ha⁻¹, and by 50% with the application of 60 kg N, 13 kg P and 25 kg K ha⁻¹ in comparison to the fields which did not receive synthetic fertilisers (Khunthasuvon *et al.*, 1998). Similar results are reported by Wihardjaka *et al.* (1999) in Indonesia. Becker & Johnson (2001) observed an increase in grain yield of upland rice by 26% when SOC content increased by 19% across several sites in West Africa. Therefore, sole application of organic matter or combined application of organic matter with synthetic fertiliser have increased the grain yield of rice and improved soils. Thus the application of organic materials with or without synthetic fertilizer should be promoted in fields low in SOC with site-specific approaches.

Widely available materials to improve the soil organic carbon pool are rice straw, animal manure, farmyard manure, and other crop residues. Returning one t ha⁻¹ of straw (rice, wheat or maize) to the soil each year can sequester about 130 kg C ha⁻¹ yr⁻¹ (Lu *et al.*, 2009). Moreover, about 40% of N, 30–35% of P, and 80–85% of K absorbed by rice remains in the straw at maturity (Dobermann & Fairhurst, 2000;

Das *et al.*, 2014). Incorporation of one t of straw per hectare accounts for the addition of 5–8 kg N, 0.7–2.1 kg P, 12–17 kg K, 0.5–1 kg S, 3–4 kg Ca, 1–3 kg Mg and 40–70 kg Si on a dry weight basis (Dobermann & Fairhurst, 2000, 2002; Dobermann & Witt, 2000; Das *et al.*, 2014). When rice straw is not added to fields, P needs to be re-applied at about 17 kg ha⁻¹ every two crops on acid sandy lowland soils to maintain the grain yield at about 2.5–3.0 t ha⁻¹ in Cambodia (Pheav *et al.*, 2003). Similar results were obtained for N in China (Pan *et al.*, 2009). Therefore, incorporation of rice straw improves soil quality including nutrient content and retention (Kumar & Goh, 1999; Wihardjaka *et al.*, 1999; Sahrawat, 2000; Suriyagoda *et al.*, 2014), and the effects are more pronounced in clayey than in sandy soils (*i.e.*,

due to greater retention), and in flooded than in aerobic soil (*i.e.*, due to higher solubility) conditions (Suriyagoda *et al.*, 2014; Haefele *et al.*, 2016). Rice straw retention also influences availability of micronutrients such as Zn and Fe (Dobermann & Fairhurst, 2000, 2002). Straw incorporation facilitates greater root system growth plasticity with the production of deeper and finer roots allowing the rice plants to explore a large volume of soil in search of nutrients (Suriyagoda *et al.*, 2014). However, straw produced from nutrient-limited soils contains low contents of nutrients, and thus the addition of straw will at best reduce the rate of decline in soil fertility (Table 1). Therefore, nutrient-dense alternative sources should be incorporated into the soil to improve its nutritional condition.

Table 1: Methods widely used to improve soil organic carbon (SOC) and nutrient pools, and the advantages and disadvantages of those methods

Method	Advantages	Disadvantages
Addition of organic matter	<ul style="list-style-type: none"> • Improve nutrient-retention capacity, water-stable aggregates and porosity • Reduce bulk density • Facilitate crop establishment, root growth and nutrient uptake 	<ul style="list-style-type: none"> • Variable in types and quality • Require large amounts (content of nutrients is low) • Labour cost for collection and application is high • May fix plant available N early in the season
Fallow or crop rotation	<ul style="list-style-type: none"> • Improve productivity, soil fertility, SOC pool • Reduce bulk density • Avoid risk in cultivation under environmental stresses 	<ul style="list-style-type: none"> • Improvement of soil quality depends on the types of crops used during fallow periods or in rotation • Depends on area specific socio-economic, climatological and ecological conditions
Minimum tillage	<ul style="list-style-type: none"> • Minimize the loss of SOC, soil structure and nutrients • Increase bulk density • Saves labour, energy and water needed for puddling 	<ul style="list-style-type: none"> • Effectiveness is variable and depends on the region, climatic condition, cropping system, rate and time of fertilizer application, rice crop establishment method and duration of minimum till adoption • May increase percolation losses during the season
Application of biochar	<ul style="list-style-type: none"> • Improve SOC and fertility • Improves soil physical characteristics 	<ul style="list-style-type: none"> • Large quantities are recommended • Largely variable in quality • Bind mobile nutrients

Soil analyses after six years of experimentation in Nepal showed that treatments receiving organic sources of nutrients (4–12 t ha⁻¹ yr⁻¹) increased total soil C and N by 18% to 62% and 15% to 48%, respectively, compared with the NPK treatment (Gami *et al.*, 2009). The authors also observed a build-up of total and plant-available P in plots receiving farmyard manure. However, total and

available soil K were similar in all treatments (Gami *et al.*, 2001). Farmyard manure also contained a bacterial population of 5.2×10^6 cfu g⁻¹ (Mukhopadhyay *et al.*, 2013). After 11 years of experimentation in Sri Lanka the lowest yield was observed in the plots which did not receive any form of nutrients (2.85 t ha⁻¹). However, the yield in the plots receiving organic manure only (straw

from the previous season, 1 t green manure and 4 t cattle manure ha^{-1}) was 3.75 t ha^{-1} and that in plots receiving both organic manure and synthetic fertilisers was 5.5 t ha^{-1} (Sirisena *et al.*, 2016). In China, there has been a prominent topsoil C accumulation of 0.1–0.4 t $\text{ha}^{-1} \text{yr}^{-1}$ when pig manure was applied compared with adding synthetic fertiliser only (Pan *et al.*, 2009). Therefore, incorporation of farmyard manure, in the absence of straw or together with straw, can be promoted to replenish soil C and nutrient pools in nutrient-limited soils. However, the general trend in rice cultivation is a decreasing use of local nutrient sources, mainly because of increasing opportunity costs of the labour needed for collection and application, and an increasing use of synthetic fertilisers when available, incurring high cost (Table 1). Between rice environments, the relative contribution from local nutrient sources increases from irrigated lowlands to rainfed lowlands to uplands, whereas synthetic fertiliser use has an opposite trend (Haefele *et al.*, 2014, 2016). This is, of course, due to the comparatively higher risk and poorer farmers in most rainfed rice environments. Another reason might be the often poor soil fertility in rainfed environments, requiring more organic fertiliser to improve/maintain soil fertility.

Although the beneficial effects of organic manure incorporation, to improve the physical, chemical, and biological properties of soil and the grain yield of rice, have been shown numerous times, they depend on the types and quality of the organic manure available, and its use. Therefore, the amount of organic residues requires to reach productivity targets in different rice-based cropping systems may vary widely and is largely unknown. The amount of organic manure needed, its form, method of application and management, and nutrient releasing patterns are material, site, or region specific. Therefore, localised quantifications are required to maximise the efficiency of an integrated nutrient management system. Based on such calculations, large amounts of organic manure were sometimes recommended, *e.g.*, 13–35 t ha^{-1} of compost in Africa (Nhamo *et al.*, 2014 and references therein). Such application rates are often far greater than what an average livestock farmer produces annually and what a rice farmer can afford and apply. A potential approach in such situations is to spread out the application of organic manure across several seasons for gradual improvement of soil fertility, as a component in an integrated plant nutrient management system (IPNS). These practices will ensure long-term development of soil nutrient pools while minimising their losses.

Uptake (*i.e.*, recovery) efficiency of N from a single fertiliser application typically decreases in proportion

to the amount of N fertiliser applied (Reddy & Reddy, 1993). The same principle applies to available N derived from organic N sources such as legume green manures, cover crops, and animal manures. Potential nitrate leaching from manures can be equal or even greater than the losses from synthetic N fertiliser when the available N supply from either source exceeds crop demand by similar amounts for comparable times (Cassman *et al.*, 2002). Similar results were reported for other nutrients across different regions (Zhang & He, 2004; Pratiwi & Shinogi, 2016; Sirisena & Suriyagoda, 2018). Moreover, substantial incorporation of organic fertiliser at once accelerates sudden release of greenhouse gases such as CH_4 and N_2O , and may contribute greatly to global warming (Yuan *et al.*, 2017). In China, the amount of CH_4 emission is greater with organic manure (77.3 kg $\text{CH}_4 \text{ha}^{-1}$) than with urea (30.2 kg $\text{CH}_4 \text{ha}^{-1}$) and the blank control treatment (17.4 kg $\text{CH}_4 \text{ha}^{-1}$) (Yuan *et al.*, 2017). Similar results were observed by Zhao *et al.* (2015). Therefore, it is important to gradually enhance the soil's nutrient pool in order to minimise nutrient losses.

Incorporation of fallow or crop rotation

One of the most prominent approaches to improving soil fertility is the incorporation of fallow periods in the cropping calendar and improvement of the quality of the fallow vegetation, *e.g.*, introduction of legume cover crops (Balasubramanian & Sekayange, 1992). Improvements in productivity, soil fertility, SOC pool, and reduction in bulk density with the incorporation of legumes in rotation, such as mungbean, pigeon pea, and forage cowpea, were observed in rice-based cropping systems in Bangladesh, India, and Sri Lanka (Timsina & Connor, 2001; Dwivedi *et al.*, 2004; Singh *et al.*, 2005; Malaviarachchi *et al.*, 2016; Hossain *et al.*, 2016) (Table 1). Das *et al.* (2014) observed that 70–106 kg N ha^{-1} , 10–14 kg P ha^{-1} and 63–94 kg K ha^{-1} could be incorporated to the soil with the recycling of 4.8–7.2 t ha^{-1} of weed biomass under subtropical conditions in Meghalaya, India. In African uplands, where rice is grown with the incorporation of *Mucuna* spp., *Canavalia* spp., and *Stylosanthes guianensis* during the dry season, consistently high N accumulation was found. Grain yields of rice, which had been preceded by a legume fallow, were on average 0.2 t ha^{-1} or about 30% greater than that preceded by a natural weedy fallow control treatment (Becker & Johnson, 1998). However, some contradictory results have also been reported about the use of legume crops in rotations. For example, although including grain legumes in rotation with cereals can reduce N-fertiliser requirements compared with continuous cereal cropping, the degree of soil-N stock replenishment depends on the

amount of N removed with the harvested seeds (Cassman *et al.*, 2002). Overall, the incorporation of legume-based fallow to rice rotation has merits, particularly when non-crop legume species are used, and the net benefit of such systems need to be assessed.

Other studies showed that the SOC pool is affected by the types of crops incorporated into the rice cropping systems; for example, rice-rice (1470 kg ha⁻¹) and rice-soybean (677 kg ha⁻¹) rotations added more dry matter to soil than rice-tobacco (29 kg ha⁻¹) or rice-onion (not detectable) as most of the biomass of onion and tobacco is removed at harvest (Ratnayake *et al.*, 2017). Moreover, intensive land preparation used in onion and tobacco crops resulted in faster decomposition of SOC, depleting the soil carbon reserves. Therefore, the incorporation of high-value cash crops into the rotation, in which either relatively little biomass is produced or most of the biomass is removed from the field, do not fit well in rice rotations in terms of the enrichment of soil quality.

Besides putting good crops in rotation, letting rice fields go fallow is a better way to improve soil fertility than growing high-risk rice crops in dry seasons or areas, even though the soil nutrition benefits are less than when legume species are added (Becker & Johnson, 1998). This can effectively lower production costs and significantly improve soil fertility and biological properties. However, many studies tended to focus on the benefits of crop rotation or fallow in terms of soil fertility, ignoring the diversity of systems, socio-economic characteristics, weeds, and farmers' perceptions and production objectives (Becker & Johnson, 1998). For these reasons, the adoption of fallow as a viable technique by rice farmers in nutrient-limited soils is less successful unless drought becomes a barrier for crop cultivation (Table 1). Thus, future endeavours should consider socio-economic, climatological and ecological concerns when introducing different crops or fallow to rice rotations.

Tillage practices

Minimum tillage can improve SOC and nutritional status compared to puddled rice culture (Table 1). The primary benefits of puddling are the creation of a soft seedbed, reduction of water and nutrient losses and weed control (So *et al.*, 2001). Moreover, deep ploughing could be advantageous in certain systems as a short-term strategy to bring the leached nutrients to top soil layers while enhancing the benefits mentioned above. However, leaching would be negligible when the soil is inherently low in nutrients. Moreover, excessive tillage in continuous rice culture destroys soil structure,

disrupts the continuity of soil pores, reduces the number of residues on the soil surface and degrades soil quality (Osunbitan *et al.*, 2005; Jiang & Xie, 2009). Therefore, puddling disrupts most of these soil qualities. Though the short-term agronomic response to tillage may be either negative (Agboola, 1981; Lal *et al.*, 1989) or positive (Awadhwai & Smith, 1989; Kawakye & Bobo, 1995), most of the long-term crop responses are either neutral or negative (Kirchhof *et al.*, 2000; So *et al.*, 2001). Long-term field experiments conducted in different countries indicate improvements in soil aggregation, higher microporosity and SOC content, and lower bulk density under no-till, compared with conventional tillage systems (Andrade *et al.*, 2010; Das *et al.*, 2014). Moreover, combining minimum tillage and organic residues with reduced synthetic fertiliser application has increased soil nutritional status (Singh & Singh, 1995). For example, instead of conventional land preparation and application of high doses of N fertiliser (120 kg ha⁻¹), a reduced-till system with the incorporation of straw from the previous season and reduced N-application (60 kg ha⁻¹) was shown to improve the SOC pool in Benin (Dossou-Yovo *et al.*, 2016). Therefore, practising reduced-tillage or no-tillage systems has been suggested as a potential approach to reducing the rates of decline of SOC and nutrient stocks (Timsina & Connor, 2001; Gami *et al.*, 2009; Jiang & Xie, 2009; Das *et al.*, 2014; Du *et al.*, 2015). However, the effectiveness of no-till rice culture in improving SOC and nutritional status is site-specific. It depends on region, climatic condition (season), cropping system, rate and time of fertiliser application, rice-crop establishment method (*i.e.*, transplanting, seedling throwing or direct seeding), and duration of no-till adoption (Gupta & Seth, 2007; Ghimire *et al.*, 2012; Suriyagoda *et al.*, 2014; Huang *et al.*, 2015) (Table 1). Despite the influence of many factors, minimum tillage shows promise in improving long-term SOC in nutrient-limited soils. Thus site-specific packages, including minimum tillage, should be developed considering the major factors mentioned above.

Use of biochar

Biochar is the product of heating biomass in the absence of or in limited air to above 250 °C, a process called charring or pyrolysis (Lehmann and Joseph 2015). In some instances, the material properties of biochar may overlap with those of charcoal as an energy carrier. Still, many types of biochar do not easily burn, and are typically made to address soil issues. A certain level of organic carbon forms, called fused aromatic ring structures, similar to charcoal, is an important defining feature of biochar. These structures are formed during

pyrolysis and are crucial to biochar properties with respect to mineralization or adsorption. Therefore, biochar is typically enriched in carbon and other mineral elements. The chemical properties of the organic carbon structure of biochar are fundamentally different from those of the material that the biochar was produced from, and it is depleted in oxygen and hydrogen. In contrast, the macro-morphological characteristics of biochar typically resemble those of the starting material, which means that it normally looks the same, apart from its black color (Lehmann and Joseph 2015).

Applying biochar to soil can improve soil carbon and fertility (Table 1). Biochar improved the saturated hydraulic conductivity of soil and the xylem sap flow of upland rice in Laos (Asai *et al.*, 2009). Biochar application results in higher grain yield at sites with low P availability and improves the response to N and P chemical fertilisers (Munda *et al.*, 2018; Bi *et al.*, 2019). Rice straw or rice husk biochar application results in increasing grain yield by 8.5–11% or 1–24%, respectively, due to increased nutrient availability and SOC content (Liu *et al.*, 2016; Munda *et al.*, 2018). Similar results were observed by Bi *et al.* (2019). However, biochar application reduces leaf SPAD values, possibly due to adsorption of $\text{NO}_3\text{-N}$ on to biochar and reduced N availability, indicating that biochar application without additional N fertiliser application could reduce grain yields in soils with a low N supply (Lehmann *et al.*, 2003; Asai *et al.*, 2009). Recently Wu *et al.* (2019) revealed that the combined application of biochar and vermicompost increased rice yield by 26.5–35.3% compared with biochar amendment alone. These limited results suggest that biochar application alone or in combination can improve productivity, but the effect of biochar application highly depends on soil fertility and fertiliser management (Haeefe *et al.*, 2011; Liu *et al.*, 2019). Moreover, the effects are more evident in highly weathered and infertile tropical soils under upland conditions (Asai *et al.*, 2009; Peng *et al.*, 2011; Liu *et al.*, 2016; Bi *et al.*, 2019).

Biochar application has its own disadvantages. When considering the amount of biochar required, incorporation of large amounts and ranges are often recommended, *e.g.*, 2 - 50 t ha⁻¹ is recommended in Madagascar (Asai *et al.*, 2009; Raboin *et al.*, 2016; Wang *et al.*, 2018; Bi *et al.*, 2019) and farmers cannot afford to apply this amount. Moreover, biochar may bind certain mobile nutrients, making them less available for rice plants, thereby aggravating the nutrient limitations (Asai *et al.*, 2009; Noguera *et al.*, 2010; Pratiwi & Shinogi, 2016; Haeefe *et al.*, 2011) (Table 1). Due to the presence of conflicting results with adding biochar, further testing is warranted before making recommendations.

Site-specific nutrient management

Uniform and/or high fertiliser applications across large regions likely result in low fertiliser-use efficiency of rice due to the heterogeneity of soil nutrient availability and crop uptake. One possibility to improve productivity and nutrient use efficiency in such situations is to use the site- or even field-specific nutrient and crop management (Dobermann *et al.*, 1998, 2003; Gupta & Seth, 2007; Haeefe & Konboon, 2009; Nhamo *et al.*, 2014; Haeefe *et al.*, 2016). The site-specific nutrient management system has gained popularity in rice production in Africa and Asia (Dobermann *et al.*, 2002; Buresh *et al.*, 2010; Saito *et al.*, 2015; Niang *et al.*, 2017). Site-specific nutrient management may increase the grain yield of rice by 300-900 kg ha⁻¹, and N, P and K accumulation by 8-21 % compared with farmers' practice or grain obtained before practising the site-specific nutrient management package in Asia (Peng *et al.*, 1996, 2006; Wang *et al.*, 2001; Dobermann *et al.*, 2002; Khurana *et al.*, 2007; Sui *et al.*, 2013). The amounts and types of nutrients required through site-specific nutrient management may vary across sites, with variable spatial and temporal scales depending on the degree and type of nutrient limitations (*i.e.*, both macro- and micro-nutrients). Moreover, this approach explicitly recognises the need to efficiently utilise organic and inorganic nutrient sources (*i.e.*, IPNS) and minimise losses.

When synchronising nutrient demand and supply to rice plants, the potential use of slow-release fertilisers should receive much attention, particularly for nutrients such as N, as the losses are high, leading to low-use efficiency, and frequent applications are not practicable (Tang *et al.*, 2007; Ye *et al.*, 2013). Site-specific and slow-release synthetic fertiliser management strategies are possible only when farmers have access to them which is unlikely for most rice farmers in low-input agricultural systems. In such situations, the application of organic manure with a site-specific focus should be considered. However, this requires constant testing of available organic materials for nutrients, nutrient releasing patterns, and deciding on application rates to specific fields, thus making it less attractive to extension officers and farmers.

In-season nutrient management

The nutrient-use efficiency can be increased by improved in-season nutrient management, *e.g.*, N-use efficiency can be increased by 30-40% by increasing the number of splits with lower rates (Cassman *et al.*, 2002; Dobermann *et al.*, 2002). This is particularly important for N, as the losses are relatively high but can only be

reduced with higher labour requirements. Leaf-colour charts or chlorophyll meters are helpful in this, as it can determine the time and rate at which N fertiliser application is required to maximise the use efficiency, based on the depletion of leaf greenness (Hussain *et al.*, 2000; Singh *et al.*, 2002; Alam *et al.*, 2005; Gupta & Seth, 2007; Nhamo *et al.*, 2014). Experiments conducted in the Philippines revealed that the N fertiliser rate could be reduced by 30–45% without affecting the grain yield of rice when N is applied based on the leaf colour chart compared with the conventional practice (Hussain *et al.*, 2000). In Bangladesh, N application based on the leaf colour chart increased the grain yield by 0.1–0.7 t ha⁻¹ compared with farmer's management (Alam *et al.*, 2005).

When leaf colour chart-based N application was combined with site-specific nutrient management for other nutrients (*i.e.*, K, P, S and Zn), further improvement in grain yield by 0.4 t ha⁻¹ was observed (Alam *et al.*, 2005). Despite the benefits of using leaf colour charts and chlorophyll meters in enhancing the N-use efficiency, they are not available for nutrients other than N, and assume that other nutrients are not limiting. Moreover, those charts are not explanatory enough to less educated farmers and thus have limited applicability. The reasons hindering simple and promising techniques should be understood when managing in-season N nutrition efficiently.

Use of microorganisms

Integrated nutrient management with biological and chemical fertilisers can improve rice crop productivity, bio-fortification, soil health, and fertility (Hoseinzade *et al.*, 2016). Contribution from biological forms of fertilisation would be either through the enrichment of nutrient pools (*e.g.*, N) or by increasing the bioavailability of fixed forms of nutrients present in the soil (*e.g.*, P). Arbuscular mycorrhizal fungi (AMF) isolated from rice and non-rice soils increased plant growth and P and Zn nutrition of wetland rice (Khan & Belik, 1995; Roy & Srivastava, 2013). However, the contribution from AMF is usually more significant in aerobic/upland soil conditions. Hoseinzade *et al.* (2016) evaluated the combined effects of biological fertilisers, including AMF (*Glomus mosseae* Nicol. & Gerd.), free-living N-fixing bacteria (*Herbaspirillum seropedicae*), and chemical fertilisers on yield, nutrient concentrations, and contents of wetland rice. Chemical and biological fertilisation interactions resulted in significant improvement in grain Zn, Fe, P, and N, and soil Fe, K, and N nutrition. According to Hoseinzade *et al.* (2016), the most excellent rates of grain and straw nutrient accumulation

and soil nutrient enrichment result from combined use of biological and chemical fertilisers. Moreover, Jeong *et al.* (2015) found that arbuscular mycorrhizal symbiosis is established after four weeks of growth under aerobic soil conditions, and P nutrition of rice was improved. Therefore, introducing beneficial AMF to the soil, and/or adopting crop management practices that enhance the activity and contribution of AMF, such as aerobic rice culture, may enhance soil fertility and plant nutrition.

Large amounts of N derived from biological fixation have been reported in rice fields, and this contribution can be as much as 50 kg N ha⁻¹ (Cassman *et al.*, 1998). According to laboratory and greenhouse experiments, a significant improvement in rice N has resulted from *Herbaspirillum seropedicae* (Baldani *et al.*, 2000; Roy & Srivastava, 2013). In a different experiment, 40% N and 25% P chemical fertiliser could be supplemented by the low-cost, natural resource-based bio-fertiliser (*Azotobacter* sp. and phosphate-solubilising bacteria *Bacillus polymyxa*) at 12 kg ha⁻¹ and organic manure at 10 t ha⁻¹, possibly enabling long-term sustainability in rice cultivation (Mukhopadhyay *et al.*, 2013). In a separate experiment, Prasanna *et al.* (2015) tested the effect of cyanobacterial inoculation on the growth and yield of rice while reducing the application of synthetic fertiliser N. The treatments *Anabaena–Nostoc* consortium, *Trichoderma*, and *Anabaena–Trichoderma* biofilm recorded the highest values of grain and straw yield, which are on par with those of plots with recommended doses of synthetic fertilisers. Therefore, the benefits of cyanobacteria in improving nutrient availability, soil fertility, and crop productivity, with savings of 60 kg N ha⁻¹ season⁻¹ from synthetic fertilizer application, seem promising. Cyanobacterial inoculation increases grain yields by 10–24% compared to the non-inoculated control in different rice-cropping systems (De Caire *et al.*, 2000; Prasanna *et al.*, 2012). Moreover, the presence, activity, and associations between microorganisms, and thereby the benefits of soil microorganisms to rice plants, is affected by the duration and intensity of inorganic and organic fertiliser management (Li *et al.*, 2019; Zhou *et al.*, 2020). Biological fertilisers can save at least a fraction of recommended chemical fertilisers and improve the growth and yield of rice, especially if combined with the appropriate rate and type of fertilisers, by enhancing nutrient uptake.

In summary, to date, the general trend in rice cultivation is a decreasing use of organic and biological nutrient sources, mainly because of increasing opportunity costs of labour needed for collection, preparation, and application, and increasing use of synthetic fertilisers

when available and affordable. As opportunities are limited for breeding varieties that acquire more nutrients from the soil or have higher internal nutrient-use efficiencies, long-term management strategies must focus on improving soil fertility by applying alternative sources of nutrients, including better management of residues. In this context, improving SOC and advocating for site-specific and integrated nutrient management are promising options, as the application and management of these practices are simple and cost-effective. Moreover, bio-fertilisers act synergistically to improve soil fertility and grain yield of rice. The combined use of nutrient management strategies under the context of cultural and socio-economic conditions is also recognised as the IPNS.

INSTITUTIONAL SUPPORT AND EDUCATION OF FARMERS

Agronomic interventions alone are insufficient to ensure improved productivity and sustainability in rice fields. Improving productivity in nutrient-limited rice fields require collective actions at both field and regional scales. At the field level, there is a need for regular evaluation of soil fertility to make decisions on nutrient management, aiming to lower the cost of inputs and improve the soil fertility status. Challenges remain, as the transfer of technological and intuitive knowledge is restricted by the lack of routine farm-level soil and crop monitoring procedures. Thus, there is a requirement to focus on agronomic advice and support at the local farm and regional scales (Rahman & Parkinson, 2007; Stewart *et al.*, 2020). As a long-term strategy, nutrient limitations/availability heterogeneity can be mapped, along with other soil limitations, at field scale. Once mapped, breeding priorities can be identified in specific areas.

Moreover, simple decision-support tools can be developed using farmers' knowledge of their fields and planned crops to determine the best nutrient management package. When considering such site-specific nutrient management approaches, the major challenge is to reduce the complexity of new technologies disseminated to farmers. Site-specific nutrient management may vary in temporal and spatial scales, and farmers are not adopting new technologies at a fast rate (Dobermann *et al.*, 2002, 2003; Haefele & Konboon, 2009). Therefore, continuous guidance and directions should be provided to farmers until they become independent in using these technologies and improve soil fertility.

Apart from the field-level factors mentioned above, other major constraints to rice production at the regional scale are weed infestation, variable rainfall patterns, low and under-developed irrigation infrastructure, and socio-economic, institutional, and political considerations (*i.e.*, lack of financial resources, labour shortages, low levels of education, weak infrastructure, lack of conducive policies) (Stewart *et al.*, 2020). Solving these constraints along with the improvement of productivity of nutrient-limited rice fields to bridge the large existing gap between current and the potential productivity is necessary (Nhamo *et al.*, 2014; Stewart *et al.*, 2020). Farmers need access to credits and markets to follow up on recommendations (Nhamo *et al.*, 2014). Subsidies, improved agro-dealer networks, and re-packaging fertilisers and/or manures into smaller portions are measures used to improve fertiliser and/or manure use by resource-poor farmers (Poulton *et al.*, 2006). However, the success of such interventions has been low except for some models such as the OneAcre Fund in Africa and the track (*Yaya*) approach in Sri Lanka. The successful collective decision-making strategy adopted in the *Yaya* approach scheme, when making in-farm and off-farm decisions, also contributed to the *Yaya* approach success.

Agronomic planning (*e.g.*, recycling/importing crop residues or other forms of organic manure and selection of crops in rotation among the farming community) should be carried out with the participation of stakeholders (Gupta & Seth, 2007; Rodenburg *et al.*, 2014; Stewart *et al.*, 2020). This is perhaps the most challenging part of the approach, as getting all stakeholders together has proven difficult in many situations (Warner, 2006). Reaching consensus among a wide range of stakeholders with different interests might be another considerable hurdle. The outcome of such multi-stakeholder platform processes should aim for a compromise between economic, social, and environmental gains (Rodenburg *et al.*, 2014). Therefore, improving the productivity in nutrient-limited rice fields cannot be achieved only through agronomic means, and success depends on the input and commitment made by all stakeholders.

CONCLUDING REMARKS

The increase in productivity of rice lands has slowed down or even declined in the long term in most rice cultivating systems in Africa and Asia. Declining soil fertility is identified as one of the major causes for this. Currently, over two-thirds of the total rice area on the globe is classified as soil with nutrient deficiencies or

toxicities (only in localised areas), and yield gaps over 5 t ha⁻¹ have been reported from many of these systems. Widely reported soil fertility constraints are deficiencies of N, available P, S, and Zn, and Fe²⁺ toxicity. In general, the occurrence of nutrient deficiencies decreases in the sequence of uplands > rainfed lowlands > irrigated lowlands > deepwater areas.

Most of the global rice breeding programs are focused on improving productivity in high potential areas under optimal management. Additionally, numerous efforts have been made to breed rice varieties for pest and disease tolerance, flood tolerance, salinity tolerance, iron toxicity, and drought tolerance. Despite these breeding efforts, most rice farmers in Africa and Asia continue to grow rice in nutrient-deficient field conditions, mainly due to their poor economic status, lack of awareness, and/or lack of access to synthetic fertilisers. This problem is likely to be aggravated further. Thus, optimal management practices recommended for nutrient unlimited soils are not applicable, and the varieties developed for those conditions are not adoptable to nutrient-limited grounds. In order to sustain rice cultivation in nutrient-limited soils, soil fertility constraints must be identified and addressed without prioritising synthetic fertiliser application as a potential option. In addition, parallel implementation of other technologies and policies must be considered. The introduction of credit facilities to farmers, knowledge and technology in weed, pest, and disease management, the selection of suitable varieties, water management, the use of quality seeds, exposure to competitive markets, and IPNS are examples.

Improving SOC and nutrient pools through incorporating organic manure or farm yard manure should be practised without allowing the amount of nutrients released by organic sources to exceed crop nutrient demand to minimise nutrient losses. Apart from that, the practice of fallow or crop rotations, minimal or zero tillage practices, use of site-specific and intensive in-season nutrient management, and introducing and promotion of the activity of microorganisms fixing nutrients, such as bacteria, cyanobacteria and fungi, are promising soil carbon and nutrient improvement strategies. This integrated approach of the above components is recognised as IPNS. Depending on the availability of the above materials, institutional support and socio-economic status, farmers need to be educated to improve the soil fertility status gradually or at the least avoid further decline in soil fertility. The time required to replenish the soil fertility will depend on the type and severity of nutrient limitation, and the balance between nutrient input and output processes. Due to the

complex nature of the soil, and the socio-economic and institutional structures in these regions, an integrated and participatory approach, including all the stakeholders, is required to implement and sustain the agronomic interventions suggested above.

Acknowledgement

The author appreciates the comments made by unknown reviewers and the Editor of the Journal to improve an earlier version of this manuscript. Financial support provided by the grant AHEAD/RA3/DOR/STEM/No16 to collect some of the literature is greatly acknowledged.

REFERENCES

- Agboola A.A. (1981). The effects of different soil tillage and management practices on the physical and chemical properties of soil and maize yield in a rainforest zone of western Nigeria. *Agronomy Journal* **73**: 247–251.
DOI: <https://doi.org/10.2134/agronj1981.0002196200730020001x>
- Alam M.M., Ladha J.K., Khan S.R., Foyjunnessa, Harunur-Rashid Khan A.H. & Buresh, R.J. (2005). Leaf color chart for managing nitrogen fertilizer in lowland rice in Bangladesh. *Agronomy Journal* **97**: 949–959.
DOI: <https://doi.org/10.2134/agronj2004.0206>
- Andrade J.A., Alexandre C.A. & Basch G. (2010). Effects of soil tillage and mulching on thermal performance of a Luvisol topsoil layer. *Folia Oecologica* **37**(1):1–7.
- Asai H., Samson B.K., Stephan H.M., Songyikhangsuthor K., Homma K., Kiyono Y., Inoue Y., Shiraiwa T. & Horie T. (2009). Biochar amendment techniques for upland rice production in Northern Laos. *Field Crops Research* **111**: 81–84.
DOI: <https://doi.org/10.1016/j.fcr.2008.10.008>
- Awadhwal N. & Smith G.D. (1989). *New Implements for Crop Production in the Semi-Arid Tropics, Information Bulletin no. 27*. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India.
- Balasubramanian V. & Sekayange L. (1992). Five years of research on improved fallow in the semi-arid highlands of Rwanda. In: *Biological Nitrogen Fixation and Sustainability of Tropical Agriculture* (eds. K. Mulongoy, M. Gueye, D.S.C. Spencer), pp. 405–422. John Wiley & Sons, New Jersey, USA.
- Baldani V.L.D., Baldani J.I. & Döbereiner J. (2000). Inoculation of rice plants with the endophytic diazotrophs *Herbaspirillum seropedicae* and *Burkholderia* spp. *Biology and Fertility of Soils* **30**: 485–491.
DOI: <https://doi.org/10.1007/s003740050027>
- Becker M. & Asch F. (2005). Iron toxicity in rice-conditions and management concepts. *Journal of Plant Nutrition and Soil Science* **168**: 558–573.
DOI: <https://doi.org/10.1002/jpln.200520504>

- Becker M. & Johnson D.E. (1998). Legumes as dry season fallow in upland rice-based systems of West Africa. *Biology and Fertility of Soils* **27**: 358–367.
DOI: <https://doi.org/10.1007/s003740050444>
- Becker M. & Johnson D.E. (2001). Cropping intensity effects on upland rice yield and sustainability in West Africa. *Nutrient Cycling in Agroecosystems* **59**: 107–117.
DOI: <https://doi.org/10.1023/A:1017551529813>
- Bi Y., Cai S., Wang Y., Xia Y., Zhao X., Wang S. & Xing G. (2019). Assessing the viability of soil successive straw biochar amendment based on a five-year column trial with six different soils: Views from crop production, carbon sequestration and net ecosystem economic benefits. *Journal of Environmental Management* **245**: 173–186.
DOI: <https://doi.org/10.1016/j.jenvman.2019.03.051>
- Bi L. *et al.* (11 authors) (2009). Long-term effects of organic amendments on the rice yields for double rice cropping systems in subtropical China. *Agriculture, Ecosystems & Environment* **129**: 534–541.
DOI: <https://doi.org/10.1016/j.agee.2008.11.007>
- Blair G.J., Momuat E. & Mamaril C. (1979). Sulfur nutrition of rice. II. Effect of source and rate of S on growth and yield under flooded conditions. *Agronomy Journal* **71**: 477–480.
DOI: <https://doi.org/10.2134/agronj1979.0002196200710030023x>
- Buresh R.J., Pampolino M.F. & Witt C. (2010). Field-specific potassium and phosphorus balances and fertilizer requirements for irrigated rice-based cropping systems. *Plant and Soil* **335**: 35–64.
DOI: <https://doi.org/10.1007/s11104-010-0441-z>
- Buri M.M., Masunaga T. & Wakatsuki T. (2000). Sulfur and zinc levels as limiting factors to rice production in West Africa lowlands. *Geoderma* **94**: 23–42.
DOI: [https://doi.org/10.1016/S0016-7061\(99\)00076-2](https://doi.org/10.1016/S0016-7061(99)00076-2)
- Cassman K.G., Dobermann A. & Walters D.T. (2002). Agroecosystems, nitrogen-use efficiency, and nitrogen management. *AMBIO* **31**: 132–140.
DOI: <https://doi.org/10.1579/0044-7447-31.2.132>
- Cassman K.G., Peng S., Olk D.C., Ladha J.K., Reichardt W., Dobermann A. & Singh U. (1998). Opportunities for increased nitrogen-use efficiency from improved resource management in irrigated rice systems. *Field Crops Research* **56**: 7–39.
DOI: [https://doi.org/10.1016/S0378-4290\(97\)00140-8](https://doi.org/10.1016/S0378-4290(97)00140-8)
- Conen F., Yakutin M.V. & Sambuu A.D. (2003). Potential for detecting changes in soil organic carbon concentrations resulting from climate change. *Global Change Biology* **9**: 1515–1520.
DOI: <https://doi.org/10.1046/j.1365-2486.2003.00689.x>
- Cooper M., Rajatasereekul S., Immark S., Fukai S. & Basnayake J. (1999). Rainfed lowland rice breeding strategies for Northeast Thailand. *Field Crops Research* **64**: 131–151.
DOI: [https://doi.org/10.1016/S0378-4290\(99\)00057-X](https://doi.org/10.1016/S0378-4290(99)00057-X)
- Das B., Chakraborty D., Singh V.K., Aggarwal P., Singh R., Dwivedi B.S. & Mishra R.P. (2014). Effect of integrated nutrient management practice on soil aggregate properties, its stability and aggregate-associated carbon content in an intensive rice-wheat system. *Soil and Tillage Research* **136**: 9–18.
DOI: <https://doi.org/10.1016/j.still.2013.09.009>
- Das D.K., Maiti D. & Pathak H. (2009). Site-specific nutrient management in rice in Eastern India using a modeling approach. *Nutrient Cycling in Agroecosystems* **83**: 85–94.
DOI: <https://doi.org/10.1007/s10705-008-9202-2>
- de Caire G.Z., de Cano M.S., Palma R.M. & de Mulé C.Z. (2000). Changes in soil enzyme activities following additions of cyanobacterial biomass and exopolysaccharide. *Soil Biology and Biochemistry* **32**: 1985–1987.
DOI: [https://doi.org/10.1016/S0038-0717\(00\)00174-7](https://doi.org/10.1016/S0038-0717(00)00174-7)
- Dhanapala M.P. (2000). Bridging the rice yield gap in Sri Lanka. In: *Bridging the Rice Yield Gap in the Asia-Pacific Region* (eds. M.K. Papametriou, F.J. Dent & E.M. Herath.), pp. 135–144. FAO, Bangkok, Thailand.
- Diallo M.D., Wood S.A., Diallo A., Mahatma-Saleh M., Ndiaye O., Tine A.K., Ngamb T., Guisse M., Seck S., Diop A. & Guisse A. (2016). Soil suitability for the production of rice, groundnut, and cassava in the peri-urban Niayes zone, Senegal. *Soil and Tillage Research* **155**: 412–420.
DOI: <https://doi.org/10.1016/j.still.2015.09.009>
- Dobermann A., Cassman K.G., Mamaril C.P. & Sheehy J.E. (1998). Management of phosphorus, potassium, and sulfur in intensive, irrigated lowland rice. *Field Crops Research* **56**: 113–138.
DOI: [https://doi.org/10.1016/S0378-4290\(97\)00124-X](https://doi.org/10.1016/S0378-4290(97)00124-X)
- Dobermann A., Cassman K.G., Cruz P.C. Sta., Adviento M.A.A. & Pampolino M.F. (1996). Fertilizer inputs, nutrient balance and soil nutrient supplying power in intensive, irrigated rice system. III. Phosphorus. *Nutrient Cycling in Agroecosystems* **46**: 111–125.
DOI: <https://doi.org/10.1007/BF00704311>
- Dobermann A. & Fairhurst T. (2000). *Rice: Nutrient Disorders and Nutrient Management*. pp. 191. International Rice Research Institute, Los Banos, Philippines.
- Dobermann A. & Fairhurst T. (2002). Rice straw management. *Better Crops International* **16**: 7–11.
- Dobermann A. & Witt C. (2000). The potential impact of crop intensification on carbon and nitrogen cycling in intensive rice systems. *Proceedings of the Workshop on Carbon and Nitrogen Dynamics in Flooded Soils*. 19–22 April. International Rice Research Institute, Los Banos, Philippines. pp. 1–25.
- Dobermann *et al.* (17 authors) (2003). Soil fertility and indigenous nutrient supply in irrigated rice domains of Asia. *Agronomy Journal* **95**: 913–923.
DOI: <https://doi.org/10.2134/agronj2003.9130>
- Dobermann *et al.* (23 authors) (2002). Site-specific nutrient management for intensive rice cropping systems in Asia. *Field Crops Research* **74**: 37–66.
DOI: [https://doi.org/10.1016/S0378-4290\(01\)00197-6](https://doi.org/10.1016/S0378-4290(01)00197-6)
- Dossou-Yovo E.R., Brüggemann N., Ampofo E., Igue A.M., Jesse N., Huat J. & Agbossou E.K. (2016). Combining no-tillage, rice straw mulch and nitrogen fertilizer application to increase the soil carbon balance of upland rice field in northern Benin. *Soil and Tillage Research* **163**: 152–159.
DOI: <https://doi.org/10.1016/j.still.2016.05.019>
- Druilhe Z. & Barreiro-Hurlé J. (2012). *Fertilizer Subsidies in*

- Sub-Saharan Africa*. FAO, Rome, Italy.
- Du Z., Ren T., Hu C. & Zhang Q. (2015). Transition from intensive tillage to no-till enhances carbon sequestration in microaggregates of surface soil in the North China Plain. *Soil and Tillage Research* **146**: 26–31.
DOI: <https://doi.org/10.1016/j.still.2014.08.012>
- Duxbury J.M., Abrol I.P., Gupta R.K. & Bronson K.F. (2000). Analysis of long-term soil fertility experiments with rice-wheat rotations in South Asia. In: *Long-Term Soil Fertility Experiments in Rice-Wheat Cropping Systems* (eds. I.P. Abrol, K.F. Bronson, J.M. Duxbury & Gupta R.K.), pp 7–22. Rice-Wheat Consortium for the Indo-Gangetic Plains, New Delhi, India.
- Dwivedi B.S., Singh V.K. & Dwivedi V. (2004). Application of phosphate rock, with or without *Aspergillus awamori* inoculation, to meet phosphorus demands of rice wheat systems in the Indo-Gangetic plains of India. *Australian Journal of Experimental Agriculture* **44**: 1041–1050.
DOI: <https://doi.org/10.1017/EA03208>
- Fageria N. (2003). Plant tissue test for determination of optimum concentration and uptake of nitrogen at different growth stages in lowland rice. *Communications in Soil Science and Plant Analysis* **34**: 259–270.
DOI: <https://doi.org/10.1081/CSS-120017430>
- Fageria N.K., dos Santos A.B. & Cobucci T. (2011). Zinc nutrition of lowland rice. *Communications in Soil Science and Plant Analysis* **42**: 1719–1727.
DOI: <https://doi.org/10.1080/00103624.2011.584591>
- Fageria N.K. & Zimmermann F.J.P. (1998). Influence of pH on growth and nutrient uptake by crop species in an Oxisol. *Communications in Soil Science and Plant Analysis* **29**: 2675–2682.
DOI: <https://doi.org/10.1080/00103629809370142>
- Fairhurst T., Witt C., Buresh R., Dobermann A. & Fairhurst T. (2007). *Rice: A Practical Guide to Nutrient Management*. International Rice Research Institute. Los Banos, Philippines.
- Fu Y.-Q., Shen H., Wu D.-M. & Cai K.-Z. (2012). Silicon-mediated amelioration of Fe²⁺ toxicity in rice (*Oryza sativa* L.) roots. *Pedosphere* **22**: 795–802.
DOI: [https://doi.org/10.1016/S1002-0160\(12\)60065-4](https://doi.org/10.1016/S1002-0160(12)60065-4)
- Gami S., Ladha J., Pathak H., Shah M., Pasuquin E., Pandey S., Hobbs P., Joshy D. & Mishra R. (2001). Long-term changes in yield and soil fertility in a twenty-year rice-wheat experiment in Nepal. *Biology and Fertility of Soils* **34**: 73–78.
DOI: <https://doi.org/10.1007/s003740100377>
- Gami S.K., Lauren J.G. & Duxbury J.M. (2009). Soil organic carbon and nitrogen stocks in Nepal long-term soil fertility experiments. *Soil and Tillage Research* **106**: 95–103.
DOI: <https://doi.org/10.1016/j.still.2009.10.003>
- Ghimire R., Adhikari K.R., Chen Z.-S., Shah S.C. & Dahal K.R. (2012). Soil organic carbon sequestration as affected by tillage, crop residue, and nitrogen application in rice-wheat rotation system. *Paddy and Water Environment* **10**: 95–102.
DOI: <https://doi.org/10.1007/s10333-011-0268-0>
- Gupta R. & Seth A. (2007). A review of resource conserving technologies for sustainable management of the rice-wheat cropping systems of the Indo-Gangetic plains (IGP). *Crop Protection* **26**: 436–447.
DOI: <https://doi.org/10.1016/j.cropro.2006.04.030>
- Haefele S.M., Kato Y. & Singh S. (2016). Climate ready rice: Augmenting drought tolerance with best management practices. *Field Crops Research* **190**: 60–69.
DOI: <https://doi.org/10.1016/j.fcr.2016.02.001>
- Haefele S.M. & Konboon Y. (2009). Nutrient management for rainfed lowland rice in northeast Thailand. *Field Crops Research* **114**: 374–385.
DOI: <https://doi.org/10.1016/j.fcr.2009.09.007>
- Haefele S.M., Konboon Y., Wongboon W., Amarante S., Maarifat A.A., Pfeiffer E.M. & Knoblauch C. (2011). Effects and fate of black carbon from rice residues in rice-based systems. *Field Crops Research* **121**: 430–440.
DOI: <https://doi.org/10.1016/j.fcr.2011.01.014>
- Haefele S.M., Nelson A. & Hijmans R.J. (2014). Soil quality and constraints in global rice production. *Geoderma* **235**: 250–259.
DOI: <https://doi.org/10.1016/j.geoderma.2014.07.019>
- Haefele S.M., Wopereis M.C.S., Ndiaye M.K. & Kropff M.J. (2003). A framework to improve fertilizer recommendations for irrigated rice in West Africa. *Agricultural Systems* **76**: 313–335.
DOI: [https://doi.org/10.1016/S0308-521X\(02\)00080-X](https://doi.org/10.1016/S0308-521X(02)00080-X)
- Haefele S.M., Wopereis M.C.S., Schloebom A.M. & Wiechmann H. (2004). Long-term fertility experiments for irrigated rice in the West African Sahel: effect on soil characteristics. *Field Crops Research* **85**: 61–77.
DOI: [https://doi.org/10.1016/S0378-4290\(03\)00153-9](https://doi.org/10.1016/S0378-4290(03)00153-9)
- He W., Yang M., Li Z., Qiu J., Liu F., Qu X., Qiu Y. & Li R. (2015). High levels of silicon provided as a nutrient in hydroponic culture enhances rice plant resistance to brown planthopper. *Crop Protection* **67**: 20–25.
DOI: <https://doi.org/10.1016/j.cropro.2014.09.013>
- Holford I.C.R. (1997). Soil phosphorus: its measurement, and its uptake by plants. *Australian Journal of Soil Research* **35**: 227–239.
DOI: <https://doi.org/10.1071/S96047>
- Homma K., Horie T., Shiraiwa T., Supapoj N., Matsumoto N. & Kabaki N. (2003). Toposequential variation in soil fertility and rice productivity of rainfed lowland paddy fields in mini-watershed (Nong) in Northeast Thailand. *Plant Production Science* **6**: 147–153.
DOI: <https://doi.org/10.1626/pp.6.147>
- Hoseinzade H., Ardakani M.R., Shahdi A., Rahmani H.A., Noormohammadi G. & Miransari M. (2016). Rice (*Oryza sativa* L.) nutrient management using mycorrhizal fungi and endophytic *Herbaspirillum seropedicae*. *Journal of Integrative Agriculture* **15**: 1385–1394.
DOI: [https://doi.org/10.1016/S2095-3119\(15\)61241-2](https://doi.org/10.1016/S2095-3119(15)61241-2)
- Hossain M.S., Hossain A., Sarkar M.A.R., Jahiruddin M., Teixeira da Silva J.A. & Hossain M. I. (2016). Productivity and soil fertility of the rice-wheat system in the high ganges river floodplain of Bangladesh is influenced by the inclusion of legumes and manure. *Agriculture, Ecosystems & Environment* **218**: 40–52.

- DOI: <https://doi.org/10.1016/j.agee.2015.11.017>
- Hu X.-F., Cheng C., Luo F., Chang Y.-Y., Teng Q., Men D.-Y., Liu L. & Yang M.-Y. (2016). Effects of different fertilization practices on the incidence of rice pests and diseases: A three-year case study in Shanghai, in subtropical southeastern China. *Field Crops Research* **196**: 33–50. DOI: <https://doi.org/10.1016/j.fcr.2016.06.004>
- Huang M., Liang T., Wang L. & Zhou C. (2015). Effects of no-tillage systems on soil physical properties and carbon sequestration under long-term wheat-maize double cropping system. *Catena* **128**: 195–202. DOI: <https://doi.org/10.1016/j.catena.2015.02.010>
- Huang M., Yang L., Qin H., Jiang L. & Zou Y. (2013). Quantifying the effect of biochar amendment on soil quality and crop productivity in Chinese rice paddies. *Field Crops Research* **154**: 172–177. DOI: <https://doi.org/10.1016/j.fcr.2013.08.010>
- Hussain F., Bronson K.F., Singh Y., Singh B. & Peng S. (2000). Use of chlorophyll meter sufficiency indices for nitrogen management of irrigated rice in Asia. *Agronomy Journal* **92**: 875–879. DOI: <https://doi.org/10.2134/agronj2000.925875x>
- Jeong K., Mattes N., Catausan S., Chin J.H., Paszkowski U. & Heuer S. (2015). Genetic diversity for mycorrhizal symbiosis and phosphate transporters in rice. *Journal of Integrative Plant Biology* **57**: 969–979. DOI: <https://doi.org/10.1111/jipb.12435>
- Jiang X.-J. & Xie D.-T. (2009). Combining ridge with no-tillage in lowland rice-based cropping system: long-term effect on soil and rice yield. *Pedosphere* **19**: 515–522. DOI: [https://doi.org/10.1016/S1002-0160\(09\)60144-2](https://doi.org/10.1016/S1002-0160(09)60144-2)
- Katsura K., Tsujimoto Y., Oda M., Matsushima K.-I., Inusah B., Dogbe W. & Sakagami J.-I. (2016). Genotype-by-environment interaction analysis of rice (*Oryza* spp.) yield in a floodplain ecosystem in West Africa. *European Journal of Agronomy* **73**: 152–159. DOI: <https://doi.org/10.1016/j.eja.2015.11.014>
- Kawakye P.K. & Bobo U.G. (1995). Tillage depth and fertilization on wheat (*Triticum aestivum* L.) on a Nigerian savanna soil. *Tropical Agriculture* **72**: 126–129.
- Kekulandara D.S., Sirisena D.N., Bandaranayake P.C.G., Samarasinghe G., Wissuwa M. & Suriyagoda L.D.B. (2019). Variation in grain yield, and nitrogen, phosphorus and potassium nutrition of irrigated rice cultivars grown at fertile and low-fertile soils. *Plant and Soil* **434**: 107–123. DOI: <https://doi.org/10.1007/s11104-018-3663-0>
- Khan A.G. & Belik M. (1995). Occurrence and ecological significance of mycorrhizal symbiosis in aquatic plants. In: *Mycorrhiza: Structure, Function, Molecular Biology and Biotechnology* (eds. A. Varma & B. Hock), pp. 627–666. Springer, Berlin, Germany. DOI: https://doi.org/10.1007/978-3-662-08897-5_27
- Khunthasuvon S., Rajastasereekul S., Hanviriyapant P., Romyen P., Fukai S., Basnayake J. & Skulkhu E. (1998). Lowland rice improvement in northern and northeast Thailand. *Field Crops Research* **59**: 99–108. DOI: [https://doi.org/10.1016/S0378-4290\(98\)00109-9](https://doi.org/10.1016/S0378-4290(98)00109-9)
- Khurana H.S., Phillips S.B., Singh B., Dobermann A., Sidhu A.S., Singh Y. & Peng S. (2007). Performance of site-specific nutrient management for irrigated, transplanted rice in northwest India. *Agronomy Journal* **99**: 1436–1447. DOI: <https://doi.org/10.2134/agronj2006.0283>
- Kirchhof G., Priyono S., Utomo W.H., Adisarwanto T., Dacanay E.V. & So H.B. (2000). The effect of soil puddling on the soil physical properties and the growth of rice and post-rice crops. *Soil and Tillage Research* **56**: 37–50. DOI: [https://doi.org/10.1016/S0167-1987\(00\)00121-5](https://doi.org/10.1016/S0167-1987(00)00121-5)
- Kumar K. & Goh K.M. (1999). Crop residues and management practices: Effects on soil quality, soil nitrogen dynamics, crop yield, and nitrogen recovery. *Advances in Agronomy* **68**: 197–319. DOI: [https://doi.org/10.1016/S0065-2113\(08\)60846-9](https://doi.org/10.1016/S0065-2113(08)60846-9)
- Ladha et al. (22 authors) (2003). How extensive are yield declines in long-term rice-wheat experiments in Asia? *Field Crops Research* **81**: 159–180. DOI: [https://doi.org/10.1016/S0378-4290\(02\)00219-8](https://doi.org/10.1016/S0378-4290(02)00219-8)
- Ladha J.K., Pathak H.P., Krupnik T.J., Six J. & Kessel Chris V. (2005). Efficiency of fertilizer nitrogen in cereal production: Retrospect and prospects. *Advances in Agronomy* **87**: 85–156. DOI: [https://doi.org/10.1016/S0065-2113\(05\)87003-8](https://doi.org/10.1016/S0065-2113(05)87003-8)
- Lal R., Logan T.J. & Fausey N.R. (1989). Long-term tillage and wheel traffic effects on a poorly drained mollic ochraqualf in northwest Ohio. 1. Soil physical properties, root distribution and grain yield of corn and soybean. *Soil and Tillage Research* **14**: 341–358. DOI: [https://doi.org/10.1016/0167-1987\(89\)90054-8](https://doi.org/10.1016/0167-1987(89)90054-8)
- Lambers H., Chapin F.S. & Pons T.L. (2008). Photosynthesis, respiration, and long-distance transport. In: *Plant Physiological Ecology*, pp. 11. Routledge, New York, USA. DOI: https://doi.org/10.1007/978-0-387-78341-3_2
- Lehmann J. & Joseph S. (2015). *Biochar for Environment Management: Science and Technology*. Routledge, Abington, England. DOI: <https://doi.org/10.4324/9780203762264>
- Lehmann J., Silva J.P.D., Steiner C., Nehls T., Zech W. & Glaser B. (2003). Nutrient availability and leaching in an archaeological anthrosol and a ferralsol of the central Amazon basin: fertilizer, manure and charcoal amendments. *Plant and Soil* **249**: 343–357. DOI: <https://doi.org/10.1023/A:1022833116184>
- Li W., Liu M., Wu M., Jiang C., Kuzyakov Y., Gavrichkova O., Feng Y., Dong Y. & Li Z. (2019). Bacterial community succession in paddy soil depending on rice fertilization. *Applied Soil Ecology* **144**: 92–97. DOI: <https://doi.org/10.1016/j.apsoil.2019.07.014>
- Liu S., Zhang Y., Zong Y., Hu Z., Wu S., Zhou J., Jin Y. & Zou J. (2016). Response of soil carbon dioxide fluxes, soil organic carbon and microbial biomass carbon to biochar amendment: a meta-analysis. *GCB Bioenergy* **8**: 392–406. DOI: <https://doi.org/10.1111/gcbb.12265>
- Liu X., Zhou J., Chi Z., Zheng J., Li L., Zhang X., Zheng J., Cheng K., Bian R. & Pan G. (2019). Biochar provided limited benefits for rice yield and greenhouse gas mitigation six years following an amendment in a fertile rice paddy.

- Catena* **179**: 20–28.
DOI: <https://doi.org/10.1016/j.catena.2019.03.033>
- Lu F.E.I., Wang X., Han B., Ouyang Z., Duan X., Zheng H.U.A. & Miao H. (2009). Soil carbon sequestrations by nitrogen fertilizer application, straw return and no-tillage in China's cropland. *Global Change Biology* **15**: 281–305.
DOI: <https://doi.org/10.1111/j.1365-2486.2008.01743.x>
- Ly P., Jensen L.S., Bruun T.B. & de Neergaard A. (2016). Factors explaining variability in rice yields in a rain-fed lowland rice ecosystem in Southern Cambodia. *NJAS - Wageningen Journal of Life Sciences* **78**: 129–137.
DOI: <https://doi.org/10.1016/j.njas.2016.05.003>
- Ma J.F. & Takahashi E. (2002). *Soil, Fertilizer, and Plant Silicon Research in Japan*, pp. 191–200 Elsevier Science, Netherlands.
DOI: <https://doi.org/10.1016/B978-044451166-9/50009-9>
- Malaviarachchi M.A.P.W.K., De Costa W.A.J.M., Kumara J.B.D.A.P., Suriyagoda L.D.B. & Fonseka R.M. (2016). Response of mung bean (*Vigna radiata* (L.) R. Wilczek) to an increasing natural temperature gradient under different crop management systems. *Journal of Agronomy and Crop Science* **202**: 51–68.
DOI: <https://doi.org/10.1111/jac.12131>
- Mayamulla S., Weerathne L., Marambe B., Sirisena D. & Suriyagoda L. (2017). Variation in seed nutrient content, seedling growth and yield of rice varieties grown in a paddy field without application of fertilisers for forty years. *Crop and Pasture Science* **68**: 337–348.
DOI: <https://doi.org/10.1071/CP17060>
- Meharg C. & Meharg A.A. (2015). Silicon, the silver bullet for mitigating biotic and abiotic stress, and improving grain quality, in rice? *Environmental and Experimental Botany* **120**: 8–17.
DOI: <https://doi.org/10.1016/j.envexpbot.2015.07.001>
- Mishra R., Joshi R.K. & Zhao K. (2018). Genome editing in rice: recent advances, challenges, and future implications. *Frontiers in Plant Science* **9**: 1361.
DOI: <https://doi.org/10.3389/fpls.2018.01361>
- Mukhopadhyay M., Datta J.K. & Garai T.K. (2013). Steps toward alternative farming system in rice. *European Journal of Agronomy* **51**: 18–24.
DOI: <https://doi.org/10.1016/j.eja.2013.06.005>
- Munda *et al.* (12 authors) (2018). Dynamics of soil organic carbon mineralization and C fractions in paddy soil on application of rice husk biochar. *Biomass and Bioenergy* **115**: 1–9.
DOI: <https://doi.org/10.1016/j.biombioe.2018.04.002>
- Nhamo N., Rodenburg J., Zenna N., Makombe G. & Luzi-Kihupi A. (2014). Narrowing the rice yield gap in East and Southern Africa: Using and adapting existing technologies. *Agricultural Systems* **131**: 45–55.
DOI: <https://doi.org/10.1016/j.agsy.2014.08.003>
- Niang *et al.* (23 authors) (2017). Variability and determinants of yields in rice production systems of West Africa. *Field Crops Research* **207**: 1–12.
DOI: <https://doi.org/10.1016/j.fcr.2017.02.014>
- Noguera D., Rondón M., Laossi K.-R., Hoyos V., Lavelle P., Cruz de Carvalho M.H. & Barot S. (2010). Contrasted effect of biochar and earthworms on rice growth and resource allocation in different soils. *Soil Biology and Biochemistry* **42**: 1017–1027.
DOI: <https://doi.org/10.1016/j.soilbio.2010.03.001>
- Osunbitan J.A., Oyedele D.J. & Adekalu K.O. (2005). Tillage effects on bulk density, hydraulic conductivity and strength of a loamy sand soil in southwestern Nigeria. *Soil and Tillage Research* **82**: 57–64.
DOI: <https://doi.org/10.1016/j.still.2004.05.007>
- Pan G., Zhou P., Li Z., Smith P., Li L., Qiu D., Zhang X., Xu X., Shen S. & Chen X. (2009). Combined inorganic/organic fertilization enhances N efficiency and increases rice productivity through organic carbon accumulation in a rice paddy from the Tai Lake region, China. *Agriculture, Ecosystems and Environment* **131**: 274–280.
DOI: <https://doi.org/10.1016/j.agee.2009.01.020>
- Peng S., Buresh R.J., Huang J., Yang J., Zou Y., Zhong X., Wang G. & Zhang F. (2006). Strategies for overcoming low agronomic nitrogen use efficiency in irrigated rice systems in China. *Field Crops Research* **96**: 37–47.
DOI: <https://doi.org/10.1016/j.fcr.2005.05.004>
- Peng S., Garcia F.V., Laza R.C., Sanico A.L., Visperas R.M. & Cassman K.G. (1996). Increased N-use efficiency using a chlorophyll meter on high-yielding irrigated rice. *Field Crops Research* **47**: 243–252.
DOI: [https://doi.org/10.1016/0378-4290\(96\)00018-4](https://doi.org/10.1016/0378-4290(96)00018-4)
- Peng S., Huang J., Sheehy J.E., Laza R.C., Visperas R.M., Zhong X., Centeno G.S., Khush G.S. & Cassman K.G. (2004). Rice yields decline with higher night temperature from global warming. *Proceedings of the National Academy of Sciences of the United States of America* **101**(27): 9971–9975.
DOI: <https://doi.org/10.1073/pnas.0403720101>
- Peng X., Ye L.L., Wang C.H., Zhou H. & Sun B. (2011). Temperature- and duration-dependent rice straw-derived biochar: Characteristics and its effects on soil properties of an Ultisol in southern China. *Soil and Tillage Research* **112**: 159–166.
DOI: <https://doi.org/10.1016/j.still.2011.01.002>
- Pheav S., Bell R.W., White P.F. & Kirk G.J.D. (2003). Fate of applied fertilizer phosphorus in a highly weathered sandy soil under lowland rice cropping, and its residual effect. *Field Crops Research* **81**: 1–16.
DOI: [https://doi.org/10.1016/S0378-4290\(02\)00191-0](https://doi.org/10.1016/S0378-4290(02)00191-0)
- Poulton C., Kydd J., Wiggins S. & Dorward A. (2006). State intervention for food price stabilisation in Africa: Can it work? *Food Policy* **31**: 342–356.
DOI: <https://doi.org/10.1016/j.foodpol.2006.02.004>
- Prasanna R., Adak A., Verma S., Bidiyaran N., Babu S., Pal M., Shivay Y.S. & Nain L. (2015). Cyanobacterial inoculation in rice grown under flooded and SRI modes of cultivation elicits differential effects on plant growth and nutrient dynamics. *Ecological Engineering* **84**: 532–541.
DOI: <https://doi.org/10.1016/j.ecoleng.2015.09.033>
- Prasanna R., Joshi M., Rana A., Shivay Y.S. & Nain L. (2012). Influence of co-inoculation of bacteria-cyanobacteria on

- crop yield and C-N sequestration in soil under rice crop. *World Journal of Microbiology and Biotechnology* **28**: 1223–1235.
DOI: <https://doi.org/10.1007/s11274-011-0926-9>
- Pratiwi E.P.A. & Shinogi Y. (2016). Rice husk biochar application to paddy soil and its effects on soil physical properties, plant growth, and methane emission. *Paddy and Water Environment* **4**: 521–532.
DOI: <https://doi.org/10.1007/s10333-015-0521-z>
- Quijano-Guerta C., Kirk G.J.D., Portugal A.M., Bartolome V.I. & McLaren G.C. (2002). Tolerance of rice germplasm to zinc deficiency. *Field Crops Research* **76**: 123–130.
DOI: [https://doi.org/10.1016/S0378-4290\(02\)00034-5](https://doi.org/10.1016/S0378-4290(02)00034-5)
- Raboin L.-M., Razafimahafaly A.H.D., Rabenjarisoa M.B., Rabary B., Dusserre J. & Becquer T. (2016). Improving the fertility of tropical acid soils: Liming versus biochar application? A long term comparison in the highlands of Madagascar. *Field Crops Research* **199**: 99–108.
DOI: <https://doi.org/10.1016/j.fcr.2016.09.005>
- Rahman S. & Parkinson R.J. (2007). Productivity and soil fertility relationships in rice production systems, Bangladesh. *Agricultural Systems* **92**: 318–333.
DOI: <https://doi.org/10.1016/j.agsy.2006.04.001>
- Ratnayake R.R., Perera B.M.A.C.A., Rajapaksha R.P.S.K., Ekanayake E.M.H.G.S., Kumara R.K.G.K. & Gunaratne H.M.A.C. (2017). Soil carbon sequestration and nutrient status of tropical rice based cropping systems: Rice-rice, rice-soya, rice-onion and rice-tobacco in Sri Lanka. *Catena* **150**: 17–23.
DOI: <https://doi.org/10.1016/j.catena.2016.11.006>
- Ray D.K., Ramankutty N., Mueller N.D., West P.C. & Foley J.A. (2012). Recent patterns of crop yield growth and stagnation. *Nature Communications* **3**: 1293.
DOI: <https://doi.org/10.1038/ncomms2296>
- Reddy G.B. & Reddy K.R. (1993). Fate of nitrogen-15 enriched ammonium nitrate applied to corn. *Soil Science Society of America Journal* **57**: 111–115.
DOI: <https://doi.org/10.2136/sssaj1993.03615995005700010021x>
- Rodenburg J., Zwart S.J., Kiepe P., Narteh L.T., Dogbe W. & Wopereis M.C.S. (2014). Sustainable rice production in African inland valleys: Seizing regional potentials through local approaches. *Agricultural Systems* **123**: 1–11.
DOI: <https://doi.org/10.1016/j.agsy.2013.09.004>
- Roder W., Phengchanh S. & Keoboulapha B. (1995). Relationships between soil, fallow period, weeds and rice yield in slash-and-burn systems of Laos. *Plant and Soil* **176**: 27–36.
DOI: <https://doi.org/10.1007/BF00017672>
- Roy M. & Srivastava R.C. (2013). Assembling BNF system in rice plant: frontier areas of research. *Current Science* **104**: 326–334.
DOI: <http://www.jstor.org/stable/24089633>
- Sahrawat K.L. (2000). Macro- and micronutrients removed by upland and lowland rice cultivars in West Africa. *Communications in Soil Science and Plant Analysis* **31**: 717–723. <https://doi.org/10.1080/00103620009370472>
- Sahrawat K.L. (2004). Organic matter accumulation in submerged soils. *Advances in Agronomy* **81**: 169–201.
DOI: [https://doi.org/10.1016/S0065-2113\(03\)81004-0](https://doi.org/10.1016/S0065-2113(03)81004-0)
- Sahrawat K.L. (2006). Organic matter and mineralizable nitrogen relationships in wetland rice soils. *Communications in Soil Science and Plant Analysis* **37**: 787–796.
DOI: <https://doi.org/10.1080/00103620600564034>
- Saito K., Asai H., Zhao D., Laborte A.G. & Grenier C. (2018). Progress in varietal improvement for increasing upland rice productivity in the tropics. *Plant Production Science* **21**: 145–158.
DOI: <https://doi.org/10.1080/1343943X.2018.1459751>
- Saito K., Dieng I., Toure A.A., Somado E.A. & Wopereis M.C.S. (2015). Rice yield growth analysis for 24 African countries over 1960–2012. *Global Food Security* **5**: 62–69.
DOI: <https://doi.org/10.1016/j.gfs.2014.10.006>
- Saito K. & Futakuchi K. (2009). Performance of diverse upland rice cultivars in low and high soil fertility conditions in West Africa. *Field Crops Research* **111**: 243–250.
DOI: <https://doi.org/10.1016/j.fcr.2008.12.011>
- Sánchez P.A. & Salinas J.G. (1981). Low-input technology for managing oxisols and ultisols in tropical America. *Advances in Agronomy* **34**: 279–406.
DOI: [https://doi.org/10.1016/S0065-2113\(08\)60889-5](https://doi.org/10.1016/S0065-2113(08)60889-5)
- Savant N.K., Datnoff L.E. & Snyder G.H. (1997). Depletion of plant-available silicon in soils: A possible cause of declining rice yields. *Communications in Soil Science and Plant Analysis* **28**: 1245–1252.
DOI: <https://doi.org/10.1080/00103629709369870>
- Shen J., Li R., Zhang F., Fan J., Tang C. & Rengel Z. (2004). Crop yields, soil fertility and phosphorus fractions in response to long-term fertilization under the rice monoculture system on a calcareous soil. *Field Crops Research* **86**: 225–238.
DOI: <https://doi.org/10.1016/j.fcr.2003.08.013>
- Shepherd J.G., Kleemann R., Bahri-Esfahani J., Hudek L., Suriyagoda L., Vandamme E. & van Dijk K.C. (2016). The future of phosphorus in our hands. *Nutrient Cycling in Agroecosystems* **104**: 281–287.
DOI: <https://doi.org/10.1007/s10705-015-9742-1>
- Singh B., Singh Y., Ladha J.K., Bronson K.F., Balasubramanian V., Singh J. & Khind C.S. (2002). Chlorophyll meter- and leaf color chart-based nitrogen management for rice and wheat in northwestern India. *Agronomy Journal* **94**: 821–829.
DOI: <https://doi.org/10.2134/agronj2002.8210>
- Singh G., Jalota S.K. & Sidhu B.S. (2005). Soil physical and hydraulic properties in a rice-wheat cropping system in India: effects of rice-straw management. *Soil Use and Management* **21**: 17–21.
DOI: <https://doi.org/10.1079/SUM2004285>
- Singh H. & Singh K.P. (1995). Effect of plant residue and fertilizer on grain yield of dryland rice under reduced tillage cultivation. *Soil and Tillage Research* **34**: 115–125.
DOI: [https://doi.org/10.1016/0167-1987\(95\)00455-2](https://doi.org/10.1016/0167-1987(95)00455-2)
- Sirisena D.N., Wanninayake W.M.M. & Silva A.G.S.D. (2016). Long term application of organic manure and chemical fertilizers on rice productivity and fertility in paddy growing

- soils in Kurunegala district. *Annals of the Department of Agriculture, Sri Lanka* **18**: 6–8.
- Sirisena D. & Suriyagoda L.D.B. (2018). Toward sustainable phosphorus management in Sri Lankan rice and vegetable-based cropping systems: A review. *Agriculture and Natural Resources* **52**: 9–15.
DOI: <https://doi.org/10.1016/j.anres.2018.03.004>
- So H.B., Kirchof G., Bakker R. & Smith G.D. (2001). Low input tillage/cropping systems for limited resource areas. *Soil and Tillage Research* **61**: 109–123.
DOI: [https://doi.org/10.1016/S0167-1987\(01\)00182-9](https://doi.org/10.1016/S0167-1987(01)00182-9)
- Somaweera K.A.T.N., Suriyagoda L.D.B., Sirisena D.N. & De Costa W.A.J.M. (2016). Accumulation and partitioning of biomass, nitrogen, phosphorus and potassium among different tissues during the life cycle of rice grown under different water management regimes. *Plant and Soil* **401**: 169–183.
DOI: <https://doi.org/10.1007/s11104-015-2541-2>
- Somaweera K.A.T.N., Suriyagoda L.D.B., Sirisena D.N. & De Costa W.A.J.M. (2017). Growth, root adaptations, phosphorus and potassium nutrition of rice when grown under the co-limitations of phosphorus, potassium and moisture. *Journal of Plant Nutrition* **40**: 795–812.
DOI: <https://doi.org/10.1080/01904167.2016.1201497>
- Stewart Z.P., Pierzynski G.M., Middendorf B.J. & Prasad P.V.V. (2020). Approaches to improve soil fertility in sub-Saharan Africa. *Journal of Experimental Botany* **71**: 632–641.
DOI: <https://doi.org/10.1093/jxb/erz446>
- Sui B., Feng X., Tian G., Hu X., Shen Q. & Guo S. (2013). Optimizing nitrogen supply increases rice yield and nitrogen use efficiency by regulating yield formation factors. *Field Crops Research* **150**: 99–107.
DOI: <https://doi.org/10.1016/j.fcr.2013.06.012>
- Suriyagoda L., De Costa W.A.J.M. & Lambers H. (2014). Growth and phosphorus nutrition of rice when inorganic fertiliser application is partly replaced by straw under varying moisture availability in sandy and clay soils. *Plant and Soil* **384**: 53–68.
DOI: <https://doi.org/10.1007/s11104-014-2049-1>
- Suriyagoda L.D.B., Dittert K. & Lambers H. (2018). Arsenic in rice agroecosystems: uptake, translocation and mechanisms to reduce arsenic accumulation in rice grain. *Agriculture, Ecosystems and Environment*. **253**: 23–37.
DOI: <https://doi.org/10.1016/j.agee.2017.10.017>
- Suriyagoda L.D.B., Sirisena D.N., Somaweera K.A.T.N., Dissanayake A., De Costa W.A.J.M. & Lambers H. (2017). Incorporation of dolomite reduces iron toxicity, enhances growth and yield, and improves phosphorus and potassium nutrition in lowland rice (*Oryza sativa* L). *Plant and Soil* **410**: 299–312.
DOI: <https://doi.org/10.1007/s11104-016-3012-0>
- Tang J.-C., Shibata A., Zhou Q. & Katayama A. (2007). Effect of temperature on reaction rate and microbial community in composting of cattle manure with rice straw. *Journal of Bioscience and Bioengineering* **104**: 321–328.
DOI: <https://doi.org/10.1263/jbb.104.321>
- Timsina J. & Connor D.J. (2001). Productivity and management of rice-wheat cropping systems: issues and challenges. *Field Crops Research* **69**: 93–132.
DOI: [https://doi.org/10.1016/S0378-4290\(00\)00143-X](https://doi.org/10.1016/S0378-4290(00)00143-X)
- Tsujimoto Y., Yamamoto Y., Hayashi K., Zakaria A.I., Inusah Y., Hatta T., Fosu M. & Sakagami J.-I. (2013). Topographic distribution of the soil total carbon content and sulfur deficiency for rice cultivation in a floodplain ecosystem of the Northern region of Ghana. *Field Crops Research* **152**: 74–82.
DOI: <https://doi.org/10.1016/j.fcr.2012.11.007>
- Vandamme *et al.* (14 authors) (2016). Genotypic variation in grain P loading across diverse rice growing environments and implications for field P balances. *Frontiers in Plant Science* **7**: 1435.
DOI: <https://doi.org/10.3389/fpls.2016.01435>
- Van Oort P.A.J. (2018). Mapping abiotic stresses for rice in Africa: Drought, cold, iron toxicity, salinity and sodicity. *Field Crops Research* **219**: 55–75.
DOI: <https://doi.org/10.1016/j.fcr.2018.01.016>
- Wang G., Dobermann A., Witt C., Sun Q., & Fu R. (2001). Performance of site-specific nutrient management for irrigated rice in Southeast China. *Agronomy Journal* **93**: 869–878.
DOI: <https://doi.org/10.2134/agronj2001.934869x>
- Wang C., Liu J., Shen J., Chen D., Li Y., Jiang B. & Wu J. (2018). Effects of biochar amendment on net greenhouse gas emissions and soil fertility in a double rice cropping system: A 4-year field experiment. *Agriculture, Ecosystems & Environment* **262**: 83–96.
DOI: <https://doi.org/10.1016/j.agee.2018.04.017>
- Warner J.F. (2006). More sustainable participation? Multi-stakeholder platforms for integrated catchment management. *International Journal of Water Resources Development* **22**: 15–35.
DOI: <https://doi.org/10.1080/07900620500404992>
- Weerahewa J. (2004). *Impacts of trade liberalization and market reforms on the paddy-rice sector in Sri Lanka*. pp. 151–188. Oxford University Press, New Delhi.
- Weerakoon W.M.W., Mutunayake M.M.P., Bandara C., Rao A.N., Bhandari D.C. & Ladha J.K. (2011). Direct-seeded rice culture in Sri Lanka: Lessons from farmers. *Field Crops Research* **121**: 53–63.
DOI: <https://doi.org/10.1016/j.fcr.2010.11.009>
- Wihardjaka A., Kirk G.J.D., Abdulrachman S. & Mamaril C.P. (1999). Potassium balances in rainfed lowland rice on a light-textured soil. *Field Crops Research* **64**: 237–247.
DOI: [https://doi.org/10.1016/S0378-4290\(99\)00045-3](https://doi.org/10.1016/S0378-4290(99)00045-3)
- Wissuwa M., Ismail A.M. & Yanagihara S. (2006). Effects of zinc deficiency on rice growth and genetic factors contributing to tolerance. *Plant Physiology* **142**: 731–741.
DOI: <https://doi.org/10.1104/pp.106.085225>
- Witt *et al.* (12 authors) (1999). Internal nutrient efficiencies of irrigated lowland rice in tropical and subtropical Asia. *Field Crops Research* **63**: 113–138.
DOI: [https://doi.org/10.1016/S0378-4290\(99\)00031-3](https://doi.org/10.1016/S0378-4290(99)00031-3)
- Wonprasaid S., Khunthasuvon S., Sittisuang P. & Fukai S. (1996). Performance of contrasting rice cultivars selected for rainfed lowland conditions in relation to soil fertility and water availability. *Field Crops Research* **47**: 267–275.
DOI: [https://doi.org/10.1016/0378-4290\(96\)00013-5](https://doi.org/10.1016/0378-4290(96)00013-5)

- Wu D., Feng Y., Xue L., Liu M., Yang B., Hu F. & Yang L. (2019). Biochar combined with vermicompost increases crop production while reducing ammonia and nitrous oxide emissions from a paddy soil. *Pedosphere* **29**: 82–94. DOI: [https://doi.org/10.1016/S1002-0160\(18\)60050-5](https://doi.org/10.1016/S1002-0160(18)60050-5)
- Xu X., Xie J., Hou Y., He P., Pampolino M.F., Zhao S., Qiu S., Johnston A.M. & Zhou W. (2015). Estimating nutrient uptake requirements for rice in China. *Field Crops Research* **180**: 37–45. DOI: <https://doi.org/10.1016/j.fcr.2015.05.008>
- Yamamoto T., Nakamura A., Iwai H., Ishii T., Ma J.F., Yokoyama R., Nishitani K., Satoh S. & Furukawa J. (2012). Effect of silicon deficiency on secondary cell wall synthesis in rice leaf. *Journal of Plant Research* **125**: 771. DOI: <https://doi.org/10.1007/s10265-012-0489-3>
- Yamaguchi J. (1999). Sulfur deficiency of rice plants in the Lower Volta area, Ghana. *Soil Science and Plant Nutrition*. **45**: 376–383. DOI: <https://doi.org/10.1080/00380768.1999.10409351>
- Ye Y., Liang X., Chen Y., Liu J., Gu J., Guo R. & Li L. (2013). Alternate wetting and drying irrigation and controlled-release nitrogen fertilizer in late-season rice. Effects on dry matter accumulation, yield, water and nitrogen use. *Field Crops Research* **144**: 212–224. DOI: <https://doi.org/10.1016/j.fcr.2012.12.003>
- Yuan J., Sha Z.-M., Hassani D., Zhao Z. & Cao L.-K. (2017). Assessing environmental impacts of organic and inorganic fertilizer on daily and seasonal greenhouse gases effluxes in rice field. *Atmospheric Environment* **155**: 119–128. DOI: <https://doi.org/10.1016/j.atmosenv.2017.02.007>
- Zhang M. & He Z. (2004). Long-term changes in organic carbon and nutrients of an Ultisol under rice cropping in southeast China. *Geoderma* **118**: 167–179. DOI: [https://doi.org/10.1016/S0016-7061\(03\)00191-5](https://doi.org/10.1016/S0016-7061(03)00191-5)
- Zhao M., Tian Y., Ma Y., Zhang M., Yao Y., Xiong Z., Yin B. & Zhu Z. (2015). Mitigating gaseous nitrogen emissions intensity from a Chinese rice cropping system through an improved management practice aimed to close the yield gap. *Agriculture, Ecosystems & Environment* **203**: 36–45. DOI: <https://doi.org/10.1016/j.agee.2015.01.014>
- Zhou G., Gao S., Lu Y., Liao Y., Nie J. & Cao W. (2020). Co-incorporation of green manure and rice straw improves rice production, and soil chemical, biochemical and microbiological properties in a typical paddy field in southern China. *Soil and Tillage Research* **197**: 104499. DOI: <https://doi.org/10.1016/j.still.2019.104499>

REVIEW

Food Safety

Heavy metals and food safety in Sri Lanka: A review

U Samarajeewa

Department of Food Science and Technologh, Faculty of Agriculture, University of Peradeniya, Peradeniya.

Submitted: 20 March 2022; Revised: 09 June 2022; Accepted: 24 June 2022

Summary: Entry of heavy metals to the food chain leads to food safety hazards. The origins of possible food safety hazards in Sri Lanka due to metalloid arsenic and the heavy metals cadmium, lead, and mercury are reviewed. Of them, arsenic and cadmium draw attention as contaminants in rice. Of the four heavy metals, cadmium in agricultural soils is of anthropogenic origin. Arsenic is of lithogenic origin. In some locations lead appears to be of anthropogenic origin, especially in commercial leafy vegetable cultivating soils. Marine fish, particularly swordfish and yellow fin tuna, occasionally carry cadmium and mercury concentrations above the tolerance limits established by the Codex and European Food Safety Authority. Heavy metals in well water are far below tolerance limits and are safe. Patterns of annual cancer incidences in Sri Lanka do not provide evidence to consider arsenic as a food safety hazard. Food safety hazards may occur with arsenic in the long term if attention is not paid to the quality of fertilizers or the current daily rice consumption level is not reduced. Arsenic being of lithogenic origin, unhealthy exposures cannot be prevented without affecting the food security of the country. High consumption of cadmium containing rice exposes Sri Lankans to health problems. Signs of hotspots of lead are visible. Food safety hazards are predicted by assessing exposure of humans based on their body weight and daily intake of hazardous constituents. Provincial Tolerable Weekly Intakes (PTWI) are calculated for Sri Lanka using information on heavy metals in foods from research publications. International food regulatory limits on heavy metals in foods are summarized. Horizontal standards for heavy metals in foods are developed to minimize food safety hazards in Sri Lanka.

Keywords: Arsenic, cadmium, food chain, lead, mercury.

INTRODUCTION

Entry of harmful agents through the food chain into the human body results in food safety hazards. Heavy metals have drawn scientific attention as a potential food safety hazard entering the human food chain. Of the heavy metals in the soils and the environment, cadmium (Cd), lead (Pb), mercury (Hg), and the metalloid arsenic (As) are of concern globally. Heavy metals are also described as toxic trace metals in the literature. The exposure of humans to heavy metals is linked to their presence in foods, as free metal particles, inorganic compounds, or organic compounds. Exposure of humans to heavy metals in agri-foods depends on the food processing, preparation, and consumption patterns associated with food cultures. Heavy metals of lithogenic origin may be naturally present in the soil and water, or get added to agricultural soils through anthropogenic activities, especially through fertilizers and pesticides. Heavy metals also may enter the agri-foods from organic waste, compost, animal dung, sewage sludges, irrigation water *etc.* used as manure. There are no mining activities, volcanic emissions, or glacial activities in Sri Lanka that could bring heavy metals from the core of the earth to surface soils. Sri Lankan agriculture depends totally on rain fed and irrigated water. Water from deep aquifers which may carry high concentrations of heavy metals are not used in agriculture. Deep well water in Sri Lanka is free of heavy metals. Heavy metals in foods is suspected

* Corresponding author (smrjee@gmail.com;  <https://orcid.org/0000-0003-4481-5204>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in any way.

as one of the probable causes of chronic diseases of unknown aetiology (CKDU) among Sri Lankans. Similar diseases are reported in a few other countries. Understanding the presence of heavy metals in the agronomic soils, their appearance in foods, and cause-effect relationship with chronic diseases in the country need in-depth multidisciplinary scientific studies. This publication examines the presence of heavy metals in Sri Lankan agricultural soils and food production environment, their probable origins, opportunities for them to enter the food chain, the Sri Lankan food consumption patterns and incidences of related effects on health. This review proposes regulatory standards to minimize food safety hazards. The focus is on arsenic, cadmium, and lead entering the food chain during crop production and mercury entering through marine fish.

HEAVY METALS IN AGRICULTURE AND FOOD SAFETY

Heavy metals are present naturally in varying concentrations in soils. They are distributed heterogeneously in the Earth's crust, resulting in differing views of their presence, and mapping of heavy metal hotspots. The heavy metal concentrations the Earth's crust is different from that of the core. The heavy metal concentrations on the crust change continuously due to environmental, agronomic, and natural factors. The suitability of soils for cultivation are decided based on the "threshold values" of heavy metals, above which there is a risk of heavy metals entering food crops. The aim is to ensure public health through a safe food supply. The average concentrations of heavy metals in the Earth's crust, their threshold values based on information from European and Indian studies, and those established by Finland, are presented in Table 1 (Ministry of Environment, Finland, 2007).

Most heavy metals of concern in Sri Lanka have been established to arise mainly from weathering of rocks and

soils and not from agrochemical contamination of soil surfaces or from deep well water (Jayawardene *et al.*, 2012). The homogeneity of arsenic in Sri Lankan paddy soils is evident from the presence of similar concentrations in 70 locations of 14 villages (Chandrajith *et al.*, 2005). In a study comparing the trace metals in Sri Lankan soils from agricultural and non-agricultural highlands and lowlands of the Anuradhapura district by principal component analysis, calculating the geo-accumulation index and pollution loading index, it is established that copper, nickel, zinc, and lead in agricultural soils are of lithogenic origin and cadmium is of anthropogenic origin (Sanjeevani *et al.*, 2017). Cadmium is one of the problematic heavy metals in human food chains in many countries.

Entry of heavy metals into the food chain

Every atom of a heavy metal or every molecule containing a heavy metal in soil does not end up in the food plate. The physical properties of the soil, the ability of soil particles to absorb, desorb or bind heavy metals, the redox potential of soils governed by oxygen availability, the soil pH, and biological properties of the plants govern the entry of heavy metals into plants. Redox potentials of paddy soils have opposing effects on the absorption of arsenic and cadmium by plants, whereas soil pH affects the bioavailability of arsenic and cadmium (Zhao & Wang, 2019). It is thus challenging to work towards food safety through adjustment of soil conditions. The heavy metals should be available in the soil at depths down to 30 cm on average, for them to be absorbed by food crops. The depth of extension of roots in the soils limit the opportunity for the heavy metals to enter plants. The plants are selective in up-taking heavy metals and in translocating them to the edible components of crops. These complexities make it difficult to scientifically predict the extent of movement of heavy metals from soil to food of plant origin. The entry of each toxic heavy metal into the human food chain needs to be understood in working towards food safety.

Table 1: Abundance of As, Cd, Pb and Hg in the Earth's crust and their threshold values (mg/kg)

Distribution	As	Cd	Pb	Hg	References
Average on Earth's crust*	1.8**	0.15	10	0.05	https://en.wikipedia.org/wiki/Earth%27s_crust
Natural abundance	8.93	0.48	29.9	0.13	Arunakumara <i>et al.</i> , 2013
Thresholds for agricultural soils	5	1	60	0.5	Govt. Decree 214.2007.doc (finlex.fi) of Finland
After sewage sludge applications in agriculture	50	3	300	1	EC Directive 86/278/EEC

*The values vary in different publications; **Range 0.1 - 40 mg/kg

Arsenic in Sri Lankan food chain

The average concentration of arsenic in the earth crust is 1.8 mg/kg with a possible range of 0.1 – 40 mg/kg in non-contaminated soils. The threshold value for arsenic in agricultural soils is 5 mg/kg. Arsenic could be present in 200 different forms in soil of which 60% are arsenates, 20% are arsenites, oxides, arsenides, silicates etc., and 20% as sulphosalts and sulphides (Lim *et al.*, 2014). Arsenites are more toxic than arsenates. Out of these, only a few forms of arsenic enter the rice grain. Though total arsenic in foods was used in assessing food safety hazards in the past, current toxicological studies focus on inorganic arsenic in foods. Of the total arsenic content in rice, 60-80% is inorganic arsenic which may increase up to 90% (Jose *et al.*, 2009).

Organic arsenic, mainly present as arsenobetaine, forming more than 90% of total arsenic in fish, pose no recognizable threat to human health. Organic arsenic is not absorbed in the human gut. No regulatory limits are established for inorganic arsenic in fish in the food regulatory system in developed countries, as its presence is low and negligible (Codex, 2015; EC- 2006; FDA, 2021).

The food safety hazards through arsenic in rice occurs due to following reasons.

1. High volumes of rice consumed by Sri Lankans make rice a major entry pathway.
2. Arsenic exists in rice as the toxic inorganic form constituting 60 – 90% of total arsenic.

3. Removal of arsenic in rice is limited to polishing and washing prior to cooking. Each washing takes away 15% of arsenic in rice.
4. There is no mechanism to avoid food safety hazards from rice cultivated in areas where arsenic content in soil or irrigation water is naturally high.
5. Reduction of arsenic regulatory limit of 0.2 mg/kg for rice to 0.1 mg/kg would eliminate 70% of rice in the market leading to food security problems culminating in undernutrition of the populations (FDA, 2016).

Absorption of arsenic by plants from soil and its accumulation finally in the edible tissues occur along a concentration gradient (Jose *et al.*, 2009). With 7.19 - 18.63 mg/kg of arsenic in rice plant roots in Bangladesh, the observed arsenic content in rice grains was 0.25 - 0.73 mg/kg (Bhattacharya *et al.*, 2009). In rice plants, the arsenic content in roots is reported to be 28-fold higher than in the shoot, and 75-fold higher than raw rice grains (Raman *et al.*, 2002). Approximately 1% of arsenic in the soil ends up in the rice grains. Raman *et al.*, (2002) postulate that rice in soils containing less than 14 mg/kg of arsenic could be considered safe for human consumption. West Bengal in India and Bangladesh continue to face food safety hazards due to higher concentrations of arsenic in rice and water. The main source of arsenic in Bangladesh is the deep well water used for irrigation, resulting in contamination of the rice soils over years. A comparison of total arsenic concentrations in soil, water, and rice from West Bengal, India, where severe exposure of populations to arsenic occurs, with that of Sri Lanka are given in Table 2.

Table 2: Total As in soil, water and rice grains in West Bengal, India and in Sri Lanka

Location	Soil arsenic (mg/kg)	Irrigated water arsenic ^a (µg/L)	Rice grain arsenic ^b (mg/kg)	References
West Bengal	1.38 - 12.27	110 - 760	0.25 - 0.73	Battacharya <i>et al.</i> , 2009
Dry zone Sri Lanka	0.45 - 1.04	0.015 - 0.361	0.09 - 0.26	Chandrajith <i>et al.</i> , 2011
	Dry zone; 0.5 – 24 (tank sediments)	Nikawewa, Girandurukotte	Girandurukotte, Nikawewa	Chandrajith <i>et al.</i> , 2005 Chandrajith <i>et al.</i> , 2008
Dry zone paddy soils	1.18 ± 0.59			Balasoorya <i>et al.</i> , 2021
Wet zone paddy soils	1.32 ± 0.85			Balasoorya <i>et al.</i> , 2021
CKDU hotspot soils	1.33 ± 0.60			Balasoorya <i>et al.</i> , 2021

a = FAO permissible limit for irrigation water is 0.10 mg/L (= 100 µg /L); Value for drinking water is 10 µg /L; b = Codex limit for arsenic in rice is 0.2 mg/kg

Arsenic in Sri Lankan soils are reported to originate from sulphide minerals of the basement rocks and the concentrations are insignificant to affect plant growth (Jayawardena, 2012). The concentrations of arsenic in

well water in Sri Lanka are far below the internationally accepted maximum tolerance limit of 10 µg/L (Table 22). The concentrations are reported by Herath *et al.*, 2018 (Table 3) and in the WHO report on CKDU, 2016 (Figure 1).

Table 3: Average As, Cd and Pb concentrations in well waters from the 25 districts in Sri Lanka (µg/L) (Herath *et al.*, 2018) Values are computed from district averages

Metal	Country average	SD	Range	MTL (µg/L)	Below MTL	High district averages ± SD & (maximum) µg/L
Total arsenic	1.38	2.68	0 - 66	10	21/25 districts	Batticaloa = 3 ± 3.2 (14) Mannar = 7 ± 11.7 (66) Mullativu = 3 ± 3.7 (13) Puttalam = 4 ± 4.1 (15)
Cadmium	0.008	0.007	0 - 0.05	3	All	
Lead	0.133	0.155	0 - 0.5	50	All	Galle: one sample had 228

Concentrations of 0.5 – 24 mg/kg of arsenic and 10 – 33 mg/kg of lead have been reported from the sediments at different levels in Malagane Tank, Deduruoya. The values probably reflect accumulation of arsenic from weathering rocks (Chandrajith, 2008). These concentrations do not reflect alarming anthropogenic contributions from paddy cultivations in the surroundings.

In pot experiments with arsenic concentrations of 60 mg/kg in soil, the corresponding arsenic concentrations in cultivated rice plants were 21 mg/kg in panicles at initiation stage and 23 mg/kg in the panicles at maturity stage (Rahman *et al.*, 2008). At these artificially high exposure levels to soil arsenic, less than 50% appears to enter the rice grains. The information in Table 2 indicates that the exposure of rice plants to arsenic through soils

Table 4: Average As, Cd, and Pb in rice from CKDU endemic and non-endemic areas (mg/kg) (Herath *et al.*, 2018)

Heavy metal [Sample number]	Polished rice		Unpolished rice		Area related to CKDU	Codex mg/kg
	Average ± SD	Maximum	Average ± SD	Maximum		
Total arsenic [67]	0.03 ± 0.04	0.20	0.03 ± 0.04	0.20	Endemic	0.2
Cadmium [67]	0.12 ± 0.19	0.87	0.16 ± 0.17	0.52	Endemic	0.4
Lead [67]	0.01 ± 0.02	0.08	0.00 ± 0.02	0.08	Endemic	0.2
Total arsenic [24]	0.03 ± 0.03	0.00	0.04 ± 0.03	0.09	Non endemic	0.2
Cadmium [24]	0.21 ± 0.24	0.65	0.18 ± 0.36	1.43	Non endemic	0.4
Lead [24]	0.00 ± 0.00	0.00	0.00 ± 0.00	0.02	Non endemic	0.2

Table 5: As, Cd and Hg in Sri Lankan rice (mg/kg) (Jayasekara & Fretas, 2005)

Heavy metals	Raw polished grains		Parboiled grains		Rice flour	
	Producer 1	Producer 2	Producer 1	Producer 2	Producer 1	Producer 2
Total arsenic	0.034 ± 0.006	0.034 ± 0.001	0.065 ± 0.012	0.092 ± 0.001	0.035 ± 0.001	0.061 ± 0.006
Cadmium	0.192	< 0.5	< 0.5	< 0.5	< 0.5	< 0.5
Mercury	< 0.01	< 0.01	0.033	< 0.01	< 0.01	< 0.01

[Codex tolerance limits for rice As = 0.2, Cd = 0.4 mg/kg]

in Sri Lanka is far below that of the pot experiments and in West Bengal. The soils and agronomic practices in Sri Lanka do not seem to result in entry of hazardous arsenic concentrations into the rice production-processing chain. Independent studies reported in Tables 4 and 5 by different groups of scientists support the view that rice arsenic concentrations cannot cause food safety hazards in Sri Lanka as the concentrations are well below Codex maximum tolerance limits.

The presence of arsenic in rice is a global phenomenon. Rice arsenic concentrations in USA and in several countries are given in Tables 6 and 7. A study of 901 polished white rice samples from 10 countries has shown 7-fold variation from the median. The lowest values were 0.04 mg/kg in Egypt and 0.07 mg/kg in India. The highest values were 0.25 mg/kg in USA and 0.28 mg/kg in France (Meharg *et al.*, 2009).

The arsenic concentrations in rice in Sri Lanka (Tables 2 and 4) are notably less than what is in Tables 6 and 7 indicating a low exposure. The arsenic concentrations reported in Tables 2 and 4 are less than the Codex tolerance limit of 0.2 mg/kg.

Correlation of the information on arsenic in the WHO study (2016) in Sri Lanka is given in Figure 1. The figure presents the food chain in blue boxes and arrows. The values given above the food chain (in red) indicate potential food safety hazards and outcomes associated with arsenic. The values given below the food chain (in green) represent values that do not point towards food safety hazards.

The evidence does not suggest a food safety hazard or health effects on humans arising from arsenic in agricultural soils or water. A long-term risk may arise from pesticides and excessive use of phosphate fertilizers containing arsenic on continuous application. The long-term effects are generally predicted by examining the exposure of agricultural lands to contaminated fertilizers over a period of 45 years. Arsenic available in added phosphate fertilizers may get diluted in the soil allowing only a fraction to reach rice as discussed by Raman *et al.*, (2002). The pesticides along with arsenic in them are absorbed through leaves and panicles leading to increased risk. However, the current evidence does not indicate food chain (rice and water) as a pathway for exposure of humans to arsenic in Sri Lanka.

Table 6: Concentrations of As in rice products in USA (Meharg *et al.*, 2009)

No	Rice type	No of samples	Mean (mg/kg)	SD (mg/kg)	Range (mg/kg)
1	All samples tested	193	0.194	0.144	0.006 - 0.723
2	Rice (non-organic)	88	0.205	0.122	0.047 - 0.559
3	Rice (organic)	13	0.174	0.142	0.086 - 0.526
4	Rice products (non- organic)	67	0.214	0.171	0.010 - 0.723
5	Rice products (organic)	25	0.125	0.142	0.006 - 0.620

Table 7: Concentrations of As in rice from several countries (Duxbury & Zavala, 2005)

Country	No. of samples	Arsenic (mg/kg)	Country	No. of samples	Arsenic (mg/kg)
Argentina	1	0.136	Korea	2	0.045
Bangladesh	3	0.046	Lebanon	1	0.169
Bhutan	1	0.032	Pakistan	3	0.033
China	2	0.146	Spain	2	0.186
Egypt	2	0.032	Thailand	9	0.093
Greece	1	0.114	USA	22	0.181
India	16	0.037	Venezuela	12	0.084
Italy	7	0.158	TOTAL	84	Mean 0.107

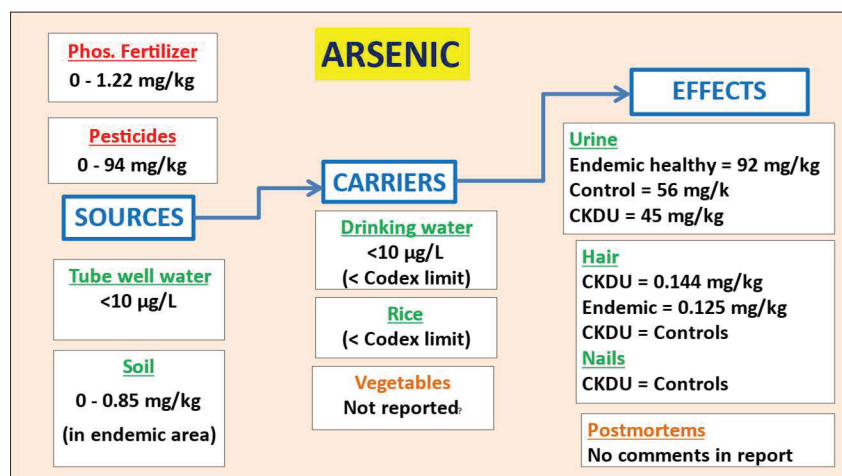


Figure 1: Concentrations of As linked to agronomic practices, appearing in food chain and in tissues of persons affected by CKDU.

While precaution is important, haste does not seem to be scientifically justifiable in assessing food safety hazards.

The major concern that may arise in Sri Lanka is the increased exposure to arsenic due to higher consumption of rice compared to developed nations, rather than high concentrations of arsenic in rice. The comparable per capita consumption values for rice in kg are Bangladesh 268; Vietnam 220; Thailand 178; Sri Lanka 160; China 125; India 98; Australia 13; USA 10; UK 8 and World average of 80 (Helgi library, 2017). The effects of volumes of rice consumed is addressed in Table 17.

Among other food sources, concerns about arsenic in canned fish has been expressed at times. Assessment of weekly, monthly or daily exposure to arsenic in foods is used in recognizing risks leading to food safety hazards. The average consumption of rice by a Sri Lankan is 438 g per day. The corresponding value for fish in Gampaha district, which get plenty of fish from Negombo is 43 g per day (Jayasinghe *et al.*, 2018). With a 10-fold lower consumption of fish than rice, there is extremely low possibility of food safety risks through arsenic in fish. The arsenic in fish is 90% arsenobetaine and other organic forms which are non-toxic. Organic arsenic is not absorbed by the human gut. The European Food Safety Authority (EFSA), Joint Expert Committee of FAO/WHO (JECFA) guidelines used by Codex Alimentarius Commission (Codex), and the Food and Drugs Administration of USA (FDA) have not established regulatory limits for arsenic in fish, considering the insignificant contribution of toxic

inorganic arsenic in fish to human diet. Arsenic in fish, fruits and vegetables are 90% organic (Mandel & Suzuki, 2005).

Cadmium in Sri Lankan food chain

Cadmium could enter the human body through food and water. The average concentration of cadmium in the Earth's crust is 0.15 mg/kg. The threshold value for agricultural soils is 1 mg/kg (Table 1). Cadmium compounds contaminate agricultural soils from sewage, manure, and phosphate fertilizers. Soils naturally contain low concentrations of cadmium. Gunadasa *et al.*, (2021) reported that the cadmium and arsenic concentrations in paddy soils and in rice are less than 0.1 mg/kg. The soil cadmium concentrations observed in uncultivated soils in Sri Lanka is 25 - 50% below the threshold value (Sanjeevani *et al.*, 2017) and vegetable soils reach threshold values (Sanjeevani *et al.*, 2017) and vegetable soils reach threshold values (Table 8).

Rice, potatoes, grains, and vegetables tend to absorb cadmium from soil more readily than other plants. Green vegetables appear to possess a tendency to accumulate cadmium. Unacceptable cadmium concentrations have been reported in sword fish from most oceans. Entry of cadmium to plants is reported to be 2-fold higher than arsenic. Cadmium accumulate in kidneys over long periods, and its half-life in the human body is around 30 - 38 years, and is 10 years in the kidney. Cadmium is a Group 1 carcinogen affecting mainly the kidneys, followed by bones (IARC, 2018).

Table 8: Cd and Pb in low-country and up-country vegetable soils (Premarathna *et al.*, 2011)

Area	Sub area	No. fields	Cadmium (mg/kg)	Lead (mg/kg)
Low country	Sedawatta	4	0.61-3.28	39-113
	Welewatta	3	0.46-1.37	34-66
	Kotuwilla	3	0.98-1.31	20-56
	Kahathuduwa	10	0.49-1.55	17-33
	Bandaragama	2	0.53-0.89	15-15
	Mean \pm SD		1.18 \pm 0.82	54 \pm 29
	Control soil		0.26	49
Up country	Sithaeliya	4	0.51-0.88	56-311
	Kandapola	6	0.39-1.96	27-97
	Haputale	3	0.51-3.86	26-242
	Bogahakumbura	3	1.30-1.42	45-75
	Rahangala	2	1.22-1.29	97-116
	Mean \pm SD		1.21 \pm 0.80	88 \pm 75
	Control soil		0.51	40

Concentrations of cadmium observed in water and rice in Sri Lanka are below the Codex tolerance limit of 3 $\mu\text{g/L}$ and 0.4 mg/kg respectively (Tables 3, 4, and 5). Meharg *et al.* (2013) reported the highest quantities of rice consumption in Sri Lanka and Bangladesh, among 12 countries, and the concentration of cadmium is the highest in Sri Lankan rice. The mean cadmium concentration

reported by them is 0.081 ± 0.024 mg/kg, with the highest value of 0.80 mg/kg in the range for Sri Lanka. Though the concentration of cadmium in Sri Lankan rice is below the Codex tolerance limits (Table 9), the exposure of Sri Lankans appears high. The study has estimated the exposure of Sri Lankans and Bangladeshis to be 100 $\mu\text{g/kg}$ of body weight per week.

Table 9: Concentrations of Cd in market rice and projected maximum weekly exposures in several countries including Sri Lanka. (Meharg *et al.*, 2013)

Country	Number of samples	Mean (mg/kg)	Range (mg/kg)	Max weekly Cd intake in $\mu\text{g/kg}$ body weight
Bangladesh	260	0.099	< 0.0005 – 1.31	100
Cambodia	14	0.006	0.0010 – 0.03	5
France	37	0.010	0.0030 – 0.10	1
Ghana	428	0.020	<0.005 – 0.27	10
India	58	0.078	0.0020 – 1.00	12
Italy	114	0.038	0.0030 – 0.16	10
Japan	18	0.059	0.0101 – 0.14	10
Nepal	12	0.050	0.0139 – 0.08	10
Spain	92	0.024	0.0008 – 0.14	10
Sri Lanka	75	0.081	<0.0005 – 0.80	100
Thailand	18	0.027	0.0057 – 0.07	8
USA	21	0.018	0.0095 – 0.04	7

Leafy vegetables possess a higher tendency to accumulate cadmium than other plants (Westfall *et al.*, 2005; Gupta *et al.*, 2019). Vegetables from commercial plots with long term cultivations and vegetables available along the roadside on the Kesbewa – Kalutara road are reported (Premarathna, *et al.*, 2011; Kananke *et al.*, 2014) to carry cadmium concentrations above the Codex tolerance limit and EU tolerance limit of 0.2 mg/kg for leafy vegetables and 0.1 & 0.05 mg/kg for other vegetables (Table 10 and Figure 2).

Premarathna *et al.*, (2011) also reported high concentrations of cadmium in upcountry and low country vegetable soils, indicating the probable origin of cadmium in leafy vegetables (Table 8). The reported concentrations are 8-fold higher than the average cadmium concentrations in earth crust and 10-20% higher than the threshold values for agricultural soils indicated in table 1.

Table 10: Cd and Pb in low-country and up-country vegetables (Premarathna, *et al.*, 2011^a; Kananke *et al.*, 2014^b)

Area	Vegetable	Cd ^a (mg/kg)	Cd ^b (mg/kg)	Pb ^a (mg/kg)	Pb ^b (mg/kg)
Low country	<i>Kankun</i>	0.28 - 0.53	0.09 - 0.19	7.4 - 11.3	0.27 - 0.45
	<i>Mukumuwenna</i>	0.17 - 1.10	0.08 - 0.90	5.6 - 10.36	0.18 - 1.32
	<i>Sarana</i>	0.48 - 0.65		8.3 - 12.7	
	<i>Niwithi</i>	0.30 - 0.51	0.18 - 0.72	6.6 - 12.3	0.44 - 0.97
	<i>Thampala</i>	0.55 - 0.62	0.11 - 0.54	8.15 - 11.72	0.54 - 1.04
	<i>Gotukola</i>	0.54		8.75	
	<i>Kohila</i>		0.24 - 0.97		1.37 - 1.59
Up country	Potato	0.22 - 0.86		3.66 - 9.95	
	Lettuce	0.3		10.12	
	Cabbage	0.37 - 2.02		5.68 - 10.10	
	Leeks	0.48 - 0.54		3.84 - 10.45	
	Carrot	0.41 - 2.05		3.81 - 10.10	
	Knol-khol	1.28		7.65	

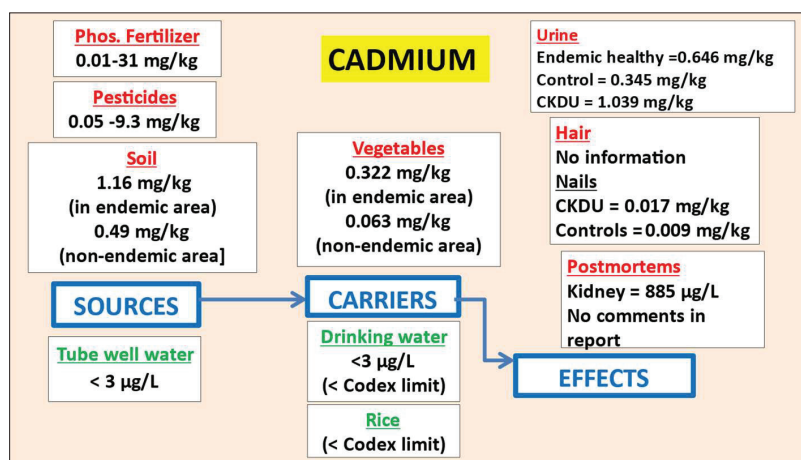


Figure 2: Concentrations of Cd linked to agronomic practices appearing in the food chain and in tissues of persons affected by CKDU.

The concentrations of cadmium in the market triple super phosphate (TSP) from locations in Anuradhapura, Medawachchiya, Medirigiriya, Girandurukotte, and Kandy are reported to be above the acceptable limits and vary widely compared to other fertilizer components urea and Nitrogen – Phosphorus – Potassium mixtures (NPK) (Chandrajith *et al.*, 2010). The results are summarized in Table 11.

Reports on cadmium concentrations in other possible sources of fertilizers and manures clearly indicate the probable origin of cadmium toxicity in Sri Lankan foods leading to food safety hazards (Table 12).

The extremely heavy concentrations of cadmium in imported triple superphosphate, against local rock phosphate and other sources of phosphate fertilizers of

Table 11: Summary of Cd and Pb concentrations in market fertilizers in 6 locations of Sri Lanka.

Fertilizer	Cadmium (mg/kg)		Lead (mg/kg)	
	Mean ± SD	Range	Mean ± SD	Range
Urea	0.3 ± 0.2	n.d – 0.4	4.2 ± 0.9	3.7 – 6.0
NPK	0.5 ± 0.1	0.4 – 0.5	3.6 ± 0.5	2.6 – 3.9
TSP	49.9 ± 29.2	39.8 – 80.2	79.2 ± 53.5	58.2 – 166.0
Tolerance limits, Texas, USA	39		300	

Table 12: Cd and Pb in phosphate fertilizers and manures in Sri Lanka (Premarathna *et al.*, 2011)

Fertilizer / manure	Number of samples	Cd (mg/kg)	Pb (mg/kg)
Triple super phosphate	5	23.50	5.15
Eppawala rock phosphate	4	1.92	13.00
Rock phosphate (USA)	2	13.54	12.00
Rock phosphate (India)	2	12.18	13.50
Rock phosphate (Bi Ru – China)	2	14.1	14.20
Rock phosphate (Lucille)	2	15.8	12.00
Dolomite	2	9.06	16.9
Poultry manure	12	0.97	3.20
Cattle manure	10	0.43	1.10

Table 13: Pb, Cd, and As in household foods in CKDU endemic area (Ananda Jayalal *et al.*, 2019) and Codex tolerance limits

Food item	Number	As (mg/kg)	Cd (mg/kg)	Pb (mg/kg)
Rice	65	0.046	0.022	0.236
Flower vegetables	88	0.045	0.018	0.232
Leafy vegetables	43	0.066	0.019	0.212
Legume vegetables	19	0.020	0.019	0.185
Fruits	24	0.020	0.015	0.188
Fat and oil	6	0.020	0.027	0.154
Inland fish	2	0.020	0.015	0.057
Root & tuber crops	27	0.020	0.015	0.186
Codex MTL		0.20	0.40	0.20

Note: These foods represent farmer cultivations for their use

global origin, indicate the possible entry of cadmium at unsafe concentrations into the food chain in Sri Lanka. Contrary to the above observations on cadmium in foods, a study on heavy metals in household foods in farmer families indicated low arsenic and cadmium concentrations, and high lead concentrations (Table 13). The samples tested perhaps represent foods produced under controlled agronomic practices for family consumption by farmers.

Correlation of the information in the WHO study (2016) points towards unacceptable cadmium concentrations in soils, in vegetables, and the tissues from persons exposed in CKDU endemic and non-endemic areas (Figure 2).

Cadmium tends to accumulate in marine fish. A study on heavy metals in fish in Sri Lanka has revealed high cadmium concentrations in sword fish compared to other

fish (Table 14). Sword fish accumulates more cadmium than others. The values also represent fish ready to be exported.

Exposure of Sri Lankans to cadmium appears to occur through several routes with a common anthropogenic origin. Sanjeevani *et al.* (2017) identified cadmium as the only toxic heavy metal of anthropogenic origin in agricultural soils from Anuradhapura district. The distribution pattern of cadmium in agricultural soils of CKDU endemic and non-endemic regions, in commercial vegetable plots and the market vegetables (in non-endemic areas) indicates high exposure of crops to cadmium leading to food safety hazards. Evidence on fertilizer analysis further confirms this origin. Cadmium, probably arising from imported TSP threatens food safety by burdening the soil and food with cadmium. Fish does not seem to carry cadmium concentrations above the tolerance limits.

Table 14: Summary of total Hg, Cd and Pb in sea fish in Sri Lanka

Fish (number of samples)	Hg (mg/kg)	Cd (mg/kg)	Pb (mg/kg)	Reference
Cooked sword fish (11)	nd – 1.47	nd – 0.029	-	Jayasinghe <i>et al.</i> , 2018
Cooked yellowfin tuna (11)	nd – 0.95	0.007 – 0.049	-	
Cooked sardinella (11)	nd – 0.09	0.007 – 0.021	-	
Sword fish (35)	1.24 ± 0.72 (0.20 – 2.58)	0.13 ± 0.08 (0.03 – 0.36)	0.03 ± 0.04 (nd – 0.15)	Jinadasa <i>et al.</i> , 2010
Yellow fin tuna (25)	0.39 ± 0.19 (0.14 – 0.88)	0.02 ± 0.02 (nd – 0.09)	0.06 ± 0.06 (nd – 0.24)	
Red snapper (12)	0.17 ± 0.06 (0.09 – 0.28)	0.02 ± 0.01 (nd – 0.04)	0.04 ± 0.05 (nd – 0.15)	
Export sword fish (176)	0.90 ± 0.51	0.09 ± 0.13	0.08	Jinadasa <i>et al.</i> , 2014
Export yellow fin tuna (140)	0.30 ± 0.18	-	0.11 ± 0.16	
Export red snapper (28)	0.16 ± 0.11	0.01 ± 0.01	0.04 ± 0.04	
Export marlin (24)	0.49 ± 0.37	0.02 ± 0.02	0.05 ± 0.05	
European Commission specifications (mg/kg)	1.0 = sword fish, tuna; 0.5 = other	0.3 = sword fish; 0.1 = tuna; 0.05 = other	0.3 = tuna, sword fish	EC 1881 (2006)
Codex	Predatory fish = 1; fish = 0.5	Bivalve mollusks = 2	Fish = 0.3	Codex 193-1995

Lead in Sri Lankan food chain

The average concentration of lead in the Earth's crust is 10 mg/kg. The threshold value for lead in agricultural soils is 60 mg/kg. Sewage sludges used in agriculture elsewhere carry up to 300 mg/kg of lead. It is not possible

to prevent the entry of lead naturally present in soil into the food chain through plants.

Lead in agricultural soils in Sri Lanka is mainly of lithogenic origin. Acceptable lead concentrations of 15 - 32 mg/kg have been reported from soils in Medirigiriya,

Talawa and Padawiya (Jayawardene *et al.*, 2012) and 4 mg/kg in rice soils recently (Balasooriya *et al.*, 2021). The average lead concentrations reported in soils of vegetable plots in the low country are within, yet close to the threshold value for agricultural soils, while the ranges of lead in upcountry vegetable soils reach 2 to 5-fold higher at the maxima of the ranges (Table 8).

The lead concentrations observed in urea and NPK mixtures of market fertilizers are around 4 mg/kg indicating a low contribution (Table 11). By contrast, the market TSP contributes 15 to 40-fold higher concentrations of lead compared to urea and NPK mixtures.

Herath *et al.* (2018) reported lead concentrations below 0.03 mg/kg in rice from CKDU endemic and non-endemic areas (Table 4). The lead concentrations observed in leafy vegetables from up country and low country are 20-fold higher compared to Codex and EU regulatory limits of 0.3 mg/kg (Table 10). These observations point more towards the probable anthropogenic origin of lead in some vegetable soils. The concentrations of lead observed in agricultural inputs, vegetables, and human test materials are correlated in Figure 3.

Ananda Jayalal *et al.*, (2019) observed lead concentrations of 0.20 mg/kg in rice and home-grown vegetables in CKDU affected families in the Anuradhapura district against the Codex tolerance limit of 0.3 mg/kg for leafy vegetables and 0.2 mg/kg for legume vegetables and pulses (Table 13). Table 14 shows

lead concentrations below the tolerance limits for sea fish (Jinadasa *et al.*, 2010; 2014a) and in Table 13 for inland fish (Ananda Jayalal *et al.*, 2020).

Though the lead concentrations in rice and vegetable were notably high in home-grown vegetables, the cadmium and arsenic concentrations were low (Table 13). This may be an indicative sign of food safety hazards in the local foods due to lead hotspots.

The average lead concentrations in water from wells in 25 districts in Sri Lanka is $0.133 \pm 0.155 \mu\text{g/L}$ against the Codex regulatory limit of $50 \mu\text{g/L}$ (Table 3). Though the lead concentrations in soils reported in the WHO study (2016) are within the threshold value, contribution from agricultural inputs appears high (Figure 3).

Lead concentrations in blood indicate human exposure. Madhavan *et al.* (1989) suggests that a concentration of 600 mg/kg of lead in soil contributes not more than $50 \mu\text{g/L}$ to blood in children under 12-years of age. In an indicative communication of blood lead levels, $100 - 220 \mu\text{g/L}$ has been reported in persons exposed to traffic polluted air in Kelaniya against $40 \mu\text{g/L}$ in control groups (Gunasekera *et al.*, 2015). If the level in the control group is considered as normal exposure through foods, it stands below the WHO accepted limit of lead in blood of $100 \mu\text{g/L}$ for adults and $50 \mu\text{g/L}$ for children. Concentrations of lead in bones of CKDU affected persons were about 60% more compared to control groups (Ananda Jayalal *et al.*, 2020) suggesting a risk through lead in diet at least to some population groups.

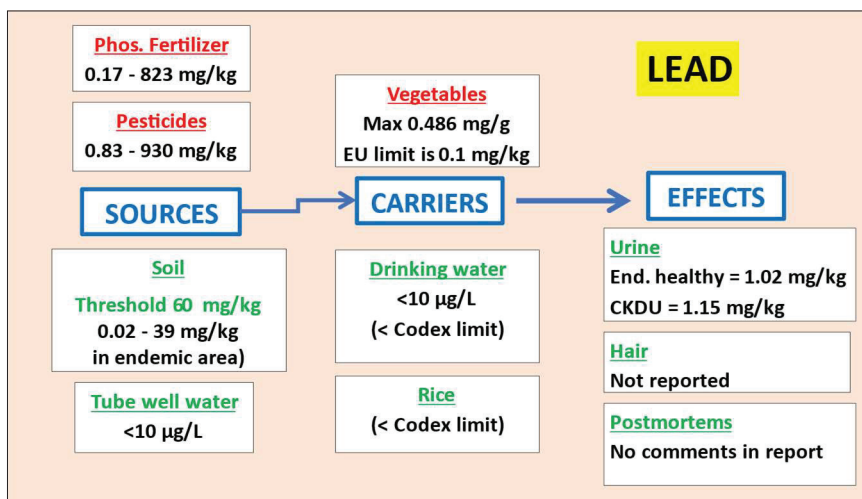


Figure 3: Concentrations of Pb linked to agronomic practices, appearing in the food chain and in tissues of persons affected by CKDU.

The high concentrations of lead observed in soils of vegetable plots (Table 8), in up-country and low-country vegetables (Table 10), in household vegetables of CKDU patients (Table 13) and the WHO study (Figure 3) together indicates probable exposure of Sri Lankans to unacceptable lead concentrations leading to food safety hazards needing attention at least in certain locations.

Mercury in Sri Lankan food chain

The average concentration of mercury in the Earth's crust is 0.05 mg/kg. The threshold value for agricultural soils is 0.5 mg/kg. Contamination of agricultural soils by mercury is not possible as there are no ores containing mercury in Sri Lanka. Mercury enters the food chain in Sri Lanka mostly through sea fish. Studies on mercury content in the Pacific and Indian Oceans have shown a mean total mercury concentration of 5.3 ng/L with a range of 3-6 ng/L in sea water with no significant variations at different depths of the oceans. Mercury occurs mostly as methyl mercury in fish and shellfish. Of the total mercury in fish, organic mercury could be 80% (Kannan *et al.*, 1998). Methyl mercury is of high toxicity (Nishimura *et al.*, 1983).

Of the different types of fish, swordfish and tuna are known to commonly accumulate mercury. Among four types of fish tested for mercury in Sri Lanka, swordfish is observed to carry high mercury concentrations, sometimes beyond the regulatory limits established by the Codex and EFSA (Table 14). Mercury concentrations in tuna fish also reach regulatory limits at times.

Mercury is not considered a serious food safety problem associated with rice, though the concentrations in rice are generally described to be high among grains. Mercury concentrations reported in rice in Sri Lanka are very low (Table 5). There is no tolerance limit established for mercury in rice in the Codex or EFSA. In China, mercury of anthropogenic origin has been reported in rice from districts where mercury is extracted from the ores.

Mercury could become a food safety hazard only with high consumption of swordfish and tuna.

FOOD SAFETY AND HUMAN HEALTH LINKED TO HEAVY METALS

Chronic exposure of humans to heavy metals could occur through air, water, and food, or even through skin absorption. The major health problems associated with

heavy metals of interest to Sri Lanka, identified from World Health Organization information, are as follows.

Arsenic (Arsenic key facts: WHO 2018)

Skin lesions leading to cancer, bladder cancer, and lung cancer. Increased risk of diabetes, pulmonary diseases, and cardiovascular diseases.

Cadmium: (Exposure to Cadmium, WHO 2019)

Primarily toxicity to kidney, leading to tubulointerstitial damage, affecting lung function inducing cancer, and causing bone demineralization due to interference with calcium metabolism.

Lead: (Lead poisoning – key facts, WHO 2021)

Cardiovascular effects, nervous disorders, decreased kidney function, and fertility problems

Mercury: (Mercury and Health, WHO 2017)

Toxic effects on nervous, digestive, and immune systems.

Health problems arising from exposure to heavy metals

Effects of heavy metals on public health are established by understanding the exposure dose and resulting toxicity or carcinogenicity.

Arsenic related health problems are associated with inorganic arsenic. The patterns of cancer reported for 2019 by the National Cancer Control Program of Sri Lanka (Cancer Incidence data book, 2019), published in November 2021, provide the following information on the crude rate of new cancer cases per one hundred thousand of population in 2019 (Table 15).

Skin, bladder or lung cancer, which are the predominant health problems arising from exposure to arsenic do not appear to be high in the classified information in the Table 15. The value for skin cancer is approximately 2% of total cancers. The incidence of lung cancer is higher with males associated with smoking. The ratio of male to female CKDU patients reported in Sri Lanka is approximately 1:1 (Jayatilake *et al.*, 2013). Bladder cancer and lung cancers are the other major type of cancers caused by arsenic in food. Comparison of bladder cancer rates in Colombo with the three agricultural areas is given in the Table 16.

The reported cancer patterns are not congruent with the districts where agrochemicals containing arsenic are applied. Review of the variations of the different cancers in Sri Lanka over the period 2005 to 2019 do not show relative differences in the variations of cancer incidences by types. The patterns suggest that the factors affecting

cancers in Sri Lanka are common to all types of cancers. Heavy metals in the food cannot be identified as a major contributor to cancer in Sri Lanka.

The incidence of CKDU in the North Central region of the country is well studied. Its symptoms, mainly the effects on the tubular cells of the kidney, are documented

(Jayatilake *et al.*, 2013). It appears to be a toxicological problem different from cancer. There are no published evidence from the health sector in the country to suspect direct food safety hazards related to lead and mercury. There are mechanisms based on food safety studies to predict the possibilities of food safety hazards associated with exposure to heavy metals under discussion.

Table 15: Crude cancer rate (CR*) in Sri Lanka by major types of cancers in 2019.

Males		Females	
Type	CR	Type	CR
Lip, tongue and mouth	20.6	Breast	39.5
Trachea, bronchi and lung	12.0	Thyroid	19.6
Colon and rectum	11.9	Colon and rectum	11.1
Oesophageal	10.4	Cervix, uteri	9.9
Prostate	9.5	Uterus	8.8
Pharynx	6.3	Oesophagus	7.3
Larynx	5.9	Lip, tongue, and mouth	8.8
Lymphoma	5.8	Ovary	8.5
Bladder	5.5	Trachea, bronchi and lung	4.1
Stomach	4.3	Lymphoma	3.9

CR* = number of new cancer cases diagnosed divided by at risk population multiplied by 100,000.

Table 16: Comparison of crude rates (CR) of bladder, tracheal, and bronchial & lung cancers in 4 districts of Sri Lanka in 2019.

CR bladder cancer among males		CR Tracheal, bronchial and lung cancer among females	
District	CR	District	CR
Colombo	5.3	Colombo	5.8
Anuradhapura	0.4	Anuradhapura	0.6
Polonnaruwa	3.7	Polonnaruwa	3.6
Batticaloa	8.4	Batticaloa	1.7

Assessment of exposure of Sri Lankans to heavy metals through food chain

Assessment of safety levels of hazardous substances are done by examining the quantity of hazardous agent that would be ingested by subjects taking into consideration the quantities of contaminated food, the body weights and durations of exposure. Durations may be a month, a week, or a day. This information is examined against the

adverse effects on public health to establish Provisional Tolerable Weekly Intakes (PTWI). The PTWI values are revised based on new evidence from the disease patterns associated with the hazardous component in the food. The units describing PTWI values are weight (mg / µg / ng) of the hazardous substance per kg of body weight per week. Si Lankan data on heavy metals in foods are examined using this concept to recognize the degree of food safety. The values are compared with international norms.

Risks through exposure to arsenic

A study predicted the global burden due to arsenic by applying models designed by the US Environment Protection Agency to WHO data on cancer. According to the predictions, increased incidence of bladder cancer from 9129 to 119,176, lung cancer from 11, 844 to 121,442, and skin cancer from 10,729 to 110,015 annually attributable to inorganic arsenic in food is expected (Oberoi *et al.*, 2014). The prediction represents an approximately 11-fold increase in the three types arsenic induced cancers. In contrast to the above global prediction, the information from the cancer records in Sri Lanka (Cancer Incidents data book – 2019, 2021)

does not indicate an increase in lung, skin or bladder cancer patterns different from other cancers in Sri Lanka. Patterns of all cancer incidences have remained the same with only a 3-fold increase of total cancer, and each type of cancer incidences over 10 years.

A risk assessment study on exposure to inorganic arsenic in rice was done by the FDA (2016) in USA. The study assessed the risks quantitatively by examining the population affected by bladder cancer and lung cancer, as influenced by inorganic arsenic in rice consumed by pregnant mothers and infants. A parallel qualitative examination looked at exposure-effect relationships linked to other health concerns.

Important information from the above study of relevance to Sri Lanka is as follows.

1. Heavy metal toxicity associated with rice is mainly due to inorganic arsenic, which is a large fraction of total arsenic in rice.
2. The availability of inorganic arsenic in rice is 0.092 mg/kg for white rice and 0.154 mg/kg for brown rice in the USA, which were used in the study.
3. It is noted that 28-60% of inorganic arsenic is lost during the washing of rice prior to cooking.
4. A single rinsing of rice with water removes up to 15% of inorganic arsenic in rice.
5. A selected population group in USA consuming 2-3 meals of rice per day are exposed to 435 ng inorganic arsenic per kilogram of body weight per day.
6. The regulatory limit of 0.2 mg/kg for arsenic in rice could reduce the cancer risk by 11%, and a limit of 0.075 mg/kg could reduce the risk by 79%.
7. Reducing regulatory limit arsenic from 0.2 to 0.1 mg/kg would decrease the market availability of rice by 4 - 93% depending on other factors.

8. A healthy balance between availability of rice (food security) and risk (food safety) needs to be established scientifically for consumer benefit.

To examine the exposure levels and predict the risks associated with arsenic in the Sri Lankan food chain, PTWI values were calculated (Table 17). The following assumptions were made in the calculations.

- a) The average body weight of a Sri Lankan is 55 kg.
- b) A Sri Lankan consumes 438 g of rice per day.
- c) Inorganic arsenic (I-As) in rice is 80% of total arsenic (T-As).

The Table 17 compares the exposure of Sri Lankans against recommendations of JECFA. An attempt to compare it with the USA situation for those consuming 2-3 meals per day, as reported by the FDA, was also made. The basis used here is the calculation of PTWI in μg of inorganic arsenic per kg of body weight over one week. Calculations and predictions here involve varying amounts of uncertainty due to approximations and assumptions.

Based on above calculations, it appears that the concentrations of arsenic naturally present in rice in Sri Lanka would not pose a significant risk to human health through weekly exposure. If the PTWI value is reduced to 10 or 12 in the future by JECFA based on new exposure-effect evidence, the calculation would show a tendency towards an increased risk through rice at the maximum reported values from Girandurukotte, but not with values in the other two studies. Scientifically it is not possible to reduce the amount of natural arsenic in soil. The concentrations of arsenic naturally present in Sri Lankan soils are about 1/5 threshold value for agricultural soils. It may be possible to minimize

Table 17: Comparison of exposure of Sri Lankans to As through rice on different published data converted to PTWI.

I – As in rice (mg/kg)	Notes	Exposure $\mu\text{g}/\text{kg}$ bw (I-As) [PTWI]
	JECFA exposure value withdrawn in 2015	15*
0.20 Codex limit	Applied to 438g rice / day in SL	11.1
	Based on 435 ng/kg bw/day (FDA, 2016)	3.0
0.07 = 80% T-As	Maximum reported (Jayasekara & Fretas, 2005)	3.9
0.21 = 80% T-As	Max. Girandurukotte (Chandrajith <i>et al.</i> , 2011)	11.7
0.03 = 80% T-As	Mean reported (Herath <i>et al.</i> , 2018)	1.7

* Protection provided by Provisional Tolerable Weekly Intake (PTWI) of 15 μg I-As/kg bw was found to be inadequate and withdrawn in 2015

risk by preventing entry of inorganic arsenic to rice soils through rigorous testing of agricultural inputs or reduction of rice consumptions perhaps from 3 meals to 2 meals a day. The fertilizer import policy needs to address this food safety issue. The arsenic, cadmium, and lead concentrations in fertilizer are concerns in most countries. The State of Texas has established a limit for arsenic as 41 mg/kg in fertilizer (Westfall *et al.*, 2005). Other countries address the heavy metal concentrations in phosphate fertilizers or by examining heavy metals added to soil over 45-year period (Agriculture and agri-food Canada, 1997).

Risk through exposure to cadmium

Food safety concerns about cadmium in the Sri Lanka arise mainly due to its nephrotoxic effects. The high quantities of rice consumed in Sri Lanka and Bangladesh are known to expose the populations to unhealthy concentrations of cadmium compared to other rice consuming countries. PTWI values are used to recognize the exposure to cadmium from different food sources based on the information from published research data. The PTWI values for Sri Lanka arising from rice, vegetables and fish are presented in Table 18.

The following assumptions were made in calculating PTWI for exposure to cadmium.

- a) The average body weight of humans is 55 kg.
- b) The average daily consumption of rice is 438 g.
- c) The average daily consumption of vegetables is 1.73 portions (approximately 130 g)
- d) The average daily leafy vegetable consumption is 50 g.
- e) Weekly fish consumption is 300 g.

Thresholds derived for “safe” exposure to cadmium by international regulatory agencies are given as the basis to understand the risks.

The PTWI values indicates high exposure of Sri Lankans to cadmium from rice and green leafy vegetables. Uruguchi and Fujiwara (2012) reported the intake of cadmium in Japanese population is 3.0 µg/kg bodyweight per week, which is above the PTWI limit of EFSA. In this respect the PTWI values from Sri Lankan foods in the Table 18 indicates a very high degree of food safety risk.

The high concentrations of cadmium observed in urine, hair and postmortem kidneys of CKDU affected persons further indicate the exposure of Sri Lankans to cadmium through the food chain (Figure 2). Accumulation of cadmium in bones, which is a feature of cadmium toxicity is however not detected in the WHO study (2016) or by Ananda Jayalal *et al.* (2020). Bone demineralization arising from exposure to cadmium is described more as an indirect influence via induction of renal dysfunction (Godt *et al.*, 2006). The observations on relationship of cadmium with bone demineralization differ, though there is evidence supporting bone demineralization with Itai-Itai disease linked to cadmium in rice.

Patterns of daily intake of cadmium as a function of rice consumption is presented in Figure 4. The exposure due to rice consumption at 438 g per day in Sri Lanka is represented by the brown upward arrow. The red dotted line indicates the exposure of Sri Lankans to cadmium through rice at different concentrations of cadmium in rice and quantities of rice consumed.

Table 18: Comparison of exposure of Sri Lankans to Cd through foods based on published research data converted to PTWI.

PTWI	Notes
0.7	US Agency for toxic substances and diseases registry
2.5	European Food Safety Authority (EFSA, 2012)
5.8	Equivalent for JEFCA value of 25 µg/kg PTMI (monthly intake)
9.5	Rice based on Herath <i>et al.</i> , 2018 (Table 4) for 0.17 (average) mg/ kg
5.3	Vegetables based on Cd observed in endemic area (Figure 2)
1.0	Vegetables based on Cd observed in non-endemic area (Figure 2)
3.2	Leafy vegetables based on (Table 9); Approximated as 0.5 Cd mg/kg
2.5 – 7.0	Rice, vegetables and water combined based on Jayalal <i>et al.</i> (2015) 60 kg body weight (consumption: 152-419 µg Cd / week)
0.23	Fish at reported maximum value of 0.049 mg/kg (Jinadasa <i>et al.</i> , 2010); for maximum exposure with consumption of 300 g per week

The curves indicate the need to either reduce the cadmium in rice or reduce consumption of rice to levels around 2/3 of the current consumption to be in line with Codex standards or to a level of 1/3 to be in line with EFSA standards.

Evidence leading to food safety hazards through cadmium in the food chain is apparent through several studies. Sanjeevani *et al.* (2017) indicated only cadmium, among the heavy metals under discussion here, to be of anthropogenic origin. High cadmium concentrations in TSP used as a fertilizer in Sri Lanka is documented in Table 11 (Chandrajith *et al.*, 2010) and the heavy cadmium concentrations associated with vegetable production in Tables 8 and 10. The observations of

Meharg *et al.*, (2013) in Table 9 provide independent evidence on high exposure of Sri Lankans to cadmium. Cadmium is a known cause of tubulointerstitial damage to the kidney. Tubulointerstitial damage is the main lesion among CKDU patients.

When the contribution of cadmium from leafy vegetables (Figure 3) is added to the contribution from rice, the need to prevent the use of TSP as a fertilizer and application of better quality phosphate fertilizer becomes evident. It must be mentioned that the cadmium concentrations in Eppawala rock phosphate is much less compared to TSP detected in Sri Lankan markets (Table 11 and 12).

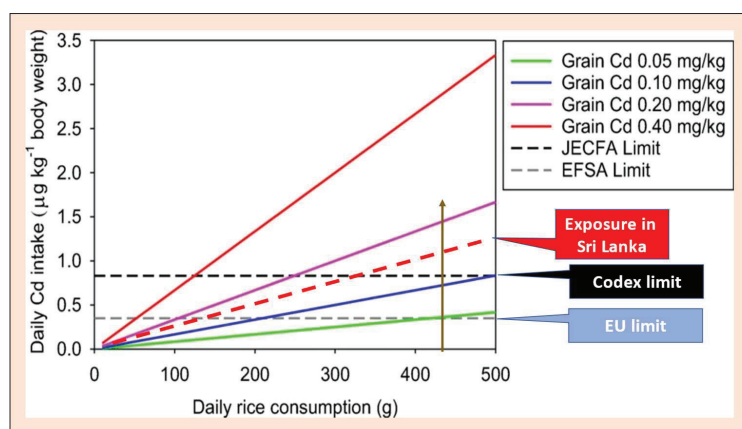


Figure 4: Daily intake of cadmium as a function of rice consumption modified from Zhao and Wang (2020).

Regulating the entry of heavy metals present in fertilizer into the human food chain is difficult. However, there are regulatory limits for heavy metals in fertilizers in other countries. California has regulations for arsenic, cadmium, and lead in fertilizers. The regulatory limits are linked with the main components in fertilizers such as 1% P_2O_5 or 1% Zn arising from minerals. The regulations in Canada specify application limiting to accumulation over a 45-year period expressed as kg per hectare (Agriculture and Agri-foods Canada, 1997). The regulations in Texas specify the limits of arsenic, cadmium, lead, and mercury as 41, 39, 300 and 17 mg/kg of fertilizer respectively (Westfall *et al.*, 2005). A mechanism to limit the entry of cadmium into agricultural soils is a need, in order to prevent food safety hazards arising through commercial agriculture. Limiting the entry of cadmium to agricultural soils is highly relevant to Sri Lanka, in the light of evidence regarding high cadmium concentrations in rice and leafy vegetables arising from use of TSP.

Risk through exposure to lead

Risks in the food chain may arise from lead present naturally in soil or added through contaminated agricultural inputs. The lead content reported in rice is less than 0.03 mg/kg on the average (Table 4), which is much below the Codex tolerance limit of 0.2 mg/kg. Norton *et al.* (2014) have reported lead concentrations of 0.020 mg/kg and 0.048 mg/kg in white and brown rice from Sri Lankan markets respectively. Based on findings of Norton *et al.* (2014) and Herath *et al.* (2018), 0.03 mg/kg was taken as a basis for estimating the contribution to the PTWI of lead from rice.

With the lead content in water reported as 0.133 µg/L in well water (Herath *et al.*, 2018) against the Codex tolerance limit of 10 µg/L (WHO 2016), the risk arising from drinking water would be insignificant.

Maximum lead concentrations of 0.486 mg/kg reported in vegetables (WHO, 2016) and in leafy vegetables (Premarathna *et al.*, 2011) are of concern, and used in calculating PTWI.

The following assumptions were made in calculating the PTWI and PTTI values for exposure to lead through food. Each contribution may be used to recognize cumulative contribution.

- a) The average body weight of humans is 55 kg.
- b) Rice consumption is 438 g per day containing 0.03 mg/kg of Pb (Herath *et al.*, 2018)
- c) A mean contribution of 8.5 mg/kg from leafy vegetables (Table 10, Premarathna *et al.*, 2011)
- d) A weekly consumption of 300 g of fish containing a mean of 0.11 mg/kg lead.
- e) The total diet of farmers from lead hot spots consists of 330 g of food with 0.200 mg/kg lead.

The analysis indicates high exposure to lead through leafy vegetables.

In estimating the provisional tolerable intake of lead through food, a factor of 10 was applied to obtain an exposure level to achieve some margin of safety based on lead concentrations in blood (Carrington & Bolger, 1992). Estimation of lead in blood is the only reasonably acceptable level of understanding exposure to lead from a variety of sources, *i.e.*, food, water and air. The FDA classifies a population into 4 groups by age. When the reported lead concentrations in Sri Lankan foods is viewed from the daily exposure angle, the resulting Provisional Tolerable Total Intake (PTTI) values could be compared (Table 20). The values reflect daily exposure in micrograms.

The risk due to lead through exposure to vegetables appear to reach the PTTI value of the FDA for adults.

Table 19: Comparison of exposure of Sri Lankans to Pb through foods based on published research data converted to PTWI

PTWI	Notes
25	This Codex value was declared inadequate in 2010
1.67	At the rate of 0.03 mg/kg in rice (Herath <i>et al.</i> , 2018; Norton <i>et al.</i> , 2014); Consumption of 438 g rice per day
8.04	Based on WHO, 2016 (0.486 mg/kg); consumption of 130 g vegetables per day
52.2	Based on Table 9 (approx. maximum of 8.2 mg/kg); consumption of 50 g leafy vegetables per day.
7.7	Household rice and vegetables 330 g/day from CKDU patients with approx. 0.200 mg/kg of lead (Ananda Jayalal <i>et al.</i> , 2019); 60 kg body weight.
4.2	Fish (observed mean) 0.11 mg/kg in tuna fish (Jinadasa <i>et al.</i> , 2014)

Table 20: Comparison of exposure of Sri Lankans to Pb through foods based on published research data converted to PTTI

PTTI (µg/day)	Notes
6	Limit for children below 7 years
15	Limit for children above 7 years
25	Limit for pregnant women
75	Limit for adults
13.4 (Rice)	At the rate of 0.03 mg/kg in rice (Herath <i>et al.</i> , 2018; Norton <i>et al.</i> , 2014); Consumption 438 g rice per day
63.2 (Vegetables)	Based on WHO (2016) (0.486 mg/kg); consumption 130 g per day
410 (Leafy vegetables)	Based on Table 9 (approximate maximum lead of 8.2 mg/kg) and leafy vegetable consumption 50 g per day
66	Household rice and vegetables 330 g/day from CKDU patients approximated to 0.20 mg/kg of lead (Ananda Jayalal <i>et al.</i> , 2019)
4.4	Fish (mean) 0.11 mg/kg in tuna (Jinadasa <i>et al.</i> , 2014); consumption 43 g per day [mean exposure may be 1/10 of this]

The same trend is visible when the PTWI levels of the Codex is examined in the Table 19. The risk due to leafy vegetables appears to be extremely high based on an assumption of 50 g consumption per day, at least in the study in the vegetable cultivating villages. TSP appears to be a major source of lead entering Sri Lankan foods (Figure 3). Ananda-Jayalal *et al.*, (2020) reported high concentrations of lead in bones of autopsy samples of CKDU affected subjects.

Risk through exposure to mercury

Risk associated with mercury in the food chain occurs mainly on exposure to methyl mercury in fish, which is of higher toxicity than its inorganic forms. Methyl mercury accounts for 83% of total mercury in fish muscle from south Florida estuaries (Kannan *et al.*, 1998). In the same location, methyl mercury in estuary sediments were only 0.77%, and that of water is less than 52% of total mercury. Separate studies by the same research group in Sri Lanka have shown mean total mercury concentrations

of 0.43 in cooked, 0.90 in export, and 1.24 mg/kg in fresh swordfish, with a maximum value of 1.47 mg/kg (Table 14). These values were used in estimating the PTWI values for comparison with international norms.

The following assumptions were made in calculating the PTWI values for exposure to total mercury through fish.

- The average body weight of humans is 55 kg.
- The concentrations of mercury in fish is 1.24 mg/kg (mean value for sword fish) and 2.58 mg/kg (maximum reported for sword fish) (Jinadasa *et al.*, 2010).
- The average consumption of fish is 300 g per week, as observed in a study in Gampaha district (Jayasinghe *et al.*, 2018).
- The average consumption of pelagic fish, mainly swordfish, is 83 g per week (Jinadasa *et al.*, 2014b).

On the above basis, PTWI levels for mercury are shown in Table 21.

Table 21: Comparison of exposure of Sri Lankans to mercury through sea fish based on published research data converted to PTWI.

PTWI	Notes
1.6	JECFA PTWI for methyl mercury (2003)
5.0	JECFA PTWI for total mercury (1978)
2.3	Assuming consumption of 300 g per week of cooked fish containing 0.43 mg/kg total mercury (Jayasinghe <i>et al.</i> , 2018)
6.7	For a mean value of 1.24 mg/kg of total mercury in swordfish (Jinadasa <i>et al.</i> , 2010)
3.9	Maximum of 2.58 mg/kg of total mercury in swordfish (Jinadasa <i>et al.</i> , 2014b) with a weekly consumption of 83 g.

Table 22: Summary of food standards for heavy metals in developed countries (mg/kg)

Regulator	Arsenic	Cadmium	Lead	Mercury
Codex - food	0.1 [rice 0.2]	0.05-0.2	0.001 – 0.3	0.5 -1.0
Codex - water	10 µg/L	3 µg/L		0.1 µg/L
European Commission (ESFA)	0.1-0.3 [rice products]	0.05 -0.5 [meats] 0.05 [root vegetables] 0.2 [leafy vegetables]	0.02 – 0.5 0.1 [vegetables]	0.5 – 1.0 [fish]
USA - FDA	10 achievable in apple juice	50 ppb suitable for juices		
USA - EPA - water	10 µg/L	7 µg/L for tap water	0 µg/L for tap water	1 µg/L for tap water
FAO/WHO	10 µg/L for water	0.2 for food	0.3 for food 100 µg/L water for adults 10 µg/L water Canada	1 µg/kg body weight per week other than seafoods

If the data in the Table 21 are examined assuming that methyl mercury is 75% of total mercury (Florida estuaries showed 80%), the calculated PTWI values would be almost the same as JEFCA value for cooked fish, and would exceed that for fresh swordfish. Jinadasa *et al.* (2014a), in a study of 140 and 176 samples of yellow fin tuna and swordfish, concluded that they contribute only 9% and 27% total mercury respectively to the PTWI of JCEFA, based on consumption of 83 g large pelagic fish per week. They also reported that 32% of swordfish samples exceeded the EU tolerance limit of 1 mg/kg for mercury, but not any of the tuna fish samples.

The results suggest that there is a food safety risk to persons consuming tuna or swordfish as the only type of fish in Sri Lanka in high volumes.

Regulations on heavy metals in foods

Each country decides on the level of protection needed for its population from food safety hazards. This results in variations in food standards. A summary of the food standards by major food safety authorities in the World is given in Table 22.

Table 23 : Proposed horizontal food standards for heavy metals in Sri Lanka.

Heavy metal	Commodities	MTL mg/kg	Notes
Arsenic (Total)	Rice and rice- based products for adults	0.2	Recommendations take into consideration what is agronomically possible based on reported levels of arsenic in soil and water, leaving out what may be contributed by agricultural inputs, and accepting what is recognized internationally as safe.
Arsenic (Total)	Rice and rice -based products for infants	0.1	Risk reductions need to be addressed for more vulnerable groups, the infants, children, and pregnant mothers consuming special foods through rice-based cereals <i>etc.</i>
Arsenic (Total)	Drinking water	10 µg/L	Codex limit would be satisfactory under Sri Lanka conditions as total arsenic levels are 7 to10-fold less than the limit.
Arsenic (Total)	Vegetables and fruits	0.2	General Codex limit for arsenic in foods as there is no evidence of arsenic in vegetables and fruits in Sri Lanka
Arsenic (Inorganic)	Fish and fish products	0.2	The bulk of arsenic in fish is in a non-toxic form as arsenobetaine.
Cadmium	Rice and rice products	0.1	This appears achievable under current cultivation conditions and would prevent development of cadmium hot spots. The need is linked with observed high consumption of rice in Sri Lanka
Cadmium	Water	3 µg/L	Codex standard
Cadmium	Vegetables and fruits	0.2	Stringent conditions are needed considering current field observations. Controls through fertilizer and pesticides is possible. The suggested value is the Codex limit.
Cadmium	Fish & fish products	0.1	In line with the limits in many countries. It is achievable.
Lead	Rice & rice products	0.1	A more reasonable tolerance limit to discourage use of contaminated agricultural inputs while accommodating natural soil concentrations.
Lead	Water	10 µg/L	This low limit is suggested as it is already reflected in reports and would be beneficial to be used as a mechanism to reduce the overall exposure and the risk through lead
Lead	Vegetables and fruits	0.3	A low tolerance limit for vegetables to discourage use of agricultural inputs contaminated with lead. Same as Codex limit
Lead	Fish & fish products	0.3	Achievable. It is in line with EFSA limits.
Mercury	Rice & rice products	0.02	Mercury is not a food safety hazard associated with rice. It is best to use the Codex standard
Mercury	Water	1 µg/L	Same as Codex standard
Mercury	Vegetables and fruits	0.02	Mercury is not a food safety problem associated with vegetables. It is best to use the Codex standard.
Mercury	Fish & fish products	1.0 methyl mercury	For predatory fish (shark, swordfish, tuna). 0.5 for other fish. Same as Codex standards

All developed countries and most of the Southeast Asian countries moved to new food safety authorities and horizontal food safety standards during 2012 to 2020. Preparation of country specific standards needs deep understanding of the exposure levels and risks arising from heavy metals in foods and dietary patterns.

Recommendations

Considering this national need, the recommendations given in Table 23 are made to develop Horizontal Food Safety Standards in Sri Lanka.

CONCLUSION

Sri Lanka needs to understand the critical points in the food chain where controls should be introduced to ensure the food safety of the nation without affecting its food security. It requires addressing the problems with deep scientific understanding. This review is meant to provide the analyzed scientific information for Sri Lanka in its efforts to ensure a safe food supply.

REFERENCES

- Agriculture and Agri-Food Canada. (1997). Standards for methods in fertilizers and supplements, memorandum. T-4-93. *Fertilizer Act Regulations*. Government of Canada.
- Ananda-Jayalal T.B. (2015). Chronic kidney disease of uncertain aetiology; adding vital piece of information to the national project team report of Sri Lanka. *BioMedical Central Nephrology* **16**(216): 1–3.
DOI: <https://doi.org/10.1186/s12882-015-0211-5>
- Ananda-Jayalal T.B., Jayaruwan Bandara T.W.M.A., Mahawithanage S.T.C., Wansapala J.M.A. & Galappaththi S.P.L. (2019). A quantitative analysis of chronic exposure of selected heavy metals in a model diet in a CKD hotspot in Sri Lanka. *BioMedical Central Nephrology* **20**: 208.
DOI: <https://doi.org/10.1186/s12882-019-1371-5>
- Ananda-Jayalal T.B., Mahawithanage S.T.C., Senanayaka S.M.H.M.K. & Dasanayaka P.B. (2020). Evidence of selected nephrotoxic elements in Sri Lankan human autopsy bone samples of patients with CKDU and controls. *BioMedical Central Nephrology* **21**: 384.
DOI: <https://doi.org/10.1186/s12882-020-02049-4>
- Arunakumara K.K.I.U., Walpola B.C. & Yoon M-H. (2013). Current status of heavy metal contamination in Asia's rice lands. *Reviews in Environmental Science and Biotechnology* **12**: 355–377.
DOI: <https://doi.org/10.1007/s11157-013-9323-1>
- Balasooriya S., Diyabalanage S., Yatigammana S.K., Ileperuma O.A. & Chandrajith R. (2021). Major and trace elements in paddy soils in Sri Lanka with special emphasis on chronic kidney disease of undetermined origin. *Environmental Geochemical Health*. **44**: 1841–1855
DOI: <https://doi.org/10.1007/s10653-021-01036-4>
- Bhattacharya P., Samal A.C., Majumdar J. & Santra S. (2009). Transfer of arsenic from groundwater and paddy soil to rice plant (*Oryza sativa L.*): A micro level study in West Bengal, India. *World Journal of Agricultural Sciences* **5**(4): 425–431.
- Ministry of Health (2021). *Cancer Incidence and Mortality Data, Sri Lanka*. National Cancer Control Programme, Ministry of Health, Colombo, Sri Lanka.
- Carrington C.D. & Bolger P.M. (1992). An assessment on the hazards of lead in food. *Regulatory Toxicology and Pharmacology* **16**(3): 265–272.
DOI: [https://doi.org/10.1016/0273-2300\(92\)90006-U](https://doi.org/10.1016/0273-2300(92)90006-U)
- Chandrajith R., Dissanayake C.B. & Tobschall H.J. (2005). The abundance of rarer trace elements in paddy (rice) soils of Sri Lanka. *Chemosphere* **58**: 1415–1420.
DOI: <https://doi.org/10.1016/j.chemosphere.2004.09.090>
- Chandrajith R., Mahatantila K., Jayasena H.A.H. & Tobschall H.J. (2008). Geochemical characteristics of sediments from a reservoir (tank) ecosystem in Sri Lanka. *Paddy Water Environment* **6**: 363–371.
DOI: <https://doi.org/10.1016/j.chemosphere.2004.09.090>
- Chandrajith R., Seneviratna S., Wickramaarachchi K., Attanayake T., Aturaliya T.N.C. & Dissanayake C.B. (2010). Natural radionuclides and trace elements in rice field soils in relation to fertilizer application: study of a chronic kidney disease area in Sri Lanka. *Environmental Earth Science* **60**:193–201.
DOI: <https://doi.org/10.1007/s12665-009-0179-1>
- Chandrajith R., Nanayakkara S. K., Aturaliya T.N.C., Dissanayake C.B., Abeysekera T., Harada, K., Watanabe T. & Koizumi A. (2011). Chronic kidney diseases of uncertain etiology (CKDU) in Sri Lanka: geographic distribution and environmental implications. *Environmental and Geochemical Health* **33**: 267–268.
DOI: <https://doi.org/10.1007/s10653-010-9339-1>
- Commission Regulations (EC) (2006). *Commission Regulations 1881/2006 Setting Maximum Levels for Certain Contaminants in Foodstuffs*. European Union.
- Duxbury J.M. & Zavala Y.J. (2005). What are safe levels of arsenic in food and soils. *Term Paper*, Cornell University, Ithaca, NY, USA.
- EFSA (2012). Cadmium dietary exposure in the European population. *European Food Safety Authority Journal* **10**(1): 2551.
DOI: <https://doi.org/10.2903/j.efsa.2012.2551>
- FAO (2015). *Codex General Standard For Contaminants and Toxins in Food and Feed (CODEX STAN 193-1995)*. FAO, Rome, Italy.
- FDA (2016). *Arsenic in Rice and Rice Products; Risk assessment report (Version released for public comments)*. Center for Food Safety and Applied Nutrition. FDA, Maryland, USA.
- FDA (2021). *Fish and Fishery Product Hazards and Control Guidelines*. Center for Food Safety and Applied Nutrition. FDA, Maryland, USA.
- Gunadasa S.G., Tighe M.K. & Wilson S.C. (2021). Arsenic and cadmium and associated risk in farm soils of the dry zone Sri

- Lanka where chronic kidney disease of unknown etiology (CKDU) is endemic. *Exposure and Health* **14**: 543–556.
DOI: <https://doi.org/10.1007/s12403-021-00433-5>
- Gunasekera D., Amarasinghe K.K.H.A. & Deeyamulla M. (2015). Determination of lead exposure in humans by biomonitoring of selected Sri Lankan population. *Postgraduate Research Conference, University of Kelaniya*.
- Godt J., Scheidig F., Grosse-Siestrup C., Esche V., Brandenburg P., Andrea Reich R. & Groneberg D.A. (2006). The toxicity of cadmium and resulting hazards for human health. *Journal of Occupational Medicine and Toxicology* **1**:22.
DOI:<https://doi.org/10.1186/1745-6673-1-22>
- Gupta N., Yadav K.K., Kumar V., Kumar S., Chadd R.P. & Kumar A. (2019) Trace elements in soil-vegetables interphase: Translocation, bioaccumulation, toxicity and amelioration – A review. *Science of the Total Environment* **651**: 2927–2942.
DOI:<https://doi.org/10.1016/j.scitotenv.2018.10.047>
- Helgi Library (2017). *Rice Consumption Per Capita in World*. Helgi Library, Czech Republic. Available at <https://www.helgilibrary.com/indicators/rice-consumption-per-capita>.
- Herath H.M.A.S., Kawakami T., Nagasawa S., Serikawa Y., Motoyama A., Chaminda T., Weragoda G.G., Yatigammana S.K. & Amarasooriya A.A.G.D. (2018). Arsenic, cadmium, lead, and chromium in well water, rice, and human urine in Sri Lanka in relation to chronic kidney disease of unknown etiology. *Journal of Water Health* **16**(2): 212–222.
DOI: <https://doi.org/10.2166/wh.2018.070>
- IARC (2018). *IARC Monographs: Cadmium and Cadmium Compounds*. IARC, Lyon, France. Available at <https://monograph.iarc.who.int/wp-content>.
- Jayasekara R. & Fretas M.C. (2005). Concentration levels of major and trace elements in rice from Sri Lanka as determined by the k_0 standardization method. *Biological Trace Elements Research* **103**: 83–95.
DOI: <https://doi.org/10.1385/BTER:103:1:083>
- Jayasinghe G.D.T.M., Sandaruwan K.P.G.L., Silva D.W.L.U. & Jinadasa B.K.K.K. (2018). Total diet study approach in estimating mercury and cadmium levels using selected fish: A case study from Sri Lanka. *Ceylon Journal of Science* **47**(3): 275–279.
DOI: <https://doi.org/10.4038/cjs.v47i3.7534>
- Jayatilake N., Mendis S., Maheepala P. & Mehta F.R. (2013). Chronic kidney disease of uncertain aetiology: prevalence and causative factor in a developing country. *BioMedical Central Nephrology* **14**: 1–13.
DOI: <https://doi.org/10.4038/cjs.v47i3.7534>
- Jayawardena D.T., Pitawala H.M.T.G.A. & Ishiga H. (2012). Geochemical assessment of arsenic and selected elements in agricultural and non-agricultural soils in Sri Lanka. *Tropical Agriculturist* **20**: 1–19.
- Jinadasa B.K.K.K., Edirisinghe E.M.R.K.B. & Wickramasinghe I. (2014a) Total mercury, cadmium and lead levels in main export fish of Sri Lanka. *Food Additives and Contaminants: Part 2* **7**(4): 309–314.
DOI: <https://doi.org/10.1080/19393210.2014.938131>
- Jinadasa B.K.K.K., Ahmad S.B.N., Edirisinghe E.M.R.K.B. & Wickramasinghe I. (2014b). Mercury content in yellowfin tuna (*Thunnus albacores*) and swordfish (*Xiphias gladius*) and estimation of mercury intake. *Journal of Food Security* **2**(1): 23–26.
- Jinadasa B.K.K.K., Rameesha L.R.S., Edirisinghe E.M.R.K.B. & Rathnayake R.M.U.S.K. (2010). Mercury, cadmium and lead levels in three commercially important marine fish species of in Sri Lanka. *Sri Lanka Journal of Aquatic Science* **15**: 39–43.
DOI: <https://doi.org/10.4038/sljias.v15i0.5481>
- Jose R., Peralta-Videa, Lopez M. L., Narayan M., Saupe G. & Gardea-Torresdey J. (2009). The biochemistry of environmental heavy metal uptake by plants: Implications for the food chain. *The International Journal of Biochemistry and Cell Biology* **41**(8): 1665–1677.
DOI: <https://doi.org/10.1016/j.biocel.2009.03.005>
- Kananke T., Wansapala J. & Gunaratne A. (2014). Heavy metal contamination in green leafy vegetables collected from selected market sites of Piliyandala area, Colombo District, Sri Lanka. *American Journal of Food Science and Technology* **2**(5): 139–144.
DOI: <https://doi.org/10.12691/ajfst-2-5-1>
- Kannan K., Smith R.G., Jr., Lee R.F., Windom H.L., Heitmuller P.T., Macauley J.M. & Summers J.K. (1998). Distribution of total mercury and methyl mercury in water, sediment, and fish from south Florida estuaries. *Archives of Environmental Contamination and Toxicology* **34**(2):109–118.
DOI: <https://doi.org/10.1007/s002449900294>
- Lim K. T., Shukor M.Y. & Wasoh H. (2014). Physical, chemical, and biological methods for the removal of arsenic compounds. *Biomed Research International*. 2014
DOI: <https://doi.org/10.1155/2014/503784>
- Madhavan S., Roseman K.D. & Skehata T. (1989). Lead in soil: recommended permissible levels. *Environmental Research* **49**(1): 136–142.
DOI: [https://doi.org/10.1016/S0013-9351\(89\)80028-3](https://doi.org/10.1016/S0013-9351(89)80028-3)
- Mandal B.K. & Suzuki K.T. (2002). Arsenic round the world – a review. *Talanta* **58**: 201–235.
DOI: [https://doi.org/10.1016/S0039-9140\(02\)00268-0](https://doi.org/10.1016/S0039-9140(02)00268-0)
- Meharg *et al.* (16 authors) (2013). Variation in rice cadmium related to human exposure. *Environmental Science and Technology* **47**: 5613–5618.
DOI: <https://doi.org/10.1021/es400521h>
- Meharg A.A., Williams N.P., Adomako E.E., Lawgali Y.Y., Deacon C., Villada A., Robert C., Cambell J., Sun G., Y-G., Feldmann J., Raab A., F-J., Rafiqul Islam R., Hossain S., & Yanai J. (2009). Geographic variation of total and inorganic arsenic content of polished white rice. *Environmental Science and Technology* **43**(5): 1612–1617.
- Ministry of Environment (2007). *Government Decree on the Assessment of Soil Contamination and Remediation Needs 214/2007*. Ministry of the Environment, Finland. Available at <http://extwprlegs1.fao.org/docs/pdf/fin113198.pdf>.
- Nishimura M, Konishi S., Matsunaga K., Hata K. & Kosuga T. (1983). Mercury concentration in the ocean. *Journal of the Oceanographical Society of Japan* **39**: 295–300.

- DOI: <https://doi.org/10.1007/BF02071824>
- Norton *et al.* (17 authors) (2014). Lead in rice: Analysis of baseline lead levels in market and field collected rice grains. *Science of the Total Environment* **485**: 428–434.
DOI: <https://doi.org/10.1016/j.scitotenv.2014.03.090>
- Oberoi S., Barchowsky A. & Wu F. (2014). The global burden of disease for skin, lung, and bladder cancer caused by arsenic in food. *Cancer Epidemiol Biomarkers Prevention* **23**(7):1187–1194.
DOI: <https://doi.org/10.1158/1055-9965.EPI-13-1317>
- Premarathna H.M.P.L., Hettiarachchi G.M. & Indraratne S.P. (2011). Trace metal concentration in crops and soils collected from intensively cultivated areas of Sri Lanka. *Pedologist* 230–240.
- Rahman M.A., Hasegawa H., Rahman M.M., Mazid-Mia M.A. & Tasmin A. (2008). Arsenic accumulation in rice (*Oryza sativa* L.): human exposure through food chain. *Ecotoxicology and Environmental Safety* **69**: 317–324.
DOI: <https://doi.org/10.1016/j.ecoenv.2007.01.005>
- Rahman M.A., Hasegawa H., Rahman M.M., Rahman M.A. & Miah M.A. (2007). Accumulation of arsenic in tissues of rice plant (*Oryza sativa* L.) and its distribution in fractions of rice grain. *Chemosphere* **69**(6): 942–948.
DOI: <https://doi.org/10.1016/j.chemosphere.2007.05.044>
- Sanjeevani U.K.P.S., Indraratne S.P., Weerasooriya R., Vitharana U.W.A. & Kumaragamage D. (2017). Identifying the sources and contamination status of potentially toxic trace elements in agricultural soils. *Communications in Soil Science and Plant Analysis* **48**(8): 865–877.
DOI: <https://doi.org/10.1080/00103624.2017.1299168>
- Toth G, Hermann T., Da Silva M.R. & Montanarella L. (2016). Heavy metals in agricultural soils of the European Union with implications for food safety. *Environment International* **88**: 299–309.
DOI: <https://doi.org/10.1016/j.envint.2015.12.017>
- Uraguchi S. & Fujiwara T. (2012). Cadmium transport and tolerance in rice: Perspectives for reducing grain cadmium accumulation. *Rice* **5**: 1–8.
DOI: <https://doi.org/10.1186/1939-8433-5-5>
- Westfall D.G., Mortvedt J.J., Peterson G.A. & Gangloff W.J. (2005). Efficient and environmentally safe use of micronutrients in agriculture. *Communications in Soil Science and Plant Analysis* **36**(1-3): 169–182,
DOI: <https://doi.org/10.1081/CSS-200043024>
- WHO (2016). *Investigation and Evaluation of Chronic Kidney Disease of Uncertain Aetiology in Sri Lanka – Final Report*. Ministry of Health, Colombo Sri Lanka.
- WHO (2017). *Mercury and Health, Fact Sheet*. WHO, Geneva, Switzerland.. Available at <https://www.who.int/news-room/fact-sheets/detail/mercury-and-health>.
- WHO (2018) *Arsenic: Key Facts*. WHO, Geneva, Switzerland. Available at <https://www.who.int/news-room/fact-sheets/detail/arsenic>.
- WHO (2019). *Preventing Disease Through Healthy Environments: Exposure To Cadmium: a Major Public Health Concern*. WHO., Geneva, Switzerland.
- WHO (2021), *Lead Poisoning: Key Facts*, WHO, Geneva, Switzerland. Available at <https://www.who.int/news-room/fact-sheets/detail/lead-poisoning-and-health>.
- Zhao F.J. & Wang P. (2020). Arsenic and cadmium accumulation in rice and mitigation strategies. *Plant Soil* **446**:1–21.

RESEARCH ARTICLE

Machine Learning

Fuzzy based congestion detection and control algorithm for energy efficient wireless sensor network (WSN)

A Ramasamy-Rajeswari^{1*}, S Ganapathy², K Kulothungan³ and A Kannan⁴

¹ Department of Computer Science and Engineering, Rohini College of Engineering and Technology, Kanyakumari, India.

² Centre for Cyber-Physical Systems & School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.

³ Department of Information and Technology, CEG Campus, Anna University, Chennai, India.

⁴ School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India.

Submitted: 02 February 2021; Revised: 27 February 2022; Accepted: 04 March 2022

Abstract: A wireless sensor network (WSN) is built by numerous tiny, resource aware sensor nodes to facilitate and develop various fields such as the military, health care, disaster management etc. One of the major design issues in WSN is congestion detection and control. Data collision, channel contention, and buffer overflow are considered to be a few important factors that may lead to a congestion problem. Congestion causes increase in packet drop rate and thus network efficiency is decreased. Hence, the network performance and quality of service may be degraded. Therefore congestion detection and control are vital and critical challenges to be addressed. In this paper, a novel energy efficient fuzzy based congestion detection and control algorithm (EE-FBCDCA) has been proposed with the objective of addressing the energy aware congestion detection and control mechanism. The proposed work consists of three phases, namely, 1) Fuzzy inference system based congestion detection, 2) Congestion notification phase and 3) Fuzzy inference system based congestion control. In this work, once the congestion is detected and upon receiving the notification message, the intermediate nodes will adjust their respective transmission rate to prevent further drop of packets. Thus, EE-FBCDCA detects and control the congestion more effectively to improve the QoS of the network. The proposed work is implemented and analysed by using MATLAB, and the performance is compared with the existing HTCCFL and FBACC. The simulation results of EE-FBCDCA shows enhanced network lifetime, energy efficiency, and reliability.

Keywords: Congestion detection, control, fuzzy inference system, energy, network lifetime, WSN.

INTRODUCTION

A wireless sensor network (WSN) consists of numerous tiny sensor nodes deployed over a vast geographical area with limited resources such as energy and bandwidth availability (Bhushan & Sahoo, 2010; Banimelhem & Khasawneh, 2012; Antoniou *et al.*, 2013; Chen *et al.*, 2016; Rao & Rama Rao, 2022). In today's era, WSNs have wide applications in various fields, namely healthcare systems, military, multimedia etc. The sensor nodes are capable of collecting, processing and transmitting the data through the intermediate nodes to reach the sink nodes (Farzaneh & Yaghmaee, 2015). In WSN many design issues such as queuing delay, packet drop, packet collision, and congestion need to be addressed. Among all the above stated issues, one of the major problems in WSN is congestion, which needs to be addressed to enhance the network lifetime and performance. Congestion leads to numerous issues especially inefficient energy consumption by nodes, buffer overflow, and high packet loss (Gholipour *et al.*, 2015). Thus, this problem will degrade the network performance and quality of service. The primary working principle of nodes is to sense, gather, process, and transmit the data to the sink nodes for analysis and decision making. In WSN, each node transmits the gathered data to the sink node for processing. Thus, vast amounts of data are forwarded towards the base station at

* Corresponding author (arrajeswari.2015@gmail.com;  <https://orcid.org/0000-0002-9893-1027>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

any given instant of time, leading to network congestion and packet loss. As WSN play a vital role in various real time application domains, congestion will jeopardize the network lifetime and performance that need to be avoided (Jan *et al.*, 2014). Thus, congestion is considered as a very important design challenge. Hence, it becomes essential to implement an efficient and reliable congestion control mechanism to address the problem (Sonmez *et al.*, 2014). Moreover, sensor nodes are power constrained devices and therefore energy is considered to be one of the major performance metrics during congestion detection (Uthra & Raja, 2014). Much research has been done towards the enhancement of energy conservation to prevent packet collision and congestion (Wang *et al.*, 2012). But one of the major drawbacks in all this work is that energy consumption is not equally distributed among all the nodes (Ahmed *et al.*, 2014; Yi & Yang, 2016; Zou & Qian, 2019). Different mechanisms and schemes were utilized by various researchers towards congestion detection and control process. Moreover, many research projects are focused towards fuzzy based congestion control and detection mechanisms (Munir & Man, 2007; Jutan *et al.*, 2012; Rekha & Gomathy, 2012). Therefore, designing a novel energy efficient congestion detection and control process with a better solution is considered as a major research challenge. In this work, a novel fuzzy based energy aware congestion detection and control algorithm for WSN has been proposed. This algorithm yields a better solution for congestion detection and control when compared to other existing algorithms.

The primary focus of the proposed work is to develop an energy aware, high reliable congestion detection and control algorithm for WSN. The proposed algorithm consists of three phases. The first phase is defined as a novel fuzzy based energy aware congestion detection phase with fuzzy input variables, such as packet delivery rate, residual energy level, and queue occupancy, with the aim of performing the task of congestion detection to enhance the energy efficiency of the network.

The next phase, the congestion notification phase, will be triggered to transmit the congestion signal to the nodes whenever congestion occurs in the network. Finally, in the third phase, *i.e.*, the fuzzy based congestion control phase, the data transmission rate of the nodes is controlled to prevent further congestion in the network. Thus, the proposed work ensures energy efficient, high throughput, minimal end-to-end delay, a smaller packet drop rate, and overall enhancement in the network performance.

Much research has been carried out in the field of congestion detection and controls in WSN (Misra *et al.*, 2009; Jenolin *et al.*, 2011; Jan *et al.*, 2014; Mahdizadeh *et al.*, 2014; Ghanavati *et al.*, 2015). Ghaffari (2015) and Rekha Chakravarthi *et al.* (2010) have carried out comparative surveys on the congestion control algorithm. Congestion control (Chen & Zhang, 2006) is defined as a technique of controlling the amount of data flow in the network without any loss of information (Rezaee *et al.*, 2014). It consists of three primary phases namely detection (Chen & Yang, 2006), notification (Wan *et al.*, 2011), and control (Yin, 2009).

Ahmed & Paulus (2017) developed congestion avoidance and mitigation for multipath routing and load balancing in WSN. Cao *et al.* (2012) proposed a novel delay based congestion control for a multipath TCP feedback congestion control protocol. The experimental results of this protocol shows better energy and throughput.

Huang *et al.* (2009) developed an FACC protocol to detect congestion based on the classification of the nodes adjacent to the source and adjacent to the base station. Tao and Yu (2010) presented an enhanced congestion detection and avoidance protocol for detecting congestion by using buffer threshold and buffer weight. This protocol supports a hop-by-hop congestion control scheme. Adjero & Yaghmaee (2008) presented a queue based congestion control protocol with priority support to congestion control by considering the queue length as the index of density degree. Thus, in this protocol the transmission rate is controlled depending upon the priority and density degree.

Basaran *et al.* (2010) developed a control of sensor queue (CONSEQ) protocol with the aim of detecting congestion by considering the queue length and data transmission among the one hop neighbouring nodes. Akyildiz *et al.* (2003) proposed an event-to-sink reliable transport protocol. In this protocol, the packet transmission rate to the sink node is observed and queue overflow is identified. The intermediate nodes will detect the presence of congestion in the network by referring to the congestion bit in the packet.

A fuzzy inference system can be utilized to implement the congestion detection and control scheme (Rezaee & Pasandideh 2017; Sangeetha *et al.*, 2019). Many fuzzy inference system based models have been proposed to address various issues in WSN (Logambigai & Kannan, 2016; Rajeswari *et al.*, 2019).

A fuzzy temporal congestion predication model was proposed to identify the congestion free path. This model consists of two phases, namely, congestion prediction and localization phases (Mukherjee & Dasgupta, 2013). Hatamian *et al.* (2016) developed a fuzzy rate controller based congestion aware routing by assigning priority to data packets. Ghanavati *et al.* (2013) presented a type 2 fuzzy logic algorithm to measure and control the congestion level, to enhance the QoS in WSN. This algorithm improves the network lifetime and decreases the packet drop rate.

Ghaffari (2014) proposed a star algorithm based routing mechanism to ensure improved link quality and reduced hop count. But in this method end-to-end delay is more and thus the complexity is greater. Sayyad and Choudhari (2014) proposed a Hierarchical Tree based congestion control using fuzzy logic (HTCCFL) using the following metrics, namely, packet service ratio, number of competitors, and buffer occupancy. The above algorithm gives increased energy efficiency by selecting an alternate path with less congestion for data transmission.

Hatamian and Barati (2015) developed a priority based congestion control schema using fuzzy rules in WSN. In this method, buffer occupancy and congestion level are considered as input parameter for a fuzzy rate controller to monitor the packet transmission rate. Damaso *et al.* (2014) proposed a reliable framework with energy as a primary constraint. Blocks and models were used to express the sensor reliability elements using the energy factor. Mostafaei (2019) proposed distributed learning automation (DLA) to identify the small group of nodes to enhance end to end reliability.

Ullah *et al.* (2017) proposed an energy and congestion aware routing for an AMI network in a smart city. In this method, residual energy and queue occupancy are considered as major factors monitoring and controlling the congestion level. Song *et al.* (2019) developed a unified modelling framework known as GLS for space terrestrial networks. Jaiswal and Yadav (2013) developed a fuzzy based adaptive congestion control (FBACC) that enable congestion estimation by considering the factors buffer occupancy and traffic rate. FBACC enhances the energy efficiency and reduces the packet drop rate.

Gharajeh and Khanmohammadi (2016) developed DFRTTP based upon the distance and numbers of neighbouring nodes, which were considered as the input parameters. DFRTTP yields better network lifetime and

packet drop rate. Chen *et al.* (2015) developed a RLAN to detect the node position under varying magnitude across various directions. Khan *et al.* (2020) proposed and implemented the ant colony optimization Technique and repetitive route configuration with reactive routing protocol for obstruction of Black Hole Attack in mobile *ad hoc* networks. Perumal and Subramaniyan (2020) developed the statistical Markov model based natural inspired glow worm multi-objective optimization (SMM-NIGMO) technique for Energy efficient data delivery in MANET.

In spite of the availability of much related work in the literature survey, an energy efficient congestion detection and control in WSN still remains as a major design goal, following the metrics buffer size, service time, and channel load. These studies do not focus much on energy efficiency based congestion detection and control process. Hence there is a need to design an energy efficient congestion detection and control algorithm for WSN for enhancing the throughput, reducing delay, and increasing energy efficiency and reliability. This novel challenge is addressed in this work. Moreover, most of the related work focuses on Energy Efficient Fuzzy Based Congestion Detection and Control Algorithm, and This work is proposed with the objective of improving the network performance and lifetime.

The rest of this paper is structured as follows. Section 2 describes in detail the Proposed Methodology. Section 3 illustrates and discusses the simulation based results. Section 4 finally concludes the paper.

MATERIALS AND METHODS

Energy efficient fuzzy based congestion detection and control algorithm (EE- FBCDCA)– energy model

The energy model considered in this proposed work, EE-FBCDCA, is similar to the work developed by Heinzelman and Balakrishnan (2002). According to the first order radio model, the amount of energy consumed during the transmission of a packet with 1 bit over d distance is given by the Equation (1). The E_{elec} indicates the amount of energy consumed per bit to run the transmitter or receiver circuitry. E_{fs} and E_{mp} are the amplifier energies in free space and multipath, respectively.

$$E_{Tran}(L) = \begin{cases} LE_{elec} + L\epsilon_{fs}d^2 f & \text{ord} < d_0 \\ LE_{elec} + L\epsilon_{mp}d^4 f & \text{ord} > d_0 \end{cases} \quad \dots(1)$$

$$\text{where } d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}$$

The energy $E_{RX}(l)$, which indicates the amount of energy consumed in receiving a packet with 1 bit, is given by Equation (2).

$$E_{RX}(l) = E_{elec} \dots(2)$$

EE-FBCDCA

An energy aware congestion control mechanism is considered to be a major design challenge for WSN.

Hence in this work, a novel EE- FBCDCA based on a fuzzy inference system has been proposed with the objective of detecting and controlling the congestion. The residual energy level and reliability are considered as major aspects for measuring the network performance. In the proposed algorithm the network congestion is detected by considering the data flow from sources to base station. Moreover, buffer capacity is considered to be an important metric for the process of congestion monitoring and detection. The proposed work consists of the following three phases, namely i) Congestion Detection Phase, ii) Congestion Notification Phase, and iii) Congestion Control Phase. Figure 1 Shows the architecture of the proposed EE- FBCDCA.

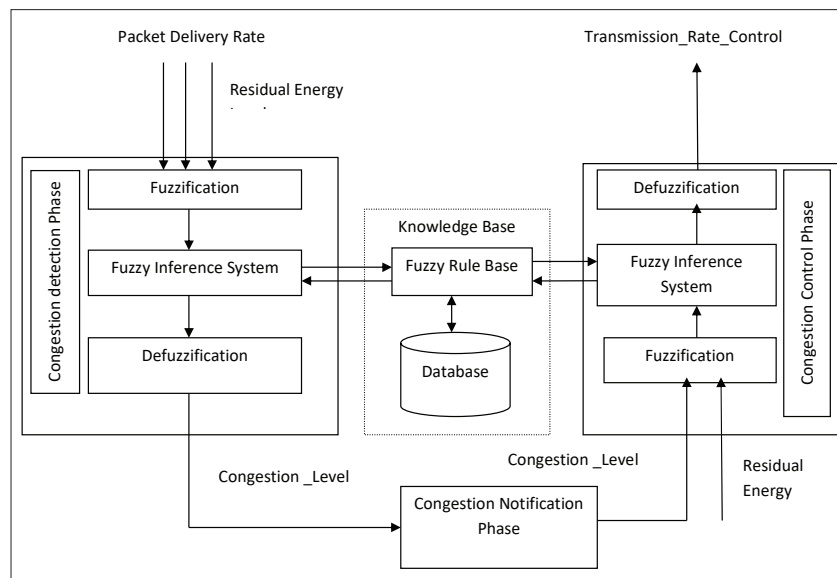


Figure 1: Architecture of the proposed EE- FBCDCA

Design of energy efficient fuzzy inference system for congestion control and detection Phases.

In this proposed work, the FIS is defined under the following two scenarios: firstly, for congestion detection phase and secondly, during the process of congestion control.

The fuzzy inference system (FIS) employed in this work consists of the following 4 main components:

- (i) Fuzzifier
- (ii) Fuzzy inference engine (FIE)
- (iii) Fuzzy rule base, and
- (iv) Defuzzifier.

Congestion detection and control are the most important design challenges in the WSN environment. Thus, to address these challenges an EE- FBCDCA is developed in this work. The major goal of the proposed work is address detection as well as controlling congestion in the network, in order to improve the network reliability and lifetime. Thus, this section elaborates in detail the proposed work, EE- FBCDCA. The main flow of the proposed EE- FBCDCA is shown below in Algorithm 1.

The EE- FBCDCA has been proposed with the following two major objectives.

1. A novel energy aware fuzzy based congestion control scheme is proposed to prevent the congestion.

2. Energy conservation is improved by congestion control using the fuzzy rules. Hence, the network lifetime is enhanced.

The proposed EE- FBCDCA consists of the following three phases:

1. Fuzzy inference system based congestion detection phase
2. Congestion notification phase
3. Fuzzy inference system based congestion control phase

Algorithm 1 - EE- FBCDCA

```

Input : Senor Nodes
Output : Congestion Free Network Environment
For each Node n do
    Random_number = rand(0,1)
    If ( $q_{len} < TH$ ) then
        Congestion_Level = Fuzzy_logic1
        (REL,PDR,Buffer_Capacity)
        If ( $Congestion\_Level < Random\_Number$ )
            Add Packet 'P' to the node's buffer
        Else
            Send Congestion_Warning_Signal (Node_ID,REL) to its Neighbour Nodes
        End If
        For each Neighbour (M) on receiving the Congestion_Warning_Signal
            Transmission_Rate_Control = Fuzzy_logic_2
            (Congestion_Level,REL)
        End For
    End if
End For
    
```

The following subsection describes each of the phases in detail.

Fuzzy inference system based congestion detection phase

In the proposed work, the fuzzy inference system is employed to detect the congestion level of the network. The main objective of this proposed work is to improve the network lifetime and performance by controlling and managing the transmission rate of the nodes, so that the congestion level will be reduced. The primary role of this phase is to detect the congestion level of every node in the network. Thus, in this phase, queue length is measured periodically to detect the traffic level in the network. Moreover, in this proposed work the fuzzy logic

inference system is used for measuring the congestion level. The three fuzzy input variables used are Buffer Capacity, Residual Energy, and Packet Delivery Rate.

Description of the fuzzy variables

Buffer capacity

The buffer capacity of any node, say (n), is defined as the maximum amount of data packets that a node can store before processing. Thus, the buffer capacity parameter plays a vital role during the congestion detection scheme. The buffer capacity is measured by using Equation (3), given below:

$$Buffer_Capacity = \frac{Buffer_Size - Buffer_Packet}{Buffer_Size} \dots(3)$$

where $Buffer_{Size}$ indicates the buffer size of a node, and $Buffer_{Packet}$ indicates the number of packets available in the buffer at given instant 't'.

Hence, if the $Buffer_{Packet}$ is zero, then it indicates that the node's buffer is empty, whereas if $Buffer_{Size}$ is equal to $Buffer_{Packet}$ it indicates that the buffer is full and thus all the arriving packets are dropped, which may lead to congestion in the network .

The threshold Buffer length (B_{TH}) is calculated by using Equation (4), given below:

$$B_{TH} = \frac{Packet_T^N}{Buffer_{Size}} \dots(4)$$

where $Packet_T^N$ is the amount of data packets arrived at node 'N' at given time 't'.

Packet delivery rate

The packet delivery rate at any node 'n' is estimated with respect to the packet service time at node 'n'. P_{ST} is defined as the time taken by the node for successfully transmission of a single packet over the communication path.

Thus, the packet average service time can be measured by considering an exponential weighted sum formula as shown below, Equation (5):

$$P_{ST} = (1 - \alpha)P_{ST} + \alpha P_{ST} \dots(5)$$

where α is constant such that $0 \leq \alpha \leq 1$

Thus, by using the measured P_{ST} , the node rate can be estimated by using Equation (6), given below:

$$PDR = \frac{1}{P_{ST}} \quad \dots(6)$$

Residual energy level (REL)

The REL of a node is calculated by using Equation (7), given below:

$$REL(N) < E_{thres} \quad \dots(7)$$

where REL (N) is the Residual Energy Level of the node N, and E_{thres} is the threshold energy. The threshold energy (E_{thres}) is calculated by using Equation (8).

$$E_{thres} = \frac{\sum_{i=1}^N REL_i}{N} \quad \dots(8)$$

where REL is the Residual Energy Level, and N is the Number of nodes in the network.

Congestion in a network leads to packet drop or collision that may increase the retransmission of same set of data packets by node. Hence, the node’s energy level drastically decreases. Thus, congestion acts as one of the primary source of decrease in network lifetime as well as energy efficiency. Table 1 shows the fuzzy rules for estimating the congestion level in the system.

The following parameters such as Buffer Length, Residual Energy Level, and Packet Delivery Rate are considered as fuzzy inputs whereas the Congestion Level act as fuzzy output. Thus, the fuzzy inputs variables and its linguistic variables used for Congestion Level estimation are shown below.

Buffer Length has the following linguistic variables, namely Low, Medium, and High. Residual Energy Level holds linguistics variables Low, Medium, and High, whereas the Packet Delivery Rate has the following linguistic variables, namely High_Flow, Medium_Flow, and Low_Flow. The output variable Congestion_Level has the following five linguistics variables: Low, Very Low, High, Very High and Medium.

Table 1: Fuzzy Rules for measuring the Congestion_Level

Set of “IF- THEN” Rules
Rule_I: IF Buffer_Length is High and Residual_Energy_Level is High and Packet_Delivery_Rate is High-Flow THEN Congestion_Level is High.
Rule_II :IF Buffer_Length is Low and Residual_Energy_Level is Low and Packet_Delivery_Rate is High_Flow THEN Congestion_Level is Low
Rule_III: IF Buffer_Length is Low and Residual_Energy_Level is Low and Packet_Delivery_Rate is High_Flow THEN Congestion_Level is Low.
Rule_IV: IF Buffer_Length is High and Residual_Energy_Level is Low and Packet_Delivery_Rate is Low_Flow THEN Congestion_Level is Very High.
Rule_V: IF Buffer_Length is Low and Residual_Energy_Level is High and Packet_Delivery_Rate is Low_Flow THEN Congestion_Level is Medium.

Congestion notification phase

Once the congestion is detected, in order to prevent the worst case of more packets begin transmitted over the network all other intermediate nodes should be notified about the congestion issue. Thus in this proposed work, through the congestion notification phase, the nodes are given a warning signal about the congestion in the network. In this phase, the congestion notification mechanism based on the ICN scheme of Yi et al.(2008) is employed. Thus, the nodes will transmit the piggyback packet holding the congestion warning signal. The main advantage of this phase is to prevent a network form getting more congested with additional control packets

and in turn to enhance the network lifetime and energy efficiency.

Fuzzy inference system based congestion control phase

In this phase, we have developed a fuzzy based rate control scheme that will perform congestion control and prevent its subsequent issues in network such as delay in transmission, inefficient energy consumption, reduced network lifetime, etc. Moreover, the primary objective of this phase is to control the congestion by initiating the intermediate nodes to reduce its transmission rate once after receiving the congestion notification signal. Hence,

this may leads to prevent further more congestion in network and enhance the network performance. Once again Fuzzy Inference System is utilized in this phases. Congestion Control phases is carried out by considering the following two metrics namely Congestion_Level and Residual Energy. Hence each node evaluates the Transmission_Rate_Control by using the fuzzy logic to determine the nodes probability to control the data transmission rate. Table 2 shows the fuzzy rules for evaluating the parameter Transmission_Rate_Control.

The two fuzzy input variables and its corresponding linguistics variables for Transmission_Rate_Control are shown below.

Congestion_Level has the following linguistic variables, namely Low, Medium and High. The Residual_Energy_Level holds linguistic variables Low, Medium, and High. The output variable Transmission_Rate_Control has the linguistics variables Increase_Low, Increase_High and Decrease.

Table 2: Fuzzy Rules to evaluate the Transmission_Rate_Control

Set of "IF- THEN" Rules
Rule_I: IF Congestion_Level is Low and Residual_Energy_Level is Low and THEN Transmission_Rate_Control is Decrease.
Rule_II: IF Congestion_Level is Low and Residual_Energy_Level is Medium and THEN Transmission_Rate_Control is Increase.
Rule_III: IF Congestion_Level is Medium and Residual_Energy_Level is High and THEN Transmission_Rate_Control is Increase.

In our fuzzy inference system, the Congestion_Level and Transmission_Rate_Control are determined using both the triangular and trapezoidal membership functions. The triangular membership functions are used to represent the intermediate variables whereas for boundary variables the trapezoidal membership functions is used.

The triangular membership functions and trapezoidal membership functions (Rezaee *et al.*, 2014) are calculated using Equation (9) and Equation (10), respectively.

$$A = \left\{ \begin{array}{ll} 0, x \leq a1 \\ \frac{x - a1}{b1 - a1}, a1 \leq x \leq b1 \\ \frac{c1 - x}{c1 - b1}, b1 \leq x \leq c1 \\ 0, x \leq a2 \end{array} \right\} \dots(9)$$

$$A = \left\{ \begin{array}{ll} 0, x \leq a2 \\ \frac{x - a2}{b2 - a2}, a2 \leq x \leq b2 \\ \frac{d2 - x}{d2 - c2}, c2 \leq x \leq d2 \\ 0, d2 \leq x \end{array} \right\} \dots(10)$$

The finale step in FIS is the defuzzification. The primary objective of this step is to obtain the crisp output equivalent to the fuzzy output. Moreover, in this work we have used the Center of Area COA method for defuzzification. This is measured by using the Equation (11).

$$COA = \frac{\int \mu_A(x) \cdot x dx}{\int \mu_A(x) dx} \dots(11)$$

where $\mu_A(x)$ denotes the fuzzy values for the membership functions.

Illustration

In this section we have illustrated the fuzzy based congestion_level estimation mechanism in detail, with a suitable example.

For example, assume that the Packet_Delivery_Rate is 50, Buffer Length is 0.35 and Residual Energy Level (REL) is 0.9. Table 3 shows the fuzzy variables and their corresponding values. Table 4 shows the corresponding minimum and maximum values for the fuzzy variables.

The input value Z = 50 indicates the packet delivery rate of the node. Thus by setting the trapezoidal member function to low, we get

$$\begin{aligned} f(z; k, l, m, n) &= \max\left(\min\left(\frac{z - k}{l - k}, 1, \frac{n - z}{n - m}\right), 0\right) \\ f(162; 70, 135, 162, 165) &= \max\left(\min\left(\frac{50 - 0}{1 - 0}, 1, \frac{65 - 50}{65 - 40}\right), 0\right) \\ &= \max(\min(50, 1, 0.6), 0) \\ &= \max(0.6, 0) \\ &= 0.6 \end{aligned}$$

Similarly, by setting the triangular and trapezoidal member functions for buffer length to low and medium, respectively, we get,

Z = 0.35 (Low)

$$f(z; k, l, m, n) = \max\left(\min\left(\frac{z - k}{l - k}, 1, \frac{n - z}{n - m}\right), 0\right)$$

$$\begin{aligned} f(162; 70, 135, 162, 165) &= \max\left(\min\left(\frac{0.35 - 0}{1 - 0}, 1, \frac{0.40 - 0.35}{0.40 - 0.30}\right), 0\right) \\ &= \max(\min(0.35, 1, 0.5), 0) \\ &= \max(0.35, 0) \\ &= 0.35 \end{aligned}$$

Z = 0.35 (Medium)

$$f(z; k, l, m) = \max\left(\min\left(\frac{z - k}{l - k}, \frac{m - z}{m - l}\right), 0\right)$$

$$\begin{aligned} f(z; k, l, m) &= \max\left(\min\left(\frac{0.35 - 0.25}{0.45 - 0.25}, \frac{0.70 - 0.35}{0.70 - 0.45}\right), 0\right) \\ &= \max(\min(0.5, 1.4), 0) \\ &= \max(0.5, 0) \\ &= 0.5 \end{aligned}$$

Similarly, by setting the triangular and trapezoidal member functions for residual energy level to medium and high, respectively, we get,

Z = 0.9 (Medium)

$$f(z; k, l, m) = \max\left(\min\left(\frac{z - k}{l - k}, \frac{m - z}{m - l}\right), 0\right)$$

$$\begin{aligned} f(z; k, l, m) &= \max\left(\min\left(\frac{0.9 - 0.1}{0.4 - 0.1}, \frac{0.9 - 0.8}{0.9 - 0.4}\right), 0\right) \\ &= \max(\min(2.66, 0.2), 0) \\ &= \max(0.2, 0) \\ &= 0.2 \end{aligned}$$

Z = 0.9 (High)

$$f(z; k, l, m, n) = \max\left(\min\left(\frac{z - k}{l - k}, 1, \frac{n - z}{n - m}\right), 0\right)$$

$$\begin{aligned} f(z; k, l, m, n) &= \max\left(\min\left(\frac{0.9 - 0.4}{0.8 - 0.4}, 1, \frac{1.2 - 0.9}{1.2 - 1}\right), 0\right) \\ &= \max(\min(1.25, 1, 1.5), 0) \\ &= \max(1, 0) \\ &= 1 \end{aligned}$$

Applying the obtained values in fuzzy rules 1 to 4,

Rule 1: $\min(0.6, 0.35, 0.2) = 0.2$

Rule 2: $\min(0.6, 0.35, 1) = 0.35$

Rule 3: $\min(0.6, 0.5, 0.2) = 0.2$

Rule 4: $\min(0.6, 0.5, 1) = 0.5$

The maximum value obtained from Rule 1 to Rule 4 is 0.5. Rule 4, corresponding to the maximum value, represents the medium probability of congestion_level in the node. The crisp value ranges from 0.25 to 0.7. During the defuzzification process, the corresponding values are substituted to obtain the value of the congestion_level as 0.61.

Table 3: Fuzzy variables and their input range

Fuzzy Input	Fuzzy variable	Input range
Packet_Delivery_Rate	Low	0-60
	Medium	55-170
	High	150-200
Buffer Length	Low	0-0.40
	Medium	0.25-0.70
	High	0.60-1
Energy	Low	0.0 – 0.5
	Medium	0.1 – 0.8
	High	0.5 – 1
Fuzzy Outputs	Fuzzy variable	Input Range
Congestion_Level	Low	0.01-0.05
	Medium	0.25-0.7
	High	0.5-1
Transmission_Rate_Control	Increase_Low	0.01-0.06
	Increase_High	0.35-0.7
	Decrease	0.65-1

Table 4: Minimum and maximum value for fuzzy variables

Fuzzy variables	Minimum crisp value	Maximum crisp value
Packet_Delivery_Rate	0	200
Buffer Length	0	1
Energy	0	1
Congestion_Level	0.01	1
Transmission_Rate_Control	0	1

RESULTS AND DISCUSSION

The proposed work has been tested using MATLAB software because the MATLAB fuzzy tool box consists of all the fuzzy membership functions. Hence, MATLAB is the most efficient and convenient software to carry out the experiments of the proposed work EE- FBCDCA. The experiment were carried out with different nodes, starting from 100 nodes up to 200, deployed over network area of (500 x 500) m². The simulation parameters are listed in Table 5.

Table 5: Simulation parameters

Parameters	Value
Simulation Area	500 m x 500 m
Number of Nodes	200 Nodes
Initial Node Energy	0.5 J
Packet Size	512 Bytes

Packet delivery rate

Figure 2 shows the packet delivery rate of proposed algorithm, and it is compared with other existing algorithms. From Figure 2, it is observed that the packet delivery rate of EE- FBCDCA is greater when compared to other existing work. This is because, in the proposed work, the congestion control phase is initiated to control the packet transmission rate of the intermediate nodes, to control the congestion level in the network. Thus, in EE- FBCDCA congestion control phases play a vital role to control the packet loss ratio and hence the network efficiency is enhanced.

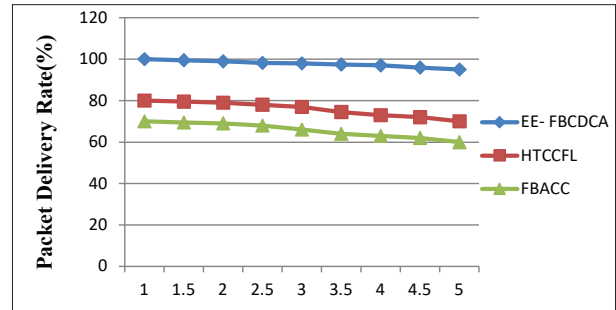


Figure 2: Packet delivery rate

Network lifetime

Figure 3 illustrates the network lifetime comparative analysis of the EE- FBCDCA and other existing algorithms. From Figure 3, it is observed that the proposed work shows better network life than other algorithms. This is because, in the proposed work, once congestion is detected, the packet forwarding rate of the nodes in the network is reduced. This may prevent furthermore congestion and loss of data in the network. Thus, EE- FBCDCA enhances the network lifetime and improves the performance of the network.

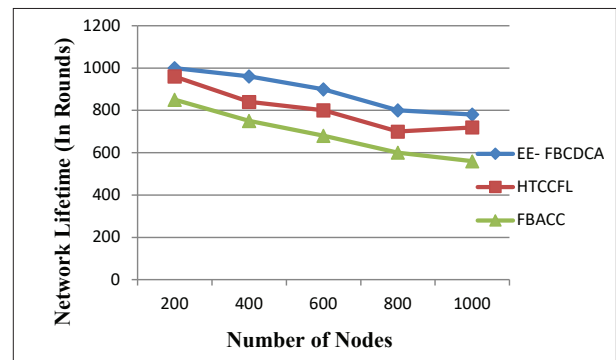


Figure 3: Network lifetime

Reliability

Figure 4 depicts the reliability comparison between the EE-FBCDCA and other existing algorithms. Reliability is the network parameter that estimates the network lifetime for a given period of time, say 't'. From

Figure 4 it is observed that the proposed algorithm gives a high reliability rate when compared with other existing algorithms. This is because in EE- FBCDCA, further congestion will be avoided by activating the congestion notification phase and hence the packet drop in the network will be prevented. Therefore the reliability rate is improved in the proposed work and network performance is enhanced.

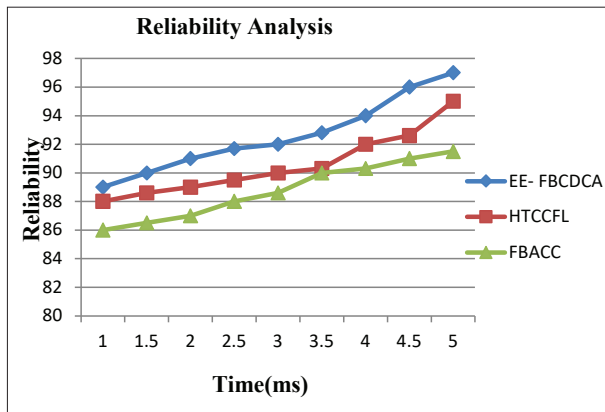


Figure 4: Reliability

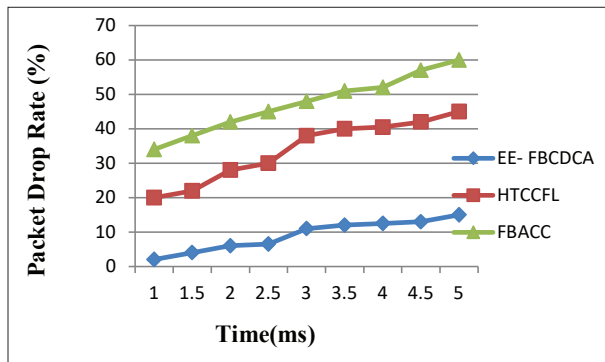


Figure 5: Packet drop rate

Packet drop rate

Figure 5 illustrate the comparative analysis of the packet drop rate of the EE-FBCDCA and other existing algorithms. The packet drop rate is measured by considering the formula given below, Equation (12).

$$PacketDropRate = \frac{P_{\alpha}}{P_{\alpha} + P_{\beta}} \dots(12)$$

Where P_{α} represents the number of packets dropped by a node

P_{β} represents the number of packets received by a node

From Figure 5 it is observed that the packet drop probability of the proposed work is less than the that of other existing algorithms. This is because in the proposed algorithm, both the fuzzy based congestion detection and control phases ensure reduced packet loss rate. Thus, the proposed work yields improved network lifetime with decreased packet drop rate.

Delay

Figure 6 shows the comparative analysis of delay of the EE-FBCDCA with the existing algorithms. End to end delay is defined as the average time taken by the packet to reach the destination node successfully. Thus, minimal delay indicates a better performance of the network. From Figure 6 it is observed that the delay of the proposed work is less than that of the other existing algorithms. This is because the EE-FBCDCA algorithm contains a very efficient congestion control process that may lead to reduced delay in the network.

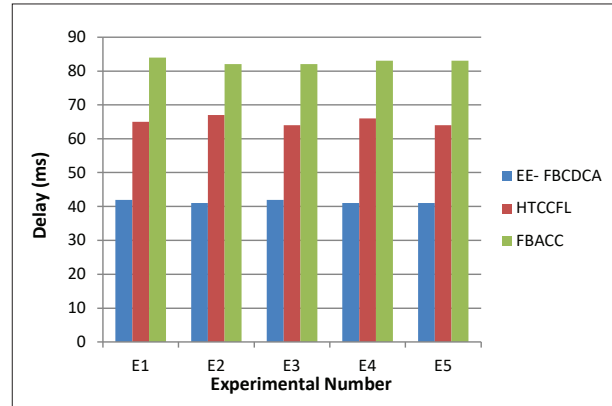


Figure 6: Delay

Energy consumption

Figure 7 depicts the average energy consumption per bit comparative analysis of the EE- FBCDCA and other existing algorithms. The average energy consumption per bit is defined to be the energy required to transmit a data packet successfully from the source to the destination.

From Figure 7, it is clear that the proposed work consumes less energy than existing algorithms. This is because the EE-FBCDCA uses fuzzy based congestion control phases that reduce the amount of data transmitted by the intermediate nodes. Since the congestion is controlled to greater extent that may lead to reduce the data drop rate and therefore energy conservation is improved.

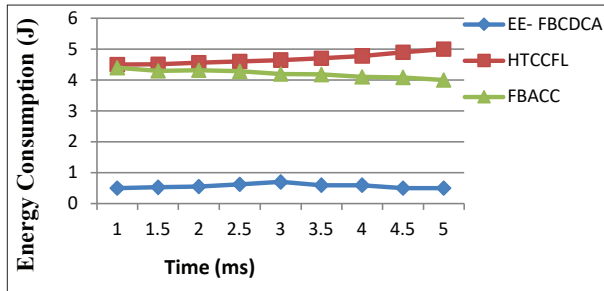


Figure 7: Energy consumption

Throughput

Figure 8 shows the throughput analysis of the proposed work and other existing algorithms. The experiments are conducted for a varying number of nodes, from 200 to 1000. From Figure 8 it is observed that network throughput increases continuously with an increase in the number of nodes. The throughput in the proposed work is improved by applying a fuzzy based congestion detection and control phase that reduces the packet loss rate. Hence, in EE-FBCDCA the number of data packets that are transmitted successfully is more, and in turn network throughput is improved to a greater extent.

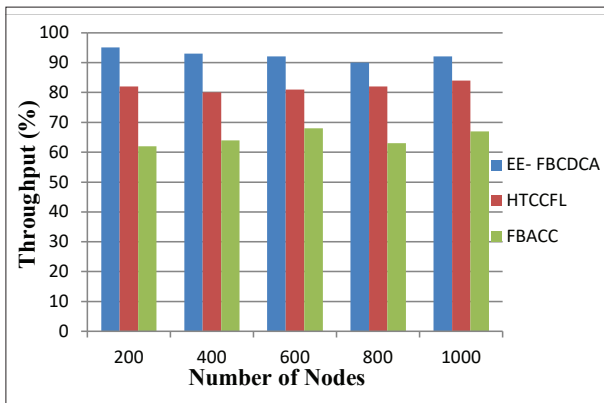


Figure 8: Throughput

Residual energy

Figure 9 depicts the residual energy analysis of the EE-FBCDCA algorithm with other existing algorithms. From Figure 9 it is clear that the proposed work gives more residual energy than other algorithms. This is due to the fuzzy based congestion detection process utilized in the proposed work. Hence the packet drop rate is reduced, and so, more energy is conserved. Therefore, the proposed work gives high residual energy.

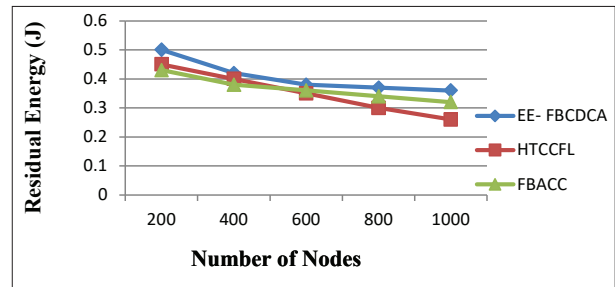


Figure 9: Residual energy

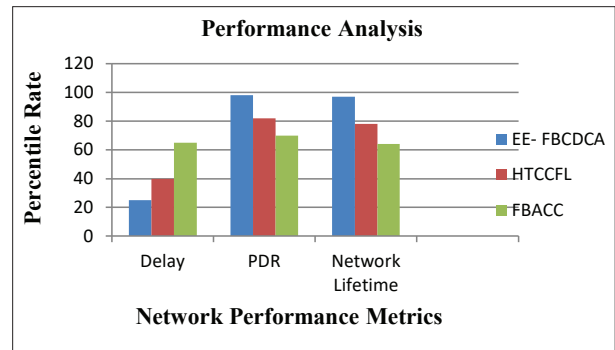


Figure 10: Network performance analysis

Network performance analysis

Figure 10 depicts the network performance analysis of the EE-FBCDCA algorithm and other two existing algorithms. From Figure 10 it is clear that the EE-FBCDCA algorithm gives better network performance in terms of the metrics Packet Delivery Rate (PDR), Delay, and Network Lifetime.

CONCLUSION

In this paper, a novel energy efficient fuzzy based congestion detection and control algorithm has been

proposed with the objective of detecting and controlling congestion in WSN for enhancing the energy efficiency and lifetime of the network. Moreover, fuzzy rules are used in the process of both detection and control of congestion. Through simulation results using MATLAB, the proposed algorithm is evaluated and it has been proved that EE-FBCDCA shows better results when compared to the existing traditional processes, such as HTCCFL and FBACC, in terms of network lifetime, stability and overall performance. Thus, from the simulation results it is observed that the proposed process, EE-FBCDCA, outperforms the other existing processes in improving the lifetime of the network by 35%. Similarly, EE-FBCDCA shows 30% - 33%, and 35% - 38% better throughput in comparison to HTCCFL and FBACC, respectively. Moreover, from the proposed work, it is observed that above 78% of the residual energy in the network is sustained, which enables the transmission of data with reduced delay. Finally, from the experiments conducted it has been proved the proposed algorithm, EE-FBCDCA, gives better congestion detection and control and reduced end to end delay, as well as increased throughput, energy efficiency, and reliability.

REFERENCES

- Abuarqoub A., Hammoudeh M., Adebisi B., Jabbar S., Bounceur A. & Al-Bashar H. (2017). Dynamic clustering and management of mobile wireless sensor networks. *Computer Networks* **117**: 62–75.
DOI: <https://doi.org/10.1016/j.comnet.2017.02.001>
- Adjeroh D. & Yaghmae H. (2008). A new priority based congestion control protocol for wireless multimedia sensor networks. In: *International Symposium on A world of Wireless, Mobile and Multimedia Networks*, 23-26 June. Newport beach, CA, USA, pp1–8.
- Ahmad A., Javaid N., Khan Z.A., Qasim U. & Alghamdi T.A. (2014). (ACH)2: routing scheme to maximize lifetime and throughput of wireless sensor networks. *IEEE Sensors Journal* **14**(10): 3516–3532.
DOI: <https://doi.org/10.1109/JSEN.2014.2328613>
- Ahmed A.M. & Paulus R. (2017). Congestion detection technique for multipath routing and load balancing in WSN. *Wireless Networks* **23**(3): 881–888.
DOI: <https://doi.org/10.1007/s11276-015-1151-5>
- Akyildiz I.F., Akan O.B. & Sankarasubramaniam Y. (2003). ESRT: Event-to-sink reliable transport in wireless sensor networks. In *Proceedings of ACM Mobihoc'03*, 1–3 June. Annapolis, Maryland. pp.177-188.
- Antoniou P., Pitsillides A., Blackwell T., Engelbrecht A. & Michael L. (2013) Congestion control in wireless sensor networks based on bird flocking behavior. *Computer Networks* **57**(5): 1167–1191.
DOI: <https://doi.org/10.1016/j.comnet.2012.12.008>
- Banimelhem O. & Khasawneh S. (2012). GMCAR: grid-based multipath with congestion avoidance routing protocol in wireless sensor networks. *Ad Hoc Networks* **10**(7): 1346–1361.
DOI: <https://doi.org/10.1016/j.adhoc.2012.03.015>
- Basaran C., Kang K.D. & Mehmet H.S. (2010). Hop-by-hop congestion control and load balancing in wireless sensor networks. In: *IEEE 35th Conference on Local Computer Networks (LCN)*, 11-14 October. Colorado, USA. pp. 448-455.
- Bhushan B. & Sahoo G. (2017). A comprehensive survey of secure and energy efficient routing protocols and data collection approaches in wireless sensor networks. In: *International Conference on Signal Processing and Communication (ICSPC)*, 28-29 July. Coimbatore, India. pp. 294–299
- Cao Y., Xu M. & Fu X. (2012). Delay-based congestion control for multipath TCP. In: *20th IEEE International Conference on Network Protocols (ICNP)*, 30 October-2 November Austin, TX, USA. pp 1–10.
- Chen L., Pang L., Zhou B., Zhang J., Liu Z., Luo Q. & Sun L. (2015). RLAN: range-free localisation based on anisotropy of nodes for WSNs. *Wireless Communications* **51**(24): 2066–2068.
DOI: <https://doi.org/10.1049/el.2015.2554>
- Chen S. & Yang N. (2006). Congestion avoidance based on lightweight buffer management in sensor networks, *IEEE Transactions on Parallel and Distributed Systems* **17**(9): 934–946.
DOI: <https://doi.org/10.1109/TPDS.2006.115>.
- Chen S. & Zhang Z. (2006). Localized algorithm for aggregate fairness in wireless sensor networks. *Proceedings of the 12th annual international conference on Mobile computing and networking*, 23-26 September. Los Angeles, CA, USA. pp. 274–285.
- Chen W., Niu Y. & Zou Y. (2016). Congestion control and energy-balanced scheme based on the hierarchy for WSNs. *IET Wireless Sensor Systems* **7**(1): 1–8.
DOI: <https://doi.org/10.1049/iet-wss.2015.0097>
- Damaso A., Rosa N. & Maciel P. (2014). Reliability of wireless sensor networks. *Sensors* **14**: 15760–15785.
DOI: <https://doi.org/10.3390/s140915760>
- Farzaneh N. & Yaghmae M.H. (2015). An adaptive competitive resource control protocol for alleviating congestion in wireless sensor networks: an evolutionary game theory approach. *Wireless Personal Communications* **82**(1): 123–142.
DOI: <https://doi.org/10.1007/s11277-014-2198-9>
- Ghaffari A. (2014). An energy efficient routing protocol for wireless sensor networks using A-star algorithm. *Journal of Applied Research and Technology* **12**(4): 815–822. DOI: [https://doi.org/10.1016/S1665-6423\(14\)70097-5](https://doi.org/10.1016/S1665-6423(14)70097-5)
- Ghaffari A. (2015). Congestion control mechanisms in wireless sensor networks: A survey *Journal of Network and Computer Application* **52**: 101–115.
DOI: <https://doi.org/10.1016/j.jnca.2015.03.002>

- Ghanavati S., Abawaji J. & Izadi D. (2015). A congestion control scheme based on fuzzy logic in wireless body area networks. In: *Network Computing and Applications (NCA), IEEE 14th international symposium*, 28-30 September. Cambridge, MA, USA. pp. 235–242.
- Ghanavati S., Abawaji J. & Izadi D. (2013). A fuzzy technique to control congestion in WSN. *Proceedings of International Joint Conference on Neural Networks*, 4–9 August. Dallas, Texas, USA. pp. 1806-1810.
- Gharajeh M.S. & Khanmohammadi S. (2016). DFRTF: dynamic 3D fuzzy routing based on traffic probability in wireless sensor networks. *IET Wireless Sensor Systems* **6**(6): 211–219.
DOI: <https://doi.org/10.1049/iet-wss.2015.0008>
- Gholipour M., Haghghat A.T. & Meybodi M.R. (2015). Hop-by-hop traffic-aware routing to congestion control in wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking* **1**(15): 1-13.
DOI: <https://doi.org/10.1186/s13638-015-0241-5>
- Hatamian M., Bardmily M.A., Asadboland M., Hatamian M. & Barati H. (2016). Congestion-aware routing and fuzzy-based rate controller for wireless sensor networks. *Radio Engineering* **25**(1): 114–123.
DOI: <https://doi.org/10.13164/re.2016.0114>
- Hatamian M. & Barati H. (2015). Priority-based congestion control mechanism for wireless sensor networks using fuzzy logic. In: *6th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 13-15 July. Denton, TX, USA. pp.1-5.
- Heinzelman W.B., Chandrakasan A.P. & Balakrishnan H. (2002). An application specific protocol architecture for wireless sensor network. *IEEE Transactions on Wireless Communications* **1**(4): 660–670.
DOI: <https://doi.org/10.1109/TWC.2002.804190>
- Huang R., Fang Y., Li S., Yin X. & Zhou X. (2009). A fairness-aware congestion control scheme in wireless sensor networks. *IEEE Transactions on Vehicular Technology* **58**(9): 5225–5234.
DOI: <https://doi.org/10.1109/TVT.2009.2027022>
- Jaiswal S. & Yadav A. (2013). Fuzzy based adaptive congestion control in wireless sensor networks. In: *Proceedings of 6th IEEE International Conference on Contemporary Computing (IC3)*, pp. 433–438.
DOI: <https://doi.org/10.1109/IC3.2013.6612234>
- Jan M.A., Nanda P., He X. & Liu R.P. (2014). PASCCC: priority-based application-specific congestion control clustering protocol. *Computer Networks* **74**: 92–102.
DOI: <https://doi.org/10.1007/s11036-018-1018-y>
- Jan M.A. (2018). A comprehensive analysis of congestion control protocols in wireless sensor networks *Mobile Networks and Applications* **23**(3): 456–468.
DOI: <https://doi.org/10.1007/s11036-018-1018-y>
- Jenolin F.D.F., Kavitha V. & Muthuselvi M. (2011). A survey on congestion control techniques in wireless sensor networks. In: *International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT)*, 23-24 March, Nagercoil, India. pp.1146–1149.
DOI: <https://doi.org/10.1109/ICETECT.2011.5760292>
- Jutan W., Bing F. & Yi S. (2012). A congestion control scheme based on fuzzy logic for wireless sensor networks. In: *9th IEEE International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 29-31 May. Sichuan, China, pp. 501–504.
DOI: <https://doi.org/10.1109/FSKD.2012.6234353>
- Khan D.M., Aslam T., Akhtar N., Qadri S., Khan N.A., Rabbani I.M. & Aslam M. (2020). Black hole attack prevention in mobile ad-hoc network (MANET) using ant colony optimization technique. *Information Technology and Control* **49**(3): 308–319.
DOI: <https://doi.org/10.5755/j01.itc.49.3.25265>
- Logambigai R. & Kannan A. (2016). Fuzzy logic based unequal clustering for wireless sensor networks, *Wireless Networks* **22**(3): 945–957.
DOI: <https://doi.org/10.1007/s11276-015-1013-1>
- Mahdizadeh A.S., Khansari M., Rabiee H.R. & Salehi M. (2014). A congestion control protocol for wireless multimedia communication in sensor networks, *Ad Hoc Networks* **13**: 516–534.
DOI: <https://doi.org/10.1016/j.adhoc.2013.10.006>
- Misra S., Tiwari V. & Obaidat M.S. (2009). Lacas: learning automata based congestion avoidance scheme for healthcare wireless sensor networks, *IEEE Journal on Selected Areas in Communications* **27**(4): 466–479.
DOI: <https://doi.org/10.1109/JSAC.2009.090510>
- Mostafaei H. (2019). Energy-efficient algorithm for reliable routing of wireless sensor networks. *IEEE Transactions on Industrial Electronics* **66**(7): 5567–5575.
DOI: <https://doi.org/10.1109/TIE.2018.2869345>
- Mukherjee S. & Dasgupta P. (2013). A fuzzy real-time temporal logic. *International Journal of Approximate* **54**: 1452–1470.
DOI: <https://doi.org/10.1016/j.ijar.2013.06.004>
- Munir S., Bin Y.W., Biao R. & Man M. (2007). Fuzzy logic based congestion estimation for qos in wireless sensor network. *2007 IEEE Wireless Communications and Networking Conference (WCNC 2007)*, pp. 4336–4341.
- Perumal T. & Subramaniyan S. (2020). Statistical Markov model based natural inspired glowworm swarm multi-objective optimization for energy efficient data delivery in MANET. *Information Technology and Control* **49**(2): 333–347.
DOI: <https://doi.org/10.5755/j01.itc.49.2.23554>
- Rajeswari A.R., Kulothungan K., Ganapathy S. & Kannan A. (2019). A trusted fuzzy based stable and secure routing algorithm for effective communication in mobile adhoc networks. *Peer-to-Peer Networking and Applications* **12**: 1076–1096.
DOI: <https://doi.org/10.1007/s12083-019-00766-8>.
- Rao S.N. & Rama R.K.V.S.N. (2022). An energy-efficient routing scheme in wireless sensor networks for life time maximization using fuzzy based decision. *Adhoc & Sensor Wireless Networks* **51**: 23-39.
- Rekha C. & Gomathy C. (2012). IFCCDC: A fuzzy control based congestion detection and control in wireless sensor networks. *International Journal of Computer Applications* **47**(19): 12-17.

- DOI: <https://doi.org/10.5120/7295-0285>
- Rekha C., Gomathy C., Sebastian P. & Mon V. (2010). A survey on congestion control in wireless sensor networks. *International Journal of Computer Science and Communication* **1**(1):161–164.
- Rezaee A.A. & Pasandideh F. (2018). A fuzzy congestion control protocol based on active queue management in wireless sensor networks with medical applications, *Wireless Personal Communications* **98**(1): 815–842.
DOI: <https://doi.org/10.1007/s11277-017-4896-6>
- Rezaee A.A., Yaghmaee M.H. & Rahmani A.M. (2014). Optimized congestion management protocol for healthcare wireless sensor networks, *Wireless Personal Communications* **75**(1): 11–34.
DOI: <https://doi.org/10.1007/s11277-013-1337-z>
- Sangeetha G., Vijayalakshmi M., Ganapathy S. & Kannan A. (2019). An improved congestion-aware routing mechanism in sensor networks using fuzzy rule sets. *Peer-to-Peer Networking and Applications* **13**: 890–904.
DOI: <https://doi.org/10.1007/s12083-019-00821-4>.
- Sayyad J. & Choudhari N.K. (2014). Hierarchical tree based congestion control using fuzzy logic for heterogeneous traffic in WSN. *International Journal of Current Engineering and Technology* **4**(6): 4136–4143.
- Song F., Zhou Y., Chang L. & Zhang H. (2019). Modeling space terrestrial integrated networks with smart collaborative theory. *IEEE Networks* **33**(1): 51–57
DOI: <https://doi.org/10.1109/MNET.2018.1800187>
- Sonmez C., Incel O.D., Isik S., Donmez M.Y. & Ersoy C. (2014). Fuzzy based congestion control for wireless multimedia sensor networks. *EURASIP Journal on Wireless Communications and Networking* **63**: 1-17.
DOI: <https://doi.org/10.1186/1687-1499-2014-63>
- Tao L.Q. & Yu F.Q. (2010). ECODA: Enhanced congestion detection and avoidance for multiple class of traffic in sensor networks. *Transactions on Consumer Electronics* **56**(3): 1387 – 1394.
DOI: <https://doi.org/10.1109/TCE.2010.5606274>
- Ullah R., Faheem Y. & Kim B. (2017). Energy and congestion-aware routing metric for smart grid AMI networks in smart city. *IEEE Access* **5**: 13799–13810.
DOI: <https://doi.org/10.1109/ACCESS.2017.2728623>
- Uthra R.A. & Raja S.V.K. (2014). Energy efficient congestion control in wireless sensor network. In: *Recent Advances in Intelligent Informatics* (eds. S. Thampi, A. Abraham, S. Pal, J. Rodriguez), pp.331-341. Advances in Intelligent Systems and Computing, vol 235, Springer, Cham.
DOI: https://doi.org/10.1007/978-3-319-01778-5_34
- Wan C.Y., Eisenman S.B. & Campbell A.T. (2011). Energy-efficient congestion detection and avoidance in sensor networks. *ACM Transactions on Sensor Networks* **7**(4): 1–32.
DOI: <https://doi.org/10.1145/1921621.1921626>
- Wang X., Han S., Wu Y. & Wang X. (2012). Coverage and energy consumption control in mobile heterogeneous wireless sensor networks. *IEEE Transactions on Automatic Control* **58**(4): 975–988.
DOI: <https://doi.org/10.1109/TAC.2012.2225511>
- Yi D. & Yang H. (2016). HEER—a delay-aware and energy-efficient routing protocol for wireless sensor networks. *Computer Networks* **104**: 155–173.
DOI: <https://doi.org/10.1016/j.comnet.2016.04.022>
- Yi S., Kappes M., Garg S., Deng X., Kesidis G. & Das C.R. (2008). Proxy-RED: An AQM scheme for wireless local area networks. *Wireless Communications and Mobile Computing* **8**(4): 421–434.
DOI: <https://doi.org/10.1109/ICCCN.2004.1401706>
- Yin X. (2009). A fairness-aware congestion control scheme in wireless sensor networks, *IEEE Transactions on Vehicular Technology* **58**(9): 5225–5234.
DOI: <https://doi.org/10.1109/TVT.2009.2027022>
- Zou Z., & Qian Y. (2019). Wireless sensor network routing method based on improved ant colony algorithm. *Journal of Ambient Intelligence and Humanized Computing* **10**(3): 991–998
DOI: <https://doi.org/10.1007/s12652-018-0751-1>
- Rao N., Srinivas R.K.V., & Rama. (2022). An Energy-Efficient Routing Scheme in Wireless Sensor Networks for Life Time Maximization Using Fuzzy Based Decision. *Adhoc & Sensor Wireless Networks* **51**: 23–39.

RESEARCH ARTICLE

Agricultural Statistics

Evaluating the potential of an open sensor network to support reservoir pre-release decision making

E Warusavitharana^{1*}, PKS Mahanama¹, M Cannata², R Ratnayake¹, BH Sudantha¹ and D Strigaro²

¹ Department of Town and Country Planning, Faculty of Architecture, University of Moratuwa, Moratuwa.

² University of Applied Sciences and Arts of Southern Switzerland, Switzerland.

Submitted: 17 May 2021; Revised: 22 December 2021; Accepted: 25 February 2022

Abstract: Still in most countries, reservoir flood warnings are threshold-based alerts issued when water levels exceed thresholds. This current practice of releasing water from reservoirs causes flash floods in downstream areas and increases the likelihood of dam failures and public outrage. Pre-release of water from reservoirs is therefore an important strategy for downstream flood mitigation. Hydrological models can simulate river flow with sufficient lead time. Thus, the resulting outputs can be effectively applied to pre-release decision making in the reservoirs. Since the beginning of computer-aided applications, many attempts have been made to establish a decision support system for reservoir flood control. However, this is hampered by manual stations, low quality data, high cost of software and data, unknown parameter values, and lack of expertise, especially in developing countries. Therefore, a total open-source solution combined with low-cost open-source hardware, free and open-source software, and open standards was seen as the only way to overcome reservoir-related flood risk. Moreover, research studies on open-source hardware, software, standards, and data are limited to a few case studies reporting real-time data on certain environmental parameters. Therefore, the application of integrated open-source technologies for reservoir flood control remains an unexplored area. In this background, a hydrological model powered by integrated open-source technology is presented in this research for reservoir pre-release decision making. The model was tested for the Deduru Oya watershed using the SWAT (Soil and Water Assessment Tool) toolkit. The calibration results appear to be satisfactory for both daily and hourly time intervals. Thus, this model helps to simulate the inflow of the reservoir and determine the level of reservoir gate opening.

Keywords: 4ONSE, Deduru Oya basin, open-source technologies, reservoir pre-release.

INTRODUCTION

A reservoir is an artificial or natural lake or pond which is used to collect and store water for versatile activities. In Sri Lanka, reservoirs are called tanks or 'wewa', which is the Sinhala word. Reservoir flood control measures can be broadly classified as structural and non-structural. Structural measures involve the mitigation of floods through physical constructions. In ancient Sri Lanka, floods associated with tanks were mainly controlled through the 'Ellanga gammana' system which is also known as the 'Cascaded tank-village' system (Figure 1). This system includes a network of small to large tanks in a micro or mesoscale catchment for storing, conveying, and utilizing water from an ephemeral rivulet (Bandara, 1995). It can be considered as one of the structural measures adopted by ancient Sri Lankans to create a buffer against seasonal flooding and to store water during droughts. However, in certain circumstances, the functioning of this system has been negatively affected by the filling of paddy lands and the availability of abandoned tanks.

Compared to structural methods, non-structural methods of flood control have been accepted in the present day as a more proven method for reducing flood risk and damages. This is mainly due to its long-term sustainability and minimal cost for operation and maintenance. Out of the different non-structural measures, the flood warning is the best measure to undertake for the areas that deserve prompt attention. Flood warnings differ from forecasts

* Corresponding author (emeshi@uom.lk;  <https://orcid.org/0000-0002-0352-2926>)



as they are issued when an event occurs or is about to occur (WMO, 2013). In the Asian context, China and Japan have established reservoir flood control systems, taking advantage of real-time hydro-meteorological data (Takeuchi *et al.*, 1998; Guo *et al.*, 2004). However, like most developing countries, Sri Lanka lacks a reliable weather station network, which could offer continuous near-real-time data for decision-making. Limited weather stations, costly and offline data, unavailability of parameter values, expensive modeling software, and limited resource persons are some of the major reasons which hinder the application of hydro-meteorological data for reservoir pre-release decision making. In many countries, even in Sri Lanka, reservoir flood warnings are issued under conditions where the reservoir is already at full capacity. This creates flash floods in downstream areas due to sudden release of water at a high speed.

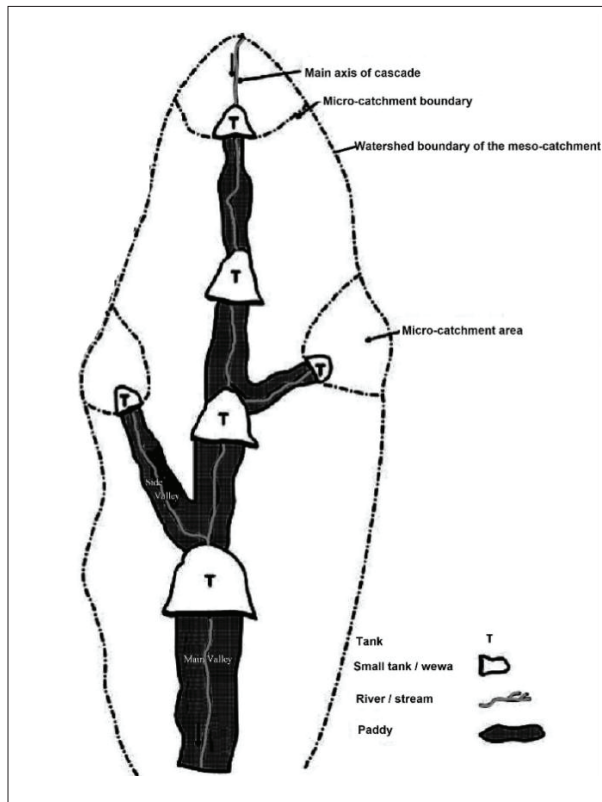


Figure 1: Schematic diagram of a tank cascade system (Panabokke *et al.*, 2002)

The project 4ONSE (4 times open and non-conventional technologies for sensing the environment) was launched in this setting to introduce an open-source weather station

network to support disaster warnings. IoT (internet of things) enabled technologies, open-source standards, and software have now made a significant turning point in the future of disaster warning from high-cost sophisticated devices to low-cost, open-source solutions. Evans (2011) defined the term internet of things (IoT) as a moment where more things are connected to the internet than people. Wireless connectivity and smart sensors are the two technologies that shape up the IoT network. Hence, IoT usually does the collecting of data through smart sensors and sharing them through the internet. Hart & Martinez (2015) have stated that most of the IoT-based applications and IoT-oriented research have been applied in cities and indoor environments where the relevant infrastructure facilities such as internet connectivity, accessibility, and electric power supply are available. Therefore, the most essential requirement to form a global environmental monitoring sensor network is to introduce IoT to remote environments where the IoT systems are powered by energy harvesting systems composed of sustainable energy source/s and energy storage units and wireless internet connectivity. Several of the world's most recognized companies such as IBM and HP have already started some initiatives in this respect (IBM,2010; HP, 2013; Liang & Huang, 2013).

The most cost-effective IoT applications have started to become popular after the addition of low-cost sensors, open-source hardware platforms, open-source software and standards in system development (Bitella *et al.*, 2014; Sadler *et al.*, 2014; Chemin *et al.*, 2015; Formisano *et al.*, 2015; Prescott *et al.*, 2016; Rao *et al.*, 2016; Saini *et al.*, 2016). However, only a few cases of integration of these open-source platforms for monitoring various environmental parameters have been reported. Valenzuela *et al.* (2018) developed a turbidity data acquisition system using Arduino as open hardware and MyOpenLap free software as open-source software. Sabatini (2017) has presented an approach of step-by-step installation of an automatic weather station in remote sites. Daniele *et al.* (2016) have developed an open hardware device based on Arduino to monitor the soil water potential for irrigation activities. A similar kind of application was developed by Bitella *et al.* (2014) to monitor the soil water content integrating the soil, vegetation, and atmosphere parameters. Prescott *et al.* (2016) discussed a hydro-climatic monitoring station that observes six water quality parameters. Mesas-Carrascosa *et al.* (2015) developed an open-source hardware device to record environmental parameters and a smartphone application to analyze the data. Sadler *et al.* (2014) developed a low-cost environmental monitoring system that measures air temperature and

relative humidity and automatically sends the collected data to Hydrologic Information System. Samourkasidis & Athanasiadis (2014) demonstrated an automated data archival system integrated with OGC-SOS, low-cost sensors, and Raspberry Pi as open hardware. Therefore, the integration of open-source technologies for reservoir

management in river basins, remains an unexplored area, especially with regard to reservoir pre-release decision making. Therefore, the main objective of this research is to evaluate the potential of such a system for reservoir pre-release decision-making with the support of a hydrological model.

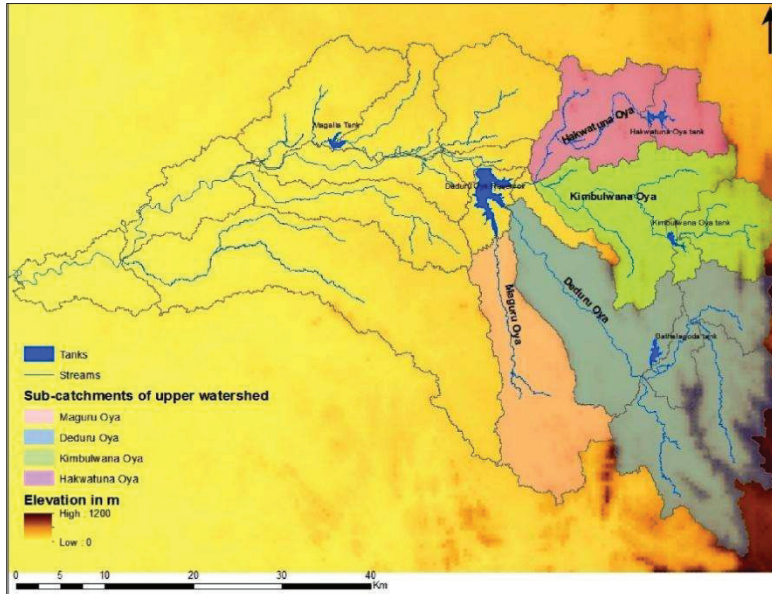


Figure 2: Hydrological network, sub-catchments & sub-basins of Deduru Oya basin

MATERIALS AND METHODS

Study site

Figure 2 shows the river basin chosen to establish the 4ONSE network, the Deduru Oya river basin. Its catchment area is about 2687 km² and the length of the main channel is 115 km. There are 8 major reservoirs and 2408 minor reservoirs in the basin. Among them, Deduru Oya reservoir is the largest and main reservoir in the basin. Thus, it has been selected for testing the reservoir management decision support system. Its significant geographical location is a major factor in flood control in the lower basin. The upper catchment area can be divided into four sub-catchments based on the four stream networks originating from the Central Highlands. Among them, the reservoir receives the largest amount of water from the Deduru Oya sub-catchment.

Open-source Framework

Arduino, istSOS (Istituto Scienze della Terra Sensor Observation Service) and OGC-SOS (Open Geospatial Consortium – Sensor Observation Service) are the open-source hardware, software and standards, respectively, used in system development. The istSOS allows access to all sensor observations from a centralized location based on OGC-SOS standard. The data are visualized in istSOS at a rate of 10 min. It has an automatic data validation procedure to identify the quality of near-real-time data. This validation procedure assigns a code for each datum after the validation test.

All the 4ONSE stations were built on Arduino Mega 2560 open hardware platforms. In addition to the weather stations, several river gauges were built using the same Arduino Mega version, to measure the water levels of the streams. Each station is powered by 30W

solar panel and 12V 35Ah rechargeable battery. The sensors of the stations, measured parameters, units, accuracy, and measuring range are shown in Table 1. Cost, WMO standards, and durability are the main factors that have been considered when selecting sensors for the stations (Cannata *et al.*, 2017; Sudantha *et al.*,

2018, 2019). Figures 3 (a) and (b) shows photographs of a weather station and a river gauge developed under the 4ONSE project. Figure 4 shows the locations where the 27 weather stations and 06 river gauges are deployed. The quality of the 4ONSE data was checked using some reference stations' data at daily and 10 minutes intervals.

Table 1: Sensors of the 4ONSE stations

Sensor	Parameter	Unit	Accuracy	Measuring range
DS18B20	Temperature	Degrees Celsius (°C)	±0.5°C	-10 to 85 °C
BME280	Relative Humidity	Percentage (%)	±3%	0% - 100%
BME280	Barometric pressure	Hectopascals (hPa)	±1 hPa	300 – 1000 hPa
ZHIPU wind speed sensor	Wind speed	Meters per second (ms ⁻¹)	±1 m/s	0-32.4 m/s
Anemometer 485 wind direction sensor	Wind direction	Degrees	±3°	16 different directions and any angle values can be identified
6465 Davis AeroCone Rain Gauge with Mountable Base	Precipitation	Millimeters (mm)	0.2mm	N/A
BH1750 light sensor module	Light intensity	Lux (lx)	1.44 times, Sensor Out / Actual lx	(1-65535lx)
Soil moisture module	Soil moisture	Percentage (%)	±2%	0 to 22%
River gauge module MB7062 XL-MaxSonar-WR1 Ultrasonic sensor	Water Level	Meters	±0.5cm	0 - 10m

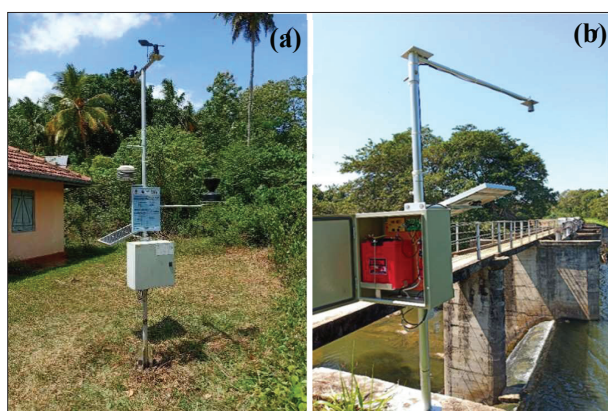


Figure 3: 4ONSE (a) weather station and (b) river gauge

In this study, the SWAT (Soil and Water Assessment Tool) open-source tool has been used to develop the hydrological model. The model calibration part has been

done using the SWAT-CUP open-source tool. Figure 5 shows the complete open-source framework used to develop the hydrological model to support pre-release decision making in the reservoir.

Development of hydrological model

The SWAT model has shown its capability in regional-scale hydrological modelling for simulation of river discharge (Ghoraba, 2015; Bailey *et al.*, 2016; Zhang *et al.*, 2016; Rafiei *et al.*, 2017; Warusavitharana, 2020). The model was developed using QSWAT plugin version 1.9 (Dile *et al.*, 2016) embedded in QGIS software. The QSWAT plugin successfully works in QGIS 2.6.1 Brighton version. The processing of input data such as DEM (Digital Elevation Model), land use, and soil was performed using QGIS Brighton version. SWAT's algorithms for infiltration, surface runoff, flow routing, impoundments, and lagging of surface runoff have been modified to allow flow simulations with a sub-daily time interval as small as one minute, and evapotranspiration,

soil water contents, base flow, and lateral flow are estimated daily and distributed equally for each time step (Jeong *et al.*, 2010). Therefore, using precipitation on a sub-daily basis and the other input data (relative humidity, temperature, wind speed and solar radiation) on a daily basis was sufficient.

Since SWAT is a continuous hydrological model, it requires a warm-up period of several years to stabilize

the model. Since 4ONSE is a newly deployed sensor network, the required data for the model’s warm-up period has been estimated through SWAT’s weather generator. The required monthly statistical data to operate the weather generator has been calculated using CFSR (Climate Forecasting System Reanalysis) data. Table 2 shows the models and methods used in the SWAT weather generator to estimate the missing weather data.

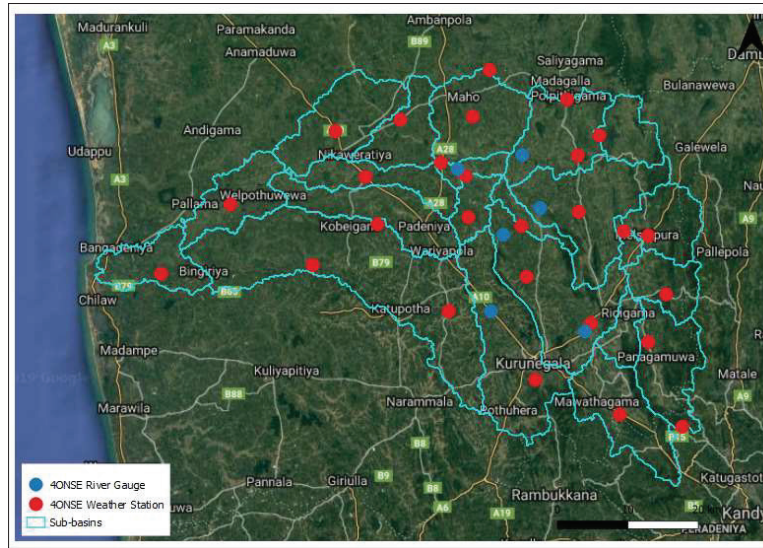


Figure 4: Locations of the 4ONSE stations in Deduru Oya basin

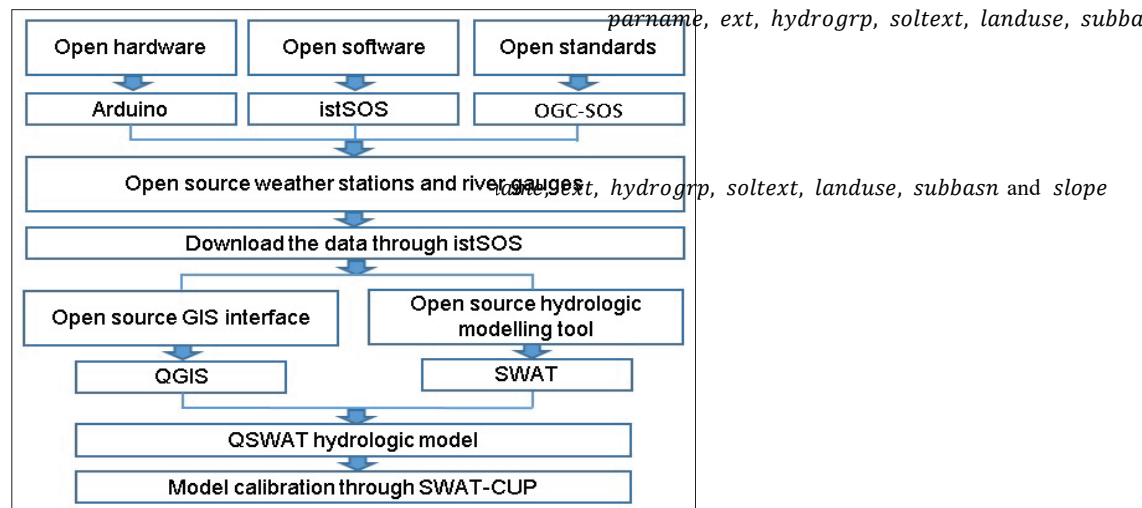


Figure 5: 4ONSE open-source framework for reservoir pre-release decision making

Table 2: Models and methods used in SWAT weather generator

Climatic variable	Model / Method
Daily precipitation	Model developed by Nicks (1974)
Sub-daily precipitation	Double exponential function
Daily maximum and minimum air temperature	Normal distribution
Daily average relative humidity	Triangular distribution
Daily solar radiation	Normal distribution
Daily mean wind speed	Modified exponential equation

Model calibration and uncertainty analysis

The model calibration and uncertainty analysis has been accomplished through the SWAT-CUP (SWAT-Calibration and Uncertainty Procedures) public domain programme. The four upper sub-catchments in the Deduru Oya basin have been separately calibrated in this study by developing four separate SWAT models. However, this research extends the results of the model associated with the Deduru Oya sub-catchment.

Primarily, the effectiveness of the model at daily time-step was checked using the default parameter values in the tool. As there was a substantial difference between the simulated flow and the observed flow, the model was regionalized first. Generally, the terms regionalization, parameterization, and calibration have similar meaning, which is adjusting the parameter values to reduce the difference between a simulated result and observed values. The parameterization / regionalization scheme of SWAT-CUP tool is as follows:

$x_{<parname>.<ext>.<hydrogrp>.<soltext>.<landuse>.<subbasn>.<slope>}$

The components in the scheme namely, *parname*, *ext*, *hydrogrp*, *soltext*, *landuse*, *subbasn* and *slope* represents the name of the parameters (as it appears in SWAT), the extension of the parameter, soil hydrologic group, land use type, sub-basin number, and the slope, respectively. As per the scheme, $x_{<parname>}$ represents the type of change to be applied to the parameter. The type of changes represented by $x_{<parname>}$ are as follows:

- 1) $V_{<parname>}$ - replacing the existing parameter
- 2) $A_{<parname>}$ - (existing parameter) + (given value)
- 3) $R_{<parname>}$ - (existing parameter) × (1+ given value)

In this study, several rules have been applied when selecting the appropriate $x_{<parname>}$ type. The usual principle in the SWAT_CUP tool to use type $R_{<parname>}$ for spatial parameters (e.g., land use and soil type). Moreover, considering the

ease of examining a wide range of values, type $R_{<parname>}$ was applied to parameters with a large range. Type $V_{<parname>}$ was applied to all the other parameters.

The sensitive/dominant parameters of the hydrological model were identified first through the OAT (One-at-a-time) sensitive analysis in SWAT-CUP. It shows the sensitivity of each parameter if all other parameters are kept constant. The model calibration and uncertainty analysis were performed according to the SUFI 2 (Sequential Uncertainties Fitting Version 2) algorithm integrated in SWAT. The fitness of the model was assessed statistically using the Nash-Sutcliffe Efficiency (NSE) method (Nash, 1957) (Equation 1), Goodness of fit linear regression model (R^2) (Equation 2), P factor (percentage of observed data simulated in the model) and R factor (average thickness of the 95% predication uncertainty (95PPU) divided by the standard deviation) (Equation 3). An NSE value between 0 and 1, $R^2 > 0.5$, P-factor > 70% and R-factor of around 1 are considered acceptable.

$$Maximize : NSE = \frac{\sum_{i=1}^n (O_i - \bar{O})^2 - \sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \dots(1)$$

$$R^2 = \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \dots(2)$$

$$R - factor = \frac{\frac{1}{n_j} \sum_{t_i=1}^{n_j} (x_s^{t_i,97.5\%} - x_s^{t_i,2.5\%})}{\sigma_{oj}} \dots(3)$$

where n is the number of observations, O_i is the i^{th} observed value, \bar{O} is the mean observed value, P_i is the i^{th} model-predicted value, \bar{P} is the mean model-predicted value, $x_s^{t_i,97.5\%}$ is the upper boundary and $x_s^{t_i,2.5\%}$ is the lower boundary of 95PPU.

As a summary, Table 3 lists all the open-source tools used in developing the hydrological model to support reservoir pre-release decision making.

Table 3: Open-source tools used in the hydrological model

Open-source tool	Function	Link
QGIS Brighton Version	<ul style="list-style-type: none"> Analyse vector and raster data GIS interface to run the SWAT model 	http://qgis.org/downloads/QGIS-OSGeo4W-2.6.1-1-Setup-x86.exe
QSWAT	<ul style="list-style-type: none"> SWAT plugin used to run the model in QGIS software 	https://swat.tamu.edu/software/qswat/
SWAT-Editor	<ul style="list-style-type: none"> Reading project databases Generating missing weather data Executing SWAT run Calibrating the model 	https://swat.tamu.edu/software/swat-editor/
SWAT-CUP	<ul style="list-style-type: none"> Identifying the dominant parameters & their ranges Calibrating the model Validating the model 	https://www.2w2e.com/home/SwatCup
istSOS	<ul style="list-style-type: none"> To view and download the data of 4ONSE stations 	https://geoservice.ist.supsi.ch/4onse/admin/

RESULTS AND DISCUSSION

Before applying the 4ONSE meteorological data in the model, its quality was checked using data from some reference stations in Sri Lanka. Table 4 shows the coefficient of determination of all key meteorological parameters measured at 10 minutes and daily intervals.

Table 4: Coefficient of determination (R^2) between 4ONSE stations and reference stations

Parameter	Interval	R^2 value
Temperature	10 minutes	0.9678
	Daily	0.9921
Rainfall – low altitude	10 minutes	0.7292
	Daily	0.7784
Rainfall – high altitude	10 minutes	0.7448
	10 minutes	0.9184
Relative humidity	Daily	0.9811
	10 minutes	0.9771
Air pressure	Daily	0.9929
	10 minutes	0.9889
Water level	Daily	0.9916

After running the model using daily 4ONSE data with SWAT’s default parameter values, several key features of

simulated flow and the actual flow were observed. They are:

1. Simulated flow peaks are higher than actual flow
2. The baseflow of the simulated flow is higher than the actual flow
3. The discharge in the simulated flow shifted to the left

To correct the above issues, the model was first regionalized. Table 5 shows the related dominant/sensitive parameters identified through the OAT analysis and applied modifications to correct the above deviations.

In the SWAT model, CN2 (curve number), CANMX (maximum canopy storage), SOL_AWC (soil available water content) and ESCO (soil evaporation compensation factor) are the parameters that contribute to the peak fluctuation of streamflow. However, CANMX was the only parameter that showed sensitivity to high peaks, especially for the dominant land use categories of coconut, rice, low-density residential, and rubber. The CANMX parameter represents the maximum amount of water that trees can be hold. This value is zero by default in the SWAT database. The peaks can be reduced by increasing the CANMX value of dominant land use classes. Since CANMX is a parameter that introduces water into the system it was taken separately and 50 runs performed, to obtain optimal values related to the four dominant land use categories. Table 6 shows the sensitive parameters and their values for daily and hourly intervals obtained during model calibration.

Table 5: Regionalized parameters

Observation	Reason	Related parameters	Applied modification
High peaks	High surface flow	CN2, SOL_AWC, ESCO, CANMX	Increase CANMX
Model over predicts the flow	High baseflow and/or little evapotranspiration	GWQMN, GW_REVAP, REVAPMN	Increase GWQMN & GW_REVAP
Discharge was shifted to left	Simulated flow leads the actual flow	SLOPE, OV_N, SLSUBBSN, CH_N2	Increase CH_N2

Table 6: Sensitive parameters and their value ranges related to daily and hourly intervals.

Sensitive parameter	Description	Type of change	Value range – daily interval	Value range – hourly interval
CN2	Initial SCS runoff curve number for moisture condition II)	R_	0.14 – 0.25	(-0.3) – 0.1
SOL_AWC	Available water capacity of the soil layer	R_	(-0.14) – 0.03	Not sensitive
ESCO	Soil evapotranspiration compensation factor	V_	0.68 – 0.95	Not sensitive
SOL_BD	Moist bulk density	R_	(-0.12) – 0.06	(-0.08) – 1.77
MSK_X	Weighting factor for wedge storage	V_	0 – 0.14	0 – 0.1
MSK_CO2	Muskingum coefficient for low flow	V_	(-0.04) – 0.59	0 – 8.1
MSK_CO1	Muskingum coefficient for normal flow	V_	0.96 – 1.31	1.0 – 5.2
ALPHA_BF	Baseflow alpha factor	V_	(-0.17) – 0.23	0 – 0.2
SOL_K	Saturated hydraulic conductivity	R_	(-1.21) – 0.06	(-0.47) – (-0.04)
CH_K2	Effective hydraulic conductivity in main channel alluvium	V_	4.27 – 6.73	1.1 – 27.3
CH_N1	Manning's "n" value for the tributary channels	V_	(-0.13) – 0.20	(-0.3) – 0.7
CH_N2	Manning's "n" value for main channel	V_	0.04 – 0.06	0 – 0.7
GW_DELAY	Ground water delay time	V_	38.90 – 89.94	Not sensitive
GW_REVAP	Groundwater "revap" coefficient	V_	0.24 – 0.45	0.1 – 0.2
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur	R_	3.14 – 4.32	0.8 – 2.0
SURLAG	Surface runoff lag coefficient	V_	Not sensitive	(-0.5) – 1.0

Except the SOL_AWC, ESCO, GW_DELAY, and SURLAG parameters, all the other parameters were received as sensitive for both daily and hourly time intervals for the Deduru Oya sub-catchment. For hourly simulation, the effect of ESCO, SOL_AWC, and GW_DELAY was not significant while SURLAG parameter was significant to simulate hourly flows.

Figure 6 shows the simulated and actual inflow in the Deduru Oya sub-catchment, after model calibration. The statistical results regarding P factor, R factor, R², and NSE presented in Table 7 appear to be satisfactory.

The Davis rain gauge used at 4ONSE weather stations typically has a margin of error of $\pm 4\%$ for rainfall rates up to 50 mm/hour and $\pm 5\%$ for rainfall rates in the 50 mm/hr to 100 mm/hr range. This is the main reason why some peaks do not reach the desired level.

Accordingly, the calibrated parameter values of the model can be easily applied to the model to simulate the inflow of the Deduru Oya reservoir. The time interval of the model can be changed according to the needs of the decision-makers who decide on water pre-release.

Customization of weather data and optimization of parameters should be done according to the selected time interval. Moreover, the application of near real-time weather data in the model is more valid when the time required for decision-making is greater than the time required to concentrate water from the upper basin to the reservoir.

The model developed in this research cannot be directly used for flood modelling. It only helps to simulate the flow into the reservoirs / tanks and to determine the level of opening of the reservoir gates, which helps in reducing downstream flood inundation. However, the inflows simulated by the model can be applied to SWAT's reservoir management tool to estimate reservoir capacity and determine reservoir outflow.

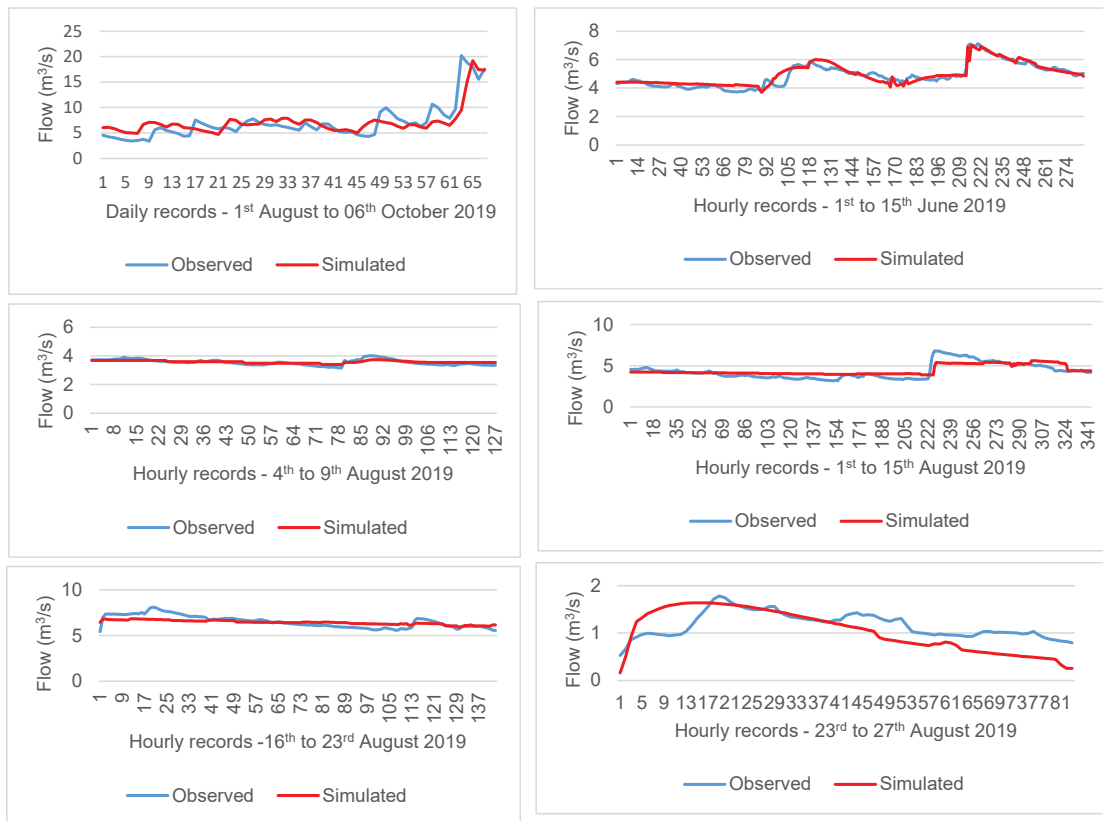


Figure 6: Simulated flow and Observed/Actual flow

Table 7: Statistical results of the hydrological model

Period	Interval	P factor	R factor	R ²	NSE
1 st August to 6 th October 2019	Daily	0.87	0.98	0.69	0.69
1 st to 15 th June 2019	Hourly	0.96	0.76	0.76	0.75
15 th to 30 th June 2019	Hourly	0.96	0.89	0.89	0.88
4 th to 9 th August 2019	Hourly	1.0	0.53	0.77	0.55
1 st to 15 th August 2019	Hourly	0.74	0.54	0.67	0.63
16 th to 23 rd August 2019	Hourly	0.83	0.00	0.67	0.43

CONCLUSION

Reservoir flood management applications require generating modelling results with the shortest lead time using near real-time meteorological data as input data. Due to the limitations of the existing setup, the incorporation of near real-time and quality hydro-meteorological input data to produce accurate hydrological estimates has never been used in Sri Lanka. Thus, this study intends to introduce a cost-effective decision support system to reservoir pre-release decision making. It is built entirely using several open source technologies for model implementation and data entry. QGIS Brighton version, QSWAT, SWAT Editor and SWAT-CUP are the open-source tools used to implement the hydrological model and Arduino, istSOS and OGS-SOS are the open-source technologies used to build meteorological stations and feed data to the model. Estimating parameter values for different time periods is tedious, time consuming, and labour and capital intensive. Therefore, the modelling approach presented in this research avoids the need for pre-specified parameter values and allows users to determine them at any time. The approach presented in this study is more appropriate for simulating reservoir inflow. If the basin has any series of tanks/cascade systems, the daily and sub-daily discharge flow of the upper basin tanks should be considered while developing the model. Furthermore, this study introduces a new approach for conducting hazard warnings in more remote environments where environmental processes cannot be realistically observed or studied due to the lack of access and facilities.

Conflict of interest statement

The authors declare that they have no conflict of interest

Data availability statement

The data that support the findings of this study are available from the corresponding author, [E.J. Warusavitharana], upon reasonable request.

Acknowledgment

The 4ONSE project was funded by the Swiss National Science Foundation (SNSF).

REFERENCES

Bailey R.T., Wible T.C., Arabi M., Records R.M. & Ditty J. (2016). Assessing regional-scale spatio-temporal patterns

of groundwater–surface water interactions using a coupled SWAT-MODFLOW model. *Hydrological Processes* **30**(23): 4420–4433.

DOI: <https://doi.org/10.1002/hyp.10933>

Bandara C.M.M. (1995). *Tank Cascade Systems in Sri Lanka: Some Thoughts on Their Development Implications*, pp 14. International Water Management Institute (IWMI), Colombo, Sri Lanka.

Bitella G., Rossi R., Bochicchio R., Perniola M. & Amato M. (2014). A novel low-cost open hardware platform for monitoring soil water content and multiple soil-air-vegetation parameters. *Sensors* **14**: 19639–19659.

DOI: <https://doi.org/10.3390/s141019639>

Cannata M., Strigaro D., Cardoso M., Sudantha B.H., Ratnayake R., Warusavitharana E.J. & Banadra P.D. (2017). 4onse D2.1 - System design report. Zenodo.

DOI: <https://doi.org/10.5281/zenodo.1244162>

Chemin Y., Bandara N. & Eriyagama N. (2015). A national upgrade of the climate monitoring grid in Sri Lanka. The place of Open design, OSHW and FOSS. *Geophysical Research Abstracts. EGU General Assembly*, 12–17 April, Vienna, Austria, pp. 17.

Daniele M., Facchi A., Depoli E.V., Renga F.M. & Gandolfi C. (2016). Irrig-OH: An open-hardware device for soil water potential monitoring and irrigation management. *Irrigation and Drainage* **65**(5): 750–761.

DOI: <https://doi.org/10.1002/ird.1989>

Dile Y., Daggupati P., George C., Srinivasan R. & Arnolde J. (2016). Introducing a new open source GIS user interface for the SWAT model. *Environmental Modelling and Software* **85**: 129–138.

DOI: <https://doi.org/10.1016/j.envsoft.2016.08.004>

Evans D. (2011). The Internet of Things: How the next evolution of the internet is changing everything. Cisco Internet Business Solutions Group. Available at http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

Formisano F., Massera E., De Vito S., Buonanno A., Di Francia G. & Delli Veneri P. (2015). Tinynose, an auxiliary smart gas sensor for RFID tag in vegetables ripening monitoring during refrigerated cargo transport. In: *Sensors: Lecture Notes in Electrical Engineering*, volume 319 (eds. D. Compagnone, F. Baldini, C. Di Natale, G. Betta & P. Siciliano), Springer, Cham, Switzerland.

DOI: https://doi.org/10.1007/978-3-319-09617-9_39

Ghoraba S.M. (2015). Hydrological modeling of the Simly Dam watershed (Pakistan) using GIS and SWAT model. *Alexandria Engineering Journal* **54**(3): 583–594.

DOI: <https://doi.org/10.1016/j.aej.2015.05.018>

Guo S., Zhang H., Chen H., Peng D., Liu P. & Pang B. (2004). A reservoir flood forecasting and control system for China. *Hydrological Sciences Journal* **49**(6): 972.

DOI: <https://doi.org/10.1623/hysj.49.6.959.55728>

Hart J. & Martinez K. (2015). Toward an environmental internet of things. *Earth and Space Science* **2**: 194–200.

DOI: <https://doi.org/10.1002/2014EA000044>

Hewlett Packard Development Company (2013). Central Nervous System for the Earth (CeNSE). Available at <http://>

- www8.hp.com/us/en/hp-information/environment/cense.html#.WTLzUmiGM2w
- IBM (2010). A Smarter Planet. Available at <http://www.ibm.com/smarterplanet/>
- Jeong J., Kannan N., Arnold J., Glick R., Gosselink L. & Srinivasan R. (2010). Development and integration of sub-hourly rainfall-runoff modeling capability within a watershed model. *Water Resources Management* **24**(15): 4505–4527.
DOI: <https://doi.org/10.1007/s11269-010-9670-4>
- Liang S. & Huang C.Y. (2013). GeoCENS: A geospatial cyberinfrastructure for the World-Wide Sensor Web. *Sensors* **13**: 13402–13424.
DOI: <https://doi.org/10.3390/s131013402>
- Mesas-Carrascosa F., Verdu Santano D., Morono J., Sanchez de la Orden M. & Garcia-Ferrer A. (2015). Open-source hardware to monitor environmental parameters in precision agriculture. *Biosystem Engineering* **137**: 73–83.
- Nash J. (1957). The form of the instantaneous unit hydrograph. *Comptes Rendus et Rapports Assemblee Generale de Toronto* **3**: 114–121.
DOI: <https://doi.org/10.1016/j.biosystemseng.2015.07.005>
- Nicks A. (1974). Stochastic generation of the occurrence, pattern and location of maximum amount of daily rainfall. *Proceedings of The Symposium on Statistical Hydrology*, pp. 154–171. Government Printing Office, Washington DC, USA.
- Prescott E., Rom C., Marchiori A. & Hayes B. (2016). HydroSense: An open platform for hydroclimatic monitoring.
DOI: <https://doi.org/10.1109/SMARTCOMP.2016.7501695>
- Rafiei Emam A., Kappas M., Linh N.H.K. & Renchin T. (2017). Hydrological modeling and runoff mitigation in an ungauged basin of central Vietnam using SWAT model. *Hydrology* **4**(1): 16.
DOI: <https://doi.org/10.3390/hydrology4010016>
- Rao B., Rao K. & Ome N. (2016). Internet of Things (IOT) based weather monitoring system. *International Journal of Advanced Research in Computer and Communication Engineering* **5**(9): 312–319.
DOI: <https://doi.org/10.3390/10.17577/IJERTCONV6IS13149>
- Sabatini F. (2017). Setting up and managing automatic weather stations for remote sites monitoring from Niger and Nepal. In: *Renewing Local Planning to Face Climate Change in the Tropics* (eds. T. Pezzoli & A. Tarchiani), pp. 21–39. Springer International Publishing, Cham, Switzerland.
DOI: https://doi.org/10.1007/978-3-319-59096-7_2
- Sadler J., Ames D. & Khattar R. (2014). Open-hardware meets open software for environmental monitoring. *7th International Congress of Environmental Modelling and Software*, San Diego, California, USA.
DOI: <https://doi.org/10.13140/2.1.4693.8241>
- Saini H., Thakur A., Ahuja S., Sabharwal N. & Kumar N. (2016). Arduino based automatic wireless weather station with remote graphical application and alerts. *Proceedings of the 3rd International Conference on Signal Processing and Integrated Networks (SPIN)*, 11–12 February, Noida, India, pp. 605–609.
DOI: <https://doi.org/10.1109/SPIN.2016.7566768>
- Samourkasidis A. & Athanasiadis I. (2014). Towards a low-cost, full-service air quality data archival system. *7th International Congress on Environmental Modelling and Software: Bold Visions for Environmental Modeling, iEMSS 2014* (eds. D.P. Ames, N.W.T. Quinn & A.E. Rizzoli), San Diego, USA, 15–19 June, pp. 1192–1199. Available at http://www.iemss.org/sites/iemss2014/papers/iemss2014_submission_300.pdf
- Sudantha B.H., Warusavitharana E.J., Ratnayake R., Mahanama P.K.S., Cannata M. & Strigaro D. (2018). Appropriateness of low-cost sensor network for environmental monitoring in a tropical country: Experience and lessons learnt from real world deployment. *PeerJ PrePrints* **6**: e27224v1
DOI: <https://doi.org/10.7287/peerj.preprints.27224v1>
- Sudantha B.H., Warusavitharana E.J., Ratnayake R., Mahanama P.K.S., Warusavitharana R.J., Tasheema R.P., Cannata M. & Strigaro D. (2019). 4ONSE as a complementary to conventional weather observation network. *Proceedings of the 4th International Conference on Information Technology Research (ICITR)*, 10–13 December. University of Moratuwa, Sri Lanka, pp. 1–6.
DOI: <https://doi.org/10.1109/ICITR49409.2019.9407798>
- Takeuchi K., Hamlin M., Kundzewicz Z., Rosbjerg D. & Simonovic S. (1998). *Sustainable Reservoir Development and Management*. IAHS Publications, Wallingford, UK.
- Valenzuela C., Sosa C., Castaneda M., Palomeque J. & Amaro I. (2018). Turbidity measurement system for aquaculture effluents using an open source software and hardware. *Nature Environment and Pollution Technology* **17**(3): 957–961.
- Warusavitharana E.J. (2020). Semi-distributed parameter optimization and uncertainty assessment for an ungauged catchment of Deduru Oya Basin in Sri Lanka. *International Journal of River Basin Management* **18**(1): 95–105.
DOI: <https://doi.org/10.1080/15715124.2019.1656221>
- World Meteorological Organization (WMO) (2013). *Flood Forecasting and Early Warning*. World Meteorological Organization.
- Zhang L., Jin X., He C., Zhang B., Zhang X., Li J., Zhao C., Tia J. & DeMarchi C. (2016). Comparison of SWAT and DLBRM for hydrological modeling of a mountainous watershed in arid northwest China. *Journal of Hydrologic Engineering* **21**(5).
DOI: [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001313](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001313)

RESEARCH ARTICLE

Ophthalmic Genetics

High prevalence of glaucoma-associated CYP1B1 mutation (p.G61E) in primary congenital and open angle glaucoma patients in Pakistan

K Yousaf, R Bashir*, K Balqees, S Naz, N Munir and F Aslam

Department of Biotechnology, Lahore College for Women University, Lahore, Pakistan.

Submitted: 17 June 2021; Revised: 22 December 2021; Accepted: 25 February 2022

Abstract: The purpose of this research study was to investigate the prevalence of the p.G61E variant of the *CYP1B1* gene among primary congenital and open angle glaucoma (PCG and POAG) patients in the province of Punjab, Pakistan. A total of 112 POAG and 50 PCG patients were enrolled in this study. Detailed clinical examination was carried out on all patients. Screening of G61E was done by direct Sanger sequencing. Different *in silico* tools, e.g., Clustal Omega, PSIPRED, and Franklin tools were used to check the conservation, secondary structure, and pathogenicity, respectively, of this variant. Sanger sequencing of the whole *CYP1B1* gene revealed a homozygous missense transition, c.182G>A, p.G61E in 25/50 (50%) of PCG and in 42/112 (37.5%) of POAG cases, which co-segregated with the disease phenotype. This study revealed that p.G61E is the relatively major contributor of PCG. However, 42 POAG patients harbouring the G61E mutation showed moderate to severe phenotype, suggesting the genetic heterogeneity of this variant in Pakistani population.

Keywords: Juvenile open angle glaucoma, intra ocular pressure, latent transforming growth factor-beta protein 2, primary congenital glaucoma.

INTRODUCTION

Glaucoma is a heterogeneous group of visual neuropathies which leads to optic nerve damage and permanent loss of vision if left untreated (Allingham *et al.*, 2009). Primary congenital glaucoma (PCG) is a more severe form of the disease categorized by atypical development of trabecular meshwork and elevated intraocular pressure (IOP) at the

time of birth or within the first three years of life, which manifest itself with photophobia, excessive tearing, and buphthalmos (Liu & Allingham, 2017). PCG mostly shows autosomal recessive form of inheritance (Firasat *et al.*, 2008). In the Western region, the incidence of PCG is 1 in 10,000–20,000 live births while in the Middle Eastern region the incidence is 1 in 2500 to 8200 (Lewis *et al.*, 2017). Primary open angle glaucoma (POAG) is a common cause of irreparable blindness in the world. It is an optic neuropathy characterized by open anterior chamber, complete cupping of nerve head, loss of visual field and elevated IOP (Liu & Allingham, 2017). On the basis of age of onset, POAG is subcategorized into adult-onset open angle glaucoma (AOAG) and juvenile-onset open angle glaucoma (JOAG). JOAG as an idiopathic glaucoma occurring in children above three years of age and it is associated with myopia and elevated intraocular pressure with fluctuation (Kwun *et al.*, 2016). The age of onset of JOAG is 3-40 years and shows autosomal dominant inheritance; however a few cases of autosomal recessive inheritance from Saudi Arabia and India have been reported (Khan *et al.*, 2011). Adult-onset POAG manifest itself after the age of 40 years and shows a complex pattern of inheritance because multiple genes and environmental factors are involved (Rauf *et al.*, 2016). The estimated prevalence of JOAG ranges from 0.38 to 2 in 100,000 in individuals between 4 and 20 years of age (Taqi *et al.*, 2011). JOAG comprises 4% of childhood glaucoma (Kwun *et al.*, 2016). POAG is the

* Corresponding author (rashidasbs@yahoo.com;  <https://orcid.org/0000-0003-1270-3742>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

most common form of glaucoma worldwide accounting for 90% of all cases reported (Qureshi *et al.*, 2006). POAG is also the common types of glaucoma in Pakistan (Taqi *et al.*, 2011).

The *CYP1B1* (OMIM 601771) gene is a major contributor in glaucoma pathogenesis (Achary *et al.*, 2006). *CYP1B* encodes a member of the Cytochrome p450 superfamily, subfamily I. The *CYP1B1* gene is localized in chromosome number 2 and consists of three exons out of which two are coding (López-Garrido *et al.*, 2015). It has been reported that the *CYP1B1* protein plays role in the development of trabecular meshwork, iridocorneal angle, and in the formation of the ciliary (Rauf *et al.*, 2016). Pathogenic and potentially pathogenic mutations in *CYP1B1* have been reported in JOAG, POAG, and PCG patients from various populations around the globe (Micheal *et al.*, 2014). Mutation p.G61E has been reported previously in POAG and PCG patients from different ethnicities (Qashqai *et al.*, 2018). Recent investigations on the functional role of the homozygous p.G61E *CYP1B1* mutation in a glaucoma family have shown that this mutation disturbs the microfibrils of the extracellular matrix (ECM), making its presence less abundant, and increases more fragmented protein in affected individuals. This study has suggested that disruption in development of the ECM-trabecular meshwork leads to increased resistance to aqueous outflow, causing elevated IOP, and eventually causes development of glaucoma (Banerjee *et al.*, 2016). The G61E mutation has been reported to be the commonest mutation in Lebanon, Iran, Egypt, and other Middle Eastern countries (Qashqai *et al.*, 2018). This variant is a founder mutation in Middle Eastern populations (Saudi Arabia, Iran, Morocco, Kuwait, Israel, Oman, and Egypt) (Campos-Mollo *et al.*, 2009; Li *et al.*, 2011). In Saudi POAG patients the frequency of this variant accounted for 64.3% (Qashqai *et al.*, 2018). Heterozygosity for the mutation p.G61E has already been reported as a cause of adult-onset POAG in two Pakistani sporadic cases (Micheal *et al.*, 2014). In heterozygous form, p.G61E has been reported with mild phenotype in PCG and POAG patients of different ethnicities (Khan *et al.*, 2011; López-Garrido *et al.*, 2015; Wiggs & Pasquale, 2017). Previously, our study demonstrated that the p.G61E variant was segregated in a consanguineous Pakistani family with co-existence of JOAG and PCG in two successive generations (Bashir *et al.*, 2015). Only one study showed heterozygous p.G61E mutation in sporadic POAG patients from Pakistan (Micheal *et al.*, 2014). Mutation Taster and PolyPhen predicted this

variant as disease causing, and the Franklin tool qualified this variant to be deleterious.

POAG and PCG are the most prevalent types of glaucoma in Pakistan (Iqbal *et al.*, 2011). However, few studies are available on the contribution of the p.G61E mutation in the Pakistani population. In the current study, the aim was to investigate the frequency of the p.G61E variant in PCG and POAG cases.

MATERIALS AND METHODS

Selection of study subjects and clinical examination

The study followed the guidelines of the Declaration of Helsinki. This study was approved by the Ethical Review Committee of the Department of Biotechnology, Lahore College for Women University, Lahore, Pakistan with reference number 2019-07. Informed consent was obtained from all participants or from the parents in case of infants.

Patients identified as having POAG and PCG were recruited from 2017-2020 (a 3-year period) from the Mughal Eye Hospital Lahore and the eye ward of the Children Hospital and Institute of Child Health (CHICH) Punjab. CHICH is the major ophthalmology hospital recommended for congenital cases in Punjab, and patients travel from all parts of Pakistan to this hospital. 112 unrelated Pakistani POAG and 50 PCG patients from consanguineous marriage were enrolled. 196 unrelated ethnically matched controls were also recruited.

For this study, the following inclusion criteria were used for POAG and PCG cases: (1) Age at the time of diagnosis, (2) Intraocular pressure > 21 mm Hg in adults and >12mm Hg in at least one eye in case of PCG, (3) Anterior chamber angle that was open to scleral spur or ciliary body, (4) Evidence of optic nerve damage and loss of visual field consistent with glaucoma in at least one eye, and (5) cup-to-disc (C/D) ratio > 0.3 of optic nerve. Megalocornea with corneal diameter >12 mm and rupture in Descemet's membrane have been considered. Inclusion criteria for POAG: ethnically matched controls (50 males, 50 females) and age ranges from 40 to 67 years were selected after ophthalmic screening, and individuals with normal range of IOP (not having anti-glaucoma medication), open angles after gonioscopy examination and healthy optic disk with no history of eye disease(s) or any ophthalmic surgeries.

For PCG, ethnically matched controls (56 males, 40 females), age ranges from 2 to 15 years with normal IOP range, normal cup/disc ratio, normal corneal diameter and intact Descemet's membrane were included.

Screening of p.G61E mutation

Primers were designed for coding exon 2 of *CYP1B1* with Primer3 including 50–80 bp of intronic region from the upstream and the downstream region of coding exon with Primer3 software (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>). Polymerase chain reaction (PCR) amplification was performed from genomic DNA. Each reaction was performed in a final volume of 50 μ L under standard conditions with 50 ng DNA in 2.5 μ L of Taq buffer, 2 mM MgCl₂ (1 μ L), 0.2 mM dNTPs (1.5 μ L), 0.24 μ L forward primer, 0.24 μ L reverse primer, 25 ng genomic DNA and then PCR water was added to raise its volume to 50 μ L. Denaturation of PCR products was performed at 95 °C for 30 s and annealed at 55 °C for 30 s. The initial extension was performed at 75 °C for 1 min and final extension was performed at 75 °C for 5 min. PCR products were then purified by gene cleaning column (Gel Extraction Kit Thermo Scientific GeneJet USA). After purification, samples were sequenced by direct sequencing with Big Dye Terminator® (V 3.1) Cycle Sequencing Kit (Applied Biosystems, Foster City, CA).

In silico analysis

ClustalW alignments were performed to determine evolutionary conservation, Figure 1C. To predict secondary structure, elements from amino acid sequence PSIPRED v.3.3 were used, Figure 1D. Franklin tool classified the variant as potentially pathogenic (<https://franklin.genoox.com/clinical-db/variant/snp/chr2-38302350-C-T?app=assessment-tools>).

RESULTS AND DISCUSSION

Table 1 shows the clinical characteristics of all PCG and POAG patients clinically examined. In 50 PCG and 112 POAG patients, the mean age in POAG was 52.8 \pm 11.1 years with range 40–70 years. In PCG cases the mean age was 12.2 \pm 9.7 years and range was 1 month to 9 years. The mean age of IOP in PCG cases was 19.5 \pm 6 years whereas in POAG the mean age of IOP was 24.3 \pm 7.8 years (Table 1). Clinical features of POAG patients showed watery eyes with enlarged globe and hazy cornea. They felt pain, headache, itching, and blurred vision. Some of them had maintained normal IOP pressure by regular treatment and prescribed medications. Clinical features of PCG patients showed hazy cornea, photophobia, complete opacification of corneas, and excessive tearing with enlarged globe. Most of the PCG patient underwent trabeculectomy in both or one eye to maintain normal IOP pressure along with the use of prescribed medications. IOP of PCG and POAG was positively associated with glaucoma development (Figure 1 and Figure 2).

Sequencing results of *CYP1B1* showed a missense mutation c.182G>A (p.G61E) in exon 2 which was homozygous in 25/50 (50%) PCG cases and in 42/112 (37.5%) POAG cases; the mutation co-segregated with the disease phenotype in 50 PCG and 42 POAG patients (Figure 2A, 2B). The variant was absent in 196 ethnically matched control chromosomes, further confirming that this mutation is pathogenic in the Pakistani population. MutationTaster and PolyPhen predicted this variant as disease causing. Franklin tool qualified this variant to be deleterious. PSIPHRED sequence plot showed that residue G61 lies in the Protein Helix region. Genetic variants in *CYP1B1* gene are major contributing factors in the pathogenesis of PCG and POAG in

Table 1: Clinical parameters of PCG and POAG patients and control groups

Parameters	Status	PCG	POAG	Control POAG	Control PCG
Gender	Male	30	66	50	56
	Female	20	46	50	40
Age	Mean \pm SD	12.18 \pm 9.7	52.8 \pm 11.1	65.9 \pm 11.6	13.2 \pm 6.4
Unit: (mth,yrs)	Range	1 mth–14 yrs	40 – 70 yrs	40 yrs – 67 yrs	2–15 yrs
Max. IOP (Unit: mmHg)	Mean \pm SD	19.46 \pm 5.96	24.32 \pm 7.8	15.0 \pm 2.7 8–21	16.5 \pm 3.7
CD ratio	Range	0.3–0.9	0.4 – 0.9	0.2–0.3	0.2–0.3

C/D: cup to disk ratio, mth: month yrs: years

various populations of the world (Qashqai *et al.*, 2018; Youngblood *et al.*, 2019; Ling *et al.*, 2020). The variant p.G61E has been found with varying frequency in different populations; however, data on genetic spectrum of this variant in Pakistani population is very limited. To explore the clinical spectrum of this variant, the present study was performed to determine the frequency and to

make genotype-phenotype comparison of the p.G61E mutation associated with PCG and POAG in the Pakistani population. This genetic study was done to demonstrate the role of the p.G61E variant of *CYP1B1* gene in Pakistani PCG and POAG patients. This cohort included 162 patients (n = 50 PCG, n = 112 POAG) as well as 196 control.

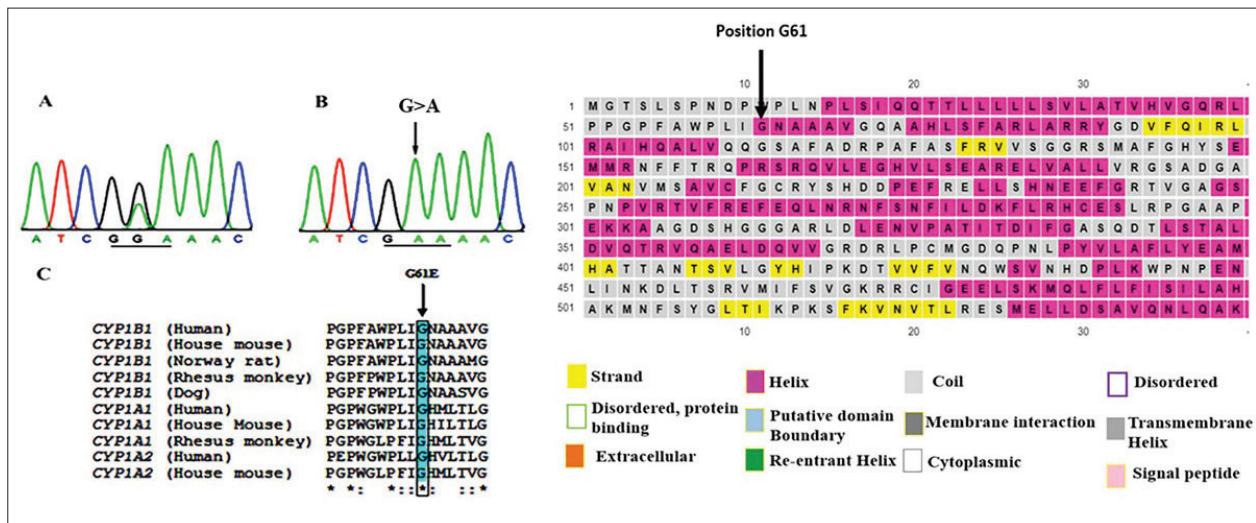


Figure 1: Electrogram of G61E mutation sequence. (A) normal sequence in control sample; (B) G>A transition in patient’s DNA; (C) ClustalW sequence alignment indicates conservation of p. G61E mutation in *CYP1B1*, *CYP1A1* and *CYP1A2* among different vertebrate species; (D) PSIPred sequence plot indicating that residue G61 lies in Protein Helix.

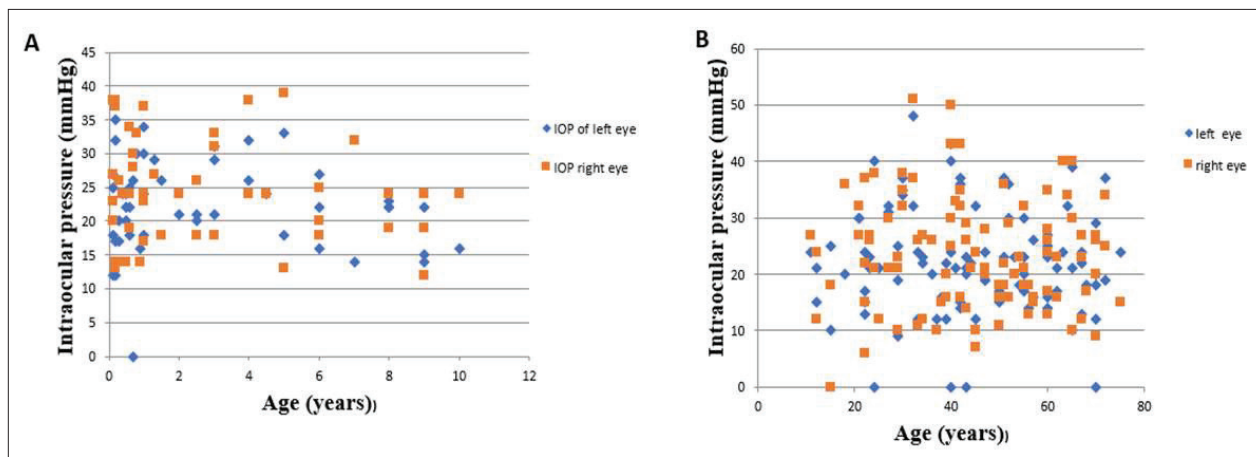


Figure 2: Graph showing IOP among both groups, *i.e.*, PCG patients (A) and POAG (B) patients.

The p.G61E mutation has already been reported to be a frequent cause of PCG and POAG in many different populations of the world including Arab, Middle Eastern and Mediterranean countries (Khan *et al.*, 2011; Micheal *et al.*, 2014; López-Garrido *et al.*, 2015), whereas this variant has been reported with incomplete penetrance as well in different populations (Campos-Mollo *et al.*, 2009 Lewis *et al.*, 2017;). This mutation has been shown to be a ‘hypomorphic’ allele, having reduced enzymatic activity of the *CYP1B1* protein (Campos-Mollo *et al.*, 2009). In previous studies, it was reported that the homozygous p.G61E mutation was pathogenic in PCG patients of Saudi Arabia (Abu-Amero *et al.*, 2011). In a recent study from Pakistan, this variant has been identified in homozygous state in a PCG patient belonging to a consanguineous family. However, in some cases the p.G61E mutation may exhibit decreased penetrance which has been reported previously in Spanish and Saudi Arabian patients (Campos-Mollo *et al.*, 2009; Youngblood *et al.*, 2019). In a cohort of Saudi POAG patients, the p.G61E variant accounted for 64.3% (Abu-Amero *et al.*, 2011). However, in another study of the Saudi population, the mutation was reported in heterozygous form and accounted for 4% of POAG cases and 2% of controls (Abu-Amero *et al.*, 2011). Mutation p.G61E has been the most frequent mutated *CYP1B1* allele in Iranian PCG patients. It is also the most commonly detected mutation among the POAG patients (Qashqai *et al.*, 2018). In adult-onset POAG patients, the variant p.G61E is mostly reported in heterozygous and compound heterozygous state, with a milder disease phenotype (López-Garrido *et al.*, 2015). In homozygous state the p.G61E variant showed severe phenotype in Indian and Saudi Arabian PCG patients. In Saudi JOAG patients, this mutation is reported to result in a severe phenotype in recessive form. In a report from the Pakistani population, the heterozygous p.G61E variant was reported in sporadic cases of POAG (Micheal *et al.*, 2014). Moreover, in our previous study from Pakistan we reported the homozygous p.G61E variant, which, co-segregating in two generations of JOAG and PCG patients from a consanguineous family, showed a severe phenotype with variable penetrance of this p.G61E variant (Bashir *et al.*, 2015).

The residue p.G61E lies in the functional part of a very highly conserved hinge region of the *CYP1B1* protein, from which region the polypeptide chain shows a sharp turn (Panicker *et al.*, 2004). It is present right after the N-terminus of a region rich in proline. The proline–proline–glycine–proline motif might make the

link between the membrane binding N-terminus and the globular region of the P450 protein. The change of glycine to glutamic acid at position 61 causes a complete loss of haem binding, resulting in minimal catalytic activity of the enzyme and results in disease manifestation (Panicker *et al.*, 2004). Furthermore, MutationTaster, SIFT, Franklin tool, and PolyPhen predicted this variant as disease causing.

In this study, screening for mutation p.G61E in the *CYP1B1* gene revealed that 50% of PCG cases and 37.5% of POAG cases carried this mutation in homozygous state. The frequency of the p.G61E variant in PCG cases is higher in the Pakistani population as compare to Chinese, Iranian, and Spanish populations (López-Garrido *et al.*, 2010; Safari *et al.*, 2016; Abu-Amero *et al.*, 2018). These variations in mutational frequency may be due to differences in geographical region, and epigenetic factors such as age, gender, race, and family history also contribute in causing PCG and POAG. IOP and CD ratio of PCG cases in p.G61E positive patients was high, resulting in opacification of cornea and need of surgical interventions. The patients affected with PCG showed severe phenotype indicating complete penetrance of this mutation in congenital cases. However, 19 patients affected with POAG showed mild phenotype whereas 23 cases showed severe phenotype, which showed the genetic heterogeneity of this variant in POAG.

CONCLUSION

We conclude that the *CYP1B1* gene is important in the aetiology of PCG as well as POAG, in the Pakistani population. This study reveals that p.G61E is more common among primary congenital glaucoma in the Pakistani cohort. However, the phenotypic variability associated with p.G61E variant in POAG patients of our population is possibly the effect of genetic modifiers or epigenetic factors that may modify the manifestation and variable penetrance. Further detailed studies applying whole-exome sequencing on the DNA of a large number of severely affected cases and those of less affected patients could reveal further genetic modifiers.

REFERENCES

- Abu-Amero K.K., Sultan T., Al-Obeidan S.A. & Kondkar A.A. (2018). Analysis of *CYP1B1* sequence alterations in patients with primary open-angle glaucoma of Saudi origin. *Clinical Ophthalmology* **12**: 1413–1416.

- DOI: <https://doi.org/10.2147/OPHTH.S169943>
- Abu-Amero K.K., Osman E.A., Mousa A., Wheeler J., Whigham B., Allingham R.R., Hauser M.A. & Al-Obeidan S.A. (2011). Screening of CYP1B1 and LTBP2 genes in Saudi families with primary congenital glaucoma: genotype-phenotype correlation. *Molecular Vision* **17**:911.
- Achary M.S., Reddy A.B., Chakrabarti S., Panicker S.G., Mandal A.K., Ahmed N., Balasubramanian D., Hasnain S.E. & Nagarajaram H.A. (2006). Disease-causing mutations in proteins: structural analysis of the CYP1B1 mutations causing primary congenital glaucoma in humans. *Biophysical Journal* **15**: 91(12): 4329-4339.
- Allingham R.R., Liu Y., & Rhee D.J. (2009). The genetics of primary open-angle glaucoma: a review. *Experimental Eye Research* **88** (4): 837–844.
DOI: <https://doi.org/10.1016/j.exer.2008.11.003>
- Bashir R., Tahir H., Yousaf K., Naz S. & Naz S. (2015). Homozygous p.G61E mutation in a consanguineous Pakistani family with co-existence of juvenile-onset open angle glaucoma and primary congenital glaucoma. *Gene* **570**(2): 295–298.
DOI: <https://doi.org/10.1016/j.gene.2015.07.014>
- Banerjee A., Chakraborty S., Chakraborty A., Chakrabarti S. & Ray K. (2016). Functional and structural analyses of CYP1B1 variants linked to congenital and adult-onset glaucoma to investigate the molecular basis of these diseases. *PLoS One* **11**(5): e0156252.
DOI: <https://doi.org/10.1371/journal.pone.0156252>.
- Campos-Mollo E., López-Garrido M.P., Blanco-Marchite C., Garcia-Feijoo J., Peralta J., Belmonte-Martínez J., Ayuso C. & Escribano J. (2009). CYP1B1 mutations in Spanish patients with primary congenital glaucoma: phenotypic and functional variability. *Molecular Vision* **15**: 417-431.
- Firasat S., Riazuddin S.A., Khan S.N. & Riazuddin S. (2008). Novel CYP1B1 mutations in consanguineous Pakistani families with primary congenital glaucoma. *Molecular Vision* **14**: 2002–2009.
- Iqbal S., Khan Z., Shah S.A., & Khan M.Y. (2011). Types and presentation of glaucoma. *Journal of Postgraduate Medical Institute* **22**(4): 324–332.
- Khan A.O., Al-Abdi L., Mohamed J.Y., Aldahmesh M.A. & Alkuraya F.S. (2011). Familial juvenile glaucoma with underlying homozygous p.G61E CYP1B1 mutations. *Journal of American Association for Pediatric Ophthalmology and Strabismus* **15**: 198–199.
DOI: <https://doi.org/10.1016/j.jaapos.2011.01.156>
- Kwun Y., Lee E.J., Han J.C. & Kee C. (2016). Clinical characteristics of juvenile-onset open angle glaucoma. *Korean Journal of Ophthalmology* **30**(2):127–133.
DOI: <http://dx.doi.org/10.3341/kjo.2016.30.2.127>
- Liu Y. & Allingham R.R. (2017). Major review: molecular genetics of primary open-angle glaucoma. *Experimental Eye Research* **160**: 62–84.
DOI: <https://doi.org/10.1016/j.exer.2017.05.002>
- López-Garrido M.P., Sánchez-Sánchez F., López-Martínez F., Aroca-Aguilar J.D., Blanco-Marchite C., Coca-Prados M. & Escribano J. (2006). Heterozygous CYP1B1 gene mutations in Spanish patients with primary open-angle glaucoma. *Molecular Vision* **12**: 748–755.
- Ling C., Zhang D., Zhang J., Sun H., Du Q. & Li X. (2020). Updates on the molecular genetics of primary congenital glaucoma. *Experimental and Therapeutic Medicine* **20**(2): 968–977.
DOI: <https://doi.org/10.3892/etm.2020.8767>
- Lewis C.J., Hedberg-Buenz A., DeLuca A.P., Stone E.M., Alward W. & Fingert J. H. (2017). Primary congenital and developmental glaucomas. *Human Molecular Genetics* **26**(R1): 28–36.
DOI: <https://doi.org/10.1093/hmg/ddx205>
- López-Garrido M.P., Blanco-Marchite C., Sánchez-Sánchez F., López-Sánchez E., Chaqués-Alepuz V., Campos-Mollo E., Salinas-Sánchez A.S. & Escribano J. (2010). Functional analysis of CYP1B1 mutations and association of heterozygous hypomorphic alleles with primary open-angle glaucoma. *Clinical Genetics* **77**(1): 70–78.
DOI: <https://doi.org/10.1111/j.1399-0004.2009.01284.x>
- Micheal S., Ayub H., Zafar S.N., Bakker B., Ali M., Akhtar F., Islam F., Khan M., Qamar R. & Hollander A.D. (2015). Identification of novel CYP1B1 gene mutations in patients with primary congenital and primary open-angle glaucoma. *Clinical and Experimental Ophthalmology* **43**(1): 31–39.
DOI: <https://doi.org/10.1111/ceo.12369>
- Panicker S.G., Mandal A.K., Reddy A.B.M., Gothwal V.K., & Hasnain S.E. (2004). Correlations of genotype with phenotype in Indian patients with primary congenital glaucoma. *Investigative Ophthalmology and Visual Science* **45**(4): 1149–1156.
DOI: <https://doi.org/10.1167/iovs.03-0404>
- Qashqai M., Suri F., Yaseri M.E. & Elahi E. (2018). P.Gly61Glu and P.Arg368His mutations in CYP1B1 that cause congenital glaucoma may be relatively frequent in certain regions of Gilan Province Iran. *Journal of Ophthalmic and Vision Research* **13**(4): 403–410.
DOI: https://doi.org/10.4103/jovr.jovr_147_17
- Rauf B., Irum B., Kabir F., Firasat S., Naeem M.A., Khan S.N., Husnain T., Riazuddin S., Akram J. & Riazuddin S.A. (2016). A spectrum of CYP1B1 mutations associated with primary congenital glaucoma in families of Pakistani descent. *Human Gene Variation* **3**: 16021.
- Safari I., Suri F., Haji-Seyed-Javadi R., Yazdani S. & Elahi E. (2016). The p.Gly61Glu mutation in CYP1B1 affects the extracellular matrix in glaucoma patients. *Ophthalmic Research* **56**(2): 98–103.
- Taqi U., Fasih U., Jafri S.F.A. & Sheikh A. (2011). Frequency of primary open angle glaucoma in Abbasi Shaheed Hospital. *Journal of Pakistan Medical Association* **61**(8): 778–781.
- Wiggs J.L. & Pasquale L.R. (2017). Genetics of glaucoma. *Human Molecular Genetics* **26**(R1): R21–R27.
DOI: <https://doi.org/10.1093/hmg/ddx184>
- Youngblood H., Hauser M.A. & Liu Y. (2019). Update on the genetics of primary open-angle glaucoma. *Experimental Eye Research* **188**:107795.
DOI: <https://doi.org/10.1016/j.exer.2019.107795>

RESEARCH ARTICLE

Environmental Toxicology

Effect of cylindrospermopsin on the hepatotoxicity in wistar rats

HASN Abeyasiri^{1,2}, JKP Wanigasuriya³, TS Suresh⁴, DH Beneragama⁵ and PM Manage^{1,2*}

¹ Centre for Water Quality and Algae Research, Department of Zoology, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda.

² Faculty of Graduate Studies, University of Sri Jayewardenepura, Gangodawila, Nugegoda.

³ Centre for Kidney Research, Department of Medicine, Faculty of Medical Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda.

⁴ Department of Biochemistry, Faculty of Medical Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda.

⁵ Department of Pathology, Faculty of Medical Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda.

Submitted: 25 October 2021; Revised: 14 February 2022; Accepted: 25 March 2022

Abstract: A naturally-derived cyanotoxin, cylindrospermopsin (CYN), present in freshwater systems, poses a threat to human health due to potential hepatotoxicity. The present study aimed to determine the hepatotoxic effects of CYN on male Wistar rats. Following ethical clearance, 35 rats were divided into five groups. Test groups were administered with pre-prepared solutions to provide three doses of CYN: 0.175 µg/kg, 0.140 µg/kg, and 0.105 µg/kg. Well-water collected from Padaviya (0.161 µg/kg of CYN) was given to the environmental-exposure group (EN) and distilled water was administered to the control. The total duration of exposure was 90 days. Blood samples were collected at 0, 7, 14, 28, 42, 60, 90 days. Aspartate amino-transferase (AST), aspartate alanine-transferase (ALT), and full blood count (FBC) were analyzed. At 90 days, hepatic tissue samples were taken for histology. The mean body weight of the treated and control groups of rats gradually increased until the 90th day. A statistically significant reduction ($p < 0.05$) in increment of body weights was observed in CYN-treated rats, compared to the control, at 12th and 13th weeks. Relative weights of the livers of treated groups were significantly lower ($p < 0.05$) than those of the control group. The highest AST and ALT concentrations were recorded in rats given a CYN dose of 0.175 µg/kg. Haematology revealed a significant ($p < 0.05$) reduction in monocyte and lymphocyte counts in the EN group given well-water compared to the control. Ballooning degeneration, Kupffer cell hyperplasia, lobular haemorrhage and necrosis, perivenular inflammation, and sinusoidal congestion were observed in the livers of treated groups. The findings show that prolonged exposure to CYN contaminated water leads to hepatotoxicity in Wistar rats.

Keywords: Cylindrospermopsin (CYN), hepatotoxicity, histopathology, Wistar rats.

INTRODUCTION

Cyanobacteria are a group of over 2000 prokaryotic organisms commonly named 'blue-green algae' (Woese, 2002). Several genera of cyanobacteria produce two major toxin groups, the microcystins (MCs) and cylindrospermopsin (CYN). The *Microcystis* spp. and *Cylindrospermopsis* spp. of cyanobacteria grow well in a wide range of fresh, brackish, and marine habitats (Sethunga & Manage, 2010; Kulasoorya, 2011), dug wells (Abeyasiri *et al.*, 2018a, 2018b; Abeyasiri & Manage, 2018), terrestrial ecosystems (Codd *et al.*, 2005), extreme habitats of hypersaline localities (Chatchawan *et al.*, 2011), hot springs (Wijesekara & Manage, 2017; Sadeepa *et al.*, 2019a; 2019b) and arid deserts (Kulasooriya, 2011). Approximately 40 cyanobacteria species have been recorded as potential producers of cyanotoxins (Merel *et al.*, 2013; Paerl *et al.*, 2013). Cyanotoxins are secondary metabolites categorized chemically as cyclic peptides and alkaloids with a stable chemical structure (Zanchett & Oliveira-Filho, 2013). These are recognized as a potential hazard in drinking water worldwide, resulting in organ toxicity in humans and animals (Falconer & Humpage, 2005).

* Corresponding author (pathmalal@sjp.ac.lk;  <https://orcid.org/0000-0003-3876-309X>)



CYN, a cyanotoxin contaminant of water supplies, is produced by freshwater, non-bloom-forming filamentous cyanobacteria; *Cylindrospermopsis raciborskii*, *Anabaena bergii*, *Anabaena lapponica*, *Lyngbya wollei*, *Aphanizomenon flos-aquae*, *Aphanizomenon ovalisporum*, *Umezaka natans*, and *Raphidiopsis curvata* (Falconer & Humpage, 2005; Sethunga & Manage, 2010; Mazmouz *et al.*, 2011). *C. raciborskii* and *Anabaena* sp. are abundantly present in lakes, reservoirs, ponds, and rivers in many tropical and subtropical areas globally, with increased colonization recorded in temperate regions of Australia, Europe, Japan, India, Sri Lanka, and South America (Sethunga & Manage, 2010; Piyathilaka *et al.*, 2015; Wijewickrama & Manage, 2019). CYN is a water-soluble, stable, tricyclic guanidine combined with uracil by the hydroxyl bridge (Runnegar *et al.*, 2002) (Figure 1).

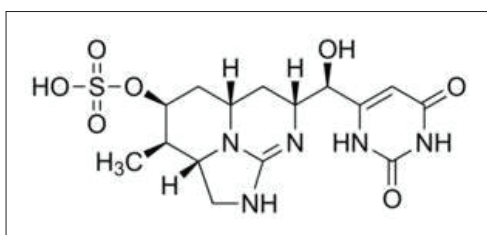


Figure 1: Chemical structure of cylindrospermopsin

The toxicity of CYN to the humans was first recorded by the “Palm Island Mystery Disease” of 1979, in Queensland, Australia, which affected 148 children and 10 adults, causing severe hepatotoxicity with renal tubular damage (Byth, 1980; Hawkins *et al.*, 1985; Ohtani *et al.*, 1992). Fatalities in haemodialysis patients due to the reverse osmosis unit of the dialysis water supply has been reported (Carmichael, 2001). *In vitro* animal experimental studies in rodents fed with MCs contaminated environmental water samples and extracts, or purified toxins of CYN have shown toxic effects on the vascular system (Hawkins *et al.*, 1985), lymphatic system (Hawkins *et al.*, 1985), reproductive system (Chen *et al.*, 2011), immune system (Lankoff *et al.*, 2004), and other organ damage: liver (Hawkins *et al.*, 1985; Manage *et al.*, 2009c; Drobac *et al.*, 2011), kidney (Hawkins *et al.*, 1985; Piyathilaka *et al.*, 2015; Abey Siri *et al.*, 2018a; Abey Siri & Manage, 2018a), lungs, heart, stomach, and adrenal glands (Hawkins *et al.*, 1985).

Thus, considering the toxic nature of CYN on human health, the World Health Organization (WHO) established a provisional guideline limit for CYN of 2 µg/L in drinking water (WHO, 2003; SLSI, 2013) and 0.02 µg/kg of body weight/day as the Tolerable Daily Intake (TDI) for humans (Guzman-Guillen *et al.*, 2014). Recently, there have been increasing reports on the number and the intensity of blooms in reservoirs in the North Central, North East, and Uva Provinces of Sri Lanka (Madhushankha *et al.*, 2013; Yatigamma & Perera, 2017; Manage, 2019). Further, scientific research has revealed contamination of the majority of drinking water reservoirs by toxin-producing cyanobacterium *M. aeruginosa*, *Cylindrospermopsis* sp., *Anabaena* sp. and a significant correlation between cyanobacterial cell density and the cyanotoxins in drinking water sources including dug wells (Hettiarachchi & Manage, 2014; Piyathilaka *et al.*, 2016; Manage, 2019). Therefore, the present study was aimed at evaluating the effects of subchronic exposure to different doses of purified CYN and CYN contaminated well water on the liver of male Wistar rats through oral gavage over 90 days.

MATERIALS AND METHODS

Ethical approval was obtained from the Ethics Review Committee of the Faculty of Medical Sciences (An.E.17/18), University of Sri Jayewardenepura. The study was carried out following the relevant guidelines and regulations given by the Ethics Review Committee.

Animals

Eight-week-old male Wistar Rats (183.2 ± 0.2 g) purchased from the Medical Research Institute (MRI) of Sri Lanka were acclimatized for one week at the Animal House, Faculty of Medical Sciences, University of Sri Jayewardenepura. The rats were provided with food and water *ad libitum* unless otherwise stated, and maintained on a 12 hour light/dark cycle at $28 \pm 2^\circ\text{C}$. Water and food consumption were measured, and the behaviour of the animals was observed throughout the study.

Compound

CYN used in these investigations was Lot # BCBS9482v supplied by Sigma-Aldrich (USA). The purified toxin ($\text{C}_{15}\text{H}_{21}\text{N}_5\text{O}_7\text{S}$; product no: 32087) had an estimated purity of >95%.

Experimental design

In this study, three CYN doses were prepared based on the WHO standard recommendation of CYN in drinking water: higher than the WHO standard (HW) 2.5 µg/L, equal to the WHO standard (W) 2.0 µg/L and lower than the WHO standard (LW) 1.5 µg/L. The pure CYN doses, according to the mean initial body weights of the rats fed with the above concentrations, were HW = 0.175 µg/kg/day, W = 0.140 µg/kg/day, and LW = 0.105 µg/kg/day. A single well water sample collected from the Padaviya area of the North Central Province of Sri Lanka contained 2.3 µg/L CYN, which was above the WHO recommended level. This was used as the environmental sample. Accordingly, the environmental exposure group (EN) received 0.161 µg/kg/day of CYN. Distilled water was given to the control group. The CYN solutions, well water, and distilled water were fed for 90 days. The mouse groups were labeled as HW, W, LW, EN, and control according to the solutions received by them. Rats behaviour, and water and food consumption were recorded at intervals of every two days throughout the experiment. Nutritionally complete solid food (Saboudry, 1988) was supplied *ad libitum* to avoid contaminated urine samples with food dust. The solutions of CYN were formulated biweekly and adjusted to reflect weight changes in the treated groups before administration. After 90 days of dosing, the animals were euthanized, weighed (g), and the liver was resected out with minimum trauma.

Collection of blood

Venous blood samples were drawn from the lateral tail vein at 0, 7, 14, 28, 42, and 60 days for clinical chemistry. At the end of the experiment (90 days), when the rats were under anaesthesia prior to euthanization, blood was collected by cardiac puncture for haematological and chemical analysis. Blood for haematological analysis was transferred to a 1 mL tube containing EDTA to prevent coagulation.

Clinical chemistry

Serum concentrations of aspartate amino-transferase (AST) and aspartate alanine-transferase (ALT) were analyzed using Bialabo diagnostic kits (France) and a fully automated analyzer (Thermo Fisher Scientific, INDIKO, Finland). All animals were used to measure AST and ALT at 0, 7, 14, 28, 42, 60, and 90 days.

Haematology

Full blood count (haemoglobin (Hb), red cell indices, platelet count, total and differential white blood cell counts) and reticulocyte count were analyzed using a fully automated 3 part haematology analyzer (BCC-3000, China).

Histopathological evaluation

At necropsy, the liver was observed for gross pathological changes, and the weight was measured. The liver was serially cut, and the cut surfaces were examined for any difference in colour, areas of necrosis or fibrosis, and photographed with the reference number. Two random slices were selected and immediately fixed in ten times the volume of the slices in 10% formalin for routine histological assessment. If any abnormal areas were present, additional slices were fixed similarly for the histological evaluation. The changes observed were recorded in the datasheet. Following processing and dehydration, prepared slides were transferred to xylene for 5 min., absolute alcohol for 5 min., 90% alcohol for 2 min., 80% alcohol for 2 min., 70% alcohol for 2 min., and 60% alcohol for 2 min., and stained with hematoxylin and eosin; special stains used included periodic acid – Schiff stain (PAS) and silver special stains. Masson trichrome to assess fibrosis was not used due to the short duration of the experimental period.

For histological changes, the prepared sections were examined under the light microscope (Olympus CX31, magnification ×40, ×100, and ×400). The scoring system consisted of evaluating the tissues and assigning a semi-quantitative score scheme, from 0 = no lesion to 3 = severe lesions, according to the changes observed in the liver architecture, hepatocytes, and the presence of inflammation. After an initial review, selected tissues were re-assessed.

Statistical evaluation

One-way ANOVA was used to statistically evaluate significant differences, by using MINITAB version 17 statistical software (MINITAB, State College, PA, USA), and differences were considered significant if $p < 0.05$.

RESULTS AND DISCUSSION

Clinical findings

The pure CYN doses received by the rats in different groups according to the mean initial body weights were HW = 0.175 µg/kg, W = 0.140 µg/kg, LW = 0.105 µg/kg and EN = 0.161 µg/kg.

The external appearance and signs of toxicity were not observed in CYN-treated animals during the study. Similar findings were reported by Diez-Quijada *et al.* (2021); clinical and visual observations during the study period did not show abnormalities in the groups studied. In general, there were no changes in behaviour or locomotor activity during the study period due to repeated exposure to CYN. Gradual mean body weight increments were noted in both treated (190.4 ± 3.8 – 300.3 ± 2.2 g) and

control (198.4 ± 1.6 – 317.7 ± 3.5 g) groups of animals until the thirteenth week (Figure 2). However, there was a statistically significant reduction (p < 0.05) in body weights of rats in the W, HW, and EN groups given CYN when compared to the control group at weeks 12 and 13. Comparable results were reported by Chernoff *et al.* (2018) after the 90-day gavage of CYN in a mice experiment. Increased body weight with low toxin doses (30 and 60 µg/kg/day) and decrease of body weight at high doses (432 and 657 µg/kg/day) of CYN in mice had been reported in a previous study (Humpage and Falconer, 2003). In another study, no statistically significant differences were observed in body weight between treated and control groups in males (Diez-Quijada *et al.*, 2021). CYN is known to affect both hepatic and renal systems (Seawright *et al.*, 1999; Humpage & Falconer, 2003; Bazin *et al.*, 2012), and data presented here specify hepatic toxicity at four dose levels.

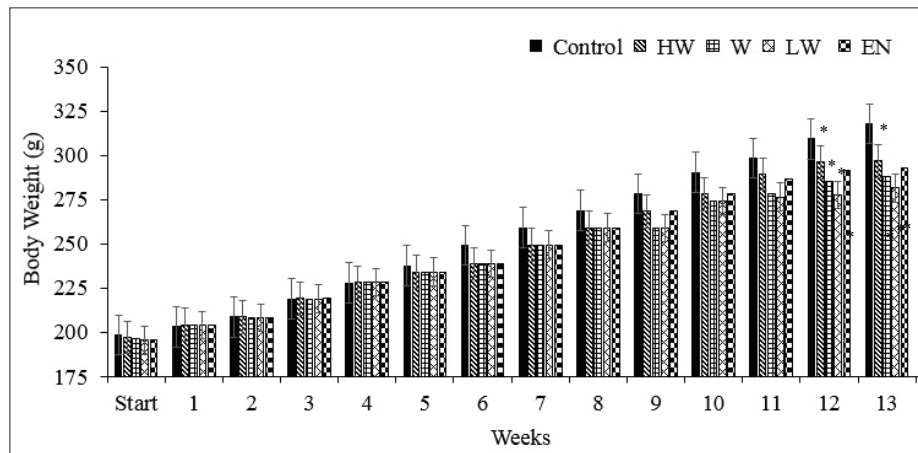


Figure 2: Mean body weight changes of Wistar rats in different treatment groups (HW = 0.175 µg/kg, W = 0.140 µg/kg, LW = 0.105 µg/kg, EN = 0.161 µg/kg) *p < 0.05

Table 1: Bodyweight (g) and relative percentages of the liver weight of male Wistar rats exposed to different doses of CYN for 90 days. Values are mean ± SD for 7 rats/group. Statistics ANOVA test; *p < 0.05

Parameters	Control		Pure CYN dose (µg/kg)		CYN environmental dose (µg/kg)
	0 N = 7	0.105 (LW) N = 7	0.140 (W) N = 7	0.175 (HW) N = 7	0.161(EN) N = 7
Body weight (g)	318 ± 0.98	284 ± 1.03 *	289 ± 1.26 *	300 ± 2.42 *	296 ± 1.33 *
Absolute weight (g)	9.02 ± 0.63	8.62 ± 0.56	10.41 ± 0.70*	9.27 ± 0.94	9.27 ± 1.30*
% Liver	3.19 ± 0.13	2.87 ± 0.03*	2.87 ± 0.05*	2.75 ± 0.02*	2.78 ± 0.01*

There were no gross pathological changes observed in the livers. The liver weights as a percentage of the body weight showed a significant decrease in treatments groups: 284 ± 1 g in LW ($p < 0.05$), 289 ± 1 g in W ($p < 0.05$), 300 ± 2 g in HW ($p < 0.05$), and 296 ± 1 g in EN ($p < 0.05$) in comparison to the control group (Table 1). No significant differences were found in the absolute and relative weights of males at any toxin concentration in the study by Diez-Quijada *et al.* (2021). However, Reisner *et al.*, 2004 found an increase in the relative weights of the livers in male mice exposed to CYN.

The AST level gradually increased up to the endpoint of the study and the differences were statistically significant in HW ($p < 0.05$), W ($p < 0.05$), LW ($p < 0.05$), and EN ($p < 0.05$) groups compared to the control at day 7, 14, 28, 42, 60, and 90 (Figure 3a).

Significant differences of the ALT activity was found in HW ($p < 0.05$), W ($p < 0.05$), LW ($p < 0.05$) and EN ($p < 0.05$) groups compared to the control at day 7, 14, 28, 42, 60, and 90 (Figure 3b). Increased serum levels of AST and ALT activity indicated hepatic toxicity, which was compatible with changes observed on histology. Elevated ALT levels in mice with a high dose of CYN exposure were reported by other investigators (Chernoff *et al.*, 2011; 2014). The AST activity was also elevated after oral exposure at the intermediate and highest dose groups (37.5 and 75 $\mu\text{g}/\text{kg}$ CYN) to the control group ($p < 0.001$) (Diez-Quijada *et al.*, 2021). The data obtained from Dordevic *et al.* (2017) showed that the alterations in serum biochemical parameters did not produce significant liver damage in rats. Moreover, increased AST activity was found in females at all doses assayed (Chernoff *et al.*, 2018).

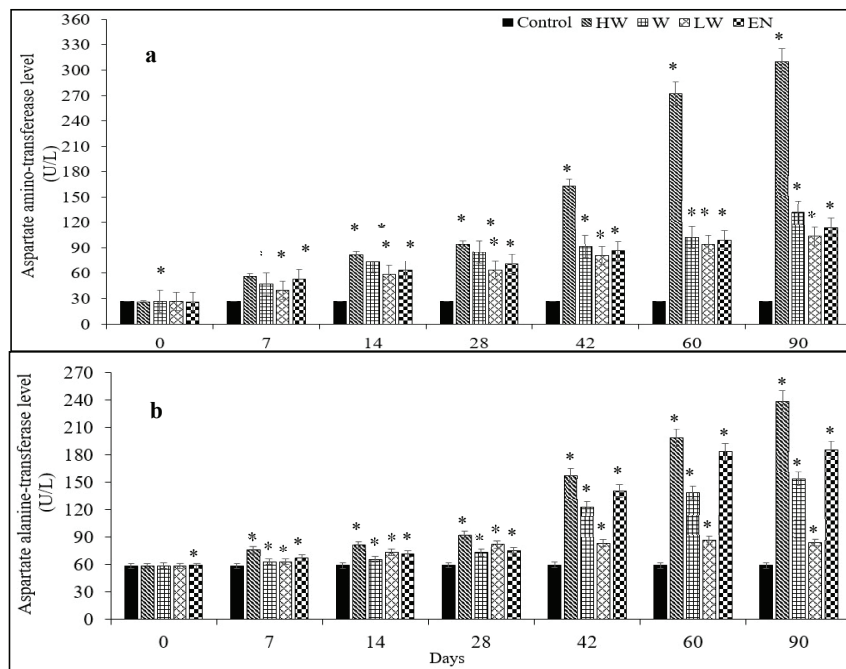


Figure 3: (a) Mean serum aspartate amino-transferase (U/L), (b) Mean serum as partate alanine-transferase (U/L) concentrations changes of male Wistar Rats in different treatment groups (HW= 0.175 $\mu\text{g}/\text{kg}$, W= 0.140 $\mu\text{g}/\text{kg}$, LW= 0.105 $\mu\text{g}/\text{kg}$, EN=0.161 $\mu\text{g}/\text{kg}$)
* $p < 0.05$

Haematology

Haematological parameters measured at day 90 are summarized in Table 2.

Non-significant reductions in both haematocrit and haemoglobin levels were reported in the HW, W, and EN groups. In the EN group, a significant decrease in lymphocyte count ($p < 0.05$) and monocyte count

($p < 0.05$) were observed in comparison to the control group. The absence of similar changes in the groups treated with pure CYN extracts makes it unlikely that the haematological changes can be attributed to the toxic effects of CYN. However, the rats in the EN group received the highest dose of CYN. Augmented haematocrit levels were reported as a response to exposure to low levels of CYN and morphological changes in the RBC of mice, which gradually developed into an acanthocyte-like form, cytoplasmic irregularly spaced projections, etc. (Reisner *et al.*, 2004). Similarly,

elevated haematocrit levels in male and female mice after 16 weeks of exposure to variable concentrations of CYN in their drinking water (10–55 $\mu\text{g}/\text{kg}/\text{day}$ during 42 weeks) were also found as the most critical effect observed (Sukenik *et al.*, 2006). By contrast, all the values remained within the normal range for the strain (Lira *et al.*, 2020). There were no significant alterations in the differential white blood cell counts of rats exposed to CYN in comparison to the untreated groups (Diez-Quijada *et al.*, 2021).

Table 2: Selected haematology parameters of Wistar Rats on the 90th day of the experiment
Statistics ANOVA test; * $p < 0.05$

Parameters	HW (0.175 $\mu\text{g}/\text{kg}$)	W (0.140 $\mu\text{g}/\text{kg}$)	LW (0.105 $\mu\text{g}/\text{kg}$)	EN (0.161 $\mu\text{g}/\text{kg}$)	Control
Haematocrit (%)	23.59 \pm 1.18	22.27 \pm 0.71	24.58 \pm 0.96	22.93 \pm 1.20	24.00 \pm 1.14
Haemoglobin (g/dL)	16.23 \pm 1.02	15.68 \pm 0.25	17.07 \pm 0.68	16.13 \pm 0.89	16.63 \pm 0.97
Reticulocyte count ($10^{12}/\text{L}$)	8.85 \pm 0.35	8.19 \pm 0.29	9.06 \pm 0.23	8.55 \pm 0.56	8.83 \pm 0.50
Mean cell volume (fL)	106.54 \pm 1.96	108.77 \pm 3.09	108.47 \pm 4.07	107.07 \pm 1.41	108.93 \pm 3.74
MCHC (g/dL)	69.11 \pm 0.74	70.46 \pm 1.38	69.43 \pm 0.41	70.53 \pm 0.61	69.27 \pm 1.00
PLT ($10^9/\text{L}$)	689.1 \pm 112.3	682.9 \pm 82.6	754.3 \pm 118.4	682.7 \pm 97.2	622.3 \pm 184.3
Lymphocyte count ($10^9/\text{L}$)	6.80 \pm 2.58	6.06 \pm 1.49	9.33 \pm 2.81	3.50 \pm 1.51*	7.03 \pm 2.71
Monocyte count ($10^9/\text{L}$)	1.31 \pm 1.65	0.74 \pm 0.22	0.77 \pm 0.27	0.40 \pm 0.18*	0.67 \pm 0.24
Granulocyte count ($10^9/\text{L}$)	1.51 \pm 2.00	0.57 \pm 0.21	0.83 \pm 0.29	0.37 \pm 0.15	0.47 \pm 0.24

Table 3: Incidence and severity of histopathologic liver lesions in Wistar rats on the day 90 of exposure to cylindrospermopsin
Lesion severity was based on a 0 – 3 grading scale

Histological features	HW (0.175 $\mu\text{g}/\text{kg}$)	W (0.140 $\mu\text{g}/\text{kg}$)	LW (0.105 $\mu\text{g}/\text{kg}$)	EN (0.161 $\mu\text{g}/\text{kg}$)	Control
Hepatocytes	Ballooning (3)	Ballooning (2)	Ballooning (1)	Ballooning (3)	Normal (0)
Inflammation-lobular	Mild/spotty inflammation < 5/lpf (1)	Mild/spotty inflammation < 5/lpf (1)	Mild/spotty inflammation < 5/lpf (1)	Mild/spotty inflammation < 5/lpf (1)	Normal (0)
Hepatocyte cell death	Lobular/spotty (2)	Lobular/spotty (2)	Lobular/spotty (2)	Lobular/spotty (2)	Normal (0)
Hepatocyte regeneration	Bi-nucleation (2)	Normal (0)	Normal (0)	Bi-nucleation (2)	Normal (0)
Kupffer cells	Prominent (3)	Mild (1)	Mild (1)	Prominent (3)	Normal (0)
Nuclear pyknosis	Mild (1)	Mild (1)	Mild (1)	Mild (1)	Normal (0)
Haemorrhage	Mild (1)	Normal (0)	Normal (0)	Mild (1)	Normal (0)

Histopathology

Histopathological changes of the liver associated with CYN exposure are summarized in Table 3. These changes

include hepatocyte ballooning, focal inflammation nuclear pyknosis, hepatocyte cell death/necrosis, features of hepatocyte regeneration, Kupffer cell hyperplasia, and haemorrhages (Table 3).

The hepatocytes showed ballooning degeneration (Figure 4b), Kupffer cell hyperplasia (Figure 4c), lobular haemorrhage and necrosis (Figure 4d) in EN (0.161 µg/kg). Lobular haemorrhage and necrosis (Figure 5b), lobular inflammation (Figure 5c), lobular haemorrhage (Figure 5d), sinusoidal congestion (Figure 5e), and centrilobular haemorrhage and necrosis (Figure 5f) were prominent in HW (0.175 µg/kg) group. Lobular inflammation and necrosis were observed in W (0.140 µg/kg) (Figure 6b). Lobular inflammation (Figure 7b) was significant in the LW (0.105 µg/kg) group. All histological features were typical in the control group (Figure 4a, 5a, 6a, 7a). In this study, hepatic lesions were seen in all treated groups. The effects were more severe in the HW and EN groups. The well water fed to the EN group contained CYN well above the safe level recommended by the WHO. These results were similar to the findings in 50 µg/kg doses of CYN administered for 5 days which produced a widespread disruption of hepatic cords, hepatocellular degeneration, necrosis, and centrilobular inflammation, as reported by Chernoff (Chernoff *et al.*, 2011; 2014). In CYN-exposed cells in a culture, prominent features included necrosis and/or apoptosis of hepatocytes possibly

resulting from a disruption of the cytoskeleton (Fessard & Bernard, 2003). Rats exposed by gavage to pure CYN (7.5–75.0 µg/kg b.w.) at 0, 24, and 45 h, showed irritation in the stomach and changes in hepatocytes (Diez-Quijada *et al.*, 2019).

Thus, Bernard *et al.* (2003) tested *in vivo* the toxicity of cell extracts of different *C. raciborskii* strains in mice and found fatal acute neurotoxicity, hepatotoxicity, and the absence of toxicity, depending on the toxins and the toxicological activity of unknown compounds detected. In the 2021 study by Leticia *et al.*, at a dose of 75.0 µg/kg b.w, only scarce microscopic changes were observed in the livers of rats. Livers displayed a very mild hepatic sinusoidal ectasia and small numbers of hepatocytes with apoptotic features or spotty necrosis, mainly localized in the surroundings of the centrilobular vein. Occasionally, hepatocytes presented eosinophilic cytoplasmic inclusions (Leticia *et al.*, 2021). The present study further confirmed the signs of hepatic injuries due to CYN. Liver/body weight ratios, the elevation of AST and ALT levels, and histopathological changes were important parameters of CYN-induced liver toxicity.

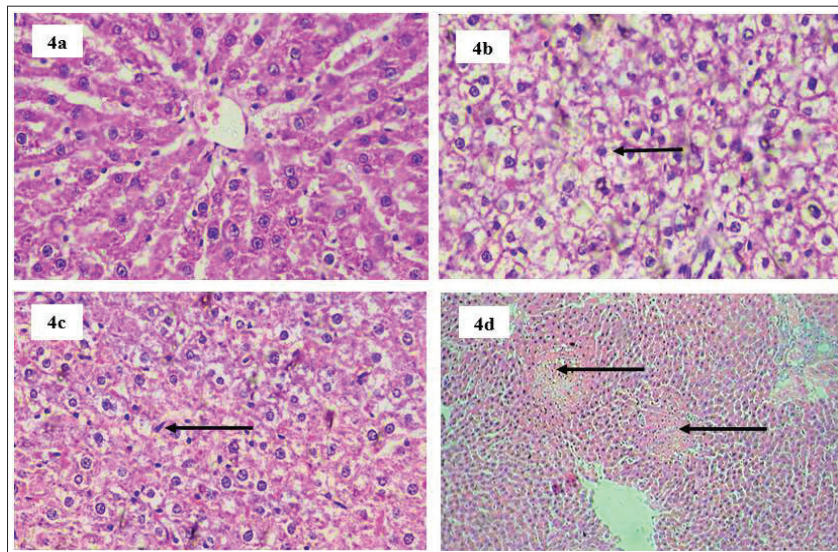


Figure 4: Environmental exposure (EN) Haematoxylin and eosin stain. Figure 4a - Control ×400; Figure 4b - Ballooning degeneration of hepatocytes ×400; Figure 4c - Kupffer cell hyperplasia (arrow) ×400; Figure 4d - Lobular haemorrhage and necrosis (arrows) ×100

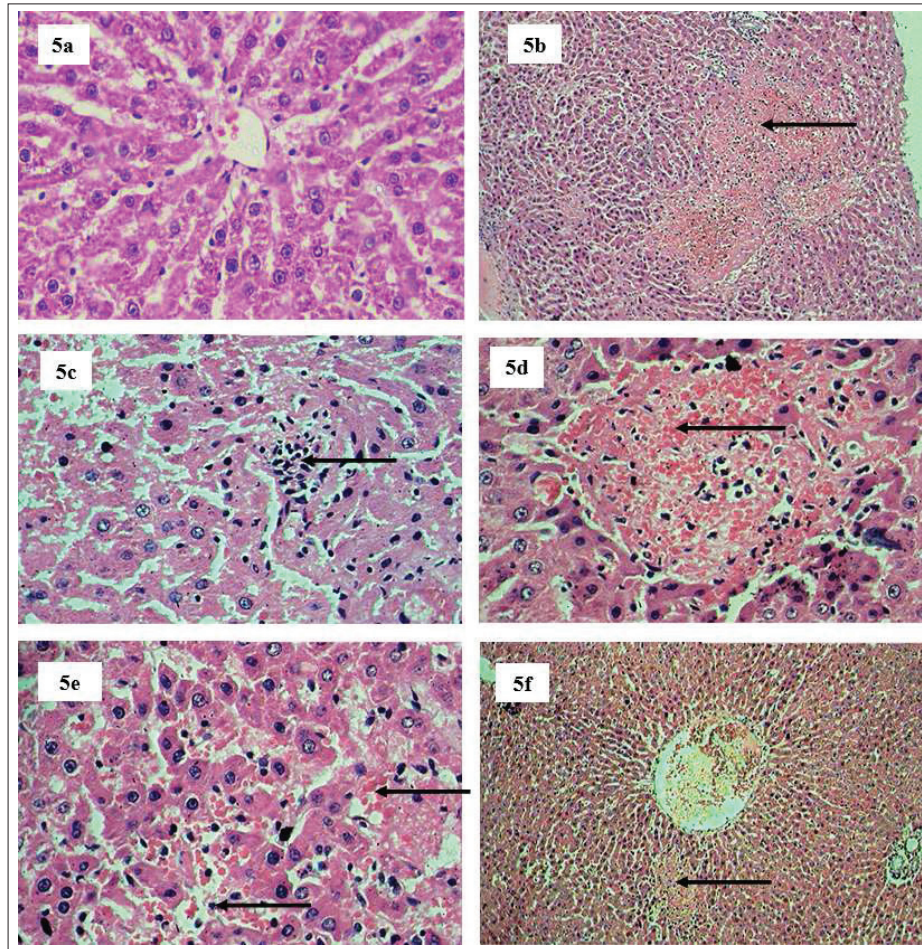


Figure 5: Above WHO level (HW) Haematoxylin and eosin stain
 Figure 5a - Control $\times 400$; Figure 5b - Lobular haemorrhage and necrosis $\times 100$; Figure 5c - Lobular inflammation (arrow) $\times 400$; Figure 5d - Focus of lobular haemorrhage $\times 400$; Figure 5e - Sinusoidal congestion $\times 400$; Figure 5f - Centrilobular haemorrhage and necrosis shown by the arrow $\times 100$

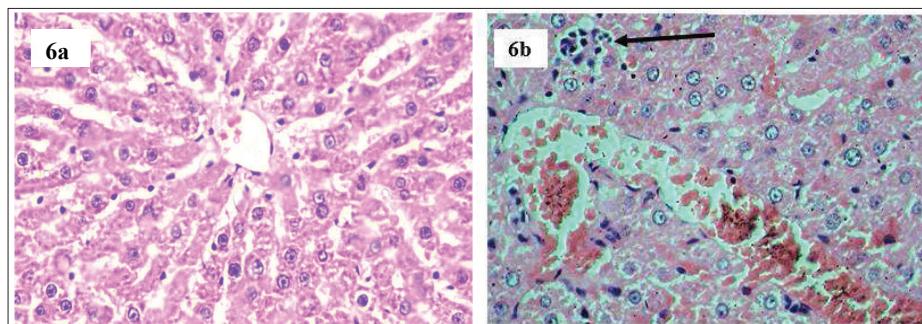


Figure 6: WHO level (W) Haematoxylin and eosin stain
 Figure 6a Control $\times 400$; Figure 6b - Lobular inflammation and necrosis (Arrow) $\times 400$

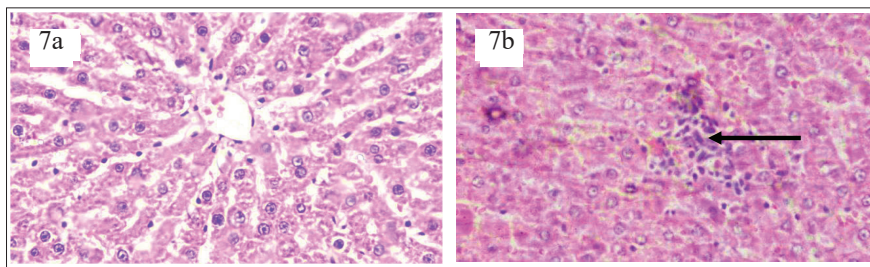


Figure 7: Below WHO level (LW). Haematoxylin and eosin stain
Figure 7a - Control $\times 400$; Figure 7b - LW - Lobular Inflammation (Arrow) $\times 400$

CONCLUSION

The study results showed that the differences in liver toxicity were associated with the administered CYN dose. It provides substantial evidence that CYN contaminated water for drinking purposes may lead to mammalian liver cell injury.

Conflict of interest statement

The authors declare that they have no competing interests

Acknowledgments

The Authors thank the National Research Council, Sri Lanka (Grant No. NRC-16-078), University Research Grant (Grant No. ASP/01/RE/SCI/2019/29), Algae and Water Quality Research Centre and Centre for Kidney Research, and the University of Sri Jayewardenepura for the financial support. Mrs. M.D.S.S. Jayathilake, Department of Pathology, Faculty of Medical Sciences, the University of Sri Jayewardenepura for technical assistance; National Centre for Primary Care and Allergy Research, Department of Family Medicine, University of Sri Jayewardenepura, S.P. Manage, I. Abinaiyan, K.T. Dilrukshi and G.D. Ramanayake for assistance.

REFERENCES

Abeyesiri H.A.S.N. & Manage P.M. (2018). Detection of Cylindrospermopsin in well water in CKDu endemic Padaviya, Medawachciya and Kebithigollewa area in Anuradhapura District, Sri Lanka. *Proceedings of the 6th International Symposium of Water Quality and Human Health: Challenges Ahead*, 13 – 14 July. PGIS, University of Peradeniya, Peradeniya, Sri Lanka.

Abeyesiri H.A.S.N., Wanigasuriya K. & Manage P.M. (2018a). Cylindrospermopsin and Microcystin-LR contaminated human urine in CKDu high prevalence Medirigiriya in Polonnaruwa District, Sri Lanka. *Proceedings of the International Conference of Chronic Kidney Disease of Unknown Aetiology: Current Aspects and Future Prospects*, 8 – 9 December. University of Peradeniya, Peradeniya, Sri Lanka, pp. 10.

Abeyesiri H.A.S.N., Wanigasuriya K. & Manage P.M. (2018b). Detection of Cylindrospermopsin and Microcystin-LR in well water of CKDu prevalent Medirigiriya, Sri Lanka, Sri Lanka. *Proceedings of the 3rd International Research Symposium of Pure and Applied Sciences*, 26 October. Faculty of Science, University of Kelaniya, Sri Lanka, pp. 17.

Bazin E., Huet S., Jarry G., Hégarat L.L., Munday J.S., Humpage A.R. & Fessard V. (2012). Cytotoxic and genotoxic effects of cylindrospermopsin in mice treated by gavage or intraperitoneal injection. *Environmental Toxicology* **27**(5): 277–284.

DOI: <https://doi.org/10.1002/tox.20640>

Bernard C., Harvey M., Briand J.F., Biré R., Krys S. & Fontaine J.J. (2003). Toxicological comparison of diverse Cylindrospermopsis raciborskii strains: evidence of liver damage caused by a French C. raciborskii strain. *Environmental Toxicology: An International Journal* **18**(3): 176–186.

DOI: <https://doi.org/10.1002/tox.10112>

Byth S. (1980). Palm Island mystery disease. *Medical Journal of Australia* **2**(1): 40–42.

DOI: <https://doi.org/10.5694/j.1326-5377.1980.tb131814.x>

Carmichael W.W. (2001). Health effects of toxin-producing cyanobacteria: 'The CyanoHABs'. *Human and Ecological Risk Assessment* **7**(5): 1393–1407.

DOI: <https://doi.org/10.1080/20018091095087>

Chatchawan T., Peerapornpisal Y. & Komárek J. (2011). Diversity of cyanobacteria in man-made solar saltern, Petchaburi Province, Thailand—a pilot study. *Fottea* **11**(1):203–14.

- Chen Y., Xu J., Li Y. & Han X. (2011). Decline of sperm quality and testicular function in male mice during chronic low-dose exposure to microcystin-LR. *Reproductive Toxicology* **31**(4): 551–557.
DOI: <https://doi.org/10.1016/j.reprotox.2011.02.006>
- Chernoff et al. (11 authors) (2018). Cylindrospermopsin toxicity in mice following a 90-d oral exposure. *Journal of Toxicology and Environmental Health, Part A* **81**(13): 549–566.
DOI: <https://doi.org/10.1080/15287394.2018.1460787>
- Chernoff et al. (11 authors) (2011). Toxicity and recovery in the pregnant mouse after gestational exposure to the cyanobacterial toxin, cylindrospermopsin. *Journal of Applied Toxicology* **31**(3): 242–254.
DOI: <https://doi.org/10.1002/jat.1586>
- Chernoff N., Rogers E.H., Zehr R.D., Gage M.I., Travlos G.S., Malarkey D.E., Brix A., Schmid J.E. & Hill D. (2014). The course of toxicity in the pregnant mouse after exposure to the cyanobacterial toxin cylindrospermopsin: clinical effects, serum chemistries, hematology, and histopathology. *Journal of Toxicology and Environmental Health, Part A* **77**(17): 1040–1060.
DOI: <https://doi.org/10.1080/15287394.2014.919838>
- Codd G.A., Morrison L.F. & Metcalf J.S. (2005). Cyanobacterial toxins: risk management for health protection. *Toxicology and Applied Pharmacology* **203**(3): 264–272.
DOI: <https://doi.org/10.1016/j.taap.2004.02.016>
- Diez-Quijada L., Moyano R., Molina-Hernández V., Cameán A.M. & Jos Á. (2021). Evaluation of toxic effects induced by repeated exposure to Cylindrospermopsin in rats using a 28-day feeding study. *Food and Chemical Toxicology* **151**: 112108.
DOI: <https://doi.org/10.1016/j.fct.2021.112108>
- Diez-Quijada L., Llana-Ruiz-Cabello M., Cătușescu G.M., Puerto M., Moyano R., Jos A., & Cameán A.M. (2019). In vivo genotoxicity evaluation of cylindrospermopsin in rats using a combined micronucleus and comet assay. *Food and Chemical Toxicology* **132**: 110664.
DOI: <https://doi.org/10.1016/j.fct.2019.110664>
- Đorđević N.B., Matić S.L., Simić S.B., Stanić S.M., Mihailović V.B., Stanković N.M., Stanković V.D. & Ćirić, A.R. (2017). Impact of the toxicity of Cylindrospermopsis raciborskii (Woloszynska) Seenayya & Subba Raju on laboratory rats in vivo. *Environmental Science and Pollution Research* **24**(16): 14259–14272.
- Drobac D., Svirčev Z., Tokodi N., Vidović M., Baltić V., Božić-Krstić V., Lazić D. & Pavlica T. (2011). Microcystins: Potential risk factors in carcinogenesis of primary liver cancer in Serbia. *Geographica Pannonica* **15**(3): 70–80.
DOI: [10.5937/GeoPan1103070D](https://doi.org/10.5937/GeoPan1103070D)
- Falconer I.R. & Humpage A.R. (2005). Health risk assessment of cyanobacterial (blue-green algal) toxins in drinking water. *International Journal of Environmental Research and Public Health* **2**(1): 43–50.
DOI: <https://doi.org/10.3390/ijerph2005010043>
- Fessard V. & Bernard C. (2003). Cell alterations but no DNA strand breaks induced in vitro by cylindrospermopsin in CHO K1 cells. *Environmental Toxicology* **18**(5): 353–359.
DOI: <https://doi.org/10.1002/tox.10136>
- Guzmán-Guillén R., Prieto A.I., Moreno I., Rios V., Vasconcelos V.M. & Cameán A.M. (2014). Effects of depuration on oxidative biomarkers in tilapia (*Oreochromis niloticus*) after subchronic exposure to cyanobacterium producing cylindrospermopsin. *Aquatic Toxicology* **149**: 40–49.
DOI: <https://doi.org/10.1016/j.aquatox.2014.01.026>
- Hawkins P.R., Runnegar M.T., Jackson A.R. & Falconer I.R. (1985). Severe hepatotoxicity caused by the tropical cyanobacterium (blue-green alga) *Cylindrospermopsis raciborskii* (Woloszynska) Seenaya and Subba Raju isolated from a domestic water supply reservoir. *Applied and Environmental Microbiology* **50**(5): 1292–1295.
DOI: <https://doi.org/10.1128/aem.50.5.1292-1295.1985>
- Hettiarachchi I.U. & Manage P.M. (2014). Cyanobacterial cell density & intracellular Microcystin-LR levels in drinking/ irrigation reservoirs in Anuradhapura, Sri Lanka. *Proceedings of the of Symposium: Global Climate Change and Sustainability Pathways*, 6 – 7 November. Thailand, pp.42.
- Humpage A.R. & Falconer I.R. (2003). Oral toxicity of the cyanobacterial toxin cylindrospermopsin in male Swiss albino mice: determination of no observed adverse effect level for deriving a drinking water guideline value. *Environmental Toxicology* **18**(2): 94–103.
DOI: <https://doi.org/10.1002/tox.10104>
- Kulasooriya S.A. (2011). Cyanobacteria: pioneers of planet earth. *Ceylon Journal of Science (Biological Sciences)* **40**(2): 71–88.
- Lankoff A., Carmichael W.W., Grasman K.A. & Yuan M. (2004). The uptake kinetics and immunotoxic effects of microcystin-LR in human and chicken peripheral blood lymphocytes in vitro. *Toxicology* **204**(1): 23–40.
DOI: <https://doi.org/10.1016/j.tox.2004.05.016>
- Lira A.C., Prieto A.I., Baños A., Guillamón E., Moyano R., Jos A. & Cameán A.M. (2020). Safety assessment of propyl-propane-thiosulfonate (PTSO): 90-days oral subchronic toxicity study in rats. *Food and Chemical Toxicology* **144**: 111612.
DOI: <https://doi.org/10.1016/j.fct.2020.111612>
- Madhushankha L., Dhammika M.A. & Naduwiladath C. (2013). Identification of Cylindrospermopsin and *Cylindrospermopsis raciborskii* from Anuradhapura District, Sri Lanka. *Journal of Ecotechnology Research* **17**(1): 23–28.
DOI: <https://doi.org/10.11190/jer.17.23>
- Manage P. (2019). Cyanotoxins: A hidden cause of Chronic Kidney Disease of unknown etiology (CKDu) in Sri Lanka-A review. *Sri Lanka Journal of Aquatic Sciences* **24**(1): 1–10.
- Manage P.M., Yasawardna S.G. & Wedage W.S. (2009(c)). Hepatotoxic effects of *Microcystis aeruginosa* (PCC7820) on Wistar Rats, *Golden Jubilee Special Issue of the Vidyodaya Journal* 2019: 23–26.
DOI: <https://doi.org/10.31357/fhss/vjhss.v00i00.4028>
- Mazmouz R., Chapuis-Hugon F., Pichon V., Méjean A. & Ploux O. (2011). The last step of the biosynthesis of the cyanotoxins cylindrospermopsin and 7-epi-

- cylindrospermopsin is catalysed by CyrI, a 2-Oxoglutarate-dependent iron oxygenase. *ChemBioChem* **12**(6): 858–862. DOI: <https://doi.org/10.1002/cbic.201000726>
- Merel S., Walker D., Chicana R., Snyder S., Baurès E. & Thomas O. (2013). State of knowledge and concerns on cyanobacterial blooms and cyanotoxins. *Environment International* **59**: 303–327. DOI: <https://doi.org/10.1016/j.envint.2013.06.013>
- Ohtani I., Moore R.E. & Runnegar M.T. (1992). Cylindrospermopsin: a potent hepatotoxin from the blue-green alga *Cylindrospermopsis raciborskii*. *Journal of the American Chemical Society* **114**(20): 7941–7942. DOI: <https://doi.org/10.1021/ja00046a067>
- Paerl H.W. & Otten T.G. (2013). Harmful cyanobacterial blooms: causes, consequences, and controls. *Microbial Ecology* **65**(4): 995–1010. DOI: <https://doi.org/10.1007/s00248-012-0159-y>
- Piyathilaka M.A., Pathmalal M.M., Tennekoon K.H., De Silva B.G., Samarakoon S.R. & Chanthirika S. (2015). Microcystin-LR-induced cytotoxicity and apoptosis in human embryonic kidney and human kidney adenocarcinoma cell lines. *Microbiology* **161**(4): 819–828. DOI: <https://doi.org/10.1099/mic.0.000046>
- Piyathilaka M.A.P.C., Tennekoon K.H., De Silva B.G.D.N.K. & Manage P.M. (2016). Microcystin contamination and potential of microcystin production in major drinking water bodies of Sri Lanka. *Proceedings of the 21st International Symposium of Forestry and Environment Science*, 23-24 September. Department of Forestry and Environmental Science, University of Sri Jayewardenepura, Sri Lanka, pp. 60.
- Reisner M., Carmeli S., Werman M. & Sukenik A. (2004). The cyanobacterial toxin cylindrospermopsin inhibits pyrimidine nucleotide synthesis and alters cholesterol distribution in mice. *Toxicological Sciences* **82**(2): 620–627. DOI: <https://doi.org/10.1093/toxsci/kfh267>
- Runnegar M.T., Xie C., Snider B.B., Wallace G.A., Weinreb S.M. & Kuhlenkamp J. (2002). *In vitro* hepatotoxicity of the cyanobacterial alkaloid cylindrospermopsin and related synthetic analogues. *Toxicological Sciences* **67**(1): 81–87. DOI: <https://doi.org/10.1093/toxsci/67.1.81>
- Saboudry M.A. (1988). *Breeding and care of Laboratory Animals*. WHO, Health Laboratory Technology Unit. Geneva, Switzerland.
- Sadeepa H.D.D., Sirisena K.A. & Manage P.M. (2019a). Isolation and characterization of thermophilic bacteria with extracellular enzymatic activities from hot springs of Sri Lanka. *Proceedings of the Symposium of ASM Microbes*, 20 – 24 June. American Society for Microbiology California, USA, pp.30.
- Sadeepa H.D.D., Sirisena K.A. & Manage P.M. (2019b). Isolation and characterization of industrially important enzyme producing bacteria from hot springs of Sri Lanka. *Proceedings of the Symposium of International Conference of Multidisciplinary Approaches*, 26-27 November. University of Sri Jayewardenepura, Colombo, Sri Lanka, pp. 116.
- Seawright A.A., Nolan C.C., Shaw G.R., Chiswell R.K., Norris R.L., Moore M.R. & Smith M.J. (1999). The oral toxicity for mice of the tropical cyanobacterium *Cylindrospermopsis raciborskii* (Woloszynska). *Environmental Toxicology* **14**(1): 135–142. DOI: [https://doi.org/10.1002/\(SICI\)1522-7278\(199902\)14:1<135::AID-TOX17>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1522-7278(199902)14:1<135::AID-TOX17>3.0.CO;2-L)
- Sethunge S. & Manage P.M. (2010). Nuisance algae in water supply projects in Sri Lanka. *Proceedings of the Symposium of Sustainable Built Environment (ICSBE-2010)*, 13th –14th December. Kandy, Sri Lanka, pp. 28.
- SLSI (Sri Lanka Standard Institute) (2013). SLS 614-Specification for Portable Water, Sri Lanka SLSI, Colombo, Sri Lanka.
- Sukenik A., Reisner M., Carmeli S. & Werman M. (2006). Oral toxicity of the cyanobacterial toxin cylindrospermopsin in mice: Long-term exposure to low doses. *Environmental Toxicology* **21**(6): 575–582. DOI: <https://doi.org/10.1002/tox.20220>
- Wijesekara W.A.M.A. & Manage P.M. (2017). *In vitro* screening of antibacterial and phytochemical properties of hot water spring cyanobacterium *Lynbya* sp. *Proceedings of the NARA Scientific Sessions*, 27 July. National Aquatic Resources Research and Development Agency (NARA), Colombo 15, Sri Lanka, pp. 38.
- Wijewickrama M.M. & Manage P.M. (2019). Accumulation of Microcystin-LR in grains of two rice varieties (*Oryza sativa* L.) and a leafy vegetable, *Ipomoea aquatica*. *Toxins* **11**(8): 432. DOI: <https://doi.org/10.3390/toxins11080432>
- Woese C.R. (2002). On the evolution of cells. *Proceedings of the National Academy of Sciences* **99**(13): 8742–8747.
- WHO (2003). *Cyanobacterial Toxins: Microcystin-LR in Drinking-Water*. World Health Organization (WHO), Geneva, Switzerland.
- Yatigammana S.K. & Perera M.B. (2017). Distribution of *Cylindrospermopsis raciborskii* (cyanobacteria) in Sri Lanka. *Ceylon Journal of Science* **46**(3): 65–80. DOI: <https://doi.org/10.4038/cjs.v46i3.7444>
- Zanchett G. & Oliveira-Filho E.C. (2013). Cyanobacteria and cyanotoxins: from impacts on aquatic ecosystems and human health to anticarcinogenic effects. *Toxins* **5**(10): 1896–1917. DOI: <https://doi.org/10.3390/toxins5101896>

RESEARCH ARTICLE

Functional Analysis

On the construction of unitaries representing minimal inner toral polynomials

CSB Dissanayake and UD Wijesooriya*

Department of Mathematics, Faculty of Science, University of Peradeniya, Peradeniya.

Submitted: 17 November 2021; Revised: 17 February 2022; Accepted: 25 March 2022

Abstract: A polynomial $p = p(z, w) \in \mathbb{C}[z, w]$ is called an inner toral polynomial if the zero set of p , contained in $\mathbb{D}^2 \cup \mathbb{T}^2 \cup \mathbb{E}^2$ where \mathbb{D}^2 is the open bidisk, \mathbb{T}^2 is the two dimensional torus and \mathbb{E}^2 is the exterior of the closed bidisk in \mathbb{C}^2 . Such a polynomial is called a minimal inner toral polynomial if it divides all the other polynomials with same zero set as itself. The bidegree of p is defined as the ordered pair (n, m) whenever p has degrees n and m in variables z and w respectively. For a minimal inner toral polynomial $p(z, w)$ of bidegree (n, m) , there exist block unitary matrices, $\begin{pmatrix} A & B \\ C & D \end{pmatrix}_{((m+n) \times (m+n))}$, such that $p(z, w)$ is a constant multiple of the determinant of $\begin{pmatrix} A - wI_m & zB \\ C & zD - I_n \end{pmatrix}$, where, blocks A, B, C and D are matrices with complex entries of sizes $(m \times m), (m \times n), (n \times m)$ and $(n \times n)$ respectively. Because the size of such unitaries solely depends on the bidegree of the considered minimal inner toral polynomial, the process of computing such unitary matrices is manageable when the bidegree is very small but become strenuous exponentially with the bidegree. In this work, we introduce a method of constructing unitary matrices representing reducible minimal inner toral polynomials in terms of the unitaries representing its factors and vice versa.

Keywords: Block matrices, distinguished varieties, inner toral polynomials, unitary matrices.

INTRODUCTION

Inner toral polynomials are polynomials in two complex variables, say in z and w , such that their zero sets lie in the open bidisk and exit the open bidisk to the exterior

of the closed bidisk through the two-dimensional torus. These polynomials became interesting, because the variety generated by them have this special form, and most importantly, a given pair of pure commuting isometries that satisfies an algebraic relationship contains an inner toral polynomial that annihilates the same pair of isometries. Therefore, studies related to such polynomials appear in many research articles in functional analysis as well as operator theory (Agler & McCarthy, 2002, 2005, 2006; Agler *et al.*, 2006, 2012; Knese, 2010; McCarthy, 2012; Jury *et al.*, 2012; Pal & Shalit, 2014; Wijesooriya, 2018; Bhattacharyya *et al.*, 2022). An inner toral polynomial is called a minimal inner toral polynomial or simply minimal, if it divides all the polynomial with the same zero set as itself. In 2005, Agler and McCarthy have proved that a minimal inner toral polynomial can be represented with a unitary block matrix in a way that the determinant of a shifted version of the unitary block matrix is a constant multiple of the polynomial.

If the inner toral polynomial we consider has bidegree (n, m) , then a unitary representing it, is a square matrix of size $n + m$. Due to the large size of these matrices, and the fact that the number of unknowns in the matrices increases significantly with the bidegree of the polynomial, it is not computationally efficient to come up with examples. In this study we paid attention to reducible minimal inner toral polynomials which are obviously factorizable into simpler irreducible minimal inner toral polynomials and we found a method to compute unitaries representing such reducible minimal inner toral polynomials in terms

* Corresponding author (udeniw@sci.pdn.ac.lk;  <https://orcid.org/0000-0003-3306-7036>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

of unitaries representing their factors and vice versa. We define a special type of a direct sum for block matrices, which we call the diagonal wise direct sum, and we construct desired unitary matrices by performing the diagonal wise direct sum on them.

PRELIMINARIES

In this section we discuss necessary background materials to discuss the main result in the preceding section. Throughout this paper we consider \mathbb{D} to be the open unit disk, \mathbb{T} to be the unit circle and \mathbb{E} to be the exterior of the closed unit disk in the complex plane.

Definition 1: Let $p = p(z, w) \in \mathbb{C}[z, w]$ be a polynomial in z and w . We say that p has bidegree (n, m) if p has degree n in z and m in w .

Definition 2: A polynomial $p = p(z, w) \in \mathbb{C}[z, w]$ is called an inner toral polynomial if the zero set of p , $Z(p)$, contained in $\mathbb{D}^2 \cup \mathbb{T}^2 \cup \mathbb{E}^2$.

In other words, for a two-variable polynomial $p = p(z, w) \in \mathbb{C}[z, w]$, in order to be inner toral, every zero (z_0, w_0) of p must satisfy the condition that either $|z_0|, |w_0| < 1$ or $|z_0| = |w_0| = 1$ or $|z_0|, |w_0| > 1$.

For example, $p_1(z, w) = z^n - w^m$ for $n, m \in \mathbb{N}$ is inner toral. However, $p_2(z, w) = z^n - 2w^m$ is not inner toral because $(\frac{1}{2^n}, 1)$ is a zero of p_2 but it is not in $\mathbb{D}^2 \cup \mathbb{T}^2 \cup \mathbb{E}^2$.

Definition 3: An inner toral polynomial $p(z, w) \in \mathbb{C}[z, w]$ is called a minimal inner toral polynomial, if it divides any $q \in \mathbb{C}[z, w]$ with $Z(p) = Z(q)$.

A minimal inner toral polynomial is minimal in the sense that it has smallest possible bidegree, and it is square free (unique up to constant multiples) with the same zero set as the original one. Therefore, clearly the minimal polynomial is also inner toral. Given an inner toral polynomial p , its minimal version always exists as it is the product of all the irreducible factors (of algebraic multiplicity one) of p (up to constant multiples). For example, $(z - w)^2(z + w)$ is inner toral and $(z - w)(z + w)$ is its minimal version.

Definition 4: For an inner toral polynomial p the set, $\mathfrak{B}(p) = Z(p) \cap \mathbb{D}^2$ is called a distinguished variety.

Definition 5: A square matrix U is called a unitary matrix if its conjugate transpose U^* is its inverse. That is, $UU^* = U^*U = I$ where I is the identity matrix.

Our work is based on the unitaries representing minimal inner toral polynomials. The Theorem given below explain the existence of such unitary matrices and how exactly they look like. This Theorem was first proved by Agler and McCarthy in 2005 and an alternative proof was given by Knese in 2010. This result is one of the breakthrough results in this setting as it eventually became a very helpful tool in the studies carried out by many mathematicians related to inner toral polynomials, distinguished varieties and pure algebraic iso-pairs (Knese, 2010; Jury *et al.*, 2012; McCarthy, 2012; Pal & Shalit 2014; Wijesooriya, 2018).

Theorem 1: (Knese, 2010) For a minimal inner toral polynomial $p(z, w)$ of bidegree (n, m) there exists an unitary matrix $U = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{matrix} m & n \\ m & n \end{matrix}$ such that

- $\det \begin{pmatrix} A - wI_m & zB \\ C & zD - I_n \end{pmatrix} = k p(z, w)$ where k is a constant and
- if $\det(D - \lambda I_n) = 0$, then $|\lambda| \neq 1$.

Here, A, B, C and D are matrices of sizes $(m \times m)$, $(n \times m)$, $(m \times n)$ and $(n \times n)$ respectively. Such unitary matrices, U , are called unitaries representing minimal inner toral polynomial $p(z, w)$.

For example, let $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}_{2 \times 2}$ be a unitary representing the minimal inner toral polynomial $z - w$. That means the determinant of $\begin{pmatrix} a - w & zb \\ c & zd - 1 \end{pmatrix}$ is equal to $k(z - w)$ for some constant $k \in \mathbb{C}$. By comparing the coefficients of the equation $(ad - bc)z - dzw + w - a = k(z - w)$, and considering the fact that U is a unitary matrix, we have $a = d = 0$, and $|b| = |c| = 1$ with $bc = 1$. Therefore, unitaries representing $z - w$ are of the form $\begin{pmatrix} 0 & \alpha \\ \bar{\alpha} & 0 \end{pmatrix}$ where α is a unimodular constant. This computation can be done easily due to the fact that U has a relatively small size and the system of equations involved is simple. However, that is not the case all the time. For example, unitaries representing the minimal inner toral polynomial $z^4 - w^4$ are 8×8 matrices and it is not quite easy to compute such unitaries as it involved finding 64 unknowns.

Remark 1: It is evident from the above example that, given an inner toral polynomial one can find more than one unitary matrix representing the considered polynomial.

Before moving to the results and discussions, we prefer to discuss the invertibility of the block matrix $zD - I$ to avoid any ambiguities that could appear in the proof of the main Theorem.

Theorem 2. (Bezout’s Theorem) If two algebraic curves, say described by $p(z, w) = 0$ and $q(z, w) = 0$ do not have any common components, then they have only finitely many points in common.

In fact, if $p(z, w)$ is a minimal inner toral polynomial and if zeroes of p and q are common except for finitely many zeroes, then p must be a factor of q .

Remark 2: Let $U = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ be a unitary representing the minimal inner toral polynomial $p(z, w)$ of bidegree (n, m) . By Theorem 1,

$$\det \begin{pmatrix} A - wI_m & zB \\ C & zD - I_n \end{pmatrix} = k p(z, w) \text{ for some constant } k. \text{ Note that } \det \begin{pmatrix} A - wI_m & zB \\ C & zD - I_n \end{pmatrix} = \det(zD - I) \cdot$$

$\det[(A - wI) - zB(zD - I)^{-1}C]$. It is immediate that this equation is true only if the matrix $zD - I$ is invertible. Since the matrix D is of size $n \times n$, the determinant of $zD - I$ can be vanished only at finitely many values of z . Therefore, $\det \begin{pmatrix} A - wI_m & zB \\ C & zD - I_n \end{pmatrix}$ vanishes at all except finitely many points in $Z(p)$. By Theorem 2, and the minimality of p , $\det \begin{pmatrix} A - wI_m & zB \\ C & zD - I_n \end{pmatrix}$ vanishes at all the points in $Z(p)$. Eventually, this leads to the fact that $\det \begin{pmatrix} A - wI_m & zB \\ C & zD - I_n \end{pmatrix} = \det(zD - I) \cdot$

$\det[(A - wI) - zB(zD - I)^{-1}C] = k p(z, w)$ [see page 16 of (Knese, 2010) for more details on this]. Therefore, the non-invertibility of $zD - I$ at finitely many $z \in \mathbb{C}$, does not harm the determinantal representation of the unitary to the polynomial. In the proof of the main theorem presented in this paper, we say the fact that $zD - I$ is invertible, with the fully understanding that $zD - I$ is invertible at all except finitely many $z \in \mathbb{C}$.

Now we present the method we found on computing unitaries representing minimal inner toral polynomials.

RESULTS AND DISCUSSION

Let $m_i, n_i \in \mathbb{N}$ and suppose U_i is a block matrix of the form

$$U_i = \begin{pmatrix} m_i & n_i \\ A_i & B_i \\ C_i & D_i \end{pmatrix} \begin{matrix} m_i \\ n_i \end{matrix}$$

for $i = 1, 2, \dots, s$ with sizes given above.

$$U_1 \oplus_D U_2 \oplus_D \dots \oplus_D U_s =$$

$$\begin{pmatrix} A_1 & 0 & 0 & \dots & 0 & B_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 & 0 & B_2 & 0 & \dots & 0 \\ 0 & 0 & A_3 & \dots & 0 & 0 & 0 & B_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & A_s & 0 & 0 & 0 & \dots & B_s \\ C_1 & 0 & 0 & \dots & 0 & D_1 & 0 & 0 & \dots & 0 \\ 0 & C_2 & 0 & \dots & 0 & 0 & D_2 & 0 & \dots & 0 \\ 0 & 0 & C_3 & \dots & 0 & 0 & 0 & D_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_s & 0 & 0 & 0 & \dots & D_s \end{pmatrix}.$$

We call this sum the diagonalwise direct sum. For the ease of notation, let us denote the block matrix

$$\begin{pmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ 0 & 0 & A_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & A_s \end{pmatrix} \text{ by } \oplus A_i \text{ and likewise } \oplus B_i, \oplus C_i$$

and $\oplus D_i$. Likewise, we will use the notation, \oplus , to denote the block diagonal matrices where it is necessary. Again, for the ease of using notations, given a block matrix $U = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, we denote $\begin{pmatrix} A - wI & Bz \\ C & zD - I \end{pmatrix}$ by $U(z, w)$.

Let $p(z, w) \in \mathbb{C}[z, w]$ be a reducible minimal inner toral polynomial. Write $p = p_1 p_2 p_3 \dots p_s$ for some $s \in \mathbb{N}$. Note that each factor p_i is also a minimal inner toral polynomial. Let (n, m) and (n_i, m_i) be the bi-degrees of the polynomial p and p_i for $i = 1, 2, \dots, s$ respectively. Note that $n = \sum_{i=1}^s n_i$ and $m = \sum_{i=1}^s m_i$. Further, let U and U_i be square matrices of order $(m + n)$ and $(m_i + n_i)$ for $i = 1, 2, \dots, s$ respectively. Now, we present our main result:

Theorem 3

If U_i is a unitary matrix representing the minimal inner toral polynomial p_i for each, $i = 1, 2, \dots, s$, then the

matrix $U = U_1 \oplus_D U_2 \oplus_D \dots \oplus_D U_s$ is a unitary matrix representing the polynomial p . Conversely, suppose $U = \begin{pmatrix} \oplus A_i & \oplus B_i \\ \oplus C_i & \oplus D_i \end{pmatrix}$ is a unitary representing the minimal inner toral polynomial p , where $i = 1, 2, 3, \dots, s$. Each matrix given by $U_i = \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix}$ is a unitary matrix representing the minimal inner toral polynomial p_i (up to re-labelling) for $i = 1, 2, \dots, s$.

Proof: Let $U_i = \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix}$ for $i = 1, 2, 3, \dots, s$ with the sizes given as above. First, we must show that U is unitary. Since each $U_i = \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix}$ is a unitary, $U_i U_i^* = U_i^* U_i = I_{n_i+m_i}$ for $i = 1, 2, \dots, s$. In fact,

$$\begin{pmatrix} A_i A_i^* + B_i B_i^* & A_i C_i^* + B_i D_i^* \\ C_i A_i^* + D_i B_i^* & C_i C_i^* + D_i D_i^* \end{pmatrix} = \begin{pmatrix} A_i^* A_i + B_i^* B_i & C_i^* A_i + D_i^* B_i \\ A_i^* C_i + B_i^* D_i & C_i^* C_i + D_i^* D_i \end{pmatrix} = I_{n_i+m_i}$$

for $i = 1, 2, \dots, s$.

Clearly $U = \begin{pmatrix} m & n \\ \oplus A_i & \oplus B_i \\ \oplus C_i & \oplus D_i \end{pmatrix} \begin{matrix} m \\ n \end{matrix}$

$$U U^* = \begin{pmatrix} \oplus(A_i A_i^* + B_i B_i^*) & \oplus(A_i C_i^* + B_i D_i^*) \\ \oplus(C_i A_i^* + D_i B_i^*) & \oplus(C_i C_i^* + D_i D_i^*) \end{pmatrix} \text{ and}$$

$$U^* U = \begin{pmatrix} \oplus(A_i^* A_i + B_i^* B_i) & \oplus(C_i^* A_i + D_i^* B_i) \\ \oplus(A_i^* C_i + B_i^* D_i) & \oplus(C_i^* C_i + D_i^* D_i) \end{pmatrix}.$$

It easily follows that $U U^* = U^* U = I_{n+m}$. Therefore U is also a unitary matrix. It is left to prove that the unitary matrix U represents the polynomial p . Since U_i represent p_i , there exist $c_i \in \mathbb{C}$ such that $\det U_i(z, w) = c_i p_i(z, w)$ for all $i = 1, 2, 3, \dots, s$. In fact, for $i =$

$$1, 2, 3, \dots, s, \det \begin{pmatrix} A_i - w I_{m_i} & z B_i \\ C_i & z D_i - I_{n_i} \end{pmatrix} = c_i p_i(z, w),$$

and hence

$$\det(z D_i - I_{n_i}) \cdot \det \left[(A_i - w I_{m_i}) - z B_i (z D_i - I_{n_i})^{-1} C_i \right] = c_i p_i(z, w).$$

Note that

$$U(z, w) = \begin{pmatrix} \oplus(A_i - w I_{m_i}) & \oplus z B_i \\ \oplus C_i & \oplus(z D_i - I_{n_i}) \end{pmatrix}.$$

Since $(z D_i - I_{n_i})$ is invertible, $(\oplus(z D_i - I_{n_i}))$ is also invertible and for $i = 1, 2, \dots, s$. Therefore,

$$\begin{aligned} \det [U(z, w)] &= \det(\oplus(z D_i - I_{n_i})) \cdot \det \left[\oplus(A_i - w I_{m_i}) - \oplus z B_i (\oplus(z D_i - I_{n_i}))^{-1} \oplus C_i \right] \\ &= \det(\oplus(z D_i - I_{n_i})) \cdot \det \left[\oplus \left((A_i - w I_{m_i}) - z B_i (z D_i - I_{n_i})^{-1} C_i \right) \right] \\ &= \prod_{i=1}^s \det(z D_i - I_{n_i}) \cdot \det \left[(A_i - w I_{m_i}) - z B_i (z D_i - I_{n_i})^{-1} C_i \right] \\ &= \prod_{i=1}^s c_i p_i(z, w) \\ &= c p(z, w) \text{ where } c = c_1 c_2 c_3 \dots c_s. \end{aligned}$$

Therefore, U represents p .

Conversely, since U represent p , $\det(U(z, w)) = c p(z, w)$ for some constant $c \in \mathbb{C}$. In other words,

$$\det(\oplus(z D_i - I_{n_i})) \cdot \det \left[\oplus(A_i - w I_{m_i}) - \oplus z B_i (\oplus(z D_i - I_{n_i}))^{-1} \oplus C_i \right] = c p(z, w).$$

Since $(\oplus(z D_i - I_{n_i}))$ is invertible, $(z D_i - I_{n_i})$ is also invertible for $i = 1, 2, 3, \dots, s$ and

$$(\oplus(z D_i - I_{n_i}))^{-1} = \oplus(z D_i - I_{n_i})^{-1}.$$

Therefore,

$$\det(\oplus(z D_i - I_{n_i})) \cdot \det \left[\oplus \left((A_i - w I_{m_i}) - z B_i (z D_i - I_{n_i})^{-1} C_i \right) \right] = c p(z, w).$$

Consequently, $\prod_{i=1}^s \det(z D_i - I_{n_i}) \cdot$

$$\det \left[(A_i - w I_{m_i}) - z B_i (z D_i - I_{n_i})^{-1} C_i \right] = c p_1 p_2 \dots p_s.$$

That is, $\prod_{i=1}^s \det(U_i(z, w)) = c p(z, w) = c p_1 p_2 \dots p_s$. Note that each p_i is irreducible and distinct. Moreover, bidegree of $\det(U_i(z, w))$ is at most (n_i, m_i) and bidegree of $p_i(z, w)$ is (n_i, m_i) . Therefore, $\det(U_i(z, w))$ divides $p_i(z, w)$ up to relabeling of p_i 's. Therefore, unitary matrix U_i represents p_i , upto relabeling of p_i 's for $i = 1, 2, \dots, s$. Since p_i is a minimal inner toral polynomial, $\det(U_i(z, w)) = c_i p_i(z, w)$ up to relabeling of p_i 's. Therefore each U_i is a unitary representing p_i , up to relabeling.

For examples, consider the reducible minimal inner toral polynomial $p = z^4 - w^4 = (z - w)(z + w)(z - iw)(z + iw)$. Unitaries representing p are of the

size 8×8 and obviously, it is difficult to evaluate all the 64 unknowns in such a matrix. However, we can easily compute the unitaries representing each factor of p . Unitaries representing $z - w$, $z + w$, $z - iw$ and $z + iw$ are of the form $\begin{pmatrix} 0 & \alpha \\ \bar{\alpha} & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & \alpha \\ -\bar{\alpha} & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & \alpha \\ i\bar{\alpha} & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & \alpha \\ -i\bar{\alpha} & 0 \end{pmatrix}$ respectively where $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. Now using the diagonal wise direct sum, and the above Theorem, we can easily obtain a collection of unitary matrices representing p , which are of the form

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha \\ \bar{\alpha} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\bar{\alpha} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & i\bar{\alpha} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i\bar{\alpha} & 0 & 0 & 0 & 0 \end{pmatrix}$$

Likewise, for the minimal inner toral polynomial $q(z, w) = (z^2 + zw) + i(zw + w^2) = (z + w)(z + iw)$, a collection of unitary matrices representing q can be obtained as:

$$\begin{pmatrix} 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha \\ -\bar{\alpha} & 0 & 0 & 0 \\ 0 & -i\bar{\alpha} & 0 & 0 \end{pmatrix}$$

In another example, a collection of unitaries representing the polynomial $z^2 - w$ is given as $\begin{pmatrix} 0 & 0 & \lambda \\ \beta & 0 & 0 \\ 0 & \bar{\lambda}\bar{\beta} & 0 \end{pmatrix}$ where $\lambda, \beta \in \mathbb{C}$ with $|\lambda| = |\beta| = 1$. Using the diagonalwise direct sum, we can construct a collection of unitary matrices representing the minimal inner toral polynomial $r(z, w) = z^3 - z^2w - zw + w^2 = (z^2 - w)(z - w)$ as

$$\begin{pmatrix} 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \alpha \\ \beta & 0 & 0 & 0 & 0 \\ 0 & 0 & \bar{\lambda}\bar{\beta} & 0 & 0 \\ 0 & \bar{\alpha} & 0 & 0 & 0 \end{pmatrix}$$

for $\alpha, \lambda, \beta \in \mathbb{C}$ with $|\alpha| = |\lambda| = |\beta| = 1$.

CONCLUSIONS

In this work, we have been able to introduce a method to construct unitaries representing reducible minimal inner toral polynomials in terms of unitaries representing its factors. This method of construction unitaries is useful

as it reduces and simplifies the number of computations involve in computing unitaries representing reducible minimal inner toral polynomials of higher bidegree.

Conflict of interest

Authors hereby declare that there is no conflict of interest to disclose.

REFERENCES

Agler J. & McCarthy J.E. (2002). *Pick Interpolation and Hilbert Function Spaces*. American Mathematical Society, Providence, Rhode Island, USA.
DOI: <https://doi.org/10.1090/gsm/044>

Agler J. & McCarthy J.E. (2005). Distinguished varieties. *Acta Mathematica* **194**(2005): 133–153.
DOI: <https://doi.org/10.1007/bf02393219>

Agler J. & McCarthy J.E. (2006). Parametrizing distinguished varieties: Recent advances in operator-related function theory. In: *Contemporary Mathematics*, Volume 393. American Mathematical Society, Providence, Rhode Island, USA.
DOI: <https://doi.org/10.1090/conm/393/07368>

Agler J., Knese G. & McCarthy J.E. (2012). Algebraic pairs of isometries. *Journal of Operator Theory* **67**: 215–236.

Agler J., McCarthy J.E. & Mark S. (2006). Toral algebraic sets and function theory on polydisks. *Journal of Geometric Analysis* **16**: 551–562.
DOI: <https://doi.org/10.1007/bf02922130>

Bhattacharyya T., Kumar P. & Sau D. (2022). Distinguished varieties through the Berger–Coburn–Lebow theorem. *Analysis and PDE* **15**(2): 477–506
DOI: <https://doi.org/10.2140/apde.2022.15.477>

Jury M.T., Knese G. & McCullough S. (2012). Nevanlinna–Pick interpolation on distinguished varieties in the bidisk. *Journal of Functional Analysis* **262**: 3812–3838.
DOI: <https://doi.org/10.1016/j.jfa.2012.01.028>

Knese G. (2010). Polynomials defining distinguished varieties. *Transactions of the American Mathematical Society* **362**: 5635–5655.
DOI: <https://doi.org/10.1090/s0002-9947-2010-05275-4>

McCarthy J.E. (2012). Shining a Hilbertian lamp on the bidisk: Topics in complex analysis and operator theory In: *Contemporary Mathematics*, Volume 561. American Mathematical Society, Providence, Rhode Island, USA.
DOI: <https://doi.org/10.1090/conm/561/11109>

Pal S. & Shalit O. (2014). Spectral sets and distinguished varieties in the symmetrized bidisc. *Journal of Functional Analysis* **266**: 5779–5800.

Wijesooriya U. (2018). Algebraic pairs of pure commuting isometries with finite multiplicity. *Journal of Operator Theory* **79**(2): 507–527.
DOI: <http://dx.doi.org/10.7900/jot.2017may12.2171>

RESEARCH ARTICLE

Applied Statistics

A coupled system of stochastic differential equations for probabilistic wind speed modelling

HMDP Bandarathilake^{1*}, GWRMR Palamakumbura² and DHS Maithripala³

¹ Faculty of Arts and Sciences, Sri Lanka Technological Campus, Padukka.

² Department of Engineering Mathematics, Faculty of Engineering, University of Peradeniya, Peradeniya.

³ Department of Mechanical Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya.

Submitted: 10 March 2021; Revised: 20 March 2020; Accepted: 25 March 2022


Abstract: Wind-based electricity generation plays a vital role in the renewable energy industry. Benefits from its use can only be obtained if it is possible to accommodate its variability and limited predictability. Stochastic differential equation (SDE) based approaches have demonstrated an improved capability of predicting temporal wind speed patterns that have statistical properties that are similar to those observed in reality. However, no standard approach for deriving such models exist due to the wide variations in the temporal statistical properties that one observes in wind speed data measured from location to location. Wind speed data which have been recorded at coastal locations, exhibit non-stationary features, like diurnal effect, seasonal effects, and temporal trends. In this work, such effects are eliminated using standard smoothing techniques. A coupled system of second-order ordinary linear differential equations driven by a white noise forcing term was used for the probabilistic modelling of the residual data. The model was then used to predict the wind speed distribution and the corresponding autocorrelation function of wind speed data, recorded at the wind measurement centre of Kokkilai in northern Sri Lanka, from February 2015 to February 2016. Finally, the results of this novel approach were compared against the probabilistic modelling methods existing in the literature.

Key words: Autocorrelation function, probabilistic modelling, smoothing techniques, stochastic differential equation, white noise.

INTRODUCTION

Environmental concerns and supply uncertainties are forcing many countries to rethink their energy mix and develop diverse sources of clean, renewable energy.

Cost-effective energy that can be produced without major negative environmental impact has become the goal worldwide. Wind energy, as a clean and renewable resource, has been under wide consideration around the world in the last decade. The benefits of wind energy are accompanied by several challenges: high variability, limited predictability, and non-storability. Thus wind can be identified as a highly volatile energy source and hence an efficient mechanism for modelling wind speed has become a mandatory requirement in power system regulations, control, and management of wind-based electricity generation (Zhu & Genton, 2012). How to reduce the uncertainty in predicting wind speeds has been the focus of studies over the last two decades. To smoothly integrate large-scale wind power into power systems, wind generation forecasting models have been developed to improve the accuracy of forecasts. Time series analysis-based methods, probability distribution-based approaches and applications of Kalman filters are found as the major statistical and mathematical modelling approaches used in modelling wind speeds. In Zhu & Genton (2012), it is pointed out that the accurate probabilistic modelling of wind speeds is more useful than point-forecasting for the power system operations of wind-based energy industry. Thus, in addition to the traditional mathematical and statistical tools for wind speed modelling, Stochastic Differential Equations (SDE) based models have become popular and have shown much promise (Zhu & Genton, 2012). Most of these SDE-based approaches have employed a single

* Corresponding author (dbthilak@gmail.com ;  <https://orcid.org/0000-0002-5214-4626>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

equation to produce forecasting trajectories of wind speeds (Bibby *et al.*, 2005; Calif, 2012; Zarate-Minano *et al.*, 2016). In addition, applications of those methods have been confined to modelling wind speed data that exhibit significant stationary features. Yet, from a practical point of view, wind speed data exhibit considerable non-stationary features. Therefore, applications of SDEs that generate stationary Ornstein-Uhlenbeck (O-U) processes are not sufficient to obtain reliable probabilistic forecasting for the wind energy industry. In this regard, the sums of diffusion method (Bibby *et al.*, 2005) has demonstrated successful results for weakly stationary real process that have several time scale effects. However, when the influence of cyclic effects becomes intensive, the method becomes ineffective due to the increased non-stationariness. In addition, there are other sophisticated approaches based on non-linear filtering and particle filtering that can be successfully adapted to probabilistic modelling of wind speed data (Fernando & Hausenblas, 2018; Hausenblas *et al.*, 2021).

In this paper, we propose a novel approach that first eliminates the non-stationary effects using a smoothing method and then uses two coupled oscillators driven by white noise to model the residual process. The approach was then used to successfully model empirical wind speed data, recorded at Kokkilai in the North Eastern coast of Sri Lanka from February 2015 to February 2016 under the Quantum Leap Wind Power Assessment Project of the Sustainable Energy Authority of Sri Lanka.

MATERIALS AND METHODS

In this section, we present a novel method for predicting wind speeds and validate it against the wind speed data measured at Kokkilai, a coastal point located on the North Eastern shore of Sri Lanka. The wind speed data that were originally recorded at 10 min intervals were converted to hourly averaged wind speeds. The empirical autocorrelation function of hourly wind speed data in Figure 1(a) exhibits a complex cyclic behaviour due to the influence of diurnal and seasonal effects (Brett & Tuller, 1991). As a coastal point, Kokkilai also comes under the influence of both North Eastern and South Western monsoonal wind patterns. In addition, due to the inter-monsoonal wind variations, short-scale temporal trends of wind speed variation were crucial factors that contributed to the non-stationariness. Figure 1(b) shows that the empirical wind speed distribution deviates significantly from a normal Gaussian distribution.

In order to eliminate the non-stationary features of wind speed data, the corresponding time plot and correlogram were carefully inspected to identify various cyclic patterns with distinct periods of temporal trends. Accordingly, the dataset was divided into distinct blocks based on the visual inspection of the time plot and correlogram. Regression models were then incorporated to eliminate the non-stationary features that were specific to each block.

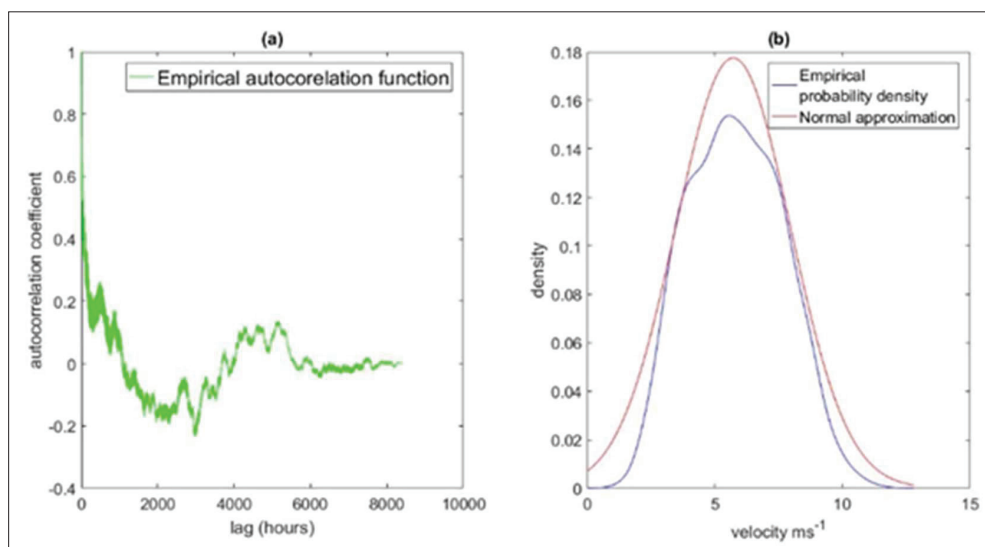


Figure 1: (a): Empirical autocorrelation of hourly wind speeds and (b): empirical wind speed distribution approximated by a normal Gaussian distribution

When a cyclic variation of fixed period is apparent in a wind speed data set, it is eliminated with a regression model of the form,

$$u_r = b_0 + b_1 \cos\left(\frac{2\pi}{T}t\right) + b_2 \sin\left(\frac{2\pi}{T}t\right), \quad \dots(1)$$

where $b_i (i = 0, 1, 2)$ are constants and T is the period of the cyclic variation (Chatfield, 2004; Shumway & Stoffer 2016). Here, b_0 captures the mean value of the data.

In the event where several cyclic effects are present a regression model of the following form can be used (Chatfield, 2004; Shumway & Stoffer, 2016).

$$u_r = b_0 + \sum_{i=1}^n \left(a_i \cos\left(\frac{2\pi}{T_i}t\right) + b_i \sin\left(\frac{2\pi}{T_i}t\right) \right). \quad \dots(2)$$

Here, $a_i, b_i (i = 0, 1, 2 \dots)$ are constants and $T_i (i = 0, 1, 2 \dots)$ is the period of the i^{th} cycle.

Furthermore, a polynomial regression model of the form shown below can be used to capture the temporal trend patterns included in a wind speed dataset (Chatfield, 2004; Shumway & Stoffer, 2016)

$$u_r = \sum_{i=0}^m c_i t^i, \quad \dots(3)$$

where $c_i (i = 0, 1, 2 \dots)$ are constants.

In the case, when both cyclic effects and temporal trends are present a sum of the regression models given by equation (2) and (3) can be used,

$$u_r = b_0 + \sum_{i=1}^n \left(a_i \cos\left(\frac{2\pi}{T_i}t\right) + b_i \sin\left(\frac{2\pi}{T_i}t\right) \right) + \sum_{i=1}^m c_i t^i. \quad \dots(4)$$

By visual inspection of the time plot of wind speed data and the corresponding correlogram, the dataset was partitioned into eleven blocks based on the observed different non-stationary features. The non-stationary features and the regression models incorporated are summarized in the following table (Table 1).

Then, the residual process (v) was obtained by eliminating the respective non-stationary component (u_r) from the empirical data (u) as in (5) and the corresponding time plots are given in Figure 2.

$$v = u - u_r \quad \dots(5)$$

Table 1: Periods of cyclic effects and the order of regression polynomial functions corresponding to each wind speed data block

Block numbers	Parameter values of regression models	Regression model
1 and 3	$T = 24$ hours	$b_0 + b_1 \cos\left(\frac{2\pi}{T}t\right) + b_2 \sin\left(\frac{2\pi}{T}t\right)$
2, 4, 5, 6, 7 and 9	$T_1 = 24$ hours, $m = 2$	$b_0 + \sum_{i=1}^n \left(a_i \cos\left(\frac{2\pi}{T_i}t\right) + b_i \sin\left(\frac{2\pi}{T_i}t\right) \right) + \sum_{i=1}^m c_i t^i$
8	$T_1 = 24$ hours, $T_2 = 350$ hours	$b_0 + \sum_{i=1}^n \left(a_i \cos\left(\frac{2\pi}{T_i}t\right) + b_i \sin\left(\frac{2\pi}{T_i}t\right) \right)$
10 and 11	$T_1 = 24$ hours, $T_2 = 144$ hours, $m = 2$	$b_0 + \sum_{i=1}^n \left(a_i \cos\left(\frac{2\pi}{T_i}t\right) + b_i \sin\left(\frac{2\pi}{T_i}t\right) \right) + \sum_{i=1}^m c_i t^i$

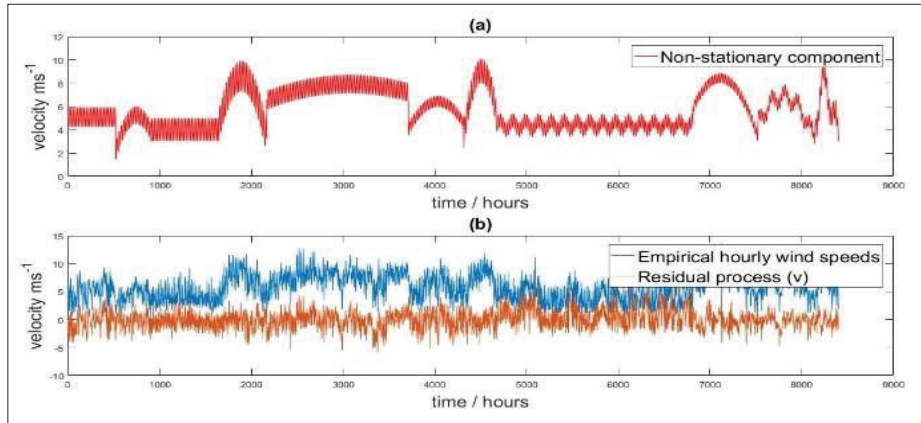


Figure 2: (a): Non-stationary component (u_t) of empirical wind speed data, (b): Time plot of empirical wind speeds and residual process (v)

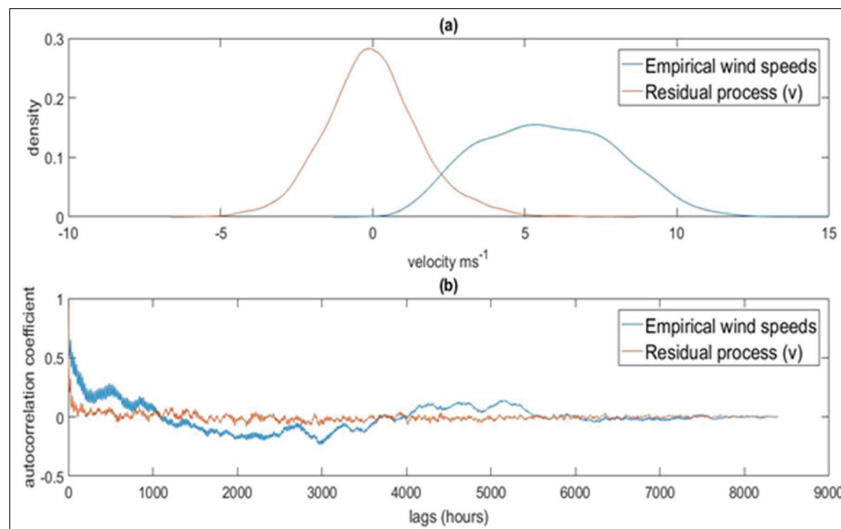


Figure 3: (a): Distributions of empirical wind speeds and residual process (v), (b): Correlograms of empirical wind speeds and residual process (v)

The correlogram of Figure 3(b) indicates that the residual process can be approximately treated as a wide sense stationary process (Chatfield, 2004).

In the rest of this section, we show how this approximately stationary residual process can be modeled by a process that is generated by a coupled pair of second-order linear SDEs.

Consider the general n -dimensional SDE,

$$\dot{X} = A(X, t) + B(X, t)\eta(t) \quad \dots(6a)$$

$$y = CX \quad \dots(6b)$$

where X is a $n \times 1$ matrix, A is a $n \times n$ matrix, B is a $n \times m$ matrix, and $\eta(\cdot)$ is a $m \times 1$ white noise process.

Definition: White noise

A white noise process $\{\eta(t) | t \in [0, T]\}$ is a random process with the following properties (Gardiner, 1985; Särkkä & Solin, 2019):

1. For $t_1 \neq t_2$, $\eta(t_1)$ and $\eta(t_2)$ are independent.
2. $t \rightarrow \eta(t)$ is a Gaussian process with zero mean and

Dirac delta correlation:

Mean: $m(t) = E[\eta(t)] = 0$

Correlation: $C_\eta(t,s) = E[\eta(t)\eta^T(s)] = \delta(t-s)Q$

where Q is the spectral density of the process.

From the above properties we can also deduce the following peculiar properties of white noise:

- a. The sample path $t \rightarrow \eta(t)$ is discontinuous almost everywhere.
- b. White noise is unbounded and it takes arbitrarily large positive and negative values at any finite interval.

White noise can be considered a weak derivative of the standard Wiener process (Särkkä & Solin, 2019). Therefore, the system (6a) - (6b) can be expressed as the following state-space model.

$$dX = A(X,t)dt + B(X,t)dW \quad \dots(7a)$$

$$y = CX \quad \dots(7b)$$

Definition: Wiener process

The standard Wiener process $\{w(t) | t \in [0, T]\}$ is a continuous stochastic process that satisfies the following conditions (Gardiner, 1985; Särkkä & Solin, 2019).

- 1. Any increment $\Delta w_k = w(t_{k+1}) - w(t_k)$ is a zero mean multivariate Gaussian random variable with covariance $T\Delta t_k$ where T is the diffusion matrix of the Wiener process and $\Delta t_k = t_{k+1} - t_k$.
- 2. The increments are independent where the time spans of increments are not overlapped.
- 3. $w(0) = 0$ with probability one.

The solution of state-space model (7a) - (7b) is a random process (Gardiner, 1985; Allen, 2007; Särkkä & Solin, 2019). Since with a different realization of the noise process we get a different solution, the particular solutions of the equations are not often of interest, but instead, the aim is to determine the statistics of the solutions over all realizations. In the context of SDEs, the term $A(X,t)$ of (7a) is called the drift function which determines the nominal dynamics of the system, and $B(X,t)$ is the dispersion matrix which determines how the noise enters the system.

In our work, we consider a coupled system of differential equations which has a white noise driven forcing term of the following form (Särkkä & Solin, 2019).

$$\ddot{x}(t) + 2\xi_1\omega_1\dot{x}(t) + \omega_1^2x(t) = \eta(t) \quad \dots(8)$$

$$\ddot{y}(t) + 2\xi_2\omega_2\dot{y}(t) + \omega_2^2y(t) = \sigma\dot{x}(t) \quad \dots(9)$$

Here, σ is a constant and standard white noise is denoted by η and represents the variation of wind speed. Then, following the relation between the white noise and the Wiener process in the context of SDEs, the above system of (8) - (9) can be expressed in the state space form of (7a) - (7b),

where

$$X = \begin{pmatrix} x \\ \dot{x} \\ y \\ \dot{y} \end{pmatrix}; A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\omega_1^2 & -2\xi_1\omega_1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \sigma & -\omega_2^2 & -2\xi_2\omega_2 \end{pmatrix};$$

$$B = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}; C = (0 \quad 0 \quad 0 \quad 1)$$

The transfer function of the above state space model is obtained by applying the Laplace transform with zero initial conditions.

$$H(s) = \frac{\sigma s^2}{(s^2 + 2\xi_1\omega_1s + \omega_1^2)(s^2 + 2\xi_2\omega_2s + \omega_2^2)} \quad \dots(10)$$

By substituting $s = i\omega$ in the equation (10), the power spectral density function of the system is given by

$$S(\omega) = H(i\omega)H(-i\omega) =$$

$$\frac{\sigma^2\omega^4}{((\omega_1^2 - \omega^2)^2 + 4\xi_1^2\omega_1^2\omega^2)((\omega_2^2 - \omega^2)^2 + 4\xi_2^2\omega_2^2\omega^2)} \quad \dots(11)$$

According to the Wiener-Khinchin theorem stated as theorem 1 in Appendix B, the autocorrelation function is the Fourier inversion of the spectral density function. The Fourier inversion was obtained by substituting $p = i\omega$ (Lee, 1970).

$$S(p) = \frac{\sigma^2 p^4}{\left[(\omega_1^2 + p^2)^2 - 4\xi_1^2\omega_1^2 p^2 \right] \left[(\omega_2^2 + p^2)^2 - 4\xi_2^2\omega_2^2 p^2 \right]}$$

By decomposing $S(p)$ into partial fractions and taking the Fourier inversion, we have

$$\mathcal{F}^{-1} \left\{ \frac{p}{p^2 + 2\zeta_1 \omega_1 p + \omega_1^2} \right\} = u(t) e^{-\zeta_1 \omega_1 t} \left(\cos \left(\omega_1 \sqrt{1 - \zeta_1^2} \right) - \frac{\zeta_1}{\sqrt{1 - \zeta_1^2}} \sin \left(\omega_1 \sqrt{1 - \zeta_1^2} \right) \right) \dots(12)$$

$$\mathcal{F}^{-1} \left\{ \frac{1}{p^2 - 2\zeta_1 \omega_1 p + \omega_1^2} \right\} = \frac{1}{\omega_1 \sqrt{1 - \zeta_1^2}} u(t) e^{-\zeta_1 \omega_1 t} \sin \left(\omega_1 \sqrt{1 - \zeta_1^2} \right) \dots(13)$$

$$\mathcal{F}^{-1} \left\{ \frac{p}{p^2 - 2\zeta_2 \omega_2 p + \omega_2^2} \right\} = -u(-t) e^{-\zeta_2 \omega_2 t} \left(\cos \left(\omega_2 \sqrt{1 - \zeta_2^2} \right) - \frac{\zeta_2}{\sqrt{1 - \zeta_2^2}} \sin \left(\omega_2 \sqrt{1 - \zeta_2^2} \right) \right) \dots(14)$$

$$\mathcal{F}^{-1} \left\{ \frac{1}{p^2 + 2\zeta_2 \omega_2 p + \omega_2^2} \right\} = \frac{1}{\omega_2 \sqrt{1 - \zeta_2^2}} u(t) e^{-\zeta_2 \omega_2 t} \sin \left(\omega_2 \sqrt{1 - \zeta_2^2} \right) \dots(15)$$

where $u(t)$ is the Heaviside step function. Hence, the autocorrelation function of y , given by equation (7b), is found to be,

$$\kappa(t) = \frac{\sigma^2}{2\pi} (A_1 e^{-\zeta_1 \omega_1 t} \cos (t\omega_{d_1} + \phi_1) + A_2 e^{-\zeta_2 \omega_2 t} \cos (t\omega_{d_2} + \phi_2)) \dots(16)$$

Each term A_1, A_2, ϕ_1 and ϕ_2 is a function of the system parameters $\zeta_1, \omega_1, \zeta_2$ and ω_2 . In order to estimate the system parameters, the normalized autocorrelation given below was used.

$$g(t) = \frac{(A_1 e^{-\zeta_1 \omega_1 t} \cos (t\omega_{d_1} + \phi_1) + A_2 e^{-\zeta_2 \omega_2 t} \cos (t\omega_{d_2} + \phi_2))}{(A_1 \cos \phi_1 + A_2 \cos \phi_2)} \dots(17)$$

The system parameters were estimated by comparing the normalized autocorrelation function $g(t)$ with the empirical autocorrelation function of the residual process v .

RESULTS AND DISCUSSION

To illustrate the validity of the developed model, the wind speed dataset measured at Kokkilai in the North Eastern coast of Sri Lanka was used. As described in section 2, the wind speed dataset was first filtered using (4) and (5) and then the residual was modelled using the coupled second-order SDEs (8) and (9). The theoretically constructed normalized autocorrelation function (17) that corresponds to (8) and (9) and the output was fitted to the normalized autocorrelation function of the residual data in order to estimate the model parameters. The empirical correlogram of Figure 3(b) exhibits rapid decay in the first 100 hours and this decides the time scale of the underlying process (Brett & Tuller, 1991; Zarate-Minano, Mele & Milano, 2016). Thus, the parameters of the normalized theoretical correlogram of equation (17) were estimated using the empirical autocorrelation values of first 100 hours. The least squares errors method was used to fit the parameters. Moreover, the variance of a stationary process is equal to the initial value of the corresponding autocorrelation function (Chatfield, 2004). Therefore, the value of the parameter σ , can be fitted using the following relationship.

$$\begin{aligned} Var(v) &\approx \kappa(0) \\ Var(v) &\approx \frac{\sigma^2}{2\pi} (A_1 \cos \phi_1 + A_2 \cos \phi_2) \dots(18) \end{aligned}$$

where $Var(v)$ is the variance of the residual process.

Using the estimated parameter values, the linear system of SDEs of equation (8) and (9) was simulated using the Euler-Maruyama numerical scheme. Then, the model-generated output was compared with the residual process using 30 sample realizations. Figure 5 shows that the probability density of the model matches that of the residual process v very closely.

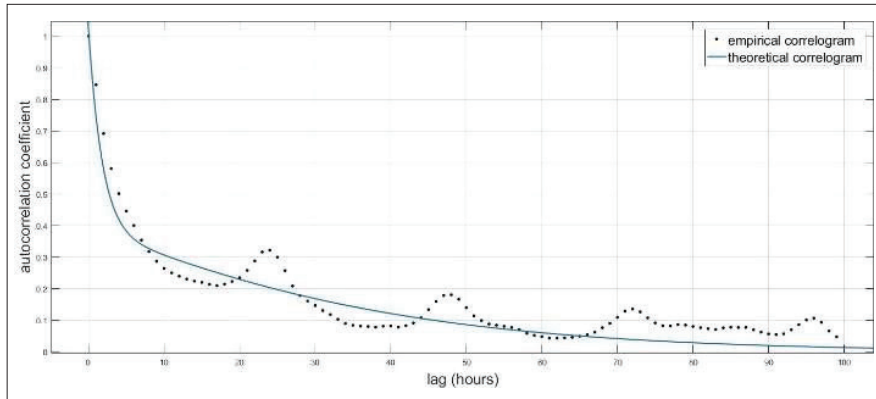


Figure 4: Approximation of empirical correlogram of residual data: $\xi_1=0.9812$, $\omega_1=0.3321$, $\xi_2=0.9907$, $\omega_2=0.02845$ and $\sigma=0.4818$.

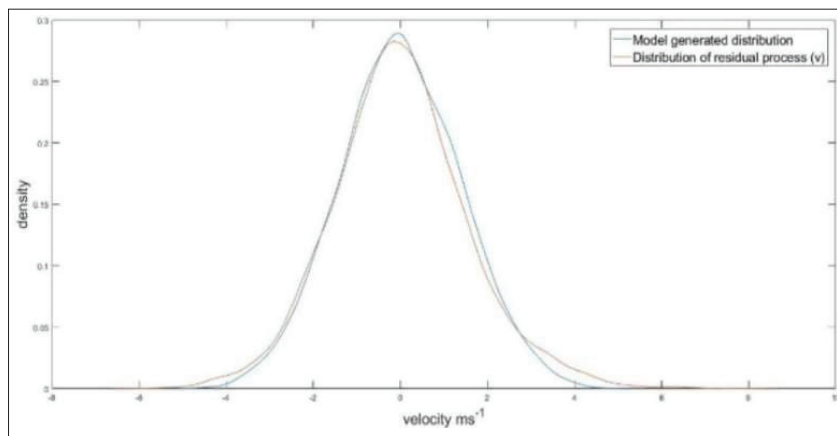


Figure 5: Comparison of probability distributions of SDE model generated stationary process and residual process

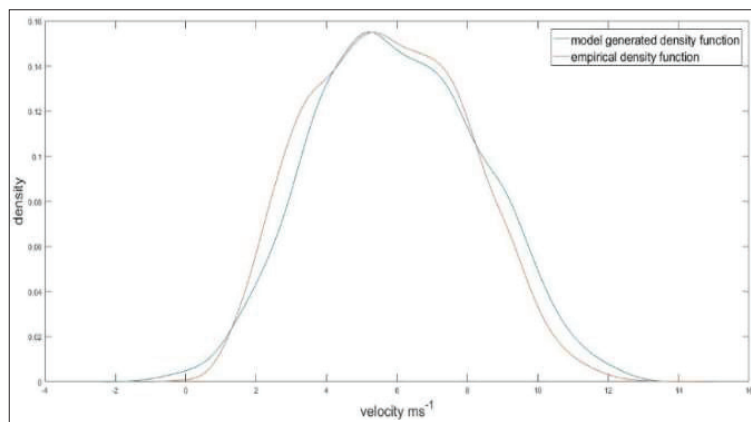


Figure 6: Comparison of final output of SDE model generated probability distribution of hourly wind speeds and that of empirical data

Next, the identified non-stationary component given by (4) was added to the model-generated outcome of the residual process by applying the inverse data transform corresponding to (5). Then the resulting wind speed distribution was compared with the empirical distribution of hourly wind speed distribution in Figure 6. It can be seen that our novel approach is capable of generating a wind speed distribution that is very close to those of the distribution of empirical data.

In addition to the comparison of the probability distributions, the comparison of the autocorrelation functions can be used to verify the level of accuracy.

The comparison of autocorrelation functions is given in Figure 7 and shows excellent agreement between the model and empirical data.

Figure 8 shows the comparison of the probability distributions based on the proposed method with a method based on a single one-dimensional SDE. Probabilistic models based on a single one-dimensional SDE have been considered in many situations (Bibby et al., 2005; Calif, 2012; Zarate-Minano et al., 2016). Figure 8 demonstrates that the model proposed in this paper outperforms the models that use a single one-dimensional SDE.

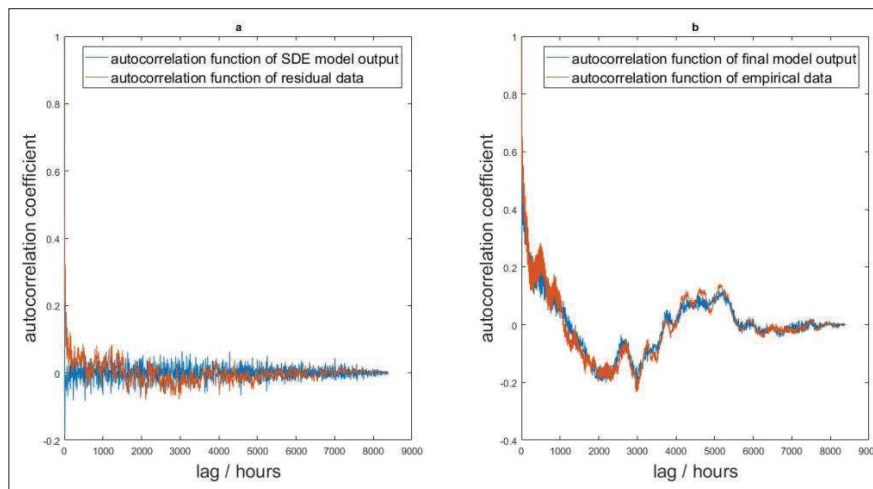


Figure 7: Comparison of correlograms

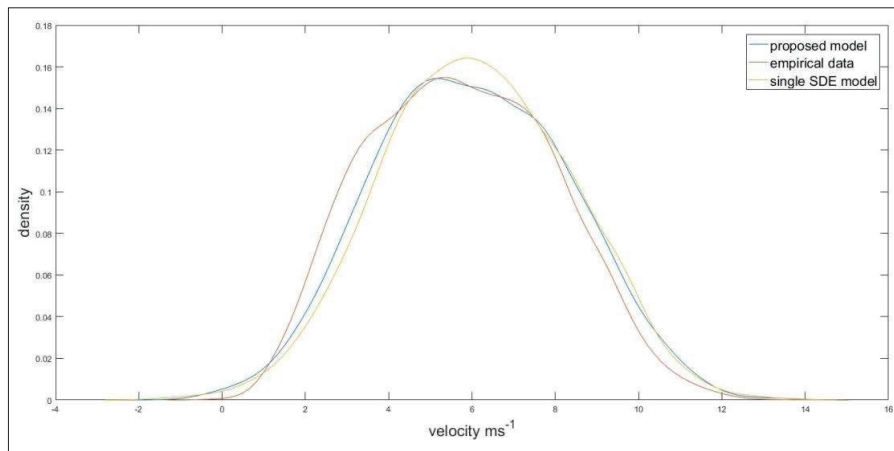


Figure 8: Comparison of distributions of single SDE model and empirical distribution and the proposed model

CONCLUSION

SDE-based approaches for modelling statistical properties of wind speed data are comparatively new. In existing work reported in the literature, probabilistic models are developed based on a single SDE of the standard form. In contrast, in this work, we have developed a novel approach that uses two coupled linear oscillators to model the statistical properties of appropriately filtered hourly wind speed data. The coupled system of SDEs generates a strictly stationary Gaussian process. Since wind speed data usually exhibits complex non-stationary features, we propose the use of a filter to remove these effects from the original wind speed data. The correlogram of the residual process is then fitted to the normalized theoretical autocorrelation function of the coupled SDE system.

The coupled SDE model consists of five unknown parameters. Four of them were estimated by curve fitting the theoretical normalized autocorrelation function to the empirical autocorrelation values, using the least squares method, while the other parameter was computed using these estimated parameter values.

The developed model was used to predict the statistical properties of wind speed data measured at Kokkilai, a coastal point located in the North Eastern shore of Sri Lanka. The simulation results show that the statistical properties of the simulated wind sequence are very close to those of the measured wind sequence. In particular, the model-generated correlogram of our method successfully captures almost all the features of the empirical correlogram. This is a significant improvement in comparison to those predicted by existing methods.

Acknowledgement

The authors acknowledge the support provided by the Sustainable Energy Authority (SEA) of Sri Lanka in obtaining wind speed data.

REFERENCES

Allen E. (2007). *Modeling with Ito Stochastic Differential Equations*. Springer, Dordrecht, The Netherlands.
DOI: <https://doi.org/10.1007/978-1-4020-5953-7>

- Bibby B.M., Skovgaard I.M. & Sørensen M. (2005). Diffusion-type models with given marginal distribution and autocorrelation function. *Bernoulli* **11**(2): 191–220.
DOI: <https://doi.org/10.3150/bj/1116340291>
- Brett A.C. & Tuller S.E. (1991). The Autocorrelation of hourly wind speed observations. *Journal of Applied Meteorology* **30**(6): 823–833.
DOI: [https://doi.org/10.1175/1520-0450\(1991\)030<0823:TAOHWS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1991)030<0823:TAOHWS>2.0.CO;2)
- Calif R. (2012). PDF models and synthetic model for the wind speed fluctuations based on the resolution of Langevin equation. *Applied Energy* **99**: 173–182.
DOI: <https://doi.org/10.1016/j.apenergy.2012.05.007>
- Chatfield C. (2004). *The Analysis of Time Series: An Introduction*. CRC Press, Florida, USA.
- Fernando B.P.W. & Hausenblas E. (2018). Nonlinear filtering with correlated Lévy noise characterized by copulas. *Brazilian Journal of Probability and Statistics* **32**(2): 374–421.
DOI: <https://doi.org/10.1214/16-BJPS347>
- Gardiner C.W. (1985). *Handbook of Stochastic Methods*, pp. 2–20, Springer, Berlin, Germany.
- Hausenblas E., Fahim K. & Fernando P.W. (2021). A particle filter for nonlinear filtering with Lévy jumps. *International Journal of Applied Mathematics* **34**(5): 817.
DOI: <https://doi.org/10.12732/ijam.v34i5.1>
- Lee S.Y. (1970). *Explicit Inverse Fourier Transformation of a Rational Function and a New Theorem*. Bellcomm Inc. Washington, DC, USA.
- Lindgren G., Rootzén H. & Sandsten M. (2013). Stationary Stochastic Processes. Theory and Applications. In: *Texts in Statistical Science Series*, pp. 347. CRC Press, Boca Raton, FL, USA.
- Mao X. (2007). *Stochastic Differential Equations and Applications*. 2nd edition, Elsevier, Amsterdam, Netherlands.
DOI: <https://doi.org/10.1533/9780857099402>
- Särkkä S. & Solin A. (2019). *Applied Stochastic Differential Equations*. Institute of Mathematical Statistics Textbooks, Cambridge University Press, UK.
DOI: <https://doi.org/10.1017/9781108186735>
- Shumway, Robert H. & Stoffer, David S. (2016). *Time Series Analysis and Its Applications with R Examples*, 4th Edition, Springer, Berlin, Germany.
DOI: <https://doi.org/10.1007/978-3-319-52452-8>
- Zárate-Miñano R., Mele F.M. & Milano F. (2016). SDE-based wind speed models with Weibull distribution and exponential autocorrelation. *IEEE Power and Energy Society General Meeting (PESGM)* **2016**: 1–5.
DOI: <https://doi.org/10.1109/PESGM.2016.7741754>
- Zhu X. & Genton M.G. (2012). Short-term wind speed forecasting for power system operations. *International Statistical Review* **80**(1): 2–23.
DOI: <https://doi.org/10.1111/j.1751-5823.2011.00168.x>

Appendix A

Linear time invariant stochastic differential equations

A time invariant SDE can be expressed in the standard form as follows,

$$dx = Ax(t)dt + Bdw(t) \quad \dots(A.1)$$

Then the formal solution of the above equation (A.1) is written as,

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}B(\tau)dW(\tau) \quad \dots(A.2)$$

where e^{At} is the matrix exponential function. Since the noise process is Gaussian forcing, a linear differential equation can be considered as a linear operator acting on the noise process (and the initial conditions) and the solution is also Gaussian (Gardiner, 1985; Allen, 2007; Särkkä & Solin, 2019).

• **Expectation and correlation**

Since the white noise process has zero mean, taking expectations from both sides of equation (A.2) gives,

$$m(t) = E(x(t)) = \exp(At)E(x(t_0)) \quad \dots(A.3)$$

which is the expected value of the SDE solutions over all realizations of noise. The corresponding covariance function is calculated by using the delta-correlation property of white noise as follows,

$$P(t) = E[(x(t)-m(t))(x(t)-m(t))^T]$$

$$P(t) = \exp(At)P_0 \exp(A^T t) + \int_{t_0}^t \exp(A(t-\tau))BQB^T w(\tau) \exp(A^T(t-\tau))d\tau \quad \dots(A.4)$$

where $P_0 = E[x(0)x(0)^T]$. By differentiating the above expressions for mean and covariance, the governing differential equations for the mean and covariance are given by

$$\frac{dm(t)}{dt} = Am(t) \quad \dots(A.5)$$

$$\frac{dP(t)}{dt} = AP(t) + P(t)A^T + BQB^T \quad \dots(A.6)$$

Moreover, it can be seen that the solutions of equation (A.1) generate a multivariate Gaussian distribution (Gardiner, 1985; Särkkä & Solin, 2019).

It should be noted when $P(t)$ is a constant, (A.6) is reduced to a Lyapunov equation. Then, the steady covariance function can be obtained by following the context of a matrix Lyapunov equation (Mao, 2007; Särkkä & Solin, 2019).

Appendix B

• **Fourier analysis of linear time invariant stochastic differential equations**

One way to study linear time invariant SDEs is the Fourier domain (Lindgren et al., 2002; Särkkä & Solin, 2019). In this case, the power spectral density which is the squared absolute value of the Fourier transform of the process is defined as follows,

$$S_X(\omega) = X(i\omega)X^T(-i\omega) \quad \dots(B.1)$$

The power spectral density Q of a multi-dimensional white noise process is also defined accordingly.

$$Q = W(i\omega)W^T(-i\omega)$$

The spectral density is an efficient tool for estimating the covariance function generated as a result of SDE. Suppose $x(t)$ is a stationary multi-dimensional stochastic process with zero mean. Then the corresponding covariance function is defined as,

$$C_x(\tau) = E[x(t)x^T(t+\tau)]$$

The covariance function of a wide sense stationary process is only dependent on the time lag (τ). The Wiener-Khinchin theorem (Lindgren et al., 2002; Särkkä & Solin, 2019) states the Fourier inversion of the power spectral density function of a given wide sense stationary process provides the corresponding covariance function as follows,

Theorem 1 (Wiener-Khinchin theorem):

The covariance of a wide sense stationary random process has a spectral decomposition given by the power spectrum of that process, i.e., $C_x(\tau) = \mathcal{F}^{-1}\{S_X(\omega)\}$.

Therefore, we have the autocorrelation function for a purely white noise process as follows,

$$C_W(\tau) = \mathcal{F}^{-1}[Q] = Q\mathcal{F}^{-1}[1] = Q\delta(\tau)$$

- **Covariance of the solution of linear time invariant stochastic differential equations**

The spectral density function of resultant process $\mathbf{x}(t)$ of equation (A.1) can be given in the matrix version with the assumption of zero initial conditions. Note that the stationary stage can only exist if the matrix corresponds to a stable system, which means that all its eigenvalues have negative real parts. The Fourier transform of $\mathbf{x}(t)$ is,

$$X(i\omega) = ((i\omega)I - A)^{-1} BW(i\omega)$$

where $W(i\omega)$ is the formal Fourier transform of the white noise $w(t)$. It should be noted that the above Fourier transform does not strictly exist, as white noise is not square integrable. Then the spectral density of $\mathbf{x}(t)$ is given by the matrix,

$$S_X(\omega) = (A - (i\omega)I)^{-1} BQB^T (A + (i\omega)I)^{-T} \quad \dots(B.2)$$

Hence the covariance function is,

$$C_X(\tau) = \mathcal{F}^{-1}\{(A - (i\omega)I)^{-1} BQB^T (A + (i\omega)I)^{-T}\} \quad \dots(B3)$$

This provides useful means to estimate the autocovariance function of a solution to a SDE without explicitly solving the equation.

RESEARCH ARTICLE

Manufacturing and Industrial Engineering

Failure prediction of solid resilient tyres due to kerb impact: A finite element modelling approach

WAAS Premarathna^{*1}, JASC Jayasinghe², CD Senanayake¹, KK Wijesundara² and P Gamage¹

¹ Department of Manufactory and Industrial Engineering, University of Peradeniya, Peradeniya, Sri Lanka.

² Department of Civil Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya, Sri Lanka.

Submitted: 21 June 2021; Revised: 25 May 2022; Accepted: 24 June 2022

Abstract: Solid resilient tyres frequently impact on kerbs or obstacles when they are operated in the construction and transportation sectors. These sudden impacts can generate high stresses and thereby damage the tyre. The factors which cause such failures are not easy to capture experimentally due to their complexity and high experimental cost. Hence, this study focused on using the finite element method to model the solid resilient tyre, generate stresses, and identify the failures and regions that are susceptible to damage. Initially, a tyre static model was developed and validated using laboratory experimental data obtained from the industry. The validation results showed that the numerical results are in good agreement with the experimental data. Subsequently, the model was extended to incorporate impact simulation. The simulation considered tyre motion and impact on three different types of kerbs - angular, circular and square-shaped. Simulation results showed that high stresses occur mainly in the side walls of the solid tyre while high contact pressure and high in-plane shear stresses were observed at the tread layer especially when it moves on square-type kerbs. Hence, there is a high tendency for tyre failures to occur due to the side wall cracks at the base layer. Furthermore, sudden wear and damage due to chunking can be expected on the tread layer of the solid resilient tyre.

Keywords: Hyper-elastic models, impact simulations, non-linear finite element model, solid resilient tyres.

INTRODUCTION

Solid resilient tyres are mainly utilized in forklift vehicles and often operate in harsh environmental conditions while bearing heavy loads. In these rugged environmental

conditions, the tyres experience high stresses, which may lead to sudden wearing and damage due to chunking and rupturing that occur when the tyre impacts on kerbs or obstacles. High experimental cost and time are the main impediments to investigating such behaviour of solid tyres using experimental setups. In this study, finite element (FE) numerical modelling approaches have been utilized to model and predict the solid tyre impact behaviour. The main components of solid resilient tyres and their functions are illustrated in Figure 1. The base layer and embedded bead bundles of the solid tyre ensure good gripping contact with the tyre rim. The internal heat generation and traction of the tyre mainly occur in the cushion layer and tread layer, respectively.

Most of the literature has focused on the numerical analyses of pneumatic tyres (Han *et al.*, 2019; Azizi, 2020). Even for those analyses, the impact analyses were performed for rectangular cleats or kerbs only. Wei and Olatunbosun (2014) studied the impact analysis of a pneumatic tyre and their findings showed that when the height of the rectangular obstacles was increased the amplitude of the vertical force and longitudinal force on the tyre and the tyre deformation were also increased. Suvanjumrat and Rugsaj (2020) investigated impact force and vertical deformation of non-pneumatic tyres when impacted on rectangular cleats with three different heights. Moreover, impact forces on the axle of a pneumatic tyre against a rectangular kerb were analysed by including the suspension system with a full

* Corresponding author (asaranga225@gmail.com;  <https://orcid.org/0000-0002-5336-2932>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

car model (Shiraishi *et al.* 2000). Furthermore, Reida *et al.* (2007) used a detailed truck numerical model to study the pneumatic tyre impacts on a rectangular kerb as well as its behaviour when it was moving over rocks and culvert grates. Kamoulakos and Kao (1998) analysed the similarities of numerical and experimental readings on vertical and horizontal spindle forces when a pneumatic tyre moved over a rectangular cleat. In Chang and Yang (2009), the generation of ductile fractures on a wheel has been observed numerically as well as experimentally, by impacting it on a solid rectangular cube. Their results showed that the damage to the wheel could be accurately predicted using maximum strain energy density based on the plastic work due to the impact. Duni *et al.* (2008) reported the contact pressure behaviour of a complete car numerical model with pneumatic tyres when it was moving on a pothole and fatigue reference track. Han *et al.* (2019) conducted a transient dynamic analysis to study the pneumatic tyre behaviour when it impacted on a rectangular cleat. The results showed that large contact pressure accumulated in the two edges of the tread layer where it made contact with the road. In a more recent study of Han *et al.* (2020), they utilized a quarter vehicle model with pneumatic tyres to investigate the structural vibration and acoustic radiation under rectangular cleat impact excitations. They observed localized contact pressure on the cleat impact area, and widely distributed contact pressure around the contact region in the tyre when the tyre was on the cleat and just before it passed

through the cleat. In Azizi (2020), a cleat impact experimental test was performed by using a pneumatic tyre and test drum. In this study, the frequency spectrum of the cleat impact was obtained as a function of tyre contact patch length, cleat size, and rolling speed.

Few studies have been performed to analyse static and dynamic behaviours of solid resilient tyres. Suripa and Chaikittiratana (2008), investigated the strain energy density and stress distribution of solid tyres under static conditions by using simplified 3D FE solid models. The results showed that the highest strain energy accumulated in the cushion layer of the solid tyre. Dechwayukul *et al.* (2010) also conducted laboratory tests to analyse solid tyre durability to evaluate tyre service life. Phromjan and Suvanjumrat (2018a) presented the development of a detailed 3D FE model to analyse the deformations and stress distributions of a solid tyre under different load levels in static conditions. In their results, high stress was observed in the base layer of the solid tyre. Phromjan & Suvanjumrat (2018b) conducted lab experiments to compare the contact forces and vibration analyses of a rolling solid tyre and a pneumatic tyre under different rolling velocities. It was identified that higher contact forces and vibration levels were generated in solid tyres than in pneumatic tyres. To the best of the authors' knowledge, there have been no dynamic numerical FE models developed to study the impact of solid resilient tyres on kerbs and predict their failures.

The structure of this paper is as follows: In the methodology section, the development procedures for static and impact models of a solid resilient tyre are presented. The next section presents the model validation, analysis of results, and discussion. The final section provides concluding remarks and possible future research directions.

MATERIALS AND METHODS

3D FE model of the solid tyre

A CAD model of a solid resilient tyre was first developed as shown in Figure 2. A detailed numerical model of the solid resilient tyre was then developed for the static analysis using Abaqus 6.14 software (ABAQUS Inc., 2014) and the procedure is shown in Figure 3.

In the detailed numerical model, the tyre constraints were applied on nodes which were in between the three rubber layers. Surface-to-surface contact was incorporated to assign interactions between contact

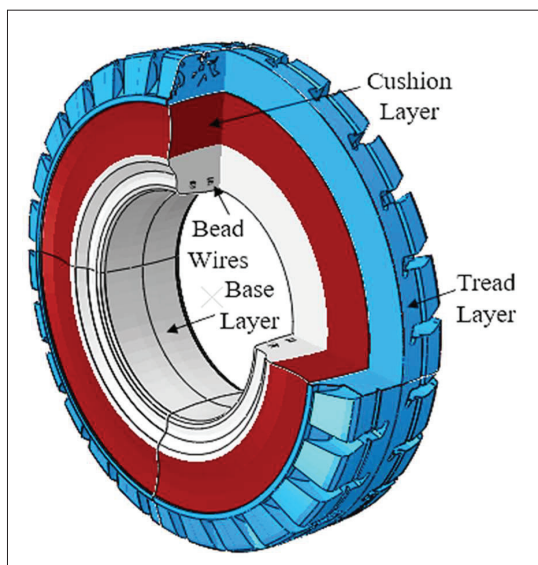


Figure 1: Components of solid resilient tyre

surface and tread layer of the tyre. A mesh convergence analysis was performed to identify the suitable mesh size of the tyre, as shown in Figure 4. The mesh of all components in the tyre was modelled by using 8-node 3D hexahedral elements and the converged model contains

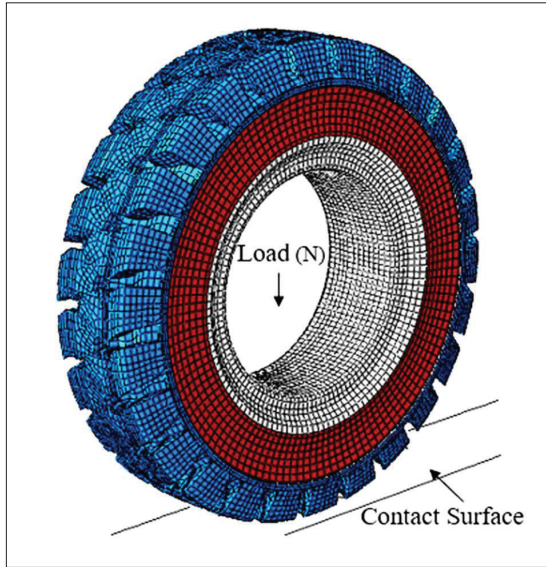


Figure 2: Static numerical model of solid resilient tyre

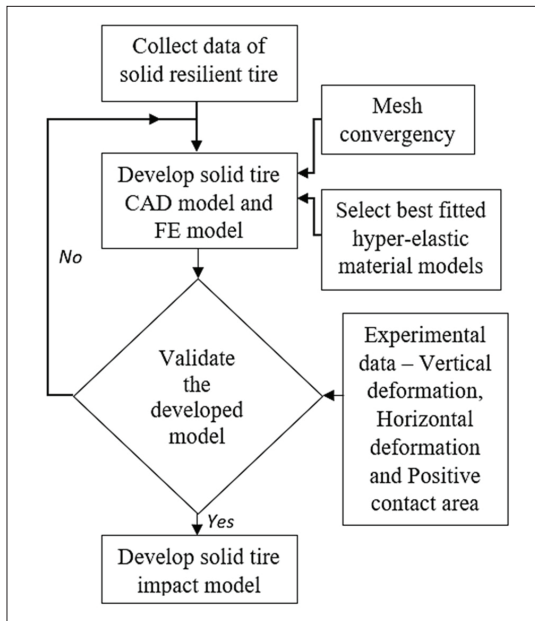


Figure 3: Methodology flow chart of static FE model development

in total 71685 solid elements. The inner surface of the base layer was assigned as a rigid body to represent the tyre rim. Loads were applied to the centre of the tyre.

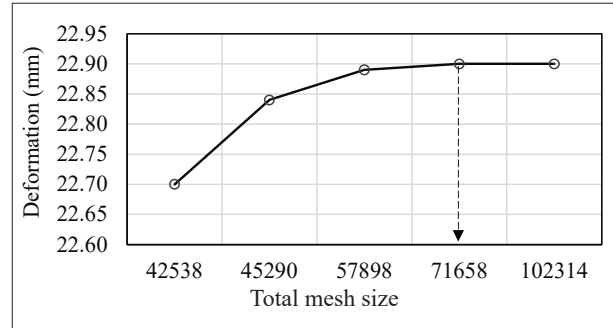


Figure 4: Mesh convergence with vertical deformation

Selection of best-fitted hyper-elastic material models

Suitable hyper-elastic material models are required to develop the static model of the solid tyre. The best-fitted material models are utilized to describe the mechanical behaviour of rubber layers in the solid tyre. In this study, the best-fitted material models for base layer, cushion layer, and tread layer were selected by combining the curve fitting method and three standard measures of accuracy (MAPE: Mean Absolute Percentage Error; MAD: Mean Absolute Deviation; and MSD: Mean Squared Deviation). Based on the results of curve fitting, the Mooney-Rivlin, Ogden, and Yeoh hyper-elastic material models were obtained as the best-fitted material models of the base layer, cushion layer and tread layer, respectively. Equations 1, 2, and 3 represent the general form of the Mooney-Rivlin material model, Yeoh material model, and Ogden material model, respectively (Mooney, 1940; Ogden, 1972; 1973; Yeoh, 1990; 1997).

$$U = C_{10}(\bar{I}_1 - 3) + C_{01}(\bar{I}_2 - 3) + \frac{1}{D_1}(J_{el} - 1)^2 \quad \dots(1)$$

$$U = \sum_{i=0}^3 C_{i0}(\bar{I}_1 - 3)^i + \sum_{i=1}^3 \frac{1}{D_i}(J_{el} - 1)^{2i} \quad \dots(2)$$

$$U = \sum_{i=1}^N \frac{2\mu_i}{\alpha_i^2}(\bar{\lambda}_1^{\alpha_i} + \bar{\lambda}_2^{\alpha_i} + \bar{\lambda}_3^{\alpha_i} - 3) + \sum_{i=1}^N \frac{1}{D_i}(J_{el} - 1)^{2i} \quad \dots(3)$$

Where:

λ_1, λ_2 and λ_3 : Principle stretches of principle directions; C_{ij} and D_{ij} : Material constants which control the shear behaviour and bulk compressibility of the material; J_{el} : Elastic volume ratio; \bar{I}_n : Invariants; μ_i and α_i : Temperature-dependent material coefficients

Table 1 presents the best-fitted hyper-elastic material models and their coefficient values for each rubber layer of the solid tyre.

Dynamic impact simulation of the solid tyre

The validated static numerical model of the solid resilient tyre was utilized to implement the impact numerical model of the tyre. Figure 5 illustrates the steps that were followed to develop the tyre impact model. The

Abaqus/Explicit solver has been utilized to generate this model. The *Abaqus* results transferring capability was employed to transfer results from *Abaqus/Std.* to *Abaqus/Explicit*. Furthermore, relevant boundary conditions and velocities were also applied to the tyre model. Since solid tyres usually move short distances at low velocities, 5 kmh⁻¹ and 10 kmh⁻¹ translational velocities were accounted as impact velocities of the tyre model. Finally, the generation of maximum stress in the rubber layers, maximum contact pressure, and shear stress due to friction were determined at the time of tyre impacts on three types of kerbs. In these impact models, the tread layer of the tyre was ignored to avoid convergence issues and to reduce the simulation time.

The relaxation property of each rubber layer was assigned by considering the viscoelastic behaviour of the rubber material. The corresponding viscoelastic properties were obtained by utilizing the dynamic mechanical analyzer (DMA) as shown in Figure 6 and frequency sweep test data.

Table 1: Coefficients of best-fitted hyper-elastic material models for tyre base, cushion and tread layers

Rubber layer of the tyre	Base			Cushion					Tread		
Material model	Mooney-Rivlin			Ogden					Yeoh		
Coefficients	C_{10}	C_{01}	μ_1	μ_2	μ_2	α_1	α_2	α_3	C_{10}	C_{20}	C_{30}
	0.6	2.5	-9.3	8.9	0.9	1.5	2.1	-3.7	0.7	-0.1	0.1

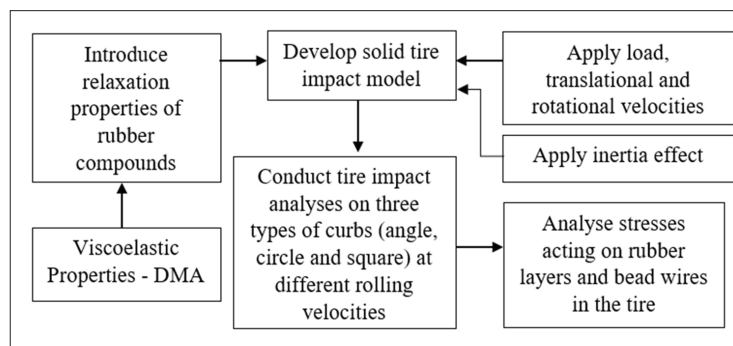


Figure 5: Flow chart of impact FE model development

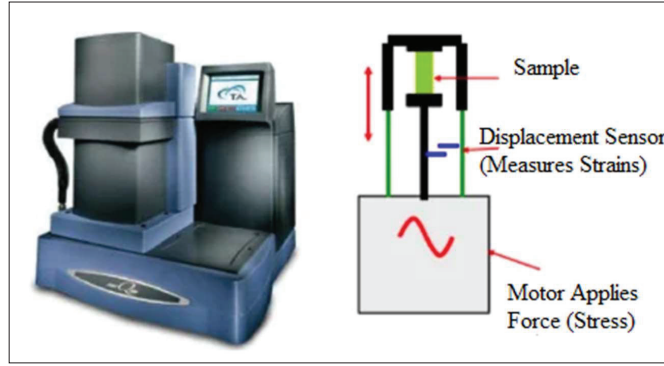


Figure 6: Dynamic mechanical analyser (Eindhoven University of Technology, 2021)

Table 2: Coefficients of Prony series for tyre base layer, cushion layer and tread layer

Rubber layer of the tyre	Three-term Prony coefficients					
	g_1	ψ_1	g_2	ψ_2	g_3	ψ_3
Base	0.2208	0.0003	0.1208	0.0146	0.0993	0.1890
Cushion	0.0016	0.0001	0.1099	0.0173	0.0603	0.2198
Tread	0.1545	0.0032	0.1223	0.0377	0.1253	1.8288

The corresponding storage modulus (G') and loss modulus (G'') data of the rubber materials were obtained using the frequency sweep test. Equations (4) and (5) express the relationship between Prony series coefficients related to storage modulus and loss modulus of the material, respectively.

$$G'(\omega) = G_0 \left\{ \left[1 - \sum_{n=1}^N g_n \right] + \left[\sum_{n=1}^N \frac{g_n \omega^2 \psi_n^2}{1 + \omega^2 \psi_n^2} \right] \right\} \quad \dots(4)$$

$$G''(\omega) = G_0 \left\{ \sum_{n=1}^N \frac{g_n \omega \psi_n}{1 + \omega^2 \psi_n^2} \right\} \quad \dots(5)$$

Where:

ω : Rolling frequency of the tyre

G_0 : Instantaneous shear modulus

g_n and ψ_n : Prony series coefficients

The obtained three-term Prony series coefficients are presented in Table 2. Subsequently, the tyre impact model was implemented using the validated static numerical model of the solid tyre as shown in Figure 7.

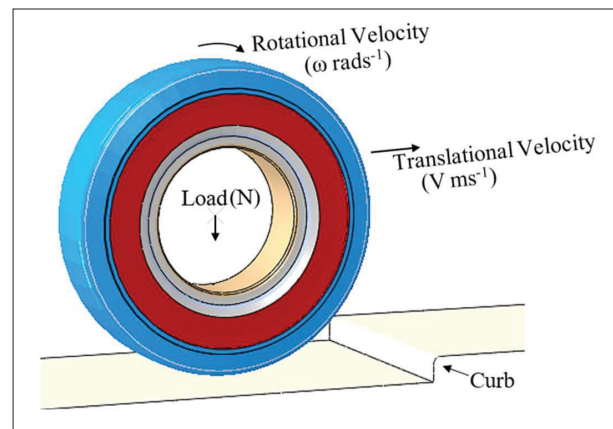


Figure 7: Impact numerical model of solid tyre

Selection of kerbs

According to the literature review, most of the impact analyses were performed for pneumatic tyres with square-type kerbs. Hence, it is necessary to select suitable kerb geometries to perform the impact analyses

on solid tyres. Equal height and three different types of kerb geometries were introduced onto the contact surface as shown in Figure 8. The kerbs were selected based on their practical usage (Hills Kerbs PTY Ltd., 2021). The following kerb types are usually utilized on the edges of pavements, roads and parking areas.

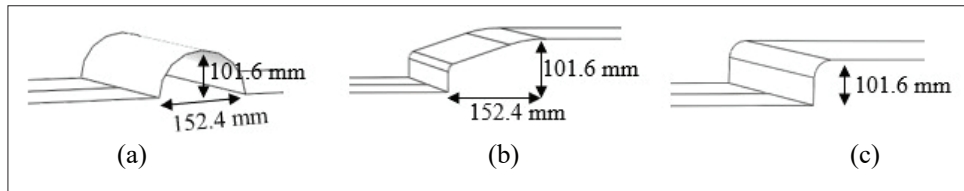


Figure 8: Three types of kerbs: (a) circular type, (b) angle type and (c) square type

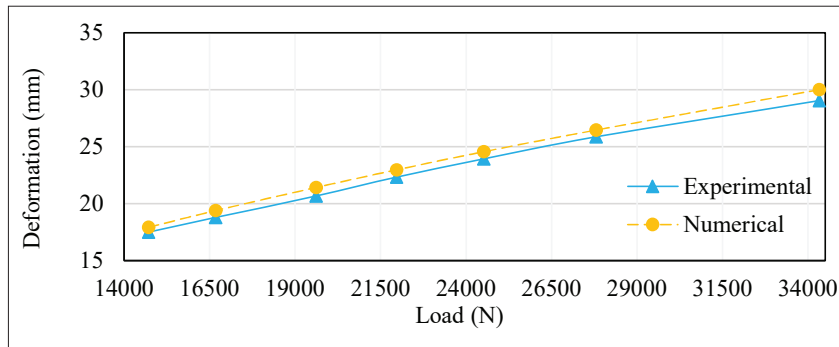


Figure 9: Experimental and numerical vertical deformation data of the solid tyre

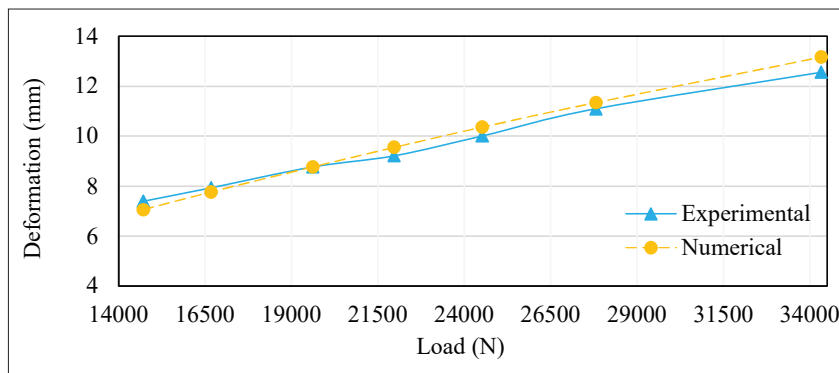


Figure 10: Experimental and numerical horizontal deformation data of the solid tyre

RESULTS AND DISCUSSION

Validation of the solid tyre numerical model

The developed static model was validated by applying different loads on it and comparing corresponding deformation readings with laboratory experimental data. Figures 9 and 10 illustrate the vertical and horizontal deformations of the numerical tyre model and actual tyre.

Table 3: Readings of three measures of accuracy

Type of the deformation	MAPE (%)	MAD	MSD
		(%)	(%)
Vertical	2.91	0.66	0.46
Horizontal	3.00	0.29	0.12

Table 3 shows the values of three measures of accuracies calculated between experimental and numerical results. According to the Table 3 results, the Mean Absolute Percentage Error (MAPE) value for vertical and horizontal deformations are 2.19% and 3.00%, respectively. The readings of Mean Absolute Deviation (MAD) and Mean Squared Deviation (MSD) also show very small values (<1%). Hence, the values of the three measures of accuracy results confirmed that the behaviour of the developed tyre model is in good agreement with the actual tyre behaviour.

Analysis of maximum stress generation on the solid tyre during the impacts

The maximum von Mises stress accumulation in the rubber components of the solid tyre under angle type, circle type, and square type kerb geometries is illustrated in Figures 11, 12, and 13, respectively. According to Figure 11, a maximum von Mises stress reading of 2.915 MPa was observed in the side wall rubber compound of the solid tyre when the tyre impacted on an angle-type kerb.

Furthermore, according to Figure 12, a maximum von Mises stress value of 2.861 MPa is detected in the side walls of the solid tyre when it impacted on a circle-type kerb.

Moreover, a maximum von Mises stress reading of 3.089 MPa is obtained in the side walls of the solid tyre when it impacted on a square type kerb, as shown in Figure 13.

Among these three types of kerbs, the solid tyre experiences a higher impact stress on the square-type kerb than on the other two kerb types. However, during the impact the fillet radius on the square-type kerb can significantly affect the generation of maximum stress in a solid tyre. Figures 14 and 15 illustrate the variation of maximum impact stress in the solid resilient tyre with 5 kmh⁻¹ and 10 kmh⁻¹ impact velocities, respectively. Figures 14 and 15 show that impact with the square-type kerb has resulted in a higher value of the maximum stress compared with angle and circle-type kerbs under both impact velocities.

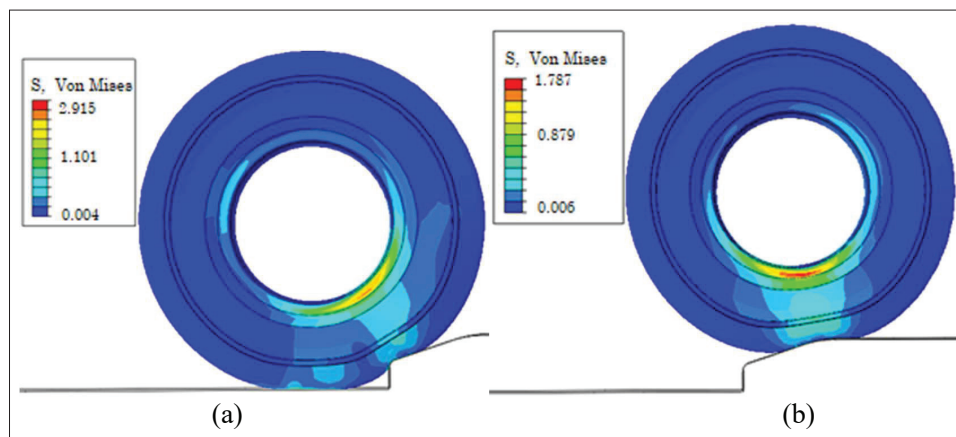


Figure 11: Stress generation of the solid tyre when it impacted on an angle type kerb (load 19614 N and velocity 10 kmh⁻¹)

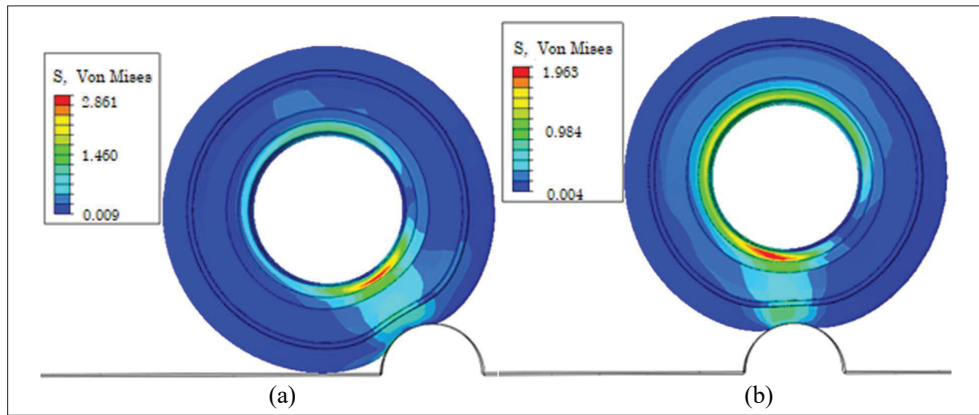


Figure 12: Stress generation of the solid tyre when it impacted on a circle type kerb (load 19614 N and velocity 10 kmh⁻¹)

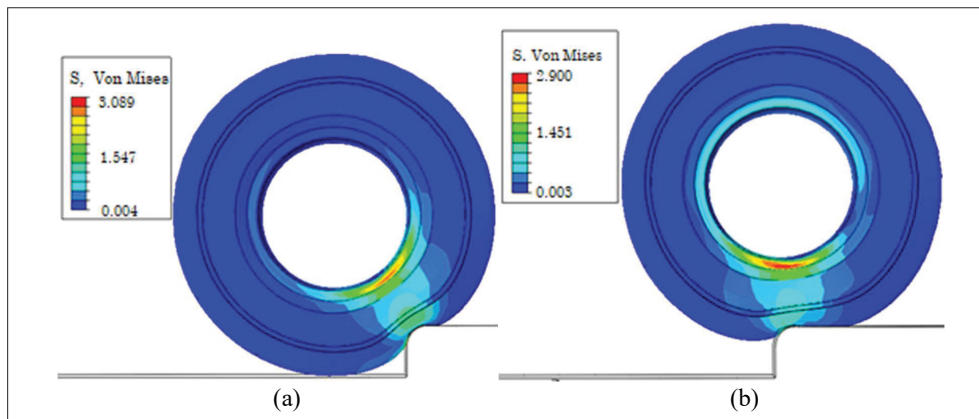


Figure 13: Stress generation of the solid tyre when it impacted on a square type kerb (load 19614 N and velocity 10 kmh⁻¹)

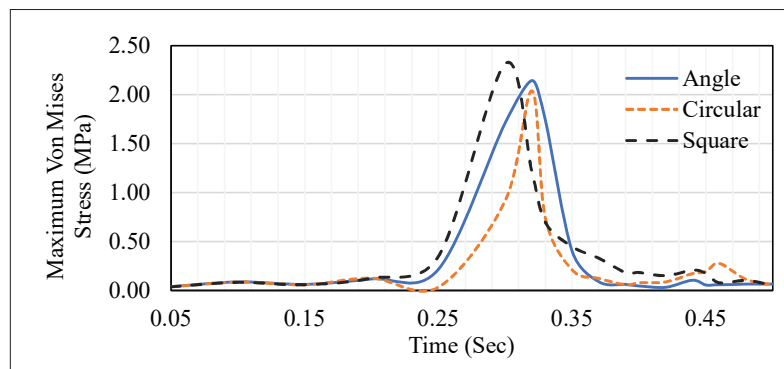


Figure 14: Maximum impact stress history of solid tyre rubber components with 5 kmh⁻¹ velocity

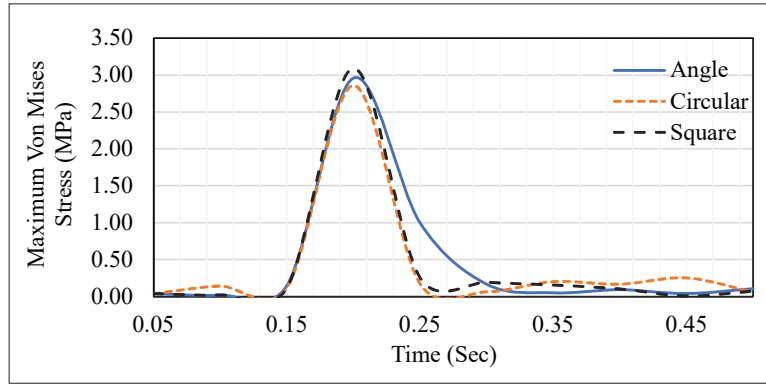


Figure 15: Maximum impact stress history of solid tyre rubber components with 10 kmh^{-1} velocity

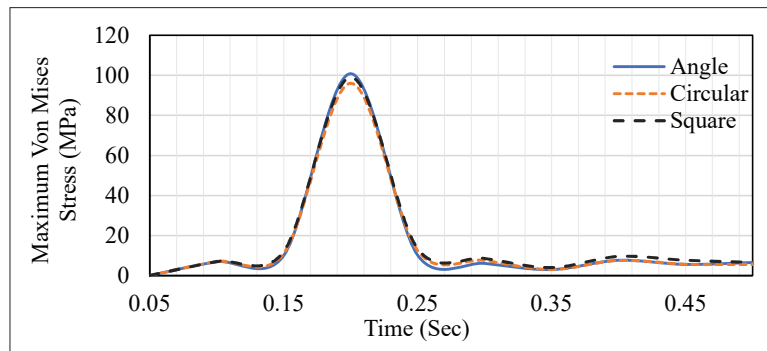


Figure 16: Maximum impact stress history of bead wires at velocity 10 kmh^{-1}

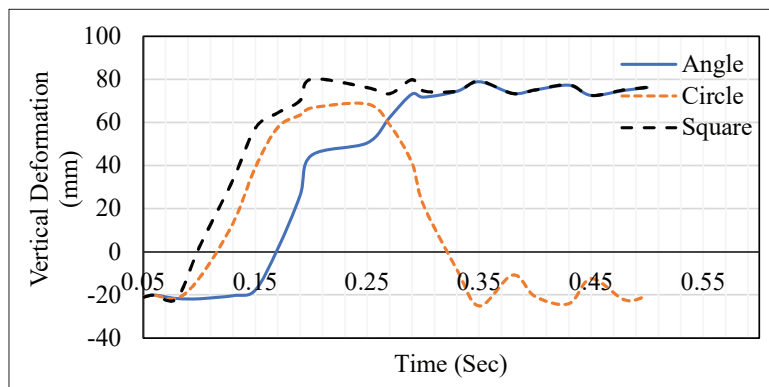


Figure 17: Deformation history of solid tyre at velocity 10 kmh^{-1}

Furthermore, it can be observed that the maximum value of the impact stress is increasing with increasing impact velocity. According to the above stress analyses, it can be seen that the high impact stresses lead to sidewall crack propagations in the solid tyre rubber components. Figure 16 illustrates the maximum stresses obtained in the bead wires of the tyre when it travels over the three types of kerbs. It is noticed that there is no significant difference among three maximum stresses generated in the bead wires based on the three types of kerbs. In addition, their peak stress values are well below the material yield stress (1765 MPa) of the bead wires.

The vertical deformations of the solid tyre for the three types of kerbs are presented in Figure 17. In the

dynamic condition of the tyre, the maximum vertical deformations that can be observed are 80.1 mm for the square type kerb, 79.2 mm for the angle type kerb, and 68.6 mm for the circular type kerb. At 0.25 seconds, the tyre experiences the highest vertical deformation (80.1 mm) when it is moving over the square-type kerb.

Analysis of contact pressure generation of the solid tyre

In this section, the contact pressure of the solid tyre is examined to investigate its effect on tyre wear and tread layer damage during impact with the three different kerbs types, as shown in Figure 18.

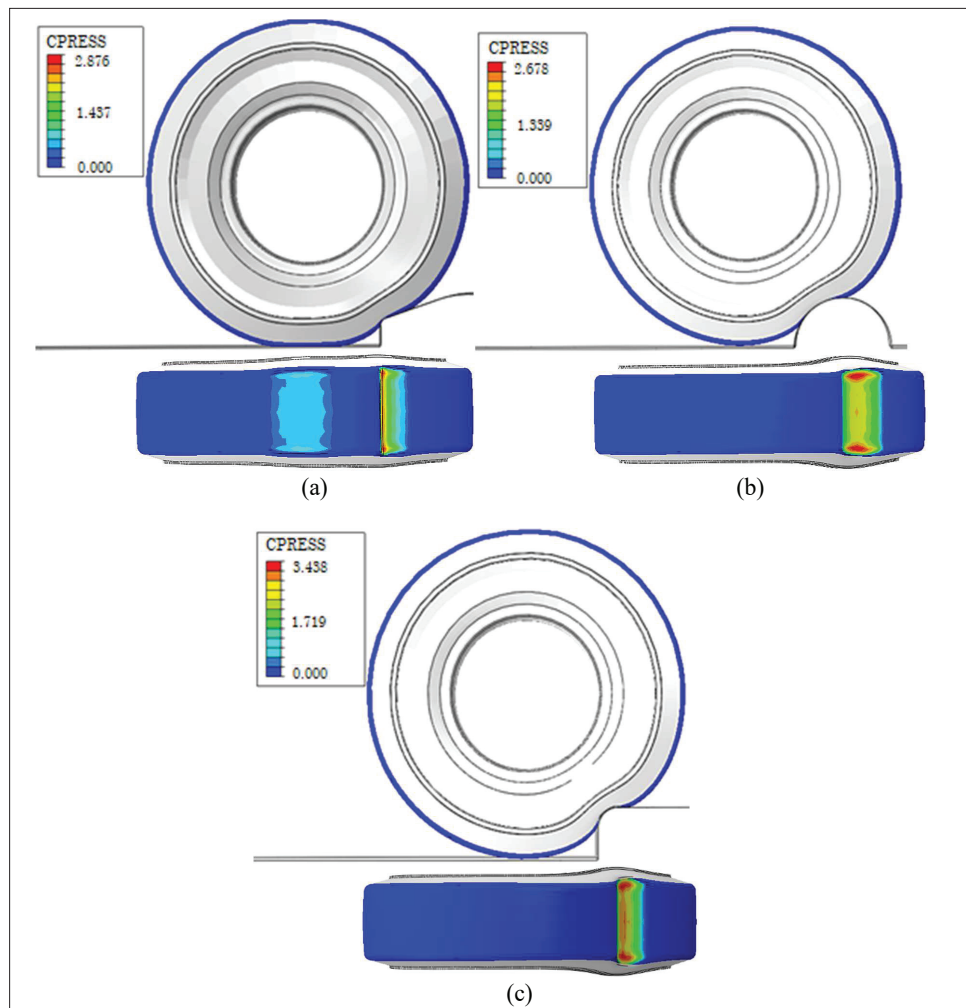


Figure 18: Solid tyre contact pressure (CPRESS in Nmm^{-2}) generation (load 19614 N and velocity 10 kmh^{-1}): (a) impact on angle type kerb; (b) impact on circle type kerb and (c) impact on square type kerb

According to the contact pressure analysis, the highest contact pressure (3.438 Nmm⁻²) is observed when the tyre impacts on the square type kerb. Furthermore, the distribution of the contact pressure throughout the tyre contact patch varies with the type of impact kerb. Based on the contact pressure analysis, the highest wear rate and damage on tread layer can be expected from the tyre when it impacts on the square type kerb compared to the other two types of kerb. In the circle type kerb, damage in the tyre tread layer occurs closer to the edges of the tyre contact patch than the middle.

Analysis of in-plane shear stress generation of the solid tyre

The in-plane shear stress generation in the tyre circumferential direction is presented in Figure 19.

According to Figure 19, high in-plane shear stress (1.297 MPa and 1.314 MPa) is observed when the tyre impacts on the circular type and square type kerbs, respectively. This leads to sudden wear and damage due to chunking in the region of the tread layer. Based on

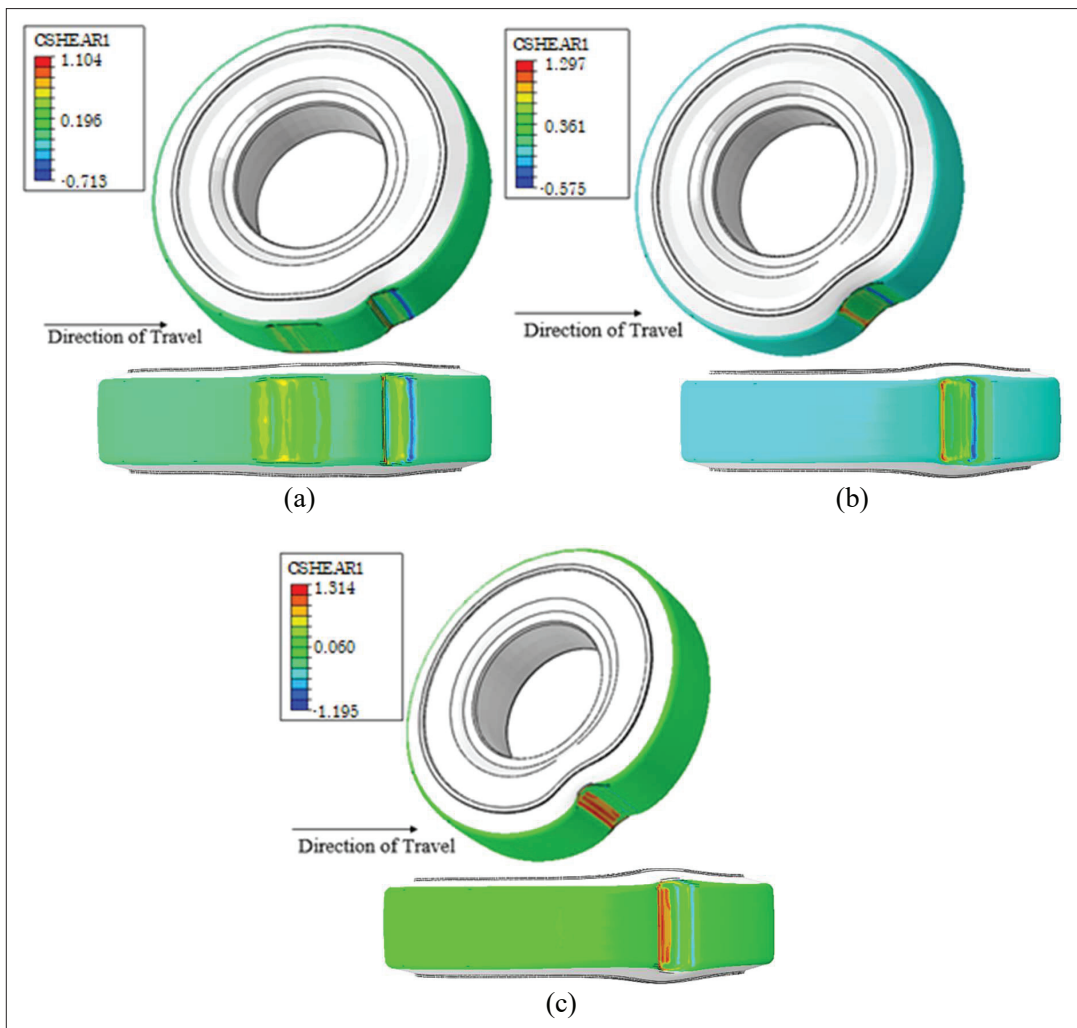


Figure 19: In plane shear stress (CSHEAR1 in MPa) generation at tread layer of the solid tyre (load - 19614 N and velocity - 10 kmh⁻¹): (a) impact on angle type kerb; (b) impact on circle type kerb and (c) impact on square type kerb

above stress analyses and contact pressure analyses, it is clear that the sudden impacts of the solid tyre on square-type kerbs have a relatively more adverse effect on the durability of the tyre.

CONCLUSIONS

This study was conducted to investigate the stress generation and failure regions of solid resilient tyres as they moved on angle, circle and square-type kerbs. The analyses were performed on stress histories, contact pressure and in plane shear stress generation of tyre rubber layers and side walls, under two different moving velocities. To obtain the behaviour of tyre impact stress histories, a validated numerical model of a tyre was developed. The numerical model was used for tyre impact analyses on the three different kerbs by importing static results into the *Abaqus/Explicit* interface. Based on the tyre impact analyses results, relatively higher stresses were observed in the tyre sidewalls (3.089 MPa) when it impacted on square type kerbs. The results show that there is a high tendency for the solid tyres to fail due to side wall crack propagation when the tyre moves on square-type kerbs. Furthermore, high contact pressure (3.438 Nmm⁻²) and high in-plane shear stress (1.314 MPa) were observed at the tread layer of the tyre when it moved on a square-type kerb. These findings highlighted that impact of solid tyres with square-type kerbs can significantly decrease the service life and durability of solid tyres compared to other kerb types. Therefore, this model can be reliably used to identify and analyse improvements to the solid tyre design in order to reduce the probability of failure due to kerb impact. This study can be further extended to observe the rolling resistance and heat generation of the solid resilient tyre as well as the behaviour of tyre impact due to changes of material properties and geometry parameters.

Acknowledgement

The authors would like to thank Elastomeric Engineering Company Ltd., Horana, for providing materials and laboratory facilities for this research study. Furthermore, the authors extend their appreciation to the Finite Element Simulation Centre of the Rubber Research Institute of Sri Lanka, Rathmalana for facilitating the use of licensed applications of the ABAQUS package for this study.

REFERENCES

ABAQUS G. (2014). *Abaqus 6.14 User's Guide*. Dassault Systemes Simulia Corporation, Providence, RI, USA.
Azizi Y. (2020). Measurement methods of tire/road noise. In:

- Automotive Tire Noise and Vibrations* (ed X. Wang), pp. 65–90. Butterworth-Heinemann, Oxford, UK.
DOI: <https://doi.org/10.1016/b978-0-12-818409-7.00005-2>
Chang C.L. & Yang S.H. (2009). Simulation of wheel impact test using finite element method. *Engineering Failure Analysis* **16**(5): 1711–1719.
DOI: <https://doi.org/10.1016/j.engfailanal.2008.12.010>
Dechwayukul C., Kao-ien W., Chetpattananondh K. & Thongruang W. (2010). Measuring service life and evaluating the quality of solid tires. *Sonklanakarin Journal of Science and Technology* **32**(4): 387–390.
Duni E., Toniato G., Saponaro R. & Smeriglio P. (2008). Vehicle fatigue load prediction based on finite element tire/road interaction implemented in an integrated implicit-explicit approach. *Proceedings of the 2008 Abaqus Users' Conference*, 19-22 May. Newport, Rhode Island, USA.
Eindhoven University of Technology. (2021). *Dynamic Mechanical Thermal Analysis – DMAT*. Eindhoven University of Technology. Available at: <https://research.tue.nl/en/equipments/dynamic-mechanical-thermal-analysis>, Accessed 4 April 2021.
Gao Q., Shan Y., Wan X., Feng Q. & Liu X. (2019). 90-degree impact bench test and simulation analysis of automotive steel wheel. *Engineering Failure Analysis* **105**: 143–155.
DOI: <https://doi.org/10.1016/j.engfailanal.2019.06.097>
Han M.J., Lee C.H., Park T.W., Kim Y.S., Shin K.D. & Sim K.S. (2019). Cleat impact analysis of a multi-body quarter vehicle system with finite element models. *International Journal of Mechanical Engineering and Robotics Research* **8**(6): 880–884.
DOI: <https://doi.org/10.18178/ijmerr.8.6.880-884>
Han M.J., Lee C.H. & Park T.W. (2020). Vibro-acoustic response in vehicle interior and exterior using multibody dynamic systems due to cleat impacts. *International Journal of Automotive Technology* **21**(3): 591–602.
DOI: <https://doi.org/10.1007/s12239-020-0056-1>
Hills Kerbs Pty Ltd. (2021). *Concrete Kerbing*. Available at <https://www.hillskerbs.com.au/kerbing/concrete-kerbing>, Accessed 1 May 2021.
Kamoulakos A. & Kao B.G. (1998). Transient dynamics of a tire rolling over small obstacles—a finite element approach with PAM-SHOCK. *Tire Science and Technology* **26**(2): 84–108.
DOI: <https://doi.org/10.2346/1.2135964>
Mooney M. (1940). A theory of large elastic deformation. *Journal of Applied Physics* **11**(9): 582–592.
Ogden R.W. (1972). Large deformation isotropic elasticity on the correlation of theory and experiment for incompressible rubberlike solids. *Proceedings of the Royal Society of London A* **328**: 565–584.
DOI: <https://doi.org/10.1098/rspa.1972.0096>
Ogden R.W. (1973). Large deformation isotropic elasticity—on the correlation of theory and experiment for incompressible rubberlike solids. *Rubber Chemistry and Technology* **46**(2): 398–416.
Neves R.R.V., Micheli G.B. & Alves M. (2010). An experimental and numerical investigation on tyre impact. *International Journal of Impact Engineering* **37**(6): 685–693.

- DOI: <https://doi.org/10.1016/j.ijimpeng.2009.10.001>
- Phromjan J. & Suvanjumrat C. (2018a). A suitable constitutive model for solid tire analysis under quasi-static loads using finite element method. *Engineering Journal* **22**(2): 141–155.
- DOI: <https://doi.org/10.4186/ej.2018.22.2.141>
- Phromjan J. & Suvanjumrat C. (2018b). Vibration effect of two different tires on baggage towing tractors. *Journal of Mechanical Science and Technology* **32**(4): 1539–1548.
- DOI: <https://doi.org/10.1007/s12206-018-0307-5>
- Phromjan J. & Suvanjumrat C. (2020). The contact patch analysis of solid tire on drum testing by finite element method. *IOP Conference Series: Materials Science and Engineering* **886**: 012049.
- DOI: <https://doi.org/10.1088/1757-899x/886/1/012049>
- Reida J.D., Boesch D.A. & Bielenberg R.W. (2007). Detailed tire modeling for crash applications. *International Journal of Crashworthiness* **12**(5): 521–529.
- DOI: <https://doi.org/10.1080/13588260701483813>
- Shiraishi M., Yoshinaga H., Iwasaki N. & Hayashi K. (2000). Making FEM tire model and applying it for durability simulation. *Proceedings of the 6th International LS-Dyna Users Conference 2000*. Detroit, USA.
- Suripa U. & Chaikittiratana A. (2008). Finite element stress and strain analysis of a solid tyre. *Journal of Achievements in Materials and Manufacturing Engineering* **31**(2): 576–579.
- Suvanjumrat C. & Phromjan J. (2020). The contact patch characterization of various solid tire testing methods by finite element analysis and experiment. *International Journal of GEOMATE* **19**(76): 25–32.
- DOI: <https://doi.org/10.21660/2020.76.9134>
- Suvanjumrat C. & Rugsaj R. (2020). The dynamic finite element model of non-pneumatic tire under comfortable riding evaluation. *International Journal of GEOMATE* **19**(76): 62–68.
- DOI: <https://doi.org/10.21660/2020.76.9135>
- Yeoh O.H. (1990). Characterization of elastic properties of carbon-black-filled rubber vulcanizates. *Rubber Chemistry and Technology* **63**(5): 792–805.
- Yeoh O.H. (1997). Hyperelastic material models for finite element analysis of rubber. *Journal of Natural Rubber Research* **12**(3): 142–153.
- Wei C. & Olatunbosun O.A. (2014). Transient dynamic behaviour of finite element tire traversing obstacles with different heights. *Journal of Terramechanics* **56**: 1–16.
- DOI: <https://doi.org/10.1016/j.jterra.2014.07.001>

RESEARCH ARTICLE

Bioremediation

Screening of potential aerobic denitrifying bacteria for nitrate removal from water

A Kirisan^{1*}, N Gnanavelrajah¹ and RR Ratnayake²

¹ Department of Agricultural Chemistry, Faculty of Agriculture, University of Jaffna.

² National Institute of Fundamental Studies, Hantana Road, Kandy.

Submitted: 08 July 2021; Revised: 09 May 2022; Accepted: 24 June 2022

Abstract: Nitrate pollution in groundwater is a common problem in areas where inorganic fertilizer is used to a large extent. This situation seriously affects communities that use ground water as their main source of drinking water and for many other purposes. Therefore, finding an efficient and cost effective system for the removal of nitrate from groundwater is an urgent necessity. The present study was aimed at identifying aerobic bacteria isolated from various soils and water sources and to test their potential for reducing nitrate in groundwater. The bacterial isolates (n = 128) were screened for nitrate reduction by various processes in nutrient broth and in mineral salt medium containing glucose and starch, using KNO₃ as the nitrate substrate. Liberated gases during nitrate reduction were analyzed using gas chromatography. Out of 128 morphologically different isolates, two strains, namely *Paracoccus* sp. (A2) and *Bacillus* sp. (A19), were selected for further analysis on the basis of their performance for water treatment. The nitrate reduction percentages of A2 and A19 were within the range of 59.63-100% and 86.67-100%, respectively. Gas chromatography results indicated that these two strains liberated a higher percentage of N₂ (68 - 90%) compared to N₂O (5-13%) and CO₂ (traces), while reducing the amount of nitrate. These results confirmed that A2 and A19 have the potential to be used in bioremediation of nitrate contaminated groundwater.

Keywords: Bioremediation, contamination, groundwater, isolates, nitrate, treatments

INTRODUCTION

Nitrate is a harmful pollutant that has become a common water contaminant in many parts of the world

(Rajta *et al.*, 2019; Zhang *et al.*, 2019b; Abascal *et al.*, 2021). Excessive use of nitrogen-rich fertilizers used for agricultural purposes, discharge of poorly treated domestic and industrial wastewater, livestock manure, and leachate from landfill sites are the main anthropogenic sources for nitrate pollution of groundwater (Gutierrez *et al.*, 2018; Tokazhanov *et al.*, 2020). The World Health Organization (WHO) reconfirmed the safe level of nitrate as below 50 mg/L for drinking water, which was set to protect against methaemoglobinaemia (WHO, 2017). Excessive consumption of nitrates can cause health effects in humans and animals alike, especially methaemoglobinaemia (blue baby syndrome) in infants and gastrointestinal cancer in adults (Ren *et al.*, 2018; Cotruvo, 2017). There are reports of other health disorders, including increased infant mortality, hypertension, central nervous system birth defects, diabetes, spontaneous abortions, respiratory tract infections, and changes to the immune system due to the consumption of high levels of nitrates (Kotopoulou *et al.*, 2022).

Nitrate contamination of groundwater is a burning issue in many areas in the world such as UK (Neal *et al.*, 2006), Australia (Rasiah *et al.*, 2013), North America (Power & Schepers, 1989), Morocco, Changshu in China (Sadeq *et al.*, 2008), and Toyserkan in western Iran (Jalali, 2011). The problem is more severe in some regions of South East Asia. Karunanidhi *et al.* (2021) reported that synthetic fertilizers, cow dung and sheep manure, industrial discharge, septic tank leakage, and

* Corresponding author (abhiramyt@univ.jfn.ac.lk;  <https://orcid.org/0000-0002-7407-2571>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

municipal solid waste disposal are the major sources of nitrate pollution in India. They reported that around 117.93 million people drink water contaminated with nitrate levels between 45-100 mg/L and 108.2 million people consume water with levels more than 100 mg/L in India. It is also reported that nitrate contamination in groundwater is a major concern in the coastal region of Bangladesh (Jannat *et al.*, 2022).

Similar to many parts of the seasonally dry tropical areas of the world, the Jaffna peninsula in Sri Lanka also experiences minimal periods of rain, while ground water is the only source of water for drinking (Prabagar *et al.*, 2020). The water quality in the area has been drastically reduced due to intensive inorganic fertilizer use, resettlement and urbanization within the last decade (Piyathilake *et al.*, 2022).

Due to the above reasons protection of groundwater quality is an important concern confronting much of the world's population. Though sophisticated technologies such as chemical denitrification (Xu *et al.*, 2017), ion exchange (Vandekerckhove *et al.*, 2018), reverse osmosis (Epsztein *et al.*, 2015), electro-dialysis and catalytic denitrification (Zhang *et al.*, 2016) can be used to remove nitrate from groundwater, proper, cost effective and environmentally friendly systems need to be adopted for remediation of nitrate from ground water in the developing countries. Therefore, finding appropriate treatment technologies for nitrate removal is critical. Biological denitrification is the most promising approach currently investigated for treatment of nitrate contaminated water.

Microbial denitrification has been proven to be an advanced, high performance, and highly selective method for complete nitrate elimination (Gomez *et al.*, 2000b). Biological denitrification is the most important and widely used method to treat nitrate wastes as it enables the conversion of nitrogen compounds into harmless dinitrogen (N₂) gas (Costa *et al.*, 2018). A microbial consortium composed of *Cellulosimicrobium* sp., *Aeromonas veronii*, *Lysinibacillus sphaericus*, and *Rhodococcus rhodochrous* was found to be the most efficient bacterial consortium for reducing nitrate in rubber latex wastewater (Dey *et al.*, 2019). Biological denitrification utilizes the anaerobic reduction of oxidized nitrogen compounds through the sequential activity of microbial reductase enzymes, finally converting them to harmless nitrogen gas. Four enzymes, namely, *nitrate reductase*, *nitrite reductase*, *nitric oxide reductase*, and *nitrous oxide reductase*, are responsible for the complete reduction of nitrate ion to dinitrogen gas (Pang & Wang,

2021). A variety of incomplete denitrification pathways exists. Some denitrifying bacteria reduce both nitrates and nitrites, while others reduce only nitrite. Some produce only dinitrogen, some produce a mixture of dinitrogen and nitrous oxide, while others produce only nitrous oxide.

Even within a single species such as *Pseudomonas fluorescens*, the biotypes differ in the end product of the pathway. Although nitrate reduction activity is exhibited by diverse microbial genera, with a range of heterotrophic and autotrophic metabolisms, the aerobic nitrate reducers belong to a variety of groups of heterotrophs (Guo *et al.*, 2013). Aerobic denitrification gained attention due to its easier operation and higher nitrate reduction efficiency compared to anaerobic denitrification (Chen *et al.*, 2012). The most predominant denitrifying bacteria that are reported in our environment belong to the genus *Pseudomonas*. There are reports on aerobic denitrifying species isolated from environmental samples such as ponds, canals, soils, and activated sludge (Patureau *et al.*, 2000; Huang *et al.*, 2020). *Pseudomonas aeruginosa* has been extensively studied genetically (Wu *et al.*, 2013) and therefore is usually considered a favorable organism to be used in studies on denitrification in wastewater treatment plants. However, as it is an opportunistic pathogen, it could not be utilized for drinking water treatment processes.

The present study investigates the possibility of utilizing microorganisms isolated from the environment to recover NO₃⁻ contaminated ground water from Jaffna, Sri Lanka. We hypothesized that microorganisms isolated from different environments would be efficient and capable for this purpose. The expected result would be applicable for the remediation of groundwater resources in Sri Lanka and other countries in the same region of South Asia.

MATERIALS AND METHODS

Sources for the isolation of bacteria

Soil and water from submerged paddy fields and ponds, and wet soil enriched with partially decomposed manure from the Jaffna peninsula (9.6615°N, 80.0255°E) were used for the isolation of bacteria.

Isolation and screening of aerobic denitrifiers

Nitrate rich modified nutrient broth consisting of (g/L): Peptone 5.0, NaCl 5.0, KNO₃ 1.0, glucose 1.0, yeast extract 5.0, and beef extract 5.0 with pH 7.2 was used for the

enrichment of nitrate reducers. The enrichment cultures were plated on modified bromothymol blue (BTB) medium containing 0.1% L-asparagine, 0.1% KNO_3 , 0.1% KH_2PO_4 , 0.005% $\text{FeCl}_2 \cdot 6\text{H}_2\text{O}$, 0.02% $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.1% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 mL of BTB (1% in ethanol), 2% agar, and 0.5% glucose at pH 7.0, to isolate and screen the denitrifier under aerobic conditions (Takaya *et al.*, 2003). Culture plates were incubated at 30 °C for 3 ds. Well defined bacterial strains on the basis of their colony and different morphological characteristics were selected for further screening. A nutrient broth with KNO_3 was used to determine denitrification activity at the initial screening stage. Each strain was inoculated into a 15 mL screw cap tube containing the sterile nutrient broth and Durham's tube, and was incubated at 30 °C for 48 hours (Guo *et al.*, 2013).

Nitrate reduction in synthetic medium

Nitrate removal activity of five selected strains based on nitrate reduction in the nitrate broth medium was evaluated in a synthetic mineral salt medium (MSM) consisting of potassium dihydrogen phosphate (0.1 g/L), dipotassium hydrogen phosphate (1 g/L), ammonium chloride (0.5 g/L), calcium chloride (0.005 g/L), magnesium sulphate (0.1 g/L), and sodium silicate (0.05 g/L), the pH being adjusted to 7.2, with either glucose or starch as the carbon source in three different percentages, *viz*; 0.25 %, 0.5 % and 1.0 % (Ayyasamy *et al.*, 2007).

Analytical methods

Nitrate content of the bacteria inoculated sample was determined according to Anderson & Ingram (1993) via reaction with salicylic acid and sodium hydroxide followed by spectrophotometry at 410 nm. The amount of nitrite was measured through the reaction with sulphanilic acid and N,N-dimethyl-1-naphthylamine, followed by spectrophotometry at 520 nm, as described by Blaszczyk (1993). Ammonium ion content was determined by the method described by Guo *et al.* (2013).

Gas chromatographic analysis

Gases evolved during denitrification were analyzed for N_2 , N_2O , and CO_2 by a gas chromatograph equipped with Shimadzu GC 9 AM analyzer and a thermal conductivity detector with helium as the carrier (Green *et al.*, 2010). Column temperature, detector temperature and injector temperature were maintained at 50-200 °C, 200 °C and 175 °C respectively and the gas flow rate was kept at 30 mL/min. The mixture of gases evolved during nitrate reduction was collected and identified by comparing the retention time of the peaks with standards.

Identification of isolated bacteria

Standard physiological and biochemical characteristics, such as colony morphology, cell shape, gram reaction, catalase reaction, oxidase reaction, motility, nitrate reduction test, anaerobic growth, glucose acid test and starch hydrolysis were used for the identification of bacteria according to the methods described by Bergey (1994).

Evaluation of selected bacterial strains for removal of nitrate from contaminated well water with starch as the carbon source

Based on efficient nitrate removal in the carbon sources and negative growth on Maconkey agar medium, two bacterial strains (A2 and A19) were selected for further study with five nitrate contaminated water samples. Based on optimization of two different carbon sources and different percentages of the synthetic medium, 0.5% starch was selected for water treatment. Conical flasks with 100 mL of 0.5% (0.25 g) starch were sterilized by autoclaving at 120 °C for 20 min. A water sample of 50 mL was filter sterilized using 0.45 μm syringe filter and was added into each flask containing 0.5% (0.25 g) sterile starch and mixed well. Bacterial isolates were cultured on nutrient agar for 24 – 48 h and cell suspensions were prepared by suspending the cultures in 5 mL of sterile distilled water and the turbidity was adjusted to OD 0.5. Each culture suspension (0.5 mL) was added aseptically into each flask and incubated at 30 °C and kept at 120 rpm in a shaking incubator for 72 h. The same conditions were provided without the cultures as the control. Nitrate and nitrite concentrations were analyzed every 12 h in all samples.

Data analysis

All the experiments were carried out in triplicate and the results were analyzed by analysis of variance (ANOVA) using SAS statistical software version 9.1. Treatment means were compared using Duncan's multiple-range test at a significance level of 0.05.

RESULTS AND DISCUSSION

Screening of aerobic denitrifying bacteria

Out of 128 morphologically different strains isolated from different sources, 70 strains were capable of forming blue colonies on the BTB agar plates (Table 1) due to an increase of pH on the medium (Wu *et al.*, 2013).

Table 1: Number of isolates and nitrate reducers from different sources

Sources of isolation	Number of strains isolated	Number of nitrate reducers
Municipal compost (GMC)	12	11
Pond soil	8	1
Fish waste (GFW)	15	1
Municipal solid waste dumping place (MSW)	17	4
Manure (swine, poultry)	17	6
Paddy soil (KPS)	23	17
Paddy water (KPW)	18	14
Unutilized well water (PSW)	8	8
Compost (COM)	10	8
Total	128	70

Among the 70 strains, 38 showed gas bubble formation in Durham's tubes. After rescreening, the nitrate removal efficiency was examined in nutrient broth with nitrate under aerobic conditions. Out of 38 strains, 5 bacterial isolates, capable of reducing either nitrate or nitrite efficiently (more than 50%) were selected by quantitative screening. Since most of the denitrifiers were heterotrophs, they required carbon sources for energy consumption (Pang & Wang 2021). Although five strains had higher nitrate reduction capacity (more than 70 %) either with glucose or starch, three of them were grown on Maconky agar medium, hence, cannot be used for treatment of water samples. The other two, A2 and A19 were tested with five nitrate contaminated water samples for their nitrate removal efficiency.

Table 2: Physical and biochemical identification of bacterial strains

Identification test	A2	A19
	(<i>Paracoccus</i> sp.)	(<i>Bacillus</i> sp.)
Shape	Round	Rod
Gram staining	Negative	Positive
Endospore formation	Negative	Positive
Motility test	Negative	Positive
Catalase test	Positive	Positive
Oxidase test	Positive	Positive
Anaerobic growth	Positive	Positive
Nitrate reduction test	Positive	Positive
Growth on Maconky agar	Negative	Negative
Glucose acid test	Positive	Positive
Starch hydrolysis test	Positive	Positive

Identification of the selected bacterial strains

According to the physical and biochemical identification, strains A2 and A19 were identified as *Paracoccus* sp. and *Bacillus* sp. (Table 2).

Optimization of the percentage of the carbon sources for aerobic nitrate removal

The two carbon sources used, namely glucose and starch at the levels of 0.25%, 0.5%, and 1%, were tested for carbon source optimization for the strains A2 and A19. The optimum carbon percentage was decided based on efficient nitrate removal as well as lowest intermediate accumulation (nitrite) in the medium.

Figure 1 shows the effect of glucose (a) and starch (b) on nitrate removal and nitrite accumulation by the strain A2. The nitrate nitrogen concentration was lowered with both carbon sources at all three levels tested. Although nitrate reduction was observed with time, after 24 hours, nitrite began to accumulate with glucose at concentrations of 0.25%, 0.5%, and 1% up to 60 hours. However, when starch was used as the carbon source, nitrite accumulation was not observed at all three levels.

Strain A2, grown with 0.5% and 1% starch reduced nitrate to 11 mg/L after 36 hours of incubation (Figure 1). Although nitrate reduction was observed with glucose as the carbon source, nitrite accumulation also was recorded at all three levels. When the two carbon sources were compared at three different levels, 0.5% starch exhibited higher nitrate reduction and nitrite was not detected. Ammonium ion was not detected at any levels of the two carbon sources. Therefore, it could be concluded that the strain A19 has a higher potential to remove nitrate with 0.5% of starch, without accumulation of the nitrite intermediate. Further, the nitrate reduction efficiency and intermediate accumulation of strain A2 could be controlled by the organic carbon sources, nitrate concentration, and C/N ratio. A study conducted by Blaszczyk (1993) clearly stated that the denitrification performance of *Paracoccus denitrificans* strongly depended on the quality of the medium.

The carbon source can provide the energy for the aerobic denitrifier and electron donors for their growth and metabolism (Li et al., 2020b). Nitrogen transformation of strain A19 with the carbon sources glucose (a) and starch (b) is shown in Figure 2. The results indicate that, after 60 hours of incubation, complete reduction of nitrate was observed with 0.5% and 1% of starch, without the accumulation of nitrite. However, with glucose,

nitrate was reduced to below the permissible level with high amount of nitrite accumulation. Intermediate accumulation mainly depends on the carbon sources. In a study, N_2O and NO accumulation in the presence of nitrite during denitrification was observed with acetate-fed denitrifying cultures, but not in methanol- or ethanol-fed denitrifying reactors with excessive carbon source supply (Lu *et al.*, 2014). At the end of 60 hours of incubation, almost all the nitrate was found to be in the form of nitrite with glucose as the carbon source. Therefore, it could be concluded that nitrite concentration also varied

with the types and levels of carbon sources for each strain. Among the two carbon sources, starch possessed significantly higher nitrate reduction efficiency for both strains. It may be due to the presence of amylolytic enzymes in the two strains (A2 and A19) and the ability to utilize starch as the carbon source. Ayyasamy *et al.* (2007) also stated that the amylolytic enzymes present in organisms can utilize the starch well. In another study conducted by Rajakumar *et al.* (2008), it is reported that the denitrification rate was higher for starch than glucose, acetic acid, cellulose, and sucrose.

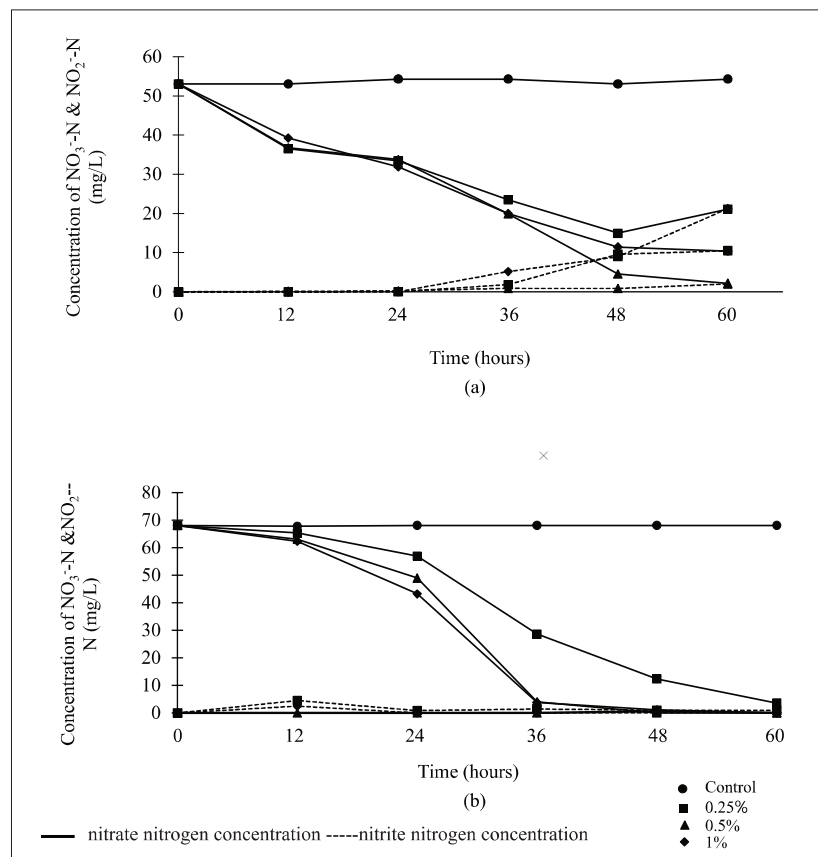


Figure 1: Effect of glucose (a) and starch (b) on nitrate and nitrite concentration in MSM inoculated with strain A2 up to 60 hours of incubation

Therefore, the results indicate that the process of complete denitrification is not a stable process and depends on several factors such as bacterial strains and types and amounts of carbon sources. This unstable nature of the process is due to the influence of these factors, which have been reported in many previous studies (Li *et al.*, 2020; Yin & Yan, 2020). Furthermore, nitrate removal is a highly dynamic process that can

be affected by temperature, pH, C/N ratio, dissolved oxygen concentration, and bacterial population (Chen *et al.*, 2006; Olaya-Abri *et al.*, 2021; Zhou *et al.*, 2021). In another bacterial denitrification study conducted by Gomez *et al.* (2000), the nitrite accumulation was higher when sucrose was used as the carbon source, but it was not found when methanol and ethanol were used.

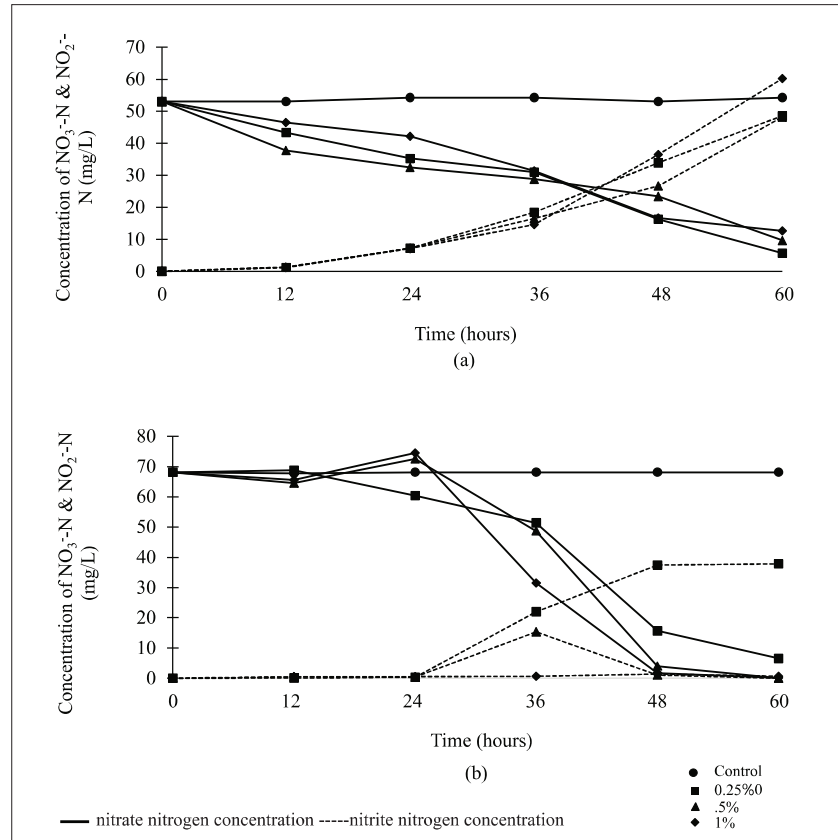


Figure 2: Effect of glucose (a) and starch (b) on nitrate and nitrite concentrations in MSM inoculated with strain A19

Aerobic nitrate removal in well water contaminated with nitrate

Five water samples having initial nitrate nitrogen concentrations of 57.37 mg/L (S1), 32.72 mg/L (S2), 26.70 mg/L (S3), 20.81 mg/L (S4) and 15.91 mg/L (S5) were tested with either A2 or A19. Changes in the nitrate nitrogen concentration of different water samples treated with A2 and A19 with 0.5% of starch are shown in Figures 3.

Figure (3a) shows the nitrate reduction profile of water sample S1, having an initial nitrate nitrogen concentration of 57 mg/L, treated either with A2 or A19. With the sample treated with A19, nitrate was found to be at a safe level after 60 hours and not detected at 72 hours. During nitrate reduction by A19, nitrite formation increased at a high rate up to 48 hours and declined thereafter. Although, nitrate was not detected at 72 hours, nitrite was not at a safe level (3.89 mg/L). Accumulation of nitrate at early stage by aerobic denitrifiers such as

Paracoccus denitrificans, which had 14.07 mg/L initially and reduced to zero at 40 hours (Zhang *et al.*, 2020) has been reported. The accumulation of NO₂⁻-N in S1 treated with strain A19 is possibly due to high initial nitrate content of the water and the consecutive lag of nitrite reduction (Chen *et al.*, 2020), which could be possibly removed with more time. Significant difference was observed in final nitrate concentration in the sample S1 treated with either A2 or A19. Further, A2 reduced nitrate nitrogen to the level of 9.1 mg/L at 72 hours without accumulation of nitrite. *Paracoccus* sp. (A2) expressed greater aerobic nitrate removal capacity similar to numerous aerobic bacteria. For instance, *Paracoccus denitrificans* strain removed 90.00% of NO₃⁻-N in a 250 mg/L initial NO₃⁻-N medium (Medhi *et al.*, 2018). Similar aerobic nitrate removal (87.63%) was achieved by *Paracoccus denitrificans* strain Z195 (Zhang *et al.*, 2020). *Paracoccus thiophilus* strain LSL 251 had an aerobic denitrification rate of 5.90 mg/L/h in nitrate rich medium (Chen *et al.*, 2020).

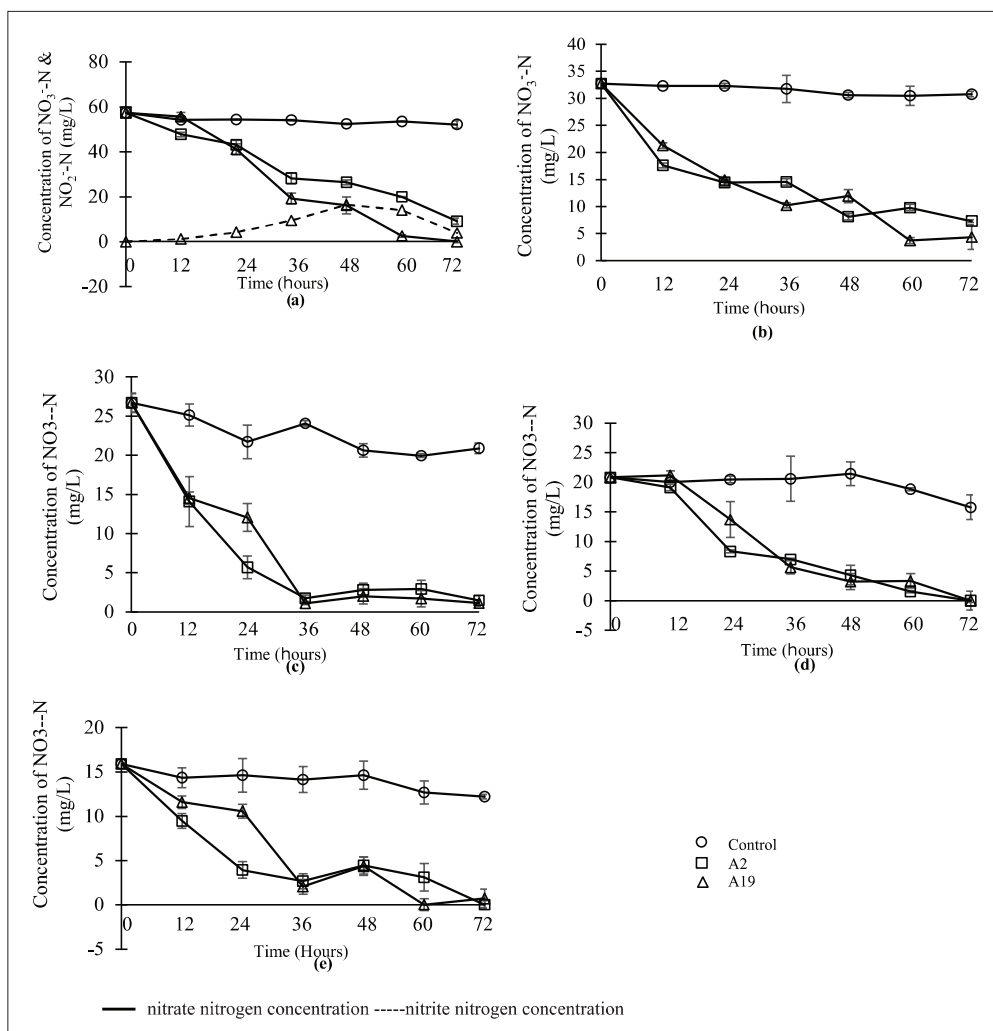


Figure 3: Bacterial reduction of nitrate nitrogen in well water in the presence of 0.5 % starch a) S1, b) S2, c) S3, d) S4 and e) S5

Both strains reduced nitrate with time without nitrite accumulation in sample S2 (Figure 3b). Reduction of the nitrate nitrogen concentration from 32.71 mg/L to 7.30 mg/L was observed in 72 hours with A2, while 4.36 mg/L was observed with A19 at the same time. In another water sample having an initial nitrate nitrogen concentration of 26.7 mg/L, the nitrate level was reduced below the safe level (11 mg/L) at 24 hours by the strain A2, while it reached 1.09 mg/L with A19 at 36 hours, without nitrite being detected (Figure 3c). Moreover, when the initial nitrate concentration was about 20.81 mg/L both strains A2 and A19 lowered the nitrate concentration below the safe level after 24 and 36 hours, respectively (Figure 3d). Changes in NO₃⁻ nitrogen concentrations of water sample S4 and S5 treated with the strains are shown in Figures

3d and 3e respectively. Initial concentrations in water samples S4 and S5 were 20.81 mg/L and 15.91 mg/L, respectively.

During incubation, fluctuation was observed in the nitrate nitrogen concentration with the treatment of either A2 or A19, however, no nitrite was detected in either sample with strain A2. This might be due to the oxidation of the negligible amount of nitrite, resulting from the brief exposure to air during sampling (Zhang *et al.*, 2011) or the simultaneous nitrification and denitrification capability of the strains (Kim *et al.*, 2008; Khardenavis *et al.*, 2007). The denitrification process can also be influenced by metal ions such as Fe³⁺ and Mo⁶⁺ (Pintathong *et al.*, 2009). As can be observed from the

figures, both strains of A2 and A19 with 0.5% of starch lowered the nitrate nitrogen concentration from various initial levels of nitrogen in water. Nevertheless, the time needed to attain the permissible level or below varied with strains and the well water samples. Although the difference is not significant, it might be due to the effect of initial nitrate concentration, pH, and availability of other nutrients in the water sample (Körner & Zumft, 1989). Denitrification of synthetic waste water having a high nitrate level was inhibited at the pH values of 6.5 and 7.0. Although, higher nitrate reduction was achieved with the increased pH values of 7.5, 8.0 and 9.0, accumulation of nitrite increased significantly (Glass & Silverstein, 1998). A *Bacillus pumilus* strain removed 99.7% of NO_3^- – nitrogen in a 70 mg/L initial NO_3^- containing medium (Elkarrach *et al.*, 2021). In another study, 89.4% of nitrate removal was reported by *Bacillus* sp. after a 48-h cultivation in a sole N-source medium with initial nitrogen approximately 20 mg/L (Huang *et al.*, 2017). *Bacillus* sp. PB8 showed excellent aerobic denitrifying ability (0.25 mg/L/h) both in artificial media and real wastewater treatment (Barman *et al.*, 2018).

Gas chromatographic analysis

During denitrification by A2, a higher percentage (90.6%) of N_2 , lower percentage of N_2O (5.7%) and traces of CO_2 were released, while A19 released 68.5% of N_2 and 12.5% of N_2O . Considering the composition of gases released, A2 would be a better strain than A19.

CONCLUSION

Among the 70 nitrate reducing bacteria strains isolated, *Paracoccus* sp. A2 and *Bacillus* sp. A19 exhibited a high nitrate reduction potential either with glucose or starch as the carbon source. However, with glucose, in all three levels (0.25%, 0.5% and 1%), nitrite accumulation was observed. Among the three percentages of starch (0.25%, 0.5% and 1%), 0.5% was the optimum level for efficient nitrate removal. Further, *Paracoccus* sp. A2 and *Bacillus* sp. A19 were capable of removing the nitrate nitrogen content in the range of 15 mg/L to 57.37 mg/L in groundwater to safe levels within 72 hours with 0.5% of starch. To the best of our knowledge, this is the first report on nitrate reduction by bacterial strains in Sri Lanka and this finding may provide useful information for the potential use of these two bacterial species. Further studies are required to validate nitrate removal efficiency of the strains under different physico-chemical conditions.

Acknowledgement

The authors would like to thank the National Science Foundation, Sri Lanka, for financial assistance through the grant number of NSF/SCH/2016/03 and the National Institute of Fundamental Studies, Sri Lanka for laboratory facilities.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- Abascal E., Gomez-Coma L., Ortiz I. & Ortiz A. (2021). Global diagnosis of nitrate pollution in groundwater and review of removal technologies. *Science of the Total Environment* **816**:152233.
DOI: <https://doi.org/10.1016/j.scitotenv.2021.152233>
- Anderson J.M. & Ingram J. (1993). *Tropical soil biology and fertility: A handbook of Method*. CAB International, Wallingford, England.
DOI: <https://doi.org/10.1097/00010694-199404000-00012>
- Ayyasamy P.M., Shanthi K., Lakshmanaperumalsamy P., Lee S.J., Choi N.C. & Kim D.J. (2007). Two-stage removal of nitrate from groundwater using biological and chemical treatments. *Journal of bioscience and bioengineering* **104**(2): 129–134.
DOI: <https://doi.org/10.1263/jbb.104.129>
- Barman P., Bandyopadhyay P., Kati A., Paul T., Mandal A.K., Mondal K.C. & Mohapatra P.K.D. (2018). Characterization and strain improvement of aerobic denitrifying EPS producing bacterium *Bacillus cereus* PB88 for shrimp water quality management. *Waste and Biomass Valorization* **9**(8): 1319–1330.
DOI: <https://doi.org/10.1007/s12649-017-9912-2>
- Bergey D.H. (1994). *Bergey's Manual of Determinative Bacteriology*. Williams & Wilkins, Baltimore, USA.
- Blaszczyk M. (1993). Effect of medium composition on the denitrification of nitrate by *Paracoccus denitrificans*. *Applied and Environmental Microbiology* **59**(11): 3951–3953.
DOI: <https://doi.org/10.1128/aem.59.11.3951-3953.1993>
- Chen F., Xia Q. & Ju L.K. (2006). Competition between oxygen and nitrate respirations in continuous culture of *Pseudomonas aeruginosa* performing aerobic denitrification. *Biotechnology and Bioengineering* **93**(6): 1069–1078.
DOI: <https://doi.org/10.1002/bit.20812>
- Chen P., Li J., Li Q.X., Wang Y., Li S., Ren T. & Wang L. (2012). Simultaneous heterotrophic nitrification and aerobic denitrification by bacterium *Rhodococcus* sp. CPZ24. *Bioresource Technology* **116**: 266–270.
DOI: <https://doi.org/10.1016/j.biortech.2012.02.050>

- Chen S., Li S., Huang T., Yang S., Liu K., Ma B. & Miao Y. (2020). Nitrate reduction by *Paracoccus thiophilus* strain LSL 251 under aerobic condition: Performance and intracellular central carbon flux pathways. *Bioresource Technology* **308**: 123301.
DOI: <https://doi.org/10.1016/j.biortech.2020.123301>
- Costa D.D., Gomes A.A., Fernandes M., da Costa Bortoluzzi R.L., Magalhaes M.D. L.B. & Skoronski E. (2018). Using natural biomass microorganisms for drinking water denitrification. *Journal of Environmental Management* **217**: 520–530.
DOI: <https://doi.org/10.1016/j.jenvman.2018.03.120>
- Cotruvo J. A. (2017). WHO guidelines for drinking water quality: first addendum to the fourth edition. *Journal of American Water Works Association* **109**(7): 44–51.
DOI: <https://doi.org/10.5942/jawwa.2017.109.0087>
- Dey S., Kumar H., Sinha S.K., Goud V.V. & Das S. (2020). Bacterial biofilm-based nitrate and phosphate removal from rubber latex wastewater for sustainable water usage. *Water and Environment Journal* **34**: 170–182.
DOI: <https://doi.org/10.1111/wej.12515>
- Elkarrach K., Merzouki M., Atia F., Laidi O. & Benlemlih M. (2021). Aerobic denitrification using *Bacillus pumilus*, *Arthrobacter* sp., and *Streptomyces lusitanus*: Novel aerobic denitrifying bacteria. *Bioresource Technology Reports* **14**: 100663.
DOI: <https://doi.org/10.1016/j.biteb.2021.100663>
- Epsztein R., Nir O., Lahav O. & Green M. (2015). Selective nitrate removal from groundwater using a hybrid nanofiltration–reverse osmosis filtration scheme. *Chemical Engineering Journal* **279**: 372–378.
DOI: <https://doi.org/10.1016/j.cej.2015.05.010>
- Glass C. & Silverstein J. (1998). Denitrification kinetics of high nitrate concentration water: pH effect on inhibition and nitrite accumulation. *Water Research* **32**(3): 831–839.
DOI: [https://doi.org/10.1016/S0043-1354\(97\)00260-1](https://doi.org/10.1016/S0043-1354(97)00260-1)
- Gomez M.A., Gonzalez-Lopez J. & Hontoria-García E. (2000 b). Influence of carbon source on nitrate removal of contaminated groundwater in a denitrifying submerged filter. *Journal of Hazardous Materials* **80**(1-3): 69–80.
DOI: [https://doi.org/10.1016/S0304-3894\(00\)00282-X](https://doi.org/10.1016/S0304-3894(00)00282-X)
- Green S.J., Prakash O., Gihring T.M., Akob D.M., Jasrotia P., Jardine P.M. & Kostka J.E. (2010). Denitrifying bacteria isolated from terrestrial subsurface sediments exposed to mixed-waste contamination. *Appl. Environ. Microbiol* **76**(10): 3244–3254. DOI: <https://doi.org/10.1128/AEM.03069-09>
- Guo L., Chen Q., Fan F., Hu Z., Wu J., Miao A., Xiao L., Chen X. & Yang L. (2013). Application potential of a newly isolated indigenous aerobic denitrifier for nitrate and ammonium removal of eutrophic lake water. *Bioresource Technology* **142**: 45–51.
DOI: <https://doi.org/10.1016/j.biortech.2013.05.021>
- Gutierrez M., Biagioni R.N., Alarcon-Herrera M.T. & Rivas-Lucero B.A. (2018). An overview of nitrate sources and operating processes in arid and semiarid aquifer systems. *Science of the Total Environment* **624**: 1513–1522.
DOI: <https://doi.org/10.1016/j.scitotenv.2017.12.252>
- Huang F., Pan L., Lv N. & Tang X. (2017). Characterization of novel *Bacillus* strain N31 from mariculture water capable of halophilic heterotrophic nitrification–aerobic denitrification. *Journal of Bioscience and Bioengineering* **124**(5): 564–571.
DOI: <https://doi.org/10.1016/j.jbiosc.2017.06.008>
- Huang X., Weisener C.G., Ni J., He B., Xie D. & Li Z. (2020). Nitrate assimilation, dissimilatory nitrate reduction to ammonium, and denitrification coexist in *Pseudomonas putida* Y-9 under aerobic conditions. *Bioresource Technology* **312**: 123597.
DOI: <https://doi.org/10.1016/j.biortech.2020.123597>
- Jannat et al. (13 authors) (2022). Hydro-chemical assessment of fluoride and nitrate in groundwater from east and west coasts of Bangladesh and India. *Journal of Cleaner Production* **372**: 133675.
DOI: <https://doi.org/10.1016/j.jclepro.2022.133675>
- Jalali M. (2011). Nitrate pollution of groundwater in Toyserkan, western Iran. *Environmental Earth Sciences* **62**: 907–913.
DOI: <https://doi.org/10.1007/s12665-010-0576-5>
- Karunanidhi D., Aravinthasamy P., Subramani T. & Kumar M. (2021). Human health risks associated with multipath exposure of groundwater nitrate and environmental friendly actions for quality improvement and sustainable management: a case study from Texvalley (Tiruppur region) of India. *Chemosphere* **265**: 129083.
DOI: <https://doi.org/10.1016/j.chemosphere.2020.129083>
- Khardenavis A.A., Kapley A. & Purohit H.J. (2007). Simultaneous nitrification and denitrification by diverse *Diaphorobacter* sp. *Applied Microbiology and Biotechnology* **77**(2): 403–409.
DOI: <https://doi.org/10.1007/s00253-007-1176-5>
- Kotopoulou S., Zampelas A. & Magriplis E. (2022). Dietary nitrate and nitrite and human health: a narrative review by intake source. *Nutrition Reviews* **80**(4): 762–773.
DOI: <https://doi.org/10.1093/nutrit/nuab113>
- Kim M., Jeong S.Y., Yoon S.J., Cho S.J., Kim Y.H., Kim M.J. & Lee S.J. (2008). Aerobic denitrification of *Pseudomonas putida* AD-21 at different C/N ratios. *Journal of Bioscience and Bioengineering* **106**(5): 498–502.
DOI: <https://doi.org/10.1263/jbb.106.498>
- Körner H. & Zumft W.G. (1989). Expression of denitrification enzymes in response to the dissolved oxygen level and respiratory substrate in continuous culture of *Pseudomonas stutzeri*. *Applied and Environmental Microbiology* **55**(7): 1670–1676.
DOI: <https://doi.org/10.1128/aem.55.7.1670-1676.1989>
- Li S., Zhang H., Huang T., Ma B., Miao Y., Shi Y. & Huang X. (2020 b). Aerobic denitrifying bacterial communities drive nitrate removal: Performance, metabolic activity, dynamics and interactions of core species. *Bioresource Technology* **316**: 123922.
DOI: <https://doi.org/10.1016/j.biortech.2020.123922>
- Lu H., Chandran K. & Stensel D. (2014). Microbial ecology of denitrification in biological wastewater treatment. *Water Research* **64**: 237–254.
DOI: <https://doi.org/10.1016/j.watres.2014.06.042>
- Medhi K., Mishra A. & Thakur I. S. (2018). Genome sequence

- of a heterotrophic nitrifier and aerobic denitrifier, *Paracoccus denitrificans* strain ISTOD1, isolated from wastewater. *Genome Announcements* **6**(15): e00210-18.
DOI: <https://doi.org/10.1128/genomeA.00210-18>
- Neal C., Jarvie H. P., Neal M., Hill L. & Wickham H. (2006). Nitrate concentrations in river waters of the upper Thames and its tributaries. *Science of the Total Environment* **365**: 15–32.
DOI: <https://doi.org/10.1016/j.scitotenv.2006.02.031>
- Olaya-Abril A., Hidalgo-Carrillo J., Luque-Almagro V.M., Fuentes-Almagro C., Urbano F.J., Moreno-Vivian C. & Roldan M.D. (2021). Effect of pH on the denitrification proteome of the soil bacterium *Paracoccus denitrificans* PD1222. *Scientific Reports* **11**(1): 1–12.
DOI: <https://doi.org/10.1038/s41598-021-96559-2>
- Pang Y. & Wang J. (2021). Various electron donors for biological nitrate removal: A review. *Science of the Total Environment* **794**: 148699.
DOI: <https://doi.org/10.1016/j.scitotenv.2021.148699>
- Patureau D., Zumstein E., Delgenes J.P. & Moletta R. (2000). Aerobic denitrifiers isolated from diverse natural and managed ecosystems. *Microbial Ecology* **39**(2): 145–152.
DOI: <https://doi.org/10.1007/s002480000009>
- Pintathong P., Richardson D.J., Spiro S. & Choirit W. (2009). Influence of metal ions and organic carbons on denitrification activity of the halotolerant bacterium, *Paracoccus pantotrophus* P16 a strain from shrimp pond. *Electronic Journal of Biotechnology* **12**(2): 4–5.
DOI: <https://doi.org/10.2225/vol12-issue2-fulltext-12>
- Piyathilake I.D.U.H., Ranaweera L.V., Udayakumara E.P.N., Gunatilake S.K. & Dissanayake C.B. (2022). Assessing groundwater quality using the Water Quality Index (WQI) and GIS in the Uva Province, Sri Lanka. *Applied Water Science* **12**(4): 1–19.
DOI: <https://doi.org/10.1007/s13201-022-01600-y>
- Prabagar J., Prabagar S. & Thuraisingam S. (2020). Nitrate contamination of groundwater: An issue for livelihood in Jaffna Peninsula, Sri Lanka. *African Journal of Agricultural Research* **16**(7): 1025–1032.
DOI: <https://doi.org/10.5897/AJAR2020.14740>
- Power J. & Schepers J. (1989). Nitrate contamination of groundwater in North America. *Agriculture, Ecosystems & Environment* **26**: 165–187.
DOI: [https://doi.org/10.1016/0167-8809\(89\)90012-1](https://doi.org/10.1016/0167-8809(89)90012-1)
- Rajakumar S., Ayyasamy P.M., Shanthi K., Thavamani P., Velmurugan P., Song Y.C. & Lakshmanaperumalsamy P. (2008). Nitrate removal efficiency of bacterial consortium (*Pseudomonas* sp. KW1 and *Bacillus* sp. YW4) in synthetic nitrate-rich water. *Journal of Hazardous Materials* **157**(2-3): 553–563.
DOI: <https://doi.org/10.1016/j.jhazmat.2008.01.020>
- Rajta A., Bhatia R., Setia H. & Pathania P. (2019). Role of heterotrophic aerobic denitrifying bacteria in nitrate removal from wastewater. *Journal of applied microbiology* **128**(5): 1261–1278.
DOI: [10.1111/jam.14476](https://doi.org/10.1111/jam.14476)
- Rasihah V., Armour J. & Nelson P. (2013). Nitrate in shallow fluctuating groundwater under sugarcane: Quantifying the lateral export quantities to surface waters. *Agriculture, Ecosystems & Environment* **180**: 103–110.
DOI: <https://doi.org/10.1016/j.agee.2012.07.002>
- Ren H.J., Su Y., Wang C., Hou Z.M., Zhou Y. & Zhou R. (2018). Application of response surface methodology to optimize nitrate removal at low temperature by aerobic denitrificator *Pseudomonas* strain An-1. *Water and Environment Journal* **32** (2): 235–241.
DOI: <https://doi.org/10.1111/wej.12319>
- Sadeq M., Moe C.L.S Attarassi B., Cherkaoui I., Elaouad R. & Idrissi L. (2008). Drinking water nitrate and prevalence of methemoglobinemia among infants and children aged 1–7 years in Moroccan areas. *International Journal of Hygiene and Environmental Health* **211**: 546–554.
DOI: <https://doi.org/10.1016/j.ijheh.2007.09.009>
- Takaya N., Catalan-Sakairi M.A.B., Sakaguchi Y., Kato I., Zhou, Z. & Shoun H. (2003). Aerobic denitrifying bacteria that produce low levels of nitrous oxide. *Applied and Environmental Microbiology* **69**(6): 3152–3157.
DOI: <https://doi.org/10.1128/AEM.69.6.3152-3157.2003>
- Tokazhanov G., Ramazanov E., Hamid S., Bae S. & Lee W. (2020). Advances in the catalytic reduction of nitrate by metallic catalysts for high efficiency and N₂ selectivity: a review. *Chemical Engineering Journal* **384**: 123252.
DOI: <https://doi.org/10.1016/j.cej.2019.123252>
- Vandekerckhove T.G., Kobayashi K., Janda J., Van Nevel S. & Vlaeminck S. E. (2018). Sulfur-based denitrification treating regeneration water from ion exchange at high performance and low cost. *Bioresource Technology* **257**: 266–273.
DOI: <https://doi.org/10.1016/j.biortech.2018.02.047>
- WHO (2017). *Guidelines for Drinking-water Quality : First addendum to the fourth edition*. WHO, Geneva, Switzerland.
- Wu Z., Huang S., Yang Y., Xu F., Zhang Y. & Jiang R. (2013). Isolation of an aerobic denitrifying bacterial strain from a biofilter for removal of nitrogen oxide. *Aerosol and Air Quality Research* **13**(3): 1126–1132.
DOI: <https://doi.org/10.4209/aaqr.2012.07.0199>
- Xu J., Pu Y., Qi W.K., Yang X.J., Tang Y., Wan P. & Fisher A. (2017). Chemical removal of nitrate from water by aluminum-iron alloys. *Chemosphere* **166**: 197–202. DOI: <https://doi.org/10.1016/j.chemosphere.2016.09.102>
- Yin Y. & Yan Z. (2020). Variations of soil bacterial diversity and metabolic function with tidal flat elevation gradient in an artificial mangrove wetland. *Science of the Total Environment* **718**: 137385.
DOI: <https://doi.org/10.1016/j.scitotenv.2020.137385>
- Zhang J., Wu P., Hao B. & Yu Z. (2011). Heterotrophic nitrification and aerobic denitrification by the bacterium *Pseudomonas stutzeri* YZN-001. *Bioresource Technology* **102**(21): 9866–9869.
DOI: <https://doi.org/10.1016/j.biortech.2011.07.118>
- Zhang Z., Xu Y., Shi W., Wang W., Zhang R., Bao X. & Cui F. (2016). Electrochemical-catalytic reduction of nitrate over Pd–Cu/γAl₂O₃ catalyst in cathode chamber: enhanced removal efficiency and N₂ selectivity. *Chemical*

- Engineering Journal* **290**: 201–208.
DOI: <https://doi.org/10.1016/j.cej.2016.01.063>
- Zhang H., Feng J., Chen S., Zhao Z., Li B., Wang Y. & Hao H. (2019 b). Geographical patterns of nirS gene abundance and nirS-type denitrifying bacterial community associated with activated sludge from different wastewater treatment plants. *Microbial Ecology* **77**(2): 304–316.
DOI: <https://doi.org/10.1007/s00248-018-1236-7>
- Zhang H., Wang Y., Huang T., Liu K., Huang X., Ma B. & Sekar R. (2020). Mixed-culture aerobic anoxygenic photosynthetic bacterial consortia reduce nitrate: Core species dynamics, co-interactions and assessment in raw water of reservoirs. *Bioresource Technology* **315**: 123817.
DOI: <https://doi.org/10.1016/j.biortech.2020.123817>
- Zhou Y., Suenaga T., Qi C., Riya S., Hosomi M. & Terada A. (2021). Temperature and oxygen level determine N₂O respiration activities of heterotrophic N₂O-reducing bacteria: Biokinetic study. *Biotechnology and Bioengineering* **118**(3): 1330–1341. DOI: <https://doi.org/10.1002/bit.27654>

RESEARCH ARTICLE

Dairy Technology

Physicochemical and sensory attributes of inulin incorporated set yoghurt prepared using cow and buffalo milk

KADC Gnanarathna¹, MTLK Jayasumama¹, KWDP Weragama¹, DCS Gunasekara², WVVR Weerasingha^{1*}, A Chandrasekara³ and PHP Prasanna⁴

¹ Department of Animal and Food Sciences, Faculty of Agriculture, Rajarata University of Sri Lanka, Mihintale.

² Food and Nutrition Research Center, CIC Agribusiness (Pvt) Ltd., Pelwehera, Dambulla.

³ Department of Applied Nutrition, Faculty of Livestock Fisheries and Nutrition, Wayamba, University of Sri Lanka, Kuliyapitiya.

⁴ Department of Animal Science, Faculty of Agriculture, University of Peradeniya, Peradeniya.

Submitted: 02 September 2021; Revised: 14 March 2022; Accepted: 22 April 2022

Abstract: The demand for probiotic food products has been increased over the last few decades. The current study aimed to evaluate the microbiological and physicochemical properties of two set yoghurt types prepared using cow milk (CM) and buffalo milk (BM) with three different concentrations of inulin (1%, 2% and 3% w/v). Two types of starter cultures BB12 (*Bifidobacterium bifidum* – BB) and YC-X11 (*Streptococcus thermophilus* – ST and *Lactobacillus bulgaricus* subsp. *Bulgaricus* – LB) were used as probiotic sources. Probiotic viable cell counts, syneresis, total acidity, pH, organoleptic properties, and viscosity were evaluated during 21 days of storage at 4 °C. Both inulins incorporated CMY and BMY showed significantly high ($p < 0.05$) ST and LB counts at the end of the enumeration period (~ 8.0 log cfu/mL) while control samples of each group showed significantly lower (~ 7.0 log cfu/mL) probiotic counts. 3% CMY and BMY reported the highest probiotic survivability. Significantly ($p < 0.05$) high viscosity was observed in 2% inulin incorporated yoghurts which were highly acceptable and showed better stability during storage. This study concludes that the addition of inulin improves the overall probiotic survivability of the tested microorganisms while ensuring consumer acceptance without drastically altering the physicochemical properties of yoghurt.

Keywords: *Bifidobacterium*, buffalo milk, inulin, *Lactobacillus*, probiotic, yoghurt.

INTRODUCTION

The consumer approach towards eating healthy foods has evolved during the last decade, enhancing the potential to develop various functional foods that provide more benefits than simply nutritional value. Among the functional foods, much attention has been focused on probiotic products. Probiotics are health-promoting viable microorganisms such as *Lactobacillus acidophilus*, *Bifidobacterium bifidum*, and *Bifidobacterium longum* that exhibit a beneficial effect on human health, by improving the intestinal microbial balance (Martin-Diana *et al.*, 2003). In addition, probiotics improve lactose utilization by producing lactase enzyme which helps to digest lactose, preventing intestinal infections by producing organic acids and other antibacterial agents, reducing cancer risk, improving immunity, lowering cholesterol and blood pressure, and protecting against diarrhoea, constipation and osteoporosis (Ranadheera *et al.*, 2010). Several studies have demonstrated the modes of actions of probiotics as antioxidants, antimicrobial compounds, and agents that improve bowel functions (Beena *et al.*, 2012).

* Corresponding author (vviraj@agri.rjt.ac.lk;  <https://orcid.org/0000-0002-5234-8266>)



In order to ensure the survival of probiotic strains, most probiotic-containing foods also include prebiotic sources. A prebiotic is defined as a 'non-digestible food ingredient that beneficially affects the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon, thus improving the health of the host' (Gibson *et al.*, 2004). Ingestion of prebiotics significantly enhances the proliferation of colon microflora by acting as a substrate for probiotics, while improving the colonization rate of bacterial strains such as *Lactobacillus* and *Bifidobacterium* which have high probiotic activity. The prebiotics can balance the composition of microflora in the human gut (Ranadheera *et al.*, 2010). In many foods, non-digestible oligosaccharides and fructooligosaccharides are used as prebiotics. The most utilized prebiotic in yoghurt manufacturing is inulin. It is a non-digestible and fermentable fructan (Ehsani *et al.*, 2016) that stimulates the growth and activity of probiotic bacteria. Inulin consumption significantly increases faeces loading and improves faecal excretion in humans. Thus, inulin is effective against colon disorders such as constipation.

In the food industry, inulin is often used as a low-calorie sweetener and a non-digestible fibre that forms gels which increase viscosity and enhance organoleptic properties (Meyer *et al.*, 2011). Considering these beneficial health effects imparted by both prebiotic compounds and probiotic microbial strains, current food manufacturers tend to produce symbiotic foods. Yoghurt is such a food that is widely consumed and produced by incorporating prebiotics with the probiotics *Lactobacillus acidophilus*, *Lactobacillus casei*, and *Bifidobacterium* to enhance their therapeutic value. The viability of the probiotic bacterial count should be as high as 10^6 - 10^7 CFU/mL at consumption to achieve the aforementioned benefits (de Souza Oliveira *et al.*, 2011). Hence, it is essential to ensure sufficient and viable probiotic bacteria throughout the shelf life of the yoghurt. Recent studies have shown that most probiotic foods have a low population of probiotic bacteria at the time of consumption. The main dairy source of yoghurt is fresh cow milk. When compared to cow milk, buffalo milk has a higher content of fat, protein, lactose and total solids (Abesinghe *et al.*, 2020). These properties of buffalo milk give a rich flavour and taste and make it very suitable for the manufacture of a wide variety of milk products (Han *et al.*, 2012b). Moreover, recent studies on the shelf life of buffalo milk yoghurt have indicated that it could be a good vehicle for probiotics (Han *et al.*, 2012a). As a result, the potential of utilizing buffalo milk to produce a symbiotic yoghurt was examined. The present study

aims to explore and compare the physicochemical and probiotic activity of cow and buffalo milk set yoghurt to identify the best probiotic carrier with a suitable level of prebiotic inulin addition.

MATERIALS AND METHODS

Preparation of set type yoghurt

Cow milk set yoghurt (CMY) and buffalo milk set yoghurt (BMY) were prepared according to the following methodology. Milk was pre-heated at 65 °C and inulin powder (Orafti® HPX, High-Performance Inulin Powder, Mannheim, Germany) was added in three concentrations of 1%, 2% and 3% (w/v). Then, skim milk powder (SMP) was added at concentrations of 3% (w/v) for the control, 2% (w/v) SMP for 1% inulin treated sample and 1% (w/v) SMP for 2% and 3% inulin treated samples respectively. The mixtures were homogenized for 15-20 min at 5 Mpa. The milk was then pasteurized at 90-95 °C for 5-10 min followed by cooling down to 36 °C. A starter culture containing *Streptococcus thermophilus* (ST) and *Lactobacillus delbrueckii* subsp. *bulgaricus* (LB) (Chr. Hansen YC-X11, Hoersholm, Denmark) and *Bifidobacterium bifidum* (BB) (Chr. Hansen, Hoersholm, Denmark) were added in 0.002% (w/v) concentration. The yoghurts were incubated in cups at 42 °C for 4 h. Finally, the samples were stored at 4 °C pending analyses.

Preparation of microbiological isolates

The pour plate method was used to evaluate the growth and viability of yoghurt bacteria, and bifidobacterial (Shori & Baba, 2012) MRS media (M641, HIMEDIA, Mumbai, India) was used to isolate *Lactobacillus*. Commercial M17 agar base (HIMEDIA, Mumbai, India) was used for isolation of the *S. thermophilus* strain, and commercial Bifidobacterium agar (HIMEDIA, Mumbai, India) was used to isolate *Bifidobacterium* strains of the samples. A dilution series was prepared by using 0.1% (w/w) peptone (CDH, India) solution for all microbiological analyses. All culture media were incubated at 35-37 °C for 48-72 h in the laboratory incubator (IN- 601/GEMMYCO) in anaerobic conditions. Colony counts were obtained at 1, 7, 14, and 21 ds by hand enumeration through a colony counter (Witeg Colony Counter Galaxy 330, Germany). For each microbial strain, three nearby dilutions were prepared and for each dilution, duplicate plates were prepared and enumerated. The microbial counts were expressed as log colony forming units per millilitre (log CFU/mL).

Physicochemical analyses

pH measurement

Changes in the pH during the storage period were measured using a multi-parameter pH meter (35 Series, Eutech Instruments, Singapore). The pH probe was immersed in 25 mL of homogenized sample and the readings were taken in relation to room temperature (27 ± 2 °C) in triplicate. Readings were collected on the 1st, 7th, 14th, and 21st ds after manufacturing.

Titrateable acidity

A slurry of each sample was made by mixing 10 g of yoghurt with 5 mL of distilled water. Samples were titrated against 0.1 N NaOH solution, in the presence of 1% phenolphthalein until a persistent pink colour was observed. Acidity is expressed as a percentage of lactic acid (Jayarathna *et al.*, 2020). Analysis was carried out in triplicate.

Syneresis

Twenty grams of each sample was placed on a Whatman No.1 filter paper and the weight of the filtrate was measured for 2 h with 30 min time intervals in triplicate. Syneresis is expressed as the percentage of the initial sample (Amal *et al.*, 2016).

Texture

Texture analyses were completed using a programmable digital viscometer (Brookfield DVE Digital Viscometer, Massachusetts, USA) at 4 °C, using spindle numbers 6 and 7 at speed of 10 rpm. Readings were recorded after 30 s and expressed in millipascal-second (mPa.s) on the 7th day after manufacturing.

Sensory evaluation

Sensory evaluation of inulin incorporated cow milk and buffalo milk yoghurt was carried out with 30 untrained panelists (aged 22-55 years, women 60%, men 40%) according to a randomized design in order to find out the degree of likeability of the probiotic yoghurt with the most preferable level of inulin. The study was carried out at least 2 h after a main meal and smokers, betel chewers, or alcohol consumers were not recruited as panelists. Each panelist evaluated the final yoghurt product according to a five-point hedonic scale (1-5, from strongly disliked to strongly liked, respectively) using a simplified and structured sensory evaluation ballot at 25 °C.

Statistical analysis

Physicochemical data of the yoghurt were analysed by using one-way ANOVA of Minitab 18 software (Minitab® 17.1.0, UK) with a 95% confidence interval (significance at the $p < 0.05$) while the Friedman test was used to analyze sensory data. Microbiological data were analysed via one way ANOVA of the SAS program (Version 9.0, SAS Inc, North Carolina, USA).

RESULTS AND DISCUSSION

Changes in the microbial count during the storage

In consideration of variations of viable ST counts in CMY during the end of storage, there was a significant difference ($p < 0.05$) of viability between 2% and 3% inulin incorporated yoghurt samples in contrast to the control and 1% inulin incorporated samples (Table 1). The highest ST count throughout the storage period in both CMY and BMY was recorded from 3% inulin incorporated samples (9.11 ± 0.07 log CFU/mL to 8.72 ± 0.14 log CFU/mL). There was no significant difference in initial viable counts among the samples, although there was a significant reduction of the viable counts in both 2 % and 3 % CMY on the 21st day of storage. The viable counts in CMY followed a gradual decline towards the end of the storage period, yet they were remaining in the range of critical microbial count which is $>10^6$ CFU/mL.

It is evident that ST has dominantly grown in both cow milk and buffalo milk set yoghurts. Oliveira *et al.* (2002), reported that ST predominated in all products, showing higher counts than 9 log CFU/mL in fermented milk prepared with mixed culture.

The yoghurt samples prepared with buffalo milk showed a significant difference ($p < 0.05$) in microbial viability of ST among inulin incorporated samples and the control sample at the end of storage. The highest viable count on the 21st day of storage was reported from the 3% inulin incorporated BMY sample (8.23 ± 0.12 log CFU/mL). Considering the BMY, they showed no significant difference ($p > 0.05$) compared to their initial viable count throughout the storage up to the 14th day, yet they showed a significant reduction ($p < 0.05$) when comparing the initial stage and the final day of storage. However, the viable ST counts were still above the critical count. This prominent and better viability of ST is likely due to its enhanced proteolytic activity combined with higher lactose utilization, compared to the other two probiotic strains (Özer & Kirmaci, 2010).

Table 1: Viable counts of *Streptococcus thermophilus* (ST) in cow and buffalo milk set yoghurt stored at 4 °C during 21 days of storage (log CFU/mL)

Treatment	Period of storage in days (Mean ± SD)			
	1	7	14	21
Cow milk yoghurt				
Control	8.34 ± 0.14 ^{Ba}	8.27 ± 0.17 ^{Ca}	8.13 ± 0.15 ^{Ba}	7.93 ± 0.12 ^{Ca}
1% Inulin	8.46 ± 0.1 ^{Ba}	8.49 ± 0.06 ^{BCa}	8.41 ± 0.12 ^{Ba}	8.27 ± 0.17 ^{Ba}
2% Inulin	8.82 ± 0.14 ^{ABa}	8.75 ± 0.13 ^{ABab}	8.65 ± 0.09 ^{Bab}	8.49 ± 0.06 ^{ABb}
3% Inulin	9.11 ± 0.07 ^{Aa}	9.08 ± 0.08 ^{Aa}	9.04 ± 0.08 ^{Aab}	8.78 ± 0.12 ^{Ab}
Buffalo milk yoghurt				
Control	8.13 ± 0.15 ^{Ba}	8.08 ± 0.09 ^{Ca}	7.78 ± 0.33 ^{Bab}	7.54 ± 0.16 ^{Bb}
1% Inulin	8.38 ± 0.09 ^{ABa}	8.31 ± 0.10 ^{BCa}	8.13 ± 0.15 ^{ABab}	7.93 ± 0.12 ^{Ab}
2% Inulin	8.63 ± 0.12 ^{ABa}	8.58 ± 0.11 ^{ABa}	8.43 ± 0.15 ^{Aab}	8.13 ± 0.15 ^{Ab}
3% Inulin	8.72 ± 0.14 ^{Aa}	8.65 ± 0.10 ^{Aa}	8.55 ± 0.14 ^{Aab}	8.23 ± 0.12 ^{Ab}

Means in the same row without a common simple letter superscript significantly ($p < 0.05$) differ for each treatment; means in the same column without a common capital letter superscript significantly ($p < 0.05$) differ among the treatments per microbial analysis.

Table 2: Viable counts of *Bifidobacterium bifidum* (BB) in cow and buffalo milk set yoghurt stored at 4 °C during 21 days of storage (log CFU/mL)

Treatment	Period of storage (Days) (Mean ± SD)			
	1	7	14	21
Cow milk yoghurt				
Control	8.31 ± 0.10 ^{Ba}	8.23 ± 0.12 ^{Ca}	8.13 ± 0.15 ^{Ba}	7.69 ± 0.21 ^{Bb}
1% Inulin	8.63 ± 0.12 ^{ABa}	8.49 ± 0.06 ^{Ca}	8.49 ± 0.14 ^{ABa}	8.31 ± 0.10 ^{Aa}
2% Inulin	8.75 ± 0.13 ^{Aa}	8.67 ± 0.07 ^{ABab}	8.60 ± 0.08 ^{Aab}	8.32 ± 0.1 ^{Ab}
3% Inulin	8.91 ± 0.15 ^{Aa}	8.86 ± 0.13 ^{Aa}	8.78 ± 0.12 ^{Aa}	8.65 ± 0.04 ^{Aa}
Buffalo milk yoghurt				
Control	8.00 ± 0.21 ^{Ba}	8.67 ± 0.10 ^{Bab}	8.46 ± 0.10 ^{Bb}	7.39 ± 0.21 ^{Bab}
1% Inulin	8.35 ± 0.04 ^{Aa}	8.31 ± 0.10 ^{Aa}	8.28 ± 0.05 ^{Aa}	8.17 ± 0.21 ^{Aa}
2% Inulin	8.58 ± 0.05 ^{Aa}	8.46 ± 0.10 ^{Aa}	8.41 ± 0.12 ^{Aa}	8.35 ± 0.04 ^{Aa}
3% Inulin	8.67 ± 0.10 ^{Aa}	8.46 ± 0.10 ^{Aa}	8.39 ± 0.21 ^{Aa}	8.27 ± 0.17 ^{Aa}

Means in the same row without a common simple letter superscript significantly ($p < 0.05$) differ for each treatment; means in the same column without a common capital letter superscript significantly ($p < 0.05$) differ among the treatments per microbial analysis.

According to Table 2, it is evident that the addition of inulin, the duration of storage, and their collective interaction have had significant ($p < 0.05$) effects upon the BB viability in both CMY and BMY samples. Especially, considering the 1st and 21st days of CMY samples, the control yoghurt sample has reported the lowest BB viability (8.31 ± 0.10 log CFU/mL and 7.69 ± 0.21 log CFU/mL, respectively) in contrast to other

samples, although the viable BB count has not remarkably varied among the inulin incorporated samples. A recent study by Kamel *et al.* (2021) concludes that increasing inulin concentrations helped to amplify the viability of BB significantly ($p < 0.05$) during storage, as well as acting as an antifungal agent of probiotic yoghurt, suggesting that inulin can be used as a suitable and a successful nutrient for BB.

Table 2 shows that significantly lower ($p < 0.05$) viable BB counts were observed with time in each BMY sample, compared to their initial counts, though the viable counts were reported above the critical microbial count up to the 21st day. Similarly, Han *et al.* (2012a) reported that *Bifidobacterium* spp. remained viable throughout 10 weeks of storage when incorporated in

low-fat symbiotic yoghurt. The highest viable counts observed in 3% inulin incorporated yoghurt sample on 1st and 21st days of storage (8.67 ± 0.1 log CFU/mL to 8.27 ± 0.17 log CFU/mL) suggesting that BM yoghurt could provide a suitable environment for the growth of probiotic microbes and maintain therapeutic levels until the end of the storage period.

Table 3: Viable counts of *Lactobacillus bulgaricus* (LB) in cow and buffalo milk set yoghurt stored at 4 °C during 21 days of storage (log CFU/mL)

Treatment	Period of storage (Days) (Mean \pm SD)			
	1	7	14	21
Cow milk yoghurt				
Control	8.28 \pm 0.05 ^{BCa}	8.08 \pm 0.09 ^{Ca}	7.93 \pm 0.12 ^{Bab}	7.69 \pm 0.21 ^{Bb}
1% Inulin	8.54 \pm 0.06 ^{ACa}	8.43 \pm 0.15 ^{BCa}	8.39 \pm 0.21 ^{Aa}	8.27 \pm 0.17 ^{Aa}
2% Inulin	8.79 \pm 0.10 ^{ABa}	8.82 \pm 0.07 ^{ABa}	8.52 \pm 0.09 ^{Aab}	8.39 \pm 0.21 ^{Ab}
3% Inulin	8.96 \pm 0.16 ^{Aa}	8.86 \pm 0.08 ^{Aab}	8.74 \pm 0.11 ^{Aab}	8.59 \pm 0.13 ^{Ab}
Buffalo milk yoghurt				
Control	8.08 \pm 0.09 ^{Ca}	7.93 \pm 0.12 ^{Cab}	7.80 \pm 0.70 ^{Bb}	7.54 \pm 0.39 ^{Bc}
1% Inulin	8.45 \pm 0.07 ^{Ba}	8.38 \pm 0.09 ^{Ba}	8.27 \pm 0.17 ^{Aa}	8.04 \pm 0.27 ^{ABa}
2% Inulin	8.64 \pm 0.02 ^{ABa}	8.44 \pm 0.07 ^{ABab}	8.31 \pm 0.10 ^{ABbc}	8.13 \pm 0.15 ^{Ac}
3% Inulin	8.73 \pm 0.09 ^{Aa}	8.59 \pm 0.13 ^{Aab}	8.50 \pm 0.16 ^{Aab}	8.31 \pm 0.10 ^{Ab}

Mean values in the same row without a common simple letter superscript significantly ($p < 0.05$) differ for each treatment; means in the same column without a common capital letter superscript significantly ($p < 0.05$) differ among the treatments per microbial analysis.

According to Table 3, there is a significant difference ($p < 0.05$) in LB viability between inulin incorporated samples and control samples made with cow milk. The viable counts of LB ranged from 8.28 ± 0.05 log CFU/mL to 7.69 ± 0.21 log CFU/mL on the 1st and 21st day in storage, while the highest viable LB counts were observed in 3% inulin incorporated yoghurt sample, on the 21st day of storage (8.59 ± 0.13 log CFU/mL). However, the viability of LB counts significantly declined with time compared to their initial values yet remained in the critical microbial count even on the final day of storage. A similar study (Saccaro *et al.*, 2009) revealed that the viability of LB was reduced by 1.51 log CFU/mL when grown in co-culture with ST and in association with probiotic microorganisms, which explains the receding nature of LB viability in these mixed culture CMY samples.

The viability of LB in BMY samples also shows a significant difference between inulin incorporated and control samples, where the viable counts of LB in the control sample gave values of 8.08 ± 0.09 log CFU/mL

and 7.54 ± 0.39 log CFU/mL on 1st and 21st day of storage, respectively. However, a significantly different ($p < 0.05$) value for the highest viability of LB was reported from the 3% inulin incorporated sample on the 21st day, which was 8.31 ± 0.1 log CFU/mL. Significantly low viable counts were observed in each treatment compared to their initial counts and the rate of reduction of viable counts was higher in the control sample though the counts remained in the range of critical microbial count up to the 21st day ($>10^6$ CFU/mL).

The results of the study showed the increase of probiotic bacterial counts with increasing inulin concentrations for both milk types. Inulin is a fructan that can be effectively used as a nutrient for microbial growth (Srisuvor *et al.*, 2013). Inulin enhances probiotic proliferation by acting as dietary fibre (Ramirez-Santiago *et al.*, 2010). In addition, lactic acid bacteria (LAB) are capable of producing exopolysaccharides (EPS) during their growth, and these EPS are essential for LAB-host interaction mechanisms such as bacterial tolerance to gut conditions and their immune-modulatory activities,

which helps them to survive in the host's gastrointestinal tract (Caggianello *et al.*, 2016), The increase in inulin levels may lead to high concentrations of EPS amounts with a positive effect on microbial growth.

Lower viability was observed for LB bacteria in both yoghurt types, which may be due to the post acidification (Sendra *et al.*, 2008) leading to lower pH and low availability of bacteria in the commercial starter culture (Chr. Hansen YC-X11, Hoersholm, Denmark). The results of the present study showed that the viability of microbes decreased with time, yet the decrease of LAB viability was less with high concentrations of inulin. These results agree with Ehsani *et al.* (2016), who also observed that counts of probiotic strains declined over four weeks of storage. Donkor *et al.* (2007) suggested that ultimate pH and accumulation of organic acids had a notable impact on the survival of LB and BB in yoghurt. Therefore, considering the results obtained in this study, reduction of energy source with time would be another reason for the declining survival rate of microbes with time, apart from previously mentioned factors.

Changes in pH, titratable acidity, syneresis and texture

Variations of pH in different yoghurts

According to the pH variation data of CMY, there were no significant differences ($p > 0.05$) among the samples throughout the storage period and a decrease of pH during storage was observed in all the treatments with time (Figure 1a). At the end of the experiment, the control sample recorded the lowest pH (4.13 ± 0.01) while the 1% inulin incorporated sample recorded the highest pH value (4.24 ± 0.01).

When considering the pH variation of BMY, there were significant differences ($p < 0.05$) in pH among the treatments throughout the storage period (Figure 1b). The pH varied from 4.86 ± 0.01 to 4.96 ± 0.00 from the 1st day to the 21st. On the 14th day, 1% CMY recorded a significantly higher ($p < 0.05$) pH value compared to other samples (4.97 ± 0.01). However, the pH at the end of the storage period of yoghurts manufactured with inulin addition showed no significant difference ($p > 0.05$) among samples.

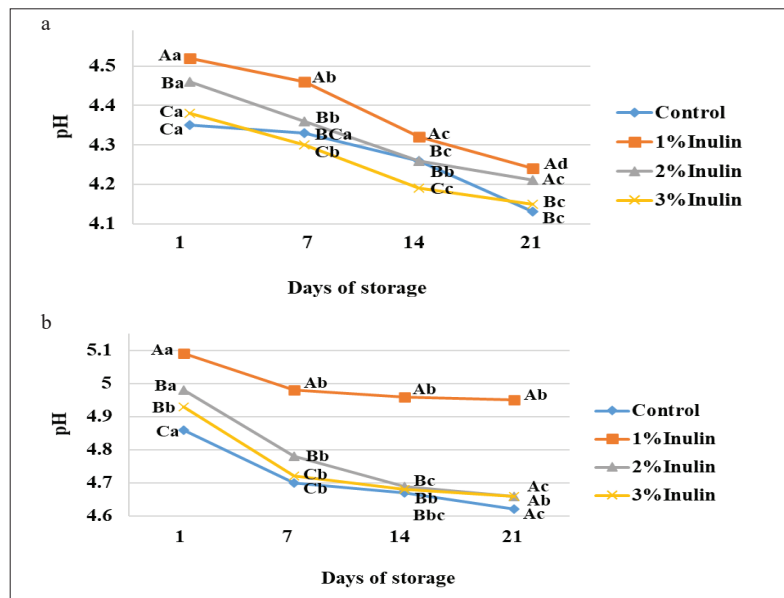


Figure 1: a) Variation in pH of CMY with different inulin concentrations stored at 4 °C during 21 days of storage b) Variation in pH of BMY with different inulin concentrations stored at 4 °C during 21 days of storage.

ABC Means with different uppercase letters are significantly different between each treatment.

abc Means with different lowercase letters are significantly different between days of storage.

In yoghurt processing, lactic acidification is the result of lactose fermentation by the associative growth of the two thermophilic, homofermentative, lactic acid bacteria, ST and LB. Increasing H⁺ concentration in the yoghurt due to acid formation leads to low pH. Similar results were reported by Tiwari *et al.* (2015). The accumulation of organic acids contributes to the continuous reduction of pH throughout the storage. Results of the present study also showed that pH decreased with increasing inulin levels. Continued acidification during the storage can be linked to a progressive transformation of lactose into lactic acid. However, in yoghurts containing high soluble fibre components, other mechanisms could also promote a further increase in acidity level (Ramirez-Santiago *et al.*, 2010). According to Sendra *et al.* (2008), fibre addition itself contributes to a decrease in pH but the increased metabolic activity of the probiotics would also contribute to drastic pH reduction. Increasing inulin levels lead to high bacterial growth, consequently increasing acid production, which tends to lower the pH over time.

Changes in titratable acidity of yoghurts

The titratable acidity (TA) of CMY showed significant differences ($p < 0.05$) between control and inulin added yoghurt samples while there were no significant differences in acidity among treatments throughout 21 days of storage period (Figure 2a). Throughout the storage, the control sample reported the highest titratable acidity levels whereas, the 1% sample reported the lowest titratable acidity values.

According to data of BMY, there was no significant difference in TA between the control sample and inulin added yoghurt samples throughout the storage period. A significant ($p < 0.05$) increase of TA during storage was observed in all the treatments. The TA values were ranged from 0.94 - 0.95% and 0.96 - 0.97% during the storage period of CMY and BMY, respectively. The 1% inulin incorporated BMY sample showed a lower TA throughout the storage period while the 3% inulin incorporated BMY sample showed the highest TA. However, there

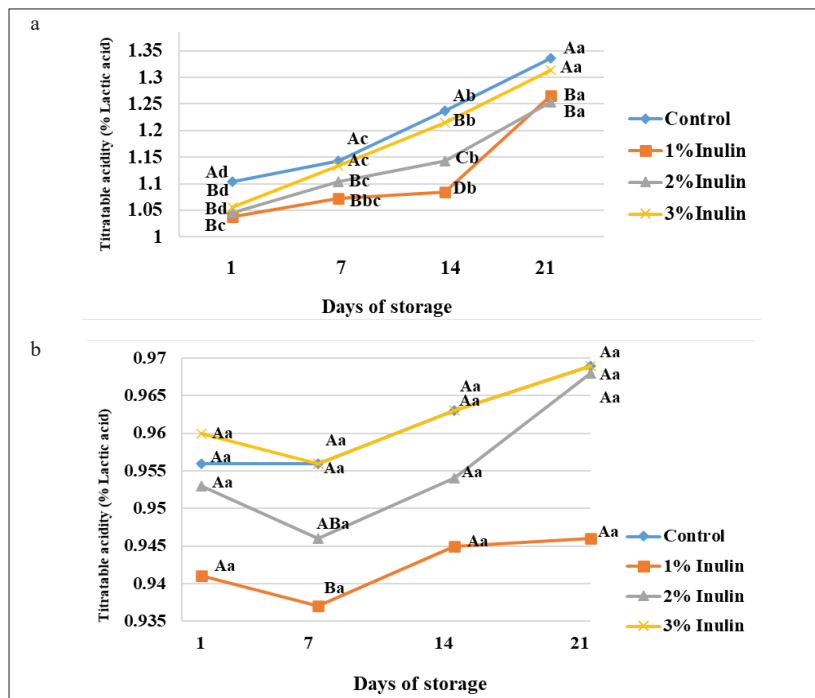


Figure 2: a) Variation in titratable acidity of CMY with different inulin concentration levels stored at 4 °C during 21 days of storage. b) Variation in titratable acidity of BMY with different inulin concentration levels stored at 4 °C during 21 days of storage.

ABC Means with different uppercase letters are significantly different between each treatment.
 abc Means with different lowercase letters are significantly different between days of storage.

was an overall increase in titratable acidity while the pH dropped marginally throughout the storage period. These phenomena indicate continuously ongoing enzymic and metabolic activities of mixed starter cultures in low-temperature storage. These results agree with the findings of Canbulat and Ozcan (2015) who also reported a similar type of pH drop and proportional increase in titratable acidity with inulin incorporated set yoghurts. These results confirmed that there is no drastic change in both pH and titratable acidity with the incorporation of inulin into CM and BM yoghurts.

Variations of syneresis in yoghurt during storage

The results of CMY showed a significant ($p < 0.05$) difference in syneresis among treatments throughout the storage period (Figure 3a). On the 1st day, the highest syneresis was observed in the 3% CMY sample ($32.51 \pm 0.50\%$) and the lowest value in the 2% CMY ($30.37 \pm 0.38\%$). On the 21st day, the lowest syneresis was observed with the yoghurt produced with the 2% inulin CMY ($14.3 \pm 0.28\%$), while the highest value was observed in the 3% inulin CMY ($16.3 \pm 0.0\%$).

With BMY, the control sample showed significantly ($p < 0.05$) lowest syneresis ($23.81 \pm 0.08\%$) while the 3% inulin BMY sample gave the highest syneresis value ($26.12 \pm 0.10\%$) on the 1st day of storage. At the end of the storage period, the highest syneresis of $14.3 \pm 0.28\%$ was observed in the 3% inulin inclusion, while the lowest syneresis value of $9.27 \pm 0.28\%$ was observed in the 2% inulin incorporated BMY sample. A significant ($p < 0.05$) decrease of syneresis during the storage time was observed in all the BMY yoghurts throughout the storage period.

Inulin, when incorporated into the protein matrices, forms micro-crystals that interact with each other to form small aggregates. These aggregates are capable of encapsulating water, decreasing the whey removal (syneresis), while creating a smooth and creamy mouthfeel (Kip et al., 2006). It was observed that the 2% inulin incorporated yoghurt samples displayed a similar pattern of behaviour to that of the control sample when compared with the other two concentrations. The 1% and 3% inulin incorporated CMY and BMY samples had higher syneresis levels than the 2% inulin yoghurts.

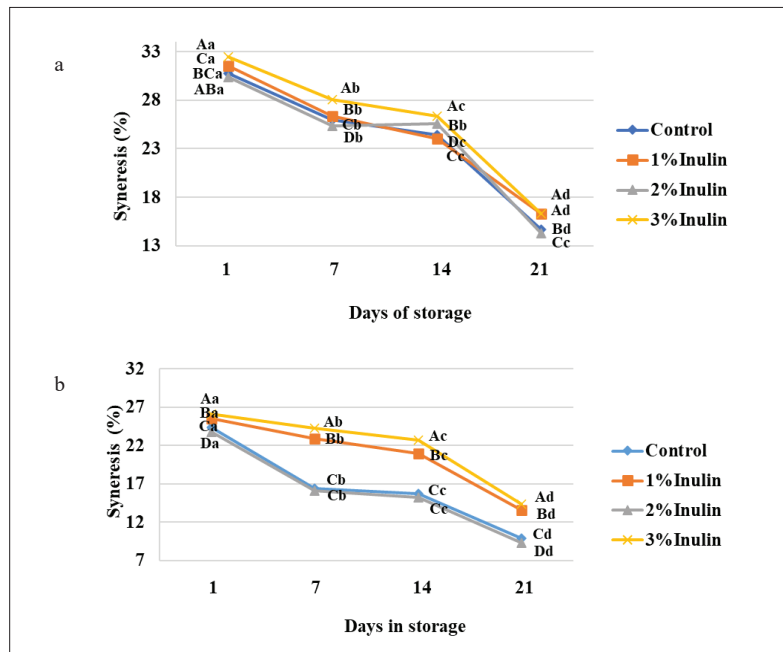


Figure 3: a) Variation in syneresis (%) of CMY with different inulin concentration levels stored at 4 °C during 21 days of storage. b) Variation in syneresis (%) of BMY with different inulin concentration levels stored at 4 °C during 21 days of storage.

^{ABC} Means with different uppercase letters are significantly different between each treatment.
^{abc} Means with different lowercase letters are significantly different between days of storage.

Perhaps, the 1% inulin concentration was insufficient to form a strong tridimensional network of insoluble submicron crystalline inulin particles that entraps large amounts of water, hence the rate of syneresis was higher. In contrast, it has been observed that gel stiffness becomes lower when the inulin concentration is increased. Therefore, the highest syneresis levels observed by 3% inulin incorporated CMY and BMY samples can be explained in terms of the weak network resulting from decreased amounts of casein-casein bonds. It was observed that 2% inulin incorporation level has the ability to retain the strength and rigidity of the yogurt gel matrix without affecting its structure or the rate of whey removal in a drastic manner. Amatayakul *et al.* (2006) observed a reduction in whey separation when using exopolysaccharide (EPS) producing LAB starter cultures. YCX-11 the commercial starter culture used during this study can produce EPS which can facilitate the water binding capacity and improve the yoghurt microstructure (Han *et al.*, 2016). The resulting increase in inulin level could enhance the EPS production reducing the syneresis process compared to the control sample.

Table 4: Changes in the viscosity of CMY and BMY with refrigerated storage (4 °C) for 21 days.

Treatment	Viscosity (x10 ³ mPa.s) (Mean ± SD)
Cow milk yoghurt	
Control	49.850 ± 0.07 ^C
1% inulin	46.75 ± 0.07 ^D
2% inulin	55.85 ± 0.07 ^A
3% inulin	54.05 ± 0.07 ^B
Buffalo milk yoghurt	
Control	70.85 ± 0.07 ^B
1% inulin	62.85 ± 0.07 ^C
2% inulin	86.45 ± 0.07 ^A
3% inulin	86.40 ± 0.06 ^A

Means in the same column without a common capital letter superscript significantly ($p < 0.05$) differ for each treatment.

Viscosity changes in yoghurt

Overall, there seems to be an increased viscosity in BMY yoghurts when compared to CMY (Table 4). This phenomenon can be linked to the differences in the two types of milk. The average amount of total solids in buffalo milk is 16.3%, whereas in cow milk it is 12.8%. In addition, buffalo milk has high amounts of protein, ash, fat and lactose content (Abesinghe *et al.*, 2020).

Viscosity is directly related to the curd stability of yoghurt. According to Srisuvor *et al.* (2013), the curd stability of the yoghurt is one of the important physical properties in yoghurt matrices. Parameters such as total solids and protein content, the heating process, homogenization, the acidity of yoghurt, and the activity of starter culture bacteria generally influence curd stability. The structure and the chemical composition of the added ingredients mainly affect the strength of the casein aggregate. The generated branching structure enhances the protein and carbohydrate interaction giving the gel better curd stability. According to the results, 2% inulin incorporated yoghurt samples had a much higher viscosity when compared to 1% inulin incorporated yoghurts, though there was only a slightly lower viscosity at the 3% incorporation level compared to the 2% incorporation level. Viscosity reduction may be due to the incompatibility of gel structure formed with the interactions of inulin. Increasing inulin levels developed the casein aggregates only up to a certain limit, if inulin addition exceeded the critical value, layer separation could happen, and leading to lower viscosity. Similar observations were made by Helal *et al.* (2018), where the addition of 1% and 2% inulin increased the viscosity of a low-fat yoghurt by 14.7 % and 17.9% respectively, yet the addition of 3% inulin only showed an 8% increase in the viscosity of the yoghurt samples.

Sensory evaluation

According to Figure 4a), 2% CMY gave the highest preference for the odour, colour, texture, taste and overall acceptability while the control sample showed the lowest acceptability. Figure 4b) shows that panelist perception for odour and taste was higher in 3% IBM while colour, texture and overall acceptability were higher in 2% IBM. The control yoghurt sample had the lowest acceptability. However, the sensory results of the present study directly implied inulin addition improved the organoleptic properties of set yoghurt without affecting negatively the taste, odour and colour of the yoghurt. The addition of inulin influences the perception of 'creaminess.' Therefore, among these samples, with increasing inulin concentrations, the perception of 'creaminess' has increased as well. Increasing inulin levels lead to high acceptability via taste and odour while excess addition of inulin lowers the acceptability due to texture degradation. The acceptability may increase due to inulin addition. Inulin enhances the taste (due to fructose), odour, and colour of the yoghurt, though it adversely affects improving the texture by adding additional total soluble solids to yoghurt. Staffolo *et al.* (2004) stated that yoghurt supplemented with inulin maintained its

colour stability during storage, and Seydim *et al.* (2005) stated that yoghurt made with inulin addition had more perceptible flavour and a smooth texture.

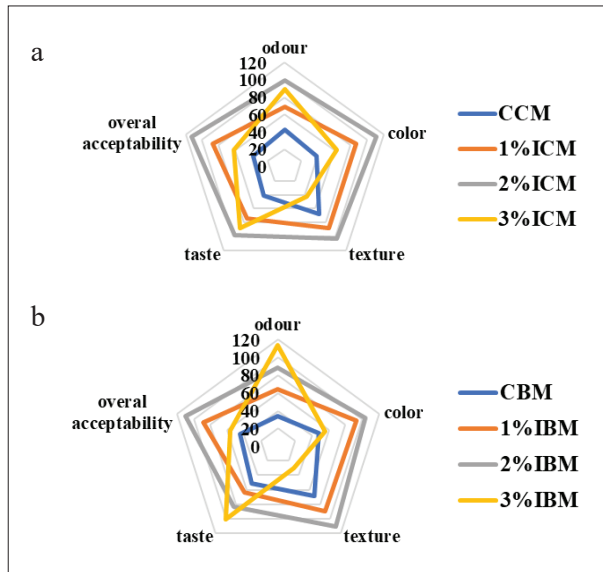


Figure 4: a) Variation of sensory properties in CMY with varying levels of inulin incorporation. CCM, 1% ICM, 2% ICM and 3% ICM samples contain varying levels of inulin, w/w = 0% (Control), 1%, 2%, and 3%, respectively. b) Variation of sensory properties in BMY with varying levels of inulin incorporation. CBM, 1% IBM, 2% IBM and 3% IBM samples contain varying levels of inulin, w/w = 0% (Control), 1%, 2%, and 3%, respectively.

CONCLUSION

The finding of the present study concludes that incorporation of inulin enhances the survival of *S. thermophilus*, *L. bulgaricus*, and *B. bifidum* in set type cow milk and buffalo milk yoghurts during refrigerated storage for 21 days. Cow milk was shown to be a better probiotic carrier than buffalo milk during the storage. It confirms that the addition of 2% inulin into cow milk and buffalo milk yoghurts results in low syneresis without altering the textural properties of the yoghurt matrices. The sensory evaluation results showed that additional inulin concentrations help to increase the ‘creaminess’ in yoghurts, hence indicating the 2% inulin incorporation as the optimal condition for yoghurt with both cow milk and buffalo milk.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- Abesinghe A.M.N.L., Priyashantha H., Prasanna P.H.P., Kurukulasuriya M.S., Ranadheera C.S. & Vidanarachchi J.K. (2020). Inclusion of probiotics into fermented buffalo (*Bubalus bubalis*) milk: an overview of challenges and opportunities. *Fermentation* **6**(4): 121.
DOI: <https://doi.org/10.3390/fermentation6040121>
- Amal A., Eman A. & Nahla S. Z. (2016). Fruit flavored yogurt: Chemical, functional and rheological properties. *International Journal of Environmental and Agriculture Research* **2**(5): 57–66.
DOI: <https://doi.org/10.3390/fermentation6040121>
- Amatayakul T., Sherkat F. & Shah N.P. (2006). Syneresis in set yogurt as affected by EPS starter cultures and levels of solids. *International Journal of Dairy Technology*: **59**(3): 216–221.
DOI: <https://doi.org/10.1111/j.1471-0307.2006.00264.x>
- Beena Divya J., Kulangara Varsha K., Madhavan Nampoothiri K., Ismail B., & Pandey A. (2012). Probiotic fermented foods for health benefits. *Engineering in Life Sciences*: **12**(4): 377–390.
DOI: <https://doi.org/10.1002/elsc.201100179>
- Caggianiello G., Kleerebezem M. & Spano G. (2016). Exopolysaccharides produced by lactic acid bacteria: from health-promoting benefits to stress tolerance mechanisms. *Applied Microbiology and Biotechnology* **100**(9): 3877–3886.
DOI: <https://doi.org/10.1007/s00253-016-7471-2>
- Canbulat Z., & Ozcan T. (2015). Effects of short-chain and long-chain inulin on the quality of probiotic yogurt containing *Lactobacillus rhamnosus*. *Journal of Food Processing and Preservation* **39**(6): 1251–1260.
DOI: <https://doi.org/10.1111/jfpp.12343>
- de Souza Oliveira R.P., Perego P., de Oliveira M.N. & Converti A. (2011). Effect of inulin as prebiotic and synbiotic interactions between probiotics to improve fermented milk firmness. *Journal of Food Engineering* **107**(1): 36–40.
DOI: <https://doi.org/10.1016/j.jfoodeng.2011.06.005>
- Donkor O. N., Nilmini S.L.I., Stolic P., Vasiljevic T. & Shah N.P. (2007). Survival and activity of selected probiotic organisms in set-type yoghurt during cold storage. *International Dairy Journal* **17**(6): 657–665.
DOI: <https://doi.org/10.1016/j.idairyj.2006.08.006>
- Ehsani A., Banihabib E.K., Hashemi M., Saravani M. & Yarahmadi E. (2016). Evaluation of various properties of symbiotic yoghurt of buffalo milk. *Journal of Food*

- Processing and Preservation* **40**(6): 1466–1473.
DOI: <https://doi.org/10.1111/jfpp.12732>
- Gibson G.R., Probert H.M., Van Loo J., Rastall R.A. & Roberfroid M.B. (2004). Dietary modulation of the human colonic microbiota: updating the concept of prebiotics. *Nutrition Research Reviews* **17**(2): 259–275.
DOI: <https://doi.org/10.1079/NRR200479>
- Han X., Lee F.L., Zhang L. & Guo M.R. (2012a). Chemical composition of water buffalo milk and its low-fat symbiotic yogurt development. *Functional Foods in Health and Disease* **2**(4): 86–106.
DOI: <https://doi.org/10.31989/ffhd.v2i4.96>
- Han X., Yang Z., Jing X., Yu P., Zhang Y., Yi H. & Zhang L. (2016). Improvement of the texture of yogurt by use of exopolysaccharide producing lactic acid bacteria. *Biomed Research International* **2016**:7945675.
DOI: <http://dx.doi.org/10.1155/2016/7945675>
- Han X., Zhang L. & Guo M. R. (2012b). Survivability of probiotics in symbiotic low fat buffalo milk yogurt. *African Journal of Biotechnology* **11**(59): 12331–12338.
DOI: <https://doi.org/10.5897/AJB12.1076>
- Helal A., Rashid N., Dyab M., Otaibi M. & Alnemr T. (2018). Enhanced functional, sensory, microbial and texture properties of low-fat set yogurt supplemented with high-density inulin. *Journal of Food Processing & Beverages* **6**(1): 1–11.
- Jayarathna S., Priyashantha H., Johansson M., Vidanarachchi J.K., Jayawardana B.C. & Liyanage R. (2021). Probiotic enriched fermented soy-gel as a vegan substitute for dairy yoghurt. *Journal of Food Processing and Preservation* **45**(1): e15092.
DOI: <https://doi.org/10.1111/jfpp.15092>
- Kamel D.G., Hammam A.R., Alsaleem K.A. & Osman D.M. (2021). Addition of inulin to probiotic yogurt: Viability of probiotic bacteria (*Bifidobacterium bifidum*) and sensory characteristics. *Food Science & Nutrition* **9**(3): 1743–1749.
DOI: <https://doi.org/10.1002/fsn3.2154>
- Kip P., Meyer D. & Jellema R.H. (2006). Inulins improve sensoric and textural properties of low-fat yoghurts. *International Dairy Journal* **16**(9): 1098–1103.
DOI: <https://doi.org/10.1016/j.idairyj.2005.10.011>
- Martín-Diana A.B., Janer C., Peláez C. & Requena T. (2003). Development of a fermented goat's milk containing probiotic bacteria. *International Dairy Journal* **13**(10): 827–833.
DOI: [https://doi.org/10.1016/S0958-6946\(03\)00117-1](https://doi.org/10.1016/S0958-6946(03)00117-1)
- Meyer D., Bayarri S., Tárrega A. & Costell E. (2011). Inulin as texture modifier in dairy products. *Food Hydrocolloids* **25**(8): 1881–1890.
DOI: <https://doi.org/10.1016/j.foodhyd.2011.04.012>
- Oliveira M. N., Sodini I., Remeuf R., Tissier J. P. & Corrieu G. (2002). Manufacture of fermented lactic beverages containing probiotic cultures. *Journal of Food Science* **67**(6): 2336–2341.
DOI: <https://doi.org/10.1111/j.1365-2621.2002.tb09550.x>
- Özer B.H. & Kirmaci H.A. (2010). Functional milks and dairy beverages. *International Journal of Dairy Technology* **63**(1): 1–15.
DOI: <https://doi.org/10.1111/j.1471-0307.2009.00547.x>
- Ramirez-Santiago C., Ramos-Solis L., Lobato-Calleros C., Peña-Valdivia C., Vernon-Carter E.J. & Alvarez-Ramirez J. (2010). Enrichment of stirred yogurt with soluble dietary fiber from *Pachyrhizus erosus* L. Urban: Effect on syneresis, microstructure and rheological properties. *Journal of Food Engineering* **101**(3): 229–235.
DOI: <https://doi.org/10.1016/j.jfoodeng.2010.06.023>
- Ranadheera R.D.C.S., Baines S.K. & Adams M.C. (2010). Importance of food in probiotic efficacy. *Food Research International* **43**(1): 1–7.
DOI: <https://doi.org/10.1016/j.foodres.2009.09.009>
- Saccaro D.M., Tamime A.Y., Pilleggi A.L.O. & Oliveira M.N. (2009). The viability of three probiotic organisms grown with yoghurt starter cultures during storage for 21 days at 4 C. *International Journal of Dairy Technology* **62**(3): 397–404.
DOI: <https://doi.org/10.1111/j.1471-0307.2009.00497.x>
- Sendra E., Fayos P., Lario Y., Fernández-López J., Sayas-Barberá E. & Pérez-Alvarez J. A. (2008). Incorporation of citrus fibers in fermented milk containing probiotic bacteria. *Food Microbiology* **25**(1): 13–21.
DOI: <https://doi.org/10.1016/j.fm.2007.09.003>
- Seydim Z., Sarikus G. & Okur O.D. (2005). Effect of inulin and Dairy-Lo (R) as fat replacers on the quality of set type yogurt. *Milchwissenschaft-Milk Science International* **60**(1):21–25.
- Shori A.B. & Baba A.S. (2012). Viability of lactic acid bacteria and sensory evaluation in *Cinnamomum verum* and *Allium sativum*-bio-yogurts made from camel and cow milk. *Journal of the Association of Arab Universities for Basic and Applied Sciences* **11**(1): 50–55.
DOI: <https://doi.org/10.1016/j.jaubas.2011.11.001>
- Srisuvor N., Chinprahast N., Prakitchaiwattana C. & Subhimaros S. (2013). Effects of inulin and polydextrose on physicochemical and sensory properties of low-fat set yoghurt with probiotic-cultured banana purée. *LWT-Food science and Technology* **51**(1): 30–36.
DOI: <https://doi.org/10.1016/j.lwt.2012.10.018>
- Staffolo M.D., Bertola N. & Martino M. (2004). Influence of dietary fiber addition on sensory and rheological properties of yogurt. *International Dairy Journal* **14**(3): 263–268.
DOI: <https://doi.org/10.1016/j.idairyj.2003.08.004>
- Tiwari A., Sharma H.K., Kumar N. & Kaur M. (2015). The effect of inulin as a fat replacer on the quality of low-fat ice cream. *International Journal of Dairy Technology* **68**(3):374–380.
DOI: <https://doi.org/10.1111/1471-0307.12176>

RESEARCH ARTICLE

High Voltage Engineering

Fractal characteristics of creeping discharges propagating on nano-epoxy composite insulators immersed in coconut oil

S Ediriweera^{1*}, PB Jayarathna², R Samarasinghe², and R Lucas³

¹ Department of Electrotechnology, Faculty of Technology, Wayamba University of Sri Lanka, Kuliyaipitiya.

² Department of Electrical Engineering, Faculty of Engineering, University of Moratuwa, Moratuwa.

³ Department of Electrical, Electronic & Telecommunication Engineering, General Sir John Kothalawala Defence University, Ratmalana.

Submitted: 16 August 2021: Revised: 16 January 2022: Accepted: 25 March 2022

Abstract : A solid/liquid dielectric interface is considered to be one of the weakest points in a composite insulation system, as it facilitates creeping discharges on the interface when the electric field strength exceeds a threshold value. This paper presents a study on the use of nano epoxy composite to minimize the effect of damage which occurs due to creeping discharge activity, when it was immersed in coconut oil. A point-plane electrode arrangement-based test apparatus, energized by a high voltage supply, was used for analysing the propagation of creeping discharges over solid/liquid interfaces using visual observation. An algorithm was used to determine the fractal dimension of creeping discharges propagating over various solid/liquid insulating interfaces. In particular, it focused on the variation in creeping discharge patterns with the use of pure epoxy and nano-composite epoxy samples. The results showed that the use of nanofillers can increase the dielectric breakdown strength of epoxy by 7%. The results also showed nanofillers of 20 nm with 1% concentration do not affect the tortuosity of the creeping discharge pattern on the solid material immersed in coconut oil. However, it can increase the possibility of flashover due to creeping discharge propagation over epoxy insulating material immersed in coconut oil by around 12%.

Keywords: Coconut oil, creeping discharges, final discharge length, fractals, nanotechnology, propagation.

INTRODUCTION

A quality electricity supply has become one of the basic necessities in modern society, demanding an efficient

and uninterrupted power supply. The failure of critical oil-filled high voltage (HV) equipment such as cables, transformers, capacitors, and circuit breakers with many solid/liquid interfaces may lead to catastrophic events, resulting in economic losses and injury to humans. Therefore, it is important for the power utilities to reliably operate and maintain their high voltage equipment (Kebbabi & Beroual, 2006; Ediriweera *et al.*, 2018a). Dielectric liquids are used in the power industry together with solid insulating material to serve as an insulation, impregnation, and cooling medium. In some cases, the solid barrier materials are used to improve the dielectric strength of gas and oil gaps (Liu *et al.*, 2016; Ediriweera *et al.*, 2018b). Dielectric liquid and solid barriers can increase the dielectric strength of oil-filled HV equipment. However, the arrangement of the composite insulation system is also one of the key points that has to be considered. When the electric field strength exceeds a threshold value, the tangential electric field over the solid/liquid interface initiates surface discharges (Yi, 2012). These discharges are commonly called creeping discharges. This discharge can cause irrecoverable damage to the insulation system and cause flashover which undermines the use of a composite insulation system (Yi, 2012).

A lot of work has recently been devoted to improving the dielectric properties of insulating materials by adding specific nano-fillers to polymer-based materials, changing the dielectric and mechanical properties. Nano-

* Corresponding author (ediriweera.sampath@gmail.com;  <https://orcid.org/0000-0002-5937-7932>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

composite polymer materials show better performance when it comes to the dielectric properties of the materials (Yanashima *et al.*, 2017; Hornak *et al.*, 2020). In addition to the electrical characteristics, they also show higher tensile strength, thermal stability, and improved chemical properties over conventional insulating materials (Liang & Wong, 2017).

Petroleum-based mineral oil has been used as the dielectric liquid in high voltage equipment for more than seven decades due to its better performance. However, due to some of its disadvantages such as low fire point, limited biodegradability, and high toxicity, attention has been paid to use alternative vegetable oils such as soya bean oil, coconut oil (Abey Bandara *et al.*, 2001; Lucas *et al.*, 2002), and sunflower oil, even though their properties such as viscosity, conductivity, acidity, and water solubility are still to be improved. Coconut oil is produced on a commercial scale in Sri Lanka and it is rich in saturated fat with medium and short chains. Most of its physical, chemical, and dielectric properties are at satisfactory levels as an alternative for the conventional mineral oil for high voltage equipment (Abey Bandara *et al.*, 2001; Matharage *et al.*, 2013), and it offers better oxidation stability compared to other vegetable oil due to the higher percentage of saturated fatty acids. However, the studies have not progressed to the analysis of creeping discharge over nanocomposite insulating materials immersed in coconut oil, which can affect the performance of the nanocomposite insulation system. Fractal analysis of creeping discharges can be used to compare the amount of discharge branch formation on the advanced solid dielectric materials with that of conventional materials.

The main objective of this study was to perform a fractal analysis of creeping discharges propagating over nanocomposite polymer insulating materials immersed in coconut oil. In particular, the discharge length and the fractal dimension are determined using a test apparatus based on a point-plane electrode arrangement. An experimental setup based on a point-plane electrode system is used to initiate the discharges in a laboratory environment. Pure epoxy and nanocomposite samples are synthesized for testing purposes and a statistical method is used to analyse the breakdown of prepared solid material samples.

MATERIALS AND METHODS

Experimental test setup

In the current study, a point-plane electrode system based test apparatus was adopted in order to initiate creeping

discharges and compare their propagating characteristics on solid/liquid interfaces. Needle plane electrode systems are usually selected to experimentally investigate the creeping discharge development (Kebbabi & Beroual, 2006; Dang *et al.*, 2012). The tangential field component of the divergent electric field drives the propagation of creeping discharge as observed inside actual high voltage apparatus (Ediriweera *et al.*, 2018a).

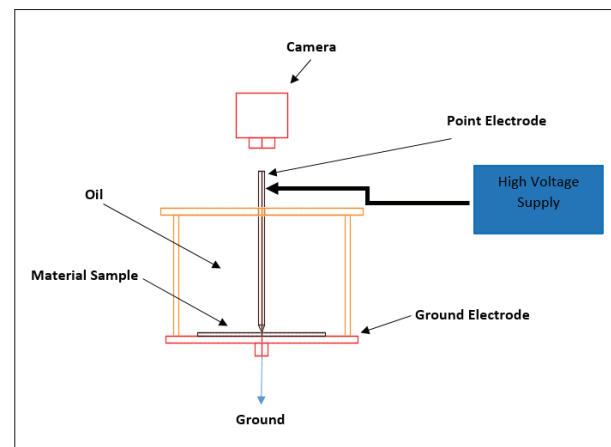


Figure 1: Schematic of the test setup

Figure 1 shows the schematic view of the test cell and this is similar to what other researchers adopted in their studies (Kebbabi & Beroual, 2006). The test cell consists of a cylindrical core with a height of 100 mm and 130 mm inner diameter, made of clear acrylic to visualize the discharge, and a square-shaped brass plate as its lower cover. The test cell is mounted on a structure made of wood. The test cell has a point-plane electrode arrangement, where the vertical point electrode is held using a movable bearing system to permit vertical movement. The point electrode, made of tungsten, has a diameter of 1 mm and the tip of the point is hemispherical so that the field between the two electrodes would be divergent (Ediriweera *et al.*, 2018a). The gap between the two electrodes can be varied from zero to 20 mm. Thin flat insulating material samples can be maintained horizontally between the two electrodes and the point electrode is configured in such a way that the tip of the point electrode nearly touches the solid material sample. The samples are immersed in an insulating liquid by filling the test cell with oil. A CCD camera, connected to a high-performance video card is mounted over the test cell and the integrated images taken by it are used for the optical observation of the discharge. The plane electrode is grounded and the point electrode is energized by a partial discharge (PD) free HVAC test transformer.

Sample preparation

There are various types of epoxy resins that are used for many engineering applications. Bisphenol - A type epoxy resins do not fall in to the category of cycloaliphatic epoxy resins. The epoxy used consists of diglycidyl ether of bisphenol-A (DGEBA) type Epon@Resin 828. The curing agent is Modified Cyclo-aliphatic Amine Adduct type 3388. The chemical structure of the epoxy resin is shown in Figure. 2. The considered epoxy type has a lower viscosity compared to other types and is commonly used in the high voltage industry. SiO_2 with a particle size of 20 nm is chosen as the filler material of the epoxy-based nano-composite. Both pure epoxy and nano-composite material samples are synthesized and a block diagram representing the synthetic procedure for the nano-composite material is shown in Figure 3 (Kochetov, 2012).

Dielectric samples having silicon dioxide nanoparticle compositions (weight to weight) of 1 % are prepared in

this study for testing purposes. The synthetic procedure is similar to what other researchers adopted in their studies (Kochetov, 2012; Karunaratna *et al.*, 2019). As the first step, required volumes of epoxy and curing agents are measured using a pipette and measuring cylinders. Sonication is a process used to disperse the agglomerate filler particles in an ethanol medium. The mechanical mixing process mixes the epoxy and the filler medium making the fillers transfer into the epoxy. The evaporation process removes the ethanol remaining in the mixture. Finally, in the mechanical process, it mixes the medium after the curing agent is added to the mixture. The fabrication process of pure epoxy consists of simple mechanical mixing of the measured volumes of epoxy and curing agents.

Breakdown voltage of solid samples

The chemical and physical properties of dielectric insulators can affect their breakdown voltage, and breakdown voltage becomes a statistically distributed

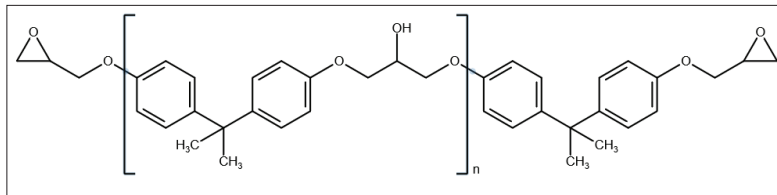


Figure 2: Chemical structure of bisphenol-A type epoxy resin

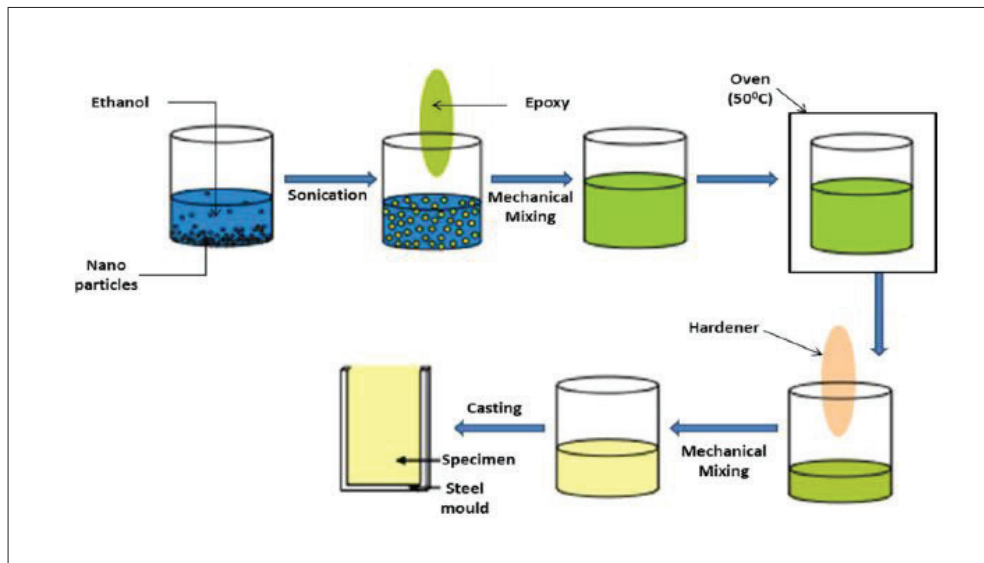


Figure 3: Fabrication process of nano-composite material

quantity. Therefore, statistics are being considered to analyse the breakdown voltage (Gupta & Das, 2017). In the current study, a two-parameter Weibull distribution was used to model and analyse the distribution of the breakdown voltage of the prepared solid insulating samples. The probability density function (PDF) and the cumulative distribution functions are given by equation 1 and equation 2, respectively (Gupta & Das, 2017; Yuan *et al.*, 2017).

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} \exp \left\{ -\left(\frac{x}{\alpha}\right)^{\beta} \right\} \quad \dots(1)$$

$$f(x) = 1 - \exp \left\{ -\left(\frac{x}{\alpha}\right)^{\beta} \right\} \quad \dots(2)$$

where β and α are the shape and scale parameters of the experimental data respectively.

The shape parameter is a measure of the range of experimental data and the scale parameter is a characteristic breakdown voltage at which the probability of failure is 63 %.

Skewness can be used to identify the symmetric or asymmetric nature of breakdown voltage values using equation 3 (Gupta & Das, 2017).

$$skw = \frac{E(x-\mu)^3}{\sigma^3} \quad \dots(3)$$

Experimental procedure

Flat square shaped solid materials of prepared pure epoxy and silica/epoxy nano-composite of side length 9 cm and thickness 3 mm were used to analyse the creeping discharge propagation over the solid/liquid interface. A type of commercially available copra coconut oil with a relative permittivity of 2.9 was selected as the insulating oil in the study, as it is an emerging alternative liquid insulating material (Matharage *et al.*, 2013). The electric field in the experimental setup used in this study was not uniform. In actual power system assets, a uniform electric field cannot be expected due to the electrodes and dielectric configurations in the equipment. Therefore, tests were carried out to make a comparison between propagation patterns on different interfaces that were subjected to the same electrode arrangement. According to the IEC 60156 standard, a test cell with 2 spherical electrodes separated by a gap of 2.5 mm was used to

measure the breakdown voltage of coconut oil samples at the University of Moratuwa, Sri Lanka. The mean value of the breakdown voltage given for five oil samples under ambient conditions was considered as the average breakdown voltage. The average breakdown voltage of 'as received' oil samples was measured at 20 kV.

The test cell was filled with coconut oil and solid materials prepared in square shape were inserted between the two electrodes, such that the point electrode touches the surface of the sample. If the flashover is observed, the test cell was refilled with new oil from the same test oil sample. The samples were kept in the laboratory in a darkened place without any UV exposure in order to avoid contamination. Therefore, the material and electrical properties of the test volumes were assumed to be the same in all the samples in the experiments. The height of the oil volume was kept at 5 cm throughout the testing procedure.

Dielectric strength tests were performed according to the ASTM-D149 standard at the University of Moratuwa, Sri Lanka, and five samples of epoxy and nanocomposite were tested. Parameters of Weibull distribution were estimated using Minitab software.

Fractal dimension

Dimension identifies the way the object fills the space. Dimensions of a point, a line, a square, and a cube are 0, 1, 2, and 3 respectively. Such dimensions are called the Euclidian dimensions. Likewise, fractal patterns have fractal dimensions. As a parameter of characterization of a complex degree of natural phenomena, a fractal dimension is used to describe the surface topography of a fracture and to give a quantitative value to each pattern (Zhou *et al.*, 2014). Several methods can be adapted to calculate the fractal dimension of creeping discharges (Beroual *et al.*, 2014; Ediriweera *et al.*, 2018a). In the current study, the box-counting method was used. Researchers have used this method for purposes of analysis of surface discharges in their findings (Kebbabi & Beroual, 2005). A greater fractal dimension means a more tortuous fracture surface.

One of the common methods called box-counting had been adopted to calculate the dimension of fractal patterns in our previous studies (Kebbabi & Beroual, 2005; Ediriweera *et al.*, 2019). The same method was used in the current study and basic steps were recalled. The original image shown in Figure 4a, obtained in a creeping discharge, was converted into a binary image as shown in Figure 4b. Before the conversion process, all the important traces of the discharge were extracted

using photo editing software. The camera did not depict the entire discharge, since the needle was in the centre of the image. However, the pattern in the middle has

the least amount of unclear information due to the point electrode, when compared to the visible area as the pattern was recorded to its full extent.

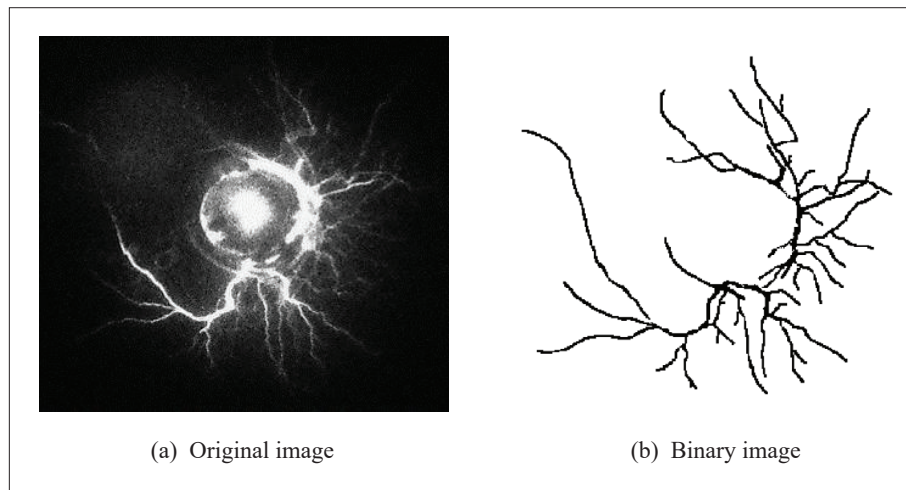


Figure 4: Application of the box-counting method

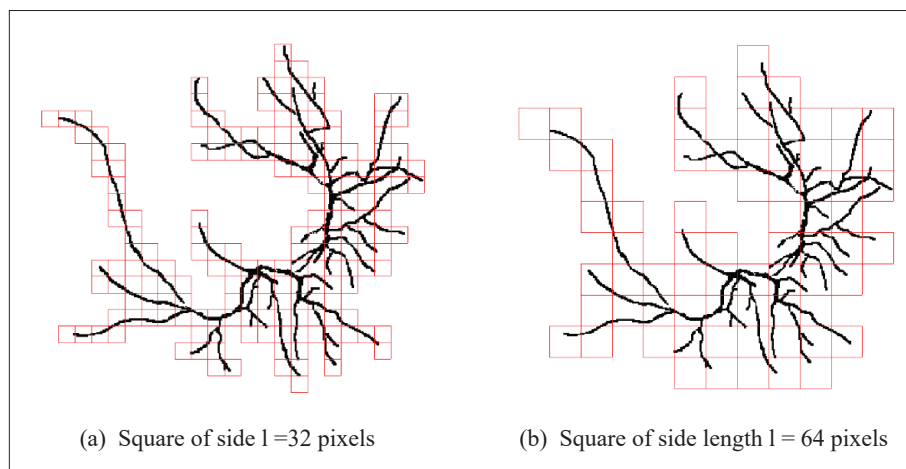


Figure 5: Application of the box-counting method

However, with the glow around the tip of the point electrode and thick streamers in the centre, individual streamers cannot be distinguished because of the noise captured by the camera. So, information in the middle was ignored in any case, and the box-counting algorithm is applied considering that the morphological variation of the pattern has the same self-similarity throughout the pattern. Fractals reflect the self-similarity of the morphological variation of the pattern. Therefore self-similarity can be regarded as implying that at all scales or in all areas of the pattern, it should show the same relative

variation of morphology or the fractal dimension. Then a code was applied to the binary image which generates a grid of square boxes over the image.

While changing the side length of the boxes at each step, the number of boxes having any parts of the discharge pattern was counted. Figures 5a and 5b show the binary images covered by squares of side $l = 32$ pixels and $l = 64$ pixels, respectively. The slope of the log-log plot of the number of boxes vs. box size gives the fractal dimension as shown in Figure 6.

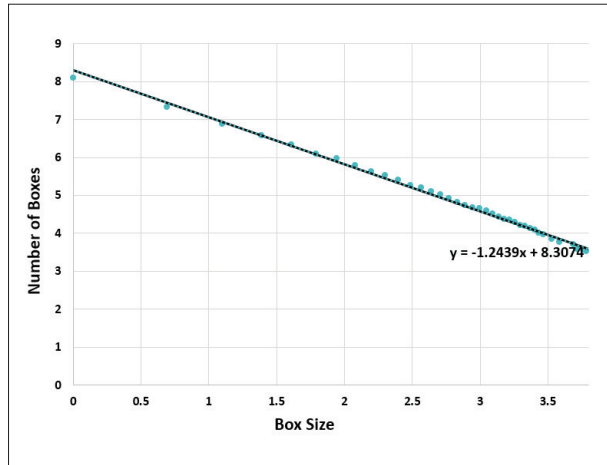


Figure 6: Total number of boxes N versus the side length l of the boxes obtained from the analysis of an example of a discharge

Discharge length

When it comes to creeping discharges, the discharge may have several mainstreamers starting from its starting point. There also can be side branches starting from mainstreamers. Therefore, each streamer may have different total discharge lengths. What is defined as the final discharge length is the maximum discharge length out of all the streamers of the discharge pattern, and the voltage required to get a streamer of a particular length is assessed. In the study, the camera setup provides the images of the propagating pattern continuously and an algorithm was used to identify the image with maximum extension to measure the final discharge length. It should be noted that each discharge length value presented in the study is an average value measurement taken from four solid dielectric samples

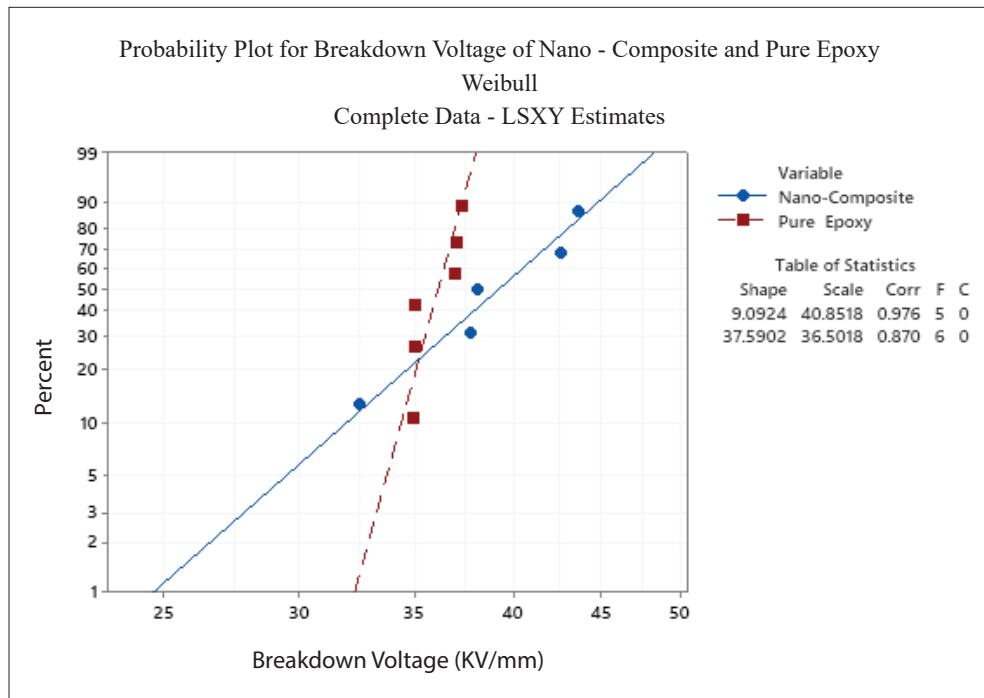


Figure 7: Probability plot for breakdown strength of nano - composite and pure epoxy

Table 1: Parameters of Weibull distribution

Material	Shape parameter	Scale parameter (kV/mm)	Mean (kV/mm)	Skewness
Pure epoxy	40.061	36.551	36.046	0.03
Nano-composite	12.034	40.667	38.975	-0.53

RESULTS AND DISCUSSION

The maximum likelihood estimator is used to estimate the parameters of the distribution of breakdown voltages. Figure 7 shows the probability plot for the breakdown voltages of the material samples.

The shape, scale, mean, and skewness of the distribution are listed in Table 1. The shape parameters are 40 and 12 for pure epoxy and nanocomposite respectively. The shape parameter gives an idea about the failure. A shape parameter lower than one means early failure and when it equal one, then it represents a constant failure rate after commissioning (Gupta & Das, 2017). The skewness of the nanocomposite is a little bit higher than that of epoxy material and the asymmetrical nature of the data points can be seen with non-zero screw values for both of the material samples.

The mean breakdown voltage of epoxy materials has been increased slightly by nano-particles as shown in Table 1. It accounts for around a 7% increment. The behavior of nano-fillers is somewhat different even though microparticles decrease the breakdown strength, acting as centres of charge concentration. The interfacial area of nano-particles is considerably higher and their interparticle distance is quite small. Therefore, they act as barriers to the flow of current through the composite increasing the breakdown strength.

Figure 8 shows the development of discharge patterns propagating over epoxy samples and Figure 9 on nano-composite samples. The basic characteristics of the propagating patterns are similar to the studies with patterns propagating on the solid/mineral oil interface (Kebbabi & Beroual, 2006). Discharge patterns propagating on the interfaces have curved streamers as a result of the electro-hydrodynamic motion of oil in the vicinity of the point electrode and the electric field distorted by space charges on the insulators under AC voltage.

Therefore, streamers try to follow these contours. Observation shows that the discharges propagating on epoxy/coconut oil and nano-composite/coconut oil interfaces have similar tortuous behavior, and, however, with longer discharges on a nano-composite/glass interface. The kind of the solid material can have a significant effect on the amount of surface discharge.

The most likely reasons for the promotion effect on discharges of the solid insulation materials considered would be the influence of hetero-charges, the permittivity of the material (Yi & Wang, 2015), and nanoparticles (Jang, 2017). Once the surface of a solid material is charged with space charges, it takes some time for charges to dissipate. The dissipation time depends on the characteristics of the material such as surface

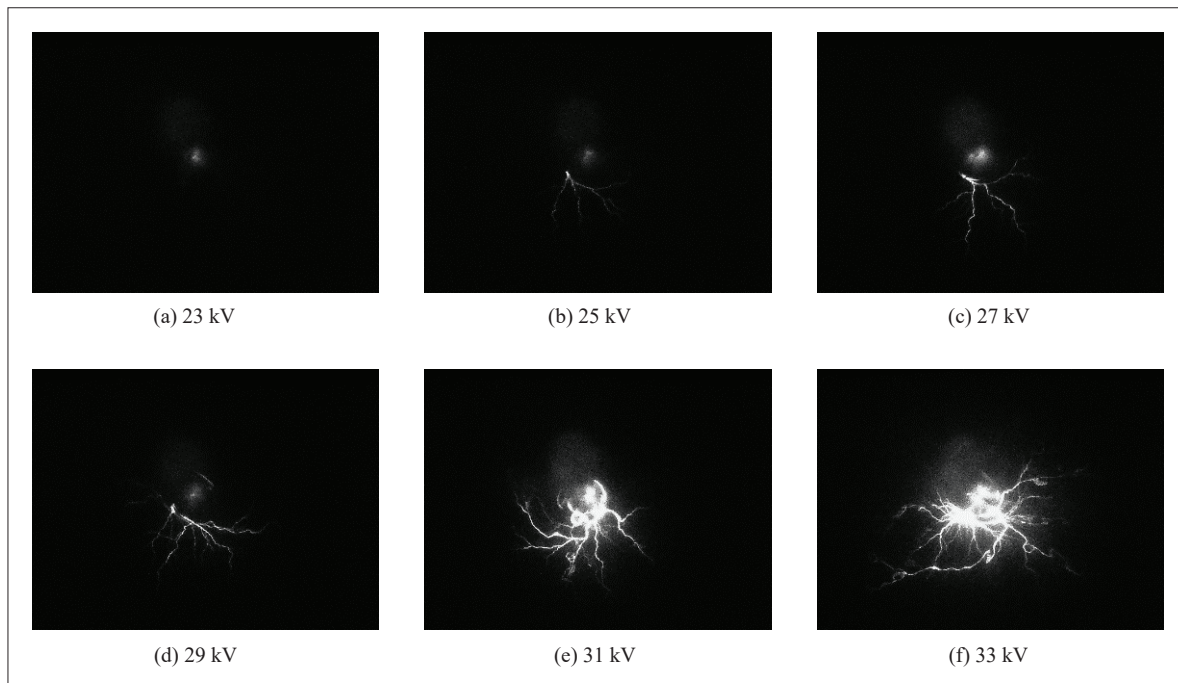


Figure 8: Stages of creeping discharge development on an epoxy surface

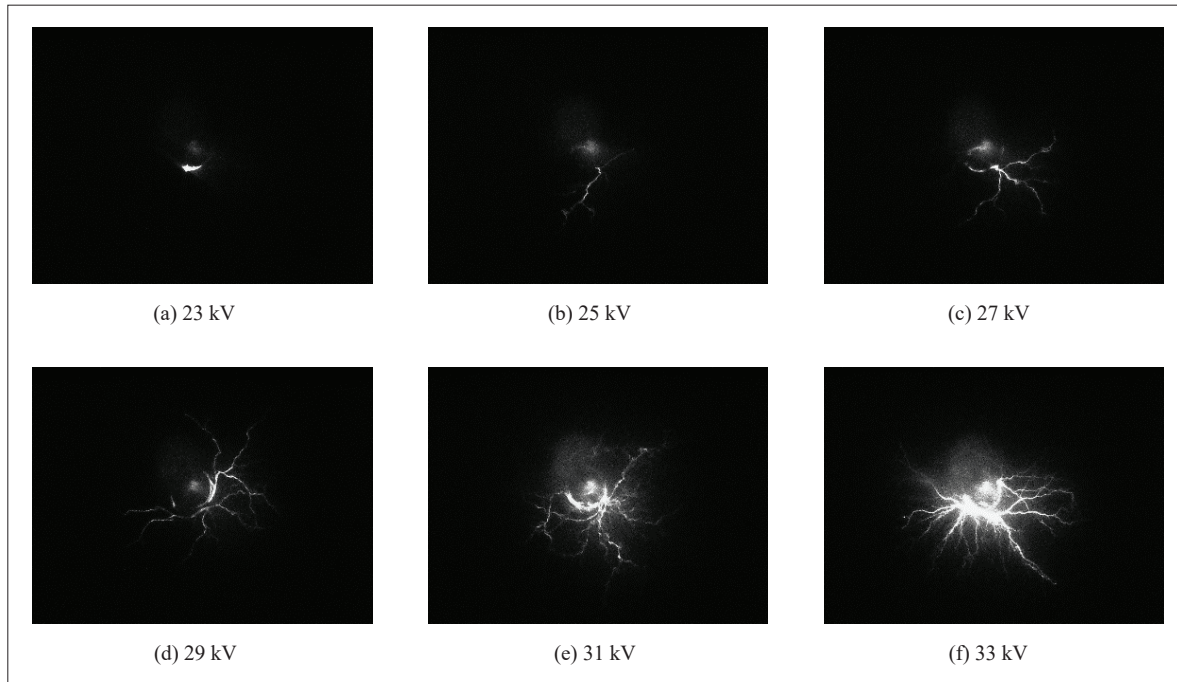


Figure 9: Stages of creeping discharge development on a nano-composite surface.

conductivity. The effect of both homo-space charges and hetero-space charges, which promote the development of subsequent discharge, should be considered in this scenario, because deposited space charges can last until the negative cycle of the supply voltage (Yi, 2012). The capacitance of the solid insulation affects the discharge inception voltage (Jang *et al.*, 2016). When a flat solid material sample is modeled as a capacitor, its value depends on the permittivity of the material. The increase in the capacitance of the solid material sample enhances the capacitive charges. Also, the permittivity mismatch between the liquid and solid affects the electric field at the tip of the point electrode, which has a direct impact on the ionization probability due to the added effect of change in total polarizability and interfacial polarization, which in turn will affect the propagation of the creeping discharge. The enhanced electric field and charges make the discharge propagate more on the surface of the material. Therefore, as the relative permittivity of the solid material increases, it increases the tendency of creeping discharges to propagate. The dielectric constants of pure epoxy and nano-composite samples are 4.89 and 5.25, respectively. The parallel plate method is used to measure the relative permittivity of the material samples at 100 Hz using a LCR meter, as in previous studies (Mandric *et al.*, 2018; Karunarathna *et al.*, 2019). Therefore, the promotional effect on the nano-composite/glass interface can be expected.

The final discharge length holds the quasi-linearity with the voltage similar to the previous experiments with mineral oil, as shown in Figure 10. It increases with the dielectric constant of the solid dielectric material. For a given value of the voltage, discharge lengths are higher with nano-composite than with pure epoxy samples. The capacitive effect is more significant over the pattern propagation on the nano-composite material samples immersed in coconut oil.

Nanoparticles have increased the final discharge length corresponding to the nano-composite samples by 4% to 17%, while having a 7.3 % increment in the dielectric constant. So there is a positive correlation between the dielectric constant and corresponding discharge length curves, with coconut oil being the dielectric liquid. Therefore, more damage on the nano-epoxy solid insulator can be expected due to the traces of the discharges.

Table 2: Fractal dimension of creeping discharges

	Epoxy	Nano-composite
Fractal dimension	1.26	1.26

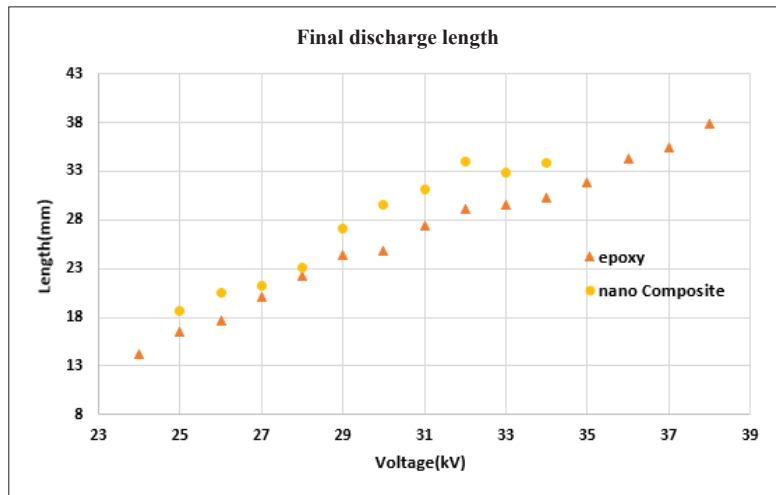


Figure 10: Variation of the final discharge length of the creeping discharges

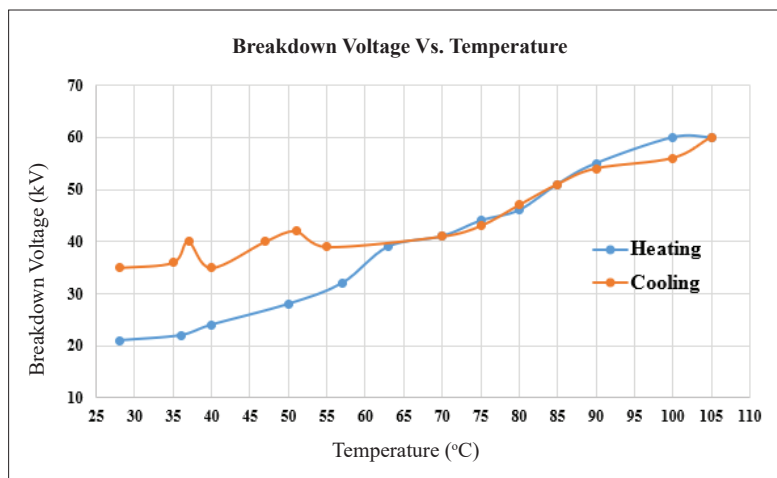


Figure 11: Variation of the breakdown voltage of coconut oil with the temperature.

Fractal dimension is more related to the amount of the divergence of the streamers and branches of a discharge pattern into smaller ones. Previous studies have shown that when the relative permittivity increases, dimension increases (Beroual *et al.*, 2014). However, according to Table. 2 it can be seen that the nanoparticles considered do not have a significant effect on the amount of ramification on the epoxy/coconut oil interface, as the fractal dimension of patterns propagating on nanocomposite and epoxy material samples have the same values. Therefore, the same amount of tortuousness is found in both cases.

According to previous studies (Abey Bandara *et al.*, 2001; Matharage *et al.*, 2013), moisture absorption can reduce the breakdown characteristics of coconut oil.

Therefore, a domestic heater was used to heat the oil samples, and at different temperature values, samples are taken and the dielectric breakdown voltage is measured. After the oil is heated up to 100 °C, it was then allowed to cool down back to room temperature. While the sample was cooling, its dielectric breakdown voltage was measured at several intermediate temperatures. An improvement of dielectric strength was observed after heating the oil sample to over 100 °C.

Figure 11 shows that the breakdown voltage increases gradually up to 60kV when the oil is heated. The water content in the sample affects the dielectric strength of the oil. Heating evaporates water in the oil sample. When the temperature of the sample is near to the boiling

point of water (100°C) the water content is very low, hence it shows a higher dielectric strength. When the oil sample is kept open to the air and allowed to cool back, a reduction in dielectric strength was observed as a result of moisture absorption as shown in Figure 11. Therefore, poor dielectric properties can affect the discharge propagation and hence the accuracy of the results. Therefore, the condition of the dielectric liquid on the creeping discharge propagation is an under-researched topic that requires further investigation.

CONCLUSION

This paper presented a study of creeping discharge propagation over epoxy/coconut oil and nanocomposite/coconut oil insulation interfaces. Experiments were carried out with a point plane electrode arrangement and an image recorder in order to investigate creeping discharges along surfaces of pure epoxy and nanocomposite epoxy samples under an AC divergent field. Results on discharge length, breakdown strength, and fractal dimension have shown that the propagation of creeping discharges depends on the properties of the solid insulator such as dielectric constant and properties of fillers.

Silicon dioxide nanoparticles with 1% concentration were used to increase the dielectric properties of the epoxy material. According to the experimental results, it was evident that it increases the dielectric breakdown voltage of epoxy by 7%. The relative permittivity of the nanocomposite was higher than that of pure epoxy. Therefore, further studies should be carried out with different filler sizes to increase the dielectric properties, especially creeping discharge performance significantly. It was shown that there is a positive correlation between the dielectric constant and corresponding discharge length curves, with coconut oil as the dielectric liquid. Final discharge length increases quasi-linearly with the applied voltage even on materials immersed in coconut oil. According to the discharge length results corresponding to the nano-composite material, it can be concluded that even though nanoparticles can increase the dielectric strength of the epoxy insulators, the capacitive effect is more significant in discharge propagation, and final discharge length has been increased by around 12%. However, the considered nanoparticle concentration of 1% does not have a significant effect on the amount of ramification of the creeping discharges, because the fractal dimension of 1.26 was observed in creeping discharges propagating on both nanocomposite and epoxy material samples. Overall, it can be concluded that the concentration and the particle size of silicon dioxide

can affect the nano dielectric properties of nano-epoxy composite, such as an increase in breakdown strength and dielectric constant by 7%. A nanocomposite with high breakdown strength is not always satisfactory when it is immersed in coconut oil as the electric field at the tip of the point electrode depends on the permittivity mismatch. The research had shown the fabricated material samples enhance the breakdown strength by 7% while increasing the possibility of damage to the solid insulator due to creeping discharges by 4% to 17%. Therefore, the materials should be fabricated to minimize the creeping discharge propagation while keeping the dielectric properties within an acceptable level.

Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- Abeyesundara D., Weerakoon C. Lucas R., Gunatunga K. & Obadage K. (2001). Coconut oil as an alternative to transformer oil. *ERU - Proceedings of the 7th Symposium on Research for Industry, Moratuwa*, November 2001, pp. A12–A22.
- Beroual A., Coulibaly M.L., Girodet A. & Aitken O. (2014). Relationship between the fractal dimension of creeping discharges propagating at solid/gas interfaces and the characteristics parameters of interfaces. *International Review of Electrical Engineering* 9(2): 460–465.
- Dang V.H., Beroual A., Coulibaly M.L. & Perrier C. (2012). Investigation on creeping discharges propagating over pressboard immersed in mineral and vegetable oils submitted to ac and dc voltages. *2012 International Conference on High Voltage Engineering and Application*, pp. 215–218.
DOI: <https://doi.org/10.1109/ICHVE.2012.6357024>
- Ediriweera W., Jayarathna K.P.B., Lucas J. & Samarasinghe R. (2018a). Effect of the shape of the insulator on fractal characteristics of creeping discharges. *2018 Moratuwa Engineering Research Conference (MERCCon)*, pp. 506–510.
DOI: <https://doi.org/10.1109/MERCCon.2018.8421974>
- Ediriweera W., Jayarathna K., Samarasinghe R. & Lucas J. (2018b). Effect of barrier and its shape over the breakdown voltage of oil gaps. *2018 2nd International Conference on Electrical Engineering (EECon)*, pp. 76–80.
DOI: [10.1109/EECon.2018.8540987](https://doi.org/10.1109/EECon.2018.8540987)
- Ediriweera S., Jayarathna P., Lucas R. & Samarasinghe R. (2019). Fractal analysis of creeping discharges propagating over solid/liquid interfaces- influence of oil level. *2019 9th International Conference on Power and Energy Systems (ICPES)*, pp. 1–6.

- DOI: <https://doi.org/10.1109/ICPES47639.2019.9105611>
 Gupta H. & Das S. (2017). Statistical analysis of oil insulation breakdown voltage. *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pp. 2044–2048.
 DOI: <https://doi.org/10.1109/IEEM.2017.8290251>
- Hornak J., Kubes T. & Trnka P. (2020). Effect of nanometric oxides on dielectric and mechanical properties of epoxy resin. *2020 International Conference on Diagnostics in Electrical Engineering (Diagnostika)*, pp. 1–4.
 DOI: <https://doi.org/10.1109/Diagnostika49114.2020.921716>
- Jang K. (2017). Creepage discharge behaviour and numerical model analysis at the oil/pressboard interface with the effect of nano composite coating. *PhD thesis*, Kyushu Institute of Technology, Japan.
- Jang K., Akahoshi T., Kozako M. & Hikita M. (2016). Nano SiO_2 /epoxy coating effect on creepage discharge characteristics in oil/pressboard composite insulation system. *2016 IEEE International Conference on Dielectrics (ICD)*, volume. 1, pp. 394–397.
 DOI: <https://doi.org/10.1109/ICD.2016.7547626>
- Karunaratna P., Chithradewa K., Kumara S., Weerasekara C., Sanarasinghe R. & Rathnayake T. (2019). Study on dielectric properties of epoxy resin nanocomposites. *2019 International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*, pp. 1–5.
 DOI: <https://doi.org/10.1109/ISAECT47714.2019.9069694>
- Kebbabi L. & Beroual A. (2005). Fractal analysis of creeping discharge patterns propagating at solid/liquid interfaces-influence of the nature and geometry of solid insulators. *CEIDP '05. 2005 Annual Report Conference on Electrical Insulation and Dielectric Phenomena*, pp. 132–135.
 DOI: <https://doi.org/10.1109/CEIDP.2005.1560638>
- Kebbabi L. & Beroual A. (2006). Optical and electrical characterization of creeping discharges over solid/liquid interfaces under lightning impulse voltage. *IEEE Transactions on Dielectrics and Electrical Insulation* **13** (3): 565–571.
 DOI: <https://doi.org/10.1109/TDEI.2006.1657969>
- Kochetov R. (2012). Thermal and electrical properties of nanocomposites, including material properties (Unpublished doctoral dissertation), Lappeenranta University of Technology, Finland.
- Liang M. & Wong K. (2017). Improving the long-term performance of composite insulators use nanocomposite: A review. *Energy Procedia* **110**: 168–173.
 DOI: <https://doi.org/10.1016/j.egypro.2017.03.123>
- Liu Y., Sasamoto R., Matsumoto T., Izawa Y. & Nishijima K. (2016). Influence of barrier thickness on discharge behavior in air gap with gfrp insulator under impulse voltage stress. *2016 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP)*, pp. 971–974
 DOI: <https://doi.org/10.1109/CEIDP.2016.7785497>
- Lucas R., Abeysundara D., Weerakoon C., Perera M., Obadage K. & Gunatunga K. (2002). Coconut oil insulated distribution transformer. *Transactions of the IEE Sri Lanka*, volume 4, no. 2, p 1–5.
- Mandric V., Rupcic S., Srnovi'c M. & Ben'si'c G. (2018). Measuring the dielectric constant of paper using a parallel plate capacitor. *International Journal of Electrical and Computer Engineering Systems* **9**: 1–10.
 DOI: <https://doi.org/10.32985/ijeces.9.1.1>
- Matharage B.S.H.M.S.Y., Fernando M.A.R.M., Bandara M.A.A.P., Jayantha G A. & Kalpage C.S. (2013). Performance of coconut oil as an alternative transformer liquid insulation. *IEEE Transactions on Dielectrics and Electrical Insulation* **20** (3): 887–898.
 DOI: <https://doi.org/10.1109/TDEI.2013.6518958>
- Yanashima R., Hirai N. & Ohki Y. (2017). Effects of addition of mgo fillers with various sizes and coaddition of nano-sized SiO_2 fillers on the dielectric properties of epoxy resin. *2017 International Symposium on Electrical Insulating Materials (ISEIM)*, volume. 2, pp. 650–653.
 DOI: <https://doi.org/10.23919/ISEIM.2017.8166574>
- Yi X. (2012). Characteristics of creepage discharges along ester-pressboard interfaces under ac stress. *PhD thesis*, University of Manchester, USA.
- Yi X. & Wang Z. (2015). The influences of solid surface on the propagation of creepage discharge in insulating liquids. *IEEE Transactions on Dielectrics and Electrical Insulation* **22** (1): 303–312.
 DOI: <https://doi.org/10.1109/TDEI.2014.004195>
- Yuan W., Wang, T., Ni H., Gao M., Ding Y., Li Y., Zhao Y. & Zhang Q. (2017). Weibull statistical analysis of size effects on the impulse breakdown strength in transformer oil. *2017 IEEE 19th International Conference on Dielectric Liquids (ICDL)* pp. 1–4.
 DOI: <https://doi.org/10.1109/ICDL.2017.8124618>
- Zhou H.W., Xue D.J. & Jiang D.Y. (2014). On fractal dimension of a fracture surface by volume covering method. *Surface Review and Letters* **21**(01).
 DOI: <https://doi.org/10.1142/S0218625X14500152>

RESEARCH ARTICLE

Meteorology

Forecasting the track and the intensity of the cyclone Burevi using WRF

TD Gamage^{1,2*}, U Sonnadara², S Jayasinghe³ and S Basnayake³

¹ Department of Electrical and Information Engineering, Faculty of Engineering, University of Ruhuna, Hapugala, Galle.

² Department of Physics, Faculty of Science, University of Colombo, Colombo 03.

³ Climate Resilience Department, Asian Disaster Preparedness Center, Thailand.

Submitted: 22 October 2021; Revised: 04 April 2022; Accepted: 04 May 2022

Abstract: The cyclone 'Burevi' was recorded in the North Indian Ocean in the month of December 2020. In this study a numerical weather forecasting model WRF-ARW was used to forecast the track, intensity, and landfall location and time of the cyclone. The forecasting was performed as 16 consecutive model runs. Each forecast was initiated with different model initializations with a 12 hour time interval. For such initializations, the lateral boundary condition was set by the global forecast system data with a horizontal resolution of $0.25^\circ \times 0.25^\circ$ at 3 hour time intervals. The forecasts were conducted up to 120 hours and evaluated using the data obtained from an international best track archive. The forecasted track was evaluated using root mean square error, direct positional error, along-track error and cross-track error. The intensity of the cyclone was evaluated by comparing the minimum sea level pressure and maximum forecasted 10-m wind speed obtained from the positions of the forecasted cyclone track. The WRF-ARW model was successful in forecasting the track and the intensity of the cyclone up to a lead time of 48 hours having a mean direct positional error of less than 106 km. This cyclone recorded three landfalls in Sri Lanka. The accuracy of the forecasted first landfall position and time were within the acceptable range, but the second and third require further investigation.

Keywords: North Indian Ocean, track forecast, tropical cyclone, WRF-ARW.

INTRODUCTION

Weather research and forecasting (WRF) is a numerical weather prediction system used for research and

operational forecasting (Skamarock *et al.*, 2019). Advanced Research WRF (ARW) is a configuration of the WRF system having ARW as the dynamic solver. WRF has been used by a number of researchers in the past to track cyclones (Islam *et al.*, 2015; Kotal & Bhattacharya, 2020; Nadimpalli *et al.*, 2020; Mohanty *et al.*, 2021). Islam *et al.* (2015) used WRF to track the cyclone 'Haiyan' recorded over West Pacific Ocean. They managed to forecast the track of the cyclone prior to 40 hours with a track error less than 100 km. Although they were successful in forecasting the track of the cyclone, the intensity of the cyclone was underestimated. Nadimpalli *et al.* (2020) carried out an assessment of WRF and hurricane weather research and forecasting (HWRF) by forecasting tropical cyclones in the Bay of Bengal. They have shown that the results obtained using WRF and HWRF are comparable up to 36 hours. For longer forecasts beyond 36 hours HWRF has given better results. Hon (2020) has forecasted 13 North Indian Ocean (NIO) cyclones. Mean positional errors for those forecasts are 38 km, 69 km and 107 km for 0, 24 and 48 hour forecast lead times respectively.

The forecast error of rainfall depth due to a cyclone was subjected to experiment using ARW by Osuri *et al.* (2020). They reported higher rainfall error with the long-range (48 - 96 hour) forecast than the short-range (24 hour). The forecast of wind field displacement and speed of the cyclone 'Vardah' formed over the Bay of

* Corresponding author (tharindu@eie.ruh.ac.lk;  <https://orcid.org/0000-0002-9770-7795>)



Bengal is improved by applying the following two methods. The two methods are relocation of the NCEP GFS forecast wind field and modification of the wind speed at each grid point over the cyclone influenced area (Kotal & Bhattacharya, 2020). Severe cyclonic storm ‘Fani’ over the Bay of Bengal was forecasted using WRF and HWRF by Mohanty *et al.* (2021). They concluded that the landfall time and position of the cyclone can be forecasted well with a 60 hour lead time. The tropical cyclone tracking and verification techniques used in weather prediction models by the UK Met Office is described elsewhere (Heming, 2017).

The cyclonic storm ‘Burevi’ was recorded in the Bay of Bengal in December 2020 (Farzan, 2020b). It is the fifth cyclone recorded in the North Indian (NI) basin during year 2020 (Knapp *et al.*, 2018). Three of those five are recorded in the Bay of Bengal subbasin. Those cyclones were named ‘Amphan,’ ‘Nivar,’ and ‘Burevi’ (Knapp *et al.*, 2018). It is worth noting that all those three cyclones have affected Sri Lanka, and cyclone ‘Burevi’ made a landfall in the east coast (Farzan, 2020a). Figures 1a and 1b are meteorological satellite observation images of cyclone ‘Burevi’ recorded at 0600 Coordinated Universal Time (UTC) on 02 December 2020. Those two images consist of Day Convective Storms RGB and Day Microphysics RGB, issued by the Japan Meteorological Agency (JMA) using their satellites. The yellow areas shown in Figure 1a are the cumulonimbus clouds with strong updrafts. Those clouds produce severe wind, tornadoes and heavy rainfall (Cotton *et al.*, 2011). In Figure 1b the red and brown colour area shows the thick clouds of cyclone ‘Burevi’ with deep precipitation.

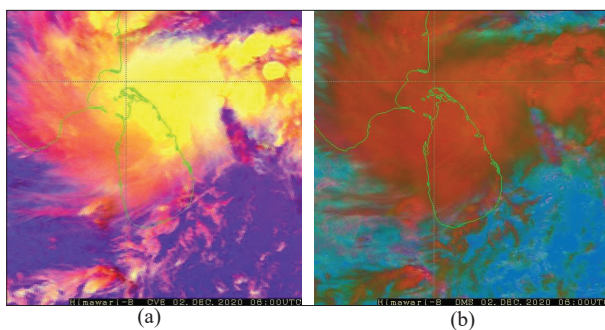


Figure 1: Satellite images captured at 0600 UTC on 02 December 2020

According to the report issued by the Disaster Management Center (DMC) of Sri Lanka there were three deaths due to cyclone ‘Burevi’ (Priyantha, 2020). The number of people affected in the Northern and

Eastern provinces due to the cyclone are 102,098 and 30,830 respectively. In those two provinces, 103 houses were fully damaged and 3,598 were damaged partially. There were 1,054 small and medium scale enterprises were also damaged due to the cyclone (Priyantha, 2020). Prediction of the track and the intensity of the cyclone is important for the disaster management teams to take early mitigation measures to minimize the damage.

The main objective of this work is to forecast the track and the intensity of cyclone ‘Burevi’ that affected Sri Lanka, using WRF. Another objective is to identify the best possible forecast lead time in forecasting the track and the intensity of the cyclone. The final objective of this work is to identify the forecasted landfall position and time of the cyclone.

MATERIALS AND METHODS

In this study the track and the intensity of the cyclone ‘Burevi’ is forecasted using WRF. WRF consist of two dynamical cores namely ‘ARW’ and ‘NMM’. In this study the dynamical core ‘ARW’ was selected. The lateral boundary condition of the selected domain was set by the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) data with a horizontal resolution of $0.25^\circ \times 0.25^\circ$ at 3 h time intervals. The dataset was released four times a day in real-time and in this study 16 such datasets were used having 12 h time intervals, starting from 0000 UTC 27 November 2020 to 1200 UTC 04 December 2020. The domain was selected to cover the entire Bay of Bengal in the NI basin. It was centred over the point 12°N , 83°E having 240 x 180 grid points with 15 km grid spacing, covering a 3600 km x 2700 km area. The terrain height of the selected domain is shown in Figure 2. In the model configuration the number of vertical levels were set to 33. Based on the literature the pressure at the top of the model was set to 5000 Pa for all the forecasts (Baki *et al.*, 2022). However with the support of the dataset used to set the lateral boundary condition, it is possible to use a lower model top pressure (Hon, 2020). Adaptive time step was used to automatically set the time step to a maximum value while keeping the model numerically stable (Hutchinson, 2007; Skamarock *et al.*, 2019).

The physics options used in this study were WRF Single-moment 6-class Scheme (WSM6) for microphysics (Hong & Lim, 2006), New Tiedtke Scheme (Zhang & Wang, 2017) for the cumulus parameterization option, Dudhia shortwave Scheme (Dudhia, 1989) for the shortwave radiation, Rapid and Accurate Radiative

Transfer Model (RRTM) Longwave Scheme (Mlawer *et al.*, 1997) for the longwave radiation and University of Washington Boundary Layer Scheme (Bretherton & Park, 2009) for Planetary Boundary Layer (PBL). Added to those options MM5 Similarity Scheme (Beljaars, 1995) is used as the Surface Layer option and the Unified Noah Land Surface Model (Tewari *et al.*, 2004) is used as the Land Surface Option.

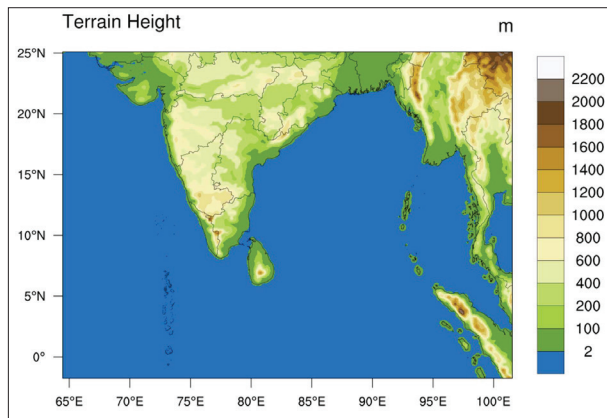


Figure 2: Terrain Height of the selected domain

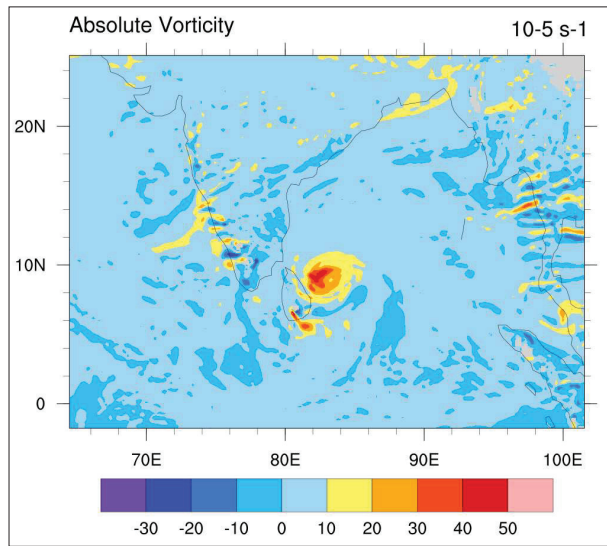


Figure 3: Absolute Vorticity at 0600 UTC 02 December 2020

The forecasting of the cyclone ‘Burevi’ is performed as 16 consecutive model runs. Each forecast is initiated with different model initialization, from 0000 UTC 27 November 2020 to 1200 UTC 04 December 2020, with

a 12 h time interval. The initialization times and the forecast durations are shown in Table 1. The forecast durations vary from 30 to 120 h. The forecast positions of the cyclone are identified in the literature using the forecasted minimum sea level pressure (MSLP) and relative vorticity (RV). The locations of maximum RV at 850 hPa and MSLP are close in a well developed cyclone over the open ocean (Heming, 2017). In this research the local MSLP locations are identified as possible positions of the cyclone. The absolute vorticity (AV) at 850 hPa is considered instead of RV for a search radius of 2° having local MSLP locations as the centres. The combination of local MSLP and the AV at 850 hPa is considered in selecting the forecasted position of the cyclone. The forecasted AV at 0600 UTC on 02 December 2020 is shown in Figure 3. All the forecasts conducted in this study are evaluated using the International Best Track Archive for Climate Stewardship (IBTrACS) (Knapp *et al.*, 2018). In that dataset the track records for the cyclone ‘Burevi’ are available from 1200 UTC 01 December 2020 to 1800 UTC 05 December 2020, and it is shown in Figure 4.

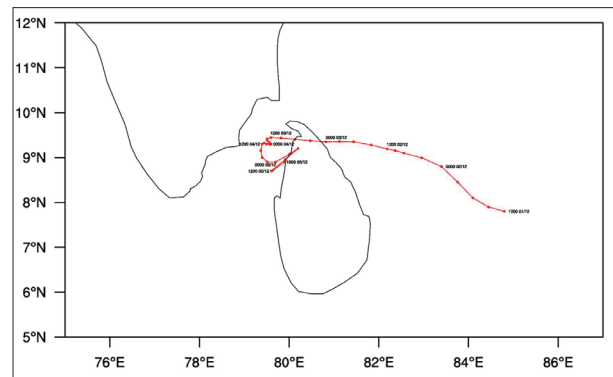


Figure 4: The track of cyclone ‘Burevi’ (Knapp *et al.*, 2018)

The forecast leads vary from 0 to 120 h with a 3 h time interval. Since the track record of the cyclone ‘Burevi’ provided by IBTrACS is from 1200 UTC 01 December 2020 to 1800 UTC 05 December 2020 the number of verifiable forecasting positions for different forecast leads are different. The root mean square error (RMSE) is calculated to evaluate the track of the cyclone ‘Burevi’ using IBTrACS dataset as a reference. The calculation is done according to the Equation 1. In that equation ‘simlat_n’ and ‘simlon_n’ are simulated latitude and longitude coordinates of the cyclone. The values for ‘btlat_n’ and ‘btlon_n’ are the latitude and longitude of the cyclone track recorded in the IBTrACS dataset.

Table 1: RMSE of the selected forecasts with different initialization times

No	Initialization time	Forecast duration (h)	Verifiable positions	Track RMSE (km)
1	00 UTC 27 November 2020	120	5	340.1
2	12 UTC 27 November 2020	120	9	293.9
3	00 UTC 28 November 2020	120	13	127.0
4	12 UTC 28 November 2020	120	17	166.7
5	00 UTC 29 November 2020	120	21	59.0
6	12 UTC 29 November 2020	120	25	164.9
7	00 UTC 30 November 2020	120	29	215.7
8	12 UTC 30 November 2020	120	33	202.8
9	00 UTC 01 December 2020	114	35	146.4
10	12 UTC 01 December 2020	102	35	90.5
11	00 UTC 02 December 2020	90	31	245.1
12	12 UTC 02 December 2020	78	27	84.7
13	00 UTC 03 December 2020	66	23	195.7
14	12 UTC 03 December 2020	54	19	72.3
15	00 UTC 04 December 2020	42	15	87.9
16	12 UTC 04 December 2020	30	11	88.3

The great-circle distance between those two points was calculated using the ‘haversine’ formula (Sinnott, 1984). N is the number of verifiable forecasting positions for each forecast as shown in Table 1.

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (\text{haversine}(\text{simlat}_n, \text{btlat}_n, \text{simlon}_n, \text{btlon}_n))^2}{N}} \quad \dots(1)$$

The verification of the tracks for different forecast leads is done using direct positional error (DPE), along-track error (ATE), and cross-track error (CTE) as mentioned by Heming (2017). The DPE is the great-circle distance between the observed and forecast positions considered at the same time instant in UTC. The DPE does not provide any information about speed and the direction of the forecast (Heming, 2017). The ATE was calculated by extrapolating the line joining observed position and the position observed 3 h earlier and drawing a perpendicular line from the forecasted point to it. The distance between the point of intersection of those two lines and the observed point is considered as the ATE. If the ATE is positive, the forecast is ahead of the observed position in the observed direction of motion having a faster movement, and if the ATE is negative, it is having a slower one. The CTE is the perpendicular distance from forecasted point to the extrapolated line joining observed position and the position observed 3 h

earlier (Heming, 2017). Since the ATE and CTE require one observed position recorded 3 h prior to the observed position, it is not possible to calculate the ATE and CTE for the first observed position. In each of the DPE, ATE, and CTE calculations, the ‘haversine’ formula is used to calculate the great-circle distance.

The forecasted intensity of the cyclone was evaluated by comparing the sea level pressure and 10-m wind speed obtained in the positions of the forecasted track with the corresponding pressure and maximum sustained wind speed values recorded in IBTrACS dataset (Knapp *et al.*, 2018). The forecasted positions have the MSLP of the considered domain. The forecasted position of the cyclone ‘Burevi’ at 0600 UTC on 02 December 2020, having an MSLP of 1001.27 hPa, was rounded to 1001 hPa and marked in Figure 5. The absolute MSLP error was calculated by using such forecasted MSLP values and using the pressure values recorded in the IBTrACS dataset as a reference.

The wind speed forecasted at 0600 UTC 02 December 2020 of the cyclone ‘Burevi’ is shown in Figure 6. A search radius of 2° is considered from the point of MSLP to obtain the maximum forecasted 10-m wind speed as used by Heming (2017). The absolute wind speed error of the forecasted maximum wind speed was calculated by using the maximum sustained wind speed of IBTrACS dataset (Knapp *et al.*, 2018) as a reference.

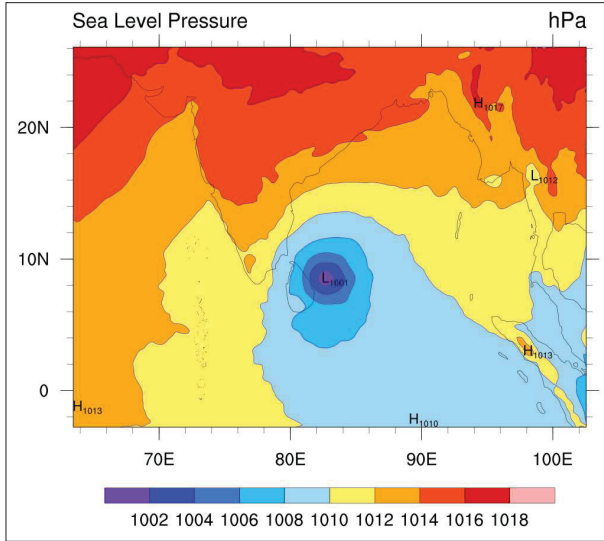


Figure 5: Forecasted sea level pressure at 0600 UTC on 02 December 2020.

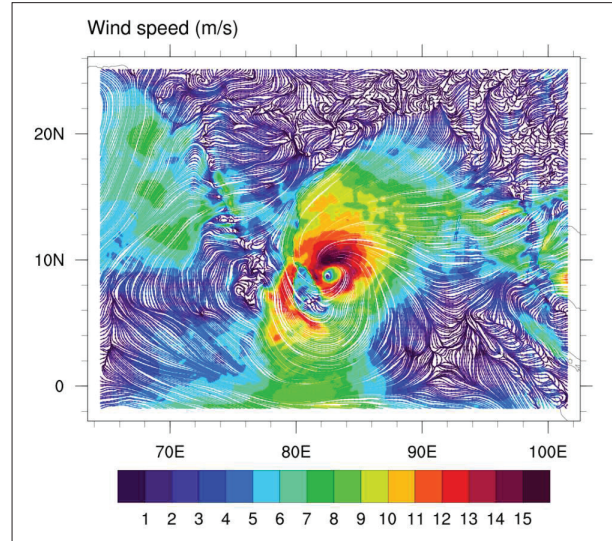


Figure 6: Forecasted 10-m wind speed at 0600 UTC on 02 December 2020.

The cyclone ‘Burevi’ produced three landfall positions. Those positions were calculated based on the observed cyclone positions recorded in IBTrACS dataset (Knapp *et al.*, 2018). The latitude and longitude coordinates of such points before and after the landfall are used to find the equation of the straight line joining those two. Once the coefficients of the equation are obtained, it is possible to find all the latitude and longitude coordinates between the two points. The terrain elevation of all those points were obtained using the calculated latitude and longitude coordinates. The difference in terrain elevation between a location on sea and a location on land is used to identify the latitude and longitude coordinate of the observed landfall location. The same process described above was applied to the forecasted cyclone positions to obtain the forecasted landfall location. Landfall error was obtained by calculating the great-circle distance between the observed and forecasted landfall positions using the ‘haversine’ formula. Landfall time was calculated by assuming that the cyclone had moved at a constant speed between the two cyclone positions considered for a period of 3 h. It was calculated according to the distance ratio between the cyclone positions and the landfall location. Landfall time error was calculated by getting the time difference between the forecasted landfall time and the observed one.

RESULTS AND DISCUSSION

The track RMSE of the selected forecasts with 16 different model initializations are shown in Table 1. From the model initialization time 0000 UTC on 27 November 2020 to 0000 UTC on 29 November 2020, the track RMSE was reduced from 340.1 km to 59.0 km. For the model initialized at 1200 UTC on 01 December 2020, the track RMSE was 90.5 km.

In Figure 7 the forecasts of up to 72 hours are plotted in blue colour and the reference provided by IBTrACS dataset (Knapp *et al.*, 2018) is plotted in red colour. The cyclone started at 1200 UTC 01 December 2020 from the location 7.80°N, 84.80°E and moved towards Sri Lanka according to the IBTrACS dataset (Knapp *et al.*, 2018). It is worth noting that all those forecasts shown in Figure 7 are initiated before the starting time of the cyclone. The observed track of the entire cyclone is plotted in red from Figure 7a to Figure 7d as a reference. According to the observed track of the cyclone, it had got trapped between Sri Lanka and India and was weakened at the end (Knapp *et al.*, 2018). It had also made its second and third landfall on Sri Lanka during that final stage. That event is not forecasted and the cyclone points have moved away from Sri Lanka at the final stage as shown in Figure 7.

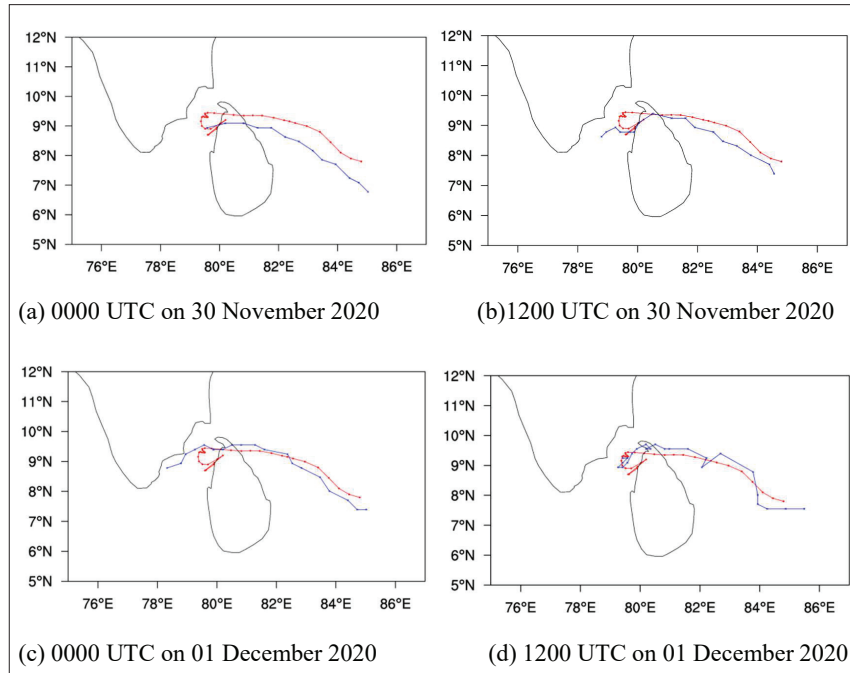


Figure 7: Forecast tracks up to 72 hour duration with different model initialization times. Observed and forecasted tracks are plotted in red and blue colours, respectively.

The DPEs of the forecasted tracks for different forecast leads are verified using the IBTrACS data (Knapp *et al.*, 2018). The mean DPEs for different forecasting times up to 72 hours are shown in Table 2. The mean DPE is less than 61 km up to a 24 hour lead time. It is less than 106 km and 131 km for lead times up to 48 and 72 hours, respectively. These results are comparable with the results obtained by Hon (2020) for their forecasts of 13 cyclones in the NIO. The ATE is calculated for different forecast leads using IBTrACS data (Knapp *et al.*, 2018) as observed positions. The ATE gets positive and negative values depending on the forecast position. The absolute value of ATE is considered to calculate the mean ATE, and mean absolute ATE records up to 72 hours are shown in Table 2. The mean absolute ATE is less than 48 km up to a 24 hour forecast lead time. It is less than 84 km and 122 km for lead times up to 48 and 72 hours respectively.

The CTE is calculated for the forecast leads based on the observed positions recorded in IBTrACS data (Knapp *et al.*, 2018). The CTE is considered positive if the forecast position of the cyclone is recorded to the right (left) of the extrapolated line joining observed position and the position observed 3 hour earlier in the northern (southern) hemisphere. The absolute value of

CTE is considered to calculate the mean CTE, and mean absolute CTE records up to 72 hours are shown in Table 2. Mean absolute CTE is less than 42 km up to a 24 hour forecast lead time. It is less than 47 km and 64 km for lead times up to 48 and 72 hours respectively. The mean DPE, mean absolute ATE, and mean absolute CTE are further illustrated in Figure 8. According to that figure the mean DPE, mean absolute ATE, and mean absolute CTE shows higher values for 0 and 3 hour forecast lead times compared to a 6 hour forecast lead time. The main reason is that, during that period, the model is in the spin-up phase to develop the waves and the clouds (Skamarock *et al.*, 2019). It is possible to provide a pre-forecast period to the model, which will provide the required time to adjust to the topography, and produce cloud fields (Skamarock *et al.*, 2019). Such a pre-forecast period is not configured in this model setup and will be considered in future work.

The use of AV with MSLP in selecting the forecasted positions of the cyclone has reduced DPE, ATE, and CTE compared to using only MSLP. When the forecasted MSLP is considered for the entire domain, there exist some MSLP points that do not belong to the cyclone. Those low pressure points may belong to another

cyclone that is forecasted in the same simulation domain. If MSLP is the only parameter considered in selecting the cyclone position it would provide an erroneous point resulting in larger DPE, ATE, and CTE. Considering AV with local MSLP points avoids such erroneous points that do not belong to the forecasted track of the cyclone. It is possible to extend this method to forecast the tracks of multiple cyclones of the same simulation domain by using the MSLP and AV.

MSLP error is calculated based on the forecasted sea level pressure along the cyclone positions and the pressure values recorded in IBTrACS data (Knapp *et al.*, 2018).

The mean absolute MSLP error up to a 72 hour lead time is shown in Table 2. The mean absolute MSLP error is less than 2.93 hPa up to a 72 hour forecast lead time. The forecasted MSLP is shown in Figure 9 for the corresponding observed MSLP values of cyclone ‘Burevi’ recorded in IBTrACS data (Knapp *et al.*, 2018). The correlation coefficient was calculated using linear regression based on the observed and forecasted MSLP. The MSLP forecast has a bias of 1.18, an RMSE of 3.07 hPa, and a correlation coefficient of +1.00 compared to the observations. In Figure 9 it is evident that the forecasted MSLP is increased based on the observed MSLP values, with a positive correlation coefficient of 1.00.

Table 2: Mean DPE, mean absolute ATE, mean absolute CTE, mean absolute MSLP error and mean absolute wind speed error for different forecasting lead times.

Forecast lead time (h)	Mean DPE (km)	Mean absolute ATE (km)	Mean absolute CTE (km)	Mean absolute MSLP error (hPa)	Mean absolute wind speed error (ms ⁻¹)
T + 0	58.2	45.4	23.8	1.15	2.71
T + 3	60.2	39.1	41.8	2.03	2.43
T + 6	46.4	27.9	36.1	2.14	1.73
T + 9	47.9	36.8	28.6	1.94	2.76
T + 12	50.4	38.7	25.6	2.85	3.02
T + 15	54.4	47.3	21.4	2.93	3.45
T + 18	54.9	39.6	27.8	2.25	3.27
T + 21	53.0	34.9	32.5	2.39	3.81
T + 24	50.2	38.6	24.5	2.54	3.85
T + 27	53.5	42.4	23.3	2.53	3.32
T + 30	68.0	53.5	33.2	2.11	3.09
T + 33	48.4	40.6	16.5	1.77	3.54
T + 36	62.1	44.0	24.7	1.85	3.92
T + 39	62.4	49.2	29.1	2.36	3.35
T + 42	83.3	60.8	36.8	2.00	3.57
T + 45	90.2	71.1	46.2	1.86	3.50
T + 48	105.7	83.1	42.9	2.20	4.13
T + 51	124.9	117.7	27.4	2.89	4.30
T + 54	130.9	121.7	33.4	2.17	3.65
T + 57	129.8	106.3	57.2	2.27	4.49
T + 60	104.1	92.4	49.6	2.58	4.51
T + 63	125.9	110.8	42.2	2.88	4.24
T + 66	124.2	92.9	63.1	1.89	3.79
T + 69	104.3	82.8	53.5	1.91	4.50
T + 72	128.8	112.9	53.2	2.24	4.55

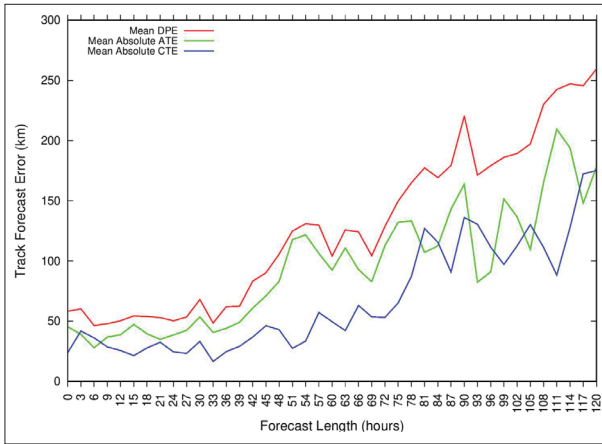


Figure 8: Mean DPE, mean absolute ATE and mean absolute CTE for different forecast lead times of the cyclone ‘Burevi’.

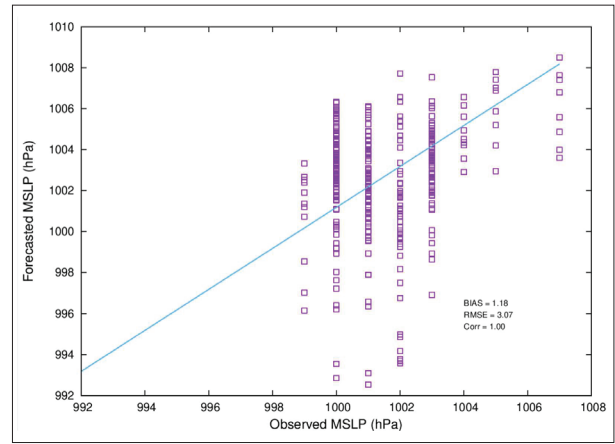


Figure 9: Scatter diagram of forecasted MSLP, line indicates the straight line fit to the data

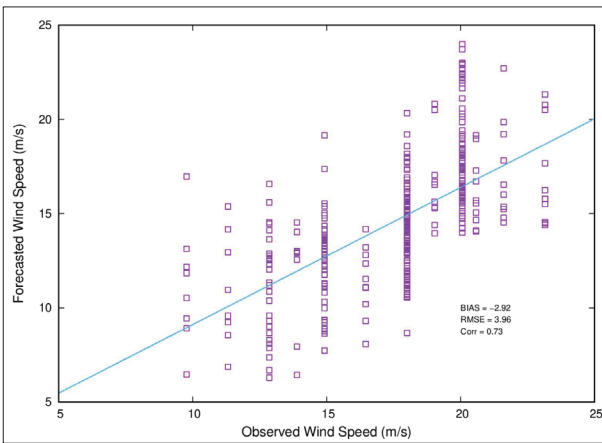


Figure 10: Scatter diagram of maximum forecasted 10-m wind speed, line indicates the straight line fit to the data

The forecast error of maximum wind is calculated based on the 10-m wind speed and the maximum sustained wind speed recorded in IBTrACS data (Knapp *et al.*, 2018). The mean absolute wind speed error up to a 72 hour lead time is shown in Table 2. The mean absolute wind speed error is less than 4.55 ms^{-1} for all the forecasts. The forecasted 10-m wind speed is shown in Figure 10 for the corresponding observed MSLP values of cyclone ‘Burevi’ recorded in IBTrACS data (Knapp *et al.*, 2018). The correlation coefficient was calculated using linear regression based on the observed and forecasted 10-m wind speed. The 10-m wind speed forecast has a bias of -2.92 , an RMSE of 3.96 ms^{-1} and a correlation coefficient of $+0.73$ compared to the observations. In Figure 10 it is evident that the forecasted 10-m wind speed is increased

based on the observed wind speed values with a positive correlation coefficient of 0.73.

The first landfall of the cyclone ‘Burevi’ was observed at the 9.35°N , 80.74°E latitude and longitude coordinate at 0041 UTC on 03 December 2020. The method used to calculate the observed landfall location and time is described in Section 2. The observed cyclone positions in the IBTrACS dataset (Knapp *et al.*, 2018) are used in that calculation. The same process is used to calculate the forecasted landfall location and time. The landfall location and time error for the first landfall were calculated and the results are shown in Table 3. The forecasted minimum landfall position error is recorded as 18.1 km with a lead time of 48 hours. The minimum forecasted landfall time error of the first landfall of the cyclone is 4 minutes with a lead time of 36 hours. Obtaining such information regarding the landfall location and time of the cyclone is an advantage when it comes to disaster management.

The intensity of the cyclone was reduced after the first landfall. This is clearly visible from the data recorded by Knapp *et al.* (2018), where the maximum sustained wind speed was 20 ms^{-1} at the first landfall, 14 ms^{-1} at the second landfall, and 10 ms^{-1} at the third landfall. The second landfall was observed at 9.08°N , 80.02°E at 0300 UTC on 05 December 2020. In all the forecasts conducted in this study the second landfall was only forecasted with a landfall error of 62 km in a forecast initialized at 0000 UTC on 03 December 2020. In that forecast, the landfall time error is minus 40 hours and 30 minutes. The third landfall was observed at 8.9°N , 79.9°E at 1800 UTC on 05 December 2020. This landfall is not

recorded in any of the forecasts conducted in this study. The intensity of the cyclone was reduced when it reached

its final stage. This is the reason for not forecasting the third landfall accurately.

Table 3: Landfall position and time error of the first landfall

Initialization time	Lead time (h)	Landfall position error (km)	Landfall time Error
00 UTC 29 November 2020	102	35.0	6 h 49 m
12 UTC 29 November 2020	96	21.0	11 h 49 m
00 UTC 30 November 2020	63	51.0	- 5 h 11 m
12 UTC 30 November 2020	48	18.1	- 11 h 33 m
00 UTC 01 December 2020	42	27.8	4 h 26 m
12 UTC 01 December 2020	36	48.8	4 m
00 UTC 02 December 2020	15	45.0	- 7 h 49 m
12 UTC 02 December 2020	0	58.8	- 9 h 56 m

CONCLUSIONS

The cyclone considered in this study has affected Sri Lanka by having its first landfall at 0041 UTC on 03 December 2020. Its track and intensity were forecasted with a maximum forecast duration of 120 hours using WRF. The track RMSE was reduced from 340.1 km to 59.0 km for model initializations from 0000 UTC on 27 November 2020 to 0000 UTC on 29 November 2020. This clearly shows the increase of accuracy of the forecasted track when the forecast lead time is reduced. For forecast lead times up to 24 hours the DPE, ATE, and CTE are less than 61 km, 48 km, and 42 km, respectively. When it comes to a forecast lead time up to 48 hours the DPE, ATE, and CTE are less than 106 km, 84 km, and 47 km, respectively. DPE, ATE, and CTE are higher for 0 and 3 hour forecast lead times compared to a 6 hour forecast lead time. The authors believe that this is due to the model spin-up phase. Providing the model with a pre-forecast period will be considered in future to minimize the effect of the spin-up phase. The forecasted mean absolute MSLP is less than 2.93 hPa up to a 72 hour forecast lead time. The forecasted MSLP has a positive correlation coefficient of 1.00 with its corresponding observations. The mean absolute wind speed error is less than 4.55 ms^{-1} for all the forecasts. The forecasted wind speed has a positive correlation coefficient of 0.73 with its corresponding observations. With all the above mentioned results, it is possible to conclude that WRF is capable of forecasting the track and the intensity of a cyclone that affects Sri Lanka with a forecast lead time up to 48 hours.

The first landfall of the cyclone ‘Burevi’ was forecasted with a minimum landfall position error of 18.1 km with a forecast lead time of 48 hours. The minimum forecasted landfall time error of the cyclone is 4 minutes with a forecast lead time of 36 hours. Obtaining such information is an advantage when it comes to disaster management. The intensity of the cyclone was reduced after the first landfall. The second landfall was forecasted by only one forecast. That had a landfall error of 62 km and a landfall time error of minus 40 hours and 30 minutes. The third landfall of the cyclone was not recorded in any of the forecasts conducted in this study. In conclusion, the first landfall position forecast of this cyclone is acceptable with the WRF configuration used in this study. The position and time forecasts for the second and third landfall require further investigation to improve their accuracy.

Acknowledgements

The authors would like to express their sincere gratitude to the National Center for Atmospheric Research (NCAR) and NCEP for providing the WRF model to the research community. The support given by the Faculty of Engineering, University of Ruhuna, Sri Lanka is highly appreciated.

REFERENCES

- Baki H., Chinta S., Balaji C. & Srinivasan B. (2022). Determining the sensitive parameters of the Weather Research and Forecasting (WRF) model for the simulation

- of tropical cyclones in the Bay of Bengal using global sensitivity analysis and machine learning. *Geoscientific Model Development* **15**(5): 2133–2155.
DOI: <https://doi.org/10.5194/gmd-15-2133-2022>
- Beljaars A.C.M. (1995). The parametrization of surface fluxes in large-scale models under free convection. *Quarterly Journal of the Royal Meteorological Society* **121**(522): 255–270.
DOI: <https://doi.org/10.1002/qj.49712152203>
- Bretherton C.S. & Park S. (2009). A new moist turbulence parameterization in the community atmosphere model. *Journal of Climate* **22**(12): 3422–3448.
DOI: <https://doi.org/10.1175/2008JCLI2556.1>
- Cotton W.R., Bryan G. & van den Heever S.C. (2011). *Chapter 8 - Cumulonimbus Clouds and Severe Convective Storms*, pp. 315–454, Academic Press.
DOI: [https://doi.org/10.1016/S0074-6142\(10\)09914-6](https://doi.org/10.1016/S0074-6142(10)09914-6)
- Dudhia J. (1989). Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of Atmospheric Sciences* **46**(20): 3077–3107.
- Farzan Z. (2020a). Cyclone Burevi makes landfall in Sri Lanka's east coast. Available at <https://www.newsfirst.lk/2020/12/02/cyclone-burevi-makes-landfall-in-sri-lankas-east-coast/>, Accessed 02 January 2021.
- Farzan Z. (2020b). Cyclone Burevi to cross Sri Lanka's east tomorrow. Available at <https://www.newsfirst.lk/2020/12/01/cyclone-burevi-to-cross-sri-lankas-east-tomorrow/>, Accessed 02 January 2021.
- Heming J.T. (2017). Tropical cyclone tracking and verification techniques for met office numerical weather prediction models. *Meteorological Applications* **24**(1): 1–8.
DOI: <https://doi.org/10.1002/met.1599>
- Hon K.-K. (2020). Tropical cyclone track prediction using a large-area WRF model at the Hong Kong Observatory. *Tropical Cyclone Research and Review* **9**(1): 67–74.
DOI: <https://doi.org/10.1016/j.tcr.2020.03.002>
- Hong S.-Y. & Lim J.-O. J. (2006). The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). *Journal of the Korean Meteorological Society* **42**(2): 129–151.
- Hutchinson T.A. (2007). An adaptive time-step for increased model efficiency. *Eighth WRF Users' Workshop*, NCAR, Boulder, USA.
- Islam T., Srivastava P., Rico-Ramirez M., Dai Q., Gupta M. & Singh S. (2015). Tracking a tropical cyclone through WRF–ARW simulation and sensitivity of model physics. *Natural Hazards* **76**: 1473–1495.
DOI: <https://doi.org/10.1007/s11069-014-1494-8>
- Knapp K.R., Diamond H.J., Kossin J.P., Kruk M.C. & Schreck C.J. (2018). International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4. NOAA National Centers for Environmental Information, USA.
DOI: <https://doi.org/10.25921/82ty-9e16>
- Kotal S. & Bhattacharya S. (2020). Improvement of wind field forecasts for tropical cyclones over the North Indian Ocean. *Tropical Cyclone Research and Review* **9**(1): 53–66.
DOI: <https://doi.org/https://doi.org/10.1016/j.tcr.2020.03.004>
- Mlawer E.J., Taubman S.J., Brown P.D., Iacono M.J. & Clough S.A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres* **102**(D14): 16663–16682.
DOI: <https://doi.org/10.1029/97JD00237>
- Mohanty S., Nadimpalli R., Mohanty U.C., Mohapatra M., Sharma A., Das A.K. & Sil S. (2021). Quasi-operational forecast guidance of extremely severe cyclonic storm Fani over the Bay of Bengal using high-resolution mesoscale models. *Meteorology and Atmospheric Physics* **133**: 331–348.
DOI: <https://doi.org/10.1007/s00703-020-00751-4>
- Nadimpalli R., Osuri K.K., Mohanty U.C., Das A.K., Kumar A., Sil S. & Niyogi D. (2020). Forecasting tropical cyclones in the Bay of Bengal using quasi-operational WRF and HWRF modeling systems: An assessment study. *Meteorology and Atmospheric Physics* **132**(1): 1–17.
DOI: <https://doi.org/10.1007/s00703-019-00669-6>
- Osuri K.K., Ankur K., Nadimpalli R. & Busireddy N.K.R. (2020). Error characterization of ARW model in forecasting tropical cyclone rainfall over North Indian Ocean. *Journal of Hydrology* **590**: 125433.
DOI: <https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125433>
- Priyantha H.D.G. (2020). Disaster Management Centre Situation Report 8 December 2020. Available at https://www.dmc.gov.lk/images/dmcreports/Situation_Report_on_2020_1607430892.pdf, Accessed 2 January 2021.
- Sinnott R.W. (1984). Virtues of the Haversine. *Sky and Telescope* **68**(2): 158.
- Skamarock W., Klemp J., Dudhia J., Liu Z., Berner J., Wang W., Powers J., Duda M., Barker D., & Huang X.-Y. (2019). *A Description of the Advanced Research WRF version 4*. NCAR Technical Notes.
- Tewari M., Chen F., Wang W., Dudhia J., LeMone M.A., Mitchell K., Ek M., Gayno G., Wegiel J. & Cuenca R. H. (2004). Implementation and verification of the unified NOAA land surface model in the WRF model. *20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*, 11–15.
- Zhang C. & Wang Y. (2017). Projected future changes of tropical cyclone activity over the Western North and South Pacific in a 20-km-mesh regional climate model. *Journal of Climate* **30**(15): 5923–5941.
DOI: <https://doi.org/10.1175/JCLI-D-16-0597.1>

RESEARCH ARTICLE

Marine Microalgae

Anti-inflammatory activity of nonyl 8-acetoxy-6-methyloctanoate, isolated from the cultured marine diatom, *Phaeodactylum tricornutum* : mediated via suppression of inflammatory mediators in LPS-stimulated RAW 264.7 macrophages

KW Samarakoon^{1,2*}, JY Ko^{2,3}, JH Lee^{2,4} and YJ Jeon²

¹ Institute for Combinatorial Advanced Research and Education (KDU-CARE), General Sir John Kotelawala Defence University, Kandawala Road, 10390, Ratmalana.

² Department of Marine Life Sciences, Jeju National University, Jeju 690-756, Republic of Korea.

³ Research Institute of Basic Sciences, Incheon National University, Incheon 22012, Republic of Korea.

⁴ Lee Gill Ya Cancer and Diabetes Institute, 7-45, Songdodong, Yesonsu, Incheon 406-840, Republic of Korea.

Submitted: 03 February 2021; Revised: 14 July 2022; Accepted: 22 July 2022

Abstract: In this study, the anti-inflammatory effects of nonyl 8-acetoxy-6-methyloctanoate (NAMO), isolated from the cultured marine diatom, *Phaeodactylum tricornutum* Bohlin, against LPS-induced RAW 264.7 macrophages were evaluated. NAMO has indicated the strongest inhibitory effects against nitric oxide (NO) and prostaglandin E₂ (PGE₂) production, dose-dependently, on lipopolysaccharide (LPS)-induced RAW 264.7 cells. The 50% NO production inhibitory concentration (IC₅₀) of NAMO was 24.8 μM with the least cytotoxic effect in both LDH and MTT assays. Inflammatory stimulators such as LPS, which induce cytokines in the process of macrophage activation and mediate tissue response in different phases of inflammation, were studied. NAMO showed significant and strong suppression of pro-inflammatory cytokine and interleukin-1β (IL-1β) production, but no significant inhibitory effect on the production of cytokines including tumor necrosis factor-α (TNF-α) and interleukin-6 (IL-6) at the tested concentrations against LPS treatment on RAW macrophages. Western blot analysis was carried out to determine the protein expressions of inducible nitric oxide synthase (iNOS) and cyclooxygenase-2 (COX-2) proteins which mediate the suppression effect of NAMO on NO and PGE₂ production. The Western blot assay confirmed the suppression of iNOS and COX-2 protein expressions against LPS-stimulated RAW264.7 cells. Collectively, NAMO isolated from the cultured marine diatom, *P. tricornutum* exhibited a

profound anti-inflammatory effect *in vitro*, suggesting that the compound might have a beneficial effect during the treatment of inflammatory diseases and can be used in functional food applications.

Keywords: Anti-inflammatory activity, cultured marine diatom, fatty alcohol esters, *Phaeodactylum tricornutum*, RAW 264.7 cells

INTRODUCTION

Marine microalgae, with the associated broad spectrum of secondary metabolites, have stimulated a vast interest due to biochemical and ecological diversity (Borowitzka, 1995). Exploration of new bioactive natural products from marine bio-resources like marine microalgae for pharmacological applications has been described as a promising aspect of this. The interest in the isolation of pharmacologically active metabolites from marine microalgae has increased throughout the world. Microalgae are rich sources of nutrients including a high content of protein, polyunsaturated fatty acids (PUFA), carbohydrates, minerals, vitamins, pigments, and secondary metabolites (Becker, 2007).

* Corresponding author (samarakoon@kdu.ac.lk;  <https://orcid.org/0000-0002-9668-3293>)



Cultured marine microalgae have been widely used as a food source for aquaculture. Diatoms are abundant components of the phytoplankton, representing major nutritional sources. *Phaeodactylum tricorutum*, a marine diatom, has been characterized due to containing different forms of PUFAs and lipids, including free sterols, and esterified and glycosylated conjugates (Veron *et al.*, 1996; Yang *et al.*, 2019). A few reports have investigated the quantification of lipids that are available in *P. tricorutum*. The distribution of lipid constituents in marine microalgae is dependent on the growing conditions, that include different temperatures, light spectral quality, salinity, and nutrients provided in the aquaculture (Samarakoon *et al.*, 2013). The possibility of changing the lipid concentration of microalgae is a very important characteristic that can be used to determine the physiological state of a microalgae population, as a population biomarker (Fabregas *et al.*, 1997).

Inflammation is a physiological process initiated by pathogenic invasion or cell and tissue injury (Wadleigh *et al.*, 2000). Inflammation can be caused by the activation of various immune cells such as macrophages, lymphocytes, and neutrophils. Overproduction of pro-inflammatory cytokines including interleukin (IL) and tumor necrosis factor (TNF)- α can be stimulated in macrophages as per the inflammatory disease. In addition, inflammatory mediators such as nitric oxide (NO) and prostaglandin (PGE₂) are also activated due to inducible nitric oxide synthase (iNOS) and cyclooxygenase-2 (COX-2) (Lee *et al.*, 2006). Therefore, finding a better insight into anti-inflammatory agents is gaining interest, in order to regulate immune-related diseases, inflammatory disorders, and infections (Jung *et al.*, 2009). Assessing the anti-inflammatory effect of the isolated lipid against LPS-stimulated macrophages from cultured marine microalgae was of great interest. Hence, in this study, we examined whether the cultured marine diatom, *P. tricorutum* is capable of producing novel compounds with pharmaceutical value. In our previous publication, the anticancer activity of a novel fatty alcohol ester; nonyl 8-acetoxy-6-methyl octanoate (NAMO) isolated from the cultured marine diatom *P. tricorutum* has been revealed (Samarakoon *et al.*, 2014). In addition, in this study, we evaluate the *in vitro* anti-inflammatory activity of the isolated novel compound from the hexane fraction of the extract from the cultured marine diatom, *P. tricorutum*, against LPS-stimulated RAW macrophages.

MATERIALS AND METHODS

Chemicals and reagents

Dulbecco's Modified Eagle's Medium (DMEM), foetal bovine serum (FBS), and penicillin-streptomycin were purchased from Gibco/BRL (Burlington, ON, Canada). 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), and dimethyl sulphoxide (DMSO) were purchased from Sigma-Aldrich (St. Louis, MO, USA). All other chemicals and reagents used in these investigations were of analytical grade.

Culturing of *Phaeodactylum tricorutum*

The marine diatom *Phaeodactylum tricorutum* Bohlin was kindly provided by the Korea Marine Microalgae Culture Center (KMMCC). The alga was inoculated in 30 L plastic cylinders at 20 °C after pre-cultivation in 5 L glass vessels (medium 4 L), and the air was continuously supplied at 5 L min⁻¹ by air-lift. The light was provided by 60 W fluorescent lamps at an intensity of 2,500 lx (Light : Darkness = 24 : 0). The alga was cultured in Conway medium (Walne, 1966) prepared from filter-sterilized seawater, and the cultures were continuously active during the 8-10 ds after inoculation. The cells were flocculated with 200 ppm Al₂(SO₄)₃ (v/v) (Ilshin biochemical, Magicpool-99) and then recovered with centrifugation at 2,000 rpm using a basket centrifuge (Hanseong Co., Ansan, South Korea). The harvested diatom biomass was frozen at -25 °C and preserved until freeze drying.

Extraction and isolation

The material (38 g) was extracted three times with 80% methanol by sonication for 90 min each time, at 25 °C. After being concentrated by evaporating off the solvent under reduced pressure using a rotary evaporator, the crude methanol extract was further subjected to solvent-solvent partition chromatography. Then, four different fractions with varying polarity, namely, *n*-hexane, chloroform, ethyl acetate, and aqueous extracts were separated. The *n*-hexane fraction (550 mg) was found to be active and subjected to fractionation using a solid-liquid phase chromatography (normal phase silica) column (3 cm × 22 cm) using hexane and ethyl acetate as an increasing hydrophilic solvent system to

give 14 fractions, labelled F1~F14. Using the thin layer chromatography (TLC) results, visualized by heating after spraying with ethanol (90%) - sulphuric acid (10%) staining system, the desired fraction was identified as eluting with hexane (10%): ethyl acetate (90%), (F7: 19 mg), with enough purity to determine the molecular structure. Therefore, the molecular mass was measured using a high-performance liquid chromatography - diode array detector coupled with an electrospray ionization mass spectrometer (HPLC-DAD-ESI/MS) (Hewlett-Packard, Waldbronn, Germany), and NMR studies using a JEOL JNM-ECX400 spectrometer ($^1\text{H-NMR}/400\text{ MHz}$ and $^{13}\text{C-NMR}/100\text{ MHz}$) used for structure elucidation.

Nonyl 8-acetoxy-6-methyloctanoate (NAMO)

Liquid (CHCl_3): $\text{C}_{20}\text{H}_{38}\text{O}_4$ in the ESI (negative mode) as (M-H) 341.14 m/z. The calculated molecular mass is 342.28 m/z (M). $^1\text{H-NMR}$ (400 MHz, CDCl_3): δ /ppm: 0.87 (3H, t, $J = 6.72\text{ Hz}$, H-9'), 0.96 (3H, s, H-11), 1.25 (2H, m, H-4, 5, 5', 6'), 1.29 (2H, m, H-4', 7'), 1.31 (2H, m, H-8'), 1.43 (2H, m, H-3'), 1.50 (2H, m, H-2'), 1.53 (2H, m, H-7), 1.65 (H, m, H-6), 1.68 (2H, m, H-3), 2.21 (3H, s, H-10), 2.32 (2H, t, $J = 7.53\text{ Hz}$, H-2), 4.08 (2H, dd, $J = 7.4, 14.6\text{ Hz}$, H-8), 4.13 (2H, dd, $J = -7.4, 14.6\text{ Hz}$, H-1'). $^{13}\text{CNMR}$ (100 MHz, CDCl_3): δ 14.3 (q, C-9'), 21.2 (q, C-10), 22.8 (q, C-11), 22.9 (t, C-8'), 24.6 (t, C-3'), 24.7 (t, C-4), 24.9 (t, C-3), 29.2 (t, C-2'), 29.4 (t, C-7), 29.6 (t, C-4'), 29.9 (t, C-5'), 29.9 (t, C-6'), 32.1 (t, C-7'), 32.9 (d, C-6), 34.1 (t, C-2), 37.5 (t, C-5), 60.6 (t, C-1'), 63.3 (t, C-8), 171.3 (s, C-9), 179.2 (s, C-1).

Cell culture

The murine macrophage cell line RAW 264.7 was purchased from the Korean Cell Line Bank (KCLB, Seoul, South Korea). RAW 264.7 cell line was cultured in Dulbecco's modified eagle medium (DMEM) supplemented with 100 U mL^{-1} of penicillin, 100 $\mu\text{g mL}^{-1}$ of streptomycin, and 10% fetal bovine serum (FBS). The cells were incubated and maintained in an atmosphere of 5% CO_2 at 37 °C. The cells were sub-cultured every 2 days and exponential phase cells were used throughout the experiments (Kim *et al.*, 2013).

Determination of nitric oxide (NO) production

RAW 264.7 cells (1×10^5 cell mL^{-1}) were placed in a 24-well plate and after 24 h the cells were pre-incubated

with various concentrations of the sample at 37 °C for 1 h. Then further incubation was done for another 24 h with LPS ($1\ \mu\text{g mL}^{-1}$) at the same temperature. After the incubation, the quantity of nitrite accumulated in the culture medium was measured as an indicator of NO production (Lee *et al.*, 2006). Briefly, 100 μL of cell culture medium were mixed with 100 μL of Griess reagent (1% sulfanilamide and 0.1% naphthyl ethylenediamine dihydrochloride in 2.5% phosphoric acid), and the mixture was incubated at room temperature for 10 min, and the optical density at 540 nm was measured using an enzyme-linked immunosorbent assay (ELISA) microplate reader (Sunrise, Tecan Co. Ltd., Australia). The fresh culture medium was used as a blank in every experiment (Kim *et al.*, 2013).

Lactate dehydrogenase (LDH) cytotoxicity assay

RAW 264.7 cells (1.5×10^5 cells mL^{-1}) were plated in a 96-well plate and after 16 h the cells were pre-incubated with various concentrations of the sample for 1 h at 37 °C. Then the cells were further incubated for another 24 h with LPS ($1\ \mu\text{g mL}^{-1}$) at the same temperature. After the incubation, the LDH level in the culture medium was determined using an LDH cytotoxicity detection kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. Briefly, 50 μL of the reaction mixture were added to each well, and the incubation was done for 30 min at room temperature in the dark. Then, 50 μL of stop solution were added to each well, and absorbance was measured at 490 nm using a microplate reader (Sunrise, Tecan Co. Ltd., Australia) (Kim *et al.*, 2013).

Cytotoxicity assessment using MTT assay

The cytotoxicity of NAMO against the RAW 264.7 cells was determined using the colorimetric MTT assay. Cells were seeded in a 24-well plate at a concentration of 1×10^5 cells/ mL^{-1} . After 24 h, the seeded cells were treated with sample. Then, all of the cells were incubated for an additional 24 h at 37 °C. MTT stock solution (50 μL ; 2 mg/ mL in PBS) was added to each well to a total reaction volume of 250 μL . After incubating for 3 h, the supernatants were aspirated. The formazan crystals in each well were dissolved in 200 μL of dimethylsulphoxide (DMSO). The resulting absorbance was measured with an ELISA plate reader set at 540 nm (Samarakoon *et al.*, 2014).

Determination of pro-inflammatory cytokines, tumor necrosis factor- α (TNF- α), interleukin-1 β (IL-1 β), and interleukin-6 (IL-6) production.

The inhibitory effect of the sample on the production of pro-inflammatory cytokines from LPS-stimulated RAW 264.7 cells was determined according to a previously described method (Cho *et al.* 2000). Briefly, RAW 264.7 cells (1×10^5 cells mL⁻¹) were pretreated with the sample for 2 h and then treated with LPS ($1 \mu\text{g mL}^{-1}$) to allow the production of pro-inflammatory cytokines for 24 h. Supernatants were used for the assay using an ELISA kit (R & D Systems, Minneapolis, MN, USA) according to the manufacturer's instructions.

Determination of prostaglandin-E₂ (PGE₂) production

RAW 264.7 cells (1×10^5 cells mL⁻¹) were pretreated with the sample for 2 h and then treated with LPS ($1 \mu\text{g mL}^{-1}$) to allow the production of pro-inflammatory cytokines for 24 h. The PGE₂ levels in the culture medium were quantified using a competitive enzyme immunoassay kit (R & D Systems, Minneapolis, MN, USA) according to the manufacturer's instructions. The release of PGE₂ was measured relative to that of the control value (Cho *et al.*, 2000).

Western blot analysis

RAW 264.7 cells (1×10^5 cells mL⁻¹) were pretreated for 16 h and then treated with LPS ($1 \mu\text{g mL}^{-1}$) in the presence or absence of the sample. After incubation for 24 h, the cells were harvested, washed twice with ice-cold phosphate-buffered saline (PBS), and the cell lysates were prepared with lysis buffer (50 mmol L⁻¹ Tris-HCl (pH 7.4), 150 mmol L⁻¹ NaCl, 1% Triton X-100, 0.1% SDS and 1 mmol L⁻¹ ethylenediaminetetraacetic acid (EDTA)) for 20 min on ice. Cell lysates were centrifuged at 14,000 x g for 20 min at 4°C. Then the protein concentrations in the supernatants were measured using a BCATM protein assay kit. Cell lysates containing 30 μg of protein were subjected to electrophoresis using sodium dodecyl sulphate-polyacrylamide gels (SDS-PAGE) at 12%, and the separated proteins were transferred onto nitrocellulose membranes (Bio-Rad, Hercules, CA, USA). The membranes were incubated with a blocking solution (5% skim milk in Tris-buffered saline containing Tween -20) for 90 min at room temperature. Then the membrane was incubated with anti-mouse iNOS (1:1000; Calbiochem, La Jolla, CA, USA) and anti-mouse COX-2 (1:1000; BD BiosciencPharmingen, San Jose, CA, USA) overnight at room temperature.

After washing, the blots were incubated with horseradish peroxidase-conjugated goat anti-mouse IgG secondary antibody (1: 5000; Amersham Pharmacia Biotech, Little Chalfont, UK) for 90 min at the ambient temperature. Signals were developed by exposure to X-ray films and visualized according to the described procedure, using a Western Lightning-ECL detection kit (PerkinElmer, MA, USA).

Statistical analysis

All the data are expressed as mean \pm standard deviation of three determinations. Statistical comparison was performed via a one-way analysis of variance followed by Duncan's multiple range test (DMRT). P values of less than 0.05 ($p < 0.05$) were considered as significant.

RESULTS AND DISCUSSION

Inhibitory effect of 8-acetoxy-6-methyloctanoate (NAMO) from *P. tricornutum* against LPS-induced nitric oxide production and cytotoxicity on RAW macrophages

The isolated compound (Figure 1), NAMO from the hexane fraction of the cultured marine microalga, *P. tricornutum*, was studied to assess the anti-inflammatory activity using LPS-induced RAW macrophages, and the NO production inhibitory and cytotoxic effects were determined. Five different concentrations (0, 5, 12.5, 25, and 50 $\mu\text{g mL}^{-1}$) of NAMO were pre-incubated with LPS for 24 h and the significant inhibitory effect on NO production level was observed dose-dependently. The 50% inhibitory concentration (IC₅₀) value of 8.50 $\mu\text{g mL}^{-1}$ (24.8 μM) was determined. Nitric oxide (NO) is a signaling molecule. The inflammatory mediator, which is stimulated by iNOS. The effects of various

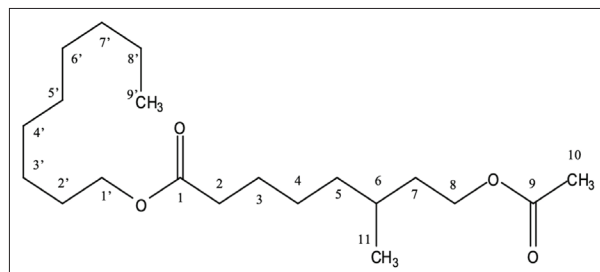


Figure 1: Structure elucidation of nonyl 8-acetoxy-6-methyloctanoate (NAMO) isolated from *Phaeodactylum tricornutum*

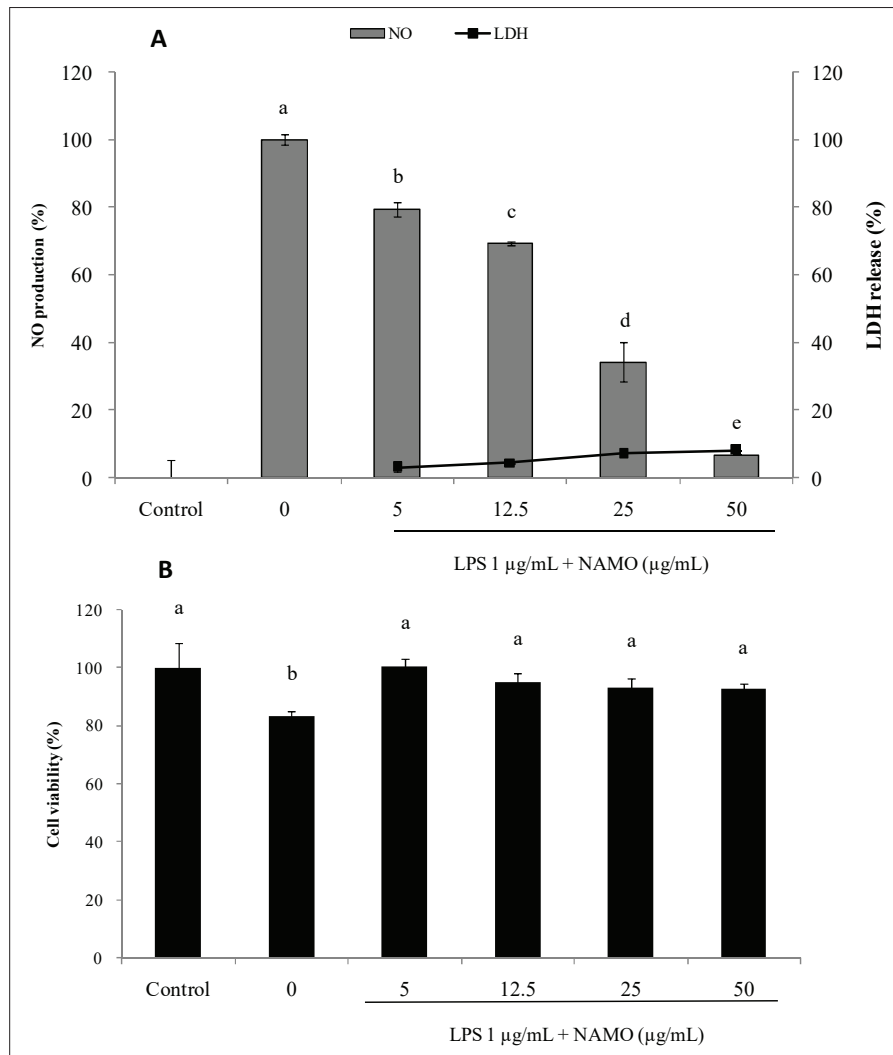


Figure 2: Inhibitory effect of the isolated NAMO compound from cultured marine diatom *Phaeodactylum tricornutum* against LPS-induced NO production (%) and cell viability (%) by LDH (A) and MTT assay in RAW 264.7 macrophages (B). Values are expressed as mean \pm SD in triplicate experiments. Values with different letters are significantly different at $p < 0.05$ as analyzed by DMRT.

concentrations of effect of NAMO isolated from the cultured marine diatom *P. tricornutum* on NO production in LPS-activated RAW 264.7 cells was evaluated *in vitro*. LPS stimulates the cells to form NO molecules and causes inflammation. The increase of NO concentration in the medium was dose-dependently suppressed with the incubation of NAMO (Figure 2). Moreover, the cell viability assay indicated that NAMO did not show a cytotoxic effect on RAW 264.7 cells at all the

concentrations. The cytotoxic effects on RAW 264.7 cells with the tested concentrations of NAMO were determined by MTT assay, and a 5.0 $\mu\text{g/mL}$ concentration of NAMO gave the highest cell viability, 100%, compared to the negative control (Figure 2B). It is also confirmed that the LDH release assay indicated little or no cytotoxicity on macrophages at the tested concentrations (Figure 2A). Therefore, NAMO is marked as a potential agent for attenuating NO production without cytotoxic effects.

Inhibitory effect of NAMO on LPS-induced PGE₂ production

In this study, NAMO inhibited the LPS-induced PGE₂ production in a dose-dependent manner. PGE₂ production

was 30% of normal at 25 µg mL⁻¹ concentration of NAMO, and 100 % without NAMO treatment. Therefore, NAMO might induce anti-inflammatory effects by strongly inhibiting PGE₂ production when incubated with RAW macrophages (Figure 3 A).

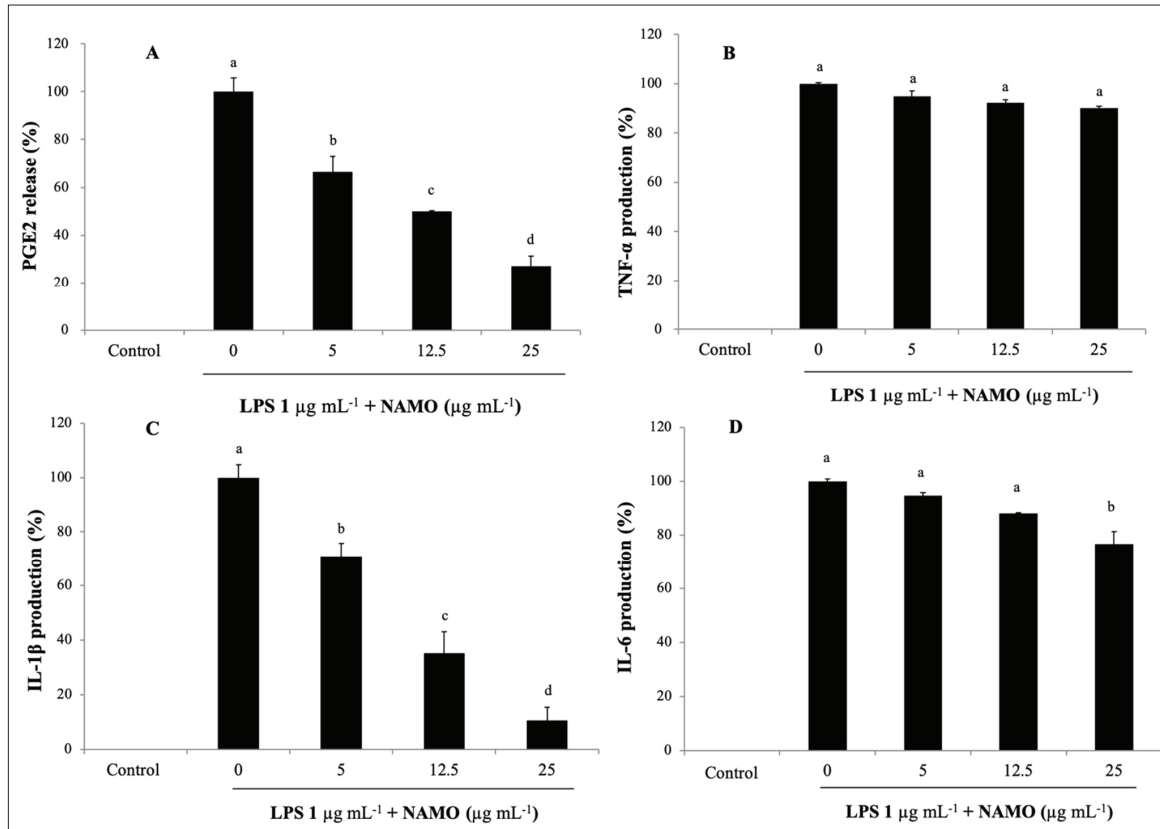


Figure 3: Inhibitory effect of the isolated NAMO compound from cultured marine diatom *Phaeodactylum tricoratum* on LPS-induced PGE₂ (A), TNF-α (B), IL-1β (C) and IL-6 (D) production in RAW 264.7 macrophages, after the incubation of cells with LPS for 24 h in the presence or absence of NAMO. Values are expressed as mean ± SD in triplicate experiments. Values with different letters are significantly different at $p < 0.05$ as analyzed by DMRT.

Inhibitory effect of NAMO on LPS-induced pro-inflammatory cytokine (TNF-α, IL-1β, and IL-6) Production

The inhibition of cytokine production or regulation is a key mechanism in the mediation of inflammation. Inflammatory stimulators such as LPS induce cytokines in the process of macrophage activation and mediate tissue response in different phases of inflammation. In this assay, the inhibitory effect of NAMO on the production

of pro-inflammatory cytokines such as TNF-α, IL-1β, and IL-6 in LPS-stimulated RAW 264.7 macrophages was studied. The NAMO did not show significant inhibitory activities on the production of cytokine TNF-α and IL-6 at the tested concentrations (Figure 3B, D). However, the strong suppression of the production of cytokine IL-1β was observed after the pretreatment of NAMO with RAW macrophages (Figure 3C). Therefore, the inhibition pathway might be mediated by the association of IL-1β formation.

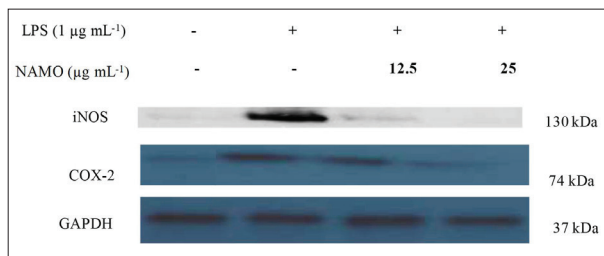


Figure 4: Inhibitory effect of the isolated NAMO compound from cultured marine diatom *Phaeodactylum tricornerutum* on LPS-induced iNOS and COX-2 protein expression in RAW 264.7 macrophages. The cells were incubated with LPS for 24 h in the presence or absence of the NAMO. Then, cell lysates were subjected to electrophoresis and the expression levels of iNOS and COX-2 were detected with specific antibodies.

Inhibitory effect of NAMO on LPS-induced iNOS and COX-2 protein expression

Western blot analysis was carried out to determine the protein expressions of iNOS and COX-2 proteins which mediate the suppression effect of NAMO on NO and PGE₂ production (Figure 4).

According to the results, iNOS and COX-2 protein expressions were markedly induced when the macrophages were treated with LPS compared to the control. However, upon incubation with NAMO followed by LPS treatment, both iNOS, and COX-2 protein expression was downregulated, and at the dose of 25 µg mL⁻¹, the COX-2 protein expression was completely suppressed. Overall, it was suggested that the strong inhibitory effect on PGE₂ production by NAMO was correlated with the down-regulation pattern of COX-2 protein expression. In addition, NAMO inhibited the production of NO levels by suppressing iNOS proteins.

P. tricornerutum is a well-known marine diatom that accumulates fatty acids with varying degrees of unsaturation. In particular, the compositions of the sterols, PUFA, and fatty acids which are associated with *P. tricornerutum* biomasses as lipids have different functional values that are of significant importance for the advancement of marine biotechnology (Radakovits *et al.*, 2011). The cultivation of marine microalgae in an aquaculture system has been practised for a few decades. Open or closed (photobioreactor) microalgae culture systems are applicable based on the energy source and type of culture. Each one of the culturing systems has either desirable or undesirable effects, since

the photosynthetic organisms and their mass production may directly correlate with the source of energy (Hong *et al.*, 2012). The controlled optimum conditions with operational inputs such as salt, dissolved CO₂, water, nutrients, pH, and O₂ provides a great opportunity for a steady environment without being contaminated by photo-bioreactors. Researchers and technologists have been attracted to cultured marine microalgae to disclose the biochemical constituents of the crude extracts and to find out which components have pharmacological effects (Plaza *et al.*, 2009). Based on the major ingredients available in cultured marine microalgae, it is a promising source for the functional food industry (Garcia *et al.*, 2006; Samarakoon & Jeon, 2012;).

The results of our previous work, culturing and harvesting the biomass of the marine diatom, *P. tricornerutum*, were discussed (Samarakoon *et al.*, 2013). In addition, biological activity-guided fractionation led to the isolation of five different compounds from the hexane fraction of cultured *P. tricornerutum*. The isolated novel fatty alcohol ester; NAMO was evaluated for inhibiting the growth of human leukemia cancer cells significantly through the p53 and caspase-3 mediated cell apoptotic pathway (Samarakoon *et al.*, 2014). In the present study, the same isolated compound from *P. tricornerutum* was evaluated for anti-inflammatory effects *in vitro*. Nitric oxide (NO) acts as an intercellular and intracellular signal molecule and plays crucial roles in many cellular functions in the nervous, cardiac, vascular, and immune systems (Wichard & Pohnert, 2006). However, the low levels of NO production are important in protecting organs such as the liver from ischaemic damage but the chronic expression of NO is associated with various inflammatory conditions (Kassim *et al.*, 2010). Over the pathological conditions, NO production is increased by iNOS (Kim *et al.*, 1999; Sahu *et al.*, 2013). In terms of the potential therapeutic value when related to inflammation, inhibition of NO production by suppression of iNOS expression can be considered. A few recent publications have confirmed that the organic solvent extractions and compounds isolated from plant-based materials attenuated the inflammations caused by LPS-induced RAW macrophages via key signaling pathways such as the NF-κB (Ahujaa *et al.*, 2019; Hao *et al.*, 2019), and the iNOS pathway (Bharadwaj *et al.*, 2020). The ethanol extract of a new microalgal strain, *Aurantiochytrium limacine* showed the significant suppression of genes for pro-inflammatory cytokines mediated by the NF-κB pathway (Takahashi *et al.*, 2019). With the profound inhibitory effect on NO production evidenced by NAMO, it was further evaluated for its ability to inhibit the LPS-induced PGE₂ production in RAW264.7 macrophages.

The present findings indicated that NAMO has anti-inflammatory properties, and the effects might be caused by inhibiting iNOS, COX-2, TNF- α , IL-1 β , and IL-6 expression via affecting the NF- κ B pathway. The accessible nature of the cultured marine microalgae is useful not only for the development of pharmacological prospects, but also to meet an increasing demand for nutraceuticals and food supplements. Exploring the sustainable use of cultured marine diatoms is considered important in future industrial and viable applications, and for bioprospecting for the novel bioactive metabolites. Developing new and safer treatment strategies using biomaterials extracted from marine microalgae to prevent and treat inflammatory diseases could be possible (Morais *et al.*, 2015). Therefore, we have evidence that the isolated NAMO from the cultured marine diatom *P. tricornutum* is a promising source for both *in vitro* anticancer and anti-inflammatory effects. The findings of the study strongly emphasize the clinical and biotechnological potential of the isolated novel fatty alcohol ester from the cultured marine diatom *P. tricornutum* for further developments.

CONCLUSION

NAMO, isolated from the cultured marine diatom *P. tricornutum*, was identified as an anti-inflammatory active compound and showed promising inhibitory effects against cytokine production or regulation, which is a key mechanism in the mediation of inflammation against 264.7 RAW macrophages. NAMO demonstrated an inhibitory effect against NO with the 50% inhibitory concentration (IC₅₀) value of 8.50 $\mu\text{g mL}^{-1}$ (24.8 μM). In addition, NAMO attenuated PGE₂ and IL-1 β production significantly in LPS-induced RAW cells. Moreover, the protein expression of iNOS and COX-2 confirmed the down-regulatory effect in a dose-dependent manner against RAW macrophages. Hence, the isolated NAMO can be a novel alternative with anti-inflammatory efficacy for future lead compounds in pharmacological applications.

Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of interest

The author(s) declare(s) that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors thank Dr. O-Nam Kwon, Marine Biology Center for Research and Education, Gangneung-Wonju National University, Gangwon-do, the Republic of Korea, for the cultivation of microalgae and for providing the cultured marine diatom for this study.

REFERENCES

- Ahujaa A., Kim M.Y. & Cho J.Y. (2019). *Protium javanicum* Burm. methanol extract attenuates LPS-induced inflammatory activities in macrophage-like RAW264.7 cells. *Evidence-Based Complementary and Alternative Medicine* **12**: 2910278.
DOI: <https://doi.org/10.1155/2019/2910278>
- Becker E.W. (2007). Micro algae as a source of protein. *Biotechnology Advances* **25**: 207–210.
DOI: <https://doi.org/10.1016/j.biotechadv.2006.11.002>
- Bharadwaj M., Padmavathy T.K., Mani S., Malarvizhi R., Sali V.K. & Vasanthi H.R. (2020). Sulfated polysaccharide from *Turbinaria ornata* suppress lipopolysaccharide-induced inflammatory response in RAW 264.7 macrophages. *International Journal of Biological Macromolecules* **164**: 4299–4305.
DOI: <https://doi.org/10.1016/j.ijbiomac.2020.09.036>
- Borowitzka M.A. (1995). Microalgae as sources of pharmaceuticals and other biological active compounds. *Journal of Applied Phycology* **7**: 3–15.
- Cho J.Y., Baik K.U., Jung J.H. & Park M.H. (2000). In vitro anti-inflammatory effects of cynaropicrin, a sesquiterpene lactone, from *Saussurea lappa*. *European Journal of Pharmacology* **398**: 399–407.
DOI: [https://doi.org/10.1016/s0014-2999\(00\)00337-x](https://doi.org/10.1016/s0014-2999(00)00337-x)
- Fabregas J., Aran J., Morales E.D., Lamela T. & Otero A. (1997). Modification of Sterol concentration in marine microalga. *Phytochemistry* **46**: 1189–1191.
DOI: [https://doi.org/10.1016/S0031-9422\(97\)80009-X](https://doi.org/10.1016/S0031-9422(97)80009-X)
- Garcia C.M.C., Camacho F.G., Miron A.S., Sevilla J.M.F., Chisti Y. & Grima E.M. (2006). Mixotrophic production of marine microalgae *Phaeodactylum tricornutum* on various carbon sources. *Journal of Microbiology and Biotechnology* **16**: 689–694.
- Hao X., Sun W., Ke C., Wang F., Xue Y., Luo Z., Wang X., Zhang J. & Zhang Y. (2019). Anti-inflammatory activities of leaf oil from *Cinnamomum subavenium* *in vitro* and *in vivo*. *BioMed Research International* **2019**: 1823149.
DOI: <https://doi.org/10.1155/2019/1823149>
- Hong J.W., Kim S.A., Chang J. Yi J., Jeong J., Kim S.H. & Yoon H.S. (2012). Isolation and description of a Korean microalga, *Asterarcys quadricellulare* KNUA020, and analysis of its biotechnological potential. *Algae* **27**: 197–203.
DOI: <https://doi.org/10.4490/algae.2012.27.3.197>
- Jung H.W., Seo U.K., Kim J.H., Leem K.H. & Park Y.K. (2009). Flower extract of *Panax notoginseng* attenuates

- lipopolysaccharide-induced inflammatory response via blocking of NF-kappaB signaling pathway in murine macrophages. *Journal of Ethnopharmacology* **122**: 313–319.
DOI: <https://doi.org/10.1016/j.jep.2008.12.024>
- Kassim M.A., Choui M., Mustafa M.R., Mohd M.A. & Yusoff K.M. (2010). Ellagic acid, phenolic acids, and flavonoids in Malaysian honey extracts demonstrate in vitro anti-inflammatory activity. *Nutrition Research* **30**: 650–659.
DOI: <https://doi.org/10.1016/j.nutres.2010.08.008>
- Kim H.K., Cheon B.S., Kim Y.H., Kim S.Y. & Kim H.P. (1999). Effects of naturally occurring flavonoids on nitric oxide production in the macrophage cell line RAW264.7 and their structure-activity relationship. *Biochemical Pharmacology* **58**: 759–765.
DOI: [https://doi.org/10.1016/S0006-2952\(99\)00160-4](https://doi.org/10.1016/S0006-2952(99)00160-4)
- Kim *et al.* (11 authors) (2013). Anti-inflammatory effects of trans-1,3-diphenyl-2,3-epoxypropane-1-one mediated by suppression of inflammatory mediators in LPS-stimulated RAW 264.7 macrophages. *Food and Chemical Toxicology* **53**: 371–375.
DOI: <https://doi.org/10.1016/j.fct.2012.12.021>
- Lee H.L., Hyun E. A., Yoon W. J., Kim B.H., Rhee M.J. & Kang H.K. (2006). In vitro anti-inflammatory and anti-oxidative effects of *Cinnamomum camphora* extracts. *Journal of Ethnopharmacology* **103**: 208–216.
DOI: <https://doi.org/10.1016/j.jep.2005.08.009>
- Morais M.G., Vaz B.S., Morais E.G. & Costa A.V. (2015). Biologically active metabolites synthesized by microalga. *BioMed Research International* 2015:835761.
DOI: <https://doi.org/10.1155/2015/835761>
- Plaza M., Herrero M., Cifuentes A. & Ibanez E. (2009). Innovative natural functional ingredients from microalgae. *Journal of Agricultural and Food Chemistry* **57**: 7159–7170.
DOI: <https://doi.org/10.1021/jf901070g>
- Radakovits R., Eduafo P.M. & Posewitz M.C. (2011). Genetic engineering of fatty acid chain length in *Phaeodactylum tricornerutum*. *Metabolic Engineering* **13**: 89–95.
DOI: <https://doi.org/10.1016/j.ymben.2010.10.003>
- Sahu A., Pancha I., Jain D., Paliwal C., Ghosh T., Patidar S., Bhattacharaya S. & Mishra S. (2013). Fatty acids as biomarker of microalgae. *Phytochemistry* **89**: 53–58.
DOI: <https://doi.org/10.1016/j.phytochem.2013.02.001>
- Samarakoon K. & Jeon Y. J. (2012). Bio-functionalities of proteins derived from marine algae-A review. *Food Research International* **48**: 948–960.
DOI: <https://doi.org/10.1016/j.foodres.2012.03.013>
- Samarakoon K.W., Ko J.Y., Lee J.H., Shah M.M.R., Kwon O.N., Kang M.C., Lee J.B. & Jeon Y.J. (2013). In vitro studies of anti-inflammatory and anticancer activities of organic solvent extracts from cultured marine microalgae. *Algae* **28**: 111–119.
DOI: <https://doi.org/10.4490/algae.2013.28.1.111>
- Samarakoon K.W., Ko J.Y., Lee J.H., Kwon O.N., Kim S.W. & Jeon Y. J. (2014). Apoptotic anticancer activity of novel fatty alcohol ester isolated from cultured marine diatom, *Phaeodactylum tricornerutum*. *Journal of Functional Foods* **6**: 231–240.
DOI: <https://doi.org/10.1016/j.jff.2013.10.011>
- Takahashi S., Yoshida M., Watanabe M.M. & Isoda H. (2019). Anti-inflammatory effects of *Aurantiochytrium limacinum* 4W-1b ethanol extract on murine macrophage RAW264 cells. *BioMed Research International* 2007: 3104057.
DOI: <https://doi.org/10.1155/2019/3104057>
- Veron B., Billard C., Dauguel J.C. & Hartmann M.A. (1996). Sterol composition of *Phaeodactylum tricornerutum* as influenced by growth temperature and light spectral quality. *Lipids* **31**: 989–994.
DOI: <https://doi.org/10.1007/BF02522694>
- Wadleigh D.J., Reddy S.T., Kopp E., Ghosh S. & Herschman H.R. (2000). Transcriptional activation of the cyclooxygenase-2 gene in endotoxin-treated RAW 264.7 macrophages. *The Journal of Biological Chemistry* **275**: 6259–6266.
DOI: <https://doi.org/10.1074/jbc.275.9.6259>
- Walne P.R. (1966). Experiment in the large scale culture of larvae of *Ostrea edulis*. L. *Fishery Investigation (London) Series 2*: 1–51.
- Wichard T.G. & Pohnert G. (2006). Formation of halogenated medium chain hydrocarbons by a lipooxygenase/hydroperoxide halolyase-mediated transformation in planktonic microalgae. *Journal of American Chemical Society* **128**: 7114–7115.
DOI: <https://doi.org/10.1021/ja057942u>
- Yang S. Wan H., Wang R. & Hao D. (2019). Sulfated polysaccharides from *Phaeodactylum tricornerutum*: isolation, structural characteristics, and inhibiting HepG2 growth activity in vitro. *PeerJ* **7**: 6409–6413.
DOI: <https://doi.org/10.7717/peerj.6409>

RESEARCH ARTICLE

Mechanical Engineering

Implementation of a fuzzy logic approach for a smart production system

D Dihovicni^{*}, M Mišćević², NR Kovačević¹ and D Kreculj¹

¹ College of Computer Mechanical Engineering, Academy of Applied Technical Sciences Belgrade, Bulevar Zorana Đinđića 152a, 11070 Belgrade, Serbia.

² Teximp d.o.o, Rajkova ulica 1.b, 11070 Belgrade, Serbia.

Submitted: 04 March 2021; Revised: 19 July 2022; Accepted: 24 July 2022

Abstract: This paper considers the monitoring of the entire process of an installation of a smart autonomous production system, and continuous analysis of its operation. The smart production system consists of an industrial robot and a computer numerical control (CNC) machine. Research is focused on the industrial robot subsystem; an implementation of a fuzzy logic method for end effectors is applied, and results regarding relevant clamping force are presented. The grippers and magnetic switches are described in more detail, and devices for measuring the clamping force are shown and analyzed. After installation and levelling of the machine, it is necessary to perform calibration – the correction of movement along the axis of the CNC machine. In theory, if the CNC machine is programmed to follow an ideal circular path, when the positioning accuracy of the machine and its tooling is perfect, then the actual curve that the machine produces, would correspond exactly to the ideally programmed circle. In practice, many factors affect CNC machine performance causing deviations from the ideal circuit. Factors influencing the accuracy are various, stemming from machine geometry, through the control system, to machine wear over many years of the operation. A possible upgrade of the existing system, in order to increase the autonomy, and reduce the need for a human operator, would include the installation of a 3D camera-sensor. The robot would take the correct pieces in accordance with production or assembly process, from the box next to it. The box is being filled with preparations from a conveyor belt.

Keywords: Artificial intelligence, automation, fuzzy logic, industrial computer systems, Industry 4.0.

INTRODUCTION

Complex working conditions and the size of machines are the key challenges facing modern industrial plants today (Mikić *et al.*, 2022). Nowadays industrial production companies strive to make the transition from automation to digitization, and reach the requirements and exploit opportunities of Industry 4.0. Due to the necessity for modernization of the companies, reduction of production costs, increase of productivity, quality of service, improvement of working conditions, timely supply of material and parts, as well as finished products, it is crucial to develop methods for the achievement of these goals, and apply these appropriately in the real work environment (Günter *et al.*, 2006; Novaković *et al.*, 2021).

The aim of this paper is to examine and evaluate the possibility of successful installation of a smart production system of the latest generation, based on fuzzy logic implemented in the decision making, and to compare the results with the production systems of the previous generations. The research includes all the necessary conditions for the successful implementation of a flexible production cell in an automated production line in a modern factory. All conditions for the implementation of the Industry 4.0 are fulfilled, with the use of robots and

* Corresponding author (djdihovicni@atssb.edu.rs;  <https://orcid.org/0000-0003-0961-2540>)



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

intelligent control, as well as trends in their development and application.

A concrete example of the application of an intelligent autonomous system in the production of machine elements is explained. The HAAS CNC machine is chosen for an experiment (HAAS, USA) and the intelligent production is described in detail, with the flexible cell, robot and smart control.

Choosing the appropriate grippers, with respect to the grasping force, is presented and it is implemented by using an artificial intelligence technique – the SWARA method (Step-wise weight assessment ratio analysis).

In order to perform fusion of the computer numerical control (CNC) machine and industrial robot, it is necessary to:

- Install an experimental setup, consisting of a CNC machine, with work-piece plate - tilting table with two axes, the industrial robot as an active part of the clamping system and work-piece;
- Connect the devices into a network so that all can communicate with each other;
- Create a numeric control program for the HAAS CNC machine in computer aided manufacturing (CAM) software;
Program a robot and synchronize it with the CNC machine program;
- Test whether pieces produced by the CNC machine are within the given tolerances.

The basic challenge is to connect these subsystems and ensure their mutual M2M (machine to machine) communication in order to determine whether the CNC cell system works well as a whole. There are many disturbances the robot is subjected to while being in operation – the weight it lifts can vary, as well as the dimensions of the work-piece, the tools are changing too, and the travel path is different from one task to another (Owen *et al.*, 2006; Chen & Dong, 2013; Cen & Melkote, 2017; Gualtieri *et al.*, 2021).

After the service department of the company had completed the installation of the components on the customer's premises, it is necessary to perform leveling and calibration of the machine and the robot. The leveling of the CNC machine and the robot is performed with the help of a special digital measuring device. This measuring device has the ability to communicate to mobile devices and computers over the Bluetooth connection. Information on the position of the work-piece, plate, or some other reference part of the machine is sent in real

time to the mobile phone of the service technician who performs the adjustment (Dumas *et al.*, 2011; Tyapin *et al.*, 2014). That way, the service technician does not need to go to the device after checking each foot rate. To add to that, there is no need for another service technician who would constantly read the data during levelling (Lehmann *et al.*, 2012).

CRITERIA

The end effectors are the crucial part of the robot subsystem, which is the focus of this project, and the appropriate choice is obtained by using a fuzzy logic approach (Pérez Armayor & Díaz Batista, 2010; Dihovični & Škrbić, 2020).

In robotics, the end effectors are devices at the end of a robotic arm designed to work with the environment. The exact nature of this device is determined by the application of the robot (Wan *et al.*, 2015; Wu *et al.*, 2015).

By strict definition, which comes from serial robotic manipulators, the end effectors represent the last connection (or end) of the robot. In a broader sense, the end effectors can be viewed as part of a robot that communicates with the work environment (Owen *et al.*, 2006).

End effectors can consist of a gripper, or a tool. In general, there are four categories of robotic gripping (Wan *et al.*, 2015):

- Impactive - jaws, or claws that directly physically grip an object;
- Ingressive - needles, or buttons that physically penetrate the surface of an object (used to handle textiles, carbon fibers, and glass fibers);
- Astrictive - attractive forces applied to the surface of an object (vacuum, magnetic, or electro adhesion);
- Contiguous - requires direct contact (such as gluing, surface tension, or freezing).

Most often, a mechanical gripper can have two, three, or even five grippers (fingers). In general, the gripping mechanism is operated by grippers, or mechanical fingers. Two-finger clamping is commonly used for industrial robots that perform specific tasks in less complex applications (Martinov *et al.*, 2012).

The mechanism used in gripping with two-finger grips is characterized by:

- 1) the shape of the surface to be gripped;
- 2) the force required to grip an object.

The shape of the grips can be selected according to the shape of the object being manipulated. For example, if the robot is designed to lift a round object, the shape of the gripper may have a concave contour.

Although there are numerous forces acting on a body raised by a robotic arm, the main force is the force of friction. The surface of the grip can be made of soft material with a high coefficient of friction, so that the surface of the object is not damaged. The robotic gripper must withstand not only the weight of the object, but also the acceleration caused by the frequent displacement of the object.

To determine the force required to grasp an object, the following formula is used:

$$F = \frac{m(a + g)}{\mu n} \tag{1}$$

where:

- F is a force required to grasp an object;
- m is a mass of the grasped object;
- a is an acceleration of the object;
- μ is a coefficient of friction;
- n is a number of grippers;
- g is an acceleration of gravity.

The clamping force, in addition to acceleration, is also affected by the length of the grippers.

The diagram in Figure 1 displays how the clamping force is changing with respect to the length of the grippers.

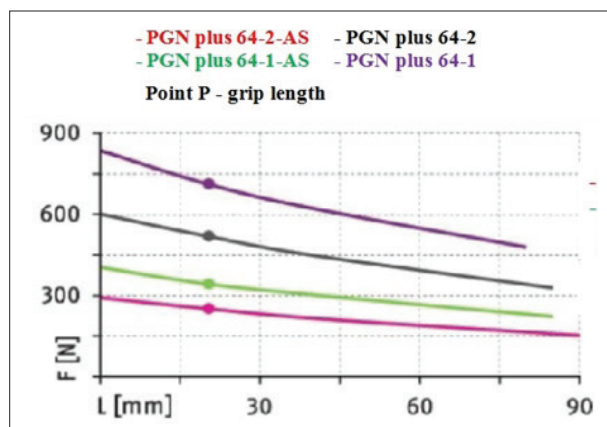


Figure 1: The diagram of the gripping force (F) with respect to the length of the grippers (L) (source: SCHUNK GmbH & Co. KG)

The grippers are pneumatically actuated. The opening and/or closing of the grippers is controlled with programmable magnetic switches SCHUNK MMS 22-PI1. A pneumatic nozzle between the grippers is installed, which is used for cleaning work-pieces and clamping accessories from sawdust and coolant.

DISCUSSION

Decision making in many areas can be very risky and difficult, affected by competing constraints and therefore cannot be individual. Multi Criteria Decision Making (MCDM) is a renowned methodology in decision making and evaluation. One of the famous MCDM tools is SWARA - Step-wise Weight Assessment Ratio Analysis method (Zolfani & Saparaukas, 2013).

Choosing end effectors can be enhanced by implementing the SWARA method in decision making. The objective of this study is to calculate relative weights of the average criteria by implementing the SWARA method (Merigó & Engemann, 2010; Zolfani & Saparaukas, 2013).

Based on their expected significances, the arithmetic mean values for the average clamping force in N are shown in the Table 1, sorted in descending order.

Table 1: The presentation of arithmetic mean of clamping force in [N] for different gripper types

Criteria – Gripper type	Mean clamping force [N]
PGN-plus 64-1	633
PGN-plus 64-2	450
PGN-plus 64-1-AS	322
PGN-plus 64-2-AS	225

Table 2 displays the relative importance of the mean criteria for average clamping force. The relative importance of the mean criteria j , denoted as \tilde{x}_j to previous criteria ($j-1$), is given in the Table 2.

Solving the weight q of the j -th criteria, denoted by \tilde{q}_j , is given by Equation 2 (Keropyan & Gil-Lafuente, 2010):

$$\tilde{q}_j = \begin{cases} 1, & j = 1 \\ \frac{\tilde{x}_{j-1}}{\tilde{k}_j}, & j > 1 \end{cases} \tag{2}$$

Table 2: The relative importance of the mean criteria for average clamping force

Criteria – Gripper type	\tilde{s}_j
PGN-plus 64-1	
PGN-plus 64-2	0.710
PGN-plus 64-1-AS	0.705
PGN-plus 64-2-AS	0.698

The relative weight of the evaluation criteria j is determined in Equation 3 (Nicolas, 2004; Pérez Armayor & Díaz Batista, 2010), and is given in the rightmost column of the Table 3:

$$\tilde{w}_j = \frac{\tilde{q}_j}{\sum_{k=1}^n \tilde{q}_k} \tag{3}$$

Table 3: The relative weight of the evaluation criteria

Criteria – Gripper type	\tilde{s}_j	\tilde{k}_j	\tilde{q}_j	\tilde{w}_j
PGN-plus 64-1		1	1	0.470
PGN-plus 64-2	0.710	1.710	0.584	0.274
PGN-plus 64-1-AS	0.705	1.705	0.342	0.160
PGN-plus 64-2-AS	0.698	1.698	0.201	0.096
			2.127	1.000

The SWARA method is often applied for two reasons. Firstly, SWARA’s perspective is different from other similar methods – this method gives the chance to decision makers to select their priority based on the current situation of environment and economy. Secondly, the role of the experts is very important in the SWARA method. Experts have a key role in process of decision making on very important projects. Finally, SWARA has the advantage of more logical calculation of weights and relative importance of criteria (Zolfani & Sapauskas, 2013).

The summary of resulting weight of the mean criteria for the clamping force for all of the four selected types of grippers is shown in Figure 2.

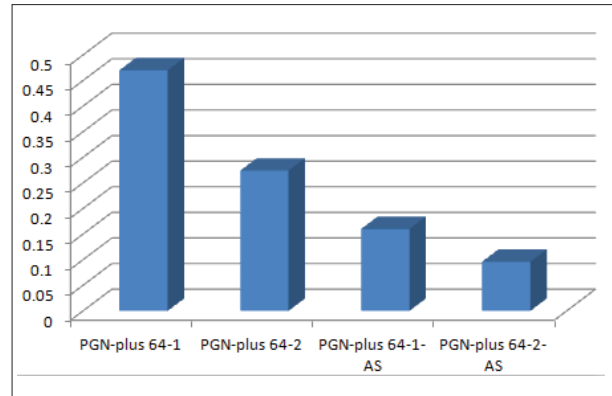


Figure 2: Graphical representation of resulting weight of the mean criteria for average clamping force

Two universal parallel grippers such as SCHUNK PGN-plus 64-1 grippers are chosen as the end effectors. The maximum force of external clamping of the work-piece is 135 N, and the step is 4 mm per grip.

The grippers are pneumatically actuated, where the control of opening and closing of the grip is done via programmable SCHUNK MMS 22-PI1 magnetic switches. A pneumatic nozzle is installed between the grippers, which is used to clean the work-pieces and clamping accessories from sawdust and coolant (Ma et al., 2007).

Two universal parallel grippers, with two grippers SCHUNK PGN-plus 64-1, in Figure 3 are presented. The grippers are made of aluminum bars having a profile that corresponds to the shape of the work piece, and the profiles are oriented towards the gripping position. The grippers are manufactured by the custom design and machined on a CNC machine on the customer’s premises, in the customer’s plant.

The solution with two grippers placed perpendicularly on the fork, at an angle of 90° to each other, is chosen (as shown in Figure 3) in order to save time. While the robot brings the preparation in one gripper, with the other it takes and lifts-up the finished work-piece. After that the robot pulls the work-piece out of the clamping accessories and rotates the last joint by 180° to swap the places of both grippers. This way, the double stroke of the robot arm from the clamping accessories to the pallet is eliminated (Ezugwu et al., 2009).

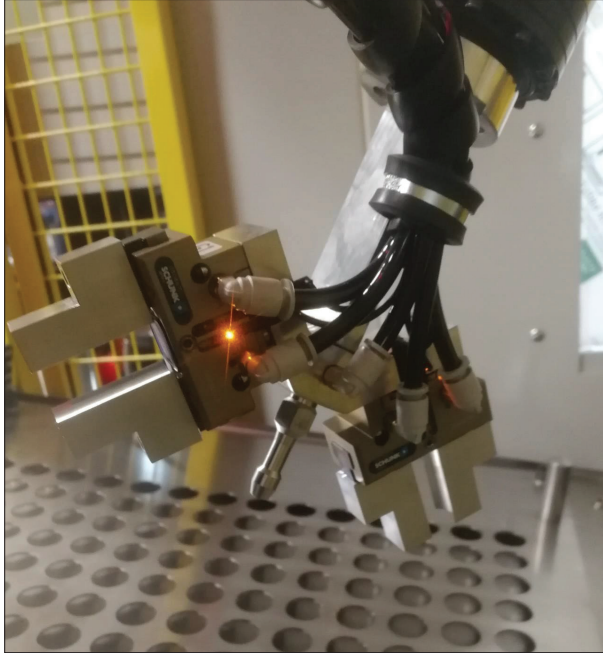


Figure 3: Parallel grippers with two grippers SCHUNK PGN-plus 64-1

The programmable magnetic switches SCHUNK MMS 22-PI1 has the function of sending the information to the robot in order to close the feedback.

The sensor is mounted on the body of the catcher, and connected with the channels provided. The sensors can also be programmed using special programming tools (plug-in or magnetic). The sensors have LED lights that send signals to operator by varying the number and frequency of flashes. The sensors also send the information to the robot via cable.

In case of any physical problem while catching the work piece (jamming, incorrectly placed work piece, etc.), the robot receives a signal to stop further execution, and signals the stop by turning on the red light on its signal traffic light, as well as by sending a message with a description of the error to the control panel. If the light signal on the robot's traffic light fails, the CNC machine recognizes that the waiting time is longer than the usual cycle (the allowed value of time delay expressed in seconds previously entered in the programme).

After reporting an error on its control unit and turning on its traffic light, the CNC machine sends an

SMS (Short Message Service) with a description of the problem to the mobile phone of the responsible person, and a message to the predefined e-mail addresses. All this is provided via the HaasConnect application. The ability to equip grippers of the robotic arm of the CNC machine with inductive sensor systems, magnetic switches, or work-piece-specific grips, improves their functionality, while increasing both the reliability and productivity of the process. The quality control of machining of the parts (with measuring and checking dimensions) is automated and performed using WIPS – Wireless Intuitive Probing System. A WIPS measuring probe manufactured by RENISHAW (RENISHAW mark - OMP40) is delivered with the CNC machine, as displayed in Figure 4. This probe enables fast and precise calculation (within the accuracy of the CNC machine) of the machined work piece on the CNC machine.



Figure 4: Measuring probe RENISHAW OMP40

The CNC machine manufacturer HAAS offers a complete system with WIPS equipment as a standard option. Considering the growing demands regarding precision and reduction of cycle time, as well as reduction of scrap, it is almost impossible to use a machining centre without this device.

Programming of a CNC machine for a cycle of inspection and quality control is done directly on the machine's control unit. The WIPS probe communicates

with the CNC machine over transceiver located in the workspace in the vicinity of the machine using an infrared (IR) radiation connection (Figure 6). The probe can be programmed to find the zero point of the work piece, measure the dimensions of the work piece, touch the holes on the piece, tactilely scan the work piece, etc.

Communication between the measuring device and the CNC machine is shown in Figure 5.

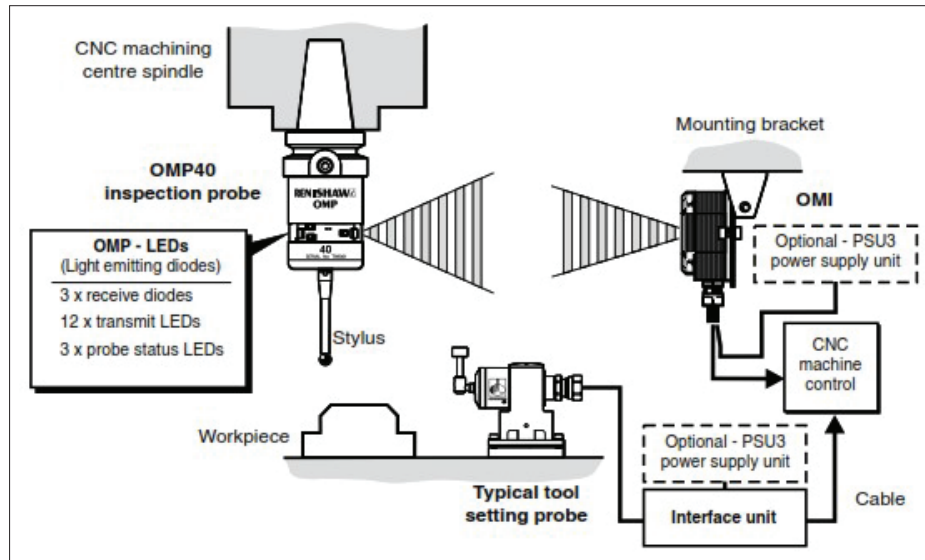


Figure 5: Communication of the measuring device with the CNC machine

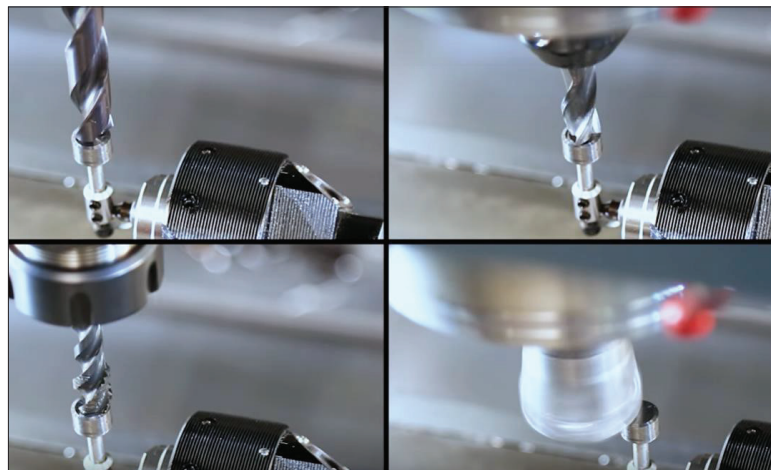


Figure 6: Measuring the tool dimension on the CNC machine

When the tool sizing cycle is started, the tool itself approaches the hard metal cylinder located on the top of the device and after touching the tip of the probe sends a signal to the machine control unit (Figure 6). At that moment, the machine compares the difference in the lengths of the axis of the movement and the previously calibrated (known) length at which the tip of the probe is located. Thus, the exact length of the tool is obtained, which is measured from the tip of the tool to the front of the spindle, and the machine automatically saves the measurement – it enters value in the tool table.

The same process is repeated along the second axis in case it is necessary to measure the diameter of the tool. Then the tip of the probe approaches the cylinder of the tool sideways, with the spindle rotating the tool in the opposite direction then while performing cutting or turning.

Measuring the tool dimension on the CNC machine is presented in Figure 6.

This measuring device is very useful in series production, because it has the ability to detect tool breakage. The tool check cycle and repetition interval are set on the machine control unit in the tool table. The principle of the operation is such that the machine has a predefined allowable path length that can be travelled when the tool approaches the measuring probe. If there is no contact between the tool tip and the measuring probe within this range, the machine signals a fault on the control unit. The cycle can be repeated periodically in predefined equal time intervals (after the tool has been in the grip for a certain period of time), or after each change of the tool.

The use of automated measuring devices for the machining quality control and inspection process reduces significantly the occurrence of scrap in automated production, where the human presence is not perpetual. The initial investment in systems of this type returns rather quickly through the decrease of the preparation time, which includes adjusting the position of the part or preparation and defining the dimensions of the cutting tools.

The classic calibration methods last approximately 11 minutes for the preparation, with 8 tools stacked to be ready to work. This system reduced that time to less than 2 minutes.

A possible plan for future work is to upgrade the existing system, described here, by installation of a 3D camera-sensor, which would provide an even higher level of autonomy and eliminate the need for a worker who serves the machine. Successful implementations of artificial vision systems exist (Roberti *et al.*, 2009). Instead of a human worker who places preparations on the pallet matrix and transports the finished pieces the robot would take the pieces from the box next to it, which would be filled with preparations by a conveyor belt. The robot would also place the finished pieces on another conveyor belt that would take these to the next operation (inspection and packaging). If this is the case, it would not be necessary to place the workpieces in the proper position on the pallet. This approach would be a significant improvement towards Industry 4.0.

In recent years fuzzy logic has provided advances in the design and control of hybrid intelligent systems based on nature-inspired optimization. Applications of fuzzy and hybrid intelligent systems in areas such as intelligent control and robotics, pattern recognition, medical diagnosis, time series prediction, and optimization of complex problems have been explored (Melin *et al.*, 2018).

Dumitrescu *et al.*, (2021) introduced a concept of distributed intelligent control (DIC) and applied fuzzy logic to the control of an autonomous robot (vehicle). From the point of view of the DIC concept, three types of components can be defined and fuzzified (Dumitrescu *et al.*, 2021):

- Fuzzy sensors, which provide a representation of measurements as fuzzy subsets;
- Fuzzy actuators, which can act and operate in a real world; depending on the fuzzy subsets they receive;
- Fuzzy components of inference, also received by the fuzzy actuators, which can perform distributed fuzzy logic.

Fuzzy components can be integrated for different applications into a compact model of components, called the fuzzy cell. The configuration of the cells in the network is done through a specific programming language (Dumitrescu *et al.*, 2021).

Increasing the variety of products that market needs impose on manufacturers requires high adaptability of production systems that can be achieved through the introduction of flexible production systems composed of

interoperable devices with variable architecture. Control and management of such a complex system requires fast and reliable real-time virtualization of real-world applications, as well as real-time feedback from real-world virtual (cyber) models (Wan *et al.*, 2015). The boundary line between the real production system and its cyber-representation is characterized by extremely high information throughput, combining these two systems into a single system, the Cyber-Physical Manufacturing System (CPMS). Recent advances in the fields of Cyber-Physical Systems (CPS) and the Internet of Things (IoT) have enabled the creation of CPMS.

CONCLUSION

In this study the SWARA method was applied on given data for the clamping force, obtained from the SCHUNK manufacturer. For the chosen criteria such as clamping force, the resulting weight was calculated. The degree of importance is presented in the appropriate tables, as well as in graphs. Taking into account the relative weight, the SCHUNK PGN-plus 64-1 grippers are chosen for the end effectors out of group of four types of the grippers.

In order to achieve adequate grip and capture of the object, the positioning and orientation of the grippers play a very important role. The industrial robot must achieve the precise relative positioning and orientation of the catcher with respect to the object, which in the case of hierarchical intelligent control can be placed in an arbitrary position.

Movement of an autonomous industrial robot in accordance with the recognized object, initiates many difficulties during completion of other defined phases of object manipulation. Thanks to visual feedback and learning based on artificial neural network systems, the robot catcher can determine and take a proper position and orientation with respect to the recognized object. This can be obtained in different ways, primarily by using systems with adaptive flexibility.

Smart behavior, which involves an autonomous industrial robot, is provided through adequate coordination between visual information about the object and surrounding space obtained from the sensor-camera and the control system. The smart control of the robot, realized by using sensor-cameras and artificial neural networks, has proven to be effective in overcoming the problems of modeling and the emergence of uncertainty in the real environment (*e.g.*, changing the position of objects).

The successor of the contemporary CNC machine could be called the ICNC machine (intelligent computer numerical control). In the future, there will be more flexible production cells controlled by artificial intelligence which will be able to 'learn', and the human presence will be decreased further and reduced only to assigning tasks.

Two decades ago, there were less than 700,000 industrial robots worldwide; however nowadays the number is over 3 million. In developed countries, there has been a simultaneous increase in production and a decrease in employment, and one, but not the only, cause is the development of technology. Today, 3D printing is used in the production of cars and airplanes, biotechnology is changing the way we grow crops and produce food, and the further development of nanotechnology and artificial intelligence will greatly reshape numerous economic activities, probably faster than we expect.

Currently the need for unskilled laborers whose work can be automated shows a downward trend and this is expected to continue, while the need for experts having skills to develop, launch and maintain highly automated systems is growing year by year.

REFERENCES

- Cen L. & Melkote S.N. (2017). Effect of robot dynamics on the machining forces in robotic milling. *Proceedings of the 45th SME North American Manufacturing Research Conference (NAMRC 45)* (eds. L. Wang, L. Fratini & A.J. Shih), Volume 10, Los Angeles, USA, 4-8 June. University of Southern California, California, USA. pp. 486–496. DOI: <https://doi.org/10.1016/j.promfg.2017.07.034>
- Chen Y. & Dong F. (2013). Robot machining: recent development and future research issues. *The International Journal of Advanced Manufacturing Technology* **66**: 1489–1497. DOI: <https://doi.org/10.1007/s00170-012-4433-4>
- Dihovicki Dj. & Škrbić S. (2020). Fuzzy approach to supply chain management for e-commerce store. *Applied Engineering Letters* **5**(2): 62–67. DOI: <https://doi.org/10.18485/aeletters.2020.5.2.4>
- Dumas C., Caro S., Garnier S. & Furet B. (2011). Joint stiffness identification of six-revolute industrial serial robots. *Robotics and Computer-Integrated Manufacturing* **27**(4): 881–888. DOI: <https://doi.org/10.1016/j.rcim.2011.02.003>
- Dumitrescu C., Ciotirnae P. & Vizitiu C. (2021). Fuzzy logic for intelligent control system using soft computing applications. *Sensors* **21**(8):2617. DOI: <https://doi.org/10.3390/s21082617>

- Ezugwu E.O., Sales W.F. & Landre Jr J. (2009). Machining dynamics in turning processes. In: *Machining Dynamics: Fundamentals, Applications and Practices* (ed. K. Cheng), pp. 151–166. Springer, Berlin, Germany.
DOI: https://doi.org/10.1007/978-1-84628-368-0_6
- Gualtieri L., Rauch E. & Vidoni R. (2021). Emerging research fields in safety and ergonomics in industrial collaborative robotics: a systematic literature review. *Robotics and Computer-Integrated Manufacturing* **67**: 101998.
DOI: <https://doi.org/10.1016/j.rcim.2020.101998>
- Günter H., Grote G. & Thees O. (2006). Information technology in supply networks: Does it lead to better collaborative planning?. *Journal of Enterprise Information Management* **19**(5): 540–550.
DOI: <https://doi.org/10.1108/17410390610703666>
- Keropyan A. & Gil-Lafuente A.M. (2010). Decision making in strategic business management, In: World Scientific Proceedings Series on Computer Engineering and Information Science. *Proceedings of the MS'10th International Conference* (eds. A. Gil-Lafuente & J. Merigó), volume 3, Barcelona, Spain, 15-17 July. University of Barcelona, Spain, pp. 297–304.
DOI: https://doi.org/10.1142/9789814324441_0037
- Lehmann C., Halbauer M., Euhus D. & Overbeck D. (2012). Milling with industrial robots: Strategies to reduce and compensate process force induced accuracy influences. In: *Proceedings of 2012 IEEE 17th International Conference on Emerging Technologies & Factory Automation (ETFA 2012)*, 17-21 September. Krakow, Poland, pp. 1–4.
DOI: <https://doi.org/10.1109/ETFA.2012.6489741>
- Ma X.B., Han Z.Y., Wang Y.Z. & Fu H.Y. (2007). Development of a PC-based Open Architecture Software-CNC System. *Chinese Journal of Aeronautics* **20**(3): 272–281.
DOI: [https://doi.org/10.1016/S1000-9361\(07\)60044-2](https://doi.org/10.1016/S1000-9361(07)60044-2)
- Martinov G.M., Ljubimov A.B., Grigoriev A.S. & Martinova L. (2012). Multifunction numerical control solution for hybrid mechanic and laser machine tool. *Procedia CIRP* **1**: 260–264.
DOI: <https://doi.org/10.1016/j.procir.2012.04.047>
- Melin P., Castillo O., Kacprzyk J., Reformat M. & Melek W.W. (2018). *Fuzzy Logic in Intelligent System Design: Theory and Applications*, Springer, Switzerland.
DOI: [10.1007/978-3-319-67137-6](https://doi.org/10.1007/978-3-319-67137-6)
- Merigó J.M. & Engemann K.J. (2010). Fuzzy decision making with probabilities and induced aggregation operators. In: World Scientific Proceedings Series on Computer Engineering and Information Science. *Proceedings of the MS'10th International Conference* (eds. A. Gil-Lafuente & J. Merigó), volume 3, 15-17 July. University of Barcelona, Spain, pp. 323–332.
DOI: https://doi.org/10.1142/9789814324441_0040
- Mikić D., Desnica E., Kiss I. & Mikić V. (2022). Reliability analysis of rolling ball bearings considering the bearing radial clearance and operating temperature. *Advanced Engineering Letters* **1**(1): 16–22.
DOI: <https://doi.org/10.46793/adeletters.2022.1.1.3>
- Nicolas R. (2004). Knowledge management impacts on decision making process. *Journal of Knowledge Management* **8**(1): 20–31.
DOI: <https://doi.org/10.1108/13673270410523880>
- Novaković B., Desnica E., Radovanović Lj., Ivetić R., Đorđević L. & Labović Vukić D. (2021). Optimization of industrial fan system using methods laser alignment. *Applied Engineering Letters* **6**(2): 62–68.
DOI: <https://doi.org/10.18485/aeletters.2021.6.2.3>
- Owen W.S., Croft E.A. & Benhabib B. (2006). Real-time trajectory resolution for a two-manipulator machining system. *Supplemental Issue of Journal of Field Robotics* **22**(S1): S51–S63.
DOI: <https://doi.org/10.1002/ROB.V22:S1>
- Pérez Armayor D. & Díaz Batista J.A. (2010). Decision support model for information systems integration in supply chains: A fuzzy approach. In: World Scientific Proceedings Series on Computer Engineering and Information Science. *Proceedings of the MS'10th International Conference* (eds. A. Gil-Lafuente & J. Merigó), volume 3, 15-17 July. University of Barcelona, Spain, pp. 281–288.
DOI: https://doi.org/10.1142/9789814324441_0035
- Roberti F., Toibero J.M., Soria C., Vassallo R.F. & Carelli R. (2009). Hybrid collaborative stereo vision system for mobile robots formation. *International Journal of Advanced Robotic Systems* **6**(4): 7241.
- Tyapin I., Hovland G., Kosonen P. & Linna T. (2014). Identification of a static tool force model for robotic face milling. *Proceedings of 10th International Conference on Mechatronic and Embedded Systems and Applications (MESA 2014)*, IEEE/ASME, 10-12 September. Senigallia, Italy, pp.1–6.
DOI: <https://doi.org/10.1109/MESA.2014.6935591>
- Wan J., Cai H. & Zhou K. (2015). Industrie 4.0: enabling technologies. In *Proceedings of 2015 International Conference on Intelligent Computing and Internet of Things (2015 ICIT)*, 17 – 18 January. IEEE, Harbin, China, pp. 135-140.
DOI: <https://doi.org/10.1109/ICAOT.2015.7111555>
- Wu K., Brueninghaus J., Johnen B. & Kuhlenkoetter B. (2015). Applicability of stereo high speed camera systems for robot dynamics analysis. In: *Proceedings of the 2015 International Conference on Control, Automation and Robotics*, 20 – 22 May. IEEE, Singapore, pp. 44–48.
DOI: <https://doi.org/10.1109/ICCAR.2015.7165999>
- Zolfani S.H. & Sapauskas J. (2013). New application of SWARA method in prioritizing sustainability assessment indicators of energy system. *Engineering Economics* **24**(5): 408-414.
DOI: <http://dx.doi.org/10.5755/j01.ee.24.5.4526>

RESEARCH ARTICLE

Biomedical Engineering

Non invasive automated approach for eczema lesions segmentation using colour space normalization

H Nisar*, Y-K Ch'ng and KH Yeap

Department of Electronic Engineering, Faculty of Engineering and Green Technology, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia.

Submitted: 09 February 2021; Revised: 17 September 2021; Accepted: 21 October 2021


Abstract: Eczema is a common type of atopic dermatitis. Eczema skin lesions can be identified visually by observing the difference between the colour and texture of the lesions and the normal skin. Dermatologists assess eczema by direct visual assessment and record their observations in specialized forms. These methods are tedious as well as time consuming and introduce inter-rater and intra-rater variability in the results. To make the assessments objective and easier for the dermatologists, we have proposed a framework for the segmentation of eczema skin lesions. Red-Green-Blue (RGB), CIELab and their normalized colour spaces were considered. For segmentation, a two step K-means algorithm was proposed. In the 1st step, a conventional K-means algorithm segments the image into three regions, *i.e.*, skin, lesion, and mixed region. This is followed by a 2nd K-means segmentation step. The performance of this method was better than the conventional methods. The algorithm was evaluated using 85 eczema images of different severity and grades. To assess the performance of the algorithm, the gold standard segmentation for eczema lesions was manually drawn and verified by a dermatologist. The Green-channel of normalized (CSN-I) RGB colour space provided the best result for a semi-supervised approach giving the segmentation accuracy of 88.28% whereas for fully automated approach a segmentation accuracy of 84.43% was achieved using support vector machine (SVM).

Keywords: Atopic dermatitis, K-means segmentation, selection modes, skin lesions.

INTRODUCTION

Atopic eczema is a chronic, relapsing and inflammatory skin condition that is widespread in all age groups and may cause social and psychological problems (Carroll *et al.*, 2005). Measuring the severity of a skin disease is the main objective in dermatology. It is essential to assess the course of the disease before, during and after the treatment. In skin diseases like atopic eczema, there are no serologic markers to identify the severity and extent of the disease. Hence evaluation of the disease is based on the visual signs and symptoms (Finlay, 1996). The general symptoms of atopic eczema are dry reddish skin with signs of inflammation. Due to scratching, the skin patches may be thickened, blistered, and sometimes infected. The different symptoms on eczema plaques are: 'erythema, edema/swelling, oozing/crusting, excoriation, and lichenification' (HOME, 2014). Subjective symptoms are itching and loss of sleep which also includes difficulty in falling asleep and/or waking up at night. Eczema is commonly found at the front side of elbows and wrists, around the neck, and back of the knees.

The prevalence of eczema appears to vary across the world as observed in key international studies (Lee &

* Corresponding author (humaira@utar.edu.my;  <https://orcid.org/0000-0003-2026-5666>)



Detzel, 2015). The latest data based on the phase three of the ‘International Study of Asthma and Allergies in Childhood’ (ISAAC) (Mallol *et al.*, 2013) showed that atopic dermatitis (AD) has reached a plateau in countries like the UK and New Zealand, but it continues to increase in prevalence, specifically in younger children (age 6–7 years) in low-income countries, in regions such as South East Asia and Latin America. This skin condition affects about 31.6 million people in the United States, and one in ten individuals will develop eczema during their lifetime (Silverberg & Hanifin, 2013). In the Asia Pacific, the 12-month prevalence of AD in children (age 13–14 years) was reported to be as high as 9% in Malaysia and Singapore, and as low as 0.9% in China, which has the lowest prevalence in the world. The reasons for the differences in prevalence are not fully understood, but industrialization and socio-economic factors have been implicated (Lee & Detzel, 2015).

The main objective of measuring the severity level of eczema is for assessing the progression of the disease in patients. The quantitative scales used for measuring eczema severity are difficult to interpret in clinical practice. Doctors use systems like ‘Eczema Area and Severity Index’ (EASI) (Barbier *et al.*, 2004), and ‘Patient Oriented Eczema Measure’ (POEM) (Charman *et al.*, 2002) to assess the severity of eczema. The process

of manual form filling is time consuming, and involves subjective assessment that results in inter-rater and intra-rater variability. Therefore, the objective of this research is to develop an automated eczema lesion segmentation system based on the EASI index, for assessment and grading of atopic eczema. The symptoms and severity levels for EASI scoring scheme are shown in Figure 1.

A comprehensive literature review of eczema has shown that very little research has been done for the development of a non-invasive, automated, objective method for the assessment and grading of the severity of eczema. A study conducted by Tremp *et al.* (2011) proposed the use of digital imaging techniques to assess the severity of atopic eczema using the EASI scoring system. They called their method EASIdig. In this method digital camera is used to take images of the patients; these images were evaluated by general physicians trained in EASI and SCORAD scoring systems. It took about thirty minutes to calculate the score using EASIdig. Another study reported by Alam *et al.* (2016) used 85 images to develop an automated approach for eczema detection and severity measurement. K-means clustering, and morphological image processing were used to segment eczema lesions using CIElab colour space. For classification ten colour and four texture features were used, reporting an accuracy of 90% using SVM classifier.



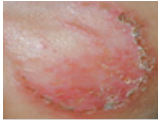













Severity level		Symptoms			
		None (score 0)	Mild (score 1)	Moderate (score 2)	Severe (score 3)
Symptoms	Redness				
	Thickness				
	Lichenification				
	Scratching				

Figure 1: Symptoms and severity levels used for assessment in EASI scoring scheme (HOME, 2014)

Based on a literature review it is observed that other skin conditions such as Psoriasis, acne vulgaris, and melanoma have been studied extensively. The research is normally conducted for the classification of skin lesions originating from different diseases. For example, Hegde *et al.* (2018) used RGB colour and gray level co-occurrence matrix (GLCM) texture features to classify different types of skin lesions, *i.e.*, chronic eczema, lichen planus, and plaque psoriasis. The machine learning methods SVM, naïve bayesian classifier (NBC), artificial neural networks (ANN), and linear discriminant analysis (LDA) were used. The classification accuracy using SVM classifier was 81.61%. Hameed *et al.* (2020) have proposed a framework for classification of healthy, benign, and malignant melanoma, and eczema images. A classification accuracy of 96.47% was achieved using colour and texture features.

Recently neural networks were investigated for melanoma lesion segmentation. Zafar *et al.* (2020) segmented dermoscopic images of melanoma using convolutional neural networks (CNN). They have combined UNet and ResNet architectures to form a novel architecture known as Res-Unet. A Jaccard Index of 77.2% was obtained. Abraham and Khan (2018) proposed U-Net and attention U-Net for segmentation of melanoma lesions. They achieved a dice score of 0.838 by U-Net and 0.856 by attention U-Net. The above literature review shows that tremendous research has been carried out for the classification of diseases where lesions are significantly different from each other in terms of surface properties, size, and shape.

In this paper our aim is to segment eczema skin lesions having the characteristic feature of no specific size and shape with a fuzzy boundary. The shape, size and colour of eczema lesions are not well defined; it can take any shape, spread to any size, and its colour varies a lot with the increase in the severity level. Some lesions are concentrated in a smaller area whereas others are scattered in a larger area. This makes the segmentation and classification of eczema lesions extremely difficult.

In the following sub-sections, we will give a brief overview of the pre-processing methods, colour spaces, segmentation, and classification techniques that we found useful for eczema lesion segmentation.

Image pre-processing

Image pre-processing removes noise and the effect of irregular illumination from an image. Ambient light condition is a very important criterion for image

acquisition. The image segmentation may be affected by over-exposure or under-exposure of light. Two illumination correction techniques are found suitable for eczema lesion image illumination correction; these are adaptive light compensation (ALC) (Ch'ng *et al.*, 2014a) and Frankle-McCann retinex (retinex) (Funt *et al.*, 2004). In ALC the value of the luminance denoted as Y is extracted from the input image and average luminance (Y_{avg}) is computed. The maximum value of 200 and minimum value of 80 for good luminance exposure is obtained empirically from the images. A luminance factor is computed for RGB image correction as given in equation (1).

$$Factor = \begin{cases} 200 / Y_{avg} ; & \text{if } Y_{avg} > 200 \\ 1 ; & \text{if } 80 < Y_{avg} < 200 \\ 80 / Y_{avg} ; & \text{if } Y_{avg} < 80 \end{cases} \quad \dots(1)$$

The retinex method is designed to correct the illumination of under-exposed (dark) areas of an image while maintaining the visual discrimination of lighter areas (Funt *et al.*, 2004).

Colour spaces

Colour is an important property in visual perception. Different colour spaces are useful for different applications. A preliminary study with four colour spaces (HSI, YCbCr, CMY, CIE Lab) for eczema lesion segmentation using k-means showed that hue (H) channel of HSI has the highest segmentation accuracy of 76.63% (Nisar *et al.*, 2013).

Yang *et al.* (2010) proposed a colour space normalization (CSN) technique to improve the face recognition rate. His analyses of different colour spaces showed that those colour spaces perform better when the sum of elements in the 2nd and 3rd row of transformation matrix are zero. The RGB colour space does not possess this property; however, the YUV colour space has this property as shown in equations (2) and (3).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \dots(2)$$

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1471 & -0.2888 & 0.4359 \\ 0.6148 & -0.5148 & -0.1000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \dots(3)$$

The CSN technique can normalize any colour space and the normalized colour space has a transformation matrix in which the sum of elements in the 2nd and 3rd rows are zero. Colour space normalization-I (CSN-I) is also known as ‘within-colour-component normalization.’ It is obtained by removing the means of 2nd and 3rd row vectors respectively. As a result, the sum of the row vectors is equal to zero in the transformation matrix. Equations (4) and (5) (Yang *et al.*, 2010) show the transformation matrix for the CSN-I RGB, \tilde{A}_{RGB-I} and CSN-I CIE-XYZ, \tilde{A}_{XYZ-I} colour spaces, respectively. Colour space normalization-II (CSN-II) is known as ‘across-colour-component normalization.’ In this technique the zero-mean vector of the 2nd and 3rd rows is obtained via linear combination. Equations (6) and (7) (Yang *et al.*, 2010) are the transformation matrices obtained for CSN-II RGB, \tilde{A}_{RGB-II} and CSN-II CIE-XYZ, \tilde{A}_{XYZ-II} colour spaces, respectively. Figure 2 gives an overview of different colour components.

$$\tilde{A}_{RGB-I} = \begin{bmatrix} 1 & 0 & 0 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix} \quad \dots(4)$$

$$\tilde{A}_{XYZ-I} = \begin{bmatrix} 0.6070 & 0.1740 & 0.2 \\ -0.0343 & 0.2537 & -0.2193 \\ -0.3940 & -0.3280 & 0.7220 \end{bmatrix} \quad \dots(5)$$

$$\tilde{A}_{RGB-II} = \begin{bmatrix} 1 & 0 & 0 \\ -0.5774 & 0.7887 & -0.2113 \\ -0.5774 & -0.2113 & 2/3 \end{bmatrix} \quad \dots(6)$$

$$\tilde{A}_{XYZ-II} = \begin{bmatrix} 0.6070 & 0.1740 & 0.2 \\ -0.0901 & 0.3631 & -0.2730 \\ -0.4600 & -0.1986 & 0.6586 \end{bmatrix} \quad \dots(7)$$

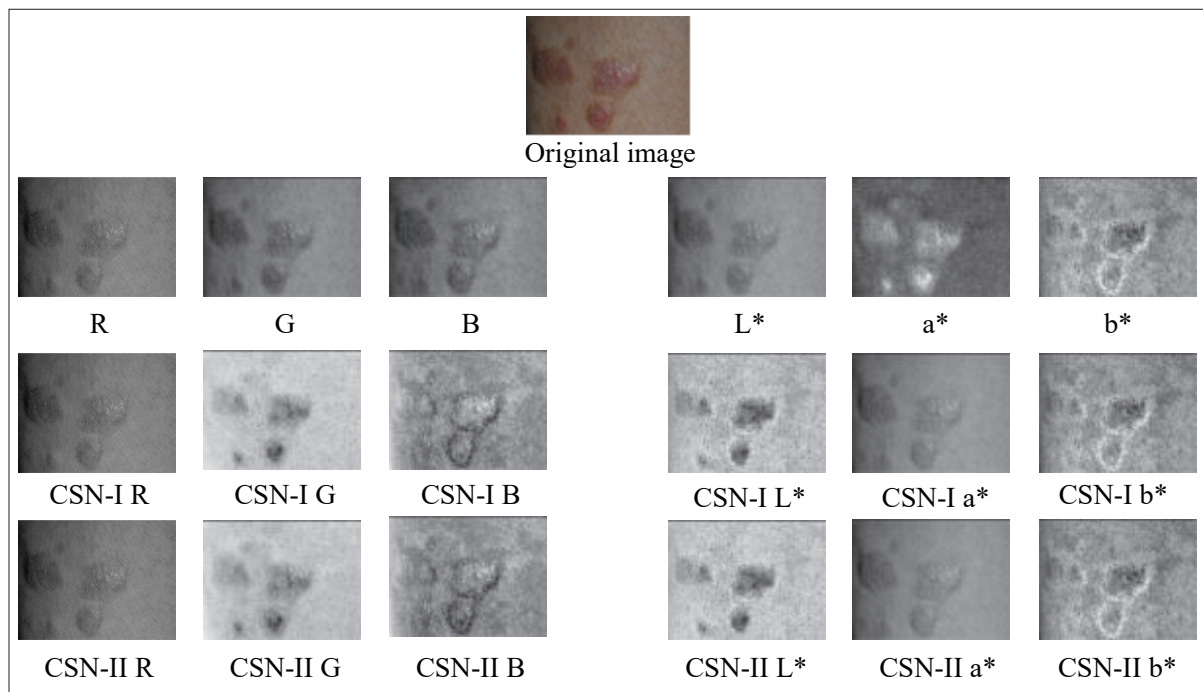


Figure 2: Illustration of different colour components for RGB, CIE Lab and their normalized CSN-I and CSN-II colour spaces

Segmentation and classification methods

In image processing, clustering methods are commonly used for segmentation of the desired regions of interest (ROI). The classification methods are divided into two categories: unsupervised learning and supervised

learning classification methods. Unsupervised learning classifiers do not require any reference data for carrying out the classification. The classifier groups the data to a class itself. Clustering requires the user to define the hypothesized number of clusters. Commonly used methods of clustering are K-means clustering (Hartigan

& Wong, 1979) and Fuzzy C-mean (FCM) (Bezdek *et al.*, 1984). However, supervised learning classifiers require a set of training data to work as a reference for classification. Supervised learning classification groups similar data based on the training data groups. Commonly used approaches are region-based, energy-based, and region and contour-based approaches. There are also some other classifiers such as SVM (Hegde *et al.*, 2018).

MATERIALS AND METHODS

The first step in any image processing-based application is image acquisition. It is followed by image pre-processing for image enhancement and noise removal. This is followed by image segmentation to obtain the

ROI. In this paper we have proposed an algorithm to segment eczema skin lesions. We will discuss all the steps in the algorithm in the following sub-sections.

Data acquisition

The images were acquired using a digital single lens reflex (DSLR) camera. A dataset of 85 images in JPEG format is available for this study, the images were voluntarily provided. The gold standard for the skin lesions was manually drawn using the GNU Image Manipulation Program (GIMP) (The GIMP Team, 2014); and verified by a dermatologist. For performance assessment of segmentation, the gold standard images were used.

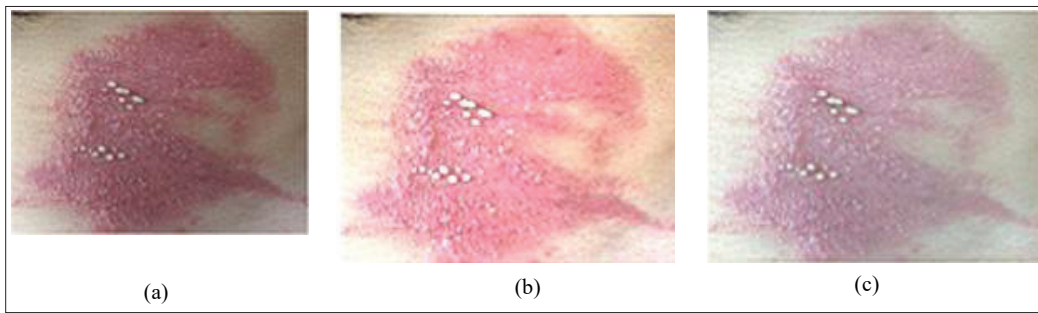


Figure 3: Comparison of pre-processing algorithms (a) input image; (b) adaptive light compensation (ALC) and (c) Frankle-McCann retinex (retinex)

Image pre-processing

Colour invariance due to uneven illumination makes the image segmentation more challenging. Hence, an ALC method (Ch'ng *et al.*, 2014a) was used to pre-process the images. We have also used another pre-processing method, Frankle-McCann retinex (Funt *et al.*, 2004) to compare the performance of pre-processing as illustrated in Figure 3.

Proposed segmentation approach

The K-means segmentation method is commonly used in unsupervised segmentation. We have proposed a bi-level K-means segmentation that divides the segmentation process into two steps (Ch'ng *et al.*, 2014b). The proposed technique is shown in in Figure 4. In the first

step, it divides the image into three regions: the skin region, the mixed region, and the lesion region. Then the intensity feature [mean (μ) \pm standard deviation (SD)], of skin or lesion region can be used as a reference value for the calculation of Mahalanobis distance to segment the mixed region into three sub-regions. We have used skin region as the reference as it gives better results.

For user supervised selection, the segmented cluster can be obtained by using one of the four different combinations as shown in the flowchart in Figure 4. The user will choose the most appropriate segmentation combination based on visual observation as shown in Figure 5. However, for the unsupervised method the SVM classifier with ranked features as training data will be used.

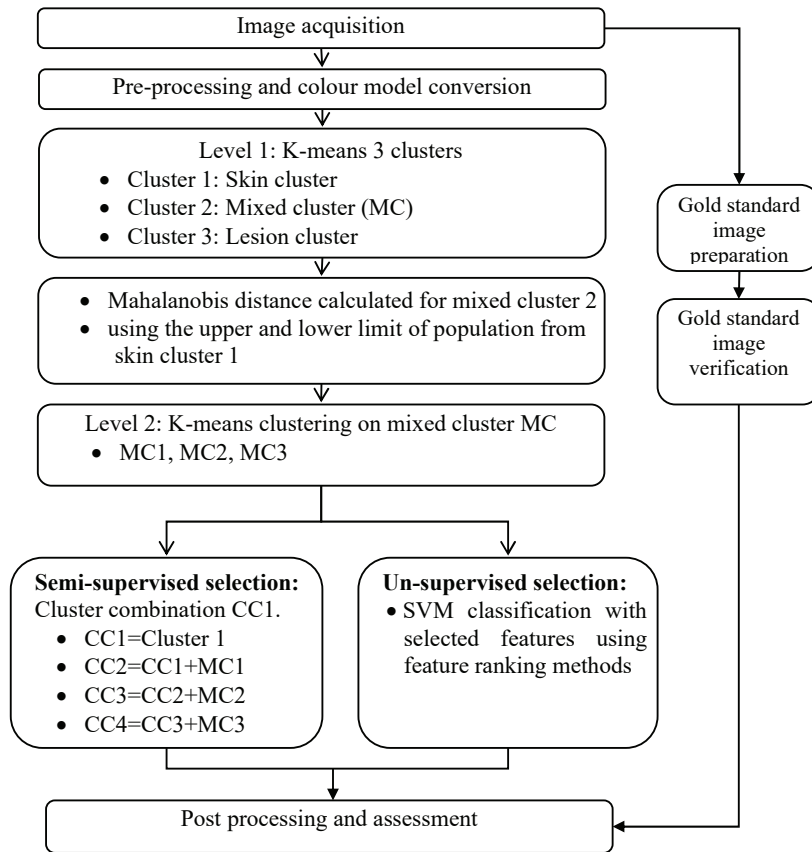


Figure 4: The proposed bi-level segmentation algorithm

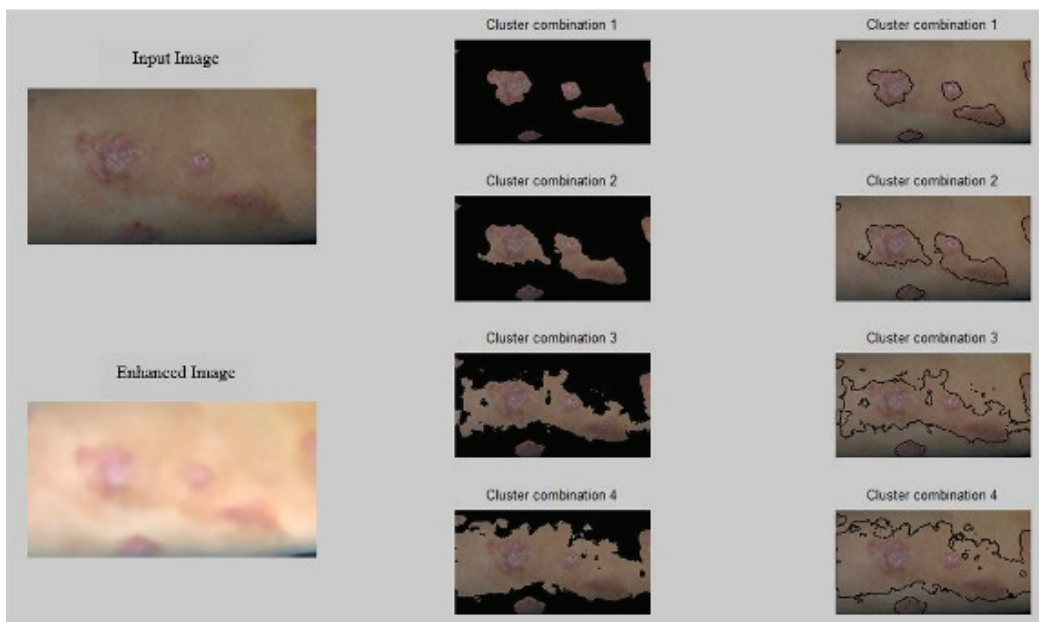


Figure 5: Window for the operator to select a suitable cluster combination for manual selection in level 2

Level 1 clustering

After the pre-processing, the image was segmented into three regions using K-means clustering. The regions were labelled as skin region (cluster 1), mixed region (cluster 2), and lesion region (cluster 3). K-means clustering approach proposed by Yao *et al.* (2013) is used. Equation (8) shows the Euclidean distance (Brindha *et al.*, 2013) used in the 1st level of K-means clustering.

$$\text{Euclidean Dist, } d(c, x) = \sqrt{\sum_{u=1}^v (x_u - c_u)^2} \quad \dots(8)$$

where d represents the distance, x and c represent the data points and centroid respectively, and u and v are the data dimensions used (here only 1 dimension of data is used).

Level 2 clustering

In Level 2 the mixed region is divided into three sub-regions (skin, mixed and lesion) using K-means clustering. The lesion is now composed of 5 clusters. Here the Mahalanobis distance (McLachlan, 1999) was used for the mixed region segmentation as it gives better results than the Euclidean distance (empirically observed). The Mahalanobis distance is given in equation (9).

$$\text{Mahalanobis Dist, } d(x, y) = \sqrt{(y - \mu_x) / \text{cov}(x) \times (y - \mu_x)^T} \quad \dots(9)$$

where y represents the data point and x represents the vector (mean \pm standard deviation) that is used as the reference in equation (9).

The segmentation of the final (mixed) cluster was the most difficult. It was done, using semi-supervised (operator input) or un-supervised (SVM) method.

Cluster combination for manual selection

For semi-supervised selection, operator input is required. The operator subjectively selects the most suitable combination of clusters that represents the segmented lesion. From Level 2 segmentation one cluster was combined with the lesion region at a time, and the option for four types of segmentation selection is present, as shown in the selection window in Figure 5. In the selection window the input image and pre-processed image are shown in the left column. The middle and right columns show the cluster combinations that were selected by the operator. In the right column, the boundary of the lesion was drawn using a black line. For the middle column

the region that does not belong to lesion region in the segmented image was excluded. The operator selects one of the four cluster combinations that best represents the segmented lesion, based on his judgement.

SVM classifier for automated selection

SVM was trained using selected features that are ranked by feature ranking methods. In this study, the T-statistic scoring method was used for feature ranking. The T-statistic scoring method calculates a score for the capability, C_T , of each feature to separate them between classes. Based on the value of C_T , the scoring of each feature was ranked in descending order. The two classes are denoted as 'a' and 'b'. For each feature, the mean, μ_a and μ_b ; standard deviation, σ_a and σ_b ; and the number of images used, n_a and n_b were calculated (Liu *et al.*, 2002). The mathematical expression for calculating C_T is given in equation (10).

$$C_T = \frac{\mu_a - \mu_b}{\sqrt{(\sigma_a^2/n_a) + (\sigma_b^2/n_b)}} \quad \dots(10)$$

The score value obtained in equation (10) was used to sort the features in descending order. The combination of features was done by increasing one feature at a time. Finally, validation was done using 53 gold standard images; the validation error rate was obtained using equation (11).

$$\text{Error rate} = \frac{\text{Falsely classified data}}{\text{Overall data}} \quad \dots(11)$$

Feature extraction and selection

For SVM training features were extracted from the skin and lesion regions of 20 gold standard images. These images were classified as mild images in context of the redness symptom. Five features extracted from the colour layers of four-colour spaces (RGB, CIELab, CSN-I RGB, CSN-I CIELab) are mean, standard deviation (SD), mode, skewness, and kurtosis, whereas the four features extracted from GLCM are contrast, correlation, energy, and homogeneity (Nisar *et al.*, 2020). Colour layers of colour spaces were used for creating the GLCM; the intensity of the colour layer was normalized to '0' to '1'. A symmetrical GLCM was used with the horizontal direction only. The number of levels used for GLCM is 256. All extracted features were ranked using the T-statistical score ranking method and the combination of ranked features with lowest error was selected for SVM training.

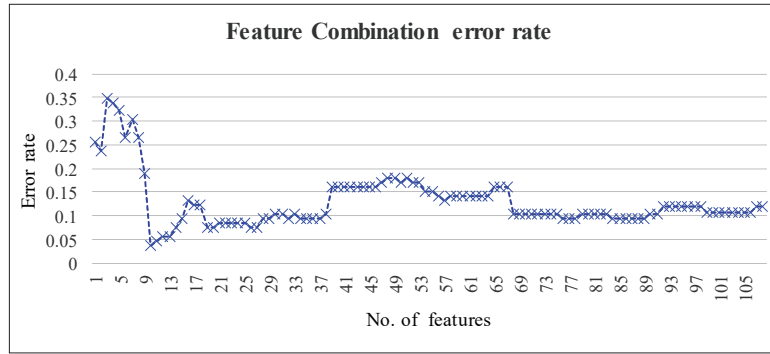


Figure 6: Feature combination error rate graph

Figure 6 shows the feature combination error rate obtained during the training and classification of gold standard data. In this study, 108 features are used, of which 60 (5 features \times 4 colour spaces \times 3 layers) are colour features and 48 (4 features \times 4 colour spaces \times 3 layers) are texture features. It is observed from Figure 6 that the lowest error rate (less than 5%) is achieved for 10 features. Hence, we have used these 10 features for segmentation and classification, which are SD and kurtosis of the R channel, the energy of the G channel, correlation and homogeneity from CSN-I R, homogeneity from CSN-I G, Mean, skewness and homogeneity from CSN-I a*, and skewness from CSN-I b*.

RESULTS AND DISCUSSION

In this paper we have proposed an algorithm for segmentation of eczema skin lesions. For pre-processing, the ALC method is used, and the results are compared with the retinex algorithm. The performance of the proposed segmentation algorithm is compared with K-means and FCM algorithms with three and five clusters. Two core colour spaces, RGB and CIELab, and two normalization techniques for colour spaces, CSN-I and CSN-II, are studied. The green channel of RGB colour space is selected for segmentation as it gives better segmentation performance when compared with red and blue channels; we denote it as RGB(G) (Ch'ng *et al.*, 2014a). Similarly, the green channel is also chosen for CSN-I RGB and CSN-II RGB, and we denote it as CSN-I RGB(G) and CSN-II RGB(G), respectively. For CIELab colour space, chrominance channel a* is selected, and we denote it as CIELab(a*); for CSN-I CIELab and CSN-II CIELab illumination channel L* is selected and we denote it as CSN-I CIELab(L*) and CSN-II CIELab(L*), respectively.

For performance assessment, different metrics are used. True positive, true negative, false positive and false negative are denoted as TP, TN, FP, and FN respectively. These are used to calculate the precision, recall/sensitivity (Sens), specificity (Spec), and accuracy (Acc), as given in equations (12)–(15). All the three parameters should have a high value for good classification results. The DICE coefficient is calculated using equation (16) to compare the similarity between the gold standard and segmented images.

$$\text{Precision} = TP / (TP + FP) \quad \dots(12)$$

$$\text{Recall/Sensitivity (Sens)} = TP / (TP + FN) \quad \dots(13)$$

$$\text{Specificity (Spec)} = TN / (TN + FP) \quad \dots(14)$$

$$\text{Accuracy (Acc)} = (TP + TN) / (TP + TN + FP + FN) \quad \dots(15)$$

$$\text{DICE Coefficient} = (2 \times TP) / (2 \times TP + FP + FN) \quad \dots(16)$$

Table 1 shows the performance comparison of segmentation algorithms using the semi-supervised method: ALC as pre-processing technique and CSN-I RGB(G) colour channel. It is observed that all methods can achieve above 98% accuracy for the best-case segmentation, image 1; however, for image 2, the proposed algorithm has the highest accuracy whereas K-means and FCM with 5 clusters have better accuracy than with 3 clusters.

Tables 2 to 4 show the performance assessment of different segmentation algorithms. Out of the six

colour spaces evaluated, results for only top three best colour spaces are shown. Table 2 shows the overall segmentation performance for the proposed technique using semi-supervised and un-supervised methods. Among the three-colour spaces, CSN-I RGB(G) has the highest accuracy. Between pre-processing methods, ALC has a higher accuracy as compared to retinex.

Tables 3 shows the segmentation assessment for K-means using five and three clusters respectively. It is observed that the ALC pre-processing technique gives better segmentation accuracy. In K-means, for both pre-processing methods CSN-I RGB(G), CSN-II RGB(G) and CIELab(a*) are the three top ranked colour spaces that give the highest segmentation accuracy. Finally,

K-means with $k = 5$ gives higher segmentation accuracy as compared to the $k = 3$ case.

Table 4 shows the segmentation performance for FCM using five and three clusters respectively. For both pre-processing techniques, ALC has a better segmentation accuracy for FCM compared to retinex. For FCM the top three ranked colour spaces are the same as with K-means, as in Table 3. For both FCM ($k = 5$) and FCM ($k = 3$), CSN-I RGB(G) has highest segmentation accuracy. However, for the semi-supervised method using the retinex pre-processing technique there is an exception, as the CIELab(a*) colour space has a slightly higher segmentation accuracy compared to other colour spaces.

Table 1: Performance comparison of segmentation algorithms (semi-supervised method) using CSN-I RGB (G channel).

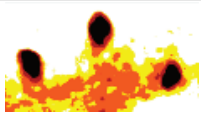


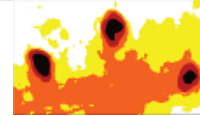






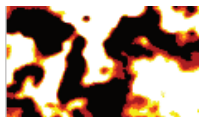
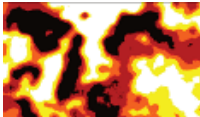

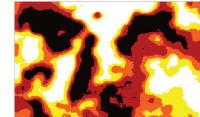






	Proposed	K-means with 5 clusters	K-means with 3 clusters	FCM with 5 clusters	FCM with 3 clusters
Clustered image 1					
Performance of segment (Best case)					
Accuracy	98.91%	98.20%	98.91%	98.32%	98.90%
Clustered image 2					
Performance of segment (random sample)					
Accuracy	91.22%	90.79%	79.15%	84.31%	78.42%

Table 2: Segmentation performance of the proposed two-level K-means algorithm (%)

	ALC				Retinex			
	Precision	Recall/Sens	Spec	Acc	Precision	Recall/Sens	Spec	Acc
Semi-supervised selection								
CSN-I RGB(G)	86.78	81.76	90.37	88.28	87.78	76.86	87.93	86.40
CSN-II RGB(G)	82.52	80.34	87.77	85.98	80.07	82.08	82.79	84.63
CIELab(a*)	79.93	82.45	84.91	85.27	76.31	85.09	78.12	83.08
Un-supervised selection								
CSN-I RGB(G)	80.05	84.98	84.28	84.43	77.91	81.80	81.96	81.99
CSN-II RGB(G)	73.52	86.80	77.87	81.04	73.73	84.51	77.91	80.28
CIELab(a*)	71.95	87.17	75.33	80.20	74.22	84.25	79.10	80.81

Table 3: Segmentation performance of the K-means with 5 and 3 clusters (%)

	ALC				Retinex			
	Precision	Recall/Sens	Spec	Acc	Precision	Recall/Sens	Spec	Acc
Semi-supervised selection (5 clusters)								
CSN-I RGB(G)	87.51	79.23	90.70	87.51	83.38	78.44	87.01	84.87
CSN-II RGB(G)	84.24	77.52	88.41	85.77	80.15	80.94	84.24	84.25
CIELab(a*)	81.04	79.94	85.33	84.67	80.06	82.73	83.65	84.79
Un-supervised selection (5 clusters)								
CSN-I RGB(G)	76.71	84.77	76.04	79.78	73.89	83.24	73.47	78.23
CSN-II RGB(G)	68.64	88.15	68.19	76.25	70.28	85.52	68.68	76.39
CIELab(a*)	68.12	88.33	67.28	75.61	71.02	85.46	71.03	77.08
Semi-supervised selection (3 clusters)								
CSN-I RGB(G)	86.88	77.84	87.37	86.23	79.91	81.30	80.41	82.49
CSN-II RGB(G)	81.33	80.08	84.07	84.32	80.70	77.95	84.13	82.84
CIELab(a*)	75.98	84.91	78.34	82.91	80.21	81.29	83.25	83.67
Un-supervised selection (3 clusters)								
CSN-I RGB(G)	80.01	82.55	81.68	82.46	78.88	79.26	79.31	80.31
CSN-II RGB(G)	72.91	86.25	74.78	79.50	76.78	80.87	78.60	79.77
CIELab(a*)	71.91	87.72	73.61	79.66	73.51	82.74	76.70	78.33

Table 4: Segmentation performance of Fuzzy C mean with 5 and 3 clusters (%)

	ALC				Retinex			
	Precision	Recall/Sens	Spec	Acc	Precision	Recall/Sens	Spec	Acc
Semi-supervised selection (5 clusters)								
CSN-I RGB(G)	87.52	78.60	91.37	87.36	84.25	78.59	88.65	85.65
CSN-II RGB(G)	85.02	77.15	89.18	86.08	81.10	81.16	84.27	84.44
CIELab(a*)	80.13	81.30	85.11	84.63	78.89	83.57	83.24	84.31
Un-supervised selection (5 clusters)								
CSN-I RGB(G)	77.79	84.16	76.71	79.90	74.48	82.57	73.76	78.10
CSN-II RGB(G)	68.17	87.80	67.96	75.94	72.21	84.55	71.62	77.02
CIELab(a*)	67.83	88.62	66.90	75.61	70.36	86.09	70.20	76.75
Semi-supervised selection (3 clusters)								
CSN-I RGB(G)	87.78	76.86	87.93	86.40	83.34	78.02	83.05	83.43
CSN-II RGB(G)	80.07	82.08	82.79	84.63	80.59	79.08	83.66	83.11
CIELab (a*)	76.31	85.09	78.12	83.08	80.49	80.92	81.73	83.60
Un-supervised selection (3 clusters)								
CSN-I RGB(G)	79.80	81.74	81.43	81.70	80.00	79.74	79.87	80.90
CSN-II RGB(G)	72.57	86.38	73.97	78.90	76.39	81.66	77.62	79.38
CIELab(a*)	71.11	87.94	72.63	79.06	72.68	83.42	74.84	77.59

Table 5: Comparison of the segmentation algorithms in terms of DICE coefficient (%)

		Semi-supervised selection				Un-supervised selection					
		Proposed	K-means (k=5)	FCM (k=5)	K-means (k=3)	FCM (k=3)	Proposed	K-means (k=5)	FCM (k=5)	K-means (k=3)	FCM (k=3)
ALC	CSN-I RGB(G)	83.47	81.72	81.74	79.96	79.73	80.17	76.74	77.01	77.41	76.44
	CSN-II RGB(G)	80.56	79.70	79.69	78.51	79.06	77.28	73.94	73.47	75.87	75.46
	CIELab(a*)	80.62	79.21	79.52	78.48	78.72	76.47	73.30	73.32	76.27	75.67
Retinex	CSN-I RGB(G)	80.41	79.34	80.29	77.75	77.88	76.99	74.09	73.65	74.97	76.02
	CSN-II RGB(G)	79.94	79.32	79.96	76.71	77.61	76.13	73.79	74.14%	75.65	75.60
	CIELab(a*)	80.61	80.09	79.77	78.46	78.13	76.31	74.00	73.76	73.63	73.34

Table 5 shows the comparison in terms of DICE coefficient. When comparing between pre-processing methods, ALC performs better than retinex. CSN-I RGB(G) colour space outperforms all the other colour spaces with highest performance accuracy. For segmentation techniques, the proposed bi-level K-means segmentation gives a better accuracy for most of the cases. There are only six cases out of 24 where bi-level K-means does not give the best accuracy; these are RGB(G) in ALC for both the semi-supervised and unsupervised selection mode, CSN-II RGB(G) in ALC, and CIELab(a*) colour space with both normalized techniques in retinex for semi-supervised selection mode. Conventional K-means seems to have a slightly better performance than the proposed method for those cases. Finally, we can say that the proposed method gives an overall best accuracy for the semi-supervised method, which is 88.28%, whereas the un-supervised selection mode has an accuracy of 84.43%.

For the semi-supervised selection mode, segmentation with a higher cluster number (5) gives a better accuracy compared to fewer clusters (3) for both K-means and FCM, whereas the accuracy of the unsupervised selection mode tends to be lower for the higher number of clusters as compared to the lower number of clusters used. For the proposed segmentation algorithm, the input image will first be segmented into three regions in the first step and then the mixed region will further be segmented into another three sub-regions. For the unsupervised selection mode, the poor performance with a higher number of clusters for K-means and FCM may be due to the features selected for training by SVM, despite using the features with lowest validation error rate.

Hence from the results it is observed that the CSN-I RGB(G) colour space has better segmentation accuracy resulting in the highest similarity between the gold standard and segmented images. The colour space normalization technique proposed by Yang *et al.* (2010) shown in equation (4) amplifies the difference between normal skin and an eczema lesion in the green channel of CSN-I RGB colour space, that leads to a better segmentation for the proposed bi-level K-means segmentation algorithm.

CONCLUSION

In this paper we have investigated the performance of colour spaces, pre-processing methods, segmentation techniques and selection modes. Investigations show that G channel of CSN-I RGB outperforms all other colour spaces for the Eczema skin lesion segmentation.

The proposed segmentation method and pre-processing technique ALC performed better in most cases. The highest segmentation accuracy achieved is 88.28% for proposed bi-level segmentation method using CSN-I RGB (G channel) colour space for semi-supervised selection. While highest accuracy achieved for unsupervised selection is 84.43% with the proposed segmentation method using the CSN-I RGB (G channel) colour space. In future we plan to increase the image database for an effective evaluation. Secondly more features should be considered for feature selection; used for training the supervised learning classifier. Deep learning methods will also be considered.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgement

This work is supported by Universiti Tunku Abdul Rahman Research Fund (UTARRF) Grant Number: IPSR/RMC/UTARRF/2020-C1/Y03. The authors would also like to thank Dr Jyh Jong Tang, Department of Dermatology, Hospital Raja Permaisuri Bainun, Ipoh, Malaysia, for his help in verifying the gold standard for eczema skin lesions.

REFERENCES

- Abraham N. & Khan N.M. (2018). A Novel Focal Tversky Loss Function with Improved Attention U-Net for Lesion Segmentation. Available at *arXiv preprint [cs.CV]*. arXiv:1810.07842
- Alam M.N., Munia T.T.K., Tavakolian K., Vasefi F., MacKinnon N. & Fazel-Rezai R. (2016). Automatic detection and severity measurement of eczema using image processing. *Proceedings of 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, USA, 16–20 August. IEEE, pp. 1365–1368.
DOI: <https://doi.org/10.1109/EMBC.2016.7590961>
- Barbier N., Paul C., Luger T., Allen R., De Prost Y., Papp K., Eichenfield L.F., Cherill R. & Hanifin J. (2004). Validation of the eczema area and severity index for atopic dermatitis in a cohort of 1550 patients from the pimecrolimus cream 1% randomized controlled clinical trials programme. *British Journal of Dermatology* **150**(1): 96–102.
DOI: <https://doi.org/10.1111/j.1365-2133.2004.05696.x>
- Bezdek J.C., Ehrlich R. & Full W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences* **10**(2–3): 191–203.
DOI: [https://doi.org/10.1016/0098-3004\(84\)90020-7](https://doi.org/10.1016/0098-3004(84)90020-7)
- Brindha M., Tamilselvan G.M., Valarmathy S., Kumar M.A. & Suryalakshmi M. (2013). A comparative study of face

- authentication using Euclidean and Mahalanobis distance classification method. *International Journal of Emerging Technology and Advanced Engineering* **3**(1): 263–268.
- Carroll C.L., Balkrishnan R., Feldman S.R., Fleischer A.B. & Manuel J.C. (2005). The burden of atopic dermatitis: impact on the patient, family, and society. *Pediatric Dermatology* **22**(3): 192–199.
DOI: <https://doi.org/10.1111/j.1525-1470.2005.22303.x>
- Ch'ng Y.K., Nisar H., Yap V.V., Yeap K.H. & Tang J.J. (2014). Segmentation and grading of eczema skin lesions. In: *Signal Processing and Communication Systems (ICSPCS), 8th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Gold Coast, Australia, 15-17 December. IEEE, pp. 1–5.
DOI: <https://doi.org/10.1109/ICSPCS.2014.7021131>
- Ch'ng Y.K., Nisar H., Yap V.V. & Tang J.J. (2014). A two level k-means segmentation technique for eczema skin lesion segmentation using class specific criteria. In: *Biomedical Engineering and Sciences (IECBES), IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, Kuala Lumpur, Malaysia, 8-10 December. IEEE, pp. 985–990.
DOI: <https://doi.org/10.1109/IECBES.2014.7047659>
- Charman C.R., Venn A.J. & Williams H.C. (2002). Reliability testing of the six area, six sign atopic dermatitis severity score. *British Journal of Dermatology* **146**(6): 1057–1060.
DOI: <https://doi.org/10.1046/j.1365-2133.2002.04644.x>
- Finlay A.Y. (1996). Measurement of disease activity and outcome in atopic dermatitis. *British Journal of Dermatology* **135**(4): 509–515.
DOI: <https://doi.org/10.1046/j.1365-2133.1996.d01-1031.x>
- Funt B., Ciurea F. & McCann J. (2004). Retinex in MATLAB™. *Journal of Electronic Imaging* **13**(1):48–57.
DOI: <https://doi.org/10.1117/1.1636761>
- Hameed N., Shabut A.M., Ghosh M.K. & Hossain M.A. (2020). Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. *Expert Systems with Applications* **141**(1a): 112961.
DOI: <https://doi.org/10.1016/j.eswa.2019.112961>
- Hartigan J.A. & Wong M.A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **28**(1): 100–108.
DOI: <https://doi.org/10.2307/2346830>
- Hegde P.R., Shenoy M.M. & Shekar B.H. (2018). Comparison of machine learning algorithms for skin disease classification using color and texture features. In: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, India, 19–22 September. IEEE, pp. 1825–1828.
DOI: <https://doi.org/10.1109/ICACCI.2018.8554512>
- HOME (Harmonising Outcome Measures for Eczema). (2014). EASI For Clinical signs. Available at <http://www.homeforeczema.org/research/easi-for-clinical-signs.aspx>
- Lee B.W. & Detzel P.R. (2015). Treatment of childhood atopic dermatitis and economic burden of illness in Asia Pacific countries. *Annals of Nutrition and Metabolism* **66**(Suppl 1): 18–24.
DOI: <https://doi.org/10.1159/000370221>
- Liu H., Li J. & Wong L. (2002). A comparative study on feature selection and classification methods using gene expression profiles and proteomic patterns. *Genome Informatics* **13**: 51–60.
DOI: <https://doi.org/10.11234/gi1990.13.51>
- Mallol J., Crane J., von Mutius E., Odhiambo J., Keil U., Stewart A. & ISAAC Phase Three Study Group (2013). The international study of asthma and allergies in childhood (ISAAC) phase three: a global synthesis. *Allergologia et Immunopathologia* **41**(2): 73–85.
DOI: <https://doi.org/10.1016/j.aller.2012.03.001>
- McLachlan G.J. (1999). Mahalanobis distance. *Resonance* **4**(6): 20–26.
DOI: <https://doi.org/10.1007/BF02834632>
- Nisar H., Ch'ng Y.K., Chew T.Y., Tang J.J., Yap V.V. & Yeap K.H. (2013). A color space study for skin lesion segmentation. In: *Circuits and Systems (ICCAS), IEEE International Conference on Circuits and Systems (ICCAS)*, Kuala Lumpur, Malaysia, 18–19 September. IEEE, pp. 172–176.
DOI: <https://doi.org/10.1109/CircuitsAndSystems.2013.6671629>
- Nisar H., Ch'ng Y.K. & Ho Y.K. (2020). Automatic segmentation and classification of eczema skin lesions using supervised learning. In: *IEEE Conference on Open Systems (ICOS)*, Kota Kinabalu, Malaysia, 17–19 November. IEEE, pp. 25–30.
DOI: <https://doi.org/10.1109/ICOS50156.2020.9293657>
- Silverberg J.I. & Hanifin J.M. (2013). Adult eczema prevalence and associations with asthma and other health and demographic factors: A US population-based study. *Journal of Allergy and Clinical Immunology* **132**(5): 1132–1138.
DOI: <https://doi.org/10.1016/j.jaci.2013.08.031>
- The GIMP Team (2014). *GNU Image Manipulation Program ver 2.8*. Available at <https://download.gimp.org/mirror/pub/gimp/v2.8/>, Accessed 24 January 2021.
- Tremp M., Knafla I., Burg G., Wüthrich B. & Schmid-Grendelmeier P. (2011). 'EASIdig'-A digital tool to document disease activity in atopic dermatitis. *Dermatology* **223**(1): 68–73.
DOI: <https://doi.org/10.1159/000330333>
- Yang J., Liu C. & Zhang L. (2010). Color space normalization: Enhancing the discriminating power of color spaces for face recognition. *Pattern Recognition* **43**(4): 1454–1466.
DOI: <https://doi.org/10.1016/j.patcog.2009.11.014>
- Yao H., Duan Q., Li D. & Wang J. (2013). An improved K-means clustering algorithm for fish image segmentation. *Mathematical and Computer Modelling* **58**(3-4): 790–798.
DOI: <https://doi.org/10.1016/j.mcm.2012.12.025>
- Zafar K., Gilani S.O., Waris A., Ahmed A., Jamil M., Khan M.N. & Kashif A.S. (2020). Skin lesion segmentation from dermoscopic images using convolutional neural network.

RESEARCH COMMUNICATION

Natural Products

Bioactive cyclo-(S-Pro-R-Leu) from *Aspergillus flavus*, the marine endophytic fungus from brown alga, *Dictyota kunthi*

Haroon M. Haniffa^{1,2*}, H. Ranjith W. Dharmaratne¹ and M Yasin Mohammad³

¹ Natural Products Programme, National Institute of Fundamental Studies, Hantana Road, Kandy.

² Department of Chemical Sciences, Faculty of Applied Sciences, South Eastern University of Sri Lanka, Oluvil.

³ Michael Sayegh Faculty of Pharmacy, Aqaba University of Technology, Aqaba-11191, Jordan.

Submitted: 25 November 2021; Revised: 14 June 2022; Accepted: 22 July 2022

Abstract: The marine environment is regarded as a rich source of bioactive secondary metabolites, and some of these compounds have been shown to be useful for novel drug discovery. In recent years, research on the chemistry of endophytic fungi from marine algae has shown them to be an important source of secondary metabolites with a wide range of biological activities. The objective of this study was to describe the isolation and identification of bioactive compounds obtained from *Aspergillus flavus*, hosted in the brown alga, *Dictyota kunthi*, together with β -glucuronidase bioassay of them. This fungal strain was cultivated on a large scale and extracted with ethyl acetate (EtOAc). Using various column chromatographic and spectroscopic techniques, we isolated and characterized two metabolites, cyclo-(S-Pro-R-Leu), 1 and genistein 2. Cyclo-(S-Pro-R-Leu) inhibited β -glucuronidase with an IC_{50} value of $83.9 \pm 0.1 \mu\text{M}$ when compared to glucosaccharo-1,4-lactone ($IC_{50} = 48.4 \pm 1.3 \mu\text{M}$). This appears to be the first report of the isolation of cyclo-(S-Pro-R-Leu) from this endophyte. To the best of our knowledge, this is the first report of significant β -glucuronidase inhibitory activity of this compound. From these findings, we can confirm that the endophyte from the brown alga, *Dictyota kunthi*, produces a β -glucuronidase inhibitory metabolite.

Keywords: β -glucuronidase, *Aspergillus flavus*, cyclo-(S-Pro-R-Leu), *Dictyota kunthi*, marine endophytes.

INTRODUCTION

Due to the discovery of unique chemical structures and their bioactivity, the search for novel chemicals from the marine environment has rapidly increased in recent years

(Teixeira *et al.*, 2019). Endophytic fungi live in the inner tissues or even cells of their hosts (Zhang *et al.*, 2009). Secondary metabolites of algal-derived endophytic fungi have proven to be a good source of novel bioactive secondary metabolites for the development of pharmaceuticals and agrochemicals, showing activities such as anticancer, antibiotic, antiviral, antioxidative, and kinase inhibitory or activating activities (Sarasan *et al.*, 2017; Handayani *et al.*, 2019). In continuation of our studies towards the discovery of biologically active compounds from the Sri Lankan seaweeds and their endophytes (Haroon *et al.*, 2013; Haniffa *et al.*, 2019) we isolated the fungal strain *A. flavus*, from the brown alga, *D. kunthi*, collected from the south coast of Sri Lanka. This fungal strain was cultivated on a large scale and extracted using EtOAc. We extracted and characterized two metabolites, 1 and 2, using various column chromatographic techniques. Compounds 1 and 2 were evaluated *in vitro* for β -glucuronidase inhibitory activity and compound 1 was found to have significant activity. This appears to be the first report of significant β -glucuronidase inhibitory activity of this compound. The activity of β -glucuronidase rises in various conditions, including cancer, some hepatic disorders, and AIDS (Awolade *et al.*, 2020). Therefore, development of specific inhibitors of β -glucuronidase could be useful. The present study describes the isolation and identification of cyclo-(S-Pro-R-Leu), 1 and genistein 2 together with β -glucuronidase bioassay for screening of 1. To our knowledge, this is the first report of the isolation of cyclo-(S-Pro-R-Leu) from this endophytic fungus.

* Corresponding author (haroonmh@seu.ac.lk;  <https://orcid.org/0000-0001-7439-0724>)



MATERIALS AND METHODS

General procedure

A Büchi 535 melting point instrument was used to determine melting points. A Jasco DIP-360 digital polarimeter was used to quantify optical rotation in MeOH. Bruker AMX 300 and 500 MHz NMR spectrometers were used to record ^1H -NMR spectra, while the same instrument was used to conduct ^{13}C -NMR studies at 75.45 and 100 MHz, respectively. Chemical shifts (δ) were measured in ppm relative to TMS, with coupling constants J measured in Hz. A Jeol JMS HX 600 mass spectrometer was used to detect EI MS.

Isolation and identification of the fungal strain

D. kunthi, a brown alga, was collected at Kirinda on the southern coast of Sri Lanka. The endophytic fungi were isolated from the blade of the seaweed using the method developed to isolate endophytes from higher plants by Petrini (Fokkema & Heuvel, 1986), with certain modifications (Haroon *et al.*, 2013). The endophytic strain (DK-11) was identified as *A. flavus*, by S.R. Premaratne (Institute of Fundamental Studies, Natural Products Programme, Kandy, Sri Lanka). A subculture has been deposited at the Natural Products Laboratory, Institute of Fundamental Studies, Kandy.

Cultivation of *Aspergillus flavus*

The growth medium (Czapex Dox) for *A. flavus* was prepared by mixing the following chemicals (g/L) in filtered (No. 3 Whatman filter paper) natural seawater (500 mL/L) and distilled water (500 mL/L): K_2HPO_4 (1.0 g), $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ (0.5 g), KCl (0.5 g), $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (0.01 g), sugar (30 g), and peptone (30 g). Culture medium (6.0 L) was prepared and distributed (250 mL, each flask) among 24 Erlenmeyer flasks of 1000 mL capacity (each), and autoclaved at 121 °C for 30 min. Then, agar blocks inoculated with mycelia were added into a 1000 mL flask containing 250 mL of culture medium and incubated at 28 °C with a control (without fungal strain) for 14 ds.

Extraction and isolation of pure compounds

After 14 ds, 250 mL of EtOAc was added into each flask containing culture media and mycelium. They were mixed thoroughly using a sonicator and the EtOAc

layer was separated from each flask and evaporated using a rotary evaporator at 50 °C. The EtOAc extract was dried in a vacuum oven to afford a white residue (1.5 g). Flash column chromatography (CC) on silica gel (230 - 400 mesh size) produced two pure compounds from the EtOAc extract. This column was initially eluted with gradient polarities of n-hexane and EtOAc, which afforded five fractions (F -1 to F-5) and their weights were given as. F-1 (0.3 g), F-2 (0.1 g), F-3 (0.7 g), F-4 (10 mg), and F-5 (20 mg). Repeated flash silica gel (SiO_2) column chromatography of F-3 (20% EtOAc:hexane) yielded pure compound 1 (40 mg, 2.6%, 15% Acetone:hexane). Fraction F-4, when subjected to column chromatography, yielded compound 2 (2 mg, 0.13%, 20% Acetone : hexane).

β -glucuronidase inhibition assay

The β -glucuronidase inhibition assay was performed according to the method developed by Collins *et al.* (Collins *et al.*, 1997; Haroon *et al.*, 2013), with slight modification.

Characterization of compounds

Cyclo-(S-Pro-R-Leu) (1)

A white amorphous solid (40 mg, 2.6%) mp 170 °C; $[\alpha]_D^{27}$ -105.5° (c = 0.5, EtOH) UV (MeOH), λ_{max} at 242 nm; IR (KBr) ν_{max} 1432, 1635, 1670, 2955, 3261 cm^{-1} ; ^1H NMR (400 MHz, CDCl_3); δ_{H} 0.92 (3H, d, J = 6.5 Hz, H_3 -12), 0.96 (3H, d, J = 6.5 Hz, H_3 -13), 1.51 (1H, m, Ha-10), 1.74 (1H, m, H-11), 1.85 (2H, m, H_2 -4), 2.06 (1H, m, Hb-10), 2.13 (1H, m, Ha-5), 2.33 (1H, m, Hb-5), 3.53-3.6 (2H, m, H_2 -3), 3.91 (1H, dd, J = 9.5, 3.4 Hz, H-9), 4.15 (1H, t, J = 7.7 Hz, H-6), 6.16 (1H, s, NH); ^{13}C NMR (75 MHz, CDCl_3); δ_{C} 21.2 (C-13), 22.7 (C-4), 23.2 (C-12), 24.6 (C-11), 28.0 (C-5), 38.6 (C-10), 45.4 (C-3), 53.4 (C-9), 59.4 (C-6), 166.2 (C-7), 170.3 (C-1); CI MS (positive mode) m/z 211 $[\text{M} + \text{H}]^+$.

Genistein (2)

A pale yellow powder (2 mg, 0.13%) mp 286°C lit (Aida *et al.*, 1995) 285 °C; ^1H NMR (500 MHz, CD_3OD); δ_{H} 6.21 (1H, d, J = 2.2 Hz, H-8), 6.31 (1H, d, J = 2.2 Hz, H-6), 6.84 (2H, d, J = 8.6 Hz, H-3'/5'), 7.35 (2H, d, J = 8.5 Hz, H-2'/6'), 8.04 (1H, d, H-2); EI MS 70 eV, m/z (rel.int, %): 270 (100) $[\text{M}^+]$, 153 (56), 137(10), 118 (21).

RESULTS AND DISCUSSION

Chemistry of isolated compounds

Repeated column chromatography on flash silica gel yielded two metabolites from the EtOAc extract of *A. flavus*. Compound 1 was found to be a white amorphous solid with a melting point of 170 °C. The compound's molecular ion peak was found at m/z 211 [$M + H$]⁺ with the aid of CI-MS (positive mode). The amide NH group absorbed at 3261 cm^{-1} in the infrared spectrum of 1, while the two amide carbonyl groups absorbed at 1635 and 1670 cm^{-1} . In the ¹H-NMR spectrum, recorded in CDCl₃, the most downfield singlet, that of the amide NH, appeared at δ_H 6.16. The other two downfield sets of protons, H-6 and H-9 appeared as triplet at δ_H 4.15 (1H, t, $J_{6,5} = 7.7$ Hz) and as double doublet at δ_H 3.91 (1H, dd, $J_{9,10a/10b} = 3.4, 9.4$ Hz), respectively. The ¹³C-NMR broad-band decoupled, HMQC and DEPT spectra of 1, in CDCl₃ indicated 11 carbon signals, including 2 methyls, 4 methylenes, 3 methines, and 2 quaternary carbons. The most downfield signals, those of amide carbons, appeared at δ_C 170.3 and 166.2 while the second most downfield carbon sandwiched between NH and CO appeared at δ_C 58.9. The methine carbon attached to C-6 and the methylene carbon attached to C-3 of 1 appeared at δ_C 53.4 and 45.4 respectively. Compound 1 was identified as cyclo-(S-Pro-R-Leu) by direct comparison with reported ¹H- and ¹³C-NMR spectral data (Adamczeski *et al.*, 1995; Pedras *et al.*, 2005; Liu *et al.*, 2012) (Figure 1). The absolute configuration was determined by the optical rotation, the $[\alpha]_D^{27}$ ($c = 0.5$, EtOH) value of 1 (-105.5°), well agreeing with reported data, $[\alpha]_D = -108$ (Adamczeski *et al.*, 1995; Li *et al.*, 2008). However, according to Yang *et al.* (2009), it showed the optical rotation of $[\alpha]_D = -78.3$.

Compound 2 was found to be a pale yellow powder with a melting point of 286 °C. The molecular ion and base peak of compound 2 was found at m/z 270, which corresponded to the chemical formula C₁₅H₁₀O₅, indicating eleven degrees of unsaturation. The most downfield singlet seen δ_H 8.04 in the ¹H-NMR spectra recorded in CD₃OD was ascribed to olefinic H. Moreover, H-2'/6' and H-3'/5' were ascribed to the para-substituted aromatic ring protons that resonated at δ_H 7.37 (2H, d, $J = 8.6$ Hz) and δ_H 6.84 (2H, d, $J = 8.6$ Hz), respectively.

The presence of tetra - substituted aromatic ring protons was suggested by meta coupling signals at δ_H 6.32 (1H, $J = 2.2$ Hz, H-8) and 6.21 (1H, $J = 2.2$ Hz,

H-6) in the ¹H-NMR spectrum of 2. Compound 2 was confirmed as genistein based on a direct comparison of physical and spectral data (Aida *et al.*, 1995) (Figure 1). Isoflavones are known to occur commonly in higher plants, but they have also been isolated from bacteria, and from a culture of the fungus *Aspergillus niger*.

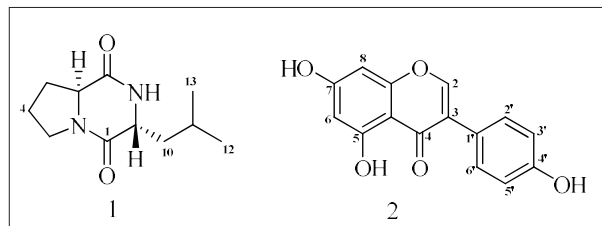


Figure 1: Structures of isolated compounds 1-2.

Table 1: β -glucuronidase inhibition of pure compounds

Compound	IC ₅₀ (μ M)
	Mean \pm S.E.M
1	83.9 \pm 0.1
2	-
glucosaccharo-(1,4)-lactone	48.4 \pm 1.3

β -glucuronidase inhibitory activity

The compounds 1 and 2 were screened for β -glucuronidase inhibitory activity at 200 μ M. When compared with the standard glucosaccharo-(1,4)-lactone (IC₅₀ = 48.4 \pm 1.3 μ M), compound 1 showed a significant β -glucuronidase inhibitory activity with an IC₅₀ value of 83.9 \pm 0.05 μ M and 2 did not show any activity (Table 1). The substantial inhibitory activity of Compound 1 is most likely due to the enzyme's active site accepting a proton from the CONH. This compound was reported to inhibit the marine pathogenic bacteria, *Vibrio anguillarum* (Fdhila *et al.*, 2003) and fungi *Plasmopara viticola* (Musetti *et al.*, 2007). Further, it also exhibits antifouling activity against microfoulers (Li *et al.*, 2006). To the best of our knowledge, cyclo-(S-Pro-R-Leu) from algae-associated fungi is uncommon compared to sponge-associated fungi, actinomycetes, and bacteria (Schmidt *et al.*, 1983; Stierle *et al.*, 1988; Adamczeski *et al.*, 1995; Fdhila *et al.*, 2003; Li *et al.*, 2006; Li *et*

al., 2008; Yang *et al.*, 2009). According to Liu *et al.* (2012), cyclo-(*S*-Pro-*R*-Leu), which has already been shown to be produced by the close species of *Aspergillus fumigatus*, showed weak inhibitory activity against the release of β -glucuronidase from rat polymorphonuclear leukocytes (PMNLs). However, as per the results of this study, cyclo-(*S*-Pro-*R*-Leu) showed a significant β -glucuronidase inhibitory activity with an IC₅₀ value of $83.9 \pm 0.1 \mu\text{M}$. To the best of our knowledge this is the first report of this compound showing significant β -glucuronidase inhibitory activity.

CONCLUSION

The present study described the isolation and characterization of cyclo-(*S*-Pro-*R*-Leu) and genistein from EtOAc extract of *A. flavus* hosted in the brown alga, *Dictyota kunthi*. When compared to the standard glucosaccharo-(1,4)-lactone (IC₅₀ = $48.4 \pm 1.3 \mu\text{M}$), the compound cyclo-(*S*-Pro-*R*-Leu) demonstrated significant β -glucuronidase inhibitory activity with an IC₅₀ value of $83.9 \pm 0.1 \mu\text{M}$. Based on these findings, we can conclude that the endophyte from *Dictyota kunthi* produces a β -glucuronidase inhibitory metabolite.

Acknowledgement

The Authors acknowledge the Third World Academy of Science for the Developing World (TWAS), Trieste, Italy, for funding (partially) of Ph.D. studies at the H. E. J. Research Institute of Chemistry, International Center for Chemical and Biological Sciences, University of Karachi, Pakistan.

REFERENCES

- Adamczeski M., Reed A.R. & Crews P. (1995). New and known diketopiperazines from the Caribbean sponge, Calyx cf. podatypa. *Journal of Natural Products* **58**: 201–208. DOI: <https://doi.org/10.1021/np50116a007>
- Aida M., Hano Y. & Nomura T. (1995). Ficusins A and B, two new cyclic-monoterpene-substituted isokavones from *ficusseptica* barm. f. *Heterocycles* **41**: 2761.
- Awolade P., Cele N., Kerru N., Gummi L., Oluwakemi E. & Singh P. (2020). Therapeutic significance of β -glucuronidase activity and its inhibitors: A review. *European Journal of Medicinal Chemistry* **187**: 111921. DOI: <https://doi.org/10.1016/j.ejmech.2019.111921>
- Collins R., Ng T., Fong W., Wan C., & Yeung H. (1997). Inhibition of glycohydrolase enzymes by aqueous extracts of Chinese medicinal herbs in a microplate format. *IUBMB Life* **42**: 1163–1169. DOI: <https://doi.org/10.1080/15216549700203631>
- Fdhila F., Vázquez V., Sánchez J. L. & Riguera R. (2003). dd-Diketopiperazines: antibiotics active against *Vibrio anguillarum* isolated from marine bacteria associated with cultures of *Pecten maximus*. *Journal of Natural Products* **66**: 1299–1301. DOI: <https://doi.org/10.1021/np030233e>
- Fokkema N.J. & Heuvel J. (1986). *Microbiology of the phyllosphere*. Cambridge University Press, UK. DOI: <https://doi.org/10.1128/aem.55.6.1340-1345.1989>
- Handayani D., Ananda N., Artasasta M. A., Ruslan R., Fadriyanti O. & Tallei T. E. (2019). Antimicrobial activity screening of endophytic fungi extracts isolated from brown algae *Padina* sp. *Journal of Applied Pharmaceutical Science* **9**: 09–13. DOI: <https://doi.org/10.7324/JAPS.2019.90302>
- Haniffa H.M., Ranjith H., Dharmaratne W., Yasin Mohammad M. & Choudhary M.I. (2019). Allelopathic activity of some Sri Lankan seaweed extracts and the isolation of a new brominated nonaromatic isolarene type sesquiterpene from red alga *Laurencia heteroclada* Harvey. *Natural Product Research* **35**(12): 1–8. DOI: <https://doi.org/10.1080/14786419.2019.1655023>
- Haroon M., Premaratne S., Choudhry M. I. & Dharmaratne H. (2013). A new β -glucuronidase inhibiting butyrolactone from the marine endophytic fungus *Aspergillus terreus*. *Natural Product Research* **27**: 1060–1066. DOI: <https://doi.org/10.1080/14786419.2012.708659>
- Li X., Dobretsov S., Xu Y., Xiao X., Hung O.S. & Qian P.-y. (2006). Antifouling diketopiperazines produced by a deep-sea bacterium, *Streptomyces fungicidicus*. *Biofouling* **22**: 187–194. DOI: <https://doi.org/10.1080/08927010600780771>
- Li Z., Peng C., Shen Y., Miao X., Zhang H. & Lin H. (2008). L, L-Diketopiperazines from *Alcaligenes faecalis* A72 associated with South China Sea sponge *Stelletta tenuis*. *Biochemical Systematics and Ecology* **36**: 230–234. DOI: <http://doi:10.1016/j.bse.2007.08.007>
- Liu Y. X., Ma S. G., Wang X. J., Zhao N., Qu J., Yu S.S., Dai J.G., Wang Y.H. & Si Y.K. (2012). Diketopiperazine alkaloids produced by the endophytic fungus *Aspergillus fumigatus* from the stem of *Erythrophloeum fordii* Oliv. *Helvetica Chimica Acta* **95**: 1401–1408. DOI: <https://doi.org/10.1002/hlca.201100521>
- Musetti R., Polizzotto R., Vecchione A., Borselli S., Zulini L., D'Ambrosio M., di Toppi L.S. & Pertot I. (2007). Antifungal activity of diketopiperazines extracted from *Alternaria alternata* against *Plasmopara viticola*: an ultrastructural study. *Micron* **38**: 643–650. DOI: <https://doi.org/10.1016/j.micron.2006.09.001>
- Pedras M.S.C., Yu Y., Liu J. & Tandron-Moya Y.A. (2005). Metabolites produced by the phytopathogenic fungus *Rhizoctonia solani*: Isolation, chemical structure determination, syntheses and bioactivity. *Zeitschrift für Naturforschung C* **60**: 717–722. DOI: <https://doi.org/10.1515/znc-2005-9-1010>
- Sarasan M., Puthumana J., Job N., Han J., Lee J.-S., & Philip R. (2017). Marine algicolous endophytic fungi - a promising drug resource of the era. *Journal of Microbiology and Biotechnology* **27**: 1039–1052.

- DOI: <https://doi.org/10.4014/jmb.1701.01036>
- Schmidtz F.J., Vanderah D.J., Hollenbeak K.H., Enwall C.E., Gopichand Y., SenGupta P., Hossain M. & Van der Helm D. (1983). Metabolites from the marine sponge *Tedania ignis*. A new atisanediol and several known diketopiperazines. *The Journal of Organic Chemistry* **48**: 3941–3945.
DOI: <https://doi.org/10.1021/jo00170a011>
- Stierle A., Cardellina J. & Singleton F. (1988). A marine *Micrococcus* produces metabolites ascribed to the sponge *Tedania ignis*. *Experientia* **44**: 1021–1021.
DOI: <https://doi.org/10.1007/BF01939910>
- Teixeira T.R., dos Santos G.S., Armstrong L., Colepicolo P., & Deboni H. M. (2019). Antitumor potential of seaweed derived-endophytic fungi. *Antibiotics* **8**: 205.
DOI: <https://doi.org/10.3390/antibiotics8040205>
- Yang B., Dong J., Zhou X., Yang X., Lee K.J., Wang L., Zhang S & Liu Y. (2009). Proline-containing dipeptides from a marine sponge of a *Callyspongia* species. *Helvetica Chimica Acta* **92**: 1112-1117.
DOI: <https://doi.org/10.1002/hlca.200800422>
- Zhang Y., Mu J., Feng Y., Kang Y., Zhang J., Gu P.-J., Wang Y., Ma L.-F. & Zhu Y.-H. (2009). Broad-spectrum antimicrobial epiphytic and endophytic fungi from marine organisms: isolation, bioassay and taxonomy. *Marine Drugs* **7**: 97–112.
DOI: <https://doi.org/10.3390/md7020097>

GUIDANCE TO CONTRIBUTORS

GENERAL INFORMATION

Scope

The Journal of the National Science Foundation of Sri Lanka publishes the results of research in all aspects of Science and Technology. It is open for publication of Research Articles, Reviews, Research Communications and Correspondence.

Categories of manuscripts

Research Articles: Research Articles are papers that present complete descriptions of original research. Research Articles should include an Abstract, Keywords, Introduction, Methodology, Results and Discussion, Conclusion and Recommendations where relevant. References should be prepared according to the “Guidelines for the preparation of manuscripts”. Maximum length of the article should be limited to 25 pages with a word count of 10,000 including references, figures and tables. Any articles above this limit will be returned.

Reviews: Reviews are critical presentations on selected topics of Science or Technology. They should be well focused and organized and avoid general “textbook” style. As reviews are intended to be critical presentations on selected topics, reviewers need to have had substantial leadership in research supported by a publication track record in the areas covered by the review. A person/s wishing to submit a Review Article should obtain prior approval from the Editorial Board by submitting a concise summary of the intended article, along with a list of the author’s publications in the related area (jnsf@nsf.gov.lk). Maximum length of the article should be limited to 40 pages with a word count of 12,000 including references, figures and tables. Any articles above this limit will be returned.

Research Communications: Research Communications are intended to communicate important new findings in a specific area of limited scope that are worthy of rapid dissemination among the scientific community. Authors are required to provide a statement justifying the suitability of the submission for a Research Communication. The article should include an Abstract, Keywords, Introduction, Methodology, Results & Discussion, Conclusion and References. Maximum length of the article should be limited to 10 pages with a word count of 2,500 including references, figures and tables. Any articles above this limit will be returned.

Correspondence: Correspondence will be accepted regarding one or more articles in the preceding four issues of the Journal, as well as Letters to the Editor. Articles covering important scientific events or any other news of interest to scientists, reviews of books of scientific nature, articles presenting views on issues related to science and scientific activity will also be considered. Publication will be made at the discretion of the Editor-in-Chief. Maximum length of the article should be limited to 05 pages with a word count of 1,500 including references, figures and tables. Any articles above this limit will be returned.

SUBMISSION OF MANUSCRIPT

Authors submitting articles to the JNSF should first create an account in the Sri Lanka Journals Online System (<https://jnsfsl.sljol.info/>). All manuscripts in MS Word format must be electronically submitted to the journal’s online platform at <https://jnsfsl.sljol.info/submit/start/>. Submissions via emails are not encouraged. Please make sure that no author information is mentioned in the article submitted. The names and details of affiliations of all authors and contact information of the corresponding author must be fed into the system during the online submission process. Authors (at least the corresponding author) are required to provide their personal, validated ORCID ID (by obtaining an ORCID ID from <https://orcid.org/>) when submitting the manuscript. No change to the authors or order of authors will be accepted after the submission. All those who have made significant contributions should be listed as co-authors. The corresponding author should ensure that all contributing co-authors are included in the author list and have approved the final version of the paper and have agreed to its submission for publication.

All submissions should be in English. If the manuscript conforms to the guidelines specified, the date received will be the date that the manuscript was submitted to the online system.

Submissions are accepted for processing on the understanding that they will be reviewed and that they have not been submitted for publication elsewhere (including publication as a full paper or extended abstract as a part of Conference Proceedings). The JNSF does not accept manuscripts that have already been submitted to pre-print servers.

Suggesting potential reviewers by authors

The authors may suggest up to three names of referees when submitting their manuscript, in the Cover Letter space provided at the bottom of the page in the first stage of online submission. Referees should not be from the institution where the work was carried out and should not have been co-authors in previous publications. The address, institutional affiliation and e-mail of the suggested referees should be supplied. Please note that the JNSF is not bound to select all or any of the suggested referees for sending the manuscript for reviewing

Authorship

All authors designated as authors should be eligible for authorship. Those who have made a substantial contribution to the concept or design of the work; or acquisition, analysis or interpretation of data are recognized as Authors.

The corresponding author should be prompt and ensure adherence to timelines when responding to requests, queries and recommendation of reviewers conveyed by or on behalf of the Editor-in Chief and Editorial Board.

Supplementary materials

Any experimental data necessary to evaluate the claims made in the paper but not included in the paper should be provided as supplementary materials. Supplementary materials will be sent to the reviewers and published online with the manuscript if accepted. The supplementary materials should conform to Journal guidelines and should be uploaded as separate files. Authors should number Supplementary Tables and Figures as, for example, 'Supplementary Table S1'. Refer to each piece of supplementary material at the appropriate point(s) in the main article. Supplementary Materials may include description of the materials and methods, controls, or tabulated data presented in Tables or Figures, and programming codes.

Peer review

The manuscripts submitted to the JNSF will initially be screened by the Editorial Board and, if suitable, will be referred to at least two subject experts in the relevant field. The peer-review process of the JNSF is double-blind.

When revision of a manuscript has been requested, the revised manuscript should be submitted on or before the stated deadline. If the revised manuscript is not received on time, the manuscript will not be processed further. The authors' response to the comments of referees should be tabulated with the comment, response and the line number/s for reference. The decision of the Editorial Board shall be final.

Accepted papers are subject to editing. The date of acceptance will be the date when the Editorial Board has decided it to be acceptable for publication.

Article publication fee and complementary copies

A publication fee of US\$ 150 will be levied for each manuscript other than where the corresponding author is affiliated to a Sri Lankan Institute, to cover the publication cost.

A complimentary copy of the Journal issue carrying the respective article will be supplied to each of the authors.

Authors' declaration

When an article is accepted for publication, the authors are required to submit the Authors' Declaration signed by all the authors.

Copyright

Articles in JNSF are published under the Creative Commons License CC-BY-ND. This license permits use, distribution and reproduction of articles for commercial and non-commercial purposes, provided that the original work is properly cited and is not changed in anyway. The copyright of the article is with the National Science Foundation of Sri Lanka. Therefore, authors are requested to check with institution's copyright and publication policy before submitting an article to the JNSF. Authors secure the right to reproduce any material that has already been published or copyrighted elsewhere. When an article is accepted for publication, the authors are required to submit the Transfer of Copyright document signed by all the authors.

Post-publication corrections

The Editorial Board reserves the right to take action on publishing an erratum or corrigendum. If serious errors are identified in a published article, the Journal may consider a retraction or publishing a correction.

STRUCTURE OF MANUSCRIPT

Manuscript

The manuscript should be free of errors and prepared in single column, using double-spaced text of Times New Roman 12 font throughout with line numbers, leaving at least 2 cm margins on both sides, and liberal spacing at the top and bottom of each page. Pages should be numbered consecutively.

a. Style

The paper should be written clearly and concisely. The style of writing should conform to scholarly writing. Slang, jargon, unauthorized abbreviations, abbreviated phrasings should not be used. In general, the impersonal form should be used. Poor usage of language will result in rejection of the manuscript during initial screening.

b. Layout

Manuscripts other than review articles should be generally organized as follows: Title, Abstract, Keywords, Introduction, Methodology, Results and Discussion, Conclusions and Recommendations (where relevant), Acknowledgements and References. Pages should be arranged in the following order:

Title page should include the title of manuscript, and no author information should be mentioned in the title page. If a major part of the research has been published as an abstract in conference proceedings, it should be cited as a footnote on the title page. Authors must also indicate the **general and specific research area** of the manuscript in the title page. In order to highlight the significance of the manuscript, authors are required to provide the following highlights in brief. (1) Why was this study conducted? (2) What are the new findings? (3) Possible applications of the findings. Please limit your answers to 25-30 words for each.

Title: Should accurately and concisely reflect the contents of the article.

Running title: Should be a shortened title (limited to a maximum of 50 characters) that could be printed at the top of every other page of the Journal article.

Abstract: Should be between 200 - 250 words for full length articles and written as a single paragraph. It should not contain any references and should be able to stand on its own. It should outline objectives and methodology together with important results and conclusions. A Review Article should carry a summary of not more than 300 words.

Keywords: Include a maximum of six keywords, which may include the names of organisms (common or scientific), methods or other important words or phrases relevant to the study.

Introduction: This should state the reasons for performing the work with a brief review of related research studies in the context of the work described in the paper. Objectives of the study should be clearly stated.

Materials and Methods: This section should give the details of how you conducted your study. New methods may be described in detail with an indication of their limitations. Established methods can be mentioned with appropriate references. Sufficient details should be included to allow direct repetition of the work by others. Where human subjects are involved, they should be referred to by numbers or fictitious names. A paper reporting the results of investigations on human subjects or on animals must include a statement to the effect that the relevant national or other administrative and ethical guidelines have been adhered to, and a copy of the ethical clearance certificate should be submitted. Methods of statistical analyses used should be mentioned where relevant.

Results and Discussion: Results: the results should be concisely and logically presented. Repetition of the same results in figures, tables or text should be avoided.

Discussion: data essential for the conclusions emerging from the study should be discussed. Long, rambling discussions should be avoided. The discussion should deal with the interpretation of results. It should logically relate new findings to earlier ones. Unqualified statements and conclusions not completely supported by data should be avoided.

Molecular sequence data, such as gene or rDNA sequences, genome sequences, metagenomic sequences etc. must be deposited in a public molecular sequence repository, such as GenBank, that is part of the International Nucleotide Sequence Database Collaboration (INSDC). The accession numbers obtained must be cited in the text, Table or on Figures of phylogenetic trees of the manuscript.

Conclusion: The conclusion should be brief, highlight the outcomes of the study and should be aligned with the objectives of the study. It should not contain references.

Conflict of interest statement: All authors should include a statement on conflict of interest disclosing any financial or other substantive conflicts of interest that may be construed to influence the results or interpretation of their research. All sources of financial support for the project should be disclosed.

Acknowledgement: Should be brief and made for specific scientific, financial and technical assistance only. If a significant part of the research was performed in an institution other than in those indicated by the authors' affiliations given in the title page, this fact should be acknowledged. All those who have made substantial contribution to the research but do not qualify to be authors should be acknowledged.

References :

All research work of other authors, when used or referred to or cited, should be correctly acknowledged in the text and in the References.

Citing references in the text:

- References to the literature must be indicated in the text and tables as per the Author-Year System, by the author's last name and year, in parenthesis (i.e. Able, 1997) or (Able & Thompson, 1998).
- Citation to work by more than two authors should be abbreviated with the use of *et al.* (i.e. Able *et al.*, 1997).
- Multiple publications by the same first author in the same year should be coded by letters, (i.e. Thompson, 1991a; b).
- Multiple citations should be made in chronological order and separated by a semi-colon, (i.e. Zimmerman *et al.*, 1986; Able *et al.*, 1997).
- Reference to unpublished work, work in preparation or work under review should be cited in italics as (*unpublished data*) or, with the author's initials and surname given; such works should not be included in the Reference section.
- Personal communications may be mentioned in the text with the date of communication as (*Personal communication*, 2 June 2000).

List of references:

- The list of References should be arranged in alphabetical order based on the last name of the first author.
- Names of all the authors should be given except when there are more than 10 authors. When there are more than 10 authors, only the name of the first author can be given followed by *et al.*
- All the initials of the author must be given after the last name and the year of publication should follow in parentheses.
- This should be followed by the full title of the referred publication.
- When journal articles are listed, the journal name should be given in full and in italics and followed by the volume number in bold type, issue number in parentheses and then the inclusive pages.
- Where there are several publications by the same author(s) and published in the same year they should be differentiated by adding a lower-case letter after the year. When books are listed, the order should be: author(s), year, book title, volume number, edition, pagination/ inclusive pages, publisher and place of publication. The book title should be in italics. When sections of a book are listed, the order should be: author(s) of chapter, year, title of the section, title of the book, edition, inclusive pages, publisher and place of publication.
- Digital object identifiers (DOIs) should be included for all references where available.
- References should only be cited as 'in press' if the paper has been accepted for publication.

Examples of correct forms of references are given below.

Journal Articles

Boutin C. & Harper J.L. (1991). A comparative study of the population dynamics of five species of *Veronica* in natural habitats. *Journal of Ecology* 79(01): 199 – 221.

DOI: <https://doi.org/10.2307/2260793>

Books

Burnham K.P. & Anderson D.R. (2002). *Model Selection and Multimodal Inference*, 2nd edition, pp. 488. Springer Science and Business Media, Inc., New York, USA.

Book Chapters

Hinrichsen R.A. & Holmes E.E. (2009). Using multivariate state-space models to study spatial structure and dynamics. In: *Spatial Ecology* (eds. R.S. Cantrell, C. Cosner & S. Ruan), pp. 145 – 166. CRC/ Chapman Hall, Florida, USA.

DOI: <https://doi.org/10.1201/9781420059861.ch8>

Edited Books

Kimati H., Amorim L., Rezende J.A.M., Bergamin Filho A. & Camargo L.E.A. (eds.) (2005). *Manual de Fitopatologia*, volume 2. Doenças das Plantas Cultivadas, 4th edition. Ceres, São Paulo, Brazil.

Conference Papers

Weaver D. (2002). Implementation of a learning management system using an integrated approach to professional development. In: Winds of change in the sea of learning. *Proceedings of the 19th Annual Conference of the Australasian Society for Computers in Learning and Tertiary Education (ASCILITE)* (eds. A. Williamson, C. Gunn, A. Young & T. Clear), volume 2, Auckland, New Zealand, 8-11 December. Unitec Institute of Technology, Auckland, New Zealand, pp. 711-720.

Agency Publications

U.S. Census Bureau (2009). *World Population: 1950 – 2050*. U.S. Census Bureau, Washington DC, USA.

Department of Health (2008). *Health Inequalities: Progress and Next Step* (pdf). Department of Health, London, UK. Available at http://PublicationsPolicyAndGuidance/DH_08_5307, Accessed 9 June 2008.

Other

Robinson L.J. (2003) Spatial scale and depletion models of farmland birds in a fragmented landscape. *PhD thesis*, University of Reading, Reading, UK.

Efford M.G. (2008). Density 4.3: software for spatially explicit capture-recapture. Available at <http://www.otago.ac.nz/density>, Accessed 15 March 2009.

Abbreviations and Symbols: Unless common, these should be defined when first used, and not included in the abstract. The SI System of units should be used wherever possible. If measurements were made in units other than SI, the data should be reported in the same units followed by SI units in brackets, e.g. 5290 ft (1610 m).

Formulae and Equations: Equations should be typewritten and quadruple spaced. They should be started on the left margin and the number placed in parentheses to the right of the equation.

Nomenclature: Scientific names of plants and animals should be printed in italics. In the first citation, genus, species and authority must be given. e.g. *Borassus flabellifer* Linn. In latter citations, the generic name may be abbreviated, for example, *B. flabellifer* L.

Tables and figures: Tables and Figures should be clear and intelligible and kept to a minimum, and should not repeat data available elsewhere in the paper. Any reproduction of illustrations, tabulations, pictures etc. in the manuscript should be acknowledged.

Tables: Tables should be numbered consecutively with Arabic numerals and placed at the appropriate position in the manuscript. If a Table must be continued, a second sheet should be used and all the headings repeated. The number of columns or rows in each Table should be minimized. Each Table should have a title, which makes its general meaning clear, without reference to the text. All Table columns should have explanatory headings. Units of measurement, if any, should be indicated in parentheses in the heading of each column. Vertical lines should not be used and horizontal lines should be used only in the heading and at the bottom of the table. Footnotes to Tables should be placed directly below the Table and should be indicated by superscript lower case italic letters (^a, ^b, ^c, etc.).

Figures: All illustrations are considered as figures, and each graph, drawing or photograph should be numbered consecutively with Arabic numerals and placed at the appropriate position in the manuscript. Any lettering to appear on the illustrations should be of a suitable size for reproduction and uniform lettering should be used in all the Figures of the manuscript. Scanned figures or photographs should be of high quality (**300 dpi**), to fit the proportions of the printed page (12 × 17 cm). Each figure should carry a legend so that the general meaning of the figure can be understood without reference to the text. Where magnifications are used, they should be stated.

Units of measurement

Length: km, m, mm, µm, nm

Area: ha, km², m²

Capacity: kL, L, mL, µL

Volume: km³, m³, cm³

Mass: t, kg, g, mg, µg

Time: year(s), month(s), wk(s),
d(s), h, min, s

Concentration: M, mM, N, %, g/L, mg/L, ppm

Temperature: °C, K

Gravity: x g

Molecular weight: mol wt

Others: Radio-isotopes: 32P

Radiation dose: Bq

Oxidation-reduction potential: rH

Hydrogen ion concentration: pH