Pertanika Journal of

# SCIENCE & TECHNOLOGY

JST

VOL. 22 (1) JAN. 2014

PERTANIKA JOURNALS

A scientific journal published by Universiti Putra Malaysia Press

## About the Journal

*Pertanika* is an international peer-reviewed journal devoted to the publication of original papers, and it serves as a forum for practical approaches to improving quality in issues pertaining to tropical agriculture and its related fields. *Pertanika* began publication in 1978 as the Journal of Tropical Agricultural Science. In 1992, a decision was made to streamline *Pertanika* into three journals to meet the need for specialised journals in areas of study aligned with the interdisciplinary strengths of the university. The revamped Journal of Science & Technology (JST) aims to develop as a pioneer journal focusing on research in science and engineering, and its related fields. Other *Pertanika* series include Journal of Tropical Agricultural Science (JTAS); and Journal of Social Sciences and Humanities (JSSH).

JST is published in **English** and it is open to authors around the world regardless of the nationality. It is currently published two times a year, i.e. in **January** and **July**.

### Goal of Pertanika

Our goal is to bring the highest quality research to the widest possible audience.

### Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions are guaranteed to receive a decision within 12 weeks. The elapsed time from submission to publication for the articles averages 5-6 months.

### Indexing of Pertanika

Pertanika is now over 33 years old; this accumulated knowledge has resulted in Pertanika JST being indexed in SCOPUS (Elsevier), EBSCO, Thomson (ISI) Web of Knowledge [CAB Abstracts], DOAJ, Google Scholar, ERA, ISC, Citefactor, Rubriq and MyAIS.

### Future vision

We are continuously improving access to our journal archives, content, and research services. We have the drive to realise exciting new horizons that will benefit not only the academic community, but society itself.

We also have views on the future of our journals. The emergence of the online medium as the predominant vehicle for the 'consumption' and distribution of much academic research will be the ultimate instrument in the dissemination of research news to our scientists and readers.

### Aims and scope

Pertanika Journal of Science and Technology aims to provide a forum for high quality research related to science and engineering research. Areas relevant to the scope of the journal include: *bioinformatics, bioscience, biotechnology and biomolecular sciences, chemistry, computer science, ecology, engineering, engineering design, environmental control and management, mathematics and statistics, medicine and health sciences, nanotechnology, physics, safety and emergency management,* and related fields of study.

### Editorial Statement

Pertanika is the official journal of Universiti Putra Malaysia. The abbreviation for Pertanika Journal of Science & Technology is *Pertanika J. Sci. Technol.*

JST
Journal of Science & Technology
AN INTERNATIONAL PEER-REVIEWED JOURNAL



**EDITORIAL OFFICE**
**JOURNAL DIVISION**
Office of the Deputy Vice Chancellor (R&I)
1st Floor, IDEA Tower II
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 Serdang, Selangor Malaysia.
Gen Enq.: +603 8947 1622 | 1619 | 1616
E-mail: executive_editor.pertanika@upm.my
URL: www.journals-jd.upm.edu.my

**PUBLISHER**
Kamariah Mohd Saidin
UPM Press
Universiti Putra Malaysia
43400 UPM, Serdang, Selangor, Malaysia.
Tel: +603 8946 8855, 8946 8854
Fax: +603 8941 6172
E-mail: penerbit@putra.upm.edu.my
URL: http://penerbit.upm.edu.my



## ABSTRACTING/INDEXING

*Pertanika* is now over 35 years old; this accumulated knowledge has resulted the journals being indexed in SCOPUS (Elsevier), Thomson (ISI) Web of Knowledge [BIOSIS & CAB Abstracts], EBSCO, DOAJ, Google Scholar, AGRICOLA, ISC, Citefactor, Rubriq and MyAIS. JST is also indexed in ERA.

**Pertanika Journal of Science & Technology**
**Vol. 22 (1) Jan. 2014**

## Contents

**Selected Articles from CUTSE International Conference 2011**

**Guest Editor:** Ashutosh Kumar Singh
**Guest Editorial Board:** Sujan Debnath and Muhammad Ekhlasur Rahman

# Foreword

Welcome to the **First Issue 2014** of the Journal of Science and Technology (JST)!

JST is an open-access journal for the studies in science and technology published by Universiti Putra Malaysia Press. It is independently owned and managed by the university and is run on a non-profit basis for the benefit of the world-wide science community.

In this issue, **26 articles** are published, out of which **one** is a review article, **one** is a short communication and **19** are regular articles. **Four articles** are from Curtin University Technology, Science and Engineering International Conference "Innovative Green Technology for Sustainable Development" (CUTSE 2011). The authors of these articles vary in country of origin, coming from Malaysia, India, Germany and Australia.

The review article discusses a potential of Bi-Ag as an alternative high temperature solder, in search of lead-free solder replacement *(Rohaizuan Rosilli, Azmah Hanim Mohamed Ariff* and *Shahrul Fadzli Muhamad Zam)*. The short communication proposes a method that uses chirp signal as a training sequence and employs fractional Fourier transform (FRFT) as a tool to localise the training sequence (chirp) received by the Orthogonal Frequency Division Multiplexing (OFDM) system *(Saxena, R.* and *Joshi, H. D.).*

The regular articles cover a wide range of study, and include the following: a new recursive circular algorithm for listing all permutations *(Sharmila Karim, Zurni Omar, Haslinda Ibrahim, Khairil Iskandar Othman* and *Mohamed Suleiman)*; a comparative study of Yttrium-doped ceria ceramics synthesised using mechanochemical and solid state methods *(Ong, P. S., Tan, Y. P., Taufiq-Yap, Y. H.* and *Zainal, Z.)*; design and development of a sweet potato digging device *(Md. Akhir, H., Ahmad, D., Rukunudin, I. H., Shamsuddin, S.* and *A. Yahya)*; vegetative swales for treatment of stormwater runoff from construction sites *(Ismail, A. F., Sapari, N.* and *Abdul Wahab, M. M.)*; The Performance of Robust Modification of Breusch-Godfrey Test in the Presence of Outliers *(Lim, H. A.* and *Midi, H.)*. physical properties of liberica coffee (*Coffea liberica*) berries and beans *(Ismail, I., Anuar, M. S.* and *Shamsudin, R.)*; a descriptive report on gycaemic control and treatment profile among 20646 adult type 2 diabetes mellitus *(Chew, B. H., Lee, P. Y., Mastura, I., Cheong, A. T., Sri Wahyu, T.* and *Zaiton, A.)*; development of internet-based instrumentation for the study of the hall effect *(Ariffin Abas, Abdul Halim Shaari, Zainal Abidin Talib* and *Zaidan Abdul Wahab)*; *in-vitro* antidermatophytic activity of methanolic fractions from *Entada spiralis* Ridl. stem bark and its bioautographic profile *(Aiza Harun, Siti Zaiton Mat So'ad, Norazian Mohd Hassan* and *Neni Kartini Che Mohd Ramli)*; assessment of digital cameras in mapping meandering flumes using close range photogrammetric technique *(Udin, W. S., Ahmad, A.* and *Ismail, Z.)*; correlation of electrical resistivity with some soil parameters for the development of possible prediction of slope stability and bearing capacity of soil using electrical parameters *(Syed Baharom Syed Osman, Mohammad Nabil Fikri* and *Fahad Irfan Siddique)*; chemically modified sago waste for oil absorption

*(Zainab Ngaini, Rafeah Wahi, Dayang Halimatulzahara* and *Nur An-Nisaa' Mohd Yusoff)*; chemical constituents of *Aglaia lanuginose (Kamarulzaman, F. A., Mohamad, K., Awang, K.* and *Lee, H. B.)*; preconditioned subspace Quasi-Newton method for large scale optimisation *(Hong Seng Sim, Wah June Leong, Malik Abu Hassan* and *Fudziah Ismail)*; GIS routing and modelling of residential waste collection for operational management and cost optimisation *(Billa, L., Pradhan, B.* and *Yakuup, A.)*; performance evaluation and characteristics of selected tube wells in the coastal alluvium aquifer, Selangor *(Fauzie, M. J., Azwan, M. M. Z., Hasfalina, C. M.* and *Mohammed, T. A.)*; optimisation of modified Fenton ($FeGAC/H_2O_2$) pretreatment of antibiotics *(Augustine Chioma Affam, Malay Chaudhuri* and *Shamsul Rahman Mohammed Kutty)*; microclimate inside a tropical greenhouse equipped with evaporative cooling pads *(Diyana Jamaludin, Desa Ahmad, Rezuwan Kamaruddin* and *Hawa Z. E. Jaafar)*; and the use of computational fluid dynamics in the determination of solar collector orientation and stack height of a solar induced ventilation prototype *(Yusoff, W. F. M., Sapian, A. R., Salleh, E., Adam, N. M., Hamzah, Z.* and *Mamat, M. H. H.).*

I conclude this issue with four articles arising from the CUTSE 2011 international conference: Full-scaled impulse turbine performance prediction using numerical simulation *(A. Sahed* and *F. B. Ismail Alnaimi)*; sorption of $SO_2$ and NO by modified palm shell activated carbon: Breakthrough curve model *(Sumathi, S., S. Bhatia, K. T. Lee* and *A. R. Mohamed)*; geochemical assessment of sediment quality using multivariate statistical analysis of Ennore Creek, north of Chennai, SE coast of India *(M. Jayaprakash, R. Nagarajan, P. M. Velmurugan, L. Giridharan, V. Neetha* and *B. Urban)*; and a preliminary investigation on electrochemical parameters of lake waters in and around Miri city, Malaysia *(M. V. Prasanna, R. Nagarajan, A. Elayaraja* and *S. Chidambaram).*

I anticipate that you will find the evidence presented in this issue to be intriguing, thought-provoking, and, hopefully, useful in setting up new milestones. Please recommend the journal to your colleagues and students to make this endeavour meaningful.

I would also like to express my gratitude to all the contributors, namely, the authors, reviewers and editors for their professional contribution towards making this issue feasible. Last but not least, the editorial assistance of the journal division staff is fully appreciated.

JST is currently accepting manuscripts for upcoming issues based on original qualitative or quantitative research that opens new areas of inquiry and investigation.

**Chief Executive Editor**
Nayan Deep S. KANWAL, FRSA, ABIM, AMIS, Ph.D.
nayan@upm.my

Pertanika J. Sci. & Technol. 22 (1): 1 - 13 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

*Review Article*

# Bi-Ag as an Alternative High Temperature Solder

**Rohaizuan Rosilli*, Azmah Hanim Mohamed Ariff and Shahrul Fadzli Muhamad Zam**

*Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

The search for a high temperature lead-free solder replacement for high temperature leaded solder eutectic alloy has been an evolving process as the threat of a regional lead ban became a reality in July 2006. The advantages and disadvantages of lead-free solder in terms of manufacturing, performance and reliability have been increasingly revealed through companies' Research and Development (R&D), industrial consortia and university researchers. Materials and component design are the primary criteria to focus on the development for the current generation of high temperature lead-free solder alloys. According to the current status of high temperature lead free soldering, there are many unsolved technical problems such as explanation on the lift-off phenomenon, establishment of high temperature lead-free plating technology, construction of a database of physical properties (solder, parts, PCBs), standardization of high temperature solder materials evaluation technology, and most importantly, the best candidate material for high temperature solder. Clearly, high temperature soldering is one of the unsolved problems of the century in lead-free soldering. Moreover, most of the questions still remain unanswered by researchers. This paper reviews research conducted on the Bi-Ag solder alloy, which is one of the candidate alloys that has been proposed as an alternative for high temperature lead-free solder.

*Keywords:* Lead-free Solder, Bi-Ag Solder Alloy, Intermetallic Compound (IMC), Microstructure

## INTRODUCTION

Soldering is a metallurgical process for joining metal parts and it uses molten filler metal to wet the surfaces of a joint that provides a conductive path required in achieving a connection from one circuit element to another. It is also important for interconnection in both level 1 (IC Packaging)

*Article history:*
Received: 12 August 2011
Accepted: 17 October 2011

*E-mail addresses:*
rohaiz87@gmail.com (Rohaizuan Rosilli),
azmah@upm.edu.my (Azmah Hanim Mohamed Ariff),
fadzli_shahrul@yahoo.com (Shahrul Fadzli Muhamad Zam)
*Corresponding Author

ISSN: 0128-7680 

and level 2 (mounting of electronic components onto printed circuit boards) processes of the modern electronics industry. Hence, in order to produce a high quality high temperature soldering process, it is important to understand the fundamental of solder and soldering. The soldering process can be divided into three important factors that must be considered, namely, spreading, base metal dissolution, and formation of intermetallic compound layer (Lea, 1988).

Lead-free solder is defined as elimination of lead in solder alloys system because of the response to the concerns for the environment and human health. Environmental Protection Agency (EPA) already mentioned lead as one of the top 17 chemicals that poses the greatest threats to human life and the environment (Wood *et al*., 1994). Disposed lead can leach into drinking water, posing a severe health risks to human. Among the common types of lead poisoning are alimentary, neuromotor and encephalic (Ir. Sax, 1984). Lead poisoning can be detected when its level in blood exceeds 50 mg/dl of blood (Napp, 1995). Occupational Safety and Health Administration (OSHA) requires that workers have no more than 50 mg/dl of lead in their blood and recommends their workers to maintain blood level below 30 mg/dl if they plan to have children (EPA, 1991). Wave soldering produces dross on the molten solder and about 90% of dross can be refined to pure metal for reuse, but the remaining acts as waste product (Nriagu *et al*., 1988).

Attempts to recycle lead in printed wiring boards of consumer electronic products have not been successful. Therefore, Japan has required that all new electronic products be lead-free solder from January 2005, while the European Union introduced legislation to ban lead from electronic products by 1st July 2006. Nevertheless, the United States government has not yet legislated against the use of lead in electronic products. In response to the new legislation, most major electronic manufacturers, including those in the U.S., have stepped up their search for the alternatives to leaded solders. Even when the use of lead in electronics industry appears to be minimal, the potential for lead exposure cannot be ignored (Monsalve, 1984).

Although Sn-Pb can be replaced by lead-free solder for low temperature applications, there is no lead-free solder available in the market that can replace the high temperature leaded solders. Most of the internal joining such as in the first level interconnection flip chip uses 90 to 95 wt. % of lead (Katsuaki *et al*., 2009). Stated in the RoHS (The Restriction on the Use of Hazardous Substances) regulation, the law gave exemptions to the industry by allowing the use of solder with above 85% lead. The challenge faced towards substituting leaded solder is finding the right combination of alloy that can withstand the high temperature working condition with the reflow peak temperature and must not overshoot the temperature range that may damage the substrate (warpage).

High-temperature solders have been widely used in various applications. It is not only used as a die-attach solder, but also in the aircraft industry, space satellite, automotive, oil and gas well explorations, assembling optical components, automobile circuit boards, circuit modules as step soldering and many more. Most importantly, it can produce reliable interconnections, in which the high-density packaging technology and high-service temperature are needed. Important characteristics of the high-temperature solders are good wetability, high ductility, low shear modulus and good resistance to thermo-mechanical fatigue. Due to environmental reasons, the development of lead-free solders has become an important issue and until now, only a few alloy candidates have been proposed for the lead-free high-temperature solders such as

Au-Sn, Bi-Ag, and Zn-Al based alloys. The Bi-Ag alloy is the most suitable because it exhibits a high melting temperature, similar hardness with Pb-5Sn and lower cost as compared to Au-Sn (Yaowu Shi *et al*., 2009). Meanwhile, silver element is widely used in hybrid microelectronics because the cost is lower compared to gold and palladium. Moreover, silver can be processed in air without having to concern about oxidation that affects its electrical resistivity as $Ag_2O$ is conductive. This unique property gives high electrical and thermal conductivity and exhibits limited fatigue.

*The Bi-Ag Binary Phase Diagram*

Phase diagram gives valuable information on the various phases of a substance and the conditions under which each phase exists. It is a powerful tool for developing new solder since a wide range of potential alloys can be easily evaluated in theory (Ursula, 2002). Bi-Ag phase diagram in Fig.1 shows that the eutectic point is 262.5°C when the composition is at 97.5 wt. % Bi and 2.5 wt. %Ag.

As in the soldering process, eutectic solder material is normally used because of its lower temperature that reduces the likelihood of component damage, as well as lower dross formation during the soldering process due to the one point melting phenomena. The Bi-Ag alloy system exhibits a higher melting point at other wt. % combination. A further investigation by Lalena has shown that increasing the Ag content of the Bi-Ag alloy to 11 wt. %, promotes an increase in thermal conductivity and ductility (Lalena *et al*., 1980).

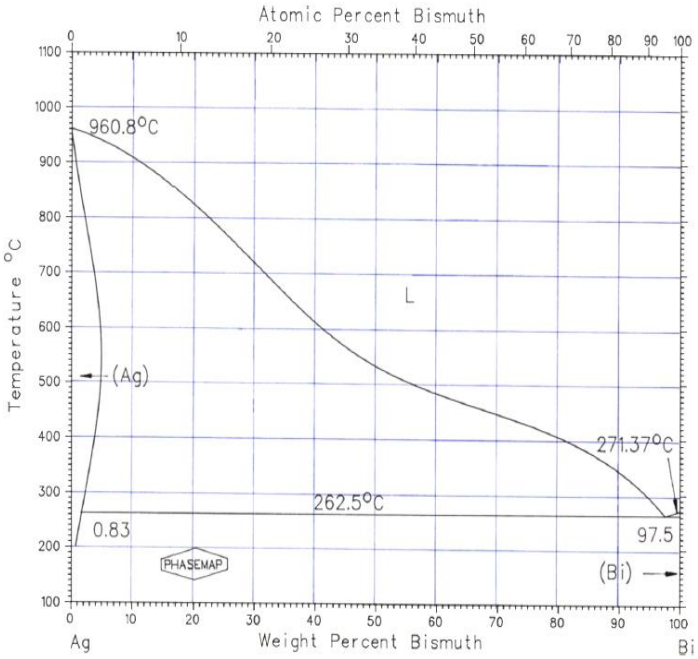



Figure 1: The Bi-Ag Binary Phase Diagram (Elliott *et al*., 1980)

*Melting Temperature Modifications*

The solidus line of the Bi-Ag alloy solder remains almost constant at a temperature of 262.5°C according to the phase diagram. The temperature between the liquidus and solidus line increases linearly with Ag wt. %. Yaowu Shi, in Table 1, shows that addition of small amounts of rare earth Ce will slightly decrease the melting temperature. This is important because, for high lead solder replacement, the lead-free solders must have solidus temperature higher than 260°C so that it will not melt during further assembly process or during usage (Katsuaki *et al*., 2009). The Bi-Ag system may have fulfil the requirement of high temperature solder even when the solidus temperature of the Bi-Ag alloys is lower than that of Pb5Sn. However, to improve the properties of high temperature lead free solders, the rare earth (RE) elements are used. Other than reducing the melting temperature, it is also known as a surface-active element that promotes refinement of microstructure, alloying, and metamorphosis of inclusions. The addition of the rare elements on the lead free solders alloys has been widely studied by researchers (Xia *et al*., 2002; Chen *et al.*, 2002; Wu *et al.*, 2007). The melting temperatures of Bi-Ag, Bi-Ag-RE, and Pb-Sn solder are as shown in Table 1:

TABLE 1
Melting temperature of Bi-Ag, Bi-Ag-RE and Pb-Sn solder (Yaowu Shi *et al*., 2009)

| Solder | Liquidus (°C) | Solidus (°C) |
|---|---|---|
| Bi2.5Ag | 262.6 | 260.9 |
| Bi2.5AgRE | 262.3 | 260.9 |
| Bi5AgRE | 301.5 | 260.9 |
| Bi7.5AgRE | 334.5 | 260.7 |
| Bi10Ag | 380.6 | 261.4 |
| Bi10AgRE | 378.2 | 260.8 |
| Pb5Sn | 311.6 | 310.6 |

*Intermetallic Compund (IMC)*

The diffusion layer between solder and substrate is known as intermetallic compound (IMC). In the soldering theory, the most important part that must be considered during a soldering process is intermetallic formation. Soldering does not only involve physical dissolution, but it is also a chemical reaction process when the solder wets on the substrate (Humpston *et al*., 1993). In other words, without intermetallic compound formation, soldering can not take place. Intermetallic compounds will enhance solder wetting on the substrate and reduce the dissolution rate of substrate in solder through the diffusion barrier role of the intermetallics layer. Intermetallic formation strictly follows stoichiometric ratios of elements, where some elements have high affinity for each other to even deny bonding with other elements (Lea, 1988).

*Microstructure of Bi-Ag/Cu Interface and Tensile Behavior of Cu/Bi-Ag/Cu Joints*

Fig.2 shows the rapidly solidified structure of Bi-Ag pellets examined using Scanning Electron Microscope (SEM). Non-equilibrium solidification structural features were identified. The Bi-

2.5Ag sample microstructure contains coarse proeutectic Bi and fine Bi-Ag eutectics, while the hypereutectic microstructure comprises of dendritic primary Ag, proeutectic Bi, and Bi-Ag eutectics that can be seen in the Bi-11Ag sample. In a study conducted by Jenn-Ming Song, some primary Ag were also found in the Bi-2.5Ag sample at the eutectic cell boundaries. Higher Ag wt% resulted in an increased population of the primary Ag and proeutectic Bi in the Bi-11Ag sample compared to the Bi-2.5Ag sample (Jenn *et al.*, 2007b).

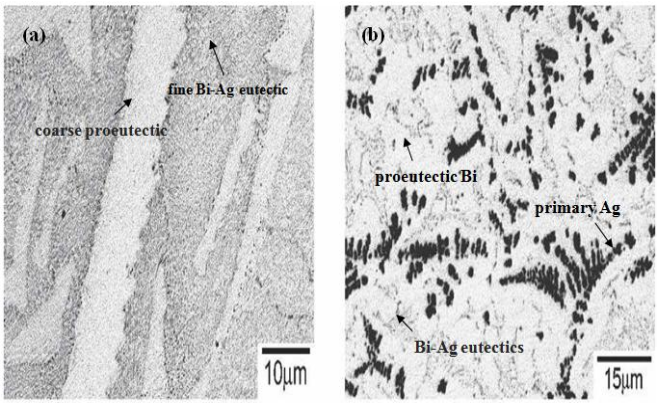



Fig.2: Microstructure of Bi-Ag solders balls; (a) Bi-2.5Ag, and (b) Bi-11Ag (Jenn *et al*., 2006)

Fig.3 shows the microstructure of the Bi-2.5Ag sample on Cu substrate after soldering for 1 min and 5 min. From the observation, there was no intermetallic compound formed but the Cu substrate became rougher, as illustrated in Fig.3a. In Fig.3a and Fig.3b, the small amount of Ag-rich compound precipitates on the Cu substrate, regardless of the reaction time. A Cu-rich phase, with acicular morphology, and the elements composition of 60 at.%Cu-35at.% Bi-5at.%Ag, were seen when the soldering time reached 5 min. Fig.2c shows the fan-like appearance of the Cu-rich needles.

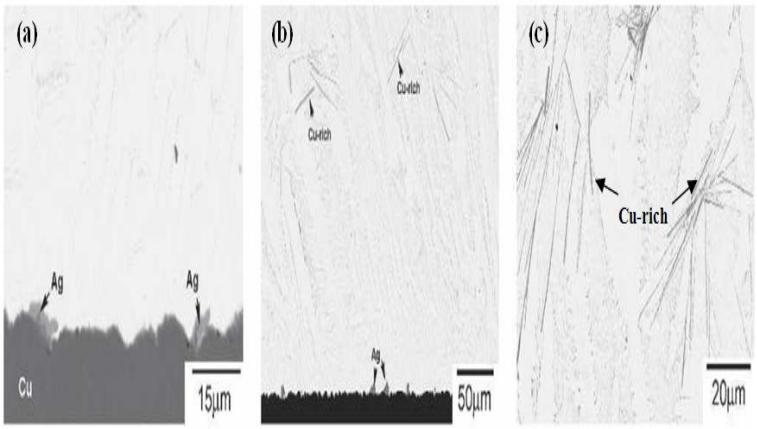



Fig.3: The structural features of the Bi-2.5Ag/Cu interface after soldering at 350°C for (a) 1 min. and (b) 5 min.; (c) Cu-rich needle within the solder after 5min (Jenn *et al*., 2006)

Fig.4 shows the structural feature of the interface between the Bi-11Ag sample and Cu substrate after the reaction at 350 °C for 1 min and 5 min. There was a similarity found between the structure and Bi-2.5Ag interface, except for the appearance of massive primary Ag and Cu-Bi-Ag needles already existing in the Bi-11Ag/Cu sample even after soldering at 350°C for only 1 min, as illustrated in Fig.4a.

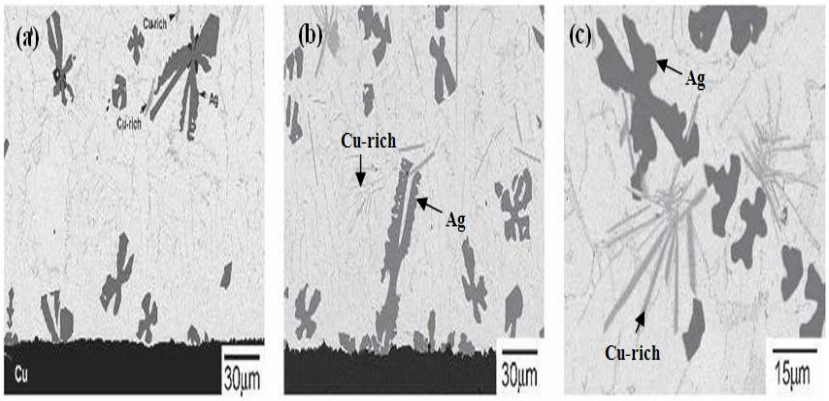



Fig.4: The structural feature of the Bi-11Ag/Cu interface after the reaction at 350°C for (a) 1 min (b) 5 min.; (c) primary Ag and Cu-rich needle within the solder (Jenn *et al*., 2006)

Fig.5 shows the magnified images of the grooved grain boundaries at Bi-11Ag/Cu interfaces, indicating the dissolution of solder at the Cu grain boundaries. The molten solder tended to penetrate into the Cu grain boundaries with the prolonged reaction time. The Ag addition accelerated the grain boundary grooving of Cu by molten solder (Jenn *et al*., 2007a).

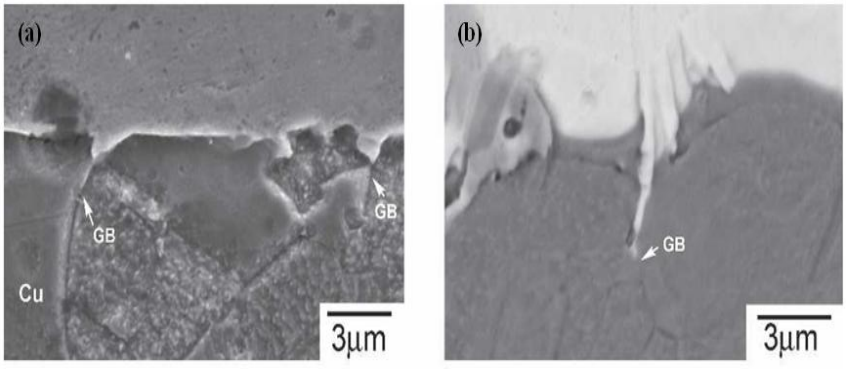



Fig.5: The grooved grain boundaries at the Bi-11Ag/Cu interface for varying reaction times: (a) 1 min. and, (b) 5 min. GB: grain boundary (Jenn *et al*., 2006)

Yaowu Shi, in his study, confirmed the findings from Jenn in 2006. He conducted a study using Bi-2.5Ag and Bi-10Ag on Cu substrate for 30 min. The microstructure of Bi-2.5Ag and Bi-10Ag solder joints shown in Fig.5 show that the Bi-2.5Ag sample mainly consists of fine Bi-Ag eutectics. Meanwhile, the non-equilibrium solidification of soldering process causes the

primary Ag particles to be found only at some visual fields and often located at the eutectic cell boundaries. However, Bi-10Ag alloy consists of a hypereutectic microstructure comprising dendritic primary Ag and Bi-Ag eutectics. Yaowu Shi also confirmed the formation of Cu-rich acicular morphology in his work. Once again, no intermetallic compound (IMC) formed at the interface between the Bi-Ag solder and substrate.

Fig.6b and Fig.6c clearly show many Ag-rich phases that are formed near the solder side at the interface. The formation of Cu-rich acicular phase results in the Ag-rich phase formation near the interface of Bi-Ag solder.

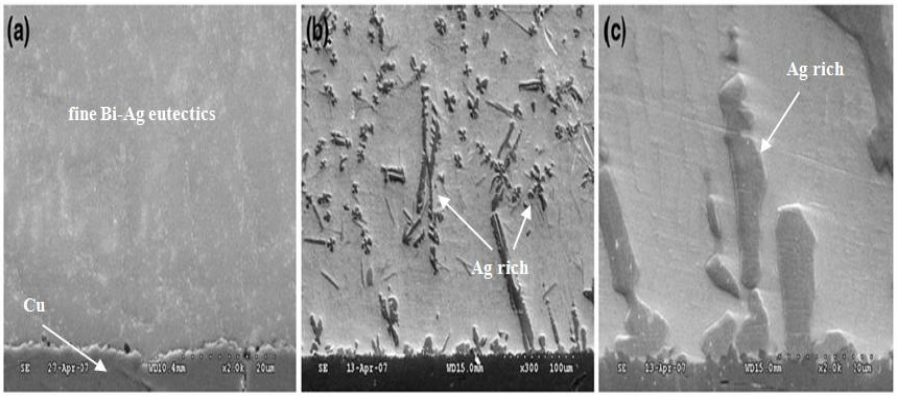



Fig.6: The microstructure of Bi-Ag solders joints; a) Bi-2.5Ag, b) Bi-10Ag, (c) Bi-10Ag with high magnification (Yaowu Shi *et al*., 2009).

The fracture surface of the Bi-11Ag/Cu sample in a study conducted by Jenn showed that fracturing mostly occurred within the solder, as shown in Fig.7. However, the Bi-2.5Ag/Cu sample, with a relatively low joint strength, showed a large area of exposed Cu, indicating an interfacial fracturing (Jenn *et al*., 2006).

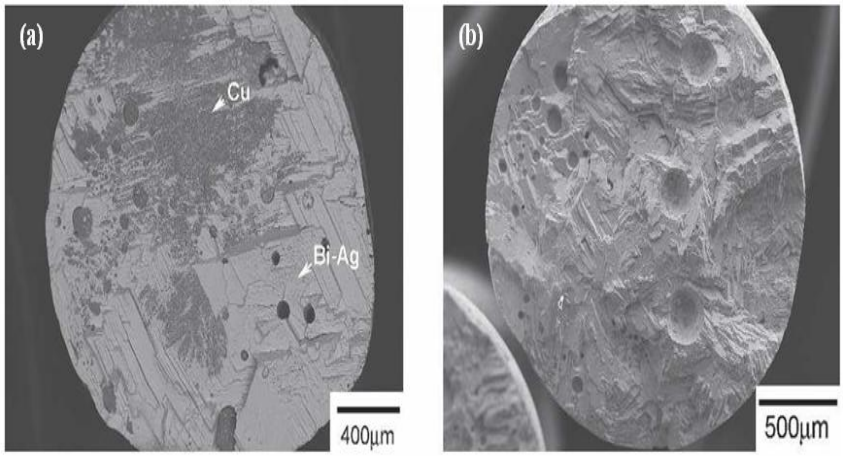



Fig.7: The tensile fracture surface of the Cu/Bi-Ag/Cu joints; (a) Bi-2.5Ag (backscattered electron image), and (b) Bi-11Ag (Jenn *et al*., 2006)

Fig.8 focuses on the cross-sectional structure and a magnified image of the fractography. Fig.8c and Fig.8d indicate that the grain boundary grooving of Cu was more pronounced in the case of the Bi-11Ag/Cu interface. It is also shown that the degree of grain boundary grooving at the Bi-11Ag interface is higher compared to the Bi-2.5Ag interface because the addition of Ag content in molten solder enhances the grooving behaviour at grain boundaries (Jenn *et al*., 2006). The surface of exposed Cu substrate is smoother for the Bi-2.5Ag interface, as illustrated in Fig.8a and Fig8.b.

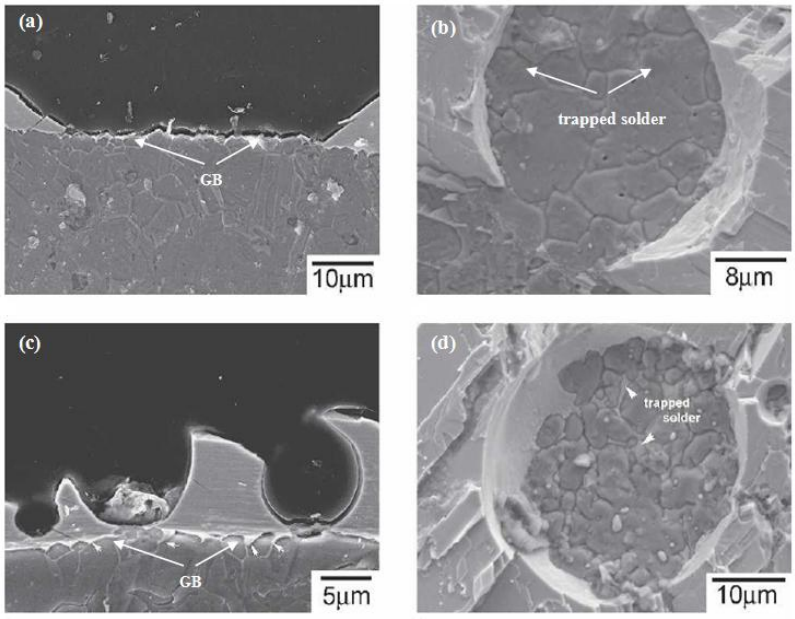



Fig.8: The cross-sectional image showing grooved grain boundaries on the exposed Cu substrate of fractured Cu/Bi-Ag/Cu joints: (a and b) Bi-2.5Ag, (c and d) Bi-11Ag. Arrows in (c) indicate grooved grain boundaries (Jenn *et al*., 2006)

*Microstructure of Bi-Ag Solder Joints with Addition of Rare Earth (RE) Element*

The effects of different Ag contents and the addition of rare earth element on the microstructure of Bi-Ag solder joints are displayed in Fig.9. In his work, Yaowu Shi describes that a higher Ag content gives an increase in the formation of dendritic primary Ag and acicular Cu-rich phase in the solder but the microstructure of the Bi-Ag solders has no obvious change with the addition of 0.1wt % rare earth element such as Ce in the solder. An acicular Cu-rich phase and the coarse Ag-rich phase are obviously refined, and this indicates that the needle-like structure is shortened by the addition of rare earth element. A few other research studies have confirmed that the addition of RE element can help to refine the microstructures of the solders, as seen in Sn-based solders such as Sn-Ag-Cu (Chen *et al*., 2002), Sn-Ag-Bi (Xia *et al*., 2002), and Sn-Bi (Dong *et al*., 2008). The unique properties of the rare earth element play an important role during the solidification process that will affect the microstructure of alloys (Wu *et al*., 2007; Zhang, 2006). The RE element tends to agglomerate at the interfaces of the primary phases

and at the grain boundaries during solidification. This phenomenon will reduce the interfacial energy and also refine grain size because RE acts as a surface active element.

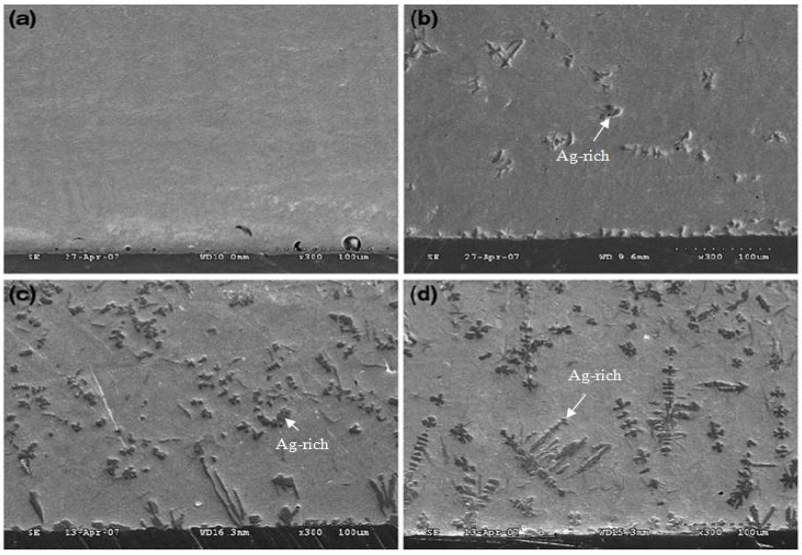



Fig.9: The effect of rare earth addition on microstructure of Bi-Ag solders joints; (a) Bi-2.5Ag-RE, (b) Bi-5Ag-RE, (c) Bi-7.5Ag-RE, (d) Bi-10Ag-RE (Yaowu Shi *et al*., 2009).

*Microstructural Evolution during Long-Time Soaking*

Fig.10a and Fig.10b display the magnified microstructure of the grain boundaries grooves formation at the interface of pure Bi/Cu and Bi-11Ag/Cu after a short period of soaking time at 350°C for 10 min and 410°C for 120 min, respectively. It is shown that a greater degree of grain boundaries grooving appeared when a higher temperature and an extended reaction time were used, as illustrated in Fig.10c and Fig.10d (Jenn *et al*., 2007a). The increasing depth of the grain boundary grooves is also dependent upon the addition of Ag related with the dissolution of the grain boundary, which is important in the interfacial reaction between Bi-Ag solders alloy and Cu substrate.

In a similar study by Jenn in 2007, soaking Cu wires in molten Bi and Bi-Ag alloys led to newborn phases in the solder with variable soaking times and temperatures (see Fig.11). The newborn phases compound elements are identified in Fig.12.

Fig.12 displays the newborn phase identified from the reaction of dissolved Cu and molten solders at 380°C for 120 min. In the interaction between pure Bi and Cu, acicular Cu-Bi phase with the composition of 59.6at. %Cu-40.4at. %Bi gathered with a fanlike appearance, as illustrated in Fig.12a and Fig.12b. Meanwhile in the Bi-11Ag and Cu wire samples, the Cu-Bi phase exhibited a round appearance and contained a certain amount of Ag (the composition was 56.3 at%Cu-39.2at. %Bi-4.5at.%Ag), as shown in Fig.12c and Fig.12d. A comparison between Fig.12a and Fig.12c reveals that the amount of Cu-Bi phase in the Bi-11Ag solder is larger than that of pure Bi because the Ag element will enhance the formation of Cu-Bi phase that reduces the Cu concentration in liquid Bi. Besides, the Cu-Bi phase has a round appearance in

Bi-11Ag solder alloys because the small amount of Cu has been replaced by the Ag element that modifies the fanlike appearance to round appearance.

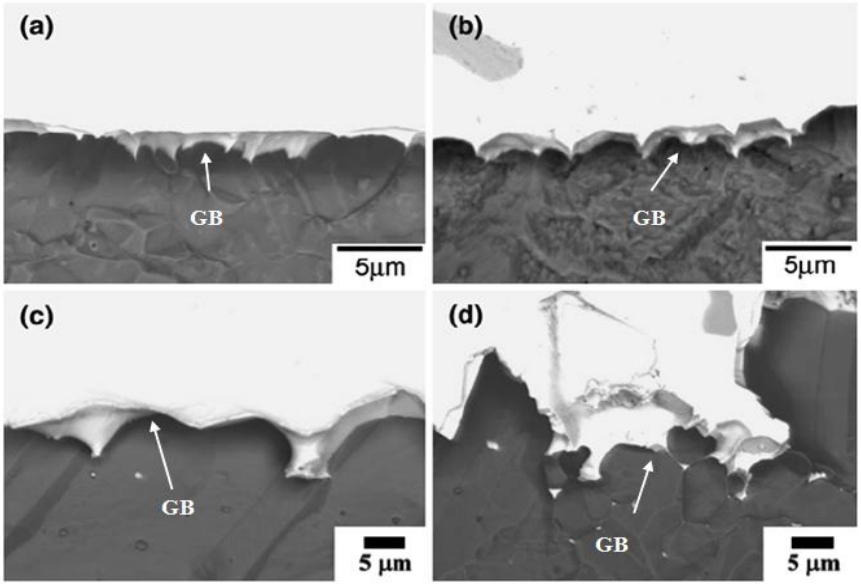



Fig.10: The interfacial morphologies of the samples after reaction under different conditions; (a) pure Bi/Cu at 350°C for 10 min, (b) Bi-11Ag/Cu at 350°C for 10 min, (c) pure Bi/Cu at 410°C for 120 min, and (d) Bi-11Ag/Cu at 410°C for 120 min (Jenn *et al*., 2007a).

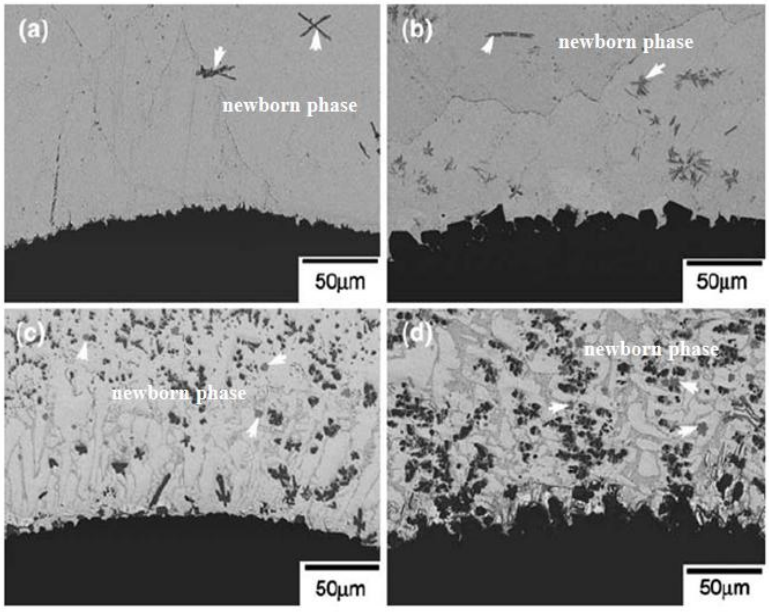



Fig.10: The microstructure in the vicinity at the interface of Bi/Cu; (a) 350°C for 120 min, and (b) 410°C for 120 min; and Bi-11Ag/Cu: (c) 350°C for 120 min, and (d) 410°C for 120 min (the gray newborn phase is indicated by the arrows) (Jenn *et al*., 2007a).

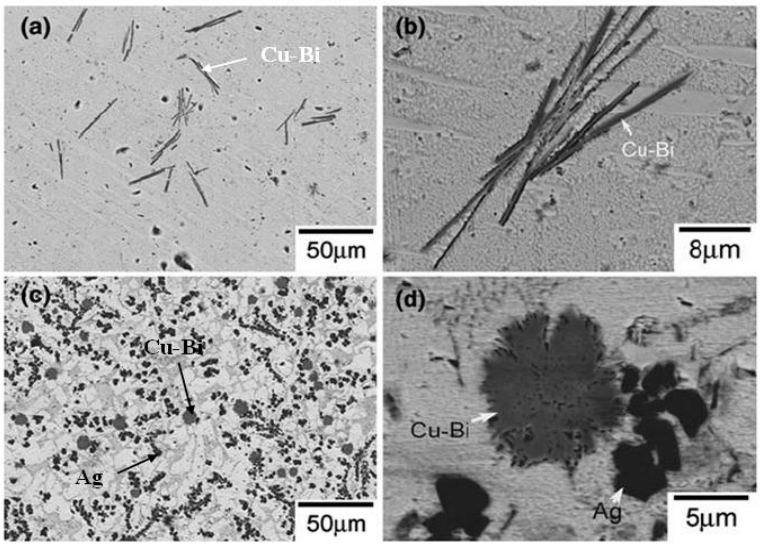



Fig.11: The micrographs of pure Bi and Bi-11Ag samples after the reaction with Cu at 380°C for 120 min; (a and b) pure Bi and (c and d) Bi-11Ag (Jenn *et al*., 2007a).

## CONCLUSION

This paper reviews the microstructure of the Bi-Ag solder alloys at the interface and tensile behaviour of Cu/Bi-Ag/Cu joints, microstructure of Bi-Ag solder joints, with the addition of Rare Earth (RE) element and microstructural evolution during the long-time soaking. The Bi-Ag alloys showed a non-equilibrium solidification structural feature. Mostly, primary Ag, Bi grains, and Bi-Ag eutectics coexisted in the microstructure that is effective for microelectronic soldering application because primary Ag enhanced the ductility of the Bi-Ag solder at high deformation rates to avoid from crack growth (Jenn *et al*., 2007b). Unfortunately, Bi-Ag alloy has some disadvantages which include low electrical conductivity and thermal conductivity with low Ag wt%. However, electrical conductivity can be increased by increasing the Ag content more than 11 wt% (Lalena *et al.,* 2002; Jenn *et al.,* 2006; Fang *et al.,* 2008). Similarly, the Bismuth solder also possesses poor wetting on substrate such as copper (Kim *et al*., 2002) and to improve the wetting properties, the solder must be alloyed with other elements such as tin and rare earth element (Katsuaki *et al*., 2009). Meanwhile, increasing Ag wt. % will improve the ductility of brittle Bi solder and effectively strengthen the solder joint. Higher Ag wt. % will also enhance the dissolution and grain boundary grooving of Cu substrate, but this also imposes a side effect, i.e. it increases the melting temperature which in real life depends on the temperature limit of the substrate (Jenn *et al*., 2007a). The Bi-Ag alloy system, coupled with the optimum level of Ag wt. % with small amount of rare earth element, has a better potential as a candidate for high leaded solders. This is because the addition of a small amount of rare earth element will not influence the melting temperature and electrical conductivity of the solder joint (Yaowu Shi *et al*., 2009). Nonetheless, further research needs to be done on Bi-Ag alloys. For example, a study is required to determine the effect of doping with other elements such as tin or rare earth element on Bi-Ag solders to substitute for the lead containing alloys.

Even with some of the limitations arising from the use of Bi-Ag solders as high temperature solder candidates such as reliability and manufacturing issues, Bi-Ag alloy is still a promising material and thus, further research work in this field is required (Kim *et al*., 2002; Rettenmayr *et al*., 2005; Song *et al*., 2006*;* Suganuma, 2007).

## REFERENCES

Chen, Z. G., Shi, Y. W., Xia, Z. D., & Yan, Y. F. (2002). *Journal of Electronic Material, 31*, 1122–1128.

Dong, W., Shi, Y., Xia, Z., Lei, Y., & Guo, F. (2008). *Journal of Electronic Material, 37*, 982–991.

Elliott, R. P., & Shunk, F. A. (1980). *Buletin of Alloy Phase Diagrams*, *1*, 62.

Environmental Protection Agency. (1991). *National Air Quality and Emission Trend Report*, 1989, EPA-450/4-91-003, Research Triangle Park, NC.

Fang, W., Shi, Y., Xia, Z., & Guo, F. (2008). *Electron Component Material*, *27*, 15-18.

Humpston, G., & Jacobson, D. (1993). *Principles of Soldering and Brazing*, ASM International, Materials Park, OH.

Jenn-Ming, S., Hsin-Yi, C., & Tien-Xiang, W. (2007b). Thermal and Tensile Properties of Bi-Ag Alloys, *Journal of Metallurgical and Materials Transactions A*, 38A, 1371-1375.

Jenn-Ming, S., Hsin-Yi, C., & Zong-Mou, W. (2006). Interfacial Reactions between Bi-Ag High-Temperature Solders and Metallic Substrates*, Journal of Electronic Materials, 35*, 1041–1048.

Jenn-Ming, S., Hsin-Yi, C., & Zong-Mou, W. (2007a). Substrate Dissolution and Shear Properties of the Joints between Bi-Ag Alloys and Cu Substrates for High-Temperature Soldering Applications, *Journal of Electronic Materials*, *36*, 1516-1523.

Kattner, U. R. (2002). *Phase Diagrams for Lead-Free Solder Alloys*. National Institute of Standards and Technology.

Kim, J. H., Jeong, S. W., & Lee, H. M. (2002). *Material Transition, 43*, 1873–1877.

Lalena, J. N., Dean, N. F., & Weiser, M. W. (2002). *Journal of Electronic Material, 31*, 1244–1249.

Lea, C. (1988). A Scientific Guide to Surface Mount Technology, *Journal of Electrochemical Publications Ltd.,* 569.

Monsalve, E. R. (1984). Lead Ingestion Hazard in Hand Soldering Environments, in: *Proceedings of the 8th Annual Soldering Technology and Product Assurance Seminar*, Naval Weapons Center, China Lake, CA.

Napp, D. (1995). Lead-free interconnect materials for the electronics industry, *Proceedings of the 27th International SAMPE Technical Conference*, Albuquerque, NM, 342.

Nriagu, J. O., & Pacyna, J. M. (1988). Quantitative Assessment of Worldwide Contamination of Air, Water and Soils by Trace Metals. *Nature, 333*(6169), 134-139.

Rettenmayr, M., Lambracht, P., & Kempf, B. (2005). *Advance Engineering Material, 7*, 965–969.

Sax, N. Ir. (1984). *Dangerous Properties of Industrial Materials* (6th edition). New York: Van Nostrand Reinhold Company.

Suganuma, K. (2007). *Polymer and Adhesives in Microelectronics and Photonics* (Piscataway, NJ: IEEE), 30–35.

Suganuma, K., Seong-Jun, K., & Keun-Soo, K. (2009). *High Temperature Lead-Free Solders: Properties and Possibilities, 61*, 64-71.

Wood, E.P. and Nimmo, K.L. (1994). In Search of New Lead-Free Electronic Solders, *Journal of Electronic Material, 23*, 709-713.

Wu, C. M. L., & Wong, Y. W. (2007). *Journal of Material Science: Material Electron, 18*, 77–91.

Xia, Z., Chen, Z., Shi, Y., Mu, N., & Sun, N. (2002) *Journal of Electronic Material, 31*, 564–567.

Yaowu, S., Weiping, F., Zhidong, X., Yongping, L., Fu, G., & Xiaoyan, L. (2009). Investigation of Rare Earth-Doped Bi-Ag High Temperature Solders. *Journal of Material Science: Material Electron, 21*, 875–881.

Zhang, Z. (2006). *Physicochemistry of Materials,* Chemical Industry Press, p. 163.

Pertanika J. Sci. & Technol. 22 (1): 15 - 24 (2014)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

*Short Communication*

# FRFT Based Timing Estimation Method for an OFDM System

**Saxena, R.[1] and Joshi, H. D.[2]***

[1]*Director, Jaypee University, Anoopshahr, District Bulandshahr– 203390 (U.P.), India*
[2]*Department of ECE, Thapar University, Patiala – 147001 (Punjab), India*

## ABSTRACT

The Orthogonal Frequency Division Multiplexing (OFDM) is a multi-carrier modulation (MCM) technique which is adopted by many wireless communication standards for transmitting data at very high rates over time dispersive radio channels. In an OFDM system, the timing estimation is extremely important for maintaining orthogonality among the subcarriers. In this paper, a method of timing estimation is proposed for an OFDM system. The proposed method used chirp signal as a training sequence and employed the fractional Fourier transform (FRFT) as a tool to localize the training sequence (chirp) at the receiver. The comparative study showed the superiority of the proposed estimator in terms of mean and MSE of timing offset. The MSE of timing offset with proposed method was found to be 76% (5 dB SNR) and 63% (8 dB SNR) lower than Awoseyila *et al.'s* method in *HIPERLAN/2 indoor channel-A* and in *Wi-Max system* (strong fading channel), respectively. However, the improvement in MSE is obtained in the proposed method at the cost of increased computational complexity, in terms of $\frac{N}{4}-1+N\log_2 N$ more complex multiplication than the Awoseyila *et al.'s* method.

*Keywords:* Timing Estimation, OFDM, FRFT, Chirp

## INTRODUCTION

The Orthogonal Frequency Division Multiplexing (OFDM) has drawn major attention over the last decade for its usefulness in broad band wireless communication. Due to advantageous features such as high spectral efficiency, robustness to fading channel and easy equalization, the OFDM has been adopted as a major data transmission technique by many wireless communication standards such as IEEE 802.11a, IEEE 802.16a and terrestrial digital video broadcasting (DVB-T) systems (Keller & Hanzo, 2000).

*Article history:*
Received: 1 November 2011
Accepted: 20 February 2012

*E-mail addresses:*
srajiv2008@gmail.com (Saxena, R.),
hemdutt@gmail.com (Joshi, H. D.)
*Corresponding Author

ISSN: 0128-7680 

However, one of the major disadvantages of OFDM system is its sensitivity to synchronization errors (time and frequency) at the receiver. Inaccurate estimation of timing and carrier frequency offset destroys the orthogonality among sub-carrier, which generates inter-carrier interference (ICI) and inter-block interference (IBI) (Morelli *et al*., 2007). Several methods which are based on the transmission of known preamble have been proposed in the past for correct timing and frequency estimation, either jointly or individually (Schmidl & Cox, 1997; Minn *et al*., 2000; Park *et al*., 2003; Shi & Serpedin, 2004; Ren *et al*., 2005; Awoseyila *et al*., 2008). The basic concept of these methods is to transmit a preamble which consists of some repetitive blocks and then, applying a sliding window correlator at the receiver to detect the maximum of timing metric. The timing metric with a sharp peak is desired for better timing estimate.

In the method by Schmidl and Cox (S&C) (1997), the timing metric exhibits a large plateau which causes a large variance in timing estimation. In order to reduce the variance, more algorithms are given by (Minn *et al*., 2000; Park *et al*., 2003; Shi & Serpedin, 2004; Ren *et al*., 2005; Awoseyila *et al*., 2008; Boumard & Mammela, 2009) with sharper timing metric and less variance. In this paper, a new timing estimation method is proposed which is based on localization of chirp signal with the help of fractional Fourier transform (FRFT).

In the proposed estimator, the emphasis is on the use of fractional Fourier transform based correlation and chirp signal. The application of chirp signal for the synchronization in OFDM system along with its advantages is very well documented by Boumard and Mammela (2009). The chirp signals have characteristics such as finite duration, finite bandwidth and better auto-correlation property. The FRFT is a generalization of the conventional Fourier transform in time-frequency plane and has found applications in non-stationary signals (chirp signal) analysis, especially in filtering, time delay estimation and radar signal processing (Almeida, 1994; Sun *et al*., 2002; Sharma & Joshi, 2007; Tao *et al*., 2009; Singh & Saxena, 2011). It is due to the fact that the FRFT nicely localizes the chirp signal in time-frequency plane (Tao *et al*., 2009). This characteristic of FRFT on chirp signal has been exploited in the proposed estimator.

Various properties of the FRFT have already been derived and established, as given by Almeida (1994). Recently, the new weighted convolution and correlation theorems in the FRFT domain have been given by Singh *et al*. (2011). The correlation theorem given by Singh and Saxena (2011) is used in the proposed timing estimation method.

The performance of proposed timing estimator is compared with the algorithms given by Schmidl and Cox (1997), Minn *et al.* (2000), Park *et al.* (2003), Shi and Serpedin (2004) and Awoseyila *et al.* (2008). It was found that the mean and mean-square error (MSE) of timing offset with proposed algorithm was better than other algorithms as mentioned above, both in ***HIPERLAN/2 indoor channel-A*** (Channel Models, 1998) *and* in strong fading channel (Wi-Max with 256 sub-carriers).

## SYSTEM MODEL

An OFDM system with N sub-carrier has been considered for the present analysis. The received signal $y(n)$ from a multipath fading channel of memory length L may be represented as:

$$y(n) = \sum_{m=0}^{L-1} h(m)\ s(n-m)\ , \quad n = 0,1,2.....N-1 \tag{1}$$

where, $h(m)$ is the channel impulse response of $m$th path and $s(n)$ is the transmitted time domain OFDM signal expressed by:

$$s(n) = \sum_{k=0}^{N-1} d_k\ e^{j2\pi k\ n/N},\ n = 0,1,2.....N-1 \tag{2}$$

where, $d_k$ is the complex data symbol modulated on the $k$th sub-carrier. At the receiver, timing offset is considered as a delay in the received signal and frequency offset is taken as phase distortion of received data in time domain. After considering the effects of timing offset, frequency offset, and AWGN noise in the received signal, the y(n) given by (1) can be re-written as:

$$r(n) = y(n-\tilde{d})\ e^{j2\pi\varepsilon n/N} + w(n) \tag{3}$$

where, $\tilde{d}$ is the integer-valued unknown arrival time of a symbol, 'ε' is the frequency offset normalized by the sub-carrier spacing ($1/NT_s$), $w(n)$ is a zero-mean, complex value Gaussian noise process with variance $\sigma_w^2$, and $T_s$ is the sampling interval. The aim of timing synchronization is to estimate the value of $\tilde{d}$. The timing estimation is performed by feeding the received time-domain samples to a sliding window correlator. The output of this correlator is expected to exhibit a peak when the sliding window is perfectly aligned with the received reference (preamble) block. The resulting timing estimate $\tilde{d}$, as given by Morelli *et al*. (2007) is:

$$\tilde{d} = \arg\left( \max_{d} \left\{ \left| M(d) \right| \right\} \right) \tag{4}$$

where, $M(d)$ is the timing metric.

## PROPOSED ESTIMATOR

The proposed timing offset estimation method has exploited the basic property of the chirp signal which showed that it peaked in the fractional Fourier domain corresponding to an optimum angle $\acute{a}_{opt}$' associated with the FRFT kernel. This particular characteristic of the FRFT on chirp signal has been used for searching the start of training sequence in received signal. Therefore, at the receiver side, the FRFT of auto-correlation of received signal is determined and its peak is observed. Almeida (1994) defines the FRFT of a signal x(t) at an angle $\alpha$ as:

$$\Im[x(t)] = X_\alpha(u) = \int_{-\infty}^{\infty} x(t)\ K_\alpha(t,u)\ dt \tag{5}$$

where, $K_{\acute{a}}(t,u)$ is the kernel of FRFT, which is given as:

$$K_{\alpha}(t,u)=\begin{cases}\sqrt{\dfrac{1-j\cot(\alpha)}{2\pi}}\; e^{(j/2)\left\{(t^2+u^2)\,Cot(\alpha)\;-\;2\,t\,u\,Cosec(\alpha)\right\}} & \\ & ;\quad if\ \alpha\neq n\pi \\ \delta(t-u) & ;\ if\ \alpha=2n\pi \\ \delta(t+u) & ;\ if\ \alpha=(2n+1)\pi\end{cases} \tag{6}$$

In order to make the transform technique compatible with the discrete signal as normally encountered in signal processing applications, the discrete version of the transform technique is needed. Therefore, the sampling type discrete FRFT (DFRFT) algorithm as given by Ozaktas *et al*. (1996) has been considered for the calculation of the FRFT of the signal in the proposed scheme.

The training sequence of proposed method is a chirp signal with chirp rate '2a', duration 'τ' and bandwidth '2a τ', defined as:

$$x(t)=e^{j\,2\,\pi\,(\,at^2+bt\,)}\;rect\left[t/\tau\right] \tag{7}$$

where, b is the centre frequency of chirp signal. It is known that a finite duration chirp signal can be concentrated maximally in the fractional Fourier domain with an angle $\alpha_{opt}$, which is determined by the chirp rate '2a'of the signal satisfying the relation as given by (Ozaktas *et al*., 1996; Tao *et al*., 2009):

$$\cot(\alpha_{opt})=-4\pi a \tag{8}$$

To illustrate this behaviour of chirp signal, a simulation exercise was performed by considering the chirp signal with chirp rate of 16 and unity duration. Thereafter, FRFT of chirp signal at optimum angle $\left(\alpha_{opt}=3.13165\right)$ and two nearby angles (α = 3.12 and α = 3.15) was determined, as shown in fig.1. It can easily be depicted from the results that at an optimum angle, the FRFT of chirp has a sharp and pronounced peak in comparison to other nearby angles. This conforms to the fact that the FRFT transforms a chirp signal into a delta function at an optimum angle corresponding to the chirp rate associated with the chirp signal.

The correlation theorem for FRFT domain as given by Singh & Saxena (2011) is:

$$r_{xx}(\tau)\;\leftrightarrow\;R_{xx}(u) \tag{9}$$

where, $r_{xx}(\tau)$ is defined as weighted auto-correlation of signal x(t) and $R_{xx}(u)$ is defined as FRFT of $r_{xx}(\tau)$, which are defined as:

$$r_{xx}(\tau)=\int_{-\infty}^{\infty}x(t)\,x(t+\tau)\;e^{j\,t\,(\,\tau+t\,)\,Cot\,(\alpha)}\;dt \tag{10}$$

$$R_{xx}(u) = \sqrt{\frac{2\pi}{1-j\,Cot(\alpha)}}\, e^{\frac{-j}{2}\left\{ u^2\, Cot(\alpha)\right\}} X_{\alpha}(-u)\, X_{\alpha}(u) \tag{11}$$

From (11), it is evident that the FRFT at an angle $\alpha$ of the time-domain weighted auto-correlation ($r_{xx}(\tau)$) is equivalent to the multiplication of signal's FRFT $(X_{\acute{a}}(u))$, the mirror image of signal's FRFT$(X_{\acute{a}}(-u))$, a chirp function and a scaling factor. The proposed timing estimation method is based on the auto-correlation of received signal. However, the timing metric is generated by using (11). The algorithm to generate timing metric is described in the following paragraph.

A sliding window (N samples) method is used to generate timing metric $M(d)$. The 'd' is a time index. First, the DFRFT of every N samples of received signal is determined at the optimum angle $\acute{a}_{opt'}$. Then, this DFRFT is used to calculate the $R_{xx}(u)$ as given in (11). The maximum of $|R_{xx}(u)|$ is used to give the value of timing metric corresponding to index 'd'. This window slides sample by sample as the receiver searches for the training symbol. Therefore, the timing metric of proposed method is defined as:

$$M(d) = \max_{u} \left| \sqrt{\frac{2\pi}{1-j\,Cot(\alpha)}}\, e^{\frac{-j}{2}\left\{ u^2\, Cot(\alpha)\right\}} \left(X_{\alpha_{opt},d}(u)\right)^2 \left(X_{\alpha_{opt},d}(-u)\right)^2 \right| \tag{12}$$

where, $X_{\alpha_{opt},d}(u)$ is a N-point DFRFT of $x$(n+d). The maximum of timing metric $M(d)$ gives the starting point of the training symbol.

The timing metric of the proposed method, under no noise and no channel condition is shown in Fig.2. The 1024 sub-carrier OFDM system with 128 cyclic prefix has been considered for the generation of this timing metric. The correct timing point is indexed as 0 in the figure. The timing metric of algorithms given by Schmidl and Cox (1997), Minn *et al.* (2000) and Park *et al.* (2003) is also shown in Fig.2 for comparison.

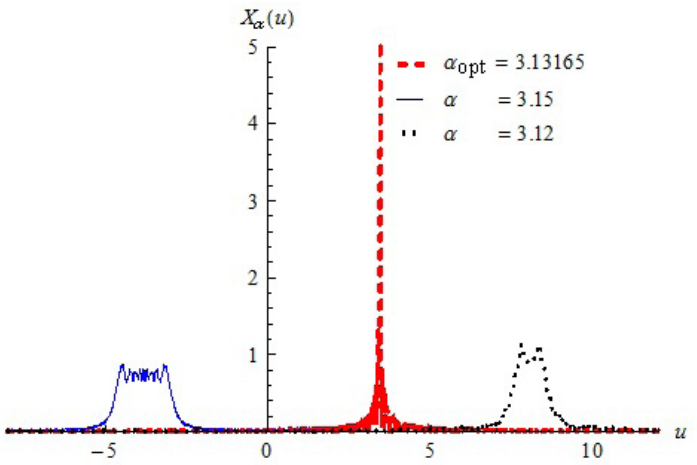



Fig.1: FRFT of a chirp signal (7) for different angle $\alpha$ with a = 8, b=55

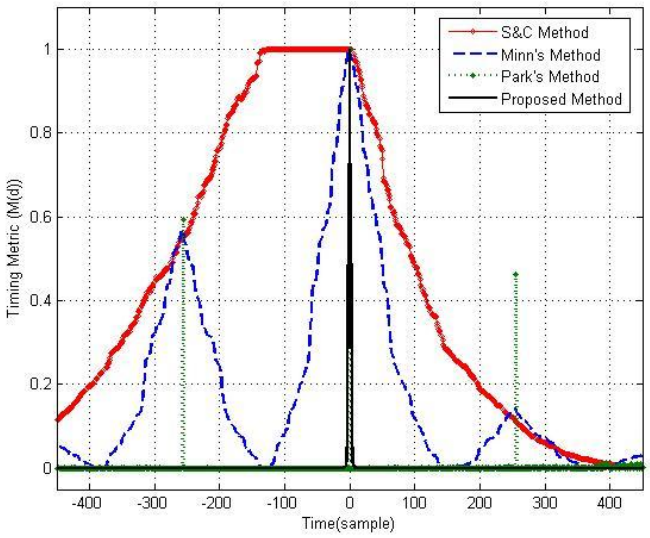



Fig.2: Timing Metric of Different Estimators

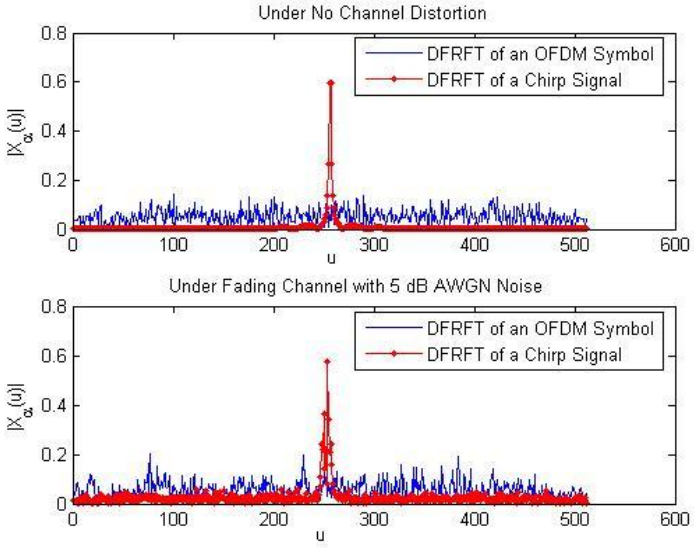



Fig.3: DFRFT of an OFDM and chirp signal

The timing metric of Schmidl and Cox's method has a plateau of length equal to length of cyclic prefix (128 in the present example). This plateau is reduced in Minn's method. The timing metric with Park's method has a sharp peak at correct time but some side peaks are also visible, whereas in the proposed method, the timing metric has a sharp peak with no side peaks and is clearly evident from the plots.

In order to explain the superiority of the proposed timing metric over other included methods in Fig.2, the DFRFT of an OFDM symbol and of a chirp signal under both conditions (no

channel distortion and channel distortion with noise) are shown in Fig.3. An OFDM symbol with 512 sub-carriers has been considered for this analysis and a chirp signal defined in (7) is taken with values of 'a' = 435 and 'b' =76. It is clearly visible from the plots (Fig.3) that the maximum value of DFRFT of chirp signal is much larger than the maximum value of DFRFT of OFDM symbol. This concept is used to search the preamble in received signal.

## PERFORMANCE EVALUATION

In this section, the performance of the proposed method is presented and compared with the methods given by Schmidl and Cox (1997), Minn *et al.* (2000), Park *et al.* (2003), Shi and Serpedin (2004) and Awoseyila *et al.* (2008). In this comparison, an OFDM system with 64 sub-carriers and 16 cyclic prefix with QPSK modulation is considered. The HIPERLAN/2 indoor channel model (Channel Models, 1998) is used for simulations. A normalized frequency offset of 0.1 is considered to explore the robustness of proposed method. The mean and mean square error was taken as performance evaluation parameter.

Fig.4 and Fig.5 show the comparison of mean and MSE of timing offset with all four estimators in the ***HIPERLAN/2 indoor channel-A***. It is clearly visible from Fig.4 that the value of mean with proposed method is very low (lies in the range of 2 to 0.1) at all the signal to noise ratios (SNR). The values of MSE of timing offset with proposed method are much better as compared to other estimators, as shown in Fig.5.

Simulation results for advance OFDM system like Wi-Max with 256 sub-carriers are also presented in Fig.6. An ISI channel consisting of $L$=8 paths with path delays of $m_i = 0,1,\ldots,L-1$ samples and an exponential power delay profile having average power of $e^{-m_i/L}$ has been considered for simulation.

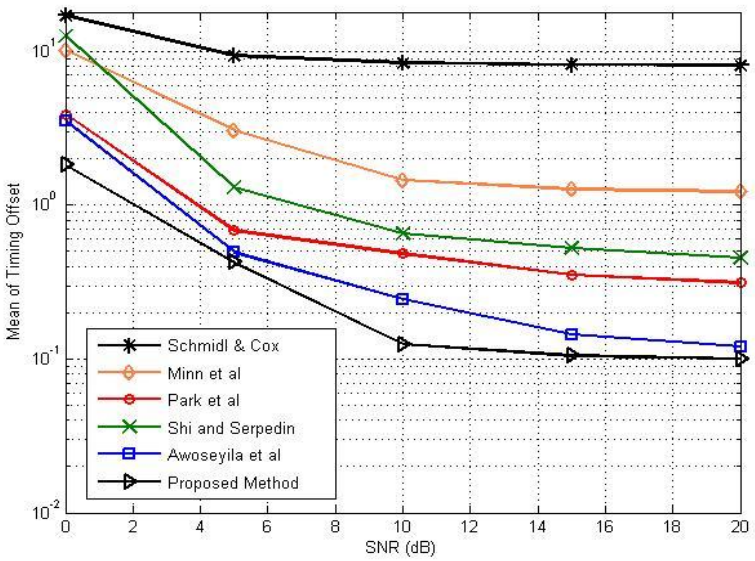



Fig.4: Mean of Timing offset estimation of different methods in HIPERLAN/2 indoor channel-A

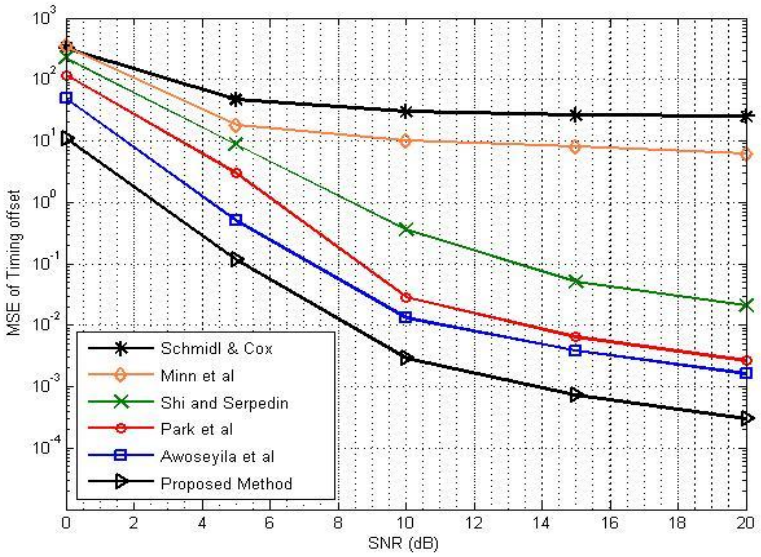



Fig.5: MSE of Timing offset estimation of different methods in HIPERLAN/2 indoor channel-A

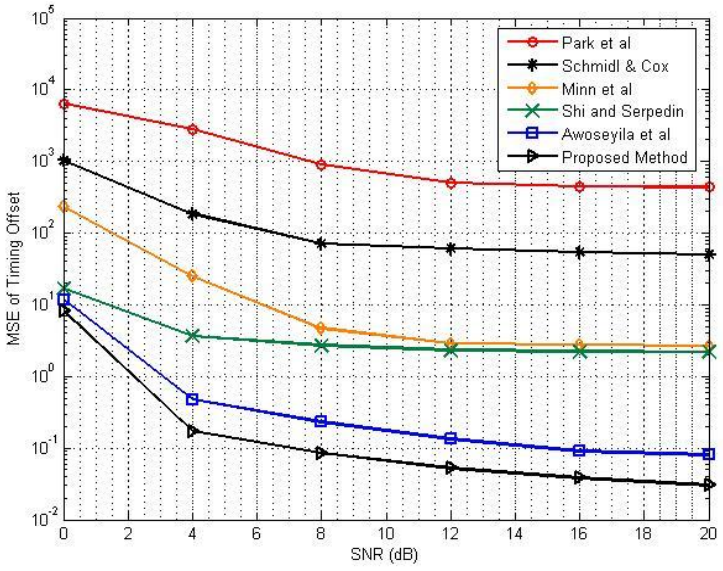



Fig.6: MSE of Timing offset estimation of different methods in Wi-Max system

This improvement in the performance is due to the impulse–shape like timing metric of proposed method in contrast to the timing metric of other methods, as discussed in the previous sections and shown in Fig.2. In addition, as shown in Fig.3, the effect of fading channel and noise is less on chirp signal. The DFRFT of chirp signal gives one sharp peak in fading channel in comparison to OFDM symbol. This concept has been used for generating timing metric. Therefore, the timing offset with proposed timing metric is lesser in both types of channel.

The cost of this improvement is complexity in terms of complex multiplications and additions. The complexity of the proposed method is high in comparison to other methods as shown in Table 1 because other methods generate timing metric by taking auto-correlation in time domain directly whereas the proposed method firstly takes the DFRFT of received signal and then generates timing metric, as given in expression (12).

TABLE 1
Approximate Computational Complexity of Various Estimators

| Methods | Complex multiplication | Complex Addition |
|---|---|---|
| Schmidl & Cox (S&C,1997) | $N/2$ | $N/2-1$ |
| Minn *et al* (2000) | $N/2$ | $N/2-1$ |
| Park *et al* (2003) | $N/2+1$ | $N/2$ |
| Shi & Serpedin (2004) | $3N/2$ | $3N/2-1$ |
| Awoseyila *et al* (2008) | $(7N+8)/4$ | $(7N-8)/4$ |
| Proposed | $(N\ \log_2 N)+2N+1$ | $N\ \log_2 N$ |

## CONCLUSION

A new method is proposed for timing offset estimation in an OFDM system. The proposed method is based on the localization of chirp signal in the FRFT domain. The suggested algorithm for timing offset estimation in both ***HIPERLAN/2 indoor channel-A*** and in strong fading channel (**Wi-Max** with 256 sub-carriers) provides the lowest MSE when compared to other available methods. However, to obtain such improvements in MSE of timing offset estimation, the computational complexity (in terms of number of complex additions and multiplications) increases. The proposed algorithm of timing offset estimation with better accuracy will certainly improve the performance metric of an OFDM system by way of making the ICI and bit error rate better. This augmentation of performance metric is a better proposition even at the cost of enhanced computational complexity.

## REFERENCES

Almeida, L. B. (1994). The fractional Fourier transform and time–frequency representations. *IEEE Trans. Signal Process, 42*(11), 3084–3091.

Awoseyila, A. B., Kasparis, C., & Evans, B.G. (2008). Improved preamble aided timing estimation for OFDM systems. *IEEE Commun. Lett*., *12*(11), 825-827.

Boumard, S., & Mammela, A. (2009). Robust and Accurate Frequency and Timing Synchronization using Chirp Signals. *IEEE Trans. Broadcasting*, *55*(1), 115-123.

Channel Models for HIPERLAN/2 in Different Indoor Scenarios (1998), *ETSI BRAN* 3ERI085B.

Keller, T., & Hanzo, L. (2000). Adaptive multicarrier modulation: A convenient framework for time-frequency processing in wireless communications. *Proc. IEEE*, *88*(5), 611–640.

Minn, H., Zeng, M., & Bhargava, V. K. (2000). On timing offset estimation for OFDM systems. *IEEE Commun. Lett.*, *4*(7), 242–244.

Morelli, M., Kuo, C.-C. J., & Pun, M.-O. (2007). Synchronization Techniques for Orthogonal Frequency Division Multiple Access (OFDMA) A Tutorial Review. *Proc. IEEE, 95*(7), 1394-1427.

Ozaktas, H. M., Arikan, O., Kutay, M.A., & Bozdagi, G. (1996). Digital computation of the fractional Fourier transform. *IEEE Trans. Signal Process, 44*(9), 2141–2150.

Park, B., Cheon, H., Kang, C., & Hong, D. (2003). A Novel Timing Estimation Method for OFDM Systems. *IEEE Commun. Lett., 7*(5), 239-241.

Ren, G., Chang, Y., Zhang, H., & Zhang, H. (2005). Synchronization methods based on a new constant envelope preamble for OFDM systems. *IEEE Trans. Broadcasting*, *51*(1), 139-143.

Schmidl, T. M., & Cox, D. C. (1997). Robust frequency and timing synchronization for OFDM. *IEEE Trans. Commun., 45*(12), 1613–1621.

Sharma, K. K., & Joshi, S. D. (2007). Time delay estimation using fractional Fourier transform. *Signal Processing, 87*(5), 853–865.

Shi, K., & Serpedin, K. (2004). Coarse frame and carrier synchronization of OFDM systems: A new metric and comparison. *IEEE Trans. Wireless Commun., 3*(4), 1271–1284.

Singh, A. K., & Saxena, R. (2011). Correlation theorem for fractional Fourier transform. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 4*(2), 31-39.

Singh, A. K., & Saxena, R. (2011). *On convolution and product theorems for FRFT, Wireless Personal Communications*. DOI 10.1007/s11277-011-0235-5.

Singh, A. K., & Saxena, R. (2011). Recent developments in FRFT, DFRFT with their applications in signal and image processing. *Recent Patents on Engineering*, *5*(2) 113-138.

Sun, H. B., Liu, G. S., Gu, H., & Su, W. M. (2002). Application of the fractional Fourier transform to moving target detection in airborne SAR. *IEEE Trans. on Aerospace and Electronics Systems*, *38*(4), 1416-1424.

Tao, R., Li, X. M., Li, Y., & Wang, Y. (2009). Time-Delay Estimation of Chirp Signals in the Fractional Fourier domain. *IEEE Trans. Signal Process, 57*(7), 2852–2855.

Pertanika J. Sci. & Technol. 22 (1): 25 - 33 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# New Recursive Circular Algorithm for Listing All Permutations

**Sharmila Karim[1]*, Zurni Omar[1], Haslinda Ibrahim[1], Khairil Iskandar Othman[2] and Mohamed Suleiman[3]**

*[1]College of Art and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*
*[2]Mathematics Department, Faculty of Information Technology and Quantitative Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia*
*[3]Mathematics Department, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

Linear array of permutations is hard to be factorised. However, by using a starter set, the process of listing the permutations becomes easy. Once the starter sets are obtained, the circular and reverse of circular operations are easily employed to produce distinct permutations from each starter set. However, a problem arises when the equivalence starter sets generate similar permutations and, therefore, willneed to be discarded. In this paper, a new recursive strategy is proposed to generate starter sets that will not incur equivalence by circular operation. Computational advantages are presented that compare the results obtained by the new algorithm with those obtained using two other existing methods. The result indicates that the new algorithm is faster than the other two in time execution.

*Keywords:* Algorithm, circular, permutation, starter sets

## INTRODUCTION

The generation of all $n!$ permutations of $n$ elements is a fundamental problem in combinatorics and important in computing. Various methods on listing all permutations have been published and can be classified into two categories: (i) exchange-based techniques; (ii) non-exchange-based techniques (Sedgewick, 1977). The exchange-based techniques generate new permutations by making possible changes among two consecutive elements such as transposition of non-adjacent elements (Well, 1961; Heap, 1963), and transposition with adjacent elements (Trotter, 1962; Johnson, 1963; Ives, 1976; Gao & Wang, 2003; Viktorov, 2007; Borisenko *et al*. 2008). Whilst non-exchange-based techniques generate new permutations

*Article history:*
Received: 7 April 2011
Accepted: 6 August 2013

*E-mail addresses:*
mila@uum.edu.my (Sharmila Karim),
zurni@uum.edu.my (Zurni Omar),
linda@uum.edu.my (Haslinda Ibrahim),
khairil17@tmsk.uitm.edu.my (Khairil Iskandar Othman),
mohameds@upm.edu.my (Mohamed Suleiman)
*Corresponding Author

ISSN: 0128-7680 

with certain restrictions such as lexicographic order (Ord-Smith, 1970), nested cyclic (Langdon, 1967), and partial reversion (Zaks, 1984; Shin, 2002; Thongchiew, 2007).

According to Sedgewick (1977), generating permutation under cycling restrictions is simpler and more powerful compared to other restriction techniques. Langdon (1967) and Iyer (1995) proposed a cycling technique where the main idea is to start with cycling interchange of $n$ elements until two elements are cycled. However, Iyer's (1995) technique is only valid for $n$ <5 because repetition of permutation occurs when $n > 4$.

In spite of that, Ibrahim *et al*. (2010) introduced a new permutation technique based on distinct starter sets that employ circular and reversing operations. The crucial task of these operations is to generate the distinct starter sets by eliminating the equivalence starter sets. Although this technique is simple and easy to use, unfortunately, eliminating the equivalence starters is quite tedious when the number of elements increases. This paper attempts to overcome this drawback by introducing a new strategy for generating distinct starter sets without eliminating the equivalence starter sets.

## MATERIALS AND METHODS

*Preliminary definition*

The following definitions will be used throughout this paper.

Definition 1. A starter set is a set that is used as a basis to enumerate other permutations.

Definition 2. An equivalence starter set is a set that can produce the same permutation from other starter sets.

Definition 3. The reverse set is a set that is produced by reversing the order of the permutation set.

Definition 4. A Latin square of order $n$ is an $n \times n$ array in which $n$ distinct symbols are arranged where each symbol occurs once in each row and column.

Definition 5. The circular permutation (CP) of order $n$ is a Latin square of order n.

Definition 6. The reverse of circular permutation (RoCP) is also a Latin square of order $n$ which is obtained by reversing arrangement elements in each row of circular permutation.

Example 1. Consider $n = 4$ and the fixed element is 1. There are two starters: (1234) and (1432).

The circular process is applied on both starters. The CP of each starter is listed as followed:

| | |
|---|---|
| 1234 | 1432 |
| 2341 | 4321 |
| 3412 | 3214 |
| 4123 | 2143 |

We may then apply the reversing process to either CP of the starter sets e.g. (1234) and RoCP as follows:

| 4321 |
|---|
| 1432 |
| 2143 |
| 3214 |

The RoCP of the starter (1234) generates the same permutation CP of (1432). Therefore, we refer to (1432) as the equivalence starter set of (1234). That equivalence starter set needs to be discarded. With a new algorithm, the equivalence starter sets will not be generated.

*The development of the algorithm*

The general algorithm for permutation generation follows:

Let $S$ be the set of $n$ elements i.e. $(1,2,3,\ldots,n-3,n-2,n-1,n)$

Step 1: Set $(1,2,3,4,\ldots,n-3,n-2,n-1,n)$ is taken as the initial permutation and it is assumed to be without loss of generality; therefore, the first element is fixed.

Step 2: Identify the last three elements of initial permutation from Step 1. By employing CP to the last three elements in initial permutation from step 1 three other distinct starter sets are produced, as shown below:

1, 2, …, $n$−3, $n$−2, $n$−1, $n$
1, 2, …, $n$−3, $n$−1, $n$, $n$−2
1, 2, …, $n$−3, $n$, n−2, $n$−1

Step 3: Identify the last four elements of each starter set in Step 2. By employing CP to the final four elements in each starter set in Step 2, 12 distinct starter sets are obtained, as shown below.

| **1, 2, …, $n$−3, $n$−2, $n$−1, $n$** | **1, 2, …, $n$−3, $n$−1, $n$, $n$−2** | **1, 2, …, $n$−3, $n$, $n$−2, $n$−1** |
|---|---|---|
| 1, 2,…, $n$−2, $n$−1, n, $n$−3 | 1, 2, …, $n$−1, $n$, $n$−2, $n$−3 | 1, 2, …, $n$, $n$−2, $n$−1, $n$−3 |
| 1, 2, …, $n$−1, $n$, $n$−3, $n$−2 | 1, 2, …, $n$, $n$−2, $n$−3, $n$−1 | 1, 2, …, $n$−2, $n$−1, $n$−3, n |
| 1, 2, …, $n$, $n$−3, $n$−2, $n$−1 | 1, 2, …, $n$−2, $n$−3, $n$−1, $n$ | 1, 2, …, $n$−1, $n$−3, n, $n$−2 |

⋮

Step $n$−2: Identify the last $n$−1 elements of each starter sets in step $n$−3 By employing to the last $n$−1 elements on each starter set in step $n$−3, the $\frac{(n-1)!}{2}$ distinct starter sets are obtained.

Step $n$−1: Perform CP and RoCP simultaneously on all $n$ elements of the $\frac{(n-1)!}{2}$ distinct starter set, and $n$! distinct permutations are obtained.

Step $n$: Display all $n!$ permutations.

There are $(n-2)$ steps needed to generate a starter set, after which, the CP and RoCP are employed on these starter sets to list down all $n!$ distinct permutations.

To illustrate this algorithm, let's consider the set of five elements, i.e. $S = (1,2,3,4,5)$.

Step 1: Take Set (1, 2, 3, 4, 5) as the initial permutation that appears without loss of generality, and the first element is fixed.

Step 2: Identify the last three elements of the initial permutation from Step 1. By employing CP to the last three elements in the initial permutation from Step 1 three other distinct starter sets are produced, as shown below:

1, 2, 3, 4, 5
1, 2, 4, 5, 3
1, 2, 5, 3, 4

Step 3: Identify the last four elements of each starter set in Step 2. By employing CP to the last four elements on each starter set in Step 2, 12 distinct starter sets are obtained, as shown below:

| | | |
|---|---|---|
| **1, 2, 4, 5, 3** | **1, 2, 5, 3, 4** | **1, 2, 3, 4, 5** |
| 1, 4, 5, 3, 2 | 1, 5, 3, 4, 2 | 1, 3, 4, 5, 2 |
| 1, 5, 3, 2, 4 | 1, 3, 4, 2, 5 | 1, 4, 5, 2, 3 |
| 1, 3, 2, 4, 5 | 1, 4, 2, 5, 3 | 1, 5, 2, 3, 4 |

Step 4: Perform CP and RoCP simultaneously to all n elements of the 12 distinct starter sets and 5! distinct permutations are obtained (see Table 1) .

Step 5: Display 5! distinct permutations ( see Table 1).

TABLE 1
The 5! Distinct Permutations

| CP | RoCP | CP | RoCP |
|---|---|---|---|
| **14532** | 23541 | **14253** | 35241 |
| 45321 | 12354 | 42531 | 13524 |
| 53214 | 41235 | 25314 | 41352 |
| 32145 | 54123 | 53142 | 24135 |
| 21453 | 35412 | 31425 | 52413 |
| **15324** | 42351 | **12534** | 43521 |
| 53241 | 14235 | 25341 | 14352 |
| 32415 | 51423 | 53412 | 21435 |
| 24153 | 35142 | 34125 | 52143 |

TABLE 1 (*continued*)

| CP | RoCP | CP | RoCP |
|---|---|---|---|
| 41532 | 23514 | 41253 | 35214 |
| **13245** | 54231 | **13452** | 25431 |
| 32451 | 15423 | 34521 | 12543 |
| 24513 | 31542 | 45213 | 31254 |
| 45132 | 23154 | 52134 | 43125 |
| 51324 | 42315 | 21345 | 54312 |
| **12453** | 35421 | **14523** | 32541 |
| 24531 | 13542 | 45231 | 13254 |
| 45312 | 21354 | 52314 | 41325 |
| 53124 | 42135 | 23145 | 54132 |
| 31245 | 54213 | 31452 | 25413 |
| **15342** | 24351 | **15234** | 43251 |
| 53421 | 12435 | 52341 | 14325 |
| 34215 | 51243 | 23415 | 51432 |
| 42153 | 35124 | 34152 | 25143 |
| 21534 | 43512 | 41523 | 32514 |
| **13425** | 52431 | **12345** | 54321 |
| 34251 | 51243 | 23451 | 15432 |
| 42513 | 31524 | 34512 | 21543 |
| 25134 | 43152 | 45123 | 32154 |
| 51342 | 24315 | 51234 | 43215 |

**Remark 2:** Permutations in **bold** represent the 12 starter sets for case $n = 5$.

## RESULTS AND DISCUSSION

*Some theoretical results*

The following lemmas and theorem are produced from the recursive circular permutation generation method.

Lemma 1. 2$n$ distinct permutations are produced by each distinct starter set.

Proof: Suppose we have a starter set of $A = (1,2,3 \ldots, n-1, n)$ with $n$ distinct elements. By using definition 5 where all the elements are cycled to the left, n distinct permutations are obtained, as given below:

| | | | | | |
|---|---|---|---|---|---|
| 1 | 2 | … | $n-2$ | $n-1$ | n |
| 2 | … | $n-2$ | $n-1$ | $n$ | 1 |

| | | | | | |
|---|---|---|---|---|---|
| $n-2$ | $n-1$ | $n$ | 1 | 2 | … |
| $n-1$ | $n$ | 1 | 2 | … | $n-2$ |
| $n$ | 1 | 2 | … | $n-2$ | $n-1$ |

Following then from definition 6 and reversing each row of CP produces other $n$ distinct permutations, as given below:

$n \quad n-1 \quad n-2 \quad \ldots \quad 2 \quad 1$
$1 \quad n \quad n-1 \quad n-2 \quad \ldots \quad 2$

$\ldots \quad 2 \quad 1 \quad n \quad n-1 \quad n-2$
$n-2 \quad \ldots \quad 2 \quad 1 \quad n \quad n-1$
$n-1 \quad n-2 \quad \ldots \quad 2 \quad 1 \quad n$

Thus, 2n distinct permutations are produced. □

Lemma 2. There are $\frac{(n-1)!}{2}$ distinct starter sets which are generated recursively for $n \geq 3$ under circular operation.

Proof: Let (1, 2, 3, ..., $n-3$, $n-2$, $n-1$, $n$) be taken as the initial starter for any $n \geq 3$. By employing CP to the last three elements, three distinct starters are produced, as shown below:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | ... | $n-3$ | $n-2$ | $n-1$ | $n$ | (*starter* 1) |
| 1 | 2 | 3 | ... | $n-3$ | $n-1$ | $n$ | $n-2$ | (*starter* 2) |
| 1 | 2 | 3 | ... | $n-3$ | $n$ | $n-2$ | $n-1$ | (*starter* 3) |

Then for each previous starter set, the last four elements are selected, and by employing CP on these elements of the previous starter sets, four distinct starters are produced, as shown below:

From starter 1,

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | ... | $n-1$ | $n-1$ | $n$ | $n-3$ |
| 1 | 2 | 3 | ... | $n-1$ | $n$ | $n-3$ | $n-2$ |
| 1 | 2 | 3 | ... | $n$ | $n-3$ | $n-2$ | $n-1$ |
| 1 | 2 | 3 | ... | $n-3$ | $n-2$ | $n-1$ | $n$ |

From starter 2,

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | ... | $n-1$ | $n$ | $n-2$ | $n-3$ |
| 1 | 2 | 3 | ... | $n$ | $n-2$ | $n-3$ | $n-1$ |
| 1 | 2 | 3 | ... | $n-2$ | $n-3$ | $n-1$ | $n$ |
| 1 | 2 | 3 | ... | $n-3$ | $n-1$ | $n$ | $n-2$ |

From starter 3,

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | ... | $n$ | $n-2$ | $n-1$ | $n-3$ |
| 1 | 2 | 3 | ... | $n-2$ | $n-1$ | $n-3$ | $n-1$ |
| 1 | 2 | 3 | ... | $n-1$ | $n-3$ | $n$ | $n-1$ |
| 1 | 2 | 3 | ... | $n-3$ | $n$ | $n-2$ | $n-1$ |

Thus, at this stage, the total starter sets are 3 × 4 = 12. The process will be continued until the last ($n-1$) element is selected.

| | |
|---|---|
| 3 last elements | 3 starter sets |
| 4 last elements | 4 starter sets |
| 5 last elements | 5 starter sets |
| | |
| ($n-2$) last elements | ($n-2$) starter sets |
| ($n-1$) last elements | ($n-1$) starter sets |

By product rule, the number of starter sets is

$$3\times 4\times 5\times \ldots \times n-2\times n-1 \tag{1}$$

$$=\frac{1}{2}\times 2\times 3\times 4\times 5\times \ldots \times n-2\times n-1 \tag{2}$$

$$=\frac{1}{2}(n-1)! \tag{3}$$

**Remark 3**: For case $n = 2$ is impossible since it has only one distinct starter set while $\frac{(2-1)!}{2}=\frac{1}{2}$.

Theorem 1. The generation of all $n!$ distinct permutations can be obtained by $\frac{(n-1)!}{2}$ distinct starter sets.

Proof: Lemma 2, there are $\frac{(n-1)!}{2}$ distinct starter sets for $n \geq 3$ while from lemma 1, $2n$ distinct permutations are obtained by employing the circular and reversing process on the starter sets.

Thus, $\frac{(n-1)!}{2}\times 2n = n!$ permutations are generated.

*Analysis of time computation*

In this section, this new algorithm is compared with Langdon's (1967) algorithm and Thongchiew's (2007) algorithm because these algorithms fall under the non-exchange restriction. The comparison over time computation between the new algorithm, Langdon's (1967) algorithm and Thongchiew's (2007) algorithm is given in Table 2. The results are given in milliseconds and all the algorithms are written in C language.

As can be observed from Table 2, the new algorithm is faster than Langdon's (1967) algorithm and Thongchiew's (2007) algorithm. At $n = 9$, the new algorithm runs the same as Langdon's (1967), but almost twice as fast as Thongchiew's(2007). In other words, Langdon's algorithm runs two times slower than the new algorithm, and Thongchiew's (2007) is the slowest among the three algorithms.

TABLE 2
Time Computation of Algorithms

| *n* | New algorithm | Langdon's (1967) algorithm | Thongchiew's (2007) algorithm |
|---|---|---|---|
| 8 | 0 | 0 | 0 |
| 9 | 15 | 15 | 63 |
| 10 | 109 | 171 | 687 |
| 11 | 983 | 2012 | 7488 |
| 12 | 12839 | 26520 | 90106 |
| 13 | 173270 | 365213 | 1672510 |
| 14 | 2590946 | 5423443 | 22448205 |
| 15 | 41885652 | 85173825 | 246139962 |

## CONCLUSION

A new approach to listing *n*! permutations that is based on recursive circular generated starter sets is proposed in this paper. Furthermore, this recursive circular algorithm is efficient as the starter sets can be generated without eliminating the equivalence starter sets, thus reducing computation time.

## REFERENCES

Barisenko, A. A., Kalashnikov, V. V., Kulik, I. A., & Goryachev, O. E. (2008). Generation of permutations based upon factorial numbers, *IEEE*, 57-61.

Gao, J., & Wang, D. (2003). *Permutation generation: Two new permutation algorithms*. Retrieved from http://arxiv.org/abs/cs/0306025

Heap, B. (1963). Permutation by interchanges. *Computer Journal*, 293-294.

Ibrahim, H., Omar, Z., & Rohni, A. M. (2010). New algorithm for listing all permutations. *Modern Sciences*, 89-94.

Ives, F. (1976). Permutation enumeration: Four new permutation algorithms. *Comm. ACM* , 68-72.

Iyer, M. (1995). Permutation generation using matrices. Retrieved January 2009, from http://www.ddj.com/184409671.

Johnson, S. M. (1963). Generation of permutations by Adjacent Transpositions. *Math. Comput.*, 282-285.

Langdon, G. (1967). An algorithm for generating permutations. *Communication of ACM*, 298-299.

Ord-Smith, J. (1970). Generation of permutation sequences Part 1. *Computer Journal*, 152-155.

Sedgewick, R. (1977). Permutation generation methods. *Computing surveys*, 137-164.

Shin, D. (2002). The permutation algorithm for non-sparse matrix determinant in symbolic computation. *Proceeding on the 15th CISL winter workshop*, (pp. 42-54). Japan.

Thongchiew, K. (2007). A computerize algorithm for generating permutation and its application in determining a determinant. *Proc. of World Academy of Science, Engineering and Technology, 21*, 178-183.

Trotter, H. F. (1962). Algorithm 115. Perm. *Comm. ACM* , 434-435.

Viktorov, O. (2007). Permutation generation algorithm. *Asian Journal of Information Technology* , 956-957.

Wells, M. (1961). Generation of permutation by transposition. *Math comp*, 192-195.

Zaks, S. (1984). A new algorithm for generation of permutations. *BIT, 24*, 196-204.

Pertanika J. Sci. & Technol. 22 (1): 35 - 42 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# A Comparative Study of Yttrium Doped Ceria Ceramics Synthesized using Mechanochemical and Solid State Methods

**Ong, P. S.[1], Tan, Y. P.[1,2]*, Taufiq-Yap, Y. H.[1,2] and Zainal, Z.[1,2]**

*[1]Department of Chemistry, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
*[2]Putra Laboratory for Catalysis Science and Technology, Department of Chemistry, Faculty of Science,Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia*

## ABSTRACT

In this work, 10 mol% yttrium-doped ceria powders, $Ce_{0.9}Y_{0.1}O_{1.95}$, were synthesised using a new mechanical technique, mechanochemical reaction, in which both impact action and shearing forces were applied for efficient fine grinding, subsequently leading to higher homogeneity of the resultant powders. $Ce_{0.9}Y_{0.1}O_{1.95}$ prepared using this new technique was systematically compared with a sample of the same prepared using conventional solid-state methodology. X-ray diffraction analysis showed all prepared samples were single phase with a cubic fluorite structure. Generally, $Y_2O_3$-doped $CeO_2$ electrolytes prepared by mechanochemical reactions were stable at a lower temperature (1100 °C) compared with a sample of the same synthesised using the conventional solid-state method. Characterisations using differential thermal analysis (DTA) and thermogravimetric analysis (TGA) showed no thermal changes and phase transitions, indicating all materials were thermally stable. The electrical properties of the samples investigated by AC impedance spectroscopy in the temperature range 200–800 ˚C are presented and discussed. Scanning electron microscopy (SEM) was used to study the morphology of the materials. Fine-grained powders with uniform grain-size distribution were obtained from the mechanochemical reaction.

*Keywords:* Oxide ion conductor, yttrium, ceria

*Article history:*
Received: 26 March 2011
Accepted: 15 August 2013

*E-mail addresses:*
poh_shing@hotmail.com (Ong, P. S.),
typ@upm.edu.my (Tan, Y. P.),
taufiq@upm.edu.my (Taufiq-Yap, Y. H.),
zulkar@upm.edu.my (Zainal, Z.)
*Corresponding Author

## INTRODUCTION

Doped ceria is a solid electrolyte which is becoming increasingly attractive for use in solid oxide fuel cells because of its high oxygen ion conductivity at low temperature. Solid oxide fuel cells continue to attract wide interest due to their high energy conversion and low pollution capacity. Currently,

ISSN: 0128-7680 

stabilised zirconia electrolytes are considered the main or established choice of electrolyte in solid oxide fuel cells under restricted conditions (Dudek. 2008). This is in spite of the short life-span of materials of cells at operation temperature range of 900–1000 ˚C. In attempting to reduce the operation temperature, several researchers have turned their attention to the application of doped-ceria electrolytes as an alternative to the conventional stabilised zirconia electrolytes. The top variants of doped-ceria systems are comprised of products resulting from a substitution of lower-valent metals such as Y, Sm, Gd and Ca (Huang *et al*., 1997). $Y_2O_3$-doped ceria is considered a potential solid electrolyte for use in intermediate temperature solid oxide fuel cells (IT-SOFC). Wang *et al*. (2006) have studied the relationship between defect structures in $Y_2O_3$-doped ceria by diffusion and electrical conduction in order to understand the conduction mechanisms of this electrolyte.

Powders of nano-scale have excellent properties suited for various applications as highlighted by Singh and Hegde (2008). This study was carried out because no systematic research has yet been conducted on the mechanochemical synthetic route, better known as the ball-milling method, due to drawbacks encountered using mechanical synthesis compared to the wet chemical approach in terms of sinterability and densification (Li *et al.,* 2002). However, this study was conducted with the main purpose of contributing to the construction of a solid-state method basis by addressing this matter, which might then enable us to draw more reliable conclusions about the potential as well as the limitations of this synthetic route.

## MATERIALS AND METHODS

Yttrium oxide-substituted ceria samples were synthesised using two methods i.e. the conventional solid state method and the mechanochemical method with starting materials $CeO_2$ (99.9% Acros Organics) and $Y_2O_3$ (98% Fluka-Garantie). All starting materials were dried at 500-600 ˚C prior to weighing. Stoichiometric amounts of these materials were mixed. The mechanochemical method used a total mass of approximately 15.0 g of stoichiometric mixtures of $CeO_2$ with $Y_2O_3$, placed in an agate bowl (99.9 % $SiO_2$) with agate balls of 10 mm diameter. Ethanol was added as a milling medium to prevent excessive abrasion. The mixture was milled using a planetary ball mill (Model Pulverisette 4 vario-Planetary mill) for one hour. The slurry was dried at 60–70 ˚C to evaporate the ethanol. In the conventional solid state method, mixtures of required molar ratios of materials were weighed and wet-mixed manually with acetone as the wetting medium using an agate mortar and pestle. The materials were subjected to heat treatments at different temperatures and durations to ensure the formation of single phase materials.

Phase purity was characterised by X-ray diffraction analysis (XRD) (Shimadzu diffractormeter XRD 6000, CuK$\alpha$ radiation) in $2\theta$ range of 10–60 ˚ at 2 ˚/minute. The CHEKCELL refinement programme was used to obtain lattice parameters of the structure. Thermal events were recorded from room temperature to 1000 ˚C on heat and cool cycles (10 ˚C/minute) using differential thermal analysis (DTA, Perkin-Elmer DTA 7) and thermogravimetric analysis (TGA, Perkin Elmer TGA 7).

The electrical conductivity of sintered ceramic pellets was measured. The pellets were pressed uniaxially and sintered at a temperature range of 1200–1500 °C. The densities of the

pellets (85 %) were determined through the Archimedes method using deionised water as the medium. The electrical properties were determined by AC impedance spectroscopy using a Hewlett Packard Impedance Analyzer, HP4192A, over the frequency range 5 Hz to 13 MHz. Measurements were made from 100 ˚C to 800 ˚C by incremental steps of 50 ˚C on a heating cycle with 30 min equilibration time. Scanning electron microscopy (SEM) analysis was carried out using SEM JEOL JSM-6400 operated at 15 kV, with a working distance of 13 mm.

## RESULTS AND DISCUSSION

### *XRD Analysis*

X-ray diffraction analysis (XRD) was used to identify the crystalline phases as well as the lattice parameters of solid solution powders. The diffracted X-rays were collected over 2θ range 10 ˚–60 ˚ using a step width of 0.02 ˚. The diffractometer was calibrated using Si standard. The XRD spectra of the 10 mol% yttrium-doped $CeO_2$ powders are illustrated in Fig.1 and Fig.2. X-ray diffraction analysis showed that all the yttrium-doped ceria samples were monophase materials. (JCPDS powder diffraction File No. 34-0394). Bragg peaks for the unreacted $Y_2O_3$ phase were found in the XRD spectra of compounds prepared using the conventional solid-state method after calcination at 800-–1200 ˚C. The Y-doped $CeO_2$ prepared using the mechanochemical method showed the presence of strong XRD Bragg peaks of a fluorite-type phase, indicating the presence of $CeO_2$ (JCPDS card No. 34-0394) at a lower temperature compared to that of the material synthesised using the conventional solid-state method. X-ray diffraction analysis also showed all doped ceria powders obtained in $Ce_{1-x}Y_xO_{2-\delta}$ ($0.05 \leq x \leq 0.3$) ceramics were monophasic material (Fig.3). These results demonstrated that successful synthesis was achieved at a temperature ~200 ˚C lower than previously reported by Zhang *et al*. (2003) using the same mechanical synthesis approach.

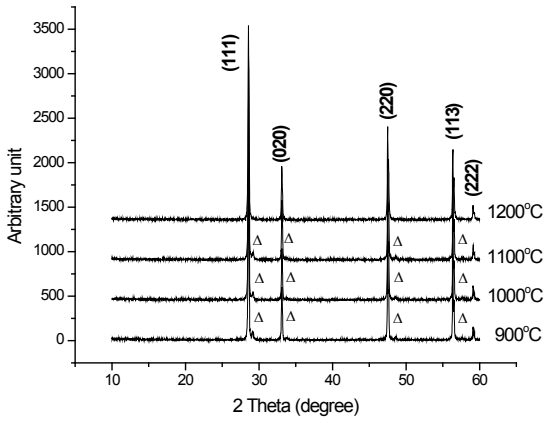



Fig.1: XRD diffraction patterns of $Y_2O_3$ doped $CeO_2$ system prepared by conventional solid-state method. Δ denotes the unreacted $Y_2O_3$ (JCPDS powder diffraction File No. 19-1448).

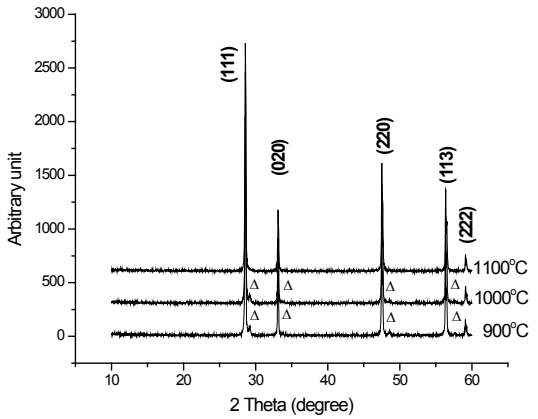



Fig.2: XRD diffraction patterns of $Y_2O_3$ doped $CeO_2$ system prepared by mechanochemical method. Δ denotes the un-reacted $Y_2O_3$ (JCPDS powder diffraction File No. 19-1448).

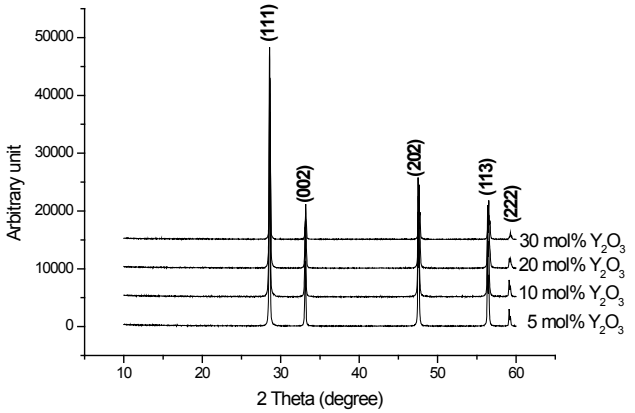



Fig.3: Powder XRD (CuK$\alpha$) patterns of $Ce_{1-x}Y_xO_{2-\delta}$ ($0.05 \leq x \leq 0.3$) prepared by mechanical syntheses.

Calculation of the cell parameters was done using the main reflection of a material having a typical fluorite structure with an fcc cell, corresponding to the (1 1 1) plane at ~ 28.5˚. Comparison of lattice parameters as a function of yttrium content is shown in Fig.4. The lattice constants of $Ce_{1-x}Y_xO_{2-\delta}$ samples synthesised by both methods decreased linearly with increasing yttrium content in the investigated substitution range $x$ = 0.05 - 0.3. The results in Fig.4 further suggest that all the doped ceria samples in this work are ceria-based solid solutions. Incorporation of $Y_2O_3$ into the $CeO_2$ system caused an almost linear decrease of cell parameter and is in good agreement with effective ionic radii for which the radius of $Y^{3+}$ (0.1019 nm) is smaller than the radius of $Ce^{4+}$ (0.1111 nm) (Shannon, 1976). It was also found that the $2\theta$ values of the doped ceria shift slightly towards higher angles when $x$ varies from 0.05 to 0.3 (Fig.3). Dissolution of a smaller radius cation into the ceria system is testimony that yttrium is soluble into the ceria crystal lattice and is in good agreement with Vegard's rule of lattice parameter linear dependence against the composition.

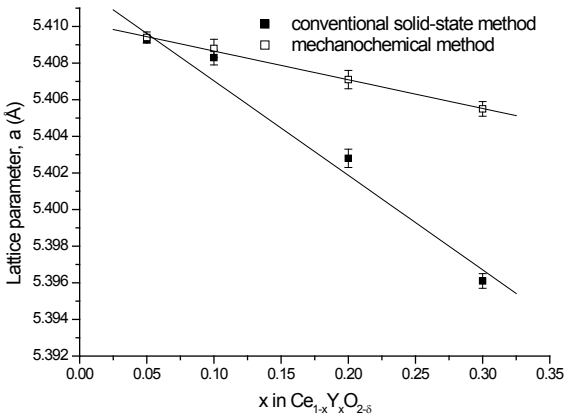



Fig.4: Variation of cell parameters with $x$ in$Ce_{1-x}Y_xO_{2-\delta}$ solid solutions prepared by both mechanochemical and solid state methods

*Thermal Analysis*

Thermal analysis showed no significant thermal events, verifying that all the compounds are thermally stable at a temperature range from room temperature to 1000 ˚C.

*Impedance Analysis*

The conductivity data measured in air below 1073 K for Y-doped ceria samples prepared by solid-state and mechanochemical methods were analysed using Arrhenius' equation:

$$\sigma = \frac{\sigma \circ}{T} \exp\left(-\frac{E}{kT}\right) \quad (1)$$

where $E$ is the activation energy of electrical conduction, $k$ Boltzmann's constant, $T$ the absolute temperature and $\sigma_o$ the pre-exponential factor.

This study showed the conductivity measurement upon $Ce_{0.9}Y_{0.1}O_{1.95}$ samples synthesised using both methods, which were sintered at 1200 – 1500 ˚C (Fig.5 and Fig.6). The optimum temperature for the highest conductivity for $Ce_{0.9}Y_{0.1}O_{1.95}$ prepared by the solid-state method was determined to be 1400 ˚C, whereas for samples prepared using the mechanochemical method, comparable conductivity occurred at 1300 ˚C. Both samples showed similar values at 600 ˚C (~ $10^{-3}$ S $cm^{-1}$) with activation energy, $E_a$ ~ 0.89 eV, in agreement with the value reported previously (Van Herle *et al*., 1996).

*Morphology Analysis*

The powder samples were pelletised and sintered at 1200 ˚C overnight with a programmed heating and cooling rate of 5 ˚C/minute. Fig.7 and Fig.8 show typical micrographs of samples prepared by the solid-state reaction and the mechanochemical method. $Ce_{0.9}Y_{0.1}O_{1.95}$ prepared by solid-state reaction revealed poor densification. It was difficult to distinguish the

grain boundaries and to estimate the corresponding grain size. The sample prepared by the mechanochemical method, however, appeared dense with low intergranular porosity. The solid-state reaction method yields higher particle size and larger intergranular porosity due to high-firing temperature. The average grain size measured from the sample prepared by the mechanochemical method is 1.04 μm compared to 4.25 μm, prepared by solid-state reaction.

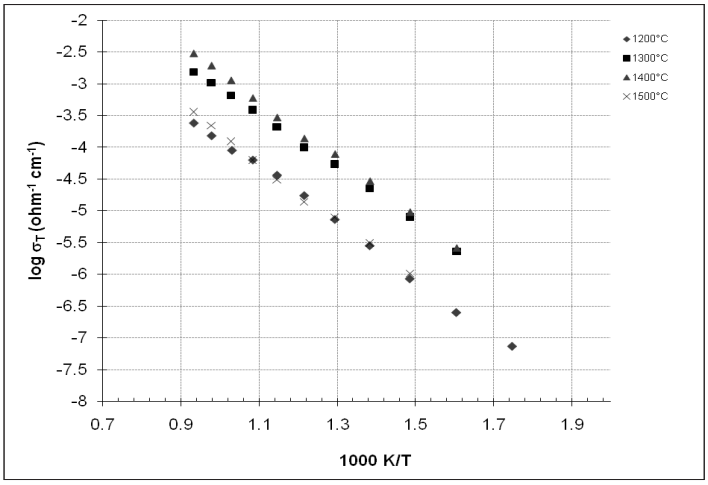



Fig.5: Arrhenius conductivity plot of the ionic conductivity of $Ce_{0.9}Y_{0.1}O_{1.95}$ prepared by conventional solid-state method.

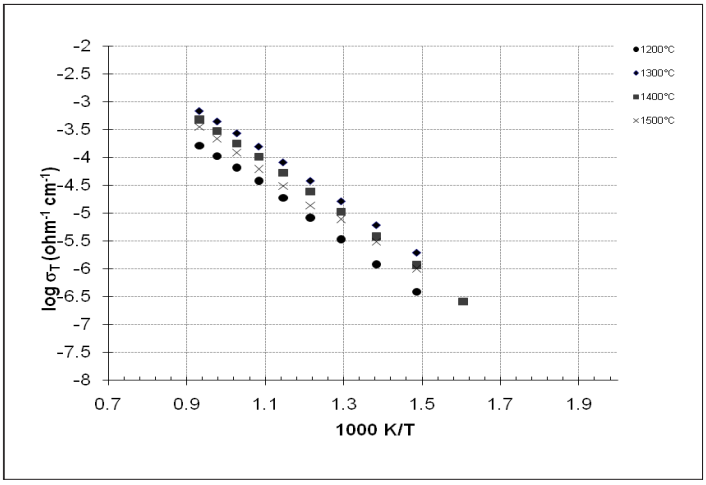



Fig.6: Arrhenius conductivity plot of the ionic conductivity of $Ce_{0.9}Y_{0.1}O_{1.95}$ prepared by mechanochemical method.

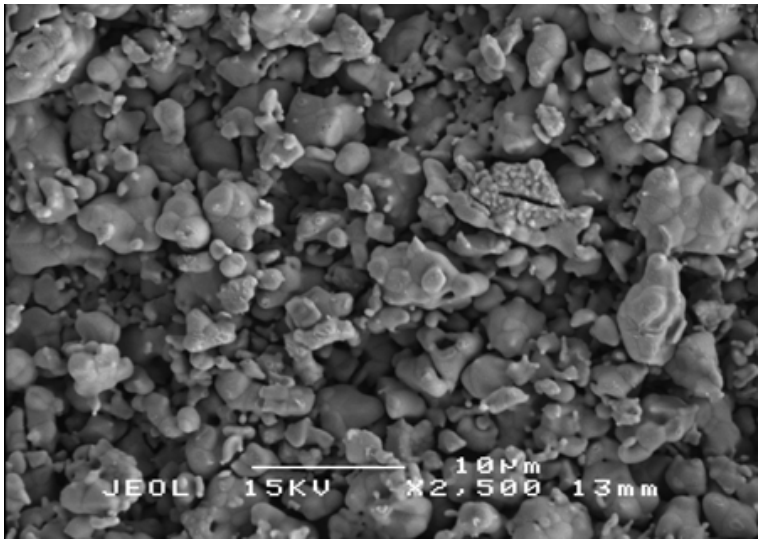



Fig.7: Scanning electron micrograph of $Ce_{0.9}Y_{0.1}O_{1.95}$ prepared by conventional solid-state method

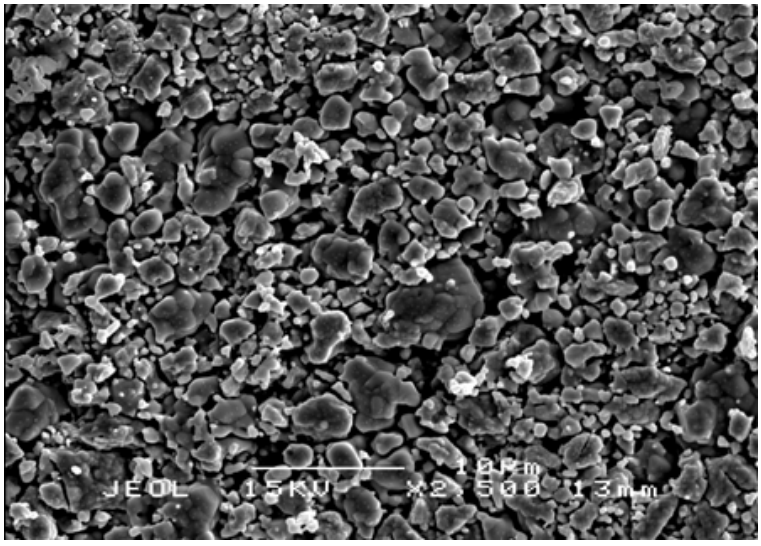



Fig.8: Scanning electron micrograph of $Ce_{0.9}Y_{0.1}O_{1.95}$ prepared by mechanochemical method

## CONCLUSION

The yttrium-doped ceria solid solutions, $Ce_{1-x}Y_xO_{2-\delta}$ ($0.05 \leq x \leq 0.3$) with fluorite structure, were successfully prepared using different mechanical syntheses. The results of X-ray diffraction showed that all samples were single phase with a cubic fluorite structure. The employment of the mechanochemical synthetic route had proven promising in terms of phase stability. The best conductivity of the $Ce_{0.9}Y_{0.1}O_{1.95}$ studied prepared by the conventional solid-state method ($\sigma_{600^\circ C} \sim 2.97 \times 10^{-3}$ S cm$^{-1}$) was close to those previously recorded by Zhang *et al*. (2003) and

Yamashita *et al*. (1995) ($\sigma_{600^{\circ}C} \sim 10^{-2}$ S cm$^{-1}$) and one order of magnitude higher than that of stabilised zirconia, the most commonly used solid electrolyte at the corresponding temperatures ($\sigma_{600^{\circ}C} \sim 10^{-4}$ S cm$^{-1}$) (Marques *et al*., 2006).

## ACKNOWLEDGEMENTS

The authors wish to thank Yayasan Khazanah for funding this research work through the Yayasan Khazanah Watan Scholarship and the Ministry of Science and Technology for financial support through Research Grants.

## REFERENCES

Dudek, M. (2008). Ceramic oxide electrolytes based on $CeO_2$ - Preparation, properties and possibility of application to electrochemical devices. *Journal of the European Ceramic Society*, *28*, 965-971.

Huang, W., Shuk, P., & Greenblatt, M. (1997). Properties of sol-gel prepared $Ce_{1-x}Sm_xO_{2-x/2}$ solid electrolytes. *Solid State Ionics*, *100*, 23-27.

Li, J. G., Ikegami, T., Wang, Y., & Mori, T. (2002). Nanocrystalline $Ce_{1-x}Y_xO_{2-x/2}$ ($0 \leq x \leq 0.35$) Oxides via Carbonate Precipitation: Synthesis and Characterization. *Journal of Solid State Chemistry*, *168*, 52–59.

Marques, F. M. B., Kharton, V. V., Naumovich, E. N., Shaula, A. L., Kovalevsky, A. V., & Yaremchenko, A. A. (2006). Oxygen ion conductors for fuel cells and membranes: Selected developments. *Solid State Ionics, 177*, 1697-1703.

Shannon, R. D. (1976). Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallographica*, *32A*, 751.

Singh, P., & Hegde, M. S. (2008). Controlled synthesis of nanocrystalline $CeO_2$ and $Ce_{1-x}M_xO_{2-\delta}$ (M = Zr, Y, Ti, Pr and Fe) solid solutions by the hydrothermal method: Structure and oxygen storage capacity. *Journal of Solid State Chemistry, 181*, 3248–3256.

Van Herle, J., Horita, T., Kawada, T., Sakai, N., Yokokawa, H., & Dokiya, M. (1996). Sintering behaviour and ionic conductivity of yttria-doped ceria. *Journal of the European Ceramic Society*, *16*, 961-973.

Wang, Y., Kageyama, H., Mori, T., Yoshikawa, H., & Drennan, J. (2006). Local structures around Y and Ce cations in 10 mol% $Y_2O_3$ doped ceria ceramics by EXAFS spectroscopy. *Solid State Ionics, 177*, 1681-1685

Yamashita, K., Ramanujachary, K. V., & Greenblatt, M. (1995). Hydrothermal synthesis and low temperature conduction properties of substituted ceria ceramics. *Solid State Ionics*, *81*, 53-60.

Zhang, T. S., Ma, J., Huang, H. T., Hing, P., Xia, Z. T., Chan, S. H., & Kilner, J. A. (2003). Effects of dopant concentration and aging on the electrical properties of Y-doped ceria electrolytes. *Solid State Sciences*, *5,* 1505–1511.

Pertanika J. Sci. & Technol. 22 (1): 43 - 53 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Design and Development of a Sweet Potato Digging Device

**Md. Akhir, H.[1]*, Ahmad, D.[2], Rukunudin, I. H.[1], Shamsuddin, S.[3] and A. Yahya[2]**

*[1]Mechanization and Automation Research Centre MARDI, P.O. Box, 12301, 50774 Kuala Lumpur, Malaysia*
*[2]Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*
*[3]Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

This paper describes a study on the design, fabrication and testing of a prototype digging device for sweet potato tubers in bris soil. The soil texture was sandy soil (fine sand 94.53%), with mean moisture content of 9.16% and mean bulk density of 1.44 g-cm$^{-3}$. The soil was prepared in a soil bin. Three types of soil digging tools were designed and fabricated to determine the optimum draft force. These were Flat or plane, V-shaped and Hoe type blades. Plane and V-shaped blades were 30 cm long, and 13 cm wide, while the Hoe type had three rods, 25 mm in diameter, 30 cm long and 6.5 cm wide with sharp cutting edge. The digging tools were tested in a soil bin filled with bris soil to determine the optimum draft force and area of soil disturbance. The results were analysed using statistical analysis of variance (ANOVA). Comparison between all blade types and blade depths to measured draft force and the area of soil disturbed showed that the highest draft of 0.54 kN-m$^{-2}$ was caused by a flat or plane blade at the optimum depth of 20 cm when the area of soil disturbed was 0.180 m$^2$. The V-shaped blade had the mean draft of 0.51 kN-m$^{-2}$, with area of soil disturbance of 0.185 m$^2$. The best choice was V-shaped blade with a rake angle of 30$^o$ at 20 cm. depth. The selected blade was fixed onto the sweet potato harvester and tested on bris soil planted with sweet potato of Telong and VitAto varieties. The harvesting efficiency of the machine in bris soil was 93.64% and 90.49% for Telong (Plot A) and VitAto (Plot B) varieties, respectively. The average ground speed and turning time during operation for plots A and B was 0.56 km-hr$^{-1}$ and 102.7 s and 0.99 km-hr$^{-1}$ and 81.22 s, respectively. The harvesting efficiencies for both plots showed no significant difference. The total productive time (harvesting time) and unproductive time (turning time) in plot A, at a tractor speed of 0.56 km.hr$^{-1}$, was 14.8 hours for harvesting a hectare of sweet potato ( 0.068 ha.hr$^{-1}$). In plot B, the total time for harvesting a hectare of sweet potato was 8.35 hours (0.12 ha.hr$^{-1}$) at a tractor speed of 0.99 km.hr$^{-1}$. The average harvesting time for both plots was 11.47 hr.ha$^{-1}$. The average field work rate was 0.087 ha.hr$^{-1}$ or 34 man-hr.ha$^{-1}$ compared to manual harvesting of 150 man-hrs.ha$^{-1}$.

*Article history:*
Received: 4 April 2011
Accepted: 4 May 2012

*E-mail addresses:*
mdakhir@mardi.gov.my (Md. Akhir, H.),
desa@upm.edu.my (Ahmad, D.),
ibnihajar@unimap.edu.my (Rukunudin, I. H.),
shamsuddin@upm.edu.my (Shamsuddin, S.),
azmiy@upm.edu.my (A. Yahya)
*Corresponding Author

ISSN: 0128-7680 

*Keywords:* Digger blade, Sweet potato, Bris soil, Harvester

## INTRODUCTION

Sweet potato is a minor crop mainly grown for local consumption. Currently, there are about 2,000 ha of the crop grown in Peninsular Malaysia (Tan *et al*., 2006). Even though it is a small industry, sweet potato is being promoted to be grown on an extensive scale to replace the tobacco crop on the coastal sand ridges (bris) of Kelantan and Terengganu on the east coast of Peninsular Malaysia. Its promotion is part of the strategy to counter the increasing competitive pressure on the tobacco crop arising from global trade liberation. At the same time, there are large areas of sand tailings on the ex-mining area which is still not re-utilised and thus, suitable for cultivating sweet potato. Sweet potato has been proven to be cultivable on this type of soil (Tan, 1998; Tan *et al*., 2000).

Cultivation in a large area will involve labour intensive work especially during the harvesting operation. Currently, sweet potato is mostly harvested manually. The manual labour cost for a harvesting operation constitutes about 30 - 40% of the total operational cost (Md. Akhir & Desa., 2005). In manual harvesting of sweet potato tubers, the farmers have to cut and pull out the vines and lay them along the furrow. The tubers will then be dug by using a hand tool such as a hoe and fork, followed by manual collection. The tubers are transported in a basket or gunny sack. The most strenuous and back-breaking task is digging the tubers. The harvesting operation of sweet potato requires about 150 man-hours per hectare (Md. Akhir & Tan, 2002; Md. Akhir *et al*., 2005).

Lately, attempts at harvesting sweet potato mechanically have been made by a few farmers. Tractor mounted tillage tools such as the chisel plough digger, mould-board plough digger and double disk harrow have been used. These diggers are simple and cheap but tend to damage the tubers and usually cause field losses. Such operations also do not leave the tubers well exposed on the ground. In contrast, the digger elevator lifts the dug-out tubers exposed on the ground surface to facilitate easier gathering of tubers. In addition, the sieving action of its elevator conveyor minimises the extent of root losses.

The potato digger-elevator for a single-row crop has also been tested for sweet potato harvesting in mineral and bris soils (Md. Akhir *et al*., 2008). However, its performance was found to be not very effective under local conditions. In view of the current need for a mechanized sweet potato harvesting system to facilitate large scale production, a sweet potato harvester needs to be developed. The major part involved in the design of tuber crop harvester is the digger blade device.

Dash *et al*. (1998) reported on the performance of four different types of bullock-drawn groundnut diggers, namely, two-row ridging type, ridging type with semi-circular blade v-type and ridger type. The results showed that the average draft for two-row ridging type, semi-circular blade, v-type blade and ridger type were 85.5, 72.3, 66.8 and 57.0 kgf, respectively. Meanwhile, the maximum digging efficiencies were 74.3, 65.5, 81.9 and 92.0%, respectively.

This paper highlights the study on the three types of digger blades in bris soil under controlled conditions. The blades were Flat blade, V-Shaped and Hoe type. The most suitable

blade among the three was selected to be fixed onto the designed and fabricated sweet potato harvester. The completed prototype was then tested in the sweet potato field to study its performance.

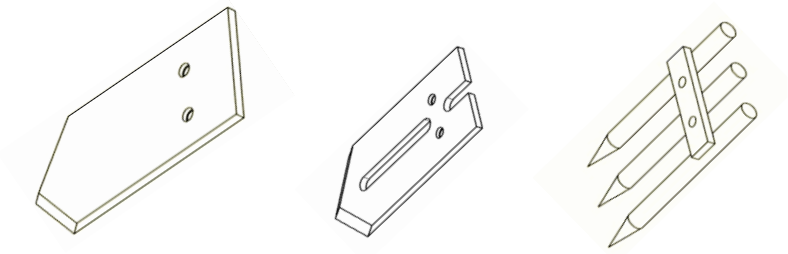
## MATERIALS AND METHODS

### *Bris Soil Properties*

Five random samples of bris soil were taken from the sweet potato fields for analysis of the properties. Bris soil texture was determined by float pipette method, while soil moisture and bulk density were determined by using the gravity method.

### *Blade Design and Fabrication*

Three types of digger blades were designed and developed to study their draft force and soil disturbance in a soil bin. The blades were Flat blade ($B_1$) 13 cm wide, 30 cm long; V- shaped ($B_2$) with 13 cm width, 30 cm long; and Hoe type ($B_3$) with three iron rods, 25 mm in diameter, 30cm long and sharp at the end (Fig.1a). All the blades were tested in a soil bin filled with bris soil with average moisture content and bulk density similar to field conditions before harvesting. The draft force was determined using different blade types ($B_1$, $B_2$ and $B_3$) at different rake angles with the horizontal ($A_1 = 30^o$, $A_2 = 45^o$ and $A_3 = 60^o$) and depths of cutting ($D_1$=10 cm, $D_2$=15 cm and $D_3$= 20 cm). The best blade with the optimum angle and depth was selected to be fixed onto the digger frame (Fig.1b.) and attached to a prototype sweet potato harvester (Fig.2). The harvester with the selected blade was tested to harvest the sweet potato in bris soil. The data on the performance of the device and prototype sweet potato harvester were analysed using one way analysis of variance (ANOVA).



Flat blade ($B_1$) V-Shaped blade ($B_2$) Hoe type ($B_3$)

Fig.1a: A schematic diagram of digger blade devices

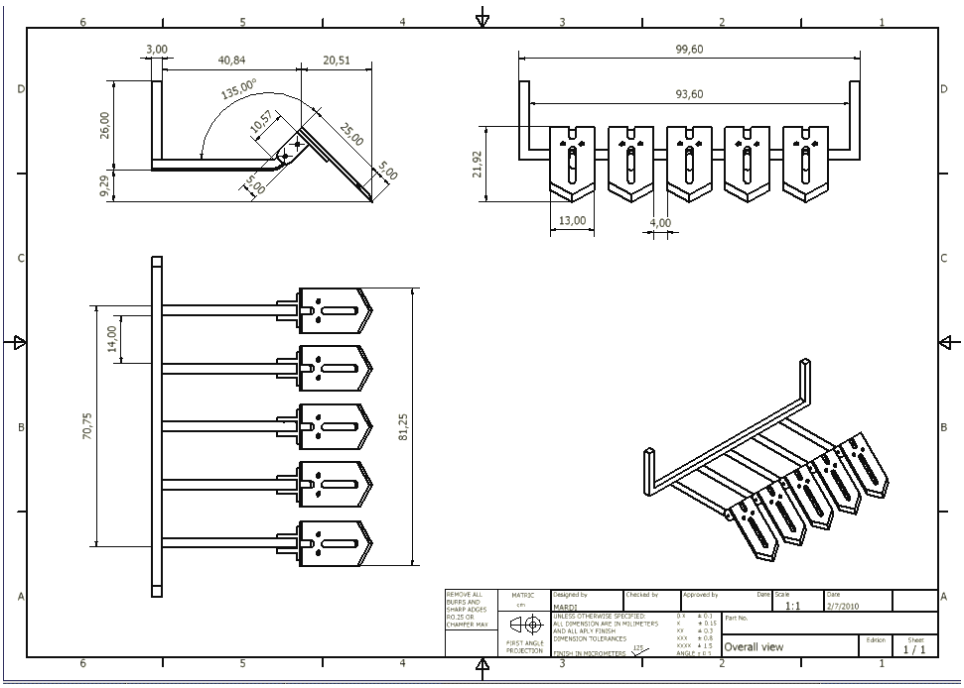



Fig.1b: A detailed design of selected digger blade (V-shape) attached to the digger frame

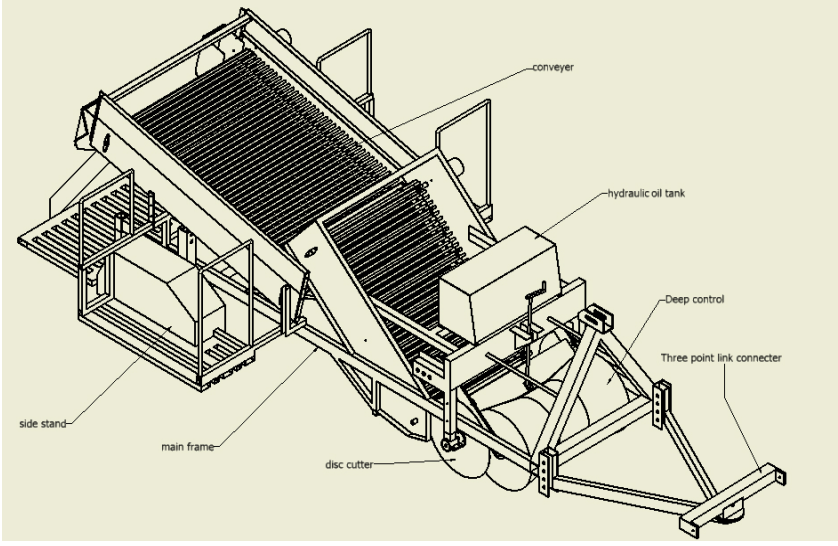



Fig.2: A schematic diagram of prototype sweet potato harvester

## RESULTS AND DISCUSSION

### *Soil Condition*

The physical properties of bris soil in this study are shown in Table 1.

TABLE 1
The Physical Properties of Bris Soil

| Soil depth (cm) | Mechanical analysis (%) | | | Other material (%) | Moisture content (%) | Bulk density (t/m$^3$) |
|---|---|---|---|---|---|---|
| | Clay | Silt | Sand | compost | | |
| 0- 15 | 0.11 | 0.13 | 94.63 | 5.13 | 9.3 | 1.49 |
| 15-30 | | | | | 9.51 | 1.56 |
| 30-45 | | | | | 10.2 | 1.57 |
| Average | | | | | 9.67 | 1.54 |

Table 1 shows that silt, clay, sand, and other material contents were 0.13%, 0.11%, 94.63% and 5.13%, respectively. The average moisture content and bulk density of the soil during harvesting operation were 9.67% and 1.54 t/m$^3$, respectively. In order to simulate the actual field condition, bris soil in soil bin was compacted and watered to achieve a bulk density and moisture content similar to that measured in the field.

### *Blade Performance in Soil Bin*

Results from statistical analysis of draft force and soil disturbance areas for different blade types, rake angles and depths in soil bin are shown in Fig.3 to Fig.8. In these figures, the means between treatments followed by similar letters do not differ significantly at $p>0.05$.

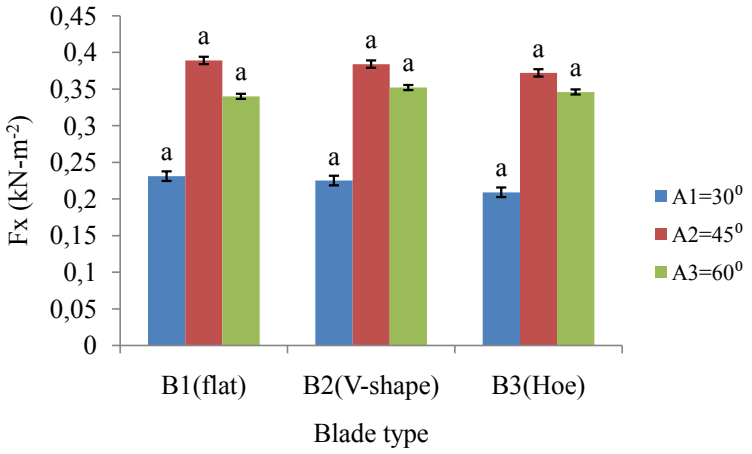



Fig.3: Draft force (*Fx*) in combination with blade type (B) and rake angle (A)

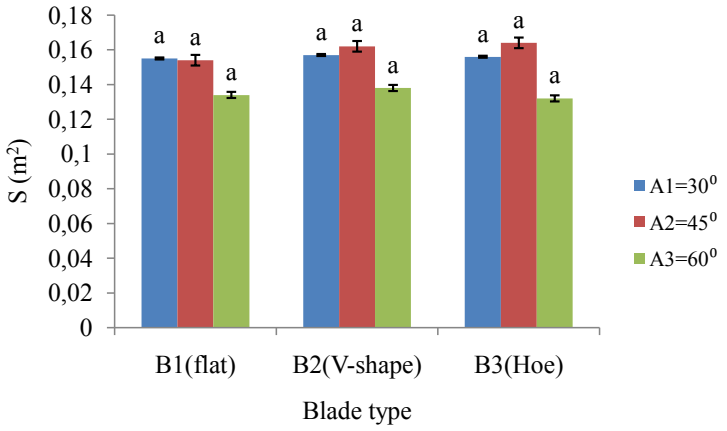



Fig.4: Soil disturbance area (*S*) in combination with blade type (B) and rake angle

Fig.3 shows that there are no significant effects of draft force (Fx) combination with the blade type and rake angle. The hoe type blade rake angle of 30° gives the lowest draft force of 0.209 kN-m$^{-2}$. The highest draft force of 0.389 kN-m$^{-2}$ was obtained by Flat blade (B1) at a rake angle of 45° (A2).

The effects on soil disturbance area(S) during the experiment in combination with blade type (B) and rake angle (A) are shown in Fig.4. It shows that there was no significant effect on soil disturbance area in combination with blade types and rake angle. The highest soil disturbance area of 0.164 m$^2$ was for Hoe type blade at a rake angle of 45° and no significant effect was found between blade types. There were no significant effects at level $p>0.05$ on soil disturbance area at the rake angle of 60°.

The effect on draft force (Fx) in combination with blade type (B) and depth (D) is shown in Fig.5. The figure shows that there was no significant effect on the draft force by the blade types, but there were significant effects on the draft force at $p> 0.01$ in combination with blade type and depth. The figure also shows that an increase of blade depth will increase the draft force, except for hoe type (B3) at the depth of 15cm(D2) and 20 cm (D3). This happened because the soil moved between the rods. The lowest draft force caused by the Flat blade at 10cm depth was 0.148 kN-m$^{-2}$. The highest draft force of 0.535 kNm$^{-2}$ was caused by the Flat blade, followed by V-Shaped blade with 0.51 kN- m$^{-2}$ and Hoe type of 0.492 kNm$^{-2}$ at a depth of 20 cm, respectively.

The effects on soil disturbance area (S), in combination with blade type (B) and depth (D), are shown in Fig.6. There was no significant effect on soil disturbance area between blade types but there were highly significant effect at $p> 0.01$ for both the depth and combination blade type and depth. The high effect of soil disturbance area of 0.183 m$^2$ was at a depth of 20 cm. The highest effect in combination with blade type and depth was caused by the V-shaped blade. The lowest effect was hoe type blade at a depth of 10 cm with soil disturbance area of 0.112 m$^2$.

Fig.7 shows that there was no significant effect of draft force between the rake angles, but highly significant effect at $p>0.01$ for both the depth and combination with the rake angle and the depth. The lowest draft force (Fx) of 0.156 kN-m$^{-2}$ was at a depth of 10 cm (D1). At

the rake angle of 30° (A1), the draft force of 0.105 kN-m$^{-2}$ was at a depth of 10 cm. The high draft force of 0.742 kN-m$^{-2}$ was for 45° rake angle at a depth of 20cm (D3).

Fig.8 shows that there were significant effects at p>0.01 of soil disturbance area for the depth and combination with the blade rake angle and depth. The highest soil disturbance area of 0.193 m$^2$ was at a depth of 20 cm (D3) and at a rake angle of 45°. For the high rake angle of 60°, and at a depth of 20 cm (D3), the soil disturbance area was 0.173 m$^2$.

Based on Fig.3 to Fig.8, the best combination was blade type (B) and rake angle (A) at the depth of 20 cm (D3) for getting the high soil disturbance area via the optimum draft force for the selected blade that was fitted to the sweet potato harvester. The V-shaped blade had the optimum draft of 0.51 kN-m$^{-2}$ at a rake angle of 30° with soil disturbance area of 0.185 m$^2$ at the depth of 20 cm. The Hoe-type tool with the lowest draft of 0.34 kN-m$^{-2}$ and soil disturbance area of 0.184 m$^2$ was not selected because soil can pass through the rods easily and can cause excessive damage to tubers.

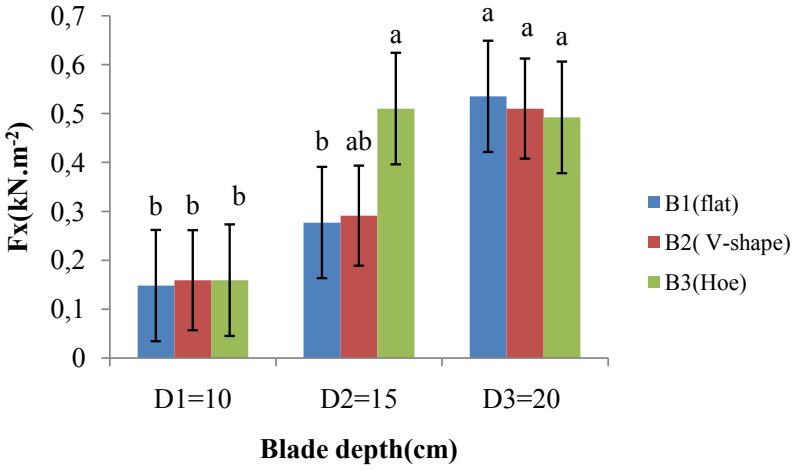



Fig.5: Draft force (Fx) in combination with blade type (B) and depth (D)

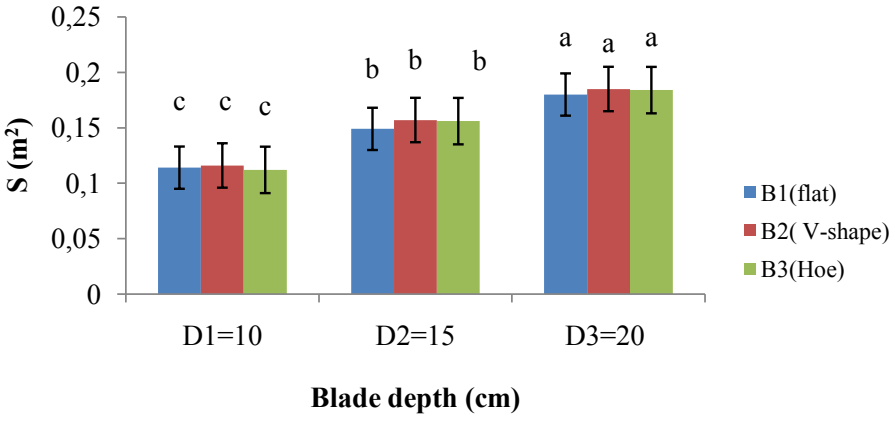



Fig.6: Soil disturbance area (S) in combination with blade type (B) and depth (D)

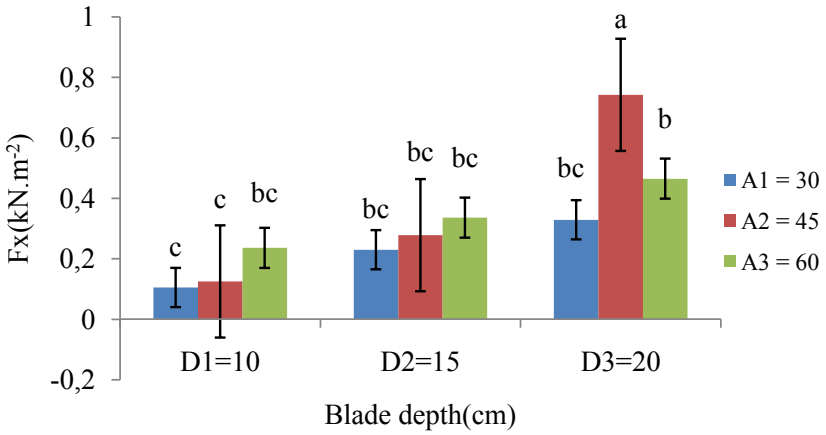



Fig.7: Draft force (Fx) in combination with rake angle (A) and depth (D)

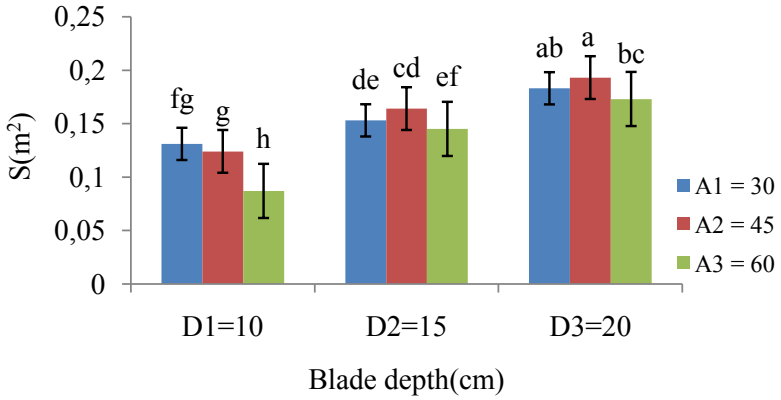



Fig.8: Area of soil disturbance (S) in combination with blade rake angle (A) and depth (D)

*Digger Blade Design and Development*

The selected V-shaped digger blade was used in this design based on the performance test in a soil bin. The data such as blade rake angle, blade depth and blade width were used to design and develop the sweet potato digging device. Fig.1b shows the detailed design of the selected digger blade (V-shaped) attached to the digger frame. The digger blade and frame (digging device) were than attached to the sweet potato harvester for performance testing (Fig.2).

*Performance of Sweet Potato Harvester*

The study was conducted on two plots, A and B, where each plot had an area of 0.25 ha. Plot A had 10 beds of 50 m length, 1.5 m from furrow to furrow and 1 m wide at the top. Plot B also had 10 beds, but these were 70 m in length. Both plots were of bris soil located at MARDI Research Station at Telong in Kelantan. Land preparation was carried out in accordance with the standard procedures as recommended by Md. Akhir *et al*. (2002). For a double-row planting,

the bed size was 1.5 m wide and 30 cm high. Sweet potato vines were planted along the bed in two rows using a mechanical transplanter. The plant spacing was 30 cm between plants and 45 cm between rows. Plot A was planted with Telong variety while plot B was planted with VitAto variety. The crop agronomies application and maintainence were done according to the standard procedure as proposed by Tan *et al*. (1998), and harvested 120 days after planting.

The field evaluation of the harvester was carried out at the experimental plots when the sweet potato plants were about 120 days old and ready for harvest. Prior to harvesting, the vines and leafs were mechanically removed from the surface to ensure smooth operation of the harvester. During harvesting operation, the time taken to cover the distance as well as the turning time for each operation was measured using a stopwatch. In addition, the functional performance of the machine such as durability, consistency during operation, conveyor operation and machine manoeuvrability were observed and noted. The machine performance data which were collected included tractor speed, digging losses, tuber damage and overall machine efficiency.

The harvesting results were analysed through simple statistical analysis RBD (Randomized block design) to determine the harvesting performance for each plot, A and B. The harvesting performance in bris soil with Telong and Vitato variety of sweet potatoes are shown in Tables 2 and 3, respectively.

TABLE 2
Harvesting performance in bris soil with Telong variety

| No. of rows | Distance (m) | Harvesting time (s) | Turning time (s) | Tractor speed (km/h) | Yield collected/row | Damaged + missing | Harvesting efficiency |
|---|---|---|---|---|---|---|---|
| | | | | | (kg) | Kg/row | % |
| Total | 500 | 3217 | 924 | 5.63 | 922 | 62.7 | 93.64 |
| Average | 50 | 321.7 | 102.7 | 0.56 | 92.2 | 6.27 | 93.64 |
| Max | 50 | 338 | 112 | 0.68 | 111 | 8.6 | 95.72 |
| Min | 50 | 265 | 95 | 0.53 | 84.0 | 3.8 | 91.00 |
| Stdev | 0.00 | 23.28 | 6.26 | 0.046 | 8.28 | 1.31 | 1.31 |
| Cv | | 0.07 | 0.06 | 0.082 | 0.09 | 0.22 | 0.014 |

TABLE 3
Harvesting performance in bris soil with VitAto variety

| No. of rows | Distance (m) | Harvesting time (s) | Turning time(s) | Tractor speed (km/h) | Yield collected/row | Damaged + missing | Harvesting efficiency |
|---|---|---|---|---|---|---|---|
| | | | | | (kg) | Kg/row | % |
| Total | 70 | 2540 | 731 | 9.99 | 999.66 | 105.60 | 90.49 |
| Average | 70 | 254.0 | 81.22 | 0.99 | 99.97 | 10.56 | 90.49 |
| Max | 70 | 297 | 103 | 1.14 | 121.38 | 16.40 | 93.22 |
| Min | 70 | 221 | 62.0 | 0.85 | 80.92 | 6.80 | 86.83 |
| Stdev | 0.00 | 22.59 | 28.66 | 0.09 | 14.38 | 3.15 | 2.13 |
| Cv | | 0.09 | 0.35 | 0.09 | 0.14 | 0.30 | 0.02 |

Table 2 shows the machine performance on bris soil for the Telong variety. There were ten beds and each bed was 50 m in length. The mean harvest time for each bed was 321.7 s at an average ground speed of 0.56 km/hr. The harvester was capable of collecting 92.2 kg of tubers per row with a corresponding loss (missing or damaged) of 6.27 kg per row. The average time for turning the tractor from row to row was 102.7 s (1.43 minutes). The harvesting efficiency was 93.6% with coefficient of variation of 0.014. Table 3 shows the statistical analysis on the performance of sweet potato harvesting machine on VitAto variety. There were ten beds and each bed was 70 m in length. The mean harvesting time for each bed was 254.0 s, with a mean turning time between beds of 81.2 s. The mean harvesting speed was 0.99 km-hr$^{-1}$. The mean yield collected was 99.97 kg of tubers per row while missing and damaged tubers were 10.56 kg per row. The total mean harvesting efficiency was 90.5%, with a coefficient of variation of 0.02.

Based on Tables 2 and 3, the total mean harvesting efficiency showed no significant difference. These results also showed that the mean effective work-rate of the machine in bris soil was 93.64% and 90.49% for Telong and VitAto varieties, respectively. The average tractor speed and turning time was 0.56 km/hr and 102.7 s and 0.99 km/hr and 81.2 s for plots A and B, respectively. Differences in results of ground speed and turning time between plots may have been attributed to different operators. Neither operator had experience with the prototype machine. Other factors included plot length as Plot A had a 50 m seedbed length while Plot B was 70 m. A comparison of harvesting efficiency shows that there was not much difference between the plots. The harvesting efficiencies for both plots were above 90%. The effective field capacity or field work rate was obtained after considering time losses for unproductive tasks such as headland turn and travel between beds. The total productive time (harvesting time) and unproductive time (turning time) resulted in a sweet potato harvest capacity of 0.068 ha/hr for plot A and nearly double the capacity or 0.12 ha/hr for plot B. The average field work rate was 0.087 ha/hr or 34 man-hr/ha compared to manual harvesting of 150 man-hr/ha.

## CONCLUSION

Based on the results of the soil bin experiments, the best digging device was the V-shaped blade with a rake angle of 30$^{o}$ to be fixed onto the prototype sweet potato harvester. In the field study, the performance of the prototype sweet potato harvester was evaluated and observed. It was found that the machine efficiency in bris soil were 93.64% and 90.49% for Telong and VitAto varieties, respectively, with a corresponding work rate of 0.28 kg/sec (1.032 ton/hr) and 0.39kg/s (1.42 t/hr). The harvested yields were 92.2 kg/row for Telong and 99.97kg/row for VitAto varieties. Sweet potato losses and damages during harvesting operations were 6.27kg (6.36%) and 10.56kg (9.51%) for each variety respectively. The coefficient of variation (CV) of the machine operation was 0.014 and 0.02 for both varieties, respectively.

## ACKNOWLEDGEMENTS

The authors are grateful to Mr. Wan Fariz Wan Azman, Mr. Saleh Bardos, Mr. Salleh Rusdi, Mr. Aris Abdullah and Mr. Rusdi for their help and support during this study, as well as to Mr Mohd. Anaur Mat of Innovasi Anggerik Sdn. Bhd. and Mr. Tajeri Hashim of Tasco Sdn. Bhd.

for their assistance in the fabrication of the machine. Special acknowledgement also goes to Dr. Aimrun, Darius El Pebrian, Mr. Ahmad Ngalim, and Mr. Hossein for sharing their ideas.

## REFERENCES

Dash, S. K., Das, D. K., Paul, J. C., Mishra, J. N., & Swain, S. K. (1998). Development and Performance evaluation of bullock-drawn grounut diggers. *Agricultural Mechanization in Asia, Africa and Latin America (AMA), 29*(3), 67-70.

Md. Akhir, H., Taufik, A., Tan, S. L., & Zaharah A. (2005). *Machinery selection for sweet potato production*. 2nd International Symposium on sweet potato & cassava, 14-17 June 2005, Kuala Lumpur.

Md. Akhir, H., Desa A., & Rukunudin, I. H. (2008). Field performance of a potato digger as sweet potato harvesting tool. Paper presented at the Symposium on Engineering and Technology '08 (SET 2008) 15 & 16 Dec. 2008, Kuching, Sarawak.

Md. Akhir, H., & Desa Ahmad. (2005). *Mechanization of sweetpotato production in Bris soil*. National Conference on Agriculture Production in Bris soil, 6-7 Sept 2005, Kota Bharu Kelantan.

Md.Akhir, H., & Tan, S. L. ( 2002). *Mechanisation possibilities for sweetpotato production, International Symposium on Machinery and Mechatronics for Agricultural and Bioproduction Systems (ISMAB)*, Nov. 3-7, 2002. TAIWAN.

Tan, S. L., Abdul Aziz, A. M., Zaharah, A., Sukra, A. B., Md. Akhir, H., & Mohsin, Y. (2006). *Manual Teknologi Penanaman Ubi Keledek*. Serdang, MARDI.

Tan, S. L., Mooi, K. C., Engku Ismail, E. A., & Vimala, P. (2000). Sweet potato in the marginal soils environment of Malaysia. In E.T. Rasco Jr., & Vd. R. Amante (Eds.), *Chapter 2: sweet potato in Tropical Asia* (pp. 22-57). Book series no. 171/2000. Los Banos, Laguna; Philippine.

Tan, S. L. (1998). Agronomic modifications for mechanized planting of sweetpotato. *Trop. Agric*

Pertanika J. Sci. & Technol. 22 (1): 55 - 64 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Vegetative Swale for Treatment of Stormwater Runoff from Construction Site

**Ismail, A. F.*, Sapari, N. and Abdul Wahab, M. M.**

*Department of Civil Engineering, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia*

## ABSTRACT

Land development, especially construction works, increase storm water volumes and pollution loads into rivers and lakes. The temporary drainage system at construction sites, particularly during the construction stage discharges a large amount of pollutants that can damage the aquatic system of the receiving water bodies. The potential of vegetative swale to alleviate this problem was evaluated. The size of the constructed vegetative swale was 7cm deep, 400cm long and 15cm wide at the bottom, and 17cm wide at the top. The experiment was conducted batch wise by filling the storage tank with the run-off water from the construction site. The water was allowed to flow through a pipe into the retention basin to maintain uniform flow before it entered the swale. The study showed that the run-off infiltrated through the soil at a rate of 489.6 mm/hr. Samples of surface run-off and infiltration water were collected at the end and the bottom of the swale. The results indicate that chemical oxygen demand (COD), total suspended solid (TSS), turbidity, iron and zinc were reduced by 85.4%, 80.8%, 36.4%, 52.8% and 96.0%, respectively, by surface flow and 91.1%, 98.8%, 58.2% 55.5% and 98.1%, respectively, by infiltration. Removal of nitrate and phosphorus by the planted vegetation was 69.4% and 21.1%, respectively, by infiltration. However, nutrient removal by surface flow was negligible. In conclusion, the vegetative swale was able to improve the water quality of the storm water run-off from the construction site from Class V to Class III, according to the Interim National Water Quality Standards for Malaysia.

*Keywords:* Construction site, Storm water run-off, Vegetative swale, Water quality

## INTRODUCTION

The rapid rate of urban development in recent decades has led to significant changes in both the quantity and quality of storm water runoff (Line & White, 2007; Walsh *et al*., 2004). Areas disturbed for construction activity have undergone soil erosion which carries

*Article history:*
Received: 26 April 2011
Accepted: 18 January 2012

*E-mail addresses:*
asmafarah@gmail.com (Ismail, A. F.),
nasiman@petronas.com.my (Sapari, N.),
mubarakwahab@petronas.com.my (Abdul Wahab, M. M.)
*Corresponding Author

ISSN: 0128-7680 

pollutants (suspended solids, asphalt, sealants, oil, etc.) at rates from 2 to 40,000 times greater than the pre-construction conditions, and are important components of nonpoint source pollution that degrades surface water quality (Harbor, 1999). The pollutants found in storm water also cause ground water contamination with metals, suspended solids (SS) and oxygen depleting material (Pitt *et al*., 1996). Over the last 50 years the effects of SS on fish and aquatic life have been studied intensively throughout the world. It is now accepted that SS is an extremely important cause of water quality deterioration leading to aesthetic issues, higher cost of water treatment, decline in the fisheries resource and serious ecological degradation of the aquatic environment (Bilotta & Brazier, 2008).

According to the Malaysia Environmental Quality Report (DOE, 2006**),** forty-two (42) of the rivers in Malaysia are categorised as being heavily polluted with suspended solids. Sediment deposition in river channesl also causes flash floods due to the reduction in the flow-carrying capacity of rivers (Zakaria *et al*., 2004). In order to protect surface and ground water quality, urban development must be guided by plans that limit run-off and reduce pollutant loading (EPA, 2008). Storm water quality is considered as improved when the quality status is at least Class IIB, based on the Interim National Water Quality Standards for Malaysia (Zulkifli, 2008). Class IIB is typically for water that is suitable for recreational use with body contact. According to the standard, Class V is the worst quality where the water is not suitable for any use (DOE, 2006).

Swales are often called grassed channels or biofilters. They are vegetated open channels for water management and have been designed specifically to treat storm water for a specified volume of run-off. Polluted storm water runs as a shallow overland flow through the grass which grows on porous soil; the water, together with dissolved pollutants, infiltrates into the soil while suspended particles settle in the grass, while the outflow from the grass is of much better quality and of lesser quantity than the incoming water (Deletic, 2005). Biofiltration is becoming widely used, due to its flexibility in terms of size, location, configuration and appearance. Biofilters may also be used as vegetated strips (Bratieres *et al*., 2008). They operate by filtering run-off through planted filtration media and provide treatment through fine filtration, extended detention and biological uptake (Melbourne Water, 2005). Most studies of biofilter performance have reported its potential for the removal of total Kjeldahl nitrogen (TKN) and ammonia ($NH_3$) (Henderson *et al*., 2007; Davis *et al*., 2006; Hseih & Davis, 2005a,b). However, in almost all studies (both laboratory and field), nitrate ($NO_3^-$) has been shown to leach out, often resulting in poor total nitrogen (TN) removal. The percolation through soil and gravel layers caused all constituents to be reduced except for nitrate (Walsh *et al*., 1997). Total phosphorus (TP) removal has generally been moderate to good (Henderson *et al*., 2007; Davis *et al*., 2006; Hseih & Davis, 2005b). Eventually, measured removal efficiencies for total suspended solids (TSS) were found to be consistently high (>90%) in all of the reviewed studies (Hatt *et al*., 2007; Hatt *et al*., 2006; Hseih & Davis, 2005a,b). Studies showed that vegetative swales are now widely employed in urban environments as an effective best management practice for controlling pollutants in storm water run-off (Kirby *et al*., 2005).

This paper presents the findings of a study on the performance of vegetative swale in removing pollutants from storm water runoff from a construction site. The main objective was to determine the efficiency of vegetative swale in removing sediment (TSS and turbidity), nutrient and metal.

## MATERIALS AND METHOD

### *Study Site*

This study was conducted using storm water run-off from a residential construction site in Batu Gajah, Perak. The study area is a typical construction site which is not only involved in construction work but also domestic activities in the workers' quarters. This study was done during the construction phase of the residential development project.

### *Field Water Sampling and Analysis*

Initially, water samples from the construction site were collected for preliminary analysis. The samples were collected from the middle of the stream that was used as a temporary drainage. The sample bottle was lowered in the stream and rinsed three times with the run-off. The bottle was then filled by lowering to 60 percent depth of water column (making sure not to disturb the bottom sediment) while facing the current. One inch of air space was left in the bottle to allow for shaking or mixing before analysis, except for the samples for chemical oxygen demand (COD) and biochemical oxygen demand (BOD) analysis. The collected water samples were analysed in terms of TSS, turbidity, pH, COD, nitrate and metals according to the Standard Methods for the Examination of Water and Wastewater (APHA, 1995). Table 1 summarizes the water quality from the construction site as compared to the water quality standard for Malaysia.

TABLE 1
Results for water quality at construction site as compared to the Interim Water Quality Standards for Malaysia

| Parameters | | Interim National Water Quality Standards for Malaysia | | | | | | Construction site |
|---|---|---|---|---|---|---|---|---|
| | | I | IIA | IIB | III | IV | V | |
| TSS, mg/L | | 25 | 50 | 50 | 150 | 300 | 300 | 652 |
| Turbidity, NTU | | 5 | 50 | 50 | N/A | N/A | N/A | 2750 |
| pH | | 6.5-8.5 | 6-9 | 6-9 | 5-9 | 5-9 | N/A | 7.96 |
| COD, mg/L | | 10 | 25 | 25 | 50 | 100 | >100 | 246 |
| Nitrate, mg/L | | Natural levels | 7 | | N/A | 5 | Levels above IV | 180 |
| Metal, mg/L | Zinc | Natural levels | 5 | | 0.4* | 2 | Levels above IV | 1.3411 |
| | Iron | | 1 | | 1 | 1(leaf) 5(others) | | 4.6143 |

The water quality shows high concentration of TSS, metals and also nutrients in the storm water run-off during the active construction activities. The average class of the discharged water is Class V. Water samples from the construction site were then used to test the effectiveness of vegetative swale in removing the pollutants.

### *Experimental Setup*

A vegetative swale was constructed using galvanized iron, plywood and wooden planks. The size of the vegetated swale was 7cm deep, 400cm long, 15cm wide at the bottom and 17cm

wide at the top with side projection of 3cm wall provided to allow run-off water containment during the experiment (Fig.1 and Fig.2). The empty-bed volume of the vegetated swale was approximately 39 200cm$^3$. The trench bottom and walls were lined with linen cloth to retain the soil inside the swale. A bottom slope of 0.44% was kept to maintain the hydraulic gradient.

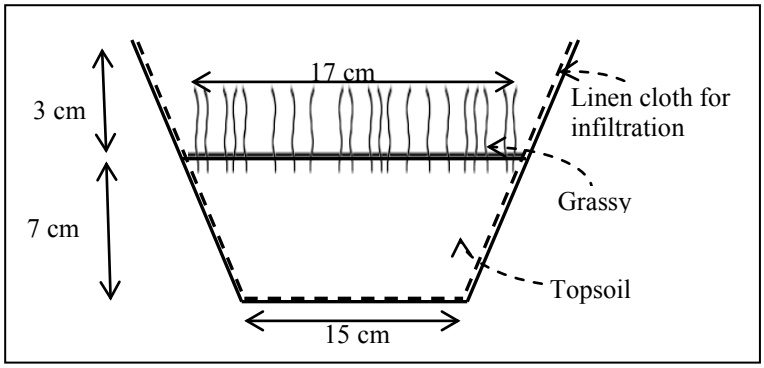



Fig.1. A cross section of vegetative swale

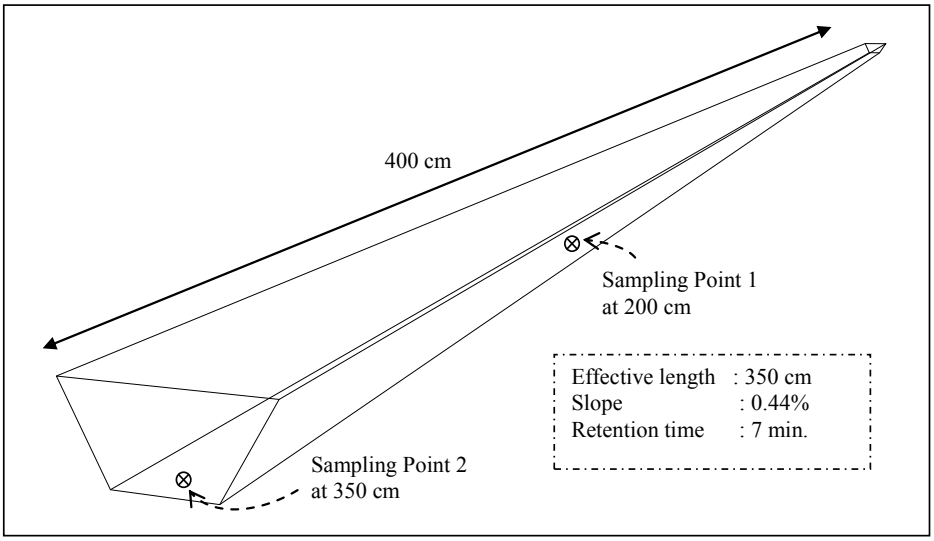



Fig.2: An overview of the infiltration points

The swale was provided with a feeding system which consisted of a storage tank and detention basin as shown in Fig.3. A detention basin was used to provide a constant inflow to the swale. A perforated PVC pipe was used to transfer the surface run-off from the storage tank to the swale. The grass species and the top soil used in this experiment were collected from Seri Iskandar wetland area.

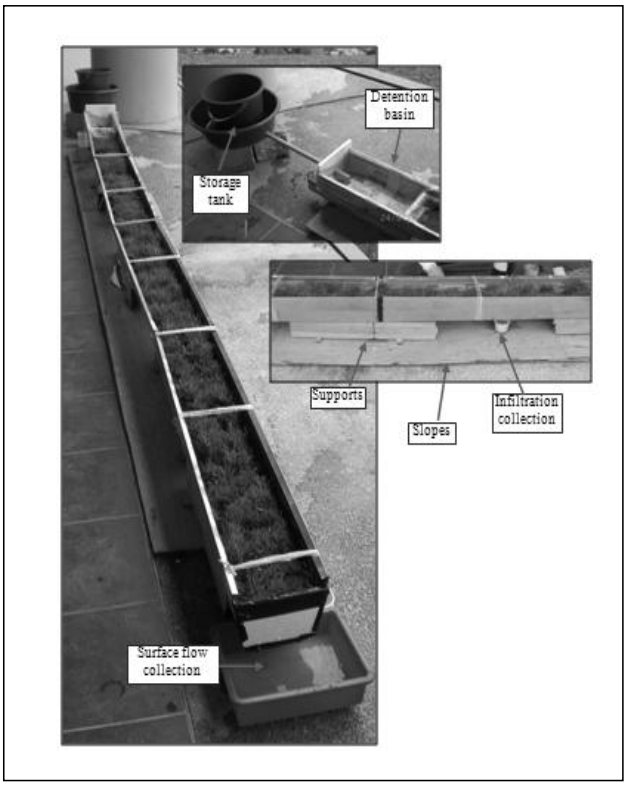



Fig.3: The constructed vegetative swale configuration

The experiment was conducted batch wise by filling the storage tank with 40 L run-off water from the construction site. The water was allowed to flow through the swale by gravity. On the other end of the swale, run-off water was allowed to overflow into a container and collected for water quality analysis. Water samples were also collected from two sampling points at the bottom of the swale, located at 200cm and 350cm from the influent point, respectively. This was done to measure the removal of pollutants through infiltration process.

The collected inflow and outflow samples were tested for total suspended solids (TSS), turbidity, pH, nitrate ($NO_3^-$), phosphorus ($PO_4^{3-}$), chemical oxygen demand (COD), and metals (zinc and iron). All samples were analysed according to the Standard Methods for the Examination of Water and Wastewater (APHA, 1995).

## RESULTS AND DISCUSSION

### *Infiltration Rate*

The infiltration rate, *F*, is the velocity or speed at which water enters into the soil. It was measured by the depth of water (mm) that can enter the soil in one hour (Abdelhak, 2009). Out of 40 L of storm water run-off, only 6 L flowed through as surface flow. Thus, the infiltration rate for this swale system was calculated based on the 34 L infiltrated in 7 minutes run-off, as follows:

$$F = \frac{volume\ of\ water\ runoff}{surface\ area}$$

$$= \frac{\left(34 \times 10^3 \times 10^3\right) mm^3}{\left(170mm \times 3500mm\right) \times 7min}$$

$$= 8.16\ mm/min$$

$$= 489.6\ mm/hr$$

*Sediment Removal*

Storm water discharges from construction sites carry large sediment loads resulting in highly turbid water (Patil *et al*., 2011). Turbidity and TSS removal occur mainly by infiltration through the soil and deposition during surface flow. The infiltration reduced turbidity and TSS of the run-off by 96.6-98.8% and 50.9-58.2%, respectively (Fig.4). The surface flow showed lower removal efficiency for turbidity (80.8%) and TSS (36.4%), mainly because solid particles were carried to the outlet by the surface water.

Turbidity of the influent was reduced from 2750 NTU to 1750 NTU (surface flow), 1350 NTU (sampling point 1) and 1150 NTU (sampling point 2). Influent TSS of 652 mg/L was reduced to 125 mg/L (surface flow), 20 mg/L (sampling point 1) and 8 mg/ L (sampling point 2). Dillaha *et al*. (1986) assessed the pollutants' removal by vegetative filter strip and found that the sediment reduction was 81-91%.

*Metal Removal*

Infiltration reduced iron and zinc of the run-off by 54.2-55.5% and 87.2-98.1%, respectively (Fig.5). The surface flow showed lower removal efficiency for iron (52.8%) and zinc (96%). Iron in the influent was reduced from 4.614 mg/L to 2.178 mg/L (surface flow), 2.113 mg/L (sampling point 1) and 2.054 mg/ L (sampling point 2). Zinc in the influent was reduced from 1.341 mg/L to 0.054 mg/L (surface flow), 0.038 mg/L (sampling point 1) and 0.026 mg/ L (sampling point 2).

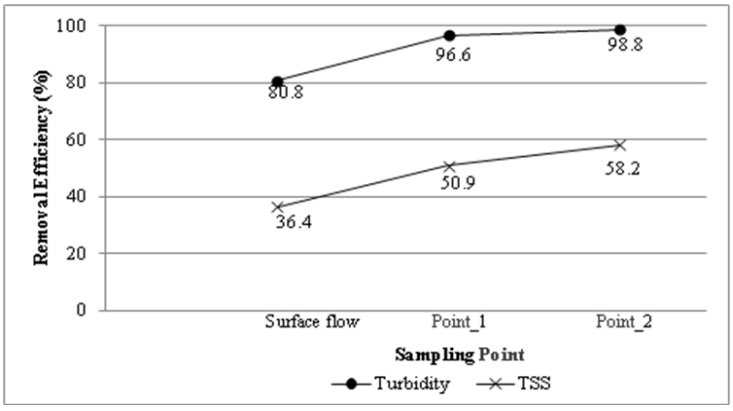



Fig.4: Turbidity and TSS removal efficiency by surface flow and infiltration through the swale

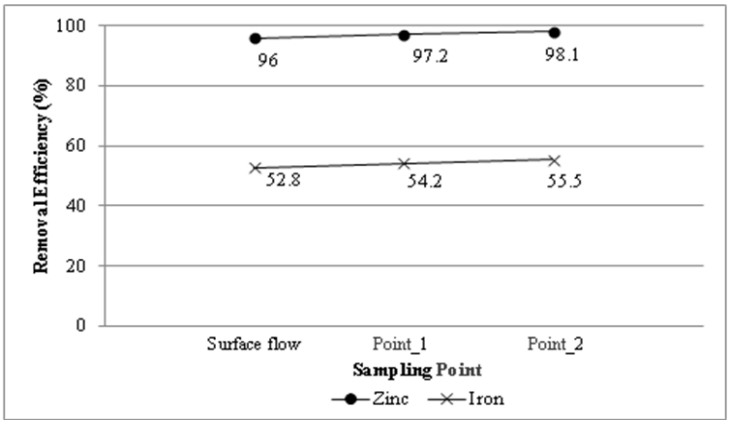



Fig.5: Metal removal efficiency by surface flow and infiltration through the swale

Similar results were obtained by Delgado *et al*. (1995), where the typical percentage reductions of zinc by vegetative filter strip was in the range of 75-84%. Yousef *et al*. (1987) deduced that the removal of metals will be greater for species present as charged ion, with the dominant removal mechanism being adsorption onto particles which are then removed by sedimentation.

*Nutrient Removal*

Results on the reduction of COD are presented in Fig.6. Significant COD reduction of more than 85% was achieved both by infiltration and surface flow. The COD in the influent was reduced from 246 mg/L to 36 mg/L (surface flow), 24 mg/L (sampling point 1) and 22 mg/L (sampling point 2). As expected, the results from surface flow indicated less reduction as compared to the results by infiltration. Removal of phosphorus and nitrate by the swale was observed in the water that infiltrated through the soil. The removal was within the range of 9.6-21.1% and 63.9-69.4% for phosphorus and nitrate. This could be due to plant uptake through the root system or by denitrification in the soil. Flow through the surface of the swale did not remove any nutrient.

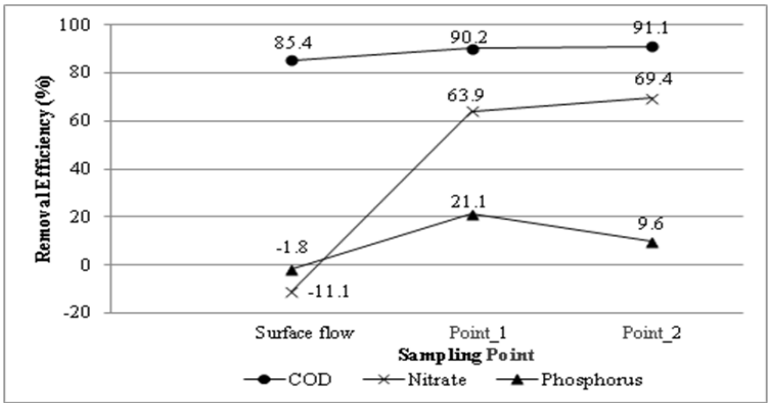



Fig.6: Nutrient removal efficiency by surface flow and infiltration through the swale

Higher value of nitrate in the surface flow (200 mg/L) as compared to the influent (180 mg/L) was probably due to nitrification of ammonia. Initial phosphorus value of 1.66 mg/L also increased to 1.69 mg/L in the surface flow and only reduced to 1.31-1.50 mg/L through infiltration. These results were similar to the findings by Dillaha *et al*. (1986), who found that soluble phosphorus was not successfully removed and in some cases, even increased as a result of solubilisation and leaching of previously accumulated phosphorus. Nutrient removal, however, can be optimised by selecting suitable species with higher capacities for assimilation of inorganic nitrogen and phosphorus and conversion into plant biomass (Vymazal, 2007; Greenway, 2003).

## CONCLUSION

In conclusion, COD, turbidity, TSS, iron and zinc were reduced by 85.4%, 36.4%, 80.8%, 52.8% and 96%, respectively by surface flow. Infiltration through the swale provided higher removals than by the surface flow. Removal by infiltration for TSS, turbidity, iron, zinc and COD were 98.8%, 58.2%, 55.5%, 98.1% and 91.1%, respectively. Removal of nutrients by infiltration was 69.4% for nitrate and 21.1% for phosphorus. However, the surface flow did not remove any nutrient. It was found that vegetative swale could improve the water quality of the storm water run-off from Class V to Class III according to the Interim National Water Quality Standards for Malaysia.

## ACKNOWLEDGEMENTS

This work was undertaken as part of a research project conducted by the Civil Engineering Department, funded by Universiti Technologi PETRONAS (UTP). The project benefited from the advice of Ar Wan Rahmah Wan Zaki, Dr. Faris Khamidi and Siti Ameera Ashahary. The researchers also gratefully acknowledge the contributions of the laboratory technologist at UTP.

## REFERENCES

Abdelhak, M. (2009). *Use of Polymers and Biopolymers for Water Retention and Soil Stabilization at Algerian Arid and Semi-Arid Soils.* (Ph.D Dissertation). Mentouri University of Constantine.

APHA (1995). *Standard Methods for the Examination of Water and Wastewater* (19th edn). Washington, DC, USA: American Public Health Association/ American Water Works Association Water Environment Federation.

Bilotta, G. S., & Brazier, R. E. (2008). Understanding the influence of suspended solids on water quality and aquatic biota. *Water Research*, *42*, 2849-2861.

Bratieres, K., Fletcher, T. D, Deletic, A., & Zinger, Y. (2008). Nutrient and sediment removal by stormwater biofilters: a large-scale design optimisation study. *Water Research*, *42,* 3930-3940.

Davis, A. P., Shokouhian, M., Sharma, H., & Minami, C. (2006). Water quality improvement through bioretention media: nitrogen and phosphorus removal. *Water Environment Research, 78,* 284–293.

Deletic, A. (2005). Sediment transport in urban runoff over grassed areas. *Hydrology*, *301*, 108-122.

Delgado, A. N, Periago, E. L., & Viqueir, F. D. (1995). Vegetated filter strips for wastewater purification: a review. *Bioresource Technology*, *51*, 13-22.

Dillaha, T. A., Sherrard, J. H., Lee, D., Shanholtz, V. O., Mostaghimi, S., & Magette, W. L. (1986). Use of vegetative strips to minimize sediment and phosphorus losses from feedlots: Phase I. Experimental plot studies. *Virginia Water Resource Research Center Bull*., 151.

DOE (2006). *Malaysia Environmental Quality Report,* 2006, Malaysia: Department of Environment. Ministry of Natural Resources and Environment.

EPA (2008). *Managing Urban Runoff.* Pointer No. 7; EPA841-F-96-004G. Retrieved June 2008, from http://www.epa.gov/nps/facts/point7.htm

Greenway, M. (2003). Suitability of macrophytes for nutrient removal from surface flow constructed wetlands receiving secondary treated sewage effluent in Queensland, Australia. *Water Science and Technology, 48*, 121-128.

Harbor, J. (1999). Engineering geomorphology at the cutting edge of land disturbance: erosion and sediment control on construction sites. *Journal of Geomorphology, 31*, 247–263.

Hatt, B. E., Deletic, A., & Fletcher, T. D. (2007). Stormwater reuse: designing biofiltration systems for reliable treatment. *Water Science and Technology*, *55*, 201–209.

Hatt, B. E., Siriwardene, N., Deletic, A., & Fletcher, T. D. (2006). Filter media for stormwater treatment and recycling: the influence of hydraulic properties of flow on pollutant removal. *Water Science and Technology, 54*, 263–271.

Henderson, C., Greenway, M., & Phillips, I. (2007). Removal of dissolved nitrogen, phosphorus and carbon from stormwater by biofiltration mesocosms. *Water Science and Technology*, *55,* 183–191.

Hsieh, C. H., & Davis, A. P. (2005a). Evaluation and optimization of bioretention media for treatment of urban stormwater runoff. *Journal of Environmental Engineering, 131,* 1521–1531.

Hsieh, C. H., & Davis, A. P. (2005b). Multiple-event study of bioretention for treatment of urban storm water runoff. *Water Science and Technology, 51,* 177–181.

Kirby, J. T., Durrans, S. R., Pitt, R., & Johnson, P. D. (2005). Hydraulic resistance in grass swales designed for small flow conveyance. *Journal of Hydraulic Engineering*, *131,* 65-68.

Line, D. E., & White, N. M. (2007). Effects of development on runoff and pollutant export. *Water Environment Research*, *79,* 185–190.

Melbourne Water (2005). *WSUD Engineering Procedure: Stormwater.* Melbourne, Australia: CSIRO Publishing.

Patil, S. S., Barfield, B. J., & Wilber, G. G. (2011). Turbidity modeling based on the concentration of total suspended solids for stormwater runoff from construction and development sites. *World Environmental and Water Resources Congress 2011: Bearing Knowledge for Sustainability - Proceedings of the 2011 World Environmental and Water Resources Congress*, pp. 477-486.

Pitt, R., Clark, S., Parmer, K., & Field, F. (1996). *Groundwater Contamination from Stormwater Infiltration* (Chapter 3). Chelsea, Michigan: Ann Arbor Press, Inc.

Vymazal, J. (2007). Removal of nutrients in various types of constructed wetlands. *Science of the Total Environment, 380*, 45-65.

Walsh, C. J., Leonard, A. W., Ladson, A. R., & Fletcher, T. D. (2004). *Urban Stormwater and the Ecology of Streams* (p. 44). Technical Report, Cooperative Research Centre for Freshwater Ecology (CRCFE) and Cooperative Research Centre for Catchment Hydrology, Canberra, Australia.

Walsh, P. M., Barrett, M. E., Malina, J. F., & Charbeneau, R. J. (1997). *Use of Vegetative Controls for Treatment of Highway Runoff*. Technical Report, Center for Research in Water Resources, Bureau of Engineering Research, University of Texas, Austin.

Yousef, Y. A., Hvitved-Jacobsen, T., Wanielista, M. P., & Harper, H. H. (1987). Removal of contaminants in highway runoff flowing through swales, *Science of the Total Environment, 59*, 391–399.

Zakaria, N. A., Ghani, A. Ab., Abdullah, R., Sidek, L. M., Kassim, A. H., & Ainan, A. (2004). MSMA - A New urban Stormwater Management Manual for Malaysia. *The 6th International Conference on Hydroscience and Engineering*. Brisbane, Australia.

Zulkifli, A. R. (2008). *Water Quality Management in Malaysia*. Slide Presentation, Department of Environment Malaysia. Retrieved July 2008, from http://www.iges.or.jp/jp/ltp/pdf/fr2.pdf

Pertanika J. Sci. & Technol. 22 (1): 65 - 79 (2014)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Physical Properties of Liberica Coffee (*Coffea liberica*) Berries and Beans

**Ismail, I.*, Anuar, M. S. and Shamsudin, R.**

*Department of Process and Food Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

Liberica coffee is the most important coffee species grown in Malaysia. However, there is little or no research at all conducted on coffee berries and green coffee beans since the plant itself is a low income crop in Malaysia. Therefore, research on Malaysian Liberica coffee can help to increase the knowledge of coffee farmers and coffee manufacturers in the processing and handling of the coffee. Physical properties of Liberica coffee berries and beans were investigated the current study. The properties investigated include the size, mass, density, coefficient of friction, angle of repose, fracture force and colour. In comparison to Arabica and Robusta coffee, Liberica coffee has the biggest size, mass, true density and fracture force values but were lower in bulk density in both berries and beans. The Liberica berries and beans were found to be orange-ish and yellowish colour respectively. Angle of repose was low and approximately similar in berries and beans while jute fibre gave the highest friction to both Liberica berries and beans.

*Keywords:* Coffee bean, Liberica, physical properties

## INTRODUCTION

Coffee is one of the most popular beverages consumed around the world. Besides coffee based beverages, coffee beans are also used to produce other products such as coffee essence, coffee syrup and coffee caramel (Zainun 1991), which are widely utilised in the confectionary industry. The major coffee species planted and traded around the world are Arabica (80%) and Robusta (20%), while the minor species are Liberica and Excelsa which accounts for less than 1% (Wintgens, 2009). In Malaysia, however, the situation is different where the major coffee species planted are Liberica (73%) and Robusta (27%) while Arabica is only planted in highland areas in negligible quantities. This is due to the climatic condition which makes it

*Article history:*
Received: 20 May 2011
Accepted: 2 March 2012

*E-mail addresses:*
iris_yamaki@yahoo.com (Ismail, I.),
mshamsul@upm.edu.my (Anuar, M. S.),
rosnahs@upm.edu.my (Shamsudin, R.)
*Corresponding Author

ISSN: 0128-7680 

unsuitable to grow Arabica in lowland (MARDI 1986). Azmil (1991) stated that Liberica and Robusta coffee can be suitably cultivated in Malaysia due to its optimum growth temperature of 18 ºC to 28 ºC (maximum at 34ºC). However, Arabica can only be grown at temperatures below 23 ºC which is applicable in highland areas only, such as Cameron Highlands in Pahang.

Before a cup of coffee is made, freshly harvested coffee berries have to undergo various processes in the post-harvest operation. Then, the coffee beans will undergo roasting, grinding, packing and transporting processes which require knowledge of basic physical properties data of the coffee beans. Hence, the basic information on physical properties of coffee berries and the green coffee beans are crucial in the design of the coffee processing machinery, and in the development of a mechanised coffee processing system (Ghosh & Gacanja 1970). The coffee berries' size, density, and crushing strength are important in classification of berries, and also in the design of pulper. In designing the handling system, grading and hulling of coffee beans, the information on size, friction, angle of repose, crushing strength and bulk density is crucial (Chandrasekar & Viswanathan 1999). Due to the greater demand for Arabica and Robusta coffee, most commercial coffee processing machinery have standard sizes according to Arabica and Robusta. Hence, it will be difficult to process the Liberica coffee in the Malaysian coffee industry since it possesses different physical attributes.

Many investigations have been done on the physical properties of Arabica and Robusta coffee berries and beans (Chandrasekar & Viswanathan, 1999; Mendonca *et al*., 2009; Franca *et al*., 2005; Afonso Jr. *et al*., 2007). However, data on physical properties of Liberica coffee is scarce, whereby the latest published data was in 1991 (Ghawas & Rubiah, 1991) and some properties which are important in evaluating quality of coffee in the literature are based on non-Malaysian Liberica coffee (Campa *et al*., 2004; Campa *et al*., 2005). Therefore, it is the aim of the current study to evaluate the physical properties of Malaysian Liberica coffee species.

## MATERIALS AND METHODS

### *Coffee Sample*

Liberica coffee berries and green beans were obtained from a factory producing coffee -Kilang Kopi FAMA Banting (Selangor, Malaysia) - in May, 2010. The coffee berries and beans were then subjected to sorting and selecting processes which were done manually. For berries, only the ripe and fully ripe berries were used whereas the black, immature and defective berries were rejected. Then, the coffee berries were separated into pulp and mucilageous parchments manually. In the current study, the coffee pulp consisted of the epicarp (skin) and mesocarp (pulp). The mucilageous coffee parchment consisted of mucilage, endocarp (parchment), integument (silverskin) and endosperm (bean), as shown in Fig.1. The coffee beans used consisted of mixed beans (which may contain sour or immature beans) but defective beans (black, partly black, broken or infested) were discarded. The beans were processed using a dry method whereby the crushed berries were sun-dried for 7-21 days, depending upon the climate conditions. The aim of the sun-drying process is to reduce the moisture content of the coffee to 12% (Zainun, 1991). The beans were wholly processed by the coffee farmers.

*Moisture Content*

The coffee pulp was cut into sizes of approximately 0.5 cm (width) x 2.5 cm (length) after being separated from the mucilageous parchment and before it can be tested for its moisture content. The mucilageous parchment and the green coffee beans were used as a whole without any size reduction procedure. Moisture content was determined by drying 3-5g of sample at 105°C in an oven (UNB400, Memmert, Germany) for 24 hours. Triplicate measurements were done (Reh *et al*., 2006).

*Water Activity*

The pre-test procedure for coffee pulp, mucilageous parchment and green coffee beans were the same as in Section 2.2, before they can be tested for their water activity. The water activity of sample was determined using a water activity meter (Fast lab, GBX, France). Five measurements were recorded (Vasconcelos *et al*., 2007).

*Size and Dimension*

A digital vernier calliper with 0.01 mm accuracy (Series 500, Mitutoyo, Japan) was used to measure the dimensions of the 100 samples of coffee berries and beans, as shown in Fig.2. The length (L), width (W) and thickness (T) measurements were taken. The L, W and T can also be referred as major diameter, intermediate diameter and minor diameter respectively. The mean diameter of geometric ($D_g$), arithmetic ($D_a$), square ($D_s$) and equivalent ($D_e$) were calculated from Equations (1), (2), (3) and (4), respectively. Sphericity and aspect ratio were calculated by Equations (5) and (6) (Mohsenin, 1986). Surface area was calculated using Equation (7), according to McCabe and others (1986).

$$D_g = (WTL)^{\frac{1}{3}} \quad (1)$$

$$D_a = \frac{1}{3}(W+T+L) \quad (2)$$

$$D_s = \left[\frac{LW+WT+TL}{3}\right]^{\frac{1}{2}} \quad (3)$$

$$D_{equiv} = \frac{D_g+D_a+D_s}{3} \quad (4)$$

$$S_{ph} = \frac{D_g}{L} \quad (5)$$

$$R_a = \frac{W}{L} \quad (6)$$

$$S_d = (\pi D_g)^2 \quad (7)$$

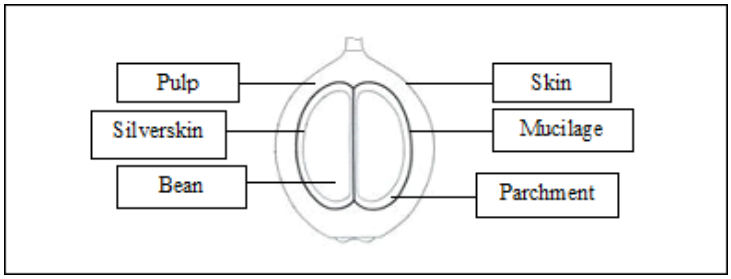



Fig.1: Longitudinal sections of coffee fruit

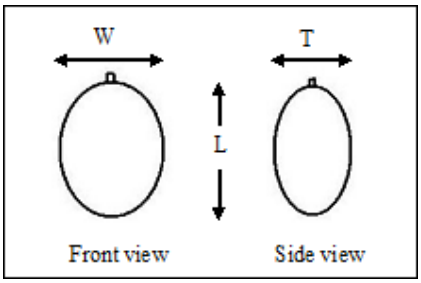



Fig.2: Dimensions of coffee berries

*Mass*

The mass was determined by a digital balance (ER-120A, AND, Japan), with an accuracy of 0.0001 g. A total of 100 samples were weighed and the values were then divided by 100 to get the average sample weight. After the berries' mass was obtained, each berry was separated into pulp and mucilageous parchments. Then, the pulp and mucilageous parchments were weighed separately to determine their proportions in the berries. Three replications were done (Bart-Plange & Baryeh 2003).

*Volume*

The volume of berries was estimated by water displacement method (Chandrasekar & Viswanathan, 1999). The test was done 10 times. Following Dutra and others (2001), the bean can be assumed as half a triaxial ellipsoid. The bean volume was calculated using Equation (8).

$$V = \frac{2}{3}\pi abc \tag{8}$$

where 2a, 2b and c are the length, width and thickness of the bean respectively.

*Density*

The volume of berries obtained in Section 2.6 was used to calculate the berry's true density. In bean true density, 100 beans were first weighed and then each bean volume was calculated by using equation (8). Bean true density was calculated by dividing the weight of 100 beans

with the total volume of the 100 beans (Franca *et al*., 2005). Bulk density was determined by filling a sample into a 500 ml measuring cylinder. By dividing weight of filled sample with the cylinder volume, bulk density was obtained. The test was done in triplicate (Chandrasekar & Viswanathan, 1999).

*Coefficient of Friction*

In determining the coefficient of friction, friction surfaces used were aluminium, plywood and jute fibre. The sample was filled into a 10 cm (diameter) x 5 cm (height) PVC cylinder which was placed on the friction surface. The PVC cylinder was then lifted 3 mm from the friction surface so that the cylinder will not touch the friction surface. The friction surface was lifted slowly until the cylinder started to slide down. The angle (θ) between the lifted friction surface and the horizontal surface is known as the static angle of friction between the sample and the friction surface. Coefficient of static friction was obtained from Equation (9) (Bart-Plange & Baryeh, 2003). Five measurements were taken.

$$\mu = \tan\theta \tag{9}$$

*Filling Angle of Repose*

By following Razavi *et al*. (2007) with a slight modification, filling angle of repose ($\theta_f$) was acquired by filling a round container with a layer of sample. Then, a PVC cylinder (16 cm x 16 cm) was placed at the centre of the container, on top of the sample layer and the cylinder was filled with sample. The cylinder was lifted slowly until the sample formed a cone. The height of the cone was measured and the angle of repose was calculated using Equation (10). The test was replicated five times.

$$\theta_f = \tan^{-1}\left[\frac{2h}{D}\right] \tag{10}$$

where h is the height of the cone and D is the cone diameter.

*Emptying Angle of Repose*

For emptying angle of repose, a plywood box with a dimension of 20 cm (L) x 15 cm (W) x 30 cm (H) was filled with the sample. The front panel of the box was quickly slid upward to let the sample flow out freely to form a natural heap. Emptying angle of repose ($\theta_e$) was determined by measuring the height of the beans at two points in the sloping bean heap and the horizontal distance between the points as in Equation (11) (Chandrasekar & Viswanathan, 1999).

$$(\theta_e) = \frac{\tan^{-1}(h_2 - h_1)}{x_2 - x_1} \tag{11}$$

where h is the height of sample at two points in the sloping heap and x is the horizontal distance between the two points in the sloping heap. Five measurements were recorded for this test.

*Fracture Force*

Samples were randomly chosen and positioned horizontally on the platform of Instron testing machine (Instron, 5566, USA). Berries were placed with its front side (Fig.2) facing upward while the beans were placed with their flat sides facing upward. Uniaxial compression was conducted at the rate of 0.83 cm/s. The maximum force recorded during the test before the sample fracture was taken as the fracture force of the sample. From the fracture force, deformation was determined. 10 samples were recorded for its values (Pittia *et al*., 2007).

*Colour*

Colour of berries and beans were observed using a colour meter (CR-10, Konica Minolta, Japan). The values of $L^*$, $a^*$ and $b^*$ obtained were used to determine the chroma and hue angle by using Equations (12) and (13) (Leite da Silveira *et al*., 2007). The test was replicated five times.

$$\text{Chroma, } c^* = [(a^*)^2 + (b^*)^2]^{1/2} \quad (12)$$

$$\text{Hue angle, } h^* = \tan^{-1}(b^*/a^*) \quad (13)$$

## RESULTS AND DISCUSSION

*Initial Moisture Content and Water Activity*

Based on Table 1, the moisture content for pulp was 20.86% and 37.78% for mucilageous parchment. Hence, the mean berries moisture content was 29.32% whereas the beans had 12.53% moisture content. Water activity is 0.943 for pulp and this is 0.946 for mucilageous parchment, with an average of 0.945 for the berries. From the berries' water activity, the berries' rate of deterioration will be fast since the value is near to 1. This also indicates that the berries need to be processed immediately after harvesting to prevent any loss of quality. The bean water activity is 0.603. Both moisture content and water activity values of Liberica beans are in the safe range from further deterioration (Reh *et al*., 2006; Quiroz *et al*., 2004).

TABLE 1
Initial moisture content and water activity of Liberica coffee

| Coffee | Moisture content (%) | Water activity |
|---|---|---|
| Berries | 29.32 | 0.945 |
| - Pulp | 20.86 | 0.943 |
| - Mucilageous parchment | 37.78 | 0.946 |
| Bean | 12.53 | 0.603 |

*Dimension*

Dimension or size is important in the screening as well as the grading process, and in evaluating the quality of the food (Sahin & Sumnu, 2006). Dimensions of Liberica coffee berries are displayed in Table 2. Based on the result, the mean length, width and thickness of Liberica berries are 23.44, 20.37 and 18.89 mm, respectively. It was found that the values of the maximum length (29.77 mm) and maximum width (26.11 mm) of Liberica berries were higher compared to 25 mm (length) and 21 mm (thickness) in previous works (Wrigley, 1988). Meanwhile, the mean length (23.44 mm) and width (20.37 mm) of Liberica coffee berries in this study were larger than Arabica (15.65mm, 13.90 mm) and Robusta (13.50 mm, 11.50 mm), as found in Chandrasekar's and Viswanathan's (1999) research. Hence, compared to Arabica and Robusta berries, Liberica berries are bigger and need different processing equipment such as screener and siever.

The dimension of Liberica beans is presented in Table 3. The mean length, width and thickness of Liberica beans are 11.99 mm, 7.67 mm and 4.67 mm, respectively. The maximum length of Liberica beans is 15.16 mm. However, the mean length (11.99 mm) is lower than in a previous MARDI (1986) study, which was 12.7 mm. This could be attributed to the usage of a mixture of random coffee bean sample in this study, where the immature or green bean may have been included. In addition, the Liberica beans' mean length is longer than Arabica (10.5 mm) and Robusta (8 mm) beans according to MARDI (1986). Therefore, Liberica beans have a bigger dimension than Arabica and Robusta beans and also need equipment (sieves, graders) suited to its bigger size.

TABLE 2
Dimension of Liberica coffee berries

| Properties | Minimum | Mean | Maximum | Std. dev. |
|---|---|---|---|---|
| Length (mm) | 19.6 | 23.44 | 29.77 | 1.84 |
| Width (mm) | 16.04 | 20.37 | 26.11 | 1.62 |
| Thickness (mm) | 15.52 | 18.89 | 24.18 | 1.40 |
| Diameter (mm) | | | | |
| - Arithmetic | 17.05 | 20.90 | 26.69 | 1.37 |
| - Geometric | 16.96 | 20.80 | 26.59 | 1.37 |
| - Square mean | 17.01 | 20.85 | 26.64 | 1.37 |
| - Equivalent | 17.01 | 20.85 | 26.64 | 1.37 |
| Sphericity | 0.87 | 0.89 | 0.90 | 0.05 |
| Aspect ratio | 0.82 | 0.87 | 0.89 | 0.07 |
| Volume ($m^3$) | $5000 \times 10^{-9}$ | $5500 \times 10^{-9}$ | $6000 \times 10^{-9}$ | $527.05 \times 10^{-9}$ |
| Surface area ($mm^2$) | 2839.97 | 4271.09 | 6978.69 | 580.76 |

TABLE 3
Dimensions of Liberica coffee beans

| Properties | Minimum | Mean | Maximum | Std. dev. |
|---|---|---|---|---|
| Length (mm) | 9.64 | 11.99 | 15.16 | 1.03 |
| Width (mm) | 6.26 | 7.67 | 9.25 | 0.57 |
| Thickness (mm) | 3.78 | 4.67 | 6.31 | 0.51 |
| Diameter (mm) | | | | |
| - Arithmetic | 6.56 | 8.11 | 10.24 | 0.49 |
| - Geometric | 6.11 | 7.53 | 9.60 | 0.46 |
| - Square mean | 6.34 | 7.82 | 9.90 | 0.47 |
| - Equivalent | 6.34 | 7.82 | 9.91 | 0.47 |
| Sphericity | 0.62 | 0.63 | 0.65 | 0.04 |
| Aspect ratio | 0.61 | 0.64 | 0.67 | 0.06 |
| Volume ($m^3$) | $119.45 \times 10^{-9}$ | $225.96 \times 10^{-9}$ | $463.37 \times 10^{-9}$ | $41.95 \times 10^{-9}$ |
| Surface area ($mm^2$) | 368.56 | 561.66 | 909.90 | 68.75 |

*Sphericity and Aspect Ratio*

Sphericity and aspect ratio has been used in heat and mass transfer calculations (Sahin & Sumnu 2006). Sphericity and aspect ratio values of Liberica berries can be observed in Table 2. Liberica berries have 0.89 for mean sphericity and 0.87 for mean aspect ratio. Based on the sphericity value of Liberica berries, it can be observed that the berries have a spherical-like shape since the sphericity exceeds 70% (Eke *et al*., 2007). Comparing to other species in the work of Afonso Jr. *et al.* (2007). Liberica berries sphericity is only slightly higher than Arabica and Robusta, whereby both berries have a 0.82 sphericity value. From Table 2, with an aspect ratio value of more than 70%, Liberica berries are also more likely to roll than to slide. Hence, Liberica berries have higher sphericity than other coffee species and move by rolling rather than to slide on top of each berry. Sphericity will affect the materials' angle of repose since rounded materials give low angle of repose which means high flowability of materials (Sahin & Sumnu, 2006).

Referring to Table 3, the mean sphericity and the mean aspect ratio of Liberica beans are 0.63 and 0.64, respectively. Liberica beans' sphericity is lower than other major species where the values are 0.68 (Arabica) and 0.72 (Robusta). From Table 3, with an aspect ratio value of less than 70%, Liberica beans are more likely to slide. Low aspect ratio of the Liberica beans also indicates the tendency of the bean shape to be oblong like (Eke *et al*., 2007). Therefore, Liberica beans have lower sphericity compared to Arabica and Robusta beans and also move on top of each bean by sliding. There is a significant difference in sphericity and aspect ratio between Liberica berries and beans.

*Volume and Surface Area*

Since coffee berries in Malaysia are processed using the dry method, the berries' volume is important in drying simulation models where the change of volume will lead to a major error in the modelling (Lang & Sokhansaj, 1993). Volume also plays an important role in the roasting process where roasting makes the beans expand (Franca *et al*., 2005). According to Table 2, the mean volume and mean surface area of Liberica berries are 5500 x $10^{-9}$ $m^3$ and 4271.09 $mm^2$, respectively. In the current study, the mean volume of Liberica berries is comparatively higher than Arabica and Robusta. Based on a previous study by Afonso Jr. *et al.* (2003), the maximum volume of Arabica and Robusta berries are 1100 x $10^{-9}$ $m^3$ and 800 x $10^{-9}$ $m^3$, respectively. Liberica berries also have a higher mean surface area of 4271.09 $mm^2$ compared to the maximum values for Arabica (600 $mm^2$) and Robusta (450 $mm^2$) (Afonso Jr. *et al.,* 2003). To conclude, Liberica berries have bigger volume and surface area values than other major coffee species.

The volume and surface area of Liberica beans are presented in Table 3. Based on the result, the mean volume and mean surface area of Liberica beans are 225.96 x $10^{-9}$ $m^3$ and 561.66 $mm^2$. The mean volume of Liberica beans is approximately twice the volume of Arabica (114.9 x $10^{-9}$ $m^3$) and triple the volume of Robusta (73.7 x $10^{-9}$ $m^3$) beans, based on a study by Mendonca *et al.* (2009). This means that Liberica beans have a bigger volume compared to Arabica and Robusta coffee beans.

*Mass*

Mass is important in calculation of densities and it can determine quality of materials. From Table 4, the mass composition of Liberica berries are shown. An average Liberica berry's mass is 5.31 g. It is approximately 4 times heavier than the Arabica (1.65 g) and Robusta (1.35g) berry, according to Chandrasekar and Viswanathan (1999). In this study, the pulp to parchment ratio is 60:40 for Liberica and 40:60 for both Arabica and Robusta from the study by Chandrasekar and Viswanathan (1999). This indicates Liberica berries are heavier than Arabica and Robusta berries. Liberica berries also have thicker epicarp (skin) and mesocarp (pulp) with a smaller portion of endocarp (parchment), integument (silverskin) and endosperm (bean) compared to other species of berries. In addition, Ghawas and Rubiah (1991) stated that Liberica berries have a higher quantity of mucilage compared to Arabica and Robusta coffee. This affects the drying rate of Liberica coffee since it is heavier and has more mucilage than other types of coffee.

The mass of an average Liberica bean is 0.26 g (from Table 4). By referring to a study by Ramalakshmi *et al*. (2007), the average bean mass of a Liberica bean evaluated in the current study is higher than Arabica (0.19 g) and Robusta (0.20 g). Hence, Liberica beans are found to be heavier than Arabica and Robusta beans.

TABLE 4
Mass composition of Liberica coffee berries and beans

| Properties (g) | Mass | Std.dev. | Percentage (%) |
|---|---|---|---|
| 100 berries | 530.60 | 5.50 | - |
| - Pulp | 318.34 | 3.30 | 60 |
| - Parchment | 212.23 | 2.20 | 40 |
| One berry | 5.31 | 0.06 | - |
| 100 beans | 25.72 | 0.90 | - |
| One bean | 0.26 | 0.01 | - |

*Density*

Density is required in a separation process such as hulling and in quality evaluation (Mohsenin, 1978). Bulk density is the material's density when packed or stacked in bulk while true density is the density of the pure substance (Sahin & Sumnu, 2006). The Liberica berries' bulk density is 497.83 kg/m$^3$ while the true density is 1112.13 kg/m$^3$, according to Table 5. Following previous literature, bulk density of Liberica berries is lower than Arabica and Robusta, where the values are 921 kg/m$^3$ and 933 kg/m$^3$. However, Liberica berries' true density is higher than Arabica and Robusta berries (1055 kg/m$^3$) (Chandrasekar & Viswanathan, 1999).

TABLE 5
Physical properties of Liberica green coffee beans and berries

| Properties | Berries | Std. dev. | Beans | Std. dev. |
|---|---|---|---|---|
| Density (kg/m$^3$) | | | | |
| - Bulk | 497.83 | 0.66 | 677.79 | 1.49 |
| - True | 1112.13 | 84.72 | 1138.25 | - |
| Coefficient of static friction | | | | |
| - Plywood | 0.41 | 0.03 | 0.34 | 0.01 |
| - Aluminium | 0.46 | 0.03 | 0.30 | 0.02 |
| - Jute fibre | 0.61 | 0.02 | 0.51 | 0.02 |
| Angle of repose (°) | | | | |
| - Filling (FAOR) | 21.86 | 3.28 | 22.49 | 0.40 |
| - Emptying (EAOR) | 21.52 | 4.44 | 19.67 | 0.45 |
| Fracture force (N) | 128.52 | 34.27 | 308.94 | 207.34 |
| Deformation (mm) | 4.62 | 0.62 | 1.16 | 0.40 |

From Table 5, Liberica beans' bulk density and true density are 677.79 kg/m$^3$ and 1138.25 kg/m$^3$, respectively. The beans' bulk density is lower than the Arabica and Robusta beans (1200-1300 kg/m$^3$) while the true density is higher following previous studies (600-650 kg/m$^3$) (Rodrigues *et al*., 2003; Dutra *et al.,* 2001). Data on bulk density of Liberica coffee berries and beans showed that Liberica coffee is packed in lesser amount which could be due to its bigger volume compared to Arabica and Robusta coffee while higher true density of Liberica coffee could be contributed by both its higher mass and volume.

*Coefficient of Friction*

Coefficient of friction is used in designing storage structures and also agricultural machinery (Mohsenin, 1978). Table 5 shows the coefficient of friction of Liberica berries and beans. From the table, Liberica beans have the highest friction on jute fibre, followed by plywood and aluminium. The surfaces' texture trend is rough (jute fibre), less rough (plywood) and smooth (aluminium). However, the trend is different for the berries where the order is jute fibre (rough), aluminium (smooth) and plywood (less rough). The reason could be the bean's silverskin is less smooth than the berries' skin. Hence, the friction between rough surfaces (plywood and bean's silverskin) is higher compared to friction between rough and smooth surfaces (plywood and berries' skin). Jute fibre offered the maximum friction for both berries and beans which could be due to the micro projections of the jute fibre which offered more resistance to the skin of berries and beans. Besides, rough surfaces have irregularities which make the real area of contact very small and almost independent of the apparent area of the surface (Mohsenin, 1978). Thus, high resistance materials such as jute fibre should be limited to ensure flowability along the tracks of coffee processing.

*Angle of Repose (AOR)*

Angle of repose can determine the materials' flow behaviour (Ileleji & Zhou, 2008). Therefore, AOR is also used in equipment design and storage structures (Mohsenin, 1978). According to Table 5, Liberica berries and beans filling angle of repose (FAOR) are 21.86° and 22.49°, while the emptying angle of repose are 21.52° and 19.67°, respectively. The FAOR is higher for both berries and beans. AOR will increase with increasing moisture content because the surface layer of moisture which surrounds each particle makes the surface tension effects become predominant in holding solids together (Mohsenin, 1978). The mean value between FAOR and EAOR of berries is 21.69° and this is 21.08° for the beans. The mean value of berries is slightly higher than the beans. This could be due to the sphere-like shape of the berries which make the berries easier to roll, thus, it contributes to an approximately higher value. Meanwhile, the Liberica beans tend to slide on top of each layer of beans making the mean values of FAOR and EAOR lower.

*Fracture Force and Deformation*

Fracture force is also important in equipment design such as the crusher. The fracture force will determine the maximum force which can be applied to a material. If the maximum force is exceeded, the crusher could damage the coffee beans rather than only the pulp. Table 5 presents the fracture force and deformation of Liberica berries and beans. The values of fracture force for Liberica berries and beans are 128.52 N and 308.94 N, respectively. Based on the result, Liberica berries have lower fracture force compared to the beans. This is because of the nature of the berries which are softer and easier to crush than the beans. According to Chandrasekar and Viswanathan (1999), fracture force values for Arabica and Robusta berries are lower than Liberica berries. The values are 30 N and 50 N, respectively. This shows that Liberica coffee berries need a special crusher for their processing. The value of Liberica berries fracture force

supports the fact from Brando (2009) where the Liberica berries' pulp adheres to the beans strongly, making it difficult to separate the berries' pulp and beans prior to sun drying. For this reason, aggressive pulpers which cut and bruised the pulp have been used in Malaysia rather than the normal pulpers which are used for Arabica and Robusta coffee. In a study by Pittia *et al*. (2007), the Liberica beans have higher fracture force (308.94 N) values compared to Arabica beans (150 N).

*Colour*

Colour can determine the materials' maturity. Hence, colour can be used in separating ripe berries from green berries. It also assists in the electronic sorting and grading processes (Mohsenin, 1978). This can be applied in sorting the good coffee beans from black or infested beans. Table 6 shows the colour of Liberica berries and beans. The $L^*$, $c^*$ and $h^*$ values for Liberica berries are 43.76, 49.86 and 39.34, respectively. Liberica beans' values of $L^*$, $c^*$ and $h^*$ are 39.94, 28.16 and 62.96 each. In comparison to Arabica and Robusta beans, Liberica beans have higher luminosity compared to Arabica (36) and slightly higher than Robusta (39). The chroma value of Liberica beans is higher than Arabica (17) but lower than Robusta (37). Furthermore, both Arabica and Robusta beans have higher values of hue angle which are 88 and 84 respectively which correspond to a yellowish colour (Mendonca *et al*., 2009).

TABLE 6
Colour values of Liberica coffee berries and bean

| Properties | Berries | Std. dev. | Beans | Std. dev. |
|---|---|---|---|---|
| $a^*$ | 37.47 | 4.77 | 12.80 | 0.50 |
| $b^*$ | 31.74 | 10.90 | 25.08 | 0.87 |
| $L^*$ - luminosity | 43.76 | 3.17 | 39.94 | 1.60 |
| $c^*$ - chroma | 49.86 | 7.01 | 28.16 | 0.94 |
| $h^*$ - hue angle | 39.34 | 11.45 | 62.96 | 0.69 |

Comparing Liberica berries and beans, the berries' luminosity (measurement of brightness) and chroma (colour intensity) values are higher than the beans'. Liberica berries' value for hue (element of the colour wheel) angle correspond to an orange shade while the beans correspond to a yellowish colour when referring to the $L^*a^*b^*$ colour system. Wrigley (1988) stated that Liberica berries have a colour ranging from yellow to a dark, reddish brown with spots or flecks of red. Thus, the colour of mature Liberica berries helps in distinguishing the mature from immature berries, while the green of the coffee beans aids in separating bad beans from the good ones.

## CONCLUSION

In this study, it was found that Liberica berries have a relatively bigger size, higher sphericity, volume, surface area, mass, true density, and fracture force value, but lower bulk density in comparison to Arabica and Robusta coffee. The order of Liberica berries' coefficient of friction

is jute fibre, aluminium and plywood. While Liberica beans have a bigger size, higher volume, mass, bean average density, fracture force, and luminosity value, but possess lower sphericity, bulk density, and hue angle in comparison to other coffee species. The order of Liberica beans' coefficient of friction is jute fibre, plywood, and aluminium. The physical properties of Liberica coffee berries and beans can be used to determine the quality of the coffee, and assist in machinery design, storage and handling in coffee-related industries.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding granted by Universiti Putra Malaysia (Project no. 05-01-10-0897RU) and the Research University Grant Scheme (RUGS) (Project no. 05-04-10-1009RU).

## REFERENCES

Afonso Junior, P. C., Correa, P. C., Pinto, F. A. C., & Queiroz, D. M. (2007). Aerodynamic properties of coffee cherries and beans. *Biosystems Engineering*, *98*, 39-46.

Afonso Jr., P. C., Correa, P. C., Pinto, F. A. C., & Sampaio, C. P. (2003). Shrinkage evaluation of five different varieties of coffee berries during the drying process. *Biosystems Engineering*, *86*(4), 481-485.

Azmil, I. A. R. (1991). Keperluan iklim dan tanah. In A. R. Azmil (Ed.), *Pengeluaran kopi* (p. 11). Malaysia: Malaysian Agricultural Research and Development Institute (MARDI).

Bart-Plange, A., & Baryeh, E. A. (2003). The physical properties of category B cocoa beans. *Journal of Food Engineering*, *60*, 219-227.

Brando, C. H. J. (2009). Harvesting and green coffee processing. In J.N. Wintgens (Ed.), *Coffee: Growing, Processing, Sustainable Production* (p. 634). Weinheim: Wiley-VCH Verlag GmbH & Co.

Campa, C., Ballester, J. F., Doulbeau, S., Dussert, S., Hamon, S., & Noirot, M. (2004). Trigonelline and sucrose diversity in wild *Coffea* species. *Food Chemistry*, *88*, 39-43.

Campa, C., Doulbeau, S., Dussert, S., Hamon, S., & Noirot, M. (2005). Qualitative relationship between caffeine and chlorogenic acid contents among wild *Coffea* species. *Food Chemistry*, *93*, 135-139.

Chandrasekar, V., & Viswanathan, R. (1999). Physical and thermal properties of coffee. *Journal of Agricultural Engineering Research*, *73*, 227-234.

Dutra, E. R., Oliveira, L. S., Franca, A. S., Ferraz, V. P., & Afonso, R. J. C. F. (2001). A preliminary study on the feasibility of using the composition of coffee roasting exhaust gas for the determination of the degree of roast. *Journal of Food Engineering*, *47*, 241-246.

Eke, C. N. U., Asoegwu, S. N., & Nwandikom, G. I. (2007). Some physical properties of jackbean seed (*Canavalia ensiformis*). *CIGR e-journal of Agricultural Engineering International* (Manuscript FP 07 014 vol. IX).

Franca, A. S., Mendonca, J. C. F., & Oliveira, S. D. (2005). Composition of green and roasted coffees of different cup qualities. *Lebensmittel-Wissenschaft und –Technologie*, *38*, 709-715.

Ghawas, M. M., & Rubiah, A. W. (1991). Botani. In A. R. Azmil (Ed.), *Pengeluaran Kopi* (p. 8). Malaysia: Malaysian Agricultural Research and Development Institute (MARDI).

Ghosh, B. N., & Gacanja, W. (1970). A Study of the Shape and Size of Wet Parchment Coffee Beans. *Journal Agricultural Engineering Resources*, *15*(2) 91-99.

Ileleji, K. E., & Zhou, B. (2008). The angle of repose of bulk corn stover particles. *Powder Technology*, *187*, 110-118.

Lang, W., & Sokhansaj. S. (1993). Bulk volume shrinkage during drying of wheat and canola. *Journal of Food Process Engineering*, *16*(4), 305-314.

Leite da Silveira, T. M., Tavares, E., & Gloria, M. B. A. (2007). Profile and levels of bioactive amines in instant coffee. *Journal of Food Composition and Analysis*, *20*, 451-457.

Malaysian Agricultural Research and Development Institute. (MARDI). 1986. *Status and prospects of the coffee industry in Malaysia* (p. 31-34). Malaysia.

McCabe, W. L., Smith, J. C., & Harriott, P. (1986). *Unit operations of chemical engineering*. New York: McGraw-Hill.

Mendonca, J. C. F., Franca, A. S., & Oliveira, L. S. (2009). Physical characterization of non-defective and defective Arabica and Robusta coffees before and after roasting. *Journal of Food Engineering*, *92*, 474-479.

Mohsenin, N. N. (1978). *Physical properties of plant and animal materials* (pp. 500-585). New York: Gordon and Breach Publishers.

Mohsenin, N. N. (1986). *Physical properties of plant and animal materials*. New York: Gordon and Breach Publishers.

Pittia, P., Nicoli, M. C., & Sacchetti, G. (2007). Effect of moisture and water activity on textural properties of raw and roasted coffee beans. *Journal of Texture Studies*, *38*, 116-134.

Quiroz, M. L. S., Rios, O. G., Barel, M., Guyot, B., Galindo, S. S., & Guiraud, J. P. (2004). Effect of chemical and environmental factors on Aspergillus ochraceus growth and toxigenesis in green coffee. *Food Microbiology*, *21*, 629–634.

Ramalakshmi, K., Kubra, I. R., & Rao, L. J. M. (2007). Physicochemical characteristics of green coffee: comparison of graded and defective beans. *Journal of Food Science*, *72*(5), 333-337.

Razavi, S. M. A., Amini, A. M., Rafe, A., & Emadzadeh, B. (2007). The physical properties of pistachio nut and its kernel as a function of moisture content and variety. Part III: Frictional properties. *Journal of Food Engineering*, *81*, 226–235.

Reh, C. T., Gerber, A., Prodolliet, J., & Vuataz, G. (2006). Water content determination in green coffee – Method comparison to study specificity and accuracy. *Food Chemistry*, *96*, 423–430.

Rodrigues, M. A. A., Borges, M. L. A., Franca, A. S., Oliveira, L. S., & Correa, P. C. (2003). Evaluation of physical properties of coffee during roasting. *Journal of Scientific Research and Development of the CIGR* 15th World Congress. (Manuscript FP 03 004. vol. V.) Chicago.

Sahin, S. & Sumnu, S. G. (2006). *Physical properties of foods*. (pp. 1, 3). New York: Springer Science+Business Media.

Vasconcelos, A. L. S., Franca, A. S., Gloria, M. B. A., & Mendonca, J. C. F. (2007). A comparative study of chemical attributes and levels of amines in defective green and roasted coffee beans. *Food Chemistry*, *101*, 26-32.

Wrigley, G., & Bridson, D. (1988). Botany. In G. Wrigley (Ed), *Coffee* (p. 74). New York: Longman Scientific & Technical.

Wintgens, J. N. (2009). The coffee plant. In J.N. Wintgens (Ed), *Coffee: Growing, Processing, Sustainable Production* (pp. 3-4). Weinheim: Wiley-VCH Verlag GmbH & Co.

Zainun, C. A. (1991). Pengendalian lepas tuai dan pemprosesan. In A.R. Azmil (Ed), *Pengeluaran kopi* (p. 46-53). Malaysia: Malaysian Agricultural Research and Development Institute (MARDI).

Pertanika J. Sci. & Technol. 22 (1): 81 - 93 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# The Performance of Robust Modification of Breusch-Godfrey Test in the Presence of Outliers

**Lim, H. A.[1,2]* and Midi, H.[1]**

[1]*Mathematics Department, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[2]*School of Applied Science and Foundation Studies, Kuala Lumpur Infrastructure University College, Jalan Ikram-Uniten, 43000 Kajang, Malaysia*

## ABSTRACT

Autocorrelation problem causes unduly effects on the variance of Ordinary Least Squares (OLS) estimates. Hence, it is very essential to detect the autocorrelation problem so that appropriate remedial measures can be taken. The Breusch-Godfrey (BG) test is the most popular and commonly used test for the detection of autocorrelation. Since this test is based on the OLS estimates, which are not robust, it is easily affected by outliers. In this paper, we propose a robust Breusch-Godfrey (MBG) test which is not easily affected by outliers. The results of the study indicate that the MBG test is more powerful than the BG test in the detection of autocorrelation problem.

*Keywords:* Autocorrelation, outliers, robust Breusch-Godfrey test

## INTRODUCTION

Many statisticians employ the Ordinary Least Squares (OLS) method to estimate the parameters of a linear model because of ease of computation. In many occasions, the assumptions of random and uncorrelated errors are taken for granted by statisticians without any rigorous check. These assumptions may not be true most of the time. The residuals may be correlated with the previous errors, which means that $E(u_i, u_j) \neq 0$ or $\text{cov}(u_i, u_j) \neq 0$ for $i \neq j$. Many statistics practitioners are not aware of the consequences of the autocorrelation problem. In specific, it ruins the important properties of OLS (Grassian & Boer, 1980; White & Brisbon, 1980). The OLS estimators are no longer the Best Linear Unbiased Estimators (BLUE) in the sense that the residual variance $\hat{\sigma}^2$ is likely to be underestimated, the true $\sigma^2$. Hence, less efficient estimates are obtained as a result of employing an incorrect model based on the erroneous assumption.

*Article history:*
Received: 30 May 2011
Accepted: 20 January 2012

*E-mail addresses:*
lim_ha@yahoo.com (Lim, H. A.),
habshahmidi@gmail.com (Midi, H.)
*Corresponding Author

ISSN: 0128-7680 

Additionally, the usual *t* and *F* tests of significance are no longer persuasive. These tests tend to be statistically significant when in fact they are not. The coefficient of determination, $R^2$, becomes inflated. As such, the estimator will look more accurate as compared to its actual value. All these problems contribute to the failure of the hypothesis testing. Hence, the autocorrelation problem will most likely give misleading conclusions about the statistical significance of the estimated regression coefficients (Gujarati & Porter, 2009). Therefore, it is very important to detect the presence of autocorrelation.

Many graphical methods have been developed and they are now available in the literature for detecting autocorrelation (Davidson & MacKinnon, 1998; Gujarati & Portter, 2009; Mirer, 1995; Murray, 2006). However, due to the fact that diagnostic plots can be very subjective, it is necessary to have some statistical methods to detect the problem of autocorrelation. Rigorous procedures for testing the autocorrelation of data have also been suggested in the literature (see Breusch, 1978; Durbin & Watson, 1951; Godfrey, 1978; Hosking, 1980; Hosking, 1981; Mirer, 1995; Murray, 2006). Most of these techniques are based on the OLS estimation.

The Breusch-Godfrey (BG) test is the most commonly used method to detect the presence of autocorrelation. It was developed by Breusch (1978) and Godfrey (1978). This test has many practical points than other existing tests of autocorrelation such as Durbin-Watson Test, Runs Test, and Portmanteau Test. First, it allows for nonstochastic regressors. Secondly, the regressors included in the regression may contain lagged values of the regressand Y, that is $Y_{t-1}$, $Y_{t-2}$, etc. These lagged values may also appear as explanatory variables in the model. Thirdly, it allows the lagged values of the regressand to follow higher-order autoregressive scheme such as AR(1), AR(2), etc. Other existing tests are not applicable in these circumstances (Breusch, 1978; Godfrey, 1978; Gujarati & Porter, 2009; Mirer, 1995; Murray, 2006).

Suppose

$$Y_t = X_t\beta + u_t, \tag{1}$$

if the error term $u_t$ follows the *p*th-order autoregressive, AR(*p*), scheme

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + .... + \rho_p u_{t-p} + \varepsilon_t \tag{2}$$

where $\varepsilon_t$ is a white noise error term that satisfies all the classical assumptions.

Then, the null hypothesis, $H_0$, to be tested is:

$$H_0 : \rho_1 = \rho_2 = ... = \rho_p = 0 \tag{3}$$

that is, there is no serial correlation between $u_t, u_{t-1}, ..., u_{t-p}$ of any order.

The procedures of the BG test are as follows:

Step 1: Estimate the coefficients of Eq. 1 by the OLS and obtain the estimated residual, $\hat{u}_t$.

Step 2: Regress $\hat{u}_t$ on the original $X_t$ and lagged values of the estimated residuals in Step (1). In summary, the following auxiliary regression is carried out:

$$\hat{u}_t = X_t \underset{\sim}{\alpha} + \hat{\rho}_1 \hat{u}_{t-1} + \hat{\rho}_2 \hat{u}_{t-2} + ... + \hat{\rho}_p \hat{u}_{t-p} + \varepsilon_t \quad (4)$$

where $\underset{\sim}{\alpha}$ is the regression coefficients of matrix $X$.

Step 3: Obtain $R^2$ from the above auxiliary regression. $R^2$ is given by:

$$R^2 = \frac{SSR}{SST}, \quad (5)$$

where SSR is the sum of the squared regression and SST is the sum of the squared total of the auxiliary regression.

When the sample size is large, the statistic $(n-p)R^2$ is asymptotically following the Chi-squared distribution with a degree of freedom of *p*, that is $(n-p)R^2 \sim \chi_p^2$. The null hypothesis is rejected if the statistic $(n-p)R^2$ exceeds the Chi-square value at the level of significant, which means at least one $\rho_i$ in Eq. 2 is statistically significantly different from zero.

In this article, a simple linear regression with autocorrelated errors are considered, as follows:

$$Y_t = \beta_1 + \beta_2 X_t + \mu_t, \quad (6)$$

and the error term is set to follow the first-order autoregressive AR(1) scheme,

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad -1 < \rho < 1. \quad (7)$$

The auxiliary regression to be examined is therefore simplified to:

$$\hat{u}_t = X_t \underset{\sim}{\alpha} + \hat{\rho} \hat{u}_{t-1} + \varepsilon_t \quad (8)$$

Since this test is based on the OLS estimates, it is suspected to be easily affected by the outliers. It is now evident that the outlier(s) have an unduly effect on the OLS estimates (Midi, 1999; Habshah *et al*., 2009; Rana *et al*., 2008; Riazosham *et al*., 2010).

In this paper, an attempt was made to robustify the Breusch-Godfrey test by incorporating the high efficient and high breakdown MM-estimator (Yohai, 1987) in the formulation of the new robust test for the identification of autocorrelation problem. We called this new test the Modified Breusch-Godfrey test (MBG). Real data and simulation experiments show that the proposed MBG outperforms the classical BG test in detecting autocorrelation in the presence of outliers.

## MATERIALS AND METHODS

We have briefly discussed the Breusch-Godfrey (BG) test for autocorrelation detection. The BG test uses the OLS to estimate the regression coefficient, so we expect it to suffer a huge setback when outliers are present in the data. Therefore, we propose a test which is robust against outliers. Here, we propose a new test which is a modification of Breusch-Godfrey test. We first identify the components of the BG test that are affected by the outliers and then

replace them with robust alternative. From the preceding procedures, we can see that the BG test requires two times of minimizing the sum of squares residuals to get the estimated coefficients. Firstly, we regress the original regression and then regress the auxiliary regression. Edgeworth (1887) has proven that squaring of the residual causes the least square to become extremely vulnerable to the presence of outliers. Therefore, the coefficients obtained are easily affected by the outlier. The MM-estimators introduced by Yohai (1987), which combined high-breakdown point and a high efficiency, are incorporated into the BG test. The robustified BG test is proven to minimize the impact of outliers on the regression model. This test is called the Modified Breusch-Godfrey test, or in short, MBG.

The proposed MBG test is summarized as follows:

Step 1: Unlike the classical BG test, we estimate the coefficients of the two variables regression by MM-estimator and get the residuals, $\hat{u}_t$.

Step 2: Regress $X_{,}$ on the original $X_t$ and $\hat{u}_{t-1}$ or run the auxiliary regression stated in equation (8) by the MM-estimator.

Step 3: Find $R^2$ from the auxiliary regression in Step 2. $R^2$ for MBG test is defined as:

$$R^2 = \frac{SSR}{(SSE + SSR)}, \tag{9}$$

where, SSE is the sum of squared errors and SSR is the sum of squared regression of the auxiliary regression.

The null hypothesis of no serial correlation between $\mu_t$ and $\mu_{t-1}$ will be rejected if the statistic $(n-1)R^2$ exceeds the Chi-square value at 0.05 significant level.

## RESULTS AND DISCUSSION

In this section, a few real world examples and a simulation study are presented to demonstrate the advantage of using the proposed Breusch-Godfrey test over the classical Breusch-Godfrey test in detecting serial autocorrelation problems.

### *Indexes of Real Compensation and Productivity Data*

The first example is the Indexes of Real Compensation and Productivity data by Gujarati and Porter (2009). The data set contains 46 observations that give the Index of Output (X) and the Index of Real Compensation per hour (Y) in U.S from 1960 to 2005. The data are shown in Table 1.

In this study, the performances of the classical BG test and the MBG test in the original data and contaminated data sets were examined. Three types of contaminated data sets were studied. The first type of the contaminated data is the data with one outlier in the *x* direction. An observation in X is replaced with an outlier; there will be a point that is in the far lower right corner. The second type of contaminated data is the data with one outlier in the *y* direction.

One observation in Y is replaced with an outlier; there will be a point that it is in the far upper left corner. The third type of the contaminated data is the data with a point that is in the far upper and far right corner, and the outlier is in both the x and y directions. For this case, a good observation is randomly replaced with an outlier. There are many definitions of outlier. In this study, outliers are considered as the values that lay outside the 3 deviation scopes from its mean. Fig.1 shows a scatter plot of the original data and the contaminated data.

Fig.2 shows the scatter plot of the current residuals (Res1) versus lagged residuals (Res(-1)) for the original data. From the plot, it is clearly seen that there is a positive serial correlation problem in the data.

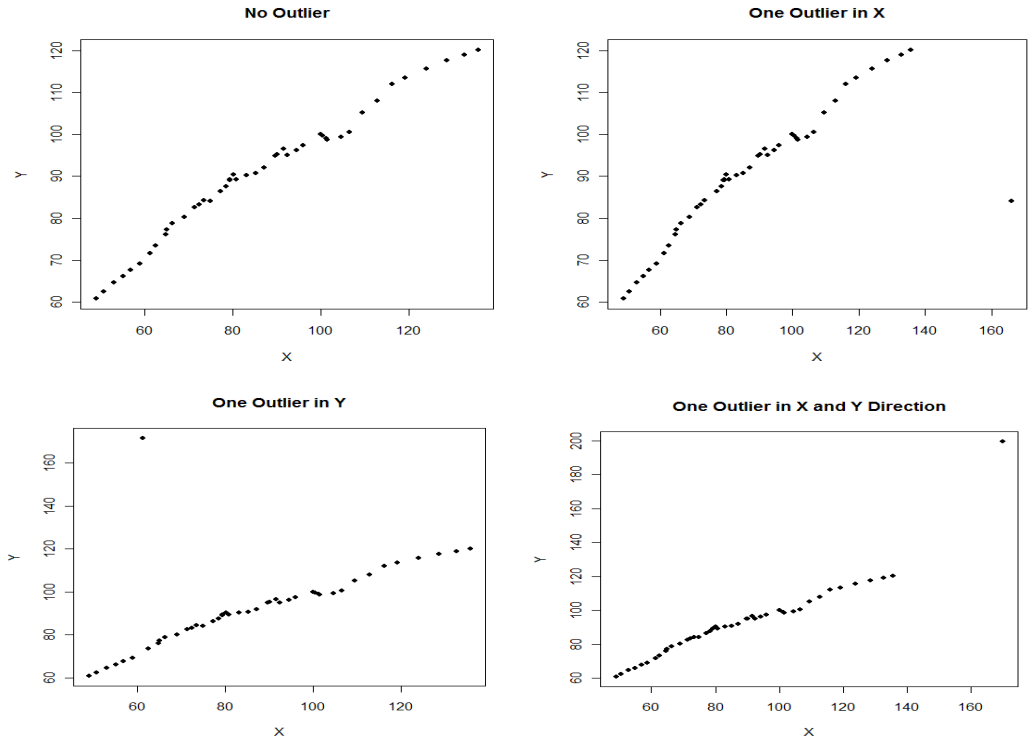



Fig.1: Scatter plot for the original and contaminated data

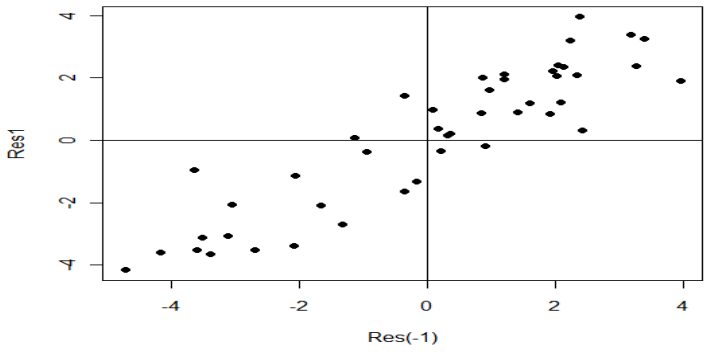



Fig.2: Current residuals (Res1) versus lagged residuals (Res(-1))

TABLE 1
Original and modified Real Compensation and Productivity Data, 1960-2005

| No | X | Y | No | X | Y |
|---|---|---|---|---|---|
| 1 | 48.9 | 60.8 | 24 | 83.0 | 90.3 |
| 2 | 50.6 | 62.5 | 25 | 85.2 | 90.7 |
| 3 | 52.9 | 64.6 | 26 | 87.1 | 92.0 |
| 4 | 55.0 | 66.1 | 27 | 89.7 | 94.95 |
| 5 | 6.8 | 67.7 | 28 | 90.1 | 95.2 |
| 6 | 58.8 | 69.1 | 29 | 91.5 | 96.5 |
| 7 | 61.2 | 71.7 [171.7] | 30 | 92.4 | 95.0 |
| 8 | 62.5 | 73.5 | 31 | 94.4 | 96.2 |
| 9 | 64.7 | 76.2 | 32 | 95.9 | 97.4 |
| 10 | 65.0 | 77.3 | 33 | 100.0 (170) | 100.0 (200) |
| 11 | 66.3 | 78.8 | 34 | 100.4 | 99.7 |
| 12 | 69.0 | 80.2 | 35 | 101.3 | 99.0 |
| 13 | 71.2 | 82.6 | 36 | 101.5 | 98.7 |
| 14 | 73.4 | 84.3 | 37 | 104.5 | 99.4 |
| 15 | 72.3 | 83.3 | 38 | 106.5 | 100.5 |
| 16 | 74.8 {166} | 84.1 | 39 | 109.5 | 105.2 |
| 17 | 77.1 | 86.4 | 40 | 112.8 | 108.0 |
| 18 | 78.5 | 87.6 | 41 | 116.1 | 112.0 |
| 19 | 79.3 | 89.1 | 42 | 119.1 | 113.5 |
| 20 | 79.3 | 89.3 | 43 | 124.0 | 115.7 |
| 21 | 79.2 | 89.1 | 44 | 128.7 | 117.7 |
| 22 | 80.8 | 89.3 | 45 | 132.7 | 119.0 |
| 23 | 80.1 | 90.4 | 46 | 135.7 | 120.2 |

Note: X = index of output
Y = index of real compensation per hour
{ }= outlier in X
[ ] = outlier in Y
( ) = outlier in X and Y direction

The performances of the BG and MBG tests are evaluated based on the *p*-values and the results are presented in Table 2.

TABLE 2
Autocorrelation diagnostics for Real Compensation and Productivity

| Test | No Outlier (*p*-value) | One Outlier in X (*p*-value) | One Outlier in Y (*p*-value) | One Outlier in X and Y Direction (*p*-value) |
|---|---|---|---|---|
| BG | 7.667e-10 | 5.664e-02 | 5.650e-01 | 2.590e-01 |
| MBG | 5.703e-10 | 1.268e-04 | 1.571e-04 | 1.363e-04 |

We observe from this table that the classical BG test is able to detect autocorrelation at 0.05 significance level if there is no outlier in the data. However, it fails to detect the problem of autocorrelation when the outlier occurs in the data set. We now observe the results of the MBG test on the original and modified Indexes of Real Compensation and Productivity data. Unlike the BG test, the MBG test can successfully detect the autocorrelation in the presence of an outlier yielding a highly significant *p*-value.

*Economic Report of the President 1982 Data*

Our next example is the economic report of the president data given by Mirer (1995). These data contain 25 observations that show the relationship between personal consumption expenditures (CON) and disposable personal income (DPI). We deliberately replace a good observation with an outlier into the data set in order to get the modified data in vertical direction, horizontal direction, as well as both vertical and horizontal directions. This data set, together with the contaminated data, is presented in Table 3.

TABLE 3
Original and modified Economic Report of the President 1982 data

| No | DPI(X) | CON(Y) | No | DPI(X) | CON(Y) |
|---|---|---|---|---|---|
| 1 | 446.1 | 405.4 | 14 | 722.5 | 657.9 |
| 2 | 455.5 | 413.8 | 15 | 751.6 | 672.1 |
| 3 | 460.7 | 418.0 | 16 | 779.2 | 696.8 |
| 4 | 479.7 | 440.4 | 17 | 810.3 | 737.1 |
| 5 | 489.7 | 452.0 | 18 | 865.3 | 768.5 |
| 6 | 503.8(1570.0) | 461.4(1461.4) | 19 | 858.4 | 763.6 |
| 7 | 524.9 | 482.0 | 20 | 875.8 | 780.2 |
| 8 | 542.3 | 500.5 | 21 | 907.4 | 823.7 |
| 9 | 580.8{1400.0} | 528.0 | 22 | 939.8 | 863.9 |
| 10 | 616.3 | 557.5[1557.5] | 23 | 981.5 | 904.8 |
| 11 | 646.8 | 585.7 | 24 | 1011.5 | 930.9 |
| 12 | 673.5 | 602.7 | 25 | 1018.4 | 935.1 |
| 13 | 701.3 | 634.4 | | | |

Note: { } = outlier in X
[ ] = outlier in Y
( ) = outlier in X and Y directions

Fig.3 shows the scatter plot of the original and modified economic reports of the president 1982 data, while Fig.4 illustrates the scatter plot of the current residuals (Res1) versus lagged residuals (Res(-1)) for the original data. Most of the residuals are bunched, in the first and the third quadrants, suggesting a positive correlation in the data.

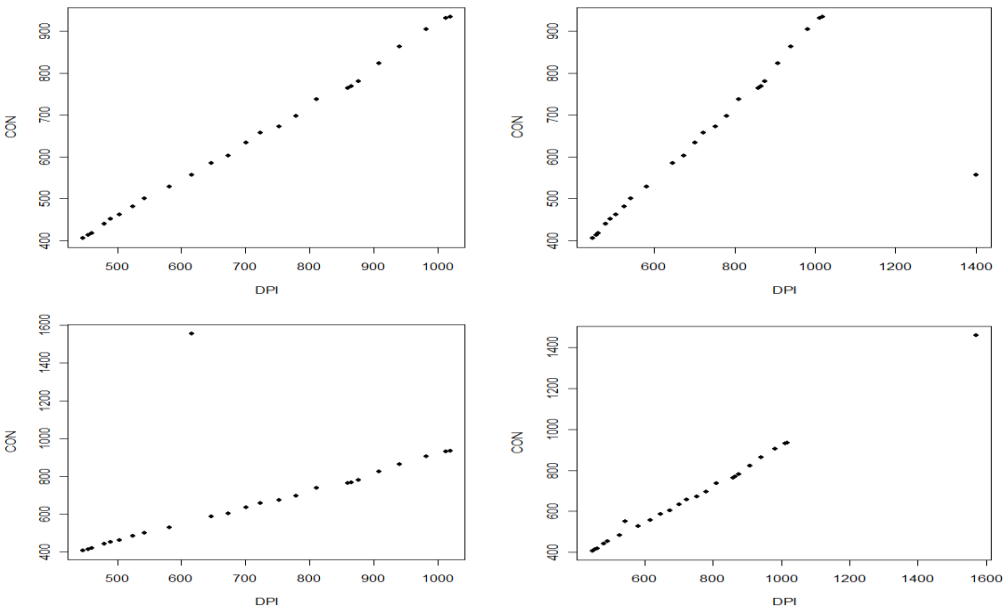



Fig.3: Scatter plots for the original and contaminated data for the Economic Report of the President data

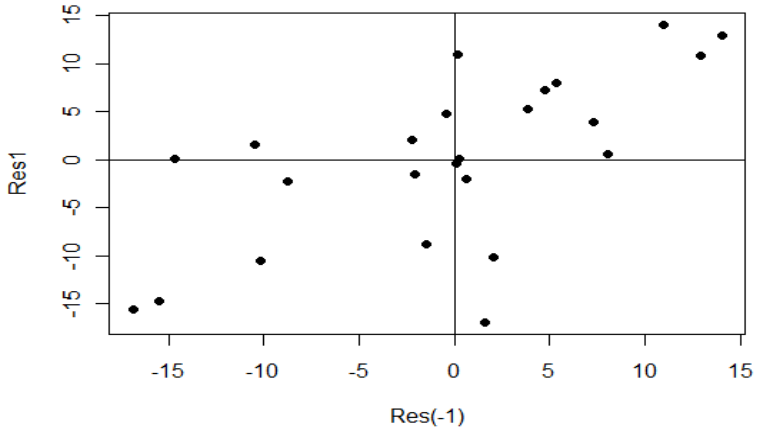



Fig.4: Current residuals (Res1) versus lagged residuals (Res(-1)) for the Economic Report of the President data

The results of the newly proposed MBG test and the classical BG test in detecting autocorrelation for the Economic Report of President data are presented in Table 4. Table 4 signifies that the classical BG test can only correctly identify the autocorrelation problem at 0.05 significance level, i.e. when the data are free from contamination although they give a false detection in the presence of outliers. The MBG test still successfully detects the presence of autocorrelation problem with and without the presence of outliers.

TABLE 4
Autocorrelation diagnostics for the real compensation and productivity data

| Test | No Outlier (*p*-value) | One Outlier in X (*p*-value) | One Outlier in Y (*p*-value) | One Outlier in X and Y Direction (*p*-value) |
|---|---|---|---|---|
| BG | 1.217e-03 | 6.284e-02 | 7.760e-01 | 8.752e-02 |
| MBG | 8.765e-04 | 1.495e-02 | 1.496e-02 | 3.711e-03 |

*Inventories and Sales in U.S. Manufacturing, 1950 – 1991 data*

For the last numerical example, we consider inventories and sales taken from Gujarati and Porter (2009). Once again, we randomly replace a good observation in the sales and inventories with the outliers and replace a coordinate paired with a contaminated pair in the sales and inventories direction. The original and contaminated data are shown in Table 5, and the scatter pot of each data set is shown in Fig.5. It can be seen by looking at the residual plot in Figure 6 that the data have positive autocorrelation problem.

TABLE 5
Original and contaminated Inventories and Sales data

| No | Sales(X) | Inventories(Y) | No | Sales(X) | Inventories(Y) |
|---|---|---|---|---|---|
| 1 | 46486 | 84646 | 22 | 224619 | 369374 |
| 2 | 50229 | 90560 | 23 | 236698 | 391212 |
| 3 | 53501 | 98145 | 24 | 242686 | 405073 |
| 4 | 52805 | 101599 | 25 | 239847 | 390950 |
| 5 | 55906 | 102567[802567] | 26 | 250394 | 382510 |
| 6 | 63027 | 108121 | 27 | 242002 | 378762 |
| 7 | 72931 | 124499 | 28 | 251708 | 379706 |
| 8 | 84790 | 157625 | 29 | 269843 | 399970 |
| 9 | 86589 | 159708 | 30 | 289973 | 424843 |
| 10 | 98797 | 174636 | 31 | 299766 | 430518 |
| 11 | 113201 | 188378 | 32 | 319558 | 443622 |
| 12 | 126905 | 211691 | 33 | 324984 | 449083 |
| 13 | 143936 | 242157 | 34 | 335991 | 463563 |
| 14 | 154391 | 265215 | 35 | 350715 | 481633 |
| 15 | 168129{579000} | 283413 | 36 | 330875 | 428108 |
| 16 | 163351(547551) | 311852(1900000) | 37 | 326227 | 423082 |
| 17 | 172547 | 312379 | 38 | 334616 | 408226 |
| 18 | 190682 | 339516 | 39 | 359081 | 439821 |
| 19 | 194538 | 334749 | 40 | 394615 | 479106 |
| 20 | 194657 | 322654 | 41 | 411663 | 509902 |
| 21 | 206326 | 338109 | | | |

Note: { } =outlier in X
[ ] =outlier in Y
( ) =outlier in X and Y directions

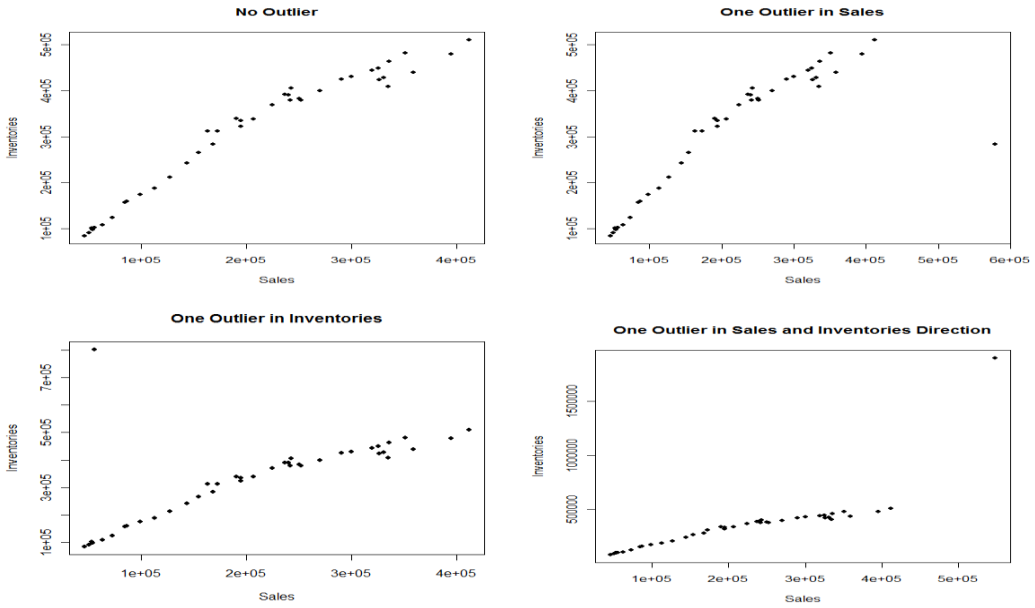



Fig.5: Scatter plot for the original and contaminated data for the Inventories and Sales data

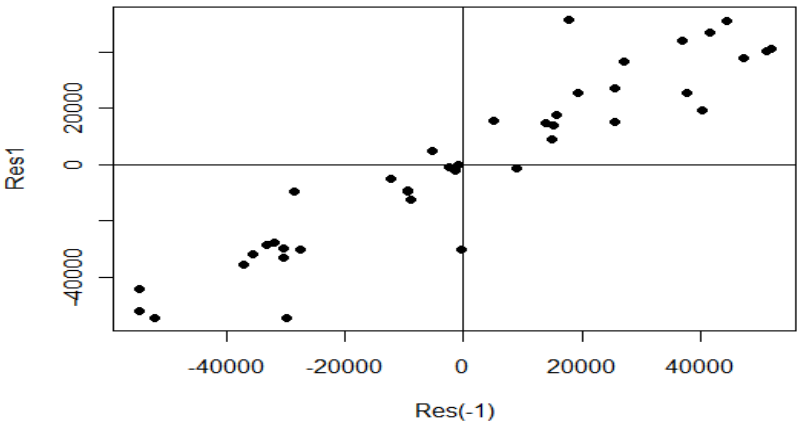



Fig.6: Current residuals (Res1) versus lagged residuals (Res(-1)) for the Inventories and Sales data

We employ the classical BG and MBG tests to the sales and inventories data. The test results are exhibited in Table 6. Similar results are obtained as in the previous examples. The power of detection of the classical BG test becomes poor when the outliers are present in the data. The MBG test is reliable in detecting the serial correlation irrespective of the presence of outliers at 0.05 significance level.

TABLE 6
Autocorrelation diagnostics for the real compensation and productivity data

| Test | No Outlier (*p*-value) | One Outlier in X (*p*-value) | One Outlier in Y (*p*-value) | One Outlier in X and Y Direction (*p*-value) |
|---|---|---|---|---|
| BG | 2.789e-09 | 5.043e-02 | 4.937e-01 | 5.047e-02 |
| MBG | 3.097e-09 | 3.996e-04 | 4.640e-04 | 2.831e-04 |

*Simulation Study*

We have seen the performance of the MBG test in the real world data. Now, we want to verify the results by checking a Monte Carlo simulation experiment. In this study, we considered three different samples sizes, $n$ = 20, 60 and 100, to represent the small, medium and large samples. For each sample, $n$ 'good" data are generated according to the following relation:

$$Y = 2 + 4X + u \tag{10}$$

where, all the values of X are generated from Uniform Distribution, U(0,10). The error term $u_t$ is generated by the first-order autoregressive scheme, as follows:

$$u_t = 0.9u_{t-1} + \varepsilon_t \tag{11}$$

with an initial value of $u_1$ equals to 2. The white noise, $\varepsilon_t$ is generated from the Normal distribution, with mean 0 and standard deviation 0.1. This autoregressive scheme is repeated for every 10 observations. Based on our experiences, the value of 0.9 is chosen in Eq. 11 to ensure the existence of a high autocorrelation problem.

We would like to compare the performance of the BG and MBG tests with 5% and 10% outlier in $x$, $y$ and both $x$ and $y$ directions. For each sample size, outliers are generated by deleting the 'good' observations and substituting them with 'bad' data points. The outliers in $x$ are represented by a uniform distributed variate $x_i$ from Uniform Distribution U(15,20), with $y_i$ being randomly selected Y values which are less than 15. Similarly, the outliers in the $y$ direction are represented by generating the $y_i$ variate from a Uniform Distribution U(50,60), with $x_i$ being randomly chosen X values which are less than 4.

Finally, the data sets with the outliers in both $x$ and $y$ directions are created by randomly replacing good observations with $x_i$ from U(15,20) and $y_i$ from U(50,60). In this study, we set the significance level to 0.05 and in each simulation run, there are 10,000 simulations.

Table 7 exhibits the classical BG and MBG tests. The classical BG test performs very poorly in the simulation. Throughout the simulation, the classical BG tests show inconsistency in detecting autocorrelation. In fact, the BG tests fail when there are outliers in the data set for all the three sample sizes. Nonetheless, the MBG test performs superbly throughout. This test is robust when the data are contaminated with the outliers. The MBG test also has higher power of detection with the increase of sample sizes. Thus, the MBG test outperforms the classical BG test in every respect of contamination.

TABLE 7
Simulation results of autocorrelation

| Sample sizes | Tests | No Outlier ($p$-value) | 5% of Outliers ($p$-value) | | | 10% of Outliers ($p$-value) | | |
|---|---|---|---|---|---|---|---|---|
| | | | X | Y | Both X and Y | X | Y | Both X and Y |
| n = 20 | BG | 1.643e-02 | 4.702e-01 | 6.635e-01 | 4.669e-01 | 4.362e-01 | 4.975e-01 | 4.491e-01 |
| | MBG | 4.214e-03 | 4.701e-02 | 4.399e-02 | 4.729e-02 | 3.584e-02 | 3.553e-02 | 3.758e-02 |
| n = 60 | BG | 5.906e-04 | 4.813e-01 | 5.947e-01 | 4.957e-01 | 4.740e-01 | 4.787e-01 | 4.781e-01 |
| | MBG | 5.099e-07 | 5.594e-05 | 5.181e-05 | 6.205e-05 | 6.211e-05 | 6.828e-05 | 7.495e-05 |
| n = 100 | BG | 1.759e-05 | 4.870e-01 | 5.290e-01 | 5.032e-01 | 4.808e-01 | 4.887e-01 | 4.815e-01 |
| | MBG | 6.958e-11 | 1.294e-07 | 1.331e-07 | 1.429e-07 | 2.163e-07 | 1.797e-07 | 1.697e-07 |

## CONCLUSION

In this research, the commonly used test for detecting autocorrelation has been shown to fail when outliers are present in any respect of the data. Hence, we formulate a simple but robust modification of the Breusch-Godfrey test to overcome the problem. Meanwhile, the comparison using the real data and Monte Carlo simulation experiments proved that the proposed Breusch-Godfrey test is consistent and reliable in offering substantial improvements over the classical Breusch-Godfrey test and also performs excellently in the detection of autocorrelation in the presence of outliers.

## REFERENCES

Breusch, T. S. (1978). Testing for autocorrelation in dynamic model. *Australian Economic Papers, 17*, 334-355.

Davidson, R. and MacKinnon, J. G. (1998). Graphical methods for investigating the size and power of test statistics. *The Manchester School, 66*, 1-26.

Durbin. J., & Watson G. S. (1951). Testing for serial correlation in least squares regression II. *Biometrika, 38*, 159-178.

Edgeworth, F. Y. (1887). On Observations Relating to Several Quantities. *Hermathena, 6*, 279-285.

Grassian, A., & Boer, E. S. (1980). Some methods of growth curve fitting. *Math. Scientist, 5*, 91-103.

Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica, 46*, 1303-1310.

Gujarati, D. N., & Porter, D. C. (2009). Autocorrelation: What Happens If the Error Terms Are Correlated? In M. A. Grove, M. Ito, H. Kim, P. V. Wunnava, & A. Paizis (Eds), *Basic Econometrics* (pp. 412-466). New York: McGraw-Hill.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P J., & Stahel, W. A. (1986). Robust Statistics: *The Approach Based on Influence Functions*. New York: John Wiley & Son, Inc.

Habshah, M., Norazan, M. R., & Imon, A. H. M. R. (2009). The performance of diagnostic-robust generalized potentials for the identification of multiple high leverage points in linear regression. *Journal of Applied Statistics, 36*(5), 62-99.

Hosking, J. R. M. (1980). The multivariate portmanteau statistics. *Journal of American Statistical Association, 75*, 602-608.

Hosking, J. R. M. (1981). Lagrange-multiplier tests of multivariate time series model. *Journal of the Royal Statistical Society, B43*, 261-262.

Midi, H. (1999). Preliminary estimators for robust non-linear regression estimation. *Journal of Applied Statistics, 26*(5), 591-600.

Mirer, T. W. (1995). Heteroscedasticity and Autocorrelation. *Economic Statistics and Econometrics*. In P. V. Wunnava, K. Lahiri, B. Bechdolt, & G. Chowdhury-Bose (Eds). United Sates of America: Prentice-Hall.

Murray, M. (2006). Autoregressive Disturbances. *Econometrics A Modern Introduction* (pp. 436-464). D. Clinton (Eds). United States of America: Pearson Addison-Wesley.

Rana, M. S., Midi, H., & Imon, A. A. H. M. R. (2008). A robust odification of the Goldfeld-Quandt test for the detection of heteroscedasticity in the presence of outliers. *Journal of Mathematics and Statistics, 4*(4), 277-283.

Riazoshams, H., Midi, H., & Sharipov, O. (2010). The performance of robust two-stage estimator in nonlinear regression with autocorrelated error. *Communications in Statistics. Simulation and Computation, 39*(6),1236-1253.

Yohai, V. (1987). High Breakdown-point and High Efficiency Estimates for Regression. *The Annals of Statistics, 15*, 642-665.

White, G. C., & Brisbon, I. L. (1980). Estimation and comparison of parameters in stochastic growth model for barn owls. *Growth, 44*, 97-111.

Pertanika J. Sci. & Technol. 22 (1): 95 - 103 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Gycaemic Control and Treatment Profile amongst 20646 Adult Type 2 Diabetes Mellitus: A Descriptive Report

**Chew, B. H.[1]*, Lee, P. Y.[1], Mastura, I.[2], Cheong, A. T.[1], Sri Wahyu, T.[3] and Zaiton, A.[1]**

*[1]Department of Family Medicine, Faculty of Medicine and Health Sciences, Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*
*[2]Klinik Kesihatan Seremban 2, Jalan S2 A2, Seremban 2, 70300 Seremban, Negeri Sembilan, Malaysia*
*[3]Klinik Kesihatan Simpang Kuala, 05400 Alor Setar, Kedah, Malaysia*

## ABSTRACT

An audit of Diabetes Control and Management-Diabetes Registry Malaysia (ADCM-DRM) was started to monitor the provision of diabetes care in the country. A total of 20,646 patients were registered in the registry until 31st December 2008. This report set out to determine the Type 2 diabetes controls and treatment profiles of these cohorts of patients. This was a registry-based observational study conducted from May to December, 2008. An online standard case record form was available for site data providers to register their diabetic patients aged 18 years old and above annually. Demographic data, diabetes duration, treatment modalities, as well as various risk factors and diabetes complications were reported. Data were analyzed using Data Analysis and Statistical Software (Stata) version 9. A total of 81 centres, 6 of which were hospitals, participated in this registry until 31st December 2008, contributing a total of 20646 patients. A majority of them (99.2%) had Type 2 diabetes mellitus. The mean HbA1c was 8.0% (SD 2.10), with 30.1% and 17.9% of the patients who attained HbA1c < 7% and HbA1c < 6.5%, respectively. Metformin was prescribed more than sulfonylurea while only 11% had insulin. A review of the diabetic care policy and strategies in the primary health care clinics is needed to implement a more effective treatment of diabetes in this country.

*Keywords:* Primary care, Registries, Glycaemic control, Hypoglycaemic agents

*Article history:*
Received: 1 July 2011
Accepted: 29 February 2012

*E-mail addresses:*
chewboonhow@yahoo.com (Chew, B. H.),
yein@upm.edu.my (Lee, P. Y.),
drmi68@yahoo.com (Mastura, I.),
cheaitheng@upm.edu.my (Cheong, A. T.),
sriwahyu2006@yahoo.com.my (Sri Wahyu, T.)
*Corresponding Author

## INTRODUCTION

The number of people with diabetes is expected to increase alarmingly in the coming decades. In 1985, an estimated 30 million people worldwide had diabetes; in 2000, a little over a decade later, the figure had risen to over 150 million. By 2030, the figure is

ISSN: 0128-7680 

expected to rise to 439 million (International Diabetes Federation, 2009). The prevalence of diabetes mellitus in Malaysian adults aged 30 years old and above has doubled over the ten year period from 8.3% in 1996 to 14.9% in 2006 (National Health Morbidity Survey,1996; Institute of Public Health, 2008). According to the International Diabetes Federation estimates for 2010 and 2030, Malaysia is among the top ten countries with the highest prevalence of diabetes mellitus (International Diabetes Federation, 2009).

Diabetes mellitus is one of the commonest chronic, non-communicable diseases which contribute to a high level of morbidity and mortality globally with its impacts coming not only from diabetes mellitus itself, but mainly arising from its complications such as coronary artery disease and chronic kidney disease (Garcia *et al*., 1974; National Health Morbidity Survey, 1996; Lim *et al*. 2009). Studies have shown that with good glycaemic control and control of other concomitant cardiovascular risk factors, there would be delay in complications (The Diabetes Control and Complications Trial Research Group, 1993; UKPDS, 1998; Berl *et al*., 2003; Sever *et al.*, 2005; Patel *et al.*, 2007; Action to Control Cardiovascular Risk in Diabetes Study Group, 2008; ADVANCE Collaborative Group, 2008). Findings from UKPDS showed that for every 1% reduction in HbA1c, there was 37% reduction in micro-vascular complications, about 40% reduction of all diabetic-related end-points, myocardial infarction and death in the metformin sub-group analysis (UKPDS, 1998). This protective effect of early intensive glycaemic control persisted into the following decade post study (Holman, 2008). However, the achievement of the control to the recommended target level is not satisfactory. Clinical audit is one of the measures which health care providers use to assess and monitor their care to the patients. By implementing changes to the shortfalls found, health care providers hope to improve on the quality of care to the patients and subsequently delay the complications.

Clinical audit is one of the measures which health care providers may use to assess and monitor their care to the patients (Costa, 2009). An online diabetes registry database called the "Audit of Diabetes Control and Management-Diabetes Registry Malaysia (ADCM-DRM)" was started in July 2008 as a pilot project in Negeri Sembilan (NS) (ADCM manual; CRC, Kuala Lumpur). The objectives of this online registry was to gather and monitor the provision of diabetes care; hence, to better inform outcomes of treatment, budget planning, health education for both the physician and patients as well as to increase awareness of the potential serious impact of this disease on the country (Stamler *et al*., 1993; American Diabetes Association, 1998; American Diabetes Association, 2001). This registry was managed by a secretariat based at Clinical Research Centre, Hospital Kuala Lumpur. The main source data providers (SDP) were those from government health clinics and hospitals throughout NS, Selangor and Perak. This report set out to determine the Type 2 diabetes controls and treatment profiles as up to 31$^{st}$ December 2008.

## MATERIALS AND METHODS

This study was approved by the Medical Research Ethics Committee (MREC), Ministry of Health, Malaysia. ADCM-DRM database from July to 31$^{st}$ December 2008 was cleaned before the analyses were carried out. All public hospitals and health clinics were invited to participate as SDP. Participation was voluntary. An online standard case record form (CRF) was available

in the ADCM website for SDP to register their diabetic patients aged 18 years old and above annually. This was done by trained physicians and paramedics. Patients with T2D were defined as when their case record fulfilled all these criteria: (1) either documented diagnosis of diabetes mellitus according to World Health Organisation (WHO) criteria, or (2) those whose current treatment consisted of life-style modification, on oral anti-diabetics or insulin. The data captured include demography, diabetes duration, treatment modalities, as well as various risk factors and diabetes complications. Each of the selected SDP site administrators was given a password guarded access to the website in order to generate local data that was analysed by the secretariat. Data were analysed using STATA version 9. The detailed methodology of this project has been described elsewhere (Mastura *et al*., 2008).

## RESULTS

A total of 81 centres, 6 of which were hospitals, participated in this registry until 31st December 2008, contributing a total of 20,646 patients. A majority of the patients (89.6%) were from Negeri Sembilan and most of the cases were Type 2 diabetes mellitus (T2D) (99.2%). In more specific, 57.2% were females, while Malays made up 57.1%, Chinese were 19.7%, Indians formed 22.6% and other races made up 0.2% (see Table 1). The mean age was 58.0 years (SD 11.49) with 77.6% of them aged 50 years old and above. About 82% of the patients had their body mass index (BMI) in the overweight category (BMI $\geq$ 23 $Kgm^{-2}$), with 42.3% of them who were actually at least in the Obese class I (BMI $\geq$ 27.5 $Kgm^{-2}$). The mean age at diagnosis was 54.8 years old (SD 11.47) and the mean duration of diabetes was 4.7 years (SD 4.20) (see Fig.1).

TABLE 1
Demographic profile

| **Profile** | | **n (%)** |
|---|---|---|
| Gender | | |
| | Male | 8687 (42.4) |
| | Female | 11722 (57.2) |
| | Missing | 72 (0.4) |
| Ethnicity | | |
| | Malay | 11694 (57.1) |
| | Chinese | 4026 (19.7) |
| | Indian | 4637 (22.6) |
| | Other Malaysian | 48 (0.2) |
| | Non – Malaysian | 7 (0.03) |
| | Missing | 69 (0.3) |
| Age group (years) | | |
| | < 30 | 178 (0.9) |
| | 30-49 | 4423 (21.6) |
| | 50-69 | 12424 (60.7) |
| | ≤ 70 | 3456 (16.9) |

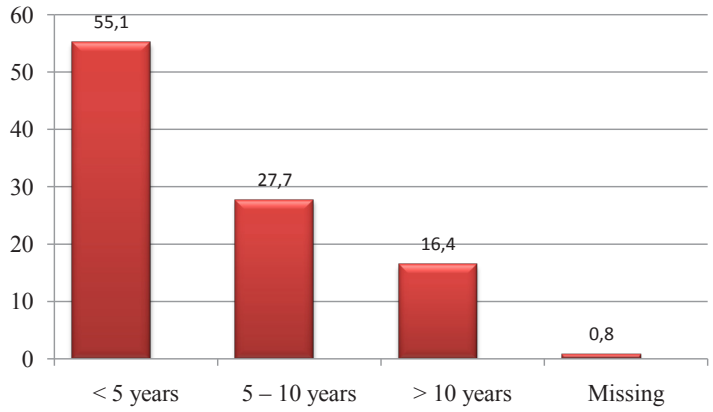



Fig.1: The percentage of patients according to duration of diabetes

The mean casual blood glucose (CBG) and fasting blood glucose (FBG) were 11.1 mmol/L (SD 4.47) and 8.4 mmol/L (SD 3.26), respectively. The mean HbA1c was 8.0% (SD 2.10) out of about two thirds of tests carried out (Table 2). From a total of 13310 HbA1c test results available for analysis, 4012 (30.1%) of the patients attained HbA1c < 7% and 2386 (17.9%) achieved HbA1c < 6.5%.

Table 3 shows the treatment modalities of the patients. There were 326 (1.6%) patients who were prescribed diet control only. Most of the patients (42.7%) were on two and more oral anti-diabetic agents (OAD) as compared to monotherapy of OAD (34.9%). The most common OAD was either Metformin (75.9%) or a sulfonylurea (60.8%). Insulin was prescribed in a total of 2254 (11.0%) patients with about two thirds of them were prescribed in combination with OAD agents (see Table 3). Amongst the patients on insulin therapy, intermediate acting insulin was the most commonly used insulin (53%), followed by pre-mixed insulin (38.9%) and the least used was long-acting insulin (0.01%).

TABLE 2
Glycaemic monitoring and profiles from July-December 2008

| Measurement | n, (% of the total patients) | Mean (SD) |
|---|---|---|
| Casual blood glucose (mmol/L) | 11200 (54.7) | 11.1 (4.47) |
| Fasting blood glucose (mmol/L) | 14214 (69.4) | 8.4 (3.26) |
| 2 hrs Post-prandial (mmol/L) | 5825 (28.4) | 13.3 (4.58) |
| HbA1c (%) | 13764 (67.2) | 8.0 (2.10) |

TABLE 3
Treatment Modalities

| Treatment Modalities | n (%) |
|---|---|
| Diet only | 326 (1.6) |
| Oral anti-diabetic agent only | |
| - Monotherapy (OAD) | 7142 (34.9) |
| - ≥2 OAD | 8753 (42.7) |
| OAD & Insulin | 1415 (6.9) |
| Insulin only | 839 (4.1) |

## DISCUSSION

There was a huge majority of T2D (99.2%) in the registry as it was expected from the mainly primary care set-ups of the site data providers. Hence, Type 1 diabetes mellitus patients were rarely seen at these health clinics, as few as an average of less than 2 patients per centre. Negeri Sembilan contributed the most number of patients because this project was initiated in this state from the beginning. The gender composition was typical of attendance to any of the Malaysian health care centres and services, whereby female attendance (57.2%) being more common than the male. However, the difference was not as large as usually seen, most probably owing to the fact that a large majority of patients were retired individuals (77.6% being older than 50 years old). The ethnicity composition did not reflect the national representation as evidenced by the Indians (22.6%) being the second most common users of these health facilities instead of the Chinese (19.7%). This phenomenon is not unexpected since the Indians are the ethnic group mostly affected by diabetes and the Chinese could be getting their health care needs from the private health care sectors (National Health Morbidity Survey, 1996; Institute of Public Health, 2008). This registry confirmed the belief that most of the T2D patients seen at the primary care level were at the early stage of the disease as there were more than half who were diagnosed less than 5 years ago.

FBG was the most commonly carried out test for the patients and this was closely followed by HbA1c; about two thirds (67.2%) of the patients managed to have HbA1c tested in the past one year. This was comparable to another study in 19 public hospitals across Peninsular Malaysia which reported annual testing rates of 67.9% for HbA1c (Mafauzy, 2006). The mean HbA1c of 8.0% in this survey was better when compared to data of 8.6% from tertiary centres and hospitals (Ismail *et al*., 2000). However, the figure is higher than in another study conducted in a government primary health centre in East Malaysia, where the average HbA1c was 7.4% (Wong & Rahimah 2004). This survey found that 30.1% of the patients attained HbA1c < 7% and 17.9% achieved HbA1c < 6.5%. The control profile was rather similar to both the local and United States academic primary care centres (Eid *et al*., 2004; Chew *et al*., 2011; McFarlane *et al*., 2002). Studies in public hospitals (Mafauzy, 2006) and another study in an urban primary health care centre (Wong & Rahimah, 2004) have shown better results. The reason for the difference could be due to the level of care. Patients from hospital and urban health clinics are mainly managed by specialist or senior medical officers whereas our survey included rural districts where the patients may be managed by medical assistants or

junior medical officers. Lack of awareness and adherence to clinical guidelines, heavy patient load and restricted access to "List A" drugs like premixed insulin and safer sulphonylurea, glicazide may be the other factors for poorer results in the public primary healthcare clinics in our survey (Tan *et al*., 2008).

We observed that a few patients were on diet control alone as their treatment. This was not surprising among the present study cohorts of patients in the early stage of diabetes mellitus who might be asymptomatic and had HbA1c < 6.5% or FBG < 6.0 mmol/L in accordance to the recommendation of the national clinical practice guidelines (Malaysian Clinical Practice Guidelines for the Management of Type 2 Diabetic Mellitus, 2009). This registry reported a positive finding that more metformin than sulfonylureas was prescribed now compared to earlier studies in the country (Sarojini *et al*., 2008) and Thailand (Kosachunhanun *et al*. 2006) that showed the reverse scenario. The first line of anti-diabetic with metformin was appropriate in this youngish cohort and contra-indication of metformin is rather rare in primary healthcare clinical practice. This similar observation was also noted in the United Kingdom general practice since 2002, when metformin use surpassed sulfonylureas (Filion *et al*., 2009). However, the insulin use, both alone and in combination with other OAD agents, was still very low (11%) and this may explain rather poor diabetic control rate amongst the T2D patients in the community. The National Health Morbidity Survey in 2006 showed only 7.2% insulin usage, both alone and in combination with oral anti-diabetics. Under-utilization of insulin is also found in other studies in primary care clinic and hospital (Wong & Rahimah, 2004; Tan *et al*., 2008). This could be due to the resistance to use insulin by the healthcare provider and the low acceptance of insulin therapy by patients due to misconception of insulin risk and interference of routine life-style (Nathan, 2002; Cefalu, 2002; Karter *et al*., 2010).

The limitations of this registry are the retrospective data retrieval from patients' records by mainly non-treating parties leading to incompleteness and inaccuracy. As the registry was not compulsory, many health centres did not participate, leading to biased patient population from only the participating centres. Thus, the data presented here could not be generalized to the whole primary care in Malaysia as a major proportion of the patients were from Negeri Sembilan.

## CONCLUSION

ADCM-DRM is gaining momentum since its inception as evidenced by the ever increasing number of site participation and patients registered into the registry. The control profile was rather fair and showed no improvement as compared to earlier glycaemic control profile in the country many years ago. There was under-utilization of insulin which could have caused poor glycaemic control in the majority and this insulin acceptance and adherence can be improved with better patient education and self-management training. A review of the diabetic care policy and strategies in the primary health care clinics is needed to implement a more effective treatment of diabetes.

## ACKNOWLEDGEMENTS

We would like to acknowledge the Director General of Health for their support in our effort in the registry and permission to publish this article using the data from public health clinics. We would like to express our sincere gratefulness to Ms Lena Yeap, Noor Akma Hassim and Tee Chin Kim from CRC, Kuala Lumpur.

## REFERENCES

Action to Control Cardiovascular Risk in Diabetes Study Group, Gerstein, H. C., Miller, M. E., Byington, R. P., & Goff, D. C. (2008). Effects of intensive glucose lowering in type 2 diabetes. *The New England Journal of Medicine*, *358,* 2545-2559.

Association of Clinical Registries, Malaysia, ACRM. (n.d.). *ADCM manual (version 2.0)* [online]. Retrieved on August 10, 2008, from https://app.acrm.org.my/ADCM/Manual.aspx.

ADVANCE Collaborative Group, Patel, A., MacMahon, S., Chalmers, J., & Neal, B. (2008). Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *The New England Journal of Medicine*, *358*, 2560-2572.

American Diabetes Association. (1998). Economic consequences of diabetes mellitus in the U.S. in 1997. *Diabetes Care*, *21*(2), 296–309.

American Diabetes Association. Diabetes. (2001). *2001 Vital Statistics*. Alexandria, VA, ADA.

Berl, T., Hunsicker, L. G., Lewis, J. B., Pfeffer, M. A., & Porush, J. G. (2003). Cardovascular outcomes in the Irbesartan Diabetic Nephropathy Trial of patients with type 2 diabetes and overt nephropathy. *Annals of Internal Medicine*, *138*, 542-549.

Cefalu, W. T. (2002). Evaluation of alternative strategies for optimizing glycemia: progress to date. *American Journal Medicine*, *113*(Suppl. 6A), 23S–35S.

Chew, B. H., Khoo, E. M., & Chia, Y. C. (2011). Does religious affiliation influence glycaemic control in primary care patients with type 2 diabetes mellitus? *Mental Health in Family Medicine*, *8*(1), 21-28.

Costa, B. M., Fitzgerald, K. J., Jones, K. M., & Dunning, A. T. (2009). Effectiveness of IT-based diabetes management interventions: a review of the literature. *BMC Family Practice*, *10*, 72.

CRC Kuala Lumpur. (n.d.). *About DRM-ADCM*. Retrieved November 16, 2009 from http://www.acrm.org.my/adcm/default.asp?page=/adcm/about.

Eid M., Mafauzy M., & Faridah A. R. (2004). Non-achievement of clinical targets in patients with type 2 diabetes mellitus. *Medical Journal of Malaysia*, *59*(2), 177-184.

Filion, K. B., Joseph, L., Boivin, J. F., Suissa, S., & Brophy, J. M. (2009). Trends in the prescription of anti-diabetic medications in the United Kingdom: a population-based analysis. *Pharmacoepidemiology and Drug Safety*, *18*(10), 973-976.

Garcia, M. J., McNamara, P. M., Gordon, T., & Kannel, W. B. (1974). Morbidity and mortality in diabetics in the Framingham population: Sixteen year follow-up study. *Diabetes*, *23*,105-111.

Holman, R. R., Paul, S. K., Bethel, M. A., Matthews, D. R., & Neil, H. A. (2008). 10-year follow-up of intensive glucose control in type 2 diabetes. *The New England Journal of Medicine, 359*(15), 1577-1589.

Institute of Public Health. (2008). *The Third National Health Morbidity Survey (NHMS III) 2006, Volume 2*. Ministry of Health, Malaysia.

International Diabetes Federation. (2009). *Diabetes Atlas* (4th Edition).

Ismail, I. S., Wan Nazaimoon, W. M., Wan Mohamad, W. B., Letchuman, R., Singaraveloo, M., Pendek, R., Faridah, I., Rasat, R., Sheriff, I. H., & Khalid, B. A. K. (2000). Socioedemographic determinants of glycaemic control in young diabetic patients in peninsular Malaysia. *Diabetes Research and Clinical Practice, 47*, 57-69.

Karter, A. J., Subramanian, U., Saha, C., Crosson, J. C., Parker, M. M., Swain, B. E., Moffet, H. H., & Marrero, D. G. (2010). Barriers to Insulin Initiation: The Translating Research into Action for Diabetes Insulin Starts Project. *Diabetes Care*, *33*(4), 733-735.

Kosachunhanun, N., Benjasuratwong, Y., Mongkolsomlit, S., Rawdaree, P., & Plengvidhya, N. (2006). Thailand Diabetes Registry Project: Glycemic Control in Thai Type 2 Diabetes and Its Relation to Hypoglycemic Agent Usage. *Journal of Medical Association of Thailand*, *89*(Suppl. 1), S66-71.

Lim, Y. N., Lim, T. O., & Lee, D. G. (2009). *16th Report of the Malaysian Dialysis and Transplant Registry 2008*. The National Renal Registry, Malaysian Society of Nephrology, Kuala Lumpur.

Mafauzy, M. (2006). Diabetes Control and Complications in Public Hospitals in Malaysia. *Medical Journal of Malaysia*, *61*(4), 477-483.

Mastura, I., Zanariah, H., Fatanah, I., M Feisul, I., Wan Shaariah, M. Y., Jamaiyah, H., & Geeta, A. (2008). An Audit of Diabetes Control and Management (ADCM). *Medical Journal of Malaysia*, *63*(suppl. C), 76-77.

Ministry of Health, Malaysia. (2009). *Malaysian Clinical Practice Guidelines for the Management of Type 2 Diabetic Mellitus* (4th Edn.).

McFarlane, S. I., Jacober, S. J., Winer, N., Kaur, J., Castro, J. P., Wui, M. L. A., Gliwa, A., Gizycki, H. V., & Sowers, J. R. (2002). Control of Cardiovascular Risk Factors in Patients with Diabetes and Hypertension at Urban Academic Medical Centers. *Diabetes Care*, *25*(4), 718-723.

Nathan, D. M. (2002). Initial Management of Glycemia in Type 2 Diabetes Mellitus. *New England Journal of Medicine*, *347*, 1342-1349.

National Health Morbidity Survey 1996. (1996). *Volume 9. Institute of Public Health*. Ministry of Health, Malaysia.

Patel, A., ADVANCE Collaborative Group, MacMahon, S., Chalmers, J., & Neal, B. (2007). Effects of a fixed combination of perindopril and indapamide on macrovasular and microvascular outcomes in patients with type 2diabetes mellitus (the ADVANCE trial): a randomized controlled trial. *Lancet*, *370*, 829-840.

Sarojini, S., Faridah, A., Lim, C. M., Sameerah, S. A., Lim, T.O., Lai, L. S., Chew, S. M. (2008). Use of prescription medicines in Malaysia 2005. *Medical Journal of Malaysia*, *63*(3), 203-206.

Sever, P. S., Poulter, N. R., Dahlof, B., Wedel, H., & Collins, R. (2005). Reduction in cardiovascular events with atorvastatin in 2,532 patients with type 2 diabetes:Anglo-Scandinavian Cardiac Outcomes Trial-Lipid-Lowering Arm (ASCOT-LLA). *Diabetes Care*, *28*, 1151-1157.

Stamler, J., Vaccaro, O., Neaton, J. D., & Wentworth, D. (1993). Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes Care*, *16*, 434-444.

Tan, F., Chan,G., Wong, J. S., & Rozario, F. (2008). Standard of care for type 2 diabetic patients in a public hospital general medical clinic: report of a self-audit. *Medical Journal of Malaysia*, *63*(3), 224-228.

The Diabetes Control & Complications Trial Research Group. (1993). The Effect of Intensive Treatment of Diabetes on The Development and Progression of Long –term Complications in insulin-dependent Diabetes Mellitus. *The New England Journal of Medicine*, *329*, 977-986.

UK Prospective Diabetes Study (UKPDS) Group. (1998). Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet*, *352*, 837-854.

UK Prospective Diabetes Study (UKPDS) Group. (1998). Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). *Lancet*, *352*(9131), 854-865.

Wong, J. S., & Rahimah, N. (2004). Glycaemic Control of Diabetic Patients in an Urban Primary Health Care Setting in Sarawak: The Tanah Puteh Health Centre Experience. *Medical Journal of Malaysia*, *59*(3), 411-417.

Pertanika J. Sci. & Technol. 22 (1): 105 - 112 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Development of Internet-Based Instrumentation for the Study of Hall Effect

**Ariffin Abas[1]*, Abdul Halim Shaari[1,2], Zainal Abidin Talib[1] and Zaidan Abdul Wahab[1]**

*[1]Department of Physics, Faculty of Science, Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*
*[2]Institute of Mathematical Research (INSPEM), Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

The computer, together with Lab View software, can be used as an automatic data acquisition system. This project deals with the development of a computer interfacing technique for the study of Hall Effect and converting the existing automation system into a Web-based automation system. The drive board RS 217-3611 with PCI 6025E card and stepper motor RS191-8340 with a resolution of 0.1mm, was used to move a pair of permanent magnets backward and forward against the sample. The General Interface Bus (GPIB) card interfaces, together with digital nano voltmeter and Tesla meter using serial port RS232 interface, are used for measuring the potential difference and magnetic field strength respectively. Hall Effect measurement on copper (Cu) and tantalum (Ta) showed negative and positive sign Hall coefficient. Therefore, the system has electron and hole charge carriers respectively at room temperature. The parameters such as drift velocity, conductivity, mobility, Hall Coefficient and charge carrier concentration were also automatically displayed on the front panel of Lab View programming and compared with standard value. The Web-based automation system can be remotely controlled and monitored by users in remote locations using only their web browsers. In addition, video conferencing through Net Meeting has been used to provide audio and video feedback to the client.

*Keywords:* Interface, Hall Effect, Web-based automation system

## INTRODUCTION

The Hall effect $V_H$ is observed when a magnetic field (B) is applied at right angles to a sample of material carrying an electric current (I). Hall Effect (or Hall voltage) that appears across the sample is due to an electric

*Article history:*
Received: 1 July 2011
Accepted: 1 December 2012

*E-mail addresses:*
ariffa70@yahoo.com (Ariffin Abas),
ahalim@upm.edu.my (Abdul Halim Shaari),
zainalat@upm.edu.my (Zainal Abidin Talib),
zaidan@upm.edu.my (Zaidan Abdul Wahab)
*Corresponding Author

ISSN: 0128-7680 

field which is at right angles to both the current and the applied magnetic field (Edward, 2006). Hall voltage can then be expressed as:

$$V_H = \frac{BI}{nqd} = \frac{R_H BI}{d} \tag{1}$$

where d= thickness of the sample, q= charge and n = charge carrier density

$$\text{The quantity } R_H = \frac{1}{nq} \text{ is called Hall Coefficient} \tag{2}$$

The Hall voltage is directly related to the magnetic field and the drift current, and it is inversely related to the thickness of the sample. The samples used for the measurement are in the form of thin strips. A plot of Hall voltage ($V_H$) as a function of magnetic field (B) at constant current will have a slope given by the following equation:

$$Slope = \frac{IR_H}{d} \tag{3}$$

From equation (3), the slope of dependency $V_H$ vs. B is defined, in particular, by the Hall coefficient $R_H$. This $R_H$ carries information about majority electron or holes concentration. The conductivity σ is given as:

$$\sigma = \frac{IL}{VA} \tag{4}$$

where, L= Length and A= Cross sectional of the sample.

The mobility μ, is given as:

$$\mu = \sigma R_H \tag{5}$$

and also the drift velocity $v_x$ of the sample is given as:

$$v_x = \frac{I}{nqA} \tag{6}$$

## INSTRUMENTATION AND SOFTWARE

The Hirst Tesla meter model GM08 was used to calibrate the magnetic field and connected to a personal computer (PC) via Serial Port (Kai Qian, 2009). A current supply of 1.0A through the sample was provided by Keithley current source model 224. Keithley nano voltmeter model 2182 was connected to a PC via General Purpose interface Bus (Sumathi, 2007) to measure the Hall Effect potential. The drive board RS217-3611 (RS Component, 2001) with PCI 6025E (Basil, 2011) card through opto isolator circuit and stepper motor (Takashi, 1985; Kalman, 1995) with a resolution of 0.1mm were used to move a pair of permanent magnets, Neodymium-Iron-

Boron (Nd-Fe-B), backward and forward against the sample. The opto isolator circuit is used to amplify 5 volt to 12 volt from DAQ Card PCI 6025E to make the stepper motor work and also served as a protection circuit for any excess current, which could backflow into the PC in any case of short circuit or current spikes. The Web designs are based on the client-server concept. The term client-server refers to a relationship between two systems or processes. The client-server is a computer architecture that divides functions into client (requestor) and server (provider) subsystems, with standard communication methods such as Transmission Control Protocol (TCP) or Internet Protocol (IP) to facilitate the sharing of information between them (Carl, 1994). A web camera also connected to a server computer running NetMeeting enables the provision of audio and video feedback to the client, on the happenings during Hall Effect experiment, as shown in Fig.1.

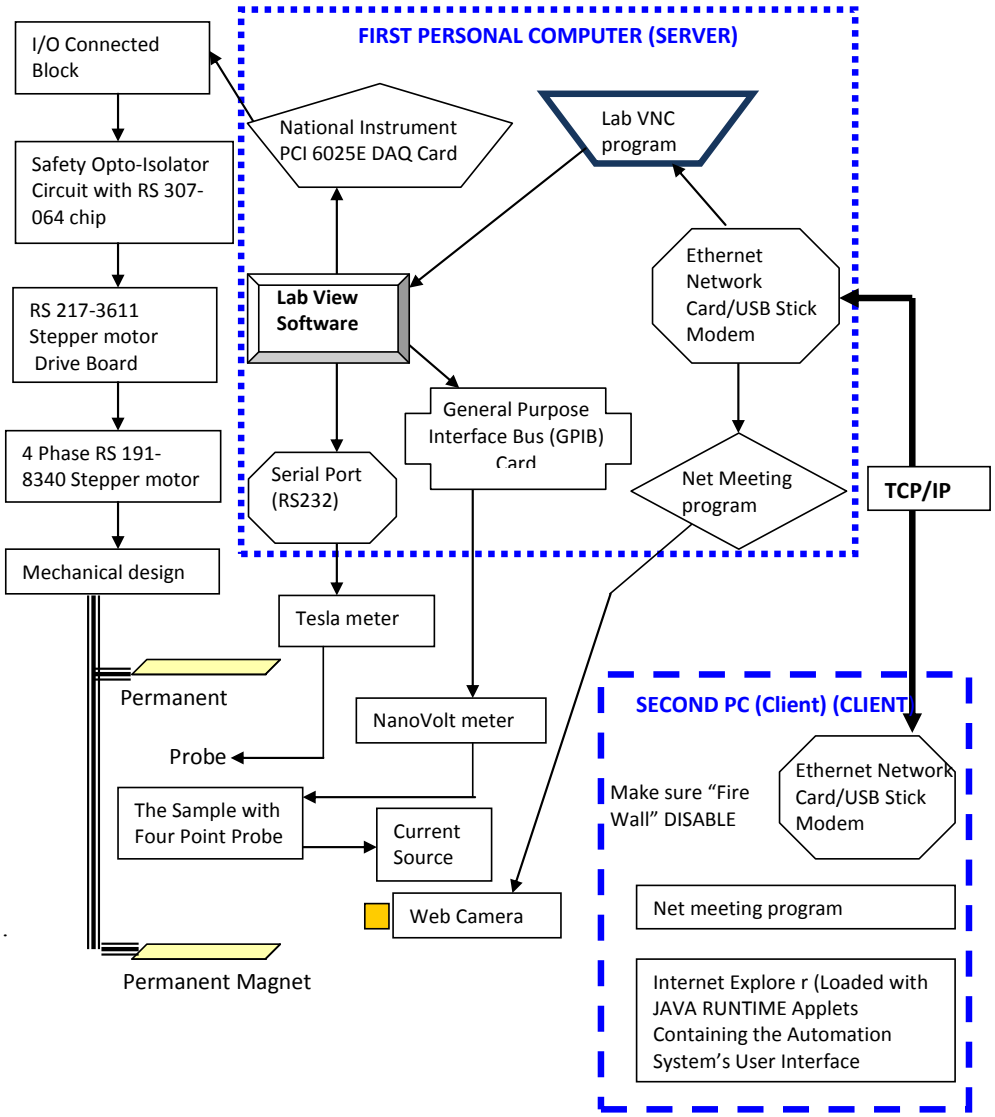



Fig.1: A Full Schematic Illustration of Web-based Automation system

Lab VNC or Lab View Virtual Network Control, a powerful third party freeware, is installed in the server computer and then configured to be a Lab View server (Travis, 2000). The server's Lab VNC will then load the user interface in the form of Java Applets into the client's web browser software such as Internet Explorer. These Lab VNC programmes run simultaneously with the automation system and enables client to control and monitor the automation system through the Internet. To achieve web-based control, the user at the server PC must NOT configure the user interface, but instead allow the user at the remote client PC to figure the user interface. Operating the automation system through web-based control involves configuring both the server and client PC as shown the flowchart in Fig.2.

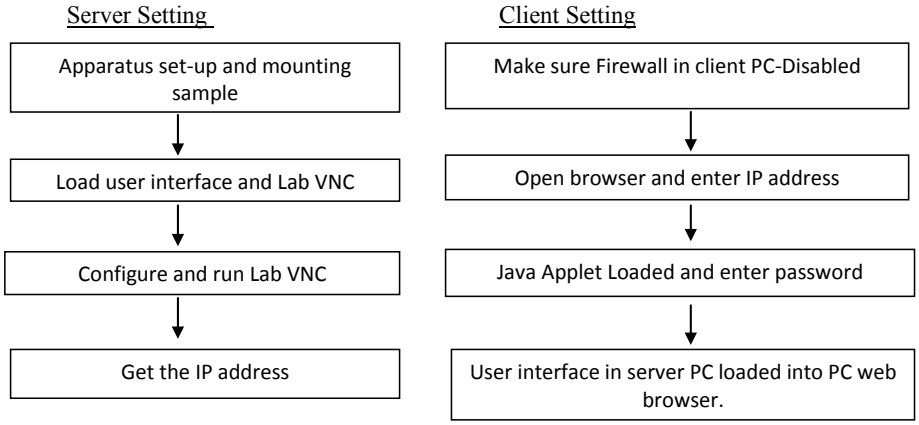



Fig.2: Flowchart Web-based control of the automation system

The sample with four point probes set up is shown in Fig.3, where two wires are connected to a current source 1.0A and the other wires are connected to a nano voltmeter. In order to have good electrical contacts, the sample surface was thoroughly cleaned, using alcohol.

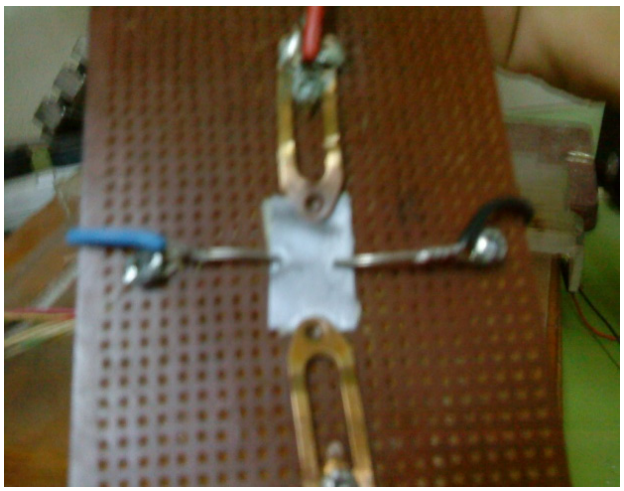


Fig.3: Sample holder with four point probes for metal sample

All the apparatus were controlled by LABVIEW programming (Wells & Travis, 1997) to acquire the data. The user interface of the LABVIEW programming containing Java Applet through Internet web is shown in Fig.4. The user interface is divided into three parts as follows:

a. The pre-setup system consisting of four features are: Test selection menu for the user to select either to choose Hall Effect experiment or to move stepper motor; Offset voltage (Brandley, 2007) for the user to provide initial voltage so that the LABVIEW programming will automatically fix this voltage as offset voltage; Dimension of the sample for providing the length, thickness and width of the sample; File selection tool that permits a user to store the acquired data.

b. The monitoring systems are a motor indicator and web camera, each one to indicate the distance between two permanent magnets and provide video and audio feedback to the client, respectively.

The data acquisition consists of parameters of the experiment that display specific values: the drift velocity, conductivity, mobility, Hall coefficient and charge carrier concentration; magnetic field and voltage reading that display the magnetic field strength and Hall voltage reading respectively; graph to display the Hall voltage as a function of the magnetic field.

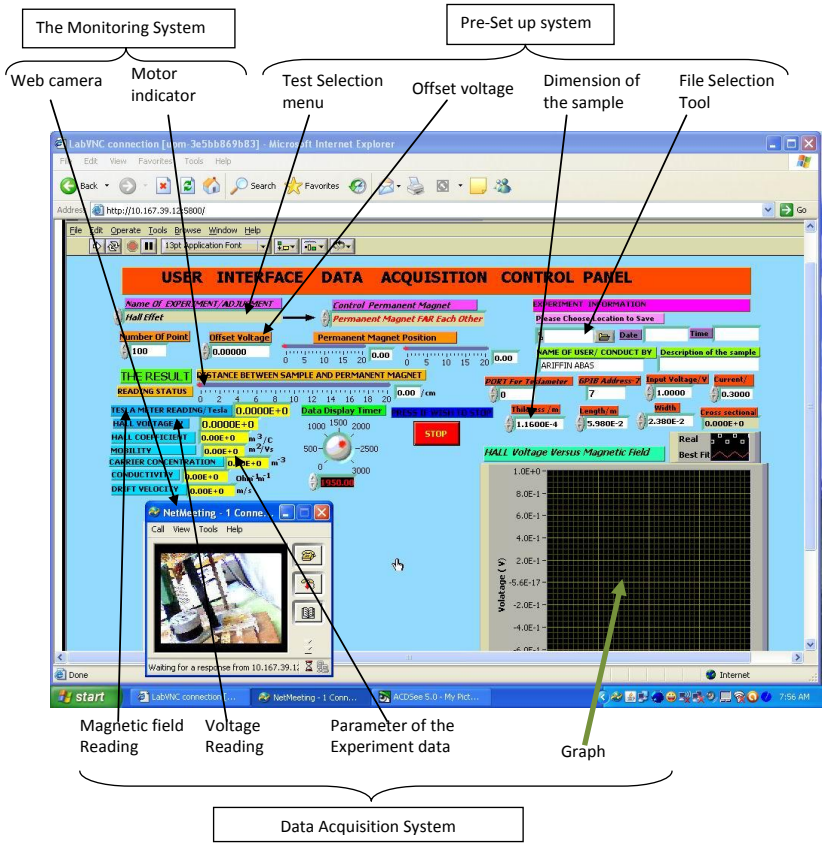



Fig.4: User Interface containing Java Applet through Internet web

The standard apparatus for Hall effect measurement, as used in the teaching laboratory and acquired from LEYBOLD, costs approximately RM36,000 per set. However, there is no facility for remote control through web-based and the data are recorded manually. An estimate of the current set-up of the whole system is approximately RM16,000.

## RESULTS AND DISCUSSION

The web-based automation system - written using Lab View programming - has been successfully developed to give instructions to the computer. The data from Tesla meter and nano voltmeter, via serial port and GPIB respectively, are automatically saved to the hard drive for further analysis and automatically plotted in real time. The Hall voltage increased linearly with increasing magnetic field strength as shown in Fig.5.

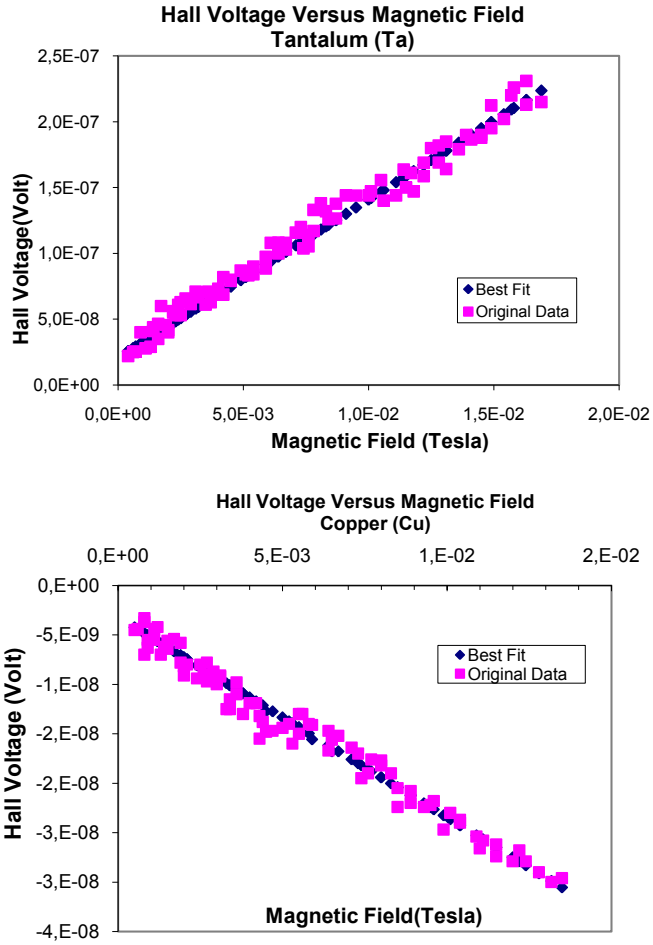



Fig.5: Hall Voltage as a function of magnetic field for tantalum and copper

Table 1 shows the results which were automatically calculated and displayed in the front panel user interface through Lab View programming and obtained by using equations 1-6.

TABLE 1
The Hall Effect result for selected materials

| Material | Thickness d±0.001 $x10^{-3}$m | Width W±0.001 $x10^{-3}$m | Length L±0.001 $x10^{-3}$m | Hall Coefficient $R_H$ $x10^{-10}$ $m^3/C$ | Charge carrier n $x10^{28}$ $m^{-3}$ | Conduction σ $x10^6$ $ohm^{-1}m^{-1}$ | Mobility μ $x10^{-3}$ $m^2/Vs$ | Drift velocity $V_d$ $x10^{-3}$ m/s |
|---|---|---|---|---|---|---|---|---|
| Copper (Cu) - 99.5% | 0.039 | 12.800 | 23.810 | -0.78 ±0.05<br>(-0.55)<br>(41.8)* | 8.01±0.52<br>(11.40)<br>(29.7)* | 47.70±2.30<br>(58.80)<br>(18.8)* | 3.72±0.29<br>(3.23)<br>(15.1)* | 0.15±0.02 |
| Tantalum (Ta) - 99.4% | 0.023 | 13.800 | 23.800 | +1.14±0.02<br>(+1.01)<br>(12.8)* | 5.48±0.09<br>(6.19)<br>(11.4)* | 7.50±0.06<br>(7.80)<br>(3.8)* | 0.88±0.02<br>(0.78)<br>(12.8)* | 0.35±0.04 |

( ) Theoretical value (Hurd, 1972; Kittel, 1971)
( )*%accuracy

Hall coefficient for copper showed that the charge carriers have a negative sign. In this case, the conduction band is dominated by electron charge carriers moving through the material and hence, making copper dominated by electrons as charge carrier. For tantalum, the Hall potential difference shows that the charge carriers have a positive sign, which means the conduction is dominated by holes, unoccupied energy levels in the valence band. The holes correspond to the absence of an electron and thus, behave like positive charge carriers moving through the material. The experimental values of Hall Effect coefficient, conductivity, mobility and charge carrier for the sample are compared to the theoretical values. The differences between the two values may be due to disturbing secondary effect particularly at the contact points or impurities in the test sample. The drift velocities for both copper and tantalum are very small. However, the drift velocity of copper is smaller than that of tantalum due to the fact that copper has a higher value of charge carrier since the drift velocity is inversely proportional to the charge carrier $V_d \propto \frac{1}{n}$.

## CONCLUSION

In this paper, the apparatus used in the experiment to measure Hall Effect are simple, cheap and easy to handle. The results obtained are in the same order of magnitude when compared to theoretical values. The Web-based automation system can be remotely controlled and monitored by users in remote locations by using only their web browsers. The advantage offered by this system is in the field of education since it can be a key component of e-learning. Students are able to view and control physics experiments carried out 'live' over the Internet and also participate over conventional web browsers in real time. Hence, the system that was developed can also be used as an affective teaching aid.

## ACKNOWLEDGEMENTS

This work was supported by Research University Grants (RUGS), Universiti Putra Malaysia, Serdang Selangor, Malaysia. A special mention and thanks are accorded to the Malaysian Ministry of Education for granting study leave to the author, thus giving the opportunity to carry out this project successfully.

## REFERENCES

Basil, H. (2011). A mimicking human arm with 5 DOF controlled by Lab View. IACSIT *International Journal of Engineering and Tech., 3*(1), 9-15.

Brandley, A. (2007). *Hall Effect Experiment*. Department of Physics and Astronomy publication. Tennessee USA.

Carl, L. H. (1994). *Technical Foundations of Client/Server Systems*. New York: John Wiley & Sons, pp. 1-20.

Edward, R. (2006). *Hall-Effect Sensors Theory and Application*. Elsevier Science (2nd Edn.) Newnes Texas USA.

Hurd, C. M. (1972). *The Hall Effect in Metals and Alloys*. New York: Plenum Press.

Kai Qian. (2009). *Embedded software development with C*. New York: Springer Dordrecht Heidelberg London New York, pp. 179-190.

Kalman, C. J. (1995). Lab View: A software for data Acquisition, Data analysis and Instrument Control. *Journal of Clinical Monitoring, 11*(1), 51-58.

Kittel, C. (1971). *Introduction to Solid State Physics* (4th Edn.). New York: John Wiley & Sons.

Wells, L. K., & Travis, J. (1997). *Lab VIEW For Everyone, Graphical Programming Made Even Easier*. Prentice Hall Virtual Instrumentation Series, New Jersey USA. Pp. 19-21.

R. S. Components (Malaysia) Sdn Bhd (1999). *Technical Manual for RS 217-3611 4 Phase Unipolar Stepper Motor Drive Board*. Issued by RS Components.

Sumathi, S. (2007). *Lab View based advanced Instrumentation system*. Springer Dordrecht Heidelberg London New York. Pp. 239-279.

Takashi Kenjo (1985). *Stepping Motors and Their Microprocessor Controls*. Monographs in Electrical and Electronic Engineering 16, Clarendon Press, Oxford. 1-3, 21, 25, 34.

Travis, J. (2000). *Internet Application in Lab View*. Upper Saddle River NJ USA Prentice Hall. pp. 350-357.

Pertanika J. Sci. & Technol. 22 (1): 113 - 121 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# *In-vitro* Antidermatophytic Activity of Methanolic Fractions from *Entada spiralis* Ridl. Stem Bark and Its Bioautographic Profile

**Aiza Harun[1]*, Siti Zaiton Mat So'ad[2], Norazian Mohd Hassan[2] and Neni Kartini Che Mohd Ramli[2]**

*[1]Faculty of Applied Sciences, Universiti Teknologi Mara Pahang, 26400 Jengka, Pahang, Malaysia*
*[2]Kulliyyah of Pharmacy, International Islamic University Malaysia, 25200 Bandar Indera Mahkota, Kuantan, Pahang, Malaysia*

## ABSTRACT

This study was performed to evaluate the antifungal activities of methanolic fractions from the stem bark of *Entada spiralis* Ridl. against human dermatophytes and yeast-like fungus *in vitro*. Three types of human dermatophyte, *Trichophyton mentagrophytes* ATCC 9533*, Microsporum gypseum* ATCC 24102 and *Trichophyton tonsurans* ATCC 28942, and one yeast-like fungus, *Candida glabrata* ATCC 66032, were tested against the methanolic fractions labelled FA1, FA4 and FA5. *T. mentagrophytes, T. tonsuran* and *M. gypseum* were susceptible to all tested fractions in a concentration-dependent manner whereas *C. glabrata* was resistant. Fraction FA1 at a concentration of 400 mg/mL was found to exhibit the highest antifungal activity with the inhibition zone diameter of 22 mm (*T. mentagrophytes*). This fraction showed a minimum inhibitory concentration MIC of 0.097 mg/mL while the MIC value for the fraction FA4 and fraction FA5 was 3.12 mg/ml and 1.56 mg/ml respectively. Agar overlay bioautography assay results showed that most of the bioactive compounds were found in the fraction FA1. Based on these findings, it can be concluded that the stem bark extracts of *E. spiralis* can be a future source of potent natural antimicrobial drugs for superficial skin diseases.

*Keywords:* Leguminoceae, *Entada spiralis* antidermatophytic activity, dermatophytes, bioautography

## INTRODUCTION

Medicinal plants contain numerous chemicals that can be employed by humans to overcome certain microbial invasions and they have been used as major sources of drug development (Kuete *et al*., 2009). Presently, their uses in traditional medicine have expanded widely and are gaining in popularity. Herbal

*Article history:*
Received:30 August 2012
Accepted: 7 June 2013

*E-mail addresses:*
aizaharun@pahang.uitm.edu.my (Aiza Harun),
dszaiton@iium.edu.my (Siti Zaiton Mat So'ad),
norazianmh@iium.edu.my (Norazian Mohd Hassan),
neni@pahang.uitm.edu.my (Neni Kartini Che Mohd Ramli)
*Corresponding Author

ISSN: 0128-7680 

medicines have been reported to serve the health needs of about 80 % of the world's population, and are especially used in the vast rural areas of developing countries (WHO, 2001). The plant extracts of herbal medicines and products have been used for the treatment of bacterial, fungal and viral infections (Bruneton, 1999; Cowan, 1999). In past decades, infection caused by dermatophytes has increased considerably. Dermatophytes are parasitic fungi that cause infections of the skin in animals and humans. These include the imperfect fungi of the genera *Epidermophyton, Microsporum* and *T. richophyton.* The infections occur as the fungi are able to obtain nutrients from keratinised material. They do not invade living tissues but colonise the outer layer of the skin. Some of the skin infections caused by *Trichophyton tonsurans, Trichophyton mentagrophytes* and *Microsporum gypseum* are tinea capitis (scalp and hair), tinea corporis (glabrous skin), tinea unguium (nails) and tinea manuum (hand). Since dermatophytes are responsible for serious human pathogenic disorders, work on alternative approaches such as the use of natural antimicrobial agents from plant extracts of herbal medicines to control such pathogens would be beneficial (Bajpai *et al*., 2009).

The antimicrobial potential of plant species of the Leguminoceae family has been reported (Fabry *et al.,* 1998; Doughari, 2006; Mboso *et al.*, 2010; Khattak *et al.*, 2010). Within this family, certain species of the genus *Entada* has also been reported to possess antimicrobial potency. For example, *Entada phaseoloides* has been reported to have been used to treat skin diseases such as eczema, itches and scabies (Ram *et al.*, 2004) While *Entada africana* is used traditionally to treat hepatitis, sores, skin-eruptions, rheumatism, cataracts, fevers, dysentery and promote wound healing (Burkill, 1995). In addition, the stem bark of *Entada spiralis* Ridl., a liana or woody climber that grows in the wild in Malaysia and locally known as 'beluru' or 'sintok', has been traditionally used for generations as a shampoo to clean the scalp and a soap for general cleaning. It has also been used to treat syphilis, insect bites and bloody defecation. Because of its wide usage and since it has not yet been studied in detail by researchers, this plant is now the subject of this study, which set out to investigate the antidermatophytic activity of this plant species against several dermatophytes that cause skin diseases. The results of this study will determine its efficacy as an antifungal agent. In the previous study, we reported that the methanol extract successfully inhibited the growth of *T. mentagrophytes, M. gypseum* and *T. tonsurans* (Harun *et al*., 2011).

## MATERIALS AND METHODS

### *Plant materials and fractionation*

The stem bark of *E. spiralis* was collected from the forest at Tasik Chini, Pahang and authenticated by a plant botanist from Universiti Kebangsaan Malaysia (voucher specimen-KMS -5228). About 3 kg of finely powdered dry stem bark of the plant was macerated in 9 L of petroleum ether for 3 days, and this process was repeated 3 times. The petroleum ether extract was filtered and evaporated to dryness. The remaining stem bark was further macerated in chloroform and methanol sequentially in a similar manner. All extracts were stored at 4 $^0$C prior to use. The methanol extract was chosen as the target extract as it exhibited moderate activity against tested dermatophytes (Harun *et al*., 2011). Thus, the screening of antifungal substances was studied in the methanol extract.

The mixture of methanol extract and silica gel (Merck, Germany) (ratio1:3) was dissolved in methanol and heated in a water bath to dryness to form a methanol extract-silica gel powder. Part of the methanol extract-silica gel powder (30 g) was subjected to vacuum liquid chromatography (VLC) on silica gel (200-430 mesh, 100 g) eluting with gradient systems of chloroform-methanol (1L for each gradient system). Fractions of 200 mL that were collected consecutively and combined on the basis of their thin layer chromatography (TLC) profiles were evaporated to dryness. The fractionation process afforded eleven main fractions and only fraction FA1 (0.15 g, eluted with chloroform-methanol 9:1), fraction FA4 (0.95 g, eluted with chloroform-methanol 6:4) and fraction FA5 (0.54 g, eluted with chloroform-methanol 5:5) fractions were further investigated. Fraction FA1, FA4 and FA5 were chosen based on their solubility in chloroform-methanol. Fractions obtained from other eluting solvents in which the composition of methanol was higher than that of chloroform were discarded because of poor solubility.

### *Microorganisms*

The microorganisms used in this study were three dermatophytes (*Trichophyton mentagrophytes* ATCC 9533, *Trichophyton tonsurans* ATCC 28942, *Microsporum gypseum* ATCC 24102) and one non-dermatophyte (*Candida glabrata* ATCC 66032). The stock cultures of dermatophytes were maintained by monthly subculturing in Saubaroud Dextrose Agar (Merck, Germany) (SDA) at room temperature.

### *Preparation of test solution and disc*

A stock solution of 400 mg/ml of fraction was prepared by dissolving 0.4 g of fraction in 1 mL of methanol and serially diluting it to achieve a concentration of 200 mg/mL, 100 mg/mL and 50 mg/mL. Sterile filter paper discs (Whatman AA disc, 6mm, England) were impregnated with 20 mL of fractions of each concentration and allowed to dry at room temperature (Chandrasekaran & Venkatesalu, 2004; Prasad *et al*., 2004). All the discs were stored at -5 $^0$C prior to use.

### *Fungal suspension*

The dermatophytes were subcultured in SDA and incubated at room temperature for 7-14 days. The mycelia was scraped aseptically, crushed and macerated thoroughly in sterile distilled water. The fungal suspension was standardised spectrophotometrically to an absorbance of 0.600 at 450 nm using a UV-VIS spectrophotometer (Secoman, France) which corresponded to 0.5 - 2.5 x $10^3$ cells/ml. It was used as an inoculum for antifungal susceptibility testing (Pankajalakhsmi *et al*., 1995; Prasad *et al*., 2004, Chandarasekaran & Venkatesalu, 2004).

### *Antifungal assay*

Antifungal activity was evaluated using the disc diffusion agar method (Bauer *et al.,* 1966) and broth microdilution. Test plates were prepared by pouring 20 mL of sterile molten SDA into petri dishes after which the agar was allowed to solidify. A sterile cotton swab was dipped

into the standardised fungal suspension and the solidified agar was then uniformly swabbed. The impregnated paper discs were applied and incubated for 48-168 hr (4-7 days) for *Candida* and the dermatophytes respectively. The antifungal activity of the fractions was determined by measuring the clear zones of growth inhibition around the paper discs. Nystatin (100 *u*g/disc) (Oxoid, England) was used as standard reference. The solvent and empty discs were used as a negative control. The clear zone around the discs indicated that the inhibition was fungicidal while the absence of a clear zone showed fungistatic properties. All assays were carried out in triplicate.

In the broth microdilution assay, the fraction was two-fold serially diluted using an appropriated solvent in a 96-well microtiter plate to give a final volume of 200 *u*L with a final concentration ranging from 50 mg/ml to 0.024 mg/mL. The plates were incubated at room temperature for 2-5 days. To indicate fungal growth, 20 *u*L 0.5% (w/v) of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) (Merck, Germany) solution was added to each well and incubated for 2 hrs. The dark-blue colour that emerged revealed microbial growth whereas the colour yellow showed inhibition of microbial growth. The lowest concentration of fraction that produced a yellow colour in a microtiter well would give the minimum inhibitory concentration (MIC) value of the fraction. Appropriate controls such as solvent used to dissolve fractions, Saubaroud Dextrose Broth (SDB) with inoculum, SDB alone and nystatin were also used in the assay. The MIC was defined as the minimum concentration of the samples that inhibited a visible growth of the tested microbes.

*Agar overlay bioautography assay*

The thin layer chromatography (TLC) of each fraction was performed on sterile 10 cm x 10 cm commercial aluminium sheets silica gel $60F_{254}$ of layer thickness 0.2 mm (Merck, Germany). Two TLC plates of each fraction were developed in the suitable developing solvent system and one of them was used as a reference chromatogram. The chromatogram was placed in sterile square petri dishes after the complete removal of solvent. The inoculum concentration was measured and adjusted to achieve 0.5 - 2.5 x $10^3$ cells/ml by diluting with sterile agar at $40^0$C. Twenty mililitres of inoculated molten agar were immediately distributed on each chromatogram to form a thin layer of agar and was then allowed to solidify. The plates were incubated at room temperature for 2-5 days. The areas where active compounds of fraction were located were detected as a clear zone against the background of living fungal mycelia. The reference chromatograms were observed under UV lights at 254 nm and 366 nm and analysed after spraying with vanillin/sulfuric acid reagent.

## RESULTS AND DISCUSSION

The results of the disc diffusion agar method indicated that all fractions possessed antifungal activity against all tested dermatophytes in a concentration-dependent manner except for *Candida glabrata* (Table 1). Fraction FA1 had a statistically bigger inhibition zone compared to the fraction FA4 and the fraction FA5. Therefore, fraction FA1 was found to be the most effective fraction with the highest inhibition zone diameter of 22 mm against *T. mentagrophytes*. All fractions displayed lower activity against all tested dermatophytes at concentration of 50

mg/mL. The standard antifungal drug test showed *T. tonsurans* was the most sensitive towards nystatin at 100 $\mu$g/disc with an inhibition zone of 42.5 mm.

The results from this current study revealed the scientific basis of the traditional usage of *E. spiralis*. The results were comparable with previous reports where methanol extract from *Psoralea corylifolia* (Leguminoceae) suppressed the growth of *T. mentagrophytes* and *M. gypseum* with inhibition zones of 28 mm and 25 mm respectively (Prasad *et al*., 2004). Likewise, the ethanol extract of *E. rheedei* (Leguminoceae) was reported to exhibit moderate activity against fungi such as *Candida albicans* (Ram *et al.,* 2004). Currently, Mutai *et al*., (2009) found the number and the nature of substituents of triterpene responsible for the antimicrobial activity of the *Acacia mellifera* extract (Leguminoceae) against *S. aureus* and *T. mentagrophytes*. The inhibition of the microbial growth could depend on the presence of hydroxyl groups and the nature of conformation in the triterpene molecule. Thus, these factors could also be the reason for the effectiveness of fraction FA1 since terpenoid compounds were found in this fraction.

The MIC of all tested fractions against dermatophytes in the broth microdilution assay is given in Table 2. Fraction FA1 showed MIC of 0.097 mg/mL for *T. mentagrophytes*, 3.125 mg/mL for *T. tonsurans* and 0.195 mg/mL for *M. gypseum*. For the fraction FA4, the MIC was determined as 3.125 mg/mL for *T. mentagrophytes* and *T. tonsurans* and 6.25 mg/mL for *M. gypseum*. Fraction FA5 showed MIC value of 1.56 mg/mL, 0.78 mg/mL and 3.125 mg/mL for *T. mentagrophytes, T. tonsurans* and *M. gypseum*, respectively. The standard antibiotic nystatin (100$\mu$g/disc) showed MIC of 0.0019 mg/mL against *T. mentagrophytes*, 0.031 mg/mL against *T. tonsurans* and 0.25 mg/mL against *M. gypseum*.

TABLE 1
*In vitro* Antifungal Activity of FA1, FA4 and FA5 Fractions of Methanol Extract of *E. spiralis* Stem Bark Against Tested Dermatophytes

| Fractions | Concentration (mg/ml) | Inhibition zone (mm) [a] | | | |
|---|---|---|---|---|---|
| | | TM | TT | MG | CG |
| FA1 | 50 | 10±0 | 9 ±0 | - | - |
| | 100 | 14.50±2.12 | 13±1 | 12±0 | - |
| | 200 | 19±0 | 13.50±3.54 | 18±0 | - |
| | 400 | 22.0±1.41* | 17.33±1.53* | 19.50±0.71* | - |
| FA4 | 50 | - | - | - | - |
| | 100 | - | 8.67±1.15 | - | - |
| | 200 | 10±0 | 10.33±0.58 | 9.0±1.73 | - |
| | 400 | 11±0 | 14.0±0 | 11.0±1.73 | - |
| FA5 | 50 | - | - | - | - |
| | 100 | 8.0±0 | 8.67±1.15 | 8.67±0.58 | - |
| | 200 | 10.0±0 | 10.67±0.58 | 10.0±1 | - |
| | 400 | 12.0±0 | 12.67±1.53 | 10.0±1 | - |
| Nystatin | | 21±0 | 42.50±0.71 | 21±0 | 24.50±0.71 |

-, No activity; TM, *Trichophyton mentagrophytes*; TT, *Trichophyton tonsurans*; MG, *Microsporum gypseum*; CG, *Candida glabrata*; ±, Standard deviation (SD); [a] Mean of three replicates; Statistical significance was determined using paired *t*-test. Differences were considered significant at (P<0.05); *, larger inhibition zone compared to FA4 and FA5

TABLE 2
MIC (mg/mL) of FA1, FA4 and FA5 Fractions of Methanol Extract of *E. spiralis* Stem Bark Against Tested Dermatophytes

| Dermatophytes | Minimum inhibitory concentration (mg/ml) | | | |
|---|---|---|---|---|
| | FA1 | FA4 | FA5 | Nystatin (*u*g/mL) |
| *Trichophyton mentagrophytes* | 0.097 | 3.125 | 1.56 | 0.0019 |
| *Trichophyton tonsurans* | 3.125 | 3.125 | 0.78 | 0.031 |
| *Microsporum gypseum* | 0.195 | 6.25 | 3.125 | 0.25 |

Based on MIC determination, it was found that fraction FA1 possessed strong inhibitory activity against *T. mentagrophytes* and *M. gypseum* but only weak activity against *T. tonsurans*, which was not consistent with the results of the disc diffusion assay. Uneven fungal plating on the surface of the agar could have been one reason for this inconsistency. Fraction FA4 showed weak inhibition against all tested dermatophytes while fraction FA5 showed moderate activity against *T. mentagrophytes* and *T. tonsurans* and weak inhibition against *M. gypseum*. The broth microdilution assay results of fraction FA4 and FA5 seemed to be consistent with the disc diffusion assay results. Generally, the disc diffusion assay is a semi-quantitative assay which is employed to screen the fractions against susceptible fungal strains and may produce equivocal results in quantitative determination. In the broth microdilution assay, the antidermatophytic activity was evaluated quantitatively by means of the MIC values which established the fungal susceptibility. Hence, MIC results confirmed the effectiveness of fraction FA1.

The MIC results were comparable to MIC results of the active compounds from *P. corylifolia*, which showed MIC of 0.0625 mg/mL for *T. mentagrophytes* and 0.125 mg/mL for *M. gypseum* (Prasad *et al*., 2004). Significantly, the MIC determination provided an important guideline for choosing an appropriate effective concentration of a dermatophyte-inhibiting substance. The suppression of a dermatophyte's growth might be due to interference by the active constituents of the fractions; this is supported by Ibrahim and Osman, 1995.

Bioautography is an assay used to facilitate the isolation of antimicrobial active compounds from natural products. It is considered to be the most efficacious technique for the detection of antimicrobial compounds as it allows the localisation of the activity even in a complex matrix and, therefore, permits a target-directed isolation of the active constituent (Rahalison *et al*., 1991). Usually, the active compounds can be seen as clear spots against the background of growing fungi. The results for the TLC agar overlay bioautography of fraction FA1, FA4 and FA5 are shown in Table 3. Generally, the results revealed the presence of antidermatophytic active compounds in all fractions according to the clear inhibition zones on bioautograms. Fraction FA1 was found to contain antidermatophytic active compounds that inhibited the growth of all tested dermatophytes, whereas active compounds that had been screened in the fraction FA4 inhibited the growth of *T. mentagrophytes* and fraction FA5 inhibited *T. tonsurans*. The results of the antifungal bioautographic profiles of fraction FA1 presented in Table 4 revealed that the inhibitory effect was due to the presence of terpenoid compounds as detected by spraying the reference chromatogram with vanillin/sulfuric acid reagent (Fig.1). The bioautogram of B, C and D in fig.1 displayed the location of the inhibition zone against *M.*

*gypseum*, *T. mentagrophytes* and *T. tosurans* respectively, while the reference chromatogram A showed various terpenoid compounds found in the fraction FA1. The terpenoid compounds labelled as 1, 2 and 3 are the responsible antidermatophytic compounds of the fraction. The clear inhibition zone detected on bioautogram B was due to the presence of terpenoid compound 3, while the clear inhibition zone on bioautogram C resulted from the presence of terpenoid compound 1 and 3. The clear inhibition zone on bioautogram D was caused by the presence of terpenoid compound 1 and 2. Therefore, the results were very encouraging as the fraction FA1 was verified as containing antidermatophytic substances. This is consistent with previous investigations reporting that triterpenoid saponins present as major compounds such as echinocystic acid, entagenic acid, acacic acid, homogentistic acid and diterpene kolavenol in *Entada sp* act as bioactive compounds (Freiburghaus *et al*., 1998; Cioffi *et al*., 2006; Nzowa *et al*., 2010). Although no phytochemical investigation has been recorded for *E. spiralis* to date, it is possible that such bioactive compounds may be responsible for the antidermatophytic properties of this plant.

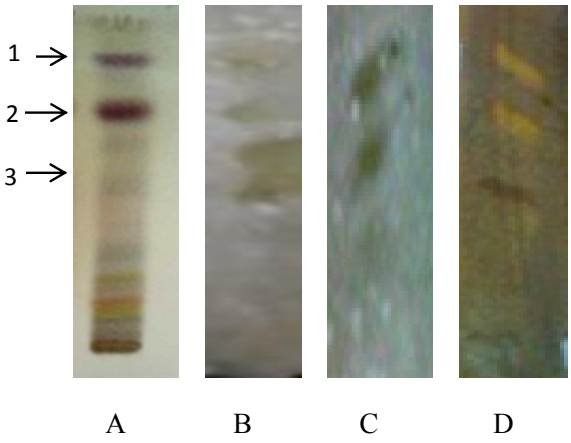



Fig.1: TLC bioautography profiles of FA1 fraction. A. Chromatogram after spraying with vanillin/ H2SO4 B. Bioautogram against *M. gypseum* C. Bioautogram against *T.mentagrophytes* D. Bioautogram against *T. tonsurans* (after exposure to iodine vapour)

TABLE 3
Agar Overlay Bioautography of FA1, FA4 and FA5 Fractions of Methanol Extract of *E. spiralis* Stem Bark Against Tested Dermatophytes

| Fraction | *T.mentagrophytes* | *T. tonsurans* | *M. gypseum* |
|---|---|---|---|
| FA1 | √ | √ | √ |
| FA4 | √ | x | x |
| FA5 | x | √ | x |

√ = Clear zone ; x = Absence of clear zone

TABLE 4
Antifungal Bioautographic Profile of Fraction FA1 of *E. spiralis* Stem Bark

| Labelled Susceptible Compound | $UV_{254}$ light | $UV_{366}$ light | Vanillin/$H_2SO_4$ | Comment | fungal |
|---|---|---|---|---|---|
| 1 | Black+ | Blue+ | Purple+++ | Terpenoid | *TT, TM* |
| 2 | Black+ | Blue+++ | Purple+++ | Terpenoid | *TT, MG* |
| 3 | Black+ | Blue+++ | Grey++ | Terpenoid | *MG, TM* |

Intensity of colour: +++ high, ++ definite, + slight

## CONCLUSION

*Entada spiralis* is a promising antidermatophytic species because of significant inhibitory activity of its fractions on dermatophytic proliferation. We believe that fraction FA1 has great potential to treat skin infections, and we have provided important scientific support for the traditional use of the stem bark of *E. spiralis*. The results also can be considered as a new finding since no antidermatophytic studies against dermatophytes have been reported for this species previously. For further work, the isolation of natural antifungal compounds need to be carried out in order to identify the active compounds responsible.

## ACKNOWLEDGEMENTS

The authors wish to express their gratitude to Kulliyyah of Pharmacy International Islamic University Malaysia (IIUM) , Universiti Teknologi Mara (UiTM) Kampus Khazanah Alam Pahang and e-science fund (project no: 06-01-08-SF0074) for their help in completing this research work.

## REFERENCES

Bajpai, V. K., Yoon, J. I., & Kang, S. C. (2009). Antioxidant and antidermatophytic activities of essential oil and extracts of *Metasequoia glyptostroboides* Miki ex Hu. *Food and Chemical Toxicology*, *47*, 1355-1361.

Bauer, A. W., Kirby W. M. J. C., & Jurk, M. (1966). Antibiotic susceptibility testing by a standardized single disc method. *American Journal of Chemical Pathology, 45*, 493-496.

Bruneton, J. (1999). Pharmacognoise: Phytochimie, *Plantes Medicinales*. Tec & Doc, Paris, pp. 309-354.

Burkill, H. M. (1995). The useful plants of Tropical Africa, Families J - L, *vol. 3*. Royal Botanic Gardens Kew, pp. 229-230.

Chandrasekaran, M., & Venkatesalu, V. (2004). Antibacterial and antifungal activity of *Syzygium jambolanum* seeds. *Journal of Ethnopharmacology*, *91*, 105-108.

Cioffi, G., Dal Piaz, F., De Caprariis, P., Sanego, R., Marzocco, S., Autore, G., & De Tommasi, N. (2006). Antiproliferative triterpene saponins from *Entada Africana*. *Journal of Natural Products*, *69*, 1323-1329.

Cowan, M. M. (1999). Plant product as antimicrobial agents. *Clinical Microbiology Review*, *12*, 564-582.

Doughari, J. H. (2006). Antimicrobial activity of *Tamarindus indica* Linn. Tropical *Journal of Pharmaceutical Research*, *5*(2), 597-603.

Fabry, W., Okemo, P. O., & Ansorg, R. (1998). Antibacterial activity of East African medicinal plants. *Journal of Ethnopharmacology*, *60*, 79-84.

Freiburghaus, F., Steck, A., Pfander, H., & Brun, R. (1998). Bioassay-guided isolation of a diastereoisomer of kolavenol from *Entada abyssinica* on *Trypanosoma brucei rhodesiense. Journal of Ethnopharmacology*, *61*, 179-183.

Harun, A., Mat So'ad, S. Z., Mohd Hassan, N., & Che Mohd Ramli, N.K. (2011). In-vitro study of antifungal activity of E*ntada spiralis* crude extract against dermatophytes of superficial skin disease. *Revelation and Science, 1*(01), 57-61.

Ibrahim, D., & Osman, H. (1995). Antimicrobial activity of *Cassia alata* from Malaysia. *Journal of Ethnopharmacology, 45*, 151-156.

Khattak, K. F., & Simpson, T. J. (2010). Effect of gamma irradiation on the antimicrobial and free radical scavenging activities of *Glycyrrhiza glabra* root. *Radiation Physics and Chemistry*, *79*, 507-512.

Kuete, V., Nana, F., Ngameni, B., Mbaveng A.T., Keumedjio, F., & Ngadjui, B. T. (2009). Antimicrobial activity of the crude extract, fractions and compounds from stem bark of *Ficus ovata* (Moraceae). *Journal of Ethnopharmacology*, *124*, 556-561.

Mboso, E. J. T., Ngouela, S., Nguedia J.C.A., Beng, V.P., Rohmer, M., & Tsamo, E. (2010). *In vitro* antimicrobial activity of extracts and compounds of some selected medicinal plants from Cameroon. *Journal of Ethnopharmacology*, *128*, 476-481.

Mutai, C., Bii, C., Vagias, C., Abatis, D., & Roussis, V. (2009). Antimicrobial activity of *Acacia mellifera* extracts and lupane triterpenes. *Journal of Ethnopharmacology*, *123*, 143-148.

Nzowa, L. K., Barboni, L., Teponno, R. B., Ricciutelli, M., Lupidi, L. Q., Bramucci, M., & Tapondjou, L. A. (2010). Rheediinosides A and B, two antiproliferative and antioxidant triterpene saponins from *Entada rheedii*. *Phytochemistry*, *71*, 254-261.

Pankajalakshmi, V., & Taralakshmi, V. V. (1995). Disk diffusion susceptibility testing of dermatophytes with imidazoles. *Indian Journal of Pathology and Microbiology*, 369-374.

Prasad, N. R., Anandi, C., Balasubranium, S., & Pugalendi, K. V. (2004). Antidermatophytic activity of extracts from *Psoralea corylifolia* (Fabaceae) correlated with the presence of flavonoid compound. *Journal of Ethnopharmacology, 91*, 21-24.

Rahalison, L., Hamburger, M., Hostettmann, K., Monod, M., & Frenk, E. (1991). A bioautographic agar overlay method for the detection of antifungal compounds from higher plants. *Phytochemical Analysis*, *2*, 199-203.

Ram, A. J., Bhakshu, L. Md., & Venkata Raju, R.R .(2004). In vitro antimicrobial activity of certain plants from Eastern Ghats, India, used for skin diseases. *Journal of Ethonopharmacology*, *90*, 353-357.

World Health Organisation. (2001). *General guidelines for methodologies on research and evaluation of traditional medicine*. WHO. Geneva, Switzerland, p. 1.

Pertanika J. Sci. & Technol. 22 (1): 123 - 138 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Assessment of Digital Camera in Mapping Meandering Flume using Close Range Photogrammetric Technique

**Udin, W. S.[1]*, Ahmad, A.[2] and Ismail, Z.[3]**

*[1]Department of Geosciences, Faculty of Earth Sciences, Universiti Malaysia Kelantan, Jeli Campus, Locked Beg No. 100, 17600 Jeli, Kelantan, Malaysia*
*[2]Department of Geoinformatics, Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Johor, Malaysia*
*[3]Department of Hydraulics and Hydrology, Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Johor, Malaysia*

## ABSTRACT

In recent years image acquisition in close range photogrammetry relies on digital sensors such as digital cameras, video cameras, CCD cameras etc that are not specifically designed for photogrammetry. This study is performed to evaluate the compatibility of the digital metric camera and non-metric camera for the purpose of mapping meandering flume, using close range photogrammetric technique and further, to determine the accuracy that could be achieved using such a technique. The meandering flume provides an opportunity to conduct an experimental study in a controlled environment. In this study, the digital images of the whole meandering flume were acquired using a compact digital camera - Nikon Coolpix S560, a Single Lens Reflex (SLR) Nikon D60 and also a metric digital camera Rollei D30. A series of digital images were acquired to cover the whole meandering flume. Secondary data of ground control points (GCP) and check points (CP), established using the Total Station technique, was used. The digital camera was calibrated and the recovered camera calibration parameters were then used in the processing of digital images. In processing the digital images, digital photogrammetric software was used for processes such as aerial triangulation, stereo compilation, generation of digital elevation model (DEM) and generation of orthophoto. The whole process was successfully performed and the output produced in the form of orthophoto. The research output is then evaluated for planimetry and vertical accuracy using root mean square error (RMSE). Based on the analysis, sub-meter accuracy is obtained. It can be concluded that the differences between the different types of digital camera are small . As a conclusion, this study proves that close range photogrammetry technique can be used for mapping meandering flume using both the metric digital camera and non-metric digital camera.

*Keywords:* Close range photogrammetry, digital camera, mapping, meandering flume, accuracy

*Article history:*
Received: 5 July 2011
Accepted: 18 December 2012

*E-mail addresses:*
wanisofia@umk.edu.my (Udin, W. S.),
anuarahmad@utm.my (Ahmad, A.),
zulhilmi@utm.my (Ismail, Z.)
*Corresponding Author

ISSN: 0128-7680 

## INTRODUCTION

Triangulation is the fundamental principle which is utilized in photogrammetry. Triangulating the coordinates within three dimensional spaces often enable the object to be measured in the photographs. Early development in the theory and science of photogrammetry occurred many years before the actual invention of a suitable means to apply the application. Over the years, photogrammetry has gone through several development phases and has now entered into digital photogrammetry phase. Digital photogrammetry technique has been used for object modelling in various close range applications. Most industrial applications require surveyors to perform dimensional measurements as three dimensional (3D) digital views. Nowadays, the generation of 3D modelling can be achieved by using non-contact systems in co-operating the close range photogrammetry technique. Additionally, images taken by non-metric digital cameras can be utilized in this technique.

The digital camera is categorized as a non-metric camera where the camera is not specifically built for photogrammetric purposes. The digital camera is not characterized with fiducial mark, unstable calibration parameter, small format and many more (Udin *et al*., 2012). "Metric cameras" are specially designed for photogrammetric tasks (Peipe & Stephani, 2003). These cameras have a robust mechanical structure, well-aligned lenses with low distortion and do not have autofocus or other technologies that can uncontrollably change the internal geometry of the camera. All other cameras are referred to as non-metric cameras (Sanz-Ablanedo et al. (2009). The most important difference between non-metric and metric cameras is the stability parameters of the camera. Images used for photogrammetry can originate from a metric camera, an ordinary camera or from digital sensors. The image can be recorded from a device mounted on a satellite, on an aeroplane, or on a tripod which is set up on the ground.

Affordable digital cameras are making photogrammetric practices useful for a variety of applications. An example of this is the river channel studies. According to Geisler *et al*. (2003), research into the hydraulic experiment place considerable emphasis on the riverbed topography measurement, because the measurement of the surface structure of the river is essential in understanding both bed roughness and sediment transport. There are several existing techniques for mapping the surface of the river in the hydraulic model tests such as wool threads, depth pointers, digital photogrammetry, and projection Moire. Optical techniques such as photogrammetry and projection Moire allowed measuring the overall surface of the river bed immediately and with the experimental work to persist without delays. There is considerable evidence of photogrammetry being used to measure river channel information.

Several studies by a number of researchers were conducted (Chandler *et al*., 2001; Chandler *et al*., 2003; Lane, 2000; Lane *et al*., 2001) to develop ideas about the relationships between river channel flow, sediment transport and bed form development. However, in Malaysia the use of close range photogrammetric for the application of mapping stream is seen as a new approach. This paper describes an experimental study conducted for mapping a meandering flume. Close range photogrammetric technique used for data acquisition of the laboratory flume and the images of the flume were acquired using both metric and non-metric digital cameras. The purpose of using stream model is to get some idea on how to gain experience of measuring real stream in the field using close range photogrammetric technique. The digital cameras used

are the metric camera Rollei D30, compact camera Nikon Coolpix S560 and Single Lens Reflex (SLR) Nikon D60 as illustrated in Fig.1(a), (b) and (c). Results are discussed and this paper concludes with a brief summary.


Fig.1: (a) Rollei D30; (b). Nikon Coolpix S560; (c). Nikon D60

## MATERIALS AND METHODS

In this study, there were seven stages that made up the methodology and it was conducted in a multi-disciplinary way. Fig.2 shows the flow chart of the research methodology.

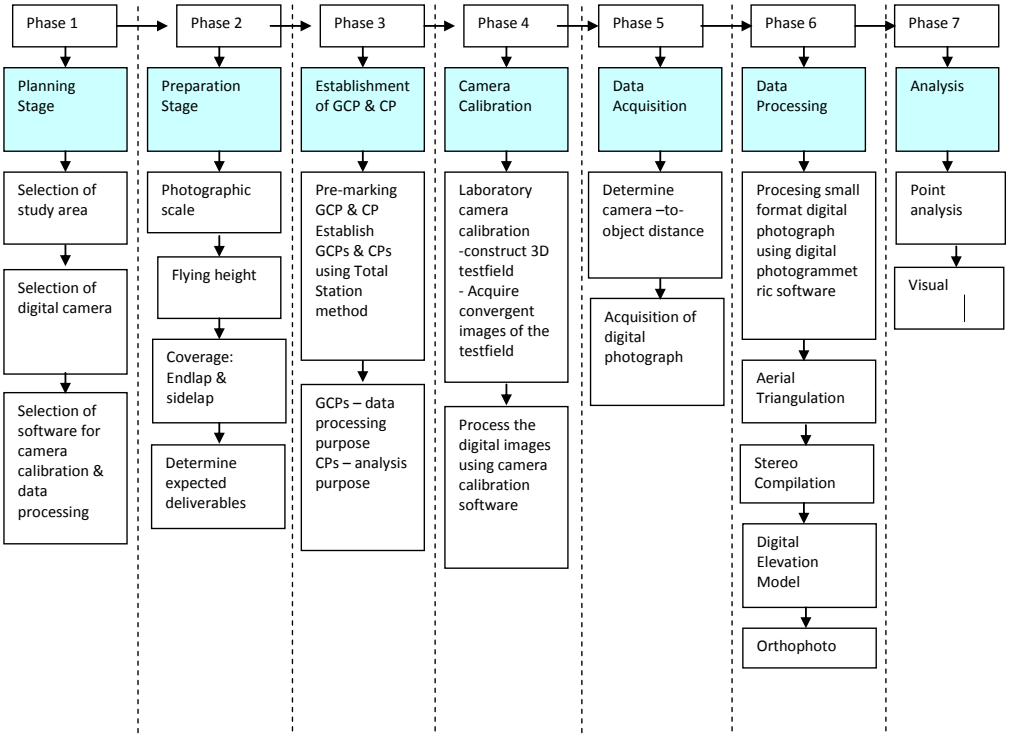



Fig.2: A flow chart of research methodology

*Planning Stage*

The planning stage included the selection of digital camera and establishment of control point and check point, method of calibrating the digital camera, planning the photography and selection of study area. The boundary of mapping area was determined. Selection of software for camera calibration and data processing was also made. The digital camera is perhaps the most important photogrammetric instrument, since it records the image on which the photogrammetric principles will be applied. Fig.3 depicts a laboratory flow channel or meandering flume. It is located at Universiti Teknologi Malaysia. The dimensions of this flume are 12.0 × 3.0m and the channel width is 0.5m. This laboratory flow channel was used as an attempt to replicate physical structures such as meandering streams found in the real world.

Fig.3: Laboratory meandering flume at UTM

*Preparation Stage*

The preparation stage is very important and it usually comprises of several steps such as converting project requirements to arrangements in terms of areas to be mapped and the desired map scale. Photogrammetric specifications include photographic height, the number of photographs, the number of strip, pixel size, photo scale, aerial triangulation, map compilation and percentage of end lap and side lap. The project planning stage is very important in as far as to obtain information related to the design and dimension of the mapping object. In addition, the expected deliverables including details on what features to be mapped and their graphic representation in the planning stage is defined.

*Establishment of Ground Control Point (GCP) and Check Point (CP)*

The ground control points (GCPs) and check points (CPs) need to be planned properly before the acquisition of digital photograph. It is necessary to relate measurements derived from the photographic images to a 3D site coordinate system. The most effective means of achieving this involves placing a number of photogrammetric target points throughout the area of interest.

Photogrammetric control targets (138) were established on the flood plain and inside the channel bed. Ninety (89) GCP were registered as a full control (XYZ) and 49 check points (CPs) were established evenly in the channel bed. The targets must be placed on the meandering flume and maintained until image acquisition is completed. The distribution of the GCPs is flexible and they need to be seen on a pair of photograph (i.e., stereo pair) at known location.

In this study, the designed blocks of the photographs consist of many stereo pairs and they are also used in aerial triangulation. Fig.4 shows where the control points were placed throughout the area of interest. These targets were 10 mm in diameter and of conventional red and black design. Horizontal and vertical angles were measured to these markers from two survey stations and their positions were coordinated using total station. Both horizontal and vertical angles were measured using a Leica TPS1100 total station. It has an automatic function of automatic target recognition (ATR). The ATR fine points to targets by itself. Manual sighting is no longer required. Surveys are completed quicker. Accuracy of total station measurement is 3mm and time for a measurement is 3 seconds.

Apart from 3D coordinates, contour line is also generated for visual purpose. The minimum contour line interval generated is 1.0 metre, while for the maximum contour interval is 5 metres. The maximum contour level is 2.263 metres. Fig.5 shows the contour generated from the survey of photo control targets.

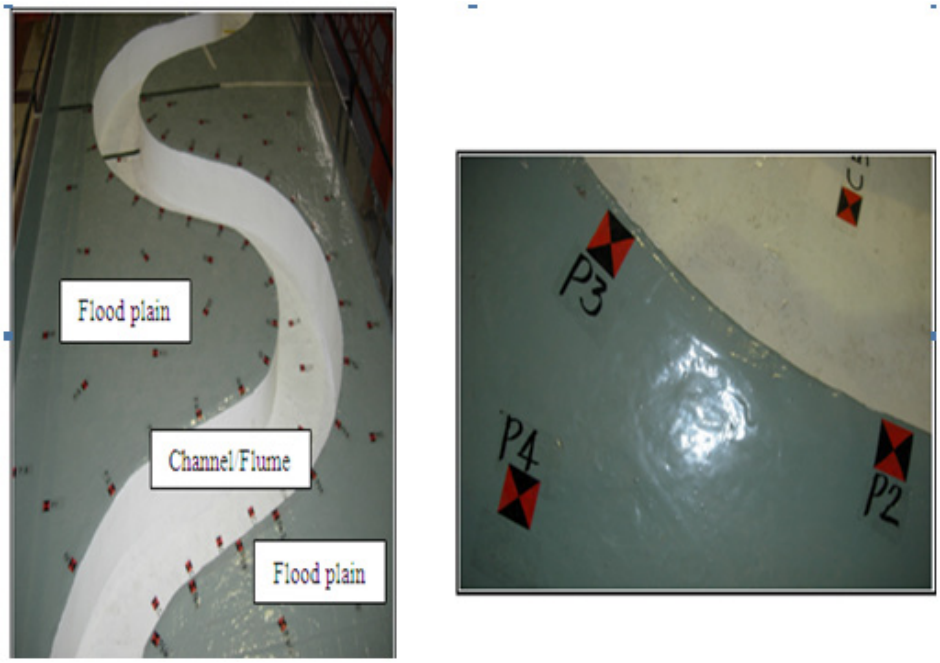



Fig.4: Photogrammetric target points

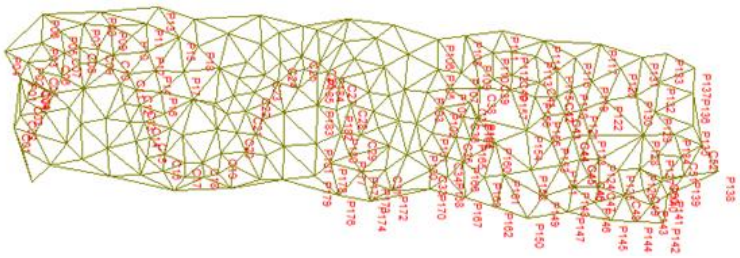


Fig.5: Contour line from installation and survey of photocontrol targets

*Camera Calibration*

Another critical aspect of photogrammetry is the calibration procedure. It is well established that accurate data can be only extracted if the parameters used to describe the internal geometry of the camera are known accurately (Chandler *et al.,* 2001; Wackrow *et al.,* 2008). The camera must be precisely calibrated to remove errors that are still present in the system. Both metric and non-metric digital cameras were calibrated in a process called self-calibration as a by-product of the actual measurement. The digital cameras were calibrated using a 3D test field. The calibration plate has a dimension of 0.4 meter x 0.4 meter and consists of 36 screws of different heights and arranged in matrix form of 6 x 6 units. Fig.6 shows an example of the calibration plate or test field. Retro-reflective target is adhered on top of each screw.

The digital cameras were used to acquire photographs of the photogrammetric test field at a constant distance (1 metre ) from the camera to the midpoint of the calibration plate. Before taking any photographs, a scale bar of known value is placed in the calibration plate or the test field (see Fig.6). The retro-reflective targets were illuminated by the built-in flash on every digital camera. Convergent photographs were taken with eight pieces of photographs for each camera. The photographs were taken with the camera in normal landscape position and then when it was rolled at 90° and the photographs were acquired from four different camera locations in space. The digital camera needs to be rotated to 90° to recover the principle point (Ahmad *et al*., 2003).

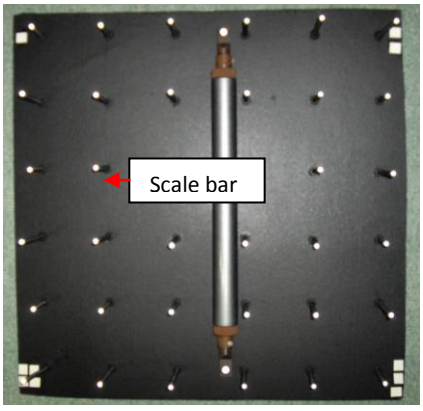



Fig.6: Test field of 0.4m x 0.4m

The images from four different positions were processed using Australis software. Australis performs the self-calibration bundle adjustment process and generates 3D coordinates of all the retro-reflective targets (Australis, 2001). Australis also provides the value of camera calibration parameters of each digital camera. This software can produce a value of precision known as posteriori variance factor (Sigma nought). The smallest value reflects the highest precision. The digital camera was calibrated three times. From the calibration process, 10 parameters were re-covered. However, only 8 parameters were used. According to Fraser (1996) and Shortis *et al*. (1998), not all the re-covered calibration parameters must be used. In this study, the 8 re-covered camera calibration parameters were then used in image processing for surface measurement of the meandering flume. The estimates and standard deviation for the camera calibration parameters are shown in Table 1.

It demonstrates that similar inner orientation parameters, comprising focal length (c), principal point offset (Xp, Yp), and correction terms for radial lens distortion (k 1, k 2, k 3), tangential lens distortion (p 1, p 2), affinity (b 1) and the differences in scale factor (b 2) were recovered for the three digital cameras. By using the convergent configuration network, additional parameters for all the three cameras can be obtained. These parameters should be tested to determine whether it is necessary to compare with stochastical properties. If the estimated value is smaller than the standard deviation, statistical parameter estimation is not required.

Typically, additional parameters that are not necessary and associated with a high correlation between estimate parameters contribute to poor geometry. In this study, every time the camera is calibrated, only a slight difference of focal length is derived for the three cameras based on the observed dataset. However, the focal length of metric Rollei D30 camera is more stable (Udin & Ahmad, 2011).

TABLE 1
Camera Calibration Parameters

| **Parameter** | **Digital Camera** | | | | | |
|---|---|---|---|---|---|---|
| | Nikon Coolpix | | Nikon SLR D60 | | Rollei D30 | |
| | Value | Std.Dev | Value | Std.Dev | Value | Std.Dev |
| c (mm) | 7.7622 | 6.302E-03 | 21.1541 | 2.251E-02 | 10.6458 | 1.001E-02 |
| Xp | -0.0643 | 7.667E-03 | 0.0896 | 1.935E-02 | 0.2721 | 9.605E-03 |
| Yp | -0.1680 | 7.039E-03 | 0.1691 | 1.792E-02 | -0.0480 | 9.859E-03 |
| k1 | 7.38201E-03 | 2.777E-04 | 6.92793E-05 | 1.068E-04 | 1.25622E-03 | 1.439E-04 |
| k2 | -8.96198E-04 | 1.244E-04 | 6.58291E-06 | 7.876E-06 | 2.50733E-05 | 3.616E-05 |
| k3 | 1.02754E-04 | 1.738E-05 | -1.32093E-07 | 1.868E-07 | -1.47722E-06 | 2.843E-06 |
| p1 | 2.70813E-05 | 4.945E-05 | -1.16266E-05 | 1.581E-05 | -2.36032E-04 | 3.397E-05 |
| p2 | 2.21556E-04 | 4.614E-05 | -7.94448E-06 | 1.549E-05 | 6.39041E-05 | 3.324E-05 |
| b1 | 2.04594E-04 | 1.098E-04 | 1.57694E-05 | 9.723E-05 | -1.21875E-04 | 1.053E-04 |
| b2 | -1.50908E-04 | 1.314E-04 | -6.95587E-05 | 1.112E-04 | 9.86266E-05 | 1.215E-04 |

*Acquisition of Digital Image of Meandering Flume*

The geometry of the meandering flume along with site constraints present several difficulties when designing the photogrammetric configuration. With digital photography, it is important to obtain adequate exposed images of the desired object using a wide range of pixel radiance values with good contrast. The provision of adequate illumination and selection of appropriate camera exposure settings is therefore critical.

According to Tahar and Ahmad (2012), the main thing that should be considered is pixel size when dealing with digital images. Pixel size will determine the smallest coverage of an area or of the object. The size of pixel engages a few aspects such as the number of pixel for object image, length of an object in real measurement, focal length of the sensor and flying height during the capturing of the images. Furthermore, each digital camera has a different pixel size and it must be calculated during flight planning phase. Pixel size will determine the ground coverage area that was covered by one digital image. The ground coverage area of the images from the digital camera could also be determined by multiplying the scale of the photography with the dimension of the digital image.

The meandering flume is located in a building and the existing lighting is adequate for photography. A moveable gantry across the meandering flume provided an ideal platform to position the digital cameras. The cameras were mounted 1.6m above the bed and were separated from each other by a distance of 0.5m. Such geometry provided approximately 70% overlap between the images with a base to distance ratio of approximately 1:3. The photogrammetric digital normal case was used, in which the camera base is parallel to the object and the camera axes intersect the object plane orthogonally. Images acquired from such geometry allow efficient coverage of an area and provided approximately 60% overlap between the images. Fig.7 depicts a strip of the digital photograph acquired using the SLR digital camera.

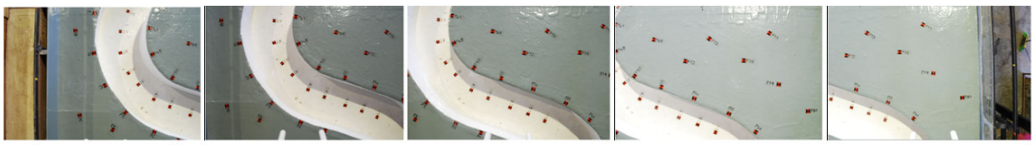


Fig.7: A strip of a digital photograph of the meandering flume

*Digital Image Processing*

All the digital images of the meandering flume were processed using Erdas Imagine software. Focal length, principal point offset (Xp, Yp) and the physical size of each pixel in the X and Y directions were an essential requirement for initial definition of primary inner orientation of the sensors in OrthoBASE Pro. Once image pyramid layers were generated, the point measurement tool was used to measure each target manually. The OrthoBASE triangulation algorithm, which implements a standard bundle adjustment, was then activated. It was essential to perform the procedures of interior, as well as exterior orientation and measure the control points before performing the triangulation process.

The GCPs are used to perform the aerial triangulation in order to produce a stereoscopic model. All GCPs were registered during exterior orientation. In the software, the 3D stereoscopic model was set up within a short period of time. The measurement of the image locations of the photogrammetric digital control points was performed manually. After the measurement of the first pair of digital images, the "automatic (x, y) drive" setting is activated to automate the measurement of corresponding points appearing on multiple frames. The digital images were added and measured in stereo pairs and the block triangulation tool was executed repeatedly to construct a valid block which was free of gross errors.

The next step was to run the bundle adjustment. This step reconstructs the geometry of the block and provides XYZ locations for the measured points from the previous step. Blunders (mis-measured points) from the previous step may have to be removed to achieve a good result. During image processing, the accuracy was maintained by checking the value of RMSE. The value of RMSE must be less than 1.0 in order to obtain good results. RMS error is reported in pixels. The amount of RMS error that is tolerated can be thought of as a window around each source coordinate, inside which a re-transformed coordinate is considered to be correct (i.e., close enough to use). In this study, the highest RMS error tolerance of 0.17 was indicated by an SLR digital camera. It showed that the re-transformed pixel was 0.17 pixels away from the source pixel and still considered to be accurate. Acceptable RMS error is determined by the end use of the data base, the type of data being used, and the accuracy of the GCPs and ancillary data being used. Acceptable accuracy depends on the image area and the particular project.

After the adjustment succeeds with a good Root mean Square Error (RMSE), the results can be checked via a report file and the stereo pairs visually inspected in stereo. Y-parallax in stereo would indicate a problem with the adjustment. At this point, images have been triangulated and can start creating data products: terrain, orthophotos and 3D features. In this study, the triangulation summary for metric camera and non-metric digital camera are shown in Fig.8(a), 8(b) and 8(c), respectively.

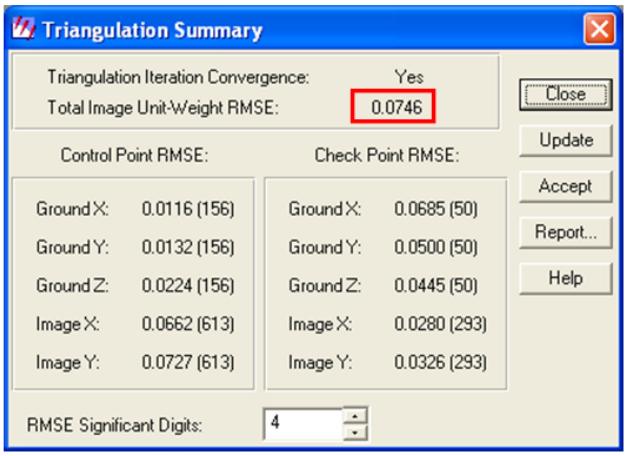



Fig.8(a): Aerial Triangulation (Nikon Coolpix S560)

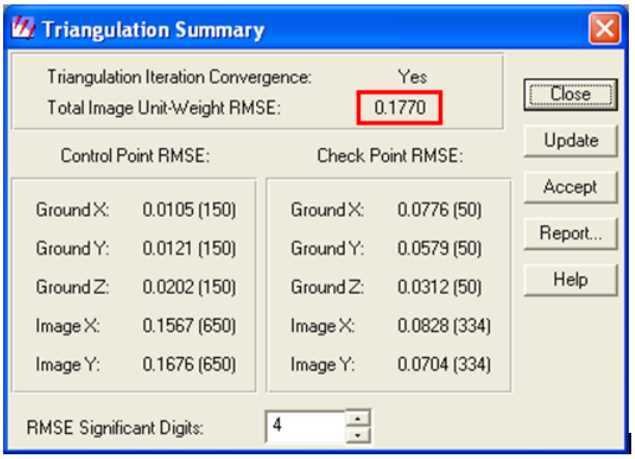



Fig.8(b): Aerial Triangulation (Nikon D60)

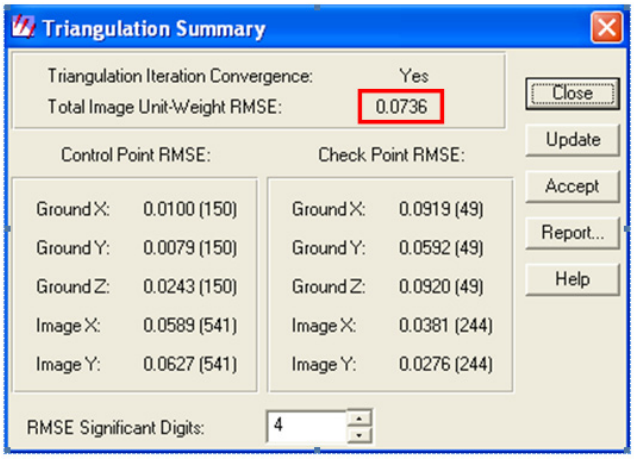



Fig.8(c): Aerial Triangulation (Rollei D30)

After performing AT, the footprint of the AT can be displayed. The foot prints of the digital photographs showing the location and names of all points (i.e., control points, check points and tie points) that participated in the adjustment. The distribution of GCP and CP for the three cameras can be viewed in Figures 9(a), 9(b) and 9(c). It shows the footprint of 100 images. It consists of 50 overlapping pairs based on the most efficient and reliable means of providing stereoscopic coverage to acquire a sequence of overlapping stereo pairs. These results indicate that the metric and non-metric camera can be used to acquire images of the meandering flume where the images were acquired in a controlled environment.

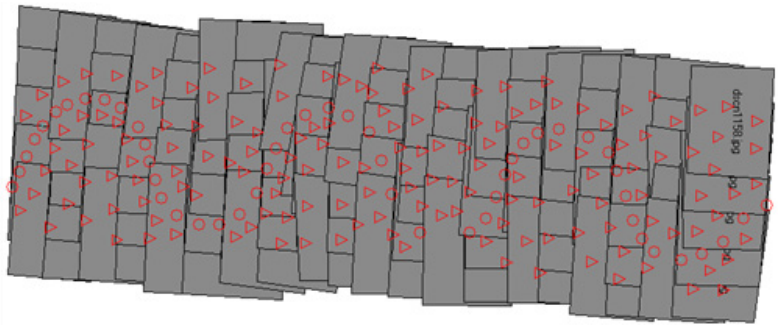


Fig.9(a): Footprint (Nikon Coolpix S560)

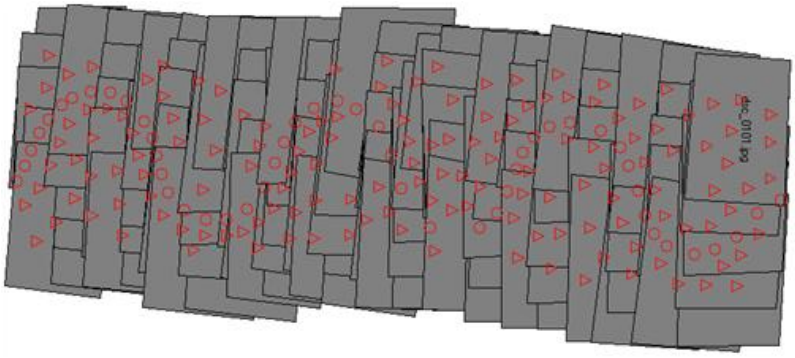


Fig.9(b): Footprint (Nikon D60)

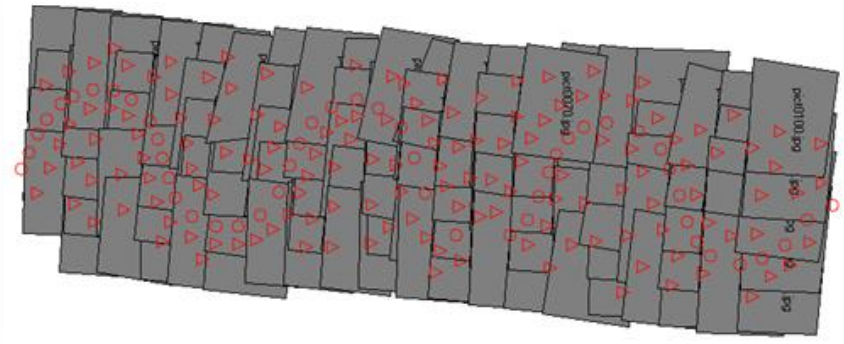


Fig.9(c): Footprint (Rollei D30)

## RESULTS

This study was conducted to assess digital cameras in mapping meandering flume (i.e.. controlled environment) and to determine the accuracy that could be achieved from close range photogrammetric technique. One of the important outputs produced after performing interior orientation, exterior orientation and aerial triangulation using Imagine OrthoBASE module is digital orthophoto of the entire flume.

*Orthophoto*

For orthophoto production, a "bare-earth" terrain model is necessary. Terrain is normally generated with DEM Terrain Extraction, which runs an auto-correlation algorithm to generate terrain points. Orthophotos are the data product created by the photogrammetric processing. Most commercial applications have orthorectification capability. After orthorectification, there may be a need to produce a final mosaic or a tiled ortho output. This is a routine procedure and standard tools within Ortho Rectification-Resampling allow the user to generate this outcome. Individual orthophoto was generated for each individual digital image. These individual orthophoto was then mosaicked together to create a composite orthophoto. Fig.10(a), Fig.10(b) and Fig.10(c) represent digital orthophotos of the entire flume produced from metric and non-metric digital imagery. Digital orthophoto only gives a two-dimensional view which generally involves x and y axis. According to Tahar and Ahmad (2012), the quality of digital orthophoto and DEM depends on the accuracy of GCP. If the quality of GCP is poor, therefore the result of digital orthophoto and DEM will be less accurate.

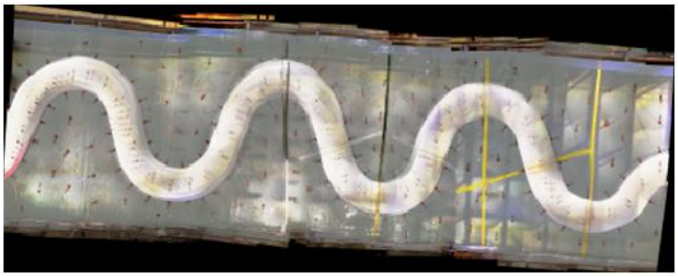


Fig.10(a): Digital orthophoto (Nikon Coolpix S560)

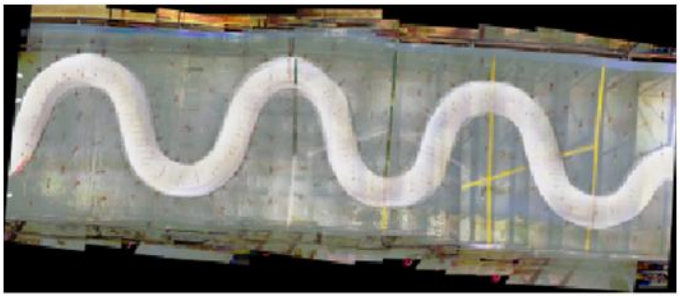


Fig.10(b): Digital orthophoto (Nikon D60)

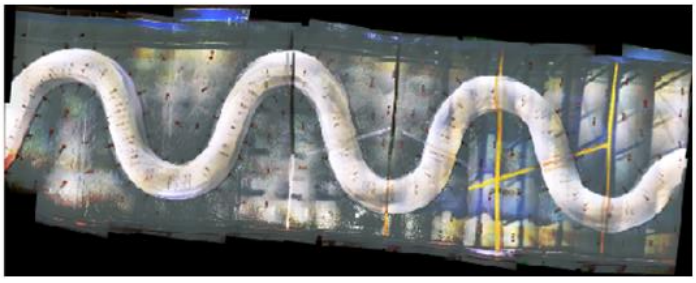


Fig.10(c): Digital orthophoto (Rollei D30)

## ANALYSIS

### *Visual Analysis*

Visual analysis is carried out by comparing the results of digital orthophotos generated from different types of camera. In this figure, it is clearly seen that the quality of the orthophoto produced from the metric digital camera was better than those proceeded by the other two non-metric digital cameras.

Based on the generated orthophotos, it can be concluded that there are no gaps or error in overlapping image regions but there are some technical difficulties of this matching process such as spatial continuity or edge matching and radiometric consistency. For spatial continuity, features that appear on more than a single image patch must be continuous. Channel flume must form a continuous meander line and show no jumps at the original photo edges where the images are connected. In the case of radiometric consistency, different photographs may have different contrast and brightness resulting from lack of uniform conditions during the photographic processing or from changes in illumination conditions.

### *Point Analysis*

Assessment of digital camera and evaluation of close range photogrammetry technique in mapping meandering flume were identified as main issues in this study. Determining the accuracy that could be achieved by analysing the residuals of the control points provides a first indication of the accuracy of the network restitution but it is not dependent. On the other hand, using independent check point data would provide a viable means of assessing accuracy. Point analysis was performed by calculating the Root Mean Square Error (RMSE).

In point analysis, the difference between the coordinates obtained from Total Station with the coordinates in Erdas Imagine is calculated to compute the RMSE of the orthophoto produced metric and non metric digital photographs. RMSE is the square root of the variance, known as standard error. However, an assumption is made whereby 3D coordinates obtained from Total Station becomes the most principal reference and it is used for comparison of 3D coordinates in deriving accuracy of measurement. The smaller the RMSE calculated, the higher the accuracy of orthophoto produced. Hence, the accuracy of orthophoto is influenced by the RMSE value. Based on the analysis, sub-meter accuracy (<1 metre) is obtained. The achievable accuracy for the entire flume using a number of check points are summarised in Table 2.

Table 2
RMSE of digital imagery orthophoto and Total Station

| Camera | Aerial Triangulation | GCP | RMSE(m) | Variance(m) |
|---|---|---|---|---|
| Nikon Coolpix | 89 GCPs<br>49 CPs | ΔX<br>ΔY<br>ΔZ | ±0.0159<br>±0.0486<br>±0.0432 | ±0.0003<br>±0.0024<br>±0.0019 |
| Nikon SLR D60 | 89 GCPs<br>49 CPs | ΔX<br>ΔY<br>ΔZ | ±0.0233<br>±0.0572<br>±0.0292 | ±0.0005<br>±0.0033<br>±0.0009 |
| Rollei D30 | 89 GCPs<br>49 CPs | ΔX<br>ΔY<br>ΔZ | ±0.0281<br>±0.0535<br>±0.0862 | ±0.0008<br>±0.0029<br>±0.0074 |

Based on Table 2, the accuracy of horizontal coordinates were low for the three types of camera but the accuracy can still be measured until centimetre level. Hence, it can be accepted in photogrammetric work. However, the accuracy of vertical coordinates was very high but it was constant for three sensors. Based on the table, it can be seen that the values of RMSE for metric and non-metric cameras were not significant. It might be affected by image matching algorithm that was used in the same software during image processing. The error might also be caused by image matching during image processing and motion movement during image acquisition.

Fig.11 shows the graph of RMSE versus ground control X, Y and Z for different types of digital camera. It was found that the residual error was not significant for all sensors. RMSE for ground control x and y was not much different but it was considerably different for ground control z , which represented the result for metric and non-metric cameras, respectively. The difference on ground control z might occur due to the effect of the automated tie point which implemented the image matching technique. The constant error of ground control z will be discussed further in a forthcoming paper and the method to improve the accuracy will be discovered.

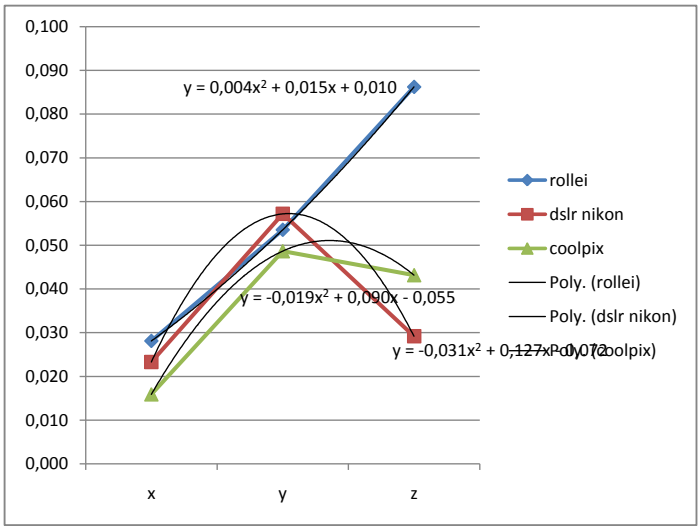



Fig.11: Root mean square error (x, y, and z)

## CONCLUSION

This study demonstrates that digital close range photogrammetry technique which utilizes a metric and non-metric digital camera is capable of mapping meandering flume successfully. The configuration of the photographs acquired has a direct impact upon the production of DTM and orthophoto. Besides, the achievable accuracy was found to be dependent upon other photogrammetric digital controls such as camera calibration and control point coordinates related to the grounds station network. When all of these errors are propagated and summed up following a valid error theory methodology, one can assess the spatial accuracy of the final product. Metric camera is already well known and proven (due to its correction parameters) to

be the best camera for the photogrammetric application. For accuracy assessment, it was found that the performances of the non-metric and metric digital cameras in terms of RMSE for ground control x and y were not much different but it was considerably different for ground control Z. Both types of digital cameras were capable of producing orthophotos of the meandering flume but the pictorial quality of orthophotos produced by non-metric cameras was not as good as that produced by the metric camera. This study shows in all certainty that good results could be produced from the metric digital camera. However, this study also demonstrates that the non-metric camera was also capable of producing good results compared to the metric digital camera. In conclusion, all the digital cameras used in this study have the potential to be used for this application and other various applications where accurate measurement is required. This study also provides a guideline for digital camera users to select the appropriate digital camera for any applications. Finally for any future work, it is hoped that this research will be expanded to determine the accuracy and cost for data acquisition in places with large areas (i.e., measurement of the real stream in an uncontrolled environment). Furthermore, it is also hoped that a variety of sensors can be explored in stream mapping. Among the factors that make close range photogrammetric technique suitable for mapping stream is that it is a practical approach in Malaysia and no in-depth study has been carried out using this method prior to this. Close range photogrammetric technique adopted in this study could be benefited by various agencies in Malaysia such as the Department of Irrigation and Drainage, the Department of Environment and other government agencies or the private sector whose work involves streams and rivers.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia, in conducting this study. The authors also would like to express their appreciation to Universiti Malaysia Kelantan (UMK) for their support in this study. Further, the authors would like to thank the Research Management Centre (RMC) of Universiti Teknologi Malaysia for their guidance and support in this study.

## REFERENCES

Ahmad, A., Busu, I., & Desa,G. (2003). Digital Close Range photogrammetry: Calibration of Different Digital Sensor Using Different Field Test and Application. *International Symposium and Exhibition on Geoinformation 2003*, 13-14 October. 2003, Shah Alam, Selangor, MALAYSIA.

AUSTRALIS (2001). *User manual*. Department of Geomatic Engineering, University of Melbourne, Australia.

Chandler, J. H., Buffin-Belanger, T., Rice, S., Reid, I., & Graham, D.J. (2003). The Accuracy of River Bed Moulding/Casting System and The Effectiveness of a Low-Cost Digital Camera for Recording River Bed Fabric. *The Photogrammatic Record*, *18*(103), 209-224.

Chandler, J. H., Lane, S. N., Shiono, K., & Rameshwaren, P. (2001). Measuring Flume Surface for Hydraulics Research Using a Kodak DCS460. *Photogrammetric Record*. *17*(97), 39–61

Chandler, J. H., Wackrow, R., Sun, X., Shiono, K., & Rameshwaren, P. (2008). Measuring Dynamic River Flooding and Surface by Close Range Digital Photogrammetry. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol.* XXXVII (B7) Beijing, China.

ERDAS. (2001). Imagine OrthoBASE User's Guide. Erdas Inc. 524 pages.

Fraser, C.S. (1996). Industrial Measurement Application in Close Range Photogrammetry and Machine Vision. Whittles Publishing, Caithness, Scotland, U.K

Geisler, T., Breitenstein, S., Hammer, A., Mayr, D., & Klasinc, R. (2003). Hydraulic Modeling-Mapping of River Bed. *XXX IAHR Congress*, AUTh, Thessaloniki, Greece.

Lane, S. N. (2000). The Measurement of River Channel Morphology Using Digital Photogrammetry. *Photogrammetric Digital Record*, *16*(96), 937-961.

Peipe, J., & Stephani, M. (2003). Performance Evaluation of a Megapixel Digital Metric Camera for Use in Architectural Photogrammetry. In *Proceedings of the XX International Congress for Photogrammetry and Remote Sensing*. Ancona, Italy, July 2003, pp. 259-262.

Samad, A. M., Ahmad, A., & Ishak, N. (2010). Assessment of Digital Camera for Mapping Stream using Close Range Photogrammetric Digital Technique. *2010 6th International Colloquium on Signal Processing & Its Application. (CSPA 2010),* 21-23 May 2010, Malaca City, Malaysia.

Sanz-Ablanedo, E., Rodriguez-Perez, J. R., Arias-Sanchez, P., & Armesto, J. (2009). Metric Potential of a 3D Measurement System Based on Digital Compact Cameras. *Sensors*, *9*, 4178-4194.

Shortis, M. R., Robson, S., & Beyer, H. A. (1998). Principal Point Behaviour and Calibration Parameter Models For Kodak DCS Cameras, *Photogrammetric Record. 16*(92), 165-186

Tahar, K. N., & Ahmad, A. (2012). A Simulation Study on the Capabilities of Rotor Wing Unmanned Aerial Vehicle In Aerial Terrain Mapping. *International Journal of Physical Sciences, 7*(8), 1300 - 1306.

Udin, W. S., & Ahmad, A. (2011). Calibration of High Resolution Digital Camera Using Self-Calibration Bundle Adjustment Method. *2011 IEEE 7th International Colloquium on Signal Processing and Its Applications* (CSPA), 4-6 March 2011, Penang, Malaysia.

Udin, W. S., Hassan, A. F., Ahmad, A., & Tahar, K. N. (2012). Digital Terrain Model Extraction Using Digital Aerial Imagery of Unmanned Aerial Vehicle. *2012 IEEE 8th International Colloquium on Signal Processing and Its Applications* (CSPA), 23-25 March 20102, Malacca, Malaysia.

Wackrow, R., Chandler, J. H., & Gardner, T. (2008). Minimising Systematic Errors in DEMs Caused by Inaccurate Lens Model. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVII(Part B5).*Beijing 2008.

Yakar, M., & Yilmaz, H. M. (2008). Using In Volume Computing of Digital Close Range Photogrammetry, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B3b.* Beijing 2008.

Pertanika J. Sci. & Technol. 22 (1): 139 - 152 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Correlation of Electrical Resistivity with Some Soil Parameters for the Development of Possible Prediction of Slope Stability and Bearing Capacity of Soil using Electrical Parameters

**Syed Baharom Syed Osman*, Mohammad Nabil Fikri and Fahad Irfan Siddique**

*Civil Engineering Department, Universiti Technologi Petronas, Bandar Seri Iskandar, 31750 Tronah, Perak, Malaysia*

## ABSTRACT

The long term objective of this research is to look into the possibility of replacing soil strength parameters such as cohesion and angle of friction with electrical resistivity value for the purpose of computing among others, factor of safety in slopes or bearing capacity of soil. This paper however is limited to the investigation of correlation between electrical resistivity with some selected soil parameters. Electrical resistivity tests, using a basic multi meter, steel moulds and other related equipment, were conducted in the laboratory on soil samples with variations in soil type, compaction energy and moisture content. The samples consisted of predominantly clay, silt and sandy size particles and were compacted in a 100 x 100 mm square mould, while the corresponding electrical resistivity tests were carried out using the disc electrode method in accordance to BS 1377. The values of the electrical parameters such as voltage, current and resistance, with the corresponding value of soil parameters such as cohesion, angle of friction and moisture content, were measured and recorded. The results of the tests produced some initial crude relationships between electrical resistivity and the selected soil parameters. The strongest correlation between electrical resistivity and angle of internal friction, $\varphi$, was obtained from the clay size samples with $R^2$ of 0.824, while the maximum correlation between electrical resistivity and moisture content again was obtained through the clay samples with $R^2$ of 0.818. From the other results and graphs analyzed, some consistencies and specific trends of behaviour observed gave some early indications that a more detail and precise correlation between electrical resistivity and soil strength parameters could be very well possible in future

*Keywords:* Electrical resistivity, correlations, shear strength, moisture content, compaction

*Article history:*
Received: 5 July 2011
Accepted: 13 September 2011

*E-mail addresses:*
sybaharom@petronas.com.my (Syed Baharom Syed Osman),
nblfkri@gmail.com (Mohammad Nabil Fikri),
fahad_03mn@hotmail.com (Fahad Irfan Siddique)
*Corresponding Author

ISSN: 0128-7680 

## INTRODUCTION

In all occasions involving the aspects of design and checking of geotechnical structures, the shear strength parameters such as cohesive ($c$) and angle of internal friction ($\varphi$) are perhaps the two most important required parameters. The calculation of skin friction and bearing capacity for shallow and deep foundations to the calculation of Factor of Safety (FOS) for slope stability (including designing of retaining walls, $c$ and $\varphi$ along with some other parameters) is normally obtained through prescribed methods before the actual design and checking. In general practice, soil investigation (SI) incorporating bore hole sampling will perhaps produce the most reliable values of the relevant soil parameters for calculation purposes. However, bore hole sampling is in general time consuming and very expensive. Conventional methods of soil analysis mostly require disturbing soil, removing soil samples and analyzing them in laboratory, where electrical geophysical methods allow rapid measurement of soil electrical properties such as electrical resistivity and conductivity directly from soil surface to any depth without soil disturbance (Pozdnyakov & Pozdnyakova, 2002).

Taking the case of standard operating procedure (SOP) for hillside development for example, among the critical element is to check the stability of the slopes which can be done by calculating FOS. For a regular checking and calculation of FOS in a certain stretch of slopes for the purpose of identification of risk/danger, bore hole sampling would not be practical due to the above mentioned reasons. This is because many bore holes are required to check the factor of safety at different locations on the stretch of slope under consideration in order to determine the risk/hazard. Hence, an alternate quick and less expensive method of assessing FOS is essential so as to enable rapid and extensive measurements and calculation of FOS at different points in slopes. Therefore, the long-term objective of this research is to produce the correlations between electrical resistivity with especially strength parameters such as $c$ and $\varphi$ which will eventually make it possible for a quick assessment for FOS of any slope on initial and regular bases. Any slope could be checked and if the FOS falls within a certain range of a "prescribed value" which indicates high risk, a further confirmation of the FOS will then be conducted if needed through the actual soil boring sampling or any other extensive method. The correlation will also enable designing and checking of any geotechnical structure, as mentioned earlier.

The work of some researchers in the past and recent years has included correlation of electrical resistivity with various soil properties. For instance, Hassanein *et al*. (1996) studied the relationship of electrical resistivity in a compacted clay with hydraulic conductivity and some index properties. An earlier research had suggested the possible correlation of electrical resistivity with hydraulic conductivity which served as a non-destructive mean of evaluating the quality of compacted soil liner (Kalinski & Kelly, 1994). Meanwhile, an extensive work by Pozdnyakova *et al*. (2001) looked into the effects of electrical resistivity in different soil types with varying water contents, humus contents, salt contents and several other parameters. Other researchers have also studied the estimation of water content of soil using electrical resistivity (Kalinski & Kelly, 1993). Others have used the knowledge of electrical resistivity to estimate liquefaction of soil (Ronald & Ronald, 1982), detecting and locating geomembrane failures (Schulz *et al*., 1984), and estimation of soil salinity for agricultural activities (Shea & Luthin, 1961).

In spite of the many research done as mentioned above, none has actually looked into the aspects of correlation of electrical resistivity with strength properties such as $c$ and $\varphi$. The general approach behind this quick assessment system is to eliminate the usage of physical soil parameters such cohesion ($c$), internal frictional angle ($\varphi$) and unit weight ($\gamma$), as is currently being practice for the calculation of FOS or bearing capacity of soil and replacing these physical parameters with their correlated electrical parameters which include resistivity, conductivity and voltage. Therefore, the future simplified method at site will require a few steel rods implanted in the soil/slope serving as the electrodes, a reel of electrical wires and an existing multimeter to generate the factor of safety or bearing capacity that is calculated through a set of empirical formula, charts, and graphs to be developed in several phases of research.

*Electrical Resistivity Measurement*

The electrical resistivity of soils varies between different geological materials and soil types, and is dependent on many factors which will be explained later in this paper. Resistivity measurement and method of soil investigation can thus be used to identify layers of zones with different electrical properties.

The electrical resistivity of the soil is determined by measuring the resistance between two points in the soil and this is done by measuring voltage across a pair of electrodes by transmitting a controlled DC or AC current ($I$) between two electrodes pushed into the ground, while measuring the potential ($V$) between two other electrodes. The setup for the measurement of electrical resistivity is shown in Fig.1.

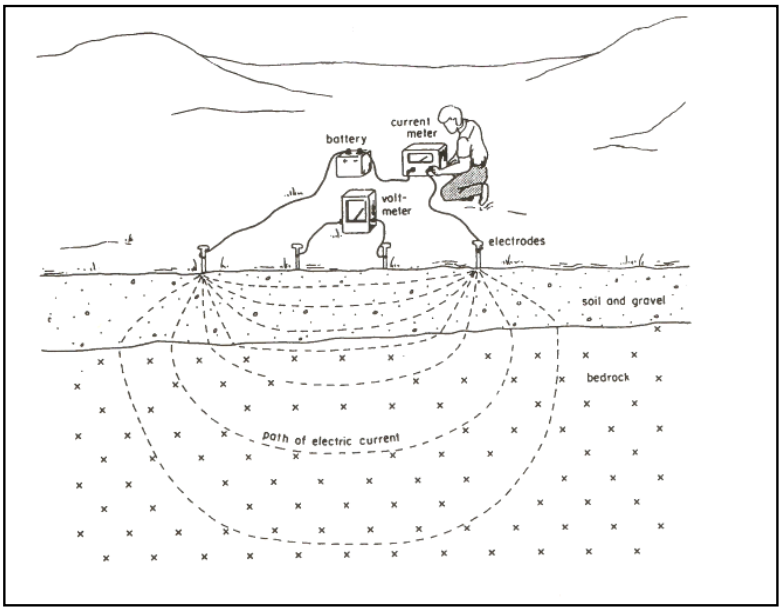



Fig.1: Principle set-up for direct current resistivity measurement (Robinson & Coruh, 1988)

The resistance ($R$) is calculated using Ohm's Law as given in Equation (1).

$$R = \frac{V}{I} \tag{1}$$

where, $V$ is voltage (V) and $I$ is current (amp).

For the case of a pair of electrodes in homogeneous, isotropic conducting media, the relationship between resistance and resistivity is linear and the material resistivity ($\rho$) can then be defined as in the following Equation (2):

$$\rho = 2\pi RL \tag{2}$$

where, $L$ is the length (m) between electrode and $R$ is resistance (Ω).

In the actual field measurement of electrical resistivity, there are many different kinds of electrode arrays or configuration that one could adopt. Some of the typical electrode arrays are Wenner, Schlumberger, Dipole-dipole and Pole-pole. In this research, however, disc electrode method in accordance to BS 1377 was adopted to enable undisturbed or disturbed samples of soil to be measured in the laboratory. By using this disc electrode method of measurement, the resistivity of the soil ($\rho$) in .m is determined by the formula given in Equation (3).

$$\rho = \left(\frac{A}{L}\right) R \tag{3}$$

where, $A$ is cross sectional area (m2) of the sample, $L$ is length (m) and $R$ is resistance (Ω).

*Factors Affecting the Electrical Resistivity of Soil*

For most common minerals forming soils and rocks, the resistivity is high in a dry condition and therefore the resistivity of soils and rocks generally depends on the amount and type of water in the pore spaces and fractures. Meanwhile, the connection between cavities and fracture is also an important factor in the final value of resistivity. The amount of water in a material depends on porosity, which may be divided into primary and secondary porosity. Primary porosity consists of pore spaces between the mineral particles, and occurs in soils and sedimentary rocks. Secondary porosity consists of fractures and weathered zones, and this is the most important porosity in crystalline rocks such as granite and gneiss (Instruction manual for LUND, 1999).

However, the basic mechanism affecting conductivity in moist soils and water bearing rocks occurs as a result of the movement of ions, while the ability to transmit ions is governed by electrical resistivity which is a basic property of all materials (Hassanein *et al*., 1996). Besides being dependant to the amount and type of water and porosity, electrical resistivity also depend on other properties such as type of material, particle shape and orientation, mineralogy, as well as the amount of clay content and electrical resistivity of the pore fluid. The presence of clay minerals strongly affects the resistivity of sediments and weathered rock. This is due to the fact

that clay minerals are electrically conductive particles having the ability to absorb and release ions and water molecules on the surface through an ion exchange process (Parasnis, 1986).

Therefore, it is worthwhile to mention here that in clean sands and gravels, electrical conduction occurs primarily in the pores (Jackson, 1975), while in clayed soils and clay-bearing rocks, electrical conduction occurs in the pores and on the surfaces of electrically charged particles (Rhodes *et al*., 1976). Mitchell (1993) supports the above statements by adding that surface conductance in clays can be a significant factor affecting the bulk electrical resistivity of soil. Other factors which indirectly affect the electrical resistivity are frequency of the current, geometry, spacing and type of electrodes used (Erzin *et al*., 2010). Temperature also plays an important role in the electrical resistivity of soil in the sense that increasing the temperature increases the mobility of the ions and this decreases the electrical resistivity of soil (Hassanein *et al*., 1996).

The statements above exhibit the complexities in correlating resistivity with the different factors associated with soil, rocks and pore fluid. However, one could start off with the variations of resistivity and some common types of material found in many tables as an initial assistance in determining what material one is working with. An example is given in Table 1 below.

TABLE 1
Variations of resistivity with some common materials (Jackson, 1975)

| Material | Ohm Meter |
|---|---|
| Clay and marl | 1 to 100 |
| Loam | 5 to 50 |
| Top soil | 50 to 100 |
| Clayey soils | 100 to 500 |
| Sandy soils | 500 to 5000 |
| Typical mine water | 1 to 10 |
| Typical surface water | 5 to 50 |
| Shale | 10 to 80 |
| Limestones | 80 to 1000 |
| Sandstones | 50 to 8000 |
| Coal | 500 to 5000 |

## MATERIALS AND METHODS

### *Soils*

Three types of soil were used for this research. The soils were purchased from a soil processing factory according to their grades, namely, KM80, KM200 and L2B20. Brief specifications for each are given in Table 2.

From Table 2, it could be seen that the three soil types used were predominantly of kaolinite and quartz mineralogy and the main variation is in the grain sizes. In this research, the author designated grade KM80 as clay, KM200 as silt and L2B20 as sand, which are in accordance to their respective predominant particle sizes.

TABLE 2
Specifications of the soil types used

| Type/Grade | Particle size | pH | Predominant Mineralogy | PI | Predominant particle size |
|---|---|---|---|---|---|
| KM80 | 1.0-3.0 µm | 3.5-5.5 | kaolinite | 30.6 | clay |
| KM200 | 44-250 µm | 3.5-5.5 | kaolinite | 10.6 | silt |
| L2B20 | 300-2000 µm | 3.5-5.5 | quartz | 3.8 | sand |

*Basic Procedure and Tests*

All soil samples were stored in airtight containers so as to reduce the absorption of moisture. Basic tests comprising of sieve analysis, plastic limit and liquid limit tests were then conducted to ascertain some basic properties of the soil samples. Following this, the samples were then prepared for the second phase tests, which consisted of the electrical resistivity and the direct shear tests.

*Electrical Resistivity Tests*

The apparatus used for this electrical resistivity test consisted of the following:

1. Standard 100 x 100 mm concrete cube mould
2. Soil mixer
3. Standard Proctor hammer
4. Two 100 mm aluminium electrodes
5. 200 volts DC power supply
6. Hand held multimeter
7. Other basic apparatus

For every specimen, 3 kg of soil were mixed with a certain amount of distilled water according to the percentage of moisture content required that ranged between 10 % - 45 % as shown in Table 3. Mixing was done by means of a soil mixer and the samples were then left aside for at least 24 hrs in the mixing bowl wrapped with plastic.

Prior to the compaction process, the internal perimeter of the mould was lined with a thick plastic material for easy removal of the specimen once the mould had been disassembled. The specimens were then compacted directly in a 100 x 100 mm square mould in three equal layers using the standard proctor hammer that delivered blows ranging from 15 to 45 blows per layer. The procedure for compaction is the same as prescribed in BS 1377, except for the fact that the mould is square instead of the standard round mould and the number of blows varies from 15 – 45 blows as mentioned earlier. The mould was disassembled upon completion of compaction and the specimen cubes were placed between two circular aluminium electrodes for the purpose of determining electrical resistivity using the disc electrode method according to BS 1377. The specimens, along with the aluminium discs, were connected to both the negative and positive

terminals of a DC power supply and also connected to a multimeter, where an initial potential with varying voltages from 30V, 60V, and 90V were applied (Fig.2). The resulting values of the current in miliampere were then recorded. The electrical resistant and resistivity of the samples were calculated from Equations (1) and (2) respectively.

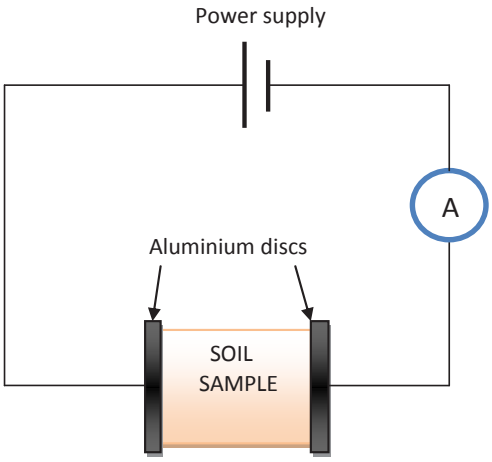



Fig.2: Electrical resistivity test set-up

The cube specimens were then sliced into 3 portions, where they were placed in a direct shear box assembly with subsequent loadings of 10 kg, 20 kg and 30 kg. Finally, the three Mohr circle diagrams from each sample were constructed and both the cohesion and angle of friction were also recorded. A summary of all the 48 tests conducted with their variations is given in Table 3 below.

For example, in tests 1 to 3, the weight of dry soil used was 3.0 kg for each sample. The soil samples were mixed with 0.75, 0.3 and 0.75 kg of water to produce the corresponding moisture content of 25%, 10% and 25% for sand, silt and clay sample respectively. The no. of blow for in tests 1 to 3 was set to 15 blows per layer. The same process was repeated for the rest of the tests in Table 3.

TABLE 3
A summary of the tests for all the samples

| Test No. | Wt. of dry soil (kg) | Moisture content for sand samples (%) | Moisture content for silt samples (%) | Moisture content for clay samples (%) | No. of blows per layer |
|---|---|---|---|---|---|
| 1-3 | 3.0 | 25 | 10 | 25 | 15 |
| 4-5 | 3.0 | 25 | 10 | 25 | 25 |
| 6-9 | 3.0 | 25 | 10 | 25 | 35 |
| 10-12 | 3.0 | 25 | 10 | 25 | 45 |
| 13-15 | 3.0 | 30 | 15 | 30 | 15 |
| 16-18 | 3.0 | 30 | 15 | 30 | 25 |
| 19-21 | 3.0 | 30 | 15 | 30 | 35 |

TABLE 3 (*continue*)

| Test No. | Wt. of dry soil (kg) | Moisture content for sand samples (%) | Moisture content for silt samples (%) | Moisture content for clay samples (%) | No. of blows per layer |
|---|---|---|---|---|---|
| 22-24 | 3.0 | 30 | 15 | 30 | 45 |
| 25-27 | 3.0 | 35 | 20 | 35 | 15 |
| 28-30 | 3.0 | 35 | 20 | 35 | 25 |
| 31-33 | 3.0 | 35 | 20 | 35 | 35 |
| 34-36 | 3.0 | 35 | 20 | 35 | 45 |
| 37-39 | 3.0 | 40 | 25 | 40 | 15 |
| 40-42 | 3.0 | 40 | 25 | 40 | 25 |
| 43-45 | 3.0 | 40 | 25 | 40 | 35 |
| 46-48 | 3.0 | 40 | 25 | 40 | 45 |

## RESULTS AND DISCUSSION

### *Variation of Electrical Resistivity with Moisture Content*

The curves of moisture content against electrical resistivity plotted in Fig.3, Fig.4 and Fig.5 are the combination of points resulting from the tests where different moisture contents and different blows were used. For simplicity reason, the effect of blows on electrical resistivity will not be discussed in detail here other than the fact that different numbers of blow were found for the same moisture content that resulted in different values of electrical resistivity for all soil types (sand, silt and clay). This is due to the changes of porosity in the soil samples. It was also found that the effect of the number of blows (which causes changes in electrical resistivity) is much more prominent in samples with lower moisture content rather than the samples with higher moisture content.

Fig.3 shows that the electrical resistivity for the sand samples, regardless of the number of blows, ranges between 188 ohm.m at 40% moisture content to the maximum 1108 ohm.m at 25% moisture content, which is within the range of the specified electrical resistivity value for sandy material (see Table 1). As for the silt samples, Fig.4 shows that the electrical resistivity values range between 78 ohm.m at 20% moisture content and 881 ohm.m at 15% moisture content, whereas the clay samples exhibit a range of electrical resistivity between 9.31 ohm.m to 37.7 ohm.m at the moisture content of 40% and 25% respectively (see Fig.5). Once again, the values of electrical resistivity for both silt and clay samples are within the specified range.

All the curves in Figures 3, 4 and 5 indicate strong correlations between moisture content and electrical resistivity for all types of soil. This is in agreement with the findings from many past researchers who found that moisture content and ionic content in pore fluids are more important than the conductivity of the constituent mineral grain of the soil or rock in governing resistivity of the sample (Kizlo & Kanbergs, 2009). By comparing the curves obtained for the sand, silt and clay samples, it is obvious that the moisture content for clay has the strongest correlation with electrical resistivity, i.e. with a regression coefficient, $R^2$, of 0.818, while silt exhibits the lowest correlation with the $R^2$ value of 0.694. The higher correlation of electrical resistivity against moisture content in clay could be well contributed by the smaller range

of electrical resistivity values for clay compared to sandy particles which limit the variation in the values of electrical resistivity in clay. In addition, the homogeneity of the sizes of the sand particles which varies from 300 – 2000 µm might again affect the range of measured electrical resistivity values and hence the correlation. However, it should be noted that the weak correlation in silt might be due to the fact that lower moisture contents (10% to 25%) were used in the silt samples and therefore, further experiments should be conducted to verify this. The combined curve for silt and clay shown in Fig.6 revealed an even better correlation, suggesting that the samples with fine grain soils produced better correlation of moisture content with electrical resistivity.

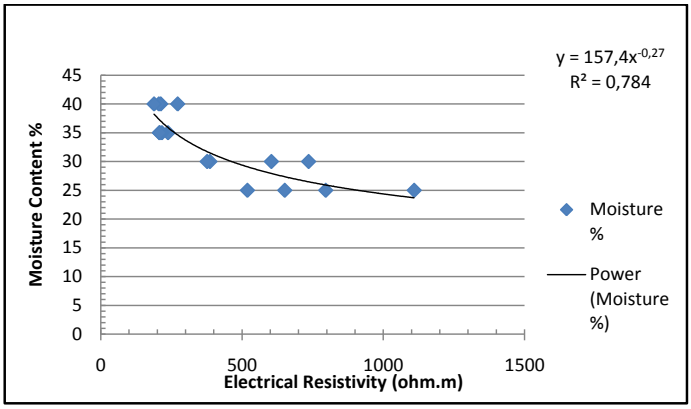



Fig.3: Moisture content vs. electrical resistivity for sand

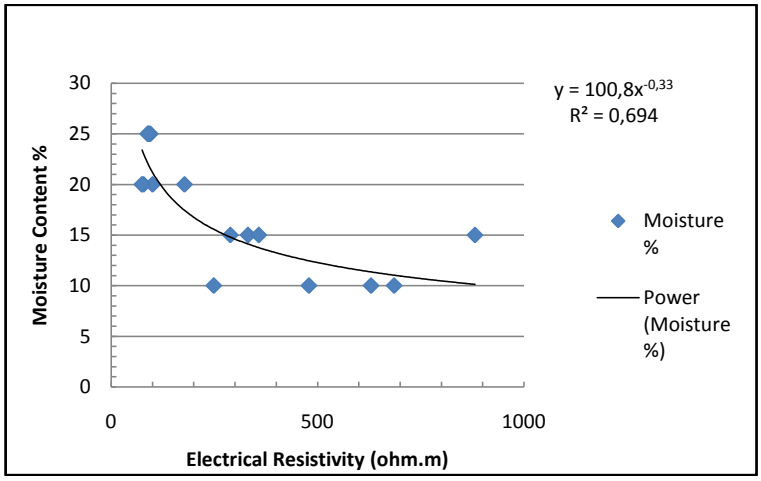



Fig.4: Moisture content vs. electrical resistivity for silt

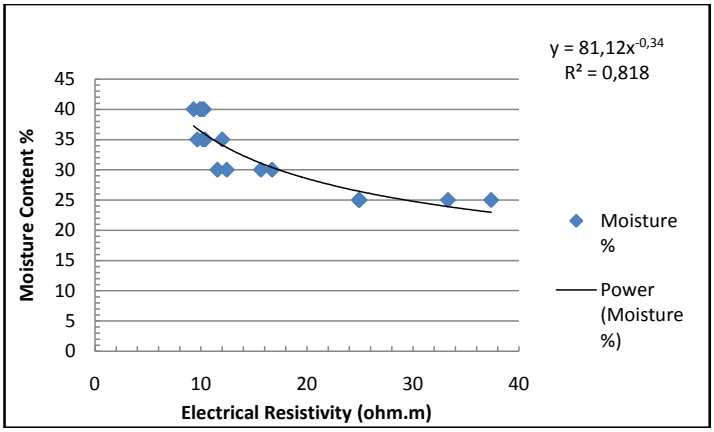



Fig.5: Moisture content vs. electrical resistivity for clay

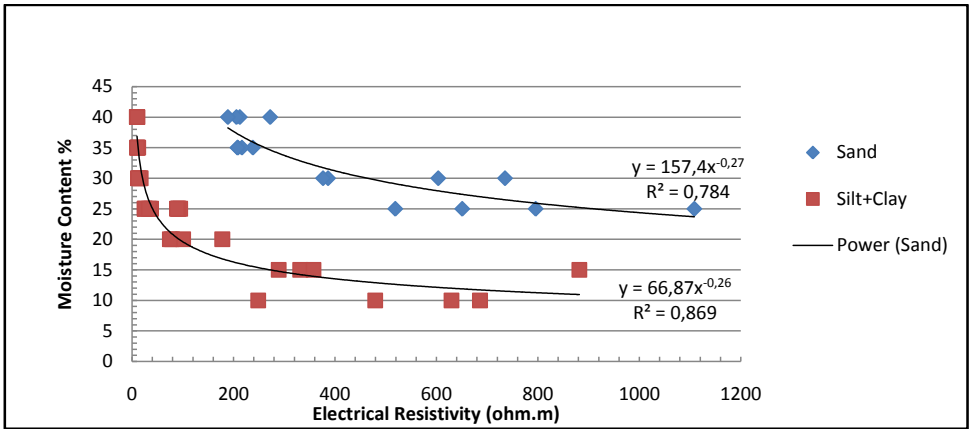



Fig.6: Moisture content vs. electrical resistivity for sand and silt+clay

*Variation of Electrical Resistivity with Angle of Friction*

All the values of angle of frictions $\varphi$ presented in the following figures were obtained from the small shear box tests conducted on the remoulded soil samples. Along with the angle of frictions, $\varphi$, cohesive values, $c$, were also obtained. However, only the results obtained for $\varphi$ are presented in this paper.

Fig.7, Fig.8, Fig.9 and Fig.10 show the correlations between the angle of friction and electrical resistivity. It is interesting to note that the trend of the curves indicates that the values of internal friction $\varphi$ increase with the increase in the electrical resistivity. It is also interesting to find out that the correlation between $\varphi$ and electrical resistivity for clay soil shown in Figure 9 gives the strongest correlation with a coefficient of regression, $R^2$, of 0.824, i.e. with the silt soil type having the least $R^2$ value of 0.012. The combined points for all the soil types are then plotted in Figure 10, with the trend of increasing $\varphi$ with the increase in the electrical resistivity persists, while the value of regression, $R^2$, is 0.338. Nevertheless, this paper does not attempt to hypothesize the reasons of such a relationship.

If what was obtained here is the true representation of the relationship between $\varphi$ and electrical resistivity, then further tests would need to be carried out to establish the governing mechanisms. Note that it is very important to establish a strong and reliable correlation between electrical resistivity and strength parameters such as $\varphi$, which is one of the main parameters used in the calculation of factor of safety (FOS) in slopes and bearing capacity of soils. A particular measurement of the electrical resistivity in the field if correlated correctly could produce a reasonable value of FOS for slope stability assessment and bearing capacity of soils for foundation design. Therefore, the related factors to look for are probably in the fabric structure or particle arrangement of the fine particles and the reduction of porosity in coarse particles, where both factors contribute to the strength of the soil samples and affect the ability in the transmission of fluid or ions in the soil, which will in turn affect electrical resistivity. Robain *et al*. (2003) and Ozcep *et al*. (2005) pointed out that solid soil components are generally insulators but electrical conductivity and resistivity lie in the fluid content in both the macro and micro voids. Porosity generally affects the pore size and volume of air voids, which will in turn increase or decrease the degree of saturation. Nearly saturated pores form bridges between the particles and greater particle-to-particle contact (Sadek, 1993). Thus, lower and higher electrical resistivities associated with $\varphi$ are the results of decreasing or increasing the electrical conductivity or resistivity in the pores and along the solid surface.

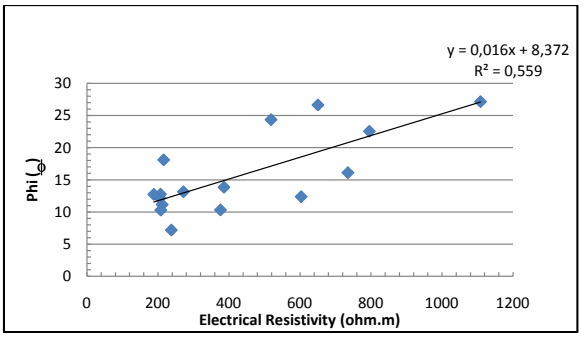



Fig.7: Angle of friction (Phi) vs. electrical resistivity for sand

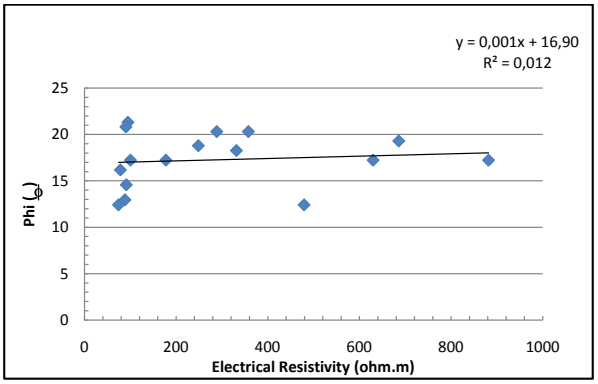



Fig.8: Angle of friction (Phi) vs. electrical resistivity for silt

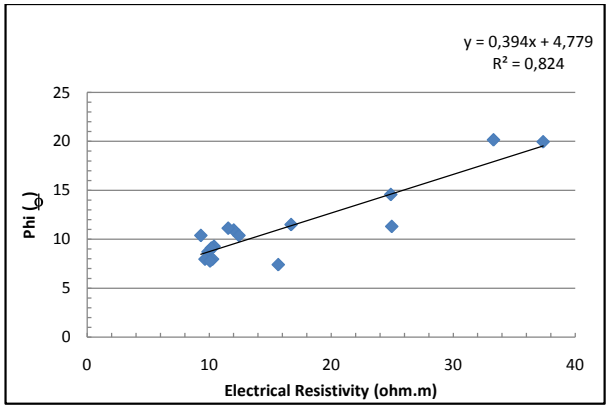



Fig.9: Angle of friction (Phi) vs. electrical resistivity for clay

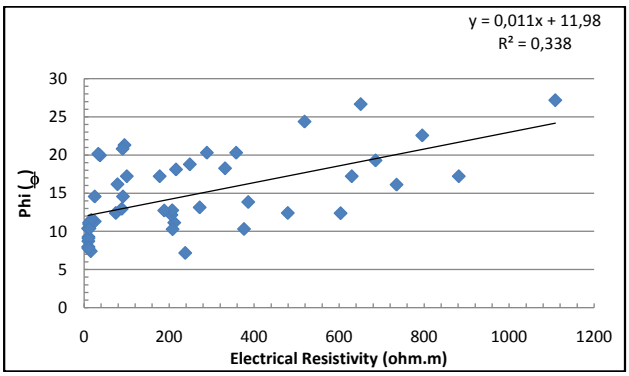



Fig.10: Angle of friction (Phi) vs. electrical resistivity for all soil samples

## CONCLUSION

The trend and reliability of relationships between moisture content and angle of friction with electrical resistivity were established from this research. Relationship between moisture content and electrical resistivity shows that higher moisture content causes electrical resistivity to be lowered with the strongest coefficient of regression, $R^2$ was obtained in the clay soil samples. On the other hand, relationship between $\varphi$ and electrical resistivity reveals that higher value of $\varphi$ caused electrical resistivity to be higher with again the strongest $R^2$ was from the clay samples. The mechanism which governs the correlations mentioned above needs to be understood and verified through further tests which hopefully will pin point to the factors contributing to such relationship or trend. For this purpose, the author suggested to initially focus on the aspects of structural arrangement in fine grained soils and porosity in course grained soils, both of which affects the transmission of ions which has direct bearing on the value of electrical resistivity.

## REFERENCES

Erzin, Y., Rao, B. H., Patel, A., Gumaste, S. D., & . Singh, D. N. (2010). Artificial Neural Network Models For Predicting Electrical Resistivity of Soils from Their Thermal Resistivity. *International Journal of Thermal Science*, *49*, 118-130.

Hassanein, Z., Benson, C., & Blotz, L. (1996). Electrical Resisitivity of Compacted Clay. *Journal of Geotechnical Engineering*, *122*(*5*)(May Issue), 397-406.

*Instruction Manual for LUND Imaging System* (1999). ABEM Instrument AB, Bromma, Sweden. Chapter 2, 1-21.

Jackson, P. (1975). An Electrical Resistivity Method for Evaluating the In-situ Porosity of Clean Marine Sands. *Marine Geotechnol*., *1*(2), 91-115.

Kalinski, R. J., & Kelly, W. E. (1993). Estimating Water Content of Soils From Electrical Resistivity. *Geotechnical Testing Journal*, *16*(3), 323-329.

Kalinski, R. J., & Kelly, W. E. (1994). Electrical Resistivity Measurements for Evaluating Compacted Soil Liner. *Geotechnical Testing Journal*, *120*(2451) 451-457.

Kizlo, M., & Kanbergs, A. (2009). The Causes of the Parameters Changes of Soil Resistivity. *The 50th International Scientific Conference, Power and Electrical Engineering*, 43-46.

Mitchell J. (1993). *Fundamentals of Soil Behavior (2nd ed.)*, New York: John Wiley and Sons.

Ozcep, F., Asci, M., Tezel, O., Yas, T., Alpaslan, N., & Gondogdu, D. (2005). Relationships Between Electrical Properties (in Situ) and Water Content (in The Laboratory) of Some Soils in Turkey. *European Geosciences Union General Assembly*, Wien, Austria.

Parasnis, D. S. (1986). *Principles of Applied Geophysics*. Chapman & Hall.

Pozdnyakova, L., Poznyakov, A., & Zhang, R. (2001). Application of Geophysical Methods of Evaluate Hydrology and Soil Properties in Urban Areas. London, UK. *Urban Water*, *3*, 205-216.

Pozdnyakov, A., & Pozdnyakova, L. (2002). *Electrical Fields and Soil Properties*. In *17th WCSS*, Thailand, Paper 1558, 1-11.

Robain, H., Camerlynck, C., Bellier C., & Tabbagh, A. (2003). Laboratory Measurements of Electrical Resistivity Versus Water Content on Small Soil Cores. *Geophysics Research Abstract*, 5.

Robinson, E. S., & Coruh, C. (1988). *Basic Exploration Geophyics*. John Wiley & Sons.

Rhodes, J., Raats, P., & Prather, R. (1976). Effect of Liquid Phase Electrical Conductivity, Water Content, and Surface Conductivity on Bulk Soil Electrical Conductivity. *Soil Science Society of American Journal, 40*, 651-655.

Ronald, A. E., & Ronald, C. G.(1982). Electrical Resistivity Used To Measure Liquefaction of Sand. *Journal of Geotechnical Engineering*, *108*, *GT5*, 779-782.

Sadek, M. (1993). *A Comparative Study of the Electrical and Hydraulic Conductivities of Compacted Clay*. *Phd Thesis*, Dept. of Civil Engineering, University of California at Berkeley, Berkeley, California.

Schultz, D. W., Duff, B. M., & Peters, W. R. (1984). Performance of An Electrical Resistivity Technique for Detecting and Locating Geomembrane Failures. *International Conference on Geomembrane*, Denver, USA, 445-449.

Shea, P. F., & Luthin, J. N. (1961) An Investigation of The Use of The Four Electrode Probe For Measuring Soil Salinity In Situ. *Soil Science*, *92*, 331-339.

Pertanika J. Sci. & Technol. 22 (1): 153 - 161 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Chemically Modified Sago Waste for Oil Absorption

**Zainab Ngaini*, Rafeah Wahi, Dayang Halimatulzahara and Nur An-Nisaa' Mohd Yusoff**

*[1]Department of Chemistry, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia*

## ABSTRACT

Oil pollution remains a serious concern especially in Malaysia. Many strategies have been employed to overcome oil pollution. In this research, sago waste material abundantly found in Sarawak was used and chemically modified into an oil adsorbent . Sago waste cellulosic residues were modified using fatty acid derivatives. The capability of the chemically modified sago waste to absorb oil from aqueous solution was studied and compared with the untreated sago waste. The modified sago waste showed higher hydrophobicity than the untreated sago waste, implying that it is less affinity for water and also an excellent affinity for oil. This chemically modified sago waste would be the most suitable for applications where engine oil (i.e., Shell Helix HX5) is to be removed from an aqueous environment. The modified sago waste selectively absorbs the oil and remains on the surface and is to be removed when the application is complete.

*Keywords:* Ester linkage, hydrophobic, modified sago waste, stearic acid.

## INTRODUCTION

Sago, which is scientifically known as *Metroxylon sagu,* comes from genus metroxylon and family palmae (Singhal *et al.,* 2008). Sago palm is commonly found in tropical lowland forests and freshwater swamps. The areas under sago cultivation in wild and semi-wild conditions are estimated to be at 19,720 hectares, with a total planted area of 28,000 hectares. Sarawak is currently one of the world's largest exporters of sago products with annual exports of approximately 43,000 tons. The mass production of sago produces residues during processing. It was estimated that from 600 logs of sago palm per day, 15.6 tons of woody bark, 237.6 tons of waste water, and 7.1 tons of starch fibrous sago pith residue are generated (Bujang & Ahmad, 1999). Sago pith residue is composed mainly of 41.7 - 65% starch and 14.8% fibre, including a fair amount of mineral (Wina *et*

*Article history:*
Received: 18 July 2011
Accepted: 14 February 2012

*E-mail address:*
nzainab@frst.unimas.my (Zainab Ngaini)
*Corresponding Author

ISSN: 0128-7680 

*al*., 1986). The residues are either incinerated or discharged into the river, which eventually contribute to serious environmental problems. Thus, there is a need to find ways to utilize these wastes into useable materials.

Sago waste comprises of cellulose and lignin, with hydroxyl functional groups (Quek *et al*., 1998). Several studies have reported on the utilization of sago wastes as fermentable sugar (Kumoro *et al.*, 2008), heavy metal remover (Quek *et al*., 1998) and some other applications (Singhal *et al*., 2008). However, no studies have reported on sago as an oil absorbent, even though sago waste is known to contain lignocellulosic material, as found in previously studied natural oil sorbents such as cotton, wool, bark, kapok, rice straw, barley straw, vegetable fibres, pith bagasse and raw bagasse (Husseien *et al*., 2008; Annunciado *et al*., 2005; Adebajo & Frost, 2004; Said *et al*., 2009).

Herein, we report on the chemical modification of sago waste with fatty acid derivatives via esterification on the cellulosic residues. The capability of the chemically modified sago wastes to absorb oil from aqueous solution was studied for potential commercial application.

## MATERIALS AND METHODS

Sago waste (SW) was obtained from Mukah, Sarawak. The engine oil used in all sorption experiments was Shell Helix HX5, Premium Multi-Grade Motor Oil. Dichloromethane was dried under calcium hydride. All chemicals were used as received and all reactions were performed under nitrogen atmosphere. The IR spectra were obtained on a Perkin Elmer Instruments Spectrum Gx1v5.0 using NaCl disc. SEM analysis was conducted to examine the surface morphology of the samples.

### *Sample Preparation of SW and Chemical Modification*

Stearic acid (0.01 mole, 3.0 g) was added to dichloromethane (60 mL). Oxalyl chloride (0.01 mole, 1.30 mL) was added drop wise in 30 minutes. Dimethylformamide (5-6 drops) was also added to initiate the reaction. The solution was stirred continuously at room temperature for 3 hours. SW (3.0 g), which had been ground and dried in an oven at 80±5 ˚C for 24 hours, was added into the prepared solution. Triethylamine (0.01 mole, 1.01 mL) was added drop wise into the solution and stirred continuously for 24 hours at room temperature. The modified sago waste (MSW) was then filtered, washed successively with cold dichloromethane followed by ethylacetate, and dried at room temperature (Cammidge *et al*., 2003).

### *Characterization of SW, MSW and Used Engine Oil (UEO)*

#### Density

A measuring cylinder was pre-weighed and SW was packed up to 1 mL measurement. The weight of SW was recorded. Three replicates were done and the density of SW was calculated using the mathematical formula shown below:

$$\text{Density (kg/m}^3\text{)} = \text{mass / volume}$$

The same steps were repeated to determine the density of MSW. A measuring cylinder was pre-weighed and UEO (10 mL) poured into it. Then, the weight of UEO was recorded. Three replicates were done and the density of UEO was calculated using the same mathematical formula.

*Water Sorption Test*

De-ionised water (70 mL) was poured into a 100 mL beaker. SW (0.1 g) was spread evenly on the surface of de-ionised water for different absorption times at 10, 20, 30 and 40 min. SW was then collected, air dried on filter paper for 10 minutes, weighed, and recorded. The same steps were applied to determine the water sorption capacity for MSW (Said *et al*., 2009). Water uptake was calculated using the formula below:

$$\text{Water uptake (\%)} = (\text{Sorption}_{\text{wet,static}} - \text{Sorption}_{\text{dry}})/\text{Sorption}_{\text{dry}} \times 100\%$$

*Hydrophobicity Test for SW and MSW*

De-ionised water (15 mL) was poured into a 30 mL vial, while SW (0.07 g) was poured onto the surface of the de-ionised water. The precipitations for these two mixtures were observed for 7 days. The same method was applied for MSW.

*UEO Sorption Tests*

Three systems applied to determine the sorption capacity of SW and MSW to UEO were the dry system, wet system (static) and wet system (dynamic). The study was performed at room temperature. The sorption capacity was calculated using the following mathematical formula (Husseien *et al*., 2008):

$$\text{Sorption (g/g)} = (S_t - S_o) / S_o$$

$S_t$ = Total mass of absorbed sample

$S_o$ = Initial weight of sample

**Dry System**

UEO (4 mL) was poured into a beaker. SW (0.1 g) was spread evenly on top of UEO. The sorption capacity test of SW was performed at 10, 20, 30 and 40 minutes. SW was collected, filtered and air dried on filter paper to remove excess oil. SW was weighed and recorded. Three replicates were carried out for each respective time. The same procedure was also performed on MSW (Husseien *et al*., 2008).

**Wet System (static)**

De-ionised water (70 mL) was poured into 100 mL beaker followed by UEO (4 mL). SW (0.1 g) was spread evenly on top of UEO. The sorption capacity of SW was performed at 10,

20, 30 and 40 minutes. SW was collected, filtered and air dried on a piece of filter paper to remove the excess oil, before it was weighed and recorded. Three replicates were performed each respective time. The same procedure was performed on MSW.

### Wet System (dynamic)

De-ionised water (70 mL) was poured into a 100 mL beaker, followed by UEO (4 mL). SW (0.1 g) was spread evenly on top of UEO. The mixture was stirred for sorption capacity at 10, 20, 30 and 40 minutes. SW was collected, filtered and air dried on a piece of filter paper to remove the excess oil, before it was weighed and recorded. Three replicates were performed each respective time. The same procedure was performed on MSW.

## RESULTS AND DISCUSSION

The lignocellulosic material of SW was modified *via* esterification of the hydroxyl groups with stearic acids. The conversion of stearic acids to acyl chlorides efficiently formed ester bonds onto SW network via hydroxyl groups, which contributed to hydrophobicity in the MSW.

IR spectra (Fig.1) showed the introduction of long alkyl ($CH_2$) groups onto the SW network at 2918 $cm^{-1}$ and 2849 $cm^{-1}$. The presence of peaks at 3416 $cm^{-1}$ and 3442 $cm^{-1}$ (Fig.1c) indicated that the hydroxyl groups were not fully substituted with stearic acid. This is due to the properties of sago waste which comprises mainly of starch, cellulose and lignin (Kumoro *et al*., 2008; Quek *et al*., 1998). In addition, it also consists of a higher number of hydroxyl groups. The peak attributed to the formation of ester is shown at 1723 $cm^{-1}$ (see Fig.1c). The peaks for carbonyl groups in MSW were fairly notable due to the massive network of MSW which overshadowed the carbonyl groups of ester linkage.

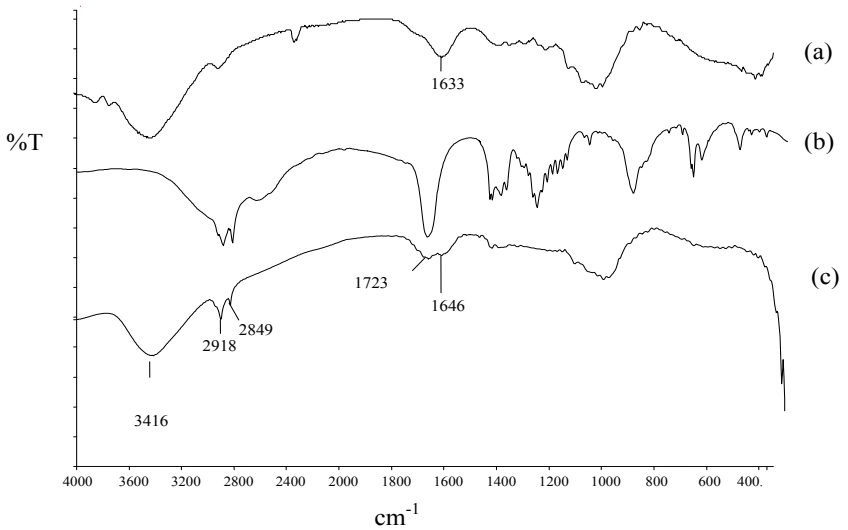



Fig.1: FT-IR spectra of (a) SW, (b) stearic acid, and (c) MSW

Scanning electron microscope (SEM) of SW and MSW is shown in Fig.2. There were no significant changes in the physical structure of SW before and after modification. This occurrence suggested that it was a chemical modification instead of physical modification (Erra *et al*., 2002).

Table 1 shows the density obtained for SW, MSW and UEO. The density of UEO was higher compared to SW and MSW. The difference in density for SW and MSW influenced their performances in the dry system, where only UEO was present. The density of water (1000 kg/m$^3$) affected the performance of SW and MSW in the wet system (static) in terms of the buoyancy (Husseien *et al*., 2008). The densities of SW and MSW were lower than water and thus, they remained on top of the water.

The water sorption capacity tests for SW and MSW were used as a control study to determine their ability to absorb oil. The sorption capacity was calculated in g/g unit and shown in Table 2. SW gave higher water sorption capacity compared to MSW.

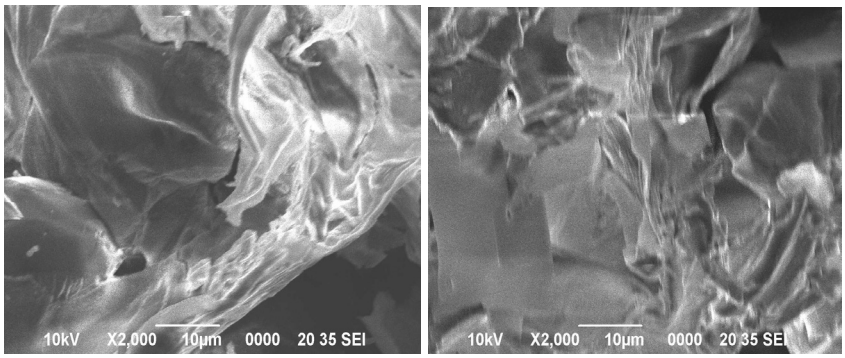



Fig.2: Scanning electron microscope of SW (left) and MSW (right)

TABLE 1
Density of SW, MSW and UEO

| Sample | Density (kg/m$^3$) |
|---|---|
| SW | 297.5±2.30 |
| MSW | 226.9±2.01 |
| UEO | 829.7±9.06 |

TABLE 2
Water sorption capacity (g/g) for SW and MSW

| Time (min) | SW | MSW |
|---|---|---|
| 10 | 2.2130±0.227 | 0.8330±0.288 |
| 20 | 2.7250±0.292 | 1.4593±0.243 |
| 30 | 3.1817±0.531 | 1.6524±0.194 |
| 40 | 3.3150±0.501 | 1.8379±0.199 |

SW and MSW were also subjected to hydrophobic tests to compare the level of hydrophobicity in both the samples. Both SW and MSW floated on the surface of water. SW was slowly submerged and completely accumulated at the bottom of the vial after 7 days (Fig.3a). However, MSW remained on the surface of water level. This observation suggests the formation of bonding between lignocellulosic hydroxyl groups with stearic acid *via* esterification.

The percentages of water uptake for SW in the wet system (static) and the wet system (dynamic) are shown in Table 3. In the wet system (static), the negative percentage values indicated that there was no water absorption. In the wet system (dynamic), the absorption of water occurred when SW was exposed to the system for 30 and 40 minutes. On the other hand, positive percentage values indicated that water absorption had occurred. The higher percentage of water uptake for SW was due to the higher number of hydroxyl groups in its lignocellulosic network (Kumoro *et al*., 2008) which easily formed hydrogen bonding with water molecules.

The percentage of water uptake for MSW is shown in Table 4. Similarly, the negative percentage values for water uptake in both the wet system (static) and wet system (dynamic) indicated that there was no absorption of water. This was due to the presence of long hydrophobic chains of stearic acid in the MSW network. The increasing hydrophobic property of MSW was explained by the negative percentage of water uptake in both wet system (static) and wet system (dynamic).

The sorption capacity of UEO onto SW and MSW is shown in Fig.4 to Fig.6. SW showed a higher sorption capacity compared to MSW in the dry system (see Fig.4) (Annunciado *et al*., 2005; Kumoro *et al*., 2008). MSW, however, showed lower sorption capacity. This was due to the lower density and higher buoyancy of MSW compared to SW (Fig.3) and UEO and thus, remained on the surface of UEO. The effectiveness of MSW to penetrate into UEO and absorb oil was lower due to the minimal contact between MSW and UEO.

Fig.3: (a) Accumulation of SW at the bottom of vial; (b) MSW remained on the surface of water after 7 days

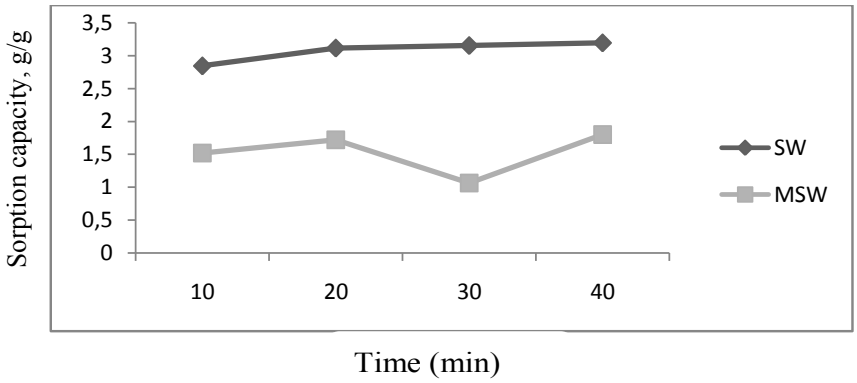



Fig.4: Sorption capacity of SW and MSW in dry system

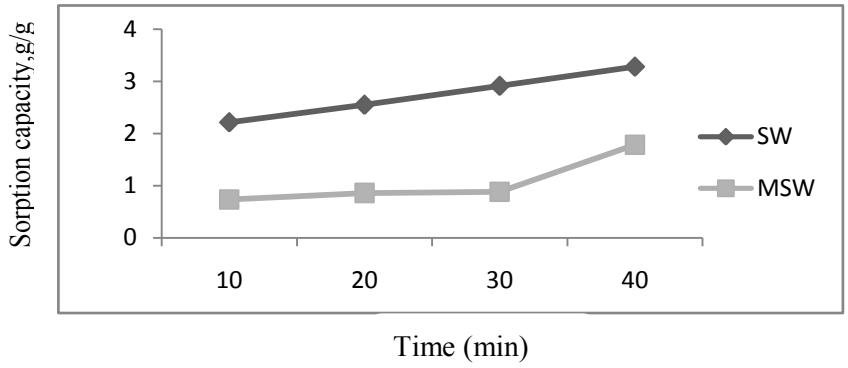



Fig.5: Sorption capacity of SW and MSW in the wet system (static)

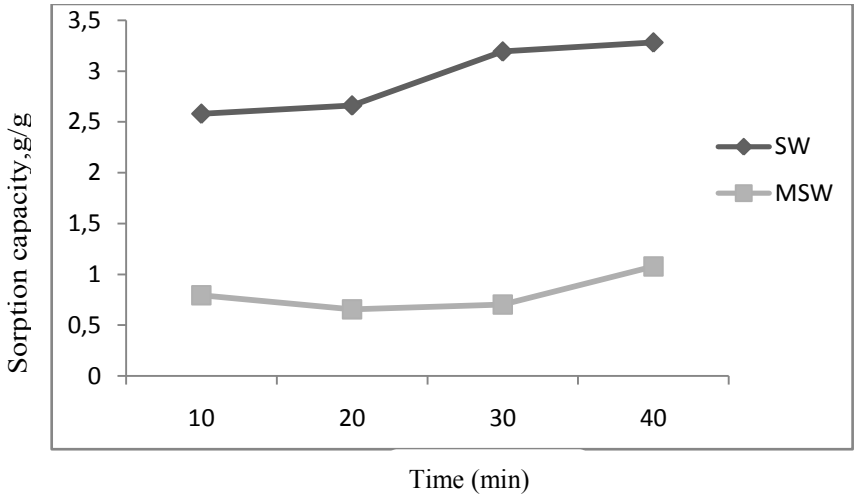



Fig.6: The sorption capacity of SW and MSW in the wet system (dynamic)

SW showed a higher sorption capacity compared to MSW in the wet system (static) (Fig.5). For the static system, the sorption capacity is inversely proportional to the buoyancy of fibres (Annunciado *et al*., 2005). MSW has higher buoyancy compared to SW and remained on the surface of UEO, thus reducing the fibre-oil contact. This might limit oil sorption ability.

Fig.6 shows the sorption capacity of SW and MSW in the wet system (dynamic). The weight gain for SW is not only contributed by the oil sorption. The incorporation of water into the fibres which ranged from 0.0-2.8% (Table 3) contributed to higher values in sorption capacity. The presence of lignocellulosic material in the SW network afforded SW to absorb hydrophobic and hydrophilic materials. Vigorous agitation and frequent water-fibres contact have affected the sorption capacity of SW (Annunciado *et al*., 2005). On the other hand, the sorption capacity of MSW is solely from UEO. This may have happened due to its higher buoyancy compared to SW and also to the presence of higher hydrophobic moieties which therefore selectively absorbed oil compared to water. Therefore, MSW showed a relatively better performance to absorb oil in dynamic aqueous condition compared to SW.

TABLE 3
Calculated percentage of water uptake for SW in wet system (static) and wet system (dynamic)

| Time (min) | Dry system (g/g) | Wet system (static) (g/g) | Calculated water uptake (%)* | Wet system (dynamic) (g/g) | Calculated water uptake (%)* |
|---|---|---|---|---|---|
| 10 | 2.8452±0.145 | 2.2144±0.238 | -22.1 | 2.580±0.376 | -9.3 |
| 20 | 3.1145±0.133 | 2.5507±0.241 | -18.1 | 2.6611±0.280 | -14.5 |
| 30 | 3.1525±0.199 | 2.9135±0.108 | -7.5 | 3.1925±0.544 | 1.3 |
| 40 | 3.1931±0.221 | 3.0190±0.284 | -5.5 | 3.2815±0.466 | 2.8 |

*Note: negative value represents "no water uptake and equiv. to 0%"

TABLE 4
The calculated percentage of water uptake for MSW in the wet system (static) and wet system (dynamic)

| Time (min) | Dry system (g/g) | Wet system (static) (g/g) | Calculated water uptake (%)* | Wet system (dynamic) (g/g) | Calculated water uptake (%)* |
|---|---|---|---|---|---|
| 10 | 1.521±0.255 | 0.7342±0.371 | -5.2 | 0.796±0.159 | -47.6 |
| 20 | 1.721±0.114 | 0.8598±0.294 | -5.0 | 0.657±0.380 | -61.8 |
| 30 | 1.067±0.165 | 0.8811±0.150 | -1.7 | 0.730±0.111 | -31.6 |
| 40 | 1.800±0.093 | 1.7844±0.093 | -0.9 | 1.079±0.197 | -40.1 |

* Note: negative value represents "no water uptake and equiv. to 0%"

## CONCLUSION

The introduction of long alkyl ($CH_2$) groups onto the SW network *via* esterification has increased the hydrophobic property. In this study, MSW afforded a better sorption capacity on UEO as compared to SW in aqueous environment (wet system). MSW showed excellent affinity for oil and low affinity for water. For oil sorption in the absence of water (dry system), SW was found to be good material.

## ACKNOWLEDGEMENTS

The authors would like to thank Universiti Malaysia Sarawak and the Ministry of Science, Technology and Innovation (MOSTI) for the financial support throughout this project, FRGS/01(03)/608/2006(41).

## REFERENCES

Adebajo, M. O., & Frost, R. L. (2004). Acetylation of raw cotton for oil spill cleanup application: an FTIR and $^{13}C$ MAS NMR spectroscopic investigation. *Spectrochimica Acta, Part A. Molecular and Biomolecular Spectroscopy, 60*(10), 2315-2321.

Annuciando, T. R., Sydenstricker, T. H. D., & Amico, S. C. (2005). Experimental investigation of various vegetables fibers as sorbent material for oil spills. *Marine Pollution Bulletin, 50*, 1340-1346.

Bujang, K. B., & Ahmad F. B. (1999). Production, properties and utilisation of sago starch in malaysia. in; sustainable small-scale sago starch extraction and utilisation: guidelines for the sago industry. *Kasertsart Agric and Agro-indust Prod Dev Institute (KAPI) and FAO*. August 9-11th, 2009 Nakhon Si Thammarat, Thailand, pp: 82-95.

Cammidge, A. N., Downing, S., & Ngaini, Z. (2003). Surface-functionalised nano-beads as novel supports for organic synthesis. *Tetrahedron Letters*, *44*, 6633-6634.

Erra, P., Jovanèiæ, P., Molina, R., Jociæ, D., & Julia, R. (2002). Study of surface modification of keratin fibres by means of SEM, AFM and light microscopy. *Science, Technology and Education of Microscopy: an Overview, 2*, 549-556.

Husseien, M., Amer, A. A., El-Maghraby, A., & Taha, N. A. (2008). Experimental modification of thermal modification influence on sorption quality of barley straw. *Journal of Applied Science Research*, *4*, 652-657.

Kumoro, A. C., Ngoh, G. C., Hasan, M., Ong, C. H., & Teoh, E. C. (2008). Conversion of fibrous sago (Metroxylon Sagu) waste into fermentable sugar via acid and enzymatic hydrolysis. *Asian Journal of Scientific Research, 1*(4), 412-420.

Quek, S. Y., Wase, D. A. J., & Forster, C. F. (1998). The use of sago waste for the sorption of lead and copper. *Water SA, 24*(3), 251-256.

Said, A. E-A. A., Ludwick, A. G., & Aglan, H. A. (2009). Usefulness of raw bagasse for oil absorption: A comparison of raw and acylated bagasse and their components. *Bioresource Technology, 100*, 2219-2222.

Singhal, R. S., Kennedy, J. F., Gopalakrishnan, S. M., Kackzmarek, A., Knill, C. J., & Akmar, P. F. (2008). Industrial production, processing, and utilization of sago palm derived products. *Carbohydrate Polymers, 72*, 1-20.

Wina, E., Evans, A. J., & Lowry, J. B. (1986). The *composition of pith* from the *sago palms Metroxylon* sagu and Arenga pinnata. *Journal of the Science of Food and Agriculture, 37*, 352-358.

Pertanika J. Sci. & Technol. 22 (1): 163 - 174 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Chemical Constituents of *Aglaia lanuginose*

**Kamarulzaman, F. A.[1,3], Mohamad, K.[2], Awang, K.[1] and Lee, H. B.[3]***

*[1]Department of Chemistry, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia*
*[2]Department of Pharmacy, Faculty of Medicine University of Malaya, 50603 Kuala Lumpur, Malaysia*
*[3]Cancer Research Initiatives Foundation (CARIF), Level 2 Outpatient Centre, Sime Darby Medical Centre, No 1 Jalan SS 12/ 1A, 47500 Subang Jaya, Selangor, Malaysia*

## ABSTRACT

Our continuing research on the *Aglaia* genus (family Meliaceace) has led us to this first study on the chemical constituents of *Aglaia lanuginose* (bark). The dichloromethane extract from the bark of *Aglaia lanuginose* showed cytotoxicity against HL-60 leukaemia cell line (45% inhibition) at 20 µg/ml and was prioritised for further investigation. Repeated chromatography of the dichloromethane extract yielded the known dammarane triterpenes which were identified as cabralealactone (1), methyl eichlerianate (2), cabraleone (3), ocotillone (4), eichleriatone (5), eichlerianic acid (6) and shoreic acid (7) together with the known sterols, sitosterol (9) and stigmasterol (10). Another isolated compound was the aromatic 4-hydroxycinnamyl-acetate (8), which has not been reported to be present in a plant from the Meliaceae family. The structures of all the compounds were elucidated on the basis of spectroscopic methods (IR, MS and NMR). Cytotoxicity testing of 1-10 showed activity only for mixtures of (3, 4), and (5, 6).

*Keywords: Aglaia lanuginose, Aglaia*, dammaranes, Meliaceae

## INTRODUCTION

*Aglaia* Lour, the largest genus of the Meliaceae family, consists of approximately 130 species that can be found in the Indo-Malaysian region in South China and on the Pacific Island. It occurs in a variety of habitats ranging from rainforests and mangrove swamps to semi deserts (Muellner *et al.,* 2003). Most of *Aglaia spp*. are trees that can reach up to 40 meters in heights. *Aglaia* spp. are important to the people in the South East Asia region because the trees are useful as sources for timber and their fruit are consumed, sometimes for their medicinal values (Simantujal *et al.,* 1999). In Vietnam, the crude extracts from the leaves and flowers of *Aglaia spp.* are used to treat inflammatory skin disease and allergic disorders such as asthma (Proksh *et al.,* 2005).

*Article history:*
Received: 15 August 2011
Accepted: 10 September 2012

*E-mail addresses:*
fadzly_adzhar@yahoo.com (Kamarulzaman, F. A.),
khalitmohamad@um.edu.my (Mohamad, K.),
khalijah@um.edu.my (Awang, K.),
hongboon.lee@carif.com.my (Lee, H. B.)
*Corresponding Author

ISSN: 0128-7680 

*Aglaia* plants are used to treat fever and as preparations in leprosy treatment by using the latex from the plant to smear on the affected parts of the skin (Christensen *et al.,* 2002). In India, these *Aglaia* plants are used to treat yellow fever, common fever and as an antidote (Kritikar *et al.,* 1975). In addition, the flowers of *Aglaia odorata* are used as perfume in China and Indonesia.

During the past few years, the genus *Aglaia* has been receiving increased focus in scientific research due to its potential in producing biologically active compounds. The most common compounds from various *Aglaia* species that have been reported so far are triterpenes and benzofurans, and many of these compounds possess biological activities. Dammaranes, cycloartanes, tirucallanes, apotirucallanes, glabretals, bacharanes and lupanes represent the most widespread triterpenes within the genus of *Aglaia* (Silvia *et al.,* 2008). As for the benzofuran compounds, studies have shown that those with the cyclopentabenzofuran core are phytochemically confined to the genus *Aglaia* (Brader *et al.,* 1998). These cyclopentabenzofurans are also commonly known as recoglate or flavagline derivatives and have been observed to be biologically active including cytotoxic activity against human cancer cells. An example of a cyclopentabenzofuran is silvestrol; isolated from *Aglaia leptantha* Miq. and collected in Borneo Island, it is currently being developed as an anti-cancer drug for chronic lymphocytic leukaemia (Lucas *et al.,* 2009).

Previous studies have indicated that many *Aglaia* plants contain toxic chemical compounds that can be used as pesticides or anticancer drugs (Simantujak *et al.,* 1999). As part of our continuing research on the *Aglaia* genus, we observed that the dichloromethane extract from the bark of *Aglaia lanuginose* showed moderate in vitro cytotoxic activity against HL60 cell line. As *Aglaia lanuginose* has not been studied for its phytochemical constituents and some of the constituents may exhibit cytotoxicity, here we describe our findings on the chemical constituents from the bark extract of *A. lanuginose* and the cytotoxicity of these compounds against HL60 cells.

## MATERIALS AND METHODS

### *Plant Material*

The bark of *Aglaia lanuginose* King was collected in Peninsular Malaysia near the town of Jeli, Kelantan, which is located 127 km along the Jeli-Gerik route from Kota Bharu to Ipoh. A voucher specimen (KL 4232) was deposited at the Herbarium of Department of Chemistry, University Malaya, Kuala Lumpur, Malaysia.

### *Experimental Design*

*Fractionation and purification.* The dichloromethane crude extract (10 g) which was cytotoxic against HL60 leukaemia cells (45% inhibition) at 20 μg/ml, was subjected to column chromatography (CC) over silica gel and eluted with dichloromethane/ methanol with 1% increments of methanol to give a total of 11 fractions.

Fraction 3 (175.9 mg), which was eluted from the previous CC ($SiO_2$) with 100% dichloromethane, was further purified by CC ($SiO_2$) using increasing amounts of ethyl acetate in hexane to give compounds 1, 2, 8, 9 and 10 (all eluted with 5% ethyl acetate in hexane). Fraction

4 (543 mg), which was eluted from the first column with 2% methanol in dichloromethane, was further purified by CC ($SiO_2$) before it was eluted using increasing amounts of ethyl acetate in hexane to obtain compounds 3 and 4. Both compounds 3 and 4 were eluted with 10% ethyl acetate in hexane. Fraction 5 (800 mg), which was eluted from the first column with 2% methanol in dichloromethane, was further purified by CC ($SiO_2$) and then eluted using increasing amounts of ethyl acetate in hexane to obtain compound 5, which was eluted at 15% ethyl acetate in hexane. Fraction 6 (3.7 g), which was eluted from the first column with 2% methanol in dichloromethane, was further purified by CC ($SiO_2$) before it was eluted using increasing amounts of acetone in hexane to obtain compounds 6 and 7.

Compounds 1, 2, 5, 8 were further purified with prep-thin layer chromatography (TLC) using a mixture of hexane and ethyl acetate or a mixture of hexane and acetone solvent systems depending on the polarity of the compounds, while compounds 6 and 7 were obtained in pure form by recrystalization from methanol. Compounds 9 and 10 were identified as the common sterols and were not purified further. Structure elucidation was carried out by $^{1}$H-NMR, $^{13}$C-NMR, COSY, HMQC, HMBC, DEPT, IR, UV and Mass-spectroscopy.

*Cytotoxicity Assay*. HL60 cells were acquired from ATCC (American Type Cell Culture) and were cultured in RPMI 1640 (Gibco) media added with 10% of fetal bovine serum (FBS, Gibco). Briefly, test compounds or extracts were prepared as stock solutions of 20 mg/ml in dimethyl sulfoxide (DMSO) and diluted accordingly in phenol red free culture medium supplemented with 5% FBS before use. In a 96-well plate, samples in 4 replicates (n = 4) at a concentration of 5 µg/ml were tested against 15,000 of HL60 cells per well for 2 hours at 37°C in 5% $CO_2$. The viability of the cells was indirectly determined by MTT dye reduction, where 15 µl of MTT (5 mg/ml in PBS) was added into each well which contained the cells. After incubating for 4 hours, 70 µl of the supernatant was pipetted out from each well. The formazan crystals formed were dissolved by the addition of 100 µl of DMSO. The absorbance was recorded at 570 nm and the viability of the cells was calculated.

## RESULTS AND DISCUSSION

Dried ground barks (2.5 kg) were extracted sequentially with hexane, dichloromethane and methanol at room temperature. The hexane, dichloromethane and MeOH extracts were evaporated in vacuum to yield 58.57 g of hexane extract, 20.20 g of dichloromethane extract and 47.05 g of methanol extract separately. Cytotoxicity assay on HL60 using 20 µg/mL of the three extracts showed that the dichloromethane extract had moderate activity (45% inhibition) while the other two extracts were inactive. The dichloromethane extract was selected for further fractionation as part of our focus on isolating chemical constituents that may have anti-cancer properties.

### *Isolation and Structural Elucidation*

Cabralealactone (1) was isolated as a white amorphous powder. The MS spectrum of compound 1 showed a $[M]^+$ peak at *m/z* 414 which corresponded to the molecular formula $C_{27}H_{42}O_3$. The triterpenoidal nature of compound 1 was revealed by $^{1}$H-NMR (Table 1) which showed six methyl groups and some complex multiplets belonging to the alicyclic protons. The $^{13}$C-NMR

showed 27 peaks assignable to six methyl ($CH_3$), ten methylene ($CH_2$), four methine (CH) groups and seven quaternary peaks which were distinguishable from the DEPT spectra. The $^{13}$C-NMR spectrum showed additional features of tertiary carbon resonances at $\delta_C$ 89.9 and $\delta_C$ 176.9, assignable to the lactone system. The $^1$H-NMR and MS data of compound 1 were similar to those of cabralealactone in the literature (Cascon *et al.,* 1972). On the basis of the evidence presented, compound 1 was identified as cabralealactone.

Methyl isoeiclerianate (2) was isolated as a white amorphous powder. The MS spectrum of compound **2** showed a $[M+H]^+$ peak at *m/z* 489 which corresponded to the molecular formula $C_{31}H_{52}O_4$. The $^1$H-NMR spectrum (Table 1) showed seven tertiary methyls, an oxomethine at $\delta_H$ 3.75 (*t*) and multiplet signals belonging to the alicylic protons. The open seco skeleton was confirmed by the presence of two broad singlets at $\delta_H$ 4.66 and 4.84 (2 *brs*, 1H each), which corresponded to the two terminal protons (H-28α and H-28β). Furthermore, a methoxy group located at C-3 proved that compound 2 is a methyl ester. The EIMS of 2 showed a characteristic fragment ion at *m/z* 143 (100%), consistent with the presence of a hydroxy-methyl tetrahydrofuran side chain $[C_8H_{15}O_2]^+$ in the molecule. The configuration of compound 2 was determined to be 24*S* based on the resonance of C-24 at $\delta_C$ 86.4 in $^{13}$C NMR. The assignments of carbons in the $^{13}$C-NMR of compound 2 were similar to those of methyl isoeichlerianate in the literature (Seger *et al.,* 2008). From these observations, compound 2 was identified as methyl isoeiclerianate.

Cabraleone (3) was isolated as a white amorphous powder. The EIMS spectrum of compound 3 showed a $[M + H]^+$ peak at *m/z* 459 which corresponded to the molecular formula $C_{30}H_{50}O_3$. The triterpenoidal nature of compound 3 was revealed in the $^1$H-NMR spectrum (Table 1) with eight tertiary methyl groups as sharp singlets, multiplet signals of the alicyclic protons and an oxomethine signal at $\delta_H$ 3.64 (*t*). The EIMS of 3 showed a characteristic fragment ion at *m/z* 143 (100%) which was similar to compound 2 and consistent with the presence of a hydroxy-methyl tetrahydrofuran side chain $[C_8H_{15}O_2]^+$ in the molecule. The configuration of compound **3** was determined to be 24*S* based on the resonance of C-24 at $\delta_C$ 86.4 in $^{13}$C NMR. The NMR and MS data of compound 3 were similar to those of cabraleone in the literature (Albersberg *et al.,* 1991). On the basis of the evidence above, the structure of compound 3 was assigned as cabraleone.

Ocotillone (4) was isolated as a white amorphous powder. the MS spectrum of compound 4 showed a $[M + H]^+$ peak at *m/z* 459 - similar to compound 3 - which corresponded to the molecular formula $C_{30}H_{50}O_3$. Signals of $^1$H- (Table 1) and $^{13}$C-NMR spectra of compound 4 were similar to those of its isomer 3, and both compounds can be distinguished by the chemical shifts of C-24. The configuration of compound 4 was determined to be 24*R* due to the resonance of C-24 at $\delta_C$ 84.5 in $^{13}$C NMR. The other proton assignments in the $^1$H-NMR of compound 4 were similar to those of ocotillone in the literature (Nuanyai *et al.* 2011). From these observations, compound 4 was identified as ocotillone.

Eichlerialactone (5) was isolated as a white amorphous powder. The MS spectrum of compound 5 showed a $[M + H]^+$ peak at *m/z* 431 which corresponded to the molecular formula $C_{27}H_{42}O_4$. The $^1$H-NMR spectrum (Table 1) showed five tertiary methyls and multiplet signals belonging to the alicylic protons. In addition, the $^1$H-NMR spectrum revealed signals of two terminal protons (H-25α and H-25β) as broad singlets at $\delta_H$ 4.67 and 4.88 (2 *brs*, each 1 H)

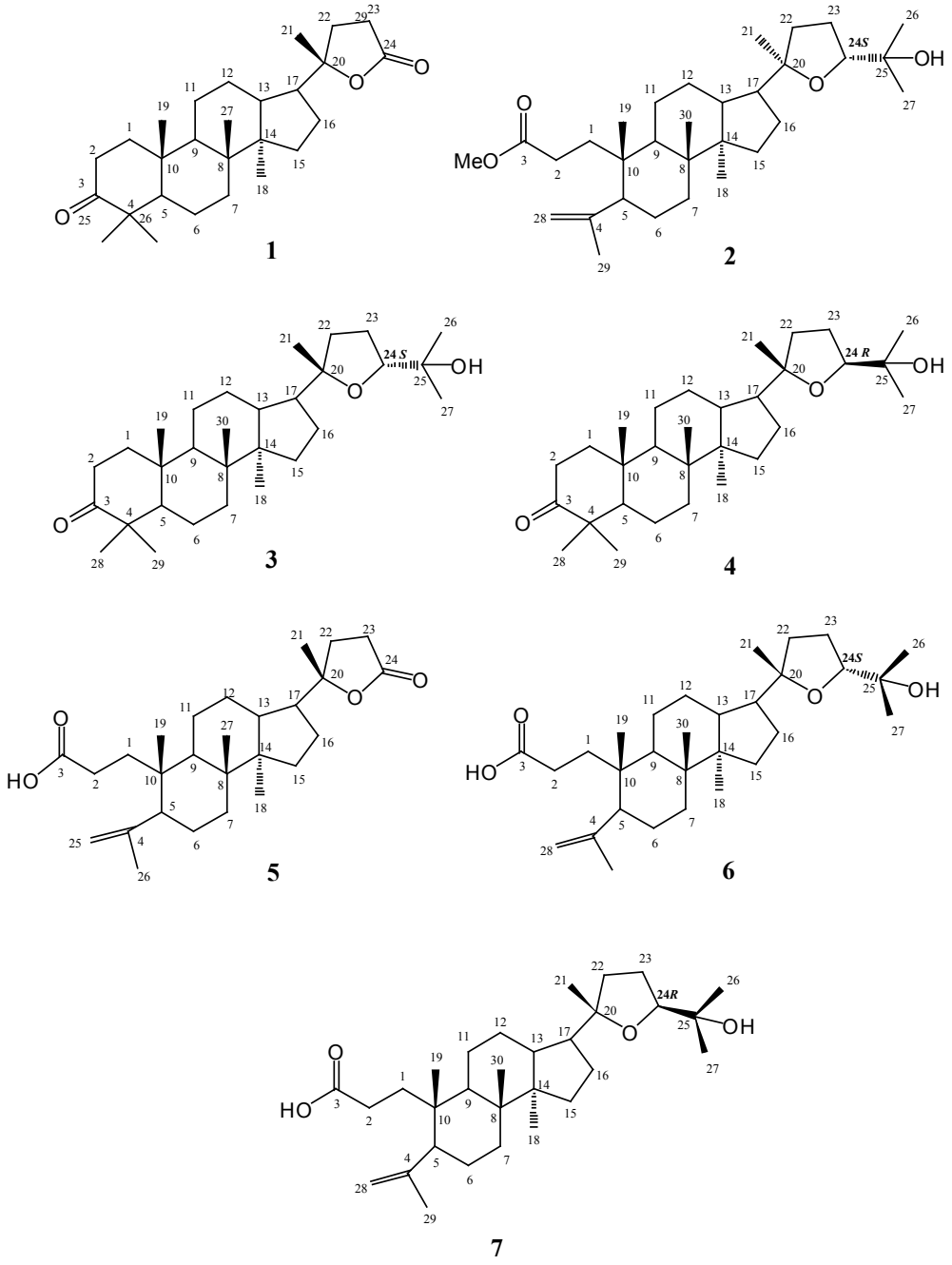



Fig.1: Structures of dammarane triterpenes isolated from *Aglaia lanuginose* King (bark)

which are typical of a seco-dammarane triterpenoid. The $^{13}$C-NMR spectrum analyzed by the aid of DEPT spectrum indicated five tertiary methyls, seven quaternary carbons and three oxygen bearing carbons at $\delta_C$ 176.9, 89.9, 177.0 for C-3, C-20 and C-24, respectively. The NMR and MS data of compound 5 were similar to those of eiclerialactone in the literature (Singh *et al.,* 1991). On the basis of the evidence above, the structure of 5 was identified as eichlerialactone.

Eichlerianic acid (6) was isolated as a white powder. The MS spectrum of compound 6 showed a molecular ion peak $[M]^+$ at *m/z* 474 corresponding to the molecular formula of $C_{30}H_{50}O_4$. The triterpenoid nature of compound 6 was revealed in the $^1$H-NMR spectrum (Table 1) by having seven tertiary methyl groups as sharp singlets and some complex multiplet signals belonging to the methylene and methine protons. Additionally, there was a signal assignable to a methine proton H-24 at $\delta_H$ 3.57 (*dd*, *J* = 10.3 and 5.0 Hz, 1H) and broad singlets at $\delta_H$ 4.60 and 4.78 (2 *brs*, each 1H) which were assignable to the terminal protons H-28α and H-28β. The configuration of compound 6 was determined to be 24*S* based on the resonance of C-24 at $\delta_C$ 86.3 in $^{13}$C NMR. The other proton assignments in the $^1$H-NMR and carbon in the $^{13}$C-NMR of compound 6 were similar to those of eiclerianic acid in the literature (Roux *et al.,* 1998). On the basis of the evidence above the structure was assigned to be eichlerianic acid.

Shoreic acid (7**)** was isolated as a white amorphous powder. The MS spectrum of compound (7) showed a molecular ion peak $[M]^+$ at *m/z* 474 corresponding to the molecular formula of $C_{30}H_{50}O_4$. The triterpenoid nature of compound 7 was revealed in the $^1$H-NMR spectrum (Table 1) by having seven tertiary methyl groups as sharp singlets, and some complex multiplet signals belonging to the alicyclic protons. Similarly, with respect to compounds 5 and 6, two doublets appeared as two broad singlets in the $^1$H-NMR spectrum at $\delta_H$ 4.60 and 4.78 (2 *brs*, each 1H), which corresponded to the two terminal protons H-28α and H-28β. The configuration of compound 7 was determined to be 24*S* based on the resonance of C-24 at $\delta_C$ 84.4 in $^{13}$C NMR. The other assignments in the $^1$H-NMR and the $^{13}$C-NMR of compound **7** were similar to those of shoreic acid in the literature (Roux *et al.* 1998). On the basis of the evidence presented, compound 7 was identified as shoreic acid.

4-hydroxycinnamyl acetate (8**)** was isolated as yellow oil. The MS spectrum of compound (8) showed a molecular ion $[M]^+$ peak at *m/z* 192 which was in agreement with the molecular formula $C_{11}H_{12}O_3$. The $^1$H-NMR spectrum (Table 2) showed the presence of a singlet peak at $\delta_H$ 2.28 assignable to C(11) methyl. The $^1$H-NMR spectra of compound **8** showed signals of H-9 methylene at $\delta_H$ 4.30 (*d*, *J* = 6.1 Hz, 2H), a H-8 *trans* olefin pair at $\delta_H$ 6.29 (*dt*, *J* = 6.1, 16.1 Hz, 1H) for H-8 and a H-7 proton at $\delta_H$ 6.57 (*d*, *J* = 16.1 Hz, 1H). There were also proton signals corresponding to a *para*-substituted aromatic ring with characteristic chemical shifts for [H-3, H-5] and [H-2, H-6] at $\delta_H$ 7.02 and 7.36 respectively (*d*, *J* = 8.6 Hz, 2H each). The assignments of protons in the $^1$H-NMR of compound **8** were similar to those of 4- hydroxycinnamic acetate in the literature (Kuichi *et al.,* 2002). On the basis of the data presented above, structure **8** was identified to be 4-hydroxycinnamyl acetate.

The MS spectrum of mixtures of sterols (9, 10) showed a molecular ion $[M]^+$ peaks at *m/z* 414 and 412, respectively. The identity of compounds 9 and 10 was established by direct comparison of $^{13}$C-NMR and $^1$H-NMR to the data published in literature (Eknmakul *et al.,* 2003).

TABLE 1a
$^1$H-NMR data of triterpene compounds 1, 2, 3, 4, 5, 6 and 7 in $CDCl_3$

| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 1.84 (m) | 1.75 (m), 1.53 (m) | 1.45 α (m), 1.72 β (m) | 1.46 α (m), 1.94 β (m) | 1.24 (m), 1.6 (m) | 1.75 (m), 1.53 (m) | 1.75 (m), 1.53 (m) |
| 2 | 2.5 (m) | 2.14 (m) 2.32 (m) | 2.46 α (m) 2.54 β (m) | 2.42 α (m) 2.50 β (m) | 1.50 β (m), 1.18 α (m) | 2.14 (m), 2.32 (m) | 2.14 (m), 2.32 (m) |
| 3 | - | - | - | - | - | - | - |
| 4 | - | - | - | - | - | - | - |
| 5 | - | - | 1.38 (m) | 1.38 (m) | 1.98 (m) | - | - |
| 6 | - | - | 1.45 α (m) 1.55 β (m) | 1.40 α (m) 1.55 β (m) | 1.38 | - | - |
| 7 | 1.49 (m) | - | 1.3 α (m) 1.58 β (m) | 1.66 α (m) 1.70 β (m) | 1.54β (m), 1.24α (m) | 1.15 (m) | - |
| 8 | - | - | - | - | - | - | - |
| 9 | 1.31 | - | 1.88 (m) | 1.80 (m) | 1.50 (m) | 1.43 (m) | 1.15 (m) |
| 10 | | - | - | - | | - | |
| 11 | - | - | 1.28 α (m) 1.50 β (m) | 1.25 α (m) 1.50 β (m) | 1.30 β (m), 1.44 α (m) | 1.35 (m) | 1.43 (m) |
| 12 | 1.56 (m), 1.21 (m) | - | 1.25 (m) | 2.00 (m) | 1.90 β (m), 1.28 α (m) | - | - |
| 13 | 1.52 (m) | - | 1.58 (m) | 1.70 (m) | 1.68 (m) | 1.60 (m) | 1.35 (m) |
| 14 | 1.31 (m) | - | - | - | - | - | - |
| 15 | - | - | 1.1α (m) 1.48 β (m) | 1.1α (m) 1.48β (m) | 1.50 β (m), 1.24 α (m) | 1.40 (m) | 1.60 (m) |
| 16 | - | - | 1.32 (m) | 1.32 | 1.82 β (m), 1.32 α (m) | 1.75 (m) | 1.40 (m) |
| 17 | 1.85 (m) | - | 1.44 (m) | 1.44 (m) | 1.94 (m) | 1.80 (m) | 1.75 (m) |
| 18 | 0.83 (s) | 0.88 (s) | 0.89 (m) | 0.88 (s) | 0.86 (s) | 0.82 (s) | 0.79 (s) |
| 19 | 0.92 (s) | 0.84 (s) | 0.95 (m) | 0.94 (s) | 0.90 (s) | 0.79 (s) | 0.82 (s) |
| 20 | - | - | - | - | - | - | - |
| 21 | 1.32 (s) | 1.14 (s) | 1.11(m) | 1.11(s) | 1.34(s) | 1.08 (s) | 1.81 (s) |
| 22 | 2.00 (m) | - | 1.7 α (m) 1.85 β (m) | 1.70 α (m) 1.85 β (m) | 2.16 β (m), 1.94 α (m) | 1.57 (m) | 1.57 (m) |
| 23 | 1.85 (m) | - | 1.3 α (m) 1.75 β (m) | 1.80 (m) | 2.56 (m) | 1.80 (m) | 1.80 (m) |
| 24 | - | 3.75 (t) | 3.64 (t) | 3.76 (t) | | 3.57 (dd, $J$ = 10.3, 5.0 Hz)) | 3.73 (t) |
| 25 | 1.06 (s) | - | - | - | 4.67 (brs), 4.88 (brs) | - | - |
| 26 | 1.01 (s) | 1.18 | 1.19 (s) | 1.15 (s) | 1.76 (s) | 1.13 (s) | 1.13 (s) |
| 27 | 0.98 (s) | 1.11 | 1.20 (s) | 1.20 (s) | 1.0 (s) | 1.05 (s) | 1.05 (s) |

TABLE 1a (*continue*)

| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 28 | - | 4.66 (brs), 4.84 (brs) | 1.08 (s) | 1.08 (s) | - | 4.60 (brs), 4.78 (brs) | 4.60 (brs), 4.78 (brs) |
| 29 | - | 1.76 (s) | 1.04 (s) | 1.04 (s) | - | 1.67 (s) | 1.67 (s) |
| 30 | - | 0.99 (s) | 1.01(s) | 1.00 (s) | - | 0.99 (s) | 0.99 (s) |
| 31 | - | 3.66 (s) | - | - | - | - | - |

The protons at H-24 of compounds 2, 3, 4 and 7 are pseudo-*dd* which appear as *t*, with approximate $J = 5$ Hz.

TABLE 1b

$^{13}$C-NMR data of triterpene compounds **1**, **2**, **3**, **4**, **5**, **6** and **7 in $CDCl_3$**

| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 39.8 | 34.4 | 39.9 | 39.9 | 34.4 | 34.3 | 34.3 |
| 2 | 34.5 | 28.3 | 34.1 | 34.1 | 28.4 | 28.5 | 28.5 |
| 3 | 218.0 | 174.6 | 218.1 | 218.1 | 176.9 | 179.5 | 179.4 |
| 4 | 47.4 | 147.5 | 47.4 | 47.4 | 147.4 | 147.5 | 147.4 |
| 5 | 55.3 | 50.7 | 55.4 | 55.3 | 50.8 | 50.8 | 55.3 |
| 6 | 19.6 | 24.6 | 19.7 | 19.7 | 24.6 | 24.6 | 19.7 |
| 7 | 34.1 | 33.9 | 34.8 | 37.6 | 33.9 | 33.9 | 37.6 |
| 8 | 42.7 | 40.0 | 40.3 | 40.3 | 40.1 | 40.0 | 40.3 |
| 9 | 50.0 | 41.0 | 49.8 | 50.1 | 41.1 | 41.2 | 50.1 |
| 10 | 36.8 | 37.5 | 36.9 | 36.9 | 39.1 | 39.0 | 39.9 |
| 11 | 21.8 | 21.7 | 22.3 | 22.0 | 21.8 | 22.3 | 22.0 |
| 12 | 25.0 | 26.9 | 27.0 | 27.0 | 26.4 | 26.9 | 27.0 |
| 13 | 43.0 | 42.9 | 43.0 | 43.0 | 42.7 | 42.9 | 43.0 |
| 14 | 49.4 | 50.6 | 50.0 | 49.9 | 50.4 | 50.3 | 49.9 |
| 15 | 31.0 | 31.2 | 31.4 | 31.2 | 31.1 | 31.4 | 31.2 |
| 16 | 28.6 | 25.8 | 25.8 | 25.9 | 25.0 | 25.8 | 25.9 |
| 17 | 49.4 | 50.0 | 50.2 | 50.1 | 49.5 | 49.7 | 49.7 |
| 18 | 16.1 | 16.0 | 16.3 | 16.1 | 16.1 | 16.3 | 16.0 |
| 19 | 15.2 | 20.1 | 16.1 | 16.1 | 20.1 | 20.2 | 20.1 |
| 20 | 89.9 | 86.0 | 86.5 | 86.5 | 89.9 | 86.6 | 86.5 |
| 21 | 22.1 | 21.9 | 27.2 | 26.8 | 22.5 | 27.1 | 24.2 |
| 22 | 33.2 | 37.5 | 34.6 | 34.6 | 33.0 | 34.7 | 34.7 |
| 23 | 26.4 | 29.6 | 26.4 | 25.8 | 28.7 | 26.3 | 26.3 |
| 24 | 176.9 | 86.4 | 86.4 | 84.5 | 177.0 | 86.3 | 84.4 |
| 25 | 26.7 | 71.1 | 70.3 | 71.1 | 147.4 | 70.3 | 76.7 |
| 26 | 21.0 | 27.6 | 24.1 | 21.7 | 23.2 | 27.8 | 27.5 |
| 27 | 16.0 | 24.3 | 27.8 | 27.7 | 15.3 | 23.2 | 21.7 |
| 28 | | 113.3 | 26.4 | 24.3 | | 113.4 | 113.4 |
| 29 | | 23.2 | 21.0 | 21.0 | | 24.0 | 23.2 |
| 30 | | 15.3 | 15.2 | 15.2 | | 15.3 | 15.3 |
| 31 | | 51.6 | | | | | |

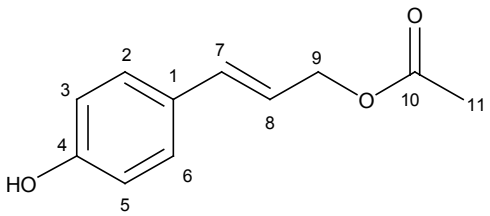



Fig.2: Structure of compound 8

TABLE 2
1D and 2D NMR of compound 8 in $CDCl_3$

| No | $\delta_C$ (ppm) | $\delta_H$ (ppm) | HMBC (H-C) |
|---|---|---|---|
| 1 | 134.5 | - | - |
| 2 | 121.7 | 7.36 (*d*, 8.6 *Hz*) | 1 |
| 3 | 127.4 | 7.02 (*d*, 8.6 *Hz*) | 1,4,5 |
| 4 | 150.6 | - | - |
| 5 | 127.4 | 7.02 (*d*, 8.6 *Hz*) | 1,3,4 |
| 6 | 121.7 | 7.36 (*d*, 8.6 *Hz*) | 2,4,8 |
| 7 | 130.1 | 6.57 (*d*, 16.1 *Hz*) | 2,6 |
| 8 | 128.7 | 6.29 (*dt*, 6.1, 16.1 *Hz*) | 1 |
| 9 | 63.6 | 4.30 (*d*, 6.1 *Hz*) | 7,8 |
| 10 | 169.6 | - | - |
| 11 | 21.1 | 2.28 (*s*) | 10 |

In this study, the compounds isolated from the bark of *Aglaia lanuginose* were the dammarane triterpenoids identified as cabralealactone 1, methyl isoeichlerianate 2, cabraleone 3, ocotillone 4, eihlerialactone 5, eichlerianic acid 6, shoreic acid 7 and two sterols identified as sitosterol 9 and sigmasterol 10. Compounds 2, 6 and 7 may be further classified as 3,4-seco-dammarane triterpene. In addition, we isolated one aromatic compound, 4-hydroxycinnamyl-acetate 8, which previously has only been isolated from *Alpinia galaga* (Zingeberaceae) (Eknmakul *et al.,* 2003). To the best of our knowledge 4-hydroxycinnamyl-acetate has never been isolated from any *Aglaia* species.

All the damarrane triterpenes isolated in this study (1-7) have similar structures, by having the 20, 24-epoxy ring. All of the compounds with *20R* configuration and isolated from *Aglaia lanuginose*, were first reported as isolates from *Cabralea eichleriana (C*ascon *et al*., 1972)*, Cabralea polytricha* (Rao *et al*., 1975) and *Dysoxylum richii* (Aalbersberg *et al*., 1991) plants from the Meliaceae family, as well as from many other plants in other families in subsequent studies. For example, compound 1 was also reported in *Betula platyphylla* (Betulaceae) (Byung *et al.,* 1977) and *Cleome africana* (Cleomaceae) (Tsichritzis *et al.,* 1993). In the case of compound 2 which has *20S* configuration, it has only been reported in recent years as a natural product from *Aglaia silvestris* (Pointinger *et al.,* 2008).

Compounds 3 and 4, as well as compounds 6 and 7 were pairs of stereoisomer where they were different in their stereochemistries at position C-24. The stereochemistries of compound (3, 4) and (6, 7) were determined by comparing the $^{13}C$ chemical shift of C-24 with those in the literature (Hisham *et al.,* 1996), whereby the *R* and *S* stereo-centres at C-24 has distinguishable C-13 chemical shifts of approximately 84.3 and 86.3 ppm respectively. In addition, compound 2 which previously has only been isolated from *Aglaia silvestris* was different from the rest of the dammarane compounds isolated so far due to its uncommon 20*R* stereochemical configuration, instead of the usual 20*S* configuration (Pointinger *et al.,* 2008). The 20*R* stereochemistry of compound 2 was determined by comparing the $^{13}C$-NMR chemical shifts of the neighboring C-21 and C-22 with those in the literature (Seger *et al.,* 2008).

We observed that a mixture of compounds 3 and 4, and a mixture of compounds 6 and 7 to have cytotoxicity against HL60 cells at 5 µg/ ml, where the mixture of 3 and 4 (1:1 ratio) showed 62.0 ± 2.7% inhibition while the mixture of 6 and 7 (1:1 ratio) showed 75.0 ± 3.0% inhibition. All the other fractions, including those that contained mixtures of compounds 1, 2, 5 and 8 until 10 were not active at 5 µg/ ml. From the literature, damarrane type triterpenoids such as compounds 1, 3, 4, 5, 6, 7 have been reported to be cytotoxic in various cell lines. Compound 2 has not been reported for its cytotoxic activity while compound 8 was only reported to have trypanocidal activity (Kuichi *et al.,* 2002). A study on phytochemicals from *Aglaia leucophylla*, Abdelilah *et al.* (1994) showed that compound 3 was cytotoxic against KB cell line. In addition, another study by Charles *et al.* (2010) which showed that the mixtures of cabraleone and ocotillone (3, 4), and eichlerianic acid and shoreic acid (6, 7) were cytotoxic towards HeLa cells. With respect to compounds 1 and 5, the disagreement between our results and the published cytotoxicity data in the literature may be due to differences in the cell lines used. Compound **1** was tested to be potent with micromolar $IC_{50}$ value in P388 leukemia cell line (Hidekazu *et al.,* 1997) and a breast cancer cell line (Jarinporn *et al.,* 2008). In the case of compound **5**, it was reported to be moderately active against NCI-H187 small-cell lung cancer cell line and weakly active towards a breast cancer cell line (Jarinporn *et al.,* 2008).

## CONCLUSION

10 compounds were obtained from the cytotoxic dichloromethane fraction from the bark of *Aglaia lanuginose*. Out of these, the aromatic compound 8, 4-hydroxycinnamyl-acetate has never been isolated from any *Aglaia* species. From the cytotoxicity, we found that some of these compounds are cytotoxic and may have potential as cytotoxic agents for cancer.

## ACKNOWLEDGEMENTS

The authors are grateful to the Malaysian IRPA grant for its financial support.

## REFERENCES

Aalbersberg, W. (1991). Dammarane triterpene from *Disoxylum richii*. *Phytochemistry, 30,* 921.

Aalbersberg, W., & Singh, Y. (1991). Dammarane triterpene from *Disoxylum richii Phytochemistry, 31,* 4033-4035.

Abdelilah, B., Pascal, R., Thierry, S., & Hadi A. H. A. (1994). Secotirucallane triterpenes from the stem bark of *Aglaia leucophylla*. *Phytochemistry, 37,* 1143-145.

Brader, G., Vajrodaya, S., Greger, H., Baher, M., Kalchause, M., & Hofer, O. (1998). Bisamide, lignans, triterpne and insecticidal cyclopenta [b] benzofuran s from *Aglaia* species. *Journal of Natural Products, 61,* 1482-1490.

Byung, H. H., & Byung, J. S. (1977). A new triterpene from betulafolienpentaol from *Betula phatyphyla*. *Phytochemistry, 16*(7), 1075-1078.

Eknmakkul, W. D., & Potduang, B. (2003). Biosynthesis of b-sitosterol and stigmasterol in *Croton sublyratus* proceeds via mixed origin of isoprene units. *Phytochemistry, 6*(2), 389- 398.

Cascon, S. C., & Brown, K. S. (1972). Biogenetically significant triterpenes in the species of Meliaceae. *Cabralea polytricha* A. Juss. *Tetrahedron, 28,* 315.

Charles, O. E. S., Gero, E., Chukwuemeka, S. N., Dennis, H., Klaus, U., & Peter, P. (2010). Dammarelonic acid, a secodammarane triterpenoid from *Aglaia sp*. shows potent anti retroviral activity *in vitro*. *Phytomedicine, 17*, 540-547.

Christiensen, H. (2002). Etnobotany of the Iban and the Kelabit. Joint Publication of Forest Department Sarawak, NEPCon and University of Aarhus. Videbaek, Denmark.

Hidekazu, H., Yonekot, T., Toshikin, N., Hideji, I., Koichi, T., Ahmed, F. H., & Osama, A. B. (1997). Cytotoxic triterpenes from *Cleome africana*. *Phytochemistry, 44*, 1115-1119.

Hisham, A., Ajitha, B. M. D., Fujinoto, Hara, N., & Shimada, H. (1966). Complete H-1 and C-13 spectral assignment of cabraleadiol, dammarane triterpene from *Dysoxylum malabaricum* Bedd. *Magnetic Resonance in Chemistry, 34,* 146-150.

Jarinporn, P., Takuya, K., Tsutomu, I., Rutt, S., & Ekarin, S. (2008). A New sesquiterpene and other terpenoid constituents *Chisochiton penduliflorus*. *Archives of Pharmacal Research, 31*(1), 21-27.

Kuichi, F., Matsuo, K., Itano, Y., Ito, M., Honda, G., Qui T. K., Nakajima, Shimoda, J., & Aoki, T. (2002). Screening of natural medicines used in Vietnam for Tryponocidal activity against epimastigotes of *Trypanosoma cruzi*. *Natural Medicine, 56*(2)*,* 64-68.

Kritikar, K. R., & Basu, B. D. (1975). Indian Medicine Plant. International Vook Distributor, Dehradun, 475.

Lucas, D. M., Edwards, R. B., Lozanski, G., West, D. A., & Shin, J. D. (2009). The novel plant – derived agent silvestrol has B cell selective activity in chronic lymphocytic leukemic and acute lymphoblastic leukaemia in vitro and in vivo. *Blood, 113*, 4656-4666.

Muellner, A. N., Samuel, R., Johnson, S. A., Cheek, M., Pennington, T. D., Chase, M. W. (2003). Molecular phylogenetics of Meliaceae (Sapindales) based on nuclear and plastid DNA sequence. *American Journal of Botany, 90,* 471-480.

Nuanyai, T., Sappapan, R., Vilaivan, T., & Pudhorn, K. (2011). Dammarane triterpenes from the epical buds of Gardenia collinsae. *Phytochem. Lett*. 2011 doi: 10.1016/j.phytol.2011.03.001.

Pointiger, S., Promdang, S., Vajrodaya, S., Pannel, C. M., Hoffer, O., Mereiter, K., & Greger, H. (2008). Silvaglins and related 2,3- secodammarane deravatives- unusual types of triterpenes from *Aglaia silvestris*. *Phytochemistry, 69,* 2696-2703.

Proksh, P., Gaisi, M., Treiber, M. K., Palfi, K., Merling, A., Spring, H., Kramer, P. H., & Li-Waber, M. (2005). Recoglamide derivatives are immunosuppressive phytochemicals that target NF-AT activity in T cells. *Journal of Immunology, 174,* 7075-7084.

Rao, M. M., Meshulam, H., Zelnik, R., Lavie, R., & Lavie, D. (1975). *Cabralea eichleriana*dc. (Meliaceae)- I structure and stereochemistry of wood extractives. *Tetrahedron, 31*, 333-339.

Roux, D., Martin M. T., Adeline, M. T., Sevenet, T., Hadi, H., & Pais, M. (1998). Faveolin A and B dammarane triterpenes from *Aglaia foveolata*. *Phytochemistry*, *49*, 1745-1748.

Seger, C., Pointiger, S., Greger, H., & Hofer, O. (2008). Isoeichlerianic acid from *Aglaia silvestris* and revision of the stereochemistry of faveolin B. *Tetrahedron Letters, 49*, 4313-4315.

Silvia, P., Somnuk, P., Sruya, V., Caroline, M. P., Otmar, H., Kurt, M., & Harald, G. (2008). Silvaglins and the related 2,3 secodammarane derivatives – unusual types of triterpenes from *Aglaia silvestris*. *Phytochemistry, 69,* 2696-2703.

Simanjuntak, P., Parwati, T., Rachmat, J., & Soeksmanto, A. (1999). Screening of some Indonesian *Aglaia* spp. Plants for insecticidal potential by Brine Shrimp Lethality Test (BSLT). *Annales Bogorienses, 6,* 61-62.

Tsichritzis, F., Mogib, M. A., & Jakupovic, J. (1993). Dammarane tritepenes from *Cleome africana. Phytochemistry, 33,* 423-425.

Pertanika J. Sci. & Technol. 22 (1): 175 - 192 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Preconditioned Subspace Quasi-Newton Method for Large Scale Optimization

**Hong Seng Sim[1]*, Wah June Leong[2], Malik Abu Hassan[2] and Fudziah Ismail[2]**

*[1]Institute for Mathematical Research, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
*[2]Department of Mathematics, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

Subspace quasi-Newton (SQN) method has been widely used in large scale unconstrained optimization problem. Its popularity is due to the fact that the method can construct subproblems in low dimensions so that storage requirement as well as the computation cost can be minimized. However, the main drawback of the SQN method is that it can be very slow on certain types of non-linear problem such as ill-conditioned problems. Hence, we proposed a preconditioned SQN method, which is generally more effective than the SQN method. In order to achieve this, we proposed that a diagonal updating matrix that was derived based on the weak secant relation be used instead of the identity matrix to approximate the initial inverse Hessian. Our numerical results show that the proposed preconditioned SQN method performs better than the SQN method which is without preconditioning.

*Keywords:* Preconditioned, subspace method, limited memory quasi-Newton methods, large scale, unconstrained optimization

## INTRODUCTION

Subspace quasi-Newton (SQN) method is generally used to solve large scale non-linear systems of equations and non-linear least square problems. This method is popular because it has the characteristic to force the next iteration in a low dimensional subspace. At each iteration, we searched for a minimum of the objective function over a subspace spanned by the current gradient and by direction of few previous steps.

The main advantage of this method is that it constructs subproblems in low dimensions so that computation cost can be reduced. It also offers a possible way to handle large scale unconstrained optimization problems. Besides, this method can be implemented extremely fast. This happens when the

*Article history:*
Received: 22 September 2011
Accepted: 11 November 2011

*E-mail addresses:*
hongseng0505@hotmail.com (Hong Seng Sim),
leongwj@upm.edu.my (Wah June Leong),
malik@science.upm.edu.my (Malik Abu Hassan),
fudziah@upm.edu.my (Fudziah Ismail)
*Corresponding Author

ISSN: 0128-7680 

objective function is a combination of expensive linear mappings with computationally cheap nonlinear functions (Yuan, 2007).

One of the famous subspace algorithms for non-linear optimization is the unbalance property shared by most line search algorithms. Any line search method is considered to have the following form:

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1}$$

Where, $d_k$ is the search direction and $\alpha_k \geq 0$ is the step-length that is computed by certain line search technique. Generally, the search direction $d_k$ is computed by solving a subproblem which is an approximation to the original non-linear optimization problem. Therefore, there are two parts combined in each iteration of a line search algorithms; the first part is to find $d_k$ in the whole $n$ dimensional space, while the other part is to search for a suitable step-length in a fixed one dimensional space spanned by the computed $d_k$. As a result, the overall algorithm swings between the $n$ dimensional search and one dimensional search alternately. Some variants of these methods can be found in [3], [4], [5], [9], [10], [11], [12], [17] and [18].

Furthermore, many well-known existing algorithms essentially have certain subspace features. For example, the conjugate gradient method uses a search direction in a two dimensional subspace spanned by the steepest descent direction and the previous step, the dog-leg method computes a step that is a convex combination of the steepest descent direction and the Newton's direction, and the limited memory quasi-Newton algorithms will also produce search directions that are spanned in a lower dimensional space to speed up the convergency and lower the computation cost.

## SUBSPACE METHOD APPROACHES

The well-known nonlinear conjugate gradient methods use a linear combination of the steepest descent direction $-g_k$ and the previous search direction $d_{k-1}$ to form the new search direction, as follows:

$$d_k = -g_k + \beta_k d_{k-1}$$

Hence, one of the important tasks is how to determine the suitable $\beta_k$ based on certain conjugate gradient principles. Instead of the conjugate property, Stoer and Yuan (1995) suggested to look at the conjugate gradient method from the subspace point of view. In the conjugate gradient method, $\beta_k$ is used to define search direction $d_k$ and the stepsize $\alpha_k$ to set $x_{k+1} = x_k + \alpha_k d_k$; thus, no matter whatever $\beta_k$ and $\alpha_k$ are used, the increment in the iterative point will be a linear combination of $-g_k$ and $d_{k-1}$. They consider a model subproblem as follows:

$$\min_{d \in span\{-g_k, s_{k-1}\}} Q_k(d) \approx f(x_k + d)$$

Let $d_k$ be the solution of the above 2-dimensional subproblem and a successive 2-dimensional search algorithm is presented, which is an example of algorithms that using subspace methods (Stoer & Yuan, 1995).

Limited memory quasi-Newton method also has the subspace nature. The Quasi-Newton updates have the following form:

$$B_k = U\left(B_{k-1}, s_{k-1}, y_{k-1}\right)$$

which satisfies

$$B_k s_{k-1} = y_{k-1}$$

where, $s_{k-1} = x_k - x_{k-1}$ and $y_{k-1} = g_k - g_{k-1}$. A famous example is the BFGS method.

$$B_k = B_{k-1} - \frac{B_{k-1} s_{k-1} s_{k-1}^T B_{k-1}}{s_{k-1}^T B_{k-1} s_{k-1}} + \frac{y_{k-1} y_{k-1}^T}{s_{k-1}^T y_{k-1}}$$

The limited memory quasi-Newton updates the approximate Hessian repeatedly:

$$B_k^{(i)} = U\left(B_k^{(i-1)}, s_{k-m-1+i}, y_{k-m-1+i}\right) \qquad i = 1, 2, \ldots, m$$

with $B_k^{(0)} = \sigma_k I$ (Liu & Nocedal, 1989). There are various formulae for $\sigma_k$, with one choice being $\frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}}$. The limited memory quasi-Newton matrix can be written as follows:

$$B_k = B_k^{(m)} = \sigma_k I + \begin{bmatrix} S_k & Y_k \end{bmatrix} T_k \begin{bmatrix} S_k^T \\ Y_k^T \end{bmatrix}$$

where, $T_k$ is a $2m \times 2m$ matrix, and

$$\begin{bmatrix} S_k & Y_k \end{bmatrix} = \left[s_{k-1}, s_{k-2}, \ldots, s_{k-m}, y_{k-1}, y_{k-2}, \ldots, y_{k-m}\right] \in \Re^{n \times 2m}.$$

We have $s_k = \alpha_k d_k = -\alpha_k B_k^{-1} g_k$ in the line search type method, while for a trust region type algorithm, we have $s_k = -\left(B_k + \lambda_k I\right)^{-1} g_k$. As a result,

$$s_k = -\left(\rho_k I + \begin{bmatrix} S_k & Y_k \end{bmatrix} T_k \begin{bmatrix} S_k^T \\ Y_k^T \end{bmatrix}\right)^{-1} g_k$$

$$\in span\left\{g_k, s_{k-1}, \ldots, s_{k-m}, y_{k-1}, \ldots, y_{k-m}\right\}.$$

It has been shown that no matter what, the limited memory quasi-Newton algorithm with line search or trust region will always produce a step in the subspace $span\left\{g_k, s_{k-1}, \ldots, s_{k-m}, y_{k-1}, \ldots, y_{k-m}\right\}$ (Wang *et al*., 2004).

A model subspace algorithm for unconstrained optimization is suggested, which is a slight modification of the standard trust region algorithm for unconstrained optimization.

**Algorithm 2.1 (A Model Subspace Algorithm for Unconstrained Optimization)**

**Step 1**: Given $x_1$, define $S_1$, $\varepsilon > 0$, $k := 1$.

**Step 2**: Solve a subspace subproblem:

$$\min_{d \in S_k} Q_k(d) = g_k^T d + \frac{1}{2} d^T B_k d \,. \tag{2}$$

to obtain $s_k$. If $\| s_k \| \leq \varepsilon$, then stop.

**Step 3**: Define

$x_{k+1} = x_k + s_k$ if $f(x_k + s_k) < f(x_k)$

$x_{k+1} = x_k$, otherwise.

**Step 4**: Generate $S_{k+1}$ and $Q_{k+1}(d)$.

**Step 5**: Set $k := k + 1$, go to step 2.

The main difference between the above algorithm and the standard whole space algorithm is the constraint for the step $S_k$ to be in the subspace $S_k$. Thus, the key issue here is how to choose the subspace $S_k$. Stoer and Yuan (1995) suggested that the choice for the subspace $S_k$ is a generalization of the 2-dimensional subspace, namely, $S_k = span\{-g_k, s_{k-1}, \ldots, s_{k-m}\}$, since all the points in $S_k$ can be expressed by:

$$d = -\sigma g_k + \sum_{i=1}^{m} \beta_i s_{k-i} \,, \tag{3}$$

using the following approximations,

$$s_{k-i}^T \nabla^2 f(x_k) s_{k-j} \approx s_{k-i}^T y_{k-j} \,, \quad s_{k-i}^T \nabla^2 f(x_k) g_k \approx y_{k-i}^T g_k \,.$$

However, the performance of a CG-like search direction can be very slow on certain types of non-linear problem such as ill-conditioned problems. Hence, the main aim of the study is to propose some preconditioners for the search direction (3), namely:

$$d_k = -D_k^{-1} g_k + \sum_{i=1}^{m} \beta_i s_{k-i} \tag{4}$$

where $D_k$ is the preconditioner in diagonal matrix form, and it is supposed to have some properties of the Hessian matrix, or a good approximation to the Hessian matrix in some sense.

## DERIVATION OF THE DIAGONAL PRECONDITIONER

In this section, we develop a preconditioner for subspace quasi-Newton algorithm in order to overcome the deficiency of the standard subspace algorithm when solving ill-conditioned optimization problems.

We shall choose a diagonal matrix $D_k$ that satisfies the weak-quasi-Newton relation, as below:

$$y_k^T D_{k+1} s_k = y_k^T y_k \tag{5}$$

where, $y_k = g_{k+1} - g_k$, and $s_k = x_{k+1} - x_k$.

Suppose that the Hessian matrix $A$ of an objective function $f(x) = \frac{1}{2} x^T Ax - b^T x$ is positive definite. We let $D_k$ be a diagonal matrix to approximate the Hessian matrix. Hence, we form our approximation as follows:

$$D_{k+1} = D_k + \Delta_k \tag{6}$$

Our purpose is to construct a $D_{k+1}$ in such a way that it is a good approximation to the actual Hessian matrix.

### Theorem 3.1

Assume that $D_k > 0$ is a positive definite diagonal matrix and $D_{k+1}$ is the updated version of $D_k$, which is also diagonal. Suppose that $s_k \neq 0$, the optimal solution of the following minimization problem will then be:

$$\text{minimize } \frac{1}{2} \left\| \Delta_k \right\|_F^2$$

$$\text{subject to } y_k^T D_{k+1} s_k = y_k^T y_k \tag{7}$$

and is given by:

$$D_{k+1} = D_k + \frac{\omega_k - \mu_k}{\gamma_k} G_k \tag{8}$$

where, $\|\Delta_k\|_F = \sqrt{tr\left(\Delta_k^T \Delta_k\right)}$ is the Frobenius norm and *tr* is the trace operator, $\omega = y_k^T y_k$, $\mu = y_k^T D_k s_k$, $\gamma = \sum_{i=1}^{n} \left( y_k^{(i)} s_k^{(i)} \right)^2$ and $G_k = diag\left( \left( s_k^{(1)} y_k^{(1)} \right), \ldots, \left( s_k^{(i)} y_k^{(i)} \right) \right)$ with $y_k^{(i)}$ and $s_k^{(i)}$ being the $i = th$ component of the $y_k$ and $s_k$ respectively.

**Proof**

Let $\Delta_k = \begin{pmatrix} a_k^{(1)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_k^{(k)} \end{pmatrix}$, $s_k = \begin{pmatrix} s_k^{(1)} \\ \vdots \\ s_k^{(n)} \end{pmatrix}$ and $y_k = \begin{pmatrix} y_k^{(1)} \\ \vdots \\ y_k^{(k)} \end{pmatrix}$ for $i = 1,2,3,\dots,n.$

From equation (7), we have:

$$\|\Delta_k\|^2 = \left(\sqrt{tr\left(\Delta_k\right)^T\left(\Delta_k\right)}\right)^2$$

$$= \left(\left(a_k^{(1)}\right)^2 + \dots + \left(a_k^{(i)}\right)^2 + \dots + \left(a_k^{(n)}\right)^2\right). \tag{9}$$

Thus, the minimization equation will become:

$$\text{minimize } \frac{1}{2}\left(\left(a_k^{(1)}\right)^2 + \dots + \left(a_k^{(i)}\right)^2 + \dots + \left(a_k^{(n)}\right)^2\right). \tag{10}$$

By substituting (6) into (7), we obtain:

$$y_k^T\left(D_k + \Delta_k\right)s_k = y_k^T y_k\,. \tag{11}$$

We expand (11) to get the following expression:

$$y_k^T D_k s_k + y_k^T \Delta_k s_k = y_k^T y_k\,.$$

Rearranging the equation, we get:

$$\mu - \omega + \sum_{i=1}^{n} y_k^{(i)} s_k^{(i)} a_k^{(i)} = 0\,, \tag{12}$$

where $\mu = y_k^T D_k s_k$ and $\omega = y_k^T y_k$.

From (12), we have:

$$\sum_{i=1}^{n} y_k^{(i)} s_k^{(i)} a_k^{(i)} = \omega - \mu\,. \tag{13}$$

Finally, we wish to solve the following:

$$\text{minimize } \frac{1}{2}\left(\left(a_k^{(1)}\right)^2 + \dots + \left(a_k^{(i)}\right)^2 + \dots + \left(a_k^{(n)}\right)^2\right)$$

$$\text{subject to } \mu - \omega + \sum_{i=1}^{n} y_k^{(i)} s_k^{(i)} a_k^{(i)} = 0\,. \tag{14}$$

Since the objective function in (14) is convex, there exists a unique solution and its Lagrange function will be:

$$L = \frac{1}{2}\left(\left(a_k^{(1)}\right)^2 + \ldots + \left(a_k^{(n)}\right)^2\right) + \lambda\left(\mu - \omega + \sum_{i=1}^{n} y_k^{(i)} s_k^{(i)} a_k^{(i)}\right), \tag{15}$$

where, $\lambda$ is the Lagrange multiplier associated with the constant. We differentiate (15) with respect to $a_k^{(i)}$, and setting the result to zero, we obtain,

$$\frac{\partial L}{\partial a_k^{(i)}} = a_k^{(i)} + \lambda y_k^{(i)} s_k^{(i)} = 0. \tag{16}$$

From (16), it is clear that,

$$\lambda y_k^{(i)} s_k^{(i)} = -a_k^{(i)}. \tag{17}$$

Multiplying (17) with $s_k^{(i)} y_k^{(i)}$ for $i = 1,2,3,\ldots,n$, respectively, we shall obtain

$$\lambda\left(y_k^{(i)} s_k^{(i)}\right)^2 = -y_k^{(i)} s_k^{(i)} a_k^{(i)}. \tag{18}$$

Summing all of the equations in (18) yields:

$$\lambda\sum_{i=1}^{n}\left(y_k^{(i)} s_k^{(i)}\right)^2 = -\sum_{i=1}^{n} y_k^{(i)} s_k^{(i)} a_k^{(i)}. \tag{19}$$

By equation (13), (19) becomes

$$\lambda\sum_{i=1}^{n}\left(y_k^{(i)} s_k^{(i)}\right)^2 = \mu - \omega, \tag{20}$$

Finally, we get

$$\lambda = \frac{\mu - \omega}{\gamma}, \tag{21}$$

where $\gamma = \sum_{i=1}^{n}\left(y_k^{(i)} s_k^{(i)}\right)^2$.

Once again, from (17), we get

$$a_k^{(i)} = -\lambda y_k^{(i)} s_k^{(i)}. \tag{22}$$

We substitute (21) into (22), the equation becomes,

$$a_k^{(i)} = \frac{\omega_k - \mu_k}{\gamma_k} y_k^{(i)} s_k^{(i)}. \tag{23}$$

Expression (23) is in the form of each component of $\Delta$. By substituting (23) into the formula of $\Delta_k$, we will get the approximation of $D_{k+1}$ as follows:

$$D_{k+1} = D_k + \frac{\omega_k - \mu_k}{\gamma_k} G_k, \tag{24}$$

where $\omega = y_k^T y_k$, $\mu = y_k^T D_k s_k$, $\gamma = \sum_{i=1}^{n} \left(y_k^{(i)} s_k^{(i)}\right)^2$ and $G_k = diag\left(\left(s_k^{(1)} y_k^{(1)}\right), \ldots, \left(s_k^{(i)} y_k^{(i)}\right)\right)$ with $y_k^{(i)}$ and $s_k^{(i)}$ being the $i-th$ component of the $y_k$ and $s_k$ respectively, and the proof is completed.

Now, we give our algorithm for solving large-scale unconstrained optimization, which is called the preconditioned subspace quasi-Newton algorithm.

**Algorithm 3.1 SQN Algorithm**

**Step 1** : Set $k = 0$; select the initial point $x_0$ and $\varepsilon$ as a stopping condition. We also set $D_0 = I$, where $I$ is $n \times n$ identity matrix.

**Step 2** : For $k \geq 0$, compute $g_k = Ax_k - b$. If $\|g_k\| \leq \varepsilon$, stop, else compute $D_k$, where $D$ is a specific diagonal preconditioner.

**Step 3** : Compute $d_{k+1} = -D_{k+1} g_{k+1} + \sum_{i=1}^{m} \beta_i s_{k+1-i}$, where $\beta_i = \frac{g_{i+1}^T A d_i}{d_i^T A d_i}$, $i \leq \min\{k, m\}$.

**Step 4** : Compute $\alpha_k = -\frac{g_k^T d_k}{d_k^T A d_k}$.

**Step 5** : Hence, $x_{k+1} = x_k + \alpha_k d_k$.

**Step 6** : Set $k := k + 1$; go to step 2.

The SQN method is tested where in Step 2, $D$ is chosen from theorem 3.1.

## CONVERGENCE ANALYSIS

In this section, we shall look at the convergence properties of the subspace quasi-Newton method. Note that all the Hessian approximations are obtained by updating a bounded matrix using the proposed preconditioned subspace quasi-Newton method. We will prove the convergence properties of our proposed methods based upon the convergence assumptions given by Liu and Nocedal (1989), since it is valid for our preconditioning formulae whose matrices are diagonal and positive definite.

**Assumption 4.1:**

(1) The objective function $f$ is twice continuously differentiable.

(2) The level set $D = \{x \in \Re^n : f(x) \leq f(x_0)\}$ is convex.

(3) There exist positive constants $M_1$ and $M_2$ such that

$$M_1\|z\|^2 \leq z^T G(x) z \leq M_2\|z\|^2 \tag{25}$$

for $\forall z \in \Re^n$ and $\forall z \in D$. This implies that the objective function $f$ has a unique minimize $x^*$ in $D$.

From (25), we can have another similar inequality, as below:

$$N_1\|z\|^2 \leq z^T G(x)^{-1} z \leq N_2\|z\|^2, \tag{26}$$

where $N_1 = \dfrac{1}{M_2}$ and $N_2 = \dfrac{1}{M_1}$ are the constants.

**Lemma 4.1**

Let $x_0$ be a starting point for which $f$ satisfies Assumptions 4.1, and we take $D_0 = I$, where $I$ is the $n \times n$ identity matrix. Assume that the matrices $D_k^0$ are chosen so that $\{\|D_k^{(0)}\|\}$ and $\{\|D_k^{(0)-1}\|\}$ are bounded. Then, $\{D_{k+1}\}$ and $\{\|D_{k+1}^{-1}\|\}$ are also bounded, where,

$$D_{k+1} = D_k + \frac{\omega_k - \mu_k}{\gamma_k} G_k \tag{27}$$

where $\omega = y_k^T y_k$, $\mu = y_k^T D_k s_k$, $\gamma = \sum_{i=1}^{n} \left(y_k^{(i)} s_k^{(i)}\right)^2$ and $G_k = diag\left(\left(s_k^{(1)} y_k^{(1)}\right), \ldots, \left(s_k^{(i)} y_k^{(i)}\right)\right)$ with $y_k^{(i)}$ and $s_k^{(i)}$ being the $i-th$ component of the $y_k$ and $s_k$, respectively.

**Proof**

Without the loss of generality, we shall assume that $D_0 = I$, where $I$ is the $n \times n$ identity matrix. It is clear that $D_0$ is bounded, as follows:

$$\mu_0 \leq \| D_0 |_F \leq \omega_0$$

Now, we need to prove $D_1$ is bounded.

Let $\nabla^2 f\left(\bar{x}\right)$ be defined as:

$$\nabla^2 f\left(\bar{x}\right) = \int_0^1 \nabla^2 f\left(x_k + \tau s_k\right) d\tau .$$

Then, we have,

$$y_k = \nabla^2 f\left(\bar{x}\right) s_k . \tag{28}$$

From (26) and (28), we get

$$N_1 \left\| y_k \right\|^2 \le y_k^T s_k \le N_2 \left\| y_k \right\|^2 , \tag{29}$$

Where, $N_1 = \dfrac{1}{M_2}$ and $N_2 = \dfrac{1}{M_1}$ are the constants.

From (27), we have

$$\left\| D_1 \right\|_F = \left\| D_0 + \frac{\omega_0 - \mu_0}{\gamma_0} G_0 \right\|_F$$

$$\left\| D_1 \right\|_F \le \left\| D_0 \right\|_F + \left\| \frac{\omega_0 - \mu_0}{\gamma_0} G_0 \right\|_F$$

$$\left\| D_1 \right\|_F \le \left\| D_0 \right\|_F + \frac{\left(\omega_0 - \mu_0\right)}{\gamma_0} \left\| G_0 \right\|_F , \tag{30}$$

where $\left\| \cdot \right\|_F^2$ is the square of Frobenius norm and let *tr* be the trace operator.

Note that

$$\left\| y_0 \right\|^2 = y_0^{(1)2} + y_0^{(2)2} + \ldots + y_0^{(n)2}$$

$$\le n\left(y_0^M\right)^2 , \tag{31}$$

where $\left(y_0^M\right)^2 = \max\left\{y_0^{(1)2}, y_0^{(2)2}, \ldots, y_0^{(n)2}\right\}$.

From (29), (30) and (31), we will get

$$\begin{aligned}\left\|G_0\right\|_F^2 &= tr\left(G_0^T G_0\right)\\ &= s_0^{(1)2} y_0^{(1)2} + s_0^{(2)2} y_0^{(2)2} + \ldots + s_0^{(n)2} y_0^{(n)2}\\ &\le \left(y_0^T s_0\right)^2\\ \| G_0 \| &\le N_2 n\left(y_0^M\right)^2 .\end{aligned} \tag{32}$$

From (27), we have

$$\begin{aligned}\omega_0 &= y_0^T y_0\\ &= \left\|y_0\right\|^2\\ &\le n\left(y_0^M\right)^2\end{aligned} \tag{33}$$

and

$$\begin{aligned}\mu_0 &= y_0^T D_0 s_0\\ \mu_0 &\le M_3 y_0^T s_0\\ \mu_0 &\le M_3 N_2 \left\|y_0\right\|^2\\ \mu_0 &\le M_3 N_2 \left(y_0^M\right)^2\end{aligned} \tag{34}$$

where, $M_3 = 1$.

Hence, from (30), (32), (33) and (34), we shall have

$$\left\|D_1\right\|_F \le \left\|D_0\right\|_F + \frac{\left(n\left(y_0^M\right)^2 - N_2 n\left(y_0^M\right)^2\right)}{\sum_{i=1}^{n}\left(s_0^{(i)} y_0^{(i)}\right)^2} N_2 n\left(y_0^M\right)^2$$

$$\left\|D_1\right\|_F \le \left\|D_0\right\|_F + \frac{\left(n - N_2 n\right)\left(y_0^M\right)^2}{\sum_{i=1}^{n}\left(s_0^{(i)} y_0^{(i)}\right)^2} N_2 n\left(y_0^M\right)^2$$

$$\left\|D_1\right\|_F \le \left\|D_0\right\|_F + \frac{\left(1 - N_2\right) N_2 n^2 \left(y_0^M\right)^4}{\sum_{i=1}^{n}\left(s_0^{(i)} y_0^{(i)}\right)^2}$$

$$\|D_1\|_F \le \|D_0\|_F + \frac{kN_2 n^2 \left(y_0^M\right)^4}{\sum_{i=1}^{n}\left(s_0^{(i)} y_0^{(i)}\right)^2}$$

$$\|D_1\|_F \le \|D_0\|_F + M_4$$

$$\|D_1\|_F \le n + M_4, \tag{35}$$

Where, $M_4 = kN_2 n^2$ and $k = \max\{(1-N_2),(1+N_2)\}$, and

$$\frac{\left(y_0^M\right)^4}{\sum_{i=1}^{n}\left(s_0^{(i)} y_0^{(i)}\right)^2} \le 1 \tag{36}$$

From (35), we can conclude that $\|D_1\|_F$ is bounded since $\|D_0\|_F$ is also bounded. Now, we assume that $D_k$ is bounded, and then, we need to prove that $D_{k+1}$ is also bounded.

From the above, we shall get a similar equation and inequality, as follows:

$$\|G_k\|_F \le N_2 n\left(y^M\right)^2, \tag{37}$$

$$\omega_k \le \|y_k\|^2, \tag{38}$$

$$\mu_k \le M_3 N_2 \|y_k\|^2, \tag{39}$$

$$\|y_k\|^2 \le n\left(y_k^M\right)^2. \tag{40}$$

From (27) and (37)-(40), we obtain

$$\|D_{k+1}\|_F \le \|D_k\|_F + M_4, \tag{41}$$

Where, $M_4 = kN_2 n^2$ and $k = \max\{(1-M_3N_2),(1+M_3N_2)\}$.

From the fact that $\|D_k\|_F$ is bounded, i.e. $\|D_k\|_F \le M_5$, and from (41),

$$\|D_{k+1}\|_F \le M_5 + M_6$$

$$\|D_{k+1}\|_F \le M_6,$$

Where, $M_6 = M_5 + M_4$ and it is a constant. Finally, we have shown that $\|D_{k+1}\|_F$ is bounded and the proof is completed.

In this section, we have shown that the proposed preconditioned subspace quasi-Newton methods are convergent on uniformly convex problems and the rate is $R$-linear. This $R$-linear convergence results obtained are based upon the assumption by Liu and Nocedal (1989).

## COMPUTATIONAL RESULTS AND DISCUSSION

The computational results and discussion on the performance of preconditioner subspace quasi-Newton (SQN) method are given in this section. All the algorithms are written in MATLAB 7.0. The total number of tested problems is 4. All the runs were terminated when

$$\|g_k\| \leq 10^{-4}$$

Where, $\|\cdot\|$ denotes the Euclidean norm. Furthermore, we also consider the number of function evaluation and gradient calls. We set our upper bound for the number of function evaluation and gradient call to be 1000.

The computational results are compared through the number of iterations, gradient evaluations as well as function evaluations. In order to test the efficiency of the proposed preconditioned methods, the number of subspaces that is considered is $m = 2$ and $m = 3$.

Meanwhile, the SQN method was tested using the following preconditioners:

1. SQN(0)-SQN method without preconditioning.
2. SQN(D1)-SQN method with diagonal preconditioner $D$, where $D$ is given by Theorem 3.1.

In order to compare the efficiency of the proposed preconditioned SQN methods with the standard SQN method, the following quadratic test problem is considered:

$$f(x) = \frac{1}{2} x^T A x - b^T x, \tag{42}$$

where, $A$ is positive definite diagonal matrix and $b = [1,1,1,1,1,\ldots,1]$.

For all the methods, the initial point is $x_0 = [0,0,0,0,\ldots,0]$. A set of unconstrained minimization quadratic problems consisting of 4 test problems were used. We now describe the 4 different quadratic test problems (43) with $n$-dimensional cases.

1. QF1, where $A = diag[a_{ii}]$, $a_{ii} = i^2 \pmod 5$, $b = [1,\ldots,1]$.
2. QF2, where $A = diag[a_{ii}]$, $a_{ii} = i^3 \pmod 5$, $b = [1,\ldots,1]$.
3. QF3, where $A = diag[a_{ii}]$, $a_{ii} = i^3 + i \pmod 5$, $b = [1,\ldots,1]$.
4. QF4, where $A = diag[a_{ii}]$, $a_{ii} = a_{i-2,i-2} + a_{i-1,i-1}$, $i \geq 3$ and $a_{11} = 1$, $a_{22} = 1$, $b = [1,\ldots,1]$.

We tested the above problems by using $m = 2$ and $m = 3$. In each table, the symbol Ite, $\|g_k\|$, and Fva mean the number of iterations, norm of the gradient and function evaluation, respectively.

TABLE 1
A comparison of the Method of $m = 2$ in solving QF1

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 10 | 106 | 9.2e-5 | -1.4636 | 40 | 6.2e-5 | -1.4636 |
| 20 | 109 | 9.8e-5 | -2.9272 | 40 | 8.8e-5 | -2.9272 |
| 40 | 113 | 9.6e-5 | -5.8544 | 41 | 6.6e-5 | -5.8544 |
| 80 | 117 | 9.3e-5 | -1.1709e+1 | 41 | 9.3e-5 | -1.1709e+1 |
| 100 | 118 | 9.5e-5 | -1.4636e+1 | 45 | 7.6e-5 | -1.4636e+1 |
| 200 | 122 | 9.3e-5 | -2.9272e+1 | 46 | 9.2e-5 | -2.9272e+1 |
| 500 | 127 | 9.2e-5 | -7.3181e+1 | 52 | 8.7e-5 | -7.3181e+1 |
| 1000 | 130 | 9.9e-5 | -1.4636e+2 | 53 | 8.9e-5 | -1.4636e+2 |
| 1500 | 133 | 9.1e-5 | -2.1954e+2 | 56 | 4.6e-5 | -2.1954e+2 |
| 2000 | 134 | 9.6e-5 | -2.9272e+2 | 56 | 5.3e-5 | -2.9272e+2 |

TABLE 2
A comparison of the Method of $m = 2$ in solving QF2

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 10 | 598 | 9.9e-5 | -1.1857 | 66 | 9.8e-5 | -1.1857 |
| 20 | 619 | 9.9e-5 | -2.3713 | 78 | 4.7e-5 | -2.3713 |
| 40 | 640 | 9.9e-5 | -4.7426 | 78 | 6.6e-5 | -4.7426 |
| 80 | 661 | 9.9e-5 | -9.4853 | 79 | 2.5e-5 | -9.4853 |
| 100 | 668 | 9.9e-5 | -1.1857e+1 | 79 | 3.5e-5 | -1.1857e+1 |
| 200 | 689 | 9.9e-5 | -2.3713e+1 | 79 | 5.5e-5 | -2.3713e+1 |
| 500 | 716 | 1.0e-4 | -5.9283e+1 | 79 | 6.2e-5 | -5.9283e+1 |
| 1000 | 737 | 1.0e-4 | -1.1857e+2 | 91 | 1.0e-5 | -1.1857e+2 |
| 1500 | 750 | 9.9e-5 | -1.7785e+2 | 102 | 7.0e-5 | -1.7785e+2 |
| 2000 | 758 | 1.0e-4 | -2.3713e+2 | 116 | 3.1e-5 | -2.3713e+2 |

TABLE 3
A comparison of the Method of $m = 2$ in solving QF3

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 10 | 311 | 9.7e-5 | -6.5573e-1 | 48 | 9.7e-5 | -6.5573e-1 |
| 20 | 322 | 9.7e-5 | -1.3115 | 54 | 8.1e-5 | -1.3115 |
| 40 | 332 | 1.0e-4 | -2.6229 | 58 | 7.4e-5 | -2.6229 |
| 80 | 343 | 1.0e-4 | -5.2459 | 64 | 9.4e-5 | -5.2459 |
| 100 | 347 | 9.8e-5 | -6.5573 | 66 | 7.2e-5 | -6.5573 |
| 200 | 358 | 9.8e-5 | -1.3115e+1 | 67 | 5.3e-5 | -1.3115e+1 |
| 500 | 372 | 9.9e-5 | -3.2787e+1 | 67 | 8.4e-5 | -3.2787e+1 |
| 1000 | 383 | 9.9e-5 | -6.5573e+1 | 68 | 5.3e-5 | -6.5573e+1 |
| 1500 | 390 | 9.7e-5 | -9.8360e+1 | 68 | 6.4e-5 | -9.8360e+1 |
| 2000 | 394 | 9.9e-5 | -1.3115e+2 | 68 | 7.4e-5 | -1.3115e+2 |

TABLE 4
A comparison of the Method of $m = 2$ in solving QF4

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 10 | 252 | 9.7e-5 | -1.6652 | 74 | 9.6e-5 | -1.6652 |
| 20 | 261 | 9.7e-5 | -3.3305 | 77 | 9.1e-5 | -3.3305 |
| 40 | 270 | 9.6e-5 | -6.6609 | 85 | 8.8e-5 | -6.6609 |
| 80 | 278 | 1.0e-4 | -1.3322e+1 | 86 | 8.1e-5 | -1.3322e+1 |
| 100 | 281 | 9.9e-5 | -1.6652e+1 | 86 | 9.1e-5 | -1.6652e+1 |
| 200 | 290 | 9.9e-5 | -3.3305e+1 | 92 | 8.2e-5 | -3.3305e+1 |
| 500 | 301 | 9.8e-5 | -8.3262e+1 | 100 | 1.0e-5 | -8.3262e+1 |
| 1000 | 311 | 9.7e-5 | -1.6652e+1 | 103 | 7.4e-5 | -1.6652e+1 |
| 1500 | 316 | 9.8e-5 | -2.4979e+1 | 103 | 9.5e-5 | -2.4979e+1 |
| 2000 | 320 | 9.7e-5 | -3.3305e+2 | 106 | 9.4e-5 | -3.3305e+2 |

TABLE 5
Comparison of the Method of $m = 3$ in solving QF1

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 10 | 81 | 9.8e-5 | -1.4636 | 36 | 8.6e-5 | -1.4636 |
| 20 | 84 | 9.6e-5 | -2.9272 | 38 | 8.1e-5 | -2.9272 |
| 40 | 87 | 9.4e-5 | -5.8544 | 39 | 7.7e-5 | -5.8544 |
| 80 | 90 | 9.3e-5 | -1.1709e+1 | 40 | 5.9e-5 | -1.1709e+1 |
| 100 | 91 | 9.2e-5 | -1.4636e+1 | 40 | 6.6e-5 | -1.4636e+1 |

TABLE 5 (*continue*)

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 200 | 94 | 9.0e-5 | -2.9272e+1 | 40 | 9.4e-5 | -2.9272e+1 |
| 500 | 97 | 9.9e-5 | -7.3181e+1 | 47 | 5.6e-5 | -7.3181e+1 |
| 1000 | 100 | 9.7e-5 | -1.4636e+2 | 47 | 8.0e-5 | -1.4636e+2 |
| 1500 | 102 | 9.4e-5 | -2.1954e+2 | 47 | 9.8e-5 | -2.1954e+2 |
| 2000 | 103 | 9.6e-5 | -2.9272e+2 | 48 | 9.5e-5 | -2.9272e+2 |

TABLE 6
A comparison of the Method of $m = 3$ in solving QF2

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 10 | 577 | 1.0e-4 | -1.1857 | 74 | 8.9e-5 | -1.1857 |
| 20 | 598 | 9.8e-5 | -2.3713 | 80 | 6.1e-5 | -2.3713 |
| 40 | 618 | 9.9e-5 | -4.7426 | 80 | 8.6e-5 | -4.7426 |
| 80 | 638 | 9.9e-5 | -9.4853 | 87 | 1.5e-5 | -9.4853 |
| 100 | 645 | 9.9e-5 | -1.1857e+1 | 87 | 6.2e-5 | -1.1857e+1 |
| 200 | 665 | 9.9e-5 | -2.3713e+1 | 89 | 5.3e-5 | -2.3713e+1 |
| 500 | 692 | 9.9e-5 | -5.9283e+1 | 89 | 8.8e-5 | -5.9283e+1 |
| 1000 | 712 | 9.9e-5 | -1.1857e+2 | 90 | 7.9e-5 | -1.1857e+2 |
| 1500 | 724 | 9.9e-5 | -1.7785e+2 | 92 | 7.6e-5 | -1.7785e+2 |
| 2000 | 732 | 9.9e-5 | -2.3713e+2 | 96 | 3.2e-5 | -2.3713e+2 |

TABLE 7
A comparison of the Method of $m = 3$ in solving QF3

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 10 | 300 | 9.8e-5 | -6.5573e-1 | 44 | 8.2e-5 | -6.5573e-1 |
| 20 | 310 | 1.0e-4 | -1.3115 | 58 | 7.0e-5 | -1.3115 |
| 40 | 321 | 9.8e-5 | -2.6229 | 58 | 1.0e-4 | -2.6229 |
| 80 | 332 | 9.7e-5 | -5.2459 | 60 | 4.6e-5 | -5.2459 |
| 100 | 335 | 9.8e-5 | -6.5573 | 60 | 5.2e-5 | -6.5573 |
| 200 | 346 | 9.7e-5 | -1.3115e+1 | 60 | 7.3e-5 | -1.3115e+1 |
| 500 | 359 | 1.0e-4 | -3.2787e+1 | 62 | 9.0e-5 | -3.2787e+1 |
| 1000 | 370 | 9.9e-5 | -6.5573e+1 | 63 | 6.1e-5 | -6.5573e+1 |
| 1500 | 376 | 9.9e-5 | -9.8360e+1 | 63 | 7.5e-5 | -9.8360e+1 |
| 2000 | 381 | 9.7e-5 | -1.3115e+2 | 63 | 8.7e-5 | -1.3115e+2 |

TABLE 8
A comparison of the Method of $m = 3$ in solving QF4

| | SQN(0) | | | SQN(D1) | | |
|---|---|---|---|---|---|---|
| N | Ite | $\|g_k\|$ | Fva | Ite | $\|g_k\|$ | Fva |
| 10 | 230 | 9.8e-5 | -1.6652 | 62 | 7.9e-5 | -1.6652 |
| 20 | 238 | 9.8e-5 | -3.3305 | 63 | 5.4e-5 | -3.3305 |
| 40 | 246 | 9.9e-5 | -6.6609 | 63 | 7.7e-5 | -6.6609 |
| 80 | 254 | 9.9e-5 | -1.3322e+1 | 64 | 8.3e-5 | -1.3322e+1 |
| 100 | 257 | 9.8e-5 | -1.6652e+1 | 64 | 9.3e-5 | -1.6652e+1 |
| 200 | 265 | 9.8e-5 | -3.3305e+1 | 68 | 3.5e-5 | -3.3305e+1 |
| 500 | 276 | 9.7e-5 | -8.3262e+1 | 68 | 9.5e-5 | -8.3262e+1 |
| 1000 | 284 | 9.8e-5 | -1.6652e+1 | 73 | 8.2e-5 | -1.6652e+1 |
| 1500 | 289 | 9.7e-5 | -2.4979e+1 | 73 | 9.6e-5 | -2.4979e+1 |
| 2000 | 292 | 9.8e-5 | -3.3305e+2 | 74 | 8.8e-5 | -3.3305e+2 |

The number of iterations is the success index in a computational method. In this study, the number of iterations was compared between the standard SQN method and the proposed SQN method.

Tables 1-4 show the comparison results between the proposed preconditioned SQN methods and the standard SQN method for $m = 2$. Generally, the computational results show that the proposed methods performed better when compared to that of the standard SQN method. As shown in the Tables, the proposed methods require less number of iterations than the standard method. Although all the methods show the same values of function evaluation, the norms of the gradient for the proposed methods are less than the norms of the gradient of the standard method. Once again, this shows that the proposed SQN methods are promising alternatives as compared to the standard SQN method.

Tables 5-8 show the comparison results between the proposed preconditioned SQN methods and the standard SQN method for $m = 3$. Once again, the results reveal that the proposed methods clearly outperform the standard method. The number of iterations and the norms of the gradient are the best evidences to show that the proposed methods generally perform better than the standard SQN method.

## CONCLUSION

The preconditioner for SQN method that is based upon variational technique and weak secant relation is proposed in this paper. The numerical results obtained suggest that the preconditioned SQN method is a good alternative for large-scale unconstrained optimization. Moreover, the preconditioned SQN method is preferred for reasons including simple implementation and it requires only function and gradient values.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledged the financial support of Graduate Research Fellowship (GRF) from Universiti Putra Malaysia and the Ministry of Higher Education Malaysia.

## REFERENCES

Farid, M., Hassan, M. A., & Leong, W. J. (2011). Improved Hessian approximation with modified secant equations for symmetric rank-one method. *J. Comput. Appl. Math*., *235*, 2423-2431.

Farid, M., Leong, W. J., & Hassan, M. A. (2010a). A new two-step gradient-type method for large scale unconstrained optimization. *Computers and Mathematics with Applications, 59*, 3301-3307.

Farid, M., Leong, W. J., & Hassan, M. A. (2010b). An improved multi-step gradient-type method for large scale optimization. *Computers and Mathematics with Applications, 61*, 3312-3318.

Leong, W. J., & Hassan, M. A. (2009). A restarting approach for the symmetric rank one update for unconstrained optimization. *Computational Optimization and applications, 43*, 327-334.

Leong, W. J., & Hassan, M. A. (2011). A new gradient method via Least change secant update. *International Journal of Computer Mathematics, 88*, 816-828.

Leong, W. J., Hassan, M. A., & Farid, M. (2010a). A monotone Gradient method via weak secant equations for unconstrained optimization. *Taiwanese J. Math*., *14*(2), 413-423.

Leong, W. J., Farid, M., & Hassan, M. A. (2010b). Improved Hessian approximation with modified quasi-Cauchy relation for gradient-type method. *Advanced Modeling and Optimization, 12*, 37-44.

Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS Method for large scale optimization. *Mathematical Programming, 45*, 503-528.

Ni, Q., & Yuan, Y. (1997). A subspace limited memory quasi-Newton algorithm for large scale nonlinear bound constrained optimization *Mathematics of Computation, 66*, 1509-1520.

Stoer, J., & Yuan, Y. (1995). A subspace study on conjugate gradient algorithms. *ZAMMZ. Angew. Math. Mech*., 75, 69-77.

Wang, Z. H. Wen, Z. W., & Yuan, Y. (2004). A subspace trust region method for large scale unconstrained ooptimization. *Numerical Linear Algebra and Optimization,* 265-274.

Waziri, M. Y., Leong, W. J., Hassan, M. A., & Monsi, M. (2010a). A new Newton's method with diagonal Jacobian approximation for systems of nonlinear equations. *J. Mathematics and Statistics, 6*, 246-252.

Waziri, M. Y., Leong, W. J., Hassan, M. A., & Monsi, M. (2010b). Jacobian computation-free Newton method for systems of non-linear equations. *Journal of numerical Mathematics and Stochastic, 2*(1), 54-63.

Yuan, Y. (2007). Subspace techniques for nonlinear optimization. *Some Topics in Industrial and Applied Mathematics (Series in Contemporary Applied Mathematics CAM), 8*, 206-218.

Pertanika J. Sci. & Technol. 22 (1): 193 - 211 (2014)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# GIS Routing and Modelling of Residential Waste Collection for Operational Management and Cost Optimization

**Billa, L.[1, 2], Pradhan, B.[2]* and Yakuup, A.[3]**

*[1]University of Nottingham Malaysia Campus, School of Geography, 43500 Semenyih, Selangor, Malaysia*
*[2]Geospatial Information Science Research Center (GISRC), Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
*[3]Unit for GIS in Spatial Planning (UGiSP), Bureau of Innovation and Consultancy, Universiti Technologi Malaysia, 81310 Skudai, Johor, Malaysia*

## ABSTRACT

In this paper, optimum routing was developed based on the travel salesman method and integrated in ArcInfo GIS using linear programming. The results of the optimized travel distances and times for residential waste collection and routing to disposal site were used to calculate the number and type of required track collection, labour requirement, costing of waste collection and to determine the overall solid waste management efficiency through waste management operation research methods. The objective of the study was to optimize residential collection and hauling to disposal site through operation cost minimization for Petaling Jaya Municipality in the state of Selangor, Malaysia. The study determined that with optimized routes and recycling possibilities, the total cost of waste collections could be reduced from RM90,372 to RM20,967, with a reduction of 76.8%. It was also revealed that optimum routes might not necessarily be the shortest distance from point A to point B as travel time maybe high on short distances due to traffic congestion and the presence of many traffic lights. Techniques and methods developed using general GIS have proven effective in route optimization and allowed management of data to suit local conditions and limitations of waste management for the studied area. Thus, scenarios of travel distances, time and waste quantity value generated from the GIS enabled appropriate determination of the number of waste trucks and labour requirements for the operation and the overall calculation of costs of waste management based on the operation research methods used in the study.

*Keywords:* Route optimization, GIS, Waste collection, Operational cost, Malaysia

*Article history:*
Received: 22 November 2011
Accepted: 27 January 2012

*E-mail addresses:*
Lawal.Billa@nottingham.edu.my (Billa, L.),
biswajeet24@gmail.com (Pradhan, B.)
*Corresponding Author

ISSN: 0128-7680 

## INTRODUCTION

Studies on waste management increasingly show that solid waste collection and transportation absorb almost 80% of municipal waste management budgets (Ogwueleka,

2009; Vijay *et al*., 2008; Meulen, 1996). This problem is experienced in Malaysia, where waste collection routes are left at the discretion of drivers, consequently impacting on costs, effectiveness and efficiency of waste management in municipal council areas such as Petaling Jaya. These financial and socio-environmental considerations such as smell and public nuisance in waste transportation have prompted waste collection planners to consider geographical information system (GIS) solutions for route selection and optimization in the collection operation. The objective of this study is to apply the travelling salesman's method, together with other heuristic routing techniques and GIS, in optimizing routing trucks and collection of residential waste so as to minimize the travel distance, time and financial costs during the collection process. Many of the waste management agencies involved in municipal waste in Malaysia incur high costs in the collection and transportation of wastes to disposal sites. Studies of their operations revealed the limited application of routing optimization and spatial information techniques.

Waste collecting routing in Malaysia is left at the discretion of drivers, resulting in abnormal time and distances travelled and thus higher operation costs. The process of residential waste collection involves door-to-door travels by collection vehicles. Studies by several researchers (see Lopez *et al*., 2008; Kim *et al*., 2006; Triantafyllou & Cherrett, 2009; Massie, 1997; Tchobanoglous *et al*., 1993; Hagget & Chorley, 1969; Mills, 1967) show that routes taken based on drivers' experiences do not necessarily mean optimum routes are followed. Although waste management software is available, it is rather expensive, while its functions and features cannot be easily customized to suite local conditions where they are being used. These problems have some limitations in their usage. In this study, data were developed in ArcView and Arc/Info GIS. Techniques and processes that had been developed for optimum routing were applied to the GIS for street network processing and analysis. Data were also assessed for separation and estimation of recyclable materials, and the results obtained were used to calculate the overall financial cost for the waste collection operation.

Network optimization techniques are used in many transport-planning cases; a typical route network optimization can be performed for waste collection, post delivery, street cleaning, snow removal, school bus routing and delivery of goods to customers (Chalkias & Lasaridi, 2009; Li *et al.*, 2009; Tavares *et al.,* 2008; Triantafyllou & Cherrett, 2009; Ghose *et al.,* 2008; Vijay, 2008). Although according to Wallace (2011), route optimization could eliminate hours spent on manual routing, maximize productivity, optimize equipment and staff allocations for better control of routing operations, a review of the routing algorithms showed that different problems required different routing models. A summary of the general routing problems and path finding algorithms revealed that the Floyd's and Short Path (SP) algorithms are suitable for identifying in a simple network from point A to point B. The Chinese Postman (CP) algorithm, on the other hand, identifies the shortest path in a network to link a series of points by passing through all the points at least once. The Travelling Salesman (TS) algorithm is similar to CP, but it has added features that allow travels through all given points and arcs in a network and arrivals at the starting point, and thus forming a circuit. These short path and routing problems have been extensively discussed in Travares *et al.* (2008), Karadimis, *et al*. (2008), Karadimis, *et al*. (2007), Belenguer *et al.* (2006), Johnson *et al.* (2002), Clarke and Wright (1964) and Mandl (1979).

Over the years, planners have established that the TS algorithm is useful for optimizing travels for street cleaning, waste collections, snow clearance and newspaper/ postal delivery that require an arc-routing solution with side constraints (Shih *et al.* 1997). In a GIS, street/ road networks are represented by arcs, intersections by nodes and directions can be assigned as one- or two-way lanes. Findings of the optimum TS routes require travels through all arcs of a given network such that the route is a circuit, where the initial and final nodes of the route are the same. A route circuit that passes through all arcs at least once guarantees that service can be carried out in the minimum time so as to minimize both travel and cost. Travares *et al.* (2008) and Belenguer *et al.* (2006) explained that municipal waste collection activities can be handled in the same way as street cleaning and the problem to be considered is how to route waste trucks to minimize the total travel time of the vehicles and determine the minimum number of trucks needed each day so that certain nodes are serviced on at least one route. The challenge of the study was to resolve the TS problem in a situation of unlimited route length, in such a way the optimum required tracks, travel distance and time can be estimated in waste collection from normal households.

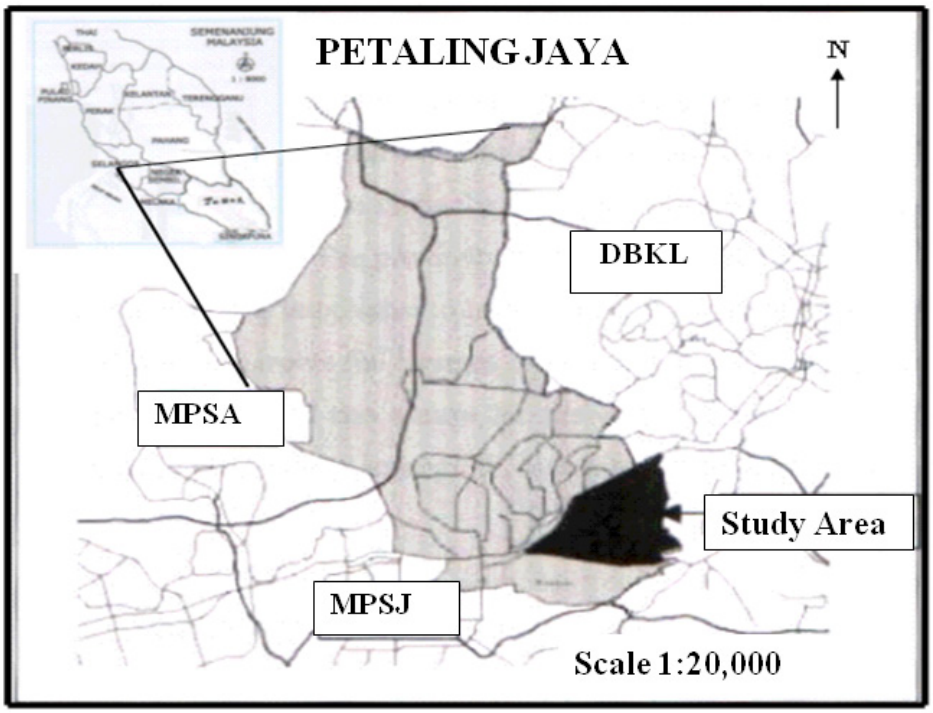
## MATERIALS AND METHODS

The waste collection route optimization study was conducted in zones eight and nine of the municipal area of Petaling Jaya, a residential suburb of Greater Kuala Lumpur in Malaysia (Fig.1). In this paper, the findings of the collection route S18 zone eight are presented and discussed. Without any effective monitoring system, there have been dissatisfactions with untimely waste collections, and the collection process in the area has also shown abnormally high travelling time for some routing areas. These abnormalities were found to be attributed to unofficial recycling activities conducted by waste collection personnel. Generally, the area has an extensive residential road network, and the limited waste management resources available make it difficult to institute a timely waste collection process and also monitor the time travel of waste collection trucks. This study was thus conducted to optimize the use of garbage trucks and improve the routing of these trucks.

There are five collection sectors/ routes for the zone with a total population of about 16,365 people. Incomplete data for the study area were collected from the municipal council (MPPJ, 1998; 1997) and the waste management authority (AFSB, 2000; 1996). Firstly, a primary survey was conducted to update information on the number and type of houses on each residential street. Housing type was categorized as semi detached, flats or others, while the income type of housing was classified as high, medium and low, based on the national classification. The extent of each residential collection sector was also confirmed using the sector map given by AFSB. Other information regarding the average waste storage container size, waste collection vehicle capacity, truck type (compaction and non-compaction), the average speed of truck during collections in residential areas and the speed of truck during travels to the landfill location was supplied by the management company (AFSB).

A thematic map was generated for the study area to show the type and quantity of wastes by collection sector. The average waste generation rate for zone 8 was calculated as 0.55 kg/ capita with the population density of 62/ha. The total daily waste production was 8,875.2 kg

and after effective recovery, the expected waste for daily disposal in the zone was 4,118.14 kg. On average, waste emanating from the medium income households accounted for over 85% of the total waste collection. Details of the waste generation rate by sectors in zone 9 are presented in Table 1. Breakdown of the number of houses by income level, the total waste collected and the type of wastes are shown in Table 2.




MPSA: Majlis Perbandaran Shah Alam
MPSJ: Majlis Perbandaran Subang Jaya
DBKL: Dewan Bandaraya Board Kuala Lumpur

Fig.1: Study area (Petaling Jaya)

TABLE 1
Waste arising for zone 8

| Col. sector | No. of Houses | Gen. Rate | Qty/day (kg) | Qty 3/week Col. (kg) | Qty after Rec. (kg) |
|---|---|---|---|---|---|
| S 2 | 556 | 0.88 | 2446.40 | 4892.80 | 1135.24 |
| S 13 | 606 | 0.56 | 1677.20 | 3354.40 | 778.10 |
| S 1 | 858 | 0.47 | 2004.55 | 4009.10 | 929.98 |
| S 5 | 838 | 0.43 | 1863.50 | 3727.00 | 864.84 |
| S 7 | 415 | 0.41 | 883.55 | 1767.10 | 409.98 |
| Total | 3273 | 0.55 | 8875.20 | 17750.40 | 4118.14 |

Col: collection; Gen: generation; Qty: quantity; kg: kilogram; Rec: recycling

TABLE 2
Waste arising in collection sector S -2

| Housing Type | NH | Qty/d (kg) | Rc. (kg) | Oc. (kg) | W. LF (kg) |
|---|---|---|---|---|---|
| High income | 0 | 0 | 0 | 0 | 0 |
| Medium | 342 | 1504.80 | 380.71 | 774.89 | 349.18 |
| Med/flats | 204 | 897.60 | 227.08 | 462.25 | 208.23 |
| Low | 10 | 44 | 11.13 | 20.66 | 10.21 |
| Low/flats | 0 | 0 | 0 | 0 | 0 |

NH: Number of houses, Qty/d: Quantity daily, Rc: Reusable content,
Oc: Organic content, W.LF: Waste for landfill, kg: kilogram

*The Existing Collection Routes*

The travel time of waste collections generally varies for each route/ sector. The influencing factors were identified as the distance of the operation depot to each routing area, the number of premises, the type of waste storage containers, and the mileage of roads to be covered within the collection area and the distance from the collection areas to the landfill site. Although travel time for the residential collection routing sector was not available, the management highly observed the collection times for the routes, particularly in S-2. Table 3 presents the travel time from the depot to collection area, the collection area to landfill, and the landfill to the depot which is located 17.5 km away.

The furthest routing sector from the landfill is about 26.2 km. At an average speed of 50 km/hr and a minimum speed of 40 km/hr, the travel time to the landfill should be between 39 and 31 min, respectively. The factors influencing collection and disposal time depend on the route size, the capacity of the collection vehicle and the travel time. Meanwhile, the routing areas vary in size and location, which are relative to the operation depot and landfill site. Open dump trucks have less waste capacity compared to compaction vehicles, which are considered to be inefficient in waste collection process. The travel time is influenced by the average speed of the vehicle, distance travelled, and accessibility.

TABLE 3
Distance and travel time for routes in km

| Adm Cd | Rte Cd | Total-time Rte-Lf-Dpt (hr) | Rte-Lf (hr) | Lf-Dpt (min) | Dist.Rte-Dpt (km) | Dist.Lf-Dpt (km) |
|---|---|---|---|---|---|---|
| **Zone** | | | | | | |
| PJ 8 - 1 | S -2 | 2:09 | 1:48 | 0:21 | 23.1 | 17.5 |
| PJ 8 - 2 | S -5 | 2:01 | 1:40 | 0:21 | 23.8 | 17.5 |
| PJ 8 - 3 | S -7 | 1:57 | 1:36 | 0:21 | 22.5 | 17.5 |
| PJ 8 - 4 | S -13 | 1:17 | 0:56 | 0:21 | 21.4 | 17.5 |
| PJ 8 - 5 | S -1 | 1:38 | 1:17 | 0:21 | 23 | 17.5 |

Adm: Administration, Cd: Code, Rte: Route, Lf: Landfill, Dpt: Depot, hr:hours, min: minute, km: kilometer
Source: MPPJ (1996) and AFSB (1997)

*Development of Waste Collection Data*

Various data were collected for the development of GIS, and these included land use (location of disposal sites), details of the roadway network of the study area, details of the residential layout and boundaries of the residential collection sectors, as demarcated by the waste management authority. Data were digitized into digital form; a total of five layers were developed in ArcView for database development. The network layer was then converted into Arc/Info coverage for cleaning and editing and prepared for network processing and analysis. Since the processing of extensive network data took up more time and computer memory, each collection sector area was dealt separately to facilitate proper routing and coordination of the road system and thus ensured an easy and quick data processing. Where an area was not clearly defined, the residential streets were demarcated as the collection area. Data were managed in such a way that enough computer space would be available to facilitate fair processing speed and analysis. Some network attributes were presented through an organized coding system. Dynamic segmentation was applied for road classification to provide an alternative to using pseudo nodes and record other attributes. The features of the route system and the event handling commands within the Arc/Info provide this dynamic segmentation capability. Then, the entire optimized routes are linked to the GIS network file using a linear programming in Arc/Info. Analysis of the database for the waste collection optimization process was carried out in Arc/Info where separate node attribute table (NAT) and an arc attribute table (AAT) were created. Using dynamic segmentation function points, line attributes were entered in one arc; the route function was then used to explore and generate various routing scenarios.

*Road Network Data Development*

The road network was the most important data needed for the study. After a detailed road network had been developed and topologized, a turntable was developed. The turntable is an ARC level command that builds or updates a coverage turntable for every possible "arc" – to - "arc" turns in the coverage by calculating the azimuth and angle for each arc to turn. A weed tolerance level is set, and this option is the distance in map units used to weed the vertices from the calculated turn angles. The road network attribute has 36 fields of attributes ranging from road name, class and traffic capacity. Others are the distances and travels where the field items are distance in kilometres and miles. Travel time is calculated at 30, 40, and 50 km/hr. Travel time which includes pick-up time for garbage is calculated at 5 to 7 km/hr plus pick-up time of 30 seconds for house to house containers, and 90 to 120 sec for 1100 kg containers located at flats, apartments and community dumps. The other category is demographic and waste quantity which comprises of the number of residences, houses, the income type of house and the quantity of waste for one to three times of collection frequency.

The waste collection truck is to follow a certain route, which should have a minimum length. This optimum route is calculated together with all the related information stored in the road segment such as waste offered, pick-up points, each pick-up time and expected delay time at traffic congestion. Using the existing routing sectors, the travelling salesman's problem was solved for each routing sector. The route enters the collection sector and follows an optimized

path while passing through all streets at least once and then back to the entry point. This ensures that the services are provided along all the streets based on the minimum distance and time, thus resulting in cost minimization.

It is crucial to note that the routes from residential or collection areas to disposal site (landfill) were not based on the travelling salesman's method. Rather, the detailed major road network was encoded with the attributes of road intersections, length, classification and traffic congestion. The choice of route was mainly on the primary roads, thus limiting the options of routing and simplifying the routing process. The route begins from the exit of the collection areas and the best possible arc link is selected on a primary road to the direction of the waste disposal site. The "near"-function was used to link the costs, directions and speed to the arcs and nodes. Together with the "near"-function, the modules calculate the shortest path between two or more points and optimize the sequence of stops making the route. The major objective is to minimize travel time and distance. Once there is a reduction in time and distance, there will be a corresponding reduction in cost of waste management. Roads are classified according to the national road classification scheme, and these include highways, primary, district, local distributors and minor roads (AFSB, 2000; Arshad, 1992). Using these attributes, it is easy to determine the road characteristics in terms of its width, size and capacity. Highways are identified separately because waste transportation on such roads makes distances shorter but imposes higher costs.

In this study, the road network of the collection sectors comprises 881 arc segments. However, when the roads from the waste management depot and the roads disposal site location are included, the total arc segments considered in the study were 3843. In the application of network analysis in Arc/Info, all the attributes of the network are calculated and included in individual arc segments. Primary, district and local distributor roads provide high vehicle capacity and allow rapid travel, and are therefore classified as mass transit roads. The travel times on these roads are low because a waste collection truck may travel at a speed between 30 to 50 km/h from the collection sectors to waste disposal site. Meanwhile, road intersections and junctions with incidence of traffic at certain time periods are identified and excluded from the mass transit. A thematic map was generated in GIS to show the road network from the waste management operation depot to the residential collection areas and to the landfill. In GIS, the roads for collection and hauling of wastes from collection sites to disposal landfill are presented separately to allow for an easy processing of data.

Waste production and distribution were analyzed for the housing file and optimum collection routes on the network file. Residential houses along the route were classified according to the national classification scheme, as follows: low, middle and high income housing type; the national average household size of five was then used to compute the population and the expected quantity of waste for collection for each house using the per capital waste generation figures collected in earlier studies for Petaling Jaya. The waste generation was calculated for different income housing levels such as low income (0.65 kg/c/d), medium income (1.28 kg/c/d) and high income (1.37 kg/c/d).

*Optimum Route Based on the Travelling Salesman (TS) Method*

The TS algorithm was used for the residential waste collection route optimization. The process of route optimization involves establishing the sequence of streets/roads to be followed in the entire road network of the selected collection area. Meanwhile, a TS cost matrix was worked out using the nodes and the distance of the roads between each two intersections. The problem considered the routing of the service trucks so as to minimize the total distance, travel time and also to determine the number of trucks and personnel needed so that the cost of waste collection in the selected area could be minimized.

According to Karadimis *et al*. (2008), solving the TS problem can be mathematically explained as a problem of combinatorial optimization. The variants of the combinatorial optimization appeared in the recent studies of Gutfraind and Hagberg (2011) and Wallace (2011). The problem is stated as follows: let $X_{ij}$ be a variable, so that $X_{ij} = 1$, if node i is assigned to j and $X_{ij} = 0$, and the coefficients $C_{ij}$ is the cost of assigning node i to j (Mandl, 1979). The governing equations presented are to find 0 - 1 variable $X_{ij}$ so as to minimize:

$$Min : z = \sum_{j=1}^{n} \sum_{i=1}^{n} C_{ij} X_{ij} \qquad \text{Eq.(1)}$$

subject to

$$\sum_{i=1}^{n} x_{ij} = \sum_{j=1}^{n} x_{ij} = 1 \text{ (for all } i \text{ and } j = 1, 2, \ldots n) \qquad \text{Eq.(2)}$$

and

$$x_{ij} = 0 \text{ or } 1 \qquad \text{Eq.(3)}$$

In solving the problem, a similar scheme was used, where $X_{ij} = 1$, indicating that a truck travels directly from node *i* to *j*, and $X_{ij} = 0$ indicating that it does not travel at all. By assuming that an assignment cost is $C_{ii} = 8$. Cij is the shortest distance from node i to j. Equation 2 in the (TS) problem involves a truck entering a node by an arc, which is also leaving this particular node. The minimum value of *z*, calculated from equations (l) - (3), is the valid lower bound for the solution to the TS problem for a network with a cost matrix $(c_{ij}) = C$. The construction of a cost matrix among the nodes requires the extraction of distance values from the network. A series of matrix iteration was constructed in Microsoft Excel based on the method of Chou (1997). The route was computed by assigning the numbers to all the intersecting nodes of the sector street network. The lengths of each arc/street between the nodes were identified and used to develop a series of cost matrixes to determine the order and sequence of node to be included in the route. The cost values for the optimum TS route were extracted from the matrix and linked to the street network coverage in the Arc/Info by linear programming.

Various waste management operation research methods were used to calculate the waste hauling time and motion, as explained by Chalkias and Lasaridi (2009), Vijay *et al.* (2008), Oliveira and Borenstein (2007), and Tchobanoglous *et al*. (1977). Collection time is an important factor that can be used to quantify collection efficiency. Optimum waste collection

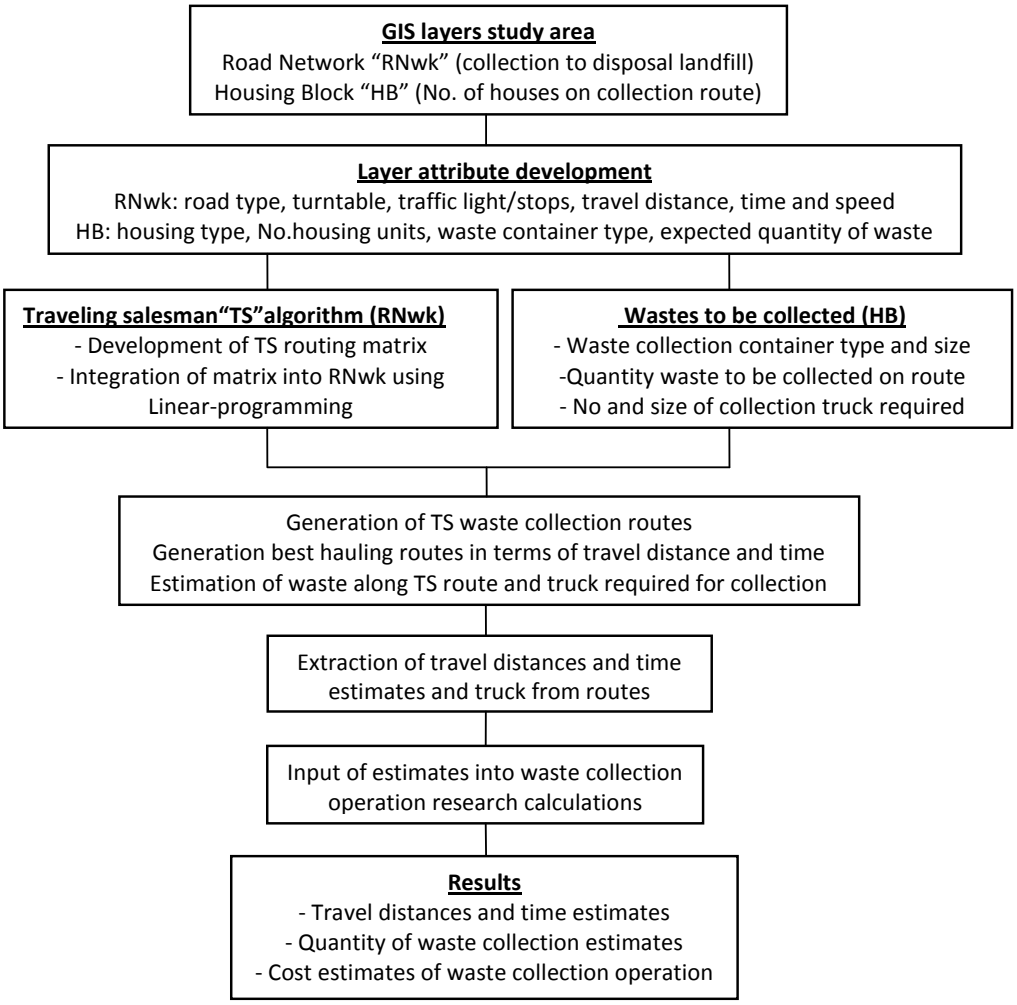



Fig.2: A flow chart of the methodological process adopted in this study

includes the concerns of the optimum filling of collection vehicles at each end of a route and the route to be followed should have travelling characteristics in terms of time and distance. Using time and motion values extracted from the system, the capital expenditure for solid waste management was computed based on Ghose *et al.* (2008) and Holmes (1983).

In this section, a detailed description of the GIS data and detailed methodological processes are given. In more specific, the section explains waste estimation based on housing type and number of units. It also gives elaborated discussions of the travelling salesman method and how it has been integrated into the GIS road network to generate an optimized TS collection route. Fig.2 shows a flow chart of the methodology process adopted in this study.

## RESULTS AND DISCUSSION

The TS optimized route means the points of entry and exit at the collection sector are the same. A route is followed from then on without backtracking so as to minimize the distance and time in the collection process. Fig.3 shows the optimum distances and travel time generated using the query function of the Arc/Info GIS. The collection time includes the time spent on picking container and the quantity of garbage to be collected along the route. The pick-up time increases by 20 seconds for every terrace house and 90 seconds for flats and apartment containers. A total of 10 flats are allocated for one container in the flats and apartment blocks.

Although sector S-2 is a large area, the routing process has the same number of housing units with that of other sectors. Hence, the optimum route was computed based on 41 intersection nodes. The travel distance and time from the optimum collection route, the number of housing units and the quantity of waste for collection are presented in Fig.4(a) and Fig.4(b), Fig.5 and Fig.6, respectively. Since the tables are generated by the GIS system as a scroll down, the values presented (in the tables) are the samples values and they do not sum up to the total value as computed by the system. At an average speed of 7 km/hr for residential waste collection, the total distance covered through the course of the collection process is 7.55 km (see Fig.4a). Travel time includes time spent on each residential pickup and the pick-up time for communal containers along the route. The collection time for S-2 is recorded as 2.24 hrs (see Fig.4b). The time recorded in S-2 was less because it has fewer number of houses, and only a few flats and apartments, where the container pick-up time of 90 seconds would have increased the collection time. The total number of residential housing units in the collection route is 524 units (see Fig.5). Fig.6 shows the expected quantity of wastes to be collected as 2199 kg.

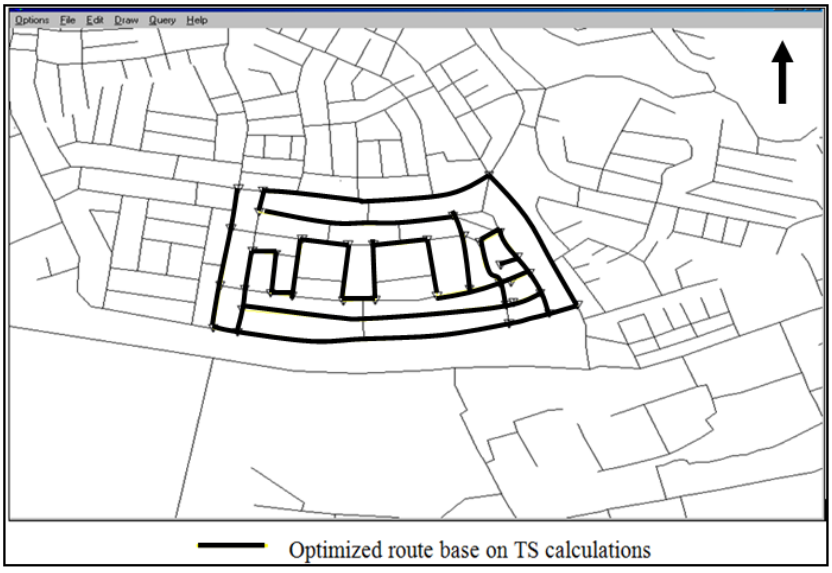



Fig.3: Optimum collection route in S-2 (zone 8)

(a)

Directions

| | | | |
|---|---|---|---|
| Travel | 0.06 | Kilometers | Go straight to JLN 1/10 |
| Travel | 0.07 | Kilometers | Turn left to JLN 1/15 |
| Travel | 0.19 | Kilometers | Turn right to JLN 1/6 |
| Travel | 0.06 | Kilometers | Turn right to JLN 1/17 |
| Travel | 0.13 | Kilometers | Turn left to JLN 1/26 |
| Travel | 0.07 | Kilometers | Turn left to JLN 1/19 |
| Travel | 0.27 | Kilometers | Turn right to JLN RAILWAY 1/2 |
| Travel | 0.07 | Kilometers | Turn right to JALAN PASAR 21 |
| Travel | 0.48 | Kilometers | |
| Total= | 7.55 | Kilometers | |

(b)

Directions

| | | | |
|---|---|---|---|
| Travel | 3.51 | Minutes | Go straight to JLN 1/10 |
| Travel | 0.00 | Minutes | Turn left to JLN 1/15 |
| Travel | 6.38 | Minutes | Turn right to JLN 1/6 |
| Travel | 0.00 | Minutes | Turn right to JLN 1/17 |
| Travel | 3.17 | Minutes | Turn left to JLN 1/26 |
| Travel | 3.53 | Minutes | Turn left to JLN 1/19 |
| Travel | 9.12 | Minutes | Turn right to JLN RAILWAY 1/2 |
| Travel | 0.00 | Minutes | Turn right to JALAN PASAR 21 |
| Travel | 0.00 | Minutes | |
| Total= | 144.82 | Minutes | |

Fig.4: Direction table of distance (a) and time travel (b) at 7km/h

Directions

| | | | |
|---|---|---|---|
| Travel | 10.00 | Units | Go straight to JLN 1/10 |
| Travel | 0.00 | Units | Turn left to JLN 1/15 |
| Travel | 18.00 | Units | Turn right to JLN 1/6 |
| Travel | 0.00 | Units | Turn right to JLN 1/17 |
| Travel | 20.00 | Units | Turn left to JLN 1/26 |
| Travel | 10.00 | Units | Turn left to JLN 1/19 |
| Travel | 35.00 | Units | Turn right to JLN RAILWAY 1/2 |
| Travel | 0.00 | Units | Turn right to JALAN PASAR 21 |
| Travel | 0.00 | Units | |
| Total= | 524.00 | Units | |

Fig.5: The sample results of the number of housing units

Directions

| | | | |
|---|---|---|---|
| Travel | 44.00 | Units | Go straight to JLN 1/10 |
| Travel | 0.00 | Units | Turn left to JLN 1/15 |
| Travel | 79.00 | Units | Turn right to JLN 1/6 |
| Travel | 0.00 | Units | Turn right to JLN 1/17 |
| Travel | 88.00 | Units | Turn left to JLN 1/26 |
| Travel | 44.00 | Units | Turn left to JLN 1/19 |
| Travel | 154.00 | Units | Turn right to JLN RAILWAY 1/2 |
| Travel | 0.00 | Units | Turn right to JALAN PASAR 21 |
| Travel | 0.00 | Units | |
| Total= | 2199.00 | Units | |

Fig.6: The sample results of the quantity of expected wastes to be collected along optimized route of S-2

*Hauling Route from Collection Sector to Disposal Site*

The optimum residential waste collection using the travelling salesman algorithm is complicated because it calculates routes from the start node and back to the start node, passing through all streets to form a circuit. However, in the hauling of waste from collection areas to the disposal site, only primary roads and highways are included, making the optimum route development process easy. The options of avoiding the highways due to tolls and also selecting routes that avoid possible traffic congestion are also included in the GIS system.

In the hauling route process, the nearest node from the origin node is located and the process continues to finding the subsequent nearest node in the network until it reaches the destination node (i.e., the nearest to the location of the disposal site). This tree building process ensures that the shortest possible path is taken in terms of distance and time. When the destination node is reached, the algorithm is terminated because the shortest path has been found. The various route scenarios for optimum hauling from sector S-2 collection to the landfill in the GIS application are illustrated in Fig.7. The system is manipulated to generate two optimum route scenarios A and B (details of this hauling route are shown in Fig.8).

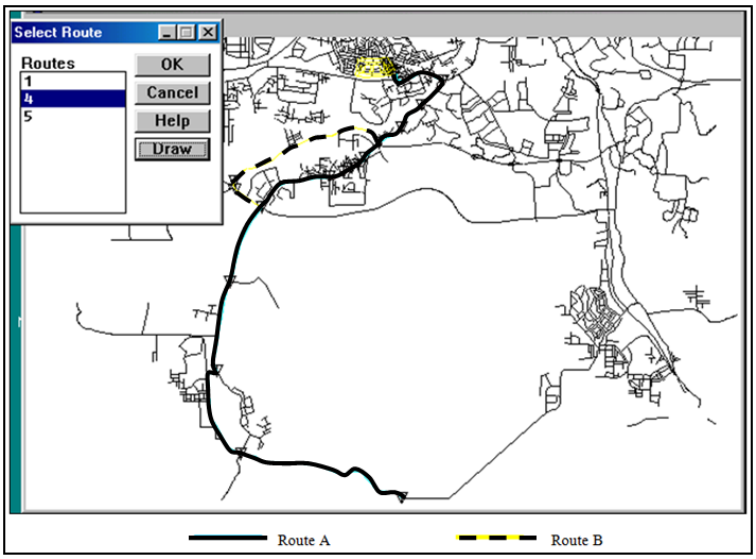



Fig.7: Hauling route A to disposal site from S-2

The travel time and motion values generated from the waste collection GIS show sample directions of two possible hauling route options to the landfill site. The unnamed features in the direction tables are the segments of road names that are not identified when conducting the aerial survey. For this reason, they could not be included in the system. Route A presents a direct route at a travel speed of 50 km/hr through the primary road. The distance to the landfill is 20.42 km, and the recorded travel time of 27.2 min is high due to the road segment being in a commercial and residential area and also the possibilities of traffic congestion along the major road. Route B (see Fig.8) shows a travel distance of 21.30 km in (Fig.9b), with an increase of

1 km on route A. Due to less traffic, however, the travel time is generated as 25.29 min, i.e. with a decrease of 2 min. The two routes were assessed in terms of time and distance travel and by taking factors such as public nuisance from smell into consideration. Route B presents a better and agreeable travel option as it makes up for the increased distance with a faster travel time that avoids possible public nuisance; thus, the best waste route is not necessarily the shortest distance.

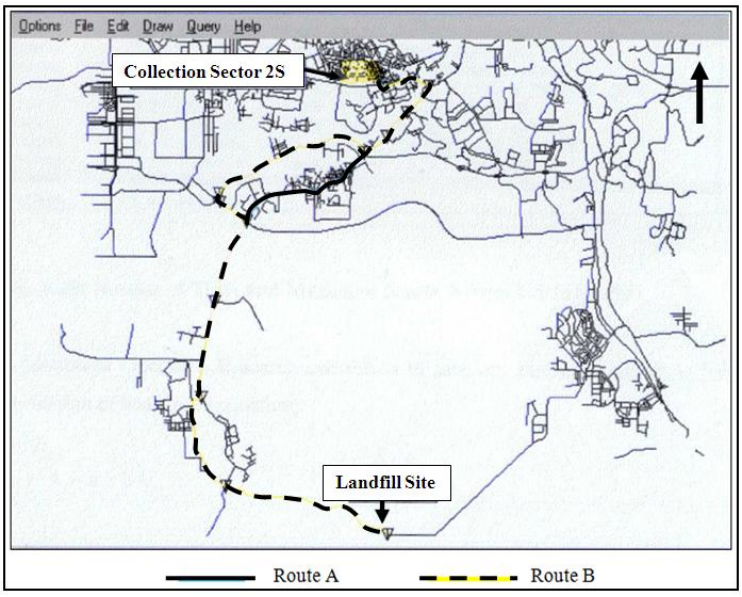



Fig.8: Hauling route B to disposal site from S–2

In this study, the time and distance values that form the optimums were used to evaluate and assess the trip hauling time. Using traditional operation research calculation of time and motion, it is shown as the basic expression of haul speed equation, as follows:

$$x/y = h = a + bx \qquad \text{Eq. (4)}$$

Thus, to find the single-trip-haul time for hauling Route A which is located 20.4 km to the disposal site, the values are substituted as:

hr = (0.080 hr / trip) + [(0.012 hr / km)(20.4 km/ trip)]
hr = 0.080 hr /trip + 0.245 h / trip
Haul time = 0.32 h / trip

Route B
hr = (0.080 hr / trip) + [ (0.012 hr / km )( 23.2 km / trip)]
Haul time = 0.36 hr / trip

(a)

Directions

| | | |
|---|---|---|
| Travel | 1.54 Minutes | Turn right to JLN PUCHONG |
| Travel | 2.92 Minutes | Turn right to S. ALAM/ KL H.W |
| Travel | 5.24 Minutes | Go straight to unnamed feature |
| Travel | 0.30 Minutes | Go straight to JLN BANDAR PUCHONG JAYA |
| Travel | 0.96 Minutes | Turn right to JLN PUCHONG |
| Travel | 5.52 Minutes | Turn right to unnamed feature |
| Travel | 0.42 Minutes | Go straight to PESIARAN PUCHONG |
| Travel | 2.80 Minutes | Go straight to JLN PUCHONG |
| Travel | 5.59 Minutes | |
| Total= | 25.29 Minutes | |

(b)

Directions

| | | |
|---|---|---|
| Travel | 1.28 Kilometers | Turn right to JLN PUCHONG |
| Travel | 2.43 Kilometers | Turn right to S. ALAM/ KL H.W |
| Travel | 4.36 Kilometers | Go straight to unnamed feature |
| Travel | 0.25 Kilometers | Go straight to JLN BANDAR PUCHONG JAYA |
| Travel | 0.80 Kilometers | Turn right to JLN PUCHONG |
| Travel | 4.60 Kilometers | Turn right to unnamed feature |
| Travel | 0.35 Kilometers | Go straight to PESIARAN PUCHONG |
| Travel | 2.34 Kilometers | Go straight to JLN PUCHONG |
| Travel | 4.66 Kilometers | |
| Total= | 21.30 Kilometers | |

Fig.9: The sample results for the collection time (a) and distance (b) of route B from S-2 to landfill

The greater part of solid waste management expenses is incurred in the waste collection and disposal processes. The possibility of reducing the costs of waste collection involves controlling the factors that influence the cost. The factors specifically include the distance travel and time travel through collection and to disposal and the quantity of waste to be collected for disposal. The development of the optimum routing scenario was difficult but with the use of GIS, the entire procedures have been simplified, and it is beneficial to determine the best route (note that the best route does not necessary be the shortest distance in terms of time). Additionally, using the value from the optimum routing, the hauling time of waste collection trip was calculated.

*The Costs of Waste Collection Management*

A total of RM5,000 per day is spent for the disposal of the entire waste collected in Petaling Jaya, while RM150 is spent on each collection vehicle per day. At three times a week, the 36.149 tons of wastes in the study area are collected at RM100/ton and RM25/ton for disposal. Thus, the total costs of waste collection and disposal was estimated at RM90,372. The potentially high recovery rate means that the quantity of waste could be reduced to 8.387 tons, and this would mean a reduction in the cost of collection and disposal to RM20,967. The evaluation of the capital expenditure for solid waste management was calculated using the mathematical formulas, as expressed by Tchobanglous *et al*. (1993; 1977). The basic expression for man-minute's labour cost per ton is formulated as the total man-minutes per trip:

$$f = b + ade + aT' + aT'' + aT'''. \quad \text{Eq. (5)}$$

The terms in Eq. 5 have been defined together with the calculation of the labour cost as follows: distance and time values were used to calculate the labour cost per ton. For a three-time per week collection frequency from curbside location, and with a three-man collection

truck crew in sector 2S, the time and motion information were substituted for the labour cost calculation, as follows:

a = 3 men per vehicle
b = 145 man-minutes per trip
de = 23.5-minute disposal site at a speed of 50 km/h
T' = 10-minute disposal time, assume time
T'' = 0 minute of route time
T''' = 7.3 travel time to first stop
c = 4.4 tons average refuse per trip
f = (145) t 3(23.5) + 3 (10) + 3(0) + 3(7.3) as the total man-minute per trip; therefore, x = f / c = 267.4 / 4.4 = 60.77man-min/ton.

Based on the average monthly income the waste workers receive i.e. RM500, the per minute wage was calculated at RM 0.04.

Labour cost per ton will then be: 61 m-m/ton x RM 0.04/m-m = 2.44/ ton.

The operational cost of the waste collection in the sectors was calculated by using the operation research in Eq. 6 to determine the number of containers per trip:

$$Ct = V\ r / cf \qquad \text{Eq. (6)}$$

The values for sector S-2 of the study area were calculated as follows:

V = 37. 5 $m^3$ / trip
C = 0.78 $m^3$
f = 0.85
r = assumed compaction ratio of 2

Thus, C t = (37.7$m^3$/ trip) ( 2 ) / (0.78$m^2$ ) ( .85) = 75 / 0.663
C t = 114 containers/ trip

The pick-up time per container was then determined by using Eq. 7.

$$Pscs = C\ t\ (u\ c) + (A\ p - 1)\ (d\ b\ c) \qquad \text{Eq. (7)}$$
$$= C\ t\ (u\ c) + (n\ p\ -1)\ (a' + b'\ x')$$

Pscs = (114 -1) (0.050 h /con) + (114 - 1 location /trip) [(0.60 h / location) + (0.041 km) + (0. 1 km / location)]
= (5.65) + (113) [(0.060 h/location + 0.0041)
= 5.65 + 113 (0.0641) = 12.89 h/trip

The number of trips required per week was then determined by using Eq. 8:

$$Nw = Vw / (vr) \qquad \text{Eq. (8)}$$

Nw = (52.4 m3/wk) / (37.5 m3/trip) (2) = 0.69 = 1 trip/ wk

The operational cost of waste collection in the study area was calculated at RM 300/week, and this cost is relatively high for the short distance to landfill. At three times a week, the 36.149 tons of waste in the area are collected and disposed of at the cost of RM90,372.00. It is also revealed that with effective waste recovery, the quantity and costs of collection and disposal can be reduced to 8.389 tons at a cost of RM20,967.00. Table 4 presents a comparative evaluation of the GIS generated hauling time to disposal site and the traditional method of calculation of time and distance. Moreover, by using GIS, the specific locations of expected traffic congestion can be identified and the delayed time input in the road segments (i.e., time travel) could be better assessed. In the traditional method, however, this factor could not be taken into account. The results of the optimum distances and time obtained from the routing were used to calculate the waste trip hauling time, pick-up time per container, number of trips required per week, basic cost of labour per ton and the overall operational cost of waste collection based on the waste collection operational research methods, as given in Equations 5, 6, 7 and 8.

TABLE 4
GIS generated and traditional time and motion measurement

| Route A | GIS generated | Traditional Calculation |
|---|---|---|
| Distance | 20. 4 km | 24.4 km |
| Travel time | 27.2 min/trip | 0.32 h = 19.2 min/trip |
| Route B | | |
| Distance | 23,2 km | 23.2 km |
| Travel Time | 25.2 min/trip | 0.36 h = 21. 6 min/trip |

## CONCLUSION

In this study, the optimum routes for residential waste collection were established using the travelling salesman route optimization techniques and network analysis in GIS. A single layer network operation was conducted to determine the optimum route for hauling waste to landfill. Although this route recorded a travel distance of 21.3 km, the travel time was low, i.e. at 25.29 min, because all the possible traffic obstructions were avoided. In the process of this routing application, many benefits of applying GIS were observed. The application did not only bring real-life waste collection routing problems for resolution on a desktop PC but it also simplified the time and means of managing the vast amount of data required. Both the spatial and network data were handled efficiently, while the calculation and estimation of wastes arising were also simplified using statistical and analytical functions in the GIS. This has allowed the evaluation of the expected quantity of waste for collection and the number of expected pick-up stops on each route in the network. Time and motion values within the optimized TS routing area were also assessed and evaluated using the system.

Although the application of GIS in Malaysia is very wide spread, its practical problem-solving application is still in its infancy. The processes and techniques of GIS and the network module in the Arc/Info for resolving the local and practical routing problem in urban residential waste collection are presented in this paper. The various network data structure and routing functions available in the Arc/Info are adequate for the necessary route planning analysis in

the study. The network module in the Arc/Info provided an opportunity to make calculations on road network and link generalized characteristics such as cost, direction and congestion to arcs and nodes of road segments. The extensive route network data of over 500 km of road length in the study was reduced to a PC desktop environment. Meanwhile, the network analysis functions in the Arc/Info were manipulated mainly on a single layer operation method to generate the optimum routing scenarios. Thus, the system allows effective and efficient visual interpretation and display of results.

## ACKNOWLEDGEMENTS

The authors acknowledge the help and contribution of Alam Flora Sdn. Bhd. (AFSB) and Majlis Prebandanran Petaling Jaya (MPPJ) for providing reports and data of their operations that enabled the implementation of this study. Thanks to the three anonymous reviewers for their critical review of the paper.

## REFERENCES

Alam Flora Sdn. Bhd. (AFSB). (2000). *GIS spatial data utility layers*. Document No AFSB /OPD/ GIS/ RPT/ 003.

Alam Flora Sdn. Bhd. (AFSB). (1997). *Waste Management Report*.

Arshad, Z. (1992). *Action plan for the national recycling program*. Paper presented at the National Program on Waste Recycling, A National Seminar on Municipal and Industrial Waste Management and Technology, 18 February 1992, Kuala Lumpur.

Belenguer, J. M., Benavent, E., Lacomme, P., & Prins, C. (2006). Lower and upper bounds for the mixed capacitated arc routing problem. *Computers and Operations Research, 33*, 3363-3383.

Chalkias, C., & Lasaridi, K. (2009). *Optimizing municipal solid waste collection using GIS*. Energy, Environment, Ecosystems, Development and Landscape architecture. ISBN: 978-960-474-125-0.

Chou, Y. H. (1997). *Exploring Spatial Analysis in Geographical Information Systems*. On Word Press 2530 Camino Entrda, Santa Fe NM 875054835 USA.

Clarke, G., & Wright, J. W. (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research, 12*, 568-581.

Ghose, M. K., Dikshit, A. K., & Sharma, S. K. (2008). A GIS based transportation model for solid waste disposal: a case study of Asansol Municipality. *Waste Management, 26*, 1287-1293.

Gutfraind, A., & Hagberg, A. (2011). *Interdiction of a Markovian Evader*. ICS 201112th INFORMS Computing Society Conference Monterey, California, January 9–11, 2011. Retrieved March 2, 2011 from http://math.lanl.gov/~hagberg/Papers/gutfraind-2010-interdiction.pdf.

Hagget, P., & Chorley, R. J. (1969). *Network analysis in geography*. Edward Arnold Ltd. 41 Moddox Street London W. l.

Holmes, J. R. (1983). *Practical waste management*. John Wiley and Sons.

Johnson, D. S., Gutin, G., McGeoch, L. A., Yeo, A., Zhang, W., & Zverovitch A. (2002). *The travelling salesman problem and its variations*. Kluwer Academic Publishers, 445-487.

Karadimas, N. V., Doukas, N., Kolokathi, M., & Defteraiou, G. (2008). Routing optimization heuristics algorithms for urban solid waste transportation management. *Journal WSEAS Transactions on Computers, 7*(12),

Karadimas, N. V., Papatzelou, K., & Loumos, V. G. (2007). Optimal solid waste collection routes identified by the ant colony system algorithm. *Waste Management and Research, 25*, 139-147.

Kim, B. I., Kim, S., & Sahoo, S. (2006). Waste collection vehicle routing problem with time windows. *Computers and Operations Research, 33*, 3624-3642.

Li, Y. P., Huang, G. H., Yang, Z. F., & Nie, S. L. (2009). 0-1 Piecewise linearization approach for interval parameter nonlinear programming: application to environmental management under uncertainty. *Can. J. Civ. Eng*., *36*, 1071-1084.

Lopez, A. J. V., Aguilar, L. M., Fernandez-Carrion, Q. S., & Jimenez del Valle, A. (2008). Optimizing the collection of used paper from small businesses through GIS techniques: The Legane´s case (Madrid, Spain). *Waste Management, 28*, 282–293.

Mandl, C. (1979). *Applied Network Optimization*. 111 Fifth Avenue. New York, New York, 10003: Academic Press Inc.

Massie, K. (1997). Solid waste planners at Portlands. In A. Mitchell (Ed.), *Regional Planning Agency, Metro use GIS to evaluate their recycling Programs. GIS at work in the community Zeroing*. ESRI, INC.

Meulen, G. (1996). *Improved solid waste collection in Asian cities*. Compuplan Foundation, Institute for Applied Spatial and Environmental Informatics. Mills & St. Hubert, the Netherlands.

Mills, G. (1967). A heuristic approach to some shortest route problems. *Journal of the Canadian Operation Research Society*, 20-25.

MPPJ. (1998). The Socio-Economic Study of Petaling Jaya. *A survey report*. No. pg. 27.

MPPJ. (1996). The Socio-Economic Study of Petaling Jaya. *A survey report*. No.pg. 34.

Oliveira, S. E., & Borenstein, D. A. (2007). Decision support system for the operational planning of solid waste collection. *Waste Management, 27*, 1286-1297.

Ogwueleka, T. C. (2009). Route optimization for solid waste collection: Onitsha (Nigeria) case study. *Journal of Applied Sciences and Environmental Management*, *13*(2), 37-40.

Shih, L. H., Liu, J. Y., & Lin, Y. T. (1997). *GIS and vehicle routing for infectious waste collection*. Proceeding of GIS AM/FM, ASIA 97 and GEOINFORMATICS. Mapping the Future of Asia Pacific. Taipei, Taiwan. 2, 505 –511.

Tavares, G., Zsigraiova, Z., Semiao, V., & Carvalho, M. (2008). A case study of fuel saving through optimization of MSW transportation routes. *Management of Environmental Quality, 19*, 444-454.

Tchobanglous, G., Theisen, H., & Eliassen, R. (1977). *Solid Wastes: Engineering Principles and Management Issues*. New York: McGraw–Hill Publishing Co.

Tchobanglous, G., Theisen, H., & Vigil, S. (1993). *Integrated Solid Waste Management, Engineering and Management Issues* New York: McGraw–Hill Inc.

Triantafyllou, M., & Cherrett, T. (2009). *The logistics of managing hazardous waste: a case study analysis in the retail sector*. 14th Annual Logistics Research Network Conference, 9th – 11th September 2009, Cardiff 613.

Vijay, R., Goutam, A., Kalamdhad, A., Gupta, A., & Davotta, S. (2008). GIS-based locational analysis of collection bin in municipal waste management system. *J Environ. Eng. Sci., 7*, 38-43.

Wallace, B. (2011). *Solid waste collection vehicle routing–route optimization – WIH Resource Group, Inc*. Retrieved April 11, 2011, from http://www.wihresourcegroup.wordpress.com.

Pertanika J. Sci. & Technol. 22 (1): 213 - 224 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Diagonal Preconditioned Conjugate Gradient Algorithm for Unconstrained Optimization

**Choong Boon Ng*, Wah June Leong and Mansor Monsi**

*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

The nonlinear conjugate gradient (CG) methods have widely been used in solving unconstrained optimization problems. They are well-suited for large-scale optimization problems due to their low memory requirements and least computational costs. In this paper, a new diagonal preconditioned conjugate gradient (PRECG) algorithm is designed, and this is motivated by the fact that a pre-conditioner can greatly enhance the performance of the CG method. Under mild conditions, it is shown that the algorithm is globally convergent for strongly convex functions. Numerical results are presented to show that the new diagonal PRECG method works better than the standard CG method.

*Keywords:* Unconstrained optimization, conjugate gradient method, preconditioning, diagonal approximation for Hessian

## INTRODUCTION

The nonlinear conjugate gradient (CG) method was designed to solve the following unconstrained optimization problem:

$$\min_{x \in R^n} f(x), \tag{1}$$

where $f : R^n \to R$ is a continuously differentiable function (Ismail Mohd *et al*., 2007). The iterative formula of the CG method is given by

$$x_{k+1} = x_k + \lambda_k d_k \tag{2}$$

and

$$d_k = -g_k + \gamma_k d_{k-1}, \tag{3}$$

where, $d_0 = -g_0$ for $k = 0,1,\ldots n$ and $g_k = \nabla f(x_k)$, $\lambda_k$ were chosen to satisfy some line search conditions along the search direction, $d_k$, and $\gamma_k$ is a scalar parameter.

The idea of incorporating a pre-conditioner to the CG method is initiated by Raydan (1997), where the spectral gradient is combined with the conjugate gradient

*Article history:*
Received: 2 February 2012
Accepted: 30 October 2012

*E-mail addresses:*
boon85@live.com.my (Choong Boon Ng),
leongwj@upm.edu.my (Wah June Leong),
mmonsi@upm.edu.my (Mansor Monsi)
*Corresponding Author

ISSN: 0128-7680 

directions. This is an iterative algorithm to generate a sequence $x_k, k = 0,1,2,\ldots,n$ as presented in (1) and (2), where $d_k$ is a spectral gradient search direction in the successive iterations generated by:

$$d_{k+1} = -\beta_k g_{k+1} + \gamma_k s_k, \tag{4}$$

where $\beta_k$ is a parameter that defines different CG methods, and $s_k = x_{k+1} - x_k$ (Andrei, 2010).

Unexpectedly, the spectral gradient choice associated in this algorithm shows more efficient results than sophisticated CG methods in many cases. It showed that spectral gradient and conjugate gradient combination produced more efficient algorithms (Raydan, 1997).

Andrei (2007) presented a new preconditioned conjugate gradient (PRECG) method, where the scaled memory-less BFGS update was used as the pre-conditioner. The scaling factor in the pre-conditioner was selected as a matrix, which was reset when the Powell restarted criterion (Powell, 1977) holds to ensure that the search directions would be descent directions. Consider (1), where the function *f* has continuous partial derivatives, and $d_k$ is a search direction generated by:

$$d_{k+1} = -\theta_k g_{k+1} + \gamma_k s_k, \tag{5}$$

For $k = 0,1,2,\ldots,n,$ where $\theta_k$ is a parameter to be computed, and $g_k$ denoting $\nabla f(x_k)$ is selected to minimize *f* along the search direction, $d_k$, with $s_k = x_{k+1} - x_k = \lambda_k d_k$, and $\gamma_k$ is a scalar parameter to be determined. $x_0 \in R^n$ is an arbitrary initial value and the iterative process is initialized with an initial point $x_0$ and $d_0 = -g_0$.

From the success of the spectral gradient method used by Raydan (1997) and Andrei's scaled memory-less BFGS method (Andrei, 2007) in the preconditioning technique, we developed a new pre-conditioner, $D_k$, which is a diagonal matrix based on both the spectral gradient and matrix preconditioning ideas.

## DIAGONAL PRECONDITIONED CONJUGATE GRADIENT ALGORITHM

To incorporate the correct inverse Hessian information into the preconditioner, $D_k$, we let the diagonal pre-conditioner $D_k$ to satisfy the weak-secant equation of Dennis and Wolkowicz (1993), as follows:

$$y_k^T s_k = y_k^T D_k y_k \tag{6}$$

With this aim, we let $\mathrm{s_k} = x_k - x_{k-1}$ and $\mathrm{y_k} = g_k - g_{k-1}$, and consider the minimization problem:

$$\min \frac{1}{2} \| D_k - I \|_F^2 \tag{7a}$$

$$\text{s.t.} y_k^T \left( D_k - I \right) y_k = y_k^T s_k - y_k^T y_k \tag{7b}$$

where $\|.\|_F$ denotes the standard Frobenius norm.

Since the objective function in (7a) and the feasible set is convex, it gives a unique solution for problems (7a) and (7b). Using the method of Lagrangian function, we can obtain:

$$D_k - I = \frac{(y_k^T y_k - y_k^T s_k)}{y_k^T G_k y_k} G_k \tag{8}$$

where $G_k = diag\left(\left((s)_k^{(1)}\right)^2, \ldots, \left((s)_k^{(n)}\right)^2\right)$.

Finally, by substituting $y_k{}^T G_k y_k = tr\left(G_k^2\right)$, where $tr(.)$ denotes the trace operator, it gives a diagonal pre-conditioner, which satisfies (6), as follows:

$$D_k = I + \frac{(y_k{}^T s_k - y_k{}^T y_k)}{tr(G_k{}^2)} G_k \tag{9}$$

It is shown that the diagonal preconditioner $D_k$ in (9) is a special class of diagonal Hessian approximation derived by Leong *et al*. (2010), Farid and Leong (2011), Leong *et al*. (2011), Farid *et al*. (2010), Leong and Hassan (2009), and Hassan *et al.* (2009).

In this case, the new PRECG method search direction $d_{k+1}$ is given by:

$$d_{k+1} = -D_{k+1} g_{k+1} + \theta_{k+1}\left(\frac{g_{k+1}^T s_k}{y_k^T s_k}\right) y_k - \left[\left(1 + \theta_{k+1}\frac{y_k^T y_k}{y_k^T s_k}\right)\frac{g_{k+1}^T s_k}{y_k^T s_k} - \theta_{k+1}\frac{g_{k+1}^T y_k}{y_k^T s_k}\right] s_k \tag{10}$$

The proposed pre-conditioner $D_k$ is in diagonal matrix form, where the storage requirements is of $O(n)$. Moreover, this pre-conditioner $D_k$ satisfies the weak-secant equation of Dennis and Wolkowicz (1993), which is a valid approximation of inverse Hessian.

The PRECG algorithm has the following steps:

Step 1. Given $x_0 \in R^n$, set $d_0 = -g_0$, $\lambda_0 = 1/\|g_0\|$ and $k = 0$. Update $x_1 = x_0 + \lambda_0 d_0$

Step 2. For $k \geq 1$, calculate $\lambda_k$ which satisfying Wolfe conditions. Compute the direction $d_k$ as in (10). Update the variables $x_{k+1} = x_k + \lambda_k d_k$. Then, compute $f(x_{k+1}), g_{k+1}$ and $s_k = x_{k+1} - x_k$, $y_k = g_{k+1} - g_k$.

Step 3. Test for the stopping condition. The iterations are stopped if stopping condition is satisfied. Else, set $k = k + 1$ and go to Step 2.

The Assumption 1 below is to guarantee the existing G is bounded.

**Assumption 1**

1. $f(x)$ is twice continuously differentiable and G denotes the matrix of second derivatives of $f(x)$.
2. G is bounded, that is $m_1 \| x \|^2 \le x^T \nabla^2 f(x) x \le m_2 \| x \|^2$, where $0 < m_1 \le m_2$.
3. Function $f(x)$ is strongly convex and has Lipschitz continuous on gradient in the level set $L_0 = \{x \in R^n : f(x) \le f(x_0)\}$, where there exists constants $\mu > 0$ and $L$ such that $(\nabla f(x) - \nabla f(y))^T (x - y) \ge \mu \| x - y \|^2$ and $\| \nabla f(x) - \nabla f(y) \| \le L \| x - y \|$, for all $x$ and $y$ from $L_0$.

This proposed pre-conditioner $D_k$ is proved to be bounded in such a way that we can expect the corresponded PRECG method to converge globally. With the aim to show that the proposed preconditioner $D_k$ is bounded, so we have the following lemma:

**Lemma 1:** Assume that $\| D_0 \|_F \le \sigma_0$, where $\sigma_0$ is a constant. Then, for all $k \ge 0$, $\| D_k \|_F \le \sigma_k$, where $\sigma_k$ is some constant and $\sigma_k \ge 0$. If we can show that the diagonal pre-conditioner $D_k$ satisfies $\| D_k \|_F \le \sigma_k$, and then the diagonal pre-conditioner $D_k$ is bounded above.

**Proof:** Let $D_k = diag\left(d_k^{(i)}\right)$, $G_k = diag(y_k^{(1)^2}, \ldots, y_k^{(n)^2})$ and $y_k^{(M)}$ be the largest component of $y_k$.

Then from (9), we have

$$d_k^{(i)} = 1 + \frac{(y_k^T s_k - y_k^T y_k)}{tr(G_k^2)} ((y)_k^{(i)})^2.$$

It follows from Assumption 1 that we have,

$$d_k^{(i)} \le 1 + \frac{\frac{1}{m_2} \| y_k \|^2 - \| y_k \|^2}{\sum (y_k^{(i)})^4} (y_k^{(M)})^2.$$

and the fact that $\| y_k \|^2 = \sum (y_k^{(i)})^2 \le n (y_k^{(M)})^2$ gives

$$d_k^{(i)} \le 1 + \frac{n \left| \frac{1}{m_2} - 1 \right|}{\sum (y_k^{(i)})^4} (y_k^{(M)})^4.$$

Finally, it leads to $d_k^{(i)} \le 1 + n \left| \frac{1}{m_2} - 1 \right|$.

Hence, we have $\| D_k \| \leq \sigma_k$ where $\sigma_k = \sqrt{n + n^2 \left| \frac{1}{m_2} - 1 \right|}$. □

Below is the convergence result of our new algorithm when the objective function $f(x)$ satisfies Assumption 1 (iii).

**Theorem 1.1** If at every step of the conjugate gradient given in (2) with the step length $\lambda_k$ selected to satisfy the Wolfe conditions (Wolfe, 1969) and $d_{k+1}$ is given by (10), then either $g_k = 0$ for some *k* or $\lim_{k \to \infty} g_k = 0$.

**Proof:** From $d_0 = -g_0$, we have $g_0^T d_0 = - \| g_0 \|^2 \leq 0$. When (10) is multiplied by $g_{k+1}^T$:

$$g_{k+1}^T d_{k+1} \leq \frac{1}{(y_k^T s_k)^2} \Big[ -\sigma_{k+1} \| g_{k+1} \|^2 + 2\theta_{k+1} \left( g_{k+1}^T y_k \right) \left( g_{k+1}^T s_k \right) \left( y_k^T s_k \right) - \ g^T{}_{+1} s \quad y^T s \ - \ {}_{+1} (y^T y \ ) (g^T$$

$$- \left( g_{k+1}^T s_k \right)^2 \left( y_k^T s_k \right) - \theta_{k+1} (y_k^T y_k) (g_{k+1}^T s_k)^2 \Big]$$

and with $u = (s_k^T y_k) g_{k+1}$ and $v = (g_{k+1}^T s_k) y_k$, we can then get the following by applying the inequality $u^T v \leq \frac{1}{2} (\| u \|^2 + \| v \|^2)$ to the second term of the right hand side of the above inequality:

$$g_{k+1}^T d_{k+1} \leq - \frac{\left( g_{k+1}^T s_k \right)^2}{y_k^T s_k}. \tag{11}$$

Therefore, by Wolfe's condition, $g_{k+1}{}^T s_k \geq \beta_2 g_k{}^T s_k$, $g_{k+1}^T d_{k+1} < 0$ for every $k = 0, 1, 2, \ldots, n.$

By strong convexity, we have $y_k^T d_k = (g_{k+1} - g_k)^T d_k \geq \mu \ \lambda_k \| d_k \|^2$.

Here $g_k \neq 0$ implies for all *k* by Theorem 1.1, where $g_k^T d_k < 0$. *f* is bounded from below due to its strongly convex over $L_0$. By summing over *k*, the Wolfe conditions $f(x_{k+1}) < f(x_k) + \beta_1 g_k{}^T s_k$, we have:

$$\sum_{k=0}^{\infty} \lambda_k g_k^T d_k \to -\infty.$$

Consider that $d_k$ is a descent direction and the lower bound for $\lambda_k$ which satisfies the Wolfe condition, $g_{k+1}{}^T s_k \geq \beta_2 g_k{}^T s_k$, then

$$\lambda_k \geq \frac{1 - \sigma_2 \left| g_k^T d_k \right|}{L \| d_k \|^2}.$$

and follows with

$$\sum_{k=1}^{\infty} \frac{\left| g_k^T d_k \right|^2}{\| d_k \|^2} < \infty. \tag{12}$$

Using the inequality of Cauchy and by strong convexity, we have $y_k^T s_k \geq \mu \| s_k \|^2$ and get

$$g_{k+1}^T d_{k+1} \leq -\frac{(g_{k+1}^T s_k)^2}{y_k^T s_k} \leq -\frac{\| g_{k+1} \|^2 \| s_k \|^2}{\mu \| s_k \|^2} = -\frac{\| g_{k+1} \|^2}{\mu}.$$

Hence, from (12) it follows that,

$$\sum_{k=0}^{\infty} \frac{\| g_k \|^4}{\| d_k \|^2} < \infty. \tag{13}$$

From (10), when $\theta_{k+1}$ is selected by spectral gradient, the direction $d_{k+1}$ will then satisfy:

$$\| d_{k+1} \| \leq \left( \frac{2}{\mu} + \frac{2L}{\mu^2} + \frac{L^2}{\mu^3} \right) \| g_{k+1} \|. \tag{14}$$

By inserting the upper bound (14) for $d_k$ in (13) will yield the following:

$$\sum_{k=0}^{\infty} \| g_k \|^2 < \infty,$$

which completes the proof. □

## NUMERICAL RESULTS

In this section, we discuss some numerical experiments that are conducted in order to test the performance of our new gradient method for unconstrained optimization against the standard CG method.

We compare the performance of a Fortran implementation of our new algorithm with the standard CG algorithm on a set of 50 large-scale unconstrained optimization test problems in extended or generalized form(Andrei, 2008). All tests are run on a 2.6 GHz Pentium IV with 512MB of RAM and all algorithms are coded in Fortran commands. We have considered a number of variables $n = 1000, 2000, \ldots, 10000$ for each problem. For all the test runs, the termination condition is $\| g_k \| \leq 10^{-6}$. The accumulated number of iterations and the average of the norm of gradient are used to compare the effectiveness of the results.

For convenience, the following abbreviations are used to identify a particular conjugate gradient method.

1. SCG: The standard conjugate gradient method (without pre-conditioner).
2. PRECG: The preconditioned conjugate gradient method with

$$D_k = I + \frac{(y_k{}^T s_k - y_k{}^T y_k)}{Tr(G_k{}^2)} G_k$$

Table 2 gives a summary of comparison results between PRECG algorithm and the standard CG method. The symbol *prob* and $\| g_k \|$ mean the number of the test problems and the norm of the gradient of the function, respectively. The $Iter$ means total iteration calls. Table 3 gives the comparison results of the number of function evaluation for all the methods. Meanwhile, Table 4 summarizes the performance of the PRECG algorithm versus the SCG algorithm on 50 problems, which achieved the least number of iteration and a lower value of gradient norms. The test problems are listed in Table 1.

From Table 4, the PRECG algorithm performs better than the SCG algorithm to achieve a minimum norm of gradients, with 31 problems out of 50 problems, as compared to the SCG algorithm, which is only achieved for 16 problems. In more specific, the percentage of efficiency for the PRECG algorithm is 30% compared to the SCG algorithm.

TABLE 1
Test Problems and their corresponding problem number (prob) (refer to Andrei, 2008)

| Problem | Test problems |
|---|---|
| 1 | Extended Freudenstein & Roth Function |
| 2 | Extended Trigonometric Function |
| 3 | Extended Rosenbrock Function |
| 4 | Extended White & Holst Function |
| 5 | Extended Beale Function |
| 6 | Extended Penalty Function |
| 7 | Perturbed Quadratic Function |
| 8 | Raydan 1 Function |
| 9 | Raydan 2 Function |
| 10 | Diagonal 1 Function |
| 11 | Diagonal 2 Function |
| 12 | Diagonal 3 Function |
| 13 | Hager Function |
| 14 | Generalized Tridiagonal 1 Function |
| 15 | Extended Tridiagonal 1 Function |
| 16 | Extended Three Expo Terms Function |
| 17 | Generalized Tridiagonal 2 Function |
| 18 | Diagonal 4 Function |
| 19 | Diagonal 5 Function |
| 20 | Extended Himmelblau Function |
| 21 | Generalized PSC1 Function |

TABLE 1 (continue)

| Problem | Test problems |
|---|---|
| 22 | Extended PSC1 Function |
| 23 | Extended Powell Function |
| 24 | Extended Block-Diagonal BD1 Function |
| 25 | Extended Maratos Function |
| 26 | Extended Cliff Function |
| 27 | Quadratic Diagonal Perturbed Function |
| 28 | Extended Wood Function |
| 29 | Extended Hiebert Function |
| 30 | Quadratic QF1 Function |
| 31 | Extended Quadratic Penalty QP1 Function |
| 32 | Extended Quadratic Penalty QP2 Function |
| 33 | Quadratic QF2 Function |
| 34 | Extended EP1 Function |
| 35 | Extended Tridiagonal 2 Function |
| 36 | BDQRTIC (CUTE) Function |
| 37 | TRIDIA (CUTE) Function |
| 38 | ARWHEAD (CUTE) Function |
| 39 | NONDIA (CUTE) Function |
| 40 | NONDQUAR (CUTE) Function |
| 41 | DQDRTIC (CUTE) Function |
| 42 | EG2 (CUTE) Function |
| 43 | DIXMAANA (CUTE) Function |
| 44 | DIXMAANB (CUTE) Function |
| 45 | DIXMAANC (CUTE) Function |
| 46 | DIXMAANE (CUTE) Function |
| 47 | Partial Perturbed Quadratic PPQ1 Function |
| 48 | BroydenTridiagonal Function |
| 49 | Almost Perturbed Quadratic Function |
| 50 | Tridiagonal Perturbed Quadratic Function |

TABLE 2
A comparison of the CG and PRECG methods in terms of total iteration calls and gradient norm

| *prob* | CG algorithm | | PRECG algorithm | |
|---|---|---|---|---|
| | *Iter* | $\|\| g_k \|\|$ | *Iter* | $\|\| g_k \|\|$ |
| 1 | 91 | 1.47e-05 | 71 | 3.85e-06 |
| 2 | 701 | 3.97e-06 | 696 | 2.83e-06 |
| 3 | 248 | 9.27e-06 | 240 | 7.08e-06 |
| 4 | 305 | 9.36e-06 | 301 | 5.79e-06 |

TABLE 2 (continue)

| *prob* | CG algorithm | | PRECG algorithm | |
|---|---|---|---|---|
| | *Iter* | $\|\| g_k \|\|$ | *Iter* | $\|\| g_k \|\|$ |
| 6 | 1067 | 1.12e-06 | 1061 | 1.93e-06 |
| 7 | 7080 | 5.72e-06 | 7076 | 5.31e-06 |
| 8 | 6057 | 4.45e-06 | 5991 | 4.32e-06 |
| 9 | 30 | 3.57e-07 | 30 | 3.57e-07 |
| 10 | 7680 | 3.84e-06 | 7256 | 4.31e-06 |
| 11 | 4773 | 1.90e-06 | 3954 | 1.77e-06 |
| 12 | 16969 | 6.97e-05 | 17245 | 6.71e-05 |
| 13 | 7896 | 1.84e-06 | 9732 | 1.50e-06 |
| 14 | 370 | 2.55e-06 | 380 | 2.34e-06 |
| 15 | 102 | 2.64e-05 | 68 | 1.78e-05 |
| 16 | 67 | 3.33e-06 | 60 | 6.12e-06 |
| 17 | 546 | 2.44e-06 | 532 | 2.71e-06 |
| 18 | 20 | 1.76e-06 | 31 | 4.32e-11 |
| 19 | 30 | 2.81e-10 | 30 | 2.81e-10 |
| 20 | 70 | 5.66e-06 | 63 | 6.29e-07 |
| 21 | 6075 | 2.17e-06 | 7134 | 1.79e-06 |
| 22 | 80 | 4.50e-09 | 89 | 3.38e-06 |
| 23 | 545 | 1.95e-05 | 573 | 2.88e-05 |
| 24 | 219 | 2.74e-05 | 2286 | - |
| 25 | 448 | 9.91e-05 | 447 | 1.75e-05 |
| 26 | 207 | 1.26e-05 | 16059 | - |
| 27 | 3026 | 9.15e-06 | 3131 | 8.83e-06 |
| 28 | 989 | 1.98e-05 | 1079 | 1.94e-05 |
| 29 | 512 | 4.12e-06 | 520 | 1.95e-06 |
| 30 | 6923 | 5.82e-06 | 7195 | 5.42e-06 |
| 31 | 528 | 1.17e-06 | 685 | 5.51e-06 |
| 32 | 128 | 1.98e-07 | 250 | 1.74e-07 |
| 33 | 7836 | 4.49e-06 | 8283 | 4.50e-06 |
| 34 | 18 | 6.56e-05 | 18 | 6.56e-05 |
| 35 | 319 | 3.44e-06 | 331 | 2.93e-06 |
| 36 | 18300 | 5.19e-05 | 18572 | 1.08e-04 |
| 37 | 19080 | 7.18e-03 | 18649 | 4.60e-02 |
| 38 | 31 | 1.29e-07 | 30 | 1.39e-07 |
| 39 | 298 | 2.51e-07 | 109 | 2.30e-08 |
| 40 | 19465 | 6.46e-05 | 18261 | 9.28e-06 |
| 41 | 51 | 2.53e-07 | 50 | 2.50e-07 |
| 42 | 16779 | 6.03e-04 | 11096 | 2.41e-04 |

TABLE 2 (continue)

| *prob* | CG algorithm | | PRECG algorithm | |
|---|---|---|---|---|
| | *Iter* | $\|\| g_k \|\|$ | *Iter* | $\|\| g_k \|\|$ |
| 44 | 149 | 2.26e-06 | 193 | 5.17e-06 |
| 45 | 174 | 4.89e-06 | 245 | 1.36e-06 |
| 46 | 4015 | 6.14e-06 | 3967 | 6.41e-06 |
| 47 | 1107 | 1.72e-05 | 1053 | 1.60e-05 |
| 48 | 634 | 2.22e-06 | 645 | 2.83e-06 |
| 49 | 7360 | 5.17e-06 | 7002 | 4.98e-06 |
| 50 | 6826 | 6.57e-06 | 7149 | 6.08e-06 |

TABLE 3
A comparison of the methods in terms of the Total Number of Function Evaluation for all n

| *Pro* | Methods (Number of functions calls) | | *Pro* | Methods (Number of functions calls) | |
|---|---|---|---|---|---|
| | CG | PRECG | | CG | PRECG |
| 1 | 265 | 178 | 26 | 300 | 16147 |
| 2 | 1076 | 1068 | 27 | 4912 | 5090 |
| 3 | 489 | 482 | 28 | 1875 | 2030 |
| 4 | 613 | 546 | 29 | 1097 | 1085 |
| 5 | 233 | 188 | 30 | 9084 | 9456 |
| 6 | 25945 | 26366 | 31 | 12937 | 15435 |
| 7 | 9281 | 9228 | 32 | 308 | 590 |
| 8 | 8468 | 8376 | 33 | 10314 | 10902 |
| 9 | 90 | 90 | 34 | 59 | 59 |
| 10 | 21881 | 17385 | 35 | 582 | 553 |
| 11 | 6999 | 5781 | 36 | 478435 | 468394 |
| 12 | 354441 | 355758 | 37 | 25118 | 24330 |
| 13 | 238400 | 298785 | 38 | 101 | 80 |
| 14 | 5306 | 5210 | 39 | 1379 | 224 |
| 15 | 209 | 156 | 40 | 33513 | 32610 |
| 16 | 138 | 140 | 41 | 132 | 130 |
| 17 | 884 | 860 | 42 | 468361 | 294247 |
| 18 | 60 | 81 | 43 | 136 | 179 |
| 19 | 90 | 90 | 44 | 248 | 2808 |
| 20 | 160 | 153 | 45 | 298 | 3677 |
| 21 | 67721 | 67912 | 46 | 5521 | 5159 |
| 22 | 180 | 189 | 47 | 1773 | 1713 |
| 23 | 1015 | 1050 | 48 | 1010 | 1052 |
| 24 | 474 | 3294 | 49 | 9649 | 9253 |
| 25 | 970 | 932 | 50 | 8986 | 9386 |

TABLE 4
A comparison of CG and PRECG (All problems)

| | CG performs better | PRECG performs better | Equal performance |
|---|---|---|---|
| function calls | 21 | 26 | 3 |
| norm of gradient | 16 | 31 | 3 |

## CONCLUSION

In this paper, we presented the performance of a new diagonal preconditioned conjugate gradient (PRECG) method. Accordingly, the PRECG method was also compared against the SCG method based on 50 benchmark problems. Based on our numerical experiments, the PRECG method has been shown to outperform the SCG method. Thus, it is concluded that the introduction of PRECG method is worthwhile.

## REFERENCES

Andrei, N. (2008). An unconstrained optimization test functions collection. *Advanced Modeling & Optimization, 10*(1), 147-161.

Andrei, N. (2007). Scaled BFGS preconditioned conjugate gradient algorithm for unconstrained optimization. *Optimization Methods & Software, 22*(4), 561-571.

Andrei, N. (2010). Accelerated scaled memoryless BFGS preconditioned conjugate gradient algorithm for unconstrained optimization. *European Journal of Operational Research*, *204*(3), 410-420.

Dennis, J. E., & Wolkowicz, H. (1993). Sizing and least-change secant methods. *SIAM Journal on Numerical Analysis*, *30*(5), 1291-1314.

Farid *et al*. (2010). A new two-step gradient-type method for large-scale unconstrained optimization. *Computers and Mathematics with Applications*, *59*(10), 3301-3307.

Farid, & Leong (2011). An improved multi-step gradient-type method for large scale optimization. *Computers and Mathematics with Applications*, *61*(11), 3312-3318.

Fletcher, R., & Reeves, C.M. (1964). Function minimization by conjugate gradients. *The Computer Journal*, *7*(2), 149-154.

Hassan *et al*. (2009). A new gradient method via quasi-Cauchy relation which guarantees descent. *Journal of Computational and Applied Mathematics*, *230*(1), 300-305.

Hestenes, M. R., & Stiefel, E.L. (1952). Methods of conjugate gradients for solving linear systems, J. *Research Nat. Bur. Standards*, *49*, 409-436.

Ismail bin Mohd *et al*. (2007). Unimodality Tests for Global Optimization of Single Variable Functions Using Statistical Methods. *Malaysian Journal of Mathematical Sciences, 1*(2), 205-215.

Leong, & Hassan (2009). A restarting approach on symmetric rank one update for unconstrained optimization. *Computational Optimization and Applications*, *42*(3), 327-334.

Leong *et al*. (2010). A monotone gradient method via weak secant equation for unconstrained optimization. *Taiwanese Journal of Mathematics*, *14*(2), 412-423.

Polak, E., & Ribiere, G. (1969). Note sur la convergence de directions conjugees, Rev. *Francaise Informat Recherche Opertionelle, 16*, 35-43.

Raydan, M. (1997). The Barzilain and Borwein gradient method for the large unconstrained minimization problem. *SIAM Journal on Optimization*, *7*, 26–33.

Wolfe, P. (1969). Convergence conditions for ascent methods. *SIAM Review, 11*, 226-235.

Pertanika J. Sci. & Technol. 22 (1): 225 - 237 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Performance Evaluation and Characteristics of Selected Tube Wells in the Coastal Alluvium Aquifer, Selangor

**Fauzie, M. J.[1]*, Azwan, M. M. Z.[1], Hasfalina, C. M.[1] and Mohammed, T. A.[2]**

*[1]Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*
*[2]Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

Alluvial aquifers can be found in most of the coastal areas of Peninsular Malaysia. Seven tube wells located in such aquifers in the west coast of Selangor state had their performance evaluated by carrying-out step drawdown tests. The performance of these wells was evaluated in terms of aquifer loss, well loss, specific capacity and well efficiency. The aquifer loss coefficient and well loss coefficient were found to be in the range of 0.0198 $hrm^{-2}$ to 0.4014 $hrm^{-2}$ and from 0.0001 $hr^2m^{-5}$ to 0.0410 $hr^2m^{-5}$, respectively. The drawdown in tube wells TW1 and TW7 is mainly influenced by well loss component as compared to the aquifer loss component, while in tube wells TW2, TW3, TW4, and TW5, the drawdown is mainly influenced by aquifer loss component. The drawdown in tube well TW6 is influenced by aquifer loss component at a low discharge rate, but at high discharge rate, it is influenced by well loss component. The specific capacity and efficiency of the tested tube wells varied from 1.329 $m^2hr^{-1}$ to 40.166 $m^2hr^{-1}$, and from 11% to 96%, respectively. Tube wells TW2 and TW4 are categorized as high productive wells, while tube wells TW1, TW3, TW5 and TW7 are categorized as moderate productive wells and tube well TW6 as low productive well.

*Keywords:* Step drawdown test, well loss, aquifer loss, specific capacity, well efficiency

## INTRODUCTION

In Malaysia, tube wells are used widely for various purposes as in agriculture and for domestic and industrial uses. Heng (2004) reported that there are about 2,466 wells drilled throughout Peninsular Malaysia starting from 1983 with a total yield of 552,000 $m^3day^{-1}$. Since the last century, extraction of groundwater in Malaysia has

*Article history:*
Received: 13 February 2012
Accepted: 18 April 2012

*E-mail addresses:*
fauzie_jsh@yahoo.com (Fauzie, M. J.),
mohdazwan@upm.edu.my (Azwan, M. M. Z.),
hasfalina@upm.edu.my (Hasfalina, C. M.),
thamer@upm.edu.my (Mohammed, T. A.)
*Corresponding Author

ISSN: 0128-7680 

increased because of many factors such as surface water depletion due to drought and increasing water demands for the domestic, agricultural and industrial sectors (Heng, 2004). Kelantan is one of the states in Malaysia which is still using groundwater as a major source for domestic water supply (Samsudin *et al.*, 2008), with a total consumption of 146 $Mld^{-1}$ in 2010 abstracted from 94 production wells at 14 well fields and treated by 7 groundwater treatment plants (Ismail *et al.*, 2011). All the production wells should be monitored and maintained each year in order to make sure that each well can produce enough quantity of water according to its design capacity.

Tube wells for water production are designed based on soil lithology and information obtained from exploration well during site investigation. When tube well has been constructed and developed, its performance evaluation is carried out. As mentioned by Shekhar (2006), groundwater users are always concerned with the performance of well structure and the relationship between discharge and drawdown at the pumping wells. Step drawdown test is widely used to identify well behaviour, determine well loss and calculate well efficiency (Kawechi, 1995). The drawdown inside the pumping well is influenced by aquifer loss component, $BQ$ and well loss component, $CQ^n$, with $n$ as the well loss exponent (Todd & Mays, 2005; Mohammed & Huat, 2004; Rahman & Dhar, 1997; Sheahan, 1971). The values of aquifer loss coefficient, $B$, and well loss coefficient, $C$ (Mishra & Sahay, 2011), as well as reliable yields estimation of water well (Misstear & Beeson, 2000) are important for a successful modelling and proper management of groundwater resources.

Drawdown that occurs at the face of the well is known as aquifer loss, whereas drawdown that occurs as water moves through the well screen and inside the well to the pump suction area is known as well loss (Mohammed & Huat, 2004). According to Todd and Mays (2005), the coefficient of well loss, $C$, is controlled by its radius, development and condition. The relationship between well loss coefficient, $C$, and well conditions is shown in Table 1, while the relationship between specific capacity and types of well productivity is shown in Table 2. Therefore, this paper aimed at evaluating tube well performance and identifying its characteristics by evaluating the parameters of well loss, aquifer loss, specific capacity, and well efficiency.

## MATERIALS AND METHODS

The study area is located in the coastal area of Selangor involving Sabak Bernam, Kuala Selangor and Kuala Langat districts as shown in Fig.1. The study area is covered with alluvium sediments which consist of peat soil, silt, sand and gravel. The wells were developed by the Department of Mineral and Geoscience Malaysia (DMGM) by using rotary mud drilling machine (BOMAG & Holy Machine Model CD80). The wells were drilled up to bedrock layer, with depths ranging between 30 m to 60 m. All tube wells use gravel pack in the sizes of 3 mm to 6 mm, except for tube wells TW5 and TW6 which use gravel pack in sizes of 3 mm to 4 mm. The diameter of the tube well casing varies between 200 mm and 250 mm. PVC material was chosen as the casing and screen at most of the developed wells, while TW5 and TW6 used mild steel and stainless steel as their casing and screen material, respectively. Further information on the tube wells is given in Table 3.

TABLE 1
Relationship between well loss coefficient, C, to well condition (Walton, 1962)

| Well loss coefficient, C ($hr^2m^{-5}$) | Well Conditions |
|---|---|
| C< 0.0001 | Great – Well is designed and developed properly |
| 0.0001< C < 0.0002 | Good – Mild deterioration due to clogging |
| 0.0002 < C < 0.0011 | Fair to Poor – Severe deterioration due to clogging |
| C > 0.0011 | Bad – Difficult to restore well to original capacity |

TABLE 2
Specific capacities values and well productivity classification (Şen, 1995)

| Specific Capacity, Sc ($m^3hr^{-1}m^{-1}$) | Well Productivity |
|---|---|
| C > 18 | High |
| 18 > C > 1.8 | Moderate |
| 1.8> C > 0.18 | Low |
| 0.18 > C > 0.018 | Very Low |
| 0.018 > C | Negligible |


Fig.1: Locations of the tube wells at the coastal area of Selangor (Selangor Map, 2012)

TABLE 3
A detailed description of the tube wells studied

| Well Name | District | Coordinate | Well Depth (m) | Gravel Pack Size (mm) | Water Table Depth (m b.g.l) | Casing Diameter (mm) | Screen Diameter (mm) | Screen Location (m) | Screen Length (m) | Aquifer Thickness (m) | Drilling Method* | Year of Testing |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TW 1 | Sabak Bernam | 412300 N 348600 E | 48 | 3-6 | 0.07 | 203 | 203 | 42-48 | 6 | 15 | 1 | 2003 |
| TW 2 | Kuala Selangor | 377100 N 372500 E | 40 | 3-6 | 0.47 | 200 | 155 | 28-40 | 12 | 27 | 1 | 2000 |
| TW 3 | Kuala Selangor | 377500 N 380300 E | 40 | 3-6 | 2.35 | 200 | 155 | 28-40 | 12 | 30 | 1 | 2000 |
| TW 4 | Kuala Langat | 328800 N 378800 E | 57 | 3-6 | 4.97 | 254 | 254 | 44-56 | 12 | 15 | 1 | 2005 |
| TW 5 | Kuala Langat | 318033 N 403740 E | 27.8 | 3-4 | 1.66 | 250 | 250 | 11.5-14.5, 18-21, 24-27 | 9 | 12 | 2 | 2009 |
| TW 6 | Kuala Langat | 318687 N 402942 E | 31.8 | 3-4 | 0.95 | 250 | 250 | 21-30 | 9 | 16 | 2 | 2009 |
| TW 7 | Kuala Langat | 312200 N 395100 E | 60 | 3-6 | 1.78 | 200 | 150 | 23-60 | 37 | 38 | 1 | 1995 |

*1: Rotary Mud Drilling (Model BOMAG), 2: Rotary Mud Drilling (Model Holy Machine CD80)
m.b.g.l- meter below ground level, N-north, E-east

Most of the tube wells were developed for the purpose of groundwater resources potential study for future groundwater development plan. Some of the water wells were constructed for the emergency plan or as an alternative source of water if forest fire happens since most of the study areas are covered with peat soil. Step drawdown test was executed between the year 1996 and 2009 and this test was carried out immediately after well construction had been completed. Submersible pumps were used to pump the water from the tube wells and the discharge rates were measured with a weir tank. The valve was installed to control and vary the discharge rates.

Step drawdown test is a single well test where the water is pumped at a low constant discharge rate until the drawdown within the well stabilizes. The lowest discharge rate is known as step 1 and the test is repeated by increasing the rate of pumping to the second pumping rate (step 2) until the drawdown within the well stabilizes once more, as in step 1 (Krusseman & de Rider, 1994). Mahajan (1989) mentioned that step drawdown tests are started at a low step; for instance, 25% of the designed capacity and increased up to 50%, 70%, 100% and 125% of the designed capacity, depending on the number of steps chosen. In this study, every step was designed with four to five steps with each step period between 1 and 2 hours. The time, water level inside the tube well, and discharge data were taken during the test. Data on water level data were taken by using an automatic water level transducer or manually by using a water level indicator. Water level data were taken every 0.5 minutes for the first 10 minutes and every 5 minutes thereafter.

The general equation for calculating total drawdown in the pumping well is given by Rorabaugh (1953), as in Equation 1. In order to solve this equation, Jacob (1947) proposed a graphical method by assuming the power of well loss, $n$, as equals to 2. Under this assumption, Equation 1 can be rewritten as Equation 2. Bierschenk (1964) mentioned that the values of $B$ and $C$ from Equation 2 could be obtained from the plot of specific drawdown, $\frac{S_w}{Q}$ ($hrm^{-2}$) against discharge rate ($m^3hr^{-1}$). Specific drawdown is defined as the ratio of drawdown to the discharge rate. From the graph, $B$ is the intercept of y-axis and $C$ is the slope of the best straight line. Equation 2 can also be applied to the confined, unconfined and leaky aquifer types.

$$S_w = BQ + CQ^n \tag{1}$$

$$\frac{S_w}{Q} = B + CQ \tag{2}$$

Where $S$ is drawdown in pumping well (m), $Q$ is discharge rate ($m^3hr^{-1}$), $B$ is aquifer loss coefficient ($hrm^{-2}$), and $C$ is well loss coefficient ($hr^2m^{-5}$). Rorabaugh (1953) argued that the power of well loss exponent is not always equal to 2, but its values are varying between 1.5 and 3.5, depending on its discharge rate. However, well loss exponent equals to 2 as proposed by Jacob is still accepted and commonly used (Todd & Mays, 2005; Bierschenk, 1964; Skinner, 1988; Ramey, 1982, as cited in Krusseman & de Rider, 1994). In the present paper, time-drawdown data were analyzed by using the regression technique, as suggested by Bierschenk (1964), as well as the trend line fitting polynomial plot (Shekhar, 2006) to obtain the values of $B$ and $C$.

Well efficiency is defined as the ratio between theoretical drawdown and actual drawdown from step drawdown test and it was calculated based on Equation 3 (Todd & Mays, 2005):

$$\eta_w = \frac{BQ}{BQ + CQ^2} \tag{3}$$

Where $\eta_w$ is well efficiency (%), $BQ$ is aquifer loss, (m) and $CQ^2$ is well loss, (m). The specific capacity of the tested well was also calculated by dividing the discharge rate with its corresponding drawdown (Mohammed & Huat, 2004), as follows:

$$S_c = \frac{Q}{S} \tag{4}$$

Where $S_c$ is specific capacity ($m^2hr^{-1}$), $Q$ is discharge rate ($m^3hr^{-1}$) and $S$ is drawdown inside the pumping well (m).

## RESULTS AND DISCUSSION

Fig.2a shows the regression equation and the plot of specific drawdown versus discharge rate from step drawdown test. The coefficient of determination, $R^2$, showed a strong relationship between drawdown and discharge rate at TW1, TW2, TW5, TW6 and TW7, with $R^2$ greater than 0.8. Tube well TW4 showed a moderate relationship between drawdown and discharge rate, with $R^2$ equals to 0.5. The lowest coefficient of determination ($R^2 = 0.1$) is recorded for tube well TW3 due to the large variation between specific drawdown and discharge rate in each step. This implied that the plot of specific drawdown and discharge data for tube well TW3 and TW4 did not show a linear trend. Thus, the trend line fitting polynomial plot (Shekhar, 2006) was applied to the drawdown-discharge data from step drawdown test for tube wells TW3 and TW4 (see Fig.2b). The polynomial plot approach shows approximation of Rorabaugh general equation with well loss exponent equals to 2 and the equation formed shows a strong relationship ($R^2$=0.9) between drawdown and discharge. The aquifer loss coefficient, B, of the tested wells ranges from 0.0198 (in TW2) to 0.4014 (in TW6), while the well loss coefficient, C, ranges from 0.0001 (in TW2) to 0.0410 (in TW6). All the tube wells have the value of aquifer loss coefficient greater than the value of well loss coefficient. Based on Table 1, tube well TW2 is properly designed and developed since the value of well loss coefficient is 0.0001 $hr^2m^{-5}$.

Fig.3 exhibits the comparison between the observed and predicted drawdown in every step for each well. The observed drawdown (taken during execution of step drawdown test) was compared to the predicted drawdown, which is calculated from drawdown discharge relationship as shown in Equation 1, based on the computed values of aquifer loss and well loss. The plots were done to verify Equation 1 between field data and model developed (drawdown-discharge equation). In general, the results showed good concordance between the observed and predicted drawdown, especially at the low discharge rate (in the first and second steps). However, tube well TW3 showed a slight difference between the observed and predicted drawdown.

The comparison between aquifer loss and well loss in each step is shown in Fig.4. From this comparison, the tested tube wells can be categorized into three groups. The first group is the tube wells with the well loss greater than the aquifer loss in every step. The tube wells that fall in the first group are TW1 and TW7. The second group of the tube wells is the tube wells with the aquifer loss greater than well loss in every step. These tube wells are TW2, TW3, TW4 and TW5. The third group of the tube well is the tube well with the aquifer loss greater than well loss at the lower discharge rate, but at the higher discharge rate, the well loss is greater than aquifer loss. Only TW6 falls in the third tube well group.

The relationship between specific capacity and discharge rate at each tube well is presented in Fig.5. Tube well TW2 has the highest value of specific capacity. The resulting drawdown in this tube well is low as compared to the volume of groundwater abstracted. Tube well TW6 has

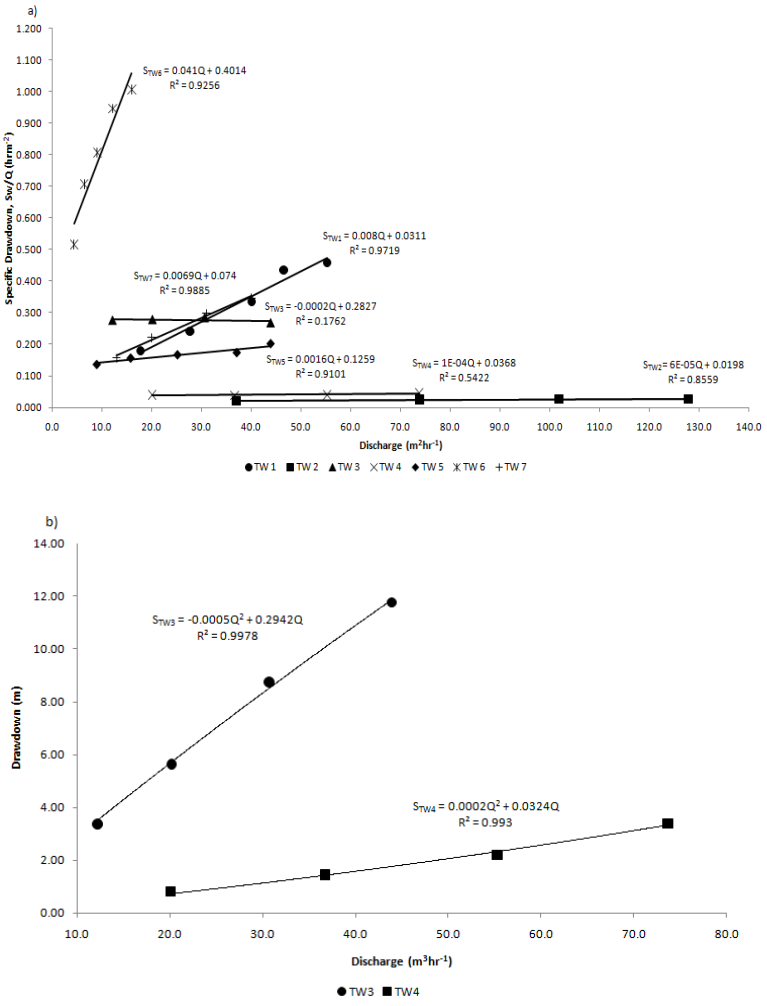



Fig.2: a) Regression equation and plot of specific drawdown versus discharge rate; b) Polynomial plot of drawdown and discharge rate (for TW3 and TW4) from step drawdown pumping test to determine aquifer loss coefficient, B and well loss coefficient, C.

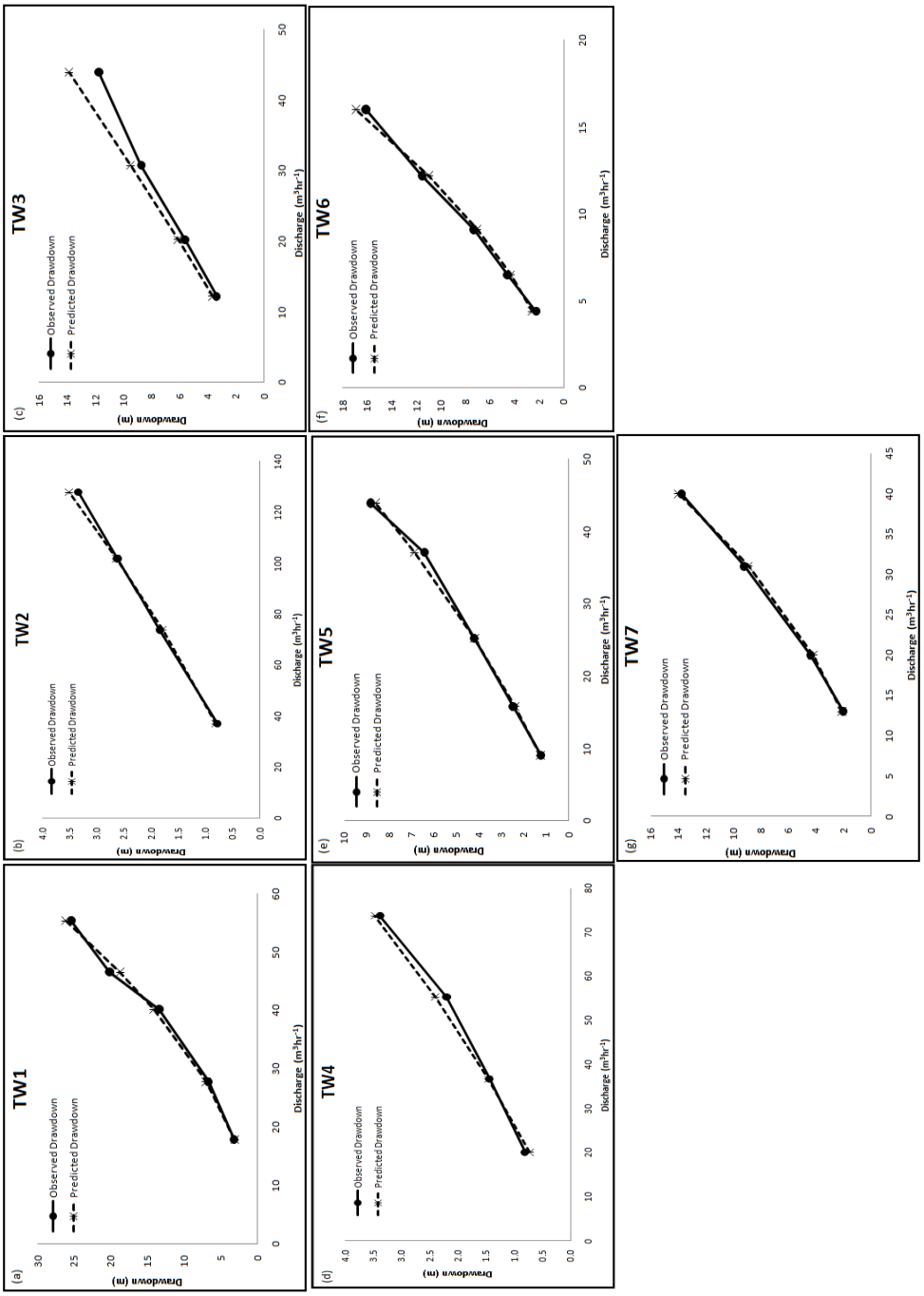



Fig.3: A comparison between the observed and predicted drawdown at: (a) TW1, (b) TW2, (c) TW3, (d) TW4, (e) TW5, (f) TW6, and (g) TW7

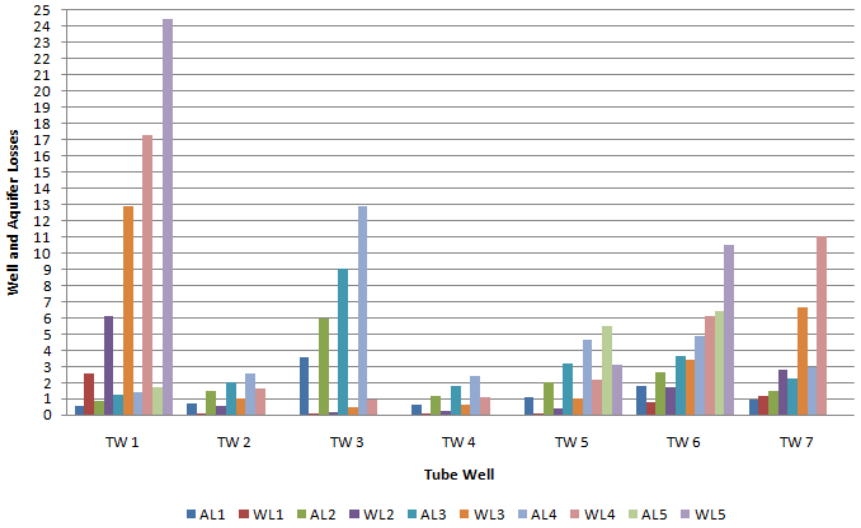



Fig.4: A comparison between aquifer loss and well loss for each step in step drawdown test (AL1 is aquifer loss from step 1, WL1 is the well loss for step 1, and so on)

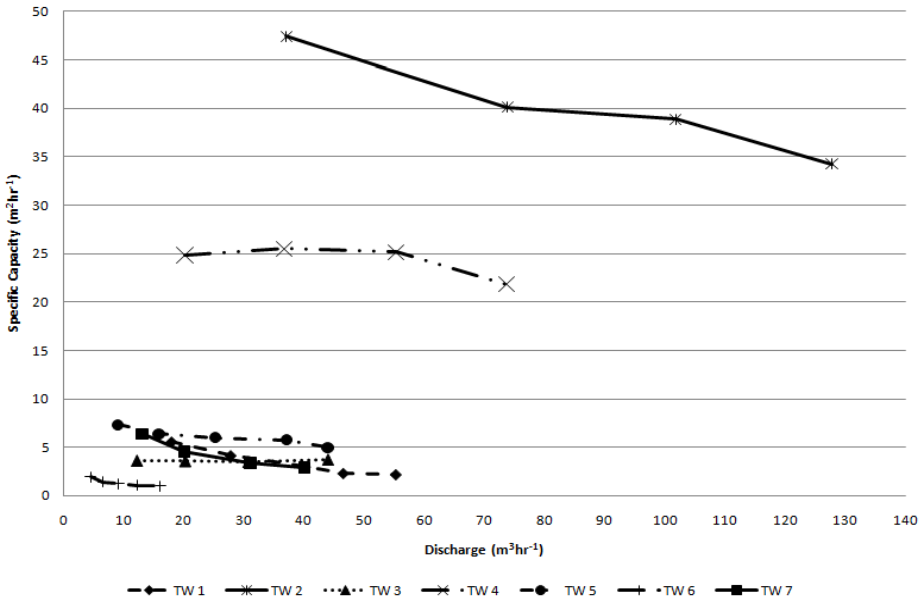



Fig.5: The plot of specific capacity versus discharge rate from step drawdown test

the lowest specific capacity due to the large amount of drawdown resulting from the increasing volume of groundwater withdrawal. Based on the specific capacity classification (see Table 2), according to Şen (1995), tube wells TW2 and TW4 are categorized as high productive wells with the specific capacity greater than 18 $m^2hr^{-1}$. Meanwhile, tube wells TW1, TW3, TW5 and TW7 are categorized as moderate productive tube wells, with the specific capacity ranging between 1.8 and 18 $m^2hr^{-1}$. Tube well TW6 is categorized as a low productive tube well, with the specific capacity having less than 1.8 $m^2hr^{-1}$.

Fig.6 shows the well efficiency of each step for every tube well. The results show that the efficiency of tube wells decreases as the discharge increases. The most efficient tube well is TW3, with an average efficiency of 96%, whereas the least efficient tube well is TW1, with an average well efficiency of 11%. TW1 has the lowest efficiency due to the poor design of well screen, particularly in terms of screen length and screen diameter, apart from the fact that it might be influenced by poor aquifer potential in that area. Nevertheless, the most efficient tube well does not mean that it is the most productive tube well. This is because well efficiency measures how much losses influence the drawdown of the pumping well. If there is no well loss and the drawdown in the pumping well is only influenced by aquifer loss, the well is the most efficient. However, this ideal condition is difficult to be achieved due to improper well construction and installation, pump factor, improper design of well screen and unsuitable screen length. The productivity of tube well is also influenced by the hydraulic characteristics of the aquifer (transmissivity and storage coefficient), as well as the effect of drawdown on pumping rate (types of aquifer either unconfined or confined aquifer). A detailed result of the step drawdown test is summarized in Table 4.

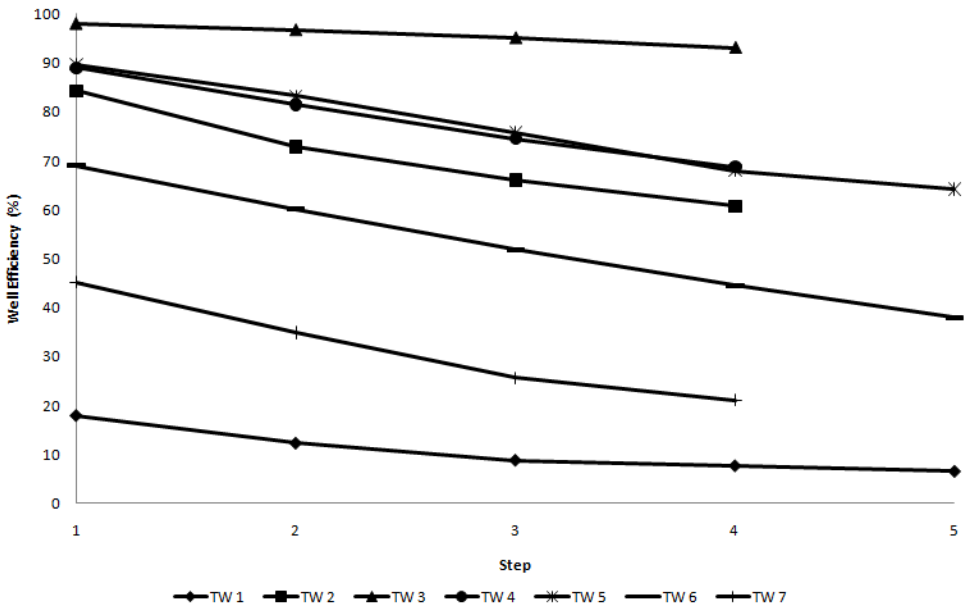



Fig.6: Well efficiency of each tube well in each step

TABLE 4
Step drawdown test result and comparison in each step

| Characteristics | | Tube Well | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | TW 1 | TW 2 | TW 3 | TW 4 | TW 5 | TW 6 | TW 7 |
| Duration/step (hr) | | 1.5 | 2 | 2 | 1.5 | 1.5 | 1.5 | 1 |
| | B | 0.0311 | 0.0198 | 0.2942 | 0.0324 | 0.1259 | 0.4014 | 0.0740 |
| | C | 0.0080 | 0.0001 | 0.0005 | 0.0002 | 0.0016 | 0.0410 | 0.0069 |
| Step 1 | Q | 17.8 | 37.0 | 12.2 | 20.1 | 9.0 | 4.4 | 13.0 |
| | Sw | 3.20 | 0.78 | 3.37 | 0.81 | 1.23 | 2.27 | 2.03 |
| | Q/Sw | 5.563 | 47.436 | 3.620 | 24.815 | 7.317 | 1.938 | 6.404 |
| | Sw/Q | 0.180 | 0.021 | 0.276 | 0.040 | 0.137 | 0.516 | 0.156 |
| | BQ | 0.55358 | 0.7326 | 3.5892 | 0.6512 | 1.1331 | 1.76616 | 0.962 |
| | $CQ^2$ | 2.5347 | 0.1369 | 0.0744 | 0.0808 | 0.1296 | 0.7938 | 1.1661 |
| | E | 17.93 | 84.26 | 97.97 | 88.96 | 89.74 | 68.99 | 45.20 |
| Step 2 | Q | 27.7 | 73.8 | 20.2 | 36.7 | 15.8 | 6.5 | 20 |
| | Sw | 6.67 | 1.84 | 5.63 | 1.44 | 2.47 | 4.59 | 4.41 |
| | Q/Sw | 4.153 | 40.109 | 3.588 | 25.486 | 6.397 | 1.416 | 4.535 |
| | Sw/Q | 0.241 | 0.025 | 0.279 | 0.039 | 0.156 | 0.706 | 0.221 |
| | BQ | 0.8615 | 1.4612 | 5.9428 | 1.1891 | 1.9892 | 2.6091 | 1.4800 |
| | $CQ^2$ | 6.1383 | 0.5446 | 0.2040 | 0.2694 | 0.3994 | 1.7323 | 2.7600 |
| | E | 12.31 | 72.85 | 96.68 | 81.53 | 83.28 | 60.10 | 34.91 |
| Step 3 | Q | 40.1 | 101.8 | 30.7 | 55.3 | 25.2 | 9.1 | 31.0 |
| | Sw | 13.42 | 2.62 | 8.75 | 2.20 | 4.20 | 7.34 | 9.22 |
| | Q/Sw | 2.988 | 38.855 | 3.509 | 25.136 | 6.000 | 1.240 | 3.362 |
| | Sw/Q | 0.335 | 0.026 | 0.285 | 0.040 | 0.167 | 0.807 | 0.297 |
| | BQ | 1.24711 | 2.01564 | 9.0319 | 1.7917 | 3.17268 | 3.65274 | 2.294 |
| | $CQ^2$ | 12.86408 | 1.036324 | 0.4712 | 0.6116 | 1.01606 | 3.39521 | 6.6309 |
| | E3 | 8.84 | 66.04 | 95.04 | 74.55 | 75.74 | 51.83 | 25.70 |
| Step 4 | Q | 46.5 | 127.8 | 43.9 | 73.7 | 37.1 | 12.2 | 40.0 |
| | Sw | 20.18 | 3.34 | 11.77 | 3.38 | 6.43 | 11.54 | 13.79 |
| | Q/Sw | 2.304 | 38.263 | 3.730 | 21.805 | 5.770 | 1.057 | 2.901 |
| | Sw/Q | 0.434 | 0.026 | 0.268 | 0.046 | 0.173 | 0.946 | 0.345 |
| | BQ | 1.4462 | 2.5304 | 12.9154 | 2.3879 | 4.6709 | 4.8971 | 2.9600 |
| | $CQ^2$ | 17.2980 | 1.6333 | 0.9636 | 1.0863 | 2.2023 | 6.1024 | 11.0400 |
| | E | 7.72 | 60.77 | 93.06 | 68.73 | 67.96 | 44.52 | 21.14 |
| Step 5 | Q | 55.3 | NA | NA | NA | 43.9 | 16.0 | NA |
| | Sw | 25.39 | NA | NA | NA | 8.83 | 16.10 | NA |
| | Q/Sw | 2.178 | NA | NA | NA | 4.972 | 0.994 | NA |
| | Sw/Q | 0.459 | NA | NA | NA | 0.201 | 1.006 | NA |
| | BQ | 1.7198 | NA | NA | NA | 5.5270 | 6.4224 | NA |
| | $CQ^2$ | 24.4647 | NA | NA | NA | 3.0835 | 10.4960 | NA |
| | E (%) | 6.57 | NA | NA | NA | 64.19 | 37.96 | NA |
| | ∑(Q/Sw) | 3.4372 | 40.1658 | 3.6118 | 24.3105 | 6.0912 | 1.3290 | 4.3005 |
| | ∑E | 10.67 | 70.98 | 95.69 | 78.44 | 76.18 | 57.68 | 31.74 |

NA: Not available, B: Aquifer loss coefficient ($hrm^{-2}$), C: Well loss coefficient ($hr^2m^{-5}$), Q: Discharge rate ($m^3hr^{-1}$), Sw: Drawdown (m), Q/Sw: Specific drawdown ($hrm^{-2}$), Sw/Q: Specific capacity ($m^2hr^{-1}$), BQ: Aquifer loss (m), BQ2: Well loss (m), ∑ (Q/Sw): Average specific capacity ($m^2hr^{-1}$), ∑E: Average well efficiency (%)

## CONCLUSION AND RECOMMENDATIONS

The performance of seven tube wells located in the Selangor coastal area in alluvium aquifer was evaluated in this study. Step drawdown test was used to assess the tube wells' performance and the data were analysed by using the graphical method and regression technique. Based on the results of this study, it can be concluded that:

1. The tested tube wells have the value of aquifer loss coefficients in the range of 0.0198 $hrm^{-2}$ to 0.4014 $hrm^{-2}$, while well loss coefficients are in the range of 0.0001 $hr^2m^{-5}$ to 0.0410 $hr^2m^{-5}$. All the tube wells have the value of aquifer loss coefficient greater than the value of well loss coefficient.
2. The drawdown in TW1 and TW7 are mainly influenced by well loss component as compared to aquifer loss component. The drawdown in TW2, TW3, TW4, and TW5 are mainly influenced by aquifer loss component compared to well loss component. Aquifer loss component is dominant in TW6 at the lower discharge rate but at the higher discharge rate, the drawdown is influenced by well loss component.
3. TW2 and TW4 are categorized as high productive wells with specific capacity greater than 18 $m^2hr^{-1}$. Meanwhile, TW1, TW3, TW5 and TW7 are categorized as moderate productive tube wells, with the specific capacity ranging between 1.8 and 18 $m^2hr^{-1}$. Only TW6 is categorized as a low productive tube well, with the specific capacity less than 1.8 $m^2hr^{-1}$.
4. The most efficient tube well in Kuala Selangor is TW3, with the average efficiency of 96%, while TW1 in Sabak Bernam is the least efficient tube well, with the average efficiency of 11%.

Continuous monitoring of the performance of these tube wells by the authorities can ensure that they will be in good conditions since all the wells are rarely used. In fact, the maintenance of the wells should be done frequently or annually to prevent tube wells from clogging and other problems which may reduce their efficiency.

## ACKNOWLEDGEMENTS

The authors are indebted to the Department of Mineral and Geoscience Malaysia (Selangor/ Wilayah Persekutuan) for providing the data resources in this paper. The financial support granted to the first author through UPM Graduate Research Fellowship, GRF, is also acknowledged.

## REFERENCES

Bierschenk, W. H. (1964). Determining well efficiency by multiple step-drawdown tests. *International Association of Scientific Hydrology*, *64*, 493-507.

Heng, C. L. (2004). *Groundwater Utilization and Management in Malaysia*. Paper presented at the meeting of the Coordinating Committee for Geoscience Programmes in East and Southeast Asia, Tsukuba. November, 2004.

Ismail, T., Anuar, S., & Saim, S. (2011). Groundwater contamination in North Kelantan: how serious? In T.F. Ng (Ed.), *Geoscientis and Ethics for a Sustainable Society*. Proceeding National Geoscience Conference, Johor Bahru, Johor, June 11-12, 2011. Geological Society Malaysia: Kuala Lumpur, 2011.

Jacob, C. E. (1947). Drawdown test to determine effective radius of artesian well. *Transactions of the American Geophysical Union*, *112*, 1047-1070.

Kawecki, M. W. (1995). Meaningful interpretation of step drawdown test. *Ground Water*, *33*(1), 23-32.

Krusseman, & de Rider (1994). *Analysis and Evaluation of Pumping Test Data* (2nd Ed.). Netherlands: International Institute for Land Reclamation and Improvement (ILRI).

Mahajan, G. (1989). *Evaluation and Development of Ground Water*. New Delhi: Ashish Publishing House.

Mishra, A., & Sahay, R. R. (2011). Well parameter's estimation using traditional and non-traditional methods. *International Journal of Earth Sciences and Engineering*, *4*(6), 235-238.

Misstear, B. D. R., & Beeson, S. (2000). Using operational data to estimate the reliable yields of water-supply wells. *Hydrogeology Journal*, *8*, 177-187.

Mohammed, T. A., & Huat, B. B. K. (2004). *Groundwater Engineering and Geotechnique*. Serdang: UPM Press.

Rahman, M. M., & Dhar, S. C. (1997). Deterioration of the yielding capacity of water production wells in Dhaka. *Journal of Civil Engineering, the Institution of Engineers, Bangladesh*, *25*(2), 157-169.

Rorabaugh, M. I. (1953). Graphical and theoretical analysis of step-drawdown test of artesian well. *Proceeding of American Society of Civil Engineers (ASCI)*, *79*, 1-14.

Samsudin, A. R., Haryono, A., Hamzah, U., & Rafek, A. G. (2008). Salinity mapping of coastal groundwater aquifers using hydrogeochemical and geophysical methods: a case study from north Kelantan, Malaysia. *Environ. Geol.*, *55*, 1737-1743.

Selangor Map. (2012). Official Website of Government State of Selangor. Retrieved January 23, 2012, from www.selangor.gov.my

Şen, Z. (1995). *Applied Hydrogeology for Scientists and Engineers:* United States of America. CRC Press.

Sheahan, N. T. (1971). Types-curve solution of step-drawdown test. *Ground Water*, *9*(1), 25-29.

Shekhar, S. (2006). An approach to interpretation of step drawdown tests. *Hydrogeology Journal*, *14*, 1018-1027.

Todd, D. K., & Mays, L. W. (2005). *Groundwater Hydrogeology* (3rd ed.). United States of America: John Wiley & Sons.

Walton, W. C. (1962). *Selected analytical methods for well and aquifer evaluation*, Bulletin 49. Urbana: Illinois State Water Survey.

Pertanika J. Sci. & Technol. 22 (1): 239 - 254 (2014)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Optimization of Modified Fenton ($FeGAC/H_2O_2$) Pretreatment of Antibiotics

**Augustine Chioma Affam*, Malay Chaudhuri and Shamsul Rahman Mohammed Kutty**

*Department of Civil Engineering, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia*

## ABSTRACT

The study examined modified Fenton ($FeGAC/H_2O_2$) pretreatment of the antibiotics amoxicillin and cloxacillin in aqueous solution for biological treatment. The treatment was optimized by the response surface methodology (RSM). The optimum operating conditions at pH3 were $H_2O_2$/COD molar ratio 2.0, FeGAC dose 3.5 g/L and reaction time 90 min for 87.53% removal of COD, 78.01% removal of TOC, and 98.24% removal of $NH_3$-N. Biodegradability ($BOD_5$/COD ratio) improved from zero to 0.36, indicating the effluent was amenable to biological treatment. Meanwhile, FTIR spectra indicated degradation of the antibiotics. Compared with Fenton or photo-Fenton, modified Fenton ($FeGAC/H_2O_2$) was more effective in the pre-treatment of the antibiotics amoxicillin and cloxacillin in aqueous solution for biological treatment.

*Keywords:* Modified Fenton (FeGAC/H2O2), antibiotic aqueous solution; amoxicillin, cloxacillin, response surface methodology (RSM)

## INTRODUCTION

The ultimate fate of pharmaceutical compounds after their intended use was not monitored until their presence in the aquatic system was detected. Consequently, these compounds have attracted increased attention as potential water pollutants (Jones *et al*., 2007). Over 60 pharmaceutical compounds from a variety of therapeutic groups (including but not limited to, analgesics, antibiotics, anti-epileptics, β-blockers, $\beta_2$-sympathomimetics, blood lipid regulators and X-ray contrast media) have been found in the aquatic matrix (Jones *et al*., 2001). These compounds have also been reportedly found in surface water, groundwater, sewage effluent, and in drinking water (Stackelberg *et al*., 2004). The concern

*Article history:*
Received: 20 February 2012
Accepted: 9 December 2012

*E-mail addresses:*
affamskii@yahoo.com (Augustine Chioma Affam),
m_chaudhuri@petronas.com.my (Malay Chaudhuri),
shamsulrahman@petronas.com.my
(Shamsul Rahman Mohammed Kutty)
*Corresponding Author

ISSN: 0128-7680 

about the presence of antibiotics in low concentrations in the aquatic environment is the development of antibiotic resistant bacteria (Walter & Vennes, 1985). Degradation of these antibiotics is a current challenge, especially as the conventional wastewater treatment plants (STPs) are not designed to remove them (Gulkowska *et al.,* 2008). Amoxicillin and cloxacillin antibiotics are broad beta-lactam antibiotic that belongs to the class of penicillin. They are used in veterinary and human medicine as initial medicine in the treatment of majority of infectious diseases (Anacona & Figueroa, 1999).

Advanced oxidation process (Fenton or photo-Fenton) pre-treatment of antibiotics in aqueous solution enhances biodegradability and produces a new effluent that is amenable to biological treatment.

Fenton and photo-Fenton pretreatments of the antibiotics amoxicillin, ampicillin and cloxacillin in aqueous solution have been reported (Elmolla & Chaudhuri, 2009a, 2009b) as advanced oxidation strategies for enhancement of antibiotics aqueous solution.

Activated carbon (AC) is known to decompose hydrogen peroxide ($H_2O_2$). Presumably, the process involves the exchange of a surface hydroxyl group with hydrogen peroxide anion (Eq. 1), according to Bansal *et al.* (1998, as cited in Khalil *et al.,* 2001). The formed surface peroxide is regarded as having an increased oxidation potential which enables the decomposition of another hydrogen peroxide molecule with release of oxygen and regeneration of the AC surface (Eq. 2).

$$AC{-}OH + H^+ OOH^- \longrightarrow AC{-}OOH + H_2O \quad (1)$$

$$AC{-}OOH + H_2O_2 \longrightarrow AC{-}OH + H_2O + O_2 \quad (2)$$

Beside this decomposition reaction, hydrogen peroxide can obviously be activated on the AC surface involving the formation of $OH^{\bullet}$. AC is considered to function as an electron-transfer catalyst similar to the Haber–Weiss mechanism known from the Fenton reaction, with AC and $AC^+$, as the oxidized and reduced catalyst states (Eqs. 3-4) (Kimura & Miyamoto, 1994).

$$AC + H_2O_2 \longrightarrow AC^+ + HO^- + OH^{\bullet} \quad (3)$$

$$AC^+ + H_2O_2 \longrightarrow AC + HO_2^{\bullet} + H^+ \quad (4)$$

In the FeGAC/$H_2O_2$ process, activated carbon acts as an adsorbent, and yet at the same time, the presence of ferrous iron enables the adsorbent (FeGAC) to interact as a catalyst to degrade adsorbed or dissolved organic matter. GAC is also involved in the catalytic activity due the graphitic structure and functional groups on their surface (Bach & Semiat, 2010). Therefore, FeGAC has high adsorption capacity due to GAC surface and coating of iron oxide. In addition, the oxidation strength of $H_2O_2$ is enhanced by the catalytic properties of FeGAC in the removal of antibiotics from the aqueous solution. There is no detailed study which has reported contaminants removal mechanism in the FeGAC/$H_2O_2$ process. However, Fenton mechanism combined with adsorption process is basic for the FeGAC/$H_2O_2$ process to proceed. Oxidation with Fenton's reagent is based on hydroxyl radicals produced by catalytic decomposition of hydrogen peroxide ($H_2O_2$) in reaction to ferrous ions ($Fe^{2+}$) (Chamarro *et al*., 2001). In the photo-Fenton process, additional reactions occur in the presence of light that

produce hydroxyl radicals or increase the production rate of hydroxyl radicals (Pignatello *et al*., 1999) and thus, increasing the efficiency of the process. Adsorption studies have been reported using the $FeGAC/H_2O_2$ process (Fan *et al.,* 2007; C-C Chen *et al.,* 2011).

Modified Fenton process could be in terms of replacement of the standard catalysts such as ferrous and/or ferric ions with novel adsorbents, and an increase in process temperature up to 80–120°C, i.e. the so-called catalytic wet hydrogen peroxide oxidation process, application of novel, highly stable and active heterogeneous catalysts ($Fe_2O_3$ or Fe supported into SBA-15 mesostructured material, etc., and the application of iron chelates (typically organic acids) can promote chemical oxidation of contaminants by formation of iron–chelate complexes among others. The primary advantages of some modifications include effective generation of hydroxyl radicals at near-neutral pH and reduction of iron sludge production in other cases (Goi *et al*., 2008).

There is insufficient literature to buttress the determination of the equilibrium proportions of ferric species as a function of pH and concentration or less presence of iron-related ionic species with respect to ratio, dosage and time in the $FeGAC/H_2O_2$ process. The species of impregnated iron were affected by impregnation methods and synthesizing conditions (Jang *et al*., 2008; Thirunavukkarasu *et al.,* 2003) and that at different temperatures, the impregnated iron oxide manifested a more crystalline form (Jang *et al*., 2008). In another study, the initial amount of iron oxides on FeGAC surface was approximately 38 mg Fe/g GAC, and after the treatment, the loss of iron oxide became less than 5% (Chen *et al.,* 2011). Acid black decolorization rate of 8.71 mg/min by the $FeGAC/H_2O_2$ process has been reported by Fan *et al.* (2006).

Generally, it is known where ferrous ions and hydrogen peroxide are used to generate OH, and production of iron sludge is inevitable. Nonetheless, this may cause additional cost with regard to handling and disposal. To circumvent the problem of iron sludge production in the processes reported, a modified Fenton process − Fe-granular activated carbon/hydrogen peroxide ($FeGAC/H_2O_2$) process − was proposed (Fan *et al*., 2007). This is possible by coating iron oxides on the surface of granular activated carbon (GAC) and could minimize the production of sludge. This has been applied in the decolourization of Acid Black 24 (Fan *et al.,* 2006) removal of humic substances, namely, humic and fulvic acids in municipal landfill leachate (Fan *et al.,* 2007) and degradation of crystal violet (Chen *et al.,* 2011). However, the application of the treatment process of antibiotics in aqueous solution has not been reported.

The present study examined the application of the $FeGAC/H_2O_2$ process in the pretreatment of the antibiotics amoxicillin and cloxacillin in aqueous solution for biological treatment. The treatment was optimized by response surface methodology for the removal of chemical oxygen demand (COD), total organic carbon (TOC), and ammonia-nitrogen ($NH_3$-N).

## MATERIALS AND METHODS

### *Fe-Granular Activated Carbon (FeGAC)*

Granular activated carbon (GAC) was obtained from the Calgon Corporation, Pittsburgh, PA, and was ground to a size of 425 µm. FeGAC was prepared by mixing GAC in a solution of ferrous sulfate ($FeSO_4 \cdot 7H_2O$) for 24 h at 40 mg ferrous sulphate/g GAC, followed by drying at 105ºC. The resulting FeGAC was washed several times with distilled water, dried at 105ºC and stored in a stopper glass bottle.

*Antibiotic Aqueous Solution*

The aqueous solution of the antibiotics amoxicillin (AMX) and cloxacillin (CLX) was prepared weekly by dissolving 150 mg each of the antibiotics in 1000 mL distilled water and stored at 4ºC. AMX and CLX were obtained from a commercial source (Farmaniaga Company, Malaysia). The characteristics of the antibiotic aqueous solution are as follows: COD 390 mg/L, TOC 168.8 mg/L, $NH_3$-N 20.6 mg/L and $BOD_5$/COD ratio zero.

TABLE 1
Characteristics of antibiotics solution

| Parameter | Range |
|---|---|
| COD | 390.0 ± 4.0 |
| TOC | 168.8 ± 2.5 |
| $NH_3$-N | 20.6 ± 1.2 |

*Analytical Methods*

Chemical oxygen demand (COD) was measured according to Method 5220D (closed reflux, colorimetric method) of the Standard Methods (APHA, 2005). If the sample contained hydrogen peroxide ($H_2O_2$), to reduce interference in COD determination, pH was increased to above 10 to decompose $H_2O_2$ to oxygen and water (Talinli & Anderson, 1992; Kang *et al*., 1999). A TOC analyzer (Model 1010; O & I Analytical) was used in determining total organic carbon (TOC). Meanwhile, $NH_3$-N was measured according to Method 8038 (Nessler Method) of the Water Analysis Handbook (Hach, 2002). Five-day biochemical oxygen demand ($BOD_5$) was measured according to Method 5210B (seeding procedure) of the Standard Methods (APHA, 2005). The bacterial seed for $BOD_5$ test was obtained from a municipal wastewater treatment plant. DO was measured using YSI 5000 dissolved oxygen meter. FTIR spectra of the untreated and treated antibiotic aqueous solution were taken by using *Shimadzu FTIR-8400S.*

*FeGAC/$H_2O_2$ Treatment*

Batch FeGAC/$H_2O_2$ treatment was performed in 250-mL conical flasks with 200 mL of the antibiotic aqueous solution adjusted to pH3, which is the optimum pH for Fenton treatment of antibiotic aqueous solution (Elmolla & Chaudhuri, 2009a). $H_2O_2$ and FeGAC were added simultaneously according to the selected $H_2O_2$/COD molar ratio and FeGAC dose and the flasks were placed on an orbital shaker. At the selected reaction time for each run 1 to 20 as obtained from the central composite design (CCD) in Table 3, a flask was removed from the orbital shaker and an aliquot of the supernatant was filtered through 0.45 µm membrane filter for measurement of COD, TOC and $NH_3$-N, and filtered through 0.20 µm membrane filter for FTIR spectra.

TABLE 2
The actual and coded values for the independent variables of the CCD design

| Independent variable values | | Coded | | |
|---|---|---|---|---|
| | | −1 | 0 | +1 |
| | Units | Actual value | | |
| A: $H_2O_2$/COD | molar ratio | 1.0 | 2.0 | 3.0 |
| B: FeGAC | g/200ml | 0.4 | 0.7 | 1.0 |
| C: Time | min | 60 | 90 | 120 |

TABLE 3
Experimental design and observed removal

| Experimental Design | | | Removal (%) | | |
|---|---|---|---|---|---|
| A: $H_2O_2$/COD (Molar ratio) | B: FeGAC (g/200 mL) | C: Reaction Time (min) | COD | TOC | $NH_3$-N |
| 2.00 (0.0) | 1.20 (1.68) | 90.00 (0.0) | 67.0 | 63.45 | 89.0 |
| 3.68 (1.68) | 0.70 (0.0) | 90.00 (0.0) | 82.1 | 43.55 | 100.0 |
| 3.00 (1.0) | 0.40 (-1.0) | 120.00 (1.0) | 77.0 | 29.32 | 96.0 |
| 2.00 (0.0) | 0.70 (0.0) | 90.00 (0.0) | 85.1 | 88.75 | 97.0 |
| 2.00 (0.0) | 0.70 (0.0) | 39.55 (-1.68) | 83.0 | 55.81 | 96.0 |
| 0.32 (-1.68) | 0.70 (0.0) | 90.00 (0.0) | 98.2 | 74.42 | 97.0 |
| 2.00 (0.0) | 0.20 (-1.68) | 90.00 (0.0) | 63.3 | 39.31 | 96.6 |
| 3.00 (1.0) | 0.40 (-1.0) | 60.00 (-1.0) | 85.8 | 56.95 | 100.0 |
| 1.00 (-1.0) | 1.00 (1.0) | 60.00 (-1.0) | 67.0 | 53.77 | 93.0 |
| 1.00 (-1.0) | 0.40 (-1.0) | 120.00 (1.0) | 75.0 | 69.65 | 97.0 |
| 2.00 (0.0) | 0.70 (0.0) | 90.00 (0.0) | 100.0 | 76.82 | 100.0 |
| 2.00 (0.0) | 0.70 (0.0) | 90.00 (0.0) | 89.0 | 74.44 | 97.0 |
| 2.00 (0.0) | 0.70 (0.0) | 90.00 (0.0) | 83.0 | 66.45 | 100.0 |
| 2.00 (1.0) | 0.70 (0.0) | 140.45 (1.68) | 97.4 | 68.85 | 98.0 |
| 3.00 (1.0) | 1.00 (1.0) | 60.00 (-1.0) | 84.0 | 47.35 | 96.0 |
| 1.00 (-1.0) | 0.40 (-1.0) | 60.00 (-1.0) | 86.7 | 57.99 | 97.0 |
| 3.00 (1.0) | 1.00 (1.0) | 120.00 (1.0) | 93.8 | 61.28 | 96.0 |
| 2.00 (0.0) | 0.70 (0.0) | 90.00 (0.0) | 83.4 | 68.53 | 97.0 |
| 1.00 (-1.0) | 1.00 (1.0) | 120.00 (1.0) | 96.0 | 66.62 | 94.0 |
| 2.00 (0.0) | 0.70 (0.0) | 90.00 (0.0) | 85.0 | 92.89 | 98.5 |

*Optimization and Response Surface Modelling*

Design expert software (version 6.0.7) was used for the statistical design of experiment and data analysis. Central composite design (CCD) of the response surface methodology (RSM) (Khuri & Cornell, 1996; Bezerra *et al*., 2008) was used to optimize the operating conditions (variables) of the FeGAC/$H_2O_2$ treatment. The variables were simultaneously changed in a central composite circumscribed design. The coded values of the variables were $H_2O_2$/COD molar ratio (A), FeGAC dose (g/200 mL) (B), and reaction time (min) (C), and they were varied over three levels: 1 (low), 0 (central) and +1 (high). The variables $H_2O_2$/COD molar ratio,

FeGAC dose and reaction time were studied in the range of 1.0-3.0, 0.4-1.0 g/200 mL and 60-120 min, respectively (Table 2). The data range for the independent variables was obtained from previous studies (Elmolla & Chaudhuri, 2009a, 2009b) and hence, no preliminary study was conducted. The significance of the chosen variables includes that $H_2O_2$/COD molar ratio relates to the approximate $H_2O_2$ concentration required based on the stoichiometric ratio with respect to COD of the antibiotics aqueous solution and was calculated by assuming that a complete oxidation of COD occurred. $H_2O_2$ ensures the oxidation of the antibiotic in the aqueous solution. The amount of $H_2O_2$ influences the modified Fenton process because excessive $H_2O_2$ will pose an inhibitory effect in the degradation of antibiotics. This could be due to $OH^{\bullet}$ scavenging and the formation of another radical ($HO_2^{\bullet}$), which has an oxidation potential considerably smaller than $OH^{\bullet}$, as in Equation (5) (Arslan-Alaton *et al*., 2009).

$$OH^{\bullet} + H_2O_2 \longrightarrow HO_2^{\bullet} + H_2O_2 \quad (5)$$

FeGAC is important for the catalysis of the process in the presence of $H_2O_2$. In addition, FeGAC serves as an adsorbent with GAC added to increase the available catalytic sites. The reaction time is a variable to observe and determine the best operating time for an optimum degradation of the antibiotics. The optimum $H_2O_2$/COD molar ratio 3.0 and reaction time 60 min for the Fenton treatment of antibiotic aqueous solution (Elmolla & Chaudhuri, 2009a) and FeGAC dose 0.8 g/200 mL for FeGAC/$H_2O_2$ treatment of a dye wastewater (Fan *et al.,* 2006) have been reported. CCD generally consists of a 2k factorial with nF factorial runs, 2k axial or star runs, and nC (i.e. centre runs). In this study, a total of 20 experiments were performed according to $2^k + 2k + 6$, where k is the number of independent variables (k=3). Fourteen experiments were enhanced with six replications to assess the pure error. The chosen response parameters of the FeGAC/$H_2O_2$ treatment were removal of COD, TOC and $NH_3$-N. Regression analyses, response surface plots, analysis of variance (ANOVA), perturbation plots and normal probability plot of the studentized residuals were carried out using the design expert software. The optimum operating conditions were identified from the response surface plots and the response equations simultaneously.

The following quadratic model Equation (6) was used to estimate the optimal values:

$$Y = \beta_o + \sum_{j=1}^{k} \beta_j X_j + \sum \beta_{jj} X^2{}_j + \sum_i \sum_{<j=2}^{k} \beta_{ij} X_i X_j + \varepsilon \quad (6)$$

where Y is the response; $X_i$ and $X_j$ are the variables; $\beta_0$ is a constant coefficient; $\beta_j$, $\beta_{jj}$, and $\beta_{ij}$ are the interaction coefficients of linear, quadratic and second-order terms, respectively; k is the number of studied factors; and $\varepsilon$ is the error. The quality of the fit was expressed by the coefficient of determination ($R^2$). The main indicators demonstrating the significance and adequacy of the model used include the F-value (Fisher variation ratio), probability value (Prob>F), and the Adequate Precision (Arslan-Alaton *et al*., 2009).

## RESULTS AND DISCUSSION

### *Statistical Analysis*

The results obtained were analyzed by ANOVA to assess the "goodness of fit". The models for COD, TOC, and $NH_3$-N removal ($Y_1$, $Y_2$ and $Y_3$) were significant by the *F*-test at 95% confidence level if Prob>*F*<0.05. The following fitted regression model (equations in terms of coded values) was obtained to quantitatively investigate the effects of $H_2O_2$/COD molar ratio (A), FeGAC dose (B), and reaction time (C) on COD, TOC and $NH_3$-N removal.

COD removal

$$Y_1 = 87.53 - 0.82A - 7.58B^2 + 7.41BC \quad (7)$$

TOC removal

$$Y_2 = 78.01 - 7.69A + 6.92A^2 - 9.61B^2 + 5.74C^2 \quad (8)$$

$NH_3$-N removal

$$Y_3 = 98.24 + 0.88A - 1.74B \quad (9)$$

In Equations 7, 8 and 9, the values of the sum of a constant ($\beta_0$), 87.53, 78.01 and 98.24 represent the percentage removal of COD, TOC and $NH_3$-N, respectively. The positive sign indicates that the variable is directly proportional to the response (COD, TOC and $NH_3$-N removal) and the negative sign indicates that the variable is inversely proportional to the response. Table 3 shows the experimental design, real and codified values (in parentheses) of the variables and response parameters (observed removal of COD, TOC and $NH_3$-N). Tables 4a-4c show the ANOVA for the response surface quadratic model. In this study, A, $B^2$ and BC; A, $A^2$, $B^2$ and $C^2$; and A and B; were found to be significant model terms for COD, TOC and $NH_3$-N removal, respectively, while others were discarded from the study to ensure an improved model. Adequate precision (AP) compares the range of the predicted values at the design points to the average prediction error. Ratios greater than 4 indicate adequate model discrimination and can be used to navigate the design space defined by CCD (Ghafari *et al.,* 2009). AP for all the responses was greater than 4. The probability of the lack of fit (PLOF) describes the variation of the data around the fitted model. This is significant when PLOF<0.05 for the overall model. PLOF for all the responses were <0.05. The $R^2$ values shown in Table 4 for the response parameters COD ($R^2 = 0.7826$), TOC ($R^2 = 0.8034$) and $NH_3$-N ($R^2 = 0.8635$) indicate that the model is good enough for the quadratic fits to navigate the design space defined by CCD (Ghafari *et al.,* 2009).

### *Process Analysis*

Fig.1, Fig.2 and Fig.3 show the response surface plots for COD, TOC and $NH_3$-N removal. The maximum COD, TOC and $NH_3$-N removals were 87.53, 78.01 and 98.24%, respectively, under $H_2O_2$/COD molar ratio 2.0, FeGAC dose 0.70 g/200mL (3.5 g/L), pH 3 and reaction time 90 min. In the Fenton treatment of the antibiotics amoxicillin, ampicillin and cloxacillin in the aqueous solution, Elmolla and Chaudhuri (2009a) reported the maximum COD and DOC removal of 81.4 and 54.3%, respectively, under the $H_2O_2$/COD molar ratio 3.0, $H_2O_2$/

TABLE 4
The ANOVA for analysis and adequacy of the quadratic model

COD removal (%)

| Source | Sum of Squares | DF | Mean Square | F Value | Prob > F |
|---|---|---|---|---|---|
| Model | 1610.11 | 9 | 178.90 | 4.00 | <0.0001 |
| A | 9.11 | 1 | 9.11 | 18.47 | 0.0016 |
| $B^2$ | 826.09 | 1 | 826.09 | 9.83 | 0.0106 |
| BC | 439.56 | 1 | 439.56 | | |
| Residual | 447.17 | 10 | 44.72 | | |
| Lack of Fit | 239.64 | 5 | 47.93 | 1.15 | <0.0001 |
| Pure Error | 207.53 | 5 | 41.51 | | |

Std. Dev. = 6.69 PRESS = 2222.2 $R^2$ = 0.7826 Adj. $R^2$ = 0.5870 Adequate Precision = 6.985

TOC removal (%)

| Source | Sum of Squares | DF | Mean Square | F Value | Prob > F |
|---|---|---|---|---|---|
| Model | 3746.01 | 9 | 416.22 | 4.54 | <0.0001 |
| A | 807.87 | 1 | 807.87 | 8.81 | 0.0141 |
| $A^2$ | 683.08 | 1 | 683.08 | 7.45 | 0.0212 |
| $B^2$ | 1330.31 | 1 | 13330.31 | 14.51 | 0.0034 |
| $C^2$ | 467.99 | 1 | 467.99 | 5.11 | 0.0474 |
| Residual | 916.59 | 10 | 91.66 | | |
| Lack of Fit | 342.17 | 5 | 68.43 | 0.60 | <0.0001 |
| Pure Error | 574.42 | 5 | 114.88 | | |

Std. Dev. = 9.57 PRESS = 3751.65 $R^2$ = 0.8034 Adj. $R^2$ = 0.6265 Adequate Precision = 6.726

$NH_3$-N removal (%)

| Source | Sum of Squares | DF | Mean Square | F Value | Prob > F |
|---|---|---|---|---|---|
| Model | 113.29 | 9 | 12.59 | 7.03 | <0.0001 |
| A | 10.62 | 1 | 10.62 | 5.93 | 0.0351 |
| B | 41.32 | 1 | 41.32 | 23.06 | 0.0007 |
| Residual | 17.92 | 10 | 1.79 | | |
| Lack of Fit | 7.04 | 5 | 1.41 | 0.65 | <0.0001 |
| Pure Error | 10.88 | 5 | 2.17 | | |

Std. Dev. = 1.34 PRESS = 70.87 $R^2$ = 0.8635 Adj. $R^2$ = 0.7406 Adequate Precision = 10.661

$Fe^{2+}$ molar ratio 10.0, pH3 and reaction time of 60 min. In the photo-Fenton treatment of the antibiotics in aqueous solution, Elmolla and Chaudhuri (2009b) reported the maximum COD and DOC removal of 80.8 and 58.4%, respectively, under the $H_2O_2$/COD molar ratio 1.5, $H_2O_2$/$Fe^{2+}$ molar ratio 20, pH3 and irradiation time of 50 min. Thus, the modified Fenton (FeGAC/$H_2O_2$) was more effective than Fenton or photo-Fenton in the pretreatment of the antibiotics in the aqueous solution. Fig.4, Fig.5 and Fig.6 show the perturbation plot for COD, TOC and $NH_3$-N removal, respectively. The adequacies of the models were also evaluated by the residuals, i.e. Normal probability plot of the studentized residuals (Fig.4 - Fig.6) and indicate

that there is no abnormality in the model as all the data were found around the line of the 'best fit'. Fig.4 - Fig.6 show the perturbation plots for COD, $NH_3$-N and TOC removals. The plots show how response (COD, TOC and $NH_3$-N removal) changes as each variable (A: $H_2O_2$/COD molar ratio, B: FeGAC dose, g/mL, and C: reaction time) moves from the chosen reference point, i.e. 0 (central) level, with other variables held constant. Thus, a comparative effect of all the independent variables on the removals was observed. A sharp curvature is seen in Figure 4, indicating that the COD removal was the most sensitive to variable B: FeGAC dose. The comparatively flat curvature from variable A: $H_2O_2$/COD molar ratio and C: reaction time shows less sensitivity on the COD removal. In Fig.5, all the variables (A: $H_2O_2$/COD molar ratio, B: FeGAC dose, g/mL and C: reaction time) had a moderately sharp curvature which showed that the TOC removal was sensitive to the three variables. Fig.6 shows that $NH_3$-N removal was most sensitive to variable B: FeGAC dose compared to A: $H_2O_2$/COD molar ratio, and C: reaction time during the treatment process. The normal probability plots of the studentized residuals for the COD, TOC and $NH_3$-N removal are shown in Fig.7- Fig.9. A normal probability plot graphically shows whether or not the residuals follow a normal distribution, hence, the points follow a straight line. The data showed a good plot although some scattering was also observed.

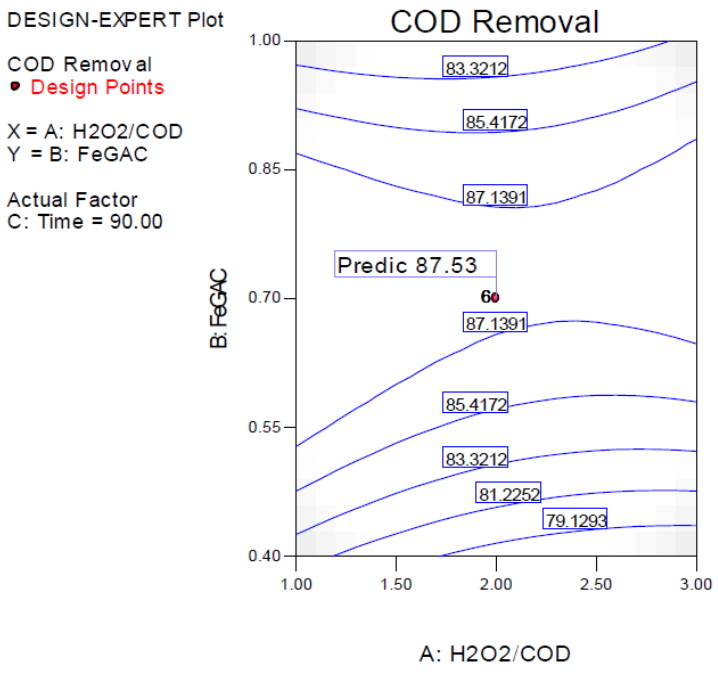



Fig.1: Response surface plot for COD removal

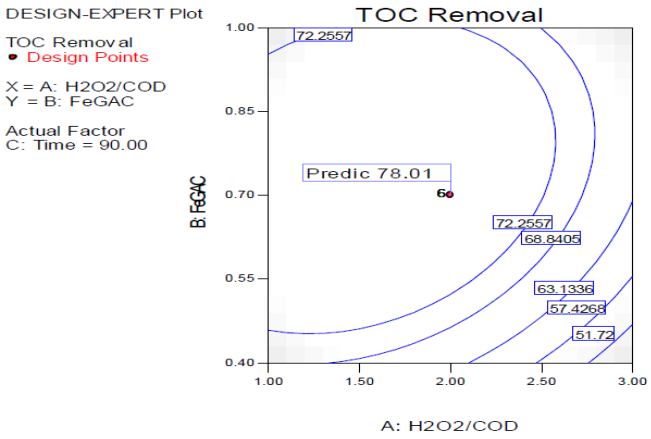



Fig.2: Response surface plot for TOC removal

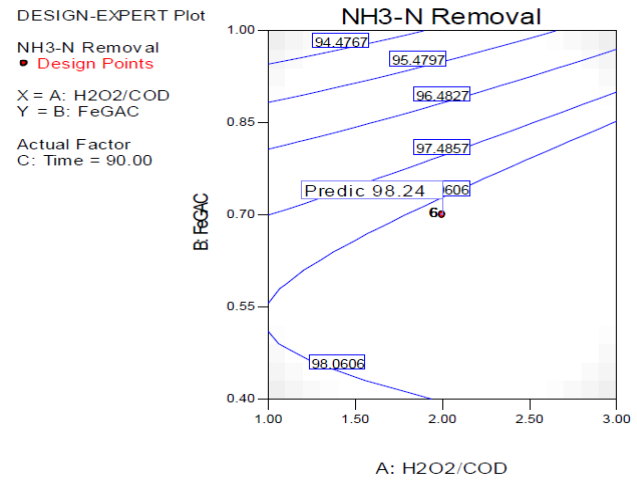



Fig.3: Response surface plot for $NH_3$-N removal

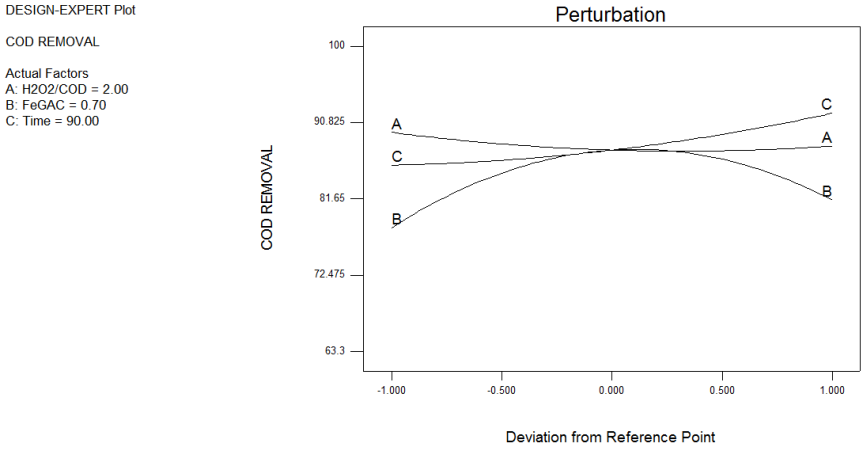



Fig.4: Perturbation plot for COD removal

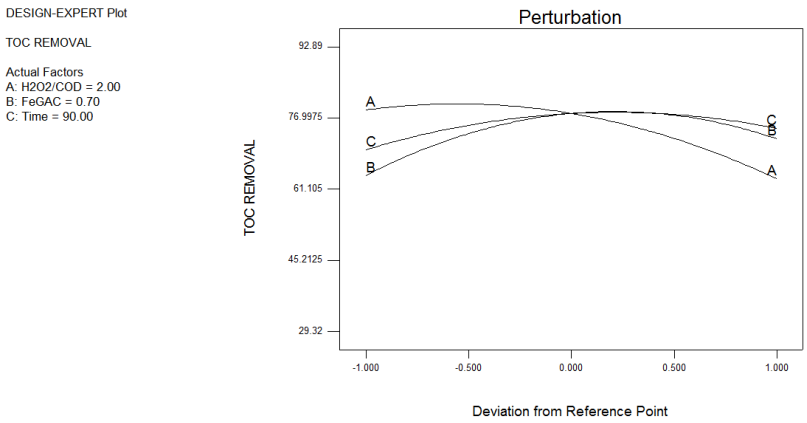



Fig.5: Perturbation plot for TOC removal

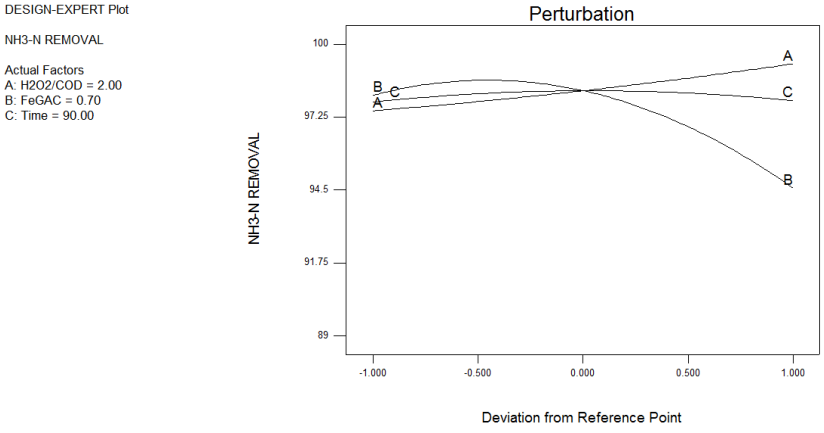



Fig.6: Perturbation plot for $NH_3$-N removal

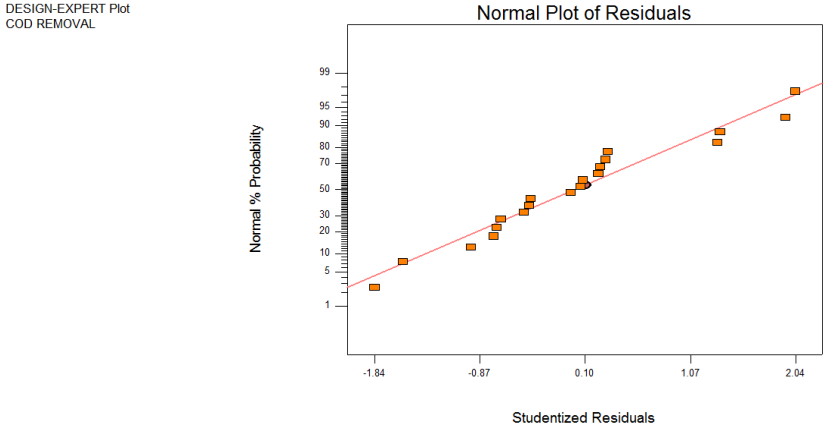



Fig.7: Normal probability plot of the studentized residuals for COD removal

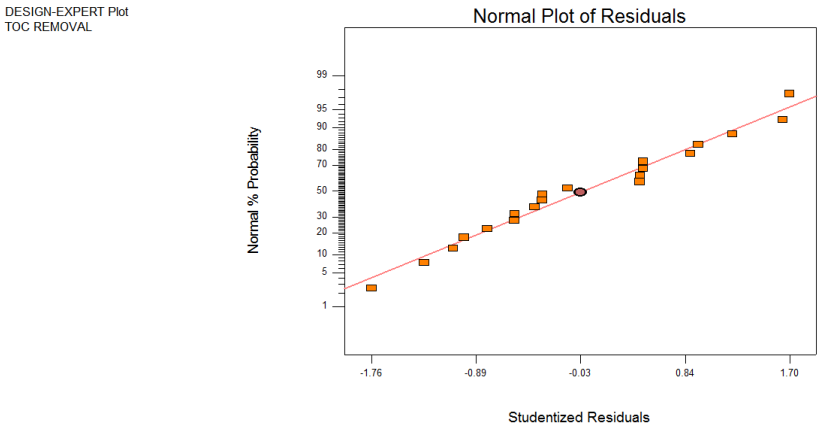



Fig.8: Normal probability plot of the studentized residuals for TOC removal

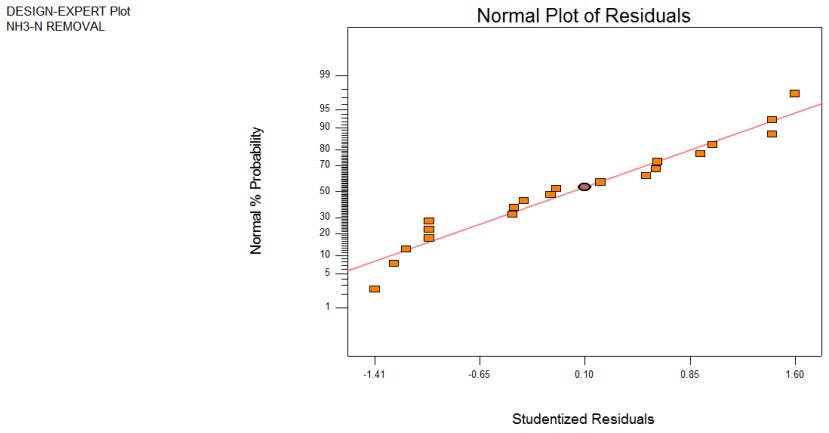



Fig.9: Normal probability plot of the studentized residuals for $NH_3$-N removal

## *Confirmatory Experiments*

In order to validate optimization by RSM, three confirmatory experiments were conducted under the optimum operating conditions to verify the model prediction. As shown in Table 5, model prediction and experimental removal efficiency were in close agreement with less than 0.40% error.

TABLE 5
Model prediction and experimental removal efficiency

| Parameter | Model prediction | Experimental removal | Error (%) |
|---|---|---|---|
| COD removal (%) | 87.53 | 85.58,88.9, 89.20 (87.89) | 0.36 |
| TOC removal (%) | 78.01 | 78.08,78.0,78.03 (78.04) | 0.03 |
| $NH_3$-N removal (%) | 98.24 | 98.89,98.4,98.58 (98.62) | 0.38 |

*Antibiotic Degradation and Biodegradability*

FTIR spectroscopy provides information on the chemical structure of organic bonds in a sample for complete structural analogy. Fig.10 and Fig.11 are important to present the analysis for the residual organic matter by the FTIR spectroscopy, and to ascertain the extent of antibiotic degradation of the organic bonds in the untreated and treated antibiotics under optimum operating conditions. Generally, when infrared light interacts or passes through an aqueous solution, the chemical bonds present in that solution stretches, contracts and/or bends. As a result, chemical functional group tends to absorb infrared radiation in a specific wavelength range, regardless of the structure of the rest of the molecule. The infrared spectra of organic bonds consisting of beta-lactam carbonyl group and aromatic ring could shift to specific bands due to degradation (Anacona & Figueroa, 1999; Rozas *et al.,* 2010). The IR features of penicillins are reported in the range of 1550 $cm^{-1}$ and 1880 $cm^{-1}$, but these usually occur at about 1770 $cm^{-1}$ as a result of the vibration of strong bands of the beta-lactam carbonyl group (Kukpa, 1997). The main difference between the beta-lactam carbonyl group and aromatic ring is located at 1596.95 $cm^{-1}$, a signal that can be attributed to conjugated carbonyl groups which only become visible in the intermediates formed by the modified Fenton reaction (Rozas *et al.,* 2010). The characteristic band which occurred at 1637.45 $cm^{-1}$ for the untreated antibiotic aqueous solution (Fig.10) shifted and modified to 1596.95 $cm^{-1}$ (see Fig.11). This can be further attributed to the pairing up of the organic group degradation intermediates (Rozas *et al.,* 2010). Meanwhile, Arjunan *et al*. (2012) reported that N–H stretching vibrations were observed between 3300-3500 $cm^{-1}$ and subsequently led to deformation vibration in the short range of 1650-1580 $cm^{-1}$ of the spectrum. In a study on amoxicillin FTIR/ATR, a stretch resulting from ν(N10H) and ν(O13H) occurred at 3451 and 3529 $cm^{-1}$, respectively (Bebu *et al.,* 2011). The band between 3361 and 3449 $cm^{-1}$ in cloxalillin was due to ν(O-H) stretching (Adedibu *et al*., 2011). The observed shift and/or modification of the band in cloxacillin occurred through the oxygen atom of the lactam carbonyl group (Anacona & Carman, 2005).

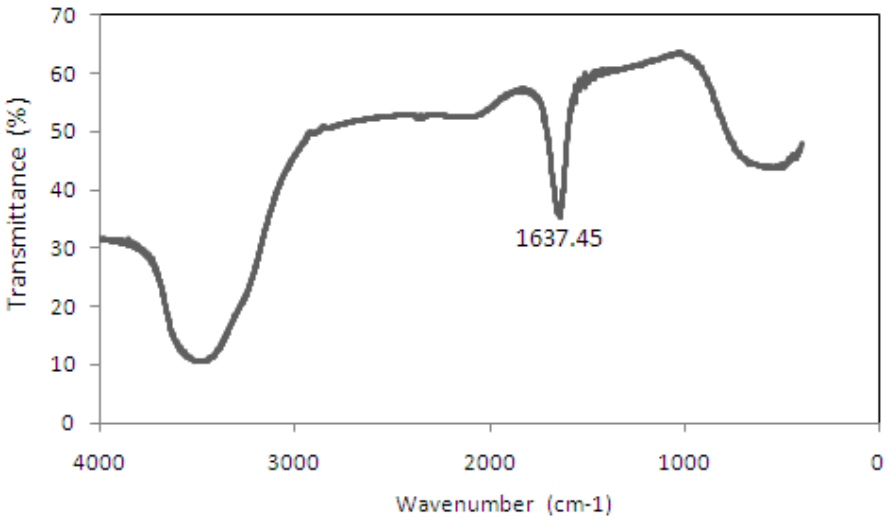



Fig.10: FTIR spectra of untreated antibiotic aqueous solution

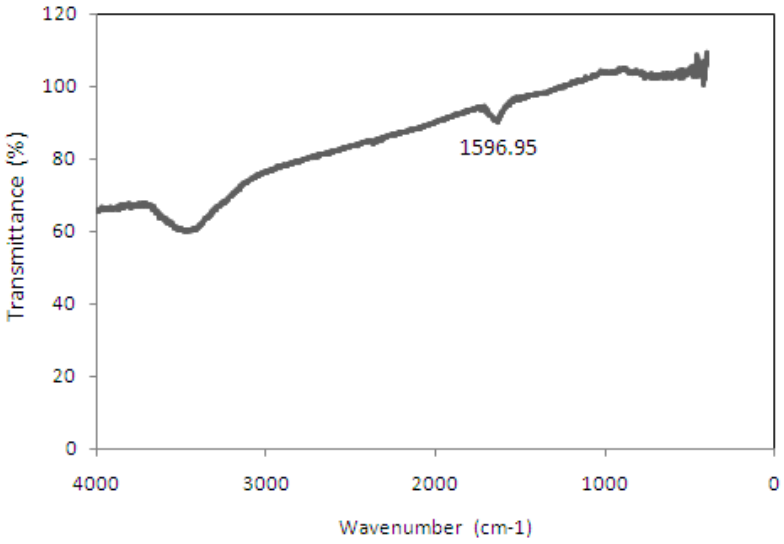



Fig.11: The FTIR spectra of treated antibiotic aqueous solution

In the modified Fenton ($FeGAC/H_2O_2$) treatment under the optimum operating conditions, biodegradability ($BOD_5$/COD ratio) of the antibiotic aqueous solution improved from zero (untreated) to 0.36 (treated), indicating that the effluent was amenable to biological treatment (Al-Momani *et al.,* 2002). A similar biodegradability improvement (zero to 0.37 and 0.40) was also observed in the Fenton and photo-Fenton treatment of the antibiotics amoxicillin, ampicillin and cloxacillin in the aqueous solution (Elmolla & Chaudhuri, 2009a; 2009b).

## CONCLUSION

The optimum operating conditions of the modified Fenton ($FeGAC/H_2O_2$) treatment of the antibiotics amoxicillin and cloxacillin in aqueous solution at pH3 were $H_2O_2$/COD molar ratio 2.0, FeGAC dose 3.5 g/L and reaction time of 90 min for 87.53% removal of COD, 78.01% removal of TOC and 98.24% removal of $NH_3$-N. The biodegradability ($BOD_5$/COD ratio) of the antibiotic aqueous solution improved from zero to 0.36. Meanwhile, FTIR spectra indicated degradation of the antibiotics. Hence, the modified Fenton ($FeGAC/H_2O_2$) process could be an alternative in the pretreatment of the amoxicillin and cloxacillin antibiotics in aqueous solution for biological treatment.

## ACKNOWLEDGEMENTS

The authors are thankful to the management and the authorities of Universiti Teknologi PETRONAS for providing the necessary facilities for this research.

## REFERENCES

Al-Momani, F., Touraud, E., Degorce-Dumas, J. R., Roussy, J., & Thomas, O. (2002). Biodegradility enhancement of textile dyes and textile wastewater by VUV photocatalysis. *Journal of Photochemistry and Photobiology A: Chemistry, 153*, 191-197.

Anacona, J. R., & Carman, C. G. (2005). Synthesis and antimbacterial activity of cefoxitin metal complexes. *Transition Metal Chemistry Journal, 30,* 605 – 609.

Anacona, J. R., & Figueroa, E. M. J. (1999). Synthesis and characterization of metal complexes with penicillin. *Journal of Coordination Chemistry*, *48*, 181-189.

APHA. (2005). *Standard Methods for the Examination of Water and Wastewater* (21st Edition). American Public Health association, American Water Works Association and Water Environment Federation, Washington DC.

Arjunan, V., Sakiladevi, S., Rani, T., Mythili, C. V., & Mohan, S. (2012) I. FTIR, FT-Raman, FT-NMR, UV-Visible and quantum chemical investigations of 2-amino-4-methylbenzothiazole *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 88,* 220-231.

Arslan-Alaton, I., Tureli, G., & Olmez-Hanci, T. (2009). Treatment of azo dye production wastewaters using photo-Fenton-like advanced oxidation processes: optimization by response surface methodology. *Journal* of *Photochemistry and Photobiology A: Chemistry*, *202,* 142-153.

Bach, A., & Semiat, R. (2010). The role of activated carbon as a catalyst in GAC/iron oxide/$H_2O_2$ oxidation process, *Desalination.* doi:10.1016/j.desal.2010.04.020.

Bansal, R. C., Donnet, J. B., & Stoeckli, F. (1998). *Active Carbon*. New York: Marcel Dekker.

Bebu, A., Szabó, L., Leopold, N., Berindean, C., & David, L. (2011). IR, Raman, SERS and DFT study of amoxicillin. *Journal of Molecular Structure*, *993*, 52–56.

Bezerra, M. A., Santelli, R. E., Oliviera, E. P., Villar, L. S., & Escaleira, L. A. (2008). Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta*, *76*, 965-977.

Chamarro, E., Marco, E., & Esplugas, S. (2001). Use of Fenton reagent to improve organic chemical biodegradability. *Water Research, 35*, 1047-1051.

Chiing-Chang, C., Wen-Ching, C., Mei-Rung, C., Sheng-Wei, C., Yao Yin, C., & Fan, H., (2011). Degradation of crystal violet by an $FeGAC/H_2O_2$ process, *Journal of Hazardous Materials*, *196*, 420– 425.

Elmolla, E., & Chaudhuri, M. (2009a). Optimization of Fenton process for treatment of amoxicillin, ampicillin and cloxacillin antibiotics in aqueous solution. *Journal of Hazardous Materials*, *170*, 666-672.

Elmolla, E., & Chaudhuri, M. (2009b). Degradation of amoxicillin, ampicillin and cloxacillin antibiotics in aqueous solution by the photo Fenton process. *Journal of Hazardous Materials*, *172*, 1476-1481.

Fan, H. J., Shu, H. Y., & Tajima, K. (2006). Decolourization of acid black 24 by the $FeGAC/H_2O_2$ process. *Journal of Hazardous Materials, B128*, 192-200.

Fan, H. J., Chen, I. W., Tzuchen, C., & Lee, M. H. (2007). Using $FeGAC/H_2O_2$ process for landfill leachate treatment. *Chemosphere, 67*, 1647-1652.

Ghafari, S., Aziz, H. A., Isa, M. H., & Zinatizadeh, A. A. (2009). Application of response surface methodology (RSM) to optimize coagulation flocculation treatment of leachate using poly-aluminium chloride (PAC) and alum. *Journal of Hazardous Materials*, *163*, 650-656.

Goi, A., Veressinina, Y., & Trapido, M. (2008). Review. Degradation of salicylic acid by Fenton and modified Fenton treatment. *Chemical Engineering Journal, 143,* 1-9.

Gulkowska, A., Leunga, H. W., Soa, M. K., Taniyasub, S., Yamashit, N., Yeung, L. W. Y., Richardson, B. J., Lei, A. P., Giesy, J. P., & Lam, P. K. S. (2008). Removal of antibiotics from wastewater by sewage treatment facilities in Hong Kong and Shenzhen, China. *Water Research, 42*, 395-403.

Hach. (2002). *Water Analysis Handbook* (4th Edn.). Loveland, CO: Hach Company.

Jang, M., Chen, W. F., & Cannon, F. S. (2008). Preloading hydrous ferric oxide into granular activated carbon for arsenic removal. *Environtal Science Technol*ogy, *42*, 3369-3374.

Jones, O. A. H., Voulvoulis, N., & Lester, J. N. (2007). The occurrence and removal of selected pharmaceutical compounds in a sewage treatment works utilizing activated sludge treatment. *Environmental Pollution, 145*, 738-744.

Jones, O. A. H., Voulvoulis, N., & Lester, J. N. (2001). Human pharmaceuticals in the aquatic environment: a review. *Environmental Technology, 22,* 1383-1394.

Kang, Y. W., Cho, M. J., & Hwang, K. Y. (1999). Correction of hydrogen peroxide interference on standard chemical oxygen demand test. *Water Research, 33*, 1247-1251.

Khalil, L. B., Girgis, B. S., & Tafwik, TA. M. (2001). Decomposition of $H_2O_2$ on activated carbon obtained from olive stones. *Journal of Chemical Technology and Biotechnology*, *76,* 622-627.

Khuri, A. I., & Cornell, J. A. (1996). *Response Surfaces: Designs and Analyses* (2nd Edn.) New York, NY: Marcel Dekker Inc.

Kimura, M., & Miyamoto, I. (1994). Discovery of the activated-carbon radical $AC^+$ and the novel oxidation-reactions comprising the $AC/AC^+$ cycle as a catalyst in an aqueous solution, *Bulletin of the Chemical Society of Japan*, 67, 2357-2360.

Kupka, T. (1997). *Beta-Lactam antibiotics. Spectroscopy and molecular orbital (MO) calculations*. Part I: IR studies of complexation in penicillin-transition metal ion systems and semi-empirical $PM_3$ calculations on simple model compounds. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 53,* 2649-2658.

Pignatello, J. J., Liu, D., & Houston, P. (1999). Evidence for an additional oxidant in the photo assisted Fenton reaction. *Environmental Science Technology, 33,* 1832-1839.

Rozas, O., Contreras, D., Mondaca, M. A., Perez-Moyac, M., & Mansilla, H. D. (2010). Experimental design of Fenton and photo-Fenton reactions for the treatment of ampicillin solutions. *Journal of Hazardous Materials*, *77*, 1025-1030.

Stackelberg, P. E., Furlong, E. T., Meyer, M. T., Zaugg, S. D., Henderson, A. K., & Reissman, D. B. (2004). Persistence of pharmaceutical compounds and other organic wastewater contaminants in a conventional drinking-water-treatment plant. *Science of the Total Environment*, *329,* 99-113.

Talinli, I., & Anderson, G. K. (1992). Interference of hydrogen peroxide on the standard COD test. *Water Research, 26,* 107-110.

Tella, A. C., Obaleye, M. O., & Akolade, E. O. (2011). Metal chelates of cellulose-antibiotics and their antimicrobial activities. *Middle-East journal of Scientific Research*, *7,* 260-265.

Thirunavukkarasu, O. Viraraghavan, T., & Subrama, K. (2003). Arsenic removal from drinking water using iron oxide-coated sand. *Water Air Soil Pollution, 142*, 95-1114.

Walter, M. V., & Vennes, J. W. (1985). Occurrence of multiple-antibiotic resistant enteric bacteria in domestic sewage and oxidative lagoons. *Applied and Environmental Microbiology, 50,* 930-933.

Pertanika J. Sci. & Technol. 22 (1): 255 - 271 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Microclimate inside a Tropical Greenhouse Equipped with Evaporative Cooling Pads

**Diyana Jamaludin[1]*, Desa Ahmad[1], Rezuwan Kamaruddin[2] and Hawa Z. E. Jaafar[3]**

*[1]Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*
*[2]School of Bioprocess Engineering, Universiti Malaysia Perlis, Kompleks Pusat Pengajian Jejawi 3, 02600 Arau, Perlis, Malaysia*
*[3]Faculty of Agriculture, Universiti Putra Malysia, 43400 Serdang, Selangor, Malaysia*

## ABSTRACT

Tropical greenhouses require active evaporative cooling system such as pad-and-fan to ensure a suitable microclimate for crop production. Excess heat causes indoor temperature to become hotter than desired resulting in detrimental effects to crop growth and production. Solar radiation intensity and outside temperature affect temperature and relative humidity level inside a greenhouse, while wide gradients in temperature and relative humidity can cause problems related to crop growth and production uniformity. A 300 $m^2$ greenhouse, equipped with evaporative pad and four exhaust fans at each end walls, was used in the study. Horizontal and vertical profiles of the temperature and relative humidity inside the greenhouse were investigated. Results showed that temperature increased from evaporative pad area to exhaust fans area in a horizontal plane, while relative humidity showed an inverse pattern from temperature. In the vertical plane, temperature increased, while relative humidity decreased from lower level to the upper level. ANOVA results showed that in overall, temperature and relative humidity inside the greenhouse were uniform as there was no significant difference at 95% confidence interval. Thus, it was concluded that greenhouse cooling system by evaporative pad and exhaust fans are suitable for application in a tropical country such as Malaysia.

*Keywords:* Tropical greenhouse, pad-and-fan, temperature, humidity

*Article history:*
Received: 2 April 2012
Accepted: 21 June 2012

*E-mail addresses:*
diyana_upm@upm.edu.my (Diyana Jamaludin),
desa@upm.edu.my (Desa Ahmad),
rezuwan@unimap.edu.my (Rezuwan Kamaruddin),
hawazej@upm.edu.my (Hawa Z. E. Jaafar)
*Corresponding Author

## INTRODUCTION

Controlled environment in a greenhouse can provide suitable conditions for temperate vegetables and flowers to be grown optimally in Malaysian climate. In the tropics, the main constraints of crop production in the open

ISSN: 0128-7680 

fields are extreme solar radiation, high rainfall, high humidity, insects and disease infestation (Hawa, 1990; Rezuwan, 2000). Excess heat in the greenhouse environment is considered a challenge in Malaysia, especially for lowland areas. This is because excess heat can cause indoor temperature to become much hotter than desired (Kittas, 2003) and thus, giving detrimental effects to crop growth and production. Malaysian climate, which is hot and humid, is not suitable for temperate crops and vegetables, causing Malaysia to import RM680 million worth of vegetables annually (Rezuwan, 2000). The major types of imported vegetables include high value temperate vegetables such as cabbage, cauliflower, broccoli, tomato and bell pepper.

Optimum temperature and humidity are essential in order to provide a suitable condition inside the greenhouse. Methods of reducing excess heat in a greenhouse are many. One is by natural ventilation, which lessens the amount of heat energy from the sun and also by shading (Miguel *et al*., 1994; Hawa, 2006). Although ventilation is probably the simplest and cheapest way to reduce heat, the method is not efficient enough to reduce temperature inside the greenhouse in the tropical climate. Based on a study by Faisal *et al*. (2006), natural ventilation alone is not enough to reduce the temperature in a tropical greenhouse in Malaysia. Hence, additional cooling system is essential. One of the methods used to reduce the temperature inside a greenhouse is by evaporative cooling. In this method, as water evaporates, energy is lost from the air and the temperature is reduced (Simmons & Lott, 1996). In the Mediterranean region, evaporative cooling is desirable to prevent plant stress and produce marketable quality of crops (Hanan *et al*., 1998).

The greenhouse in Malaysia requires cooling system to reduce the inside temperature. Most of the heat load inside the greenhouse comes from solar radiation (Walker, 1983). A major greenhouse operational cost comes from electrical consumption, whereby the cooling system contributes a significant part (Fang, 1995). Optimum temperature and humidity are essential in order to provide a suitable condition inside the greenhouse. Moreover, uniformity of temperature and humidity distribution inside a greenhouse are also crucial to gain more productivity in an effective way (Arbel *et al*., 2003). Most previous studies on the greenhouse cooling system were carried out for temperate and dry areas. For instance, Kittas et al. (2001) presented sensible and latent heat profiles along a 60-m long greenhouse in Mediterranean area. It was reported that a 4°C rise in temperature is tolerated across the greenhouse (Nelson, 2003). However, data of the horizontal and vertical microclimate profiles inside the greenhouse with cooling pads in the tropical areas are rather limited.

This paper highlights a study on the evaluation of microclimate inside the greenhouse using the pad-and-fan cooling system. Both temperature and humidity profiles inside the greenhouse were evaluated in order to find uniformity of the cooling system.

## EXPERIMENTAL SETUP

### *Study Area*

The experiments were carried out at the Malaysian Agricultural Research and Development Institute (MARDI) in Serdang, Selangor. The Institute is located at the latitude 2°59' N and the longitude 101°42' E. A quonset shape greenhouse, galvanized steel tube with polyethylene covering material, equipped with the pad-and-fan evaporative cooling system, was used for the

study. The dimension of the greenhouse is 30 m in length, 10 m in width and 4.3 m in height. The greenhouse has an area of 300 $m^2$ and a volume of 1020 $m^3$. The greenhouse elevation is 37.8 m above the mean sea level.

The covering material used for the greenhouse is a clear polyethylene thermic film. The film has a transmissivity value of more than 95% with U.V. stabilizer, anti-static and anti-condensation. Polyethylene (P.E.) is used as the covering material as it has low transmission coefficient compared to glass, which can reduce the amount of solar radiation heat transmitted inside greenhouse, and it is also less expensive compared to acrylic.

*Cooling System*

The cooling system adopted was the evaporative pad-and-fan type. Evaporative pads were attached at the end south-wall of the greenhouse (Fig.1). Ambient air was forced through a 9-m width by 2-m height and 1.5-mm thick, wet cellulose cooling pads of the 10-m wide south-wall. The pads have 85% humidification efficiency as claimed by its manufacturer, with 50 Pa pressure lost at 2.0 m $s^{-1}$ air velocity. Evaporative pads were operated only during daytime from 9 am to 5 pm. During night time, the system was switched off to avoid excess moisture inside the greenhouse, as well as to avoid pathogen and disease problems. As air flows past the moist pad surfaces, some of the moisture evaporates into the air stream. The cool air from the pads flows across the 30-m length of the greenhouse before it is being exhausted by the fans placed at the opposite end of the north wall.

The diameter of each fan is about 1.2 m, using 1.12kW with 750 cubic meters per minute (cmm). Louvers will close when the fans shut down to avoid insects from getting into the greenhouse. The distance between each fan is 1.5 m and it is placed 1 m above the ground (Fig.2). These heavy duty welded steel frames have belt driven panels for low speed operation. Different fans are used each day to reduce wear, overelative humidity eating and machine failure. The cooling system is controlled by a control panel, where the cooling stage can be set up to run either automatically or manually.

Fig.1: Evaporative pad constructed at the south wall of the greenhouse

Fig.2: Exhaust fans placed at the north wall of the greenhouse

## METHODOLOGY

### *Sensors and Equipment*

WatchDog 2000 series weather stations (Fig.3) are used for the greenhouse setup to provide real-time, local weather information and enable monitoring site-specific growing condition. One unit was placed outside the greenhouse located at the north side, and the other unit was put at the centre of the greenhouse. The weather stations are meant for measuring temperature, relative humidity, rainfall, solar radiation, wind speed, and wind plane. Measurement interval was selected from 1, 10, 15, 30, or 60 minutes. A 30-minute interval was selected to record the microclimate for 183 days before the station's memory is full. An LCD screen displays current and high or low readings. It is able to review the past 30 days' data and to confirm that the station and sensors are functioning. Built-in data logger that stores measurements in a file-safe, non-volatile memory gives an ideal solution to collect data without having the researcher present at the site during the period of the experiment or study. These data were then transferred into a PC via direct PC interface cable.


Fig.3: WatchDog weather station

Spectrum's solarimeter was used to measure solar radiation intensity. The sensor sub-samples solar radiation between 300 and 1100 nanometers. The solarimeter was positioned in an appropriate area, without being shadowed or blocked by any other sensor or structure.

In order to determine the temperature and humidity profiles and gradients inside the greenhouse, temperature and humidity sensors incorporated into the WatchDog 200 Series Data Logger (Fig.4) were placed horizontally and vertically inside the greenhouse at a height of 1.0 m above the ground. The data logger sensor provides two sensor channels with a capacity of 7000 measurements. This sensor is mounted in an aspirated enclosure to avoid direct effects from solar radiation and high air velocity.


Fig.4: WatchDog 2000 Series data logger

*Location of the Sensors*

The length of the greenhouse was divided into three designated areas, namely, pad area (0-10 m), middle area (11-20 m) and fan area (21-30 m) (see Table 1). There were eight sensors to collect data on the temperature and humidity inside the greenhouse and one weather station to collect data on the temperature, humidity, solar radiation, and wind speed outside the greenhouse every 30-minute interval. All the above measurements were recorded on a data logger system. The greenhouse was empty with no crops planted. Six sensors were installed horizontally along the middle of the greenhouse, with 5 m distance from pad to fan (see Fig.5). Using the Cartesian coordinate system, these sensors were placed at (0, 5) - Sensor 1, (5, 5) - Sensor 2, (10, 5) - Sensor 3, (15, 5) - Sensor 4, (20, 5) - Sensor 5, and (25, 5) - Sensor 6. In the vertical plane, 3 data loggers were placed at 0.3 m (Sensor A), 1.0 m (Sensor B), and 2.5 m (Sensor C) from the ground (Fig.6) to refer to the low, middle and top positions. The sensors in the vertical plane were located at the centre of the greenhouse, i.e. at (15, 5) in the Cartesian coordinate. For the purpose of this study, a 60-minute interval (1 hour) was selected and data collection was running for two months.

TABLE 1
Area segment and representation

| *Area* | *Segment (m)* | Represent |
|---|---|---|
| Pad | 0-10 | Cool area |
| Middle | 11-20 | Medium area |
| Fan | 21-30 | Warm area |

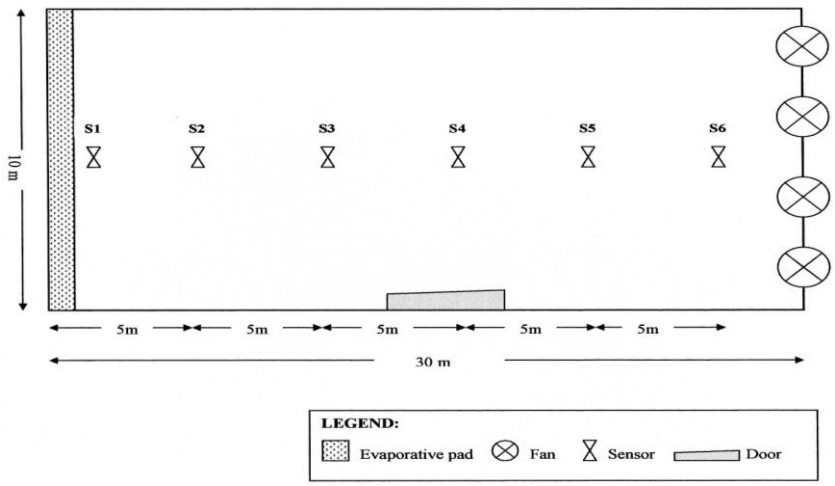



Fig.5: Greenhouse floor plan of the sensors in a horizontal plane

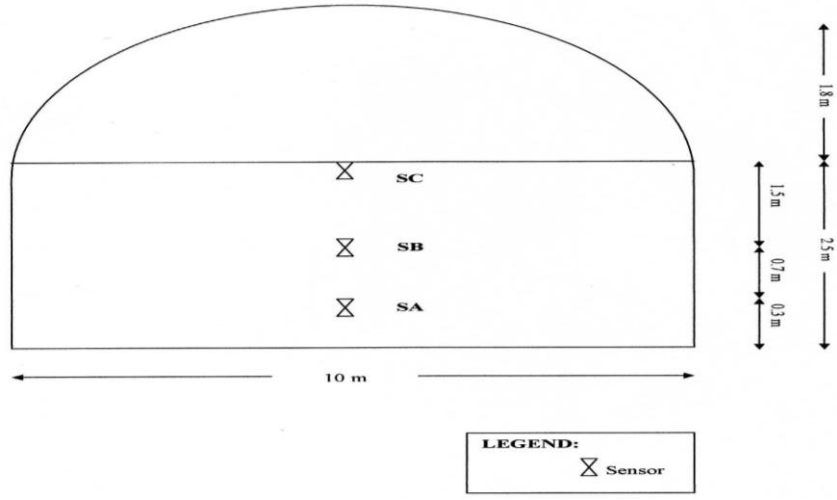



Fig.6: Greenhouse side plan of the sensors in a vertical plane

## RESULTS AND DISCUSSION

*Inside Temperature versus Outside Temperature*

Fig.7 shows that the temperature inside the greenhouse increased when the temperature on the outside increased, and this increment was with a regression coefficient of 0.996. This was an average temperature from eight sensors. The result suggests that with the increase in the ambient (outside) temperature, heat transfer into the greenhouse also increased and led to an increase in the temperature inside the greenhouse. The highest average temperature inside the greenhouse was 33.0°C. The general regression equation for the relationship between temperatures inside and outside the greenhouse shown in the graph can be formulated as follows:

$$\text{Inside temperature} = 0.9976^{*} [\text{Outside temperature}] - 0.0129;\ R^2 = 0.996$$

*Inside Relative Humidity versus Outside Relative Humidity*

The relationship between inside relative humidity and outside relative humidity is shown in Fig.8. This is the average humidity taken from eight sensors. The results show a linear relationship between inside and outside relative humidity with a strong correlation of $R^2$= 0.997, and an average inside relative humidity of 53.6%. The general regression equation for the graph is formulated as follows:

Inside relative humidity= 1.064 *[Outside relative humidity] _ 8.2844; $R^2$ = 0.997

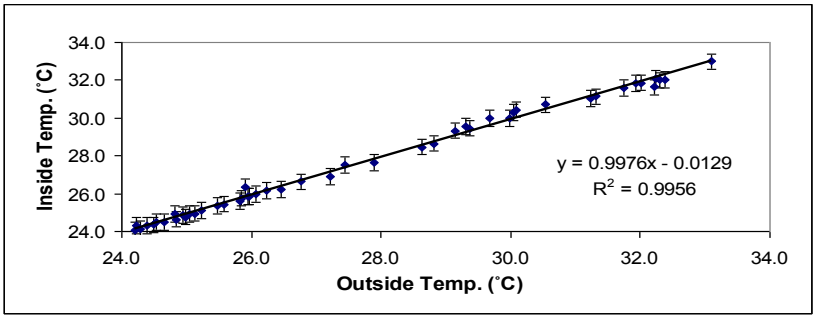



Fig.7: Inside temperature versus outside temperature in the greenhouse

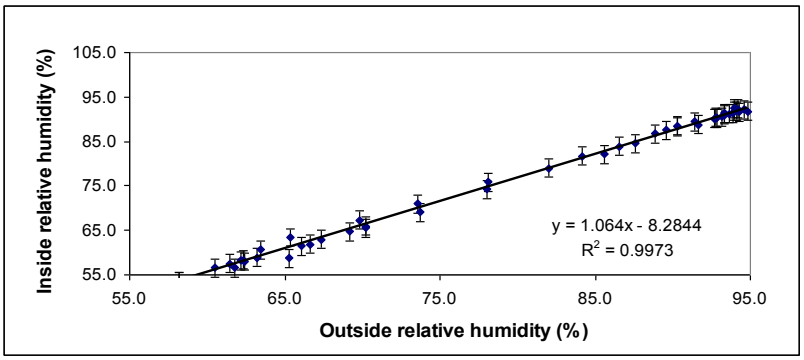



Fig.8: Inside relative humidity versus outside relative humidity

*Temperature and Relative Humidity inside the Greenhouse versus Time*

The hourly variation of the vertical temperature profile inside the greenhouse is shown in Fig.9. Each sensor showed that the temperature started to increase from 8:30 am and achieved the peak value at 14:00 pm, before it decreased until 20:00 pm. The average day temperature (7.30 am – 6.30 pm) inside the greenhouse is 30.7°C, which is lower than the average outside temperature, with the mean of cooling system (Table 2). Meanwhile, the average night temperature (7.00 pm - 7.00 am) is 25.7°C, indicating that the outside temperature is much lower than the inside temperature because the evaporative pad system was not running to avoid high relative humidity in the greenhouse. However, the average temperature inside the greenhouse during the night is acceptable for crops.

TABLE 2
Average temperature and relative humidity inside and outside the greenhouse at vertical plane

| | Average Day Temperature (°C) | Average Night Temperature (°C) | Average Day Relative Humidity (%) | Average Night Relative Humidity (%) |
|---|---|---|---|---|
| Inside Greenhouse | 30.7 | 25.7 | 71.2 | 83.5 |
| Outside Greenhouse | 31.1 | 25.4 | 68 | 87.6 |

Fig.10 shows the hourly variation of vertical inside relative humidity in an empty greenhouse. Each sensor showed that the relative humidity started to decrease from 8:30 am and achieved the lowest value at 14:00 pm. Then, the relative humidity started to increase through the day. This showed that during the morning and night, the humidity inside the greenhouse was high as the environment was condensed with vapour. When solar radiation intensity got higher, much of the vapour started to evaporate causing lower relative humidity. The average day relative humidity (7.30 am – 6.30 pm) inside the greenhouse is higher than average outside relative humidity with the mean of cooling system (Table 2). The average night relative humidity (7.00 pm- 7.00 am) shows that outside relative humidity is not much different than the inside relative humidity because the evaporative pad system was not running to avoid the high relative humidity in the greenhouse.

The average day temperature (7.30 am – 6.30 pm) inside the greenhouse at horizontal plane is 32.1°C, which is higher by 3.1% from the average outside temperature through the cooling system. This shows that the cooling system is insufficient to reduce the temperature inside the greenhouse in the horizontal plane to the range of 24 - 30˚C. Hence, more fans are needed to be run so as to decrease the temperature through convection heat transfer. The average night temperature (7.00 pm - 7.00 am) is 26.2°C, indicating that the outside temperature is much lower by 3.8% than the inside temperature because the evaporative pad system was not running to avoid the high relative humidity in the greenhouse (Table 3). However, the average temperature inside the greenhouse during night time is acceptable for crops.

In Fig.11, each sensor showed that relative humidity started to decrease from 8:00 am and achieved the lowest value at 13:00 pm. Relative humidity then started to increase through the day. Relative humidity inside the greenhouse was comparatively lower than ambient relative humidity. The graph also shows an inverse pattern with the graph of temperature versus time (Fig.12), whereby relative humidity was at the lowest value when the temperature achieved the peak value. This result supports the theory that relative humidity has an inverse relationship with temperature. Table 3 summarizes the average relative humidity inside and outside the greenhouse during at day and night times. It shows that the inside relative humidity at day time is lower than the outside relative humidity by 11.8% and at night time, the relative humidity inside the greenhouse is also lower than the outside by 4.1%.

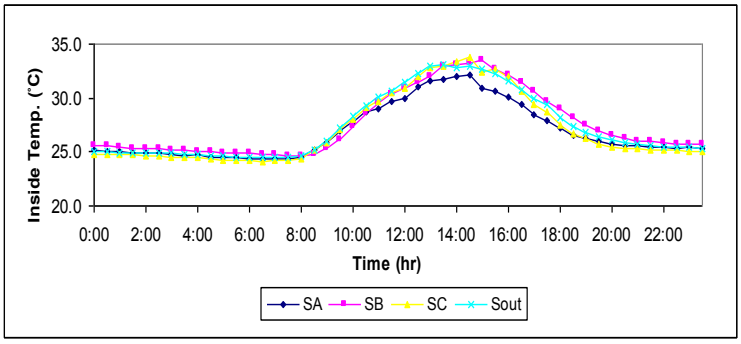



Fig.9: Changes in the vertical temperature profile with time inside the greenhouse. A represents low sensor, B represents middle sensor and C represents top sensor

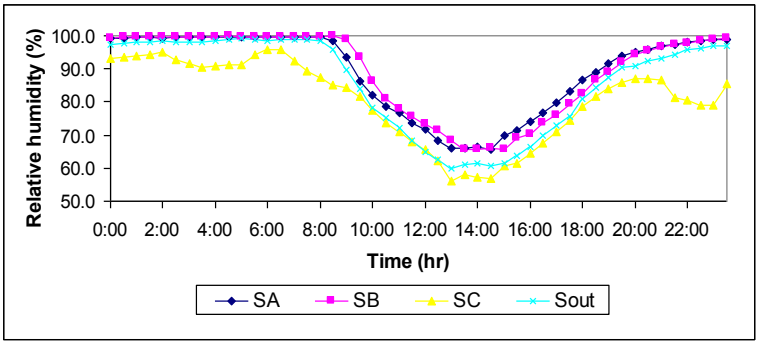



Fig.10: Changes in the vertical relative humidity profile with time inside the greenhouse. A represents low sensor, B represents middle sensor and C represents top sensor

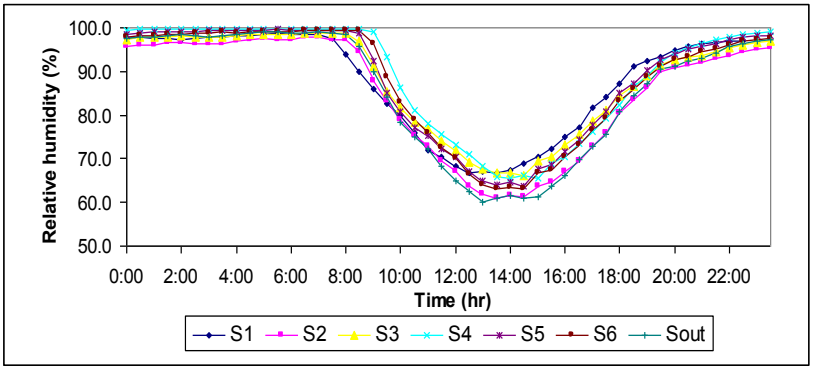



Fig.11: Changes in the horizontal relative humidity profiles with time inside the greenhouse

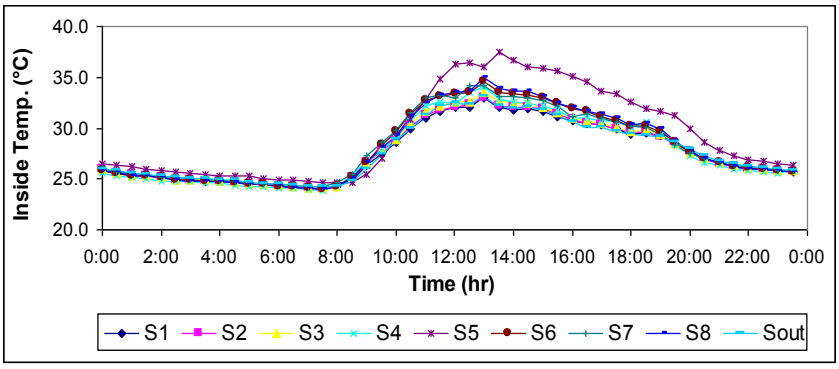



Fig.12: Changes in the horizontal temperature profile with time inside an empty greenhouse

TABLE 3
Average day and night temperature inside and outside the greenhouse at horizontal plane

| | Average Day Temperature (°C) | Average Night Temperature (°C) | Average Day Relative Humidity (%) | Average Night Relative Humidity (%) |
|---|---|---|---|---|
| Inside Greenhouse | 32.1 | 26.2 | 56.2 | 83.5 |
| Outside Greenhouse | 31.1 | 26.1 | 68 | 87.6 |

*Vertical Temperature and Relative Humidity inside the Greenhouse*

There were three temperature sensors placed in the vertical plane of the greenhouse at 0.3 m (Sensor A), 1.0 m (Sensor B) and 2.5 m (Sensor C) height. As shown in Fig.13, Sensor C had the highest average temperature (27.55°C), followed by Sensor B (27.10°C) and Sensor A (26.85°C). This shows that the convection heat transfer was from the lower level to the upper level as hot air moves to a higher altitude when it gets more energy and less density, especially when solar radiation is high. The temperature recorder at 2.5 m height (SC) was higher than the ambient temperature by 0.9°C at 14:00 hr, while at the temperature at the bottom level of the greenhouse (0.3 m – SA) was lower than ambient temperature by 0.8°C. Sensor C (SC) received higher solar radiation intensity compared to the bottom level and thus had the highest temperature than Sensors A and B. From the ANOVA results, the difference between each point in each hour was found to be insignificant as $F<F_{crit}$ at 95% confidence interval (Table 4). This was because the greenhouse was a closed system, and this caused it to have more uniform temperature between each elevation.

Fig.13 illustrates that Sensor C had the lowest relative humidity (80.56%), followed by Sensor B (89.1%) and Sensor A (89.21%). These results show that relative humidity has an opposite pattern from temperature as it will decrease relative humidity increases. Table 5 shows the ANOVA result on each sensor, where it reveals the significant difference within each sensor with p-value less than 0.05.

*Horizontal Temperature and Relative Humidity inside the Greenhouse*

Fig.14 shows that the temperature inside the greenhouse was much higher near the exhaust fan area (26.8°C), followed by the middle area (26.5°C) and near the evaporative pad area (25.9°C). This confirmed that the temperature increased along the greenhouse from the evaporative pad to the exhaust fan due to solar heat gain. Nevertheless, the ANOVA results show no significant difference in the temperature in each area as $F < F_{critical}$ with p-value= 0.09 at 95% confident interval (Table 6). Meanwhile, temperatures near the evaporative pad are cooler than the outside temperature by an average of 1°C, while the temperature in the middle of the greenhouse is hotter than the outside temperature by 5°C. This proves that the temperature increases from the evaporative pads to exhaust fans as a result of conduction and radiation heat transfer. Similar trends have also been observed by Montero *et al*. (1981), Kittas *et al*. (2001), and Al-Helal (2006).

Fig.14 also shows that relative humidity inside the greenhouse is much higher near the evaporative pad area (82.8%), followed by the middle area (82.3%) and near exhaust fans area (80.9%). This confirms that relative humidity decreases along the greenhouse from the evaporative pad to the exhaust fan due to solar heat gain which reduces moisture content inside the greenhouse. The data were statistically analyzed and the ANOVA results are summarized in Table 7. The findings show that the relative humidity between each area is not significant where $F< F_{critical}$, with p-value = 0.39 at 95% confidence interval. This means that there is no significant difference in the relative humidity between each location.

Meanwhile, the ANOVA results show that the temperature inside the greenhouse in the vertical and horizontal planes has no significant difference at 95% confidence interval (Table 9). It also shows that relative humidity in the vertical plane is significantly different between each sensor (Table 10).

TABLE 4
Distribution and ANOVA for inside temperature between sensors in vertical plane

| Groups | Count | Sum | Average | Variance |
|---|---|---|---|---|
| Sensor A | 48 | 1288.94 | 26.85 | 6.31 |
| Sensor B | 48 | 1300.96 | 27.10 | 10.04 |
| Sensor C | 48 | 1322.34 | 27.55 | 8.68 |

| Source of Variation | SS | df | MS | F* |
|---|---|---|---|---|
| Between Groups | 11.923 | 2 | 5.96 | 0.71 |
| Error | 1176.247 | 141 | 8.34 | |
| Total | 1188.171 | 143 | | |

* insignificant at 0.05 probability level

TABLE 5
Distribution and ANOVA for inside relative humidity between sensors in vertical plane

| Groups | Count | Sum | Average | Variance |
|---|---|---|---|---|
| Sensor A | 48 | 4282.2 | 89.21 | 153.04 |
| Sensor B | 48 | 4279.94 | 89.17 | 161.69 |
| Sensor C | 48 | 3867.11 | 80.56 | 151.38 |

| Source of Variation | SS | df | MS | F** |
|---|---|---|---|---|
| Between Groups | 2380.05 | 2 | 1190.03 | 7.66 |
| Error | 21907.57 | 141 | 155.38 | |
| Total | 24287.62 | 143 | | |

** significant at 0.05 probability level

TABLE 6
Distribution and ANOVA for inside temperature between sensors in horizontal plane

| Groups | Count | Sum | Average | Variance |
|---|---|---|---|---|
| Pad | 48 | 1244.11 | 25.92 | 2.11 |
| Middle | 48 | 1274.01 | 26.54 | 3.84 |
| Fan | 48 | 1285.21 | 26.78 | 5.36 |

ANOVA

| Source of Variation | SS | Df | MS | F* |
|---|---|---|---|---|
| Between Groups | 18.82 | 2 | 9.41 | 2.49 |
| Error | 531.85 | 141 | 3.77 | |
| Total | 550.67 | 143 | | |

* insignificant at 0.05 probability level

TABLE 7
Distribution and ANOVA for inside relative humidity between sensors in horizontal plane

| Groups | Count | Sum | Average | Variance |
|---|---|---|---|---|
| Pad | 48 | 3972.06 | 82.75 | 38.66 |
| Middle | 48 | 3987.51 | 83.07 | 85.60 |
| Fan | 48 | 3876.73 | 80.77 | 111.99 |

ANOVA

| Source of Variation | SS | df | MS | F* |
|---|---|---|---|---|
| Between Groups | 149.98 | 2 | 74.99 | 0.95 |
| Error | 11103.67 | 141 | 78.75 | |
| Total | 11253.64 | 143 | | |

* insignificant at 0.05 probability level

TABLE 8
ANOVA statistical test for inside temperature factors

| Regression Statistics | |
|---|---|
| R Square | 0.992 |
| Standard Error | 0.184 |
| Observations | 47 |

| | df | SS | MS | F | Significance F |
|---|---|---|---|---|---|
| Regression | 3 | 189.074 | 63.025 | 1864.996 | 1.56657E-45 |
| Residual | 43 | 1.453 | 0.034 | | |
| Total | 46 | 190.527 | | | |

| | Coefficients | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
|---|---|---|---|---|---|---|
| Intercept | 3.8454 | 0.708 | 5.428 | 2.46E-06 | 2.417 | 5.274 |
| Outside temperature (X1) | 0.8309 | 0.028 | 29.451 | 3.88E-30 | 0.774 | 0.888 |
| Solar radiation (X2) | -0.0024 | 0.0003 | -6.320 | 1.25E-07 | -0.003 | -0.002 |
| Rainfall (X3) | -0.0352 | 0.049 | -0.705 | 0.485 | -0.135 | 0.065 |

TABLE 9
ANOVA for temperature inside the greenhouse between sensors

| Inside Temperature | F | $F_{critical}$ | p-value |
|---|---|---|---|
| Vertical plane | 2.49 | 3.06 | 0.49 |
| Horizontal plane | 1.35 | 2.03 | 0.23 |

TABLE 10
ANOVA for relative humidity inside the greenhouse between sensors

| Inside Relative Humidity | F | $F_{critical}$ | p-value |
|---|---|---|---|
| Vertical plane | 0.95 | 3.06 | <0.001 |
| Horizontal plane | 1.42 | 2.03 | 0.19 |

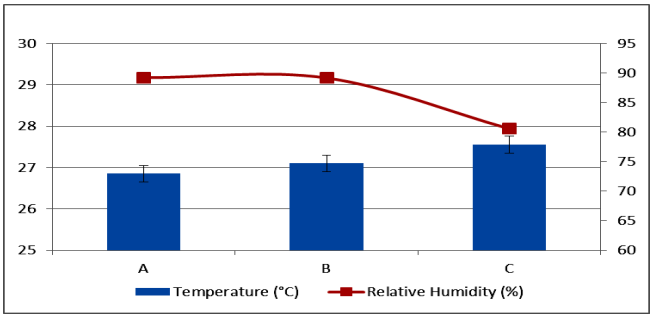



Fig.13: Inside temperature and relative humidity at vertical plane. A represents low sensor, B represents middle sensor and C represents top sensor

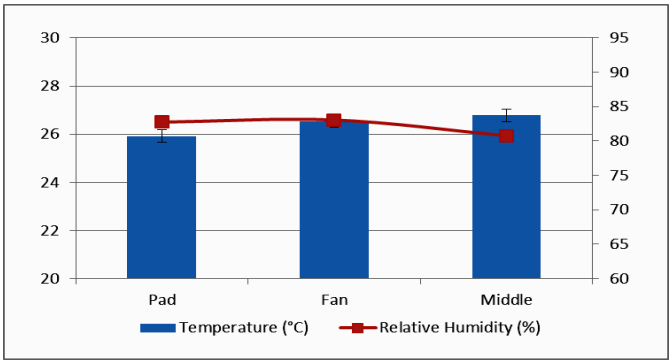



Fig.14: Inside temperature and relative humidity at horizontal plane

*Inside Temperature versus Solar Radiation*

Solar radiations outside the greenhouse gave a direct impact on the temperature inside the greenhouse due to radiation heat transfer process. Fig.15 shows linear relationships between the inside temperature and solar radiation, with a good regression coefficient of $R^2 = 0.77$. This result suggests that solar radiation heats up the greenhouse and increases heat transfer through radiation which leads to the temperature increase inside the greenhouse.

Inside temperature = 0.0179* Solar radiation + 26.652

*Inside Temperature versus Rainfall Intensity*

Rainfall intensity shows low correlation with the temperature inside the greenhouse. Fig.16 shows low regression coefficient with $R^2 = 0.14$. This concludes that the cooling effect by rainfall is not enough to reduce the temperature inside the greenhouse.

Inside temperature = 5.2483* Rainfall intensity + 28.982

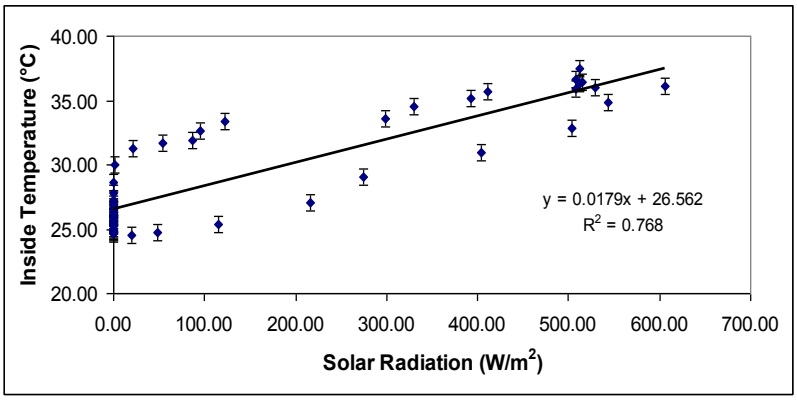



Fig.15: Inside temperature versus solar radiation

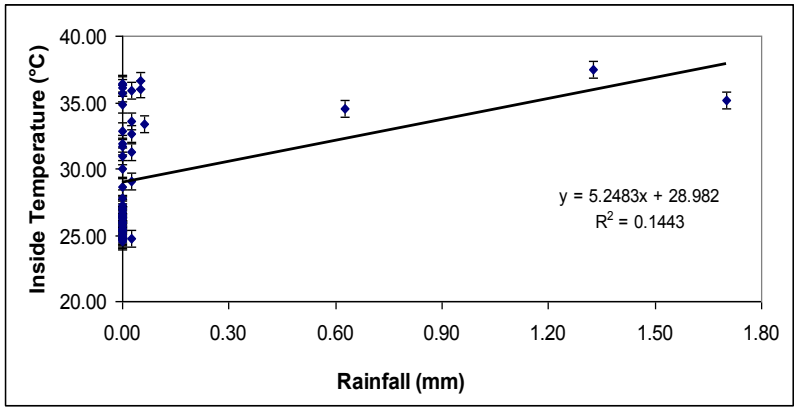



Fig.16: Inside temperature versus rainfall

Results of the Analysis of Variance (ANOVA) are presented in Table 8. The regression value between the outside temperature, solar radiation and rainfall with the inside temperature is 0.992 and the standard error is 0.184. These show that the outside temperature and solar radiation have p-value less than 0.05, while rainfall has p-value = 0.484622 > 0.05. Hence, it is concluded that rainfall has less effect on the temperature inside the greenhouse.

## CONCLUSION

The temperature inside the greenhouse increased from early morning until afternoon and then started to decrease until night time. The graph pattern of the inside temperature versus time is significantly similar with the graph of the solar radiation versus time due to the increase in temperature when solar radiation intensity increases. In the vertical plane, the temperature at the upper level (2.5 m) was found to be higher, followed by the temperature at the middle (1.0 m) and bottom levels (0.3 m). These results support the theory that hot air moves from lower level to upper level. However, since the greenhouse was fully covered, greenhouse pressure driven flow produced by convection was unlikely to differ at each level.

The temperature gradient was much obvious when the solar radiation intensity was high and the outside temperature was at the maximum in the day. Meanwhile, the inside temperature had a linear relationship with the outside temperature, with a strong regression coefficient. This concludes that the temperature inside the greenhouse with evaporative cooling system and exhaust fans can give uniform conditions inside the closed greenhouse.

Relative humidity values inside the greenhouse show opposite trends from temperature. In the vertical plane, relative humidity value at the upper level (Sensor C) was lower than at the middle (Sensor B) and bottom levels (Sensor A) of the greenhouse. In the horizontal plane, relative humidity inside the greenhouse decreased from near the evaporative pad to the exhaust fans. These results support the theory that relative humidity has an inverse relationship with temperature. Relative humidity inside the greenhouse was shown to be higher than outside the greenhouse as evaporative pads had added in moisture inside the greenhouse.

## ACKNOWLEDGEMENTS

The authors wish to thank and are grateful to the Malaysian Agriculture Research and Development Institute for providing the experimental site, cultivation materials and equipment, as well as sensors and skilled technical assistance throughout the field experiments.

## REFERENCES

Al-Helal, I. M. (2007). Effects of ventilation rate on the environment of a fan-pad evaporatively cooled, shaded greenhouse in extreme arid climates. *Journal of Applied Engineering in Agriculture, 23*(2), 221-230.

Arbel, A., Barak, M., & Shklyar, A. (2003). Combination of forced ventilation and fogging systems for cooling greenhouses. *Journal of Agricultural Engineering Research*, *84*(1), 45-55.

Faisal, M. S. A., Shariff, A. R. M., Rezuwan, K., Ahmad, D., Janius, R., & Mohamad, M. Y. (2006). Microclimate inside tunnel-roof and jack-roof tropical greenhouses structures. *Acta Horticulturae, 710*, ISHS.

Fang, W. (1995) Greenhouse cooling in subtropical regions. *Acta Hort., 399*.

Hanan, J. J. (1998). Advanced Technology for Protected Horticulture. In *Ventilation and Cooling* (pp. 236-260). CRC Press.

Hawa, Z. E. J. (2006). Carbon dioxide enriched production technology for controlled environment system in the lowland tropics. *Acta Hort. (ISHI)*, *710.*

Hawa, J. (1990). Performance and feasibility of high value vegetables cultivation under rainshelters. *MARDI Scientific Council*, *66*.

Kittas, C., Bartzanas, T., & Jaffrin, A. (2001). Greenhouse evaporative cooling: measurement and data analysis. *Transaction of ASAE, 44*(3), 663-689.

Kittas, C., Bartzanas, T., & Jaffrin, A. (2003). Temperature gradients in a partially shaded large greenhouse equipped with evaporative cooling pads*. Biosystem Engineering*, *85*(1), 87-94.

Miguel, A. F., Silva, A. M., & Rosa, R. (1994). Solar irradiation inside a single-span greenhouse with shading screens. *Journal of Agricultural Engineering Research*. *59*, 61-72.

Nelson, P. V. (2003). *Greenhouse operation and Management*. New Jersey: Prentice Hall Inc.

Rezuwan, K. (2000). *Naturally ventilated greenhouse structure for temperate vegetable production in tropical lowlands*. Presented to Institution of Engineers, Malaysia. IEM Conference Hall, Petaling Jaya, 26th July 2000.

Simmons, J. D., & Lott, B. D. (1996). Evaporative cooling performance resulting from changes in water temperature. *Appled Engineering of Agriculture, ASAE, 12*(4), 497-500.

Walker, J. N. (1983).Quantity of Air Flow for Greenhouse Structures. A.H. Mylo (Ed.), *Ventilation of Agricultural Structures* (pp. 257-277). ASAE Monograph.

Pertanika J. Sci. & Technol. 22 (1): 273 - 288 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Using Computational Fluid Dynamics in the Determination of Solar Collector Orientation and Stack Height of a Solar Induced Ventilation Prototype

**Yusoff, W. F. M.[1]*, Sapian, A. R.[2], Salleh, E.[3], Adam, N. M.[4], Hamzah, Z.[1] and Mamat, M. H. H.[5]**

*[1]Department of Architecture, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*
*[2]Department of Architecture, Kulliyyah of Architecture and Environmental Design, International Islamic University Malaysia, P.O. Box 10, 50728 Kuala Lumpur, Malaysia*
*[3]Solar Energy Research Institute, Level 3, Perpustakaan Tun Sri Lanang, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*
*[4]Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
*[5]Department of Operations and Manufacturing, TPM Engineering Sdn. Bhd, Technology Park Malaysia, Lebuhraya Puchong-Sg Besi, Bukit Jalil, 57000 Kuala Lumpur, Malaysia*

## ABSTRACT

Stack ventilation in the hot and humid climate is inherently inefficient due to minimal air temperature differences between indoor and outdoor environment of a naturally ventilated building. Solar induced ventilation is a viable alternative in enhancing this stack ventilation. This paper aims to demonstrate investigations on the effective solar collector orientation and stack height for a solar induced ventilation prototype that utilizes roof solar collector and vertical stack. The orientation of the solar collector is significant as it determines the amount of solar radiation absorbed by the solar collector. Meanwhile, the height of the vertical stack influences the creation of the stack pressure in inducing air movement. Investigations were executed using a simulation modelling software called FloVENT. The validation of the simulation modelling against physical experiment indicated a good agreement between these two results. Analyses were executed on the air temperature increments inside the solar collector. A high increment of the air temperature resulted in the effective orientation. Meanwhile, the air temperature and mass flow rate of the various heights of the vertical stack were also analyzed. The findings concluded that the recommended orientation for the prototype's solar collector is the west-facing orientation. It was also found that the higher the vertical stack, the lower the air temperature inside the stack would be, but with greater induced mass flow rate.

*Article history:*
Received: 18 May 2011
Accepted: 18 June 2012

*E-mail addresses:*
wardahyusoff@gmail.com (Yusoff, W. F. M.),
arazaks@iium.edu.my (Sapian, A. R.),
elsall06@gmail.com (Salleh, E.),
mariah@upm.edu.my (Adam, N. M.),
zh.akitek@gmail.com (Hamzah, Z.),
intelligentmove@gmail.com (Mamat, M. H. H.)
*Corresponding Author

ISSN: 0128-7680 

Keywords: Height, hot and humid climate, orientation, roof solar collector, solar induced ventilation, vertical stack

## INTRODUCTION

Solar induced ventilation studies have been gaining interest due to the strategy's potential in enhancing stack ventilation. The configuration normally consists of a glass cover, an air cavity and an absorber plate or wall. Solar radiation is used to heat up the absorber plate or wall. The heat from the absorber plate or wall, as well as the glass cover, is transferred to the air inside the cavity through convective heat transfer. Consequently, high air temperature differences between the air inside the cavity and the ambient air develops. The solar induced ventilation strategies that have been widely investigated are Trombe wall, solar chimney and roof solar collector (Awbi, 2003). Although they have similar operating concept, their configurations are somehow different. While the Trombe wall and solar chimney have vertical air cavity, the roof solar collector's air cavity is inclined and normally follows the roof slope.

Many studies have been carried out on the configurations of solar induced ventilation (see Bouchair, 1994; Chen *et al.*, 2003; Ding *et al.*, 2005; Gan, 2006; Hamdy & Fikry, 1998; Li *et al.*, 2004; Mathur *et al.*, 2006; Miyazaki *et al.*, 2006; Susanti *et al.*, 2008). The configuration parameters investigated were height, length, cavity width, tilt angle and opening sizes. The purpose of the studies was to enhance the performance of the solar induced ventilation. The investigations on the correlations between the cavity width and the inlet size showed that friction loss dominated in a small cavity width, whilst the pressure loss dominated in a small inlet size (Miyazaki *et al.*, 2006). Despite this, a simultaneous increase in both the cavity width and the inlet size resulted in no optimum cavity width, as there was no air flow rate reduction (Chen *et al.*, 2003; Gan, 2006).

The correlation between the cavity width and the stack height resulted in the effective ratio of 1:10 (Bouchair, 1994; Gan 2006; Li *et al.*, 2004). However, this correlation was investigated for solar induced ventilations with a vertical stack, namely, Trombe wall and solar chimney. Meanwhile, the height study that excluded this correlation indicated that the higher the stack, the greater the induced air flow rate (Ding *et al.*, 2005). The performance of solar induced ventilation is also affected by tilt angle and opening sizes. The effective tilt angle was found to be various, depending on the location's latitude. This is due to the different solar altitude with different latitudes. The optimum tilt angle for Jaipur, India, (27°N latitude) during the summer months was 45º (Mathur *et al.*, 2006). Meanwhile, for 32ºN latitude, the optimum tilt angle was 60º (Hamdy & Fikry, 1998). For the opening sizes, the ratio of 1 (ratio of inlet area to outlet area) was recommended in inducing high air flow rate (Susanti *et al.*, 2008).

The configuration studies mentioned above were executed for Trombe wall, solar chimney and roof solar collector. In this research, the configuration of the proposed solar induced ventilation is slightly different from those strategies. The solar induced ventilation proposed consists of two parts, which are the roof solar collector and the vertical stack. Roof solar collector has the advantage of collecting more solar radiation when the sun altitude is high, as compared to Trombe wall and solar chimney (Awbi, 2003; Mathur *et al.*, 2006). However,

its drawback is the height restriction of the stack, which is caused by the roof slope (Awbi, 2003; Harris & Helwig, 2007). Hence, the purpose of utilizing vertical stack is to increase the stack height, and consequently enhance the stack pressure. The proposed strategy has shown the potential in enhancing the stack ventilation in the hot and humid climate. For instance, the highest air temperature differences between the stack air and the ambient air attained by the proposed strategy was 9.9ºC for 877 W/m² solar radiation incidents (Yusoff *et al.*, 2010).

Due to its potential, this study was carried out with the aim to enhance the performance of the proposed strategy by investigating the orientation of the solar collector and the height of the vertical stack. These two parameters are significant as they affect the induced ventilation of the strategy. The solar collector orientation influences the catchment of solar radiation, whereas the vertical stack height affects the development of the stack pressure. Although there are studies on the effective solar collector orientation for the northern hemisphere countries, they are only conducted for inclined solar collector (Bari, 2001; Gunerhan & Hepbasli, 2007; Shariah *et al.*, 2002) and vertical solar chimney (Nugroho, 2007). The investigation of an effective solar collector orientation for the proposed strategy in the northern hemisphere countries, particularly Malaysia, is yet to be conducted. Meanwhile, the vertical stack of the proposed strategy is not used for collecting solar radiation. Hence, it is assumed that there is convective heat loss from the induced air to the stack walls. For this reason, the investigations were executed to determine whether height affects air temperature and mass flow rate inside the vertical stack.

This paper is divided into five sections. The first section reviews on the previous studies of solar induced ventilation configuration and briefly describes the present research. The second section describes the solar induced ventilation prototype investigated in this paper. The third section presents the methodology employed, which is, simulation modelling. This section also elaborates on the prototype model, simulation set-up and procedures, simulation conditions, and simulation validation. Section four discusses the results of the investigations. This section is divided into two sub-sections, which are, the results of the solar collector orientation and the vertical stack height studies. Finally, the overall findings of the research are concluded in section five.

## THE SOLAR INDUCED VENTILATION PROTOTYPE

The solar induced ventilation prototype investigated in this paper utilized a roof solar collector and a vertical stack (see Fig.1). The roof solar collector functioned as the solar radiation absorber and heater of the air inside the solar collector's cavity. The vertical stack functioned as a conventional chimney, without solar radiation collection. It provided a significant height in developing a sufficient stack pressure. The walls of the vertical stack were insulated all around in minimizing the heat gains and losses. This created air temperature differences and hence, pressure differences between the air inside the solar collector's cavity and the stack air. Consequently, the heated air inside the roof solar collector rose and flowed into the vertical stack. The inlet was located at the bottom of the roof solar collector, whilst the outlets were situated at the top of the vertical stack (Fig.1). There were obstructions placed at a distance of 0.15 m from the vertical stack outlets. These obstructions functioned as windbreakers that reduced the wind effects at the outlets. Without these obstructions, the induced air at the vertical stack outlets would be greatly affected by wind (Yusoff *et al.*, 2010).

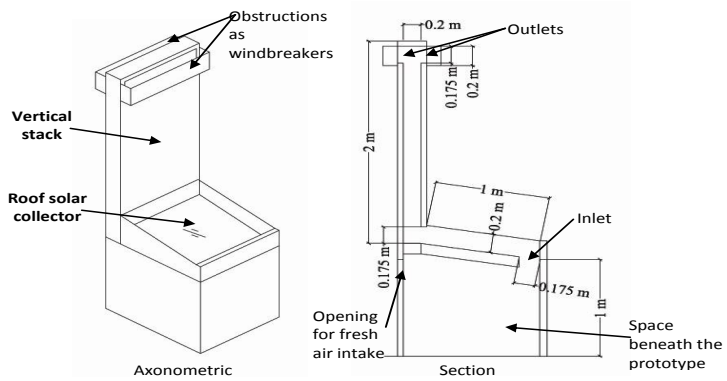



Fig.1: The prototype developed for the simulations

## METHODOLOGY

The methodology employed in this study was simulation modelling using Computational Fluid Dynamic (CFD) software FloVENT version 9.1. The feasibility of FloVENT for natural convection studies had been proven by previous research (see Manz, 2003; Nugroho, 2007). In this study, the parameters considered were the prototype model, the simulation set-up and procedures, the simulation conditions and the validation of simulation modelling.

### *Prototype Model*

The proposed prototype utilized a roof solar collector (RSC) and a vertical stack, as shown in Fig.1. The roof solar collector comprised of a glass cover, an air cavity and an absorber, while the vertical stack consisted of an air cavity and surrounded by insulated walls. The dimensions of the roof solar collector were 1 m x 1 m x 0.2 m (length x width x cavity width), whilst for the vertical stack the dimensions were 1 m x 2 m x 0.2 m (width x height x cavity width). There were obstructions made of plywood (1 m length x 0.2 m high), which were placed at a distance of 0.15 m from the vertical stack outlets.

The materials assigned to the prototype were a piece of black painted aluminium for the absorber, a clear glass for the top cover, and a plywood wall with insulations in the inner lining of the walls (Fig.2). The insulations were then covered with aluminium foil in reducing the radiative heat transfer between the walls' internal surfaces. The walls' external surfaces were also finished with aluminium foil to reduce the solar radiation absorption. The space beneath the prototype was fully covered to reduce the wind effects. However, openings were provided at one of the walls in allowing airflow into the space (Fig.1).

The orientation study utilized the prototype that had 2 meter high vertical stack. Meanwhile, for the height investigations, prototypes with three different vertical stack heights (2 m, 3 m and 4 m) were developed (Fig.3). The maximum height was 4 m, as the prototype was developed for application in a single storey industrial building; hence, it was deemed to be proportionate to the building typology applied. Other configurations which were the length and the tilt angle of the roof solar collector's channel, the cavity width and the opening areas (inlet and outlet)

were not investigated as they were derived from the literature review. In the literature, the recommended roof solar collector's channel length was 1 m, as no significant increment in the air temperature was obtained for the length greater than 1 m (Khedari *et al.*, 1997; Zhai *et al.*, 2005), and the suggested tilt angle for solar collector in Selangor, Malaysia was 10º. The cavity width of 0.2 m was utilized as the smaller the size (such as 0.1 m) would only increase the friction loss inside it (Bouchair, 1994; Miyazaki *et al.*, 2006), whereas a larger size would reduce the exit air temperature (Yousef, 2007). The inlet and outlet opening areas of the roof solar collector and vertical stack were of equivalent size, which was 0.1575 m² (0.9 m x 0.175 m). The purpose was to achieve the suggested effective ratio of 1:1 (inlet area:outlet area) for the opening areas (Khedari *et al.*, 2000; Susanti *et al.*, 2008).

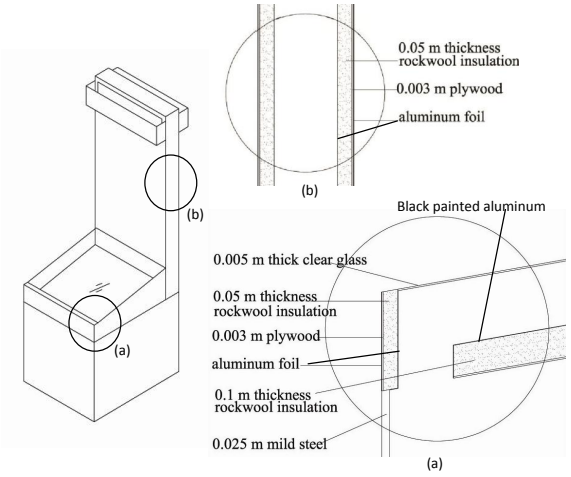



Fig.2: Materials of the prototype

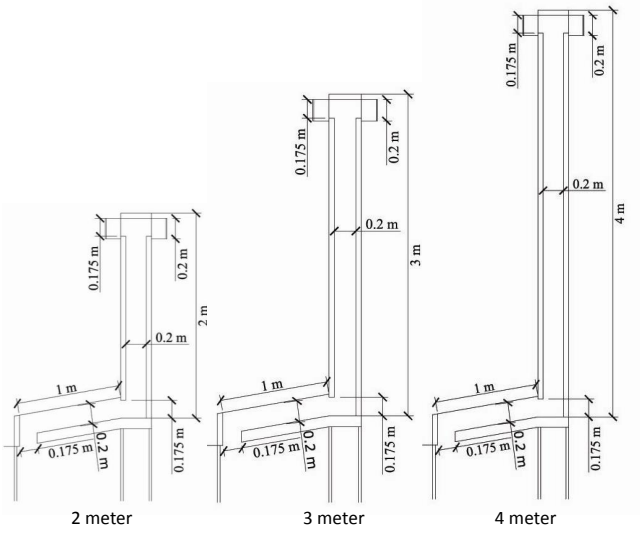



Fig.3: Prototypes with the different vertical stack heights

*General Simulation Set-Up and Procedures*

The prototype was placed in a solution domain which had a size of 7 m length x 7 m width x 6 m high. Its position was 3 m from x-plane, 3 m from z-plane and 0 m from y-plane (Fig.4). FloVENT version 9.1 provided three solution types, namely, the flow and heat transfer, the flow only and the conduction only. Since the investigations involved buoyancy induced ventilation, the solution type selected was the flow and heat transfer. Besides, the simulations also employed a steady state and three dimensional flow analyses. It also utilized Boussinesq approximation in solving the conservation equations. The flow regime was considered turbulence. There were three turbulence models available in FloVENT version 9.1, namely, the Capped LVEL, the LVEL Algebraic and the LVEL K-Epsilon. The simulations used the LVEL K-Epsilon turbulence model as it was the most appropriate for the built environment as compared to the others. This turbulence model utilized $k$-$\varepsilon$ (the turbulent kinetic energy ($k$) and turbulent dissipation rate ($\varepsilon$)) approach in the calculation (Gan, 2006). The fluid applied in the simulations was air at 30ºC, which had the following properties: conductivity was 0.02643 W/mK, viscosity was 1.872E-005 Ns/m², density was 1.149 kg/m³, and specific heat was 1007 J/kgK. The simulations used the Cartesian grid system. The outer iterations set for the simulations were 4000. Nevertheless, the simulation stopped automatically when it had attained convergence, even if the numbers set were not reached yet. Hence, the precise setting of outer iteration numbers was not critical (Mentor Graphics, 2010). The simulations also employed solar radiation calculation, in which the required input data were site latitude, simulation day, solar time and solar intensity. Meanwhile, the Atmospheric Boundary Layer (ABL) was applied in having the wind effects. The ABL was downloaded from the FloVENT website. It used Log Law model in generating the required wind profile. The north direction was set to be in the positive x direction.

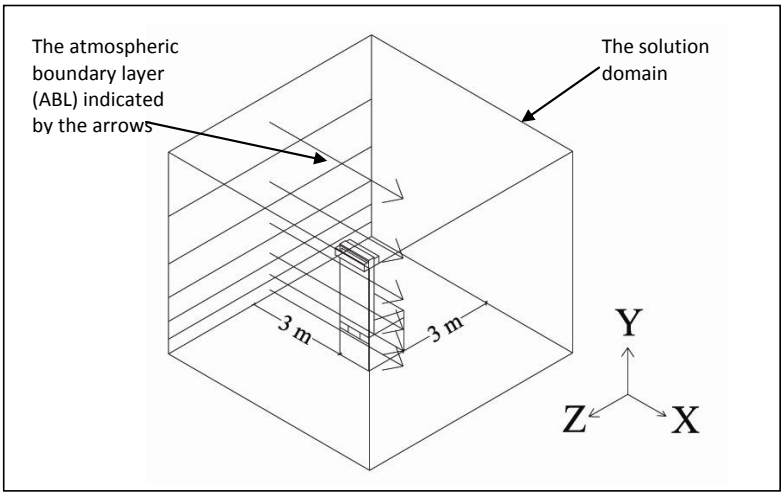



Fig.4: The position of prototype in the overall solution domain. Example is given for west orientation

*Simulation Set-Up and Procedures for Orientation Study*

The solar collector was oriented towards north, south, east and west. The effective orientation of solar collector was determined by the air temperature increment inside the RSC. This air temperature increment was calculated by the difference in the values between the monitor points at the RSC inlet and the vertical stack inlet. The locations of these monitor points are depicted in Fig.5.

*Simulation Set-Up and Procedures for Height Study*

In the investigations of vertical stack height, the prototype was oriented towards west, which was resulted from the orientation study. The variables investigated were air temperature and mass flow rate inside the vertical stack. Therefore, the monitor points were positioned at the inlet, middle and outlet of the vertical stack. The locations of monitor points are illustrated in Fig.5.

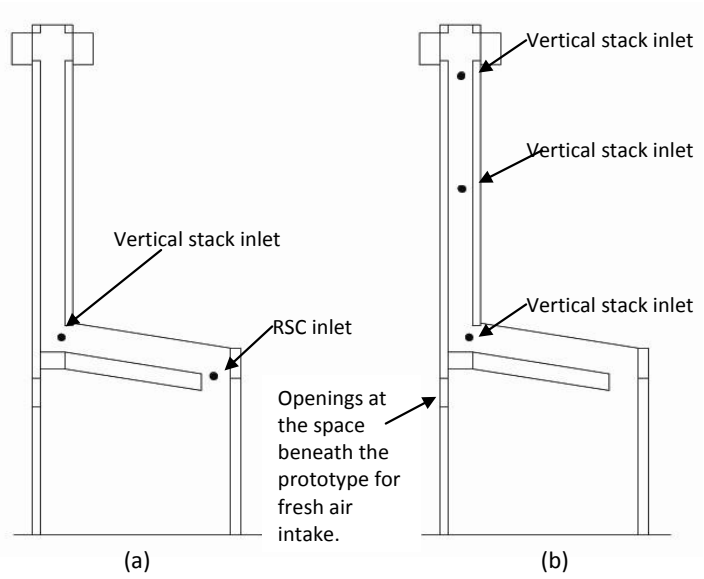



Fig.5: The location of monitor points at the prototype for solar collector orientation (a) and vertical stack height (b) studies

*General Simulation Conditions*

The initial boundary conditions employed the average of three years (2006-2008) environmental data from Subang weather station. The weather station is located at an airport formerly known as Sultan Abdul Aziz Shah (SAAS) airport, which is located in Selangor, Malaysia.

*Simulation Conditions for Orientation Study*

For the orientation study, the air temperature and solar radiation were the average data gathered on the 21 March, 22 June, 23 September and 21 December. These dates were selected as they represent the equinoxes and solstices. Meanwhile, the data utilized for wind speed and

wind direction were the annual average data. The purpose was to reduce the variability of the environmental parameters input data. The wind speed data of Subang weather station was corrected to a reference height of 1.6 m. The annual prevailing wind directions were north, northwest and south. However, the simulations utilized wind direction from south (180º) for ease of comparison. The selected times for solar collector orientation study were 10 a.m., 12 p.m. and 2 p.m. These times represented the morning, noon and evening hours.

*Simulation Conditions for Height Study*

The investigations of the vertical stack height employed the average data on the 21 March. March was selected as it was the month with the highest monthly mean solar radiation recorded for 20 years (1988-2008) by Subang weather station. Similar to orientation study, the wind speed data of Subang weather station was corrected to a reference height of 1.6 m. The prevailing wind direction in March was northwest. However, the simulations utilized annual prevailing wind direction which was south, as downward flow occurred inside the vertical stack for simulations using northwest wind direction. This downward flow hindered the investigations of height effects to the air temperature and mass flow rate inside the vertical stack. The selected times for the investigations were 12 p.m. and 2 p.m. only. No simulations were executed during morning hours because the west facing prototype had caused downward flow inside the vertical stack during this time.

*Validation of Simulation Modelling*

The simulation modelling was validated against physical experiment by Yusoff *et al.* (Mentor Graphics, 2010). Tables 1 and 2 depict the deviation percentage at each point's location from 9 a.m. to 4 p.m. The variables validated were air temperature and velocity. The air temperature was validated at six locations, namely, the RSC inlet, RSC 1, RSC 2, vertical stack inlet, vertical stack middle, and vertical stack outlet. Meanwhile, the air velocity was validated at three locations, which were the RSC inlet, vertical stack inlet and vertical stack outlet. The validation between the simulation modelling and physical experiment indicated a good agreement between the two results. The total average air temperature deviation for six

TABLE 1
The percentage of deviations for air temperature at six locations

| | Time (h) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 | Average deviation (%) |
| Monitor points' Locations | Deviation (%) | | | | | | | | |
| RSC inlet | 0 | 2 | 1 | 4 | 2 | 1 | 4 | 5 | 2 |
| RSC 1 | 12 | 15 | 11 | 7 | 12 | 11 | 9 | 6 | 10 |
| RSC 2 | 13 | 17 | 12 | 7 | 14 | 12 | 10 | 10 | 12 |
| V. stack inlet | 11 | 13 | 11 | 0 | 8 | 10 | 10 | 8 | 9 |
| V. stack middle | 5 | 9 | 6 | 2 | 8 | 8 | 7 | 11 | 7 |
| V. stack outlet | 2 | 7 | 4 | 2 | 3 | 2 | 0 | 6 | 3 |

TABLE 2
The percentage of deviations for air velocity at three locations

| Monitor points' Locations | Time (h) | | | | | | | | Average deviation (%) |
|---|---|---|---|---|---|---|---|---|---|
| | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 | |
| | Deviation (%) | | | | | | | | |
| RSC inlet | 28 | 12 | 10 | 18 | 18 | 15 | 4 | 13 | 15 |
| V. stack inlet | 21 | 7 | 1 | 2 | 17 | 14 | 15 | 5 | 10 |
| V. stack outlet | 10 | 5 | 16 | 17 | 18 | 17 | 20 | 11 | 14 |

locations was 7%, with maximum average deviation of 12% at RSC 2 (Table 1). Meanwhile, the total average deviation for air velocity at three locations was 13%, with the maximum average deviation of 15% at RSC inlet (Table 2).

## RESULTS AND DISCUSSION

The results are presented and discussed in two sub-sections, which are the effective solar collector orientation and the effects of vertical stack height to the air temperature and mass flow rate.

### *Effective Orientation of the Prototype's Solar Collector*

Fig.6 depicts the air temperature increments at three separate times, which were 10 a.m., 12 p.m. and 2 p.m., attained by the prototype for the particular orientations and on the particular dates. The results showed that the most effective solar collector orientation at 10 a.m. for all dates was east. At 2 p.m., the effective orientations on the 22 June and 21 December were west and south, respectively. Meanwhile, on the 21 March and 23 September, it could be either west or south. At 12 p.m., the effective solar collector orientations on the 22 June were east and west, whilst on the 21 December, it was south. However, the effective solar collector orientation at 12 pm on the 21 March and 23 September could be south, east or west.

Similarly, Fig.7 depicts the average air temperature increment at three separate times, namely, 10 a.m., 12 p.m. and 2 p.m. that was achieved by the prototype. The air temperature increments determined the orientation effectiveness, in which a high increment resulted in the effective orientation. The results indicated that the highest air temperature increment on the 22 June was attained by the west orientation. Meanwhile for the remaining three dates (namely, the 21 December, 21 March and 23 September), the highest air temperature increment was achieved by the south orientation. It is also apparent from Fig.7 that the most ineffective orientation for the prototype's solar collector was the north.

In summary, the suggested effective orientation for the prototype's solar collector throughout the year was west. However, it was emphasized that the west orientation might be effective for the proposed prototype only, as it had different configuration compared to the vertical solar chimney (Nugroho, 2007) and inclined solar collector (Bari, 2001; Gunerhan & Hepbasli, 2007; Shariah *et al.*, 2002), as described in the literature. The west orientation was recommended instead of the south because the air temperature increment in June for the south

orientation was very poor (Fig.7), especially at 12 p.m. (Fig.6). The sun path in June caused the vertical stack to shade the south facing solar collector, as shown in Fig.8. The prototype's solar collector was essential to attain a significant air temperature increment in June, as it was the hottest month compared to the others. The 20 years (1988-2008) monthly average weather data by Subang weather station indicated that the average ambient air temperature in June was 28ºC. It was the highest compared to December (26.8ºC), March (27.8ºC) and September (27.4ºC). Moreover, the monthly mean solar radiation received in June was also higher than in December.

Fig.7 also shows that the air temperature increments for the south orientation in December, March and September were in the average of 2ºC higher than the west orientation. In June, however, the air temperature increment of the west orientation was 3.95ºC higher than the south orientation. The results also indicated that the average air temperature increments at 12 p.m. and 2 p.m. of the south orientation was almost equal to the west orientation in March and September, 1ºC higher in December, but 7ºC lower in June (Fig.6). Meanwhile, the comparison between the east and west orientation showed that higher air temperature increment was attained by the west orientation in all months, except in September (Fig.7).

The drawback of orientating the prototype's solar collector to the west was the poor collection of solar radiation in the morning. This is due to the morning sun path which caused the vertical stack to shade the west facing solar collector, as shown in Fig.9. This phenomenon also resulted in no air temperature increment in the prototype's solar collector at 10 a.m. in March and September, as shown in Fig.6. Consequently, downward flow occurred inside the prototype.

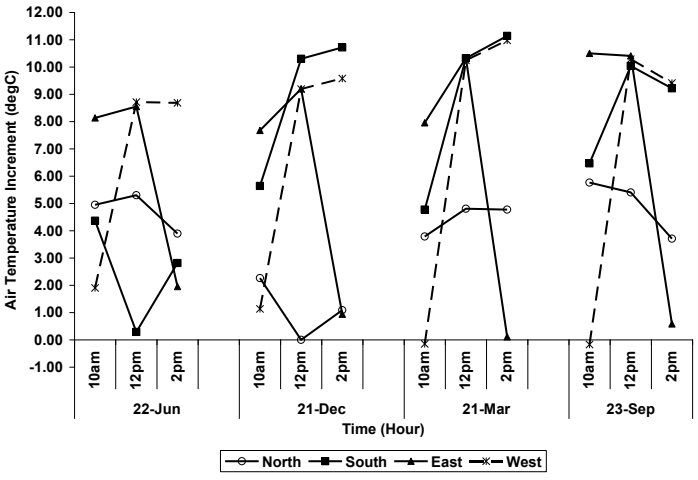



Fig.6: The air temperature increments at 10 a.m., 12 p.m. and 2 p.m. for the north, south, east and west orientations on the 22 June, 21 December, 21 March and 23 September

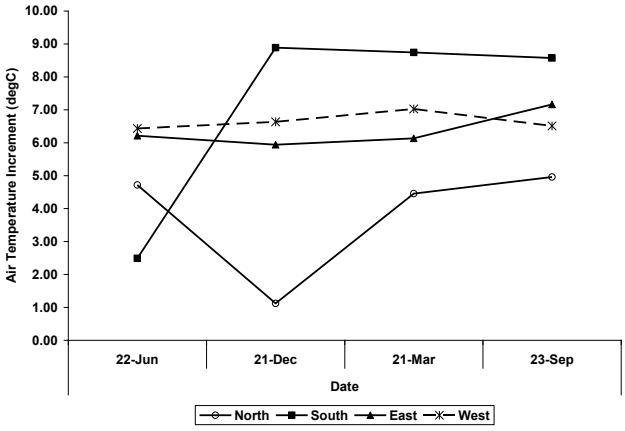



Fig.7: The average air temperature increments for the north, south, east and west orientations on the 22 June, 21 December, 21 March and 23 September

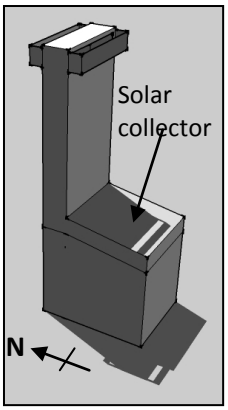



Fig.8: The south facing of the prototype's solar collector as at 12 p.m. on the 22 June

Although the west orientation resulted in a poor collection of solar radiation in the morning (Fig.6), the need for the prototype to induce more ventilation rate at noon and evening hours was more critical compared to the morning hours due to the higher indoor air temperature. Investigations into Malaysian single storey terrace houses showed that from the morning to noon hours, the indoor air temperature was almost equal to the outdoor air temperature. However, in the evening (i.e., from 3 p.m. to 7 p.m.), the indoor air temperature was 2 to 3ºC higher than the outdoor air temperature. Meanwhile, the PMV analysis of the master bedroom indicated that the PMV value was below +1.5 from 8 a.m. to 11 a.m., whereas from 1 p.m. to 3 p.m., the PMV value was +1.5 and above. The highest PMV value of +2 was achieved at 12 p.m. (Nugroho *et al.*, 2007).

In this research, comparisons of comfort condition between 10 a.m. and 2 p.m. were executed using the indoor predicted comfort temperature equations by Nicol and Humphreys (2010), as follows:

$$T_c = 13.5 + 0.54T_o \quad [1]$$

Where $T_c$ was the indoor predicted comfort temperature and $T_o$ was the monthly mean outdoor air temperature. The comparisons showed that the indoor predicted comfort temperature ($T_c$) was relatively closer to the monthly mean outdoor air temperature ($T_o$) at 10 a.m. compared to 2 p.m., except in September where the values were equal (Fig.10). In March, the monthly mean outdoor air temperature ($T_o$) was 0.6ºC lower than the indoor predicted comfort temperature ($T_c$) at 10 a.m. At 2 p.m., the monthly mean outdoor air temperature ($T_o$) was 1.2ºC higher than the indoor predicted comfort temperature ($T_c$). In June, the monthly mean outdoor air temperature ($T_o$) was higher than the indoor predicted comfort temperature ($T_c$) at both 10 a.m. (0.2ºC) and 12 p.m. (1ºC). Meanwhile in December, the comparisons showed that the monthly mean outdoor air temperature ($T_o$) was 0.3ºC lower at 10 a.m. and 0.9ºC higher at 2 p.m. than the indoor predicted comfort temperature ($T_c$).

*Height of the Vertical Stack*

It is apparent from the results obtained in Fig.11 that the higher the vertical stack was, the lower the air temperature inside the stack. The 2 m high vertical stack had the highest air temperature at all the monitor points. Meanwhile, the air temperature profiles were similar for all heights, in which the highest air temperature was attained at the vertical stack inlet. The air temperature decreased at the middle of the vertical stack and increased again when approaching the vertical stack outlet. The possible explanation to this phenomenon was due to the convective heat gain and loss of the air inside the prototype. The air at the vertical stack inlet had the highest temperature as it was heated inside the RSC. As the heated air rose and flowed upward inside the vertical stack, it experienced convective heat loss to the stack walls, which caused lower air temperature at the middle of the stack. However, at certain a height inside the vertical stack, the air temperature might be lower than the stack wall's temperature. Hence, the air experienced convective heat gain from the stack walls. This might explain the increase of air temperature from the middle to the outlet of the stack.

Fig.12 shows the mass flow rate at the inlet, middle and outlet of the 2 m, 3 m and 4 m high vertical stack. The mass flow rate results were in reverse to the air temperature results, in which the higher the vertical stack was, the greater the mass flow rate. However, there was no significant difference of the induced mass flow rate at the vertical stack inlet of all the heights compared to the middle and outlet of the stack. This finding showed that the stack pressure developed with the stack height. Another important finding was that the increase of mass flow rate from the inlet to the middle of the stack was higher than from the middle to the outlet. Moreover, at 12 p.m., there was a decrease in the mass flow rate from the middle to the outlet for the 2 m high vertical stack. This might be related to the air temperature inside the vertical stack. Besides the height, the stack pressure was also influenced by the air temperature. The higher the air temperature inside the stack a greater stack pressure is implied (2002). The air

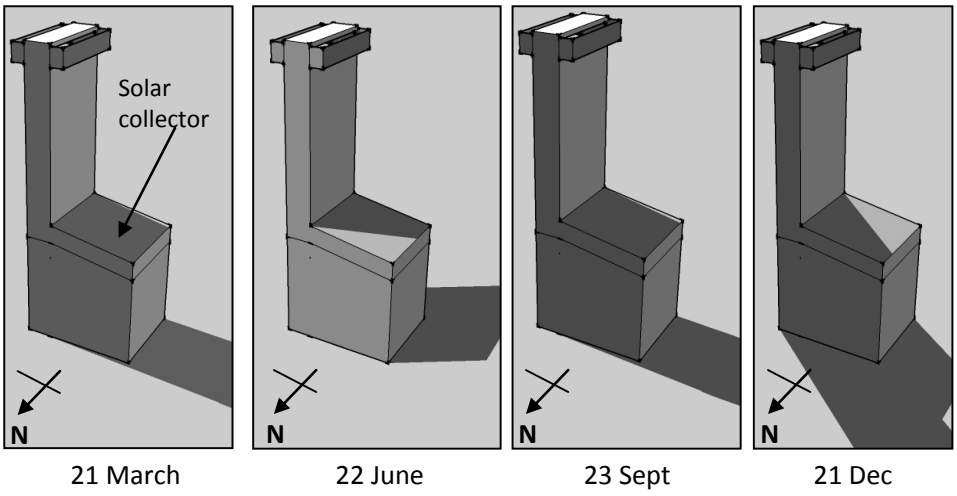



Fig.9: The west facing of the prototype's solar collector as at 10 a.m. on the 21 March, 22 June, 23 September and 21 December

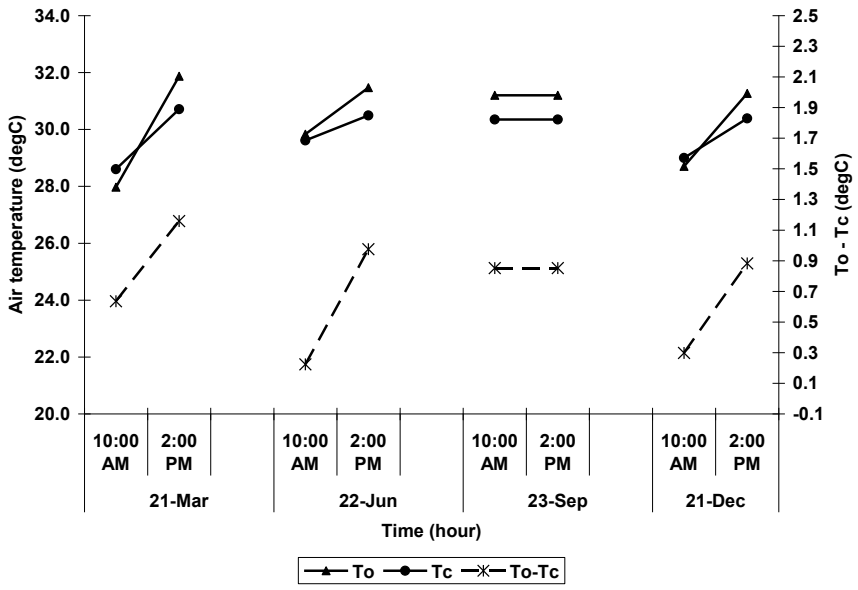



Fig.10: The comparison of comfort conditions between 10 a.m. and 2 p.m.

temperature at the inlet was the highest compared to the other part inside the vertical stack as it had just been heated inside the roof solar collector. As it flowed upwards, it experienced the heat gain and loss inside the stack. The reduction in the air temperature caused the lower increment of mass flow rate from the middle to the outlet compared from the inlet to the middle of the stack.

In summary, the higher the stack was, the lower the air temperature, but with a greater induced mass flow rate. Although the air temperature experienced more heat loss with the increase of stack height, the increase in stack pressure enabled the prototype to induce a higher mass flow rate.

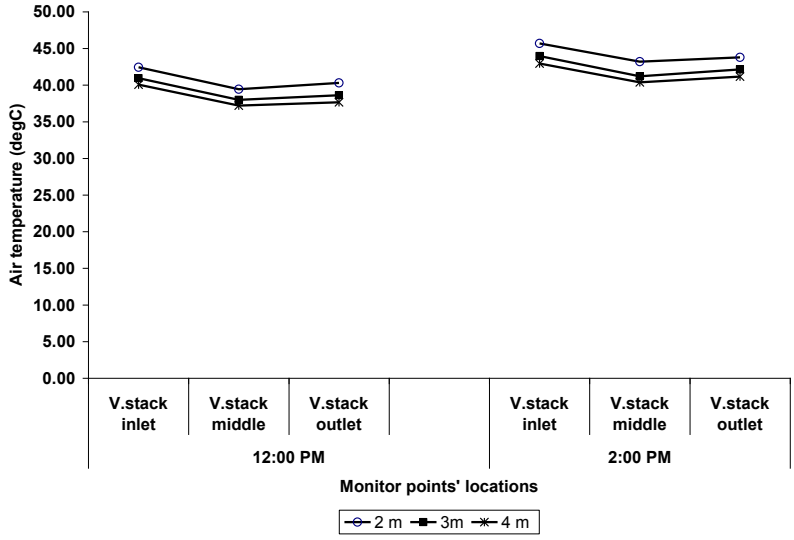



Fig.11: The air temperature inside the 2 m, 3 m and 4 m high vertical stack at 12 p.m. and 2 p.m.

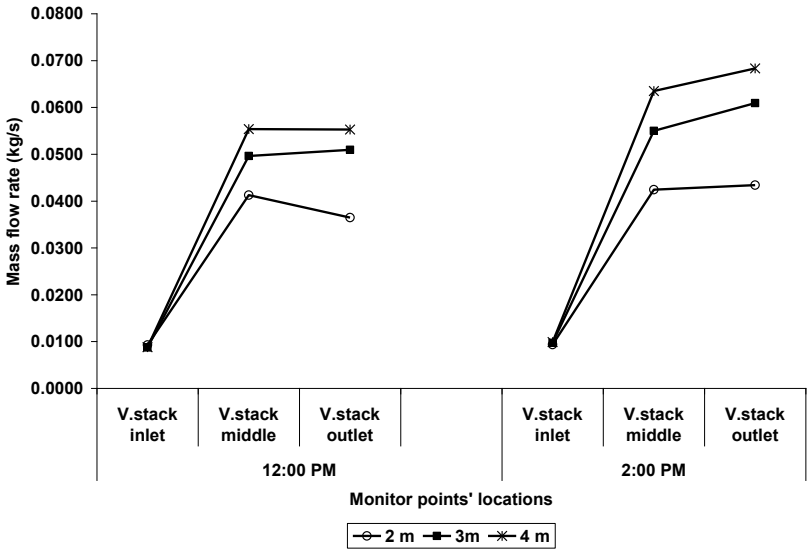



Fig.12: The mass flow rate inside the 2 m, 3 m and 4 m high vertical stack at 12 p.m. and 2 p.m.

## CONCLUSION

This study investigated the solar collector orientation and vertical stack height for solar induced ventilation that utilized roof solar collector and vertical stack using simulation modelling. The validation of simulation modelling against physical experiment had shown a positive agreement between the two results. The study concluded that the effective solar collector orientation throughout the year for the prototype in the northern hemisphere countries, particularly in Malaysia, is the west. Hence, the suggested orientation is not in conformity with the findings from the literatures, which recommended south as the effective solar collector orientation throughout the year for the northern hemisphere countries (Bari, 2001; Gunerhan & Hepbasli, 2007; Nugroho, 2007; Shariah *et al.,* 2002). The different finding is due to the different configurations of the prototype, in which the vertical stack has caused shading to the south facing solar collector in June. The study also concluded that the higher the vertical stack is, the lower the air temperature, but the greater the induced mass flow rate. In summary, the investigations confirmed the potential of the proposed solar induced ventilation for the application in the hot and humid climate. Therefore, it would be interesting to examine the effects of the proposed solar induced ventilation to the indoor air temperature and ventilation rates for buildings in future research undertakings.

## ACKNOWLEDGEMENTS

The authors would like to thank the International Islamic University Malaysia for the endowment of research grant (EDW B 1001-345) that enabled the purchasing of the CFD software.

## REFERENCES

Awbi, H. B. (2003). *Ventilation of Buildings* (2nd Edition). London: Spon Press.

Bari, S. (2001). Optimum orientation of domestic solar water heaters for the low latitude countries. *Energy Conversion and Management, 42*(10), 1205-1214.

Bouchair, A. (1994). Solar chimney for promoting cooling ventilation in Southern Algeria. *Building Services Engineering Research and Technology, 15*(2), 81-93.

Chen, Z. D., Bandopadhayay, P., Halldorsson, J., Byrjalsen, C., Heiselberg, P., & Li, Y. (2003). An experimental investigation of a solar chimney model with uniform wall heat flux. *Building and Environment, 38*(7), 893-906.

Ding, W., Hasemi, Y., & Yamada, T. (2005). Natural ventilation performance of a double-skin façade with a solar chimney. *Energy and Buildings, 37*(4), 411-418.

Gan, G. (2006). Simulation of buoyancy-induced flow in open cavities for natural ventilation. *Energy and Buildings, 38*(5), 410-420.

Gunerhan, H., & Hepbasli, A. (2007). Determination of the optimum tilt angle of solar collectors for building applications. *Building and Environment, 42*(2), 779-783.

Hamdy, I. F., & Fikry, M. A. (1998). Passive solar ventilation. *Renewable Energy, 14*(1-4), 381-386.

Harris, D. J., & Helwig, N. (2007). Solar chimney and building ventilation. *Applied Energy, 84*(2), 135-146.

Khedari, J., Hirunlabh, J., & Bunnag, T. (1997). Experimental study of a roof solar collector towards the natural ventilation of new houses. *Energy and Buildings, 26*(2), 159-164.

Khedari, J., Mansirisub, W., Chaima, S., Pratinthong, N., & Hirunlabh, J. (2000). Field measurements of performance of roof solar collector. *Energy and Buildings, 31*(3), 171-178.

Li, A., Jones, P., Zhao, P., & Wang, L. (2004). Heat transfer and natural ventilation airflow rates from single-sided heated solar chimney for buildings. *Asian Architecture and Building Engineering, 3*(2), 233-238.

Manz, H. (2003). Numerical simulation of heat transfer by natural convection in cavities of facade elements. *Energy and Buildings, 35*(3), 305-311.

Mathur, J., Mathur, S., & Anupma (2006). Summer-performance of inclined roof solar chimney for natural ventilation. *Energy and Buildings, 38*(10), 1156-1163.

Mentor Graphics (2010). *FloVENT® User Guide-Software Version 9.1*.

Miyazaki, T., Akisawa, A., & Kashiwagi, T. (2006). The effects of solar chimneys on thermal load mitigation of office buildings under the Japanese climate. *Renewable Energy, 31*(7), 987-1010.

Nicol, J. F., & Humphreys, M. A. (2002). Adaptive thermal comfort and sustainable thermal standards for buildings. *Energy and Buildings, 34*(6), 563-572.

Nugroho, A. M. (2007). *Solar Chimney Geometry for Stack Ventilation in Malaysia Terrace House.* (Doctoral dissertation). Universiti Teknologi Malaysia, Johor.

Nugroho, A. M., Hamdan, M., & Ossen, D. R. (2007). A preliminary study of thermal comfort in Malaysia's single storey terraced houses. *Asian Architecture and Building Engineering, 182*, 289-296.

Shariah, A., Al-Akhras, M. A., & Al-Omari, I. A. (2002). Optimizing the tilt angle of solar collectors. *Renewable Energy, 26*(4), 587-598.

Susanti, L., Homma, H., Matsumoto, H., Suzuki, Y., & Shimizu, M. (2008). A laboratory experiment on natural ventilation through a roof cavity for reduction of solar heat gain. *Energy and Buildings, 40*(12), 2196-2206.

Yousef, B. A.-R. A. A. A. (2007). *Development of a Mathematical Model to Predict Thermal Performance and Cost Effectiveness of Solar Air Heaters.* Doctoral dissertation, Universiti Putra Malaysia, Selangor.

Yusoff, W. F. M., Salleh, E., Adam, N. M., Sapian, A. R., & Yusof Sulaiman, M. (2010). Enhancement of stack ventilation in hot and humid climate using a combination of roof solar collector and vertical stack. *Building and Environment, 45*(10), 2296-2308.

Zhai, X. Q., Dai, Y. J., & Wang, R. Z. (2005). Experimental investigation on air heating and natural ventilation of a solar air collector. *Energy and Buildings, 37*(4), 373-381.

Pertanika J. Sci. & Technol. 22 (1): 289 - 305 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Full-Scaled Impulse Turbine Performance Prediction using Numerical Simulation

**A. Sahed* and F. B. Ismail Alnaimi**

*Mechanical Department, Curtin University, Sarawak Campus CDT 250, 98009 Miri, Sarawak, Malaysia*

## ABSTRACT

This paper presents the numerical modelling techniques for the simulation of the energy conversion chain from wave to electricity in an Oscillating Water Column (OWC) equipped with a full-scaled self rectifying turbine. The performance of the OWC device has been assessed for the stand-alone power system of a typical Irish climatic wave condition. The results showed that the overall performance of the complete device depends on the level of turbine damping, which in turn depends on the wave climate, especially the significant wave height. Furthermore, turbine efficiency predicted using quasi-steady showed qualitatively favourable agreement with the experimental results.

*Keywords:* Wave energy, applied damping, numerical simulation, irregular, unsteady, real sea conditions

## INTRODUCTION

The OWC device wave energy harnessing method is considered as one of the best techniques for converting wave energy into electricity. It is an economically viable design due to its simple geometrical construction, and it is also strong enough to withstand against the waves with different heights, periods and directions. The design (see Fig.1) consists of an OWC chamber and a circular duct, which reciprocally move the air from and into the chamber as the wave enters and intercedes from the chamber. The wave energy is converted into air pneumatic energy inside the chamber. A self rectifying impulse turbine mounted inside the duct was designed to turn in one direction only although the airflow moves bi-directionally. This converts the air pneumatic energy to a mechanical power. A matching generator is coupled to the turbine to produce electricity (de O. Falca~o, 2003).

The performance prediction of the impulse turbine under real sea conditions is very important for designers to get a feel for the behaviour of the power converting device under unsteady, irregular flow conditions. This present study deals with the performance

*Article history:*
Received: 26 December 2011
Accepted: 15 March 2012

*E-mail addresses:*
amazighihi@gmail.com (A. Sahed),
fbalnaimi@curtin.edu.my (F. B. Ismail Alnaimi)
*Corresponding Author

ISSN: 0128-7680 

prediction of 1.6m diameter impulse turbine with fixed guide vanes under these conditions using numerical simulation. The small scale (0.6m diameter) turbine performance has already been investigated by a few authors (Thakker *et al.,* 2001, 2004) and the numerical simulation technique has also been validated.

The numerical simulation procedure used in this analysis is similar to that used in [3] who investigated the performance of a small scale turbine for various fixed speeds during a time frame of around 260s operating condition. As a next step, the full-scaled turbine performance was analyzed numerically under simulated irregular and unsteady conditions based on real sea data, taking into account the effects due to applied damping provided by the turbine for a period of a year. The effect of the turbine rotational speed on the device efficiency was investigated. The turbine power output was also investigated for various rotational speeds.

Apart from the obvious function of quantifying the performance of turbines and providing a basis for comparing performance, experimental data and the associated dimensional analysis allow the designer to size a turbine to a given wave power application. In contrast to the general design problem familiar in the industry, where a turbine is expected to operate at a single design point for the majority of the time, designing a turbine to a wave power application requires that a design range be taken into account, rather than a single design point (Thakker & Fergal, 2004).

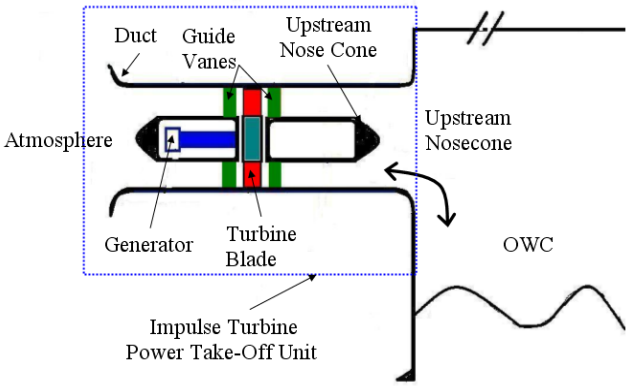



Fig.1: Impulse turbine power take-off with OWC

## EXPERIMENTAL AND BACKGROUND

### *Turbine Design and Manufacture*

The basic turbine design parameters were based on the optimum design parameters given by Setoguchi *et al.* [5], but with an H/T ratio of 0.6 [6]. The details, for both 0.6*m* and 1.6*m*, are given in Table 1 and a 2D sketch at the mid radius is shown in Figure 2. The turbine consists of 30 blades and 26 fixed angle mirror image guide vanes on both sides of the rotor. The guide vanes inlet/outlet angle is fixed at 30°. The next step is the design analysis of the next generation impulse turbine for which a preliminary diameter of 1.6*m* was established, and G.F.R.P was selected as the optimum material to manufacture the blades (Thakker *et al*., 2008).

*Experimental Analysis*

The overall performance of the turbines was evaluated by the turbine angular velocity, $\omega$, torque generated, $T$, flow rate, $Q$, and total pressure drop, $\Delta P$ across the rotor. The results were expressed in the forms of torque coefficient, $C_T$, input power coefficient, $C_A$ and efficiency, $\eta$ in terms of flow coefficient, $\varphi$, which are defined by Setogushi *et al.,* 2006. and the test Reynolds number based on the chord was $0.4\times10^4$.

$$C_T = \frac{T}{\frac{1}{2}\rho a(\upsilon a^2 + U_R{}^2)bZL_r r_R} \tag{1}$$

$$C_A = \frac{\Delta PQ}{\frac{1}{2}\rho a(\upsilon a^2 + U_R{}^2)bZL_r \upsilon_a} \tag{2}$$

$$\phi = \frac{\upsilon_a}{U_R} \tag{3}$$

$$R_e = \frac{\rho a\sqrt{\upsilon a^2 + U_R{}^2 L_r}}{\mu} \tag{4}$$

$$\eta = \frac{C_T}{C_A\ \phi} \tag{5}$$

The 0.6*m* turbine was tested for different constant axial velocities between a rotational speed range of 1200*rpm* and 100*rpm*, thus, giving a flow coefficient range between 0.15 and 3.58 (Ryan, 2005). The lowest Reynolds number achieved at peak efficiency was $0.46\times10^5$ at the axial velocity of 4.91*m=s* and the highest *Re* was 1.23 105 at $v_a = 12.6m/s$, respectively. These upper and lower limits were defined by the limitations of the test rig; the experiments were carried out up to a flow rate of $2.11m^3/s$ and the instruments used for data collection. It is believed that there are no prominent effects of Reynolds number on this type of turbine and constant performance can be achieved over a wide range of Reynolds number ($0.46\times10^5$ to $1.23\times10^5$), as the mean efficiency is within ±2.5% of the maximum and the minimum efficiencies achieved within this range of Reynolds number. Furthermore, the critical Reynolds number found experimentally for this turbine was $0.65\times10^5$ (Ryan, 2005; Hammad, n.d). In general, it can be said that this type of impulse turbine has no apparent effects due to the change in the Reynolds number as the performance was found to be stable over a wide range of Re.

TABLE 1
Rotor and guide vane parameters

| Parameter | Symbol | 0.6∅ | 1.6∅ |
|---|---|---|---|
| Blade profile: elliptical | | | |
| Hub-to-tip ratio | HT | 0.6 | 0.6 |
| Number of blades | z | 30 | 30 |
| Chord length (mm) | $l_r$ | 100 | 268 |
| Pitch (mm) | $S_r$ | 50 | 134 |
| Blade inlet angle | $\gamma$ | 60° | 60° |
| Guide vanes profile: plate type | | | |
| Pitch (mm) | $S_g$ | 58 | 154 |
| Chord length (mm) | $l_g$ | 131 | 352 |
| Number of guide vanes | g | 26 | 26 |
| Guide vane inlet/outlet angle | $\gamma$ | 30° | 30° |

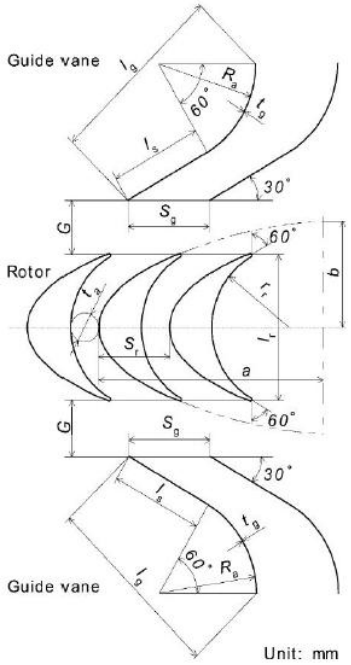



Fig.2: Rotor blade and guide vanes geometry

Therefore, the performance for the large diameter impulse turbines for actual wave energy power plants, running at higher Reynolds number can be predicted using the experimental data from model testing (Hammad, n.d; Setoguchi *et al*., 2004).

## NUMERICAL PROCEDURE

In order to analyze the overall performance of the OWC device fitted with a full-scaled impulse turbine, iterations were carried out for various turbine rotational speeds between 60 and 900*rev/min*.

Considering the quasi-steady flow conditions, typical turbine characteristics arrived at from the uni-directional experimental data for the said turbine were used for this simulation as using such uni-directional flow data in the actual quasi-steady conditions has been validated by previous studies on the Wells turbine (Inoue *et al*., 1998) and a study on the impulse turbine (Setoguchi *et al.,* 2006). A simple geometry of the OWC (Setoguchi *et al.,* 2006) was considered for the simulation with a turbine duct area to air chamber area ratio, $m = 0.0234$ (see Fig.3). The techniques and performance parameters used by Inoue (1998) and Setoguchi *et al.* (2006) for the numerical simulation of 0:3m diameter Wells and impulse turbines under irregular flow conditions were adopted by Thakker *et al.* (2004). The latter analysis simulated the 0:6*m* impulse turbine using the existing equations with the damping term included so that it completed the general equation of motuion for column of water. It should be noted here that the calculation of OWC efficiency has been simplified to show how the damping applied by the turbine affects the overall efficiency of the device and is not reacting the true behaviour of an actual OWC. As

such, the simulation technique used does not take into account the effects due to the secondary damping which further consists of two components, namely, radiation damping due to waves by the column and loss damping due to energy losses in the system (Curran *et al*., 2000). This simplified technique was used to simulate irregular, unsteady flow conditions for the turbine, based on the sea wave data. The equations were solved using the standard numerical techniques using Runge-Kutta Fehlberg Method in Matlab (Thakker *et al.,* 2004).

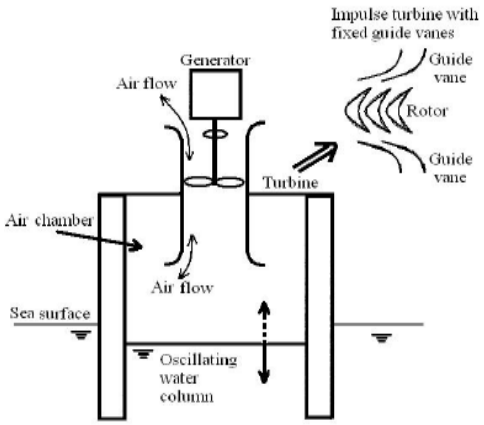



Fig.3: A schematic view of the OWC device used in the numerical simulation

*Efficiency of OWC*

The relationship between the incident wave height (H) and the wave height (h) in the air chamber is given by Setoguchi *et al.* (2006) as:

$$\frac{d}{dt}(\rho_s h A_C \frac{dh}{dt}) = \{\rho_s g (H-h) - \Delta p\} A_C \tag{6}$$

This is an approximate equation of the motion of OWC because the equation of motion about OWC is generally expressed by using linear water wave theory (Chatry *et al*., 1999).

Eq. 6, which basically originates from Newton's second law of motion i.e. $F = m\,a$, was modified to include the damping applied by the turbine. The standard equation of motion for a mass-spring-damper system is given in the basic form as follows:

$$\frac{d}{dt}\{m\dot{X}\} + C\{\dot{X}\} + K\{X\} = F(t) \tag{7}$$

where

$F(t) = \{\rho_s g (H-h) - \Delta p\} A_C$ oscillating force on water column surface

$m = \rho_s h A_C = M_E$ effective mass of water column

$C = B_A$ damping

$K = \rho_s g A_C$ the buoyancy of the column

$X = h$ displacement of water surface.

Considering the system analogous to the mechanical vibration system with single degree of freedom with damper, the equation of motion (Eq. 6) can be modified to include the damping applied by the turbine. This results in the following equation:

$$M_E = \frac{d^2h}{dt_2} + B_A \frac{dh}{dt} + K\ h = F(t) \tag{8}$$

The damping applied on the OWC by the turbine, $B_A$, is given by:

$$B_A \frac{F(t)}{\upsilon_c} = \Delta p \frac{A_C}{\upsilon_c} = \Delta p \frac{{A_C}^2}{A_T\ \upsilon_a} \tag{9}$$

Given the flow coefficient,

$$\phi = \frac{\upsilon_a}{U_R} \tag{10}$$

which is the ratio of axial velocity to tangential blade velocity at mid-span ($r_R$), and

$$p^* = \frac{\Delta p}{\rho_a\ \omega^2 (2r_R)^2} \tag{11}$$

which is the definition of the pressure coefficient (Thakker & Fergal, 2004), substituting in $B_A$ for $p^*$ and φ gives:

$$B_A = 4\rho_a \frac{{A_C}^2}{A_T} U_R \left( \frac{p^*}{\phi} \right) \tag{12}$$

It has been reported that the optimum damping is a function of wave period, incident wave power, and that no single damping level will optimize the performance of the column in all seas (Curran *et al*., 1995). Substituting the values of $M_E$, $B_A$, $K$, $X$ and $F(t)$ in Eq. 8 and simplifying, we will get:

$$\frac{d}{dt}(h\frac{dh}{dt}) + 4\rho_a \frac{A_C}{A_T} U_R \left( \frac{p^*}{\phi} \right) \frac{dh}{dt} + \frac{\Delta p}{\rho_s} + g(2h - H) = 0 \tag{13}$$

The above equation was solved using Runge-Kutta Fehlberg algorithm Matlab. The turbine performance characteristics were taken from experimental data under steady conditions. Generally, it is believed that the characteristics of the irregular motion of the OWC are different from those of the incident waves due to the energy absorbing characteristics, losses due to

radiation, and other effects due to the random nature of the waves. As the main purpose of this simulation is to evaluate turbine performance under given irregular conditions, the method was simplified and only damping applied by the turbine was taken into account, while other factors were not taken in consideration here.

The conversion efficiency of OWC can be defined as the ratio of incident wave power and the power of OWC. It is important to note here that the efficiency of the air chamber is also dependent on the compressibility effects within the chamber (Thakker *et al*., 2003). However, as the objective of this analysis is to evaluate the performance of the turbine under real sea conditions, the efficiency of the air chamber in this case can therefore be obtained by the ratio of incident wave power and the power of OWC (Setoguchi *et al*., 2006). The incident wave power is defined as the power in watts delivered by each meter width of a wave (Duckers, 1996) and is given by:

$$W = \frac{\rho_s g^2 H^2 T}{32\pi} \tag{14}$$

Therefore, the incident wave power $\overline{W_i}$ and power of OWC $\overline{W_o}$ are defined as:

$$\overline{W_i} = \frac{\sum_{i=1}^{N} \frac{1}{32\pi} \rho_s g^2 H_i^{\,2} T_i^2}{\sum_{i=1}^{N} T_i} \tag{15}$$

and

$$\overline{W_o} = \frac{\sum_{i=1}^{N} \frac{1}{32\pi} \rho_s g h_i^{\,2} T_i^2}{\sum_{i=1}^{N} T_i} \tag{16}$$

Thus, the efficiency of air chamber can be defined as:

$$\eta_c = \frac{\overline{W_o}}{\overline{W_i}} \tag{17}$$

*Efficiency of the Turbine*

If compressibility effects are neglected, the air flow rate can be obtained by calculating the volume of water displaced above (or under) the mean water level (MWL) per unit time. Consequently, the axial flow velocity is directly proportional to a variation of the wave height *h*. Therefore, the non-dimensional axial flow velocity through the turbine $v_a$* is written as:

$$\upsilon_a * = \frac{d(h / H_s)}{d(t / \overline{T})} = \frac{dh*}{dt*} \tag{18}$$

The running characteristics of the turbine under irregular, unsteady flow conditions were calculated through numerical simulation. The steady flow characteristics of the turbine are assumed to be valid for computing performance under unsteady flow conditions. Such a quasi-steady analysis has been validated in some previous studies for both Wells turbine (Inoue *et al*., 1998) and impulse turbine (Setoguchi *et al*., 2006). The running characteristics are obtained by keeping rotational speed constant. In this case, the mean output $\overline{C_o}$ and input coefficient $\overline{C_i}$ can be defined respectively as:

$$\overline{C_o} = \frac{1}{t^*}\int_0^{t^*} C_T(\phi)\frac{\left(k\omega^*\right)^2+\left(\upsilon_a^*\right)^2}{2}\sigma_r\frac{4(1-v)}{(1+v)}\omega^* dt^* \tag{19}$$

$$\overline{C_i} = \frac{1}{t^*}\int_0^{t^*} C_A(\phi)\frac{\left(k\omega^*\right)^2+\left(\upsilon_a^*\right)^2}{2k}\sigma_r\frac{4(1-v)}{(1+v)}\upsilon_a^* dt^* \tag{20}$$

It can be noted from Eq. 14 and 15 that the performance of the turbine can be calculated as a function of $k\omega^*$ and $\upsilon_a^*$, when torque coefficient, $C_T(\varphi)$, input coefficient, $C_A(\varphi)$, solidity, $\sigma_{rR}$, hub to tip ratio, $\upsilon$ and non-dimensional angular speed $\omega^*$ are specified.

The mean efficiency of the turbine can be defined as:

$$\eta_t = \frac{\overline{C_o}}{\overline{C_i}} \tag{21}$$

Therefore, the mean conversion efficiency for the OWC device at a given constant rotational speed under irregular, unsteady flow conditions can be calculated as follows:

$$\eta = \eta_t \bullet \eta_c \tag{22}$$

## METHODOLOGY

In the above section, it was mentioned that the turbine performance under irregular conditions depends on a unique parameter $k\overline{\omega^*}$. This parameter includes the characteristic parameters of the irregular wave ($H_s$ & $\overline{T}$), turbine speed ($\omega$) and the dimensions of the turbine and air chamber ($r_R$ & $m$). Therefore, once an optimum value of $k\overline{\omega^*}$ is determined by numerical simulation for a given test wave and dimensionless turbine characteristics, the combination of optimum design values (m, $r_R$ and $\omega$) can be obtained for a site where significant wave height, *Hs* an d mean time period $\overline{T}$ are known (Setoguchi *et al*., 2006).

The reciprocal $1/k\ \omega^*$ represents the flow coefficient for ordinary fluid machines. To calculate different $1/k\ \omega^*$ values, assuming quasi-steady flow conditions uni-directional steady flow experimental data at constant axial velocity was used to generate turbine performance data for different constant rotational speeds. The $C_T(\varphi$ *and* $C_A(\varphi)$ characteristics trends from experimental results were used to generate the turbine performance data for different constant rotational speeds. The turbine performance data were generated for different rotational speeds

between *900 rpm* to *60 rpm* in order to get a wide range of flow coefficient through the turbine.

After generating the performance data, the incident wave height was converted into wave height in the chamber for each constant rotational speed using Runge-Kuta-Fehlberg algorithm. Following this conversion, the wave height within the chamber was translated into the volume flow and non-dimensional axial velocity, $v_a$*, over the turbine. The variation of the damping ($B_A$) was calculated using instantaneous flow coefficient and pressure drop through the turbine. Subsequent calculations lead to the mean turbine output and input characteristics (Eq. 19 and Eq. 20, respectively). The mean efficiency of the turbine was achieved using Eq. 21. Similarly, the efficiency of the OWC was also calculated for each individual $1/k\ \omega$* by using Eq. 15 and Eq. 16. Finally, the overall mean efficiency of the device was achieved by using Eq. 17 for each constant rotational speed.

### *Application To Typical Irish Sea State*

The numerical simulation technique was used in this work to evaluate the energetic performance of an OWC device with reference to the sea state typical of an Irish location. The study was carried out using actual sea data based on the water surface elevation time history (Thakker *et al*., 2008) measured during the years 2002-2003. It was referred as wave site-2, which has an overall mean time period, $\overline{T}$ = 9.2s, and the mean significant height $H_s$ = 3.1m. The wave site-2 data had been initially assessed for the reference year and 9 significant sea states were found (summarized in Table 2).

TABLE 2
The nine sea states for the climatic Irish wave site-2

| Sea states | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $H_s(m)$ | 0.8 | 2 | 2.9 | 3.8 | 4.6 | 5.2 | 6.6 | 7.4 | 8.9 |
| $\overline{T}(s)$ | 7.4 | 8.8 | 9.7 | 10.3 | 11.7 | 12.0 | 12.7 | 13.2 | 13.5 |
| $Z(\%)$ | 12.2 | 31.0 | 26.6 | 14.8 | 7.1 | 4.0 | 2.0 | 1.5 | 0.3 |

The wave frequency is defined in terms of its energetic period, $\overline{T}$. In Table 2, $Z$ is the sea state frequency of occurrence in a year. Moreover, the input data used are shown as a diagrammatic view of both significant wave height ($m$) and wave period ($s$) versus time. In Fig.4 and Fig.5, the seasonal trend of wave data is shown for a typical winter and summer conditions, respectively, in terms of hourly distribution of $H_s$ & $\overline{T}$. It is remarkable that the Irish sea conditions feature almost unchanged wave period distributions in terms of average values, but the winter significant height is usually higher. With respect to the reference year, the annual average level of the power released on the west coast of Ireland is about 35÷ 40 *kw/m.*

As far as the input quantities (significant height and mean period) are concerned, they are given to the numerical technique Matlab programme with an hour time step during a year, from 1.00 am of September 1st to 12.00 pm of August 31st. In the following section, the January results are shown.

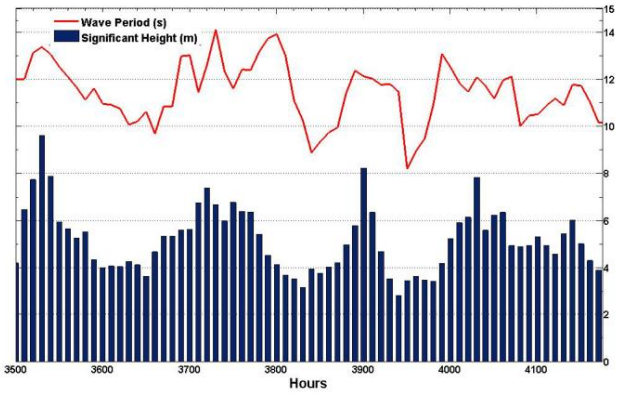



Fig.4: Wave input distributions in January 2002

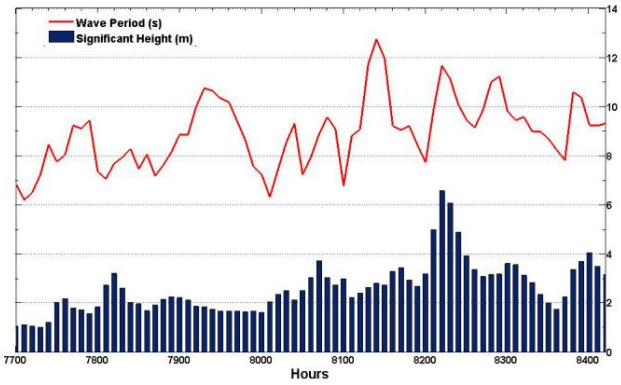



Fig.5: Wave input distributions in July 2002

*Damping Applied by the Turbine*

Fig.6 demonstrates the incident wave height (*H*) and wave height inside the chamber (*h*) plotted against time (*Hours*) for different turbine speeds (125, 200, 300, 400, 600, 700*rev/min*). It can be seen that there is a far greater variance in the scatter of the values of h at higher wave incidences. It can also be observed that as *H* decreases, this trend tends to die out and the values of h tend to converge towards a single value and closer to *H*. However, as *H* increases, it can be seen that the energy loss in converting to *h* increases dramatically. This conclusion is reached as the percentage of *h* retained after the conversion from H decreases accordingly as *H* increases.

Fig.6 also demonstrates the changing effect the turbine rotational speed has on the amount of damping applied by the turbine and hence the generated flow coefficient. It can be seen that as the turbine speed increases, the applied damping increases and the flow coefficient range of the OWC decreases.

Fig.7 shows the mean converted power incident wave to pneumatic, hourly integrated. The simulated system pneumatic power has a trend similar to the incoming wave conditions. It is worth noting that its values are irregular because of the large variability of the incident waves. This confirms that the output power from the OWC is higher for higher waves and vice

versa. This also resembles the typical behaviour of the turbine, as shown during uni-directional steady and unsteady flow conditions in terms of air axial velocity. Furthermore, it is noticed from Fig.7 that local maxima related to a particular speed exist at different levels of incident wave power. The OWC device converts more power when the turbine is running at 300rpm than any other speed when higher wave power is available. This shows that the damping applied by the turbine is optimum at 300rpm for such higher wave power. At a lower incident power, the OWC device converts more power at a higher turbine rotational speed. This is a clear indication that a higher level of damping does not indicate a lower performance of the OWC device; on the contrary, this could be optimum for a particular significant wave height.

In order to give more hints on the performance of the OWC device for different turbine speeds (200, 300, 400, 600*rev/min*), it is convenient to analyze the shaft torque output diagram. Fig.8 demonstrates the marked conversion discontinuities owing to the non-deterministic wave energy character. It can be seen that there is a far greater variance in the scatter of the values of turbine torque at the higher wave incident. It can also be observed that accordingly, as the incident power decreases, this trend tends to die out and the values of the turbine torque tend to converge towards a single value.

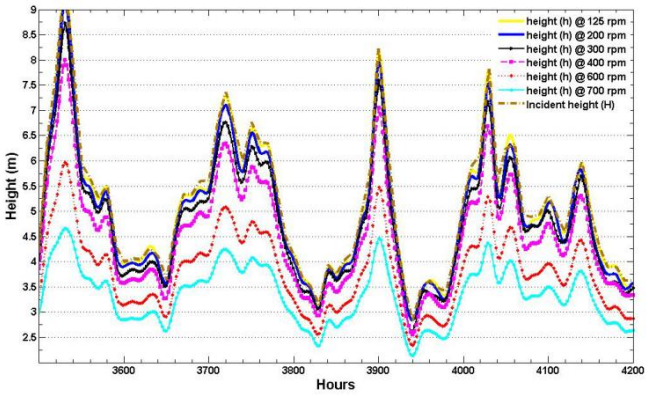



Fig.6: Incidence height (H) and height inside the chamber (h) versus time at different rotational speeds

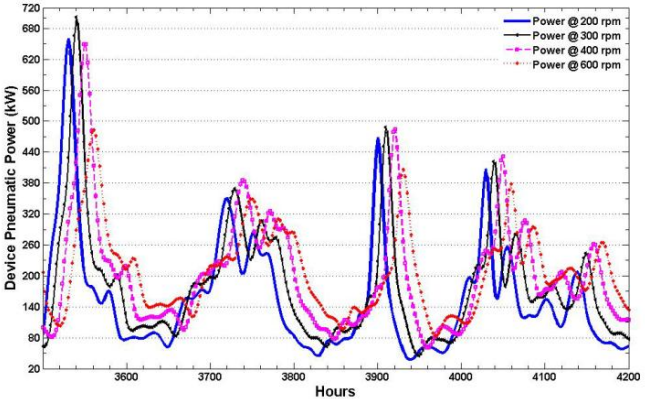



Fig.7: The OWC pneumatic power distribution in January

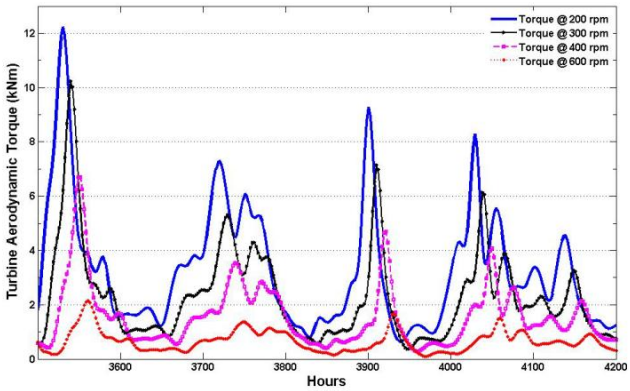



Fig.8: Turbine Torque distribution in January

However, as the incident power increases, it can be seen that the torque and hence the energy loss increases. Though the output torque at 200*rpm* is shown to be the highest, the shaft power is higher when the turbine is running at 300*rpm* because of the difference in the rotational speed. Furthermore, as shown previously, the converted pneumatic power available to the turbine is also higher.

*Optimum Turbine Rotational Speed*

One of the major parameters for the design of the impulse turbine for a wave power application is the rotational speed that the turbine is to be operated at. In Fig.9, different instantaneous efficiency profiles can be seen for the turbine over the range of input power when the turbine is operated at different rotational speeds. The higher the efficiency value (Fig.9) is about 46%, which corresponds to the optimum operating condition. This value of the peak conversion efficiency is in accordance with the available experimental data. In Fig.10, it can be seen that the corresponding shaft power produced at these rotational speeds. These last two figures are based essentially on the same information, but expressed in two different ways. Perhaps Fig.10 has a more immediate and intuitive meaning since it gives the output power directly.

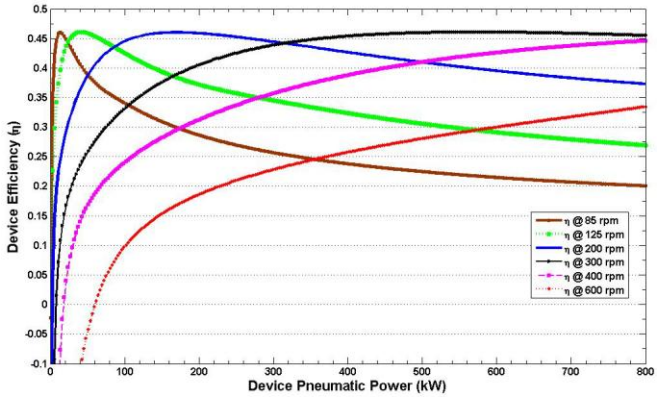



Fig.9: Device efficiency vs. available power

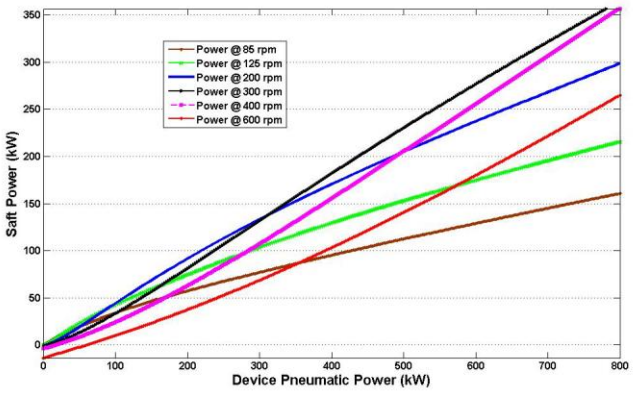



Fig.10: Shaft power vs. available power

The effect of efficiency peak occurring low in the input power range is apparent. Taking the efficiency curve corresponding to 85rpm shown in Fig.9 as a case in point, it can be seen that the efficiency characteristic of the turbine climbs rapidly to a peak at a low input power level. The impulse turbine with fixed guide vanes has a benign efficiency fall off characteristic after the peak efficiency point in comparison with the Wells turbine, which has a pronounced stall characteristic. There is no such characteristic in the impulse turbine with fixed guide vanes; this turbine still delivers useful amounts of power up to 800*kw* after it has reached its peak efficiency at a mere 20*kw*. However, this is clearly not the ideal manner in which to operate the turbine as can be seen from the corresponding curve in Fig.10.

Having the efficiency of the turbine peak low in the range of input power results in an inferior performance. In the case of the turbine being operated at 300*rpm*, Fig.10 shows that more power is delivered over a range of input power. In Fig.9, it is clear that the efficiency of the turbine peaks much closer to the mid range of input power. This results in the turbine converting a greater proportion of a larger input power, resulting in higher overall converted power. The integral OWC simulated performance is presented in Table 3, in terms of the monthly converted energy. These data are computed in time of the total power produced by the OWC device.

From the integration results, i.e. the total amount of energy converted by the OWC during a year is about 54 and 348*Mwh* when the turbine is running at 200*rpm* and 400*rpm*, respectively, with 40% of the total productivity from January and February. Critically, it was noticed that the productivity from July and August accounted for 20% of the total energy converted during the reference year, which is typical of the Irish climatic conditions. Finally, the mean turbine efficiency predictions, with and without damping along with the bi-directional unsteady flow experimental results, are shown in Fig.11. The predicted efficiency matches qualitatively the experimental efficiency results. From the curves in this plot, the effect of turbine damping can be observed, especially at higher values of $1/k\ \omega^*$. It can be noted here that in the flow conditions from this wave site, the magnitude of peak efficiency was comparable with that achieved under bi-directional unsteady flow conditions. However, the efficiency dropped significantly at higher flow coefficients as compared to that observed in the case without damping. The reason might be due to the fact that the pressure drop offered by the turbine in the region of higher flow

coefficient is comparatively very high and the curve is steeper than that of the experimental one. The peak efficiency dropped by a value of 2.5% from the experimental result of 43.5% (Rehil, 2007) at $1/k\ \omega* = 1.2$, which was achieved under uni-directional unsteady flow conditions. It can be noted that the value of $1/k\ \omega*$ corresponding to 300*rpm* at peak efficiency is different from the experimental value. This can be attributed to the fact that the turbine performance in the irregular wave conditions depends on the level of turbine damping, which in turns depends on the wave climate, especially the significant wave height. It can be said that turbine damping has a significant effect on the performance of turbine and also the overall performance of complete device.

TABLE 3
Monthly energy converted by the OWC device

| | Energy (MWh) | |
|---|---|---|
| Month | 200 rpm | 400 rpm |
| Sept | 1.35 | 11.38 |
| October | 2.42 | 32.74 |
| November | 1.74 | 6.27 |
| December | 1.59 | 13.52 |
| January | 11.72 | 70.06 |
| February | 8.55 | 76.16 |
| March | 1.37 | 14.71 |
| April | 4.78 | 33.02 |
| May | 3.44 | 22.16 |
| June | 2.64 | 22.34 |
| July | 4.18 | 10.31 |
| August | 10.40 | 35.98 |
| Tot | 54.23 | 348.72 |

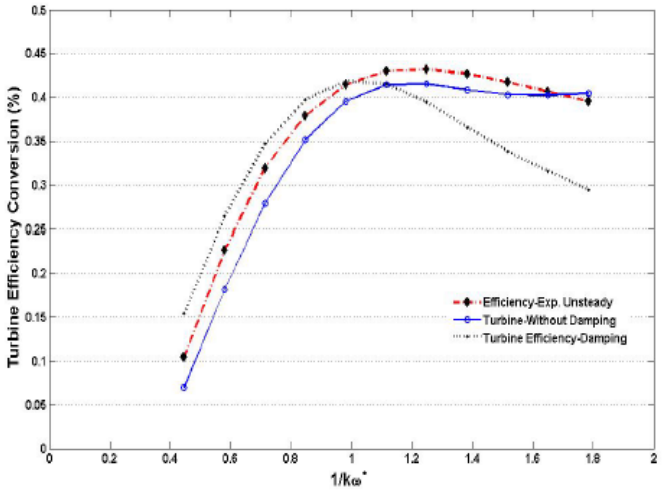



Fig.11: A comparison of turbine efficiency (with and without damping) and efficiency from the experiment

## CONCLUSION

The performance of the turbine under irregular, unsteady flow conditions using real sea data shows that the damping applied by the turbine is a major factor affecting the energy conversion from raw sea waves. The results achieved using this numerical simulation technique show that it can be used as an indicative tool to predict turbine performance under irregular, unsteady flow conditions for specific wave conditions. The use of real sea data enhances the practical use of this numerical simulation for predicting performance of the turbine for an actual power plant for a given wave site. In order to exploit the maximum power from the waves, the turbine should be matched so as to give an appropriate level of damping for the prevailing wave conditions. The typical stable behaviour, associated with this type of turbine was observed under irregular, unsteady real sea conditions. No stall point was observed and therefore the turbine was able to perform consistently at varying wave conditions. This characteristic gives it an edge over other self-rectifying turbines used for wave energy extraction. Therefore, it can be said that this impulse turbine is capable of converting energy for a wide range of flow coefficient under irregular, unsteady flow conditions. The simplified technique was used to simulate irregular, unsteady flow conditions for the turbine based on sea wave data do not react the true behaviour of an actual OWC.

## ACKNOWLEDGEMENTS

We acknowledge the support and the real Sea Wave Data given by the Marine Institute Rinville, Oranmore Galway, Ireland. http://www.marinedataonline.ie/

## REFERENCES

de O. Falca~o, A. F. (2003). First-generation wave power plants: Current status and R & D requirements. *J. of O-Shore Mechanics and Arctic Engineering*, November, 2003.

Thakker, A., Khaleeq, H. B., & Ansari, A. R. (2001). Numerical Simulation of 0.6m Impulse Turbine for Wave Power Conversion under different Flow Conditions. Proceedings of the 11th (ISOPE 2001) International O_shore and Polar Engineering Conference, June 17-22, 2001, Stavanger, Norway, *1*, 634-637.

Thakker, A., Usmani, Z., & Dhanasekaran, T. S. (2004). Effects of turbine damping on performance of an impulse turbine for wave energy conversion under different sea conditions using numerical simulation techniques. *Renewable Energy, 29*, 2133-2151.

Thakker, A., & Fergal, H. (2004). Modeling and scaling of the impulse turbine for wave power applications. *J. of Renewable Energy, 29,* 305-317.

Setoguchi, T., Takao, M., & Kaneko, K. (2006). Current status of self rectifying air turbines for wave energy conversion. *J. of Energy Conversion and Management, 47*, 2382-2396.

Thakker, A., Sheahan, C., Frawley, P., & Khaleeq H.B. (2001). The concurrent engineering approach to the manufacture of impulse turbine blades using rapid prototyping. *The Rapid Prototyping Journal, 7*(3), 15. Proceedings ofthe RPA ofthe Society ofMechanical Engineers, USA.

Thakker, A., Jarvis, J., Buggy, M., & Sahed, A. (2008). A novel approach to materials selection strategy Case study: wave energy extraction impulse turbine blade. *J. of Material Design*.

Ryan J. (2005). *Experimental analysis of irregular unsteady ow on performance of impulse turbine for wave energy conversion*. MEng thesis, University of Limerick.

Hammad, B. K. (n.d.). *Design Analysis of the Impulse Turbine with Fixed Guide Vanes forWave Energy Power Conversion*. Doctoral??

Setoguchi, T., Takao, M., Santhakuma, S., & Kaneko, K. (2004). Study of an Impulse Turbine for Wave Power Conversion: Effects of Reynolds Number and Hub-to-Tip Ratio on Performance. *J. O_shore Mech. Arct. Eng. 126*, 137-143.

Inoue, M., Kaneko, K., Setoguchi, T., & Saruwatari, T. (1998). Studies on the Wells Turbine for Wave Power Generation (Turbine characteristics and design parameters for irregular wave). *JSME International Journal, 31*, 676-682.

Curran, R., Denniss, T., & Boake, C. (2000). *Multidisciplinary design for performance: ocean wave energy conversion*. In Proceedings of the International O_shore and Polar Engineering Conference, p. 434-41. Seattle, Washington, USA, May 2000.

Chatry, G., Clement, A. H. & Sarmento, A. J. N. A. (1999). Simulation of a self-Adaptively Controlled OWC in a Nonlinear Wave Tank. Proceedings of the 9th (ISOPE-1999) International O_shore and Polar Engineering Conference, Brest, 3, 290-296.

Curran, R., Raghunathan, R. S., Stewart, T. P., & Whittaker, T. J. T. (1995). Matching a Wells Turbine to the Islay oscillating water column wave power converter. In Proceedings of the International Conference on O_shore Mechanics and Arctic Engineering 1995, p. 14754 (Copenhagen, Denmark, June 1822).

Thakker, A., Dhanasekaran, T. S., Takao, M. & Setoguchi, T. (2003). Effects of Compressibility on the Performance of a Wave- Energy Conversion Device with an Impulse Turbine Using a Numerical Simulation Technique. *International Journal of Rotating Machinery, 9*, 443-450.

Duckers, L. (1996). Chapter 8 - Wave Energy. In G. Boyle (Ed.), *Renewable Energy - Power for a Sustainable Future* (pp. 315 – 352). Oxford University Press.

Rehil O.A. (2007). *Experimental analysis of sel-rectifying turbines operating uni-directional steady flow and bi-directional unsteady flow for wave energy conversion.* Doctoral Thesis, UL (WERT).

## APPENDICES

| | |
|---|---|
| $A_C$ | area of chamber |
| $A_T$ | area of turbine |
| $b$ | blade height |
| $B_A$ | applied damping |
| $C_A$ | input coefficient |
| $C_T$ | torque coefficient |
| $\Delta p$ | pressure drop |
| $G$ | gap between rotor and guide vane |
| $H$ | incident wave height |
| $h$ | wave height in the chamber |
| $l_g$ | chord length of guide vane |
| $H_s$ | significant incident wave height |
| $\overline{T}$ | mean wave period |
| $l_r$ | chord length of rotor blade |
| $m$ | $A_T/A_C$ |
| $P^*$ | pressure coefficient |
| $Re$ | Reynolds number |
| $r_R$ | mid span radius |
| $S_g$ | guide vane pitch |
| $S_r$ | rotor blade pitch |
| $T$ | torque |
| $U_R$ | circumferential velocity at $r_R$ |
| $v_a$ | axial flow velocity |
| $v_c$ | air velocity at water column surface |
| $z$ | number of rotor blades |
| $\phi$ | flow coefficient |
| $\gamma$ | rotor blade inlet angle |
| $\eta$ | turbine efficiency under steady flow |
| $\nu$ | hub to tip ratio |
| $\theta$ | setting angle of fixed guide vane |
| $\rho_a$ | density of air |
| $\rho_s$ | density of sea water |
| $\omega$ | angular velocity of turbine rotor |
| $\omega^* = \omega\overline{T}$ | non-dimensional angular velocity |
| $k = \frac{m\, r_R}{H_{1/3}}$ | |
| $t^* = \frac{t}{\overline{T}}$ | non-dimensional time |
| $\phi = \frac{v_a^*}{K\omega^*}$ | flow coefficient |
| $\nu$ | hub to tip ratio |
| $\sigma_{rR}$ | solidity of rotor at $r_R$ |
| $v_a^* = \frac{m\overline{T}v_a}{H_s}$ | |
| $k\omega^* = \frac{\omega m r_R \overline{T}}{H_s}$ | |
| $\overline{\omega^*} = \omega\overline{T}$ | |

Pertanika J. Sci. & Technol. 22 (1): 307 - 314 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Sorption of $SO_2$ and NO by Modified Palm Shell Activated Carbon: Breakthrough Curve Model

**Sumathi, S.[1]*, S. Bhatia[2], K. T. Lee[2] and A. R. Mohamed[2]**

*[1]Faculty of Engineering and Green Technology, Department of Environmental Engineering, UTAR, Jalan Universiti, Bandar Barat, 31900 Kampar, Perak, Malaysia*
*[2]School of Chemical Engineering, Engineering Campus USM, 14300 Nibong Tebal, Pulau Pinang, Malaysia*

## ABSTRACT

Simultaneous removal of $SO_2$ and NO from simulated flue gas by cerium oxide supported over palm shell activated carbon (Ce/PSAC) was studied in a fixed bed adsorber. In this study, the adsorption breakthrough of $SO_2$ and NO on Ce/PSAC at different reaction temperatures was manipulated to test their applicability to a model developed by Yoon and Nelson (1984) for breakthrough curves. Yoon and Nelson (1984) developed a relatively simple model addressing the adsorption and breakthrough of adsorbate vapour with respect to activated charcoal. This model was based on the assumption that the rate of decrease in the probability of adsorption for each adsorbate molecule is proportional to the probability of adsorbate adsorption and the probability of adsorbate breakthrough on the adsorbent. A regression analysis (least square method) has been used to give the model parameters of $k$ and $t_{1/2}$. The results showed that the agreement between the model and the experimental results is satisfactory. From the observation, it is concluded that the simple two-parameter model of Yoon and Nelson's model can be applied for modelling the breakthrough curves of $SO_2$ and NO gas adsorption over Ce/PSAC.

*Keywords:* Palm shell, sulfur doxide ($SO_2$), nitrogen oxide (NO), breakthrough model

## INTRODUCTION

Currently, interests are growing in the use of simultaneous sorption processes to remove pollutant gases such as sulphur dioxide ($SO_2$) and nitrogen oxide (NO) simultaneous to avoid air pollution and green house effect. One of such process is adsorption. Gas adsorption is a separation process in which adsorbate molecules are adsorbed to the pore surface of solid adsorbents. Activated carbon has been recognized worldwide for its usage as an effective adsorbent for gas phase applications. This is due to its large surface area and pore volume (Barton *et al.*, 1997;

*Article history:*
Received: 26 December 2011
Accepted: 15 March 2012

*E-mail address:*
sumathi@utar.edu.my (Sumathi, S.)
*Corresponding Author

ISSN: 0128-7680 

Bansal *et al*., 1988). Literatures have shown that activated carbon can be used to remove $SO_2$ and NO (Davini, 2001; Qiang *et al*., 2005; Zhu *et al*., 2005). In previous work, the potential of PSAC (palm shell activated carbon) as a modified (impregnated with cerium oxide) sorbent was tested for its possibility in sorbing both $SO_2$ and NO gas simultaneously from a simulated flue gas at different temperatures. It was found that the modified PSAC (PSAC-Ce) successfully adsorbed $SO_2$ and NO simultaneously (Sumathi *et al*., 2010; Sumathi *et al*., 2010) and the adsorption result is presented as breakthrough curves.

Many theoretical or empirical equations have been proposed to model the adsorption breakthrough curves (Wood, 1987; Wood & Moyer, 1989) These equations were theoretically addressed to describe the adsorbate diffusion in the porous adsorbent such as activated carbon. It has been reported that breakthrough times (i.e. the time to reach defined effluent concentrations) and adsorption capacities (amounts held at breakthrough) of adsorbent beds for a given adsorbate gas are functions of the concentration of that gas in air (Yoon & Nelson, 1984). From the perspective of process modelling, the dynamic behaviour of a fixed bed column is described in terms of the effluent concentration–time profile, *i.e.* the breakthrough curve. There are a number of models suggested in the literature, such as Bohart and Adams model, Bed Depth Service Time model (BDST), Yoon and Nelson model, Clark model and Wolborska model (Hamdaoui, 2006). However, one of the most common models used for concentration profile is the Yoon and Nelson model (Yoon & Nelson, 1984).

In this paper, the theoretical breakthrough curves at five different equilibrium temperatures are generated and studied to be compared with the corresponding experimental data using Yoon and Nelson model to give an equation predicting the whole breakthrough curve for sorption of $SO_2$ and NO on PSAC-Ce.

## EXPERIMENTAL

PSAC was impregnated with cerium metal nitrate ($Ce(NO_3)_3 \cdot 6H_2O$) of an appropriate concentration to obtain around 10wt% of metal content per gram of PSAC (10 mL of 10wt % metal solution/gram of PSAC). The preparation method has been reported previously (Sumathi *et al*., 2010). The simultaneous removal activity of the prepared sorbent was carried out in a fixed bed adsorber. The prepared PSAC-Ce sorbent (1.0 g) was placed on borosilicate glass wool (0.05 g) in the centre of the adsorber. A stream of gaseous mixture representing the simulated flue gas, containing $SO_2$ (2000 ppm) (50%), NO (500 ppm) (11%), oxygen ($O_2$) (10%), and nitrogen ($N_2$)(29%) as the balance, was passed through the prepared sorbents. The schematic diagram and parameter control have been reported previously (Sumathi *et al*., 2010). The adsorption temperature was varied from 100ºC to 300ºC. The experimental results have been reported elsewhere (Sumathi *et al*., 2010). This data were used to model the breakthrough curve using Yoon and Nelson Model.

## BREAKTHROUGH MODEL

Yoon and Nelson developed a relatively simple model addressing the adsorption and breakthrough of adsorbate vapour with respect to activated charcoal (Yoon & Nelson, 1984). This model was based on the assumption that the rate of decrease in the probability of adsorption

for each adsorbate molecule is proportional to the probability of adsorbate adsorption and the probability of adsorbate breakthrough on the adsorbent. Yoon and Nelson's equation is not only less complicated than other equations, such as the Wheeler equation by Wheeler and Robell in 1963 (Tsai *et al.*, 1999) and Mecklenburg equation by Klotz in 1946 (Tsai *et al.*, 1999), but it also requires no detailed data concerning the characteristics of adsorbate, the type of adsorbent, and the physical properties of adsorption bed.

The Yoon and Nelson's equation (1984) is expressed as:

$$t = t_{1/2} + \frac{1}{K_{YN}} \ln \frac{C_e}{C_o - C_e} \quad (1)$$

$$k = K_{YN} t_{1/2} \quad (2)$$

where $K_{YN}$ is the rate constant ($min^{-1}$), $t_{1/2}$ the time required for 50% sorbate breakthrough (min), $t$ is the time (min), $C_e$ is the breakthrough concentration of adsorbate in ppm, $C_o$ is the initial inlet concentration of adsorbate in ppm and $k$ is the proportionality constant.

The calculation of the theoretical breakthrough curves for a single-component system requires the determination of the parameters $K_{YN}$ and $t_{1/2}$ for the sorbate of interest. These values may be determined from the available experimental data. The approach involves a plot of $\ln[C_e/(C_o-C_e)]$ versus $t$, time according to Eq. (1). If the theoretical model accurately characterizes the experimental data, this plot will result in a straight line with slope of $K_{YN}$ and intercept $-t_{1/2}K_{YN}=(-k)$. The proportionality constant, $k$ value can be obtained from Eq. (2).

## RESULTS AND DISCUSSION

In this study, the adsorption breakthrough of $SO_2$ and NO on PSAC-Ce at different reaction temperatures was used as the dataset to analyze the model. The experimental data were obtained from the breakthrough data of different temperature effects on $SO_2$ and NO sorption which have been reported elsewhere (Sumathi *et al*., 2010). The temperatures were in the range of 100°C to 300°C on each adsorbate respectively. Table 1 shows the saturation time and breakthrough time for $SO_2$ and NO at different temperatures extracted from the reported data earlier (Sumathi *et al*., 2010). The results show that $SO_2$ saturates faster than NO. NO takes longer time to achieve breakthrough than $SO_2$. Using these results, the rate constants were calculated.

In order to find the rate, constant plots of $\ln [C_e/(C_o-C_e)]$ vs. sampling time ($t$) according to Eq. (1) at various temperatures were done. The plots yield straight lines as expected (Yoon & Nelson, 1984) with the slope of $K_{YN}$ and the intercept of $-k$. The values were used to calculate $t_{1/2}$ and $k$ at each temperature. A regression analysis (least square method) was used to give the model the parameters of $k$ and $t_{1/2}$.

Table 2 shows the tabulated data at each temperature. The model appears to fit the experimental data reasonably well with correlation coefficient factor, ($R^2$) above 0.95. Both $K_{YN}$ and $t_{1/2}$ are dependent on the adsorbent adsorption capacity. According to Yoon and Nelson, the $K_{YN}$ and parameter values are dependent on the adsorption capacity, whereby the value of

$t_{1/2}$ decreases with decreasing adsorption capacity, while $K_{YN}$ increases (Yoon & Nelson, 1984). This phenomenon was predicted exactly in this case study.

In the case of $SO_2$, the $t_{1/2}$ values increased from 100°C to 150°C and then decreased from 200°C to 300°C. This is because the adsorption capacity of $SO_2$ by PSAC-Ce increased when the temperature increased from 100°C to 150°C, whereas when the temperatures were more than 200°C, the adsorption capacity of $SO_2$ decreased. According to Zhang *et al.,* increment in the operating temperature could enhance the chemical reaction rate and ionic diffusion rate of $SO_2$ (Zhang *et al*., 2006). Hence in this case at temperature 150°C, $SO_2$ had a higher ionic diffusion rate into the pores of Ce/PSAC. Furthermore, metal cerium possesses high oxygen mobility at higher temperature, which can easily be oxidized to $SO_3$ and this scenario indirectly gives a a higher breakthrough time (Trovarelli *et al.*, 1999). However, when the temperature was further increased from 200°C to 300°C, the $SO_2$ sorption was decreased. This is due to the kinetic energy of $SO_2$. It is known that at a higher temperature, $SO_2$ molecules lose their kinetic energies, which make the adsorption an exothermic process (Guo & Lua, 2002). This condition indirectly lessens the amount of $SO_2$ adsorbed into the pores of Ce/PSAC. The calculated $K_{YN}$ value increased when the temperature was increased from 100°C to 150°C and then decreased accordingly from 200°C to 300°C, as shown in Table 2. These findings indicate that Yoon and Nelson's model fits well with the current $SO_2$ experimental data vs. different temperature.

As for NO, it was observed that higher temperature (100°C -250°C) increases the adsorption capacity. Higher temperature shows a better removal of NO comparatively because Ce/PSAC is in a more active state to oxidize and reduce NO since cerium oxide has high oxidation activity at high temperature conditions (Waqif *et al*., 1997; Rodas-Grapain *et al*., 2005). Furthermore, at a lower temperature (150°C - 250°C), metal doped carbons showed high catalytic activity due to the disassociation of NO chemisorption, accompanied by $N_2O$ and $N_2$ evolution and oxygen accumulation on the catalyst surface (Mehandjiev *et al*., 1997). Besides that, when the temperatures were lower (<150°C), the impregnated cerium oxide was most likely in a less active state (whereby lower temperature, the capability of $CeO_2$ to store and release $O_2$ via the redox shift between $Ce^{4+}$ and $Ce^{3+}$ under oxidizing and reduction conditions respectively was lesser (Qi & Yang, 2003), thus giving only a moderate removal of NO from the simulated flue gas. Moreover, it was reported that at a higher temperature, AC itself could decompose NO and reduce it to $N_2$ (Guo *et al*., 2001; Mehandjiev *et al*., 1996).

TABLE 1
Saturation time and breakthrough time for $SO_2$ and NO at different temperatures (Sumathi *et al*., 2010)

| Temp. (°C) | $SO_2$ | | NO | |
|---|---|---|---|---|
| | Breakthrough Time | Saturation Time | Breakthrough Time | Saturation Time |
| 100 | 180 | 275 | 115 | 265 |
| 150 | 195 | 285 | 140 | 290 |
| 200 | 85 | 185 | 170 | 320 |
| 250 | 45 | 140 | 180 | 335 |
| 300 | 35 | 120 | 85 | 200 |

TABLE 2
Values of the parameters for $SO_2$ and NO adsorption at various temperatures

| | Temp. | Values of Parameters | | | $R^2$ | $R^2$ |
|---|---|---|---|---|---|---|
| | (°C) | $K_{YN}$ ($min^{-1}$) | $t_{1/2}$ (min) | k | | B/TCurve |
| $SO_2$ | 100 | 0.0754 | 228.06 | 17.20 | 0.9567 | 0.9962 |
| | 150 | 0.0714 | 240.99 | 17.21 | 0.9799 | 0.9978 |
| | 200 | 0.0632 | 133.81 | 8.46 | 0.9644 | 0.9933 |
| | 250 | 0.0653 | 90.97 | 5.94 | 0.9500 | 0.9921 |
| | 300 | 0.0832 | 74.30 | 6.18 | 0.9918 | 0.9987 |
| NO | 100 | 0.0346 | 195.33 | 6.76 | 0.9567 | 0.9940 |
| | 150 | 0.0324 | 199.56 | 6.47 | 0.9799 | 0.9937 |
| | 200 | 0.0308 | 251.95 | 7.76 | 0.9529 | 0.9947 |
| | 250 | 0.0304 | 260.15 | 7.91 | 0.9537 | 0.9932 |
| | 300 | 0.0443 | 144.43 | 6.40 | 0.9855 | 0.9964 |

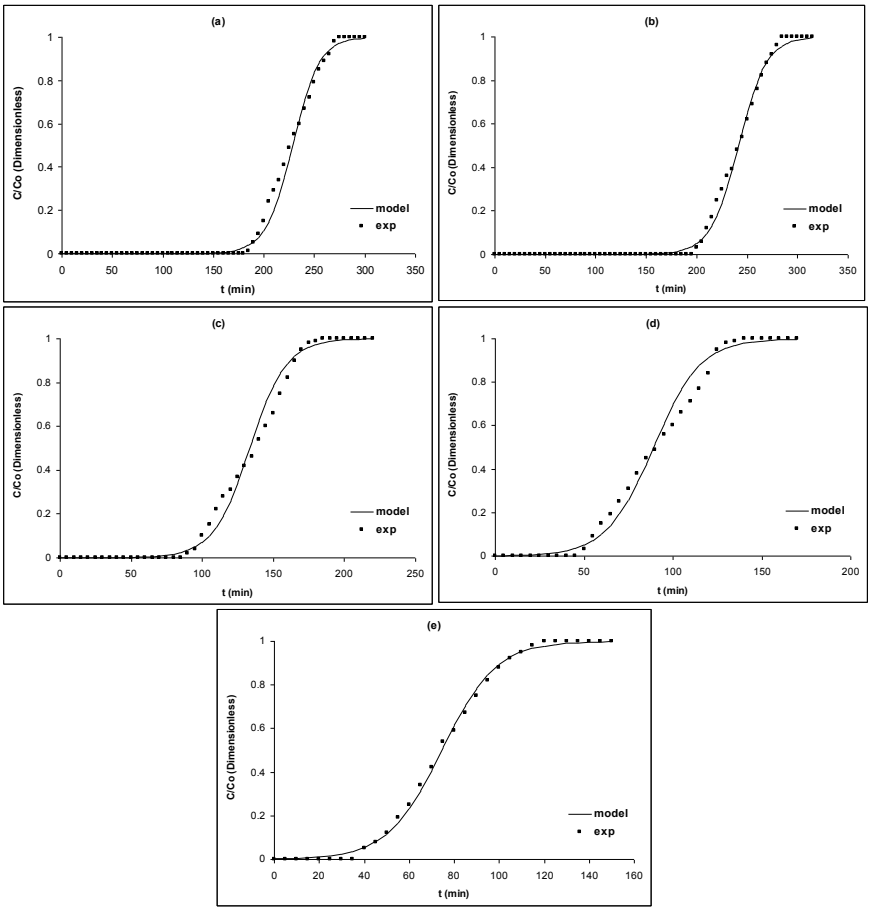



Fig.1: Breakthrough curves for various temperatures of $SO_2$ adsorption on PSAC-Ce; (a) 100ºC, (b) 150ºC, (c) 200ºC, (d) 250ºC, and (e) 300ºC by Yoon and Nelson's model

In the literature, it is indicated that a catalytic reaction and chemisorption of NO on active centres must occur at a temperature of more than 200°C (Muniz *et al*., 1999). As a result, the value of $t_{1/2}$ increased as the temperature increased. The $K_{YN}$ value decreased consequently. The value of $k$ (proportionality constant), on the other hand, is theoretically independent of the adsorption capacity. This is fairly well-demonstrated by the experimental results in this study. It is noted that different adsorbate gases are characterized by different values of $k$ (Yoon & Nelson, 1984). The $k$ values for NO range from 6.0 to 8.0 at different temperatures. Whereas for $SO_2$ the first two temperatures were around 17.2 and for the temperatures more than 150°C, it was around 6.0 to 8.5, respectively. The data fit well with the experimental data at different temperatures.

Following the determination of $K_{YN}$ and $t_{1/2}$, one can easily construct the complete breakthrough curves (Fig.1 and Fig.2) for the given set of experimental conditions by applying Eq. 1 and using the determined values in Table 2.

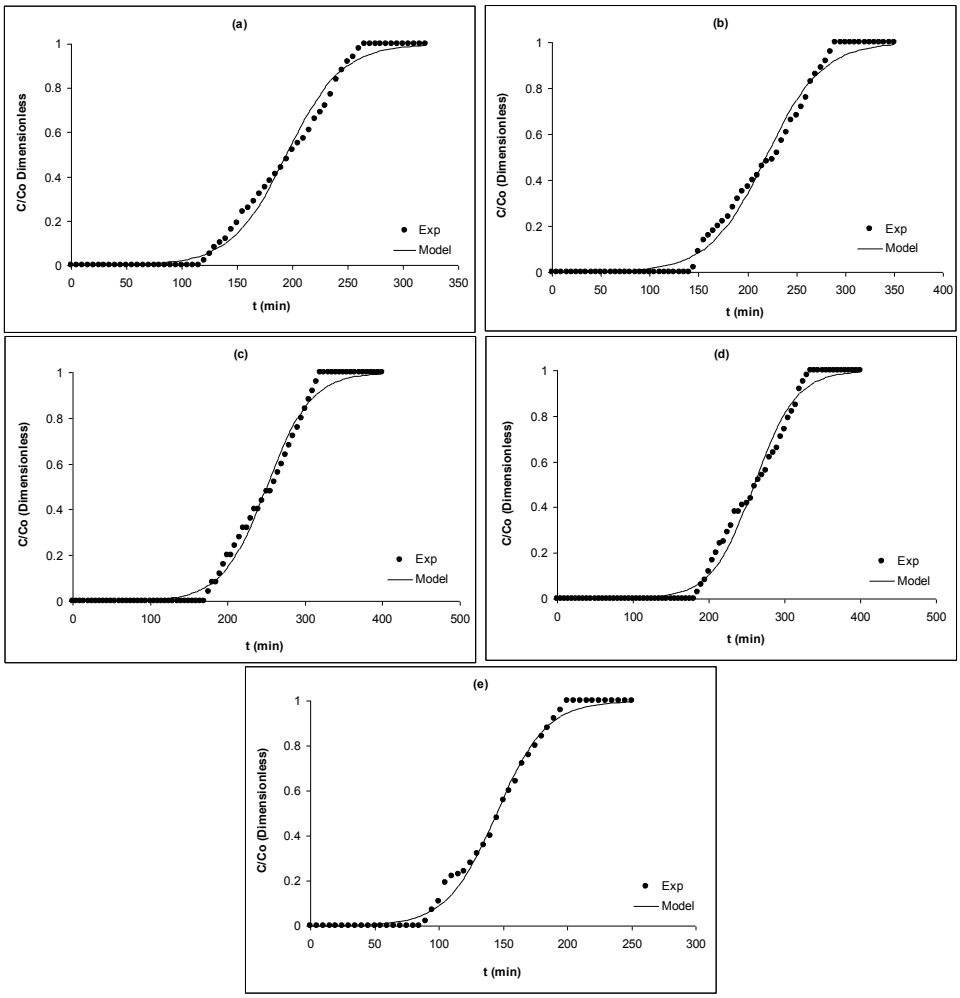



Fig.2: Breakthrough curves for various temperatures of NO adsorption on PSAC-Ce; (a) 100ºC, (b) 150ºC, (c) 200ºC, (d)250ºC, and (e) 300ºC by Yoon and Nelson's model

The calculated breakthrough curves based on Yoon and Nelson's equation were compared with the corresponding experimental data shown in Fig.1 for $SO_2$ and Fig.2 for NO. It was noted that each solid dot in Fig.1 and Fig.2 represents almost the model data of $SO_2$ and NO and it fit well with the experimental data at different temperatures. This shows that the agreement between the model and experimental results is very satisfactory. The corresponding coefficient factor was very high ($R^2$>0.9900) and near 1.0 agreement with the corresponding experimental data at different adsorber temperatures.

This further proves that the calculated theoretical breakthrough curves were in good. Even though the experimental data did show some inconsistent data due to the effects of temperature in this system, the model could still identify the changes and development of the model data accordingly.

## CONCLUSION

A simple two-parameter model, *i.e.* rate constant, *k*', and $t_{1/2}$ time required for 50% adsorbate breakthrough, $t_{1/2}$ of Yoon and Nelson's model could be applied for modelling the breakthrough curves of $SO_2$ and NO gas adsorption through PSAC-Ce. The parameters obtained from the fitting of the experimental data with the model were used to generate the theoretical breakthrough curves. The calculated theoretical breakthrough curves were in good agreement with the corresponding experimental data, whereby the $R^2$ of the model breakthrough curves was more than 0.9900.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the Ministry of Science, Technology and Innovation (MOSTI) Malaysia, Universiti Sains Malaysia's (814004) RU Grant and Yayasan Felda's (6050095) Long Term Grant for funding and supporting this research project.

## REFERENCES

Barton, S. S., Evans, M. J. B., & Halliop, E. (1997). Acidic and basic sites on the surface of porous carbon. *Carbon, 35*(9), 1361-1366.

Bansal, R. C., Donnet, J. B., & Stoeckli, F. (1988). Active Carbon, New York. *Marcel Dekker*, 450-482.

Davini, P. (2001). $SO_2$ and NOx adsorption properties of activated carbons obtained from a pitch containing iron derivatives. *Carbon, 39*, 2173–2179.

Qiang, T., Zhigang, Z., Wenpei, Z., & Zidong, C. (2005). $SO_2$ and NO selective adsorption properties of coal-based activated carbons. *Fuel, 84,* 461–465.

Zhu, J. L., Wang, Y. H., Zhang, J. C. & Ma, R. Y. (2005). Experimental investigation of adsorption of NO and $SO_2$ on modified activated carbon sorbent from flue gases. *Energy Conversion Management, 46*, 2173–2184.

Sumathi, S. Bhatia, S., Lee, K. T., & Mohamed, A. R. (2010). $SO_2$ and NO simultaneous removal from simulated flue gas over cerium-supported palm shell activated at lower temperatures-role of cerium on NO removal. *Energy and Fuels, 24*, 427-431.

Sumathi, S., Bhatia, S., Lee, K. T., & Mohamed, A. R. (2010). Cerium impregnated palm shell activated carbon (Ce/PSAC) sorbent for simultaneous removal of $SO_2$ and NO - Process study. *Chemical Engineering Journal, 162*(1), 51-57.

Wood, G. O. (1987). A model for adsorption capacities of charcoal beds: II. Challenge concentration effects. *American Industrial Hygiene Association Journal, 48,* 703-709,.

Wood, G. O., & Moyer, E. S. (1989). A review of the wheeler equation and comparison of its applications to organic vapor respirator cartridge breakthrough data. *American Industrial Hygiene Association Journal, 50*, 400-407.

Yoon, Y. H., & Nelson, J. H. (1984). Application of gas adsorption kinetics. I. A theoretical model for respirator cartridge service time. *American Industrial Hygiene Association Journal, 45*, 509-516.

Hamdaoui, O. (2006). Dynamic sorption of methylene blue by cedar sawdust and crushed brick in fixed bed columns. *Journal of Hazardous Materials, 138*, 293-303.

Tsai, W. T., Chang, C. Y., Ho, C. Y., & Chen, L. Y. (1999). Adsorption properties and breakthrough model of 1,1-dichloro-1-fluoroethane on activated carbons. *Journal of Hazardous Materials, B69*, 53–66.

Zhang, H., Tong, H., Wang, S., Zhuo, Y., Chen, C., & Xu, X. (2006). Simultaneous removal of $SO_2$ and NO from flue gas with calcium-based sorbent at low temperature. *Industrial Engineering Chemistry Research, 45*, 6099-6103.

Trovarelli, A., de Leitenburg, C., Boaro, M., & Dolcetti, G. (1999). The utilization of ceria in industrial catalysis", Catalysis Today, 50, 353-367.

Guo, J., & Lua, A. C. (2002). Microporous activated carbons prepared from palm shell by thermal activation and their application to sulphur dioxide adsorption. *Journal of Colloid Interface Science, 251*, 242-247.

Waqif, M., Bazin, P., Saur, O., Lavalley, J. C., Blanchard, G., & Touret, O. (1997). Study of ceria sulfation. *Applied Catalysis B: Environment, 11*, 193-205.

Rodas-Grapain, A., Arenas-Alotorre, J., Gomez-Cortes, A., & Diaz, G. (2005). Catalytic properties of a Cuo-$CeO_2$ sorbent-catalyst for de-SOx reaction. *Catalysis Today, 107-108*, 168-174.

Mehandjiev, D., Bekyarova, E., & Khristova, M. (1997). Study of Ni-impregnated active carbon II. Catalytic behaviour in NO conversion. *Journal of Colloid Interface Science, 192*, 440-446.

Qi, G., & Yang, R. T. (2003). Performance and kinetics study for low-temperature SCR of NO with $NH_3$ over MNOx-$CeO_2$ catalyst. *Journal Catalysis, 217*, 434-441.

Guo, Z., Xie, Y., Hong, I., & Kim, J. (2001). Catalytic oxidation of NO to $NO_2$ on activated carbon. *Energy Conversion Management, 42*, 2005-2018.

Mehandjiev, D., Khristova, M., & Bekyarova, E. (1996). Conversion of NO on Co- impregnated active carbon catalysts. *Carbon, 34*, 757-762, 1996.

Muniz, J., Marban, G., & Fuertes, A. B. (1999). Low temperature selective catalytic reduction of NO over polyarylamide-based carbon fibers. *Applied Catalysis B: Environment, 23*, 25-35.

Pertanika J. Sci. & Technol. 22 (1): 315 - 328 (2014)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Geochemical Assessment of Sediment Quality using Multivariate Statistical Analysis of Ennore Creek, North of Chennai, SE Coast of India

**M. Jayaprakash[1], R. Nagarajan[2]*, P. M. Velmurugan[3], L. Giridharan[1], V. Neetha[4] and B. Urban[5]**

*[1]Department of Applied Geology, School of Earth and Atmospheric Sciences, University of Madras, Guindy Campus, Chennai 600025, India*
*[2]Department of Applied Geology, School of Engineering and Science, Curtin University, CDT 250, Miri, 98009, Sarawak, Malaysia*
*[3]Centre for Earth and Atmospheric Sciences, Satyabama University, Jeppiaar Nagar, Chennai 600119, India*
*[4]Libinz Instutite for Applied Geophysics, Geochoronology and Isotope Hydrology, Stilleweg 2, D-30655, Hannover, Germany*
*[5]Leuphana University of Lüneburg, Environmental Sciences and Engineering, Herbert-Meyer-Str. 7, D 29556, Suderburg, Germany*

## ABSTRACT

The composite nature of Ennore Creek, receiving polluted waters from Buckingham canal, River Kortalaiyar, and the presence of numerous industries rapidly degrade the coastal environment. In this study, multivariate statistical technique was used to assess the nature of pollution and to identify the factors responsible for the enrichment of trace metals in the creek sediments. Forty samples were collected during pre- and post-monsoon periods to evaluate the seasonal variations on the concentration of trace metals in the sediments. The results indicate that not much seasonal variation exists in the concentration of trace metals in the sediments of Ennore creek. Results of the cluster analysis illustrate that the enrichment of trace metals was mainly from anthropogenic sources. Meanwhile, correlation coefficient among the metals reveals that some of the metals (Fe and Al) derived were of natural origin. The complex data matrix of the sediment were interpreted after reduction to three factors and the results illustrate the extent of the influence of anthropogenic activities. The spatial distribution diagrams demonstrate and demarcate the region of enrichment of metals in the sediments of Ennore creek.

*Keywords:* Cluster analysis, Factor scores, Ennore creek, heavy metal pollution

*Article history:*
Received: 26 December 2011
Accepted: 15 March 2012

*E-mail addresses:*
em_jaypee@yahoo.com (M. Jayaprakash),
nagarajan@curtin.edu.my (R. Nagarajan),
girilogu@yahoo.com (L. Giridharan)
*Corresponding Author

ISSN: 0128-7680 

## INTRODUCTION

Estuarine environments have traditionally served as recipients of domestic and industrial effluents from the adjacent metropolis (Jayaprakash *et al.,* 2005). The effect of contamination is dependent on the nature and quantum of effluents, and on whether they are discharged directly into the estuary from a point source, or indirectly through river systems. Heavy metal pollution in the estuarine environment is very serious due to their toxicity, persistence and ability to accumulate in biota (Beldi *et al.,* 2006). Heavy metals can remain in the environment unchanged for years and may pose threat to humans and other organisms (Ahmad *et al.,* 2009). Toxicity and bioavailability of heavy metals in an ecosystem depends on the way they are fixed in the sediments. The mobile fraction of the metal adsorbed at the surface of solid bodies is the most labile and known as exchangeable fraction. In this fraction, metals are bound by processes of physical and chemical sorption and they interchange easily with other ions in the environment. Thus, evaluation of heavy metal contamination is vital in order to determine the level of its contamination in the environment. The data obtained from such studies are significant in the decision making of environmental planning strategies.

Hydrodynamic and physico-chemical conditions of the estuary have great influence on the horizontal distribution of metals in creek water and sediments. Physical and chemical partitioning behaviour and speciation of metals, within and between different environmental compartments, are controlled by various physicochemical factors (Padmini *et al.,* 2007). Heavy metal pollution in the water and sediment of Ennore creek is influenced by both point and non-point pollution sources. The high pollution load in Ennore Creek has drastically changed the ecosystem (Azariah, 1997; Jayaprakash *et al.,* 2005; 2008). The study of the geochemistry of sediment requires handling of a large data set, which includes the concentrations of various ions. At the same time, classification, quantification and interpretation of the data are important steps in the assessment of sediment quality. In order to achieve this objective, multivariate statistical techniques have been successfully used by various researchers. Sediment quality data of the area were subjected to factor analysis (FA) to interpret, understand and identify the mechanisms, processes (both natural and anthropogenic) and specific source of water quality deterioration and contamination in the area. In particular, FA explains the correlations between the variables in terms of the underlying factors, which are not apparent otherwise (Yu *et al.,* 2003). In this work, R-mode FA is used to identify and quantify the factors responsible for the enrichment of metals in the sediments of Ennore creek.

## STUDY AREA

The study area (Ennore Creek) is located the north of Chennai, where most of the area consist of alluvial tracts and the remaining, in the eastern part, is occupied by beach dunes, tidal flats and creek (Jayaprakash *et al.,* 2008). The Creek is an aperture of the brackish water system into the Bay of Bengal. It is almost 800m wide and elongated in a NE-SW direction. Ennore Creek merges with the backwater bodies and the north-south trending channels connecting it to Pulicat Lake in the north and to the tributaries of Kortalaiyar River in the south (Fig.1). The depth of the creek ranges from 2 to 3m and is shallow near the mouth and on the northwestern part merging with the tidal flats. Mangrove swamps are also noticed in the area. The southern

limb of the creek, fringing the northern areas of Chennai city, is well-developed with industries, utilities, suburban residential areas and fishing hamlets. The northern section of the creek or Kosistalaiyar backwater is connected to Pulicat Lake and has two major developmental areas – the North Chennai Thermal Power plant and the recently built Ennore port. Waste water enters the creek through the Buckingham canal, a waterway that was built for navigation. The canal section that traverses between Chennai and Ennore currently serves as an open sewer receiving municipal and industrial wastewaters (Jayaprakash *et al*., 2012). The creek also receives wastewater from the Manali industrial area. It is crucial to note that rapid development of Chennai city in the last three decades has put additional stress on the surrounding aquatic environment.

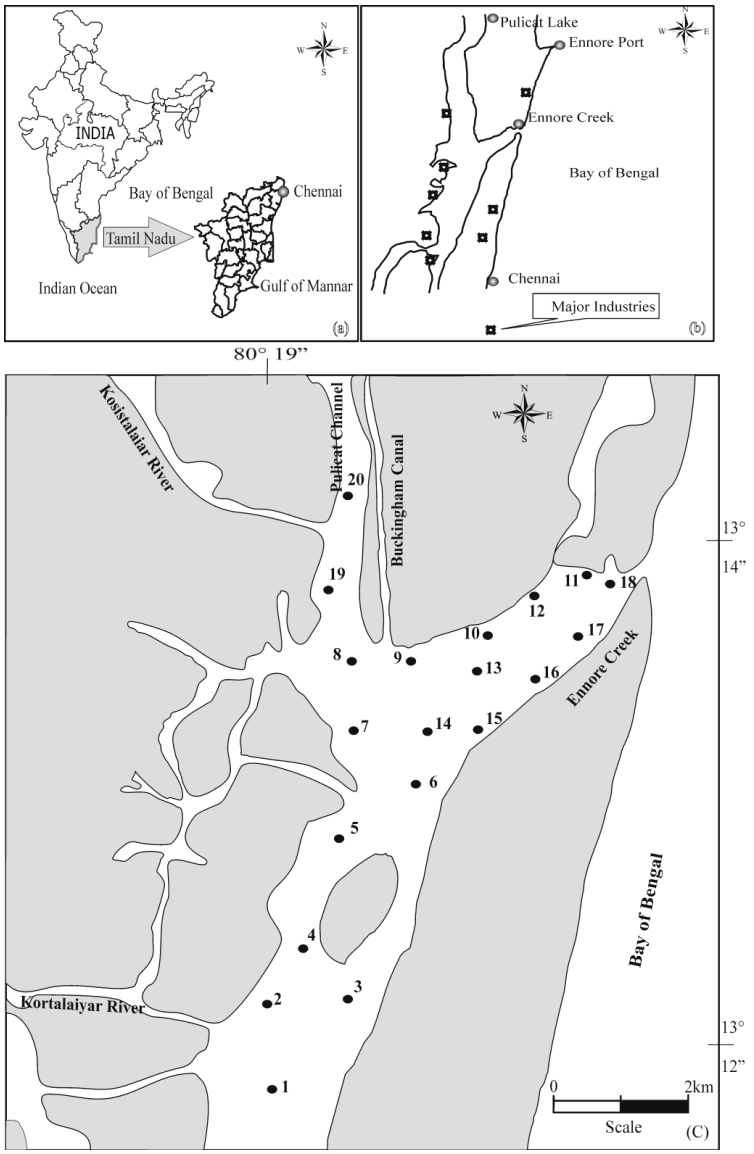



Fig.1: Location map of the study area

## MATERIALS AND METHOD

### *Analytical Methodology*

Forty surface sediment samples were collected from a boat using Van Veen grab during the pre- (PRM) and post-monsoon (POM) seasons of 2009 (20 samples in each season). The samples were collected from the entire creek region, which has a variety of branches spread over a wide area, with marshy conditions in many places. The sampling sites were selected on the basis of possible local and point sources. The samples were taken from the central part of the grab sampler to avoid any metallic contamination from the sampler itself and were frozen at -4°C immediately onboard until further analysis. The samples were collected from a depth of 2 m to a maximum depth of 3 m during both seasons. The samples were separated following the cone and quartering method and were powdered in an agate mortar to the 230 um size.

For total metal analysis, a known quantity of sediment was digested with an acid mixture of $HClO_4$ and HF and the final residue was leached with HCl and made up to the required quantity (Tessier *et al.,* 1996). Trace metal concentrations (Fe, Mn, Cr, Cu, Ni, Co, Pb and Zn) were measured using flame atomic absorption spectrophotometer (Perkin-Elmer AA700) that was equipped with a deuterium background corrector. Suitable internal chemical standards (Merck chemicals, Germany) were used to calibrate the instrument. All the reagents used were of analytical grade and high purity. The accuracy of the analytical procedures was assessed using certified reference material MESS-1 (Table 1) from the National Research Council of Canada for the present study.

TABLE 1
Published and obtained analytical results of MESS-1

| Elements | MESS-1 | Present study | Recovery % |
|---|---|---|---|
| Fe | 3.1 ± 0.38 | 2.95 | 95.16 |
| Mn | 513 ± 25 | 472 | 92.00 |
| Cr | 71 ± 1.1 | 69.8 | 98.30 |
| Cu | 25.11 ± 3.88 | 22.9 | 91.19 |
| Ni | 29.5 ± 2.7 | 28.3 | 95.93 |
| Co | 10.8 ± 1.9 | 9.9 | 91.66 |
| Pb | 34 ± 6.1 | 31.8 | 93.52 |
| Zn | 191 ± 17 | 173.3 | 90.73 |

### *Statistical Methodology*

Cluster analysis (CA) is an effective statistical tool for identifying and evaluating similar groups from the data matrix. In this study, the classification based on the sampling site was performed through CA using Ward's method (Ward, 1963), with Euclidian distance as a similarity measure and synthesis into dendogram plots. Euclidean distance is the geometric distance in multidimensional space. Ward's method is known to be distinct as it uses an analysis of variance approach to evaluate the distances between clusters. This particular method minimizes the sum of squares (SS) of any two (hypothetical) clusters that have been derived at each step (Mcgarial *et al.,* 2000; Vega *et al.,* 1998; Zeng & Rasmussen, 2005).

FA is unique since the patterns of relationship among many dependent variables are studied with the goal of determining the nature of the independent variables that affect them, even though those independent variables are not measured directly. FA was applied to the data matrix in order to reduce the data to an easily interpretable form. Before applying FA, the data were standardized according to the criteria presented by Davis (2002). Computation of the correlation coefficient matrix is the first step in FA between the standardized variables. The Eigen values quantify the contribution of a factor to the total variance. The contribution of a factor is significant when the Eigen value is greater than unity (Kaiser, 1960). Initial factors are extracted and they are subjected to mathematical rotation. Varimax rotation procedure was used to maximize the difference between the variables facilitating easy interpretation of the data. The first factor accounts for as much variance as possible in the data set. The second factor accounts for as much residual variance as possible, and so forth. The factor loading indicates the degree of closeness between the variables and the factor. The highest loading, either positive or negative, suggests the meaning of the dimension, while the positive loading indicates that the contribution of the variables increases with the increasing loading in a dimension, and the negative loading indicates a decrease (Lawrence & Upchurch, 1982). The study of factor scores reveals the extent of influence of each factor on the overall sediment chemistry at all locations of the sampling stations. Extreme negative scores reflect areas that are essentially unaffected by the processes and positive scores reflect the most affected areas. Near-zero scores indicate areas affected to an average degree (Giridharan *et al.,* 2009). In the present study, the variation of factors in each station is indicated by spatial distribution diagrams.

## RESULTS AND DISCUSSION

The statistical parameters of the analytical data, such as maximum, minimum, mean and standard deviation, are presented in Table 2.

TABLE 2
Composition of metals and texture study results of Ennore creek sediments

| Elements | Pre monsoon | | | | Post monsoon | | | |
|---|---|---|---|---|---|---|---|---|
| | Min | Max | Mean | σ | Min | Max | Mean | σ |
| Al | 48693 | 119085 | 86255 | 21005 | 39325 | 110193 | 81231 | 18355 |
| Fe | 26700 | 52000 | 40737 | 7732 | 23292 | 53500 | 41774 | 8652 |
| Mn | 422.0 | 949.0 | 666.3 | 151.6 | 391.6 | 903.0 | 686.4 | 153.0 |
| Cr | 160.0 | 700.0 | 408.6 | 136.1 | 211.0 | 598.0 | 392.0 | 110.4 |
| Cu | 46.0 | 222.0 | 112.9 | 53.0 | 42.0 | 200.0 | 106.3 | 48.6 |
| Ni | 17.3 | 52.3 | 39.0 | 10.1 | 23.8 | 48.0 | 36.5 | 8.0 |
| Co | 2.8 | 9.8 | 7.1 | 1.9 | 2.2 | 6.1 | 4.6 | 1.3 |
| Pb | 19.5 | 49.8 | 36.7 | 8.3 | 16.0 | 42.0 | 30.3 | 6.3 |
| Zn | 100.0 | 226.0 | 163.9 | 41.8 | 109.0 | 214.4 | 160.2 | 34.1 |
| Mud | 16.0 | 44.6 | 29.2 | 9.2 | 14.0 | 48.5 | 29.8 | 10.0 |
| OM | 2.5 | 5.0 | 3.7 | 0.7 | 2.0 | 6.6 | 4.0 | 1.3 |
| $CaCO_3$ | 2.8 | 5.0 | 3.8 | 0.6 | 2.7 | 5.1 | 4.0 | 0.8 |

*Temporal Similarity and Grouping of Parameters*

CA based on linear pairs of coefficient of correlation among the concentrations of geochemical parameters indicates three different clusters during both PRM and POM periods (Fig.2a-b). Cluster 1 includes all trace metals along with mud, $CaCO_3$ and Organic Matter (OM), which clearly indicates that the origin of these metals should have been from anthropogenic activities. Meanwhile, enrichment of these metals in the sediments would have occurred due to the precipitation/coagulation of the pollutants in the water column (e.g., Thangadurai *et al.,* 2005). Inter-elemental association of Cr, Cu, Zn, Pb, and Mn suggests that their origin was mainly from the long-term dumping of solid waste and sewage disposal. Moreover, close linkage distance in these trace metals further suggests that corrosion from heavy transportation vehicles and burning of tires brings in considerable amounts of Cr, Pb, Zn, and Cu which accumulate in the surface sediments (e.g., Turner *et al.,* 2001; Wang *et al.,* 2005).

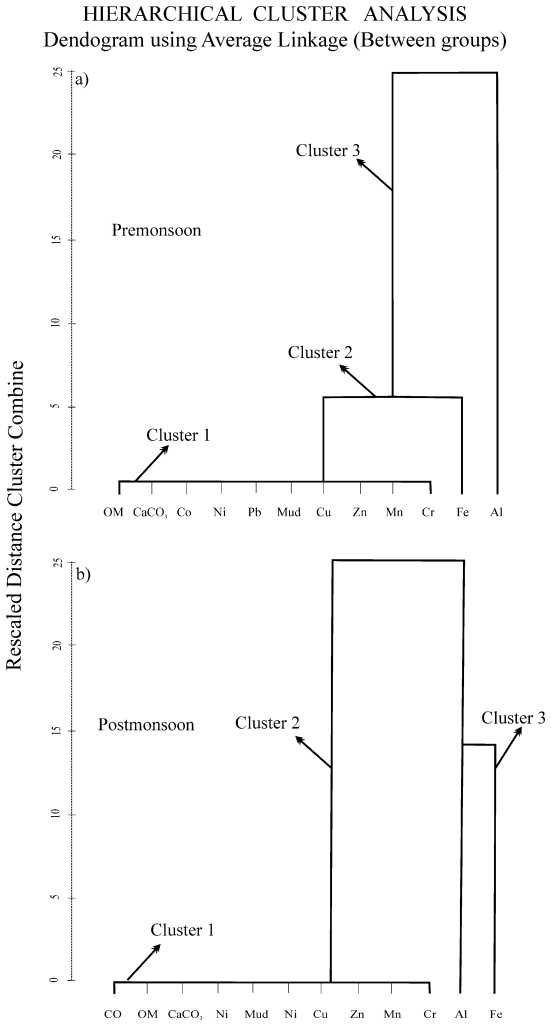



Fig.2: a-b: Dendrogram based on complete linkage method for surface sediments of Ennore creek, SE coast of India

OM has a significant role in the geochemical cycles of trace elements that accumulate in sediments; it may be used as an index of depositional environment and sedimentary processes (Seralathan *et al.,* 1993). The finer particles (mud) get bound to organic carbon and also facilitate the coagulation of trace metals from the water column. Sewage effluents enter the Ennore creek through Buckingham canal and also from the rivers Kortalaiyar and Kosistalaiyar which have increased the OM content in the water and the finer sediments in this region supporting the deposition of organic debris (Rajamanickam & Setty, 1973). Enrichment of the trace metals in the sediment of the study area through anthropogenic activities was clearly demonstrated by the cluster diagram by transporting the trace metals along with mud and OM in one group.

During PRM, Fe and Al fall separately but distinctly from the trace metals. CA shows few relationships between cluster 1 and cluster 2 (Fe); however, cluster 3 (Al) distinguishes discretely. Indeed, these results clearly demonstrate that some quantum of Fe would have originated from anthropogenic activities, whereas, Al content in the sediments may be enriched in the bed sediments by means of lithology of the region or by stream sedimentation process (Jayaprakash *et al.,* 2010). In the case of POM, cluster 1 demonstrates the same parameters as those of PRM, whereas clusters 2 and 3 fall discretely, indicating that their nature and origin may be from the weathering of aluminosilicates (Calvert & Price, 1983).

### *Correlation Studies*

Trace metals and the strong relationships between most of them indicate that they are primarily controlled by Fe-Mn oxyhydroxides (Buckley *et al.,* 1995; Zwolsman *et al.,* 1993). In the present study, majority of the trace metals were observed to be strongly associated with both Fe and Mn, indicating that total trace metals (TTMs) in the region are mainly controlled by Fe and Mn oxyhydroxides, with OM as the functional element (Tessier *et al.,* 1979). The correlation between Co and Cr, Cu, Ni, Pb and Zn during PRM indicates that the chemical cycling of metals has occurred in the study area (e.g., Takematsuuu *et al.,* 1984). In PRM sediments, Fe has strong positive correlation with Mn, Cr, Ni, Co, Pb, OM and mud. The strong correlation of Fe with Ni and Co only indicates that these metals sre not associated with oxide-oxyhydroxides of Fe-Mn phase (Praysers *et al.,* 1991). Ni-Cr-Cu (r=0.8) exhibits a strong positive correlation among them (Table 3a). OM derived from decaying organic remains has a high capability of accumulating Ni, Cu and Cr by adsorption and by the formation of chelating compounds (Wedepohl, 1974). During POM, inter-elemental relationship of TTMs shows that Cr vs Cu (r=0.80); Cr vs Ni (r=0.82); Cu vs Co (r = 0.51), Zn vs Pb (r = 0.82) and Ni vs Co (0.76) in the sediment indicate their nature and origin (Table 3b). The electroplating industries located near the Ennore creek are involved in the processes such as surface preparation and pickling which produce acidic and alkaline wastewater with elevated Zn concentration (Forstner & Wittmann, 1981; Selvaraj *et al.,* 2003). Moreover, effluents from these industries contain significant amounts of Ni, Cr and Cd. Zn shows a strong correlation with Ni, reflecting the association of Ni with the source of Zn from the industrial effluents. Cr, Zn and Cu are known as pollution indicators of paint industries (Lin *et al.,* 2002), and the nature of origin was confirmed from the strong correlation among them.

*Factor Analysis*

The first three factors accounting for 88% of the variance during PRM and 86% of the variance during POM were extracted from the principal factor matrix after varimax rotation (Table 4). The spatial distribution diagrams of factor scores reveal the extent of influence of each factor on the overall sediment chemistry of the sample stations.

TABLE 3
Correlation matrix (r) of total trace elements in surface sediments of Ennore Creek, SE coast of India (n = 20)

| | Al | Fe | Mn | Cr | Cu | Ni | Co | Pb | Mud | OM |
|---|---|---|---|---|---|---|---|---|---|---|
| Fe | 0.84 | 1 | | | | | | | | |
| Mn | 0.58 | 0.78 | 1 | | | | | a. PRM | | |
| Cr | | 0.51 | 0.78 | 1 | | | | | | |
| Cu | | 0.54 | 0.78 | 0.90 | 1 | | | | | |
| Ni | 0.56 | 0.68 | 0.73 | 0.83 | 0.82 | 1 | | | | |
| Co | | 0.50 | 0.67 | 0.81 | 0.72 | 0.76 | 1 | | | |
| Pb | | | | 0.57 | 0.58 | 0.54 | 0.72 | 1 | | |
| Zn | | | 0.45 | 0.61 | 0.68 | 0.56 | 0.65 | 0.79 | | |
| Mud | 0.90 | 0.82 | 0.59 | | | 0.50 | | | 1 | |
| OM | 0.90 | 0.73 | 0.53 | | | 0.52 | | | 0.89 | 1 |
| $CaCO_3$ | -0.77 | -0.79 | -0.56 | | | -0.58 | | | -0.81 | -0.76 |
| Fe | 0.86 | 1 | | | | | | | | |
| Mn | 0.79 | 0.81 | 1 | | | | | b. POM | | |
| Cr | | 0.55 | 0.73 | 1 | | | | | | |
| Cu | | | | 0.80 | 1 | | | | | |
| Ni | | 0.58 | 0.60 | 0.82 | 0.78 | 1 | | | | |
| Co | 0.56 | 0.61 | 0.61 | 0.56 | 0.51 | 0.76 | 1 | | | |
| Pb | | 0.45 | | | | | | 1 | | |
| Zn | | | | | | | | 0.87 | | |
| Mud | 0.84 | 0.77 | 0.82 | 0.49 | | | 0.59 | | 1 | |
| OM | 0.83 | 0.76 | 0.72 | | | | 0.57 | | 0.89 | 1 |
| $CaCO_3$ | -0.68 | -0.65 | | | | | | -0.47 | -0.60 | -0.68 |

TABLE 4
Results of the Principal component analysis - Rotated Component Matrix

| PARAMETERS | Factor 1 | | Factor 2 | | Factor 3 | |
|---|---|---|---|---|---|---|
| | PRM | POM | PRM | POM | PRM | POM |
| Al | 0.932 | 0.945 | 0.188 | 0.126 | 0.068 | 0.043 |
| Fe | 0.791 | 0.824 | 0.513 | 0.344 | -0.004 | 0.275 |
| Mn | 0.474 | 0.739 | 0.782 | 0.537 | 0.159 | 0.035 |
| Cr | 0.160 | 0.320 | 0.847 | 0.863 | 0.406 | 0.061 |
| Cu | 0.120 | -0.135 | 0.866 | 0.965 | 0.397 | -0.045 |

TABLE 4 (*continue*)

| PARAMETERS | Factor 1 | | Factor 2 | | Factor 3 | |
|---|---|---|---|---|---|---|
| | PRM | POM | PRM | POM | PRM | POM |
| Ni | 0.434 | 0.249 | 0.715 | 0.899 | 0.394 | 0.166 |
| Co | 0.277 | 0.543 | 0.551 | 0.636 | 0.670 | -0.182 |
| Pb | 0.014 | 0.215 | 0.211 | -0.027 | 0.954 | 0.939 |
| Zn | -0.085 | 0.065 | 0.459 | 0.103 | 0.763 | 0.958 |
| Mud | 0.953 | 0.889 | 0.111 | 0.239 | 0.146 | 0.090 |
| OM | 0.942 | 0.924 | 0.079 | 0.113 | 0.110 | 0.079 |
| $CaCO_3$ | -0.850 | -0.750 | -0.289 | 0.144 | 0.148 | -0.402 |
| Eigen value | 6.910 | 6.215 | 2.920 | 2.491 | 0.770 | 1.684 |
| % variance | 57.567 | 51.791 | 24.363 | 20.755 | 6.419 | 14.032 |

### *1. PRM Factors*

Factor 1 of PRM explains 57% of the total variance, has high loadings of the ions Al, Fe, mud and OM. The results show that these metals originate from a common source of aluminosilicate weathering of the catchment area. Meanwhile, associations of major metals with mud and OM indicate that the metals incorporated in the mud fractions get bound with OM in the water column and settle onto the sediment in Ennore creek (Lin *et al.,* 2002; Zang *et al.,* 2009). Factor score diagram shows high values in the south eastern part of the study area. In this scrub region, settlements near the creek are denser compared to the other parts of the study area and thus, the anthropogenic activities may have increased the OM content in the creek water which further settle onto sediments (Fig.3a).

Factor 2 of PRM accounts for 24% of the total variance and is characterised by high loadings of Fe, Mn, Cr, Cu, Ni and Co. The areal distribution diagram of factor 2 reveals that major parts of the study area show high values with regard to this particular factor. River Kosistalaiyar and Buckingham canal bring polluted water consisting of both industrial and domestic effluents from the metropolitan city, and the region adjacent to these water sources are shown to contain the highest factor scores (Fig.3c). This illustrates that sediments are enriched with these metals because of the anthropogenic activities in the source region of River Kosistalaiyar and Buckingham canal. The southern part of the study area also shows significant factor scores, where River Kortalaiyar mixes with the creek waters. Since the adjoining industries use this river as an outlet for their effluents, the heavy metal contents in the Ennore creek sediments continuously get augmented by precipitation or coagulation/ flocculation from the polluted water column.

Factor 3 of PRM that accounts 6% of the total variance has high loadings of the ions - Co, Pb and Zn. The spatial distribution map of factor 3 illustrates that about 70% of the study area was affected with regard to this factor (Fig.3e). The southern, northern and northeastern parts of the Ennore creek particularly show highly significant factor scores, indicating that the sediments in these regions are highly enriched with these heavy metals. Since these metals are not associated with either Fe-Mn or OM, the origin of these metals can be attributed to the surface run-off of contaminated soil from the adjoining industrial areas.

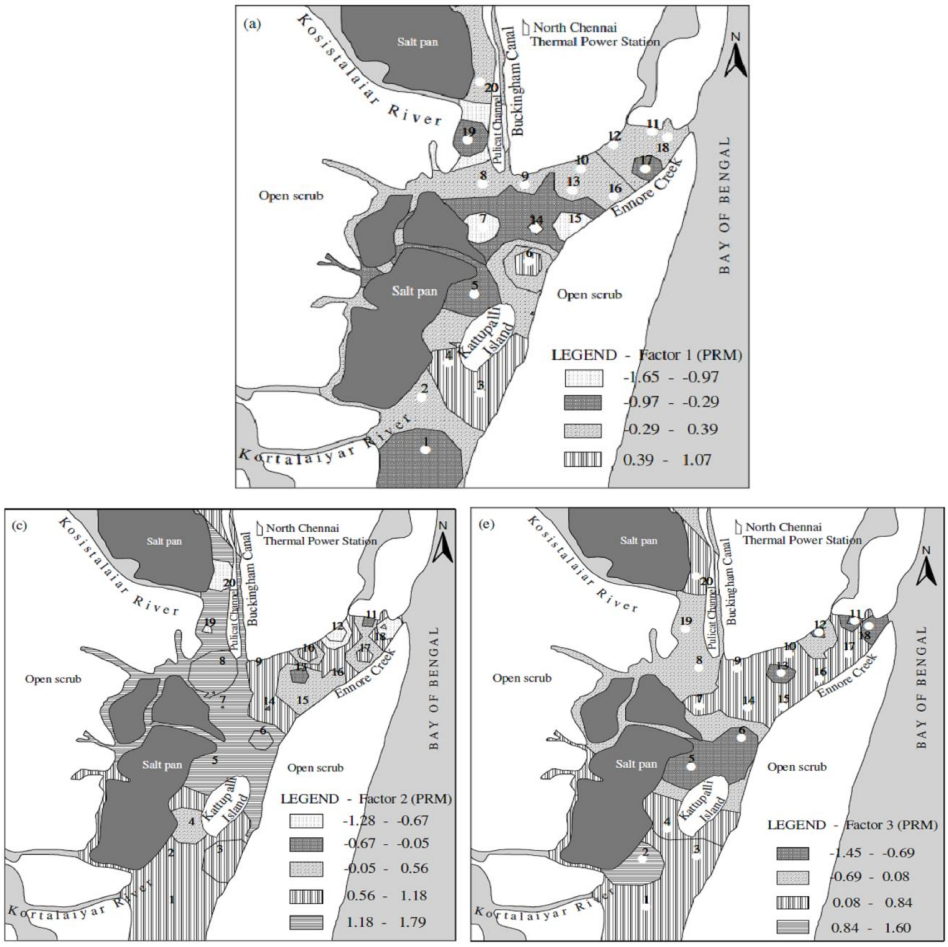



Fig.3(a, c, e): The spatial distribution diagram of factors 1-3 scores for the surface sediments of Ennore creek, SE coast of India during PRM

## 2. *POM Factors*

Factor 1 of POM explains 51% of the total variance and has high loadings of Al, Fe, Mn, Co, mud and OM. The high abundance of these metals indicates that their origin must be attributed to anthropogenic and silicate weathering. The factor score diagram shows high values in major parts of the study area (see Fig.3b). Fe-Mn hydroxides and OM act as scavengers and bind to the heavy metals by adsorption/complexation, thereby enriching the heavy metal content in sediments (Tessier *et al.* 1996). Factor 2 of POM accounts for 20% of the total variance and it has high loadings on Mn, Cr, Cu, Ni and Co. The areal distribution diagram of factor 2 reveals that major parts of the study area show high values with regard to this particular factor (Fig.3d). Domestic and industrial effluents are directed into the creek through River Kosistalaiyar, River Kortalaiyar and Buchingham canal, which increase the heavy metal contents in the creek waters. Hence, results of POM fall in line with the PRM with respect to the distribution and concentration of heavy metals. Factor 3 of POM which accounts for 14% of the total variance

shows high loadings on Pb and Zn. The spatial distribution map of factor 3 demonstrates that the northern and northeastern parts of the study area show high factor score values, indicating that these regions are affected with regard to this factor (Fig.3f). When compared to PRM, the distribution of these metals during POM was found to be reduced in area and concentrated in the northeastern direction. Seasonal effect was also seen in the distribution and concentration of these metals although the overall effect was minimal.

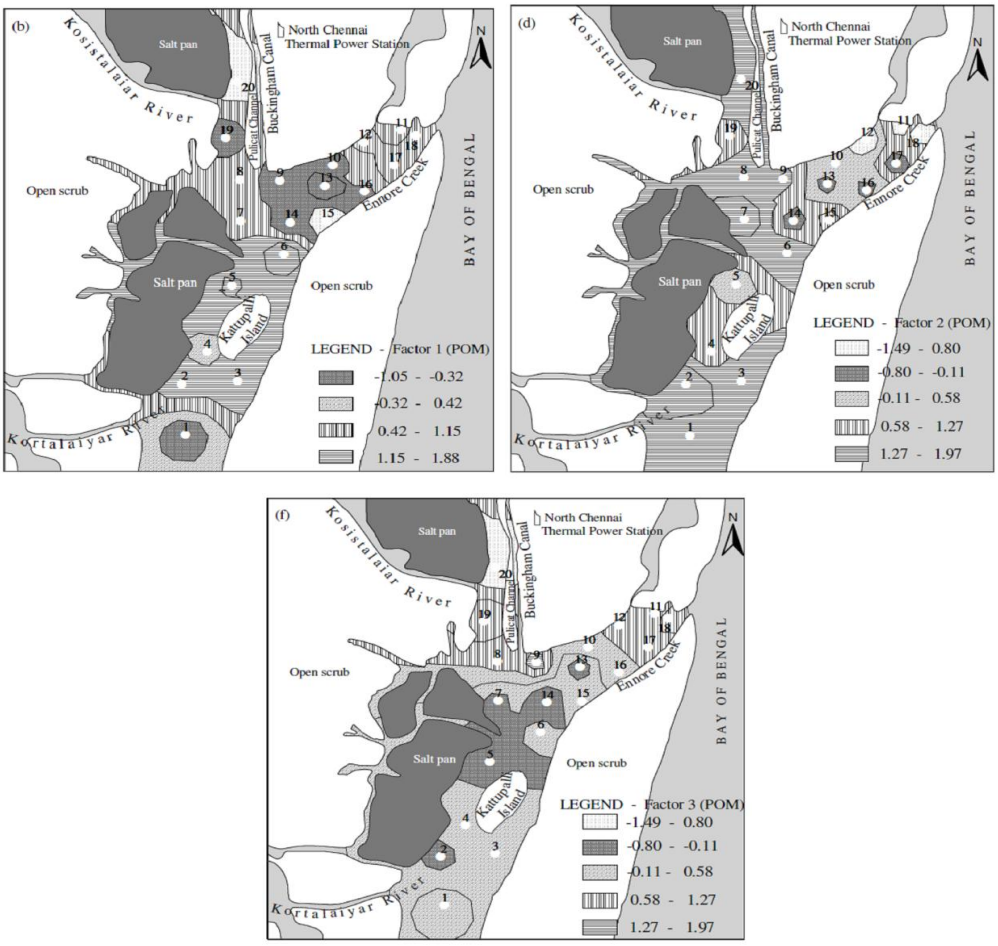



Fig.3(b, d, f): The spatial distribution diagram of factors 1-3 scores for the surface sediments of Ennore creek, SE coast of India during POM

## CONCLUSION

The multivariate statistical techniques were applied on the sediment data from Ennore creek and the comprehensive nature of the results from CA and FA clearly demonstrate the effective application of these methods. In addition, the nature and origin of the trace metals were also identified and quantified. The results of CA clearly indicate the associations of all the trace metals with mud and OM, illustrating that the enrichment of metals is mainly from the

anthropogenic activities. Meanwhile, the seasonal effects on the trace metal content of the sediments were found to be minimal. The correlation studies have shown that Fe has a strong association with most of the trace metals, along with mud and OM. The industrial activities near Ennore creek are responsible for the enrichment of the trace metals in the sediments. The spatial distribution diagram of the factor scores demonstrates a high concentration of the trace metals near the region, where the highly polluted effluents originate from Buckingham canal. Similarly, the enrichment of metals was observed near the area where the rivers Kosistalaiyar and Kortalaiyar transport both domestic and industrial effluents into the creek waters.

## ACKNOWLEDGEMENTS

The first author is grateful to the Department of Science and Technology, Govt. of India for the award of the advanced research fellowship under the Better Opportunities for Young Scientists in Chosen Areas of Science and Technology (BOYSCAST' 2009) programme. The authors also thank the authorities of Leuphana University of Luneburg, Suderburg, Germany for granting permission to carry out the research. This article is the 62nd contribution from Earth System Science Group (ESSG), Chennai, India.

## REFERENCES

Ahmad, A. K., Mushrifah, I., & Othman, M. S. (2009). Water quality and heavy metal concentrations in sediment of Sungai Kelantan, Kelantan, Malaysia: A Baseline Study. *Sains Malaysiana*, 38, 435-442.

Azariah, J. (1997). *Biomanagement of biogeoresources*. Bioethics in India. pp.16-19.

Beldi, H., Gimbert, F., Maas, S., Scheifler, R., & Soltani, N. (2006). Seasonal variations of Cd, Cu, Pb and Zn in the edible mollusc Donax trunculus (Mollusca, Bivalvia) from the gulf of Annaba, Algeria. *African J Agri Res, 1*, 85-90.

Buckley, D. E., Smith, J. N., & Winters, G. V. (1995). Accumulation of contaminant metals in marine sediments of Halifax Harbour, Nova Scotia: environmental factors and historical trends. *App Geochem, 10*, 175-195.

Calvert, S. E. & Price, N. B. (1983). Geochemistry of nambian shelf sediments. In E. Suess & J. Thiede (Eds.), *Coastal upwelling and its sediment record* (pp. 337–375). New York: Plenum.

Davis, J. C. (2002). *Statistics and data analysis in geology* (pp. 526-540). New York: John Wiley & Sons

Forstner, U., & Wittmann, G. T. W. (1981). *Metal pollution in the aquatic environment.* Springer: Berlin, 476p.

Giridharan, L., Venugopal, T., & Jayaprakash, M. (2009). Assessment of Water Quality Using Chemometric Tools: A Case Study of River Cooum, South India. *Arch Environ Contam Toxicol*, 56, 654–669. DOI 10.1007/s00244-009-9310-2.

Jayaprakash, M., Jonathan, M. P., Srinivasalu, S., Muthuraj, S., Ram-Mohan, V., & Rajeshwara-Rao, N. (2008). Acid-leachable trace metals in sediments from an industrialized region (Ennore Creek) of Chennai city, SE coast of India: An approach towards regular monitoring. *Estuar Coast Shelf Sci., 76*, 692–703.

Jayaprakash, M., Nagarajan, R., Velmurugan, P. M., Sathiyamoorthy, J., Krishnamurthy, R. R., & Urban, B. (2012). Assessment of trace metal contamination in a historical freshwater canal (Buckingham Canal), Chennai city, India. *Environ Monit Assess, 184, 7407-7424*. DOI: 10.1007/s10661-011-2509-5.

Jayaprakash, M., Srinivasalu, S., Jonathan, M.P., & Ram Mohan, V. (2005). A baseline study of physico-chemical parameters and trace metals in waters of Ennore Creek, Chennai, India. *Mar Poll Bull, 50*, 583-608.

Jayaprakash, M., Urban, B., Velmurugan, P. M., & Srinivasalu, S. (2010). Accumulation of total trace metals due to rapid urbanization in microtidal zone of Pallikaranai marsh, South of Chennai, India. *Environ Monit Assess, 170*, 609-629.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurment, 20*, 141-151.

Lawrence, F. W., & Upchurch, S. B. (1982). Identification of recharge areas using geochemical factor analysis. *Ground Water, 20*, 680-687.

Lin, S., Hsieh, I. J., Huang, K. M., & Wang, C. H. (2002). "Influence of the Yangtze River and grain size on the spatial variations of heavy metals and organic carbon in the East China Sea continental shelf sediments. *Chem Geol., 182*, 377-394.

Mcgarial, K., Cushman, S., & Stafford, S. (2000). *Multivariate statistics for wildlife & ecology research*. New York: Springer.

Padmini, E., & Geetha, V. B. (2007). A comparative seasonal pollution assessment study on Ennore Estuary with respect to metal accumulation in the grey mullet, Mugil cephalus. *Oceanological and Hydrobiological studies, 36*(4), 91-103.

Praysers, P. A., DeLange, G. J., & Middleburh, J. J. (1991). Geochemistry of eastern Medditerranean sediments: primary sediment composition diagenetic alterations. *Mar Geol., 100,* 137-154.

Rajamanickam, G. V., & Setty, M. G. A. P. (1973). Distribution of phosphorus and organic carbon in the near shore Sediments of Goa. *Indian J. Mar. Sci., 2*, 84-89.

Selvaraj, K., Ram Mohan, V., Jonathan, M. P., Siddartha, R., & Srinivasalu, S. (2003). Distribution of nondetrital trace metals in sediment cores from Ennore Creek, Southeast coast of India. *J. Geol. Soc. India, 62*, 191-204.

Seralathan, P., Meenakshikutty, N. R., Asaref, K. V., & Padmalal, D. (1993). Sediment and organic carbon distribution in the Cochin harbour area. *Ind. J. Mar. Sci., 22*, 252-255.

Takematsu, N., Sato, Y., & Okabe, S. (1984). The formation of todorokite and birnessite in seawater pumped from underground. *Geochim. Cosmochim. Acta, 48*(5), 1099-1106.

Tessier, A., Campbell, P. G. C., & Bisson, M. (1979). Sequential Extraction Procudure for the Speciation of Particulate Trace Metals. *Anal. Chem.,51*, 844-851.

Tessier, A., Fortin, D., Belzile, N., DeVitre, R. R., & Leppard, G. G. (1996). Metal sorption to diagenetic iron and manganese oxyhydroxides and associated organic matter: Narrowing the gap between field and laboratory measurements. *Geochim Cosmochim Acta, 60*(3), 387-404.

Thangadurai, N., Srinivasalu, S., Jonathan, M.P., Rajeshwara Rao N., & Santhosh Kumar, R. (2005). Pre-Tsunami chemistry of sediments along the inner continental shelf off Ennore, Chennai, southeast coast of India. *Ind. J. Mar. Sci., 34*(3), 274-278.

Turner, D., Maynard, J. B., & Sansalone, J. J. (2001). Heavy metal contamination in soils of urban highways: Comparison between runoff and soil concentrations at Cincinnati, Ohio. *Water Air Soil Pollut., 132*, 293-314.

Vega, M., Pardo, R., & Deban, L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res., 32*, 3581-3592.

Wang, X. S., Qin, Y., & Sang, S. X. (2005). Accumulation and source of heavy metals in urban top soils: A case study from the city of Xuzhou China. *Environ. Geol., 48*, 101-107.

Ward, J. H. (1963). Hierachical grouping to optimize an objective function. *J. America Statistical Association,* 236-244.

Wedepohl, K. H. (1974). Copper: Abundance in common sediments and sedimentary rocks". In K. H. Wedepohl (Ed.), *Handbook of geochemistry* (pp. 29/K/1–29/K/10). New York: Springer

Yu, S., Shang, J., Zhao, J., & Guo, H. (2003). Factor analysis and dynamics of water quality of the Songhua River, Northeast China. *Water Air Soil Pollut., 144*, 159-169.

Zeng, X., & Rasmussen, T. C. (2005). Multivariate statistical characterization of water quality in Lake Lanier, Georgia, USA. *J. Environ Qual., 34*, 1980-1991.

Zhang, W. Feng, H. Chang, J. Qu, J. Xie, H., & Yu, L. (2009). Heavy metal contamination in surface sediments of Yangtze River intertidal zone: An assessment from different indexes. *Environ Poll., 157*, 1533-1543.

Zwolsman, J. J. G., Berger, G. W., & Van Eck, G. T. M. (1993). Sediment accumulation rates, historical input and retention of major elements and trace metals in salt marsh sediments of the Scheldt estuary, SW Netherlands. *Mar. Chem., 44*, 73-94.

Pertanika J. Sci. & Technol. 22 (1): 329 - 335 (2014)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Preliminary Investigation on Electrochemical Parameters of Lake Waters in and around Miri City, Malaysia

**M. V. Prasanna[1]*, R. Nagarajan[1], A. Elayaraja[2] and S. Chidambaram[3]**

*[1]Department of Applied Geology, School of Engineering and Science, Curtin University, CDT 250, 98009 Miri, Sarawak, Malaysia*
*[2]Department of Mathematics and Statistics, Faculty of Science and Engineering, Curtin University, Perth, Western Australia*
*[3]Department of Earth Sciences, Annamalai University, India*

## ABSTRACT

Surface water samples were collected from 16 Lakes in and around Miri City to assess the electrochemical parameters includes pH, Electrical conductivity (EC), Total dissolved solid (TDS), redox potential (Eh), resistivity and salinity. Sampling locations for monitoring were selected in the vicinity of major roads, industries, settlements and agricultural region. Interpretation of data shows that the surface water in the central region of the study area is polluted by various anthropogenic activities, while in the southern part is within the limits of guideline values. This kind of investigation is essential in the study area to save the resources for future perspective. Further detailed studies are also needed to get a clear picture of the surface water quality in Miri city and for future sustainable management of this resource.

*Keywords:* Lake water, electrochemical parameters, pollution, Miri City

## INTRODUCTION

Lakes are distinctive and valued ecosystems, which forms the lifeline of many biota. Lakes are afftected by various factors like pollution, loss of biodiversity, incursion of foreign species, watershed disposal, as well as declining water levels (Kaverina & Pogozheva, 2005).

In fact, lakes are subjected to continuous evolutionary changes and finally ends up with the remnants of lake organisms and with soil carried in by floods. These physical and chemical changes of a lake affect the plant and animal populations. However, human activities also give stress to the lakes. The variations in the physico-chemical properties of lakes is due to number of factors, but mainly by rainwater dilution gives seasonal changes

*Article history:*
Received: 26 December 2011
Accepted: 15 March 2012

*E-mail addresses:*
prasanna@curtin.edu.my (M. V. Prasanna),
nagarajan@curtin.edu.my (R. Nagarajan),
elayarajah@yahoo.com (A. Elayaraja),
chidambaram_s@rediffmail.com (S. Chidambaram)
*Corresponding Author

ISSN: 0128-7680 

(Odo & Ijere, 1997) and variety of pollutants discharged into the lakes from surrounding activities (Manahan, 2005). Many works have been focused on human impact on the various ecosystem (Niemi *et al*., 1990; Issa *et al*., 1996; Szymanowska *et al*., 1999; Prasanna *et al*., 2012a). The spatial and temporal variations in surface waters provides a holistic picture of water quality (Dixon & Chiswell, 1996). The lakes also serve as reservoir for the recharging aquifers (Prasanna *et al.,* 2011), hence the quality of water in these lakes are significant.

Miri city is covered by beaches, industries, residental colonies and agricultural areas. Increasing population and economic development in Miri leads to the scarcity for freshwater resource. Recent literature surveys showed high contamination of the lakes in and around Miri city (Prasanna *et al.,* 2012a&b). Hence, this current research work provides a baseline monitoring information on lake water quality affected by natural and anthropogenic impacts in and around Miri city.

## STUDY AREA

The study area falls between the latitudes N4°21'18'' – N4°35'42'' and longitudes E113°57'54'' – E114°05'6'' and located in Miri City, the Sarawak State of East Malaysia (Fig.1). The lithology of the area consist mainly of sandstones, and laminated to thin bedded siltstones and mudstones. The residual soil cover derived from these rock types is silty or clayey sand and silty clay, weathered from sandstones and mudstones/shales, respectively. The climate of this region is controlled by northeast and southwest monsoons and the average temperature of a year is 26°C. Sarawak experiences more rainfall than west Malaysia, ranging from 3,000 to 4,000 mm/yr.

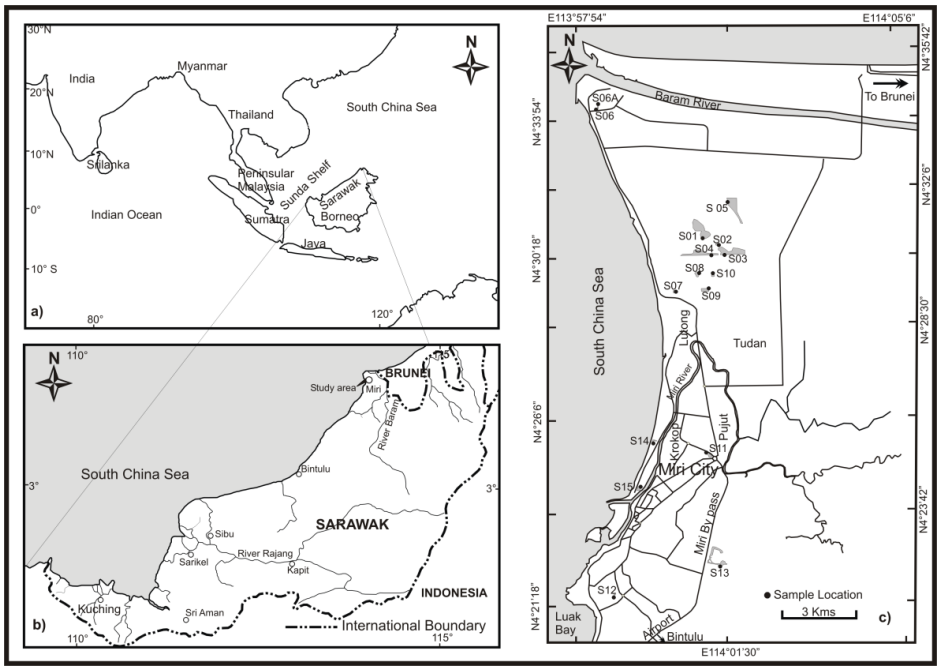



Fig.1: Location map and sampling points of the study area (Prasanna *et al*., 2012a)

## METHODOLOGY

Sixteen surface water samples were collected in June, 2011 from various lakes in and around Miri city to assess the water quality (Fig.1). Most of the monitoring sampling points are located in the vicinity of major roads, industries, settlements and agricultural areas. pH, total dissolved soilds (TDS), Electrical conductivity (EC), redox potential (Eh), resistivity, salinity and temperature of lake waters were determined *in situ* using portable pH and conductivity meter (model-Thermo Scientific Orion Star, 4 Star Plus Meter). The instrument was calibrated with standard solutions before each parameter measurement.

## RESULTS AND DISCUSSION

A summary of the measured parameters of samples is given in Table 1. The pH values of the lake water were found to ranges from 4.20 to 8.72, with an average of 7.24. This indicates that the lake water is mainly from neutral to alkaline, except at the sampling point (S5), which shows a low pH value (4.20), indicating acidic in nature. A record of a high pH value (8.72) was observed at S15, which is higher than the permissible limit of the standard outlined by the World Health Organization (WHO, 2004). The spatial distribution of pH map (Fig.2) shows that the highest values were noted in the southern part and a small pocket in the northern side, while the lowest values were observed in the north eastern part of the study area. The highest pH values were observed in lake waters that are situated near recreational centres, restaurants and settlements. Waste disposal and discharge of domestic wastes from this region may have increased the pH values in the surface water. Meanwhile, the lowest pH value of 4.20 was observed at S5, which is mainly due to the presence of sulfide minerals in the shale formation and minor coal seams in sandstones and shales (Siddharth *et al*., 2004). This was also confirmed by the low turbidity observed in that sample. Intensive agricultural activities and leaching of peat soils in this region are also responsible for the transport of salts into the lakes (Prasanna *et al.,* 2010).

TABLE 1
Maximum, minimum and average values of the parameters

| | Temp °C | pH | EC (µS/cm) | Salinity (ppt) | Resistivity (MΩ-cm) | Eh (mV) | TDS (mg/l) |
|---|---|---|---|---|---|---|---|
| Min | 27.00 | 4.20 | 49.50 | 0.10 | 0.0001 | 13.70 | 24.00 |
| Max | 34.60 | 8.72 | 13770.00 | 8.00 | 0.0200 | 196.90 | 6749.00 |
| Average | 31.04 | 7.24 | 1208.44 | 0.66 | 0.0060 | 41.34 | 592.25 |

On contrary to the general decrease in the pH values in the north eastern side of the study area, results a general increase in the Eh values. The Eh values ranged from 13.70 to 196.90, with an average of 41.34, indicating that surface water could be characterized as a low oxidizing

environment. The high values of the Eh was recorded in the north eastern part of the study area (Fig.3) indicating high oxidation conditions and also serves as a favorable condition for aquatic biota. It is also inferred that the increase of $O_2$ content in water increases the Eh value. Interestingly at S6, the Eh value is higher (40.4mV), but it is very near to S6A which shows a comparatively low value (29mV). This is due to the aquaculture activity in Lake S6, the continuous process of aeration increases $O_2$ in this lake.

In general, electrical conductivity (EC) depends on various factors like temperature, ionic concentration and type of ions in the water (Hem, 1985). EC is used to monitor the water quality for irrigation and drinking purposes, depending on salinity, ionic concentration and temperature. The EC values range from 49.50 µs/cm to 13770 µs/cm, with an average of 1208.44 µs/cm. The EC value of the surface waters indicate that all the samples fall within the permissible limit of WHO standard, except at S7 Lake. The spatial distribution map of EC (Fig.4) shows that higher values (particularly at S7 with EC value 13770 µs/cm) were observed in the central part of Miri, where is highly occupied by settlements. Such anomalous values arise from various anthropogenic activities near the residential area. Interestingly, the salinity at S7 is 8 ppt, which is also quite higher than the salinity of other lakes due to the increasing ionic concentration due to the discharge of domestic sewage. The TDS values of the lake waters ranged from 24mg/l to 6749mg/l, with an average value of 592.25 mg/l. All the samples fall within the permissible limit of WHO standard, except at S7.

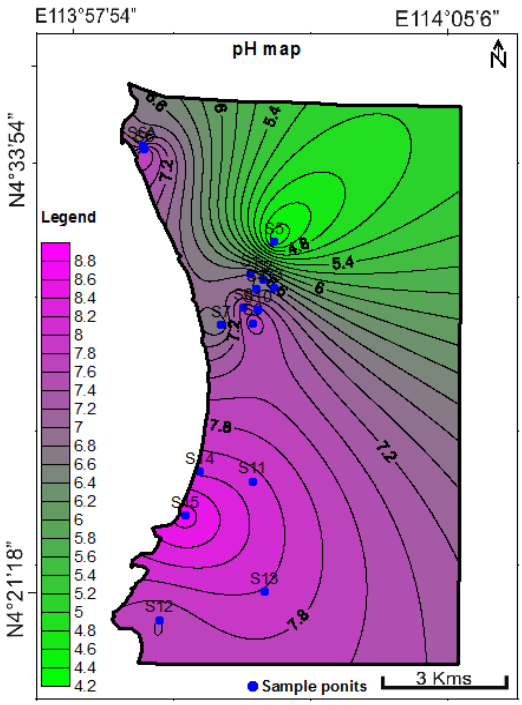



Fig.2: Spatial distribution map for pH

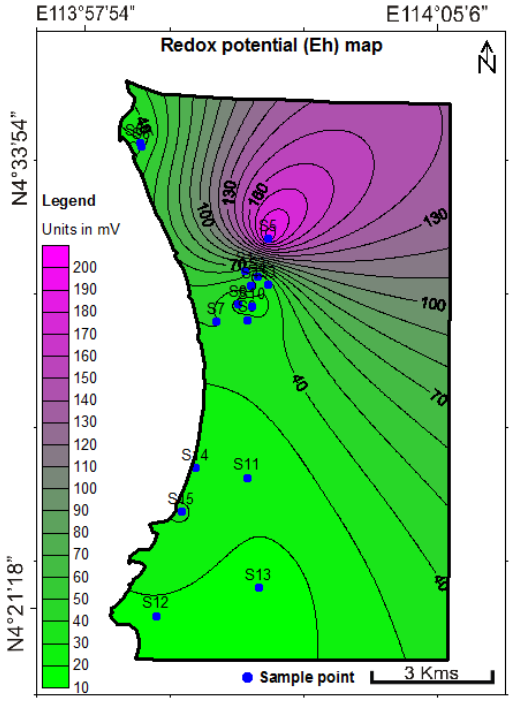



Fig.3: Spatial distribution map for Eh

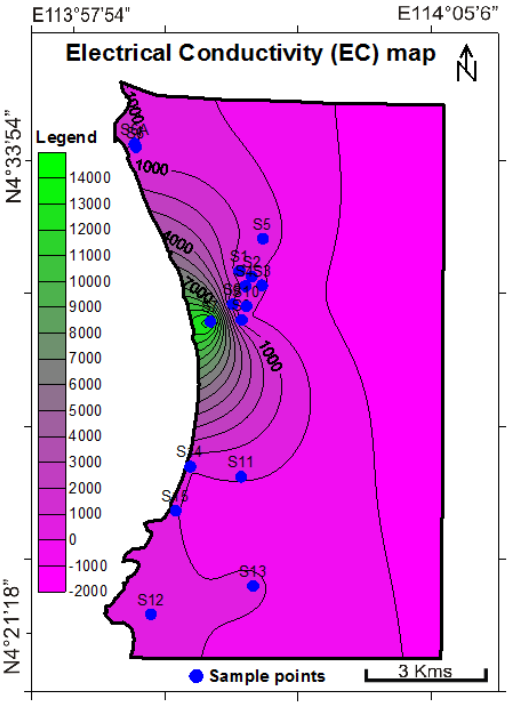



Fig.4: The spatial distribution map for EC in µS/cm

Resistivity value is dependent upon the conductance of water. In surface waters, resistivity values range from 0.0001 MΩ-cm to 0.0200 MΩ-cm, with an average of 0.0060 MΩ-cm. The spatial distribution of the resisitivity map (Fig.5) shows that the low resistivity values were observed in the residental area, which determined the poor water quality zone caused by various anthropogenic activities. In the southern part of the study area, the resistivity values were higher, indicating the non-polluted zone. It was also substantiated with the low values of EC and TDS observed in these regions.

## CONCLUSION

This study shows that the lake water in and around Miri City is neutral to alkaline in nature. All the samples are within the permissible limit, except at S7, which is highly contaminated by waste disposal from the residents. This research study also reveals that the lakes in the central part of the study area are contaminated by various anthropogenic activities and safe in the southern part. In future studies, seasonal monitoring approach is required to explicate the degree of pollution using major ions, metals and organic parameters. This also helps us to judisiously manage the fresh water lakes and to adopt strategies to prevent further pollution in these water bodies.

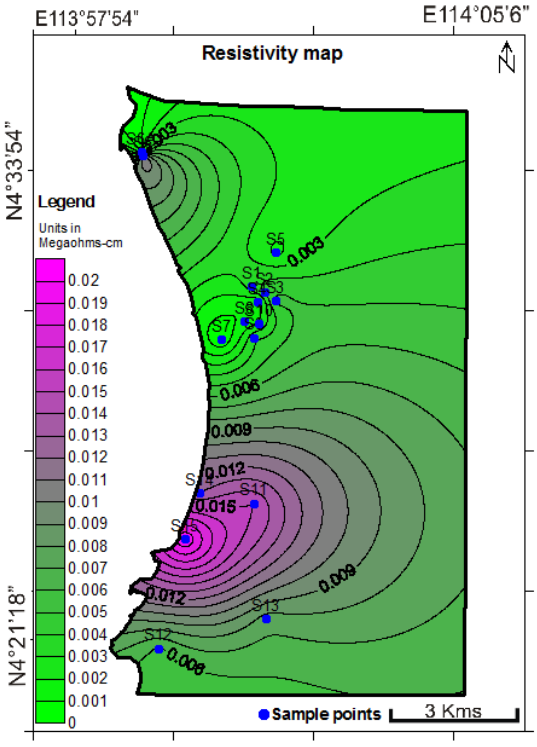



Fig.5: Spatial distribution map for Resistivity

## ACKNOWLEDGEMENTS

This work was supported by Curtin Sarawak Research Fund (CSRF - 2086) Project, Curtin University, Sarawak Malaysia.

## REFERENCES

Dixon, W., & Chiswell, B. (1996). Review of aquatic monitoring program design. *Water Research, 30*, 1935-1948.

Hem, J. D. (1985). *Study and interpretation of the chemical characteristics of natural water* (3rd Edn., p. 2254). Jodhpur, India: Scientific.

Issa, Y. M., Elewa, A. A., Rizk, M. S., & Hassouna, A. F. A. (1996). Distribution of some heavy metals in Qaroun lake and river Nile, Egypt, Menofiya. *Journal of Agricultural Research, 21*, 733-746.

Kaverina, I., & Pogozheva, V. (2005). *Geochemical study of Lake Imandra, Russia. Evaluation of effect of mining industries on the water quality.* Master Thesis, Lulea University of Technology, Sweden, 56p.

Manahan, S. E. (2005). *Environmental Chemistry* (8th Edn). CRC Press LLC, USA. pp. 169-180; 687.

Niemi, G. J., Devore, P., Detenbeck, N., Taylor, D., & Lima, A. (1990). Overview of case studies on recovery of aquatic systems from disturbance. *Environmental Management, 14*, 571-587.

Odo P. E., & Ijere, J. A. (1997). Quality analysis of water from Alau dam area in the north eastern Nigeria for irrigation suitability and some related agronomic and environmental implications. Issues in environmental monitoring in Nigeria. *Nig. Geographical Ass*., 63-67.

Prasanna, M. V., Chidambaram, S., Gireesh, T. V., & Jabir Ali, T. V. (2011). A study on hydrochemical characteristics of surface and sub-surface water in and around Perumal Lake, Cuddalore district, Tamil Nadu, South India. *Environmental Earth Sciences, 63,* 31-47.

Prasanna, M. V., Chidambaram, S., Shahul Hameed, A., & Srinivasamoorthy, K. (2010). Study of evaluation of groundwater in Gadilam basin using hydrogeochemical and isotope data. *Environmental Monitoring and Assessment, 168*, 63-90.

Prasanna, M. V., Nagarajan, R., Chidambaram, S., & Elayaraja, A. (2012a). Assessment of metals distribution and microbial contamination at selected lake waters in and around Miri City, East Malaysia. *Bulletin of Environmental Contamination and Toxicology, 89,* 507-511.

Prasanna, M. V., Praveena, S. M., Chidambaram, S., Nagarajan, R., & Elayaraja, A. (2012b). Evaluation of water quality pollution indices for heavy metal contamination monitoring: a case study from Curtin Lake, Miri City, East Malaysia. *Environmental Earth Sciences, 67,* 1987-2001.

Siddharth, S., Jamal, A., Tiwary, R. K., & Dhar, B. B. (2004). Pridicting drainage quality for sustainable exploitation of coal. *Journal of Scientific & Industrial Research, 63*, 919-926.

Szymanowska, A., Samecka-Cymerman, A., & Kempers, A. J. (2009). Heavy metals in three lakes in West Poland. *Ecotoxicology and Environmental Safety*, *43*, 21-29.

WHO (2004). Guidelines for drinking water quality recommendations (2nd Edn). Geneva: WHO.

# REFEREES FOR THE PERTANIKA
## JOURNAL OF SCIENCE AND TECHNOLOGY

## VOL. 22(1) JAN. 2014

The Editorial Board of the Journal of Science and Technology wishes to thank the following for acting as referees for manuscripts published in this issue of JST.

Abd Manan Samad
*(UiTM, Malaysia)*

Abd Nasir Matori
*(UTP, Malaysia)*

Abdul Ghani Md Rafek
*(UKM, Malaysia)*

Abdul Halim Ghazali
*(UPM, Malaysia)*

Abdul Razak Ismail
*(UTM, Malaysia)*

Adem Kilicman
*(UPM, Malaysia)*

Agbonnaya Chukwu
*(Federal University of Technology, Nigeria)*

Ahmad Zuhairi Abdullah
*(USM, Malaysia)*

Anuar Kassim
*(UPM, Malaysia)*

Anwar Fitrianto
*(UPM, Malaysia)*

Asnuzilawati Asari
*(UMT, Malaysia)*

Baharudin Yatim
*(UKM, Malaysia)*

Buhri Arifin
*(UMT, Malaysia)*

Chan Kit Lam
*(USM, Malaysia)*

Cheah Swee Hung
*(UM, Malaysia)*

Chen Soo Kien
*(UPM, Malaysia)*

Farah Saleena Taip
*(UPM, Malaysia)*

Goh Bean San
*(UM, Malaysia)*

Gwendoline Ee Cheng Lian
*(UPM, Malaysia)*

Gwendoline Ee Chiang Lian
*(UPM, Malaysia)*

Hamid Assilzadeh
*(The University of Calgary, Canada)*

Hasfalina Che Man
*(UPM, Malaysia)*

John Kuna Raj
*(UM, Malaysia)*

Kaniraj Shenbaga
*(UNIMAS, Malaysia)*

Karen Anne Crouse
*(UPM, Malaysia)*

Khali Aziz Hamzah
*(FRIM, Malaysia)*

Kulbir Singh
*(ACED, Thapar University, India)*

Loong Yik Yee
*(UPM, Malaysia)*

Mafauzy Mohamed
*(USM, Malaysia)*

Malay Chaudhuri
*(UTP, Malaysia)*

Md Sohel Rana
*(UPM, Malaysia)*

Megat Mohamad Hamdan Megat Ahmad
*(UPNM, Malaysia)*

Mohamed Hasnain Isa
*(UTP, Malaysia)*

Mohamed Othman
*(UPM, Malaysia)*

Mohammud Che Husain
*(UniMAP, Malaysia)*

Mohd Aspollah Sukari
*(UPM, Malaysia)*

Mohd Nordin Adlan
*(USM, Malaysia)*

Mustafa Mamat
*(UMT, Malaysia)*

Nayan Deep S. Kanwal
*(UPM, Malaysia)*

Nor Ashidi Mat Isa
*(USM, Malaysia)*

Nor Haniza Sarmin
*(UTM, Malaysia)*

Ravi Prakash
*(Jaypee University of Information Technology, India)*

Saied Pirasteh
*(UPM, Malaysia)*

Sanjeev Narayan Sharma
*(Samrat Ashok Technological Institute, India)*

Suchada Chantrapromma
*(Prince of Songkla University, Thailand)*

Tan Wee Tee
*(UPM, Malaysia)*

Tay Meng Guan
*(UNIMAS, Malaysia)*

Teang Shui Lee
*(UPM, Malaysia)*

Umar Hamzah
*(UKM, Malaysia)*

Wan Ishak Wan Ismail
*(UPM, Malaysia)*

Wong Ming
*(UKM, Malaysia)*

Yi Li
*(Chinese Academy of Sciences, China)*

Zuhra Hamzah
*(UKM, Malaysia)*

---

*UPM- Universiti Putra Malaysia*
*USM- Universiti Sains Malaysia*
*UM- Universiti Malaya*
*UKM- Universiti Kebangsaan Malaysia*
*UTM- Universiti Teknologi Malaysia*
*UiTM- Universiti Teknologi MARA Malaysia*
*UNIMAS- Universiti Malaysia Sarawak*
*UPNM- Universiti Pertahanan Nasional Malaysia*
*UTP- Universiti Teknologi PETRONAS*
*UMT- Universiti Malaysia Terengganu*
*UniMAP- Universiti Malaysia Pahang*
*FRIM - Forest Research Institute Malaysia*

---

While every effort has been made to include a complete list of referees for the period stated above, however if any name(s) have been omitted unintentionally or spelt incorrectly, please notify the Chief Executive Editor, *Pertanika* Journals at nayan@upm.my.

Any inclusion or exclusion of name(s) on this page does not commit the *Pertanika* Editorial Office, nor the UPM Press or the University to provide any liability for whatsoever reason.

# *Pertanika*

*Our goal is to bring high quality research to the widest possible audience*

**Journal of Science & Technology**

## INSTRUCTIONS TO AUTHORS

(Manuscript Preparation & Submission Guidelines)

Revised: January 2013

*We aim for excellence, sustained by a responsible and professional approach to journal publishing. We value and support our authors in the research community.*

Please read the guidelines and follow these instructions carefully; doing so will ensure that the publication of your manuscript is as rapid and efficient as possible. The Editorial Board reserves the right to return manuscripts that are not prepared in accordance with these guidelines.

## About the Journal

*Pertanika* is an international peer-reviewed journal devoted to the publication of original papers, and it serves as a forum for practical approaches to improving quality in issues pertaining to tropical agriculture and its related fields. *Pertanika* began publication in 1978 as Journal of Tropical Agricultural Science. In 1992, a decision was made to streamline *Pertanika* into three journals to meet the need for specialised journals in areas of study aligned with the interdisciplinary strengths of the university. The revamped Journal of Science and Technology (JST) is now focusing on research in science and engineering, and its related fields. Other *Pertanika* series include Journal of Tropical Agricultural Science (JTAS); and Journal of Social Sciences and Humanities (JSSH).

JST is published in **English** and it is open to authors around the world regardless of the nationality. It is currently published two times a year i.e. in **January** and **July**.

### Goal of *Pertanika*

Our goal is to bring the highest quality research to the widest possible audience.

### Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions are guaranteed to receive a decision within 12 weeks. The elapsed time from submission to publication for the articles averages 5-6 months.

### Indexing of *Pertanika*

Pertanika is now over 33 years old; this accumulated knowledge has resulted in Pertanika JST being indexed in SCOPUS (Elsevier), EBSCO, Thomson (ISI) Web of Knowledge [CAB Abstracts], DOAJ, Google Scholar, ERA, ISC, Citefactor, Rubriq and MyAIS.

### Future vision

We are continuously improving access to our journal archives, content, and research services. We have the drive to realise exciting new horizons that will benefit not only the academic community, but society itself.

We also have views on the future of our journals. The emergence of the online medium as the predominant vehicle for the 'consumption' and distribution of much academic research will be the ultimate instrument in the dissemination of the research news to our scientists and readers.

## Aims and Scope

*Pertanika* Journal of Science and Technology aims to provide a forum for high quality research related to science and engineering research. Areas relevant to the scope of the journal include: *bioinformatics, bioscience, biotechnology and bio-molecular sciences, chemistry, computer science, ecology, engineering, engineering design, environmental control and management, mathematics and statistics, medicine and health sciences, nanotechnology, physics, safety and emergency management*, and related fields of study.

## Editorial Statement

*Pertanika* is the official journal of Universiti Putra Malaysia. The abbreviation for *Pertanika* Journal of Science & Technology is *Pertanika J. Sci. Technol.*

## Guidelines for Authors

**Publication policies**

*Pertanika* policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. It prohibits as well publication of any manuscript that has already been published either in whole or substantial part elsewhere. It also does not permit publication of manuscript that has been published in full in Proceedings. Please refer to Pertanika's **Code of Ethics** for full details.

**Editorial process**

Authors are notified on receipt of a manuscript and upon the editorial decision regarding publication.

*Manuscript review*: Manuscripts deemed suitable for publication are sent to the Editorial Board members and/or other reviewers. We encourage authors to suggest the names of possible reviewers. Notification of the editorial decision is usually provided within to eight to ten weeks from the receipt of manuscript. Publication of solicited manuscripts is not guaranteed. In most cases, manuscripts are accepted conditionally, pending an author's revision of the material.

*Author approval*: Authors are responsible for all statements in articles, including changes made by editors. The liaison author must be available for consultation with an editor of *The Journal* to answer questions during the editorial process and to approve the edited copy. Authors receive edited typescript (not galley proofs) for final approval. Changes **cannot** be made to the copy after the edited version has been approved.

**Manuscript preparation**

*Pertanika* accepts submission of mainly four types of manuscripts. Each manuscript is classified as **regular** or **original** articles, **short communications**, **reviews**, and proposals for **special issues**. Articles must be in **English** and they must be competently written and argued in clear and concise grammatical English. Acceptable English usage and syntax are expected. Do not use slang, jargon, or obscure abbreviations or phrasing. Metric measurement is preferred; equivalent English measurement may be included in parentheses. Always provide the complete form of an acronym/abbreviation the first time it is presented in the text. Contributors are strongly recommended to have the manuscript checked by a colleague with ample experience in writing English manuscripts or an English language editor.

Linguistically hopeless manuscripts will be rejected straightaway (e.g., when the language is so poor that one cannot be sure of what the authors really mean). This process, taken by authors before submission, will greatly facilitate reviewing, and thus publication if the content is acceptable.

The instructions for authors must be followed. Manuscripts not adhering to the instructions will be returned for revision without review. Authors should prepare manuscripts according to the guidelines of *Pertanika*.

**1. Regular article**

*Definition:* Full-length original empirical investigations, consisting of introduction, materials and methods, results and discussion, conclusions. Original work must provide references and an explanation on research findings that contain new and significant findings.

*Size:* Should not exceed 5000 words or 8-10 printed pages (excluding the abstract, references, tables and/or figures). One printed page is roughly equivalent to 3 type-written pages.

**2. Short communications**

*Definition:* Significant new information to readers of the Journal in a short but complete form. It is suitable for the publication of technical advance, bioinformatics or insightful findings of plant and animal development and function.

*Size:* Should not exceed 2000 words or 4 printed pages, is intended for rapid publication. They are not intended for publishing preliminary results or to be a reduced version of Regular Papers or Rapid Papers.

**3. Review article**

*Definition:* Critical evaluation of materials about current research that had already been published by organizing, integrating, and evaluating previously published materials. Re-analyses as meta-analysis and systemic reviews are encouraged. Review articles should aim to provide systemic overviews, evaluations and interpretations of research in a given field.

*Size:* Should not exceed 4000 words or 7-8 printed pages.

**4. Special issues**

*Definition:* Usually papers from research presented at a conference, seminar, congress or a symposium.

*Size:* Should not exceed 5000 words or 8-10 printed pages.

**5. Others**

*Definition:* Brief reports, case studies, comments, Letters to the Editor, and replies on previously published articles may be considered.

*Size:* Should not exceed 2000 words or up to 4 printed pages.

With few exceptions, original manuscripts should not exceed the recommended length of 6 printed pages (about 18 typed pages, double-spaced and in 12-point font, tables and figures included). Printing is expensive, and, for the Journal, postage doubles when an issue exceeds 80 pages. You can understand then that there is little room for flexibility.

Long articles reduce the Journal's possibility to accept other high-quality contributions because of its 80-page restriction. We would like to publish as many good studies as possible, not only a few lengthy ones. (And, who reads overly long articles anyway?) Therefore, in our competition, short and concise manuscripts have a definite advantage.

**Format**

The paper should be formatted in one column format with at least 4cm margins and 1.5 line spacing throughout. Authors are advised to use Times New Roman 12-point font. Be especially careful when you are inserting special characters, as those inserted in different fonts may be replaced by different characters when converted to PDF files. It is well known that 'μ' will be replaced by other characters when fonts such as 'Symbol' or 'Mincho' are used.

A maximum of eight keywords should be indicated below the abstract to describe the contents of the manuscript. Leave a blank line between each paragraph and between each entry in the list of bibliographic references. Tables should preferably be placed in the same electronic file as the text. Authors should consult a recent issue of the Journal for table layout.

Every page of the manuscript, including the title page, references, tables, etc. should be numbered. However, no reference should be made to page numbers in the text; if necessary, one may refer to sections. Underline words that should be in italics, and do not underline any other words.

We recommend that authors prepare the text as a **Microsoft Word** file.

1. Manuscripts in general should be organised in the following order:

    - **Page 1: Running title.** (Not to exceed 60 characters, counting letters and spaces). This page should **only** contain the running title of your paper. The running title is an abbreviated title used as the running head on every page of the manuscript.

      In addition, the **Subject areas m**ost relevant to the study **must be indicated on this page.** Select the appropriate subject areas from the Scope of the Journals provided in the Manuscript Submission Guide.

      A **list** of n**umber of black and white / colour figures and tables should also be indicated on this page.** Figures submitted in color will be printed in colour. See "*5. Figures & Photographs"* for details.

    - **Page 2: Author(s) and Corresponding author information.** This page should contain the **full title** of your paper with name(s) of all the authors, institutions and corresponding author's name, institution and full address (Street address, telephone number (including extension), hand phone number, fax number and e-mail address) for editorial correspondence. The names of the authors **must** be abbreviated following the international naming convention. e.g. Salleh, A.B., Tan, S.G., or Sapuan, S.M.

      **Authors' addresses.** Multiple authors with different addresses must indicate their respective addresses separately by superscript numbers:

      George Swan[1] and Nayan Kanwal[2]
      [1]Department of Biology, Faculty of Science, Duke University, Durham, North Carolina, USA.
      [2]Office of the Deputy Vice Chancellor (R&I), Universiti Putra Malaysia, Serdang, Malaysia.

    - **Page 3**: This page should **repeat** the **full title** of your paper with only the **Abstract** (the abstract should be less than 250 words for a Regular Paper and up to 100 words for a Short Communication). **Keywords** must also be provided on this page (Not more than eight keywords in alphabetical order).

    - **Page 4 and subsequent pages**: This page should begin with the **Introduction** of your article and the rest of your paper should follow from page 5 onwards.

    **Abbreviations.** Define alphabetically, other than abbreviations that can be used without definition. Words or phrases that are abbreviated in the introduction and following text should be written out in full the first time that they appear in the text, with each abbreviated form in parenthesis. Include the common name or scientific name, or both, of animal and plant materials.

    **Footnotes.** Current addresses of authors if different from heading.

2. **Text.** Regular Papers should be prepared with the headings **Introduction**, **Materials and Methods**, **Results and Discussion, Conclusions** in this order. Short Communications should be prepared according to "*8. Short Communications."* below.

3. **Tables.** All tables should be prepared in a form consistent with recent issues of *Pertanika* and should be numbered consecutively with Arabic numerals. Explanatory material should be given in the table legends and footnotes. Each

table should be prepared on a separate page. (Note that when a manuscript is accepted for publication, tables must be submitted as data - .doc, .rtf, Excel or PowerPoint file- because tables submitted as image data cannot be edited for publication.)

4. **Equations and Formulae.** These must be set up clearly and should be typed triple spaced. Numbers identifying equations should be in square brackets and placed on the right margin of the text.

5. **Figures & Photographs.** Submit an original figure or photograph. Line drawings must be clear, with high black and white contrast. Each figure or photograph should be prepared on a separate sheet and numbered consecutively with Arabic numerals. Appropriate sized numbers, letters and symbols should be used, no smaller than 2 mm in size after reduction to single column width (85 mm), 1.5-column width (120 mm) or full 2-column width (175 mm).

6. Failure to comply with these specifications will require new figures and delay in publication. For electronic figures, create your figures using applications that are capable of preparing high resolution TIFF files acceptable for publication. In general, we require **300 dpi or higher resolution for coloured and half-tone artwork** and **1200 dpi or higher for line drawings**. For review, you may attach low-resolution figures, which are still clear enough for reviewing, to keep the file of the manuscript under 5 MB. Illustrations may be produced at extra cost in colour at the discretion of the Publisher; the author could be charged Malaysian Ringgit 50 for each colour page.

7. **References.** Literature citations in the text should be made by name(s) of author(s) and year. For references with more than two authors, the name of the first author followed by *'et al*.' should be used.

   Swan and Kanwal (2007) reported that …
   The results have been interpreted (Kanwal *et al.* 2009).

   - References should be listed in alphabetical order, by the authors' last names. For the same author, or for the same set of authors, references should be arranged chronologically. If there is more than one publication in the same year for the same author(s), the letters 'a', 'b', etc., should be added to the year.

   - When the authors are more than 11, list 5 authors and then et al.

   - Do not use indentations in typing References. Use one line of space to separate each reference. The name of the journal should be written in full. For example:

     - Jalaludin, S. (1997a). Metabolizable energy of some local feeding stuff. *Tumbuh, 1,* 21-24.
     - Jalaludin, S. (1997b). The use of different vegetable oil in chicken ration. *Malayan Agriculturist, 11,* 29-31.
     - Tan, S. G., Omar, M. Y., Mahani, K. W., Rahani, M., & Selvaraj, O. S. (1994). Biochemical genetic studies on wild populations of three species of green leafhoppers *Nephotettix* from Peninsular Malaysia. *Biochemical Genetics*, *32,* 415 - 422.

   - In case of citing an author(s) who has published more than one paper in the same year, the papers should be distinguished by addition of a small letter as shown above, e.g. Jalaludin (1997a); Jalaludin (1997b).

   - Unpublished data and personal communications should not be cited as literature citations, but given in the text in parentheses. 'In press' articles that have been accepted for publication may be cited in References. Include in the citation the journal in which the 'in press' article will appear and the publication date, if a date is available.

8. **Examples of other reference citations**:

   ***Monographs***: Turner, H. N., & Yong, S. S. Y. (2006). *Quantitative Genetics in Sheep Breeding*. Ithaca: Cornell University Press.

   ***Chapter in Book***: Kanwal, N. D. S. (1992). Role of plantation crops in Papua New Guinea economy. In Angela R. McLean (Ed.), *Introduction of livestock in the Enga province PNG* (p. 221-250). United Kingdom: Oxford Press.

   ***Proceedings***: Kanwal, N. D. S. (2001). Assessing the visual impact of degraded land management with landscape design software. In Kanwal, N. D. S., & Lecoustre, P. (Eds.), *International forum for Urban Landscape Technologies* (p. 117-127). Lullier, Geneva, Switzerland: CIRAD Press.

9. **Short Communications** should include **Introduction, Materials and Methods, Results and Discussion, Conclusions** in this order. Headings should only be inserted for Materials and Methods. The abstract should be up to 100 words, as stated above. Short Communications must be 5 printed pages or less, including all references, figures and tables. References should be less than 30. A 5 page paper is usually approximately 3000 words plus four figures or tables (if each figure or table is less than 1/4 page).

   *Authors should state the total number of words (including the Abstract) in the cover letter. Manuscripts that do not fulfill these criteria will be rejected as Short Communications without review.

**STYLE OF THE MANUSCRIPT**

Manuscripts should follow the style of the latest version of the Publication Manual of the American Psychological Association (APA). The journal uses American or British spelling and authors may follow the latest edition of the Oxford Advanced Learner's Dictionary for British spellings.

**SUBMISSION OF MANUSCRIPTS**

All articles should be submitted electronically using the ScholarOne web-based system. ScholarOne, a Thomson Reuters product provides comprehensive workflow management systems for scholarly journals. For more information, go to our web page and click "**Online Submission**".

Alternatively, you may submit the electronic files (cover letter, manuscript, and the **Manuscript Submission Kit** comprising *Declaration* and *Referral* forms) via email directly to the Executive Editor. If the files are too large to email, mail a CD containing the files. The **Manuscript Submission Guide** and **Submission Kit** are available from the *Pertanika*'s home page at http://www.pertanika.upm.edu.my/ or from the Executive Editor's office upon request.

All articles submitted to the journal **must comply** with these instructions. Failure to do so will result in return of the manuscript and possible delay in publication.

Please do **not** submit manuscripts to the editor-in-chief or to any other office directly. All manuscripts must be **submitted through the executive editor's office** to be properly acknowledged and rapidly processed at the address below:

Dr. Nayan KANWAL
The Executive Editor
Pertanika Journals, UPM Press
Office of the Deputy Vice Chancellor (R&I)
IDEA Tower II, UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 UPM, Serdang, Selangor
Malaysia

E-mail: executive_editor.pertanika@upm.my; tel: + 603-8947 1622.
or visit our website at http://www.pertanika.upm.edu.my/home.php for further information.

Authors should retain copies of submitted manuscripts and correspondence, as materials can not be returned. Authors are required to inform the Executive Editor of any change of address which occurs whilst their papers are in the process of publication.

**Cover letter**

All submissions must be accompanied by a cover letter detailing what you are submitting. Papers are accepted for publication in the journal on the understanding that the article is original and the content has not been published or submitted for publication elsewhere. This must be stated in the cover letter.

The cover letter must also contain an acknowledgement that all authors have contributed significantly, and that all authors are in agreement with the content of the manuscript.

The cover letter of the paper should contain (i) the title; (ii) the full names of the authors; (iii) the addresses of the institutions at which the work was carried out together with (iv) the full postal and email address, plus facsimile and telephone numbers of the author to whom correspondence about the manuscript should be sent. The present address of any author, if different from that where the work was carried out, should be supplied in a footnote.

As articles are double-blind reviewed, material that might identify authorship of the paper should be placed on a cover sheet.

**Peer review**

*Pertanika* follows a **double-blind peer-review** process. Peer reviewers are experts chosen by journal editors to provide written assessment of the **strengths** and **weaknesses** of written research, with the aim of improving the reporting of research and identifying the most appropriate and highest quality material for the journal.

In the peer-review process, three referees independently evaluate the scientific quality of the submitted manuscripts. Authors are encouraged to indicate in the **Referral form** using the **Manuscript Submission Kit** the names of three potential reviewers, but the editors will make the final choice. The editors are not, however, bound by these suggestions.

Manuscripts should be written so that they are intelligible to the professional reader who is not a specialist in the particular field. They should be written in a clear, concise, direct style. Where contributions are judged as acceptable for publication on the basis of content, the Editor reserves the right to modify the typescripts to eliminate ambiguity and repetition and improve communication between author and reader. If extensive alterations are required, the manuscript will be returned to the author for revision.

**The Journal's review process**

What happens to a manuscript once it is submitted to *Pertanika*? Typically, there are seven steps to the editorial review process:

1. The executive editor and the editorial board examine the paper to determine whether it is appropriate for the journal and should be reviewed. If not appropriate, the manuscript is rejected outright and the author is informed.

2. The executive editor sends the article-identifying information having been removed, to three reviewers. Typically, one of these is from the Journal's editorial board. Others are specialists in the subject matter represented by the article. The executive editor asks them to complete the review in three weeks and encloses two forms: (a) referral form B and (b) reviewer's comment form along with reviewer's guidelines. Comments to authors are about the appropriateness and adequacy of the theoretical or conceptual framework, literature review, method, results and discussion, and conclusions. Reviewers often include suggestions for strengthening of the manuscript. Comments to the editor are in the nature of the significance of the work and its potential contribution to the literature.

3. The executive editor, in consultation with the editor-in-chief, examines the reviews and decides whether to reject the manuscript, invite the author(s) to revise and resubmit the manuscript, or seek additional reviews. Final acceptance or rejection rests with the Editorial Board, who reserves the right to refuse any material for publication. In rare instances, the manuscript is accepted with almost no revision. Almost without exception, reviewers' comments (to the author) are forwarded to the author. If a revision is indicated, the editor provides guidelines for attending to the reviewers' suggestions and perhaps additional advice about revising the manuscript.

4. The authors decide whether and how to address the reviewers' comments and criticisms and the editor's concerns. The authors submit a revised version of the paper to the executive editor along with specific information describing how they have answered' the concerns of the reviewers and the editor.

5. The executive editor sends the revised paper out for review. Typically, at least one of the original reviewers will be asked to examine the article.

6. When the reviewers have completed their work, the executive editor in consultation with the editorial board and the editor-in-chief examine their comments and decide whether the paper is ready to be published, needs another round of revisions, or should be rejected.

7. If the decision is to accept, the paper is sent to that Press and the article should appear in print in approximately three months. The Publisher ensures that the paper adheres to the correct style (in-text citations, the reference list, and tables are typical areas of concern, clarity, and grammar). The authors are asked to respond to any queries by the Publisher. Following these corrections, page proofs are mailed to the corresponding authors for their final approval. At this point, only essential changes are accepted. Finally, the article appears in the pages of the Journal and is posted on-line.

**English language editing**

Pertanika **emphasizes** on the linguistic accuracy of every manuscript published. Thus all authors are required to get their manuscripts edited by **professional English language editors**. Author(s) **must provide a certificate** confirming that their manuscripts have been adequately edited. A proof from a recognised editing service should be submitted together with the cover letter at the time of submitting a manuscript to Pertanika**. All costs will be borne by the author(s)**.

This step, taken by authors before submission, will greatly facilitate reviewing, and thus publication if the content is acceptable.

## Author material archive policy

Authors who require the return of any submitted material that is rejected for publication in the journal should indicate on the cover letter. If no indication is given, that author's material should be returned, the Editorial Office will dispose of all hardcopy and electronic material.

**Copyright**

Authors publishing the Journal will be asked to sign a declaration form. In signing the form, it is assumed that authors have obtained permission to use any copyrighted or previously published material. All authors must read and agree to the conditions outlined in the form, and must sign the form or agree that the corresponding author can sign on their behalf. Articles cannot be published until a signed form has been received.

**Lag time**

A decision on acceptance or rejection of a manuscript is reached in 3 to 4 months (average 14 weeks). The elapsed time from submission to publication for the articles averages 5-6 months.

**Hardcopies of the Journals and off prints**

Under the Journal's open access initiative, authors can choose to download free material (via PDF link) from any of the journal issues from Pertanika's website. Under "Browse Journals" you will see a link entitled "Current Issues" or "Archives". Here you will get access to all back-issues from 1978 onwards.

The **corresponding author** for all articles will receive one complimentary hardcopy of the journal in which his/her articles is published. In addition, 20 off prints of the full text of their article will also be provided. Additional copies of the journals may be purchased by writing to the executive editor.

## Why should you publish in *Pertanika*?

### BENEFITS TO AUTHORS

**PROFILE**: Our journals are circulated in large numbers all over Malaysia, and beyond in Southeast Asia. Our circulation covers other overseas countries as well. We ensure that your work reaches the widest possible audience in print and online, through our wide publicity campaigns held frequently, and through our constantly developing electronic initiatives such as Web of Science Author Connect backed by Thomson Reuters.

**QUALITY**: Our journals' reputation for quality is unsurpassed ensuring that the originality, authority and accuracy of your work is fully recognised. Each manuscript submitted to Pertanika undergoes a rigid originality check. Our double-blind peer refereeing procedures are fair and open, and we aim to help authors develop and improve their scientific work. Pertanika is now over 35 years old; this accumulated knowledge has resulted in our journals being indexed in SCOPUS (Elsevier), Thomson (ISI) Web of Knowledge [BIOSIS & CAB Abstracts], EBSCO, DOAJ, Google Scholar, AGRICOLA, ERA, ISC, Citefactor, Rubriq and MyAIS.

**AUTHOR SERVICES**: We provide a rapid response service to all our authors, with dedicated support staff for each journal, and a point of contact throughout the refereeing and production processes. Our aim is to ensure that the production process is as smooth as possible, is borne out by the high number of authors who prefer to publish with us.

**CODE OF ETHICS**: Our Journal has adopted a Code of Ethics to ensure that its commitment to integrity is recognized and adhered to by contributors, editors and reviewers. It warns against plagiarism and self-plagiarism, and provides guidelines on authorship, copyright and submission, among others.

**PRESS RELEASES**: Landmark academic papers that are published in Pertanika journals are converted into press releases as a unique strategy for increasing visibility of the journal as well as to make major findings accessible to non-specialist readers. These press releases are then featured in the university's UK-based research portal, ResearchSEA, for the perusal of journalists all over the world.

**LAG TIME**: The elapsed time from submission to publication for the articles averages 4 to 5 months. A decision on acceptance of a manuscript is reached in 3 to 4 months (average 14 weeks).


An **Award-Winning** International-Malaysian Journal —MAY 2014

## About the Journal


***Pertanika*** is an international multidisciplinary peer-reviewed leading journal in Malaysia which began publication in 1978. The journal publishes in three different areas — Journal of Tropical Agricultural Science (JTAS); Journal of Science and Technology (JST); and Journal of Social Sciences and Humanities (JSSH).

**JTAS** is devoted to the publication of original papers that serves as a forum for practical approaches to improving quality in issues pertaining to **tropical agricultural research-** or related fields of study. It is published four times a year in **February**, **May**, **August** and **November**.

**JST** caters for **science and engineering research-** or related fields of study. It is published twice a year in **January** and **July**.

**JSSH** deals in **research or theories in social sciences and humanities research**. It aims to develop as a flagship journal with a focus on emerging issues pertaining to the social and behavioural sciences as well as the humanities, particularly in the Asia Pacific region. It is published four times a year in **March**, **June**, **September** and **December**.

PERTANIKA JOURNALS

## Call for Papers 2014-15

now accepting submmissions...

***Pertanika*** invites you to explore frontiers from all key areas of **agriculture**, **science and technology** to **social sciences and humanities**.

Original research and review articles are invited from scholars, scientists, professors, post-docs, and university students who are seeking publishing opportunities for their research papers through the Journal's three titles; JTAS, JST & JSSH. Preference is given to the work on leading and innovative research approaches.

***Pertanika*** is a fast track peer-reviewed and open-access academic journal published by **Universiti Putra Malaysia**. To date, Pertanika Journals have been indexed by many important databases. Authors may contribute their scientific work by publishing in UPM's hallmark SCOPUS & ISI indexed journals.

Our journals are open access - international journals. Researchers worldwide will have full access to all the articles published online and be able to download them with **zero subscription fee**.

***Pertanika*** uses online article submission, review and tracking system for quality and quick review processing backed by Thomson Reuter's ScholarOne™. Journals provide rapid publication of research articles through this system.

For details on the Guide to Online Submissions, visit
http://www.pertanika.upm.edu.my/guide_online_submission.php

Questions regarding submissions should only be directed to the **Chief Executive Editor**, Pertanika Journals.

Remember, ***Pertanika*** is the resource to support you in strengthening research and research management capacity.


*Address your submissions to:*
**The Chief Executive Editor**
Tel: +603 8947 1622
nayan@upm.my

Journal's profile: www.pertanika.upm.edu.my

**Selected Articles from CUTSE International Conference 2011**

**Pertanika Journal of Science & Technology**
**Vol. 22 (1) Jan. 2014**

## Contents



**Pertanika Editorial Office**
Office of the Deputy Vice Chancellor (R&I),
1st Floor, IDEA Tower II,
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan
Malaysia

http://www.pertanika.upm.edu.my/
E-mail: executive_editor.pertanika@upm.my
Tel: +603 8947 1622/1620

PENERBIT UPM PRESS
http://penerbit.upm.edu.my
E-mail : penerbit@putra.upm.edu.my
Tel : +603 8946 8855/8854
Fax : +603 8941 6172

ISSN 0128-7680
01>