Journal Information

Maejo International Journal of Science and Technology (ISSN 1905-7873 © 2012), the international journal for preliminary communications in Science and Technology is the first peerrefereed scientific journal of Maejo University (www.mju.ac.th). Intended as a medium for communication, discussion, and rapid dissemination of important issues in Science and Technology, articles are initially published online in an open access format, which thereby gives authors the chance to communicate with a wide range of readers in an international community.

Publication Information

MIJST is published triannually. Articles are available online and can be accessed free of charge at http://www.mijst.mju.ac.th. Printed and bound copies of each volume are produced and distributed to authors and selected groups or individuals.

Abstracting/Indexing Information

MIJST is covered and cited by Science Citation Index Expanded, SCOPUS, Journal Citation Reports/Science Edition, Zoological Record, Chemical Abstracts Service (CAS), SciFinder Scholar, Directory of Open Access Journals (DOAJ), CAB Abstracts, ProQuest, and Google Scholar.

Contact Information

Editorial office: Maejo International Journal of Science and Technology (MIJST), 1st floor, Orchid Building, Maejo University, San Sai, Chiang Mai 50290, Thailand

Tel: +66-53-87-3880
E-mails: duang@mju.ac.th

MAEJO INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY

MAEJO INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY

The International Journal for the Publication of Preliminary Communications in Science and Technology

http://www.mijst.mju.ac.th
ISSN 1905-7873

MAEJO INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY

Volume 6, Issue 3 (September - December 2012)

CONTENTS

Page

http://www.mijst.mju.ac.th
ISSN 1905-7873

MAEJO INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY

Volume 6, Issue 3 (September - December 2012)

Author Index

Instructions for Authors

A proper introductory e-mail page containing the title of the submitted article and certifying its originality should be sent to the editor (Duang Buddhasukh, e-mail : duang@mju.ac.th). The manuscript proper together with a list of suggested referees should be attached in separate files. The list should contain at least 5 referees with appropriate expertise. Three referees should be non-native from 3 different countries. Each referee's academic/professional position, scientific expertise, affiliation and e-mail address must be given. The referees should not be affiliated to the same university/institution as any of the authors, nor should any two referees come from the same university/institution. The editorial team, however, retain the sole right to decide whether or not the suggested referees are approached.

Failure to conform to the above instructions will result in non-consideration of the submission.

Please also ensure that English and style is properly edited before submission. UK style of spelling should be used. Authors who would like to consult a professional service can visit www.proof-reading-service.com, www.editage.com, www.enago.com, www.bioedit.co.uk (bioscience and medical papers), www.bioscienceeditingsolutions.com, www.scribendi.com, www.letpub.com, www.papersconsulting.com or www.stickleredit ing.com.

Important : Manuscript with substandard English and style will not be considered.

Warning : Plagiarism (including self-plagiarism) may be checked for at *the last* stage of processing and, if detected, will result in a rejection and blacklisting.

Manuscript Preparation

Manuscripts must be prepared in English using a word processor. MS Word for Macintosh or Windows, and .doc or .rtf files are preferred. Manuscripts may be prepared with other software provided that the full document (with figures, schemes and tables inserted into the text) is exported to a MS Word format for submission. Times or Times New Roman font is preferred. The font size should be 12 pt and the line spacing 'at least 17 pt'. A4 paper size is used and margins must be 1.5 cm on top, 2.0 cm at the bottom and 2.0 cm on both left and right sides of the paper. Although our final output is in .pdf format, authors are asked NOT to send manuscripts in this format as editing them is much more complicated. Under the above settings, a manuscript submitted should not be longer than **15 pages** for a full paper or **20 pages** for a review paper.

A template file may be downloaded from the *Maejo Int. J. Sci. Technol.* homepage. (DOWNLOAD HERE)

Authors' full mailing addresses, homepage addresses, phone and fax numbers, and e-mail addresses homepages can be included in the title page and these will be published in the manuscripts and the Table of Contents. The corresponding author should be clearly identified. It is the corresponding author's responsibility to ensure that all co-authors are aware of and approve of the contents of a submitted manuscript.

A brief (200 word maximum) Abstract should be provided. The use in the Abstract of numbers to identify compounds should be avoided unless these compounds are also identified by names.

A list of three to five keywords must be given and placed after the Abstract. Keywords may be single words or very short phrases.

Although variations in accord with contents of a manuscript are permissible, in general all papers should have the following sections: Introduction, Materials and Methods, Results and Discussion, Conclusions, Acknowledgments (if applicable) and References.
Authors are encouraged to prepare Figures and Schemes in colour. Full colour graphics will be published free of charge.

Tables and Figures should be inserted into the main text, and numbers and titles supplied for all Tables and Figures. All table columns should have an explanatory heading. To facilitate layout of large tables, smaller fonts may be used, but in no case should these be less than 10 pt in size. Authors should use the Table option of MS Word to create tables, rather than tabs, as tab-delimited columns are often difficult to format in .pdf for final output.
Figures, tables and schemes should also be placed in numerical order in the appropriate place within the main text. Numbers, titles and legends should be provided for all tables, schemes and figures. Chemical structures and reaction schemes should be drawn using an appropriate software package designed for this purpose. As a guideline, these should be drawn to a scale such that all the details and text are clearly legible when placed in the manuscript (i.e. text should be no smaller than 8-9 pt).

For bibliographic citations, the reference numbers should be placed in square brackets, i.e. [ ], and placed before the punctuation, for example [4] or [1-3], and all the references should be listed separately and as the last section at the end of the manuscript.

Format for References

Journal :

1. D. Buddhasukh, J. R. Cannon, B. W. Metcalf and A. J. Power, "Synthesis of 5-n-alkylresorcinol dimethyl ethers and related compounds *via* substituted thiophens", *Aust. J. Chem.*, **1971**, *24*, 2655-2664.

Text :

2. A. I. Vogel, "A Textbook of Practical Organic Chemistry", 3rd Edn., Longmans, London, **1956,** pp. 130-132.

Chapter in an edited text :

3. W. Leistritz, "Methods of bacterial reduction in spices", in "Spices: Flavor Chemistry and Antioxidant Porperties" (Ed. S. J. Risch and C-T. Ito), American Chemical Society, Washington, DC, **1997**, Ch. 2.

Thesis / Dissertation :

4. W. phutdhawong, "Isolation of glycosides by electrolytic decolourisation and synthesis of pentinomycin", *PhD. Thesis*, **2002**, Chiang Mai University, Thailand.

Patent :

5. K. Miwa, S. Maeda and Y. Murata, "Purification of stevioside by electrolysis", *Jpn. Kokai Tokkyo Koho 79 89,066* (**1979**).

Proceedings :

6. P. M. Sears, J. Peele, M. Lassauzet and P. Blackburn, "Use of antimicrobial proteins in the treatment of bovine mastitis", Proceedings of the 3rd International Mastitis Seminars, **1995**, Tel-Aviv, Israel, pp. 17-18.

(Download Submission Checklist)

Manuscript Revision Time

Authors who are instructed to revise their manuscript should do so within **45** days. Otherwise the revised manuscript will be regarded as a new submission.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 334-343

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Short Communication

Integrating association rules and case-based reasoning to predict retinopathy

Vimala Balakrishnan [1,*], Mohammad R. Shakouri [2] and Hooman Hoodeh [2]

[1] Department of Information System, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[2] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

* Corresponding author, e-mail: vim.balakrishnan@gmail.com

Received: 19 August 2011 / Accepted: 11 September 2012 / Published: 11 September 2012

Abstract: This study proposes a retinopathy prediction system based on data mining, particularly association rules using Apriori algorithm, and case-based reasoning. The association rules are used to analyse patterns in the data set and to calculate retinopathy probability whereas case-based reasoning is used to retrieve similar cases. This paper discusses the proposed system. It is believed that great improvements can be provided to medical practitioners and also to diabetics with the implementation of this system.

Keywords: predictive system, data mining, association rules, Apriori algorithm, case-based reasoning

INTRODUCTION

Diabetes is a major chronic metabolic disorder which is characterised by persistent hyperglycaemia (elevated blood glucose). It has been estimated that the total number of people in the world with diabetes will rise from 171 million in 2000 to 366 million in 2030 [1]. With a growing number of people diagnosed with diabetes, Malaysia is also experiencing the same phenomenon, as prevalence of the disease stands at 14.9% of adult population. It has been reported that one in six adult Malaysians above 30 years of age have diabetes. Interestingly, the World Health Organisation (WHO) has projected that Malaysia will have a total of 2.48 million people with diabetes by 2030 [2].

Diabetics are prone to develop various complications, especially when the disease is not well controlled. The major complications are generally categorised as micro-vascular (diabetic

nephropathy, retinopathy and neuropathy) and macro-vascular (coronary artery disease, stroke and peripheral arterial disease) [3]. Diabetic retinopathy (DR) or simply retinopathy is one of the most common micro-vascular complications of diabetes. It is a sight threatening complication that affects the retina and is a leading cause of blindness among the working age population [4]. According to WHO, approximately 4.8% of cases of blindness globally are due to retinopathy [5]. A study conducted at the Ophthalmology Clinic, Medical Centre, University of Malaya showed the overall prevalence of retinopathy to be 51.6% [6], which was higher than the previous prevalence rate of 48.6% reported from University Sains Malaysia Hospital in 1996 [7]. The severity of retinopathy can be categorised into five levels: no-DR, mild non-proliferative diabetic retinopathy (NPDR), moderate NPDR, severe NPDR and proliferative diabetic retinopathy (PDR) [8]. In 2007, the National Eye Database in Malaysia reported that 36.8% of 10,856 registered diabetic patients were inflicted with at least one of these severity levels [9].

Diabetes and its complications are becoming increasingly prevalent worldwide and they impose a heavy burden on the health system in any country. In Malaysia about 14.5 billion MYR (USD 4.75 billion) was estimated for 60,000 diabetic patients per year that were registered with the Ministry of Health [10]. Current methods of detecting, screening and monitoring retinopathy are based on subjective human evaluation, which is slow and time-consuming. The high prevalence and severity of retinopathy suggests the need for a screening programme that can recognise it as early as possible. Early detection becomes more important since retinopathy can be asymptomatic even in its more advanced stages. Medical guidelines suggest that Type 2 diabetics should have a comprehensive eye examination shortly after the diagnosis as retinopathy is often already present then [11]. This paper aims to propose a prediction system for retinopathy among diabetic patients. The system is mainly intended for the physicians as it is expected to simplify their decision-making process.

LITERATURE REVIEW

The majority of work involving retinopathy were related to improving algorithms to perform image processing in order to diagnose retinopathy. The current retinopathy diagnostic method involves the use of seven-field stereo fundus photography reviewed by a trained reader. Research has shown that combining fundus photography and computer algorithms improves diagnostic performance. For example, Sanchez et al. [12] came up with an automatic image processing algorithm to detect hard exudates (bright lipids leaked from a blood vessel) based on Fisher's linear discriminant analysis. Hann et al. [13] developed a computer-vision method of isolating and detecting two of the most common retinopathy dysfunctions, i.e. dot haemorrhages and exudates, using specific colour channels and segmentation methods to separate these retinopathy manifestations from physiological features in the digital fundus images. Other studies related to enhancing algorithms for detecting retinopathy were done using artificial neural network method [14], computer-based classification methods [15], and rule-based classifiers [16], among others.

As for prediction, most work in the literature focused on determining significant predictors of retinopathy in a diabetic patient. Statistical tests were found to be commonly used for this purpose. For instance, Semeraro et al. [17] used various statistical measures such as Kaplan-Meier method to

generate univariate survival curves, which were then compared among themselves using the logrank test, U-statistics, regression analysis and Cox analysis, among others, to identify patients who are at a higher risk for retinopathy. Their results showed that the duration of diabetes, glycosylated hemoglobin, systolic blood pressure, gender (male), albuminuria and diabetes therapy were significantly associated with the occurrence of retinopathy. Similar work was carried out by Cho et al. [18] who assessed the diagnostic efficacy of macular and peripapillary retinal thickness measurements for the staging of retinopathy and the prediction of its progression.

Another study that focused on determining retinopathy predictors was conducted in Taiwan. Chan et al. [19] compared the performance of two data-mining methods, namely C5.0 and neural network, in identifying key predictors for diabetic retinopathy, nephropathy and neuropathy. In the C5.0 method, data with diabetes duration of more than seven years were used to generate 22 rules needed for prediction. On the other hand, with the neural network method, retinopathy predictions were made based on a hidden layer with 52 neurons. The sensitivity and specificity for retinopathy prediction was found to be 58.62 and 74.73 respectively using C5.0, whereas the values were 59.48 and 99.86 respectively for neural network, indicating that the latter method was better in determining retinopathy predictors. Their results revealed ten key predictors based on the number of occurrences in the rules, with creatinine emerging as the most important predictor, followed by diabetes duration and family history.

To the best of our knowledge, only the work of Skevofilakas et al. [20] focused on developing a decision support system to predict the risk of retinopathy occurrences among Type-1 diabetic patients. The system was built by combining a feed-forward neural network, a classification and regression tree and a rule induction C5.0 classifier with an improved hybrid wavelet neural network. The data from 55 Type-1 diabetic patients were used to test the system, which resulted in a performance with an accuracy of 98%. It is to be noted that in determining retinopathy occurrences the authors only used seven risk factors, i.e. age, diabetes duration, glycated hemoglobin, cholesterol, triglycerides, hypertension incidence rate and their treatment duration. Various studies have revealed that factors such as gender, smoking habit and even race or ethnicity play major roles in retinopathy occurrences [6, 17-19]. In addition, the authors only focused on Type-1 patients.

The literature also revealed some other methods of identifying diabetic patients at risk of developing retinopathy, e.g. electroretinogram (ERG) [21], multifocal ERG [22] and visual evoked potentials [23]. However, none of these methods have fully entered the clinical practice due to the high cost of instrumentation and the difficulty of carrying out an examination by both the examiners and the patients [17].

In summary, the literature review has revealed that most previous studies related to retinopathy predictions were more inclined in determining the significant risk factors/predictors associated with retinopathy occurrences [17-19]. Study related to retinopathy prediction and machine learning techniques was limited to that of Skevofilakas and coworkers [20], although it is not without its shortcomings as indicated above. The rest of the techniques were reported to be either costly or difficult for both examiners and patients [21-23]. Therefore, the present study proposes to develop a system for prediction of retinopathy resulting from both Type-1 and Type-2 diabetics by integrating data mining, particularly association rules generated using Apriopri

algorithm, and case-based reasoning (CBR). In addition, more important risk factors are identified and included so as to ensure an accurate prediction. We have conducted a similar study in predicting retinopathy by focusing on the use of C5.0, K-nearest neighbour and Hamming algorithms to make the prediction [24]. In this communication, we integrate association rules generated using Apriopri algorithm instead of decision trees generated by C5.0.

PROPOSED METHODOLOGY

Figure 1 illustrates the proposed methodology for this study, which can be segregated into three phases. Phase one involves the data collection, followed by data analysis and data pre-processing in phase two. In addition, rules required to predict retinopathy are also generated in this phase. Finally, phase 3 involves the use of the system to predict retinopathy. The case database is accessed to retrieve similar cases for the retinopathy predictions. The following sections elaborate these phases.

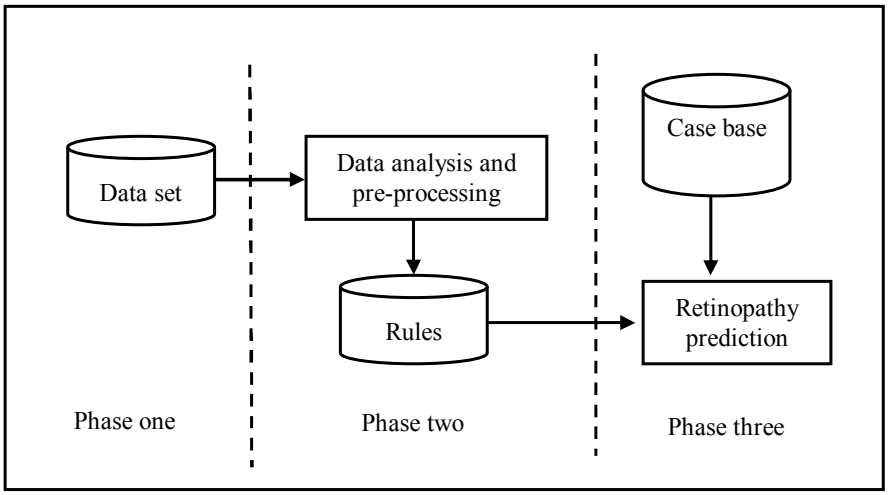


Figure 1. Proposed research methodology

Data Set

The data collection for the study has been accomplished. These data were obtained from the University of Malaya Hospital, comprising Type-1 and Type-2 diabetic patients. The data are needed to generate the necessary rules for retinopathy prediction. Medical experts were interviewed to help assist in understanding the nature of the data collected and also to determine the important risk factors for predicting retinopathy. The literature shows many risk factors that can be used to predict retinopathy, such as body mass index (BMI), smoking history, age and diabetes duration [6, 17-20]. These variables were then reviewed and validated by three medical experts who confirmed that 16 risk factors are crucial for retinopathy prediction. These are BMI, high-density lipoprotein, triglyceride, diabetes duration, glycated hemoglobin, hypertension, age, cholesterol, low-density lipoprotein, alanine aminotranferease, aspartate aminotranferease, cardiac complication, gender, race, smoking and alcohol consumption.

Data Pre-processing

The next step in the study is data pre-processing. Figure 2 shows the overall process involved in preparing the data for rule generation. Before using the data mining algorithm, the data collected may need to be cleaned and filtered to avoid the creation of inappropriate or inaccurate rules [25]. Some of the actions in the data pre-processing are removing duplicate records, normalising the values used to represent information in the database, accounting for missing data points and removing unneeded data fields [25]. The literature revealed four different approaches to clean a data set [26]: parsing, data transformation, duplication elimination and statistical methods. The current study intends to use data transformation, duplication elimination and statistical methods for data cleaning, if necessary. In data transformation the subjected variables in a database are transformed into a right format to fit in the data mining process. For example, in this study continuous attributes such as hypertension are categorised into classes for the purpose of association mining and building decision trees. Duplication elimination deletes duplicate records from the data sets whereas statistical methods analyse the data sets to identify false data. Once the data set is cleaned, the system specifies the explicit relationship between the variables in the data set using association rule mining algorithm.

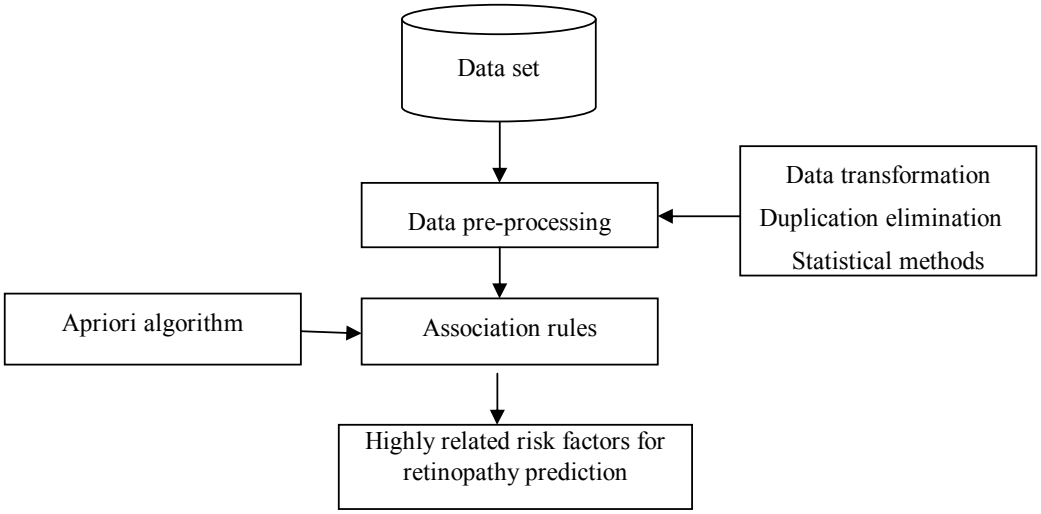


Figure 2. Overall flow of association rules generation

Association Rules

Association rules find interesting associations and/or relationships among a large set of data items. They capture all possible rules that explain the presence of some attributes in relation to the presence of other attributes. An association rule can be expressed in the form of X → Y, whereby X and Y are disjointed item sets. An association rule must satisfy two important measures, namely support and confidence. Support indicates how frequently the items in the rule occur together whereas confidence determines how frequent items in Y appear in transactions containing X [27].

It is important to analyse the relationships among the risk factors and Apriori algorithm is used for this purpose. It is an powerful algorithm for mining frequent item sets for association rules.

It involves two stages. The first stage identifies frequent item sets and the second derives the association rules from the identified frequent item sets [28]. Frequent item sets identification in the first stage is accomplished by scanning over the data multiple times and counting the support of the individual item. Support is calculated by dividing the number of rules in which the item set is found by the total number of transactions. In each subsequent pass a set of item sets found to be frequent in the previous pass are used to generate new frequent item sets, and their support is calculated again. At the end of the pass, item sets satisfying minimum support thresholds are collected and they become the seed for the next pass. This process is repeated until no new frequent item sets are found.

The second stage is to generate the desired association rules from the frequent item sets. This is accomplished by calculating the confidence for each frequent item set. The confidence is the number of times a condition in the most frequent item sets is true over the whole number of frequent item sets [28]. The result of these stages will be a set of highly related risk factors that can be used to predict retinopathy. These association rules are to be first selected based on their support and confidence values (e.g. above 0.9) and will then be verified by the medical experts to ensure an accurate retinopathy prediction.

Case-based Reasoning

Case-based reasoning (CBR) is to be used once the system has calculated the risk of retinopathy occurrence in a patient. CBR is a technique which operates based on older cases. When a new case arrives, a CBR system retrieves similar cases and adapts or justifies the new case according to old cases. Systems developed using CBR can usually predict, diagnose and even suggest solutions for a problem [29]. In CBR systems, retrieving valuable cases from the database is very important. In this study the cases obtained from the University of Malaya Hospital are to be stored in a case database. It is important for these cases to contain at least all the 16 variables needed for the prediction. The indexing function helps to index cases according to their critical risk factors (e.g. demographic data, duration of diabetic and HbA1c) and the risk factors' weights. The better the indexing is, the more accurate the results will be.

The overall flow of prediction is depicted in Figure 3. The retinopathy risk evaluator receives its input when the medical expert enters the patient's data (age, duration of diabetics, HbA1c, etc.). These inputs will be analysed by the risk evaluator and the probability of retinopathy occurrence will be displayed if the risk is deemed to be low (e.g. < 25%). On the other hand, if the risk is high, then the CBR will retrieve similar cases from the existing case database. Once the most similar case is identified from the database, the system will modify (if necessary) the retrieved cases to solve the new case, a process known as adaptation. The outputs of this phase are the probability of the incidence of retinopathy with the proximate time of occurrence (in years).

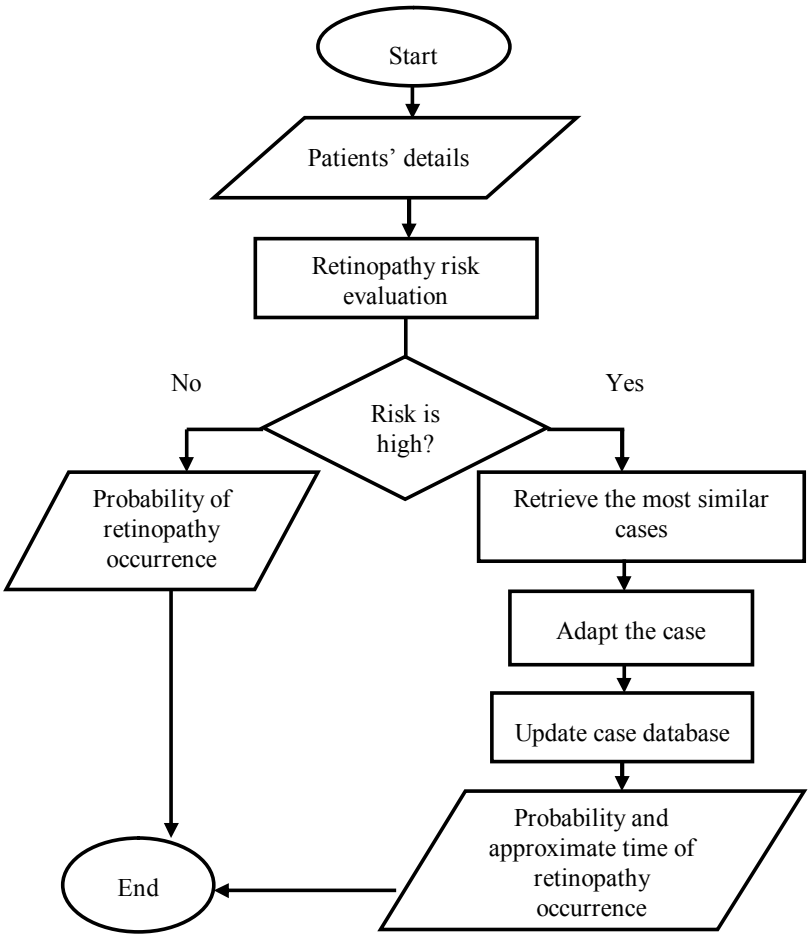


Figure 3. Retinopathy prediction flowchart

CONCLUSIONS AND FUTURE WORK

This paper has proposed to develop a prediction system for retinopathy using two popular approaches: data mining (association rules from Apriori algorithm) and CBR. The data set in the study goes through the data mining process first before being analysed by the CBR technique. The data set and cases can be obtained by interviewing some domain experts. Apriopri algorithms are used to generate the required association rules for the risk factors. These rules cam be used to make the prediction of retinopathy. Similar cases are then retrieved using CBR technique. It is believed that the implementation of the system will enable medical practitioners to predict the chances of retinopathy occurrences among their diabetic patients, and hence to be able to start the treatment or control measures early.

ACKNOWLEDGEMENTS

This study is funded and supported by the University of Malaya Research Grant (Grant no. RG052/11ICT). The authors also wish to extend their gratitude to Professor Umah Kuppusamy,

Associate Professor Chua Kek Heng and Associate Professor Tajunisah Begam from the Faculty of Medicine, University of Malaya for their valuable comments and assistance.

REFERENCES

1. S. H. Wild, G. Roglic, A. Green, R. Sicree and H. King, "Global prevalence of diabetes: Estimates for the year 2000 and projections for 2030", *Diabetes Care*, **2004**, *27*, 1047-1053.
2. M. Nor, N. Safiza, G. K. Lin, S. Suzana, C. C. Kee, H. Jamaiyah, A. Geeta, R. Rahmah, N. F. Wong, A. A. Zainuddin, A. R. Jamaluddin, A. T. Ruzita and A. F. Yusoff, "The third national health and morbidity survey (NHMS III) 2006: Nutritional status of adults aged 18 years and above", *Malaysian J. Nutr.*, **2008**, *14*, 1-87.
3. M. J. Fowler, "Microvascular and macrovascular complications of diabetes", *Clin. Diabetes*, **2008**, *26*, 77-82.
4. P. Zimmet, K. G. M. M. Alberti and J. Shaw, "Global and societal implications of the diabetes epidemic", *Nature*, **2001**, *414*, 782-787.
5. S. Resnikoff, D. Pascolini, D. Etya'ale, I. Kocur, R. Pararajasegaram, G. P. Pokharel and S. P. Mariotti, "Global data on visual impairment in the year 2002", *Bull. World Health Organ.*, **2004**, *82*, 844-851.
6. I. Tajunisah, H. Nabilah and S. C. Reddy, "Prevalence and risk factors for diabetic retinopathy-A study of 217 patients from University of Malaya Medical Centre", *Med. J. Malaysia*, **2006**, *61*,451-456.
7. S. R. Shriwas, A. B. R. Isa, S. C. Reddy, M. Mohammad, W. B. Mohammad and M. Mazlan, "Risk factors for retinopathy in diabetes mellitus in Kelantan, Malaysia", *Med. J. Malaysia*, **1996**, *51*, 447-452.
8. American Academy of Ophthalmology Retina/Vitreous Panel, "Preferred practice pattern: Diabetic retinopathy 2003", Technical Report of American Academy of Ophthalmollogy, San Francisco, **2003**, www.aao.org/ppp (Accessed: June 2011).
9. P. P. Goh, "Status of diabetic retinopathy among diabetics registered to the diabetic eye registry, National Eye Database, 2007", *Med. J. Malaysia*, **2008**, *63*, 24-28.
10. M. Y. Tan, "Self-care practices of adults with poorly controlled diabetes mellitus in Malaysia", *PhD Thesis,* **2009**, University of Adelaide, Australia.
11. Amercian Diabetes Association, "Executive summary: Standards of medical care in diabetes—2008", *Diabetes Care*, **2008**, *31*, S5-S11.
12. C. I. Sanchez, A. Mayo, M. Garcia, M. I. Lopez and R. Hornero, "Automatic image processing algorithm to detect hard exudates based on mixture models", Proceedings of 28th Annual International Conference of the IEEE, **2006**, New York, USA, pp.4453-4456.
13. C. E. Hann, J. A. Revie, D. Hewett, J. G. Chase and G. M. Shaw, "Screening for diabetic retinopathy using computer vision and physiological markers", *J. Diabetes Sci. Technol.*, **2009**, *3,* 819-834.
14. G. G. Gardner, D. Keating, T. H. Williamson and A. T. Elliott, "Automatic detection of diabetic retinopathy using an artificial neural network: A screening tool", *Br. J. Ophthalmol.*, **1996**, *80*, 940-944.

15. A. J. Frame, P. E. Undrill, M. J. Cree, J. A. Olson, K. C. McHardy, P. F. Sharp and J. V. Forrester, "A comparison of computer based classification methods applied to the detection of microaneurysms in ophthalmic fluorescein angiograms", *Comput. Biol. Med.*, **1998**, *28*, 225-238.
16. A. W. Reza and C. Eswaran, "A decision support system for automatic screening of non-proliferative diabetic retinopathy", *J. Med. Syst.*, **2011**, *35*, 17-24.
17. F. Semeraro, G. Parrinello, A. Cancarini, L. Pasquini, E. Zarra, A. Cimino, G. Cancarini, U. Valentini and C. Costagliola, "Predicting the risk of diabetic retinopathy in type 2 diabetic patients", *J. Diabetes Complicat.*, **2011**, *25*, 292-297.
18. H. Y. Cho, D. H. Lee, S. E. Chung and S. W. Kang, "Diabetic retinopathy and peripapillary retinal thickness", *Korean J. Ophthalmol.*, **2010**, *24*, 16-22.
19. C. L. Chan, Y. C. Liu and S. H. Luo, "Investigation of diabetic microvascular complications using data mining techniques", Proceedings of International Joint Conference on Neural Networks, **2008**, Hong Kong, China, pp.830-834.
20. M. Skevofilakas, K. Zarkogianni, B. G. Karamanos and K. S. Nikita, "A hybrid decision support system for the risk assessment of retinopathy development as a long term complication of type 1 diabetes mellitus", Proceedings of 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, **2010**, Buenos Aires, Argentina, pp.6713-6716.
21. C. D. Luu, J. A. Szental, S. Y. Lee, R. Lavanya and T. Y. Wong, "Correlation between retinal oscillatory potentials and retinal vascular caliber in type 2 diabetes", *Invest. Opththalmol. Vis. Sci.*, **2010**, *51*, 482-486.
22. Y. Han, M. E. Schneck, M. A. Jr Bearse, S. Barez, C. H. Jacobsen, N. P. Jewell and A. J. Adams, "Formulation and evaluation of a predictive model to identify the sites of future diabetic retinopathy", *Invest. Opththalmol. Vis. Sci.*, **2004**, *45*, 4106-4112.
23. A. Verrotti, L. Lobefalo, D. Trotta, G. D. Loggia, F. Chiarelli, C. Luigi, G. Morgese and P. E. Gallenga, "Visual evoked potentials in young persons with newly diagnosed diabetes: A long-term follow-up", *Dev. Med. Child Neurol.*, **2000**, *42*, 240-244.
24. V. Balakrishnan, M. R. Shakouri, H. Hoodeh and H. S. Loo, "Predictions using data mining and case-based reasoning: A case study for retinopathy", *World Acad. Sci. Eng. Technol.*, **2012** ,*63,* 573-576.
25. W. H. Mong, W. Hsu, M. L. Lee, B. Liu and L. W. Tok, "Exploration mining in diabetic patients databases: Findings and conclusions", Proceedings of 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, **2000**, Boston, USA, pp. 430-436.
26. H. Müller and J. C. Freytag, "Problems, methods, and challenges in comprehensive data cleansing", Technical Report HUB-IB-164, **2003**, Institut für Informatik, Humboldt University, Berlin, Germany.
27. P. N. Tan, M. Steinbach and V. Kumar, "Association analysis: Basic concepts and algorithms", in "Introduction to Data Mining", Addison Wesley, **2006,** Ch.6, www.users.cs.umn.edu/ ~kumar/dmbook/ch6.pdf (Accessed: August 2012).

28. M. J. Zaki, "Mining non-redundant association rules", *Data Mining Knowl. Disc.*, **2004**, *9*, 223-248
29. J. L. Kolodner, "An introduction to case-based reasoning", *Artif. Intell. Rev.*, **1992**, *6,* 3-34.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 344-355

Maejo International Journal of Science and Technology

ISSN 1905-7873

Available online at www.mijst.mju.ac.th

Full Paper

New exact travelling wave solutions of generalised sinh-Gordon and (2 + 1)-dimensional ZK-BBM equations

Rajesh Kumar Gupta[1], Sachin Kumar [2,*] and Bhajan Lal [1]

[1] School of Mathematics and Computer Applications, Thapar University, Patiala-147004, Punjab, India

[2] Department of Applied Sciences, Bahra Faculty of Engineering, Technical University, Patalia-147001, Punjab, India

* Corresponding authors, e-mail: sachin1jan@yahoo.com (S. Kumar)
rajeshgupta@thapar.edu (R. K. Gupta)

Received: 1 August 2011 / Accepted: 25 September 2012 / Published: 1 October 2012

Abstract: Exact travelling wave solutions have been established for generalised sinh-Gordon and generalised (2+1) dimensional ZK-BBM equations by using $\left(\frac{G'}{G}\right)$–expansion method where $G = G(\xi)$ satisfies a second-order linear ordinary differential equation. The travelling wave solutions are expressed by hyperbolic, trigonometric and rational functions.

Keywords: $\left(\frac{G'}{G}\right)$–expansion method, travelling wave solutions, generalised sinh-Gordon equation, (2+1) dimensional ZK-BBM equation

---

INTRODUCTION

Non-linear partial differential equations (PDEs) are widely used as models to describe complex physical phenomena in various fields of sciences, especially fluid mechanics, solid state physics, plasma physics, plasma wave and chemical physics. Particularly, various methods have been utilised to explore different kinds of solutions of physical models described by non-linear PDEs. One of the basic physical problems for these models is to obtain their exact solutions. In recent years various methods for obtaining exact travelling wave solutions to non-linear equations have been presented, such as the homogeneous balance method [1], the tanh function method [2, 3], the Jacobi elliptic function method [4, 5] and the F-expansion method [6, 7].

In this paper, we use $\left(\frac{G'}{G}\right)$–expansion method [8, 9] to establish exact travelling wave solutions for generalised sinh-Gordon and generalised (2+1) dimensional ZK-BBM equations. The main idea of this method is that the travelling wave solutions of non-linear equations can be expressed by a polynomial in $\left(\frac{G'}{G}\right)$, where $G = G(\xi)$ satisfies the second-order linear ordinary differential equation (LODE): $G''(\xi) + \lambda G'(\xi) + \mu G(\xi) = 0$, where $\xi = x - ct$ and λ, μ and c are arbitrary constants. The degree of this polynomial can be determined by considering the homogeneous balance between higher-order derivatives and the non-linear term appearing in the given non-linear equations. The coefficients of this polynomial can be obtained by solving a set of algebraic equations resulting from the process of using the proposed method.

The sinh-Gordon equation, viz.

$$u_{xt} = \sinh(u), \tag{1}$$

appears in many branches of non-linear science and plays an important role in physics [10], being an important model equation studied by several authors [11–14].

Wazwaz [14] studied the generalised sinh-Gordon equation given by

$$u_{tt} - au_{xx} + b\sinh(nu) = 0, \tag{2}$$

where a and b are arbitrary constants and n is a positive integer. By using the tanh method, Wazwaz derived exact travelling wave solutions of Eq.(2), which provides a more powerful model than the Eq.(1).

The generalised form of the (2 + 1) dimensional ZK-BBM equation is given as:

$$u_t + u_x - a(u^2)_x + (bu_{xt} - ku_{yt})_x = 0, \tag{3}$$

where a, b and k are arbitrary constants. It arises as a description of gravity water waves in the long-wave regime [15, 16]. A variety of exact solutions for the (2 + 1) dimensional ZK-BBM equation [17, 18] are developed by means of the tanh and the sine-cosine methods. In this paper we construct new travelling wave solutions of Eqs.(2) and (3).

(*G'/G*)-EXPANSION METHOD

In this section we describe the $\left(\frac{G'}{G}\right)$–expansion method [8, 9, 19] to find the travelling wave solutions of non-linear PDEs. Suppose that a non-linear equation with two independent variables x and t is given by

$$P(u, u_x, u_t, u_{xx}, u_{xt}, u_{tt}, \ldots\ldots\ldots\ldots) = 0, \tag{4}$$

where $u = u(x,t)$ is an unknown function, P is a polynomial in $u = u(x,t)$ and its various partial derivatives, in which the highest order derivatives and non-linear terms are involved. In the following the main steps of the $\left(\frac{G'}{G}\right)$–expansion method are given.

Step 1. Combining the independent variables x and t into one variable $\xi = x - Vt$,

we suppose that

$$u(x,t) = u(\xi),\ \xi = x - Vt. \tag{5}$$

The travelling wave variables (5) permits us to reduce Eq.(4) to an ordinary differential equation (ODE):

$$P(u, -Vu', u', V^2u'', -Vu'', u'', ...) = 0. \tag{6}$$

Step 2. Suppose that the solution of the ODE (6) can be expressed by a polynomial in $\left(\frac{G'}{G}\right)$ as follows:

$$u(\xi) = \alpha_m (\frac{G'}{G})^m + ..., \qquad \alpha_m \neq 0, \tag{7}$$

where $G = G(\xi)$ satisfies the second-order LODE in the following form:

$$G'' + \lambda G' + \mu G = 0. \tag{8}$$

The constants α_m, $\alpha_{m-1} ... \lambda$ and μ are to be determined later. The unwritten part in Eq.(7) is also a polynomial in $\left(\frac{G'}{G}\right)$, the degree of which is, however, generally equal to or less than *m*-1. The positive integer m can be determined by considering the homogenous balance between the highest order derivatives and the non-linear terms appearing in ODE (6).

Step 3. By substituting (7) into Eq.(6) and using second-order LODE (8), the left-hand side of Eq.(6) is converted into another polynomial in $\left(\frac{G'}{G}\right)$. Equating each coefficient of this polynomial to zero yields a set of algebraic equations for $\alpha_m, ..., V, \lambda, \mu$.

Step 4. The constants $\alpha_m, ..., V, \lambda, \mu$ can be obtained by solving the system of algebraic equations obtained in Step 3. Since the general solutions of the second-order LODE (8) is well known depending on the sign of the discriminant $\Delta = \lambda^2 - 4\mu$, the exact solutions of the given Eq.(4) can be obtained.

GENERALISED SINH-GORDON EQUATION

To find the explicit exact solutions of the sinh-Gordon Eq.(2), we proceed with the methodology explained in the above section. First we make the transformation:

$$u(x,t) = u(\xi) = u(x - ct), \tag{9}$$

where *c* is the wave speed. Substituting the above travelling wave transformation into (2), we get the following ODE:

$$du_{\xi\xi} + b\sinh(nu) = 0, \tag{10}$$

where $d = c^2 - a$ and b is arbitrary constant. Appling the transformation $v = e^{nu}$ to Eq.(10) and using relations

$$\sinh(nu)=\frac{v-v^{-1}}{2},\ \cosh(nu)=\frac{v+v^{-1}}{2},\ u=\frac{1}{n}\arccos h\frac{v+v^{-1}}{2},\ u^{"}=\frac{v^{"}v-v^{'2}}{nv^{2}}$$

in Eq.(10), we have

$$2d(v^{"}v-v^{'2})+bnv^{3}-bnv=0. \tag{11}$$

Suppose that the solution of ODE (11) can be expressed by a polynomial in $\left(\frac{G'}{G}\right)$ as follows:

$$v(\xi)=\alpha_m(\frac{G^{'}}{G})^{m}+..., \tag{12}$$

where $G(\xi)$ satisfies the second-order LODE in the form:

$$G^{"}+\lambda G^{'}+\mu G=0\,. \tag{13}$$

From Eqs.(12) and (13), we have

$$\begin{aligned} v^{3}&=\alpha_m{}^{3}(\frac{G^{'}}{G})^{3m}+... \\ v^{'}&=-m\alpha_m(\frac{G^{'}}{G})^{m+1}+... \\ v^{"}&=m(m+1)\alpha_m(\frac{G^{'}}{G})^{m+2}+.... \end{aligned} \tag{14}$$

Considering the homogeneous balance between $v"v$ and v^3 in Eq.(11), it is required that $m=2$. So we can write (12) as:

$$v(\xi)=\alpha_2(\frac{G^{'}}{G})^{2}+\alpha_1(\frac{G^{'}}{G})\ +\alpha_0;\ \alpha_2\neq 0\,. \tag{15}$$

By using Eq.(15) and (13) in Eq.(11) and collecting all terms with the same power of $\left(\frac{G'}{G}\right)$ together, the left-hand side of Eq.(11) is converted into another polynomial in $\left(\frac{G'}{G}\right)$. Equating each coefficient of this polynomial to zero yields a set of simultaneous algebraic equations for $\alpha_{1,}\alpha_2,\alpha_0$ and d as follows:

$$4d\alpha_2^{2}+bn\alpha_2^{3}=0$$
$$8d(-2\alpha_2\lambda-\alpha_1)\alpha_2+3bn\alpha_1\alpha_2^{2}+2d(2\alpha_1+10\alpha_2\lambda)\alpha_2+12d\alpha_2\alpha_1=0$$
$$bn\left(\alpha_0\alpha_2^{2}+2\alpha_1^{2}\alpha_2+\alpha_2+\alpha_2(2\alpha_0\alpha_2+\alpha_1^{2})\right)+2d(8\alpha_2\mu+3\alpha_1\lambda+4\alpha_2\lambda^{2})+2d(2\alpha_1+10\alpha_2\lambda)\alpha_1+12d\alpha_2\alpha_0$$
$$-2d(-4(-2\alpha_2\mu-\alpha_1\lambda)\alpha_2+(-2\alpha_2\lambda-\alpha_1)^{2})=0$$
$$bn(4\alpha_1\alpha_2\alpha_0+\alpha_1(2\alpha_0\alpha_2+\alpha_1^{2}))+2d(6\alpha_2\mu\lambda+2\alpha_1\mu+\alpha_1\lambda^{2})\alpha_2+2d(8\alpha_2\mu+3\alpha_1\lambda+4\alpha_2\lambda^{2})\alpha_1$$
$$+2d(2\alpha_1+10\alpha_2\lambda)\alpha_0-2d(4\alpha_1\mu\alpha_2+2(-2\alpha_2\mu-\alpha_1\lambda)(-2\alpha_2\lambda-\alpha_1))=0$$

$$bn(\alpha_0(2\alpha_0\alpha_1+\alpha_1^{2})+2\alpha_1^{2}\alpha_0+\alpha_2\alpha_0^{2})-bn\alpha_2+2d(2\alpha_2\mu^{2}+\alpha_1\lambda\mu)\alpha_2+2d(6\alpha_2\lambda\mu+2\alpha_1\mu+\alpha_1\lambda^{2})\alpha_1$$
$$-2d(8\alpha_2\mu+3\alpha_1\lambda+4\alpha_2\lambda^{2})\alpha_0-2d(-2\alpha_1\mu(-2\alpha_2\lambda-\alpha_1)+(-2\alpha_2\mu-\alpha_1\lambda)^{2})=0$$

$$3bn\alpha_0^2\alpha_1 - bn\alpha_1 + 2d(2\alpha_2\mu^2 + \alpha_1\lambda\mu)\alpha_1 + 2d(6\alpha_2\lambda\mu + 2\alpha_1\mu + \alpha_1\lambda^2)\alpha_0 + 4d\alpha_1\mu(-2\alpha_2\mu - \alpha_1\lambda) = 0$$
$$2d(2\alpha_2\mu^2 + \alpha_1\lambda\mu)\alpha_0 + bn\alpha_0^3 - 2d\alpha_1^2\mu^2 - bn\alpha_0 = 0 . \tag{16}$$

Solving the algebraic equations above yields

$$\alpha_0 = \pm\frac{\lambda^2}{4\mu - \lambda^2},\ \alpha_1 = \pm\frac{4\lambda}{4\mu - \lambda^2},\ \alpha_2 = \pm\frac{4}{4\mu - \lambda^2},\ d = \mp\frac{6n}{4\mu - \lambda^2}, \tag{17}$$

where λ, μ and n are arbitrary constants.

By using (17), expression (15) can be written as:

$$v(\xi) = \pm\frac{4}{4\mu - \lambda^2}\left(\frac{G'}{G}\right)^2 \pm \frac{4\lambda}{4\mu - \lambda^2}\left(\frac{G'}{G}\right) \pm \frac{\lambda^2}{4\mu - \lambda^2}, \tag{18}$$

where $\xi = x - t\sqrt{\mp\frac{6n}{4\mu - \lambda^2} + a}$.

Substituting the general solution of Eq.(13) into Eq.(18), we have two types of travelling wave solution of Eq.(2) as follows:

Case 1: When $\lambda^2 - 4\mu < 0$,

$$u(x,t) = \frac{1}{n}\log\left(\pm\frac{4}{4\mu - \lambda^2}\left(\frac{G'}{G}\right)^2 \pm \frac{4\lambda}{4\mu - \lambda^2}\left(\frac{G'}{G}\right) \pm \frac{\lambda^2}{4\mu - \lambda^2}\right), \tag{19}$$

where $\left(\frac{G'}{G}\right) = -\frac{\lambda}{2} + \frac{1}{2}\frac{\left(C_1\sinh\left(\frac{\sqrt{\lambda^2 - 4\mu}\xi}{2}\right) + C_2\cosh\left(\frac{\sqrt{\lambda^2 - 4\mu}\xi}{2}\right)\right)\sqrt{\lambda^2 - 4\mu}\xi}{C_1\cosh\left(\frac{\sqrt{\lambda^2 - 4\mu}\xi}{2}\right) + C_2\sinh\left(\frac{\sqrt{\lambda^2 - 4\mu}\xi}{2}\right)}$ and

$\xi = x - t\sqrt{\mp\frac{6n}{4\mu - \lambda^2} + a}$. C_1 and C_2 are arbitrary constants. Profile of solution (19) for $n = a = C_2 = 1, C_1 = \mu = 0$ and $\lambda = \sqrt{2}$ is shown in Figure 1.

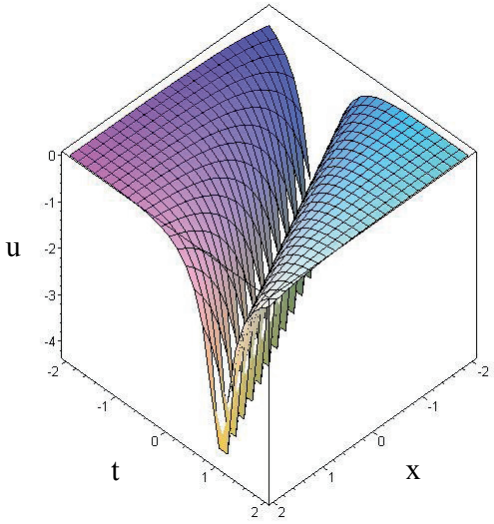


Figure 1. Profile of solution (19) when $n = a = C_2 = 1, C_1 = \mu = 0$ and $\lambda = \sqrt{2}$

Case 2: When $\lambda^2 - 4\mu > 0$,

$$u(x,t) = \frac{1}{n}\log\left(\pm\frac{4}{4\mu-\lambda^2}\left(\frac{G'}{G}\right)^2 \pm\frac{4\lambda}{4\mu-\lambda^2}\left(\frac{G'}{G}\right)\pm\frac{\lambda^2}{4\mu-\lambda^2}\right), \tag{20}$$

where $\left(\frac{G'}{G}\right) = -\frac{\lambda}{2} + \frac{1}{2}\frac{\left(-C_1\sin\left(\frac{\sqrt{\lambda^2-4\mu}\xi}{2}\right)+C_2\cos\left(\frac{\sqrt{\lambda^2-4\mu}\xi}{2}\right)\right)\sqrt{\lambda^2-4\mu}\xi}{C_1\cos\left(\frac{\sqrt{\lambda^2-4\mu}\xi}{2}\right)+C_2\sin\left(\frac{\sqrt{\lambda^2-4\mu}\xi}{2}\right)}$ and

$\xi = x - t\sqrt{\mp\frac{6n}{4\mu-\lambda^2}+a}$. C_1 and C_2 are arbitrary constants. In Figure 2 we have periodic solution (20) when $n = a = \lambda = C_1 = C_2 = 1$ and $\mu = 2\sqrt{2}$.

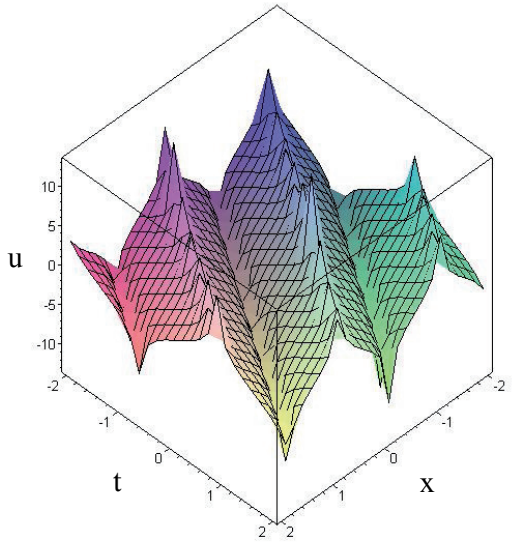


Figure 2. 3D plot of periodic solution (20) when $n = a = \lambda = C_1 = C_2 = 1$ and $\mu = 2\sqrt{2}$

(2+1) DIMENSIONAL ZK-BBM EQUATION

In finding exact solutions of ZK-BBM Eq.(3), first we make the transformation:

$$u(\xi) = u(\xi),\ \xi = x + y - ct, \tag{21}$$

which permits us to convert Eq.(3) into an ODE for $u = u(\xi)$ as follows:

$$u'(1-c) - 2auu' + cu'''(b-k) = 0, \tag{22}$$

where a, b and k are arbitrary parameters and c is the wave speed.

Integrating it with respect to ξ once yields:

$$V + u(1-c) - au^2 + cu''(b-k) = 0, \tag{23}$$

where V is an integration constant that is to be determined later.

Proceeding in a similar manner as in the above section and considering the homogeneous balance between u'' and u^2 in Eq.(23), we have $m = 2$. So we can assume the solution of Eq.(23) as follows:

$$u(\xi) = \alpha_2\left(\frac{G'}{G}\right)^2 + \alpha_1\left(\frac{G'}{G}\right) + \alpha_0;\quad \alpha_2 \neq 0, \tag{24}$$

where α_0, α_1 and α_2 are to be determined later.

By using (24) and (13) in Eq.(23) and collecting all terms with the same power of $\left(\frac{G'}{G}\right)$ together, the left-hand side of Eq.(23) is converted into another polynomial in $\left(\frac{G'}{G}\right)$. Equating each

coefficient of this polynomial to zero yields a set of simultaneous algebraic equations for c and V as follows:

$$\begin{aligned}
&-a\alpha_2^2 + bc\alpha_2(b-k) = 0\\
&c(2\alpha_1 + 10\alpha_2\lambda)(b-k) - 2a\alpha_1\alpha_2 = 0\\
&\alpha_2(1-c) + c(8\alpha_2\mu + 3\alpha_1\lambda + 4\alpha_2\lambda^2)(b-k) - a(2\alpha_0\alpha_2 + \alpha_1^2) = 0\\
&\alpha_1(1-c) + c(b\alpha_2\lambda\mu + 2\alpha_1\lambda + 4\alpha_2\lambda^2)(b-k) - 2a\alpha_0\alpha_1 = 0\\
&V + \alpha_0(1-c) - a\alpha_0^2 + c(2\alpha_2\mu^2\alpha_1\lambda\mu)(b-k) = 0.
\end{aligned} \tag{25}$$

Solving the algebraic equations above yields

$$\begin{aligned}
&\alpha_2 = \frac{bc(b-k)}{a}, \alpha_1 = \frac{b\lambda c(b-k)}{a}, \alpha_0 = \frac{b\lambda^2 c + 8bc\mu - k\lambda^2 c + 1 - c - 8kc\mu}{2a},\\
&V = \frac{\begin{array}{l}(2c - 8b^2\lambda^2c^2\mu - 8k^2\lambda^2c^2\mu + 16b\lambda^2c^2\mu k + 16c^2k^2\mu^2 + 16c^2k^2\mu^2 - 32bc^2k\mu^2\\ -2bkc^2\lambda^4 + b^2\lambda^4c^2 + k^2\lambda^4c^2 - 1 - c^2)\end{array}}{4a},
\end{aligned} \tag{26}$$

where λ, μ, a, b and k are arbitrary constants.

By using (26), expression (24) can be written as:

$$u(\xi) = \frac{bc(b-k)}{a}\left(\frac{G'}{G}\right)^2 + \frac{b\lambda c(b-k)}{a}\left(\frac{G'}{G}\right) + \frac{b\lambda^2 c + 8bc\mu - k\lambda^2 c + 1 - c - 8kc\mu}{2a} \tag{27}$$

where λ, μ, b, c and k are arbitrary constants.

Substituting the general solution of Eq.(13) into Eq.(27), we have the following travelling wave solutions of Eq. (3):

Case 1. When $\lambda^2 - 4\mu > 0$,

$$u(\xi) = \frac{bc(b-k)}{a}\left(\frac{G'}{G}\right)^2 + \frac{b\lambda c(b-k)}{a}\left(\frac{G'}{G}\right) + \frac{b\lambda^2 c + 8bc\mu - k\lambda^2 c + 1 - c - 8kc\mu}{2a}, \tag{28}$$

where $\left(\frac{G'}{G}\right) = -\frac{\lambda}{2} + \frac{1}{2}\frac{\left(C_1 \sinh\left(\frac{\sqrt{\lambda^2 - 4\mu}\xi}{2}\right) + C_2 \cosh\left(\frac{\sqrt{\lambda^2 - 4\mu}\xi}{2}\right)\right)\sqrt{\lambda^2 - 4\mu}\xi}{C_1 \cosh\left(\frac{\sqrt{\lambda^2 - 4\mu}\xi}{2}\right) + C_2 \sinh\left(\frac{\sqrt{\lambda^2 - 4\mu}\xi}{2}\right)}$

and $\xi = x + y - ct$. C_1 and C_2 are arbitrary constants. The solution (28) furnishes a bell-shaped soliton solution (28) when $a = k = c = C_1 = 1, b = 2, t = 0.2, C_2 = \mu = 0$ and $\lambda = \sqrt{2}$, as shown in Figure 3.

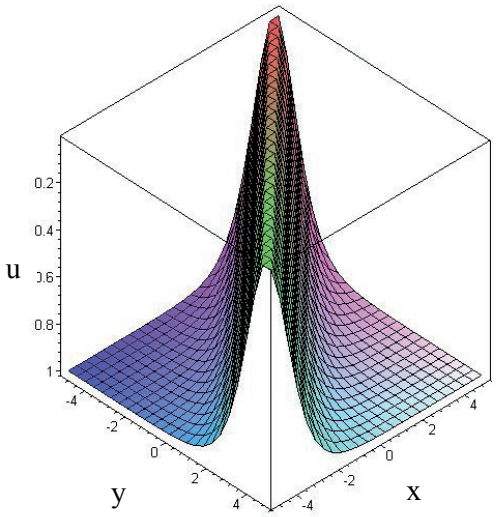


Figure 3. Bell-shaped soliton solution (28) when $a = k = c = C_1 = 1, b = 2, t = 0.2, C_2 = \mu = 0$ and $\lambda = \sqrt{2}$

Case 2. When $\lambda^2 - 4\mu < 0$,

$$u(\xi) = \frac{bc(b-k)}{a}\left(\frac{G'}{G}\right)^2 + \frac{b\lambda c(b-k)}{a}\left(\frac{G'}{G}\right) + \frac{b\lambda^2 c + 8bc\mu - k\lambda^2 c + 1 - c - 8kc\mu}{2a}, \tag{29}$$

where $\dfrac{\mathrm{G}'(\xi)}{\mathrm{G}(\xi)} = -\dfrac{1}{2}\lambda + \dfrac{1}{2}\dfrac{\left(-C_1 \sin\left(\frac{1}{2}\sqrt{-\lambda^2+4\mu}\xi\right) + C_2 \cos\left(\frac{1}{2}\sqrt{-\lambda^2+4\mu}\xi\right)\right)\sqrt{-\lambda^2+4\mu}}{C_1 \cos\left(\frac{1}{2}\sqrt{-\lambda^2+4\mu}\xi\right)_2 \sin\left(\frac{1}{2}\sqrt{-\lambda^2+4\mu}\xi\right)}$

and $\xi = x + y - ct$. C_1 and C_2 are arbitrary constants. For $a = 1.5$, $k = c = C_1 = 1$, $b = 2$, $t = 0.2$, $\lambda = C_2 = 0$ and $\mu = 1$, solution (29) gives the periodic solution as shown in Figure 4.

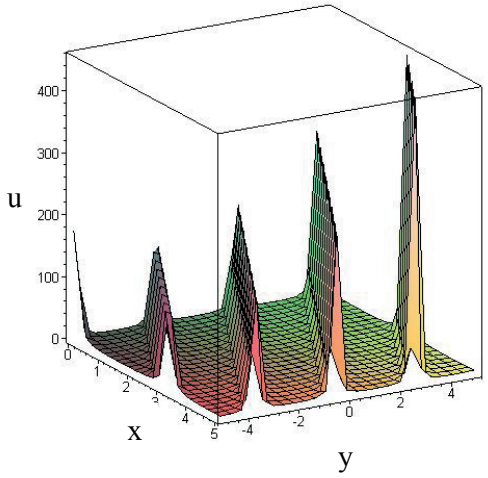


Figure 4. Periodic solution (29) when $a=1.5, k=c=C_1=1, b=2, t=0.2, \lambda=C_2=0$ and $\mu=1$

Case 3. When $\lambda^2-4\mu=0$,

$$u(\xi)=\frac{bc(b-k)}{a}\left(\frac{2C_2-C_1\lambda-C_2\lambda\xi}{2(C_1+C_2\xi)}\right)^2+\frac{b\lambda c(b-k)}{a}\left(\frac{2C_2-C_1\lambda-C_2\lambda\xi}{2(C_1+C_2\xi)}\right)+\frac{b\lambda^2c+8bc\mu-k\lambda^2c+1-c-8kc\mu}{2a}, \tag{30}$$

where $\xi=x+y-ct$. C_1 and C_2 are arbitrary constants. In Figure 5, we have a soliton solution (30) for $a=1.5, k=c=C_1=C_2=1, b=2, t=0.2, \lambda=1$ and $\mu=\frac{1}{4}$.

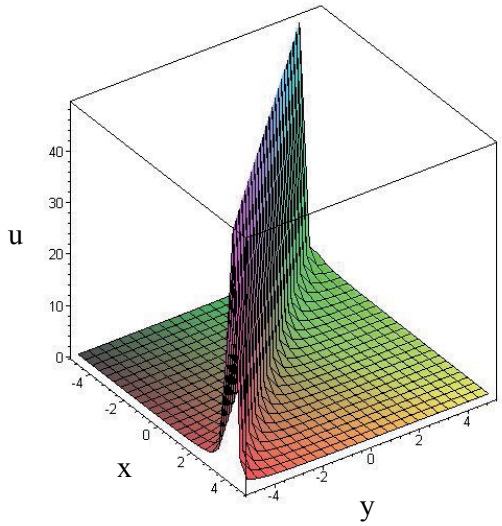


Figure 5. Soliton solution (30) for $a=1.5, k=c=C_1=C_2=1, b=2, t=0.2, \lambda=1$ and $\mu=\frac{1}{4}$

CONCLUSIONS

In this paper, the travelling wave solutions of the generalised sinh-Gordon and (2 + 1) dimensional ZK-BBM equations are found successfully through the use of $\left(\frac{G'}{G}\right)$-expansion method, which includes hyperbolic function solutions and trigonometric function solutions. One can see that this method is direct, concise and effective.

ACKNOWLEDGEMENTS

The authors thank the unanimous referees for several valuable suggestions which have considerably improved the presentation of the paper.

REFERENCES

1. M. Wang, Y. Zhou and Z. Li, "Application of a homogeneous balance method to exact solutions of non-linear equations in mathematical physics", *Phys. Lett. A*, **1996**, *216,* 67-75.
2. W. Malfliet and W. Hereman, "The tanh method: I. Exact solutions of non-linear evolution and wave equations", *Phys. Scripta*, **1996**, *54,* 563-568.
3. M. A. Abdou, "The extended tanh method and its applications for solving non-linear physical models", *Appl. Math. Comput*, **2007**, *190*, 988-996.
4. D. Lu, "Jacobi elliptic function solutions for two variant Boussinesq equations", *Chaos Solitons Fract.*, **2005**, *24*, 1373-1385.
5. Z. Yan, "Abundant families of Jacobi elliptic function of the (2+1)-dimensional integrable Davey- Stawartson-type equation via a new method", *Chaos Solitons Fract.*, **2003**, *18*, 299-309.
6. M. Wang and X. Li, "Extended F-expansion and periodic wave solutions for the generalized Zakharov equations", *Phys. Lett. A*, **2005**, *343*, 48-54.
7. M. Wang and X. Li, "Applications of F-expansion to periodic wave solutions for a new Hamiltonian amplitude equation", *Chaos Solitons Fract.*, **2005**, *24*, 1257-1268.
8. M. Wang, X. Li and J. Zhang, "The $\left(\frac{G'}{G}\right)$−expansion method and travelling wave solutions of non-linear evolution equations in mathematical physics", *Phys. Lett. A*, **2008**, *372*, 417-423.
9. E. M. E. Zayed and K. A. Gepreel, "The $\left(\frac{G'}{G}\right)$−expansion method for finding travelling wave solutions of non-linear PDEs in mathematical physics", *J. Math. Phys.*, **2009**, *50*, 013502-013513.
10. A. Grauel, "Sinh-Gordon equation, Painleve property and Backlund transformation", *Physica A*, **1985**, *132*, 557-568.
11. Sirendaoreji and S. Jiong, "A direct method for solving sine-Gordon type equations", *Phys. Lett. A*, **2002**, *298*, 133-139.
12. Z. Yan, "A sinh-Gordon equation expansion method to construct doubly periodic solutions for non-linear differential equations", *Chaos Solitons Fract.*, **2003**, *16*, 291-297.
13. H. Zhang, "New exact solutions for the sinh-Gordon equation", *Chaos Solitons Fract.*, **2006**, *28*, 489-496.
14. A. M. Wazwaz**,** "Exact solutions for the generalized sine-Gordon and the generalized sinh-Gordon equations", *Chaos Solitons Fract.*, **2006**, *28*, 127-135.
15. D. H. Peregrine, "Long waves on a beach", *J. Fluid Mech.*, **1967**, *27*, 815-827.
16. T. B. Benjamin, J. L. Bona and J. J. Mahony, "Model equations for long waves in nonlinear dispersive systems", *Philos. Trans. Roy. Soc. London Ser. A*, **1972**, *272*, 47-78.

17. A. M. Wazwaz**, "**The extended tanh method for new compact and noncompact solutions for the KP-BBM and the ZK-BBM equations", *Chaos Solitons Fract.*, **2008**, *38,* 1505-1516.
18. A. M. Wazwaz, "Compact and noncompact physical structures for the ZK-BBM equation", *Appl. Math. Comput*., **2005,** *169*, 713-725.
19. S. Kumar, K. Singh and R. K. Gupta, "Coupled Higgs field equation and Hamiltonian amplitude equation: Lie classical approach and $\left(\frac{G'}{G}\right)$-expansion method", *Parmana J. phys.,* **2012**, *79*, 41-60.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 356-371

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Non-differentiable second-order mixed symmetric duality with cone constraints

Navdeep Kailey [1] and Shiv Kumar Gupta[2,*]

[1] Department of Applied Sciences, Gulzar Institute of Engineering and Technology, Khanna-141 401, India
[2] Department of Mathematics, Indian Institute of Technology, Patna-800 013, India

* Corresponding author; tel: +91-612-2552025, fax: +91-612-2277384, e-mail: skgiitr@gmail.com

Received: 22 July 2011 / Accepted: 6 October 2012 / Published: 11 October 2012

Abstract: A pair of mixed non-differentiable second-order symmetric dual programmes over cones is considered. Weak, strong and converse duality theorems are established under second-order (F,ρ) convexity/pseudo-convexity assumptions. Special cases are also discussed to show that this paper extends some known results in the literature.

Keywords: symmetric duality, mixed second-order duality, second-order (F,ρ) convexity/pseudo-convexity, non-differentiable programming, cone constraints

INTRODUCTION

The study of symmetric duality for non-linear problems is well developed by many researchers, notably Dantzig et al. [1], Mond [2], Bazaraa and Goode [3], and Mond and Weir [4]. Mangasarian [5] was the first to identify second-order dual formulations for the non-linear programmes. Wolfe type second-order symmetric duality has been discussed by Ahmad and Husain [6] and Yang et al. [7] for single-objective non-differentiable functions, and by Yang et al. [8] and Ahmad and Husain [9] for multi-objective programming problems. The duality results for a pair of Mond-Weir type second-order multi-objective symmetric dual programmes have been considered by Suneja et al. [10] under η-bonvexity/η-pseudo-bonvexity assumptions. Ahmad and Husain [9, 11], Khurana [12], Mishra and Lai [13], and Padhan and Nahak [14] studied symmetric duality over arbitrary cones.

Bector et al. [15] and Yang et al. [16] formulated mixed symmetric dual models for multi-objective differentiable and single-objective non-differentiable programming problems respectively. Ahmad [17] introduced a mixed-type symmetric duality in multi-objective programming problems,

ignoring non-negativity restrictions of Bector et al. [15]. Recently, Ahmad and Husain [11] studied duality results for a pair of multi-objective mixed symmetric dual programmes over arbitrary cones under *K*-preinvexity/*K*-pseudo-invexity assumptions. Li and Gao [18] obtained duality relations for a mixed symmetric dual model in non-differentiable multi-objective non-linear programming problems involving generalised convex functions. Recently, Gupta and Kailey [19] formulated a second-order mixed symmetric dual programme for a class of non-differentiable multi-objective programming problems and proved usual duality results under second-order *F*-convexity/pseudo-convexity assumptions.

In this paper, we formulate a pair of mixed non-differentiable second-order symmetric dual programmes over arbitrary cones. Weak, strong and converse duality results are proved under second-order (F,ρ) convexity/pseudo-convexity assumptions. Several known results are obtained as special cases.

NOTATIONS AND PRELIMINARIES

Definition 1 [11, 12, 20]**.** Let C be a closed convex cone in R^n with non-empty interior. The positive polar cone C^* of C is defined as

$$C^* = \{z : x^T z \geq 0, \text{ for all } x \in C\}.$$

Now we consider a sub-linear functional $F : X \times X \times R^n \to R$ (where $X \subseteq R^n$).

Definition 2. A twicely differentiable function $\psi : X \to R$ is said to be second-order (F,ρ) convex at $u \in X$, if there exists a real-value function $d : X \times X \to R$ and a real number ρ, such that for all $q \in R^n$ and $x \in X$,

$$\psi(x) - \psi(u) + \frac{1}{2} q^T \nabla_{xx}\psi(u)q \geq F\left(x,u;\nabla_x \psi(u) + \nabla_{xx}\psi(u)q\right) + \rho d^2(x,u).$$

Definition 3. A twicely differentiable function $\psi : X \to R$ is said to be second-order (F,ρ) pseudo-convex at $u \in X$, if there exists a real-value function $d : X \times X \to R$ and a real number ρ, such that for all $q \in R^n$ and $x \in X$,

$$F\left(x,u;\nabla_x \psi(u) + \nabla_{xx}\psi(u)q\right) \geq 0 \Rightarrow \psi(x) \geq \psi(u) - \frac{1}{2} q^T \nabla_{xx}\psi(u)q + \rho d^2(x,u).$$

Generalised Schwartz Inequality

The following generalised schwartz inequality shall be made use of:

$$l^T Am \leq (l^T Al)^{\frac{1}{2}} (m^T Am)^{\frac{1}{2}},$$

where $l, m \in R^n$, and $A \in R^n \times R^n$ is a positive semi-definite matrix. Equality holds if for some $\lambda \geq 0$, $Al = \lambda Am$.

PROBLEM FORMULATION

For $N = \{1,2,\ldots,n\}$ and $M = \{1,2,\ldots,m\}$, let $J_1 \subseteq N, K_1 \subseteq M$, $J_2 = N \setminus J_1$ and $K_2 = M \setminus K_1$. Let $|J_1|$ denote the number of elements in J_1. The other symbols, $|J_2|, |K_1|$ and $|K_2|$, are defined similarly. Let $x^1 \in R^{|J_1|}$ and $x^2 \in R^{|J_2|}$. Then any $x \in R^n$ can be written as (x^1, x^2).

Similarly for $y^1 \in R^{|K_1|}$ and $y^2 \in R^{|K_2|}, y \in R^m$ can be written as (y^1, y^2). It may be noted here that if $J_1 = \varnothing$, then $|J_1| = 0, J_2 = N$ and therefore $|J_2| = n$. In this case, $R^{|J_1|}, R^{|J_2|}$ and $R^{|J_1|} \times R^{|K_1|}$ will be zero-dimensional, *n*-dimensional and $|K_1|$-dimensional Euclidean spaces respectively. Similarly we can describe the cases $J_2 = \varnothing, K_1 = \varnothing$ or $K_2 = \varnothing$. Let C_1, C_2, C_3 and C_4 be closed convex cones with non-empty interiors in $R^{|J_1|}$, $R^{|J_2|}$, $R^{|K_1|}$ and $R^{|K_2|}$ respectively.

Now, consider the following pair of mixed second-order symmetric dual programs:

Primal Problem (SMP)

Minimise $L(x^1, y^1, x^2, y^2, z^2, p, r) = f(x^1, y^1) + ((x^1)^T D_1 x^1)^{\frac{1}{2}} + g(x^2, y^2) + ((x^2)^T D_2 x^2)^{\frac{1}{2}}$

$$-(y^2)^T E_2 z^2 - (y^1)^T [\nabla_{y^1} f(x^1, y^1) + \nabla_{y^1 y^1} f(x^1, y^1) p] - \frac{1}{2} p^T \nabla_{y^1 y^1} f(x^1, y^1) p$$

$$-\frac{1}{2} r^T \nabla_{y^2 y^2} g(x^2, y^2) r,$$

subject to

$$-[\nabla_{y^1} f(x^1, y^1) - E_1 z^1 + \nabla_{y^1 y^1} f(x^1, y^1) p] \in C_3^*, \tag{1}$$

$$-[\nabla_{y^2} g(x^2, y^2) - E_2 z^2 + \nabla_{y^2 y^2} g(x^2, y^2) r] \in C_4^*, \tag{2}$$

$$(y^2)^T [\nabla_{y^2} g(x^2, y^2) - E_2 z^2 + \nabla_{y^2 y^2} g(x^2, y^2) r] \geq 0, \tag{3}$$

$$(z^1)^T E_1 z^1 \leq 1, \tag{4}$$

$$(z^2)^T E_2 z^2 \leq 1, \tag{5}$$

$$x^1 \in C_1, x^2 \in C_2. \tag{6}$$

Dual Problem (SMD)

Maximise $M(u^1, v^1, u^2, v^2, w^2, q, s) = f(u^1, v^1) - ((v^1)^T E_1 v^1)^{\frac{1}{2}} + g(u^2, v^2) - ((v^2)^T E_2 v^2)^{\frac{1}{2}}$

$$+(u^2)^T D_2 w^2 - (u^1)^T [\nabla_{x^1} f(u^1, v^1) + \nabla_{x^1 x^1} f(u^1, v^1) q] - \frac{1}{2} q^T \nabla_{x^1 x^1} f(u^1, v^1) q$$

$$-\frac{1}{2} s^T \nabla_{x^2 x^2} g(u^2, v^2) s,$$

subject to

$$\nabla_{x^1} f(u^1, v^1) + D_1 w^1 + \nabla_{x^1 x^1} f(u^1, v^1) q \in C_1^*, \tag{7}$$

$$\nabla_{x^2} g(u^2, v^2) + D_2 w^2 + \nabla_{x^2 x^2} g(u^2, v^2) s \in C_2^*, \tag{8}$$

$$(u^2)^T [\nabla_{x^2} g(u^2, v^2) + D_2 w^2 + \nabla_{x^2 x^2} g(u^2, v^2) s] \leq 0, \tag{9}$$

$$(w^1)^T D_1 w^1 \leq 1, \tag{10}$$

$$(w^2)^T D_2 w^2 \le 1, \tag{11}$$

$$v^1 \in C_3, v^2 \in C_4. \tag{12}$$

where

1. $f : R^{|J_1|} \times R^{|K_1|} \to R$ and $g : R^{|J_2|} \times R^{|K_2|} \to R$ are differentiable functions,
2. D_1, D_2, E_1 and E_2 are positive semi-definite matrices in $R^{|J_1|} \times R^{|J_1|}$, $R^{|J_2|} \times R^{|J_2|}$, $R^{|K_1|} \times R^{|K_1|}$ and $R^{|K_2|} \times R^{|K_2|}$ respectively, and
3. $p, z^1 \in R^{|K_1|}, r, z^2 \in R^{|K_2|}, q, w^1 \in R^{|J_1|}$ and $s, w^2 \in R^{|J_2|}$.

RESULTS AND DISCUSSION

In this section, we prove weak, strong and converse duality results for the dual pair, SMP and SMD, formulated above.

Theorem 1 (Weak duality)

Let $(x^1, y^1, x^2, y^2, z^1, z^2, p, r)$ be feasible for (SMP) and $(u^1, v^1, u^2, v^2, w^1, w^2, q, s)$ be feasible for (SMD). Let the sublinear functionals $F_1 : R^{|J_1|} \times R^{|J_1|} \times R^{|J_1|} \mapsto R$, $F_2 : R^{|K_1|} \times R^{|K_1|} \times R^{|K_1|} \mapsto R$, $G_1 : R^{|J_2|} \times R^{|J_2|} \times R^{|J_2|} \mapsto R$ and $G_2 : R^{|K_2|} \times R^{|K_2|} \times R^{|K_2|} \mapsto R$ satisfy the following conditions:

$$F_1(x^1, u^1; a^1) + (a^1)^T u^1 \ge 0, \text{ for all } a^1 \in C_1^*, \tag{A}$$

$$F_2(v^1, y^1; a^2) + (a^2)^T y^1 \ge 0, \text{ for all } a^2 \in C_3^*, \tag{B}$$

$$G_1(x^2, u^2; b^1) + (b^1)^T u^2 \ge 0, \text{ for all } b^1 \in C_2^*, \tag{C}$$

$$G_2(v^2, y^2; b^2) + (b^2)^T y^2 \ge 0, \text{ for all } b^2 \in C_4^*. \tag{D}$$

Further, let

(i) $f(.,v^1) + (.)^T D_1 w^1$ be second-order (F_1, ρ_1) convex at u^1, and $-(f(x^1,.) - (.)^T E_1 z^1)$ be second-order (F_2, ρ_2) convex at y^1,

(ii) $g(.,v^2) + (.)^T D_2 w^2$ be second-order (G_1, σ_1) pseudo-convex at u^2, and $-(g(x^2,.) - (.)^T E_2 z^2)$ be second-order (G_2, σ_2) pseudo-convex at y^2,

(iii) either $\rho_1 d_1^2(x^1, u^1) + \rho_2 d_2^2(v^1, y^1) \ge 0$ or $\rho_1, \rho_2 \ge 0$, and

(iv) either $\sigma_1 d_3^2(x^2, u^2) + \sigma_2 d_4^2(v^2, y^2) \ge 0$ or $\sigma_1, \sigma_2 \ge 0$.

Then

$$L(x^1, y^1, x^2, y^2, z^2, p, r) \ \ge \ M(u^1, v^1, u^2, v^2, w^2, q, s).$$

Proof. By the second-order (F_1, ρ_1) convexity of $f(.,v^1) + (.)^T D_1 w^1$ at u^1 and the second-order (F_2, ρ_2) convexity of $-(f(x^1,.) - (.)^T E_1 z^1)$ at y^1, we have

$$f(x^1, v^1) + (x_1)^T D_1 w^1 - f(u^1, v^1) - (u_1)^T D_1 w^1 + \frac{1}{2} q^T \nabla_{x^1 x^1} f(u^1, v^1) q$$
$$\ge \ F_1(x^1, u^1; \nabla_{x^1} f(u^1, v^1) + D_1 w^1 + \nabla_{x^1 x^1} f(u^1, v^1) q) + \rho_1 d_1^2(x^1, u^1) \tag{13}$$

and

$$f(x^1,y^1)-(y^1)^T E_1 z^1 - f(x^1,v^1)+(v^1)^T E_1 z^1 - \frac{1}{2}p^T \nabla_{y^1y^1} f(x^1,y^1)p$$
$$\geq F_2(v^1,y^1;-(\nabla_{y^1} f(x^1,y^1) - E_1 z^1 + \nabla_{y^1y^1} f(x^1,y^1)p)) + \rho_2 d_2^2(v^1,y^1). \quad (14)$$

Adding the inequalities (13) and (14), we obtain

$$f(x^1,y^1)-f(u^1,v^1)+(x_1)^T D_1 w^1 - (u_1)^T D_1 w^1 - (y^1)^T E_1 z^1 + (v^1)^T E_1 z^1$$
$$+\frac{1}{2}q^T \nabla_{x^1x^1} f(u^1,v^1)q - \frac{1}{2}p^T \nabla_{y^1y^1} f(x^1,y^1)p \geq F_1(x^1,u^1;\nabla_{x^1} f(u^1,v^1) + D_1 w^1 + \nabla_{x^1x^1} f(u^1,v^1)q)$$
$$+ F_2(v^1,y^1;-(\nabla_{y^1} f(x^1,y^1) - E_1 z^1 + \nabla_{y^1y^1} f(x^1,y^1)p)) + \rho_1 d_1^2(x^1,u^1) + \rho_2 d_2^2(v^1,y^1). \quad (15)$$

Since $(x^1,y^1,x^2,y^2,z^1,z^2,p,r)$ is feasible for primal problem (SMP) and $(u^1,v^1,u^2,v^2,w^1,w^2,q,s)$ is feasible for dual problem (SMD), by the dual constraint (7), the vector $a^1 = \nabla_{x^1} f(u^1,v^1) + D_1 w^1 + \nabla_{x^1x^1} f(u^1,v^1)q \in C_1^*$, and so from the hypothesis (A), we obtain

$$F_1(x^1,u^1;a^1) + (a^1)^T u^1 \geq 0. \quad (16)$$

Similarly,

$$F_2(v^1,y^1;a^2) + (a^2)^T y^1 \geq 0, \quad (17)$$

for the vector $a^2 = -[\nabla_{y^1} f(x^1,y^1) - E_1 z^1 + \nabla_{y^1y^1} f(x^1,y^1)p] \in C_3^*$.

Using (16) and (17) and hypothesis (iii) in (15), we have

$$f(x^1,y^1)-f(u^1,v^1)+(x_1)^T D_1 w^1 - (u_1)^T D_1 w^1 - (y^1)^T E_1 z^1 + (v^1)^T E_1 z^1$$
$$+\frac{1}{2}q^T \nabla_{x^1x^1} f(u^1,v^1)q - \frac{1}{2}p^T \nabla_{y^1y^1} f(x^1,y^1)p \geq -(u^1)^T a^1 - (y^1)^T a^2.$$

Substituting the values of a^1 and a^2, we get

$$f(x^1,y^1)+(x^1)^T D_1 w^1 - (y^1)^T[\nabla_{y^1} f(x^1,y^1) + \nabla_{y^1y^1} f(x^1,y^1)p] - \frac{1}{2}p^T \nabla_{y^1y^1} f(x^1,y^1)p$$
$$\geq f(u^1,v^1) - (v^1)^T E_1 z^1 - (u^1)^T[\nabla_{x^1} f(u^1,v^1) + \nabla_{x^1x^1} f(u^1,v^1)q] - \frac{1}{2}q^T \nabla_{x^1x^1} f(u^1,v^1)q.$$

Applying the Schwartz inequality and using (4) and (10), we have

$$f(x^1,y^1)+((x^1)^T D_1 x^1)^{\frac{1}{2}} - (y^1)^T[\nabla_{y^1} f(x^1,y^1) + \nabla_{y^1y^1} f(x^1,y^1)p] - \frac{1}{2}p^T \nabla_{y^1y^1} f(x^1,y^1)p$$
$$\geq f(u^1,v^1) - ((v^1)^T E_1 v^1)^{\frac{1}{2}} - (u^1)^T[\nabla_{x^1} f(u^1,v^1) + \nabla_{x^1x^1} f(u^1,v^1)q] - \frac{1}{2}q^T \nabla_{x^1x^1} f(u^1,v^1)q. \quad (18)$$

By hypothesis (C) and the dual constraint (8), we obtain

$$G_1(x^2,u^2;\nabla_{x^2} g(u^2,v^2) + D_2 w^2 + \nabla_{x^2x^2} g(u^2,v^2)s) \geq -(u^2)^T[\nabla_{x^2} g(u^2,v^2) + D_2 w^2 + \nabla_{x^2x^2} g(u^2,v^2)s],$$

which on using the dual constraint (9) yields

$$G_1(x^2,u^2;\nabla_{x^2}g(u^2,v^2)+D_2w^2+\nabla_{x^2x^2}g(u^2,v^2)s)\geq 0.$$

Since $g(.,v^2)+(.)^TD_2w^2$ is second-order (G_1,σ_1) pseudo-convex at u^2, we have

$$g(x^2,v^2)+(x^2)^TD_2w^2\geq g(u^2,v^2)+(u^2)^TD_2w^2-\frac{1}{2}s^T\nabla_{x^2x^2}g(u^2,v^2)s+\sigma_1d_3^2(x^2,u^2). \qquad (19)$$

Similarly, from (2) and (3) and hypothesis (D), along with second-order (G_2,σ_2) pseudo-convexity of $-(g(x^2,.)-(.)^TE_2z^2)$ at y^2, we get

$$g(x^2,y^2)-(y^2)^TE_2z^2\geq g(x^2,v^2)-(v^2)^TE_2z^2+\frac{1}{2}r^T\nabla_{y^2y^2}g(x^2,y^2)r+\sigma_2d_4^2(v^2,y^2). \qquad (20)$$

Adding inequalities (19) and (20) and using hypothesis (iv), we obtain

$$g(x^2,y^2)+(x^2)^TD_2w^2-(y^2)^TE_2z^2-\frac{1}{2}r^T\nabla_{y^2y^2}g(x^2,y^2)r$$
$$\geq g(u^2,v^2)+(u^2)^TD_2w^2-(v^2)^TE_2z^2-\frac{1}{2}s^T\nabla_{x^2x^2}g(u^2,v^2)s.$$

Applying the Schwartz inequality and using (5) and (11), we have

$$g(x^2,y^2)+((x^2)^TD_2x^2)^{\frac{1}{2}}-(y^2)^TE_2z^2-\frac{1}{2}r^T\nabla_{y^2y^2}g(x^2,y^2)r$$
$$\geq g(u^2,v^2)+(u^2)^TD_2w^2-((v^2)^TE_2v^2)^{\frac{1}{2}}-\frac{1}{2}s^T\nabla_{x^2x^2}g(u^2,v^2)s. \qquad (21)$$

The expressions (18) and (21) together yield

$$f(x^1,y^1)+((x^1)^TD_1x^1)^{\frac{1}{2}}+g(x^2,y^2)+((x^2)^TD_2x^2)^{\frac{1}{2}}-(y^2)^TE_2z^2-(y^1)^T[\nabla_{y^1}f(x^1,y^1)$$
$$+\nabla_{y^1y^1}f(x^1,y^1)p]-\frac{1}{2}p^T\nabla_{y^1y^1}f(x^1,y^1)p-\frac{1}{2}r^T\nabla_{y^2y^2}g(x^2,y^2)r\geq f(u^1,v^1)-((v^1)^TE_1v^1)^{\frac{1}{2}}$$
$$+g(u^2,v^2)-((v^2)^TE_2v^2)^{\frac{1}{2}}+(u^2)^TD_2w^2-(u^1)^T[\nabla_{x^1}f(u^1,v^1)+\nabla_{x^1x^1}f(u^1,v^1)q]$$
$$-\frac{1}{2}q^T\nabla_{x^1x^1}f(u^1,v^1)q-\frac{1}{2}s^T\nabla_{x^2x^2}g(u^2,v^2)s,$$

that is,

$$L(x^1,y^1,x^2,y^2,z^2,p,r)\geq M(u^1,v^1,u^2,v^2,w^2,q,s).$$

Theorem 2 (Weak duality)

Let $(x^1,y^1,x^2,y^2,z^1,z^2,p,r)$ be feasible for SMP and $(u^1,v^1,u^2,v^2,w^1,w^2,q,s)$ be feasible for SMD. Let the sub-linear functionals $F_1:R^{|J_1|}\times R^{|J_1|}\times R^{|J_1|}\mapsto R$, $F_2:R^{|K_1|}\times R^{|K_1|}\times R^{|K_1|}\mapsto R$, $G_1:R^{|J_2|}\times R^{|J_2|}\times R^{|J_2|}\mapsto R$ and $G_2:R^{|K_2|}\times R^{|K_2|}\times R^{|K_2|}\mapsto R$ satisfy the following conditions:

$$F_1(x^1,u^1;a^1)+(a^1)^Tu^1\geq 0,\ \text{for all } a^1\in C_1^*, \qquad (A)$$

$F_2(v^1, y^1; a^2) + (a^2)^T y^1 \geq 0$, for all $a^2 \in C_3^*$, (B)

$G_1(x^2, u^2; b^1) + (b^1)^T u^2 \geq 0$, for all $b^1 \in C_2^*$, (C)

$G_2(v^2, y^2; b^2) + (b^2)^T y^2 \geq 0$, for all $b^2 \in C_4^*$. (D)

Suppose that

(i) $f(.,v^1) + (.)^T D_1 w^1$ is second-order (F_1, ρ_1) convex at u^1, and $-(f(x^1,.) - (.)^T E_1 z^1)$ is second-order (F_2, ρ_2) convex at y^1,

(ii) $g(.,v^2) + (.)^T D_2 w^2$ is second-order (G_1, σ_1) convex at u^2, and $-(g(x^2,.) - (.)^T E_2 z^2)$ is second-order (G_2, σ_2) convex at y^2,

(iii) either $\rho_1 d_1^2(x^1, u^1) + \rho_2 d_2^2(v^1, y^1) \geq 0$ or $\rho_1, \rho_2 \geq 0$, and

(iv) either $\sigma_1 d_3^2(x^2, u^2) + \sigma_2 d_4^2(v^2, y^2) \geq 0$ or $\sigma_1, \sigma_2 \geq 0$.

Then

$$L(x^1, y^1, x^2, y^2, z^2, p, r) \geq M(u^1, v^1, u^2, v^2, w^2, q, s).$$

Proof. By the second-order (F_1, ρ_1) convexity of $f(.,v^1) + (.)^T D_1 w^1$ at u^1 and the second-order (F_2, ρ_2) convexity of $-(f(x^1,.) - (.)^T E_1 z^1)$ at y^1, we have

$$f(x^1, v^1) + (x_1)^T D_1 w^1 - f(u^1, v^1) - (u_1)^T D_1 w^1 + \frac{1}{2} q^T \nabla_{x^1 x^1} f(u^1, v^1) q$$
$$\geq F_1(x^1, u^1; \nabla_{x^1} f(u^1, v^1) + D_1 w^1 + \nabla_{x^1 x^1} f(u^1, v^1) q) + \rho_1 d_1^2(x^1, u^1) \quad (22)$$

and

$$f(x^1, y^1) - (y^1)^T E_1 z^1 - f(x^1, v^1) + (v^1)^T E_1 z^1 - \frac{1}{2} p^T \nabla_{y^1 y^1} f(x^1, y^1) p$$
$$\geq F_2(v^1, y^1; -(\nabla_{y^1} f(x^1, y^1) - E_1 z^1 + \nabla_{y^1 y^1} f(x^1, y^1) p)) + \rho_2 d_2^2(v^1, y^1). \quad (23)$$

Adding the inequalities (22) and (23), we obtain

$$f(x^1, y^1) - f(u^1, v^1) + (x_1)^T D_1 w^1 - (u_1)^T D_1 w^1 - (y^1)^T E_1 z^1 + (v^1)^T E_1 z^1$$
$$+ \frac{1}{2} q^T \nabla_{x^1 x^1} f(u^1, v^1) q - \frac{1}{2} p^T \nabla_{y^1 y^1} f(x^1, y^1) p \geq F_1(x^1, u^1; \nabla_{x^1} f(u^1, v^1) + D_1 w^1 + \nabla_{x^1 x^1} f(u^1, v^1) q)$$
$$+ F_2(v^1, y^1; -(\nabla_{y^1} f(x^1, y^1) - E_1 z^1 + \nabla_{y^1 y^1} f(x^1, y^1) p)) + \rho_1 d_1^2(x^1, u^1) + \rho_2 d_2^2(v^1, y^1). \quad (24)$$

Since $(x^1, y^1, x^2, y^2, z^1, z^2, p, r)$ is feasible for primal problem (SMP) and $(u^1, v^1, u^2, v^2, w^1, w^2, q, s)$ is feasible for dual problem (SMD), by the dual constraint (7), the vector $a^1 = \nabla_{x^1} f(u^1, v^1) + D_1 w^1 + \nabla_{x^1 x^1} f(u^1, v^1) q \in C_1^*$, and so from hypothesis (A), we obtain

$$F_1(x^1, u^1; a^1) + (a^1)^T u^1 \geq 0. \quad (25)$$

Similarly,

$$F_2(v^1, y^1; a^2) + (a^2)^T y^1 \geq 0, \tag{26}$$

for the vector $a^2 = -[\nabla_{y^1} f(x^1, y^1) - E_1 z^1 + \nabla_{y^1 y^1} f(x^1, y^1) p] \in C_3^*$.

Using (25) and (26) and hypothesis (*iii*) in (24), we have

$$f(x^1, y^1) - f(u^1, v^1) + (x_1)^T D_1 w^1 - (u_1)^T D_1 w^1 - (y^1)^T E_1 z^1 + (v^1)^T E_1 z^1$$
$$+ \frac{1}{2} q^T \nabla_{x^1 x^1} f(u^1, v^1) q - \frac{1}{2} p^T \nabla_{y^1 y^1} f(x^1, y^1) p \geq -(u^1)^T a^1 - (y^1)^T a^2.$$

Substituting the values of a^1 and a^2, we get

$$f(x^1, y^1) + (x^1)^T D_1 w^1 - (y^1)^T [\nabla_{y^1} f(x^1, y^1) + \nabla_{y^1 y^1} f(x^1, y^1) p] - \frac{1}{2} p^T \nabla_{y^1 y^1} f(x^1, y^1) p$$
$$\geq f(u^1, v^1) - (v^1)^T E_1 z^1 - (u^1)^T [\nabla_{x^1} f(u^1, v^1) + \nabla_{x^1 x^1} f(u^1, v^1) q] - \frac{1}{2} q^T \nabla_{x^1 x^1} f(u^1, v^1) q.$$

Applying the Schwartz inequality and using (4) and (10), we have

$$f(x^1, y^1) + ((x^1)^T D_1 x^1)^{\frac{1}{2}} - (y^1)^T [\nabla_{y^1} f(x^1, y^1) + \nabla_{y^1 y^1} f(x^1, y^1) p] - \frac{1}{2} p^T \nabla_{y^1 y^1} f(x^1, y^1) p$$
$$\geq f(u^1, v^1) - ((v^1)^T E_1 v^1)^{\frac{1}{2}} - (u^1)^T [\nabla_{x^1} f(u^1, v^1) + \nabla_{x^1 x^1} f(u^1, v^1) q] - \frac{1}{2} q^T \nabla_{x^1 x^1} f(u^1, v^1) q. \tag{27}$$

By second-order (G_1, σ_1) convexity of $g(., v^2) + (.)^T D_2 w^2$ at u^2 and the second-order (G_2, σ_2) convexity of $-(g(x^2, .) - (.)^T E_2 z^2)$ at y^2, we have

$$g(x^2, v^2) + (x^2)^T D_2 w^2 - g(u^2, v^2) - (u^2)^T D_2 w^2 + \frac{1}{2} s^T \nabla_{x^2 x^2} g(u^2, v^2) s$$
$$\geq G_1(x^2, u^2; \nabla_{x^2} g(u^2, v^2) + D_2 w^2 + \nabla_{x^2 x^2} g(u^2, v^2) s) + \sigma_1 d_3^2(x^2, u^2). \tag{28}$$

and

$$g(x^2, y^2) - (y^2)^T E_2 z^2 - g(x^2, v^2) + (v^2)^T E_2 z^2 - \frac{1}{2} r^T \nabla_{y^2 y^2} g(x^2, y^2) r$$
$$\geq G_2(v^2, y^2; -(\nabla_{y^2} g(x^2, y^2) - E_2 z^2 + \nabla_{y^2 y^2} g(x^2, y^2) r)) + \sigma_2 d_4^2(v^2, y^2). \tag{29}$$

Adding the inequalities (28) and (29), we get

$$g(x^2, y^2) + (x^2)^T D_2 w^2 - (y^2)^T E_2 z^2 - g(u^2, v^2) - (u^2)^T D_2 w^2 + (v^2)^T E_2 z^2$$
$$+ \frac{1}{2} s^T \nabla_{x^2 x^2} g(u^2, v^2) s - \frac{1}{2} r^T \nabla_{y^2 y^2} g(x^2, y^2) r \geq G_1(x^2, u^2; \nabla_{x^2} g(u^2, v^2) + D_2 w^2 + \nabla_{x^2 x^2} g(u^2, v^2) s)$$
$$+ G_2(v^2, y^2; -(\nabla_{y^2} g(x^2, y^2) - E_2 z^2 + \nabla_{y^2 y^2} g(x^2, y^2) r)) + \sigma_1 d_3^2(x^2, u^2) + \sigma_2 d_4^2(v^2, y^2). \tag{30}$$

Since $(x^1, y^1, x^2, y^2, z^1, z^2, p, r)$ is feasible for primal problem (SMP) and $(u^1, v^1, u^2, v^2, w^1, w^2, q, s)$ is feasible for dual problem (SMD), by the dual constraint (8), the vector $b^1 = \nabla_{x^2} g(u^2, v^2) + D_2 w^2 + \nabla_{x^2 x^2} g(u^2, v^2) s \in C_2^*$, and so from hypothesis (C), we obtain

$$G_1(x^2,u^2;\nabla_{x^2}g(u^2,v^2)+D_2w^2+\nabla_{x^2x^2}g(u^2,v^2)s) \geq -(u^2)^T[\nabla_{x^2}g(u^2,v^2)+D_2w^2+\nabla_{x^2x^2}g(u^2,v^2)s],$$

which on using the dual constraint (9) yields

$$G_1(x^2,u^2;\nabla_{x^2}g(u^2,v^2)+D_2w^2+\nabla_{x^2x^2}g(u^2,v^2)s) \geq 0. \tag{31}$$

Similarly, from (2), (3) and hypothesis (D), we have

$$G_2(v^2,y^2;-(\nabla_{y^2}g(x^2,y^2)-E_2z^2+\nabla_{y^2y^2}g(x^2,y^2)r)) \geq 0. \tag{32}$$

Using (31), (32) and hypothesis (iv) in (30), we obtain

$$g(x^2,y^2)+(x^2)^TD_2w^2-(y^2)^TE_2z^2-\frac{1}{2}r^T\nabla_{y^2y^2}g(x^2,y^2)r$$
$$\geq g(u^2,v^2)+(u^2)^TD_2w^2-(v^2)^TE_2z^2-\frac{1}{2}s^T\nabla_{x^2x^2}g(u^2,v^2)s.$$

Applying the Schwartz inequality and using (5) and (11), we have

$$g(x^2,y^2)+((x^2)^TD_2x^2)^{\frac{1}{2}}-(y^2)^TE_2z^2-\frac{1}{2}r^T\nabla_{y^2y^2}g(x^2,y^2)r$$
$$\geq g(u^2,v^2)+(u^2)^TD_2w^2-((v^2)^TE_2v^2)^{\frac{1}{2}}-\frac{1}{2}s^T\nabla_{x^2x^2}g(u^2,v^2)s. \tag{33}$$

Inequalities (27) and (33) together yield

$$f(x^1,y^1)+((x^1)^TD_1x^1)^{\frac{1}{2}}+g(x^2,y^2)+((x^2)^TD_2x^2)^{\frac{1}{2}}-(y^2)^TE_2z^2-(y^1)^T[\nabla_{y^1}f(x^1,y^1)$$
$$+\nabla_{y^1y^1}f(x^1,y^1)p]-\frac{1}{2}p^T\nabla_{y^1y^1}f(x^1,y^1)p-\frac{1}{2}r^T\nabla_{y^2y^2}g(x^2,y^2)r \geq f(u^1,v^1)-((v^1)^TE_1v^1)^{\frac{1}{2}}$$
$$+g(u^2,v^2)-((v^2)^TE_2v^2)^{\frac{1}{2}}+(u^2)^TD_2w^2-(u^1)^T[\nabla_{x^1}f(u^1,v^1)+\nabla_{x^1x^1}f(u^1,v^1)q]$$
$$-\frac{1}{2}q^T\nabla_{x^1x^1}f(u^1,v^1)q-\frac{1}{2}s^T\nabla_{x^2x^2}g(u^2,v^2)s,$$

that is,

$$L(x^1,y^1,x^2,y^2,z^2,p,r) \geq M(u^1,v^1,u^2,v^2,w^2,q,s).$$

Theorem 3 (Strong duality)

Let $f:R^{|J_1|}\times R^{|K_1|}\to R$ and $g:R^{|J_2|}\times R^{|K_2|}\to R$ be differentiable functions and let $(\bar{x}^1,\bar{y}^1,\bar{x}^2,\bar{y}^2,\bar{z}^1,\bar{z}^2,\bar{p},\bar{r})$ be a local optimal solution of SMP. Suppose that

(i) the matrix $\nabla_{y^1y^1}f(\bar{x}^1,\bar{y}^1)$ is non-singular,

(ii) $\nabla_{y^2y^2}g(\bar{x}^2,\bar{y}^2)$ is positive definite and $\bar{r}^T(\nabla_{y^2}g(\bar{x}^2,\bar{y}^2)-E_2\bar{z}^2)\geq 0$ or $\nabla_{y^2y^2}g(\bar{x}^2,\bar{y}^2)$ is negative definite and $\bar{r}^T(\nabla_{y^2}g(\bar{x}^2,\bar{y}^2)-E_2\bar{z}^2)\leq 0$,

(iii) $\nabla_{y^2}g(\bar{x}^2,\bar{y}^2)-E_2\bar{z}^2+\nabla_{y^2y^2}g(\bar{x}^2,\bar{y}^2)\bar{r}\neq 0$, and

(iv) one of the matrices $\frac{\partial}{\partial y_i^1}\left(\nabla_{y^1y^1} f(\bar{x}^1,\bar{y}^1)\right), i=1,2,\ldots,|K_1|,$ is positive or negative definite.

Then $\bar{p}=0,\ \bar{r}=0$ and there exist $\bar{w}^1 \in R^{|J_1|}$ and $\bar{w}^2 \in R^{|J_2|}$ such that $(\bar{x}^1,\bar{y}^1,\bar{x}^2,\bar{y}^2,\bar{w}^1,\bar{w}^2,\bar{q}=0,\bar{s}=0)$ is feasible for SMD and the objective function values of SMP and SMD are equal. Furthermore, if the assumptions of weak duality theorem (1 or 2) are satisfied for all feasible solutions of SMP and SMD, then $(\bar{x}^1,\bar{y}^1,\bar{x}^2,\bar{y}^2,\bar{z}^1,\bar{z}^2,\bar{p},\bar{r})$ and $(\bar{x}^1,\bar{y}^1,\bar{x}^2,\bar{y}^2,\bar{w}^1,\bar{w}^2,\bar{q},\bar{s})$ are global optimal solutions for SMP and SMD respectively.

Proof. Since $(\bar{x}^1,\bar{y}^1,\bar{x}^2,\bar{y}^2,\bar{z}^1,\bar{z}^2,\bar{p},\bar{r})$ is a local solution of SMP, there exist $\alpha \in R_+$, $\beta \in C_3$, $\gamma \in C_4$, $\delta \in R_+$, $\mu \in R_+$ and $\nu \in R_+$ such that the following by Fritz John optimality conditions studied in Suneja et al. [20] and in Schechter [21] are satisfied at $(\bar{x}^1,\bar{y}^1,\bar{x}^2,\bar{y}^2,\bar{z}^1,\bar{z}^2,\bar{p},\bar{r})$:

$$\{\alpha^T(\nabla_{x^1} f(\bar{x}^1,\bar{y}^1)+D_1\bar{w}^1)+\nabla_{y^1x^1} f(\bar{x}^1,\bar{y}^1)\left[\beta-\alpha\bar{y}^1\right] + \nabla_{x^1}(\nabla_{y^1y^1} f(\bar{x}^1,\bar{y}^1)\bar{p})[\beta-\alpha(\bar{y}^1+\frac{1}{2}\bar{p})]\}(x^1-\bar{x}^1) \geq 0,\ \forall\, x^1 \in C_1, \tag{34}$$

$$\{\alpha^T(\nabla_{x^2} g(\bar{x}^2,\bar{y}^2)+D_2\bar{w}^2)+\nabla_{y^2x^2} g(\bar{x}^2,\bar{y}^2)\left[\gamma-\delta\bar{y}^2\right] + \nabla_{x^2}(\nabla_{y^2y^2} g(\bar{x}^2,\bar{y}^2)\bar{r})[\gamma-\delta\bar{y}^2-\frac{1}{2}\alpha\bar{r}]\}(x^2-\bar{x}^2) \geq 0,\ \forall\, x^2 \in C_2, \tag{35}$$

$$\nabla_{y^1y^1} f(\bar{x}^1,\bar{y}^1)[\beta-\alpha(\bar{y}^1+\bar{p})]+\nabla_{y^1}(\nabla_{y^1y^1} f(\bar{x}^1,\bar{y}^1)\bar{p})[\beta-\alpha(\bar{y}^1+\frac{1}{2}\bar{p})]=0, \tag{36}$$

$$(\nabla_{y^2} g(\bar{x}^2,\bar{y}^2)-E_2\bar{z}^2)\left[\alpha-\delta\right]+\nabla_{y^2y^2} g(\bar{x}^2,\bar{y}^2)[\gamma-\delta(\bar{y}^2+\bar{r})] + \nabla_{y^2}(\nabla_{y^2y^2} g(\bar{x}^2,\bar{y}^2)\bar{r})[\gamma-\delta\bar{y}^2-\frac{1}{2}\alpha\bar{r}]=0, \tag{37}$$

$$(-\beta E_1+\mu E_1\bar{z}^1)=0, \tag{38}$$

$$\alpha E_2\bar{y}^2+E_2(\gamma-\delta\bar{y}^2)=2\nu E_2\bar{z}^2, \tag{39}$$

$$\nabla_{y^1y^1} f(\bar{x}^1,\bar{y}^1)[\beta-\alpha(\bar{y}^1+\bar{p})]=0, \tag{40}$$

$$\nabla_{y^2y^2} g(\bar{x}^2,\bar{y}^2)[\gamma-\delta\bar{y}^2-\alpha\bar{r}]=0, \tag{41}$$

$$\beta^T[\nabla_{y^1} f(\bar{x}^1,\bar{y}^1)-E_1\bar{z}^1+\nabla_{y^1y^1} f(\bar{x}^1,\bar{y}^1)\bar{p}]=0, \tag{42}$$

$$\gamma^T[\nabla_{y^2} g(\bar{x}^2,\bar{y}^2)-E_2\bar{z}^2+\nabla_{y^2y^2} g(\bar{x}^2,\bar{y}^2)\bar{r}]=0, \tag{43}$$

$$\delta(\bar{y}^2)^T[\nabla_{y^2} g(\bar{x}^2,\bar{y}^2)-E_2\bar{z}^2+\nabla_{y^2y^2} g(\bar{x}^2,\bar{y}^2)\bar{r}]=0, \tag{44}$$

$$(\bar{x}^1)^T D_1\bar{w}^1=((\bar{x}^1)^T D_1\bar{x}^1)^{\frac{1}{2}}, \tag{45}$$

$$(\bar{x}^2)^T D_2\bar{w}^2=((\bar{x}^2)^T D_2\bar{x}^2)^{\frac{1}{2}}, \tag{46}$$

$$(\bar{w}^1)^T D_1\bar{w}^1 \leq 1, \tag{47}$$

$$(\bar{w}^2)^T D_2 \bar{w}^2 \le 1, \tag{48}$$

$$\mu((\bar{z}^1)^T E_1 \bar{z}^1 - 1) = 0, \tag{49}$$

$$\nu((\bar{z}^2)^T E_2 \bar{z}^2 - 1) = 0, \tag{50}$$

$$(\alpha, \beta, \gamma, \delta, \mu, \nu) \neq 0. \tag{51}$$

Because of the non-singularity of $\nabla_{y^1 y^1} f(\bar{x}^1, \bar{y}^1)$, (40) yields

$$\beta = \alpha(\bar{y}^1 + \bar{p}). \tag{52}$$

Since $\nabla_{y^2 y^2} g(\bar{x}^2, \bar{y}^2)$ is positive or negative definite, (41) gives

$$\gamma = \delta \bar{y}^2 + \alpha \bar{r}. \tag{53}$$

Now, we claim that $\alpha > 0$. If possible, let $\alpha = 0$; then (53) gives $\gamma = \delta \bar{y}^2$.

Using (53) in (37), we get

$$(\alpha - \delta)[\nabla_{y^2} g(\bar{x}^2, \bar{y}^2) - E_2 \bar{z}^2 + \nabla_{y^2 y^2} g(\bar{x}^2, \bar{y}^2)\bar{r}] + \frac{1}{2}\nabla_{y^2}(\nabla_{y^2 y^2} g(\bar{x}^2, \bar{y}^2)\bar{r})\left[\gamma - \delta \bar{y}^2\right] = 0,$$

which, on using hypothesis (*iii*) and $\gamma = \delta \bar{y}^2$, yields $\alpha = \delta$. As $\alpha = 0$, therefore the equations $\alpha = \delta$ and $\gamma = \delta \bar{y}^2$ give $\delta = 0$ and $\gamma = 0$ respectively. Equation (52) gives $\beta = 0$. Also from equations (38) and (49), we have

$$\mu = \mu((\bar{z}^1)^T E_1 \bar{z}^1) = (\bar{z}^1)^T (\mu E_1 \bar{z}^1) = (\bar{z}^1)^T (E_1 \beta) = 0.$$

From (39) and (50), we get $\nu = 0$. Consequently, $(\alpha, \beta, \gamma, \delta, \mu, \nu) = 0$, contradicting (51). Hence

$$\alpha > 0. \tag{54}$$

Subtracting (44) from (43) yields

$$[\gamma - \delta(\bar{y}^2)]^T [\nabla_{y^2} g(\bar{x}^2, \bar{y}^2) - E_2 \bar{z}^2 + \nabla_{y^2 y^2} g(\bar{x}^2, \bar{y}^2)\bar{r}] = 0.$$

Using (53) and (54) in above equation, we get

$$\bar{r}^T(\nabla_{y^2} g(\bar{x}^2, \bar{y}^2) - E_2 \bar{z}^2) + \bar{r}^T \nabla_{y^2 y^2} g(\bar{x}^2, \bar{y}^2)\bar{r} = 0, \tag{55}$$

which contradicts hypothesis (*ii*) unless

$$\bar{r} = 0. \tag{56}$$

Equation (53) yields

$$\gamma = \delta \bar{y}^2. \tag{57}$$

Using (56) and (57) in (37), we obtain

$$(\alpha - \delta)(\nabla_{y^2} g(\bar{x}^2, \bar{y}^2) - E_2 \bar{z}^2) = 0,$$

which on using hypothesis (*iii*) and (56) gives

$$\alpha = \delta. \tag{58}$$

Since $\alpha > 0$, then obviously

$$\delta > 0. \tag{59}$$

Now, using (52) and (54) in (36), we get

$$(\nabla_{y^1}(\nabla_{y^1 y^1} f(\bar{x}^1, \bar{y}^1)\bar{p}))\bar{p} = 0,$$

which by hypothesis (iv) implies

$$\bar{p} = 0. \tag{60}$$

By equations (52) and (60), we have

$$\beta = \alpha \bar{y}^1. \tag{61}$$

From (54) and (61), we obtain

$$\bar{y}^1 = \frac{\beta}{\alpha} \in C_3.$$

Using (58)-(61) in (34), we get

$$(x^1 - \bar{x}^1)^T (\nabla_{x^1} f(\bar{x}^1, \bar{y}^1) + D_1 \bar{w}^1) \geq 0, \text{ for all } x^1 \in C_1. \tag{62}$$

Let $x^1 \in C_1$. Then $x^1 + \bar{x}^1 \in C_1$ as C_1 is a closed convex cone, and so (62) implies

$$(x^1)^T (\nabla_{x^1} f(\bar{x}^1, \bar{y}^1) + D_1 \bar{w}^1) \geq 0, \text{ for all } x^1 \in C_1.$$

Therefore,

$$\nabla_{x^1} f(\bar{x}^1, \bar{y}^1) + D_1 \bar{w}^1 \in C_1^*. \tag{63}$$

From (35) and (56)-(59), we have

$$(x^2 - \bar{x}^2)^T (\nabla_{x^2} g(\bar{x}^2, \bar{y}^2) + D_2 \bar{w}^2) \geq 0, \text{ for all } x^2 \in C_2. \tag{64}$$

Let $x^2 \in C_2$. Then $x^2 + \bar{x}^2 \in C_2$ as C_2 is a closed convex cone, and so (64) implies

$$(x^2)^T (\nabla_{x^2} g(\bar{x}^2, \bar{y}^2) + D_2 \bar{w}^2) \geq 0, \text{ for all } x^2 \in C_2.$$

Therefore,

$$\nabla_{x^2} g(\bar{x}^2, \bar{y}^2) + D_2 \bar{w}^2 \in C_2^*. \tag{65}$$

Also from (57) and (59), we have

$$\bar{y}^2 = \frac{\gamma}{\delta} \in C_4.$$

Now, letting $x^2 = 0$ and $x^2 = 2\bar{x}^2$ in (64), we get

$$(\bar{x}^2)^T (\nabla_{x^2} g(\bar{x}^2, \bar{y}^2) + D_2 \bar{w}^2) = 0. \tag{66}$$

Thus $(\bar{x}^1, \bar{y}^1, \bar{x}^2, \bar{y}^2, \bar{w}^1, \bar{w}^2, \bar{q} = 0, \bar{s} = 0)$ satisfies the dual constraints from (7) to (12) and so it is a feasible solution for the dual problem (SMD).

Now let $\frac{2\nu}{\alpha} = a$. Then $a \geq 0$ and from (39) and (57)

$$E_2\bar{y}^2 = aE_2\bar{z}^2, \tag{67}$$

which is the condition for equality in the Schwartz inequality. Therefore

$(\bar{y}^2)^T E_2\bar{z}^2 = ((\bar{y}^2)^T E_2\bar{y}^2)^{\frac{1}{2}}((\bar{z}^2)^T E_2\bar{z}^2)^{\frac{1}{2}}$.

In the case of $\nu > 0$, (50) gives $(\bar{z}^2)^T E_2\bar{z}^2 = 1$ and so $(\bar{y}^2)^T E_2\bar{z}^2 = ((\bar{y}^2)^T E_2\bar{y}^2)^{\frac{1}{2}}$. In the case of $\nu = 0$, (67) gives $E_2\bar{y}^2 = 0$ and so $(\bar{y}^2)^T E_2\bar{z}^2 = ((\bar{y}^2)^T E_2\bar{y}^2)^{\frac{1}{2}} = 0$. Thus, in either case,

$$(\bar{y}^2)^T E_2\bar{z}^2 = ((\bar{y}^2)^T E_2\bar{y}^2)^{\frac{1}{2}}. \tag{68}$$

By putting $x^1 = 0$ and $x^1 = 2\bar{x}^1$ in (62), we obtain

$$(\bar{x}^1)^T(\nabla_{x^1} f(\bar{x}^1, \bar{y}^1) + D_1\bar{w}^1) = 0. \tag{69}$$

Also, (45) yields

$$(\bar{x}^1)^T \nabla_{x^1} f(\bar{x}^1, \bar{y}^1) = -(\bar{x}^1)^T D_1\bar{w}^1 = -((\bar{x}^1)^T D_1\bar{x}^1)^{\frac{1}{2}}. \tag{70}$$

From (42), (54), (60) and (61) we get

$$(\bar{y}^1)^T \nabla_{y^1} f(\bar{x}^1, \bar{y}^1) = (\bar{y}^1)^T E_1\bar{z}^1. \tag{71}$$

Equation (38), implies

$E_1\beta = \mu E_1\bar{z}^1$.

Using (61) in the above equation,

$$E_1\bar{y}^1 = \frac{\mu}{\alpha} E_1\bar{z}^1. \tag{72}$$

Since equation (72) is the condition for the Schwartz inequality to hold as equality, so

$(\bar{y}^1)^T E_1\bar{z}^1 = ((\bar{y}^1)^T E_1\bar{y}^1)^{\frac{1}{2}}((\bar{z}^1)^T E_1\bar{z}^1)^{\frac{1}{2}}$.

In the case of $\mu > 0$, the equation (49) implies $(\bar{z}^1)^T E_1\bar{z}^1 = 1$ and so $(\bar{y}^1)^T E_1\bar{z}^1 = ((\bar{y}^1)^T E_1\bar{y}^1)^{\frac{1}{2}}$.

In the case of $\mu = 0$, the equation (72) gives $E_1\bar{y}^1 = 0$ and so $(\bar{y}^1)^T E_1\bar{z}^1 = ((\bar{y}^1)^T E_1\bar{y}^1)^{\frac{1}{2}} = 0$.

Thus, in either case $(\bar{y}^1)^T E_1\bar{z}^1 = ((\bar{y}^1)^T E_1\bar{y}^1)^{\frac{1}{2}}$.

Now equation (71) becomes

$$(\bar{y}^1)^T \nabla_{y^1} f(\bar{x}^1, \bar{y}^1) = (\bar{y}^1)^T E_1\bar{z}^1 = ((\bar{y}^1)^T E_1\bar{y}^1)^{\frac{1}{2}}. \tag{73}$$

Therefore, using (46), (56), (60), (68), (70) and (73), we obtain the following:

$$f(\bar{x}^1,\bar{y}^1)+((\bar{x}^1)^T D_1\bar{x}^1)^{\frac{1}{2}}+g(\bar{x}^2,\bar{y}^2)+((\bar{x}^2)^T D_2\bar{x}^2)^{\frac{1}{2}}-(\bar{y}^2)^T E_2\bar{z}^2-(\bar{y}^1)^T[\nabla_{y^1}f(\bar{x}^1,\bar{y}^1)$$

$$+\nabla_{y^1y^1}f(\bar{x}^1,\bar{y}^1)\bar{p}]-\frac{1}{2}\bar{p}^T\nabla_{y^1y^1}f(\bar{x}^1,\bar{y}^1)\bar{p}-\frac{1}{2}\bar{r}^T\nabla_{y^2y^2}g(\bar{x}^2,\bar{y}^2)\bar{r}\ =\ f(\bar{x}^1,\bar{y}^1)-((\bar{y}^1)^T E_1\bar{y}^1)^{\frac{1}{2}}$$

$$+g(\bar{x}^2,\bar{y}^2)-((\bar{y}^2)^T E_2\bar{y}^2)^{\frac{1}{2}}+(\bar{x}^2)^T D_2\bar{w}^2-(\bar{x}^1)^T[\nabla_{x^1}f(\bar{x}^1,\bar{y}^1)+\nabla_{x^1x^1}f(\bar{x}^1,\bar{y}^1)\bar{q}]$$

$$-\frac{1}{2}\bar{q}^T\nabla_{x^1x^1}f(\bar{x}^1,\bar{y}^1)\bar{q}-\frac{1}{2}\bar{s}^T\nabla_{x^2x^2}g(\bar{x}^2,\bar{y}^2)\bar{s},$$

that is, the two objective function values are equal.

Finally, from Theorem 1 or 2, we get that $(\bar{x}^1,\bar{y}^1,\bar{x}^2,\bar{y}^2,\bar{z}^1,\bar{z}^2,\bar{p},\bar{r})$ and $(\bar{x}^1,\bar{y}^1,\bar{x}^2,\bar{y}^2,\bar{w}^1,\bar{w}^2,\bar{q},\bar{s})$ are global optimal solutions for SMP and SMD respectively.

Theorem 4 (Converse duality)

Let $f:R^{|J_1|}\times R^{|K_1|}\to R$ and $g:R^{|J_2|}\times R^{|K_2|}\to R$ be differentiable functions and let $(\bar{u}^1,\bar{v}^1,\bar{u}^2,\bar{v}^2,\bar{w}^1,\bar{w}^2,\bar{q},\bar{s})$ be a local optimal solution of SMD. Suppose that

(i) the matrix $\nabla_{x^1x^1}f(\bar{u}^1,\bar{v}^1)$ is non-singular,

(ii) $\nabla_{x^2x^2}g(\bar{u}^2,\bar{v}^2)$ is positive definite and $\bar{s}^T(\nabla_{x^2}g(\bar{u}^2,\bar{v}^2)+D_2\bar{w}^2)\geq 0$ or $\nabla_{x^2x^2}g(\bar{u}^2,\bar{v}^2)$ is negative definite and $\bar{s}^T(\nabla_{x^2}g(\bar{u}^2,\bar{v}^2)+D_2\bar{w}^2)\leq 0$,

(iii) $\nabla_{x^2}g(\bar{u}^2,\bar{v}^2)+D_2\bar{w}^2+\nabla_{x^2x^2}g(\bar{u}^2,\bar{v}^2)\bar{s}\neq 0$, and

(iv) one of the matrices $\frac{\partial}{\partial x_i^1}\left(\nabla_{x^1x^1}f(\bar{u}^1,\bar{v}^1)\right), i=1,2,\ldots,|J_1|,$ is positive or negative definite.

Then $\bar{q}=0,\ \bar{s}=0$ and there exist $\bar{z}^1\in R^{|K_1|}$ and $\bar{z}^2\in R^{|K_2|}$ such that $(\bar{u}^1,\bar{v}^1,\bar{u}^2,\bar{v}^2,\bar{z}^1,\bar{z}^2,\bar{p}=0,\bar{r}=0)$ is feasible for SMP and the objective function values of SMP and SMD are equal. Furthermore, if the assumptions of weak duality theorem (1 or 2) are satisfied for all feasible solutions of SMP and SMD, then $(\bar{u}^1,\bar{v}^1,\bar{u}^2,\bar{v}^2,\bar{w}^1,\bar{w}^2,\bar{q},\bar{s})$ and $(\bar{u}^1,\bar{v}^1,\bar{u}^2,\bar{v}^2,\bar{z}^1,\bar{z}^2,\bar{p},\bar{r})$ are global optimal solutions for SMD and SMP respectively.

Proof. It follows on the lines of Theorem 3.

Special Cases

In this section, we consider some of the special cases of our problems: SMP and SMD. For all these cases, $D_1=\{0\}$, $D_2=\{0\}$, $E_1=\{0\}$, $E_2=\{0\}$, $C_1=R_+^{|J_1|}, C_2=R_+^{|J_2|}, C_3=R_+^{|K_1|}$ and $C_4=R_+^{|K_2|}$.

1. If $J_2=\varnothing$ and $K_2=\varnothing$, then our problems: SMP and SMD reduce to the programmes: SP and SD studied by Gulati et al. [22] and if $J_1=\varnothing$ and $K_1=\varnothing$ in SMP and SMD, then the programmes SP1 and SD1 in Gulati et al. [22] are obtained.

2. By eliminating the second-order terms, our problems: SMP and SMD reduce to the mixed symmetric dual programmes studied by Chandra et al. [23].

3. If $J_1 = \varnothing$ and $K_1 = \varnothing$, then SMP and SMD reduce to programmes studied by Yang [24] with the omission of non-negativity constraints from SMP and SMD.
4. If $p = 0$, $q = 0$, $J_2 = \varnothing$ and $K_2 = \varnothing$ in SMP and SMD, then the programmes WP and WD in Chandra et al. [25] are obtained and if we take $r = 0$, $s = 0$, $J_1 = \varnothing$ and $K_1 = \varnothing$, then our problems become the programmes MP and MD studied by Chandra et al. [25].

CONCLUSIONS

Weak, strong and converse duality theorems have been established for a pair of non-differentiable second-order mixed symmetric dual programmes with cone constraints under second-order (F, ρ) convexity/pseudo-convexity assumptions. It is to be noted that previously known results are special cases of our study. However, it is not clear whether the second-order mixed symmetric duality in mathematical programming can be further extended to second-order multi-objective symmetric dual programmes.

ACKNOWLEDGEMENTS

The authors wish to thank the referees for several valuable suggestions which have considerably improved the presentation of the paper.

REFERENCES

1. G. B. Dantzig, E. Eisenberg and R. W. Cottle, "Symmetric dual non linear programs", *Pacific J. Math.*, **1965,** *15*, 809-812.
2. B. Mond, "A symmetric dual theorem for non linear programs", *Quart. J. Appl. Math.,* **1965,** *23*, 265-269.
3. M. S. Bazaraa and J. J. Goode, "On symmetric duality in non-linear programming", *Oper. Res.* **1973,** *21*, 1-9.
4. B. Mond and T. Weir, "Generalized concavity and duality", in: S. Schaible, W.T. Ziemba (Eds.), *Generalized concavity in Optimization and Economics*, Academic Press, New York, **1981,** pp. 263-279.
5. O. L. Mangasarian, "Second- and higher-order duality in non linear programming", *J. Math. Anal. Appl.*, **1975,** *51*, 607-620.
6. I. Ahmad and Z. Husain, "On nondifferentiable second order symmetric duality in mathematical Programming", *Indian J. Pure Appl. Math.*, **2004,** *35*, 665-676.
7. X. M. Yang, X. Q. Yang and K. L. Teo, "Non-differentiable second order symmteric duality in mathematical programming with *F*-convexity", *Euro. J. Oper. Res.,* **2003,** *144*, 554-559.
8. X. M. Yang, X. Q. Yang, K. L. Teo and S. H. Hou, "Multiobjective second-order symmetric duality with F -convexity", *Euro. J. Oper. Res.*, **2005,** *165*, 585-591.
9. I. Ahmad and Z. Husain, "On multiobjective second order symmetric duality with cone constraints", *Euro. J. Oper. Res.*, **2010,** *204*, 402-409.
10. S. K. Suneja, C. S. Lalitha and S. Khurana, "Second order symmetric duality in multiobjective programming", *Euro. J. Oper. Res.*, **2003,** *144*, 492-500.

11. I. Ahmad and Z. Husain, "Multiobjective mixed symmetric duality involving cones", *Computers Math. Appl.,* **2010,** *59*, 319-326.
12. S. Khurana, "Symmetric duality in multiobjective programming involving generalized cone-invex functions", *Euro. J. Oper. Res.*, **2005,** *165*, 592-597.
13. S. K. Mishra and K. K. Lai, "Second order symmetric duality in multiobjective programming involving generalized cone-invex functions", *Euro. J. Oper. Res.*, **2007,** *178*, 20-26.
14. S. K. Padhan and C. Nahak, "Higher-order symmetric duality in multiobjective programming problems under higher-order invexity", *Appl. Math. Comput*., **2011**, *218*, 1705-1712.
15. C. R. Bector, S. Chandra and Abha, "On mixed symmetric duality in multiobjective programming", *Opsearch*, **1999,** *36*, 399-407.
16. X. M. Yang, K. L. Teo and X. Q. Yang, "Mixed symmetric duality in nondifferentiable mathematical programming", *Indian J. Pure Appl. Math.*, **2003**, *34*, 805-815.
17. I. Ahmad, "Multiobjective mixed symmetric duality with invexity", *New Zealand J. Math.,* **2005,** *34*, 1-9.
18. J. Li and Y. Gao, "Non-differentiable multiobjective mixed symmetric duality under generalized convexity", *J. Inequal. Appl.* **2011**, *23*, doi: 10.1186/1029-242X-2011-23.
19. S. K. Gupta and N. Kailey, "Multiobjective second-order mixed symmetric duality with a square root term", *Appl. Math. Comput.*, **2012**, *218*, 7602–7613.
20. S. K. Suneja, S. Aggarwal and S. Davar, "Multiobjective symmetric duality involving cones", *Euro. J. Oper. Res.,* **2002,** *141*, 471-479.
21. M. Schechter, "More on subgradient duality", *J. Math. Anal. Appl.,* **1979,** *71*, 251-262.
22. T. R. Gulati, I. Ahmad and I. Husain, "Second-order symmteric duality with generalized convexity", *Opsearch*, **2001,** *38*, 210-222.
23. S. Chandra, I. Husain and Abha, "On mixed symmetric duality in mathematical programming", *Opsearch*, **1999,** *36*, 165-171.
24. X. M. Yang, "Second order symmetric duality for nonlinear programs", *Opsearch*, **1995,** *32*, 205-209.
25. S. Chandra, A. Goyal and I. Husain, "On symmetric duality in mathematical programming with F -convexity", *Optim.,* **1998,** *43*, 1-18.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 372-378

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

An alternative synthesis of (±)-propranolol and (±)-atenolol

Rachaneebhorn Inkum, Aphiwat Teerawutgulrag*, Pakawan Puangsombat and Nuansri Rakariyatham

Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200 Thailand

* Corresponding author, e-mail: aphiwattee@yahoo.co.uk

Received: 14 February 2012 / Accepted: 12 October 2012 / Published: 15 October 2012

Abstract: Herein, a simple synthesis pathway of beta-blockers starting from allyl amine is presented. This synthesis features the opening of an epoxide ring with phenol derivatives followed by *N*-alkylation with iso-propylbromide to produce racemic propranolol and atenolol.

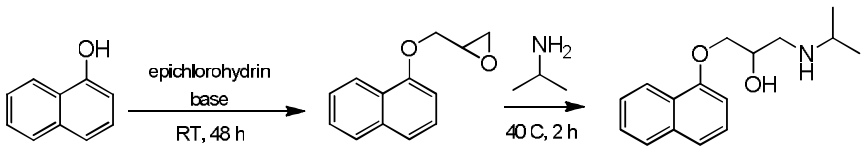
Keywords: beta-blockers, propranolol, atenolol

INTRODUCTION

Propranolol and atenolol are prescribed medicines belonging to a class of compounds known as beta-blockers, which are used to treat hypertension, angina pectoris, glaucoma, anxiety, obesity and other cardiovascular diseases [1,2]. Nowadays, these drugs are available in the market in racemic form, in which only the *S*-enantiomer possesses beta-adrenergic blocking activity [3-6] , while the *R*-form merely has a membrane stabilising effect and is 130 times less active than the *S*-analogue [3].

While various methods have been published for the purpose of synthesising racemic propranolol [7-11], the disadvantages are employing harsh conditions, multiple steps synthesis or complicated catalyst preparation, while the atenolol synthesis pathway requires a high temperature [12,13]. Several methods have been reported on the synthesis of *(S)*-propranolol and *(S)*-atenolol including the use of enzymes for resolution [14], asymmetric hydrogenation with chiral metal complex catalysts [15], asymmetric epoxidation of allyl alcohol [16] and sorbitol [17], employing a polymer supported reagent [9], as well as using $Zn(NO_3)_2$ and (+)-tartaric acid induction in the ring opening step [18]. Several researchers have reported on the synthesis of *(S)*-propranolol *via* lipase catalysed reaction [19-22] and in the presence of cyclodextrins [23]. However, the multiple steps in each procedure and the high cost of starting materials have increased the expense of manufacturing.

In pharmaceutical manufacturing, racemic propranolol is synthesised using epichlorohydrin (scheme 1) [14b]. A straightforward method for the preparation of racemic propranolol and atenolol as an inexpensive procedure is reported.



Scheme 1. Industrial synthesis of racemic propranolol

MATERIALS AND METHODS

Molecular sieves (4A, Fluka) were activated by heating in an oven at 120°C for 12 hr. α-Naphthol (May&Baker Ltd Dagenham England), was purified by sublimation. *p*-Hydroxyphenyl acetamide was purchased from Aldrich. Allyl amine and isopropyl bromide were distilled and stored in a desiccator. All solvents were distilled prior to use. Thin-layer chromatography (TLC) was utilised on silica plates 60 F_{264} and column chromatography was carried out on 0.063-0.20 mm silica gel.

Melting points were determined on MEL-TEMP Laboratory Devices INC., USA and have not been corrected. HRMS analysis was performed on ESI-Q-TOF-MS (Micromass, Manchester, UK). IR spectra were reported on a FT-IR spectrometer (Tensor 27). ^{1}H-NMR spectra were recorded on a Bruker Avance 400 MHz NMR spectrometer, using trimethylsilane as an internal standard.

t-Butyl Allyl Carbamate (2)

To a stirred solution of allyl amine **1** (3.00 g, 52.5 mmol) and triethylamine (9.5 mL, 68.0 mmol) in dichloromethane (10 mL) at 0°C, Boc_2O (15.6 mL, 68.0 mmol) was added. The mixture was warmed to room temperature and stirred overnight. The reaction was washed with 10% NaOH and then extracted with ethyl acetate (3x50 mL). The organic layer was dried with anhydrous sodium sulphate then concentrated under reduced pressure. The residue was purified by column chromatography with neat hexane as an eluent to give *t*-butyl allyl carbamate **2** (8.1 g, 98%) as a colorless solid. mp. 35-36° C. IR (neat), ν_{max} 3347, 2979, 1690, 1244; ^{1}H-NMR (400 MHz , $CDCl_3$) 1.43 (9H, s), 3.66-3.78 (2H, m), 4.54-4.72 (1H, br), 5.05-5.20 (2H, m), 5.76-5.88 (1H, m); ^{13}C-NMR (100 MHz, $CDCl_3$) 155.8, 134.9, 115.6, 79.3, 43.0, 28.4; HRMS calcd for $C_8H_{15}NO_2Na$ $[M+ Na]^+$ 180.1000, found 180.0997.

t-Butyl Oxiran-2-ylmethylcarbamate (3)

t-Butyl allyl carbamate **2** (3.00 g, 19.0 mmol) was dissolved in dichloromethane (10 mL) and *m*-chloroperbenzoic acid (6.59 g, 38.0 mmol) was added at 0°C and the reaction mixture was warmed to room temperature. After heating at reflux for 5 hr, the mixture was cooled and saturated sodium thiosulphate was added dropwise. The mixture was washed with saturated sodium bicarbonate. The aqueous layer was extracted with ethyl acetate (3x50 mL). The combined organic phase was dried, concentrated under reduced pressure and purified by column chromatography to give *t*-butyl oxiran-2-ylmethylcarbamate **3** (3.3 g, 88%) as a pale yellow liquid. IR (neat) , ν_{max} 3352, 2980, 1700, 1249; ^{1}H-NMR (400 MHz ,$CDCl_3$) 1.42 (9H, s), 2.56-2.59 (1H, m), 2.75-2.78 (1H, m), 3.04-3.10 (1H, m),

3.15-3.23 (1H, m), 3.45-3.57 (1H, m), 4.70-4.85 (1H, br); ^{13}C-NMR (100 MHz, $CDCl_3$) 155.9, 79.6, 50.7, 44.9, 41.8, 28.3; HRMS calcd for $C_8H_{16}NO_3$ $[M+H]^+$ 174.1130, found 174.1132.

_t_-Butyl 2-Hydroxy-3-(naphthalen-1-yloxy)propylcarbamate (4a)

α-Naphthol (0.28 g, 1.98 mmol) was added to the solution of potassium hydroxide (0.11g, 1.98 mmol) in water (3 mL) and epoxide 3 (0.28 g, 1.65 mmol) in tetrahydrofuran at 0°C. The reaction mixture was stirred at room temperature overnight. After neutralisation with 2 N HCl at 0°C, the resulting mixture was extracted with ethyl acetate (3x20 mL). The obtained layer was evaporated to dryness to give a crude product. The resulting crude product was purified by column chromatography eluted with a gradient of ethyl acetate-hexane (3% -30%) to give *t*-butyl 2-hydroxy-3-(naphthalen-1-yloxy) propylcarbamate **4a** (0.40 g, 77%) as a pale brown solid. mp. 80-82°C. IR (KBr), ν_{max} 3367, 3055, 2978, 1694, 1270, 1170; ^{1}H-NMR (400 MHz, $CDCl_3$) 1.46 (9H, s), 3.30-3.48 (1H, m), 3.48-3.63 (2H, m), 4.12-4.17 (2H, m), 4.25-4.35 (1H, br), 5.00-5.10 (1H, br), 6.82 (1H, d, *J*= 7.2 Hz), 7.37 (1H, t, *J*= 7.2 Hz), 7.44-7.52 (3H, m), 7.81(1H, d, *J*= 8.0 Hz), 8.18 (1H, d, *J*= 8.0 Hz); ^{13}C-NMR (100 MHz $CDCl_3$) 157.3, 154.1, 134.5, 127.6, 126.5, 125.8, 125.4, 121.7, 120.8, 105.0, 80.0, 70.0, 69.7, 43.9, 28.4; HRMS calcd for $C_{18}H_{23}NO_4Na$ $[M+Na]^+$ 340.1525, found 340.1519.

_t_-Butyl 3-(4-(2-Amino-2-oxoethyl)phenoxy)-2-hydroxypropylcarbamate (4b)

p-Hydroxyphenyl acetamide (0.21 g, 1.38 mmol) was added to the solution of potassium hydroxide (0.08 g, 1.38 mmol) in water (2.2 mL) and epoxide **3** (0.20 g, 1.16 mmol) at 0°C. The reaction mixture was stirred at room temperature overnight. After neutralisation with 2N HCl at 0°C, the resulting mixture was extracted with ethyl acetate (3x20 mL). The obtained layer was evaporated to dryness to give a crude product. The resulting crude product was purified by column chromatography eluted with a gradient of ethyl acetate-hexane (10%-50%) to give *t*-butyl 3-(4-(2-amino-2-oxoethyl)phenoxy)-2-hydroxypropylcarbamate **4b** (0.37 g, 84 %) as a white solid. mp. 128-130°C. IR(KBr), ν_{max} 3361, 3179, 2978, 1694, 1246; ^{1}H-NMR (400 MHz , DMSO-d_6) 1.37 (9H, s), 2.94-3.14 (2H, m), 3.27 (2H, s), 3.74-3.90 (3H, m), 5.08-5.14 (1H, br), 6.82 (2H, d, *J*= 8.3 Hz), 7.15 (2H, d, *J*= 8.3 Hz), 7.32-7.40 (1H, br); ^{13}C-NMR (100 MHz, DMSO-d_6) 172.6, 157.3, 155.8, 129.9, 128.5, 114.2, 77.7, 70.4, 68.2, 43.5, 41.4, 28.2; HRMS calcd for $C_{16}H_{24}N_2O_5Na$ $[M+Na]^+$ 347.1583, found 347.1581.

1-Amino-3-(naphthalen-1-yloxy)propan-2-ol (5a)

Trifluoroacetic acid (2 mL) was added to a solution of *t*-butyl 2-hydroxy-3-(naphthalen-1-yloxy) propylcarbamate **5** (0.40 g, 1.28 mmol) in dichloromethane (2 mL) at 0°C. After the solution was stirred for 2 hr, the mixture was concentrated under reduced pressure. Then the resulting crude mixture was basified with 10% NaOH, and was extracted with ethyl acetate (3x20 mL). The organic layer was dried with anhydrous sodium sulphate and was concentrated in vacuo to give 1-amino-3-(naphthalen-1-yloxy) propan-2-ol **5a** (0.32 g, quantitative yield) as a pale yellow solid. mp. 78-80°C. IR (KBr), ν_{max} 3363, 3275, 1581, 1265; ^{1}H-NMR (400 MHz, DMSO-d_6) 2.68-2.76 (1H, m), 2.80-2.88 (1H, m), 3.87-3.94 (1H, m), 4.02-4.14 (2H, m), 6.95 (1H, d, J= 7.3 Hz), 7.37-7.55 (4H, m), 7.85 (1H, d, *J*= 7.3 Hz), 8.22 (1H, m); ^{13}C-NMR (100 MHz, DMSO-d_6) 154.7, 134.5, 127.8, 126.9, 126.7, 125.6, 122.3, 120.3,106.0, 70.9, 70.8, 45.2; HRMS calcd for $C_{13}H_{16}NO_2$ $[M+H]^+$ 218.1180, found 218.1181.

2-(4-(3-Amino-2-hydroxypropoxy)phenyl) Acetamide (5b)

HCl gas was bubbled through a solution of *t*-butyl 3-(4-(2-amino-2-oxoethyl)phenoxy)-2-hydroxypropylcarbamate **4b** (0.40 g, 1.23 mmol) in methanol (2 mL) at 0°C. After the reaction was completed, the mixture was concentrated under reduced pressure. Then, the product was basicified with triethylamine and the solution was concentrated in vacuo. The resulting crude mixture was purified by column chromatography with a gradient of dichlorimethane-methanol (10-30%) to give 2-(4-(3-amino-2-hydroxypropoxy)phenyl) acetamide **5b** (0.29 g, quantitative yield) as a pale yellow solid. mp. 218-220 C. IR (KBr), ν_{max} 3356, 3176, 1639, 1247; ^{1}H-NMR (400 MHz, DMSO-d_6) 2.76-2.84 (1H, dd, J=12.9 Hz), 2.99 (1H, dd, J= 12.9 Hz), 3.29 (2H, s), 3.90-3.94 (2H, m), 3.98-4.06 (1H, m), 6.80-6.84 (1H, br), 6.87 (2H, d, *J*= 8.6 Hz), 7.17 (2H, d, *J*= 8.6 Hz) 7.40-7.46 (1H, br); ^{13}C-NMR (100 MHz, DMSO-d_6) 172.7, 156.9, 130.1, 128.8, 114.3, 69.6, 65.9, 41.9, 41.3; HRMS calcd for $C_{11}H_{17}N_2O_3$ $[M+H]^+$ 225.1239, found 225.1240.

Propranolol (6a)

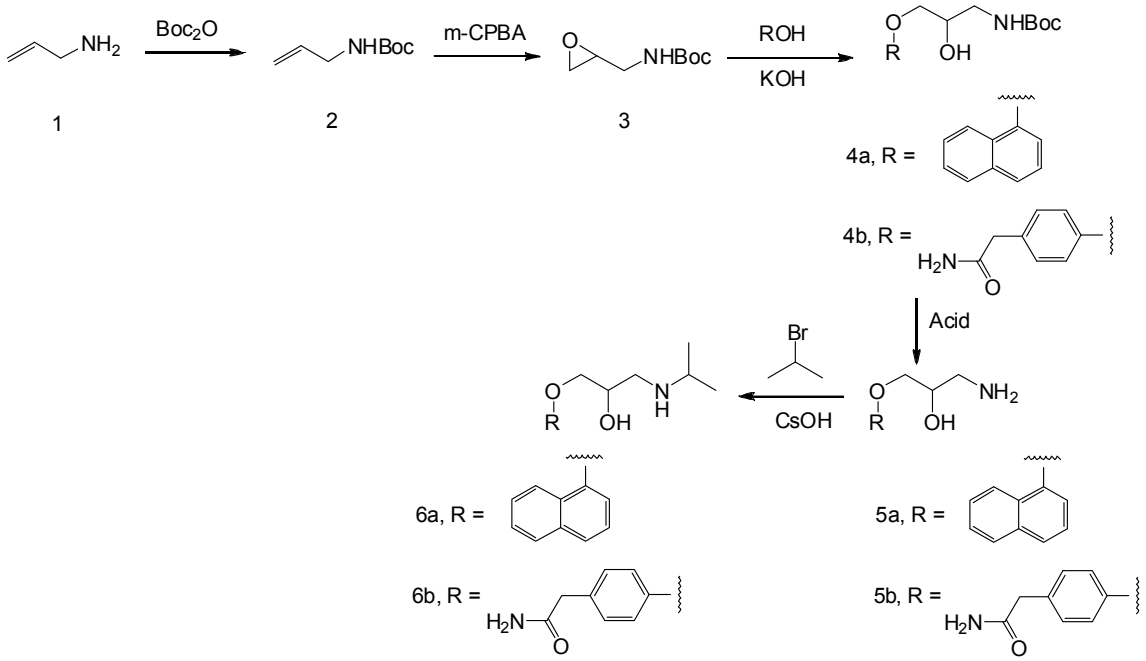
A suspension of 1-amino-3-(naphthalen-1-yloxy) propan-2-ol **5a** (0.10 g, 0.46 mmol), molecular sieves 4 (0.135 g) and cesium hydroxide (0.15 g, 0.92 mmol) in dimethylformamide (2.24 mL) was stirred at room temperature for 30 min. Then isopropyl bromide (0.43 mL, 4.60 mmol) was added dropwise. After the reaction was completed, 1 N HCl was added and the mixture was extracted with ethyl acetate (3x5 mL). The organic layer was dried with anhydrous sodium sulfate and concentrated under reduced pressure. The crude mixture was purified by column chromatography and eluted with a gradient of dichlorimethane-methanol (0-3%) to give propranolol **6a** (0.0838 g, 70%) as a white solid. mp. 90-92 C (Lit [8] mp. 93-94 C). IR (KBr), ν_{max} 3477, 3275, 2964, 1628, 1265; ^{1}H-NMR (400 MHz, $CDCl_3$) 1.14 (6H, d, *J*= 6.3 Hz), 2.85-2.96 (2H, m), 2.89-2.95 (1H, m), 4.09-4.15 (1H, m), 4.16-4.28 (2H, m), 6.75 (1H, d, *J*= 7.6 Hz), 7.35 (1H, t, *J*= 7.6 Hz), 7.42-7.52 (3H, m), 7.78-7.82 (1H, m), 8.15-8.19 (1H, m); ^{13}C-NMR (100 MHz ,$CDCl_3$) 154.4, 134.5, 127.5, 126.4, 125.8, 125.2, 121.8, 120.6, 104.9, 70.7, 68.6, 49.5, 48.9, 23.2, 23.0; HRMS calcd for $C_{16}H_{22}NO_2$ $[M+H]^+$ 260.1650, found 260.1651.

Atenolol (6b)

A solution of 2-(4-(3-amino-2-hydroxypropoxy) phenyl) acetamide **5b** (0.20 g, 0.89 mmol), molecular sieves 4 (0.135 g) and cesium hydroxide (0.18 g, 1.07 mmol) in dimethylformamide (4.48 mL) was stirred at room temperature for 30 min. Then isopropyl bromide (0.84 mL, 8.92 mmol) was added dropwise. After completion, the mixture was filtered and concentrated under reduced pressure. The crude mixture was purified by column chromatography with a gradient of dichloromethane-methanol (10-30%) to give atenolol **6b** (0.13 g, 60%) as a white solid. mp. 148-150 C (Lit [24] 148-150 C). IR (KBr), ν_{max} 3360, 2916, 1635, 1242; ^{1}H-NMR (400 MHz , DMSO-d_6) 0.98 (6H, d, *J*= 6.3 Hz), 2.54-2.58 (1H, m), 2.64-2.74 (2H, m), 3.29 (2H, s), 3.80-3.86 (2H, m), 3.88-3.94 (1H, m), 6.84 (2H, d, *J*= 8.8 Hz), 7.14 (2H, d, *J*= 8.8 Hz), 7.38-7.42 (1H, br); ^{13}C-NMR (100 MHz, DMSO-d_6) 172.6, 156.8, 130.1, 128.9, 114.3, 69.8, 65.4, 49.8, 46.8, 41.3, 19.1, 18.5; HRMS calcd for $C_{14}H_{23}N_2O_3$ $[M+1]^+$ 267.1709, found 267.1707.

RESULTS AND DISCUSSION

Treatment of allyl amine **1** with *t*-butyl dicarbonate (Boc_2O) in the presence of triethylamine at room temperature gave *t*-butyl allylcarbamate **2** in 98% yield. After epoxidation of **2** with *m*-chloroperbenzoic acid, *t*-butyl oxiran-2-ylmethylcarbamate **3** was attained with 88% yield. The epoxide ring opening of **3** with phenol derivatives and an aqueous solution of potassium hydroxide provided **4a** and **4b** in yields of 77% and 84% respectively. Deprotection of the Boc group in **4a** and **4b** was performed under acidic conditions (trifluoroacetic acid for **4a** and HCl gas for **4b**) in quantitative yields. The desired amines were reacted with iso-propylbromide and cesium hydroxide in dimethylformamide yielding propranolol **6a** (70%) and atenolol **6b** (60%) (Scheme 2).



Scheme 2. A synthetic route to beta-blockers

CONCLUSIONS

An alternative route for the synthesis of racemic propranolol and atenolol has been developed compared to previous reports. This pathway is simple, inexpensive and mild conditions could be used.

ACKNOWLEDGEMENTS

Financial assistance for this work was provided by Rajamangala University of Technology Lanna Nan. We thank The Graduate School and Department of Chemistry, Faculty of Science, Chiang Mai University for chemicals and use of their facilities.

REFERENCES

1. T. Seki, T. Takezaki, R. Ohuchi, H. Ohuyabu, T. Ishimori and K. Yasuda, "Studies on agents with vasodilator and β-blocking activities I", *Chem. Pharm. Bull.*, **1994**, *42*, 1609-1616.

2. T. Fujioka, S. Teramoto and T. Mori, "Novel positive inotropic agents: synthesis and biological activities of 6-(3-amino-2-hydroxypropoxy)-2(1H)-quinolinone derivatives", *J. Med. Chem.,* **1992**, *35*, 3607-3612.
3. A. M. Barrett and V. A. Cullum, "The biological properties of the optical isomers of propranolol and their effects on cardiac arrhythmias", *Br. J. Pharmacol.,* **1968**, *34*, 43-55.
4. R. Howe and R. G. Shanks, "Optical isomers of propranolol", *Nature*, **1966**, *210*, 1336-1338.
5. K. H. Rahn, A. Hawlina, F. Kersting and G. Planz, "Studies on the antihypertensive action of the optical isomers of propranolol in man", *Naunyn Schmiedebergs Arch. Pharmacol.*, **1974**, *286*, 319-323.
6. T. Walle, U. K. Walle, M. J. Wilson, T. C. Fagan and T. E. Gaffney, "Stereoselective ring oxidation of propranolol in man", *Br. J. Chin. Pharmacol.,* **1984**, *18*, 741-748.
7. A. F. Crowther and L. H. Smith, "β-Adrenergic blocking agents II. Propranolol and related 3-amino-1-naphthoxy-2-propanols", *J. Med. Chem*., **1968**, *11*, 1009-1013.
8. C. Giuliana, O. Mario and S. Sergio, "Oxazolidin-2-ones from allylic amine by means of iodine and carbonate anion on polymeric support. A convenient synthesis of (±)-propranolol", *J. Org. Chem.,* **1986**, *51*, 713-717.
9. V. D. Subhash, N. P. Prashant and M. S. Manikrao, "One pot synthesis of (±)/*S*-atenolol and (±)/*S*-propranolol by employing polymer supported reagent", *Synth. Commun.,* **1999**, *29*, 1639-1644.
10. J. Slobodanka, V. S. Milanka, U. K. Dragana, P. Dragica, S. Slobokanda and D. Dijamanda, "Catalysis in the alkylation reaction of 1-naphthol with epichlorohydrin", *J. Serb. Chem. Soc.,* **2006**, *71*, 867-877.
11. I. K. Rukhsana, A. Irshad, P. Kavita, N. H. Khan, H. R. A. Sayed, H. C. Bajai and S. Eringathodi, "Solvent-free microwave synthesis of aryloxypropanolamine by ring opening of aryloxy epoxides", *Research Letters in Organic Chemistry*, **2009**, *10*, 1-5.
12. (a) A. M. Barrett, J. Carter, R. Hull, D. J. Le Count and C. J. Squire, US patent US 3,836,671, 17 Sept **1974**. (b) A. M. Barrett, J. Carter, R. Hull, D. J. Le Count and C. J. Squire, US patent US 3,934,032, 20 Jan **1976**.
13. S. M. Jang and T. S. Shich, US patent US 5,290,958, 1 March **1994**.
14. (a) M. Noritada and O. Nobuo, "Preparation of optically active 1-acetoxy-2-aryloxyproprionitrile and its application to a facile synthesis *(S)*-(-)-propranolol", *Tett. Lett*., **1985**, *26*, 5533-5534.
    (b) H. S. Bevinakatti and A. A. Banerji, "Practical chemoenzymatic synthesis of both enantiomers of propranolol", *J. Org. Chem.,* **1991**, *56*, 5372-5375.
    (c) H. S. Bevinakatti and A. A. Banerji, "Lipase catalysis in organic solvents. Application to the synthesis of (*R*)- and *(S)*-atenolol", *J. Org. Chem.,* **1992**, *57*, 6003-6005.
15. H. Takahashi, S. Sakuraba, H. Takeda and K. Achiwa, "Asymmetric reaction catalyzed by chiral metal complexes. 41. Highly efficient asymmetric hydrogenation of amino ketone derivatives leading to practical syntheses of *(S)*-propranolol and related compounds", *J. Am. Chem. Soc.,* **1990**, *112*, 5876-5878.
16. J. M. Klunder, S. Y. Ko and K. B. Sharpless, "Asymmetric epoxidation of allyl alcohol: Efficient routes to homochiral *β*-adrenergic blocking agents", *J. Org. Chem.,* **1986**, *51*, 3710-3712.

17. R. A. Veloo and G. J. Koomen, "Synthesis of enantiomerically pure *(S)*-(-)-propranolol from sorbitol", *Tett. Asym.,* **1993**, *4*, 2401-2404.
18. H. Eshghi and H. P. Yazdi, "A facile synthesis of *(S)*-(-)-propranolol", *J. Sci. Islamic Republic of Iran.,* **2003**, *14*, 17-19.
19. B. Silber and S. Riegelman, "Stereospecific assay for (-)- and (+)-propranolol in human and dog plasma", *J. Pharmacol. Exp. Ther.,* **1980**, *215*, 643-648.
20. J. Hermansson and C. Von Bahr, "Simultaneous determination of *d*- and *l*-propranolol in human plasma by high-performance liquid chromatography", *J. Chromatogr.,* **1980**, *221*, 109-117.
21. Y. Terao, M. Murata, K. Achiwa, T. Nishino, M. Akamatsu and M. Kamimura, "Highly efficient lipase-catalyzed asymmetric synthesis of chiral glycerol derivatives leading to practical synthesis of *S*-propranolol", *Tett. Lett.,* **1988**, *29*, 5173-5176.
22. Y. F. Wang, S. T. Chen, K. C. Liu and C. H. Wong, "Lipase-catalyzed irreversible transesterification using enol ester: resolution of cyanohydrins and syntheses of ethyl (*R*)-2-hydroxy-4-phenylbutyrate and *(S)*-propranolol", *Tett. Lett*., **1989**, *30*, 1917-1920.
23. A. G. Rizette, P. G. Manuela and G. C. Francisco, "Lipase-catalyzed preparation of *S*-propranolol in presence of hydroxypropyl *β*-cyclodextrins", *J. Biosci. Bioeng.,* **2005**, *100*, 423-428.
24. D. S. Bose and A. V. Narsaiah, "An efficient asymmetric synthesis of *(S)*-atenolol: using hydrolytic kinetic resolution", *Bioorg. Med. Chem.,* **2005**, *13*, 627-630.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 379-396

Maejo International

Journal of Science and Technology

ISSN 1905-7873

Available online at www.mijst.mju.ac.th

Full Paper

Refining the definition of sustainable agriculture: An inclusive perspective from Malaysian vegetable sector

Yeong-Sheng Tey [1, 2, *], Elton Li [1], Johan Bruwer [1], Amin M. Abdullah [3], Jay Cummins [4], Alias Radam [5], Mohd M. Ismail [2, 3] and Suryani Darham [2]

[1] School of Agriculture, Food and Wine, the University of Adelaide, Australia
[2] Institute of Agricultural and Food Policy Studies, Universiti Putra Malaysia, Malaysia
[3] Faculty of Agriculture, Universiti Putra Malaysia, Malaysia
[4] Department of Primary Industries Victoria, Australia
[5] Faculty of Economics and Management, Universiti Putra Malaysia, Malaysia

* Corresponding author, e-mail: tyeong.sheng@gmail.com

Received: 11 January 2012 / Accepted: 14 October 2012 / Published: 15 October 2012

Abstract: Skepticism about the longevity of conventional agriculture has resulted in the quest for sustainable agriculture. Like many developing countries, a homogenous definition of the term 'sustainable agriculture' is yet to be developed in Malaysia. To fill this gap, using an inclusive perspective, this study posits a refined definition of 'sustainable agriculture' for Malaysia. Cognizant of relevant past studies, which were built on rather narrow viewpoints, this study integrates qualitative insights from selected up-stream stakeholders in Malaysian vegetable sector. The structured results suggest that sustainable agriculture can be defined as the process by which an integrative balanced agricultural system is realised through a dynamic set of practices that are (1) environmentally enhancing, (2) resource optimal, (3) economically viable, (4) socially justifiable and (5) functionally feasible over time. Though derived from Malaysia, this definition can be adapted to fit local nuances in other countries and sectoral emphases in agriculture. With these five inter-related operational attributes, which are capable of enhancement and/or modification periodically, this flexible definition can progressively provide potential direction towards academic understanding and development of agricultural sustainability.

Keywords: sustainable agriculture, Malaysian vegetable sector, inclusive perspective, sustainability concept

INTRODUCTION

Improving sustainability is of paramount importance to both developed and developing countries. The term 'sustainable development' was originally defined by the Brundtland Commission as 'the development that meets the needs of the present without compromising the ability of future generations to meet their own needs' [1]. Though the initial purpose is to link poverty alleviation to environmental and natural resource management, the definition has culminated in a consensus of the need for economic growth without degrading natural endowments. This consensus has attracted the highest socio-political support at the United Nations Earth Summit in Rio de Janeiro in 1992 [2]. It has been further endorsed by the United Nations World Sustainable Development (Millennium) Summit in Johannesburg 2002 [3].

The diffusion of sustainability concepts has created a new paradigm in agricultural development. This new paradigm seriously questions the contribution of conventional agricultural practices to industry sustainability [4]. Among others, chemical inputs are widely criticised for their undesirable impacts on soil, water, biodiversity, human health, food safety and economy [5-7]. While these findings are not new, Carson [8] in her classical "Silent Spring", presented irrefutable arguments to posit that chemical pesticides cause multiple destructions in exchange for a mono-beneficial crop protection. In the long run, such externalities can potentially strike at the heart of food security and poverty.

Skepticism about the longevity of conventional agriculture has resulted in the perceived need for sustainable agriculture, which was internationally addressed in the Agenda 21 [9], the Millennium Ecosystem Assessment [10] and the International Assessment of Agricultural Science and Technology for Development [11].

Many studies [12-16] have attempted to define the term 'sustainable agriculture'. However, there is neither a definitive nor a standardised definition of 'sustainable agriculture' [17-21]. This is mainly because definitions invariably differ in individual contexts. The term 'sustainable agriculture' should therefore be refined, at least, in the local context.

Like many developing countries, 'sustainable agriculture' has not been officially defined by the Malaysian government for a specific agricultural sector. Similarly, past studies on local agricultural sustainability issues [22-25] have not attempted to define 'sustainable agriculture' in Malaysia. The definition is crucially needed in responding to the pressing issue of agricultural sustainability while, at the same time, guiding the development of the industry and its sub sectors [26-27].

To fill the aforementioned gap, this study is intended to refine the definition of 'sustainable agriculture' through the use of an inclusive perspective from farm-level. The context of Malaysian vegetable sector forms the basis of the work.

Responding to Pimbert's [28] call for inclusive research, this study departs from past studies, which have arisen from a basis of rather narrow standpoints. Ultimately, the realisation of sustainable agricultural systems for food production requires the participation and understanding of all up-stream contributors [29]. To serve this purpose, insights from selected up-stream stakeholders

of the vegetable sector in Malaysia were integrated and operational attributes that can potentially serve as directional guidelines for future actions were also posited.

BACKGROUND

Early agricultural policies in Malaysia were economically orientated. The First National Agricultural Policy (1984-1991) and the Second National Agricultural Policy (1992-1997) emphasised the maximisation of farm income via efficient utilisation of local resources [26].

With emerging concerns centred on sustainability, the Third National Agricultural Policy (1998-2010) took a slightly different approach. The difference, however, arises as a consequence of a rather shallow vision of sustainable development [27]. Though the vision is to ensure that the present needs are not met at the expense of future generations [25], the overarching emphasis is still on income maximisation through responding to market information and optimal resource utilisation.

Under the Third National Agricultural Policy, the vegetable sector has undergone the holistic promotion of two prominent programmes of relevance to sustainable development. Firstly, the "Malaysia's Organic Scheme" was launched in 2001. According to the Department of Agriculture [30], the scheme asserts that organic production methods constitute one of the best means for the production of safe, quality foods. Organic production does not use chemical inputs and attempts to avoid environmental degradation. Secondly, the "Malaysia's Good Agricultural Practices" (GAPs) Scheme was introduced in 2002. This scheme focuses on integrated systems which aim to manage all farm resources in a sustainable format [31]. It is intended to increase farm productivity as well as produce safe, quality foods. At the same time, the scheme also seeks to assist the welfare, safety and health of farm workers by preserving a safer and a more natural environment.

These programmes have, so far, had only limited success. Up to the end of 2010, the number of adopters was less than 30 vegetable farmers for the Malaysia's Organic Scheme and 100 vegetable farmers for Malaysia's GAPs Scheme [32]. The number of adopters represents less than one percent of the 46,040 vegetable farmers in Malaysia [33]. Nevertheless, a starting point for sustainable agriculture has been made. This is supported by Tey et al.'s study [34] that recorded adoption rates of cover crops/mulches (35%-45%), organic fertilizers/composts (35%-45%), intercropping (35%-45%), crop rotation (30%-40%), conservation tillage (25%-35%), and integrated pest management (25%-35%).

One strategy, which conceivably would result in more successful implementation, is agreement on the definition of 'sustainable agriculture' at farm level. Without such agreement, relevant agricultural programmes end up promoting previous standard practices under a new name [13]. Any definition should include some notions of operational attributes for sustainable agriculture. This is not an easy task but we can progress our understanding by extracting and integrating multiple insights from industry contributors while, at the same time, being guided by the compass of past studies.

LITERATURE REVIEW

There are a number of definitions from various philosophical and technical attempts in defining the term 'sustainable agriculture'. Despite these varying definitions, past studies were mostly built on the three common components of sustainable development: the environmental, the economic and social aspects.

Environmental Aspect

The concept of environmental sustainability is traditionally important in the definition of sustainable agriculture [5, 35-37]. Typically, this component addresses the conservation and management of the environment and natural resources for present and future generations [38]. Natural resources include, among others, soil, water, ecosystems and biodiversity. Conservation seeks to ensure that agricultural practices do not degrade the environment and natural resources [14]. By reducing reliance on off-farm inputs, such management promotes more efficient use of renewable and non-renewable resources through the utilisation of on-farm integrative resource systems [13].

In addition, there is an expansionary component. Beyond management and conservation, it stresses the enhancement of environmental quality and the resource base on which agricultural activities depend. This expanded component is officially adopted by the United States [39] and Canada [40], but remains largely unacknowledged in other countries. McIsaac [41] is a notable exception.

It is important to acknowledge that the expanded component is built upon the presumption of the conservation and management of the environment and natural resources [37]. However, the maintenance of natural endowments in the current form is less than optimal, if they have already been degraded. Rodale [42] asserts that sustainable agriculture should improve the status quo of soils (among many other resources). Enhancement is one of the most plausible strategies to improve the sustainability level of the environmental component.

Economic Aspect

Economic sustainability is another important component in the definition of 'sustainable agriculture' [13-14, 41, 43]. The component refers to the economic viability of carrying out sustainable agriculture over time [38]. It considers the maintenance of economy-wide factors, such as costs-benefits, inputs-outputs, and investment-returns. Not only does it deal with monetary concerns, it also weighs on farm operation and employment factors [17]. The notion of economic viability is formally accepted by the United States [39] and Canada [40]. It is also echoed by some past studies [44-47].

Because a farm enterprise cannot become sustainable overnight, the transitional stage does not offer immediate economic returns to farmers. Sustainable agriculture is, hence, impeded by economic hurdles since it offers less attractive marketing returns and diminishes economic benefit during the transitional stage.

Against the aforementioned concern, Lehman et al. [12] provide a useful perspective which can be borrowed. In their review of Canada's [40] definition of 'sustainable agriculture', they question whether ensuring economic viability is really an essential component in sustainable

agriculture. For simplicity, they translate economic viability to farm profitability over time. They point out that farmers could engage in sustainable agriculture even though the recommended practices might not directly result in a profit. Among many initiatives, subsidies, tax reduction, and cuts in interest rates could indirectly improve farm profitability.

It can be readily conceived that the economic viability of sustainable agriculture is possible in one or many ways, both in the short- and long-run; the "economic" component in sustainable agriculture should not be neglected.

Social Aspect

"Social sustainability" is an imperative component in the definition of 'sustainable agriculture' [44-47]. Fundamentally, the component stresses social acceptable attainment and the continued satisfaction of present and future human need [38]. Among basic needs, food and fiber are commonly highlighted in connection with the crucial role of ensuring food security in society [48]. That means sustainable agriculture should endeavor to maintain or improve farm yield for the stability of food security.

Food and fiber must also be produced in safe ways. This concept is inclusive, stretching from production to consumption, i.e. to assure the health and safety of agricultural producers at farm level and ensure food safety for consumers at market level [13]. The United States [39] agreed that health, safety and food safety will enhance the quality of life and society as a whole.

The importance of social component is generally overlooked. As an example, the need to ensure the health and safety of farm workers is not included by the Science Council of Canada [40] and a number of past studies [5, 36]. Karami & Keshavarz [4] note that the social component has been relatively neglected by various studies [46, 49-50]. This situation arises because no operative definition of the social component has been developed [51].

METHODS

Three core components are indicated in the definition of sustainable agriculture. These components are environment, economic, and social. They can be used as a guiding framework to refine the definition of 'sustainable agriculture'. A similar framework has also been used by Norman et al. [15] for defining sustainable agriculture in the United States.

Our qualitative data collection method was inclusive [52]. The method involved seven selected up-stream stakeholders of the vegetable sector in Malaysia (May-June 2011). Focus group discussions (FGDs) were conducted individually for (1) the Department of Agriculture (DoA), (2) the Federal Agriculture Marketing Authority (FAMA), (3) the Cameron Highlands Vegetable Growers Association, and (4) the Vegetable Farmers Association of Selangor. Each FGD was made up by at least four volunteer respondents. With their consent, the 90 minute FGDs were audio and video taped. Interviews were conducted separately for K-Farm, Malaysian AgriFood Corporation (MAFC), and Centre for Environment, Technology & Development, Malaysia (CETDEM). As recording was not allowed, the 45 minute English interviews were documented using shorthand.

The qualitative information collected was transcribed and saved as raw datasets for content analysis which involves automated or manual coding, reflecting the frequency with which concepts appear in texts. This generates easily understandable categories of coded concepts [53]. This analysis has been commonly used on qualitative data in a variety of agricultural settings, such as the

definition of agricultural literacy [54], the concept of sustainable development [55], and the public perception of sustainable agriculture [56].

From the review of past studies and the datasets, 18 concepts were identified for coding purposes. These 18 concepts are presented in Table 1. The exercise was based on keywords and the meanings of conversational content. e.g. 'environment' and 'health and safety' were coded for a conversation "…sustainable agriculture is not just to protect the environment. It is also for the health of people. It is a holistic approach that integrates (requires) your commitment…"

Table 1. Concepts for refining the definition of 'sustainable agriculture'

No.	Components	Sub-components	Concepts	Sustainability goals
1	Environment	Environmentally enhancing	Ecosystem*	Sustainable agriculture maintains and enhances local ecosystems.
2			Environment*	Sustainable agriculture does not degrade the environment.
3		Resource optimal	Renewable resources*	Sustainable agriculture optimizes the use of renewable farm resources.
4			Non-renewable resources*	Sustainable agriculture optimizes the use of non-renewable farm resources.
5	Economic	Economically viable	Cost**	Sustainable agriculture enables reduction in production cost.
6			Price**	Sustainable agriculture results in higher ex-farm prices.
7			Income*	Sustainable agriculture increases farm incomes.
8			Accessibility**	Sustainable agriculture increases accessibility / supply points to modern markets.
9			Marketability**	Sustainable agriculture increases the demand for produce.
10	Social	Socially justifiable	Food*	Sustainable agriculture satisfies human food needs.
11			Food safety*	Sustainable agriculture produces safe foods.
12			Health and safety*	Sustainable agriculture ensures the safety and does not harm the health of farm workers.
13			Life quality – consumers*	Sustainable agriculture enhances the life quality of consumers.
14			Life quality – workers*	Sustainable agriculture enhances the life quality of farm workers.
15		Functionally feasible	Compatibility**	Sustainable agriculture is compatible with local conditions.
16			Knowledge**	Knowledge of sustainable agriculture is transferrable to farm workers.
17			Technical*	Sustainable agriculture is technically appropriate for farm workers to carry out.
18			Technology**	Sustainable agriculture is assisted by technology.

(Sources: * our literature review; ** the raw datasets)

RESULTS AND DISCUSSION

The outputs of content analysis, which showed the number of times each concept coded in texts, were summarized in Table 2. It was difficult to interpret the table meaningfully in its current form. Therefore, these coded concepts were grouped into attributes (environment, resource, economic, social, and function) according to their similarities. Following that, these attributes were categorized within their related core component of sustainable agriculture. Together, the sub-components were graphically illustrated.

Figure 1 depicts our framework of 'sustainable agriculture'. The figure illustrates, through an integrative approach, sustainable agriculture in various farms. It is a consolidated result of three balanced core components, namely of environment, economic, and social. Each of them neither is independent nor represents one third of sustainable agriculture. Instead, they are complimentarily and form the wholeness.

Though the core components are not implicitly portrayed, they are represented by their own attribute(s): (1) "environmentally enhancing" and "resource optimal" for the component of environment; (2) "economically viable" for the economic component; (3) "socially justifiable and "functionally feasible" for the social component. Each attribute is distilled from similarities shared with other concepts (Table 2).

From the above it is clear that a large part of our findings agrees with past studies [13-14, 41] and includes components of environment, economic, and social in the definition of 'sustainable agriculture'. More notably, our findings also suggest a number of refined attributes in an integrative balanced agricultural system.

Table 2. Frequency of coded concepts

No.	Components	Sub-components	Concepts	Frequency count (number of times coded)							Total
				Official segment		Farmers segment		Private & NGO segment			
				DoA	FAMA	Cameron Highlands	Selangor	K-Farm	MAFC	CETDEM	
1	Environment	Environmentally enhancing	Ecosystem	1	3	-	-	-	-	-	4
2			Environment	4	3	1	3	2	2	1	16
3		Resource optimal	Renewable resources	1	3	2	1	1	-	3	11
4			Non-renewable resources	3	4	2	1	2	1	1	14
5	Economic	Economically viable	Cost	-	2	3	1	1	1	2	10
6			Price	-	3	4	2	1	-	-	10
7			Income	-	3	9	1	1	1	1	16
8			Accessibility	3	1	-	1	1	-	-	6
9			Marketability	-	-	2	1	-	-	-	3
10	Social	Socially justifiable	Food	-	1	-	-	-	-	1	2
11			Food safety	9	2	7	2	3	2	1	26
12			Health safety	5	-	1	2	1	-	2	11
13			Life quality – consumers	1	1	-	-	1	-	4	7
14			Life quality - farmers	1	1	-	-	1	-	3	6
15		Functionally feasible	Compatibility	-	3	7	1	2	1	-	14
16			Knowledge	4	-	3	-	5	-	-	12
17			Technical	3	1	9	2	4	1	-	20
18			Technology	5	1	-	-	-	-	-	6
Total				40	32	50	18	26	9	19	

Notes:
DoA – the Department of Agriculture; FAMA – the Federal Agriculture Marketing Authority; Cameron Highlands – the Cameron Highlands Vegetable Growers Association; Selangor – the Vegetable Farmers Association of Selangor; MAFC – Malaysian AgriFood Corporation; CETDEM – Centre for Environment, Technology & Development, Malaysia

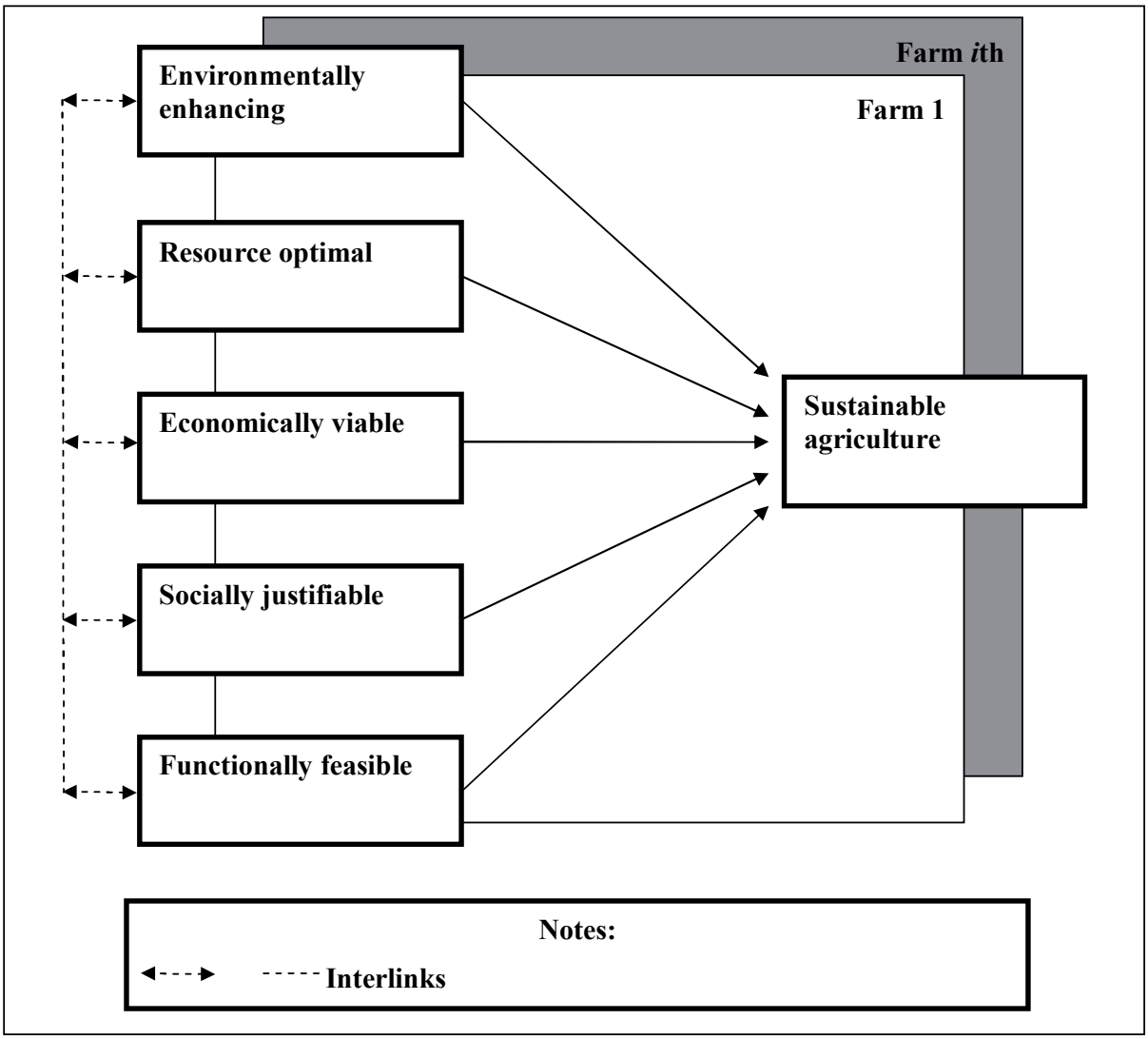


Figure 1. Definitional framework of 'sustainable agriculture'

The Refined Definition

From our structured results, 'sustainable agriculture' could be defined as *the process by which an integrative balanced agricultural system is realized through a dynamic set of practices that are environmentally enhancing, resource optimal, economically viable, socially justifiable and functionally feasible over time.*

The Integrative Balanced Agricultural System

An integrative, balanced, agricultural system is formed by three interacting and equally important components. The three components are environment, economic, and social. They operate together for the common purpose of producing sufficient food for meeting human basic need.

In understanding the core components, past studies [59-61] have characterized the economic and social components as socio-economic subsystems and the environmental component as

ecological subsystems. They are inexorably integrated with the dynamic interaction processes, which result in sustainability [62]. In other words, agricultural activities depend on ecological conditions and these conditions depend upon various agricultural activities [63]. As such, the development of the sustainability in the integrative balanced agricultural system involves feedback loops, which maintain the sustainability within and between the subsystems as a whole [64].

Agriculture is a managed system [61]. Humans in socio-economic subsystems are the integral participants, who manage agricultural systems [4, 65]. These people make decisions on agricultural activities. These decisions are determined by a complex consideration of social and economic goals [63]. In return, their induced changes influence their future decisions [66]. Therefore, improvement in the sustainability of the integrative balanced agricultural system is possible with enhanced ecological and socio-economic feedback processes to the actors over time [64].

Zooming in, these subsystems are characterised by attributes [67]. Attributes are inter-related, crucially and determine the quality of the agricultural system as a whole [68]. In our interest to improve sustainability in the integrative balanced agricultural system, five identifiable attributes emerge. They are (1) environmentally enhancing, (2) resource optimal, (3) economically viable, (4) socially justifiable and (5) functionally feasible.

Environmentally Enhancing

As a key attribute of the environmental component, implicit in the attribute "environmentally enhancing" is the concept of protecting, maintaining and improving environmental quality in sustainable agriculture. This concept emphasises reparation of on- and off-farm environmental issues. Major on-farm environmental issues include soil erosion and nutrient loss, compounding of the impairment of productive capacity [69]. Beyond the farm gate, pollution of surface, ground, and downstream water resources as well as loss of biodiversity are common off-farm issues [70].

This attribute places value on the concept that environmental quality should not be placed at risk by agricultural activity [71]. An important step in realisation is the multi-functional conservation of soil resources, preservation of soil nutrients, control of water quality and the rehabilitation of biodiversity in ensuring environmental quality as a whole. In order to achieve this, damaging practices can be substituted with environmentally sound counters [14]. One should, however, bear in mind that such substitutions require serious consideration in respect to their possibly complex management techniques and long-term environmental consequences. These substitutes do not necessarily refer to either traditional or modern approaches. Hybrid approaches may integrate beneficial practices of both of the aforementioned i.e. to bring them into consonance for the reconciliation of environmental persistence. Even in the face of disturbance, the environment has the ability to produce continuous functionality for ongoing agricultural activities.

Resource Optimal

Another environmental attribute, "resource optimal" refers to the knowledge-based efficiency in the use of available resources in sustainable agriculture. The need for both resource conservation and resource improvement, simultaneously, are embedded in the notion. General resources include soil, water and biomass in the renewable segment and fossil based chemical

inputs in the non-renewable segment. Optimisation promotes intelligent decision making to rationalise the need, the choice and the way to get the most out of available resources.

Contrary to the common view on resource attributes, we do not assert that sustainable agriculture must imply a net reduction in resource use. Our reasoning is based on the logic that the net reduction might cause yield loss, which in turn would necessitate the use of more land to produce food requirements. In contrast, appropriate resource use in sustainable agriculture should be rationalised in a manner which considers underlying need or unanticipated problems. Identified resources are best allocated and used on an optimal basis.

It should be cautiously noted that few farmers are well formally educated. Pretty [72] suggests that the education which they have gained, based on their practical experience, makes productive use of their knowledge and problem-solving skills. It improves their judiciousness and self-reliance by making better use of available resources. Therefore, it makes them less reliant on costly external inputs. In return, these create an accord of resource resilience. When facing some shock e.g. price hikes of chemical fertilisers, agricultural activities can continue to function in the same essential ways e.g. by substitution with cheaper, readily available organic fertilisers.

Economically Viable

"Economic viability" implies a need to ensure a farm's normal economic growth and development whilst realising sustainable agriculture objectives. Sustainable agriculture should not, jeopardise farm profits and the ability to improve profitability levels over the long-run. These goals might be realised through farm structure expansions or technological investments. In probing the meaning of "economic viability" further, Lehman et al. [12] have suggested that the notion represents a complete economic self-sufficiency: if a farmer is said to be economically self-sufficient, he does not get any direct and indirect monetary assistances from other parties.

Positive impacts due to knowledgeable use of resources are likely to spillover from production costs to farm profits. Assuming that sustainable agriculture does not compromise yield, a farmer still could increase profitability levels. This is possible through the reduced reliance on external inputs. Their reduction is replaced by efficient utilisation of available resources. Such a concept of profitability is realised when the net savings offsets the costs of production [73].

When resources are used in a knowledgeable manner, the farmer also makes gains from positive market responses. Sustainable agriculture is rewarded with greater market access since the possibility of supplying to modern marketing channels e.g. supermarkets and hypermarkets, which increasingly demand healthier products.

Additionally, when prices for product produced in an agriculturally sustainable manner are almost the same as those produced by more conventional farming enterprises, consumers would rationally prefer sustainably produced items over the conventional ones. In other words, healthier products are more in demand and consequently more sellable. As such, though the market prices might remain unchanged, farm profits are readily conceived to benefit in one or many ways from the introduction of a sustainable agricultural regime.

Socially Justifiable

The attribute of "socially justifiable" is based on the belief that the social functions of sustainable agriculture should be reasonably based and adequately grounded. Sustainable

agriculture is fundamentally necessitated by the need to satisfying human food needs, which is to responsibly produce sufficient nutritious food to ensure food security. This responsibility clarifies that not only is such a basic function acceptable but its consequences, both positive and negative, might well be justified when viewed through the prism of social well-being.

Food must be produced and consumed safely. This is not limited within the confines of the common concern for food safety. Its focus is also the health of farm workers. One would argue that the reduced reliance on external (chemical) resources has already enhanced the safety levels of farm workers and consumers. The situation is unlikely to be so if their substitutes (available resources) are not used in a knowledgeable manner. For example, improper handling of composts might expose farm workers to prospective chronic diseases [74]. At consumer end, the use of unstable or immature composts might endanger food safety through their contamination by pathogens [75].

To safeguard this social function, government plays a crucial role. Some kind of government support and regulation would be required to develop some rewarding certification programmes e.g. Good Agricultural Practices Scheme or Organic Scheme. These programmes may serve as a blueprint for farm operation and act as credence for public confidence.

Altogether, sustainable agriculture should enhance the quality of life for both farmers and consumers. When society as a whole is healthy, its members are more likely to perform their social functions. Therefore, the abovementioned safety concerns are two pre-requisites to maintain quality of life. Other attributes, as they are inter-related, also play a significant role in determining a farmer's quality of life e.g. lower financial returns are likely to discount a farmer's ability to improve his current qualitys of life.

Functionally Feasible

"Functionally feasible" refers to the capability which enables the practitioner to fulfill the purpose of sustainable agriculture within their current means and conditions. It goes beyond philosophy and considers the quality of being do-able. As such, successful functionality within sustainable agriculture must account for knowledge transferability, compatibility, techniques and technologies.

At its most fundamental level, the knowledge of how to carry out recommended sustainable agricultural practices must be made available and transferrable to farm government support. Given that most farmers are not highly formally educated special considerations should be given to the implementation of education programmes and extension services. Receipt of a good education should see that a farmer would (1) have acquired the necessary skills and competencies, (2) have gained ideas to adapt these practices to the farm, (3) be equipped to share that knowledge with his farm workers and (4) find himself able to realise sustainable agriculture.

Education is meaningless if recommended sustainable agricultural practices are not compatible with a dynamic and harmonious combination. This combination may include farmer value, need and local underlying conditions. Inconsistency of a practice with any one factor in a combination will, in all probability, reduce adoption and diffusion rates.

Recommended sustainable agricultural practices should be technically appropriate for farm workers to carry out. Some of them are readily understood by farm workers; others are relatively complex. In both cases, suitability of these practices is derived from practice and familiarity. However, simpler practices are preferable than those that require new understandings and skill development.

Technologies can ease the implementation of recommended sustainable agricultural practices with efficiency [76]. There are technologies that are ideologically grounded without causing undue damage to the environment [72]. Not only these technologies are meant to maintain environment but they also increase productivity, improve food quality and enhance environmental quality [77].

CONCLUSIONS

To make it more reasonable, acceptable, adaptable and more generally applicable at farm-level, the definition of 'sustainable agriculture' is refined using an inclusive approach. This work is based on a study of the vegetable sector in Malaysia. The refinements are built upon three components of sustainable development: environment, economic and social. These components are commonly embedded in previous research attempts to define the term. The inclusive paradigm involves qualitative data collection from seven selected up-stream stakeholders in the vegetable sector in Malaysia.

The results of content analysis are structured to shed light on a more meaningful definition of 'sustainable agriculture'. They point to a process whereby the realisation of an integrative balanced agricultural system is invariably circumscribed by a definitional framework containing five common structural elements: (1) environmentally enhancing, (2) resource optimal and (3) economically viable, (4) socially justifiable, and (5) functionally feasible. While they have individual emphases, they are nonetheless inter-related in determining the quality of sustainable agricultural systems. As a whole, they provide operational direction that is potentially useful in the planning of agricultural sustainability.

Our definition of 'sustainable agriculture' is flexible. Since it is empirically grounded on reasons and acceptance in an inclusive paradigm, it can be adapted to different agricultural sectors. Its operational attributes can be modified to accommodate varying local nuances and sectoral emphases as they are fluid. They evolve in tandem with changes both in the concept and local sustainability issues.

Future studies should look into the applicability of sustainable agricultural practices. This should also be considered by those research efforts which search for potential innovation in practice and technology. Environmental maintenance practices suitable in one area might be unsuitable in other areas whose focus is on environmental enhancement. Beyond physical suitability, practices which return marginal profits at the current time might be undesirable in future.

Under changing conditions, periodical review, modification and improvement should be made in respect to sustainable agriculture in dynamic patterns. Such complex permutations can be best done via a more inclusive research paradigm [28] i.e. to get more participatory involvement of upstream stakeholders who have practical knowledge and can contribute significantly to the research [78]. Farmers are the ones who must evaluate the suitability of recommended practices and, in consequence, make adoptive decisions.

ACKNOWLEDGEMENTS

This study is part of a PhD research project at the University of Adelaide. The realisation of the project is made possible by Adelaide Scholarship International from the University of Adelaide to Yeong-Sheng Tey. The research project is also partly funded by the Universiti Putra Malaysia's

Research University Grant Scheme (Project No. 9199741). We thank Mark Brindal for proofreading earlier versions of this paper. We are grateful to the Department of Agriculture, the Federal Agriculture Marketing Authority, the Cameron Highlands Vegetable Growers Association, the Vegetable Farmers Association of Selangor, K-Farm, Malaysian AgriFood Corporation and Centre for Environment, Technology & Development, Malaysia for their contribution to the success of this study.

REFERENCES

1. World Commission on Environment and Development, "Our Common Future", Oxford University Press, Oxford, **1987**.
2. S. von Wirén-Lehr, "Sustainability in agriculture - an evaluation of principal goal-oriented concepts to close the gap between theory and practice", *Agr. Ecosyst. Environ.*, **2001**, *84*, 115-129.
3. G. M. Robinson, "Towards sustainable agriculture: current debates", *Geo. Compass*, **2009**, *3*, 1757-1773.
4. E. Karami and M. Keshavarz, "Sociology of sustainable agriculture", in "Sociology, Organic Farming, Climate Change and Soil Science", Sustainable Agriculture Reviews, **2010**, *3*, 19-40.
5. N. Schaller, "The concept of agricultural sustainability", *Agr. Ecosyst. Environ.*, **1993**, *46*, 89-97.
6. A. J. Alonge and R. A. Martin, "Assessment of the adoption of sustainable agriculture practices: Implications for agricultural education", *J. Agr. Educ.*, **1995**, *36*, 34-42.
7. V. W. Ruttan, "The transition to agricultural sustainability", Proceedings of the National Academy of Sciences of the United States of America, **1999**, *96*, 5960–5967.
8. R. Carson, "Silent spring", Harmondsworth, Penguin Books, United Kingdom, **1962**.
9. M. Keating, "The Earth Summit's Agenda for Change: A Plain Language Version of Agenda 21 and the Other Rio Agreements", Geneva, Center for Our Common Future, **1993**.
10. Millennium Ecosystem Assessment, "Ecosystems and Well-Being", Island Press, Washington DC, **2005**.
11. International Assessment of Agricultural Science and Technology, "Agriculture at a Crossroads", Island Press, Washington, DC, **2008**.
12. H. Lehman, E. A. Clark and S. F. Weise, "Clarifying the definition of sustainable agriculture", *J. Agr. Envir. Ethics.*, **1993**, *6*, 127-143.
13. R. R. Weil, "Defining and using the concept of sustainable agriculture", *J. Agron. Educ.*, **1990**, *19*, 126-130.
14. P. Allen, D. V. Dusen, J. Lundy and S. Gliessman, "Expanding the definition of sustainable agriculture", *Am. J. Altrn. Agr.*, **1991**, *6*, 34-39.
15. D. Norman, R. Janke, S. Freyenberger, B. Schurle and H. Kok, "Defining and implementing sustainable agriculture", *Kansas Sustain. Agr. Ser.*, **1997**, *1*, 1-14.
16. C. den Biggelaar and M. Suvedi, "Farmers' definitions, goals, and bottlenecks of sustainable agriculture in the North-Central Region", *Agr. Hum. Val.*, **2000**, *17*, 347-358.

17. D. E. Ervin, L. L. Glenna and R. A. Jussaume Jr., "Are biotechnology and sustainable agriculture compatible?", *Renew. Agr. Food Syst.*, **2010**, *25*, 143-157.
18. R. B. Johnson, "Sustainable agriculture: competing visions and policy avenues", *Int. J. Sustain. Dev. World Ecol.*, **2006**, *13*, 469-480.
19. D. J. Pannell and S. Schilizzia, "Sustainable agriculture: a matter of ecology, equity, economic efficiency or expedience?", *J. Sustain. Agr.*, **1999**, *13*, 57-66.
20. E. B. Barrios and G. M. F. Sarte, "Monitoring sustainable agriculture in Southeast Asia", *Int. J. Sustain. Dev. World Ecol.*, **2008**, *15*, 95-102.
21. A. Bagheri, "Potato farmers' perceptions of sustainable agriculture: the case of Ardabil province of Iran", *Procedia Soc. Behav. Sci.*, **2010**, *5*, 1977-1981.
22. C. J. Barrow, J. Clifton, N. W. Chan and Y. L. Tan, "Sustainable development in the Cameron highlands, Malaysia", *Malays. J. Environ. Manage.*, **2005**, *6*, 41-57.
23. C. J. Barrow, N. W. Chan and T. B. Masron, "Farming and other stakeholders in a tropical highland: towards less environmentaly damaging and more sustainable practices", *J. Sustain. Agr.*, **2010**, *34*, 365-388.
24. T. S. Hansen and O. Mertz, "Extinction or adaptation? Three decades of change in shifting cultivation in Sarawak, Malaysia", *Land Degrad. Dev.*, **2006**, *17*, 135-148.
25. M. N. Shamsudin, H. M. Amir and A. Radam, "Economic benefits of sustainable agricultural production: the case of integrated pest management in cabbage production", *Environ. Asia*, **2010**, *3*, 168-174.
26. M. W. Murad, N. H. N. Mustapha and S. Chamhuri, "Review of Malaysian agricultural policies with regards to sustainability", *Am. J. Environ. Sci.*, **2008**, *4*, 608-614.
27. M. W. Murad, C. Siwar, N. H. N. Mustapha, N. F. N. M. Kamil, S. Muhamad and A. A. Aziz, "Emergence to develop an appraisal system for agricultural practices in Malaysia", *J. Hum. Ecol.*, **2009**, *28*, 191-198.
28. M. P. Pimbert, "The need for another research paradigm", *Seedling*, **1994**, *11*, 20-26.
29. V. Pidlisnyuk, L. Sokol and T. Stefanovska, "Perspectives on sustainable agriculture in Ukraine: the public view", in "Application of Phytotechnologies for Cleanup of Industrial, Agricultural, and Wastewater Contamination", (Eds. P. A. Kulakow, and V. V. Pidlisnyuk), Springer, Netherlands, **2010**, pp. 25-34.
30. Department of Agriculture, "Implementation Guidelines of Good Agricultural Practices toward Certification of Malaysia's Good Agricultural Practices Scheme (Malay version)", Department of Agriculture, Putrajaya, **2009**.
31. Department of Agriculture, "Implementation Guidelines of Organic Farming for the Certification of Malaysia's Organic Scheme (Malay version)", Department of Agriculture, Putrajaya, **2009**.
32. Department of Agriculture, Schemes & certificate. http://www.doa.gov.my/web/guest/skim_dan_pensijilan (Accessed 31 October **2010**).
33. Ministry of Agriculture and Agro-Based Industry "Perangkaan Agromakanan 2010", Ministry of Agriculture and Agro-Based Industry, Putrajaya, **2010**.

34. Y. S. Tey, E. Li, J. Bruwer, A. M. Abdullah, J. Cummins, A. Radam, M. M. Ismail and S. Darham, "Adoption rate of sustainable agricultural practices: a focus on Malaysia's vegetable sector for research implications", *Afr. J. Agr. Res.*, **2012**, *7*, 2901-2909.
35. T. Gips, "What is sustainable agriculture?", Proceedings of the 6th International Scientific Conference of the International Federation of Organic Agriculture Movements (IFOAM), **1986**, Santa Cruz.
36. D. R. Keeney, "Toward a sustainable agriculture: need for clarification of concepts and terminology", *Am. J. Altern. Agr.*, **1989**, *4*, 101-105.
37. I. Lewandowski, M. Haerdtlein and M. Kaltschmitt, "Sustainable crop production: definition and methodological approach for assessing and implementing sustainability", *Crop Sci.*, **1999**, *39*, 184-193.
38. Food and Agriculture Organization of the United Nations, "Dimensions of Need - An Atlas of Food and Agriculture", Food and Agriculture organization of the United Nations, Rome, **1995**.
39. United States Congress, "Food, Agriculture, Conservation, and Trade Act of 1990", in Title XVI, Subtitle A, (Ed. United States Congress), U.S. Government, Washington DC, **1990**.
40. Science Council of Canada, "sustainable agriculture", Science Council of Canada, The Research Challenge Ottawa, **1992**.
41. G. McIsaac, "Sustainability: what can we learn from the past?", *J. Sustain. Agr.*, **1996**, *9*, 3-7.
42. R. Rodale, "Breaking new ground: the search for a sustainable agriculture", *The Futurist*, **1983**, *17*, 15-20.
43. R. J. MacRae, "Strategies for overcoming the barriers to the transition to sustainable agriculture", McGill University, Montreal, **1990**.
44. W. Jackson, W. Berry and C. Colman, "Meeting the Expectations of the Land: Essays in Sustainable Agriculture and Stewardship", North Point Press, San Francisco, **1984**.
45. R. R. Harwood, "A history of sustainable agriculture", in "Sustainable Agricultural Systems" (Eds. C. A. Edwards, R. Lai, P. Madden, R. H. Miller and G. House), Soil and Water Conservation Society, Iowa, **1990**, 3-19.
46. C. Francis and G. Youngberg, "Sustainable agriculture - an overview", in "Sustainable Agriculture in Temperate Zones" (Eds. C. A. Francis, C. Flora, and L. King), John Wiley and Sons, New York, **1990**.
47. D. O. Chikwendu and J. O. Arokoyo, "Women and sustainable agricultural development in Nigeria", *J. Sustain. Agr.*, **1997**, *11*, 53-69.
48. C. Yunlong and B. Smith, "Sustainability in agriculture: a general review", *Agr. Ecosyst. Environ.*, **1994**, *49*, 299–307.
49. P. Crosson, "sustainable agriculture", *Resources*, **1992**, *106*, 14-17.
50. M. R. Bellon, "Farmers' knowledge and sustainable agroecosystem management: an operational definition and an example from Chiapas, Mexico", *Hum. Org.*, **1995**, *54*, 263-272.
51. G. Rasul and G. B. Thapa, "Sustainability analysis of ecological and conventional agricultural systems in Bangladesh", *World Dev.*, **2003**, *31*, 1721-1741.

52. Y. S. Tey, E. Li, J. Bruwer, A. M. Abdullah, J. Cummins, A. Radam, M. M. Ismail and S. Darham, "Qualitative methods for effective agrarian surveys: a researh note on focus groups", *Am.-Eurasian J. Sustain. Agr.*, **2012**, *7*, 60-65.
53. H. F. Wolcott, "Writing Up Qualitative Research", Sage Publications, London, **2009**.
54. M. J. Frick, A. A. Kahler and W. W. Miller, "A definition and the concepts of agricultural literacy", *J. Agr. Educ.*, **1991**, *32*, 49-57.
55. B. E. Aguirre, "Sustainable development as collective surge", *Soc. Sci. Quart.*, **2002**, *83*, 101-118.
56. S. M. Rogers, "Exploring the public's potential perception of sustainable agriculture through the communications in print media", *Sci. Commun. Bull.*, **2010**, *2*, 2-14.
57. G. A. J. Churchill and D. Iacobucci, "Marketing Research: Methodological Foundations", 9th Eds., Thomson, Ohio, **2005**.
58. D. Eribon, "Lévi-Strauss interviewed, part 1", *Anthropol. Today*, **1988**, *4*, 5-8.
59. H. F. Alrøe and E. S. Kristensen, "Towards a systemic research methodology in agriculture: Rethinking the role of values in science", *Agr. Hum. Val.*, **2002**, *19*, 3-23.
60. C. Folke, T. Hahn, P. Olsson and J. Norberg, "Adaptive governance of social-ecological systems", *Annu. Rev. Environ. Resour.*, **2005**, *30*, 441-473.
61. S. Snapp, "Agroecology: principles and practice", in "Agricultural Systems: Agroecology & Rural Innovation for Development" (Eds. S. Snapp, and B. Pound), Academic Press, London, **2008**.
62. J. Park and R. A. F. Seaton, "Integrative research and sustainable agriculture", *Agr. Syst.*, **1996**, *50*, 81-100.
63. G. R. Conway, "Agroecosystems", in "Systems Theory Applied to Agriculture and The Food Chain" (Eds. J. G. W. Jones, and P. R. Street), Elsevier, London, **1990**, pp. 205-233.
64. W. J. Lewis, J. C. van Lenteren, S. C. Phatak and J. H. Tumlinson, "A total system approach to sustainable pest management", Proceedings of the National Academy of Sciences of the United States of America, **1997**.
65. R. Matthews and P. Selman, "Landscape as a focus for integrating human and environmental processes", *J. Agr. Econ.*, **2006**, *57*, 199-212.
66. A. Sundkvist, R. Milestad and A. M. Jansson, "On the importance of tightening feedback loops for sustainable development of food systems", *Food Pol.*, **2005**, *30*, 224-239.
67. A. D. Hall and R. E. Fagen, "Definition of system", in "Modern systems research for the behavioral scientist" (Eds. W. Buckley. Chicago), Aldine Publishing Company, the United States of America, **1968**.
68. C. R. W. Spedding, "An Introduction to Agricultural Systems", 2nd Eds, Elsevier Applied Science, London and New York, **1988**.
69. R. Lal, F. P. Miller and T. J. Logan, "Are intensive agricultural practices environmentally and ethically sound?", *J. Agr. Environ. Ethics*, **1988**, *1*, 193-210.
70. A. C. Weinberg, "Law-input agriculture reduces nonpoint-source pollution", *J. Soil Water Conserv.*, **1990**, *45*, 48-50.

71. K. P. George, "Do we have a moral responsibility to practice a sustainable agriculture?", *J. Sustain. Agr.*, **1990**, *1*, 81-95.
72. J. Pretty, "Overview to four volumes: sustainable agriculture and food", in Earthscan Reference Collection: Sustainable Agriculture and Food" (Ed. J. Pretty), Earthscan, London, **2008**, pp. 15-55.
73. Y. S. Tey and M. Brindal, "Factors influencing the adoption of precision agricultural technologies: a review for policy implications", *Pre. Agr.*, **2012**, *in press*.
74. S. J. Pravinkumar, G. Edwards, D. Lindsay, S. Redmond, J. Stirling, R. House, J. Kerr, E. Anderson, D. Breen, O. Blatchford, E. McDonald and A. Brown, "A cluster of Legionnaires' disease caused by Legionella longbeachae linked to potting compost in Scotland, 2008-2009", *Euro Surveill.*, **2010**, *15*, 1-3.
75. O. O. Cofie, G. Kranjac-Berisavljevic and P. Drechsel, "The use of human waste for peri-urban agriculture in Northern Ghana", *Renew. Agr. Food Syst.*, **2005**, *20*, 73-80.
76. T. Serra, D. Zilberman and J. M. Gil, "Differential uncertainties and risk attitudes between conventional and organic producers: the case of Spanish arable crop farmers", *Agr. Econ.*, **2008**, *39*, 219-229.
77. N. Kingsbury, "Hybrid: The History & Science of Plant Breeding", University of Chicago Press, Chicago, **2009**.
78. J. A. Moles, "Agricultural sustainability and traditional agriculture: learning from the past and its relevance to Sri Lanka", *Hum. Org.*, **1989**, *48*, 70-78.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 397-414

Maejo International Journal of Science and Technology

ISSN 1905-7873

Available online at www.mijst.mju.ac.th

Report

Optimising steel hub location in Thailand

Sakaradhorn Boontaveeyuwat* **and Cherdvong Saengsupavanich**

International Maritime College, Kasetsart University, 199 Moo 6 Sukhumvit Rd., Tungsukla, Si Racha Chonburi, 20230, Thailand

* Corresponding author, e-mail: sakaradhorn@gmail.com, imcskb@src.ku.ac.th

Received: 6 October 2011 / Accepted: 7 November 2012 / Published: 7 November 2012

Abstract: The optimal location of a steel hub in Thailand was analysed by applying a specific research methodology designed to evaluate locations near the seaports. The growth of Thailand's steel industry has become a centre of attention in the last decade, resulting in substantial efforts to form a distribution service centre to minimise the logistic costs associated with handling large steel flows in the future. The main analysis of the steel hub location focused on areas situated near Laem Cha Bang, Map Ta Phut and Prachuab ports since these top three ports are considered important in terms of their steel throughput in Thailand. The transport costs associated with the shipment and inland transport together with port tariffs were calculated for the proposed scenarios of steel hub establishment and these were compared with the existing situation without steel hub. The findings showed that a steel hub located near Laem Cha Bang port was the optimal option involving a saving of 9.4% on the total system costs incurred under the existing situation.

Keywords: steel hub, steel industry, logistics, Laem Cha Bang seaport, Map Ta Phut seaport, Prachuab seaport

INTRODUCTION

The expansion of iron and steel industries in Thailand has become a centre of attention in the last decade due to the fact that it provides materials for major industries in the construction, machinery, automobile and appliance sectors. The steel industry is a substantially competitive business in terms of price and quality, which has resulted in the Thai steel industry experiencing fluctuations in imported raw material prices, production quality and transport costs. Serious competition among contributors in the market regarding time, costs and product quality especially has produced inefficient logistics management and weak coordination among all units of the supply chain. The obvious competitive nature of contributors in the steel supply chain and the need to cut

costs in transportation has led to the concept of forming a distribution service centre as a steel hub in Thailand with the aim of serving semi-finished and fabricated steel products to domestic steel producers and consumers at cheaper costs.

The aim of this paper is to assess the total cost benefits resulting from an alternative steel hub in Thailand—one that would primarily service the existing domestic end users. Therefore, the objectives are to model the incidence of steel flows entering Thailand and the alternative location of steel hub and to compare these with the existing condition (base case) in the evaluation process.

BACKGROUND INFORMATION

Integrated Steel Production Process

The steel production process is generally divided into four phases [1] as shown in Figure 1. In Phase 1, the iron ore procured from an ore mine is mixed with coke or natural gas and fed into a blast furnace to produce pig iron and hot metal whereas if the input is processed by direct smelting reduction, the result will be sponge iron. Pig iron, hot metal and sponge iron are the fundamental materials for the next phase of steelmaking and casting. In Phase 2, hot metal and pig iron are inserted into a basic oxygen furnace to produce liquid steel as part of the improvement process. Sponge iron and scrap can be fed into an electric arc furnace and converted into liquid steel. Then the steel liquid is sent for casting to generate billets, blooms and slabs, which are the basic components for all imported crude steel.

In Thailand, the steel producers provide steel at this stage of the process for steel forming in Phase 3 where the billets, blooms and slabs are manufactured using a hot forming process into hot-rolled products and they can be continuously transformed by cold-rolled forming to produce cold-rolled products. Further processing produces steel products such as long products (bar, wire and hot formed section) derived from billets and flat products (hot-rolled coil, hot-rolled plate and cold-rolled coil) derived from slabs [1]. The steel products in this phase can be value-added by fabrication with the processes of galvanisation by both hot-dip and electro methods, aluminisation, colour coating and tin plating. All the steel types produced in Phase 3 are delivered to end users in the construction, automotive and electronic sectors.

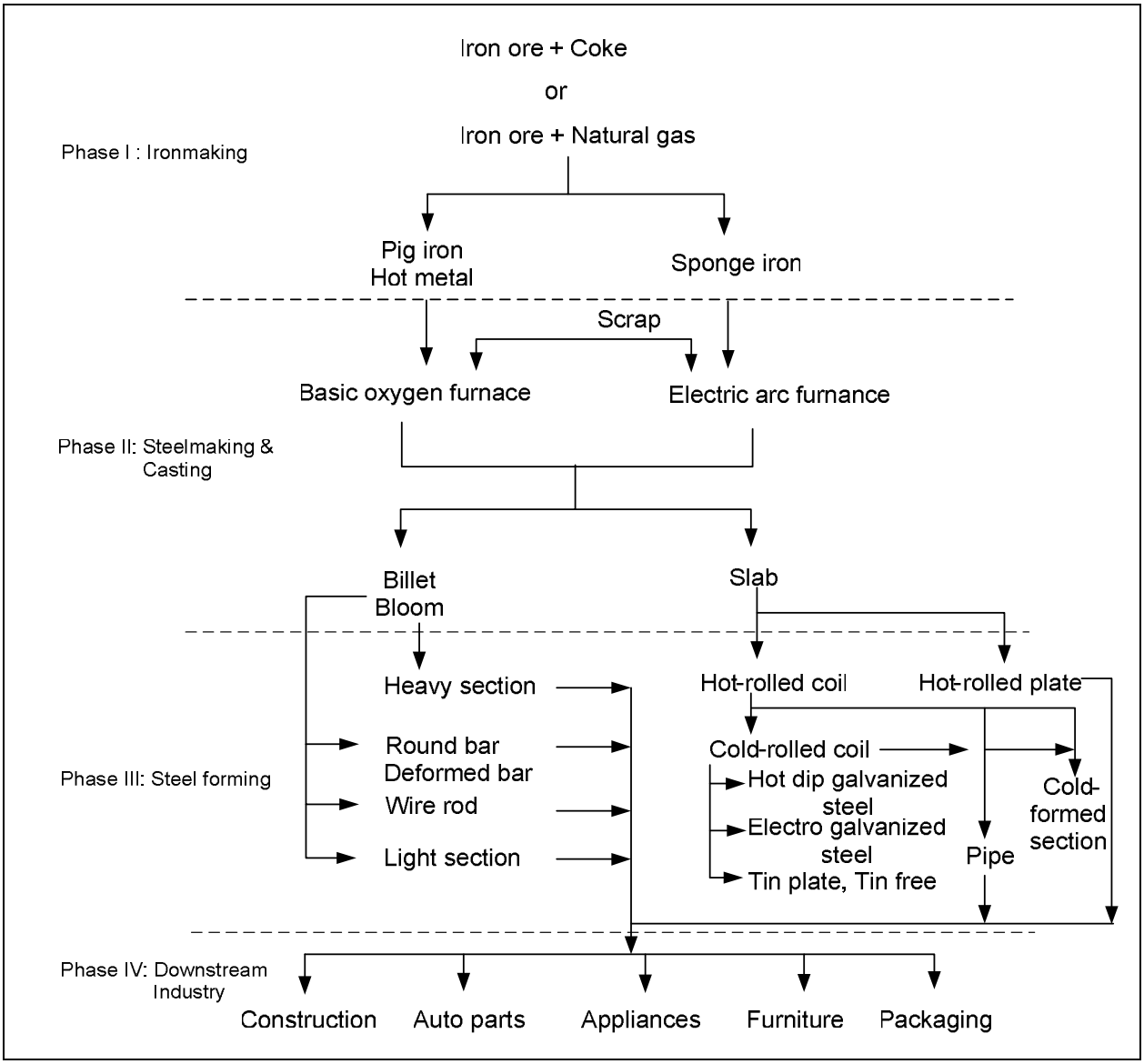


Figure 1. Integrated steel production process [2]

Overview of Global Patterns in Steel Production, Consumption and Trade

China has strengthened its position as the largest steel producing country in the world, increasing crude steel production (covering continuous casting and ingot casting processes) by 77.4% from 353.2 million ton in 2005 to 626.7 million ton in 2010. Crude steel production in Japan was the second highest in the world after China in 2010. It rose marginally from 116.2 to 120.2 million ton during 2005-2007 and gradually decreased to 109.6 million ton in 2010. Crude steel production in India in 2010 was 68.32 million ton, up 49.2% from 2005, making it the fourth largest steel producer in the world, ahead of South Korea whose crude steel production was 58.4 million ton in 2010, up 22.0% from 2005. In Thailand, over the same period, crude steel production decreased by 19.7% from 5.2 to 4.1 million ton. In other Asian countries, production in Taiwan was relatively unchanged from 18.9 slightly up to 19.7 million ton, while production in Malaysia slightly increased by 7.5% from 5.3 to 5.6 million ton and production in Indonesia was relatively

unchanged during the same period. The top 10 steel producing countries and their steel consumption, exports and imports are shown in Figure 2. The top 10 steel producers dominated world steel production, accounting for 80.7% in 2010, whereas the top four major producers in Asia, i.e. China, Japan, India and South Korea produced 750.3 million ton of crude steel, amounting to 75.3% of the world's and 95.3% of Asia's total crude steel output in 2010. China alone accounted for 44.2% of total world steel and 72.8% of steel production in Asia in 2010.

On the steel consumption side, China was also the global leader in 2010 followed by USA and India. Figure 2 shows that in many countries, the level of production was roughly equal to the level of consumption [3]. Nonetheless, the notable exceptions are China, Japan, Russia and Ukraine which had substantially higher production than domestic consumption and this implies that they can export a considerable volume of their steel products. Japan and China were the top two exporting countries in the world, accounting for 42.7 and 41.6 million ton of semi-finished and finished steel products respectively, whilst jointly Russia and Ukraine exported 27.4 million ton of steel products. On the other hand, USA, South Korea and Germany relied on imports to substitute for their lack of domestic production capacity.

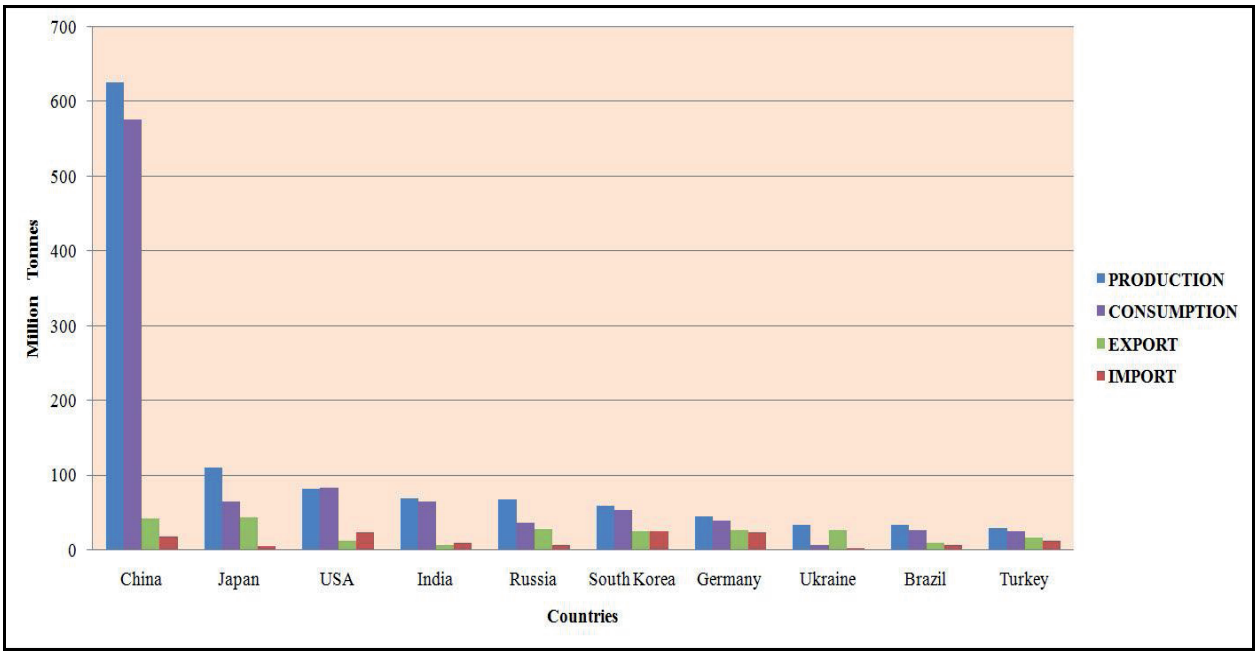


Figure 2. Top 10 crude steel producing countries and their steel consumption, imports and exports in 2010 [3]

Production, Imports, Exports and Consumption of Steel Products in Thailand

In 2010, steel supply chain in Thailand consumed about 14.0 million ton of steel products. Consumption during 2005-2010 ranged between 10.8-14.0 million ton. Production of hot-rolled products, the initial products that will be further processed into other products, however, gradually decreased from 14.4 million ton in 2005 to 11.1 million ton in 2010 with Thailand importing more than 12 million ton of iron and steel products in 2010 [3] (Figure 3). The majority of the imported steel products from Japan were those that could not be manufactured in Thailand such as the high-grade, hot-rolled products which were produced from pig iron and used in the automotive industry

and electrical appliance parts. Thailand exports of semi-finished and finished steel products gradually decreased from 2.2 million ton in 2005 to 1.8 million ton in 2010 [3].

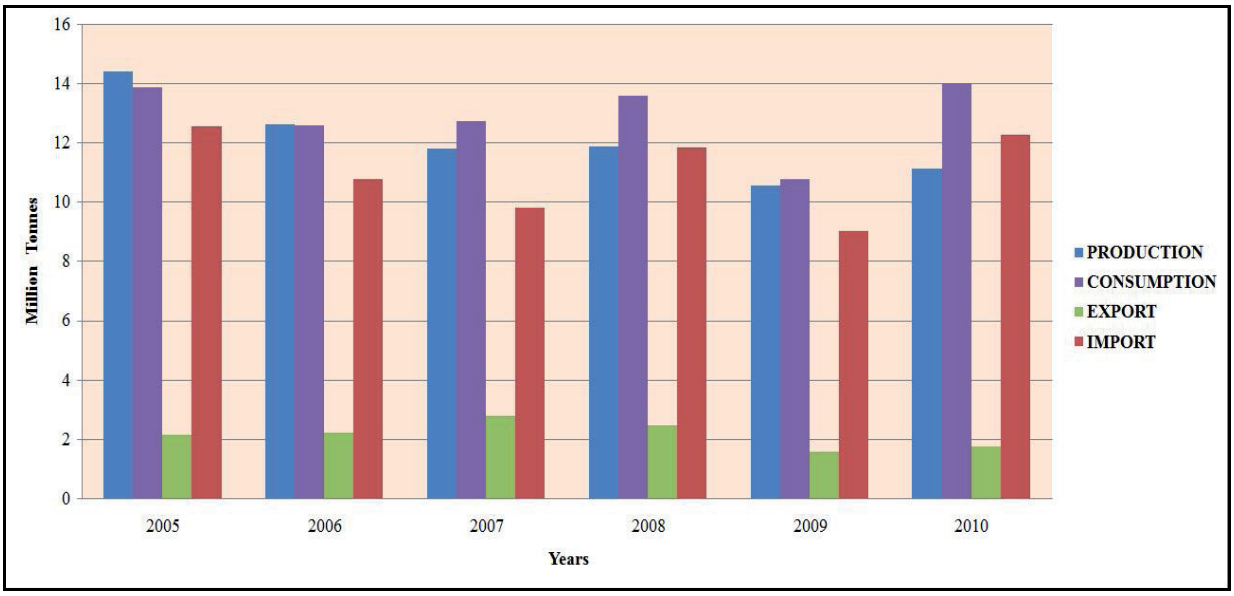


Figure 3. Steel production in Thailand, apparent steel import and export in 2005-2010 [3]

Steel Hub Model

In the steel logistics and distribution process, a steel hub is the intermediary which provides services between steel mills and end users. End users in the steel industry are the original equipment manufacturers (OEMs) that assemble finished products from parts, modules or components supplied by intermediaries in order to operate their own businesses such as construction and auto parts manufacturing. The intermediaries act as steel service centres, stockists and contract manufacturers as well as component suppliers. The Centre for Maritime Studies [4] classifies the steel distribution model into four types (Figure 4). The first type involves the steel mill supplying the crude or semi-finished steel directly to the OEM, who needs to manage its own stocks of steel. It will also process fabricated steel (coating, welding and drawing) and make steel into parts and components which will be used in its product assembly. In this case, the end user conducts the stockholding and processing functions on its own without utilising any functions of a steel hub.

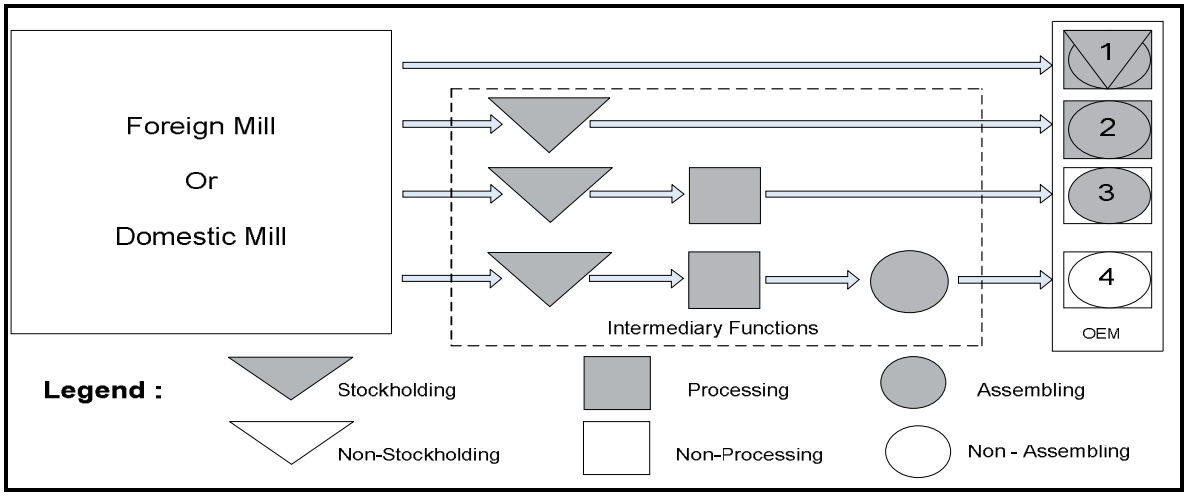


Figure 4. Steel production and distribution chain [4]

The second type of steel distribution takes the form of steel mill supplying steel to the OEM via the steel hub, which performs the functions of a steel stockist and supplies the steel to the end user. The end user purchases steel supplies in order to process them into parts and components for assembly into final products. The third type of steel distribution involves the steel mill supplying crude or semi-finished steel to the OEM via the two operations of stockholding and processing. The internal operation in the steel hub could involve a stockist who holds the stocks and sends them to a steel-processing centre. However, a steel service centre sometimes could handle both the stockholding and processing of the steel and then supplies the processed product to the end user, who still retains the assembly function in this model. Lastly, the steel hub could control the three main functions, namely stockholding, processing and assembling. The operating function in the steel hub would incorporate a stockist, processing centre and/or contract manufacturer, which then assembles the components for the OEM, who only needs to perform product testing or packaging and labelling of the assembled product. Alternatively, a steel service centre could undertake the functions of stockholding and processing and then deliver the processed steel components to a contract manufacturer to undertake the next step in the steel hub operation.

A successful steel distribution model in China provides a useful example [4]. Le Cong is a town situated in the Shunde district of Foshan city in China near the port of Nansha. In terms of steel distribution, Le Cong is the largest steel distribution centre in China; there are more than 1,600 steel distribution companies in Le Cong with about 10,000,000 ton of steel throughput handled per year. The steel in Le Cong is exported using mainly land transport. Le Cong uses open-air sites for the majority of its functions to sell the intermediate steel products such as steel rods and plates and focuses on value-added production by receiving crude steel from the mills and undertaking steel cutting and making steel parts in the mill for car manufactures. Furthermore, Le Cong has built up a strong base of knowledge and expertise in steel trading and distribution with the help of an electronic commerce (e-commerce) system. This facility can serve Le Cong as a one-stop steel trading centre for China in the future. The distribution model of the Le Cong hub is shown in Figure 5.

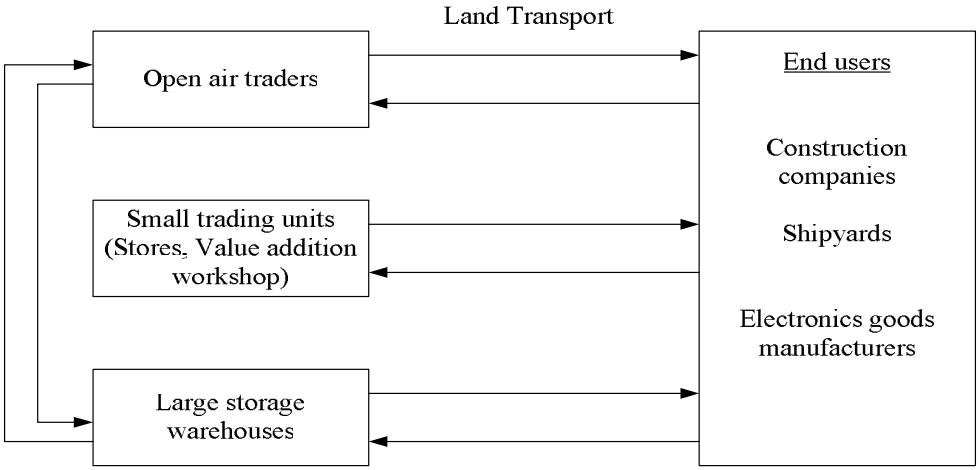


Figure 5. Steel Distribution Model of Le Cong Hub [4]

The steel hub basically contains three key operators: steel service centres (SSC), steel stockists and marketing parts. The Centre of Maritime Studies [4] defines SSC as an operation that buys finished steel, often processes it in some way and then sells it in a slightly different form. Technically, SSC distributes the steel and other metal products which have been processed from an original form into a more value-added form required by customers. Nonetheless, SSC is less capital-intensive than a steel mill as it does not need furnaces, casters and rolling mills. Stockists are not involved in the processing of base steel products but rather focus on ordinary stockholding of steel products and provide low or no value-added services. Stockists mainly operate under a break-bulk-consolidate principle at the regional and local levels and their key business is the timely delivery of products to the customers. Comparatively, SSC has high capabilities in value-adding processing services and breaking up bulk deliveries, whereas a stockist has core capabilities in stockholding and break-bulk deliveries. As part of its marketing services, it will facilitate steel trading between the manufacturers and the end users. The key business is to make strong connections with its traders.

ANALYSIS OF THE STEEL HUB MODEL IN THAILAND

The opportunities are considered for restructuring the current point-to-point routes of steel deliveries from both foreign and domestic production sources in Thailand to one based on a hub-and-spoke network centre. This involves the use of large vessels of around 35,000–50,000 dry weight metric ton (dwt) per shipment to handle the steel products and to reduce the cost of steel import through economies of scale. The initial analysis focuses on the ports with the greatest throughput of steel products. Statistics of steel throughput were gathered from the Marine and Customs Department for 9 alternative transhipment ports in Thailand (Figure 6). The top three customs checkpoints considered important in terms of steel throughput in 2008 were located at ports in Laem Chabang (LCB) consisting of the LCB port, Sri Racha Harbour and the Siam Seaport in Chon Buri (4.2 million ton), Map Ta Phut (MTP) port in Rayong (3.4 million ton) and Prachuab Port in Prachuab Khiri Khan (2.3 million ton). The top three custom checkpoints dominate annual

steel facilitation and amount to almost 100% of the nation's throughput. The residual steel is distributed through Phuket (18,184 ton), Ranong (6,601 ton), Klong Yai, Trat (5,254 ton), Songkla (4,730 ton), Nakhon Si Thammarat (2,500 ton) and Satun (0.4 ton). The steel hub analysis pays specific attention to the top-three locations and investigated them in detail. To assess the alternative steel hub options in Thailand, four possible scenarios are considered:

1. Base case: the current situation
2. Scenario 1: steel hub is established around the LCB ports.
3. Scenario 2: steel hub is established around MTP port.
4. Scenario 3: steel hub is established around Prachuab port.

Figure 6. Locations of Thai steel handling ports

Only steel product flows (hot-forming and cold-forming products) which currently do not pass through the steel hub area are assumed to be transshipped at the steel hub in each scenario (but not for the base case). For instance, there are substantial volumes of steel products shipped to the ports around LCB, MTP and Prachuab ports. In the base case, all steel flows of both crude steel, semi-finished products and hot-forming and cold-forming products were transshipped in their current unloaded ports (that could be ports around the LCB, MTP and Prachuab ports which are the focused transhipment ports in this study) and forwarded to the steel customers. Scenario 1 assumed that all these flows are actually transshipped in ports around LCB and forwarded to the steel hub situated in the industrial zone around LCB, whereas any crude steel and semi-finished products shipped still remain at their transhipment ports (in all cases), since the conceptual operation of a steel hub does not include the steel producing process. This suggests that there is indeed potential for additional steel flow transhipments through the LCB ports, especially if there are cost savings to

be gained. For a better understanding, the conceptual flow chart of the methodology is summarised in Figure 7.

The total cost shown in Figure 7 was derived from foreign sources of steel purchased by the steel producers and customers in Thailand and was calculated from the ocean shipping cost, port tariffs and inland transport costs. It makes use of cost, time and distance components. Therefore, it can be considered as a supply chain system problem to determine the minimum total cost of cargo movement along the supply chain, which is particularly significant in international trade. It has been accepted globally as a standard methodology for analysing supply chain effectiveness in a range of operational and commercial circumstances for general cargo [5-6]. Subsequently, the total costs of alternative steel hub locations and the base case were compared in order to find the optimal steel hub location. A number of studies in the last decade can be found [7-10]. The next section considers scenario modelling and cost analysis after the hub is established.

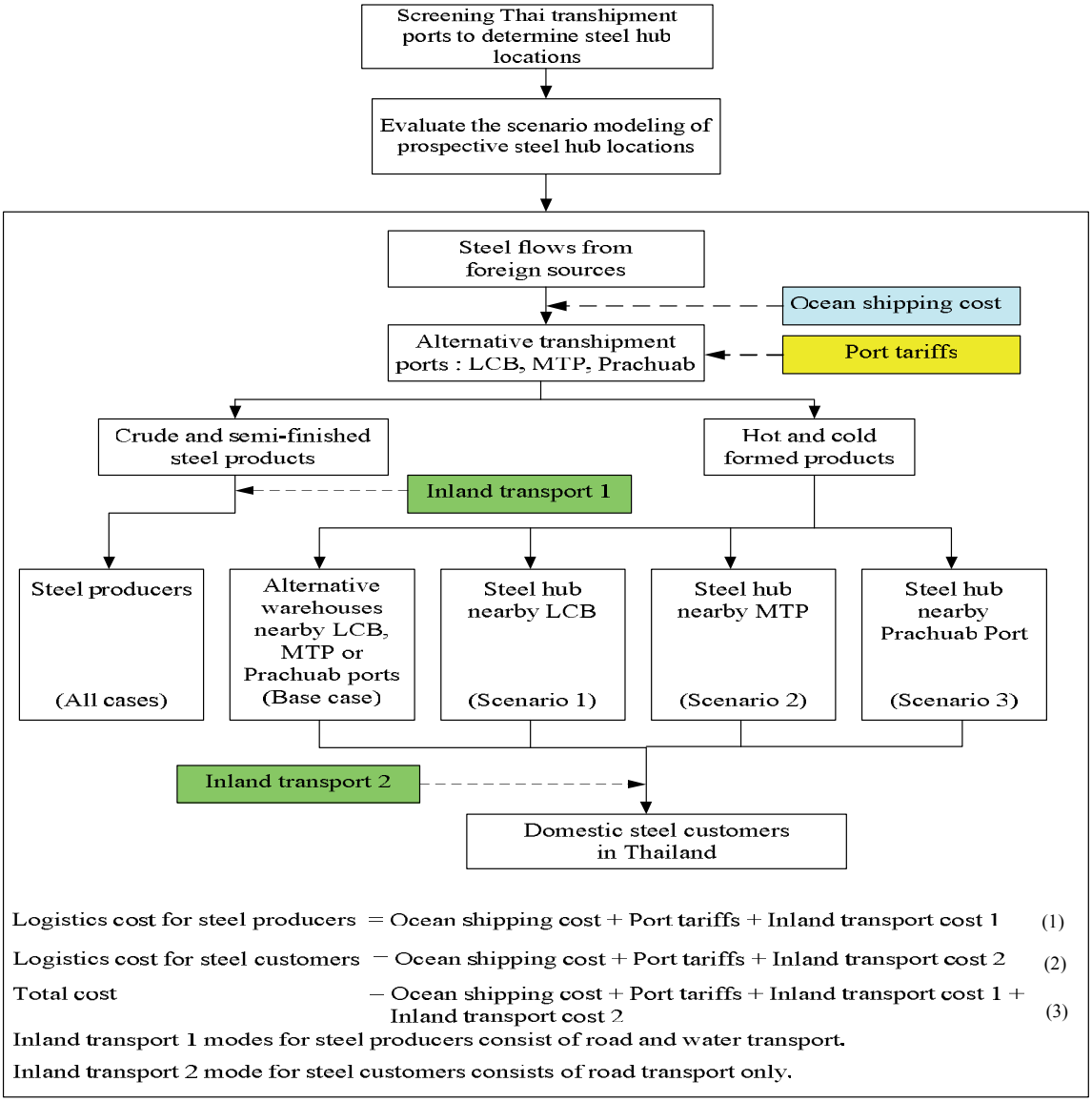


Figure 7. Conceptual flowchart of methodology

MODELED COST ANALYSIS

Mainline Ship Deviation Distance

The initial task in the evaluation process was to determine the deviation distance from the original source of material from foreign countries to each prospective steel hub (Table 1). The estimation of the mainline ship deviation distance is important, although this forms just one part of the overall transhipment distance/cost-assessment process.

Table 1. Mainline ship deviation distance from steel trading countries to prospective steel hub [11]

Steel Trading Countries/Port Transhipment via	LCB	MTP	Prachuab
	Deviation distance (Nautical miles)		
Japan	2,914	2,871	2,891
China	2,202	2,158	2,179
Russia	10,390	10,354	10,335
Korea	2,491	2,448	2,469
Australia	4,560	4,522	4,506
Ukraine	7,042	7,005	6,986
Taiwan	1,624	1,581	1,602
India	3,203	3,167	3,147
Malaysia	980	943	924
USA	7,690	7,646	7,667
Vietnam	1020	977	998
Singapore	774	741	700
Indonesia	1,238	1,200	1,184
Philippines	1,403	1,360	1,381

The targeted destinations in this analysis were divided in two groups: the steel producers and the end users who consume the steel products from the hot-forming, cold-forming and fabricating processes operated by the steel producers. The crude steel and semi-finished steel amounts transported via the candidate transhipment ports before reaching their final destination at steel producers' mills are shown in Table 2. Since the functions of the steel hub do not require furnaces, casters and rolling mills, the flow direction and amount of crude steel and semi-finished products, such as pig iron, billets, slabs, bloom and ingots, are still shipped from foreign sources to their current destinations. This suggests that the establishment of a steel hub does not need to compete directly with the existing major steel producers in Thailand but does seek to facilitate the possibility of reducing costs, in particular those associated with transporting products to serve the end users. Although the location of a steel hub near the existing mills seems to be a good option for steel producers, the steel hub really performs a transition process from manufacturing to serving end user processes. This means that steel distribution from the hub to the end users also needs to be seriously taken into account. From an investigation of the imports by steel producers in Thailand carried out by the Transport Institute [2] and the results of conducting field interviews with the major steel

producers, it can be concluded that the main steel producers are located in three provinces: Rayong, Samut Prakan and Prachuab Khiri Khan. The demand distribution of crude steel, semi-finished steel products and hot- and cold-formed products to steel producers were specified (Table 3). The results indicated that the main destination of steel import from the LCB transhipment ports was Samut Prakan (91%) with the mode of transport being 90% by ship and 10% by road. The destination of steel import from the MTP transhipment port was almost 90% to Rayong and the rest was delivered to Samut Prakan. For Prachuab port, the destination of steel import was 100% to Prachuab Khiri Khan. The demand distribution from the MTP and Prachuab transhipment ports was mostly by road. Although this part of the crude and semi-finished steel transportation process did not depend on the selection of steel hub location since this depends on the hot- and cold-formed products transportation instead, the results of different total costs between existing situation (base case) and other scenarios involved in the establishment of a steel hub are clearer if these costs are included.

Table 2. Steel quantities through custom check points [12]

Customs checkpoint	Crude steel and semi-finished steel products (ton)	Hot- and cold- formed products (ton)
LCB	2,937,463	2,354,528
MTP	2,354,398	1,887,171
Prachuab	1,598,688	1,281,431

Table 3. Steel imports by steel producers [2]

Location of steel producer	Transhipment port	Quantity (%)	Transport mode (%)	
			Ship	Road
Rayong	MTP	86	0	100
	Prachuab	14	0	100
Samut Prakan	LCB	91	90	10
	MTP	9	0	100
Prachuab Khiri Khan	Prachuab	100	0	100

Mainline Ship Deviation Cost plus Inland Transport

This cost analysis considered a 50,000 dwt ship to transport steel from various sources in 14 countries to the three transhipment ports that are nearby each of the candidate steel hubs. The mainline ship deviation costs can be considered using two approaches. The first approach establishes the actual freight rate for a number of steel products currently being quoted by the freight market. The second approach calculates the voyage cost based on the fundamental costs regarding vessels and time charts, which includes fuel expenses, to provide a better estimation of the underlying ship deviation cost structure. As the cost structure of the specified ship size was not available to include in the first approach, no market rates were available. Consequently, the second approach was used based on estimation of actual voyage costs.

Baird [8] supported this second approach asserting that the value of a ship's time can be determined by the prevailing daily time or charter rate or, for owned ships, the daily ship capital and operating costs. Furthermore, time charter rates (dependent on the voyage distance and ship size) being more visible and more standard, provide for much greater clarity and scrutiny as appropriate and representative measures of ship provision costs. The shipping costs from 14 countries to the three transshipment ports were analysed based on the estimation of actual voyage costs regarding the distance and the specified ship size (Table 4).

Based on this approach, at this stage, the port tariffs for each port have not yet been included. Therefore, in order to give a full account of costs, the port tariffs at the three candidate transshipment points were investigated because they are one of the most important factors in the selection of the transshipment location [7, 13]. Table 5 presents port tariffs which consider the main items in this analysis—conservancy dues, berth hire and conventional cargo wharfage—for a 50,000 dwt ship entering port.

The next step in the process was to calculate the inland transport distance between each of the respective transshipment ports and the main locations of steel customers. The distances between the transshipment ports and the steel hubs were not taken into account in this analysis since the analysis assumes that the three hubs are located very close to their respective transhipment ports. Table 6 presents the distance to 22 end-user destinations and demand distribution of hot- and cold-formed products after the fabricating processes from the three prospective steel hubs. This analysis considered road transport, since the steel stockists mainly used road transport in the base case. The operating cost of a vehicle was $US 0.05692/ton.km which already included any product handling costs. The discounted transport cost from each steel hub to each demand point which was assumed as the hub point in each province of Thailand for receiving steel products from the steel hub was set at 0.9 leading to a realistic configuration that every link of the inland transport network carries large amounts of steel between hubs [14].

Table 4. Mainline ship costs derived from steel trading countries to each prospective steel hub [15]

Steel trading country/Port transhipment via	LCB	MTP	Prachuab
	Steel shipping cost for 50,000 dwt ship (US$/ton)		
Japan	32.53	32.07	32.42
China	29.12	28.91	29.01
Russia	68.42	68.25	68.16
Korea	30.50	30.90	30.40
Australia	40.18	40.25	40.18
Ukraine	52.35	52.17	52.08
Taiwan	26.34	26.14	26.24
India	33.92	33.75	33.65
Malaysia	23.25	23.07	22.98
United States	55.46	55.25	55.35
Vietnam	23.44	23.24	23.34
Singapore	22.26	22.10	21.91
Indonesia	24.49	24.31	24.23
Philippines	25.28	25.08	25.18

Table 5. Port tariffs for a 50,000 dwt ship collected and developed from [16-18]

Steel trading countries/Port transhipment via	LCB	MTP	Prachuab
	Port tariff for a 50,000 dwt ship ($US)		
Conservancy dues	14,820	11,856	14,820
Berth hire	4,631	4,631	6,175
Conventional cargo wharfage	50,000	50,000	66,667

Table 6. Deviation distances and demand distribution from steel hubs to domestic destinations [2]

LCB	MTP	Prachuab	Destination	Demand distribution (%)
Distance from hub to destination (km)				
785	838	1020	Chiang Mai	2.06
151	203	329	Bangkok	26.98
379	427	500	Kanchanaburi	1.51
448	498	684	Kamphaeng Phet	0.77
426	456	747	Khon Kaen	3.14
114	145	485	Chachoengsao	2.01
28	84	446	Chon Buri	4.46
298	346	511	Chai Nat	2.1
191	239	319	Nakhon Pathom	1.46
364	413	621	Nakhon Ratchasima	4.72
352	406	572	Nakhon Sawan	3.83
143	199	513	Prachin Buri	2.01
472	522	714	Phitsanulok	3.75
66	12	535	Rayong	4.9
238	284	277	Ratchaburi	2.46
101	174	365	Samut Prakan	17.46
161	210	319	Samut Sakhon	8.28
210	259	470	Saraburi	2.1
218	267	435	Suphan Buri	0.62
764	813	285	Surat Thani	0.75
185	234	418	Ayutthaya	1.8
626	657	971	Ubon Ratchathani	2.87

RESULTS AND DISCUSSION

The enumeration method was applied to all variables. The model was composed of 14 foreign sources, 3 transhipment ports near which steel hubs were proposed to be established, 3 steel producers whose demand was crude and semi-finished steel products and 22 steel consumers whose demand was steel products fabricated from hot- and cold-formed steel products. Table 7 shows the logistics costs derived from Equation (1) in Figure 7 for the movement of the crude and semi-finished steel products through the three transhipment ports. The costs depended on the amount of steel moved through each port. Nonetheless, it can be observed that the LCB ports provide the minimum portion of ocean shipping and port tariffs (73.5% and 2.75% respectively) compared with MTP (91.58%; 3.3%) and Prachuab (92.03%; 4.35%). Furthermore, the inland transport costs by road and coastal transportation at the LCB ports were the highest components of the total system costs (23.73%) compared with MTP (5.13%) and Prachuab (3.62%) because no crude steel or semi-finished steel was forwarded to the main steel producers in Chon Buri. The main destinations of

steel transshipped using the LCB ports were Samut Prakan and Rayong. The interview results from a representative steel producer in Chonburi revealed that generally they do not import steel by themselves but rather purchase the hot-formed products from the mills in Rayong.

Table 7. Logistics costs of moving crude steel and semi-finished steel products to steel producers (US$ million)

Cost component	Transhipment port			Total cost (US$ million)
	LCB (2.94 MMT)	MTP (2.35 MMT)	Prachuab (1.6 MMT)	
Steel foreign sources to port transhipment	109.20 (73.5%)	86.97 (91.58%)	59.21 (92.03%)	255.38 (82.96%)
Port tariffs	4.08 (2.75%)	3.13 (3.3%)	2.80 (4.35%)	10.01 (3.25%)
Inland transport	35.24 (23.73%)	4.87 (5.13%)	2.33 (3.62%)	42.44 (13.79%)
Total costs	**148.53**	**94.97**	**64.34**	**307.84**

MMT = Million metric ton

Table 8 shows the logistics costs for hot- and cold-formed steel products through the three transhipment ports derived from Equation (2) in Figure 7. It can be observed that Prachuab port provides the lowest portion of total system costs in ocean shipping (57.84%) compared with the MTP (69.68%) and the LCB (73.44%) ports. Tariffs made up about the same portion of costs for all three ports, with a range of 2.5–2.74%. Nevertheless, the LCB ports offered the cheapest cost portion for inland transport with 23.82% versus MTP at 27.81% and Prachuab port at 39.43%.

Table 8. Logistics costs of moving hot- and cold-formed steel products to steel consumers (US$ million)

	Transhipment port			Total cost (US$ million)
	LCB (2.35 MMT)	MTP (1.89 MMT)	Prachuab (1.28 MMT)	
Steel foreign sources to port transhipment	87.53 (73.44%)	69.71 (69.68%)	47.46 (57.84%)	204.7 (67.94%)
Port tariffs	3.27 (2.74%)	2.51 (2.51%)	2.25 (2.74%)	8.03 (2.67%)
Inland transport	28.39 (23.82%)	27.82 (27.81%)	32.36 (39.43%)	88.56 (29.39%)
Total costs	**119.19**	**100.04**	**82.06**	**301.29**

MMT = Million metric ton

The total costs considered for deciding the optimal steel hub location focused on hot- and cold-formed steel products and the logistics cost derived from moving crude and semi-finished steel products for steel producers derived from Equation (3) in Figure 7. As a result (shown in Table 9), the LCB ports in Scenario 1 were selected as the optimal solution involving the establishment of a steel hub near the so-called LCB ports.

Table 9. Total costs (US$ million) of base case and three scenarios for a steel hub evaluation

Hot- and cold-formed steel products	**Base case**	**Scenario 1 LCB steel hub**	**Scenario 2 MTP steel hub**	**Scenario 3 Prachuab steel hub**
Steel origin sources to port transhipment	204.70 (67.49%)	205.33 (75.23%)	204.02 (71.67%)	204.56 (60.21%)
Port tariffs	8.03 (2.67%)	7.67 (2.81%)	7.34 (2.58%)	9.68 (2.85%)
Inland transport	88.56 (29.39%)	59.93 (21.96%)	73.28 (25.74%)	125.51 (36.94%)
Total costs	**301.29**	**272.93**	**284.64**	**339.75**

A number of interesting observations can be made from Table 9. Firstly, 75.23% of the total costs in Scenario 1 is related to ocean shipping costs derived from various sources to transhipment ports. This was the highest cost in every case, amounting to as much as US$ 205.33 million, which was slightly higher (0.31%) than the base case scenario of US$ 204.70 million. Secondly, the minimum cost for ocean shipping (US$ 204.02 million) was at the MTP steel hub (the steel hub located near the MTP port) in Scenario 2, which was slightly cheaper than the base case and Scenarios 1 and 3, by amounts of US$ 0.68 million, US$ 1.31 and US$ 0.54 million respectively. Thirdly, the most expensive port tariffs were at Prachuab steel hub (the steel hub located near Prachuab port) in Scenario 3, amounting to US$ 9.68 million (2.85% of total costs) whilst the cheapest port tariffs were at MTP (US$ 7.34 million, 2.58% of total costs). The inland transport costs for the LCB steel hub in scenario 1 were the cheapest compared to the values of US$ 88.56 million (29.39% of total costs), US$ 73.28 million (25.74% of total costs) and US$ 125.51 million (36.94% of total costs) for the base case, MTP and Prachuab steel hubs respectively. It can be observed that the parameter of inland transport costs for the LCB steel hub in Scenario 1 played a key role in the LCB steel hub achieving the optimal solution by providing the lowest total system costs of US$ 272.93 million, which represented a saving of US$ 28.36 million (9.41%) from the base case.

CONCLUSIONS

Compared with the reference base case and competing steel hub locations, the optimal steel hub location must offer the lowest total system costs involving a combination of mainline ocean shipping costs, port tariffs and inland transport costs. Based on these three measures, the LCB steel

hub offered the optimal location with savings of almost 10% compared to the base case. It was shown that the transport costs (a combination of shipping and inland transport costs) dominated, accounting for 97.19% whilst port tariffs represented only 2.81% of total system costs. This explains the significance of the location at which steel imports are received for distribution to domestic markets.

Since the steel hub does not aim to compete with the major steel mills in Thailand, the steel hub function focuses on imported hot- and cold-formed steel products and fabrication for value-added processing to the imported steel before supplying it to the domestic market. Another possible function of a steel hub, not dealt with here, is import and re-export services for neighboring countries such as Laos, Vietnam and Cambodia or other countries where the cost penalty for transhipment via a steel hub and the economies of scale associated with using large ships are influential points to consider when comparing with direct shipment. Such an additional function requires the hub to be established in a "free zone" or "bonded zone" where, with the agreement of the Customs Department, the establishment costs for setting up the free zone hub mean that no tax is levied on the imported and re-exported steel products.

By assessing simultaneously the impact of shipping costs, port tariffs and inland transport costs on steel flows and using the information to determine the optimal steel hub location based on these parameters, this paper has applied a specific method for the analysis of a steel hub location in Thailand. These findings appear to fit well with the steel industry trend towards increased transhipment based on the development of a steel hub near the selected port. The establishment of a steel hub offers a substantial reduction in inland transport costs, thereby resulting in minimisation of the overall transport costs. Nonetheless, further improvements to the steel hub system should consider increasing the facilities of the selected loading port servicing the hub so that larger ships can be handled to accommodate the structural shifts of increased trade flow to enhance service efficiencies and generate further reductions in overall transport costs.

ACKNOWLEDGEMENTS

The authors wish to thank United Thai Shipping Corporation Ltd. for supporting the research fund and providing important data utilised in this paper.

REFERENCES

1. Unithai Group, "Steel Study in Thailand", Iron and Steel Institute of Thailand, Bangkok, **2008**, pp**.** 4-5.
2. Transport Institute, "The Development Plan in Transportation Network for Supporting the Potential in the Steel Industry Competition" (in Thai), Iron and Steel Institute of Thailand, Bangkok, **2008**, pp. 4-28 – 4-35.
3. World Steel Association, "Steel Statistical Yearbook 2011", World Steel Committee on Economic Studies, Brussels, **2011**, pp. 5-97.
4. Centre for Maritime Studies, "Assessment of Singapore as a Steel Hub", National University of Singapore, Singapore, **2008**, pp. 89-92.

5. UNESCAP, "Transit Transport Issues in Land-locked and Transit Developing Countries", UNESCAP, Bangkok, **2003**.
6. UNESCAP, "Integrated International Transport and Logistic System for Northeast Asia", UNESCAP, New York, **2006**.
7. M. S. Ha, "A comparison of service quality at major container ports: implications for Korean Ports", *J. Transp. Geogr.*, **2003**, *11,* 131-137.
8. A. J. Baird, "Optimising the container transhipment hub location in Northern Europe", *J. Transp. Geogr.*, **2006**, *14*, 195-214.
9. A. J. Baird, "The economics of container transhipment in Northern Europe", *Int. J. Maritime. Econ.*, **2002**, *4*, 249-280.
10. Napier University, "Orkney container terminal study", Unpublished report for Orkney Islands Council, Halifax Port Authority and Highlands and Islands Enterprise, Napier University, Edinburgh, **2000**.
11. Netpas, "Netpas distance", https://netpas.net/index.php (Accessed 1 June **2011**).
12. The Thai Customs Department, "Steel quantities through custom check points", Unpublished raw data, **2008**.
13. M. S. Ha, "A study on the evaluation on service quality of major container ports in North-East Asian region and their cooperative scheme", *Korea Int. Comm. Rev.*, **2001**, *16*, 143-171.
14. J. F. Campbell, "Integer programming formulations of discrete hub location problems", *Eur. J. Oper. Res.*,**1994**,*72*, 387-405.
15. Metallic Steel Center Corporation Limited, "The charter rate for steel products shipment", Unpublished raw data, **2010**.
16. Laem Cha Bang port, "Port tariffs", www.laemchabangport.com/index.php?option=com_content&view=article&id=162&Itemid=102&lang=th (Accessed 2 June **2011**).
17. Map Ta Phut port, "Map Ta Phut port information", http://www.maptaphutport.com/maptaphut/index.php?option=com_content&view=article&id=55&Itemid=114 (Accessed 2 June **2011**).
18. Prachuab Port Corporation Limited, "Port tariffs", Unpublished raw data, **2011**.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 415-429

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

General variable strength t-way strategy supporting flexible interactions

Rozmie Razif Othman [1], Kamal Zuhairi Zamli [2,*] and Lukito Edi Nugroho [3]

[1] School of Computer and Communication, Universiti Malaysia Perlis (UniMAP), PO Box 77, d/a Pejabat Pos Besar, 01007 Kangar, Perlis, Malaysia
[2] Faculty of Computer Systems and Software Engineering, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Kuantan, Pahang, Malaysia
[3] Electrical Engineering and Information Technology Department, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia
* Corresponding author, e-mail: kamalz@ump.edu.my

Received: 10 February 2012 / Accepted: 6 November 2012 / Published: 7 November 2012

Abstract: To ensure conformance and establish quality, software testing is an integral part in software engineering lifecycle. However, because of resource and time-to-market constraints, testing all exhaustive possibilities is impossible in nearly all practical testing problems. Considering the aforementioned constraints, much research now focuses on a sampling technique based on interaction testing (termed as t-way strategy). Although helpful, most t-way strategies (e.g. AETG, In-Parameter-Order General (IPOG), and GTWay) assume that all parameters have uniform interaction. In reality, the interaction among parameters is rarely uniform. Some parameters may not even interact, wasting the testing efforts. As a result, a number of newly developed t-way strategies that consider variable-strength interaction based on input–output relationships have been developed, e.g. Union, ParaOrder, and Density. Although useful, these strategies often suffer from lack of optimality in terms of the generated test size. Furthermore, no single strategy is dominant because the optimal generation of t-way interaction test suite is considered an Nondeterministic Polynomial (NP) hard problem. Motivated by the above-mentioned challenges, this paper proposes and implements a new strategy, called General Variable Strength (GVS). GVS has been demonstrated, in some cases, to produce better results than other competing strategies.

Keywords: interaction testing, t-way test generation, variable strength interaction, software testing

INTRODUCTION

Nowadays, we are increasingly dependent on software to facilitate our daily chores, from mobile phone applications to sophisticated airplane control system. To ensure quality and reliability, we have to consider many combinations of possible input parameters, hardware/software environments, and system conditions, tested and verified for conformance. Because of resource and time-to-market constraints, testing all exhaustive possibilities is practically impossible. As a result, many t-way strategies (where t identifies the interaction strength) have been proposed in the literature for the past 20 years. All strategies help in searching, as well as minimising, the final test cases, i.e. to form a complete suite that deals with all required interactions.

Although helpful, most existing t-way strategies, e.g. GTWay [1, 2], In-Parameter-Order (IPO) General (IPOG) [3], IBM's Test Case Handler (ITCH) [4], and Jenny [5], assume that all parameters have uniform interaction. In reality, interaction among parameters is rarely uniform. Actually, some parameters may not even interact, wasting the testing efforts. To address the aforementioned issues, several newly developed t-way strategies have been developed, which consider variable-strength interaction based on input-output relationships. Schroeder proposed two strategies, called Union [6] and Greedy [7]. Meanwhile, Wang et al. proposed three strategies, namely ReqOrder [8], ParaOrder [8] and Density [9]. Finally, a public-domain tool available from SourceForge, called Test Vector Generator (TVG) [10], also supports variable-strength interaction based on the input-output relationships. Although useful, these newly proposed strategies suffer from lack of optimality (i.e. in terms of test size). Furthermore, no single strategy is dominant because the optimal generation of t-way interaction test suite is considered an Nondeterministic Polynomial (NP) hard problem [11, 12]. Motivated by the aforementioned challenges, this paper discusses the design and evaluation of a new strategy, called General Variable Strength (GVS).

Problem Definition Model

Exhaustive testing is impossible because the number of test cases can be exorbitantly large, even for simple software and hardware products. Let us consider a hardware product with 20 on/off switches. Testing all possible combinations would require 2^{20} = 1,048,576 test cases. If the time required for one test case is 5 min, then the test would take nearly 10 years to complete.

The same argument is applicable in any software system. As an illustration, let us consider the option dialog in the Microsoft Excel software (Figure 1). Even if only the View tab option is considered, 20 possible configurations have to be tested. Except for the gridline colour that takes 56 possible values, each configuration can take two values, namely checked or unchecked. Here, we must evaluate $2^{20} \times 56$ or 58,720,256 combinations of test cases. Using the same calculation assumption, a complete test of the View tab option would require nearly 559 years.

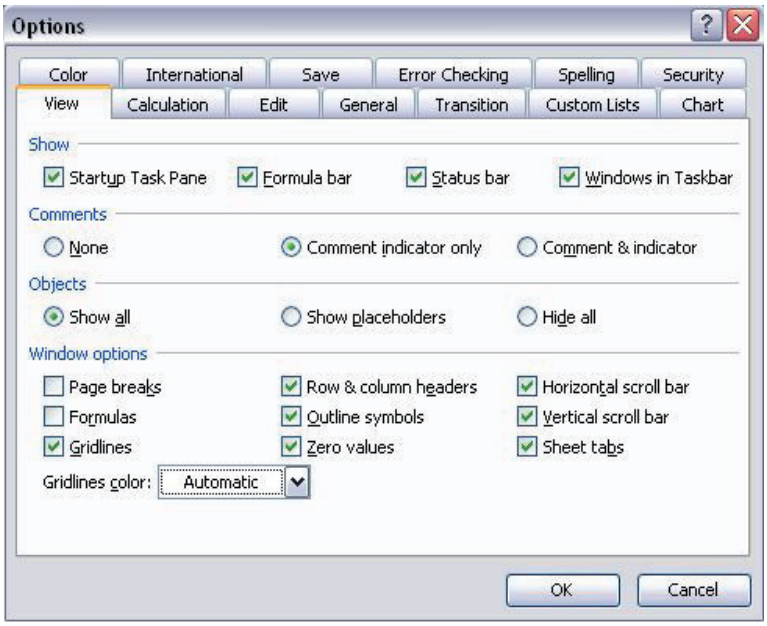


Figure 1. Microsoft excel view tab options

The above-mentioned examples highlight the common combinatorial explosion problem in software testing. Given limited time and resources, the main research questions are as follows:

- What is the minimum number of (sample) tests to be considered?
- How can one decide (i.e. the strategy) which combination of values to choose over the large combinatorial data sets?

Background

Over the years, many sampling-based testing strategies, e.g. equivalence partitioning, cause and effect analysis, decision table, and boundary value analysis, have been developed [13]. Although helpful, many strategies were not sufficiently effective in dealing with the faults due to interaction. Thus, t-way strategies have been proposed to address this issue. Briefly, the t-way strategies offer four possible interactions to generate the test suite: uniform strength, variable strength, input-output based relationship, and mixed interactions. Figure 2 shows the features of the possibilities of each interaction using a software system with five parameter inputs (P_0, P_1, P_2, P_3 and P_4) and three outputs (f_0, f_1 and f_2).

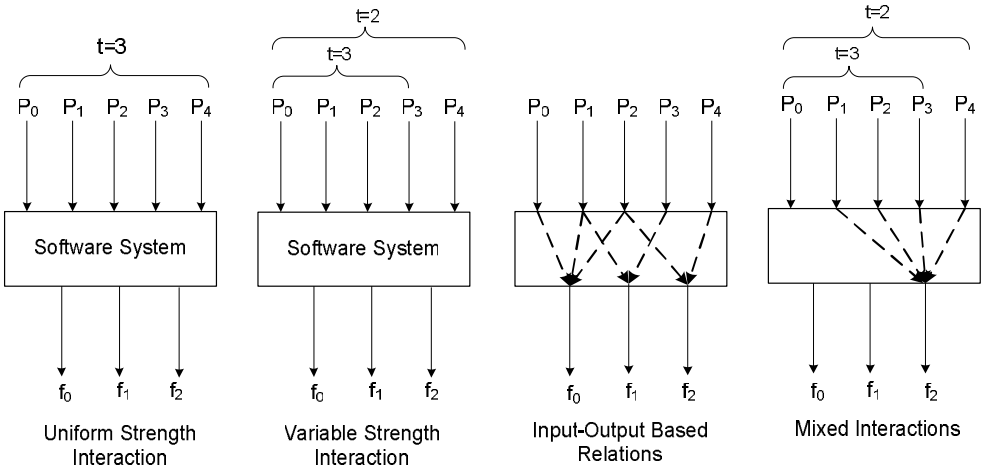


Figure 2. Interaction possibilities within the software system

Uniform-strength interaction is the basis of interaction testing, where all input parameters are assumed to be uniformly interacting (i.e. with constant interaction strength (*t*) throughout). To test all interacting parameters, the test suite must cover all the t-way combinations at least once. In this manner, all possible uniform strength interactions can be tested and, hence verified for correctness. Mathematically, the uniform-strength test suite can be represented using the covering array notation as

$$F = CA(N,t,C) \tag{1}$$

where: N is the final test suite size,

t is the interaction strength,

C is the value configuration, which can be represented as $v_0^{p_0}, v_1^{p_1}, \ldots, v_n^{p_n}$, indicating p_0 parameters with v_0 values and p_1 parameters with v_1 values, and so on.

In contrast to the uniform-strength interaction counterpart, the variable-strength interaction considers more than one interaction strength for the test suite generation. Here, a particular subset of input parameters can have higher interaction dependence than the other parameters, which indicates that failure due to the interaction of that subset may have more significant effects on the overall system. Thus, stronger interaction strength can be assigned accordingly. Using the covering array notation, the variable-strength test suite F can be represented as

$$F = VCA(N,t,C,S) \tag{2}$$

where: N is the final size of the test suite,

t is the dominant interaction strength,

C is the value configuration, which can be represented as $v_0^{p_0}, v_1^{p_1}, \ldots, v_n^{p_n}$,

S is the multi-set of the disjoint covering array with strength larger than t, as given in Eq. (1).

Although the uniform- and variable-strength interactions assume that all input parameters interact with one another, the input-output-based relationships use the knowledge of the input-

output relationships for the test suite generation. Using similar covering array notations as that in the uniform- and variable-strength interactions, the input-output-based relationship F can be represented as

$$F = IOR(N,C,Rel) \quad (3)$$

where: N is the final size of the test suite,

C is the value configuration, which can be represented as $v_0^{p_0}, v_1^{p_1}, \ldots, v_n^{p_n}$,
Rel is the input-output relationship definition set based on the combination index $p_0 \ldots p_n$,
$|Rel|$ is the number of input-output relationship definition sets

By combining the uniform strength, variable strength, and input-output-based relationships, the mixed interaction represents the amalgam of knowledge on the input and output behaviours of the system under test. Mathematically, the mixed interaction adopts the same covering array notation as the input-output-based relationships. The uniform and variable strengths can also be represented in the same manner. Table 1 summarises the input-output conversion for the earlier example shown in Figure 2.

Table 1. Input-output conversions for Figure 2

F= IOR (N,C, Rel)	**Input–output interaction representations**
Uniform-strength interaction ($\|Rel\| = 10$) Rel = {{0,1,2}, {0,1,3}, {0,1,4}, {0,2,3}, {0,2,4}, {0,3,4}, {1,2,3}, {1,2,4}, {1,3,4}, {2,3,4}}	{P_0,P_1,P_2}, {P_0,P_1,P_3}, {P_0,P_1,P_4}, {P_0,P_2,P_3}, {P_0,P_2,P_4}, {P_0,P_3,P_4}, {P_1,P_2,P_3}, {P_1,P_2,P_4}, {P_1,P_3,P_4}, {P_2,P_3,P_4}
Variable-strength interaction ($\|Rel\| = 10$) Rel = {{0,1,2}, {0,1}, {0,2}, {0,3}, {0,4}, {1,2}, {1,3}, {1,4}, {2,3}, {2,4}}	{P_0,P_1,P_2}, {P_0,P_1,P_3}, {P_0,P_2,P_3}, {P_1,P_2,P_3}, {P_0,P_1}, {P_0,P_2}, {P_0,P_3}, {P_0,P_4}, {P_1,P_2}, {P_1,P_3}, {P_1,P_4}, {P_2,P_3}, {P_2,P_4}
Input-output-based relationships ($\|Rel\| = 3$) Rel = {{0,1,2}, {1,3}, {2,4}}	{P_0,P_1,P_2}, {P_1,P_3}, {P_2,P_4}
Mixed interactions ($\|Rel\| = 11$) Rel = {{1,2,3,4}, {0,1,2}, {0,1}, {0,2}, {0,3}, {0,4}, {1,2}, {1,3}, {1,4}, {2,3}, {2,4}}	{P_1,P_2,P_3,P_4}, {P_0,P_1,P_2}, {P_0,P_1,P_3}, {P_0,P_2,P_3}, {P_1,P_2,P_3}, {P_0,P_1}, {P_0,P_2}, {P_0,P_3}, {P_0,P_4}, {P_1,P_2}, {P_1,P_3}, {P_1,P_4},{P_2,P_3}, {P_2,P_4}

The most general representation of any form of parameter interactions is the input-output-based relationship. Thus, GVS is developed as a general strategy to integrate seamlessly and support all interaction possibilities.

Related Work

Many t-way strategies have been proposed for the past 20 years. In a nutshell, existing t-way strategies can be categorised either as a one-parameter-at-a-time (OPAT) or a one-test-at-a-time (OTAT) strategy.

The OPAT strategy initially generates an exhaustive test for a few selected parameters. Then it iteratively adds OPAT until all parameters are covered, i.e. horizontal extension. Upon completion, new test cases may be added to ensure complete interaction coverage, i.e. vertical extension.

In-Parameter-Order (IPO) is the precursor of the OPAT strategy developed by Lei and Tai [14]. Because IPO is limited to pairwise interaction [15], IPOG was developed as the general version of IPO to support higher order interactions. As far as interaction support is concerned, IPOG addresses the uniform-strength, as well as the variable-strength, interaction. No support is provided for the input-output-based relationships.

A number of IPOG variants exist in the literature, including TConfig [16], ParaOrder, and ReqOrder [8]. Similar to IPOG, TConfig adopts variation in the horizontal and vertical extensions as part of its algorithm. In contrast to IPOG, TConfig only addresses the uniform-strength interaction. ParaOrder and ReqOrder differ from their predecessor in terms of how the initial test case is generated [8]. In IPOG, the initial test case is generated in the defined order of parameters found, whereas in ParaOrder, the initial test case is generated based on the first defined input-output relationship. In ReqOrder, the selection of the initial test case does not necessarily follow the first defined input-output relationship. Additionally, in contrast to IPOG, ParaOrder and ReqOrder address uniform strength, variable strength, and input-output-based relationships.

In contrast to the OPAT strategy, the OTAT strategy greedily generates one complete test case into the final test suite per iteration until all tuples are covered. Based on the main approaches of each strategy, the OTAT strategy can be further characterised into three categories: artificial intelligence (AI)-based, iterative-based, and heuristic-based strategies.

The AI-based OTAT strategy adopts an AI technique. Simulated annealing (SA) [17], ant colony-based strategy (ACS) [18], and variable-strength particle swarm optimisation (VS-PSTG) [19] are some of the AI-based techniques adopted in generating interaction test suite.

Concerning SA, the strategy is based on the annealing process, i.e. maximising the crystal size of the material via heating and slow cooling. Heating excites the atom to move from its initial position to avoid a local minimum of internal energy, whereas slow cooling allows the atom to settle for lower internal energy configurations for better crystal size. Analogous to the physical process, the SA strategy starts with a randomly generated test suite, i.e. initial state, and applies a series of transformations according to a probability equation, which depends heavily on parameter T (the controlling temperature of the simulation to simulate heating and cooling).

In ACS, the candidate test cases are searched by colonies of ants for some possible paths. The path qualities are evaluated in terms of the pheromones which signify convergence. The optimum paths correspond to the best test candidate included in the final test suite. For the VS-PSTG, the search process is inspired by the behaviour of flocks of birds. Internally, the strategy iteratively combines local and global searches to find the best test cases that cover the given interaction tuples. We should note that SA, ACS and VS-PSTG address uniform- and variable-strength interactions.

Regarded as the most popular approach, the iterative-based OTAT strategy often performs systematic iterative search to generate the final test suite. GTWay [1], ITCH [20], Jenny [5], TVG [10], PICT [21], Union [6, 22] and Greedy [7] are few examples of the iterative-based OTAT strategy.

As far as implementation is concerned, GTWay starts by generating all the required interaction tuples using its tuple generation algorithm. Then the strategy iterates all tuples and tries to merge any 'combinable' tuples based on its backtracking algorithm. Although it adopts similar merger algorithm as GTWay does, ITCH relies heavily on its exhaustive search algorithm to find the best combinable tuples. Both GTWay and ITCH address uniform-strength interaction.

With regard to Jenny, PICT and TVG, their implementations can be downloaded from the developer's website. Jenny starts by constructing a test suite that covers one-way interaction first. The strategy then extends the test suite to cover two-way interaction, and the process is repeated until the test suite covers the required t-way interactions. In contrast to Jenny, PICT generates the test suite by selecting one uncovered tuple and iteratively fills the 'don't care' parameters (parameters that do not contribute to the current tuple of interest) with the best found value to cover the most uncovered tuples. TVG adopts three algorithms for test suite generation, namely T-Reduced, Plus-One, or Random Set algorithm. Because of limited literature, how each algorithm (T-Reduced, Plus-One, or Random Sets) works is yet unclear. Based on our experiences with TVG, T-Reduced often produces the most optimal test suite compared with Plus-One and Random Sets. Relative to interaction support, Jenny addresses uniform-strength interaction whereas PICT and TVG support uniform strength, variable strength and input-output based relationships.

Union [6, 22] and Greedy [7] are two related iterative-based OTAT strategies. In the case of Union, partial test cases are first generated based on the defined input-output relationships. Then random values are assigned to all parameters that do not contribute to the defined input-output relationships to complete the test cases. Upon completion, union operations are performed for all test cases to remove repetition. Based on the Union strategy, the Greedy strategy also works in the same manner. In contrast to the Union strategy, however, the Greedy strategy completes the partial test cases greedily to cover the most uncovered interactions. In this manner, the Greedy strategy often generates a more optimal test size than Union does. Both Union and Greedy address uniform-strength interaction, variable-strength interaction, and input-output-based relationships.

The last category of the OTAT strategy is the heuristic-based strategy. The heuristic-based OTAT strategy typically uses some form of heuristic models to decide on the test case selection. Bryce's Density strategy [23, 24] and Wang's Density strategy [9] are examples of heuristic-based OTAT strategy. Bryce's Density strategy pioneers the use of density calculation model [23, 24] in constructing the test suite. For each test case, the 'parameter density' of every unassigned parameter is calculated and the parameters with the highest value are selected. Then, the 'value density' that corresponds to the selected parameters is calculated, and the highest value density is fitted in accordingly. This process is repeated OTAT until all parameters have valid value assignments and the complete test suite is formed. In contrast to Bryce's strategy, which addresses only the uniform-strength interaction, the Wang's Density strategy extends its support to the input-output-based relationships. To enable the support, Wang introduced 'local density' and 'global density' calculations. During the test generation process, the Wang's strategy chooses one input-output relationship with the highest local density value. Then for each exhaustive combination of the selected input-output relationship, the strategy selects the combination with the highest global density value to fit into the current test case. The process continues until the complete test suite is formed.

METHODS

The GVS search algorithm is inspired by earlier work in GTWay by Klaib and Zamli et al. [1, 2]. The GVS search algorithm works as follows: in contrast to the GTWay's search algorithm, which generates all tuples before iterating them, GVS generates one tuple at a time before the start of the iteration to minimise memory requirements. After requesting one tuple from the tuple

generator, the search algorithm selects one 'don't care' parameter at a time (indicated by *X*) to establish a value that can produce another uncovered tuples.

Let us consider four 2-valued parameter systems, and the interaction strength required is three. The first uncovered tuples generated are "0", "0", "0" and *X*. Here, the first 'don't care' parameter is parameter X_3, which consists of two values (i.e. "0" and "1"). The search algorithm attempts to fit "0" first into the incomplete test case and check whether uncovered tuples produced by "0" exist. In this case, because the algorithm has just started its search, no tuple is covered yet, i.e. the covered tuple list is empty. Thus, "0" is selected; following the selection of "0" for X_3, {"0", "0", "0", *X*} is produced as one uncovered tuple (for the first input-output relationship), and {"0", *X*, *X*, "0"} is produced as another uncovered tuple (for the second input-output relationship). Then the same process is repeated for the other parameter selections in case more 'don't care' parameters are found. After completing all parameters, the search algorithm checks the generated test case and determines whether the generated test case has covered the most uncovered tuples. If it does, the generated test case is selected in the final test suite list, and the covered tuples are added to the covered tuple list. This process is repeated until all tuples are covered by test cases in the final test suite. The GVS search algorithm is further summarised in Figure 3.

```
Output: Final Test Suite, T

Begin:
Initialise k as the total tuple involve
Initialise Ct as covered tuple list

While (no of tuples in Ct != k)
  P = get next tuple
  if(P not in Ct)
    for every don't care in P
      select value that can produce uncovered tuples
    if(P has the most uncovered tuples)
      store the tuples covered by P in Ct
      store P in T
    if(P still consist don't care)
       replace don't care with the first value of the parameter
    store the tuples covered by P in Ct
    store P in T
End
```

Figure 3. GVS search algorithm

RESULTS AND DISCUSSION

The GVS evaluation was divided into three parts. In the first part, the performance of GVS (in terms of generated test suite size) against the other competing uniform-strength strategies was compared based on the experimental results [2]. In the second part, the performance of GVS against existing variable-strength strategies was evaluated based on the experimental results [18, 19]. Finally, the experimental results obtained by Wang [8, 9] were adopted to benchmark GVS against the existing input-output-based strategies. In all parts, the generated test suite size, rather than the

execution time for the test generation, was compared because access to all the strategy implementations was not available. Comparing the execution time of each strategy is impossible, even from published results, which provided different running environments. Attempting to do so is counterproductive because the execution time is directly affected by the computer hardware performance, operating system and data structure, as well as language implementation.

For each part, the best test suite size for GVS on a single run was reported. GVS is a deterministic strategy, viz. multiple runs always produce identical test suite. Hence, no change occurred as far as the test size is concerned. The running environment consisted of a desktop PC with Windows XP, 2.8 GHz Core 2 Duo CPU, and 1 GB RAM. The GVS strategy was coded and implemented in Java (JDK 1.6). The results are presented in Tables 2-8. The darkened cells indicate the best obtained result for the configuration of interests. Cells marked as NA indicate that the results are not available in the publications.

GVS as Uniform Strength t-Way Strategy

Based on the benchmarking experiments [2], four groups of experiments were conducted and each group has the following system configurations:

i. Group 1: The number of parameters (P) and the value (V) were constant (10 and 5 respectively), but the interaction strength (*t*) varied from two to six.
ii. Group 2: The interaction strength (*t*) and the value (V) were constant (4 and 5 respectively), but the number of parameters (P) varied from 5 to 15.
iii. Group 3: The number of parameters (P) and the interaction strength (*t*) were constant (10 and 4 respectively), whereas the value (V) varied from 2 to 10.
iv. Group 4: The common traffic and collision avoidance system (TCAS), which consisted of 12 multi-valued parameters (two 10-valued parameters, one 4-valued parameter, two 3-valued parameters, and seven 2-valued parameters) and interaction strength (*t*) varied from 2 to the exhaustive testing (i.e. 12-way testing).

The results for Groups 1-4 are shown in Tables 2-5 respectively. The results for the other strategies are obtained from Zamli et al. [2].

Table 2. Generated Test Size For $CA(N, t, 5^{10})$

t	*N*						
	IPOG	**ITCH**	**Jenny**	**TConfig**	**TVG**	**GTWay**	**GVS**
2	48	45	45	48	50	46	44
3	308	225	290	312	342	293	288
4	1843	1750	1719	1878	1971	1714	1701
5	10119	NA	9437	NA	NA	9487	9237
6	50920	NA	NA	NA	NA	44884	45732

Table 3. Generated Test Size For $CA(N, 4, 5^P)$

P	*N*						
	IPOG	**ITCH**	**Jenny**	**TConfig**	**TVG**	**GTWay**	**GVS**
5	784	625	837	773	849	731	733
6	1064	625	1074	1092	1128	1027	1012
7	1290	1750	1248	1320	1384	1216	1215
8	1491	1750	1424	1532	1595	1443	1398
9	1677	1750	1578	1724	1795	1579	1556
10	1843	1750	1719	1878	1971	1714	1701
11	1990	1750	1839	2038	2122	1852	1837
12	2132	1750	1964	NA	2268	2022	1955
13	2254	NA	2072	NA	2398	2116	2088
14	2378	NA	2169	NA	NA	2222	2193
15	2497	NA	2277	NA	NA	2332	2294

Table 4. Generated Test Size For $CA(N, 4, V^{10})$

V	*N*						
	IPOG	**ITCH**	**Jenny**	**TConfig**	**TVG**	**GTWay**	**GVS**
2	46	58	39	45	40	46	45
3	229	336	221	235	228	224	217
4	649	704	703	718	782	621	688
5	1843	1750	1719	1878	1971	1714	1701
6	3808	NA	3519	NA	4159	3514	3502
7	7061	NA	6482	NA	7854	6459	6405
8	11993	NA	11021	NA	NA	10850	8263
9	19098	NA	17527	NA	NA	17272	17188
10	28985	NA	26624	NA	NA	26121	25927

Table 5. Generated Test Size For TCAS Module, $CA(N, t, 10^2 4^1 3^2 2^7)$

t	*N*						
	IPOG	**ITCH**	**Jenny**	**TConfig**	**TVG**	**GTWay**	**GVS**
2	100	120	108	108	101	100	100
3	400	2388	412	472	434	402	404
4	1361	1484	1536	1476	1599	1429	1302
5	4219	NA	4580	NA	4773	4286	4255
6	10919	NA	11625	NA	NA	11727	10530
7	NA	NA	27630	NA	NA	27119	28760
8	NA	NA	58865	NA	NA	58584	59477
9	NA	NA	NA	NA	NA	114411	119040
10	NA	NA	NA	NA	NA	201728	206000
11	NA	NA	NA	NA	NA	230400	230400
12	NA	NA	NA	NA	NA	460800	460800

GVS produces the best test size for the 2-, 4- and 5-way interactions. ITCH produces the best test size for the 3-way interaction, and GTWay produces the best test size for the 6-way interaction (Table 2). GVS also produces the best test size for the 7-, 8-, 9- and 10-parameter systems. ITCH produces the best test size for the 5-, 6-, 11- and 12-parameter systems, whereas Jenny produces the best test size for the rest of the cases (Table 3). GVS outperforms all other strategies in most cases except for the system with two and four values, where Jenny and GTWay outperform all other strategies (Table 4). For the TCAS system in Table 5, GTWay outperforms the other strategies in almost all cases, i.e. for 2-, 7-, 8-, 9-, 10-, 11- and 12-way interactions. GVS and IPOG obtain the same test size as that of GTWay for the 2-way interaction. In the 4-, 6- and 11-way interactions, GVS outperforms all other strategies. Similarly, IPOG outperforms all other strategies for the 3- and 5-way interactions.

As far as the algorithmic complexity analysis of GVS is concerned, the test size grows exponentially with the interaction strength (*t*) (Tables 2 and 5). Additionally, the test suite grows logarithmically with the number of parameters (P) and quadratically with the number of values (Tables 5 and 6). Theoretically, these results are consistent with those in the existing literature with O (v^t log p) [25].

GVS as Variable Strength t-Way Strategy

The benchmark experiments using the test size results were adopted from Chen et al. [18] and Ahmed and Zamli [19]. Three basic system configurations are defined as follows:

i. Fifteen 3-valued parameter systems: $VCA(N, 2,3^{15},\{C\})$,
ii. Three 4-valued parameter, three 5-valued parameter, and two 6-valued parameter systems: $VCA(N, 2,4^3 5^3 6^2,\{C\})$,
iii. Twenty 3-valued parameter and two 10-valued parameter systems: $VCA(N, 2,3^{20}10^2,\{C\})$

The generated test size for GVS is shown in Table 6, along with the other existing variable-strength strategies.

Table 6. Generated test size for different variable strength t-way strategies

{C}	*N*							
	SA	**Density**	**Para Order**	**PICT**	**TVG**	**ACS**	**VS-PSTG**	**GVS**
				VCA(*N*, 2,3^{15},{*C*})				
ϕ	16	21	33	35	22	19	19	19
CA(3,3^3)	27	28	27	81	27	27	27	27
CA(3,3^3)^2	27	28	33	729	30	27	27	27
CA(3,3^3)^3	27	28	33	785	30	27	27	27
CA(3,3^4)	27	32	27	105	35	27	30	28
CA(3,3^5)	33	40	45	121	41	38	38	39
CA(4,3^4)	NA	NA	NA	245	81	NA	81	81
CA(4,3^5)	NA	NA	NA	301	103	NA	97	99
CA(4,3^7)	NA	NA	NA	505	168	NA	158	157
CA(5,3^5)	NA	NA	NA	730	243	NA	243	243
CA(5,3^7)	NA	NA	NA	1356	462	NA	441	445
CA(6,3^6)	NA	NA	NA	2187	729	NA	729	729
CA(6,3^7)	NA	NA	NA	3045	1028	NA	966	947
CA(3,3^4) *CA*(3,3^5) *CA*(3,3^6)	34	46	44	1376	53	40	45	42
CA(3,3^6)	34	46	49	146	48	45	45	45
CA(3,3^7)	41	53	54	154	54	48	49	48
CA(3,3^9)	50	60	62	177	62	57	57	58
CA(3,3^{15})	67	70	82	83	81	76	74	75
				VCA(*N*, 2,$4^3$$5^3$$6^2$,{*C*})				
ϕ	36	41	49	43	44	41	42	40
CA(3,4^3)	64	64	64	384	67	64	64	64
CA(3, $4^3$$5^2$)	100	131	141	781	132	104	124	127
CA(3,5^3)	125	125	126	750	125	125	125	125
CA(4, $4^3$$5^1$)	NA	NA	NA	1920	320	NA	320	320
CA(5, $4^3$$5^2$)	NA	NA	NA	9600	1600	NA	1600	1600
CA(3,4^3) *CA*(3,5^3)	125	125	129	8000	125	125	125	125
CA(4, $4^3$$5^1$) *CA*(4, $5^2$$6^2$)	NA	NA	NA	288000	900	NA	900	900
CA(3,4^3) *CA*(4, $5^3$$6^1$)	NA	NA	NA	48000	750	NA	750	750
CA(3,4^3) *CA*(5, $5^3$$6^2$)	NA	NA	NA	288000	4500	NA	4500	4500
CA(4, $4^3$$5^2$)	NA	NA	NA	2874	496	NA	472	463
CA(5, $4^3$$5^3$)	NA	NA	NA	15048	2592	NA	2430	2380
CA(3, $4^3$$5^3$$6^1$)	171	207	247	1266	237	201	206	202
CA(3, $5^1$$6^2$)	180	180	180	900	180	180	180	180
CA(3, $4^3$$5^3$$6^2$)	214	256	307	261	302	255	260	237
				VCA(*N*, 2,$3^{20}$$10^2$,{*C*})				
Φ	100	100	100	100	101	100	102	100
CA(3,3^{20})	100	100	103	940	103	100	105	102
CA(3,$3^{20}$$10^2$)	304	401	442	423	423	396	481	413
CA(4,$3^3$$10^1$)	NA	NA	NA	810	270	NA	270	270
CA(5,$3^3$$10^2$)	NA	NA	NA	NA	2700	NA	2700	2700
CA(6,$3^4$$10^2$)	NA	NA	NA	NA	8100	NA	8100	8100

Table 6 shows that SA produces the best test size in all system configurations with low interaction strength ($t \leq 3$). For high interaction strength ($3 < t \leq 6$), GVS, VS-PSTG and TVG

regularly outperform all other strategies in most configurations. ACS, Density, and ParaOrder also display competitive results: in some cases, some of their results also match the best test size. PICT produces the worst overall results.

GVS as Input-Output Based Relations Strategy

Two experiments were adopted, which involved 60 input-output relationships for the 10-parameter system taken from Wang et al. [8, 9]. The input-output relationship definitions for both experiments are *Rel* = {{1, 2, 7, 8}, {0, 1, 2, 9}, {4, 5, 7, 8}, {0, 1, 3, 9}, {0, 3, 8}, {6, 7, 8}, {4, 9}, {1, 3, 4}, {0, 2, 6, 7}, {4, 6}, {2, 3, 4, 8}, {2, 3, 5}, {5, 6}, {0, 6, 8}, {8, 9}, {0, 5}, {1, 3, 5, 9}, {1, 6, 7, 9}, {0, 4}, {0, 2, 3}, {1, 3, 6, 9}, {2, 4, 7, 8}, {0, 2, 6, 9}, {0, 1, 7, 8}, {0, 3, 7, 9}, {3, 4, 7, 8}, {1, 5, 7, 9}, {1, 3, 6, 8}, {1, 2, 5}, {3, 4, 5, 7}, {0, 2, 7, 9}, {1, 2, 3}, {1, 2, 6}, {2, 5, 9}, {3, 6, 7}, {1, 2, 4, 7}, {2, 5, 8}, {0, 1, 6, 7}, {3, 5, 8}, {0, 1, 2, 8}, {2, 3, 9}, {1, 5, 8}, {1, 3, 5, 7}, {0, 1, 2, 7}, {2, 4, 5, 7}, {1, 4, 5}, {0, 1, 7, 9}, {0, 1, 3, 6}, {1, 4, 8}, {3, 5, 7, 9}, {0, 6, 7, 9}, {2, 6, 7, 9}, {2, 6, 8}, {2, 3, 6}, {1, 3, 7, 9}, {2, 3, 7}, {0, 2, 7, 8}, {0, 1, 6, 9}, {1, 3, 7, 8}, {0, 1, 3, 7}}.

For the first experiment, a system with ten 3-valued parameters was adopted. The experiment started with |*Rel*| = 10, viz. only the first ten relationships in *Rel* were used in generating the test suite. Subsequently, the first 20 relationships in *Rel* were used, until all 60 relationships were finally used. The test size obtained from the first experiment is shown in Table 7.

The second experiment involved the same relationships but using a system with multi-value parameters consisting of three 2-valued parameters, three 3-valued parameters, three 4-valued parameters, and one 5-value parameter. The result obtained from the second experiment is shown in Table 8.

Table 7. Generated Test suite Size for $IOR\{N, 3^{10}, |Rel|\}$

|*Rel*|	*N*						
	Density	**ReqOrder**	**ParaOrder**	**Union**	**Greedy**	**TVG**	**GVS**
10	86	153	105	503	104	86	104
20	95	148	103	858	110	105	98
30	116	151	117	1599	122	125	116
40	126	160	120	2057	134	135	117
50	135	169	148	2635	138	139	127
60	144	176	142	3257	143	150	140

Table 8. Generated Test suite Size for $IOR\{N, 2^3 3^3 4^3 5^1, |Rel|\}$

|*Rel*|	*N*						
	Density	**ReqOrder**	**ParaOrder**	**Union**	**Greedy**	**TVG**	**GVS**
10	144	154	144	505	137	144	144
20	160	187	161	929	158	161	162
30	165	207	179	1861	181	179	169
40	165	203	183	2244	183	181	170
50	182	251	200	2820	198	194	200
60	197	250	204	3587	207	209	200

For |*Rel*| = 10, Density and TVG produce the best test size. For |*Rel*|= 20, Density produces the best result. For |*Rel*| = 30, both Density and GVS produce the best test size. Concerning |*Rel*| = 40 until |*Rel*| = 60, GVS produces the most optimum result (Table 7). We note that in all cases, Union produces the worst result.

In Table 8, Greedy produces the best test size for |*Rel*| = 10 and |*Rel*| = 20, whereas Density produces the best test size for all other configurations. Although it does not produce the best result, GVS nevertheless produces acceptable results in all cases, i.e. second to Density in almost all values of *R*. Union produces the worst result for all configurations.

CONCLUSIONS

A new variable strength t-way test suite-generation strategy called GVS, which is based on input-output relationships, has been proposed and evaluated. The evaluation was encouraging because GVS produces good results for uniform number of parameter values and system with high interaction strength (i.e. $t > 3$). As an area for further research, we are investigating new searching algorithm for integration into GVS to produce better test size, especially where the parameter values are non-uniform. Additionally, we are also considering automating the process of determining the input-output relationships among system parameters.

REFERENCES

1. M. F. J. Klaib, "Development of an automated test data generation and execution strategy using combinatorial approach", *PhD Thesis*, **2009**, Universiti Sains Malaysia, Malaysia.
2. K. Z. Zamli, M. F. J. Klaib, M. I. Younis, N. A. M. Isa and R. Abdullah, "Design and implementation of a t-way test data generation strategy with automated execution tool support", *Inform. Sci.*, **2011**, *181*, 1741-1758.
3. Y. Lei, R. Kacker, D. R. Kuhn, V. Okun and J. Lawrence, "IPOG/IPOG-D: Efficient test generation for multi-way combinatorial testing", *Softw. Test. Verif. Reliab.,* **2008**, *18*, 125-148.
4. A. Hartman, T. Klinger and L. Raski, "IBM intelligent test case handler", http://www.alphaworks.ibm.com/tech/whitch, **2005,** (Accessed: March 2011).
5. B. Jenkins, "Jenny test tool", http://www.burtleburtle.net/bob/math/jenny.html, **2010**, (Accessed: April 2011).
6. P. J. Schroeder and B. Korel, "Black-box test reduction using input-output analysis", Proceedings of 2000 ACM SIGSOFT International Symposium on Software Testing and Analysis, **2000**, Portland, USA., pp. 173-177.
7. P. J. Schroeder, P. Faherty and B. Korel, "Generating expected results for automated black-box testing", Proceedings of 17th IEEE International Conference on Automated Software Engineering, **2002**, Edinburgh, UK, pp. 139-148.
8. Z. Wang, C. Nie and B. Xu, "Generating combinatorial test suite for interaction relationship", Proceedings of 4th International Workshop on Software Quality Assurance, **2007**, Dubrovnik, Croatia, pp. 55-61.
9. Z. Wang, B. Xu and C. Nie, "Greedy heuristic algorithms to generate variable strength combinatorial test suite", Proceedings of 8th International Conference on Quality Software, **2008**, Oxford, UK, pp. 155-160.
10. J. Arshem, "Test vector generator (TVG)", http://sourceforge.net/projects/tvg, **2010**, (Accessed: March 2011).
11. D. M. Cohen, S. R. Dalal, M. L. Fredman and G. C. Patton, "The AETG System: An approach to testing based on combinatorial design", *IEEE Trans. Software Eng.,* **1997**, *23*, 437-444.

12. C. Nie and H. Leung, "A survey of combinatorial testing", *ACM Comput. Surv.,* **2011**, *43*, Article 11.
13. K. Z. Zamli, M. I. Younis, S. A. C. Abdullah and Z. H. C. Soh, "Software Testing", 1st Ed., Open University Malaysia, Kuala Lumpur, **2008**.
14. Y. Lei and K. C. Tai, "In-parameter-order: A test generation strategy for pairwise testing", Proceedings of 3rd IEEE International Conference on High Assurance Systems Engineering Symposium, **1998**, Washington DC, USA, pp. 254-261.
15. Y. Lei, R. Kacker, D. R. Kuhn, V. Okun and J. Lawrence, "IPOG: a general strategy for t-way software testing", Proceedings of 14th Annual IEEE International Conference and Workshops on The Engineering of Computer-Based Systems, **2007**, Tucson, USA, pp. 549-556.
16. A. W. Williams, "*TConfig*", http://www.site.uottawa.ca/~awilliam/, **2010**, (Accessed: Macrch 2011).
17. M. B. Cohen, P. B. Gibbons, W. B. Mugridge, C. J. Colbourn and J. S. Collofello, "Variable strength interaction testing of components", Proceedings of 27th Annual International Computer Software and Applications Conference, **2003**, Dallas, USA, pp. 413-418.
18. X. Chen, Q. Gu, A. Li and D. Chen, "Variable strength interaction testing with an ant colony system approach", Proceedings of 16th Asia-Pacific Software Engineering Conference, **2009**, Penang, Malaysia, pp. 160-167.
19. B. S. Ahmed and K. Z. Zamli, "A variable strength interaction test suites generation strategy using particle swarm optimization", *J. Syst. Softw.,* **2011**, *84*, 2171-2185.
20. A. Hartman, T. Klinger and L. Raski, "IBM Intelligent Test Case Handler", http://www.alphaworks.ibm.com/tech/whitch, **2010**, (Accessed: January 2011).
21. J. Czerwonka, "Pairwise testing in real world", Proceedings of 24th Pacific Northwest Software Quality Conference, **2006**, Portland, USA, pp. 419-430.
22. P. J. Schroeder, "Black-box test reduction using input-output analysis", *PhD Thesis*, **2001**, Illinois Institute of Technology, Chicago, USA.
23. R. C. Bryce and C. J. Colbourn, "A density-based greedy algorithm for higher strength covering arrays", *Softw. Test. Verif. Reliab.*, **2009**, *19*, 37-53.
24. R. C. Bryce and C. J. Colbourn, "The density algorithm for pairwise interaction testing", *Softw. Test. Verif. Reliab.*, **2007**, *17*, 159-182.
25. M. B. Cohen, "Designing test suites for software interaction testing", *PhD Thesis*, **2004**, University of Auckland, New Zealand.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 430-448

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Review

A critical assessment and new research directions of rice husk silica processing methods and properties

Iyenagbe B. Ugheoke [*] **and Othman Mamat**

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

[*]Corresponding author, email: ugheokeb@gmail.com

Received: 2 November 2011 / Accepted: 7 November 2012 / Published: 8 November 2012

Abstract: A review of production and properties of silica produced from rice husk is carried out with assessment of processing and manufacturing methods and suggestion of new research directions with respect to the processing methods. It was revealed that the structural nature of the silica produced from rice husk is independent of the purification methods but largely dependent on the incineration temperature used in the production process. Also, it was established that without pre-treatment, incineration of rice husk results in the production of silica of low purity, surface area and whiteness. The paper concludes by advocating the use of a novel process called hydro thermo-baric process, for producing high-purity reactive nano-silica from rice husk. This novel process, has the advantages of high volume production, versatility of application of its product through varying any of its process parameters and being environmentally benign compared to other processes.

Keywords: rice husk silica, opaline silica, nanosilica, thermo-baric process

INTRODUCTION

It was estimated that world paddy production in 2010 was over 700 million tons [1]. Around 22% of the paddy mass is husk [2], which consists of opaline silica in combination with a large amount of a phenyl propanoid structural polymer called lignin, hemicellulose and cellulose [3]. Consequently, over 154 million tons of rice husk were generated in 2010. The average composition of rice husk is given in Table 1.

Table 1. Composition of rice husk [4]

Constituent	%
Cellulose	35
Hemicellulose	25
Lignin	20
Crude protein (N x 6.25)	3
Ash	17 (silica 94%)

When rice husk is incinerated, it generates between 17-20% ash, made up of about 87-93% opaline silica and other metallic oxide impurities depending on the source of the husk. This kind of high percentage of silica intermingling with plant fibres is quite rare in nature. The close intermingling of silica and lignin has two consequences: the rice hull is not only made resistant to water penetration and fungal decomposition, it is also resistant to the best efforts of man to dispose it. Since the hull represents an average about 22% of the rough harvested weight of rice (paddy), our planet is speedily being filled up with an abundance of this scaly residue. Juliano [5] has also pointed to the fact that of all cereal by-products, the rice hull has the lowest percentage of total digestible nutrients (less than10%). Also reporting a drawback of rice husk usefulness, Olivier [3] stated: "Nowhere could we ever find a cereal by-product so low in protein and available carbohydrates and yet, at the same time, so high in crude fibre, crude ash and silica." This statement corroborates the fact that rice husk is not even good as a fodder for animals. Though some organic products such as furfural have been produced from rice husk [6], its richness in ash and silica has been of great research interest because many value-added products could be obtained from this agro-waste.

When rice husk is incinerated, two types of products can result, depending on whether the combustion is complete or incomplete. They are white rice husk ash (WRHA) and carbonised rice husk (CRH) respectively. In either forms, rice husk has some applications. Among the uses reported are pozzolan in cement industry [7-13], aerogels [14, 15], SiC [16, 17], porous carbon [18], zeolites [19] and cordierite [20]. In earlier work, we reported an indirect use of rice husk in making insulating refractories [21, 22]. These uses are related to the processing route or method used, which often affect the product characteristics. For example, it has been shown that different methods of preparation of rice husk silica produce different morphology, structure and reactivity [23].

Previous reviews on this subject were in 2001 and 2003 [24, 25] and since then there has not been another review again even though there have been dramatic developments in the technology. This paper is therefore an attempt to review the research depth and directions of the processing methods as well as the characteristics and utilisation of rice husk silica in these intervening years. The paper presents a critical review of the literature by assessing the current production facilities and processes in the field as well as the influence each process has on the product characteristics. It ends by making case for a process currently being developed, which is called the hydro thermo-baric process. It is our hypothesis that this process would be inexpensive and versatile with regard to the effects it would have on the quality of rice husk silica produced by varying the process parameters.

Composition of Rice Husk Ash

There are many reports [26-28] on the varying composition of rice husk ash (Table 2) which is largely dependent on many factors. These include but are not limited to agricultural practices such as the use of fertilisers during rice cultivation, type of fertilisers employed and climatic or geographical factors. Thus, depending on the geographical location, different authors have published different values for the composition of rice husk. The methods used to evaluate the composition quoted are also indicated.

From Table 2, it is clear that though the composition of rice husk may be dependent on several factors, the percentage of silica (SiO_2) in the ash ranges between 87-92%. A trace amount of titania (TiO_2) was also reported [29]. Due to this compositional variation, the suitability of rice husk silica for any given application will therefore be largely dependent on the purity level needed and if it is not met from the beginning for a given application, then a purification process will be carried out to reduce or eliminate unwanted impurities.

Table 2. Composition of rice husk ash by geographical location (Percentage of constituent as per authors in square bracket and method used outside bracket)

Composition	[27]XRF	[28]ICP	[29]*	[**]XRF
SiO_2	91.56	91.5	87.79	91.25
K_2O	4.76	1.23	1.69	3.829
P_2O_5	-	0.30	4.74	2.45
CaO	0.78	0.57	1.24	0.875
SO_3	0.29	-	-	0.661
MgO	-	0.30	1.59	0.573
Al_2O_3	2.36	0.62	0.4	0.18
Fe_2O_3	0.11	0.42	0.37	0.0866
MnO	0.07	0.04	-	0.0726
Rb_2O	-	-	-	0.0143
ZnO	0.01	-	-	0.0111
CuO	0.01	-	-	-
Na_2O	-	0.18	-	-
LOI	N.D.	3.05	2.08	N.D.

* Method not specified; ** From the authors' laboratory; N.D. = Not determined

PROCESSING OF RICE HUSK FOR SILICA

Different processes have been used by different researchers to obtain silica from rice husk. The following discussion reviews these processes, highlighting their advantages and disadvantages.

Direct Incineration without Pre-treatments

Rice husk is directly incinerated to produce silica of varying purity, with or without the use of pre-treatments [30-36]. In an overall process, the temperature of incineration, holding time and pre-treatment techniques employed affect the character, especially the surface area and brightness (whiteness), of the silica produced. The transformation of raw husk to clear white, grey or pale grey ash is critically dependent on the temperature of incineration [31]. For instance, a temperature between 300-450^{o}C only transforms fresh rice husk (Figure 1(a)) to carbonised husk (Figure 1(b)), while that between 500-650^{o}C produces white or grey ash (Figure 1(c)), depending on soaking time, which is the duration for which incineration is allowed to proceed at the stated temperature range.

(a) (b) (c)

Figure 1. Rice husk (a); carbonised rice husk (b); completely incinerated rice husk (c)

As the incineration temperature increases, there appears to be some accompanying phase changes. The findings of these phase changes will be discussed later. It is noted, however, that rice husk silica produced between 500-650^{o}C with incineration holding (soaking) time of 2.5-6 hr is considered ideal for producing white amorphous silica while crystallinity sets in when incineration temperature increases beyond 700^{o}C. The quantity of the operational phase, whether cristobalite or tridymite, is dependent on the applied temperature range and the impurity level in the rice husk. Also, it was reported [31] that the incineration temperature grossly affects the surface area and hence the reactivity of silica produced from direct incineration process.

Direct incineration of rice husk can be accomplished in open air as reported by Hamdan et al. [23] or in a muffle furnace [37]. Another method used by some researchers [34-36] is the fluidised bed combustion technique to produce rice husk silica, even though the reported purity was not more than 95%. So whether in static or flowing air, complete incineration can be achieved with some varied effects on the properties of the silica produced. A new technology was recently developed in India (Figure 2) for direct incineration of rice husk [38] and it works like a TORBED reactor [39].

Figure 2. A pilot plant for rice husk incineration in India [38]

Pre-treatment Effects on Silica Production from Rice Husk

One of the reasons why it is difficult to obtain silica with purity in excess of 97% from rice husk by the direct incineration process is a consequence of the effects of the metallic impurities the husk contains. For instance, Chandrasekhar et al. [40] reported that oxides, especially K_2O, impart black colour on the particles. Some explanations to support this phenomenon is that there exists a strong interaction between oxides, especially those of potassium and sodium, contained in rice husk and the silica therein, such that it can result in the surface melting of SiO_2 particles and accelerate an early crystallisation of amorphous SiO_2 into cristobalite, as implied by research results [30, 41-43]. This is one of the reasons why Kalapathy et al. [44] could not achieve a purity of up to 98% even after 14 hr of their sol-gel treatment of rice husk ash with bases and acids. The surface melting of these oxides on the silica grossly reduces the surface area, thereby reducing the reactivity of the particles. Thus, it is often necessary to use some pre-treatment methods, which can either be done through acidic or basic medium, to reduce or remove metallic impurities in order to increase the chances of obtaining silica of higher purity and surface area than is achievable in the direct incineration method.

Three main pre-treatment methods have generally been used in the production of high purity silica from rice husk. These are acid leaching, basic pre-treatment and microbiological pre-treatment, usually in combination with some acids. Several kinds of acids, both mineral and organic, have been reported to be used to pre-treat rice husk before other value adding processes such as incineration begin [32, 33, 40, 42, 43, 45]. However, HCl has proved to be most effective in removing metallic impurities from the husk and so it is by far the most widely used. Chakraverty et al. [46, 47] found that leaching of rice husk in 1N HCl is effective in removing most of the metallic impurities. Other researchers [40] used organic acids and compared results with those obtained by using other different mineral acids in pre-treating rice husk and concluded that HCl is better. Their order of efficiency is HCl> H_2SO_4>

HNO_3. However, Umeda and Kondo [48] reported a very high purity (>99.5%) silica from rice husk on pre-leaching it with citric acid.

While acid leaching affects the chemical composition of the husk, it does not affect the structure, whether crystalline or amorphous, of the silica. Thus, the change of phase from amorphous to crystalline is not affected by the pre-treatment method employed. An insight into this dynamics was presented in the research report of Real and coworkers [43]. They found that preliminary leaching of rice husk with a solution of HCl before incineration at 600°C, if properly done, can result in a high-purity silica (approximately 99.5%) with high specific surface area (approximately 260 m^2/g). They indicated that the high-surface-area silica produced was unaffected even after being heated at 800°C. They also performed the HCl leaching on the white ashes obtained from incineration of untreated rice husk at 600°C and obtained an amorphous silica with the same purity, although its specific surface area decreased to as low as 1 m^2/g. They explained the kinetics of this drastic change in surface area, attributing it to the interaction between alkali oxides, specifically K_2O and SiO_2.

Other acids such as H_2SO_4, HNO_3 and their mixture, have also been used in the acid pre-treatment [32, 42, 46, 49-51]. The general leaching effects of H_2SO_4, HNO_3 and HCl are similar, but HCl is superior to H_2SO_4 and HNO_3 in removing the metallic ingredients [46]. Some researchers [42, 44] also attempted chemical post-treatment of incinerated rice husk using HCl but the results were inferior to those of the pre-treatment.

Some alkalis such as NaOH and NH_4OH have been used to pre-treat rice husk [33, 37, 49, 50]. However, the effects of alkali pre-treatment were not as obvious or satisfactory as those of acid pre-treatment. Attempts have been made on the use of microbial fermentation as a pre-treatment of rice husk in order to obtain silica [49, 52]. Although the results were similar to those obtained from acid pre-treatment, the method is disadvantageous in that the time required for the fermentation process to complete is too long, making it unfeasible for practical applications.

Hydrothermal Method

Hydrothermal synthesis has been defined as a process that utilises single- or heterogeneous-phase reactions in aqueous media at elevated temperature (T>25 ^oC) and pressure (P>100 kPa) to crystallise ceramic materials directly from solution [53]. As stated in the introduction, rice husk contains organic compounds and oxides of metals. Under high temperature, high pressure and acidic or basic medium with strong oxidising activity, the organic compounds are decomposed and the trace metals turned into soluble ions; then, silica is obtained. This processing method can achieve the purification of silica from the husk with only the use of water. However, achieving complete dissolution of the organic matter in the rice husk is a task that is near impossible. So, practically this process still requires an incineration step, though the soaking time may be less compared to incinerating the untreated or pre-treated rice husk. The method does not affect the amorphicity of the silica in rice husk. Some acids with strong oxidising activity such as H_2SO_4 and HNO_3 are used and sometimes H_2O_2 is also used as the oxidative medium. This method had been used by Mochidzuki et al. [26] and Wu [54].

Other Methods

Some researchers [55, 56] reacted carbonised rice husk with Na_2CO_3 solution in a proper ratio for 3 hr, followed by incineration step at 600-650^oC for varied soaking time between 3-7 hr to obtain silica. The silica made from this method has good reinforcing properties in rubber. In an earlier review [24], it was reported that this method involved the mixing of rice husk ashes with NaOH to produce sodium silicate which would be reacted with NH_4HCO_3, $(NH_4)_2SO_4$ or H_2SO_4 to produce SiO_2.

Summary of Processing Methods

Apart from the production of rice husk silica done with the TORBED incinerator or the fluidised bed combustion, all other methods were mostly done on a laboratory scale. If they were to be scaled up to commercial level, the cost and personnel risks involved would be quite high, since acids and other corrosive media would be worked with at high temperatures. The quality of the silica obtained from fluidised bed combustion or TORBED incinerator is not more than 95%, which restricts its application to chemically insensitive areas like cement and concrete admixtures where high-purity silica is not essential. For this reason, therefore, it is necessary to evolve a system that can produce high-purity silica at volumes of production capable of supporting industrial needs. This is the current research in our laboratory where we have developed a hydro thermo-baric process for volume production of high purity nanosilica for industrial applications. Some of the work in this research will be briefly explained later.

PROPERTIES OF RICE HUSK SILICA

In this section the properties of rice husk silica are examined, as well as how the production process employed affects such properties.

Structure

Different research reports have presented the structural state of rice husk silica and shown that it is dependent on the processing temperature. Often, the structure of the silica is investigated by X-ray diffraction (XRD) and the state is revealed by the shape of the diffractogram obtained. A study done by Hamdan et al. [23] showed that over various temperature ranges rice husk silica can exist in either crystalline or amorphous state as shown in Figure 3.

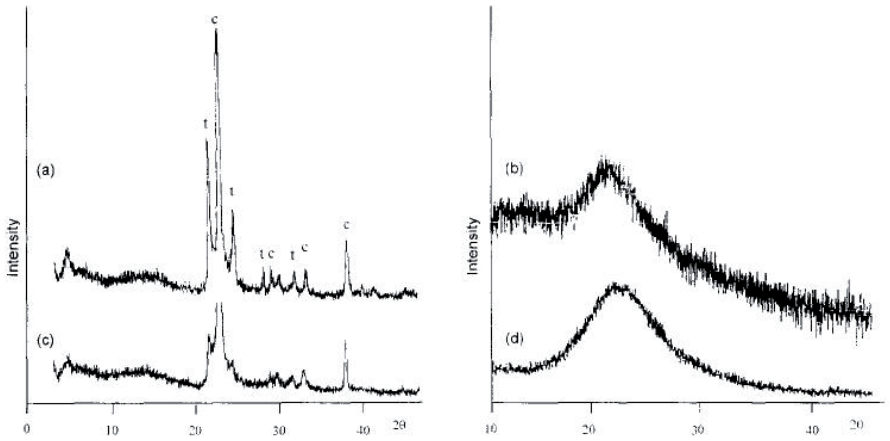


Figure 3. X-ray diffractograms of rice husk silica processed at different temperatures [23]

Figure 3(a) is a diffractogram of rice husk silica processed through open-field burning where temperatures can reach 900°C, while Figure 3(c) represents samples prepared in furnace at 1000°C for 4 hr. Figures 3(b) and 3(d) are from a sample prepared in muffle furnace at 700°C for 4 hr and sample hydrothermally extracted respectively. As is evident from the diffractograms, the samples represented by Figure 3(b) and (d) are amorphous while the other two are crystalline in nature with the formation of crystobalite and traces of tridymite. These findings are in agreement with the work of Kapur [31], who studied the structural behaviour of silica over a temperature range of 400-1500°C and reported that at combustion temperature above 900°C, the silica in rice husk ash consisted of cristobalite and a small amount of tridymite. Other researchers [57] reported similar results. Thus, to obtain amorphous silica from rice husk, the processing temperature should not exceed 700°C, as phase transition to the crystalline structure of crystobalite would soon follow, although no specific temperature has been reported for this transformation.

Surface Area and Pore Volume

The surface area and pore volume of rice husk silica is dependent on the processing temperature, since this affects the surface melting of the silica due to the presence of alkali metal oxides. Different values of surface area and pore volume have been reported [58]. Kapur [31] reported an initial increase in the surface area from 60 m^2/g to 80 m^2/g when the husk was incinerated at 350°C and 600°C respectively. This increase was perhaps due to the burn-off of the residual carbon and the opening of new pores. His work further revealed that between 700-900°C, a sharp drop occurred in the surface area from 40 to only 1 m^2/g as a result of the alkali metal oxides.

Morphology, Particle Size and Chemical Species

It is a generally accepted fact that silica which is formed by incinerating rice husk below 800°C is amorphous [31]. It seems that particles of silica in rice husk ash are agglomerates of small nano-range particles. Thus, it is very common to find aggregates of silica forming fine globules or platelets

of varied sizes as can be seen in Figure 4. The first report on the production of nanosilica from rice husk was made by Conradt et al. [49] and later by Liou [58]. Since then other researchers [59] have always stressed the term in reporting their work. From the investigations in our laboratory, we have found that not much work or processing is required to obtain nanosilica since the natural form in which it occurs in the husk is already a wide distribution of sizes in the nano-range, which is less than 100 nm.

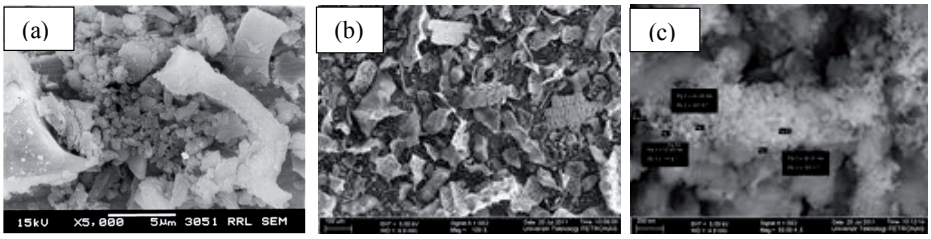


Figure 4. Morphology of rice husk silica (a) SEM micrograph showing mixed platelets and globules [25]; (b) FESEM micrograph of platelets; (c) agglomerated nanoparticles [from our laboratory]

Also studied were the chemical species that exist within the silica produced from rice husk. To demonstrate that the available chemical species in rice husk silica is process dependent, Hamdan et al. [23] and Mochidzuki et al. [26] reported similar data on the species present in rice husk silica using ^{29}Si nuclear magnetic resonance (NMR) when they produced rice husk silica via the hydrothermal process. The ^{29}Si MAS NMR spectra of rice husk silica presented in Figure 5 clearly show the varying chemical species with processing temperature.

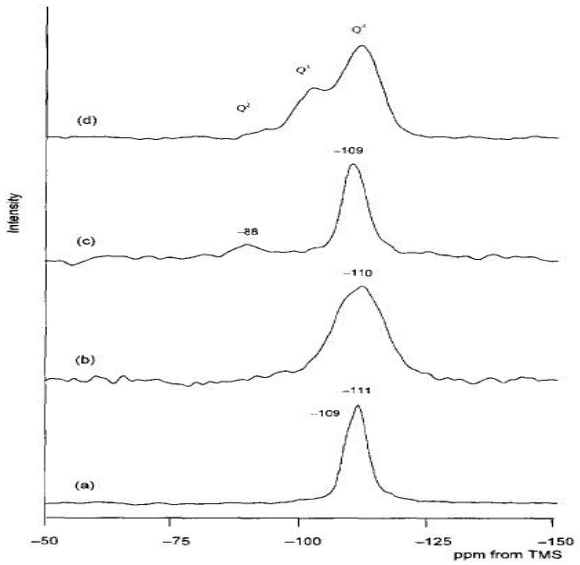


Figure 5. ^{29}Si MAS NMR spectra of rice husk silica samples: (a) open field burning; (b) from muffle furnace at 700°C; (c) from muffle furnace at 1000°C; (d) from hydrothermal processing [23]

A detailed examination of the spectra reveals important features. As explained by Hamdan et al. [23], the spectrum of the sample in Figure 5(a) consists of a single intense and narrow peak at -111 ppm, which corresponds to the presence of Q4 line of siloxane bonds of the crystalline trydimite and a shoulder at -109 ppm corresponding to β-cristobalite. This observation is also true for the sample in Figure 5(c) and is expected since the processing route involved exposure of the rice husk to temperatures in excess of 700°C. The ^{29}Si MAS NMR spectrum of the sample in Figure 5(d) consists of peaks with chemical shifts of around δ 90, 100 and 110 ppm, which correspond to the presence of Q2, Q3 and Q4 units, i.e. silanediol unit [$(OH)_2Si(OSi)_2$], silanol unit [$(OH)Si(OSi)_3$] and silicon-oxygen tetrahedral framework [$Si(OSi)_4$] respectively.

Mochidzuki et al. [27] did a more detailed study of the species change with temperature and remarked that the original rice husk silica consisted of 8.9% Q2, 60.3% Q3 and 30.8% Q4 units and that the percentage of Q4 units existing in the samples tended to increase with increasing treatment temperature, though Q2 and Q3 units remained at non-negligible percentages since the maximum temperature they worked with was 243^{o}C. Hamdan et al. [23] reported that there was a complete absence of Q2 and Q3 units with the Q4 bandwidth increasing when the spectrum for the sample in Figure 5(b) was examined. This suggests that the silicates with Q2 and Q3 structures were converted to Q4 silicates by thermal treatment and since traces of Q2 and Q3 units existed in samples in Figures 5 (a), (c) and (d), the temperature around 700^{o}C is regarded as optimal for the production of high purity silica devoid of silane-diol and silanol units.

Whiteness

One important property worthy of mentioning is the colour of the silica produced at the end of each process. When working with an open-field pile burning of fresh rice husk (Figure 1(a)), there are often spots of incomplete combustion within the pile, resulting in pockets within the pile having particles of carbonised husk as depicted in Figure 1(b). Once this happens, black coloration due to the unburnt carbon remains within the product. Chandrasekhar et al. [20, 40] did extensive studies on the optical properties of rice husk silica processed using different acid pre-treatment methods and concluded that in the processes where there has been substantial leaching of the alkali metal oxides, especially the oxide of potassium, the brightness and whiteness increases. This is due to the fact that when the husk is not substantially treated, the alkali oxides enhance the surface melting of silica, thereby entrapping carbon within the silica; the more residual carbon entrapped, the darker the colour of the husk is.

Guide to Selection of Rice Husk Silica Production Process for Specific Applications

The impurities contained in rice husk silica can have implication in some of the applications. It is therefore important to know what characteristics are needed in an intended application, which will therefore affect the choice of the purification degree that is necessary for such an application. Table 3 gives a summary of applications of rice husk silica. The levels of purity required and the recommended processes are also indicated.

Table 3. Different applications and purity levels required of rice husk silica and recommended processes

Application	Desirable qualities	Authors	Purity level	Recommended process
Filler in polymers	Ability to retard thermo-oxidative and photo degradations. Possession of some silanol group to enhance coupling and good level of residual carbon to inhibit photo degradation	[60-70]	95-98% and possessing Q2 and Q3 groups	Hydrothermal process
Cement and Concrete	High reactivity, high surface area, absence of crystallinity	[7, 9-12, 30, 71-91]	95-98%	Acid pre-treatment before incineration at temperature less than 700^{o}C
Zeolites	High purity and high porosity	[92-105]	> 99.5%	Acid pre-treatment before incineration at temperature less than 700^{o}C in regulated environment
Aerogels	High purity	[14, 15]	> 99.5%	Acid pre-treatment before incineration at temperature less than 700^{o}C in regulated environment
Cordierite	High purity/ reactivity	[20, 106]	> 99.5%	Acid pre-treatment before incineration at temperature less than 700^{o}C
SiC	Complex	[16, 17, 107]	95-98% and possessing Q2 and Q3 groups	Complex

CURRENT RESEARCH AND DEVELOPMENT IN THE PRODUCTION OF RICE HUSK SILICA

The hydro thermo-baric process is currently receiving attention in our laboratory and has a great promise for the production of rice husk silica with varied degree of purity. The process refers to that which utilises single or heterogeneous phase reactions in aqueous media at high temperature (>243^{o}C) and pressure (>3 MPa) to cause leaching or solubilising of oxide impurities as well as degradation of

organic compounds in rice husk. The principle behind the process in which water is the sole aqueous medium is that water can dissociate at high temperature and pressure, forming hydronium ions (or hydrated protons or protonised water: H_3O^+) as well as hydroxyl (OH^-) ions, thus behaving like an acid-base system capable of reacting with basic and acidic oxides. This process can leach away or reduce the metal impurities of rice husk to the levels that are acceptable in several industries utilising silica as raw material. Also, the organic components (hemi-cellulose, cellulose and lignin) of the husk are expected to be converted to sugars and acids or other low molecular weight organic compounds to a large degree. This leaves only a small portion of these cellulosic components in the post-treated rice husk, thus reducing the incineration time, smoke generation, cost and rigour required to obtain silica from the husk.

The equipment and set-up used for this process is shown in Figure 6. The reactor vessel was designed and fabricated specially for this purpose from stainless steel grade 304 and consisted of a 150-ml cavity, in which the purification reactions take place. The cavity was bored in a 6-mm-thick, closed-end cylinder, with the closed end having a thickness of 30 mm. The cylinder cover and flange had the same thickness of 15 mm. In this process, a certain amount of the dry rice husk sample is placed in the reactor cavity with one-third of the vessel filled with deionised water. The vessel is sealed tightly and placed in a furnace with a constant temperature of 300^{o}C to heat up and maintain the hydro thermo-baric conditions. The reactor is then soaked at this temperature for a varied period depending on the level of purity desired. At the end of each soaking period, the non-liquid residue in the reactor is dried and heated in a furnace at 650^{o}C for 2-3 hr to get rid of any remnants of cellulosic materials remaining in the residue. After the incineration procedure, the ash obtained is pulverised in a ceramic mortar to afford nanosilica.

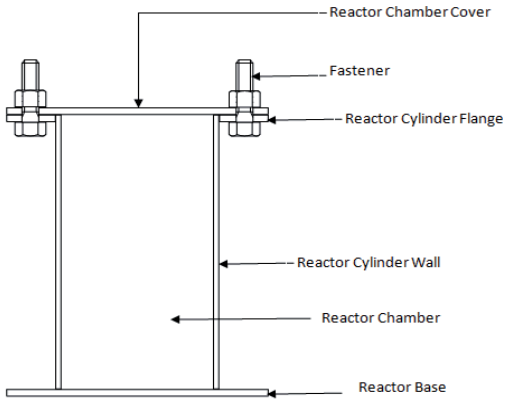


Figure 6. Schematic diagram of a hydro thermo-baric reactor

The advantage of this method is that it can reduce processing time and by controlling process variables, different degrees of purity can be obtained. The process is capable of selective leaching, depending on which process variables are altered. The principal process variables include temperature, soaking time and ratio of husk to water. This process has been optimised in our laboratory and the

purity level of silica produced approaches 99.5%. In addition, the silica obtained is porous with high surface area and can be made to retain residual carbon if so desired.

CONCLUSIONS

A review of the processing methods of rice husk and properties of the silica product obtained has been presented. It is clear from the processing methods that, apart from the TORBED system for the production of silica of purity around 95% which has been commercialised, all other processes still have the problem of scalability from laboratory scale to industrial level. This situation is still the same as in a previous review years ago. It is hoped, however, that the research evolving from our laboratory would be able to bridge this gap and take high-purity silica production from rice husk to a level that can support industrial demands at low cost and low environmental and personnel risks.

ACKNOWLEDGEMENTS

This work was supported in part by the Malaysian Ministry of Science, Technology and Innovation, E-Science fund under the project code 03-02-02-SF0089 and UTP graduate assistantship scheme. The authors appreciate these kind gestures.

REFERENCES

1. Trade and Markets Division, "Food and Agriculture Organization of the United Nations", *FAO Rice Market Monitor. Bull.*, **2010**, *15*, 1, http://www.fao.org/economic/est/publications/rice-publications/rice-market-monitor-rmm/en/.
2. E. C. Beagle, "Rice-husk conversion to energy", *FAO Agric. Serv. Bull*., **1978**, *31*, 8.
3. P. Olivier, "*The rice hull house*", **2004,** http://www.thelaststraw.org/backissues/articles/Rice%20Hull%20House.pdf (Accessed 25/08/2011).
4. National Commission for Schedule Tribe (NCST), "Utilization and recycling of agricultural wastes/by-products: A country report", **1974,** New Delhi, Department of Science and Technology.
5. B. O. Juliano, "Rice: Chemistry and Technology", 2nd Ed., American Association of Cereal Chemists, St Paul, MN (USA), **1985**, pp. 695.
6. V. Punsuron, P. Vaithanomsat and K. Iiyama, "Simultaneous production of α-cellulose and furfural from bagasse by steam explosion pretreatment", *Maejo Int. J. Sci. Technol.,* **2008**, *2,* 182-191.
7. R. M. R. de-Gutierrez and S. A. Delvasto, "Production of high strength cements from rice husk ash", *J. Resour. Manage. Technol*., **1994,** *22*, 127-134.
8. V. T. Nguyen, Y. Guang, V. B. Klaas, A. L. A. Fraaij and D. D. Bui, "The study of using rice husk ash to produce ultra high performance concrete", *Constr. Build. Mater.,* **2011,** *25*, 2030-2035.

9. M. Nehdi, J. Duquette and A. El-Damatty, "Performance of rice husk ash produced using a new technology as a mineral admixture in concrete", *Cem. Con. Res.,* **2003,** *33*, 1203-1210.
10. A. M. Shazim, A. S. Muhammad and A. Hassan, "Utilization of rice husk ash as viscosity modifying agent in self compacting concrete", *Constr. Build. Mater.,* **2011,** *25*, 1044-1048.
11. W. Tangchirapat, R. Buranasing, C. Jaturapitakkul and P. Chindaprasirt, "Influence of rice husk–bark ash on mechanical properties of concrete containing high amount of recycled aggregates", *Constr. Build. Mater.,* **2008,** *22*, 1812-1819.
12. M. F. M. Zain, M. N. Islam, F. Mahmud and M. Jamil, "Production of rice husk ash for use in concrete as a supplementary cementitious material", *Constr. Build. Mater.,* **2011,** *25*, 798-805.
13. M. H. Zhang, R. Lastra and V. M. Malhotra, "Rice-husk ash paste and concrete: Some aspects of hydration and the microstructure of the interfacial zone between the aggregate and paste", *Constr. Build. Mater.,* **1996,** *26*, 963-977.
14. J. P. Nayak and J. Bera, "Preparation of silica aerogel by ambient pressure drying process using rice husk ash as raw material", *Trans. Ind. Ceram. Soc.,* **2009,** *68,* 1-4.
15. Q. Tang and T. Wang, "Preparation of silica aerogel from rice hull ash by supercritical carbon dioxide drying", *J. Supercrit. Fluids,* **2005,** *35*, 91-94.
16. A. S. Vlasov, A. I. Zakharov, O. A. Sarkisyan and N. A. Lukasheva, "Obtaining silicon carbide from rice husks", *Refract. Ind. Ceram.,* **1991,** *32*, 521-523.
17. D. R. Dimitar and U. Ivan, "Nanosized silicon carbide obtained from rice husks", *Solid State Phenom.,* **2010,** *159*, 153-156.
18. Y. Guo, J. Zhao, H. Zhang, S. Yang, J. Qi, Z. Wang and H. Xu, "Use of rice husk-based porous carbon for adsorption of Rhodamine B from aqueous solutions", *Dyes Pigments,* **2005,** *66*, 123-128.
19. W. Panpa and S. Jinawath, "Synthesis of ZSM-5 zeolite and silicalite from rice husk ash", *Appl. Catal. B-Environ.,* **2009,** *90*, 389-394.
20. S. Chandrasekhar, P. N. Pramada and J. Majeed, "Effect of calcination temperature and heating rate on the optical properties and reactivity of rice husk ash", *J. Mater. Sci.: Mater. Electron.,* **2006,** *41*, 7926-7933.
21. B. I. Ugheoke, E. O. Onche, O. N. Namessan and G. A. Asikpo, "Property optimization of kaolin -rice husk insulating fire-bricks", *Chin. Refract.,* **2006,** *4*, 17-23.
22. E. O. Onche, B. I. Ugheoke, S. A. Lawal and U. M. Dickson, "Effect of rice husk and diatomite on the insulating properties of kaolin-clay firebricks", *Leonardo Electron. J. Pract. Technol.,* **2007,** *11*, 81-90.
23. H. Hamdan, M. N. M. Muhid, S. Endud, E. Listiorini and Z. Ramli, "^{29}Si MAS NMR, XRD and FESEM studies of rice husk silica for the synthesis of zeolites", *J. Non-Cryst. Solids,* **1997,** *211*, 126-131.
24. L. Sun and K. Gong, "Silicon-based materials from rice husks and their applications", *Ind. Eng. Chem. Res.,* **2001,** *40*, 5861-5877.
25. S. Chandrasekhar, K. G. Satyanarayana, P. N. Pramada and P. Raghavan, "Processing, properties and applications of reactive silica from rice husk—An overview", *J. Mater. Sci.,* **2003,** *38*, 3159-3168.

26. K. Mochidzuki, A. Sakoda, M. Suzuki, J. Izumi and N. Tomonaga, "Structural behavior of rice husk silica in pressurized hot-water treatment processes", *Ind. Eng. Chem. Res,* **2001,** *40*, 5705-5709.
27. K. K. Larbi, "Synthesis of high purity silicon from rice husk", *MSc Thesis*, **2010,** University of Toronto, Canada.
28. R. M. Mohamed, R. M. Radwan, M. M. Abdel-Aziz and M. M. Khattab, "Electrical and thermal properties of c-irradiated nitrile rubber/rice husk ash composites", *J. Appl. Polym. Sci.,* **2010,** *115*, 1495-1502.
29. G. R. Rao, A. R. K. Sastry and P. K. Rohatgi, "Nature and reactivity of silica available in rice husk and its ashes", *Bull. Mater. Sci.,* **1989,** *12*, 469-479.
30. J. James and M. S. Rao, "Characterization of silica in rice husk ash", *Am. Ceram. Soc. Bull.,* **1986,** *65*, 1177-1180.
31. P. C. Kapur, "Production of reactive bio-silica from the combustion of rice husk in a tube-in-basket (TiB) burner", *Powder Technol.,* **1985,** *44*, 63-67.
32. Y. Nakata, M. Suzuki, T. Okutani, M. Kikuchi and T. Akiyama, "Preparation and properties of SiO_2 from rice hulls", *J. Ceram. Soc. Jpn.,* **1989,** *97*, 842-849.
33. E. Natarajan and S. E. Ganapathy, "Pyrolysis of rice husk in a fixed bed reactor", *World Acad. Sci. Eng. Technol.,* **2009,** *32*, 504-508.
34. M. Rozainee, S. P. Ngo, A. A. Salema and K. G. Tan, "Fluidized bed combustion of rice husk to produce amorphous siliceous ash", *Ener. Sustain. Develop.,* **2008,** *12*, 32-42.
35. T-C. Luan and T-C. Chou, "Recovery of silica from the gasification of rice husk/coal in the presence of a pilot flame in a modified fluidized bed", *Ind. Eng. Chem. Res.,* **1990,** *29*, 1922-1927.
36. N. Wang and M. J. D. Low, "Spectroscopic studies of carbons XVIII: The charring of rice hulls", *Mater. Chem. Phys.,* **1990,** *26*, 117-130.
37. N. Yalcin and V. Sevinc, "Studies on silica obtained from rice husk", *Ceram. Int.,* **2001,** *27*, 219-224
38. D. S. Subbukrishna, K. C. Suresh, P. J. Paul, S. Dasappa and N. K. S. Rajan, "Precipitated silica from rice husk ash by IPSIT process", Proceedings of 15th European Biomass and Exhibition, **2007,** Berlin, Germany, pp. 2091-2093.
39. C. E. Dodson and V. I. Lakshmanan, "An innovative gas-solid torbed reactor for the recycling industries", *J. Miner. Met. Mater. Soc.,* **1998,** *50*, 29-31.
40. S. Chandrasekhar, P. N. Pramada and L. Praveen, "Effect of organic acid treatment on the properties of rice husk silica", *J. Mater. Sci.* **2005,** *40*, 6535-6544.
41. R. V. Krishnarao and M. M. Godkhindi, "Studies on the formation of SiC whiskers from pulverized rice husk ashes", *Ceram. Int.,* **1992,** *18*, 35-42.
42. A. Proctor, "X-ray diffraction and scanning electron microscope studies of processed rice hull silica", *J. Am. Oil Chem. Soc.,* **1990,** *67*, 576-584.
43. C. Real, M. D. Alcala and J. M. Criado, "Preparation of silica from rice husks", *J. Am. Ceram. Soc.,* **1996,** *79*, 2012-2016.

44. U. Kalapathy, A. Proctor and J. Shultz, "A simple method for production of pure silica from rice hull ash", *Bioresour. Technol.,* **2000,** *73*, 257- 262.
45. R. V. Krishnarao, J. Subrahmanyam and T. J. Kumar, "Studies on the formation of black particles in rice husk silica ash", *J. Eur. Ceram. Soc.,* **2001,** *21*, 99-104.
46. A. Chakraverty, P. Mishra and H. D. Banerjee, "Investigation of combustion of raw and acid-leached rice husk for production of pure amorphous white silica", *J. Mater. Sci.,* **1988,** *23*, 21-24.
47. A. R. Chakraverty, H. D. Banerjee and P. Mishra, "Production of amorphous silica from rice husk in a vertical furnace", *AMA Agric. Mech. Asia Afr. Lat. Am.,* **1990,** *21*, 69-75.
48. J. Umeda and K. Kondoh, "High-purification of amorphous silica originated from rice husk by combination of polysaccharide hydrolysis and metallic impurities removal", *Ind. Crop Prod.,* **2010,** *32*, 539-544.
49. R. Conradt, P. Pimkhaokham and U. Leela-Adisorn, "Nano-structured silica from rice husk", *J. Non-Cryst. Solids,* **1992,** *145*, 75-79.
50. M. Patel, A. Karera and P. Prasanna, "Effect of thermal and chemical treatments on carbon and silica contents in rice husk", *J. Mater. Sci.,* **1987,** *22*, 2457-2464.
51. P. Sidheswaran and A. N. Bhat, "Recovery of amorphous silica in pure form from rice husk", *Trans. Ind. Ceram. Soc.,* **1996,** *55*, 93-97.
52. K. Rohatgi, S. V. Prasd and P. K. Rohatgi, "Release of silica-rich particles from rice husk by microbial fermentation", *J. Mater. Sci. Lett.,* **1987,** *6*, 829-831.
53. W. L. Suchanek and R. E. Riman, "Hydrothermal synthesis of advanced ceramic powders", *Adv. Sci. Technol.,* **2006,** *45*, 184-193.
54. X. C. Wu, "Studies on the hydrothermal conditionsof the extraction of high purity silica from rice husks", *Bull. Chin. Ceram. Soc. ,* **1996,** 15, 36-40.
55. X. W. Yu, G. H. Xu, Y. Y. Zhou, G. P. Zhao and S. N. Shang, " A new technology of producing high quality white carbon black with rice husk, rice straw and wheat straw", *Bull. Chin. Ceram. Soc.,* **1996,** *15*, 48-53.
56. X. W. Yu, G. H. Xu, Y. Y. Zhou, G. P. Zhao and S. N. Shang, "Study of factors influencing extraction rate of white carbon black produced from rice husk", *Chem. Eng. (China),* **1998,** *26*, 51-56.
57. Y. Shinohara and N. Kohyama, "Quantitative analysis of tridymite and cristobalite crystallized in rice husk ash by heating", *Ind. Health,* **2004,** *42*, 277-285.
58. T.-H. Liou, "Preparation and characterization of nano-structured silica from rice husk", *Mater. Sci. Eng.,* **2004,** *364*, 313-323.
59. N. Thuadaij and A. Nuntiya, "Preparation of nanosilica powder from rice husk ash by precipitation method", *Chiang Mai J. Sci.,* **2008***, 35*, 206-211.
60. N. Chand, T. K. Dan, S. Verma and P. K. Rohatgi, "Rice husk ash filled-polyester resin composites", *J. Mater. Sci. Lett.,* **1987,** *6*, 733-735.
61. M. Y. A. Fuad, M. Jamaludin, Z. A. M. Ishak and A. K. M. Omar, "Rice husk ash as fillers in polypropylene: A preliminary study", *Int. J. Polym. Mater.,* **1993,** *19*, 75-92.

62. M. Y. A. Fuad, M. J. Zaini, Z. A. M. Ishak and A. K. M. Omar, "Rice husk ash (RHA)-based silica as filler in polypropylene composite: Effect of peroxide-containing silane on the rheological behaviour", *J. Elast. Plast.,* **1994,** *26*, 252-264.
63. M. Y. A. Fuad, R. Shukor, Z. A. M. Ishak and A. K. M. Omar, "Rice husk ash as filler in polypropylene — Effect of wax and silane coupling agents", *Plast. Rubber Compos. Process. Appl.,* **1994,** *21*, 225-235.
64. M. Y. A. Fuad, Z. Ismail, M. S. Mansor, Z. A. M. Ishak and A. K. M. Omar, "Mechanical properties of rice husk ash/ polypropylene composites", *Polym. J.*, **1995,** *27*, 1002-1015.
65. M. Y. A. Fuad, J. Mustafah, M. S. Mansor, Z. A. M. Ishak and A. K. M. Omar, "Thermal properties of polypropylene/rice husk ash composites", *Polym. Int.,* **1995,** *38*, 33-43.
66. M. Y. A. Fuad, Z. Ismail, Z. A. M. Ishak and A. K. M. Omar, "Application of rice husk ash as fillers in polypropylene: Effect of titanate, zirconate and silane coupling agents", *Eur. Polym. J.,* **1995,** *31*, 885-893.
67. H. Ismail, J. M. Nizam and H. P. S. A. Khalil, "The effect of a compatibilizer on the mechanical properties and mass swell of white rice husk ash filled natural rubber/linear low density polyethylene blends", *Polym. Test.,* **2001,** *20*, 125-133.
68. H. D. Rozman, M. H. Lee, R. N. Kumar, A. Abusamah and Z. A. M. Ishak, "The effect of chemical modification of rice husk with glycidyl methacrylate on the mechanical and physical properties of rice husk-polystyrene composites", *J. Wood Chem. Technol.,* **2000,** *20*, 93-109.
69. S. Siriwardena, H. Ismail and U. S. Ishiaku, "Effect of mixing sequence in the preparation of white rice husk ash filled polypropylene/ethylene-propylene-diene monomer blend", *Polym. Test,* **2000,** *20*, 105-113.
70. L. Sun, M. Xiao, P. Xiao, J. Song, W. Wang, Y. Zhang and K. Gong, "A Preliminary study on rice husk filled polypropylene composite", *Mater. Res. Soc. Symp. Proc.*, **2000**, *661*, KK5.14.1-KK5.14.6.
71. V. I. E. Ajiwe, C. A. Okeke and F. C. Akigwe, "A preliminary study of manufacture of cement from rice husk ash", *Bioresour. Technol.,* **2000,** *73*, 37-39.
72. M. N. Al-Khalaf and H. A. Yousif, "Use of rice husk ash in concrete", *Int. J. Cement Compos. Lightweight Concr.,* **1984,** *6*, 241-248.
73. A. A. Boateng and D. A. Skeete, "Incineration of rice hull for use as a cementitious material: The guyana experience", *Cement Concr. Res.,* **1990,** *20*, 795-802.
74. A. Dass, "Pozzolanic behaviour of rice husk-ash", *Batim. Int. Build. Res. Pract.,* **1984,** *12*, 307-311.
75. G. R. de Sensale, "Effect of rice-husk ash on durability of cementitious materials", *Cement Concr. Compos.,* **2010,** *32*, 718-725.
76. F. I. El-Hosiny, S. A. Abo-El-Enein, I. M. Helmy and K. A. Khalil, "Effect of thermal treatment of rice husk ash on surface properties of hydrated Portland cement-rice husk ash pastes", *J. Therm. Anal.,* **1997,** *48*, 809-817.
77. K. J. Folliard, M. Ohta, E. Rathje and P. Collins, "Influence of mineral admixtures on expansive cement mortars", *Cement Concr. Res. ,* **1994,** *24,* 424-432.

78. K. Ganesan, K. Rajagopal and K. Thangavel, "Rice husk ash blended cement: Assessment of optimal level of replacement for strength and permeability properties of concretes", *Constr. Build. Mater.,* **2008,** *22,* 1675-1683.
79. R. Jauberthie, F. Rendell, S. Tamba and I. Cisse, "Origin of the pozzolanic effect of rice husk", *Constr. Build. Mater.,* **2000,** *14*, 419-423.
80. P. R. Khangaonkar, A. Rahmat and K. G. J. Kutty, "Kinetic study of the hydrothermal reaction between lime and rice-husk-ash silica", *Cement Concr. Res.,* **1997,** *22*, 577-588.
81. F. M. Kilinckale, "The effect of $MgSO_4$ and HCl solutions on the strength and durability of pozzolan cement mortars", *Cement Concr. Res.,* **1997,** *27*, 1911-1918.
82. V. M. Malhotra, "Fly ash, slag, silica fume, and rice husk ash in concrete: A review", *Concr. Int.*, **1993,** *15*, 23-28.
83. P. K. Mehta, "Properties of blended cements made from rice husk ash", *J. Am. Concr. Inst.,* **1977,** *74*, 440-442.
84. P. K. Mehta and D. Pirtz, "Use of rice hull ash to reduce temperature in high-strength mass concrete," *J. Am. Concr. Inst.,* **1978**, *75*, 60-63.
85. V. S. Ramachandran, "Alkali-aggregate expansion inhibiting admixtures", *Cement Concr. Compos.,* **1998**, *20*, 149-161.
86. J. James and M. S. Rao, "Reaction product of lime and silica from rice husk ash", *Cement Concr. Res.,* **1986,** *16*, 67-73.
87. M. V. B. R. Sastry, M. S. Reddy and C. P. Gangaraju, "Comparative study of effect of addition of rice-husk ash and cinderash to soil-lime mixtures", *J. Ind. Highways,* **1986,** *14*, 5-14.
88. Y. Yamamoto and S. M. Lakho, "Production and utilization of active rice husk ash as a substitute for cement", *Trans. Jpn. Soc. Civ. Eng.,* **1982,** *14*, 157-166.
89. Q. Yu, K. Sawayama, S. Sugita, M. Shoya and Y. Isojima, "The reaction between rice husk ash and $Ca(OH)_2$ solution and the nature of its product", *Cement Concr. Res.,* **1999,** *29*, 37-43.
90. M. H. Zhang, R. Lastra and V. M. Malhotra, "Rice-husk ash paste and concrete: Some aspects of hydration and the microstructure of the interfacial zone between the aggregate and paste", *Cement Concr. Res.,* **1996,** *26*, 963-977.
91. M. H. Zhang and V. M. Malhotra, "High-performance concrete incorporating rice husk ash as a supplementary cementing material", *ACI Mater. J.,* **1996,** *93*, 629-636.
92. P. K. Bajpai, M. S. Rao and K. V. G. K. Gokhale, "Synthesis of mordenite type zeolite", *Ind. Eng. Chem. Prod. Res. Dev.,* **1978,** *17*, 223-227.
93. P. K. Bajpai, M. S. Rao and K. V. G. K. Gokhale, "Thermal characteristics of modernite type zeolite", Proceedings of 7th International Conference on Thermal Analysis, **1982**, Kingston, Canada, pp. 558-564.
94. A. K. Dalai, M. S. Rao and K. V. G. K. Gokhale, "Synthesis of NaX zeolite using silica from rice husk ash", *Ind. Eng. Chem. Prod. Res. Dev.,* **1985,** *24*, 465-468.
95. S. A. Deng, "Characteristics and utilizations of zeolite", *Environ. Pollut. Control,* **1991**, *3*, 40-46.
96. J. Dwyer, "Zeolite structure, composition and catalysis", *Chem. Ind.,* **1984**, *7*, 258-268.
97. E. R. Fan, "New uses of zeolite", *Conserv. Utiliz. Min. Resour.,* **1999,** *3*, 19- 27.

98. K. V. G. K. Gokhale, A. K. Dalai and M. S. Rao, "Thermal characteristics of synthetic sodium zeolites prepared with silica from rice-husk ash", *J. Therm. Anal.,* **2005,** *31*, 33-39.
99. D. Prasetyoko, Z. Ramli, S. Endud, H. Hamdan and B. Sulikowski, "Conversion of rice husk ash into zeolite beta", *Waste Manage.*, **2006**, *26*, 1173-1179.
100. N. L. Kumar and D. Das, "Zeolite (ZSM-5) synthesis from rice husk for xylene isomerization", *Res. Ind. ,* **1992,** *37*, 141-142.
101. A. V. Rawtani, M. S. Rao and K. V. G. K. Gokhale, "Synthesis of ZSM-5 zeolite using silica from rice-husk ash", *Ind. Eng. Chem. Res.,* **1989,** *28*, 1411-1414.
102. J.-J. F. Saceda, R. L. de-Leon, K. Rintramee, S. Prayoonpokarach and J. Wittayakun, "Properties of silica from rice husk and rice husk ash and their utilization for zeolite y synthesis", *Quim. Nova,* **2011,** *34*, 1394-1397.
103. P. Stroeven, D. D. Bui and E. Sabuni, "Ash of vegetable waste used for economic production of low to high strength hydraulic binders", *Fuel,* **1999,** *78,* 153-159.
104. H. P. Wang, K. S. Lin, Y. J. Huang, M. C. Li and L. K. Tsaur, "Zeolite (ZSM-5) synthesis from rice husk for xylene isomerization", *J. Hazard. Mater.,* **1998,** *58*, 147-152.
105. X. S. Zhao, Q. X. Wang and H. Y. Li, "Evolution of zeolite synthesis", *Chem. Ind. Eng. Prog.,* **1994,** *13,* 32-45.
106. S. Kurama and H. Kurama, "Particle size effect on the strength of rice husk ash blended gap-graded Portland cement concrete", *Ceram. Int*., **2008**, *34*, 269-272.
107. K. Sujirote and P. Leangsuwan, "Silicon carbide formation from pretreated rice husk", *J. Mater. Sci.,* **2003,** *38*, 4739-4744.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 449-460

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Diphenylmaleimide derivatives and their efficiency in off-on Hg^{2+} fluorometric sensing

Chantana Wainiphithapong[1], Oranual Hanmeng[1], Vannajan Sanghiran Lee[2], Kate Grudpan[3] and Nantanit Wanichacheva[1,*]

[1]Department of Chemistry, Faculty of Science, Silpakorn University, Nakorn Pathom, 73000, Thailand
[2] Department of Chemistry, Faculty of Science, University of Malaya, Kuala Lumpur, 50603, Malaysia
[3] Department of Chemistry, Centre for Innovation in Chemistry, Faculty of Science, and Centre of Excellence in Innovation for Analytical Science and Technology, Chiang Mai University, Chiang Mai 50200, Thailand

* Corresponding author, e-mail: wanichacheva.nantanit@gmail.com, nantanit@su.ac.th

Received: 16 April 2012 / Accepted: 16 November 2012 / Published: 19 November 2012

Abstract: Two novel fluoroionophores (Sensors **1** and **2**) possessing one and two units of diphenylmaleimide fluorophore covalently bound to 2-[3-(2-aminoethylsulfanyl)propylsulfanyl] ethanamine were prepared for the selective detection of Hg^{2+} ions. The binding ability with Hg^{2+} was investigated by fluorescence spectroscopy. Sensor **1** exhibited highly sensitive and selective off-on fluorescence enhancement at 500 nm upon binding to Hg^{2+} and was shown to discriminate various competing metal ions, particularly Cu^{2+} and Pb^{2+}, as well as Li^{+}, Na^{+}, Mg^{2+}, Cd^{2+}, K^{+}, Al^{3+}, Fe^{3+}, Ca^{2+}, Ba^{2+}, Co^{2+}, Mn^{2+} and Zn^{2+} with a detection limit of 6.72×10^{-7} M. On the other hand, Sensor **2** was found to be inferior fluoroionophore to Sensor **1** in terms of selectivity in the presence of competitive ions such as Mn^{2+} and Al^{3+}.

Keywords: mercury sensor, fluoroionophores, diphenylmaleimide, Hg^{2+}-selectivity

INTRODUCTION

Mercury is one of the most highly toxic and hazardous pollutants with recognised accumulative characters in the environment and biota [1-3]. Mercury can cause serious human health problems since it can easily pass through the skin, respiratory and cell membrane, leading to DNA damage, mitosis impairment and permanent damages of the central nervous system, including Minamata disease [4-7].

Recent techniques for the determination of Hg^{2+}, including atomic absorption spectroscopy [8], inductively coupled plasma mass spectrometry [9] and electrochemistry [10], often require a large amount of samples, high cost and sophisticated instrumentation, which pose serious limitations on the detection of Hg^{2+} in biological samples and the tracking of Hg^{2+} for environmental monitoring. Alternatively, fluorometric sensing for the detection of Hg^{2+} offers many advantages since it is highly sensitive and allows nondestructive, prompt determination and real time tracking for the detection of Hg^{2+}.

While many fluorescent sensors have been designed for Hg^{2+}-sensing, many lack the suitability for practical uses due to multi-step syntheses, synthetic difficulty, high costs of starting materials or high detection limits for the determination of Hg^{2+} [11-21]. In addition, they often suffer from cross-sensitivity towards other ions, particularly potential competitors such as copper (Cu^{2+}) and lead (Pb^{2+}), due to their similar chemical behaviours to Hg^{2+} [15-16, 18-19, 22-25]. Notably, most of the reported Hg^{2+} fluorescent chemosensors reveal a fluorescent quenching "turn-off" mechanism due to the quenching characteristic of Hg^{2+} ions, which limits the number of fluorescent enhancement "turn-on" Hg^{2+} sensors reported so far [11, 14, 26].

In the present study, the major motivation is the design and synthesis of new fluorescent enhancement "turn-on" Hg^{2+} sensors which are expected to provide high sensitivity and selectivity to Hg^{2+}, but with a significantly reduced synthetic effort.

These novel sensors were fabricated from the structure of 2-[3-(2-aminoethylsulfanyl) propylsulfanyl]ethanamine, which consists of two sulphur and nitrogen atoms into the platform. It is expected that the sensor can provide appropriately located sulphur and nitrogen atoms as donor atoms that can self-assemble around the Hg^{2+} ions due to the favorable electrostatic interactions [26-28]. This study also focused on the effect of utilising diphenylmaleimide as a fluorophore to increase the sensitivity of the sensor system due to its photostability, relatively high fluorescence quantum yield and long emission wavelengths (~ 500 nm) in the visible region [29-31]. Although many modified structures of diphenylmaleimide fluorophores have been utilised for optoelectronic applications such as organic light emitting diodes [32, 33] and fluorescence photopatterned images materials [34], there are no known reports of utilisation of diphenylmaleimide fluorophores for Hg^{2+} fluorometric sensing applications.

MATERIALS AND METHODS

All reagents and solvents were purchased from Fluka Chemical Corporation and were used as received. All of the metal salts used in this study were perchlorate salts and were purchased from Strem Chemicals, Inc. NMR spectra were obtained with a Bruker Avance 300 spectrometer operating at 300 MHz for ^{1}H and 75 MHz for ^{13}C. All NMR spectra were obtained in $CDCl_3$ solutions with TMS as internal standard. Mass spectra were performed by a ThermoElectron LCQ-DECA-XP, electrospray ionisation trap mass spectrometer. Fluorescence measurements were performed on a Perkin Elmer Luminescence spectrometer LS 50B. Samples were measured in a 1x1 cm quartz cuvette. The excitation and emission slit widths were 5.0 nm. The scan rate was 300 nm/min. Molecular modelling was performed with the Discovery Studio 2.5 program package.

Synthesis: 2-[3-(2-Aminoethylsulfanyl)propylsulfanyl]ethanamine

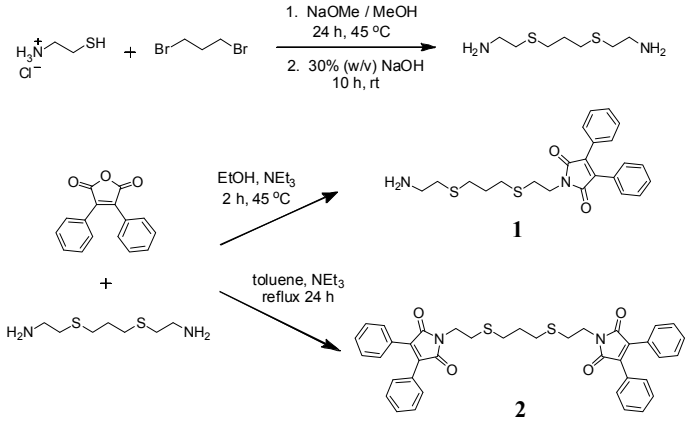
The synthesis of the titled compound was performed in the same manner as described previously [28] by alkylation of cysteamine hydrochloride with 1,3-dibromopropane and the synthetic steps are outlined in Scheme 1.

Synthesis of Sensor 1

Sensor **1** was prepared according to the synthetic pathway in Scheme 1. In a round bottom flask, 2-[3-(2-aminoethylsulfanyl)propylsulfanyl]ethanamine (0.155 g, 0.80 mmol) and triethylamine (0.20 mL, 3.4 mmol) were dissolved in dry ethanol (15.0 mL) under an argon atmosphere. Diphenylmaleic anhydride (0.100 g, 0.40 mmol) was added and the mixture was stirred for 2 hr at 45°C. The solvent was then removed under vacuum to give a yellow oil. The crude product was purified by preparative thin layer chromatography using silica gel as a stationary phase and 10% methanol in dichloromethane as mobile phase (R_f = 0.63) to yield 39 mg of a yellow oil (23%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.84-1.94 (m, 2H), 1.95 (s, NH_2), 2.59-2.63 (m, 4H), 2.69-2.75 (m, 2H), 2.80-2.89 (m, 4H), 3.87 (t, *J* = 7.5, 2H), 7.32-7.49 (m, 10H) ppm. ^{13}C NMR (75 MHz, $CDCl_3$): δ 29.2 (CH_2), 30.0 (CH_2), 30.4 (CH_3), 30.5 (CH_2), 36.0 (CH_2), 37.5 (CH), 42.5 (CH_2), 128.5 (2CH), 129.8 (8CH), 136.2 (4C), 170.5 (2C) ppm. HRMS calculated for $C_{23}H_{27}N_2O_2S_2^+$ $(M+H)^+$ 427.1508 (found 427.1487).

Synthesis of Sensor 2

Sensor **2** was obtained according to the synthetic pathway in Scheme 1. In a round bottom flask, 2-(3-(2-aminoethylsulfanyl)propylsulfanyl)ethanamine (0.078 g, 0.40 mmol) and triethylamine (0.20 mL, 3.4 mmol) were dissolved in dry toluene (8.0 mL) under an argon atmosphere. Diphenylmaleic anhydride (0.200 g, 0.80 mmol) was added and the mixture was refluxed for 24 hr. The solvent was then removed under vacuum to give a yellow oil. The crude product was purified by preparative thin layer chromatography using silica gel as a stationary phase and 15% ethyl acetate in hexane as mobile phase (R_f = 0.42) to yield 123 mg of a yellow oil (47%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.94 (quintet, *J* = 7.2, 2H), 2.74 (t, *J* = 7.2, 4H), 2.84 (t, *J* = 6.9, 4H), 3.87 (t, *J* = 7.5, 4H), 7.33-7.51 (m, 20H) ppm. ^{13}C NMR (75 MHz, $CDCl_3$): δ 28.9 ($2CH_2$), 29.9 (CH_2), 30.4 ($2CH_2$), 37.5 ($2CH_2$), 128.6 (4CH), 129.9 (4C), 130.0 (16CH), 136.2 (4C), 170.5 (4C) ppm. HRMS calculated for $C_{39}H_{34}N_2O_4S_2Na^+$ $(M+Na)^+$ 681.1852 (found 681.1773).



Scheme 1. Syntheses of Sensors **1** and **2**

RESULTS AND DISCUSSION

The target sensors were prepared using a conventional two-step synthesis (Scheme 1). Sensors **1** and **2** contain two sulphur atoms and two nitrogen atoms for the binding sites which are covalently bound to diphenylmaleimide fluorophore(s). The selective binding was expected to take place through favorable electrostatic interactions between the carbonyl carbon as well as sulphur and nitrogen atoms of the sensors and Hg^{2+} ions.

Sensitivity Studies of Sensors 1 and 2

The sensing properties of Sensors **1** and **2** were investigated in order to elucidate the quantitative binding affinity of the sensors to Hg^{2+}. Herein, the sensitivity studies of Sensors **1** and **2** were examined by measuring the fluorescence signals in the presence of various concentrations of Hg^{2+} ions. Figures 1 and 2 show the fluorescence spectra of Sensors **1** and **2** respectively, in the presence and absence of different concentrations of Hg^{2+}, which exhibited fluorescence emission maximum at 500 nm when excited at 367 nm.

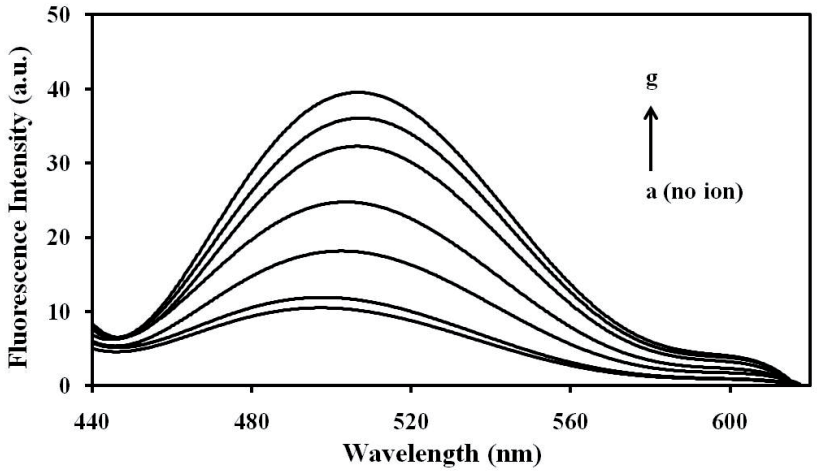


Figure 1. Fluorescence emission spectra (λ_{ex}=367 nm) of Sensor **1** (1.0 μM) in dichloromethane as a function of [Hg^{2+}]. a: 0 μM, b: 0.7 μM, c: 1.0 μM, d: 1.3 μM, e: 1.7 μM, f: 2.0 μM, g: 2.7 μM

When an ion-complexation was operative, the fluorescence behaviour of Sensor **1** clearly demonstrated the off-on switching mechanism that occurred in response to Hg^{2+} ion complexation, as demonstrated in Figure 1. In the absence of Hg^{2+} ions, the fluorescence response was at a minimum and the fluorescence "turn on" as the Hg^{2+} concentration was increased. When the added mercury perchlorate attained a concentration 2.7 times higher than that of Sensor **1**, the fluorescence response reached a maximum point and reached a plateau. The fluorescence quantum yield (ϕ_f) of Sensor **1** with 6.7 equiv. of Hg^{2+} was determined to be 0.021, using quinine sulphate standard with a ϕ_f of 0.54 in 0.1 M H_2SO_4 as a reference [35]. The association constant, K_{assoc}, was obtained by Benesi-Hildebrand plot of the signal changes in the fluorescence titration results [14, 36] and was found to be 3.27 x 10^5 M^{-1} and the 1:1 complex formation of **1**-Hg^{2+} was suggested. The 1:1 complex formation was consistent with molecular modelling studies. The detection limit of Sensor **1** for the analysis of Hg^{2+} was determined from the plot of fluorescence intensity as a function of the Hg^{2+} concentrations [37]. It was found that Sensor **1** had a detection limit of 6.72x10^{-7} M for Hg^{2+} ions, which is sufficient for the detection of sub-micromolar concentrations of Hg^{2+} ions found in many environmental systems such as edible fish [38]. In addition, Sensor **1** offered long-wavelength emission and the change in fluorescence signals in the visible regions, which could be employed to fabricate an economical Hg^{2+} testing tool.

In a similar study, the fluorescence titrations of Sensor **2** with Hg^{2+} were carried out and Sensor **2** acted as an off-on fluorescence switch upon Hg^{2+} binding as illustrated in Figure 2. The sensor showed a high Hg^{2+}-sensitivity and the emission intensity of Sensor **2** was effectively enhanced upon the addition of Hg^{2+} ions. However, Sensor **2** was found to be a comparable sensor to Sensor **1** in terms of sensitivity. It was found that Sensor **2** provided a detection limit of 6.67 x 10^{-7} M for Hg^{2+} ions. The fluorescence quantum yield (ϕ_f) of Sensor **2** with 6.7 equiv. of Hg^{2+} was found to be 0.06, based on quinine sulphate standard. The association constant was found to be 5.81 x 10^5 M^{-1} and the 1:1 complex formation of Sensor **2**-Hg^{2+} was suggested.

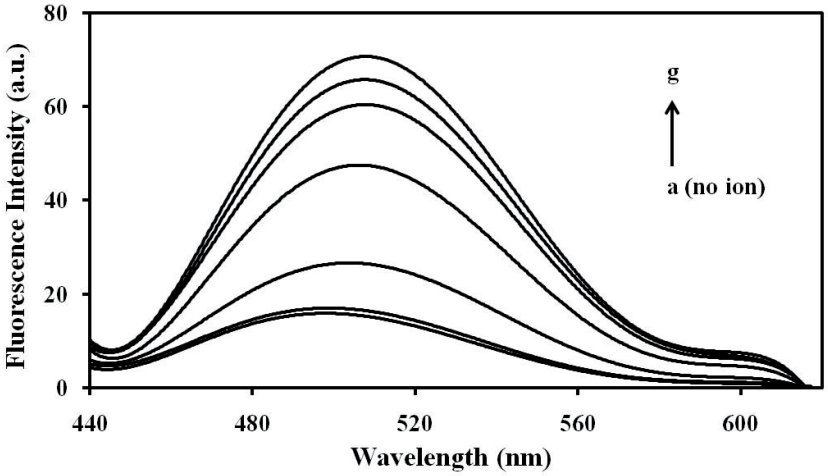


Figure 2. Fluorescence emission spectra (λ_{ex}=367 nm) of Sensor **2** (1.0 μM) in dichloromethane as a function of [Hg^{2+}]. a: 0 μM, b: 0.7 μM, c: 1.0 μM, d: 1.3 μM, e: 1.7 μM, f: 2.0 μM, g: 2.7 μM

Binding Modes of the Sensors

To explain the coordination geometry of Sensors **1** and **2** and Hg^{2+} upon binding, molecular modelling was performed using the Discovery Studio 2.5 program. The structures of Sensors **1** and **2** were initially modified from the X-ray crystal structure of N,N'-(3,7-diazanonylene)-bis-napthalimide in the protein databank PDB ID = 1CX3 and diphenylmaleic anhydride from PubChem compound (CID 78530), and optimised using density functional theory with local density approximation (LDA) of local functional PWC [39]. Then, the initial structures were optimised using CHARMm force field. The complexation energy of the host-guest structure was calculated from the Energy of complex – Energy of compound – Energy of Hg^{2+} using density functional theory with LDA of local functional PWC with implicit distance-dependent dielectrics.

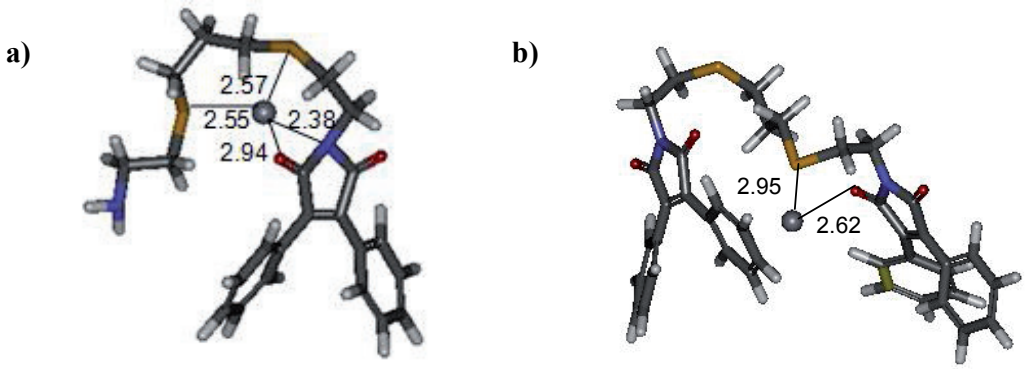


Figure 3. Optimised structures of 1:1 complexes of (a) **1**-Hg^{2+} and (b) **2**-Hg^{2+} from molecular dynamic with LDA of local functional PWC. The Hg^{2+} ions were shown in ball model. The C, N, O, S, H were in grey, blue, red, yellow and white respectively. The distances are shown in Angstrom

The optimised structures of the host-guest complexes are shown in Figure 3, indicating that ion-recognition of the sensors originated from self-assembly processes of the sensors and Hg^{2+} from the favorable electrostatic interactions (ion-dipole interactions) of the carbonyl carbon as well as the sulphur and nitrogen atoms to Hg^{2+} to form a wrapping structure. The distances to indicate the binding sites of Hg^{2+} bound to Sensors **1** and **2** are illustrated in Figure 3. For Sensor **1**, Hg^{2+} was coordinated to the carbonyl oxygen, two sulphur atoms and nitrogen atom with the distances of 2.94, 2.55, 2.57 and 2.38 Å respectively. As to Sensor **2**, Hg^{2+} was bound by carbonyl carbon and sulphur atom with the distances of 2.62 Å and 2.95 Å respectively.

Selectivity Studies of Sensors 1 and 2

The selectivity studies of Sensors **1** and **2** were performed in dichloromethane solutions by recording the fluorescence spectra of the solutions before and after the addition of each representative metal ion. In this study, the selectivity studies were obtained by a similar method to the separate solution method (SSM) used in ion-selective electrode applications [40]. This method involves the measurement of a series of separate solutions, with each solution containing only a salt of the determined ion. Figure 4 represents the dependence of the fluorescence intensity of Sensors **1** and **2** as a function of cation concentrations of Hg^{2+}, Cu^{2+}, Pb^{2+}, Li^{+}, Na^{+}, Mg^{2+}, Cd^{2+}, K^{+}, Al^{3+}, Fe^{3+}, Ca^{2+}, Ba^{2+}, Co^{2+}, Mn^{2+} and Zn^{2+}.

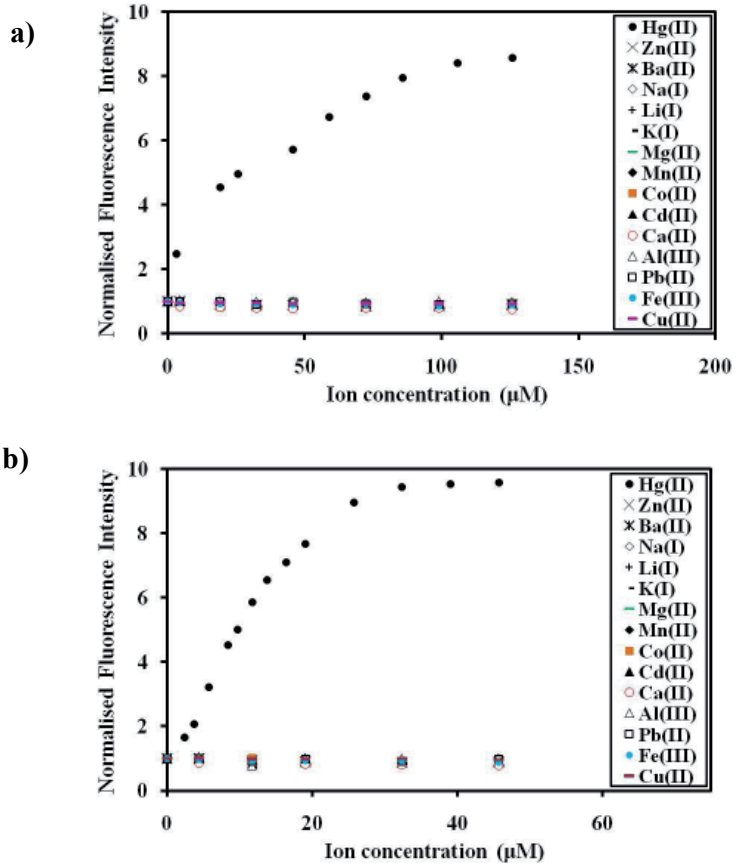


Figure 4. a) Normalised fluorescence intensity of Sensor **1** (1.0 μM) and b) normalised fluorescence intensity of Sensor **2** (1.0 μM) at 500 nm versus the concentrations of various metal ions, i.e. Hg^{2+}, Cu^{2+} Li^{+}, Na^{+}, Mg^{2+}, Cd^{2+}, K^{+}, Al^{3+}, Fe^{3+}, Ca^{2+}, Ba^{2+}, Co^{2+}, Mn^{2+}, Zn^{2+} and Pb^{2+}

The values in the plots were normalised to the fluorescence intensity at 500 nm. The selectivity studies clearly exhibited the excellent selectivity of Sensors **1** and **2** to Hg^{2+} ions in comparison with other metal ions. The results showed that the fluorescence responses at 500 nm increased as a function of added Hg^{2+} until it reached the maximum points. On the other hand, the responses of Sensors **1** and **2** did not cause any significant changes after the addition of Cu^{2+}, Pb^{2+}, Li^{+}, Na^{+}, Mg^{2+}, Cd^{2+}, K^{+}, Al^{3+}, Fe^{3+}, Ca^{2+}, Ba^{2+}, Co^{2+}, Mn^{2+} and Zn^{2+} under identical conditions.

To explore further utilisation of **1** and **2** as Hg^{2+}-selective sensors, competitive studies of Sensors **1** and **2** were performed. Figures 5 and 6 demonstrated the competitive signalling behaviours of Sensors **1** and **2** respectively, with Hg^{2+} in the presence of environmentally important metal ions (Cu^{2+}, Pb^{2+}, Li^{+}, Na^{+}, Mg^{2+}, Cd^{2+}, K^{+}, Al^{3+}, Fe^{3+}, Ca^{2+}, Ba^{2+}, Co^{2+}, Mn^{2+} and Zn^{2+}).

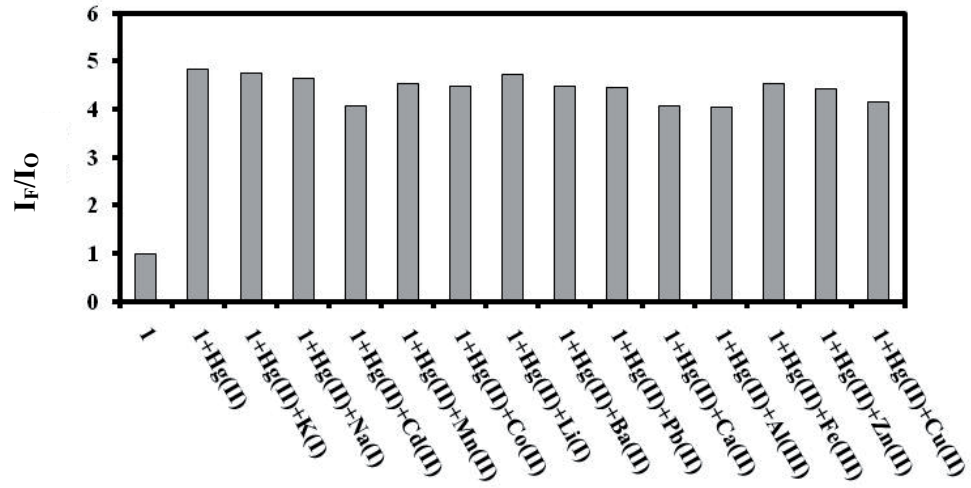


Figure 5. Competitive experiment in the **1**-Hg^{2+} system with common foreign metal ions: [**1**] = 1.0 μM, $[Hg^{2+}] = [M^{n+}]$ = 1.0 μM in dichloromethane solutions (λ_{ex} 367 nm)

The bars in the Figure represent the final fluorescence emission response (I_F) over the initial fluorescence emission response (I_0) at 500 nm. I_F was the fluorescence emission of Sensor **1** in the presence of competitive cations (1.0 μM each of Cu^{2+}, Li^{+}, Na^{+}, Mg^{2+}, Cd^{2+}, K^{+}, Al^{3+}, Fe^{3+}, Ca^{2+}, Ba^{2+}, Co^{2+}, Mn^{2+}, Zn^{2+} and Pb^{2+}) and Hg^{2+}(1.0 μM). I_F/I_0 (where I_F was the fluorescence intensity of Sensor **1** in the presence of Hg^{2+} only) was used as a reference and the I_F/I_0 reference value was equal to 4.8. The I_F/I_0 values were found to lie between 4.1-4.8, indicating that a relatively consistent Hg^{2+}-induced fluorescence enhancement was observed in the presence of equimolar amounts of competing ions. It should be noted that the sensing ability of Sensor **1** showed the sensitivity for Hg^{2+} in the presence of Cu^{2+} and Pb^{2+}, which are potential competitors. The observed selectivity for Hg^{2+} was remarkable compared to many multidentate thioether-containing ligands, i.e. calixarenes, cyclams and cyclens, in previous reports [15, 16, 18, 19, 22-25].

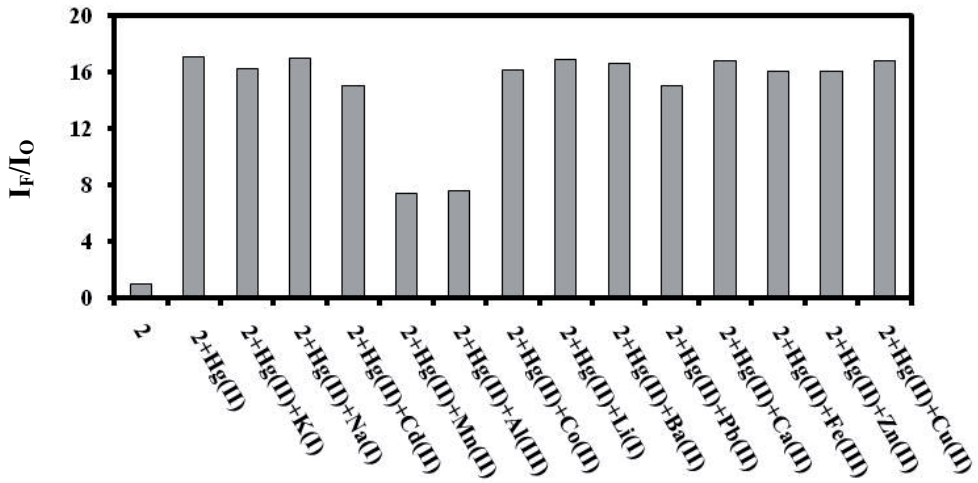


Figure 6. Competitive experiment in the **2**-Hg^{2+} system with common foreign metal ions: [**2**] = 1.0 μM, $[Hg^{2+}] = [M^{n+}]$ = 1.0 μM in dichloromethane solutions (λ_{ex} 367 nm)

Figure 6 shows the competitive signalling behaviours of Sensor **2** with Hg^{2+} in the presence of equimolar amounts of competing ions. Sensor **2** was found to be inferior to Sensor **1** in terms of selectivity in the presence of competitive ions since the sensing ability of Sensor **2** showed poor selectivity for Hg^{2+} in the presence of equimolar amounts of Mn^{2+} and Al^{3+}. The lower selectivity of Sensor **2** might be due to the steric effect from two diphenylmaleimide fluorophores upon ion binding. From the computational data, these two bulky groups give rise to a larger binding site from the self-assemble process. Therefore, in a competitive experiment, the Hg^{2+} binding could be interfered and replaced by some other ions such as Mn^{2+} and Al^{3+} in this case.

CONCLUSIONS

The first use of diphenylmaleimide fluorophore as new fluorescent enhancement "turn-on" Hg^{2+} sensors was successfully demonstrated. Two new sensors based on the 2-[3-(2-aminoethylsulfanyl)propylsulfanyl]ethanamine ligand covalently bound to one and two units of the diphenylmaleimide fluorophore, Sensors **1** and **2**, were prepared by a conventional two-step synthesis. Especially, Sensor **1** showed highly sensitive and selective fluorescence "turn-on" behaviour toward Hg^{2+} in solutions and was shown to discriminate various foreign ions, i.e. Cu^{2+}, Pb^{2+}, Li^{+}, Na^{+}, Mg^{2+}, Cd^{2+}, K^{+}, Al^{3+}, Fe^{3+}, Ca^{2+}, Ba^{2+}, Co^{2+}, Mn^{2+} and Zn^{2+}. The molecular design presented could serve as an alternative mercury fluorometric sensor due to the advantages of synthetic simplicity, cost-efficient synthetic route, high sensitivity to Hg^{2+} by "turn-on" fluorescence response in the visible region and high selectivity with particular discrimination of the potential competitors including Cu^{2+} and Pb^{2+}. The new off-on type fluorescence enhancement sensor based on diphenylmaleimide fluorophore could serve as a new potential design for future development of sensor systems.

ACKNOWLEDGEMENTS

This work was supported by Grant RGJ 2554-07 from Faculty of Science, Silpakorn University, Thailand, Grant MRG 5380093 from the Thailand Research Fund and the Commission on Higher Education, Ministry of Education of Thailand. The authors are grateful to the Computational Nanoscience Consortium, Nanotechnology (NANOTEC), Thailand for the access to the Discovery Studio 2.5 program package.

REFERENCES

1. A. Renzoni, F. Zino and E. Franchi, "Mercury levels along the food chain and risk for exposed populations", *Environ. Res.*, **1998**, *77*, 68-72.
2. S. Hardy and P. Jones, "Capillary electrophoresis determination of methylmercury in fish and crab meat after extraction as the dithizone sulphonate complex", *J. Chromatogr. A*, **1997**, *791*, 333-338.
3. H. H. Harris, I. J. Pickering and G. N. George, "The chemical form of mercury in fish", *Science*, **2003**, *301*, 1203.
4. J. Gutknecht, "Inorganic mercury (Hg^{2+}) transport through lipid bilayer membranes", *J. Membrane Biol.*, **1981**, *61*, 61-66.

5. P. B. Tchounwou, W. K. Ayensu, N. Ninashvili and D. Sutton, "Environmental exposure to mercury and its toxicopathologic implications for public health", *Environ. Toxicol.*, **2003**, *18*, 149-175.
6. P. W. Davidson, G. J. Myers, C. Cox, C. F. Shamlaye, D. O. Marsh, M. A. Tanner, M. Berlin, J. Sloane-Reeves, E. Cernichiari, O. Choisy, A. Choi and T. W. Clarkson, "Longitudinal neurodevelopmental study of Seychellois children following in utero exposure to methylmercury from maternal fish ingestion: outcomes at 19 and 29 months", *Neurotoxicology*, **1995**, *16*, 677-688.
7. P. Grandjean, P. Weihe, R. F. White and F. Debes, "Cognitive performance of children prenatally exposed to 'safe' levels of methylmercury", *Environ. Res.*, **1998**, *77*, 165-172.
8. N. S. Bloom and W. F. Fitzgerald, "Determination of volatile mercury species at the picogram level by low-temperature gas chromatography with cold-vapour atomic fluorescence detection", *Anal. Chim. Acta.*, **1988**, *208*, 151-161.
9. J. A. Moreton and H. T. Delves, "Simple direct method for the determination of total mercury levels in blood and urine and nitric acid digests of fish by inductively coupled plasma mass spectrometry", *J. Anal. Atom. Spectrosc.*, **1998**, *13*, 659-665.
10. K. H. Chen, H. W. Wang, B. S. Kang, C. Y. Chang, Y. L. Wang, T. P. Lele, F. Ren, S. J. Pearton, A. Dabiran, A. Osinsky and P. P. Chow, "Low Hg(II) ion concentration electrical detection with AlGaN/GaN high electron mobility transistors", *Sens. Actuat. B: Chem.*, **2008**, *134*, 386-389.
11. Y. Shiraishi, S. Sumiya, Y. Kohno and T. Hirai, "A rhodamine-cyclen conjugate as a highly sensitive and selective fluorescent chemosensor for Hg(II)", *J. Org. Chem.*, **2008**, *73*, 8571-8574.
12. B. K. McMahon and T. Gunnlaugsson, "Lanthanide luminescence sensing of copper and mercury ions using an iminodiacetate-based Tb(III)-cyclen chemosensor", *Tetrahedron Lett.*, **2010**, *51*, 5406-5410.
13. S. H. Kim, K. C. Song, S. Ahn, Y. S. Kang and S.-K. Chang, "Hg^{2+}-selective fluoroionophoric behaviour of pyrene appended diazatetrathia-crown ether", *Tetrahedron Lett.*, **2006**, *47*, 497-500.
14. T.-H. Ma, A.-J. Zhang, M. Dong, Y.-M. Dong, Y. Peng and Y.-W. Wang, "A simply and highly selective "turn-on" type fluorescent chemosensor for Hg^{2+} based on chiral BINOL-Schiff's base ligand", *J. Lumin.*, **2010**, *130*, 888-892.
15. S. H. Kim, J. S. Kim, S. M. Park and S.-K. Chang, "Hg^{2+}-selective OFF−ON and Cu^{2+}-selective ON−OFF type fluoroionophore based upon cyclam", *Org. Lett.*, **2006**, *8*, 371-374.
16. S. M. Park, M. H. Kim, J. I. Choe, K. T. No and S.-K. Chang, "Cyclams bearing diametrically disubstituted pyrenes as Cu^{2+}- and Hg^{2+}-selective fluoroionophores", *J. Org. Chem.*, **2007**, *72*, 3550-3553.
17. N. J. Youn and S.-K. Chang, "Dimethylcyclam based fluoroionophore having Hg^{2+}- and Cd^{2+}-selective signaling behaviours", *Tetrahedron Lett.*, **2005**, *46*, 125-129.
18. K.-C. Song, M. H. Kim, H. J. Kim and S.-K. Chang, "Hg^{2+}- and Cu^{2+}-selective fluoroionophoric behaviours of a dioxocyclam derivative bearing anthrylacetamide moieties", *Tetrahedron Lett.*, **2007**, *48*, 7464-7468.
19. R. Metivier, I. Leray and B. Valeur, "Lead and mercury sensing by calixarene-based fluoroionophores bearing two or four dansyl fluorophores", *Chem. Eur. J.*, **2004**, *10*, 4480-4490.

20. Y. H. Lee, M. H. Lee, J. F. Zhang and J. S. Kim, "Pyrene excimer-based calix[4]arene FRET chemosensor for mercury(II)", *J. Org. Chem.*, **2010**, *75*, 7159-7165.
21. Q.-Y. Chen and C.-F. Chen, "A new Hg^{2+}-selective fluorescent sensor based on a dansyl amide-armed calix[4]-aza-crown", *Tetrahedron Lett.*, **2005**, *46*, 165-168.
22. S. Y. Moon, N. R. Cha, Y. H. Kim and S.-K. Chang, "New Hg^{2+}-selective chromo- and fluoroionophore based upon 8-hydroxyquinoline", *J. Org. Chem.*, **2004**, *69*, 181-183.
23. R. Martinez, A. Espinosa, A. Tarraga and P. Molina, "New Hg^{2+} and Cu^{2+} selective chromo- and fluoroionophore based on a bichromophoric azine", *Org. Lett.*, **2005**, *7*, 5869-5872.
24. J. Tan and X.-P. Yan, "2,1,3-Benzoxadiazole-based selective chromogenic chemosensor for rapid naked-eye detection of Hg^{2+} and Cu^{2+}", *Talanta*, **2008**, *76*, 9-14.
25. Y.-B. Chen, Y.-J. Wang, Y.-J. Lin, C.-H. Hu, S.-J. Chen, J.-L. Chir and A.-T. Wu, "A water-soluble ribosyl-based fluorescent sensor for Hg^{2+} and Cu^{2+} ions", *Carbohydr. Res.*, **2010**, *345*, 956-959.
26. N. Wanichacheva, K. Setthakarn, N. Prapawattanapol, O. Hanmeng, V. S. Lee and K. Grudpan, "Rhodamine B-based "turn-on" fluorescent and colorimetric chemosensors for highly sensitive and selective detection of mercury (II) ions", *J. Lumin.*, **2012**, *132*, 35-40.
27. N. Wanichacheva, A. Kamkaew, S. Watpathomsub, V. S. Lee and K. Grudpan, "2-[3-(2-Aminoethylsulfanyl)propylsulfanyl]ethanamine bearing dansyl subunits: an efficient, simple, and rapid fluorometric sensor for the detection of mercury(II) ions", *Chem. Lett.*, **2010**, *10*, 1099-1101.
28. N. Wanichacheva, M. Siriprumpoonthum, A. Kamkaew and K. Grudpan, "Dual optical detection of a novel selective mercury sensor based on 7-nitrobenzo-2-oxa-1,3-diazolyl subunits", *Tetrahedron Lett.*, **2009**, *50*, 1783-1786.
29. D. H. Suh, N. C. Yang and Y. H. Park, "Synthesis and characterization of polyaryleneethylenes having 1,2-diphenylmaleimide in the backbone", *React. Funct. Polym.*, **2002**, *51*, 121-127.
30. D. H. Suh, J. K. Lee, N. C. Yang and H. W. Choi, "The fluorescence effects on the *N*-substituents of polyaryleneethylenes having 1,2-diphenylmaleimide moieties", *Macromol. Res*., **2003**, *11*, 92-97.
31. D. H. Suh and N. C. Yang, "A class of cis-polyarylenevinylene:preparation and optical properties of the oligomeric polyarylenevinylene derivative having an 1,2-diphenylmaleimide moiety", *Polymer*, **2001**, *42*, 7987-7992.
32. C. T. Chen, Y. D. Lee and L. H. Chan, "Achieving saturated red photoluminescence and electroluminescence with readily synthesized maleimide-arylamine copolymers", *Tetrahedron Lett.*, **2006**, *62*, 9541-9547.
33. C. T. Chen, H. C. Yeh and W. C. Wu, "The colourful fluorescence from readily-synthesised 3,4-diaryl-substituted maleimide fluorophores", *Chem. Commun.*, **2003**, 404-405.
34. D. H. Suh, J. K. Lee, N. C. Yang, H. W. Choi and J. I. Hwang, "A novel diphenylmaleimide dye for fluorescence photopatterned images", *Chem. Lett.*, **2002**, *31*, 824-825.
35. S. Yoon, A. E. Albers, A. P. Wong and C. J. Chang, "Screening mercury levels in fish with a selective fluorescent chemosensor", *J. Am. Chem. Soc.*, **2005**, *127*, 16030-16031.

36. W. H. Melhuish, "Quantum efficiencies of fluorescence of organic substances: effect of solvent and concentration of the fluorescent solute", *J. Phys. Chem.*, **1961**, *65*, 229-235.
37. X.-L. Wang, W.-Y. Zheng, G.-C. Liu and H.-Y. Lin, "A turn-on fluorescent chemosensor for Hg^{2+} based on phenanthroline fluorophore", *J. Lumin.*, **2010**, *130*, 52-55.
38. M. Shortreed, R. Kopelman, M. Kuhn and B. Hoyland, "Fluorescent fiber-optic calcium sensor for physiological measurements", *Anal. Chem.*, **1996**, *68*, 1414-1418.
39. J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy", *Phys. Rev. B*, **1992**, *45*, 13244-13249.
40. E. Bakker, P. Buhlmann and E. Pretsch, "Carrier-based ion-selective electrodes and bulk optodes. 1. General characteristics", *Chem. Rev.*, **1997**, *97*, 3083-3132.

Maejo Int. J. Sci. Technol. **2012**, *6*03), 461-469

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Bioconversion of biomass residue from the cultivation of pea sprouts on spent *Pleurotus sajor-caju* compost employing *Lumbricus rubellus*

Azizi Abu Bakar[1],*, Noor Zalina Mahmood[2], Noorlidah Abdullah[2] and Rosna Mat Taha[1]

[1]Institute of Biological Sciences, Faculty of Science University of Malaya, Kuala Lumpur 50603, Malaysia
[2]Mushroom Research Centre, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia

* Corresponding author, e-mail: azieaxis@gmail.com, azizi.bkr@um.edu.my

Received: 11 August 2011 / Accepted: 16 November 2012 / Published: 19 November 12

Abstract: Vermicomposting is a green technology for the purpose of nutrient enrichment from a variety of organic waste products. In this study, saw dust-based spent mushroom compost (SMC), an organic waste and biomass residue, was used as a medium for the cultivation of pea sprouts. After harvesting the pea sprouts, the growth medium was reused to culture earthworms, *Lumbricus rubellus*. The culturing activity was conducted for 50 days without any pre-composting or thermocomposting. Thus duration of vermicomposting process was shortened as opposed to previous work on vermicomposting of saw dust-based SMC (no amendment) for 70 days. The culturing treatments were conducted in triplicate, including one treatment without earthworms as the control. The analysis showed that concentrations of macronutrients in vermicompost were higher compared to controls, in which N = 4.12%, P = 2.07% and K = 1.56%. The C:N ratio was 11.77, which indicates a stabilisation and maturity of the organic waste compost, compared with the C:N ratio for the control, which was 59.34. At the end of the experiment, increment of total biomass and number of earthworms were observed and no mortality was recorded. The results suggested that vermicomposting could be used as an environmentally valuable technology to convert saw dust used for mushroom and pea sprouts cultivation into vermicompost or bio-fertiliser by employing *L. rubellus*.

Keywords: bio-fertiliser, epigeic earthworms, spent mushroom compost, vermicomposting, waste management

INTRODUCTION

The current trend of utilising waste products as resources forms important part of green technology policies in many countries. To achieve sustainable development in any food-based industry, several novel technologies have been introduced in recent years with the aim of recycling the waste generated. Mushrooms has become an increasingly important industrial crop in Malaysia. Hence the amounts of organic waste generated from mushroom farming have multiplied accordingly. Spent mushroom compost (SMC) was generated when fruiting bodies were no longer produced, which typically occurs after six months of mushroom cultivation [1]. SMC is commonly discarded whereby more than 4000 tonnes per year were sent to landfills or openly burnt at mushroom farms. A proactive profit-earning solution to reduce the dumping of organic waste has been put forward such as thermo-composting, co-composting, vermicomposting, ruminant feed production, biofuel formation and biomass briquetting.

Vermicomposting of SMC has given the waste a value instead of having to be discarded. Vermicomposting has recently been tested with many types of organic waste [Table 1] with the aim of producing a high-quality end product in terms of nutrient content or of remediating pollutants. However, data from studies on the epigeic *Lumbricus rubellus* are limited. Vermicomposting is technically affordable and environmentally safe, compared to conventional industrial methods that require chemicals and expensive machinery. It is an eco-biotechnological process in which earthworms and associated microflora convert the organic waste into organic fertiliser utilising available forms of nutrients [2]. Moreover, vermicompost is fragmented and microbiologically active due to humification [3] and contains important plant nutrients (N, P, K and Ca) in the forms that are soluble and more easily available to plants than ordinary compost [4].

Therefore, the aims of this study are to investigate the potential of *L. rubellus* to convert SMC after cultivation of high-priced vegetables, i.e. pea sprout seeds, into a stabilised and nutrient-rich organic fertiliser, as well as to study the multiplication in number and biomass of worms during the process of vermicomposting and to measure the nutrient elements in the end product (vermicompost) compared to control (compost).

Table 1. Recent vermicomposting studies on different types of waste

Type of waste	Amendment	Earthworm species	Vermicomposting duration	Findings	Reference
Sewage sludge	Spent mushroom compost	*L. rubellus*	91 days	No foul smell and fine texture of vermicompost. Heavy metal content in vermicompost was higher compared to its initial levels due to breakdown of organic matter (explained by heavy metal and mass balance), but the content is below US and EU biosolid compost limits. It is safe to use as bio-fertiliser or soil conditioner.	[1]

Table 1. (Continued).

Type of waste	Amendment	Earthworm species	Vermicomposting duration	Findings	Reference
Human faeces	Soil and vermicompost	*Eisenia fetida*	1year and 6 months (545 days)	Complete inactivation of total coliform	[27]
Beverage sludge	Cattle dung	*E. fetida*	4 months (120 days)	Degradation of 50:50 mixture (bio-sludge:cattle dung) achieved in 75 days when worms were inoculated at 25 g/kg feed mixture, but the best quality product was obtained after 105–110 days with 7.5 g worms/kg feed mixture.	[28]
Agro-industry sludge	Cow dung, biogas plant slurry and wheat straw	*E. fetida*	15 weeks (105 days)	40:60 (industrial sludge:cow dung) and 40:60 (industrial sludge:biogas plant slurry) showed highest mineralisation rate and earthworm growth pattern.	[10]
Municipal sewage sludge	Oyster shell	*E. andrei*	25 days	Powdered oyster shell sludge blend provided stable pH due to its buffering capacity because of the effects on the release of Ca^{2+} and OH^{-}.	[29]
Vegetable-market solid waste	Wheat straw, cow dung and biogas plant slurry	*E. fetida*	15 weeks (105 days)	Waste mineralisation and humification rate were higher in bedding of those containing easily digestible bulky agents, i.e. biogas slurry and cow dung.	[30]
Non-recyclable paper waste	Cow dung	*E. fetida*	91 days	FT-IR spectroscopy of vermicompost showed reduction in aliphatic compounds during vermicomposting process.	[31]
Coffee grounds and kitchen waste	Cow dung	*L. rubellus*	70 days	Coffee grounds can be decomposed through vermicomposting and help to enhance the quality of vermicompost produced compared to sole use of kitchen waste in vermicomposting.	[32]

MATERIALS AND METHODS

Cultivation of Pea Sprouts on Spent *Pleurotus Sajor-Caju* Compost (SMC)

Spent *P. sajor-caju* (grey oyster mushroom) compost was removed from plastic bags and broken up. A layer of the compost was filled into a tray to 5-cm height and moistened by spraying with tap water. Pea sprout seeds were soaked in tap water and then arranged on the moistened compost

at 1 cm apart. The trays were incubated at room temperature with regular spraying of tap water. Pea sprouts produced were harvested after 7-10 days. Two harvests were obtained before the compost was subjected to vermicomposting.

Vermicomposting Experiment

Clitelated earthworms (*L. rubellus*) were selected from a stock culture in the Earthworm Reservoir, Institute of Biological Sciences, University of Malaya. The treatment was carried out in an artificially designed microcosm (360 mm × 280 mm × 200 mm) with a net (250 mm × 100 mm) covering the centre of the lid to allow for aerobic exchange, to prevent any form of interruption and to ensure that the microclimate was maintained. Experiments were conducted in triplicate (T_A, T_B and T_C) with one control. The control received an identical treatment with no earthworms.

Thirty clitelated earthworms of approximately the same size were introduced into each bin containing 300 g of SMC and in the fourth week, another 300 g of the same substrate was added. The pH and temperature of the SMC were measured and an optimum level of pH 7±1 and a temperature of 27±1°C were achieved. Due to the optimum and stabilised pH and temperature at the initial period of vermicomposting, no pre-composting period was required. A pre-composting period is normally included to avoid the exposure of earthworms to high temperature during the initial thermophilic stage of microbial decomposition [5].

Vermicomposting lasted for 50 days. During this process, the moisture content of the feed materials was maintained at 60–70% by constantly spraying distilled water onto the surface, in combination with a manual turning of the feed material over once every few days to remove any stagnant water. At the end of the study period, the upper layer of the vermicompost produced in the plastic bin was sampled (~100 g with moisture content, i.e. 60%) for analysis of nutrient elements before all the earthworms were removed [6]. The upper layer was sampled because this was the first layer converted into vermicompost, which was classified by its fine, odourless texture. The number of living earthworms was determined after hand sorting and removal of all extraneous material. The biomass gain of earthworms was calculated as:

$$\frac{(\text{Biomass on day 50} - \text{Biomass on day 0})}{\text{Biomass on day 0}} \times 100$$

Nutrient Element Analysis

The production of organic C in the vermicompost was determined using the partial-oxidation method [7]. Kjeldahl digestion with concentrated H_2SO_4 (1:20, w/v) followed by distillation was used to estimate N content [8]. P was detected by a colorimetric method using ammonium molybdate in HCl [9]. K, Mg and Ca were measured by the ignition method using a Perkin-Elmer model 3110 double beam atomic absorption spectrophotometer [5]. The maturity of the vermicompost was calculated from the C:N ratio.

Statistical Analysis

Statistical analysis was carried out using SPSS v. 16.0. A paired samples t-test was performed to analyse the significance in the difference between the earthworms' biomass and number in percentage during vermicomposting at 0.05% level of significance.

RESULTS AND DISCUSSION

Multiplication of earthworms is an important indicator in determining the vermicomposting performance. The biomass and number of earthworms increased noticeably from day 0 to day 50 of the experiment (Table 2). The average of the triplicates revealed that the highest gain in biomass and number was in T_A with 178.28% and 143.33% respectively. The paired-samples t-test showed a significant difference (df=2, t=4.86, $P<0.05$) between the earthworms biomass and number. The rate of biomass gain of the earthworms ranged between 94.52 and 180.38 mg day^{-1}. No loss of biomass or numbers and no mortalities of earthworms were recorded at the end of experiment. The results clearly suggested that the multiplication of earthworms was directly related to the quality of feed materials and nutrients available in the SMC after cultivation of pea sprouts. This is supported by Suthar [10] who studied the recycling of agro-industrial sludge with organic bulky agents as amendment by vermicomposting.

Furthermore, the residue from the pea sprouts harvest, e.g. roots and the prevalent mycelium were additional food supplement for the earthworms and indirectly affected the earthworms' palatability. Due to the homogenous and non-foul-smelling feed materials, no pests were introduced into the earthworms' bin and this accelerated the vermicomposting process. In the natural environment, plant residues are quickly colonised by microorganisms. These microorganisms are common constituent of the earthworms' food resources, in particular protozoa and fungi that formed a substantial part of their diet [11-13]. However, maintaining high and stable moisture content is an important factor for the earthworms' mobility, ensuring successful feeding and copulation. The increase in the earthworms' growth may also be attributed to the low C:N ratio of the pre-decomposed substrate, i.e. the SMC used in this study [14].

Table 2. Earthworms multiplication in biomass and number

Treatment (T)	Biomass of earthworms (mg)		Earthworms biomass gain/loss (%)	Biomass gain rate (mg day^{-1})
	Initial	Final		
T_A	5059.0	14078.0	+178.28	180.38
T_B	5761.0	14054.0	+143.95	165.86
T_C	4649.0	9375.0	+101.66	94.52
	Number of earthworms		Earthworms number gain/loss (%)	Mortality rate (%)
T_A	30	73	+143.33	Nil
T_B	30	55	+83.33	Nil
T_C	30	38	+26.67	Nil

The nutrient elements of the vermicompost are presented in Table 3 after 50 days of vermicomposting. The three salient macronutrients of the vermicompost, i.e. N, P and K, were relatively higher compared to control and its initial contents. The N content in the vermicompost was ~3% higher compared to that in the control. This might originate from the addition of N by the

earthworms itself in the form of mucus, nitrogenous excretory substances, growth-stimulating hormones and enzymes [15]. Nitrogen fixed by free-living N-fixing bacteria can also result in increased N content in the vermicompost [16] and the level depends on the initial N content present in the feed materials and on the degree of decomposition [17]. Furthermore, an updated perspective attributed the contribution of N in the vermicompost not only to the status of the feed mixture, excretory products, mucus, body fluid and enzymes, but also from the decaying tissue of the dead earthworms [35].

Table 3. Nutrient elements in saw dust, SMC, vermicompost and control (compost)

Nutrient element	Saw dust (%)	SMC[a] (%)	SMC[b] (%)	Vermicompost[c] (%)	Control[d] (%)
Nitrogen (N)	0.33	0.58	2.35	4.12 ± 0.248	1.10
Phosphorous (P)	0.02	0.27	1.23	2.07 ± 0.485	1.34
Potassium (K)	0.17	0.25	0.98	1.56 ± 0.155	1.24
Calcium (Ca)	0.01	0.01	0.11	0.08 ± 0.013	0.10
Magnesium (Mg)	0.01	0.01	0.01	0.01 ± 0.003	0.03
Organic Carbon (C)	32.52	34.40	56.30	47.89 ± 5.882	65.27
C:N ratio	107.41	83.30	23.96	11.79 ± 1.955	59.34

[a] SMC prior to cultivation of pea sprouts
[b] SMC after harvesting of pea sprouts
[c] Vermicompost from SMC after 50 days of bioconversion. Values are mean and standard error (mean ± S.E.M.; $n = 3$)
[d] SMC after harvesting of pea sprouts after 50 days

The P and K contents in the vermicompost were higher than those in the compost. They were also higher than those in saw dust, SMC prior to and after pea sprouts cultivation (Table 3). These were probably due to the direct action of the earthworms gut enzymes and indirectly to the stimulation of the microflora [18]. Barois and Lavelle [19] reported that earthworms produced a huge amount of intestinal mucus, a mixture of glycoproteins and small glucidic and proteic molecules which is rapidly incorporated into the microbial biomass in the gut. The higher K content in the vermicompost was due to a higher mineralisation rate as a result of enhanced microbial and enzymic activities in the earthworms gut [15]. Nevertheless, according to a recent report by Deka et al. [33], there were many contradictory reports regarding the reduction in K content in vermicompost obtained from different feedstock. It was difficult to find any conclusive explanations for this decrease. The considerable increase in P was ascribed to changes in sorption complexes induced by competition for sorbing sites between orthophosphates and carboxyl groups of glycoproteins within the mucus produced in the earthworms gut [20]. Referring to Edwards and Lofty [21], the rise in P during vermicomposting is probably due to P mineralisation and mobilisation because of the bacterial and faecal phosphate activity of earthworms. Meanwhile, Suthar [30] suggested that the P content in the final product may vary depending on the earthworms' metabolism and available P is contributed partly by earthworm gut and partly by further release of P through P-solubilising microorganisms present in the worm cast.

The Ca and Mg contents seemed to differ only slightly although their concentrations were low, these exchangeable nutrients contribute significantly to the sustainability of agro-ecosystem in food source cultivation. According to West et al. [22], *L. rubellus* accumulates Ca in their anterior alimentary canal in order to maintain their body Ca concentration and in some cases, the Ca metabolism in the earthworms' gut enzymes and bacterial communities in vermicast result in an increase of Ca. The Mg content made up the least amount of the nutrients tested and it is categorised as

a trace nutrient. However, this plant nutrient potentially increases agriculture-ready materials. No direct contribution of earthworms to the Mg metabolism is known. It is hypothesised that fungi and microalgae which easily colonise freshly deposited worm casts contribute to the trace level of Mg in ready vermicompost [10].

The organic C content from saw dust (32.52%) increased after cultivation of mushroom (34.40%) and pea sprouts (56.30%) but decreased after vermicomposting by *L. rubellus* (47.89 ± 5.882%). According to Suthar [23], earthworms promote microclimatic conditions in vermireactors that increase the loss of organic C from the substrates through microbial respiration. Moreover, loss of organic C due to mineralisation of organic matter during vermicomposting may be the reason for the increased N in the end product [34]. The C:N ratio is used as an index for the maturity of organic wastes. In this study, the C:N ratio for the vermicompost was less than 20 (Table 3). According to Senesi [24], a C:N ratio of less than 20 indicates an advanced degree of organic matter stabilisation and reflects a satisfactory degree of maturity of organic waste. A high C:N ratio, as in the control, reflects a reduction in biological activity and consequently a slow degradation [25]. Compared to a previous study by Sailila et al. [26], this study has shown stabilisation of the C:N ratio compared to the previous forms of organic waste or biomass residue (saw dust and SMC).

CONCLUSIONS

Reuse and recycle saw dust-based SMC need to be comprehensive and practical yet environmentally sound. The amounts of nutrients available in the saw dust-based SMC after harvest enable the cultivation of pea sprouts and subsequent conversion of the waste generated postharvest of pea sprouts into a valuable product i.e. vermicompost. Hence, the saw dust-based SMC can be reused to cultivate other vegetation i.e. pea sprouts and yet recycled into a bio-fertiliser via vermicomposting. In this study, the pre-composting period was excluded and there was no amendment of the substrate with other bulky organic waste. Thus, the process can be shortened and inclusively utilise the same substrate while still yielding material that is rich in nutrient elements. Vermicomposting using *L. rubellus* has succeeded to convert a reusable biomass residue into a nutrient-rich end product for sustainable agricultural farming as an alternative to synthetic chemical fertilisers.

ACKNOWLEDGEMENTS

The authors are thankful to Institute of Research Management & Monitoring (IPPP), University of Malaya under the University of Malaya Research Grant No.: RG007/09AFR managed by Research Grant Management Unit (UPGP) for financial support.

REFERENCES

1. A. A. Bakar, N. Z. Mahmood, J. A. T. da Silva, N. Abdullah and A. A. Jamaludin, "Vermicomposting of sewage sludge by *Lumbricus rubellus* using spent mushroom compost as feed material: Effect on concentration of heavy metals", *Biotechnol. Bioproc. Eng.*, **2011**, *16*, 1036-1043.
2. S. Suthar, "Pilot-scale vermireactors for sewage sludge stabilization and metal remediation process: Comparison with small-scale vermireactors", *Ecol. Eng.*, **2010**, *36*, 703-712.

3. M. S. Maboeta and L. van Rensburg, "Vermicomposting of industrially produced woodchips and sewage sludge utilizing *Eisenia fetida*", *Ecotoxicol. Environ. Saf.*, **2003**, *56*, 256-270.
4. P. M. Ndegwa and S. A. Thompson, "Integrating composting and vermicomposting in the treatment and bioconversion of biosolids", *Bioresour. Technol.*, **2001**, *76*, 107-112.
5. T. C. Loh, Y. C. Lee, J. B. Liang and D. Tan, "Vermicomposting of cattle and goat manures by *Eisenia foetida* and their growth and reproduction performance", *Bioresour. Technol.*, **2005**, *96*, 111-114.
6. N. A. Nik Nor Izyan, A. A. Jamaludin and N. Z. Mahmood, "Potential of spent mushroom substrate in vermicomposting", *Dyn. Soil Dyn. Plant*, **2009**, *3*, 87-90.
7. A. Walkley and I. A. Black, "An examination of the degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method", *Soil Sci.*, **1934**, *37*, 29-38.
8. J. M. Bremner and R. G. Mulvaney, "Nitrogen total", in "Methods of Soil Analysis" (Ed. A. L. Page, R. H. Miller and D. R. Keeney), American Society of Agronomy, Madison, **1982**, pp. 575-624.
9. M. K. John, "Colorimetric determination of phosphorus in soil and plant materials with ascorbic acid", *Soil Sci.*, **1970**, *109*, 214-220.
10. S. Suthar, "Recycling of agro-industrial sludge through vermitechnology", *Ecol. Eng.*, **2010**, *36*, 1028-1036.
11. C. A. Edwards and K. E. Fletcher, "Interactions between earthworms and microorganisms in organic-matter breakdown", *Agric. Ecosyst. Environ.*, **1988**, *24*, 235-247.
12. G. G. Brown, "How do earthworms affect microfloral and faunal community diversity?", *Plant Soil*, **1995**, *170*, 209-231.
13. M. Bonkowski and M. Schaefer, "Interactions between earthworms and soil protozoa: A trophic component in the soil food web", *Soil Biol. Biochem.*, **1997**, *29*, 499-502.
14. P. M. Ndegwa and S. A. Thompson, "Effects of C-to-N ratio on vermicomposting of biosolids", *Bioresour. Technol.*, **2000**, *75*, 7-12.
15. G. Tripathi and P. Bhardwaj, "Comparative studies on biomass production, life cycles and composting efficiency of *Eisenia fetida* (Savigny) and *Lampito mauritii* (Kinberg)", *Bioresour. Technol.*, **2004**, *92*, 275-283.
16. R. D. Kale, K. Bano and R. V. Krishnamurthy, "Potential of *Perionyx excavatus* for utilizing organic wastes", *Pedobiol.*, **1982**, *23*, 419-425.
17. J. H. Crawford, "Review of composting", *Process Biochem.*, **1983**, *18*, 14-15.
18. J. E. Satchell and K. Martein, "Phosphatase activity in earthworm faeces", *Soil Biol. Biochem.*, **1984**, *16*, 191-194.
19. I. Barois and P. Lavelle, "Changes in respiration rate and some physiochemical properties of a tropical soil during transit through *Pontoscolex corethrurus* (glossoscolecidae, oligochaeta)", *Soil Biol. Biochem.*, **1986**, *18*, 539-541.
20. D. López-Hernández, P. Lavelle, J. C. Fardeau and M. Nino, "Phosphorus transformations in two P-sorption contrasting tropical soils during transit through *Pontoscolex corethrurus* (Glossoscolecidae: Oligochaeta)", *Soil Biol. Biochem.*, **1993**, *25*, 789-792.
21. C. A. Edwards and J. R. Lofty, "Biology of Earthworms", Chapman and Hall, London, **1972**.

22. H. K. West, M. S. Davies, A. J. Morgan and R. J. Herbert, "Intraspecific variation in calcium and strontium accumulation/depuration in an epigeic earthworm species", *Eur. J. Soil Biol.*, **2001**, *37*, 329-332.
23. S. Suthar, "Potential utilization of guar gum industrial waste in vermicompost production", *Bioresour. Technol.*, **2006**, *97*, 2474-2477.
24. N. Senesi, "Composted materials as organic fertilizers", *Sci. Total Environ.*, **1989**, *81-82*, 521-524.
25. R. T. Haug, "The Practical Handbook of Compost Engineering", 3rd Ed., Lewis Publishers, CRC Press Inc, Florida, **1993**, p. 744.
26. N. Sailila, A. A. Bakar, N. Z. Mahmood, J. A. T. da Silva, N. Abdullah and A. A. Jamaludin, "Nutrient elements of different agricultural wastes from vermicomposting activity", *Dyn. Soil Dyn. Plant*, **2010**, *4*, 155-158.
27. K. D. Yadav, V. Tare and M. M. Ahammed, "Vermicomposting of source-separated human faeces for nutrient recycling", *Waste Manage.* **2010**, *30*, 50-56.
28. J. Singh, A. Kaur, A. P. Vig and P. J. Rup, "Role of *Eisenia fetida* in rapid recycling of nutrients from bio sludge of beverage industry", *Ecotoxicol. Environ. Saf.*, **2010**, *73*, 430-435.
29. Y. T. Kwon, C. W. Lee and J. H. Yun, "Development of vermicast from sludge and powdered oyster shell", *J. Clean. Prod.*, **2009**, *17*, 708-711.
30. S. Suthar, "Vermicomposting of vegetable-market solid waste using *Eisenia fetida*: Impact of bulking material on earthworm growth and decomposition rate", *Ecol. Eng.*, **2009**, *35*, 914-920.
31. R. Gupta and V. K. Garg, "Vermiremediation and nutrient recovery of non-recyclable paper waste employing *Eisenia fetida*", *J. Hazard. Mater.*, **2009**, *162*, 430-439.
32. A. J. Adi and Z. M. Noor, "Waste recycling: Utilization of coffee grounds and kitchen waste in vermicomposting", *Bioresour. Technol.*, **2009**, *100*, 1027-1030.
33. H. Deka, S. Deka, C. K. Baruah, J. Das, S. Hoque and N. S. Sarma, "Vermicomposting of distillation waste of citronella plant (*Cymbopogon winterianus* Jowitt.) employing *Eudrillus eugeniae*", *Bioresour. Technol.*, **2011**, *102*, 6944-6950.
34. A. Yadav and V. K. Garg, "Recycling of organic wastes by employing *Eisenia fetida*", *Bioresour. Technol.*, **2011**, *102*, 2874-2880.
35. H. Deka, S. Deka, C. K. Baruah, J. Das, S. Hoque, H. Sarma and N. S. Sarma, "Vermicomposting potentiality of *Perionyx excavatus* for recycling of waste biomass of java citronella – an aromatic oil yielding plant", *Bioresour. Technol.*, **2011**, *102*, 11212–11217.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 470-482

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Effects of organic and conventional rice on protein efficiency ratio and pesticide residue in rats

Wanpen Mesomya[1,*], Pakawadee Sutthivaiyakit[2], Yaovadee Cuptapun[1] and Duangchan Hengsawadi[1]

[1] Institute of Food Research and Product Development, Kasetsart University, 50 Ngamwongwan Road, Chatuchak, Bangkok 10900, Thailand
[2] Department of Chemistry and Centre of Excellence for Innovation in Chemistry, Faculty of Science, Kasetsart University, 50 Ngamwongwan Road, Chatuchak, Bangkok 10900, Thailand
* Corresponding author, e-mail : ifrwpm@ku.ac.th ; Tel : 660 2942 8629 - 35 ext 922, Fax : 660 2940 6455

Received: 6 December 2011 / Accepted: 23 November 2012 / Published: 23 November 2012

Abstract: The comparative effects of organic rice and conventional rice on the protein efficiency ratio (PER) in rats were investigated by feeding 40 male Sprague-Dawley rats for four weeks with three experimental diets containing polished conventional rice (PCR), unpolished conventional rice (UCR), unpolished organic rice (UOR) and a control protein diet (casein) under standardised conditions. All diets were prepared according to AOAC guidelines. The results showed no statistically significant difference ($P > 0.05$) among the values of PER (2.75 ± 0.14 - 2.80 ± 0.09) in rats fed with diets containing PCR, UCR or UOR. Similar growth was also observed among the three groups fed with different experimental diets. Additionally, residues of pesticides, viz. carbofuran, methyl parathion, p-nitrophenol and β-cyfluthrin, in rat blood and rice samples were determined using liquid chromatography–electrospray ionisation tandem mass spectrometry. Pesticide residues were not detected in all serum samples of experimental rats and only p-nitrophenol was found (8.23 ± 0.65 - 12.84 ± 2.58 mg/kg) in all samples of the cooked rice diets, indicating that organic rice produced similar effect as conventional rice on PER and growth in rats.

Keywords: organic rice, conventional rice, protein efficiency ratio, pesticide residue

INTRODUCTION

Rice (*Oryza sativa*) is the main food used by half of the world's population [1]. It is one of the most important cereal crops widely used in human nutrition as a source of energy due to its high starch content (approximately 90% in polished white grains) [2]. Total starch between 72-82% has been observed in brown rice grains of six cultivars grown in Philippines [3]. However, the level of this constituent can vary among grains of different varieties due to genetic and environmental factors. Furthermore, the rate and extent of starch digestion can be influenced by various factors including amylose/amylopectin ratio, grain processing, physicochemical properties (particularly gelatinisation characteristics), particle size and presence of lipid-amylose complexes [4].

Organic farming is a form of agriculture that relies on crop rotation, green manure, compost, biological pest control, and mechanical cultivation to maintain soil productivity and to control pests. The method excludes or strictly limits the use of synthetic fertilisers, synthetic pesticides, plant growth regulators, livestock feed additives and genetically modified organisms. Since 1990, the market for organic products has grown at a rapid pace, averaging 20-25 % annually to reach US$ 33 billion in 2005. This demand has driven a similar increase in organically managed farmland. Approximately 306,000 square kilometers (30.6 million hectares) worldwide are now farmed organically, representing approximately 2% of the world's total farmland. In addition, as of 2005, organic wild products were farmed on approximately 62 million hectares [5]. As for rice, only specially selected high-quality Jasmine rice is planted organically on a very limited area, although this type of agricultural practice is becoming more widespread as the number of health-conscious consumers is growing rapidly [6].

The protein quality of roasted wheat, maize and rice has been studied in feeding trials with rats. The protein efficiency ratio (PER), true digestibility, biological value and net protein utilisation decrease significantly with roasting. Relative nitrogen utilisation also decreases on roasting but the effect is greater with rice and maize than with wheat [7]. Kaur and Sekhon [8] studied in albino rats for 28 days using a reference casein diet, a raw rice diet and diet based on pressure parboiled rice as well as traditionally parboiled rice. The PER did not vary significantly among different groups but was highest for rats fed on the casein diet, followed by raw rice, pressure parboiled rice and traditionally parboiled rice diets respectively. Growth study and nutritional evaluation of 'Basmati-370', 'Pusa Basmati-1' and 'Haryana Basmati-1' have also been carried out using rats. The feed efficiency ratio and PER values did not differ among the rice diets. 'Pusa Basmati-1' had better apparent protein digestibility, true protein digestibility, biological value, net protein utilisation, utilisable protein and liver enzymes level than the other varieties [9].

Lam-Sonhez et al. [10] measured the protein content of the rice grain and found that it decreases from the outer to the inner layers of the grain. When weaning male Wistar rats were fed on diets with 7.9% protein supplied by different parts of the grain, the residue, which was mainly the central part of the endosperm (protein content 8.54 g/100 g), had a PER of 1.68, compared with 1.57 and 1.04 for the outer layers. This may be due to the presence of grain embryo in the grain residue. When the embryo was detached from the grain at the beginning of processing, there was a decrease in the nutritional quality of the grain. However, there has been no previous report on the

PER of organic rice compared with conventional rice, so one of the objectives of the present study was to compare the PER of organic rice with that of conventional rice.

Since organic rice is grown and processed without the use of any synthetic chemicals whereas conventional rice farming involves the use of synthetic chemicals such as fertilisers, insecticides, pesticides and herbicides. Consequently rats fed with organic rice should have better growth and contain less toxic chemicals in their blood than those fed with conventional rice. The objectives of this study are thus extended to test the above hypotheses.

MATERIALS AND METHODS

Preparation of Powdered Steamed Rice and Experimental Diets

One type of unpolished organic rice supplied by Progressive Farmer Association, Ubol Ratchathani province and two types of polished and unpolished conventional rice supplied by Ubol Agri Coop Federation Ltd., Ubol Ratchathani province were weighed, washed, steamed for 30 min., dried in an oven at 70^{o}C for 7 hours and then powdered using a pin mill.

The above three kinds of rice: unpolished organic rice, polished conventional rice and unpolished conventional rice, and casein were used for the preparation of experimental diets by AOAC methods [11]. The diets consisted of 10 ± 0.3% protein, 8% soy oil, 5% mineral mixture (Table 1), 1% vitamin mixture (Table 2), 1% cellulose, 5% moisture, 35% sucrose and 35% corn starch. The composition of the experimental diets is shown in Table 3. Proximate analysis of rice and experimental diets were determined by AOAC methods [12].

Animals

Three-week old weanling male Sprague-Dawley rats were obtained from the National Laboratory Animal Centre, Mahidol University, Thailand. The rats with a mean initial weight of 50 - 60 g were used. They were divided randomly into four groups of 10 rats (one control group and three test groups). All rats were housed in individual stainless steel metabolic cages in an experimental controlled environment at 20-22^{o}C, 60% relative humidity and 12-hours light-dark cycle and were given free access to the diet and water for a 28-day feeding period. Daily food intake and weekly body weight were recorded. The experimental protocol was developed according to the guidelines of the Committee on Care and Use of Experimental Animal Resources, Institute of Food Research and Product Development, Kasetsart University.

Table 1. Composition of mineral mixture [11]

Mineral	g/kg
NaCl	139.300
KI	0.790
KH_2PO_4	389.000
Mg SO_4 anhydrous	57.300
$CaCO_3$	381.400
$FeSO_4 .7H_2O$	27.000
$MnSO_4 .H_2O$	4.010
$ZnSO_4 .7H_2O$	0.548
$CuSO_4 .5H_2O$	0.477
$CoCl_2 .6H_2O$	0.023

Table 2. Composition of vitamin mixture [11]

Vitamin	mg / 100 g ration
Vitamin A	2000 (IU)
Vitamin D	200 (IU)
Vitamin E	10 (IU)
Menadione	0.500
Choline	200.000
p-Aminobenzoic acid	10.000
Inositol	10.000
Niacin	4.000
Ca D-pantothenate	4.000
Riboflavin	0.800
Thiamine. HCl	0.500
Pyridoxine. HCl	0.500
Folic acid	0.200
Biotin	0.040
Vitamin B_{12}	0.003
Glucose, to make	1000.000

Table 3. Composition (g) of four experimental diets of 10 kg [11]

	Cooked polished conventional rice	Cooked unpolished conventional rice	Cooked unpolished organic rice	Casein
Rice	7002.80	6337.14	6693.44	-
Casein	614.63	614.63	614.63	1229.20
Cellulose	95.08	88.48	91.44	97.98
Vitamin mixture	100.00	100.00	100.00	100.00
Mineral mixture	495.70	488.65	489.42	495.14
Soy oil	728.00	771.58	771.02	798.77
Water	443.08	460.03	459.68	487.14
Corn starch	260.36	569.75	390.19	3395.89
Sugar	260.35	569.74	390.18	3395.88

Determination of Pesticide Residues in Rice and Rat Serum

Reagents

Carbofuran (99.5%) and ß-cyfluthrin were obtained from Dr. Ehrenstorfer GmbH, Germany. Methyl parathion (99.6%), p-nitrophenol (> 99.5%) and triphenyl phosphate (98%) were obtained from Chemical Service, Fluka and APS Ajax Fine Chemicals respectively.

High purity water was obtained from a Maxima Water Purification System (USF-Elga, High Wycombe, Bucks, U.K.). Standard stock solutions (1000 µg/mL in methanol) of p-nitrophenol, carbofuran, methyl parathion and ß-cyfluthrin were prepared and kept in amber bottles at 4ºC. Working solutions were diluted appropriately with methanol as required.

Separation was performed by liquid chromatography (LC) using an Agilent 1100 LC binary pump (Agilent Technology, Waldbronn, Germany) equipped with a 1100 Agilent autosampler. Lichrosphere RP18 (4.0 x 125 mm, 5µm, Merck) was used as a column. The eluents A and B were methanol and 1 mM ammonium acetate respectively. The gradient program started with 40% methanol and increased linearly to 60% over 1 min., to 75% over 2 min., to 85% over 2 min. and to 100% over 5 min. The column was equilibrated for 10 min. prior to the next injection. The injection volume was 20 µL.

Mass spectrometric measurement was performed using an Applied Biosystem API 2000 triple-quadrupole instrument equipped with an electrospray ionisation interface (ESI). Analyst software (version 1.1) from Applied Biosystem was used for system control and data acquisition. The mass analyser, the first quadrupole (Q_1) and the last quadrupole (Q_3) were operated at unit resolution. The optimal parameters for MS/MS are summarised in Table 4.

Table 4. Optimised parameters and multi-reaction monitoring (MRM) transition for tested compounds

Period	Compound	Mode	MRM	Source paramerter						Compound parameter					
				CUR	CAD	IS	TEM	GS1	GS2	DP	FP	EP	CEP	CE	CXP
1	p-Nitrophenol	NI	137.9/92	13.8	2.8	-4200	350	41.4	27.6	-66	-130	7.5	-8	-23	-8
			137.9/107.9											-17	-5
2	Carbofuran	PI	222.2/165	17.2	2.8	5500	350	37.9	31.0	16	360	-11	16	29	12
			222.2/123.1											17	15
3	Methyl	NI	247.8/137.9	13.8	2.8	-4300	350	41.4	24.1	-21	-350	4.5	-14	-26	-14
	parathion		247.8/123.1											-26	-38
4	Triphenyl	PI	327.22/153	13.8	2.8	5000	350	24.1	31.0	56	320	-11.5	16	36	25
	phosphate		327.22/215											35	29
5	β-Cyfluthrin	NI	432.1/405	13.8	1.4	-4200	350	41.4	27.6	-6	-320	6.5	-24	-6.5	-25
			434.1/407											-7.5	-25

Mode : NI = negative ionisation mode, PI = positive ionisation mode

Source parameters : CUR (N/cm^2) = curtain gas; CAD (N/cm^2) = collision gas; IS(V) = ionspray voltage; TEM = temperature (°C); GS1 (N/cm^2) = source gas 1; GS2 (N/cm^2) = source gas 2

Compound parameters : DP(V) = declustering potential; FP(V) = focusing potential; EP(V) = entrance potential; CEP(V) = collision cell entrance potential; CE (V) = collision energy; CXP(V) = collision cell exit

Determination of p-Nitrophenol, Carbofuran, Methyl Parathion and ß-Cyfluthrin in Rice and Plasma

The extraction process for rice was a modification of that developed by Hirahara et al. [13]. An aliquot of steamed rice in powder form (2 ± 0.02 g) was placed in a conical flask and 0.5 mL of 20 mg/kg (0.0020%) triphenyl phosphate (internal standard) [14, 15] was spiked in the sample prior to the addition of 20 mL of methanol-acetone (1:1). The mixture was vortex-mixed for 1 min. and then ultrasonicated for 15 min. The process was repeated once. The combined extract was evaporated in a rotary evaporator and dried under a gentle stream of nitrogen. The residue was redissolved in 500 µL of methanol and filtered through a 0.2 µm nylon membrane prior to injection into LC-ESI-MS/MS.

The extraction process for plasma was a modification of that developed by Zhang et al. [16]. An aliquot of 50 µL of rat plasma sample was mixed with 220 µL of 45 ppb triphenyl phosphate (internal standard) in a 1.5 mL microcentrifuge tube. The solution was vortex-mixed for 2 min. and centrifuged at 13,400 rpm, 4ºC for 10 min. The clear supernatant was removed and diluted to 500 µL with methanol and filtered through a 0.2 µm nylon membrane filter prior to injection into LC-ESI-MS/MS.

Determination of Percentage Recovery of Pesticides and Statistical Analysis

An experiment was performed to determine the percentage recovery of pesticides by the method employed. Serum sample from rats fed with unpolished conventional rice diet and another sample of polished conventional rice diet were taken as representatives for evaluation.

Data were analysed statistically using analysis of variance (ANOVA) and Duncan's new multiple range test. A value of $P < 0.05$ was considered significant [17].

RESULTS AND DISCUSSION

Proximate Analysis of Conventional Rice, Organic Rice and Experimental Diets

The results of proximate analysis of conventional rice, organic rice and experimental diets is shown in Tables 5 and 6. The protein, fat, ash, crude fibre and carbohydrate contents of raw rice were lower than those of the cooked rice, except the fat content of raw polished conventional rice, which was higher than that of the cooked polished conventional rice. The moisture contents of all raw rice types were higher than those of the cooked rice samples because after being steamed, the cooked rice was heated to dry in an oven. The loss of water through evaporation also caused the protein, fat, ash, crude fibre and carbohydrate contents of the cooked rice to be higher than those of the raw rice. The fat content of the polished conventional rice was the lowest because it had been polished. Separate samples of the raw polished and unpolished conventional rice, unpolished organic rice were steamed, dried and powdered. Using the AOAC method [11], three experimental diets were prepared from the three types of cooked powdered rice, i.e. Diet 1- from cooked polished conventional rice, Diet 2- from cooked unpolished conventional rice, Diet 3- from cooked unpolished organic rice and Diet 4 (control)- casein diet containing 10 ± 0.3 % protein. Proximate analysis of Diets 1-4 is shown in Table 6. The protein contents of the experimental diets and control

casein diet were 9.05 - 9.45% wet weight or 9.59 - 10.14% on a dry weight basis at 5% moisture. The level of total dietary fibre in Diet 4 was the lowest (1.89%) because this control diet was composed of casein that lacked fibre found in the rice.

Table 5. Proximate analysis of conventional and organic rice (g/100 g wet weight)

Rice	Protein	Moisture	Fat	Ash	Crude fibre	Carbohy-drate
Raw rice						
Polished conventional rice	6.33	11.54	1.16	0.23	0.31	80.43
Unpolished conventional rice	7.08	11.13	3.20	1.11	1.11	76.37
Unpolished organic rice	6.79	11.97	3.37	1.00	1.09	75.78
Cooked rice						
Polished conventional rice	7.14	7.22	1.02	0.27	0.56	83.79
Unpolished conventional rice	7.89	5.30	4.39	1.41	1.66	79.35
Unpolished organic rice	7.47	5.07	4.24	1.22	1.13	80.87

Table 6. Proximate analysis of experimental diets (g/100 g wet weight)

Experimental Diet	Protein	Moisture	Fat	Ash	Total dietary fibre
Diet 1	9.42	7.17	7.87	4.49	5.32
Diet 2	9.44	6.56	9.22	4.94	4.17
Diet 3	9.45	6.38	9.94	5.08	4.20
Diet 4	9.05	5.71	5.71	4.16	1.89

Growth of Rats Fed with and PER of Conventional and Organic Rice Diets

The final body weight of experimental rats fed with experimental diets composed of organic rice and the corrected PER were not significantly different from those fed with experimental diets from conventional rice (Table 7). PER of sample is calculated from weight gained of test animal (g) divided by protein consumed (g). Corrected PER is calculated from PER of test sample multiplied by 2.50 and divided by PER of casein. PER shown in the study was corrected PER, which were standardises to PER of 2.50 for casein (as standard) to eliminate laboratory variation [18]. Organic rice thus produced similar effect on PER and the growth rate of rats compared with conventional rice. Interestingly, the final body weight of and PER obtained from experimental rats fed with the diets composed of conventional and organic rice were higher than those fed with control casein diet, although the amount of casein in conventional and organic rice diets (614.63 g) were lower than that in the control casein diet (1229.20 g) (Table 3), indicating that the quality of protein in the three experimental rice diets was better than that of the casein diet. High quality protein in the rice diet thus gave rise to higher PER and consequent higher growth rate of the experimental rats.

Table 7. Initial body weight (IBW), final body weight (FBW) and corrected PER of experimental rats fed with four experimental diets

Diets	IBW (g)	FBW (g)	Corrected PER
Diet 1	66.83 ± 1.57^{a}	196.07 ± 9.55^{a}	2.76 ± 0.10^{a}
Diet 2	66.57 ± 1.18^{a}	190.66 ± 12.89^{a}	2.75 ± 0.14^{a}
Diet 3	66.45 ± 2.85^{a}	189.43 ± 11.79^{a}	2.80 ± 0.09^{a}
Diet 4	66.62 ± 2.89^{a}	$157.60 \pm 9.19^{b.}$	2.50 ± 0.12^{b}

Note: Values are means ± standard deviation, N = 10.
Values in the same column with different superscript letters are significantly different at $P < 0.05$.

This study illustrated that both organic and conventional rice diets could produce similar PER and growth rate in rats. The nutritional value of rice was not affected by the farming systems [19, 20] and organic farming can produce similar, if not better, rice quality as that obtained from conventional farming.

Pesticide Residues in Rice and Serum of Rats

In the determination of p-nitrophenol, carbofuran, methyl parathion and ß-cyfluthrin in rice and in serum samples, multireaction monitoring was used in the detection process, as this method is more selective compared with other techniques such as diode array detection, thereby simplifying the sample preparation [16]. A typical total ion chromatogram is shown in Figure 1. The detection of the compounds was divided into four periods to improve the sensitivity of the technique. The regression data of the compounds are presented in Table 8. Internal standard calibration was performed using triphenyl phosphate. The ratio of the peak area of analyte to that of triphenyl phosphate was plotted against the concentration of analyte. Good linear calibration curves were obtained for all compounds with correlation coefficients better than 0.99. The limit of detection (LOD) of each analyte was defined as the concentration giving a signal to noise ratio of 3.

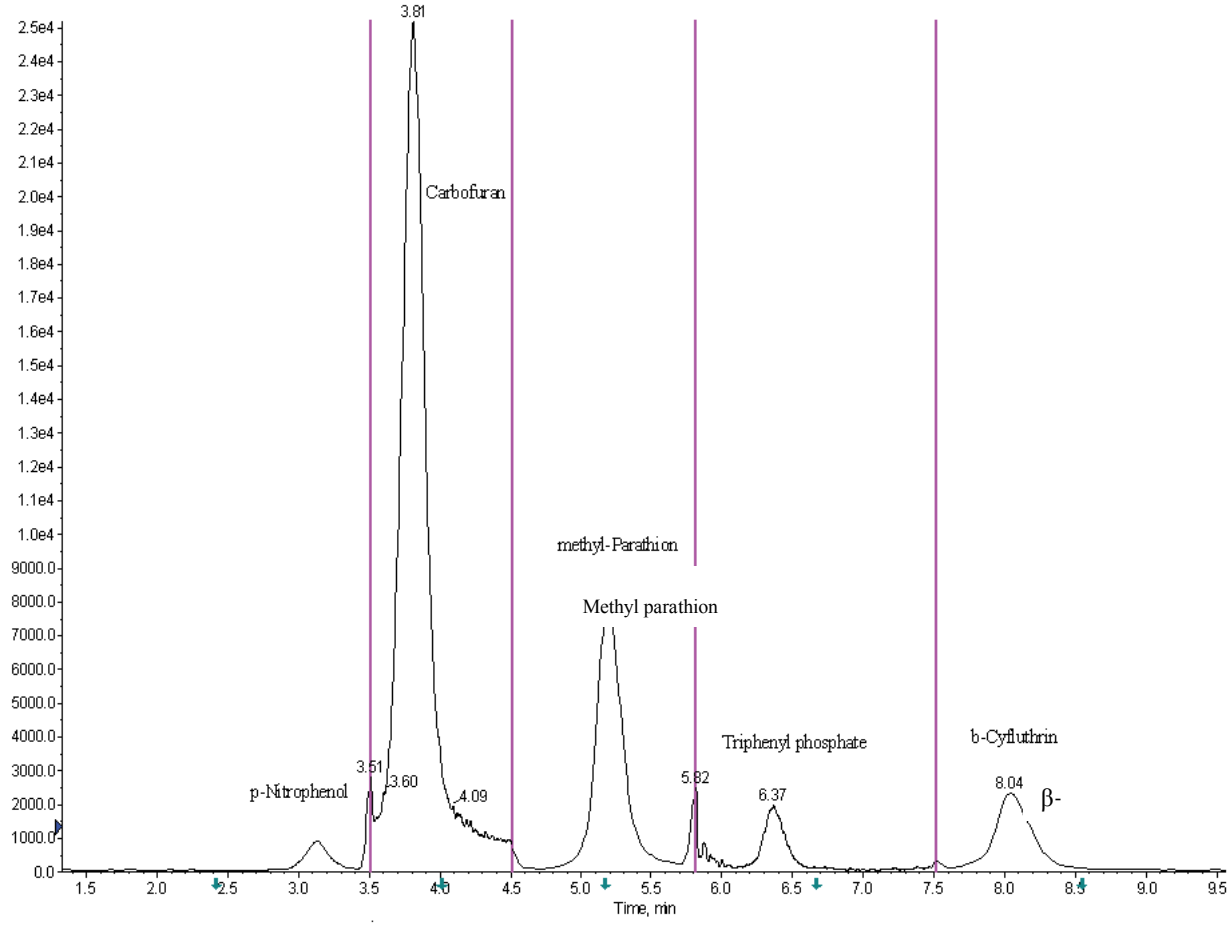


Figure 1. Typical total ion chromatogram of pesticide residues

Table 8. Regression data and detection limit of the tested compounds

Pesticide	Period	t_r *(min)	Linear range (µg/L)	R^2**	LOD*** (µg/L)
p-Nitrophenol	1	3.10	0.80 - 8.00	0.9905	1.162
Carbofuran	2	3.85	200 - 10000	0.9894	2.824
Methyl parathion	3	5.21	490 - 4920	0.9999	443.5
β-Cyfluthrin	4	8.19	1 - 20	0.9973	561.9

*Retention time, **Correlation-coefficients, ***Limit of detection

Results from LC-ESI-MS/MS indicated that, p-nitrophenol, carbofuran, methyl parathion and β-cyfluthrin were not present in any serum sample of the experimental rats. In the rice samples, only p-nitrophenol (a metabolite of methyl parathion) was detected in all samples of the cooked rice diets (Table 9). The presence of p-nitrophenol in organic rice sample might be due to contamination of nitroaromatic compounds and pesticides in the drinking water source used [21, 22]. Another interesting finding from the present study was that although p-nitrophenol was found in all samples

of the cooked rice diets (Table 9), it was not detected in any serum sample of the rats. This might reflect the ability of the animal to metabolise or excrete the residue of p-nitrophenol contaminated in the cooked rice diets. Future study should focus on the effect of organic rice on the health aspects of humans.

Table 9. Concentration (mg/kg) of p-nitrophenol, carbofuran, methyl parathion and β-cyfluthrin in cooked conventional and organic rice diets

Cooked rice diet	p-Nitrophenol	Carbofuran	Methyl parathion	β-Cyfluthrin
Polished conventional rice	12.84 ± 2.58	nd*	nd	nd
Unpolished conventional rice	8.23 ± 0.65	nd	nd	nd
Unpolished organic rice	10.13 ± 0.26	nd	nd	nd
Casein diet	4.28 ± 0.66	nd	nd	nd

* not detectable

Percentage Recovery of Pesticides in Serum and Rice Diets

Percentage recoveries of carbofuran and methyl parathion were found to be high in both samples. However, variations between samples were rather high for p-nitrophenol and β-cyfluthrin (Table 10). The recovery of each analyte was calculated from the peak area of analyte in spiked rice and serum sample, compared to expected concentration based on the combination of spiked and original amounts present in the sample.

Table 10. Percentage recoveries of pesticides in serum samples and polished conventional rice diet

Pesticide	Serum sample*	Polished conventional rice diet
p-Nitrophenol	68.82 ± 2.14	119.46 ± 0.81
Carbofuran	94.62 ± 5.69	99.09 ± 7.75
Methyl parathion	86.92 ± 0.03	82.48 ± 4.45
β-Cyfluthrin	101.03 ± 4.25	43.21 ± 12.18

* obtained from rats fed with the unpolished conventional rice diet

CONCLUSIONS

Apart from providing information that may be used in the production of low-cost novel food products with sufficiently high PER values from mixtures of animal and rice proteins, this study also shows that the effect of organic rice on the growth of rat is not significantly different from that of conventional rice and that pesticide residues found in the rice samples are not detected in the serum of rats fed with either conventional or organic rice diets. Another practical benefit is that it may be used as one of the factors in the compilation of supporting data in planning for a national or international organic rice policy.

ACKNOWLEDGEMENTS

The authors would like to express sincere thanks to Kasetsart University Research and Development Institute (KURDI) for financial support. In particular, the authors are grateful to Dr. Laddawan Kunnoot (Director of the Bureau of Rice Products Development), Mr. Montri Gosalawat (Secretary General of the Progressive Farmer Association) and Mr. Ueychai Viravan for their kind provision of conventional and organic rice from Ubol Ratchathani province. Finally, the authors wish to express their deep gratitude to Dr. Warunee Varanyanond (former Director of the Institute of Food Research and Product Development, Kasetsart University).

REFERENCES

1. Food and Agriculture Organization (FAO), "All about rice", **2004**, http://www.fao.org/rice2004/en/aboutrice.htm (Accessed: September 2005).
2. C. C. Denandin, M. Walter, L. P. da Silva, G. D. Souto and C. A. A. Fagundes, "Effect of amylose content of rice varieties on glycemic metabolism and biological responses in rats", *Food Chem*., **2007**, *105*, 1474-1479.
3. M. Frei, P. Siddhuraju and K. Barker, "Studies on in vitro starch digestibility and the glycemic index of six different indigenous rice cultivars from the Philippines", *Food Chem*., **2003**, *83*, 395-402.
4. P. Hu, H. Zhao, Z. Duan, Z. Linlin and D. Wu, "Starch digestibility and the estimated glycemic score of different types of rice differing in amylose contents", *J. Cereal Sci*., **2004**, *40*, 231-237.
5. Wikipedia, "Organic farming", **2009**, http://en.wikipedia.org/wiki/organic_ farming (Accessed: March 2009).
6. STC group, "Organic rice. Health food", **2001**, http://www.capitalrice.com/products.asp?page=organic_jasmine (Accessed: March 2009).
7. N. Chopra and C. K. Hira, "Effect of roasting on protein quality of cereals", *J. Food Sci. Technol*. *India*, **1986**, *23*, 233-235.
8. A. Kaur and K. S. Sekhon, "Evaluation of nutritional quality of parboiled rice", *J. Food Sci. Technol*., **1993**, *31*, 162-164.
9. S. C. Deka, D. R. Sood and K. R. Gupta, "Nutritional evaluation of basmati rice (*Oryza sativa* L.) genotypes", *J. Food Sci. Technol.* (Mysore), **2000**, *37*, 272-276.
10. A. Lam-Sanchez, J. E. dos Santos, K. Takamura, R. M. de O. Treptow and J. E. Dutra de Oliveira, "Estudos nutricionais com arroz (*Oryza sativa*, L.)", *Alimentos e Nutricao*., **1993**, *5*, 37-48.
11. W. Horwitz and G. W. Latimer, "Official Methods of Analysis of AOAC International", 18th Edn. 2005, Association of Official Analytical Chemists, Gaithersburg, **2006**, Ch.45.
12. W. Horwitz and G. W. Latimer, "Official Methods of Analysis of AOAC International", 18th Edn. 2005, Association of Official Analytical Chemists, Gaithersburg, **2006**. Ch.4.
13. Y. Hirahara, M. Narita, K. Okamoto, T. Miyoshi, M. Miyata, S. Koiguchi, M. Hasegawa, K. Kamakura, T. Yamana and Y. Tonogai, "Simple and rapid simultaneous determination of various

pesticides of polished rice by gas chromatography", *J. Food Hyg. Soc. Japan*, **1994**, *35*, 517-529.

14. Y. Ding, C. A. White, S. Muralidhara, J. V. Bruckner and M. G. Bartlett, "Determination of deltamethrin and its metabolite 3-phenoxybenzoic acid in male rat plasma by high-performance liquid chromatography", *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **2004**, *810*, 221-227.
15. K. B. Kim, M. G. Bartlett, S. S. Anand, J. V. Bruckner and H. J. Kim, "Rapid determination of the synthetic pyrethroid insecticide, deltamethrin, in rat plasma and tissues by HPLC", *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **2006**, *834*, 141-148.
16. H. Zhang, Y. Wang and R. Wang, "HPLC determination of DCJW in rat plasma and its application to pharmacokinetics studies", *Chromatographia*, **2007**, *66*, 493-497.
17. J. N. Miller and J. C. Miller, "Statistics and Chemometrics for Analytical Chemistry", 5th Edn., Pearson Education Limited, Harlow, **2005**, pp. 54-61.
18. P. L. Pellet and V. R. Young, "Nutritional Evaluation of Protein Foods", The United Nations University, Tokyo, **1980**, pp. 41-57.
19. W. Mesomya, P. Sutthivaiyakit, Y. Cuptapun and D. Hengsawadi, "Effects of organic rice compared with conventional rice on serum lipids in rats", *Kasetsart J.*, **2009**, *43*, 703-708.
20. P. Mader, D. Hahn, D. Dubois, L. Gunst, T. Alfoldi, H. Bergmann, M. Oehme, R. Amado, H. Schneider, U. Graf, A. Velimirov, A. Flieβbach and U. Niggli, "Wheat quality in organic and conventional farming : Results of a 21 year field experiment", *J. Sci. Food Agric.*, **2007**, *87*, 1826-1835.
21. N. Ghaemia, S. S. Madaeni, M. D. Nobili and A. Alizadeh, "Ultrafiltration behavior of nitrophenols in the presence of humic substances", *J. Membr. Sci.*, **2009**, *331*, 126-136.
22. Y. Kaya, I. vergili, Z. B. Gönder and H. Barlas, "Investigation of organic matter removal from waters with adsorption polymers", *Fresenius Environ. Bull.*, **2006**, *15*, 437-440.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 483-504

Maejo International Journal of Science and Technology

ISSN 1905-7873

Available online at www.mijst.mju.ac.th

Full Paper

Fire and the production of *Astraeus odoratus* (Basidiomycetes) sporocarps in deciduous dipterocarp-oak forests of northern Thailand

Keegan H. Kennedy [1,*], James F. Maxwell [2] and Saisamorn Lumyong [1]

[1]Department of Biology, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

[2]Chiang Mai University Herbarium, Department of Biology, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

* Corresponding author, email: keegan_kennedy@yahoo.com

Received: 27 February 2012 / Accepted: 30 November2012 / Published: 3 December 2012

Abstract: The genus *Astraeus* (Diplocystidiaceae) forms ectomycorrhizal associations with many tree species and is a common gasteromycete in tropical and temperate ecosystems worldwide. In Thailand, *Astraeus* is most prevalent in deciduous dipterocarp-oak forest (DOF) in the north and north-east and its ecology is uniquely associated with fire. Rural villagers often burn the seasonally dry DOF ground vegetation causing significant environmental disturbance to promote the growth of *Astraeus* sporocraps—a local culinary delicacy and important source of household income. The purpose of this work is to investigate whether the practice of burning DOF stimulates the production of A*straeus* sporocarps in DOF. Burned and unburned *Astraeus* habitat was surveyed over two years at two sites in Chiang Mai province and one site in Mae Hong Son province. Changes in soil fungi after a fire as well as vascular vegetation growing with *Astraeus* were studied. All sporocarps collected were identified as *Astraeus odoratus*. *Astraeus* sporocarps were found in both burned and unburned areas in 2010. In 2011, an unusually wet year, no sporocarps were found in burned or unburned areas. The top 2 cm of soil experienced high temperatures which killed fungi, but lower depths were well insulated from the heat. A wide range of vascular flora grew in *Astraeus* habitat, the most common tree species being *Dipterocarpus tuberculatus* var. *tuberculatus* and *Dipterocarpus obtusifolius* var. *obtusifolius*. This study shows that *Astraeus* can produce sporocarps without fire and future work can focus on more environmentally benign methods of harvesting this popular mushroom.

Keywords: *Astraeus odoratus*, deciduous dipterocarp-oak forest, ectomycorrhizal fungi

INTRODUCTION

Deciduous dipterocarp-oak forest (DOF) is the most common forest type in South-east Asia, covering more area than any other forest type and extends from north-eastern India and Myanmar through north and eastern Thailand, southern Laos, Cambodia and southern Vietnam [1]. This forest type is mainly a result of anthropogenic disturbance and is usually associated with a hot-dry season, poor soil, and significant disturbance including logging, grazing and seasonal fires fueled by a layer of dry accumulated plant material [2-6] (Figure 1). Deciduous dipterocarp-oak forest is an important resource for local communities providing timber, grazing, wild plants and edible mushrooms.

In the north and north-eastern Thailand fire in DOF is frequently ignited by mushroom hunters who believe that it promotes the production of *Astraeus odoratus* Phosri, Watling, M.P. Martín, & Whalley sporocarps ("earthstars" or "het tawp"), which are a popular and expensive culinary delicacy [7]. Fire is lit in the dry season (February-April) and the burned soil is scraped and scoured by gatherers looking for the immature, submerged sporocarps produced at the beginning of the following rainy season (end of May-June). The genus *Astraeus* is a cosmopolitan ectomycorrhizal (ECM) fungus found throughout temperate and tropical regions of the world and has been shown to form a relationship with several tree families including Dipterocarpaceae, Fagaceae, Betulaceae and Pinaceae [8-10]. Outside of Asia *Astraeus* is not considered edible and is not associated with fire. Frequent burning by forest gatherers has deleterious ecological consequences, causing an increase in seedling and sapling mortality, loss of biodiversity and primary forest, an increase in the proportion of grassy ground flora and a decrease in soil nutrients [3,11-13]. In addition, smoke haze from fires can be an acute public health hazard, particularly in dry years [5,14]. If *Astraeus* sporocarps can be produced without fire, then a significant motivation for burning DOF can be eliminated which will contribute reforestation and restoration goals across the north and northeast of Thailand.

Despite the importance of *Astraeus* to the local economy and DOF, both as an ectomycorrhizal symbiont and as a major reason for forest fires, the ecology of this mushroom in relation to fire is not well understood. This study examines several aspects of fire-*Astraeus* dynamics in DOF, including the amount of *Astraeus* collected in burned and unburned DOF in 2010 and 2011, a molecular analysis of *Astraeus* collected from 3 previously unsurveyed areas in Mae Hong Son and Chiang Mai provinces, the effects of fire on soil characteristics and fungi in areas where *Astraeus* sporocarps were found, as well as a survey of the vascular vegetation associated with *Astraeus*. The overall objective is to test the widely held hypothesis that fire has a positive effect on *Astraeus* sporocarp production.

Figure 1. Deciduous dipterocarp-oak forest on 14 January, 2010 , before a fire with a thick layer of combustible dry vegetation (A) and on 7 June 2011 after burning and the start of the rainy season with quickly emerging herbs and leafing, coppicing trees (B). Photo by K. Kennedy (Pa Daeng National Park)

METHODS

Collection and Survey

Fresh immature *Astraeus* sporocarps were collected from three research sites: Pa Daeng National Park, Chiang Dao district, Chiang Mai province (CD); Huay Hong Krai, Doi Saket district, Chiang Mai province (DS); and Tham Pla-Namtok Pha Suea National Park, Muang district, Mae Hong Son province (MHS). Site descriptions are shown in (Table 1). Samples were collected with local villagers in May and June in 2010 and 2011 and were brought to the Sustainable Development of Biological Resources Laboratory, Chiang Mai University, where they were cleaned and their morphological characteristics (microscopic and macroscopic) were recorded and studied. The sporocarps were then dried in an oven at 45–50°C overnight and stored at the Chiang Mai University Herbarium, Faculty of Science, Chiang Mai University.

Table 1. Data for the vegetation survey plots at three sites

Plot	Elevation (m)	Easting	Northing	Ground cover (%)	Substrate
CD	625	47494739	2170774	60	Limestone/granite
MHS	500	4751456	2150325	45	Limestone/sandstone
DS	350	47522446	2090735	40	Limestone

Vegetation surveys were conducted in *Astreaus* habitat at CD (2 December 2010), MHS (8 January 2011) and DS (28 June 2011) research sites in order to identify the *Astraeus* host range. Five-meter radius plots were established at six points at each site that had been identified by villagers as areas where *Astraeus* was collected the previous year. For each plot, all woody vegetation taller than 1.5 m was measured for diameter at breast height and total height, identified and recorded. Later the Shannon and Simpson indices of diversity were calculated. Plants and cover abundance of understory vegetation shorter than 1.5 m within the plots were recorded using the Braun-Blanquet scale (x=sparse; 1=small cover; 2=5-25%; 3=25-50%). Fire history was

established based on burn evidence and villager input. The vegetation condition and cover, elevation, bedrock and UTM coordinates for each plot were also noted.

Identification of *Astraeus*

Astraeus speciation in Thailand is still unclear and morphologically similar species have been recently separated into at least three species: *A. odoratus*, *A. asiaticus* and *A. hygrometricus*. This study was the first survey of *Astraeus* sporocarps from these study sites. *Astraeus* sporocarps were collected and identified based on morphological characteristics and molecular analysis using DNA extracted following the method of Phosri et al. [9, 10]. Polymerase chain reaction (PCR) amplification was conducted using internal transcribed spacer regions of nuclear rDNA and ITS4 and ITS5 primers. The thermal conditions were 95°C for 2 minutes, 30 cycles at 95°C for 30 seconds, 50°C for 30 seconds and 72°C for 1 minute, followed by cycling at 72°C for 10 minutes. Amplicons were then examined under UV light on 1% agarose gels stained with ethidium bromide. The specimens were cleaned using the NucleoSpin^{R} Extract II Purification Kit (Macherey-Nag, Dueren, Germany) according to the manufacturer's protocol. Purified products were sequenced and determined in a genetic analyser (1st Base, Selangor, Malaysia). Sequences were used to query GenBank via BLAST (http://blast.ddbj.nig.ac.jp/top-e.html) and a phylogenetic tree was constructed using the PUAP beta 10 software version 4.0 [15].

***Astraeus* Yield at Chiang Dao and Doi Saket**

A two-year survey of *Astraeus* yields from burned and unburned DOF was conducted in May-June 2010 and May-June 2011. One large plot approximately 0.5 km^2 was established at the CD and DS research sites in DOF identified as *Astraeus* habitat by local collectors. At the end of the dry season (February-May) the burned and unburned areas within both plots were mapped by walking the boundaries of burned areas with a handheld GPS unit and the paths were uploaded to a digital elevation model using ARCGIS (ESRI, California, USA). During the period of *Astraeus* sporocarp production, three surveys were conducted by villagers and park employees on 30 May, 3 June and 9 June 2010. Leaf litter in unburned plots was removed using a rake in order to expose the soil surface and facilitate observation and collection of submerged *Astraeus* sporocarps (Figure 2). The wet and dry weights of all *Astreaus* sporocarps collected for the day were calculated for both the burned and unburned plots as well as the sizes of 50 specimens. Information on rainfall and air temperatures for 2010 and 2011 at the two sites were obtained from the Thai Meteorological Department, Chiang Mai (Figure 3).

Figure 2. Method of *Astraeus odoratus* harvesting using a metal claw, with subterranean (A, B, C) and a mature, dehisced specimen (D) from Pa Daeng National Park, 30 May, 2010. Photo by K. Kennedy

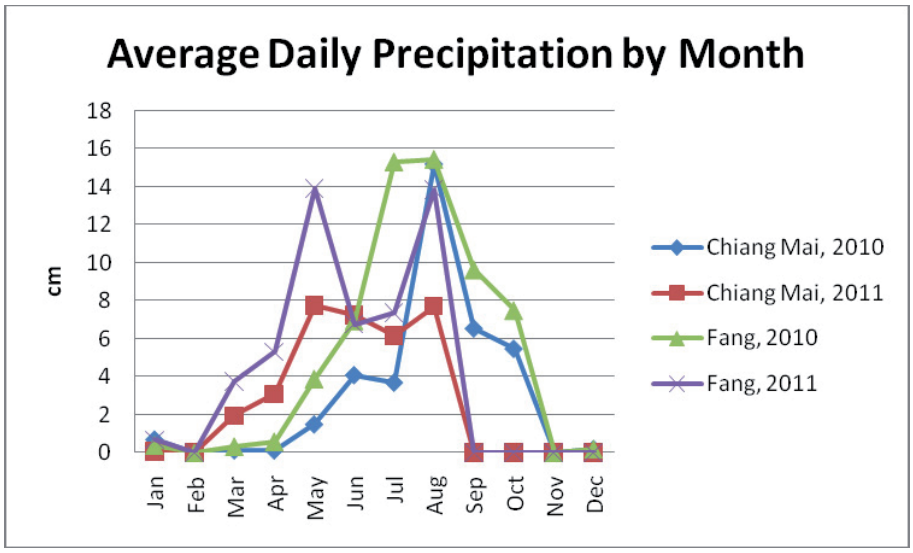


Figure 3. Average daily precipitation measured for Chiang Mai city and Fang district reflects the rainfall for the Pa Daeng and Doi Saket sites respectively

Fire and Soil Characteristics of *Astraeus* Habitat in Chiang Dao

An analysis of soil temperature during a fire event was undertaken on 15 March 2010 at the CD site in order to assess whether fire damages existing fungal hyphae in the soil. Pyrometers were constructed using Tempilaq ™ heat sensitive labels that were compressed between small sheets of aluminum and fastened with paper clips. The pyrometers covered a temperature range of 40-249°C and five sensors were placed buried 6 cm below the soil surface in a line. The sensors

were buried, the soil compacted and 6 cm of leaf litter was placed on top and burned. This method was repeated with sensors buried at the depths of 4, 2, 1, and 0 cm below the ground surface

In a separate study, the effect of heat from a fire on soil microbes was examined. A spread plate technique was used to enumerate soil microbes on potato dextrose agar (PDA). Soil samples were taken from the CD site approximately 12 hours after a fire that occurred on 27 April 2010. Soil was collected from six random sampling sites within both burned and unburned areas. Ash was removed and soil was collected from the top 2 cm of soil. The six samples were then combined and mixed into a composite sample. One gram of the composite soil was added to 10 ml of sterile water and vigorously shaken in order to make the initial solution. Serial ten-fold dilutions up to 10^{-6} were then prepared. PDA was autoclaved at 121°C and 15 psi for 15 minutes. A trace amount of 0.01% chloramphenical was added to the PDA media in order to suppress bacterial growth. Approximately 25 ml of the medium was added to each sterile Petri dish and allowed to cool to room temperature. A 1-ml pipette was then used to transfer 0.1 ml of each soil dilution onto a Petri dish and was spread over the surface with a sterile glass spreader. Each dilution was spread in triplicate for both burned and unburned soil samples. All the plates were incubated at 25°C for 48 hours and the number of colony forming units per milliliter was determined. Plates with no soil solutions were used as a control.

Locations of *Astraeus* collection points and the number of sporocarps within a 3-meter radius of the collection point was recorded with a GPS to see if there was a correlation between fire and location of *Astraeus* sporocarps**.** These locations were uploaded to a digital elevation model of each research site obtained from the Geographic Information System Centre, Geography Department, Faculty of Humanities, Chiang Mai University and analysed using ArcMap software (ESRI, Redlands, California, USA). Spatial autocorrelation was calculated using the Moran's I index.

RESULTS

Collection and Survey Results

Astraeus was found in a wide range of DOF conditions, from relatively intact forest with many trees to sparse, open canopies, with thick ground cover and eroded rocky soil at the base of 20-m tall trees. In all, 169 plant species were recorded in the 18 plots (Table 2). *Dipterocarpus tuberculatus* Roxb. var. *tuberculatus* (56 trees) followed by *Dipterocarpus obtusifolius* Teijsm. *ex* Miq. var. *obtusifolius* (34 trees) were the most common trees and at least one of these two species was present in every plot. *Shorea obtusa* Wall. *ex* Bl. (21 trees), *Gluta usitata* (Wall.) Hou, (17 trees), and *Tristaniopsis burmanica* (Griff.) Wils.var. *rufescens* (Hance) Parn. & Lug. (16 trees) were also prevalent (Table 3). The most diverse site was DS with average Shannon and Simpson indices of 1.62 and 0.22 respectively, indicating a high species richness and diversity. The CD site had the lowest diversity with a Simpson index value of 0.94 and the MHS site had the lowest richness with a Shannon index value of 1.50.

Table 2. Species and abundance of ground flora growing with *Astraeus* sporocarps by site and plot: CD in blue, MHS in red and DS in brown. For habit: a= annual; d= deciduous; e= evergreen; v= vine; h= herbaceous; wc= woody climber; t= tree; tl= treelet; sc= scandent. For abundance: x=sparse; 1=small cover; 2=5-25%; 3=25-50%

SPECIES			CD						MHS						DS					
Botanical Name	**Family**	**Habit**	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Abrus precatorius L.	Leguminosae, Papilionoideae	a, v			x		x													
Abrus pulchellus Wall. *ex* Thw. ssp. *pulchellus*	Leguminosae, Papilionoideae	a, v																x	x	x
Acacia megaladena Desv. var. *megaladena*	Leguminosae, Mimosoideae	d, wc										x								
Adiantum zollingeri Mett. *ex* Kuhn	Parkeriaceae	d, h			x			x												
Albizia chinensis (Osb.) Merr	Leguminosae, Mimosoideae	d, t																x		
Albizia odoratissima (L. f.) Bth.	Leguminosae, Mimosoideae	d, t									x									
Alpinia galanga (L.) Willd. var. *galanga*	Zingiberaceae	e, h														x	x			
Alysicarpus bupleurifolius (L.) DC.	Leguminosae, Papilionoideae	a, h								x										
Amalocalyx microlobus Pierre *ex* Spire	Apocynaceae	d, v													x					
Amorphophallus sp.	Araceae	d, h													x		x			x
Amphioneuron marginatum (Roxb.) Midd.	Apocynaceae	d, wc									x						x			
Anneslea fragrans Wall.	Theaceae	d, t				x									x		x			x
Antidesma acidum Retz.	Euphorbiaceae	d, tl									x	x	X							
Antidesma ghaesembilla Gaertn.	Euphorbiaceae	d, tl		x											x					
Antidesma soothepensis Craib	Euphorbiaceae	d, tl									x									
Apluda mutica L.	Gramineae	d, h	3	3	2		2	2			1	2								
Aporosa octandra (B.-H. *ex* D. Don) Vick. var.*octandra*	Euphorbiaceae	d, tl																x		
Aporosa villosa (Lindl.) Baill.	Euphorbiaceae,	d, tl		x		x	x				x				x	x	x	x	x	
Ardisia crenata Sims var. *crenata*	Myrsinaceae	d, wc													x	x				
Aristolochia pierrei Lec.	Aristolochiaceae	d, v	x																	
Barleria cristata L.	Acanthaceae	d, h			x															
Blumea lacera (Burm. f.) DC.	Compositae	a, h		x	1				x	x	x			x						
Blumeopsis flava (DC.) Gagnep.	Compositae	a, h	x	x	x															
Bombax anceps Pierre var. *anceps*	Bombacaceae	d, t																		
Breynia glauca Craib	Euphorbiaceae	d, tl													x	x		x	x	x

Table 2. (continued)

SPECIES			CD						MHS						DS					
Botanical Name	**Family**	**Habit**	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Buchanania lanzan Spreng.	Anacardiaceae	d, t		x		x			x	x		x	X							
Buchnera cruciata B.-H. *ex* D. Don var. *cruciata*	Scrophulariaceae	a, h				x														
Cajanus volubilis (Blanco) Blanco	Leguminosae, Papilionoideae	a, v									x	x								
Canarium subulatum Guill.	Burseraceae	d, t										x					x	x		
Capillipedium parviflorum (R. Br.) Stapf	Gramineae	d, h			x		x													
Carex continua Cl.	Cyperaceae	d, h		2																
Cassia fistula L.	Leguminsae, Caesalpiniodeae	d, t									x									
Cassytha filiformis L.	Lauraceae	e, v																x		
Catunaregam spathulifolia Tirv.	Rubiaceae	d, l		x				x	x											
Chionanthus ramiflorus Roxb.	Oleaceae	e, t															x			
Chloranthus nervosus Coll. & Hemsl.	Chloranthaceae	d, h										x								
Chukrasia tabularis A. Juss.	Meliaceae	d, t									x	x								
Cissus hastata Miq.	Vitaceae	d, wc													x	x	x			x
Clerodendrum serratum (L.) Moon var. *wallichii* Cl.	Verbenaceae	d, tl										x								
Colona floribunda (Wall. *ex* Kurz) Craib	Tiliaceae	d, t									x	x								
Commelina diffusa Burm. f.	Commelinaceae	d, h							x											
Costus speciosus (Koeh.) J.E. Sm. var. *speciosus*	Zingiberaceae	d, h		x																
Craibiodendron stellatum (Pierre) W.W. Sm.	Ericaceae	e, tl				x							X					x	x	
Cratoxylum formosum (Jack) Dyer *ssp. pruniflorum* (Kurz) Gog.	Guttiferae	d, t	l		x		x													
Crotalaria alata D. Don	Leguminosae, Papilionoideae	a, h							x	x	x									
Crotalaria albida Hey. *ex* Roth	Leguminosae, Papilionoideae	a, h	x	x	x		x	x					X	x						
Crotalaria bracteata Roxb. *ex* DC.	Leguminosae, Papilionoideae	a, h									x									
Crotalaria kurzii Baker *ex* Kurz	Leguminosae, Papilionoideae	a, h							x	x	x			x						
Crotalaria neriifolia Wall. *ex* Bth.	Leguminosae, Papilionoideae	d, h			x															

Table 2. (continued)

SPECIES			CD						MHS						DS					
Botanical Name	**Family**	**Habit**	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Croton oblongifolius Roxb.	Euphorbiaceae	d, t													x					
Curcuma plicata Wall. ex Bak. complex	Zingeberaceae	d, h							x	x			X		x	x				
Dalbergia rimosa Roxb.	Leguminosae, Papilionoideae	d, wc																		
Dalbergia cana Grah. *ex* Bth. var. *cana*	Leguminosae, Papilionoideae	d, t									x									
Dalbergia cultrata Grah. *ex* Bth.	Leguminosae, Papilionoideae	d, t	x	x	x	x	x	x							x	x	1	x		x
Dalbergia oliverii Gamb. *ex* Prain	Leguminosae, Papilionoideae	d, t							x								x			
Dalbergia velutinum DC. var. *velutinum*	Leguminosae, Papilionoideae	d, s										x								x
Desmodium heterocarpon (L.) DC. ssp. *heterocarpon* var. *birmanicum* (Watt. ex Prain) Oha.)	Leguminosae, Papilionoideae	d, h	x																	
Desmodium longipes Craib	Leguminosae, Papilionoideae	d, tl			x													x		
Desmodium motorium (Houtt.) Merr.	Leguminosae, Papilionoideae	a, h					x		x	x	x	x						x		
Desmodium oblongum Wall. *ex* Bth.	Leguminosae, Papilionoideae	d, l																		
Desmodium velutinum (Willd.) DC. var. *velutinum*	Leguminosae, Papilionoideae	d, s									2							x		
Dioscorea bulbifera L.	Dioscoreaceae	d, v													x	x	x	x	x	x
Dioscorea glabra Roxb.	Dioscoreaceae	d, v			x															
Dioscorea pentaphylla L.	Dioscoreaceae	d, v		x																
Diospyros ehretioides Wall. *ex* G. Don	Ebenaceae	d, t					x							x						
Dipterocarpus obtusifolius Teijsm. *ex* Miq. var *obtusifolius*	Dipterocarpaceae	d, t	1	1	1	x	1	1						1		x		x	x	x
Dipterocarpus tuberculatus Roxb. var. *tuberculatus*	Dipterocarpaceae	d, t	x						1	1	x	x	1		x		x			
Dumasia leiocarpa Bth.	Leguminosae, Papilionoideae	a, v									x									
Dunbaria bella Prain	Leguminosae, Papilionoideae	d, v			x												x	x		
Embelia tsjeriamcottam (Roem. & Schult.) A. DC. var. *tsjeriamcottam*	Myrsinaceae	d, wc	x																	
Eugenia cumini (L.) Druce var. *cumini*	Myrtaceae	d, t	1				x			x										x
Eulalia leschenaultiana (Decne.) Ohwi	Gramineae	d, h								x	x		X	x						
Eulalia siamensis Bor	Gramineae	d, h	3	3	2	3	2	3												

Table 2. (continued)

SPECIES			CD						MHS						DS					
Botanical Name	**Family**	**Habit**	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Eupatorium odoratum L.	Compositae	a, h	x												x		x			
Eurycoma longifolia Jack	Simoroubaceae	d, l				x												x	x	x
Flacourtia indica L.	Flacourtiaceae	d, t															x			
Flemingia sootepensis Craib	Leguminosae, Papilionoideae	d, h		x	x				x	x			X	x						
Gardenia obtusifolia Roxb. *ex* Kurz	Rubaceae	d, l								x			X	x						
Gardenia sootepensis Hutch.	Rubiaceae	d, t		x	x				x											
Geniosporum coloratum (D. Don) O. K.	Labiatae	d, h									x						x		x	
Globba nuda K. Lar.	Zingiberaceae	d, h																x	x	x
Globba reflexa Craib	Zingiberaceae	d, h															x			
Glochidion eriocarpum Champ.	Euphorbiaceae	d, t(l)									x									
Gluta usitata (Wall.) Hou	Anacardiaceae	d, t													x	x	l		x	x
Grewia abutilifolia Vent. *ex* Juss.	Tiliaceae	d, s										x			x		x			
Grewia eriocarpa Juss.	Tiliaceae	d, t										x								
Hedyotus auricularia L.	Rubiaceae	a, h						x												
Hedyotus capitellata Wall. *ex* G. Don	Rubiaceae	a, h															x			
Hedyotus tenelliflora Bl. var. *kerrii* (Craib) Fuku.	Rubiaceae	a, h	x	x	x	x		x	x	x	x	x	X	x						
Helicteres elongata Wall. *ex* Boj.	Sterculiaceae	d, ls									x	x								
Hyparrhenia rufa (Nees) Stapf var. *siamensis* Clayton	Gramineae	d, h		2																
Hypoxis aurea Lour.	Amaryllidaceae	d, h	x													x				
Inula cappa (Ham. *ex* D. Don) DC.	Compositae	a, h	2	x	x	l	l	x					X	x				x	x	
Inula indica L.	Compositae	a, h		l	x		x	x	x	x	x		X	x						
Irvingia malayana Oliv. *ex* Benn.	Irvingiaceae	e, t																x		
Kaempferia rotunda L.	Zingiberaceae	d, h																x		x
Kaempferia siamensis Siri.	Zingiberaceae	d, h													x	x	x			
Leea indica (Burm. f.) Merr.	Leeaceae	d, hs	x	x	x	l					x	x			l	x	x			
Lithocarpus polystachyus (Wall. *ex* A. DC.) Rehd.	Fagaceae	e, t																		x
Lophopetalum walichii Kurz	Celastraceae	d, t					x			x					x	x			x	

Table 2. (continued)

SPECIES			CD						MHS						DS					
Botanical Name	**Family**	**Habit**	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Lygodium polystachyum Wall. *ex* Moore	Schizaeaceae	d, v															x	x		
Lygodium flexuosum (L.) Sw.	Schizaeaceae	d, v	x		x	x	1	x	x	x	x	x	X	x	x				x	
Memecylon scutellatum (Lour.) Hk. & Arn.	Melastomataceae	e, tl																	x	
Meyna pubescens (Kurz) Roby	Rubiaceae	d, l										x								
Millettia extensa (Bth.) Bth. *ex* Baker	Leguminosae, Papilionoideae	d, wc	x	2		x		x	x	x	x	1	X	x						
Mnesithea striata (Nees *ex.* Steud.) Kon. & Sos.	Gramineae	d, h								x	x	x	X	x						
Mucna bracteata A. DC. *ex* Kurz	Leguminosae, Papilionoideae	a, v									x									
Murdannia edulis (Stokes) Fad.	Commelinaceae	d, h								x										
Mussaenda parva Wall. *ex* G. Don	Rubiaceae	d, wc																x		x
Nervilia aragoana Gaud.	Orchidaceae	d, h													x					
Ochna Integerrima (Lour.) Merr.	Ochnaceae	d, tl						x									x	x		x
Oroxylum indicum (L.) Bth. *ex* Kurz	Bignoniaceae	d, tl									x									
Paederia pallida Craib	Rubiaceae	a, v							x		x	x					x			
Panicum notatum Retz.	Gramineae	a, h									x	x								
Parinari anamensis Hance	Rosaceae	d, t																		
Pavetta fruticosa Craib	Rubiaceae	d, s		x	x		x	x												
Pavonia repanda (Roxb. *ex* J. E. Sm.) Spr.	Malvaceae	d, l,sh	x								x	x								
Pennisetum polystachyon (L.) Schult.	Gramineae	d, h			x															
Pheonix loureiroi Kunth var. *loureiroi*	Palmae	e, tl													2	1	2			
Phoebe lanceolata (Nees) Nees	Lauraceae	e, t														x	x	x		
Phyllanthus emblica L.	Euphorbiaceae	d, t		x																
Polytoca digita (L. f.) Druce	Gramineae	d, h		x	x	x		1	1	1			x		x	2	2	1	x	
Premna herbacea Roxb.	Verbenaceae	d, h							x											
Psuedopogonatherum irritans (R. Br.) A. Camus	Gramineae	d, h					x	x												
Pterocarpus macrocarpus Kurz	Leguminosae, Papilionoideae	d, t	x	x	x											x		x	x	
Pueraria stricta Kurz	Leguminosae, Papilionoideae	d, sc						x												
Quercus brandisiana Kurz	Fagaceae	d, t																x		
Quercus kerrii Craib	Fagaceae	d, t	1	1	x	x		1							x	x		x	x	1

Table 2. (continued)

SPECIES			CD						MHS						DS					
Botanical Name	**Family**	**Habit**	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Sacciolepis indica (L.) A. Chase	Gramineae	a, h				x														
Sauroporus quadrangularis (Willd.) M.-A.	Euphorbiaceae	d, h									x									
Scleria levis Retz.	Cyperaceae	d, h							x	x			x	x						
Scleria lithosperma (L.) Sw.	Cyperaceae	e, h																	x	x
Scleria terrestris (L.) Fass.	Cyperaceae	d, h		2	x	1	2													
Selaginella ostenfeldii Hier.	Selaginellaceae	d, h							x	x			x	x						
Shorea obtusa Wall. *ex* Bl.	Dipterocarpaceae	d, t	2	1	1			1	x	x		x	x	x	x		x		x	
Shorea siamensis Miq. var. *siamensis*	Dipterocarpaceae	d, t	x	x	1		1	1										x		
Smilax lanceifolia Roxb. var. *lanceifolia*	Smilacaceae	e, v															x			
Smilax ovalifolia Roxb.	Smilacaceae	d, v								x		x				x				
Smilax verticalis Roxb.	Smilacaceae	d, v		x																
Sonerila erecta Jack	Melastomataceae	a, h				x														
Sorghum verticilliflorum (Steud.) Stapf	Gramineae	d, h										x								
Spatholobus parviflorus (Roxb.) O.K.	Leguminosae, Papilionoideae	d, wc	x	2				x							x	x		x		
Sterculia balanghas L.	Sterculiaceae	d, t	x																	
Streptocaulon juventas (Lour.) Merr.	Asclepiadaceae	e, v			x	2														
Strobilanthes apricus (Hance) T. And. var. *pedunculatus* (Craib)Ben.	Acanthaceae	d, s							x		x		x	x						
Strobilanthes glaucescens Nees	Acanthaceae	d, h										x								
Strychnos nux-vomica L.	Loganiaceae	d, t													x	x	x			
Symplicos racemosa Roxb.	Symplocaceae	d, tl	x				x													
Tectona grandis L. f.	Verbinaceae	d, t										x								
Terminalia alata Hey. *ex* Roth	Combretaceae	d, t							x					x						
Terminalia chebula Retz. var. *chebula*	Combretaceae	d, t																		
Terminalia mucronata Craib & Hutch.	Combretaceae	d, t	x	x																

Table 2. (continued)

SPECIES			CD						MHS						DS					
Botanical Name	**Family**	**Habit**	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Thespesia lampas (Cav.) Dalz. & Gibs. var. *lampas*	Malvaceae	d, l										x								
Thunbergia alata Boj. *ex* Sims	Acanthaceae	e, v	x																	
Thysanolaena latifolia (Roxb. *ex* Horn.) Honda	Gramineae	e, h																x		
Tristaniopsis burmanica (Griff.) Wils. & Wat. var. *rufescens* (Hance) Parn. & Lug.	Myrtaceae	e, tl	1		x		2						x					x	1	
Vaccinium sprengelii (D. Don) Sleum.	Ericaceae	e, tl	x				x													
Vernonia parishii Hk. f.	Compostae	d, lh	1	x														x	x	
Vernonia squarrosa (D. Don) Less. var. *orientalis* Kit.	Compostae	d, h							x	x		x	x							
Vigna dalzelliana (O.K.) Verd. var. *dalzelliana*	Leguminosae, Papilionoideae	a, v									x	x								
Vitex peduncularis Wall. *ex* Schauer	Verbenaceae	d, t	x	x											x					
Walsura trichostemon Miq.	Meliaceae	e, t					x													
Wendlandia tinctoria (Roxb.) DC.	Rubiaceae	e, tl	x		x	x		x	x	x			x		1	x	x			
Xylia xylocarpa (Roxb.) Taub. var. *kerrii* (Craib & Hutch.) Niels.	Leguminosae, Mimosoideae	d, t										x								
Zingiber bradleyanum Craib	Zingiberaceae	d, h	x																	
Ziziphus rugosa Lmk. var. *rugosa*	Rhamnaceae	d, t													x	x				

Table 3. Most common tree species based on number of trees in 5-m radius plots

		Site			
Species	Family	Chiang Dao	Mae Hong Son	Doi Saket	Total
Dipterocarpus tuberculatus var. *tuber.*	Dipterocarpaceae	0	40	16	56
Dipterocarpus obtusifolius var. *obtus.*	Dipterocarpaceae	13	0	21	34
Shorea obtuse	Dipterocarpaceae	5	10	6	21
Gluta usitata	Anacardiaceae	0	0	17	17
Tristaniopsis burmanica	Myrtaceae	7	0	9	16
Dalbergia cultrate	Leguminosae, Papilionoideae	0	0	13	13
Aporusa villosa	Euphorbiaceae	3	0	6	9
Quercus kerrii	Fagaceae	7	0	2	9
Pterocarpus macrocarpus	Leguminosae, Papilionoideae	7	0	0	7
Buchanania lanzan	Anacariaceae	4	2	0	6
Simpson Index		0.94	0.35	0.22	
Shannon Index		1.55	1.50	1.62	
Total species per site		49	64	98	

The ground flora was also significantly represented by saplings (mostly coppices) of *D. tuberculatus* var. *tuberculatus, D. obtusifolius* var. *obtusifolius, Shorea siamensis* Miq. var. *siamensis, S. obtusa,* and *Quercus kerrii* Craib, which were all present as coppices and seedlings comprising 1-5% of the total cover of more than one plot. *Dalbergia cultrata* Grah. *ex* Bth., *Lygodium flexuosum* (L.) Sw. and *Polytoca digita* (L f.) Druce, a deciduous tree, deciduous vine and deciduous grass respectively were present as ground cover in 12 of the 18 plots though in low abundances. The most commonly represented plants in the plots were deciduous herbs, with 39% of all plot records, followed by deciduous trees with 25%. Most species had low abundances (≤1% cover area). Many grasses were consistently present in high densities, including *Eulalia siamensis* Bor and *Apluda mutica* L., which had densities of 25-50% in several plots. The percentage of ground cover in the plots varied greatly with a range of 10-95% and an average of 51%.

Astraeus Identification Results

The morphological characteristics of sporocarps collected were consistent at all three collection sites. The sub-globose fruiting bodies were approximately 2.0 x 2.0 x 1.5 cm. The outer peridium was light brown in color changing to darker brown after maturing. The endoperidium was dark brown with no peristome. The gleba of immature sporocarps was white or cream changing to dark brown-black in mature specimens. Basidiospores were globose and 7-15 µm in diameter (Table 4).

Table 4. Morphology of *Astraeus odoratus* sporocarps

Basidiomes	Sub-globose, approximately 2.0 x 2.0 x 1.5 cm. Strongly hydroscopic and emits a musty soil odor.
Outer peridium	Light brown surface, smooth and rubbery when submerged, hardening and splitting radially after emergence
Endoperidium	Dark brown to black with no peristome
Gleba	White or cream when immature, changing to dark brown-black
Basidiospores	globose, 7.5-15 μm
Habitat	Deciduous dipterocarp-oak forest on sandy, rocky and clay soils from May to July
Distribution	North and north-eastern Thailand

For molecular identification, the internal transcribed spacer 1, 5.8S ribosomal DNA gene and internal transcribed spacer 2 sequences of samples from the CD (CMU-HP9), DS (CMU-HP11), and MHS (CMU-HP8) sites were deposited in GenBank as JQ292818, JQ292819 and JQ292817 respectively. The aligned dataset of these 17 sequences consisted of 820 characters, of which 413 characters were constant, 180 variable characters were parsimoniously uninformative, and 227 characters were parsimoniously informative. Heuristic searches resulted in nine equally parsimonious trees with a length of 635 steps (CI=0.863, RI=0.881, RC=0.7603 and HI=0.137), one of which is shown in Figure 4. Phylogenetic analyses by Phosri et al. [9, 10] indicated that *Astraeus* species were divided into 4 clades including *A. ordoratus* (clade A), *Astraeus pteridis* (Shear) Zeller (clade B), *Astraeus hygrometricus* Pers. Morgan (clade C) and *Astraeus asiaticus* Phosri, Watling, M.P.Martín, & Whalley (clade D). All *Astraeus* samples collected in this study (Kennedy 1, 2 and 3, from the CD, DS and MHS sites respectively) belonged to *A. odoratus* in clade A, which is a sister group of *A. pteridis* with 82% bootstrap support.

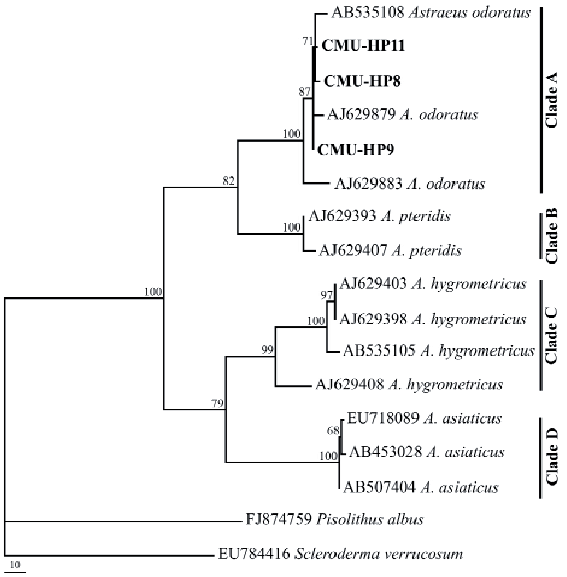


Figure 4. One of the six parsimonious trees inferred from a heuristic search of the internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2 sequence alignments of 17 isolates

Astraeus Yield at Chiang Dao and Doi Saket

Many immature *Astraeus* sporocarps were found at the CD and DS sites in 2010 and none were found at either site in 2011. The weather, particularly in March-May, was significantly different during these two years. Year 2010 was an unusually dry year with a long fire season and a late rainy season starting mid-May. In contrast, 2011 exhibited an early and prolonged rainy season starting at the beginning of March and there were no fires.

In 2010 sporocarps were found in abundance in burned areas (2,233 g and 1,015 g total cumulative wet weight for CD and DS respectively, or 9.75 x 10^{-3} g/m^2) and with a much lower frequency in unburned areas (22 g and 14 g wet weight for Pa Daeng and Doi Saket, respectively, or 3.99 x 10^{-3} g/m^2). In 2010, 97% of the plots were burned (Figure 5). *Astreaus* sporocarps were first observed at the CD site at the end of May and could be found in abundance for 10 days, after which the plots had been completely scoured and dug up by mushroom hunters. During this period, three surveys were conducted with the help of the national park employees. The highest yield was obtained in burned areas on the first survey day and decreased thereafter (1028 g, 682 g and 523 g wet weight for 30 May, 3 June and 9 June 2010 respectively). Only one cluster of *Astraeus* sporocarps with a total wet weight of 22 g was found in the unburned area (30 May) (Figure 6).

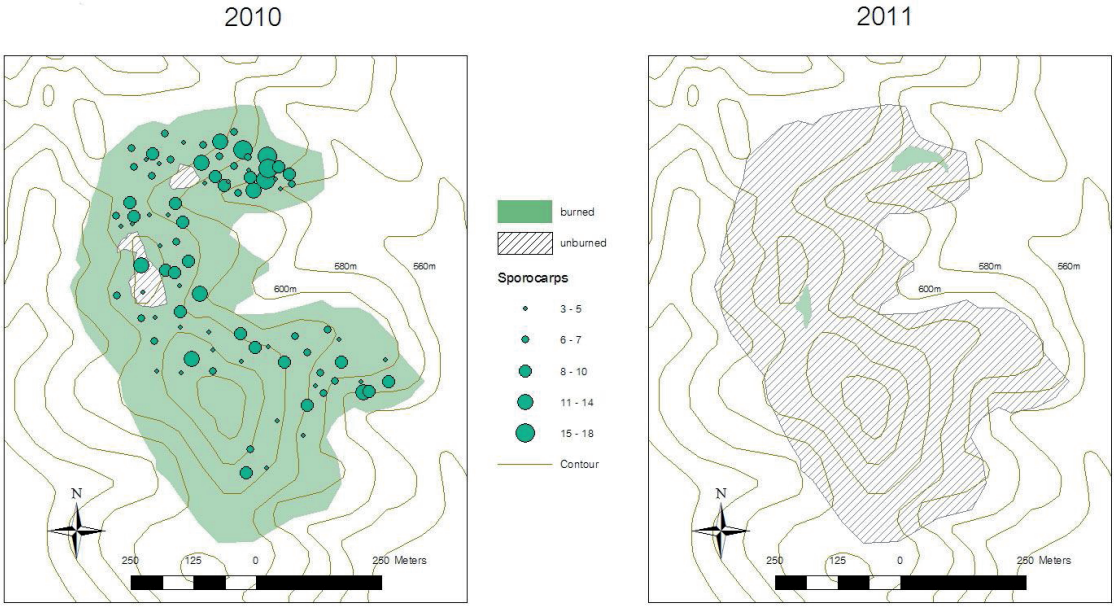


Figure 5. The extent of burning and *Astraeus odoratus* sporocarps collected in Pa Daeng National Park, Chiang Dao district, Chiang Mai province in May and June 2010 and 2011

The DS site collection profile was similar to that at CD. Most (94%) of the plot was burned in 2010. The rainy season in the more southern Doi Saket site started several weeks later and *Astreaus* sporocarps first appeared at the end of June. Two surveys were conducted with the help of locals on 20 June (1.7x10^{-3} and 4.1x10^{-5} g/m^2) and 25 June 2010 (1.2x10^{-3} and 0 g/m^2) in burned and unburned areas respectively.

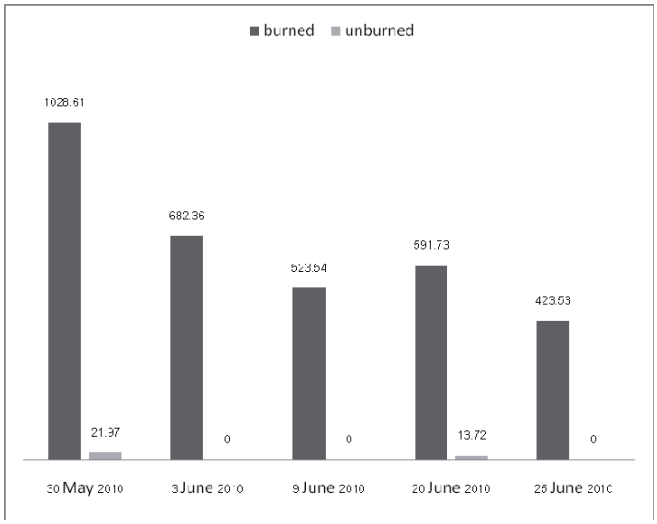


Figure 6. Yields of *Astraeus* in burned and unburned areas in grams for 2010

The sporocarp size from 50 random specimens taken from burned areas each day averaged 6.3 cm^3 and 6.16 cm^3 for CD and DS respectively. Average sizes for sporocarps collected in unburned areas were smaller, 1.89 cm^3 and 0.245 cm^3 for CD and DS respectively. Sporocarps from burned areas were heavier. The average fresh weight of individual sporocarps from CD were 2.5 and 1.3 g for burned and unburned areas respectively. The average fresh weight of individual sporocarps from DS was 3.6 g and 1.4 g for burned and unburned areas respectively.

Fire and Soil Characteristics of *Astraeus* Habitat at Chiang Dao

Soil temperature experiment showed that the surface of the soil profile became extremely hot during a fire (greater than 249°C, which was the maximum limit of the pyrometer). Below 2 cm there was very little change in the soil temperature from fire. The average temperatures were >249, 122.4, 66.3, 41.6 and <40°C at the depths of 0, 1, 2, 4 and 6 cm respectively (Figure 7). Soil temperature without fire never exceeded 40°C during the experiment. Soils were compact eroded limestone with some granite.

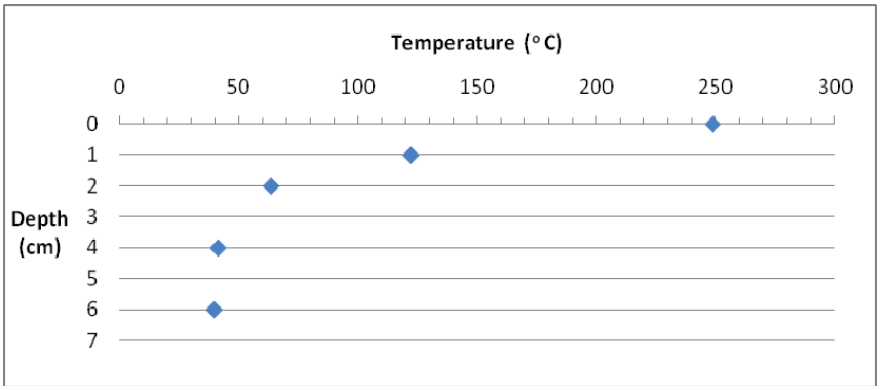


Figure 7. Soil temperatures at different depths during a fire

The soil fire-microbe analysis showed that fire has a negative effect on fungi, which is significant as during a dry year areas are sometimes burned many times. An average of 883 cfu's were found in unburned soil and 162 in burned soil. The control plates containing no soil solutions were blank indicating no contamination. Dilution levels of $1x10^{-3}$ and $1x10^{-4}$ elicited distinct and countable fungal colonies for the unburned soil, while less diluted levels ($1x10^{-2}$ and $1x10^{-3}$) were satisfactory to count the colony forming units in burned soil samples. This indicates that fire has a direct and immediate negative effect on fungi growing at the top layer of soil.

Astraeus sporocarps were always found in clusters of 3 or more. In addition, spatial autocorrelation analysis of *Astraeus* sporocarp collection points and the amount collected showed that there is a clustering pattern. The Moran I Index, which measures dispersion, was 0.14 with a z score of 2.11 indicating that there is less than 5% likelihood that the distribution of sporocarps bunches are themselves clustered.

DISCUSSION

The genus *Astraeus* is associated with a wide range of tree species around the world and particularly dipterocarps in Thailand. However, knowledge of its hosts is far from complete. In vitro experiments have shown that *Astraeus* can form associations with *Dipterocarpus alatus* Roxb. *ex* G. Don, *Eucalyptus camaldulensis* Dehnh. (Myrtaceae) and *Pinus densiflora* Sieb. and Zucc. (Pinaceae) [16, 17]. Our plots were dominated by *Dipterocarpus tuberculatus* var. *tuber*culatus, *Dipterocarpus obtusifolius* var. *obtusifolius* and *Shorea obtusa* as well as *Quercus kerrii* Craib, *Gluta. usitata* (Wall.) Hou and *Dalbergia cultrata* Grah. *ex* Bth. Our study also included many grasses and other herbs, treelets and vines consistently growing in areas that produce *Astraeus* sporocarps. Most of the focus has been on tree species as *Astraeus* hosts. However, it is possible that non-arboreal species are also *Astraeus* hosts and, having faster growth rates, are better candidates for inoculation trials and can potentially lead to commercial *Astraeus* production.

The genus *Astraeus* in Asia has been taxonomically studied in the last decade. *Astraeus* in Thailand was thought to be only *A. hygrometricus*. Phosri et al. [9, 10] showed that *A. odoratus* is a separate species and is commonly found throughout the north and north-east. Morphologically the two species are very similar but *A. hygrometricus* sporocarps tend to have a whiter colour and more basidial rays once they open (5-12 compared to 3-9 rays for *A. hygrometricus* and *A. odoratus* respectively). At a microscopic scale, the basidiospores of *A. odoratus* are slightly smaller (8.75-15.2 µm versus 7.5-12 µm for *A. hygrometricus* and *A. odoratus* respectively) and have longer and narrower spines. *Astraeus asiaticus* is present in Thailand but was not encountered in this study.

This survey of *A. odoratus* sporocarp yields over two years with very different rainfall and consequent fire regimes allows for interesting insights into the ecology of this species. It was found that fire is not an absolute requirement for *A. odoratus* sporocarp production, although this fungus is likely adapted to produce sporocarps in dry soil conditions which can result from a fire. Sporocarps were found in unburned patches of DOF in 2010, but in 2011 they were nonexistent in the small patches of burned forest. This suggests that other factors are influencing sporocarp production along with fire. Rainfall for 2010 and 2011 were very different. Between January and May 2010, 76.3 cm of rain had fallen, while at that same time in 2011, 396.5 cm had fallen. This rainfall variation had a significant effect on the fire regime since in 2010 fires were pervasive and many areas were burned twice in the same year. In contrast, fires in 2011 were almost absent as the leaf litter never reached a critically low moisture level that would allow it to sustain a fire. The amount,

average fresh weight, and dimensions of sporocarps from the burned area were significantly larger than those from the unburned area in 2010. The burned area surveyed for 2010 was much larger than the unburned area so it is difficult to make direct comparisons between the two conditions.

Soil at CD from *A. odoratus* habitat showed a sharp decline in nitrogen after a fire (Figure 8). Trappe and Cormack [18] conducted soil analyses and examined fungal communities in Ponderosa pine forests in Oregon subjected to controlled burns and found that groups of fungal species were present at either above or below a C:N threshold ratio of 26:1. Of 24 species of mushrooms collected only one species was present in soils both above and below the C:N ratio of 26:1; all other species were present in either one or the other, but not both. Autumnal burns significantly reduced the C:N ratio and produced specific fungal species that were also present in unburned areas with a similarly low C:N ratio. In this study, the C:N ratio changed in burned *Astraeus* habitat immediately after a fire as nitrogen was lost. Soil nutrient changes and moisture as catalysts for *Astraeus* sporocarp production are topics that can be explored in future research.

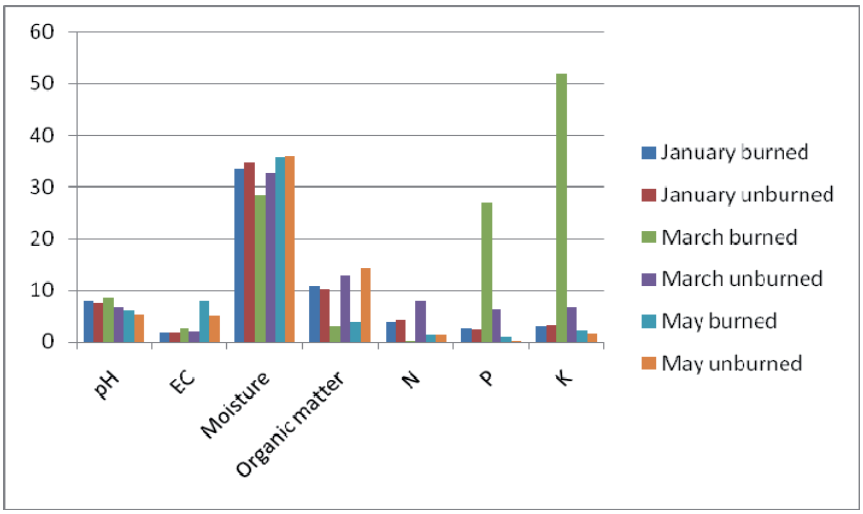


Figure 8. Changes in soil characteristics before and after a fire at CD

This survey of *Astraeus* sporocarp production in burned and unburned areas was complicated by many other people collecting *Astraeus* in the surveyed area. Locals were employed to assist with collection, but many people came from other places and entire tracts of forest were quickly scoured during the short period in which *Astraeus* sporocarps are underground and still edible. The collection method involves digging up areas with a metal claw or crowbar (Figure 2). This can have an adverse effect on the underlying host roots, soil structure and hyphae. Furthermore, only immature sporocarps are collected which do not have a chance to release spores before they are removed from the forest. Only sporocarps that are overlooked or emerge after an area has been harvested have a chance to open and spread their spores. Whether or not harvesting causes recruitment limitation is a topic for further research.

Spatial autocorrelation analysis of *Astraeus* sporocarp clusters collected at CD were not randomly distributed since the method of collection was not entirely random as we were collecting with locals who had specific places that they knew were areas of high yield. At the CD site gatherers tended to stay along the ridge running through the plot and generally avoided the steep

slopes of the western and eastern sides, or ascending the steep knoll in the southern end of the plot. Locals generally avoided looking extensively in unburned areas. There was a consensus that *Astraeus* does not need fire to produce sporocarps, although the yield was much greater in burned areas. Whether this is a result of specific conditions created by fire that induce sporocarp formation or whether fire simply makes it easier to find subterranean sporocarps once the leaf litter has been removed can be examined in future research.

A small amount of *Astraeus* was found in unburned patches with particularly gravely soil at the CD site in 2010. At the DS site *Astraeus* was found under a thick leaf litter layer with rich soil. Locals highlighted areas under burned logs as places likely to produce a high yield and this is something that can be further examined. Based on our initial survey the soil and vegetation conditions where *Astraeus* sporocarps were found varies substantially. Helfferich [19] mapped the location of individual *Morcella* sp. basidia after a fire in Alaska using a GPS accurate to 5 cm. They found that the distribution was very clustered though the mechanism that caused sporocarp development in the specific areas was unclear.

Fire, a frequent and characterising component of DOF, has been previously studied. Stott [6] found that a heat sensitive paint with a threshold of 35°C at a depth of 5 cm was not triggered during a fire. Wanthongchai et al. [12] showed that the length of period between fires did not significantly affect soil temperatures in DOF. They measured ambient soil temperature with thermocouples 5 minutes after a fire and found that soil at 1 cm did not exceed 48^0C in either frequently or infrequently burned stands. Our experiment using gradient indicators at a range of depths showed that at depths lower than 2 cm heat from a fire will not always be significant, but at 2 cm and up the soil can reach temperatures that kill mycelia. The soil structure of DOF, which is often very compact, rocky, eroded and has little organic matter, causes minimal heat transfer into the soil profile and allows for the survival of underground parts of many DOF plants and fungi.

Though the effects of fire on soil microbe communities has been extensively examined in other parts of the world, particularly in boreal forests in North America, little research has been conducted on this topic in South-east Asian DOF. Treseder et al. [20] used hyphae length of soil fungi and microbial respiration to estimate abundance in chronosequenced stands of boreal forests in Alaska. They found that a year after a fire there was no significant difference in fungal biomass, although after 6-25 years of a fire there was a significantly lower fungal biomass than the control site which had not been burned in 85 years. Our results showed that the amount of fungi in DOF surface soil decreased immediately after a fire. Widden and Parkinson [21] also showed that the fungi composition in boreal forests changed by fire. They grew soil fungal cultures in Petri dishes with aqueous extracts of burned and unburned leaf litter to show that certain fungi including *Trichoderma polysporum* and *Penicillium janthinellum* were inhibited by the fire. *Cylindrocarpon destructans* was not inhibited and *Gelasinospora* sp. was only found in burned plots. Our research showed that initially after a fire, there was a substantial decrease in the overall level of fungi in surface soil, but further research is needed in order to determine the long-term trends in DOF and the specific changes in fungal composition resulting from fire.

CONCLUSIONS

This study is the first analysis of how fire affects *Astraeus*, an important genus found throughout much of the world. The flora associated with ectomycorrhizal *Astraeus*, its taxonomy in

northern Thailand, a comparison of *Astraeus* sporocarp yields in burned and unburned DOF, the soil temperature during a fire and the change in surface soil fungi resulting from fire are quantified.

It was shown that fire is not necessary for the production of *Astraeus* sporocarps but it seems to promote more and larger sporocarps. Heat from fire can have negetative affects on soil fungi and soil quality in the upper soil layer where *Astraeus* hyphae can be found. It is possible that fire indirectly increases harvest yield by eliminating the leaf litter layer and making submerged sporocarps easier to find. The years 2010 and 2011 had drastically different rainfall conditions and *Astraeus* yields, suggesting that other factors such as soil moisture and length of the hot-dry season are also important. A long term survey of the yield and spatial distribution of *Astraeus* clusters in relation to weather conditions and with greater control over the size of burned and unburned areas should be continued at the CD site.

ACKNOWLEDGEMENTS

This work was supported by a grant from The Higher Education Commission, Thailand, under the National Research University (A1) Program and the Graduate School of Chiang Mai University. Jaturong Kumla and Nakarin Suwannarach are thanked for their help with molecular analysis.

REFERENCES

1. P. Rundel and K. Boonpragob. "Dry forest ecosystems of Thailand," in "Seasonally Dry Tropical Forests", (Ed. S. Bullock, H. Mooney and E. Medina), Cambridge University Press, Cambridge, **2005**, pp. 112-113.
2. J. F. Maxwell and S. Elliott, "Vegetation and vascular flora of Doi-Sutep-Pui National Park, northern Thailand", *Thai Stud. in Biodiv.*, **2001,** *5*, 1-15, 40-45.
3. J. F. Maxwell, "Lowland vegetation of Doi Chiang Dao Wildlife Sanctuary, Chiang Mai province, Thailand", *Tiger Paper (FAO),* **1993**, *19*, 21-25.
4. J. F. Maxwell, "A synopsis of the vegetation in Thailand", *Nat. Hist. J. Chulalongkorn U.*, **2004**, *4*, 19-29.
5. A. Putnam, "Burning as usual: The causes and effects of deliberate fire burning", Privately published, Chiang Mai, **2009**.
6. P. Stott, "The forest as *Phoenix*: Towards a biogeography of fire in mainland South East Asia", *Geogr. J.,* **1988**, *154*, 337-350.
7. B. Dell, R. Sanmee, P. Lumyong and S. Lumyong, " Ectomycorrhizal fungi in dry and wet dipterocarp forests in northern Thailand - diversity and use as food", Proceedings of Asia Pacific Association of Forestry Research Institution's 8th Roundtable Conference on Dipterocarps: Enhancing Capacities in Sustainable Development and Poverty Alleviation, **2005**, Ho Chi Minh City, Vietnam.
8. E. R. Nouhra and L. Dominguez de Toledo, "The first record of *Atraeus hygrometricus* from Argentina", *Mycologist*, **1998**, *12*, 112-113.
9. C. Phosri, R. Watling, M. P. Martin and A. J. S. Whalley, "Molecular study of the genus *Astraeus*", *Mycological Res.,* **2004**, *3*, 275-286.

10. C. Phosri, R. Watling, M. P. Martin and A. J. S. Whalley, "The genus *Astraeus* in Thailand", *Mycotaxon*, **2007**, *89*, 453-463.
11. S. K. Kafle, "Effects of forest fire protection on plant diversity in a tropical deciduous dipterocarp-oak forest, Thailand", *Int. Forest Fire News,* **2006**, *34*, 64-71.
12. K. Wanthongchai, J. Goldammer and J. Bauhus, " Effects of fire frequency on prescribed fire behavior and temperatures in dry dipterocarp forests*", Int. J. Wildland Fire,* **2011**, *20*, 35-45.
13. C. Vaidhayakarn and J. F. Maxwell, "Ecological status of the lowland deciduous forest in Chang Kian Valley, Chiang Mai, northern Thailand", *Maejo Int. J. Sci. Technol.,* **2010**, *4*, 268-317.
14. U. Vinitketkumnuen, K. Kalayanamitra, T. Chewonarin and R. Kamens**,** " Particulate matter, PM 10 and PM 2.5 levels, and airborne mutagenicity in Chiang Mai, Thailand", *Mutat. Res. Genetic Toxicol. Env. Mutagen.*, **2002**, *519*, 121-131.
15. D. L. Swofford, "Phylogenetic analysis using parsimony (and other methods). Version 4", Sunderland, MA: Sinauer Associates, **2002**.
16. U. Sangwanit and T. Sangthien, "Ectomycorrhizae of *Dipterocarpus alatus* Roxb", Proceedings of the BIO-REFOR Pre-Workshop, **1991**, Bogor, Indonesia.
17. W. Fangfuk, K. Okada, R. Petchang, C. To-anun, M. Fukada and A. Yamada. "*In vitro* mycorrhization of edible *Astraeus* mushrooms and their morphological characterization", *Mycoscience,* **2010**, *51*, 234-241.
18. M. Trappe and K. Cormack, "The forest, the fire and the fungi: Studying the effects of prescribed burning on mycorrhizal fungi in Crater Lake National Park", *Fire Sci. Brief*, **2009**, *38*, 1-6.
19. D. Helfferich, "Morels: A morsel after the fire", *Agroborealis,* **2005**, *35*, 4-9.
20. K. Treseder, M. Mack and A. Cross, "Relationships among fires, fungi, and soil dynamics in Alaskan boreal forests", *Ecol. Applic.*, **2004**, *14*, 1826-1838.
21. P. Widden and D. Parkinson, "The effects of a forest fire on soil microfungi", *Soil Bio. Biochem.,* **2011**, *78*, 125-138.

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 505-513

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Development of ginger-flavoured soya milk ice cream : Comparison of data analysis methods

Wiwat Wangcharoen

Faculty of Engineering and Agro-Industry, Maejo University, Chiang Mai 50290, Thailand

E-mail: wiwat@mju.ac.th

Received: 1 October 2012 / Accepted: 25 December 2012 / Published: 28 December 2012

Abstract: Sucrose at the concentration of 6 and 7% (w/w) and ginger extract at the concentration of 4 and 5% (w/w) were added to ginger-flavoured soya milk ice cream recipes to determine the consumer acceptablility. One hundred consumers were requested to taste four samples of ginger-flavoured soya milk ice cream and rate the intensity of ginger flavour, sweetness, richness and smoothness, together with their indication for ideal intensity of these attributes as well as their preference of each sample by using the 10 cm-line scale. Ideal ratio scores, principal component analysis and analysis of variance with mean comparison were used for data analysis. It was shown that the recipe with 7% sucrose and 4% ginger extract was close to the ideal product ($p>0.05$) and it had the highest preference mean score ($p\leq0.05$). Total phenolic content of this recipe was 91.6 $\pm$ 6.8 mg gallic acid equivalent per 100 grams and antioxidant capacity values including ferric reducing/antioxidative power (FRAP), 1,1-diphenyl-2-picryl-hydrazyl (DPPH) and 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) were 37.9$\pm$3.7, 13.4$\pm$1.2 and 49.0$\pm$5.1 mg vitamin C equivalent per 100 grams respectively.

Key words: ice cream, ginger-flavoured ice cream, soya milk ice cream, principal component analysis

INTRODUCTION

Ice-cream is normally defined as a frozen mixture of milk components, sweeteners, stabilisers, flavourings and other ingredients [1]. There are a number of related products such as milk ice, water ice and sherbets, etc., which primarily differ in the relative quantities of ingredients. Ice cream ingredients, especially milk fat and milk solid, are used for product classification in accordance with legislation [1,

2]. According to the Ministry of Public Health Announcements, Issue 222 and 257 (Thailand), there are 5 categories, viz. dairy ice cream, modified fat ice cream, mixed ice cream, ice cream in liquid or dried or powder forms, and sweet and cold ice cream from non-dairy products [3, 4]. Ice cream consumption by Thais is increasing and the market size is forecasted to be 15,000-16,500 million Baht in 2012 [5, 6].

Thai people are familiar with non-dairy ice cream. Coconut milk ice cream is well known and preferred by Thais and foreigners [7, 8]. Other kinds of non-dairy ice cream have been developed such as brown rice ice cream [9], Job's-tears ice cream [10] and soya milk ice cream [9, 11]. Soya milk can be used as a good dairy substitute for ice cream making because it is creamy but soya milk ice cream from the freezer may be hard and need about 15 minutes to soften before serving [12]. In addition, soya milk is a liquid extract of soya bean (*Glycine max* (L) Merrill), a good dietary source containing almost all components of soya bean which are beneficial to health, such as peptide and protein, lectin, trypsin inhibitor, dietary fibre, oligosaccharide, phytin, saponin, isoflavone, linoleic acid, α-linolenic acid, lecithin, tocopherol, plant sterol, vitamin K and magnesium [13, 14]. Several approaches have been undertaken to include more soya bean in the diet for better health of the people and new soya recipes have been developed.

A recipe for ginger-flavoured soya milk ice cream was first developed by Leelarattanakul and Hasun [15]. Ginger (*Zingiber officinale* Rosc.) is used as a cooking spice and herbal remedy. For therapeutic purposes, fresh ginger is used as antiemetic, antiussive, expectorant and is used to induce perspiration and dispel cold, whereas dried ginger is used for stomachache, vomiting and diarrhoea [16]. Fresh ginger is juicy, spicy, refreshing and slightly sweet and has lemon-like aroma and strong bite and it is more aromatic than dried ginger [17]. Ginger contains volatile essential oil and non-volatile compounds such as oleoresins (gingerol and shogaol, the pungent principles of ginger), and other usual organic and inorganic compounds found in food, especially vitamin C, manganese and iron [18].

This study is aimed to improve the ginger-flavoured soya milk ice cream of Leelarattanakul and Hasun [15] by varying the proportion of ginger extract and sucrose to make it more acceptable by consumers. Sensory rating was used to evaluate the products and data analysis was done by three different methods: ideal ratio scores, principal component analysis and analysis of variance, together with mean comparison. Total phenolic content and antioxidant capacity of the selected product were also determined.

MATERIALS AND METHODS

Ingredients and Ice Cream Mix Preparation

Soya bean (Aro brand), coconut milk (Chaokoh brand), tapioca flour (Pagoda brand), glucose (Pure Chem), sucrose (Mitr Phol brand) and mature fresh ginger were used as ingredients.

Soya beans were washed and soaked in water for 6 hours before blending with water at a ratio of 1:4 (soya bean:water). The blended mixture was filtered and the collected soya milk was boiled for 20 minutes before the addition of coconut milk, dispersed tapioca flour, glucose and varied amounts of sucrose and ginger extract (prepared from mature fresh ginger by a juice extractor) as shown in Table 1. The mixture was stirred and heated at 75°C for 15 minutes before rapid cooling and keeping overnight

in the refrigerator. Ice cream was made by a compressor ice cream maker (JCS Technic Line Co., Ltd.) and it was packed in plastic boxes and kept in the freezer for at least 1 week [15].

Table 1. Composition of ice cream mix (gram (%w/w))

Recipe	Soya milk	Coconut milk	Tapioca flour	Glucose	**Sucrose**	**Ginger extract**
1	3,500 (61)	1,200 (21)	120 (2)	350 (6)	**350 (6)**	**250 (4)**
2	3,500 (60)	1,200 (21)	120 (2)	350 (6)	**400 (7)**	**250 (4)**
3	3,500 (60)	1,200 (21)	120 (2)	350 (6)	**350 (6)**	**300 (5)**
4	3,500 (60)	1,200 (20)	120 (2)	350 (6)	**400 (7)**	**300 (5)**

Sensory Measurement and Chemical Analysis of Selected Recipe

One hundred volunteers were requested to evaluate the flavour and texture, viz. ginger flavour, sweetness, richness [1] and smoothness [19], including their own ideal intensity of each attribute and their preference of each recipe by using the 10 cm-line scale (0 = least, 10 = most).

Total phenolic content was determined by Folin-Ciocalteau micro method [20] and antioxidant capacity of the selected recipe was evaluated by 3 different methods, namely ferric reducing/antioxidative power (FRAP) assay [21], 1,1-diphenyl-2-picryl-hydrazyl (DPPH) free radical scavenging activity [22] and improved 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) radical cation decolourisation assay [23]. All analyses were done with some modifications [10].

Statistical Analysis

Sensory data was analysed by 3 different methods: ideal ratio scores, principal component analysis [24] and analysis of variance, by 2^2 factorial experiments with an ideal point [25] in randomised complete block design, together with mean comparison by Duncan's new multiple range test. All statistical analyses were done by SPSS 16.0 Family.

RESULTS AND DISCUSSION

Sensory attributes are the main key for consumers' acceptance of the product. For ice cream, rich and creamy, sweet, flavourful [1], smooth and velvety [19] products are desired. These sensory attributes were used for evaluating consumers' acceptance of ginger-flavoured soya milk ice cream. Consumers were requested to taste the products and rate the intensity of sensory attributes before being asked to express their ideal intensity compared to scores of tasted products. These ideal points would be the direction to improve the products.

Sensory Analysis

To calculate ideal ratio scores of each consumer, sample scores were divided by its ideal score for each attribute. The ideal ratio scale of 1.0 indicated that the intensity of the attribute was desired by

that consumer. The attribute was too weak if the ideal ratio score was less than 1.0 and it was too strong if the ideal ratio score was more than 1.0. Mean ideal ratio scores of all consumers could be interpreted as relative or percentage changes from the ideal. However, mean ideal ratio scores could not be trusted if their standard deviation was higher than 0.5 [24]. The result of mean comparison for ideal ratio scores is shown in Table 2. Almost all of the standard deviation was lower than 0.5 and all distribution could be considered as normal curve which showed the agreement of consumers on the ideal.

For ginger flavour and sweetness, the mean ideal ratio scores of Recipes 2 and 4 were not significantly different ($p>0.05$) from 1.0, the ideal, whilst those of Recipes 1 and 3 were more than the ideal for ginger flavour but less than the ideal for sweetness. The mean ideal ratio score of richness of Recipe 2 was not significantly different ($p>0.05$) from the ideal whilst other recipes were less than the ideal. The mean ideal ratio scores of smoothness of all recipes were significantly lower ($p\leq0.05$) than the ideal. These results showed that Recipe 2 was closest to the ideal product because its intensity of ginger flavour, sweetness and richness were accepted by consumers. Only its degree of smoothness, which was less than the ideal, should be improved by specific ingredients such as emulsifiers and stabilisers and better freezing process [19].

Table 2. Ideal ratio scores (mean $\pm$ standard deviation) of sensory attributes

Recipe	Ginger flavour	Sweetness	Richness	Smoothness
1 (S6, G4)	$1.13^{a} \pm 0.63$	$0.60^{b} \pm 0.17$	$0.60^{c} \pm 0.25$	$0.60^{c} \pm 0.20$
2 (S7, G4)	$\mathbf{0.91^{b}} \pm 0.24$	$\mathbf{1.03^{a}} \pm 0.21$	$\mathbf{0.97^{ab}} \pm 0.14$	$0.92^{b} \pm 0.20$
3 (S6, G5)	$1.17^{a} \pm 0.66$	$0.56^{b} \pm 0.18$	$0.61^{c} \pm 0.21$	$0.63^{c} \pm 0.21$
4 (S7, G5)	$\mathbf{0.95^{b}} \pm 0.28$	$\mathbf{1.03^{a}} \pm 0.23$	$0.93^{b} \pm 0.16$	$0.88^{b} \pm 0.24$
Ideal product	$\mathbf{1.00^{b}}$	$\mathbf{1.00^{a}}$	$\mathbf{1.00^{a}}$	$\mathbf{1.00^{a}}$

Note : S6 = 6%(w/w) sucrose; S7 = 7%(w/w) sucrose; G4 = 4%(w/w) ginger extract;
G5 = 5%(w/w) ginger extract
Means with different letters in the same column were significantly different ($p\leq0.05$).
(1 = ideal point, < 1 = too weak, > 1 = too strong)

The principal component analysis was used to reduce the sensory data of recipes to 2 principal components (PC) [24] with 82.7% variance explained. PC1 (57.7% variance explained) was a representative of sweetness, richness and smoothness, and PC2 (25.0% variance explained) was a representative of ginger flavour (Figure 1). The position of Recipes 1 and 3 on the principal component map (Figure 1) was close to each other but they were far from Recipes 2 and 4 and ideal product because of their stronger ginger flavour (PC1) and weaker sweetness, richness and smoothness (PC2), whilst Recipes 2 and 4 and ideal product were close to one another. The difference between Recipe 2 or 4 and ideal product was not obvious because of some information loss in this method.

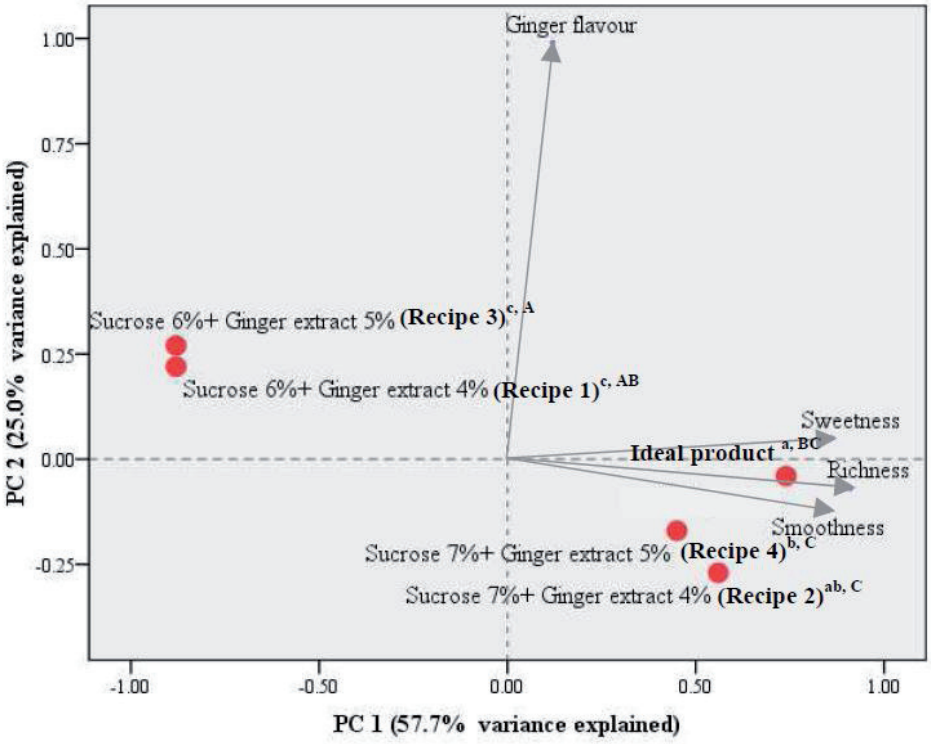


Figure 1. Principal component map of ginger-flavoured soya milk ice cream. Recipes with different letters (a, b, c for PC1 and A, B, C for PC2) were significantly different ($p \leq 0.05$).

The analysis of variance with mean comparison showed that the studied attributes of the developed recipes were close to the ideal point (Table 3). The ginger flavour of Recipes 1, 3 and 4 and ideal point was not significantly different ($p>0.05$) whilst the sweetness of Recipes 2 and 4 and ideal point, and the richness of Recipe 2 and ideal point were not significantly different ($p>0.05$). The smoothness of Recipes 2 and 4 were closer to ideal point than that of Recipes 1 and 3. In the case of preference, Recipe 2, with mildest ginger flavour, ideal sweetness, ideal richness and close-to-ideal smoothness, had the highest mean score. The second highest was Recipe 4, with ideal ginger flavour and sweetness, and close-to-ideal richness and smoothness. Recipes 1 and 3, with ideal ginger flavour but poor in sweetness, richness and smoothness, had lower mean preference scores ($p \leq 0.05$).

Table 3. Intensity (mean ± standard deviation) of sensory attributes

Recipe	Ginger flavour	Sweetness	Richness	Smoothness	Preference
1 (S6, G4)	$4.95^{a} \pm 2.19$	$3.44^{b} \pm 1.07$	$3.62^{c} \pm 1.46$	$3.94^{c} \pm 1.58$	$4.23^{c} \pm 1.88$
2 (S7, G4)	$4.30^{c} \pm 1.63$	$5.87^{a} \pm 1.29$	$5.87^{ab} \pm 1.30$	$5.98^{b} \pm 1.54$	$6.93^{a} \pm 1.63$
3 (S6, G5)	$5.11^{a} \pm 2.24$	$3.21^{b} \pm 1.06$	$3.66^{c} \pm 1.32$	$4.15^{c} \pm 1.54$	$4.19^{c} \pm 1.66$
4 (S7, G5)	$4.44^{bc} \pm 1.70$	$5.85^{a} \pm 1.27$	$5.65^{b} \pm 1.28$	$5.72^{b} \pm 1.57$	$6.46^{b} \pm 1.76$
Ideal product	$4.84^{ab} \pm 1.59$	$5.79^{a} \pm 1.10$	$6.11^{a} \pm 1.25$	$6.62^{a} \pm 1.34$	

Note : S6 = 6%(w/w) sucrose; S7 = 7%(w/w) sucrose; G4 = 4%(w/w) ginger extract; G5 = 5%(w/w) ginger extract
Means with different letters in the same column were significantly different ($p \leq 0.05$). (0 = least, 10 = most)

This preference result accorded with both previous analyses because Recipe 2, which was closer to the ideal product, was more preferred by consumers. However, the mean ideal score of ginger flavour indicated that consumers preferred higher intensity of ginger flavour than that of Recipe 2 (Table 3). This requirement might be attributable to the difference in scale uses of consumers as well as to consumer awareness on the health benefits of ginger. In the first case, the difference in scale could be eliminated by transforming raw data to standardised or related scores. Therefore, this disagreement was not found in the ideal ratio scores. In the second case, ginger is a general spice and a Thai herb recommended for primary health care system [26]. Therefore, consumers tended to rate it high if the decision was made on ginger flavour only, because they wanted to have ginger for health benefit. In this case the ideal information provided by consumers did not correspond to the ideal product in terms of preference but to the ideal in terms of health benefit.

When consumers were asked to evaluate their preference of the products, the overall perception of the products would be considered and the impact of each perception on product preference was not similar. Normally, the satisfaction of overall flavour and texture of the products highly contributed to overall preference [27]. According to a preference estimation model [28], in this study it was: Preference = 1.365 – 0.105 (ginger flavour) + 0.546 (sweetness) + 0.387(smoothness). This model showed a lower impact of ginger flavour compared to sweetness and smoothness whilst the impact of richness was not significantly different ($p>0.05$) and the impact of ginger flavour was negative. Therefore, the intensity of ginger flavour should not be strong for the overall acceptance of product flavour.

In the case of main effect analysis, only sucrose percentage significantly affected the intensity of the four sensory attributes ($p\leq0.05$) (Table 4). Ginger flavour was reduced whilst sweetness, richness and smoothness were increased by adding sucrose. For ginger flavour, it could be explained by taste-taste interaction in which a high intensity of sweetness tends to result in the suppression of other tastes [29]. For tactual properties, the sugar content could affect the consistency and texture of products [19]. Although sucrose and ginger extract significantly influenced consumer preference ($p\leq0.05$), the effect of sucrose content was significantly more pronounced (Table 4).

Table 4. Intensity (mean ± standard deviation) of sensory attributes at each level of sucrose and ginger extract

	Ginger flavour	Sweetness	Richness	Smoothness	Preference
Sucrose					
6% (w/w)	$5.03^{a} \pm 2.21$	$3.33^{b} \pm 1.07$	$3.64^{b} \pm 1.39$	$4.04^{b} \pm 1.56$	$4.21^{b} \pm 1.77$
7% (w/w)	$4.37^{b} \pm 1.67$	$5.86^{a} \pm 1.28$	$5.76^{a} \pm 1.29$	$5.85^{a} \pm 1.56$	$6.70^{a} \pm 1.70$
Ginger extract					
4% (w/w)	4.62 ± 1.95	4.66 ± 1.70	4.74 ± 1.78	4.96 ± 1.86	$5.58^{a} \pm 2.21$
5% (w/w)	4.78 ± 2.01	4.53 ± 1.77	4.65 ± 1.63	4.94 ± 1.74	$5.33^{b} \pm 2.05$

Note : Means with different letters in the same column were significantly different ($p\leq0.05$). (0 = least, 10 = most)

Chemical Analysis of Selected Recipe

The most preferred recipe with 7%(w/w) sucrose and 4%(w/w) ginger extract was selected for chemical analysis. Its total phenolic content and antioxidant capacity values are shown in Table 5. The phenolic content of the product comes from soya milk and ginger extract. Examples of phenolic compounds found in soya bean are *p*-hydroxybenzoic acid, salicylic acid, *p*-coumaric acid, ferulic acid [30] and isoflavone [13, 14, 30] while in ginger, salicylic acid, cinnamic acid, catechin, rutin, quercetin, flavonoids, among others, are present [31]. The total phenolic content of this soya milk ice cream recipe is higher than that of Job's tears ice cream recipes (9.06-25.14 mg gallic acid equivalent per 100 grams of product) [10, 32] and tofu ice cream recipe (31 mg flavone per kilogram of product) [33]. As for milk ice cream, it contains trace amounts of phenolic compounds probably originated from the milk used but the formation of phenolic compounds by Maillard reaction could occur during pasteurisation of ice cream mixes [34].

Values of antioxidant capacity depend on the mechanism of the methods used [21-23]. However, all antioxidant capacity values which are more than 20% of the daily value (60 mg) of vitamin C [35] show antioxidant potential of this selected recipe owing to phenolic compounds and other antioxidant substances present in the raw materials used in the recipe.

Table 5. Total phenolic content and antioxidant capacity (mean $\pm$ standard deviation) of selected ginger-flavoured soya milk ice cream (3 replications)

	Recipe with 7%(w/w) sucrose and 4%(w/w) ginger extract
Total phenolic content (mg gallic acid equivalent per 100 grams of product)	91.6 + 6.8
Antioxidant capacity (mg vitamin C equivalent per 100 grams of product)	
FRAP	37.9 + 3.7
DPPH	13.4 + 1.2
ABTS	49.0 + 5.1

CONCLUSIONS

The overall results of ideal ratio scores, principal component analysis and analysis of variance, together with mean comparison in this study, were equivalent but the result of a single attribute such as ginger flavour might be misleading if analysis of variance was used only. The attributing ideal point showed that the intensity of ginger flavour should be high whereas the overall preference revealed that ginger flavour should be in the low level because it was more preferred. Since food was a complex matrix, many interactions could occur and the impact of each perception on consumer preference was not equal. Both the attribute and overall preference should be comparatively studied. This contradiction of ginger flavour was not found in the ideal ratio scores because dividing sample scores by ideal score diminished the difference of scale used for each consumer. Consumer disagreement on ideal ratio scores could be observed from a high standard deviation but it was not found for the selected sample in this study. For principal component analysis, it was found to be a good technique because the detail might be disregarded.

REFERENCES

1. R. T. Marshall and W. S. Arbuckle, "Ice Cream", 5th Edn., Chapman and Hall, New York, **1996**, pp. 1-9, 18-21.
2. A. H. Varnam and J. P. Sutherland, "Milk and Milk Products: Technology, Chemistry and Microbiology", Chapman and Hall, London, **1994**, pp. 387-431.
3. Announcement Issue 222, "Ice cream" (in Thai), Ministry of Public Health, Bangkok, **2001**, pp.1-6.
4. Announcement Issue 257, "Ice cream (Issue 2)" (in Thai), Ministry of Public Health, Bangkok, **2002**, pp. 1-2.
5. Naewna, "Swensens go ASEAN", **2012**, http://www.naewna.com/business/23042, (Accessed: December 2012).
6. Kasikorn Research Center, "Home-made ice cream busisness", **2012**, http://www.ksmestartup.com/FileUpload/WebContentDownload/120420052530103.pdf, (Accessed: December 2012).
7. M. Moore, "Ice cream: Thailand's version", (no date), http://www.wheretoeat-bangkok.com/Assets/ Articles/thai_ice_cream.pdf, (Accessed: December 2010).
8. M. Badcock, "Ice cream, Thai-style", **2010**, http://thaifoodandtravel.com/blog/thai-ice-cream, (Accessed: December 2010).
9. W. Wangcharoen, "Brown rice ice cream and soya milk ice cream", Workshop on Processing of Health Food Products (in Thai), **2007**, Maejo University, Thailand.
10. W. Khongjeamsiri, W. Wangcharoen, S. Pimpilai and W. Daengprok, "Development of Job's-tears ice cream recipe", *Maejo Int. J. Sci. Technol*., **2009**, *3*, 388-399.
11. The Recipe Link, "Lots of soya milk ice cream recipes", **2000**, http://www.recipelink.com/mf/0 / 47704, (Accessed: May 2011).
12. J. Charatan, "Here's the scoop- dairy free ice cream", **2000**, http://findarticles.com/p/articles/mi_m0820/is_2000_August/ai_63902652, (Accessed: May 2011).
13. M. Sugano, in "Soya in Health and Disease Prevention" (Ed. M. Sugano), CRC Press, Boca Raton, **2006**, pp. 1-16.
14. P. Golbitz and J. Jordan, in "Soy Applications in Food" (Ed. M. N. Riaz), CRC Press, Boca Raton, **2006**, pp. 1-22.
15. K. Leelarattanakul and A. Hasun, "Ginger-flavoured soy milk ice cream", Special problem for Bachelor of Science Program in Food Science and Technology, **2007**, Maejo University, Thailand.
16. H. Tsuneki, I. Kimura and L. R. Pancho, in "Ginger – the Genus Zingiber" (Ed. P. N. Ravindran and K. Nirmal Babu), CRC Press, Boca Raton, **2005**, Ch. 13.
17. S. R. Uhl, "Handbook of Spices, Seasonings, and Flavorings", CRC Press, Boca Raton, **2000**, pp. 113-116.
18. G. Vernin and C. Parkanyi, in "Ginger – the Genus Zingiber" (Ed. P. N. Ravindran and K. Nirmal Babu), CRC Press, Boca Raton, **2005**, Ch. 3.
19. W. J. Harper and C. W. Hall, "Dairy Technology and Engineering", AVI Publishing Company, New York, **1976**, pp.133-134.
20. A. Waterhouse, "Folin-Ciocalteau micro method for total phenol in wine", (no date), http://waterhouse.ucdavis.edu/phenol/folinmicro.htm, (Accessed: May 2011).

21. I. F. F. Benzie and J. J. Strain, "Ferric reducing / antioxidative power assay: Direct measure of total antioxidant activity of biological fluids and modified version of simultaneous measurement of antioxidant power and ascorbic acid concentration", *Methods Enzymol.,* **1999**, *299*, 15-27.
22. W. Brand-William, M. Cuelier and M. E. Berset, "Use of free radical method to evaluate antioxidant activity", *Lebensm. Wiss. Technol.*, **1995**, *28*, 25-30.
23. M. J .T. J. Arts, G. R. M. M. Haenen, H. P. Voss and A. Bast, "Antioxidant capacity of reaction products limits the applicability of the Trolox equivalent antioxidant capacity (TEAC) assay", *Food Chem. Toxicol*., **2004**, *42*, 45-49.
24. H. R. Cooper, M. D. Earle and C. M. Triggs, in "Product Testing with Consumers for Research Guidance" (Ed. L. S. Wu), American Society for Testing and Materials International, West Conshohocken, **1989**, pp. 54-63.
25. Biometrics Information, "ANOVA: Factorial designs with a separated control", **1989**, http://www.for.gov.bc.ca/hre/biopamph/pamp14.pdf, (Accessed: May 2011).
26. N. R. Farnsworth and N. Bunyapraphatsara, "Thai Medicinal Plants: Recommended for Primary Health Care System", Medicinal Plant Information Center, Mahidol University, Bangkok, **1992**, pp. 253-260.
27. H. R. Moskowitz and B. Krieger, "The contribution of sensory liking to overall liking: An analysis of six food categories", *Food Qual. Prefer.*, **1995**, *6*, 83-90.
28. T. Worch, S. Lê, P. Punter and J. Pages, "Extension of the consistency of the data obtained with the ideal profile method: Would the ideal products be more liked than the tested products?", *Food Qual. Prefer.*, **2012**, *26*, 74-80.
29. R. S. J. Keast and P. A. S. Breslin, "An overview of binary taste-taste interactions", *Food Qual. Prefer.*, **2002**, *14*, 111-124.
30. E. H. Kim, S. H. Kim, J. I. Chung, H. Y. Chi, J. A. Kim and I. M. Chung, "Analysis of phenolic compounds and isoflavones in soybean seed (*Glycine Max* (L.) Merill) and sprouts grown under different conditions", *Eur. Food Res. Technol*., **2006**, *222*, 201-208.
31. A. Ghasemzadeh, H. Z. E. Jaafar and A. Rahmat, "Synthesis of phenolics and flavonoids in ginger (*Zingiber officinale* Roscoe) and their effects on photosynthesis rate", *Int. J Mol. Sci*., **2010**, *11*, 4539-4555.
32. W. Khongjeamsiri, W. Wangcharoen, S. Pimpilai and W. Daengprok, "Development of Job's tears ice cream recipes with carrot juice and pumpkin paste", *Maejo Int. J. Sci. Technol*., **2011**, *5*, 390-400.
33. Z. Wang and F. B. Murphy, "Isoflavone content of commercial soybean foods", *J. Agric. Food Chem.*, **1994**, *42*, 1666-1672.
34. O. Sagdic, I. Ozturk, H. Cankurt and F. Tornuk, "Interaction between some phenolic compounds and probiotic bacterium in functional ice cream production", *Food Bioproc. Technol*., **2012**, *5*, 2964-2971.
35. Centre for Food Safety and Applied Nutrition, **"**Guidance for industry: A food labeling guide**"**, **2008**, http://www.cfsan.fda.gov/guidance.html (Accessed: April 2008).

Maejo Int. J. Sci. Technol. **2012**, *6*(03), 514-527

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Ciphering Indicator approaches and user awareness

Iosif Androulidakis [1,*], Dionisios Pylarinos[2] and Gorazd Kandus [3]

[1] Jožef Stefan International Postgraduate School, Jamova 39, Ljubljana SI-1000, Slovenia
[2] Electrical and Computer Engineering Dpt, University Of Patras, 26504, Rio, Patras, Greece
[3] Communication Systems Dpt, Jožef Stefan Institute, Jamova 39, Ljubljana SI-1000, Slovenia
* Corresponding author, e-mail: sandro@noc.uoi.gr

Received: 13 November 2011 / Accepted: 28 December 2012 / Published: 28 December 2012

Abstract: One of the fundamental mobile phone security problems in GSM is the absence of base station authentication, which allows man-in-the-middle attacks. During such attacks, a third party activates a fake base station, which acts as a bypass to the network, thus switching off the encryption and intercepting the user's communications. 3G mobile networks enforce mutual authentication but this can be circumvented if the 3G band is jammed by the attacker, forcing the phone to connect using GSM. GSM and newer standards provide a user alert indicating that the encryption has been switched off, which is called a Ciphering Indicator. In the present paper, different approaches followed by various manufacturers concerning the Ciphering Indicator are investigated. A total of 38 different mobile phones ranging from old to new and from simple to smart-phones that were produced by 13 different manufacturers were intercepted using a GSM testing device in order to document their reactions. Four approaches were identified with some manufacturers choosing not to implement the feature at all. It was also found that in the cases in which the feature was actually implemented, no universal indication was used and it was seldom documented in the phones' manuals. User awareness regarding the Ciphering Indicator and security issues was also investigated via an empirical survey employing more than 7,000 users from 10 countries and was found to be significantly low.

Keywords : Ciphering Indicator, graphical user interface, mobile phone, fake base station

INTRODUCTION

One of the fundamental security problems and a basic shortcoming regarding GSM security planning is the fact that mobile telephone base stations do not have to authenticate themselves to the user [1]. A user wishing to gain access to a provider's GSM mobile telephone network must own the proper SIM card and have it inserted in his or her phone device. The user's authentication is therefore

performed by comparing the SIM's credentials with the data stored in the network's database [2]. A base station authentication mechanism, however, is not employed and mobile phones are not capable of assessing the legitimacy of the system they are connecting to, nor certifying whether this system is indeed part of their provider's network. Therefore, a fake base station can easily present itself as a part of the victim provider's network.

Furthermore, mobile phones constantly monitor a special data transmission beacon from the nearby base stations (through the broadcast control channel - BCCH) in order to choose the one offering the strongest signal for their communication [3]. This way, they can achieve better communication quality, economise the amount of energy consumed and increase their autonomy time. Hence, if the attacker installs his or her equipment in a nearby area and starts transmitting, masquerading as a legitimate operator and overlapping the authentic base stations' signals, mobile phones of that specific operator located nearby will choose the fake base station for their communication.

The next stage of the attack would be to neutralise the encryption. GSM uses an A5 algorithm for voice encryption [4]. There exist various versions of this algorithm that offer different levels of security (A5/2, A5/1, A5/3 – listed in strength order from lowest to highest), as well as a version with no encryption at all (A5/0) [5]. Under normal circumstances, the network has stored in its Authentication Centre's (AuC) database a secret key, Ki, which is also stored in the user's SIM card and is never transmitted in the network. These keys are compared by using a signed response (SRES) with the help of algorithm A3 and thus, the mobile phone is authenticated [6]. Finally, using the Ki and other data, the A8 algorithm produces the session encryption key, Kc, which is used in the speech encryption algorithm, A5 [6]. In the case of the fake base station, the basic information of the Ki key is not known to the attacker. Hence, the attack cannot proceed. However, the system planning prioritises usability instead of security in this case. As such, the corresponding protocols allow the negotiation and agreement between the mobile phone and the base station regarding whether they will use an encryption algorithm and which one they will use [6]. Therefore, sending the proper signal, the fake base station may inform the mobile phone that it does not have any encryption capabilities (A5/0) and that communication should take place without the use of encryption.

With encryption switched off, the attacker can act as a man-in-the-middle and intercept the communication of the target phone. Following this, using a simple mobile or fixed telephone, he can relay the call back to the genuine network and to the intended recipient, recording the communication in the process [7]. It is worth noting that 3G mobile networks enforce the mutual authentication scheme [8], which means that an authentication of the base station is required, eliminating fake base station attacks in practice. However, this can be circumvented if the 3G band is jammed by the attacker. Indeed, when a multi-band-capable mobile handset loses 3G signal connectivity, it will try to connect to older networks (2.5G and 2G) present in the area, thanks to backward compatibility. Therefore, even 3G users may fall victim to a fake base station attack. In addition, many users still prefer not to connect to 3G networks because of the increased power consumption and the shorter autonomy time of the handset [9]

Even though the industry and researchers have shown an active interest in enhancing the mobile phone user experience, offering more and more services and a wealth of applications [10-13], the user notification issue regarding encryption being switched off, as well as user awareness of the matter, has not yet been thoroughly investigated. In this paper, the history of GSM standards regarding this issue,

the different approaches followed by manufacturers and user awareness regarding this issue are widely investigated by employing a data set of 38 different models, 13 different manufacturers and 7,172 users.

METHODS

A GSM tester was used to intercept 38 different phones from 13 different manufacturers in order to identify their reaction when under attack (implementing the Ciphering Indicator or not) and also to investigate whether the matter has been documented in their manuals. User awareness was also investigated via a survey of 7,172 university students from 10 different EU countries. To formulate the process, the history of GSM standards regarding this issue was researched.

History of Ciphering Indicator in GSM Standards

It took several years for an alert informing the user of the loss of encryption to be included in the GSM standards. The first notion of this alert for a lack of encryption (albeit not explicitly stated as 'encryption') in the GSM standards was in 1997 [14], when a cryptic operational feature monitor (OFM) bit was mentioned. That bit controlled the OFM attribute, as shown in Figure 1, but the meaning of the term OFM was not explained in the abbreviations or elsewhere in the text.

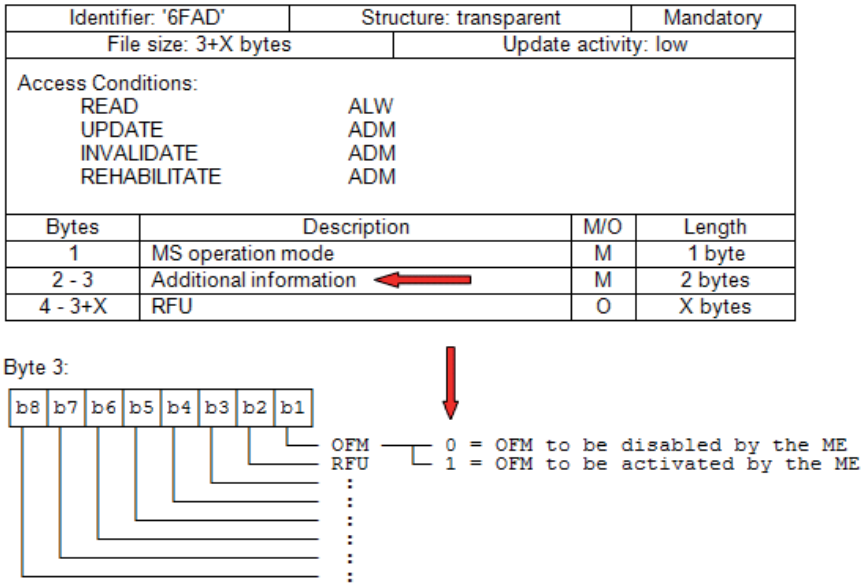


Figure 1. The first occurrence of the OFM bit in the standards

A Ciphering Indicator was introduced a few months after the first mention of the OFM and it was clearly stated that a notification should show the user the lack of data confidentiality [15]. It was also stated that the Ciphering Indicator feature should be mandatory, enabled by default, and potentially switched off via the respective SIM setting controlled by the network operator [15, 16]. As such, even if a handset has implemented the feature, the operator is able to instruct it not to alert the user in the case of a loss of encryption.

In 1999, the OFM term was described as an Operational Feature Monitor, and it was also explained that the OFM bit is indeed used to turn the Ciphering Indicator on and off [17]. In 2004, the

OFM term was abandoned in favour of the more straightforward Ciphering Indicator term [18], as shown in Figure 2. In 2009, further clarification of the feature was provided, and it was stated that phones with a suitable user interface should offer the user the capability to override Ciphering Indicator setting set by the operator [19]. This standard seems to be the first step towards actually empowering the user to overcome the control of the operator regarding Ciphering Indicator, although such technology has yet to be widely embraced.

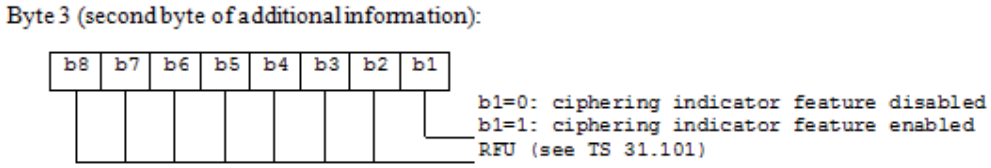


Figure 2. OFM is officially replaced by the Ciphering Indicator term

Phone Interception

In order to test mobile phones' behaviours with regard to Ciphering Indicator, a professional GSM testing device [20], shown in Figure 3, was used. The GSM tester provides all necessary signalling to the mobile phone in the same way as a base station does, enabling us to perform our tests. The experimental set-up consisted of the GSM tester, properly configured, and an antenna. The required settings for the parameterisation of the GSM tester are mobile country code (MCC), mobile network code (MNC), channel number (ARFCN) on which the broadcast control channel (BCCH) is transmitted and the traffic channel (TCH) on which the actual communication takes place. It should be noted that all MCCs are publicly available information [21]. In order not to interfere with any legitimate networks, a specially designated test-only network was used (a combination of MCC 001 and MNC 01), as shown in Figure 4. The encryption capability was deactivated by choosing algorithm A5/0, and a test call was initiated by the handset in order to test for the presence of an indication. The international mobile equipment identity (IMEI) of the target phone was logged, as shown in Figure 5, in order to deduce the exact phone model.

For every handset tested, two SIM cards from two different providers were used: one with Ciphering Indicator feature enabled and the other with Ciphering Indicator feature disabled. The former was used in order to test the manufacturer's approach regarding the indicator (whether implemented or not) and the latter was used to test whether the operator's setting was actually followed by the handset.

By adding a mobile handset with monitoring software installed, a second mobile phone, in order to channel the interception communication through it and a PC, the setup could be used to actually intercept communication before relaying it to the original recipient. More details about the experimental set-up and the process can be found in the author's previous paper [7].

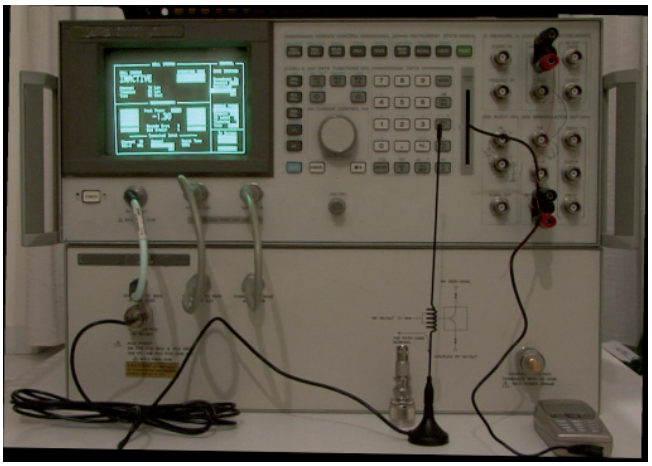


Figure 3. Mobile phone tester

Figure 4. Designated test-only network in the available networks list

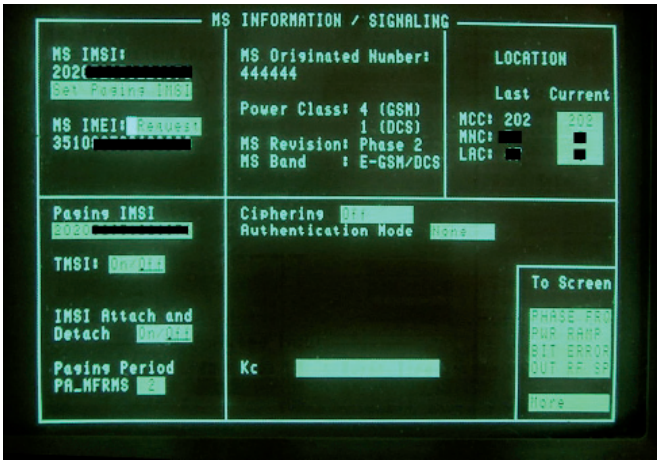


Figure 5. Fake base station has extracted the mobile subscriber and mobile equipment identification as well as the number that the user intends to call (444444). Ciphering is switched off, as shown.

User Awareness Survey

To investigate user awareness regarding Ciphering Indicator, a large-scale survey of user security habits and trends was employed. The survey started in 2009 with a small sample of students from the University of Ioannina, Greece [22] and eventually resulted in a large sample of 7,172 university students from 17 different universities located in 10 different European countries, i.e. Hungary, Czech, Estonia, Latvia, Lithuania, Bulgaria, Greece, Romania, Slovakia and Slovenia [23, 24].

Multiple-choice questionnaires were used, employing an in-person delivery technique. Data entry took place using custom optical mark recognition (OMR) software, which enabled the processing of the questionnaires in a very rapid and accurate manner, avoiding human data entry mistakes [25]. Statistical processing took place using the SPSS analysis tool [26]. The aspect of the questionnaire considered in this paper is the awareness of Ciphering Indicator along with the brand used, because different brands follow different approaches.

RESULTS AND DISCUSSION

Implementation of Ciphering Indicator in Different Brands

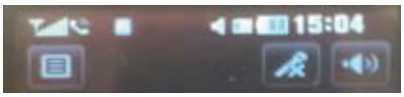
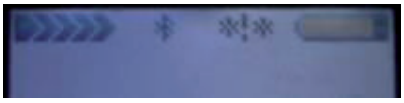
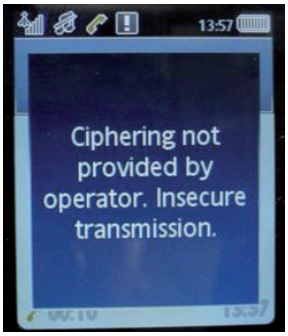
To acquire a wide view of the behaviour of phones during man-in-the-middle attacks, a group of 38 different mobile phones from 13 different manufacturers covering a large time span (from 2002 to 2010) was investigated.

In the case of SIMs with Ciphering Indicator feature enabled, four different approaches followed by the manufacturers were identified and are shown in Table 1. Nine different manufacturers in the considered dataset (Sharp, Samsung, Qtek, HTC, Motorola, LG, Huawei, Chinabuye and Apple) did not employ a Ciphering Indicator, although this is required by the standards, an approach which was identified as "Approach 0". Furthermore, a universal indication was not employed by the manufacturers that did incorporate a Ciphering Indicator. One manufacturer in the considered dataset (Siemens) used stars and an explanation mark (an approach identified as "Approach 1"). Two others (ZTE and Nokia) used an open-padlock (an approach identified as "Approach 2"). Another manufacturer (Sony Ericsson) used an exclamation mark inside a red triangle or a grey square, along with an explanatory text message, an approach identified as "Approach 3". Although there is inconsistency regarding the icon used in this approach (triangle or square), the presence of the accompanying text should be able to clear up any confusion. The text itself has been amended in more recent models to be more descriptive, as shown in Table 1. However, this informative message would disappear after a few seconds in eight out of the eleven cases considered, leaving only the icon present for the rest of the call. Only in three of the cases studied did the manufacturer choose to follow a more robust implementation, requiring that the user specifically acknowledge reading the message.

The documentation of Ciphering Indicator in manuals of the considered phones was also investigated. It was found that the presence or the meaning of the possible icons used as a Ciphering Indicator was mostly not documented. As shown in Table 1, proper documentation for the Ciphering Indicator was rather rare and only one manufacturer (Sony Ericsson) in the considered data set included it while this was done in only three out of its eleven models in the considered group. Quite interestingly, this manufacturer also used an explanatory message in addition to the icon and therefore it was the one that least needed to include such documentation in the phones' manuals. Another manufacturer (Nokia) documented, in three different cases, that a closed padlock icon shows that the data services (e.g. WAP

and WiFi) use encryption and that the absence of this icon indicates a lack of encryption, but Nokia has provided no documentation regarding the open padlock icon for voice and text communication. As an example, the way the icon of Approach 3 is explained in the respective phone manual is shown in Figure 6.

Table 1. Examples of different implementations of the Ciphering Indicator

Approach	Brand	Ciphering Indicator	Example of phone GUI	Documented
0	Sharp Samsung Qtek HTC LG Motorola Huawei Chinabuye Apple	No indicator		n/a (19 cases)
1	Siemens	Stars and an exclamation mark		0/2 cases
2	ZTE Nokia	Open padlock		0/6 cases
3	Sony Ericsson	Exclamation mark inside a grey square or a red triangle, along with a text message	 	3/11 cases

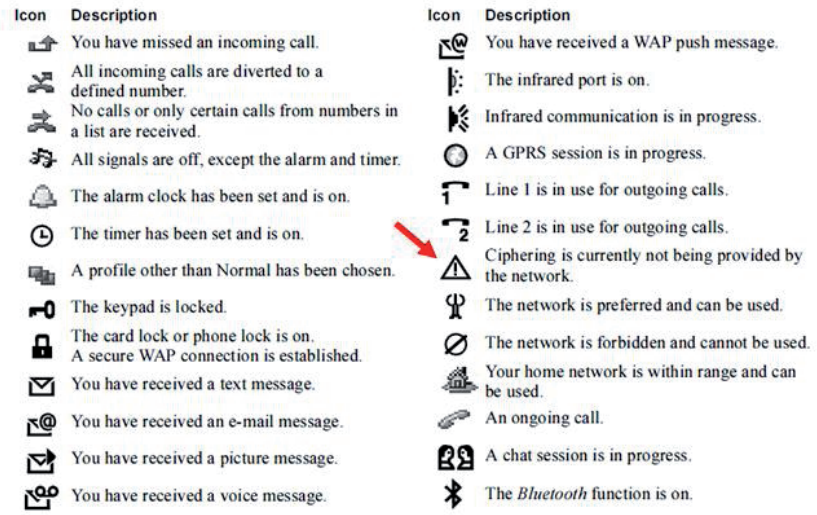


Figure 6. Examples of the documented Ciphering Indicator (Approach 3)

At this point, it is interesting to examine the behaviour of the so-called 'smartphones'. These phones have advanced operating systems that allow a myriad of applications to be installed, minimising the gap between mobile phones and PCs. Our sample encompassed smartphones with four main operating systems (Android, iOS, Windows Mobile, and Symbian). It was found that only Symbian had implemented a Ciphering Indicator. This could possibly be attributed to the fact that Symbian is closely related to Nokia, a telecom manufacturer that has been using a Ciphering Indicator since its early models, whereas the other advanced operating systems have evolved out of the computer community (i.e. Windows Mobile, Android and iOS). It should also be noted that in the case of Android, a similar icon with the Approach 3 implementation of the Ciphering Indicator (a triangle with an exclamation mark) is used as a general notification icon, as shown in Figure 7, but it is not used to alert the user when the encryption is switched off.

Handsets using a SIM with the Ciphering Indicator feature switched off were also tested. It was found that all phones obeyed the network operator setting, with the exception of one. This occurrence should probably be attributed to a bug and not to a general manufacturer approach, because a later model from the same manufacturer had no such issues. Further, it should be noted that a phone that allowed the user to override the network operator setting for the Ciphering Indicator was not found in the considered group. A detailed report for each phone tested which contained the manufacturer, the model, the IMEI**,** the manufacturer's approach and an indication whether the Ciphering Indicator was documented can be found in Table 2.

Table 2. Manufacturers' approach regarding Ciphering Indicator

	Brand	Model	IMEI	Approach	Ciphering Indicator documented	Year Launched
1	Nokia	1600	35 89 5801…	2	No	2006
2	Nokia	3510i	35 14 6280…	2	No	2002
3	Nokia	6510	35 11 0510…	2	No	2002
4	Nokia	5000	35 67 9702…	2	No	2008
5	Nokia	E71	35 82 4003	2	No	2008
6	Sony Ericsson	T610	35 12 5300…	3	Yes	2003
7	Sony Ericsson	T200	35 04 0345…	3	Yes	2002
8	Sony Ericsson	K810i	35 94 5101…	3	No	2007
9	Sony Ericsson	W810i	35 94 5701…	3	No	2006
10	Sony Ericsson	K770i	35 61 7902…	3	No	2007
11	Sony Ericsson	K750i	35 93 0200…	3	No	2005
12	Sony Ericsson	W595	3529 6503 …	3*	No	2008
13	Sony Ericsson	W700i	35 52 7101…	3	Yes	2006
14	Sony Ericsson	K630i	35 88 0101…	3*	No	2007
15	Sony Ericsson	W705	35 18 0603…	3	No	2008
16	Sony Ericsson	C902	35 87 9002...	3*	No	2008
17	Sharp	GX17	35 97 9100…	0	n/a	2005
18	Samsung	E1080	35 80 3703…	0	n/a	2010
19	Samsung	SGH-E570	35 49 9201…	0	n/a	2006
20	Samsung	E1310	35 42 3703…	0	n/a	2009
21	Samsung	C3050	35 55 2803…	0	n/a	2009
22	Samsung	SGH-J700	35 26 9302…	0	n/a	2008
23	Samsung	SGH-E250	35 60 7501…	0	n/a	2006
24	Siemens	S55	35 10 8352…	1	No	2002
25	Siemens	S65	35 39 1200…	1	No	2004
26	Qtek (HTC)	S200	35 70 3600…	0	n/a	2006
27	HTC	Wildfire	35 90 2803…	0	n/a	2010
28	Motorola	C115	35 64 9800…	0	n/a	2004
29	Motorola	U9	35 87 9801…	0	n/a	2007
30	LG	KP500 cookie	35 91 3103…	0	n/a	2008
31	LG	KP105	35 79 4002…	0	n/a	2008
32	LG	GB108	35 71 4503…	0	n/a	2009
33	LG	KU990 Viewty	35 90 3603…	0	n/a	2007
34	LG	GU230	35 72 4503…	0	n/a	2010
35	ZTE	340	35 59 2203…	2	No	2009
36	Huawei	Joy 845	35 16 0204…	0	n/a	2010
37	Chinabuye	H969	35 73 6903…	0	n/a	2010
38	Apple	iPhone 3G	01 20 2300…	0	n/a	2008

*Alerting text accompanying Ciphering Indicator should be acknowledged by user

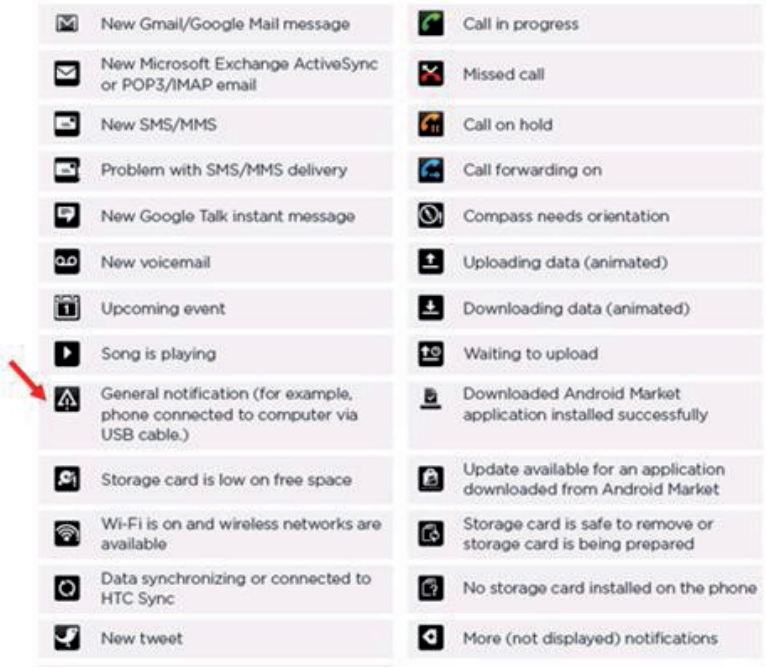


Figure 7. Use of a triangle with an exclamation mark as a general notification icon (Android)

User Awareness

Examining the user element, the awareness of the issue was found to be significantly low in the considered sample. This can be attributed to various manufacturers not employing the feature and also to the lack of proper documentation in the case of the manufacturers that do employ it. Since different manufacturers follow different approaches regarding the Ciphering Indicator, the market share of each brand is an important issue in terms of investigating user awareness. As our survey revealed, the two most popular brands (Nokia and Sony Ericsson) in the considered sample, which were used by 64.1% of the students, employ a Ciphering Indicator (Figure 8). However, the user awareness in this sample was significantly low (only 24,9% were aware of the indicator feature). This can be partly attributed to manufacturers not employing this feature and also to the lack of proper documentation by the manufacturers that do employ a Ciphering Indicator, because a large percentage of users that had a Ciphering Indicator feature enabled in their phones were still unaware of the meaning of the icon used.

Our fundamental empirical questions involved whether students are informed about how the options and the technical characteristics of their mobile phones affect their security and how secure they consider communication using mobile phones. Students answered those two questions subjectively. We also used some objective questions in regard to security practices (noting IMEI, using a PIN, using a password-protected screen saver, using antivirus software, and making backups). In this way, we were able to conclude whether their subjective answers were actually aligned with the objective facts.

When answering the question "Are you informed about how the options and technical characteristics of your mobile phone affect its security?", the majority of students (30.8%) stated that they were 'moderately' informed about the security options and characteristics, while 15.8% believed that they were 'not at all' informed. We proceeded to weigh the responses with the following weights: Very Much: 4, Much: 3, Moderately: 2, Not much: 1, Not at all: 0. Then, we divided them by the number of occurrences in order to obtain an arithmetic value and better compare the results (Figure 9). It was proved that LG and Samsung users are the most in need of security education because they

It was proved that LG and Samsung users are the most in need of security education because they scored the lowest on the 0-4 scale (1.74 and 1.73 respectively). Nokia (1.85) is around the total mean (1.86). iPhone and Ericsson users are the most informed ones (1.97 and 1.95 respectively).

Continuing with a general question about how 'secure' users felt mobile phone communication is, the majority (36.9%) replied 'moderately' , followed by 'much' at 28.6%. On the other hand, some (21.36%) felt not too much or not at all sure they were safe. Weighting with the same scale (0-4), the results were obtained as shown in Figure 10. It can be seen that iPhone users were the ones that are most 'suspicious' in regard to how safe they consider mobile phone communication to be. Sharp users were more relaxed.

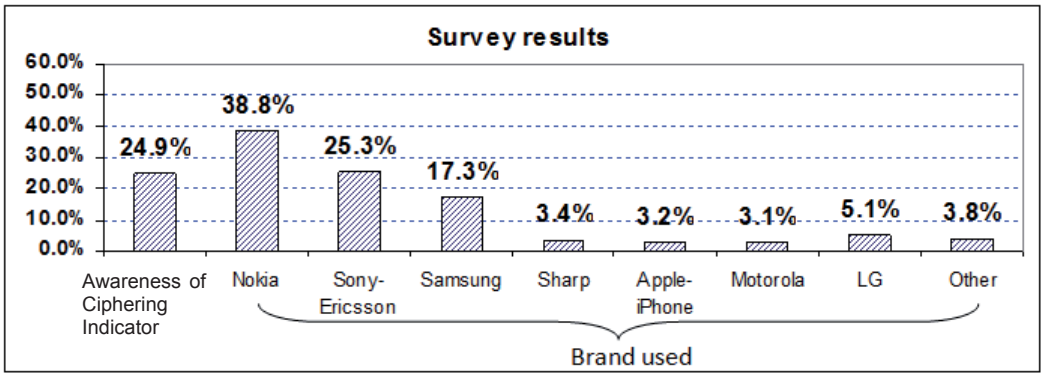


Figure 8. User awareness and percentages of the brands used in the considered samples

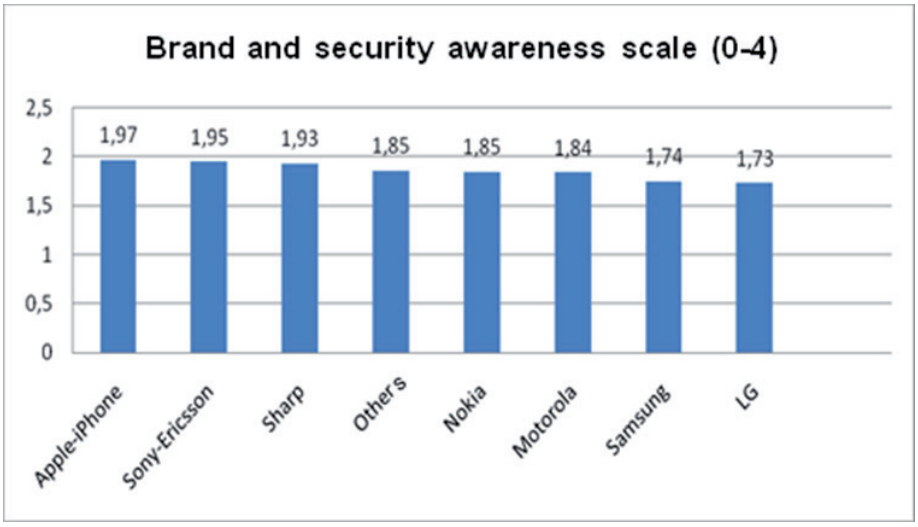


Figure 9. Brand and security awareness

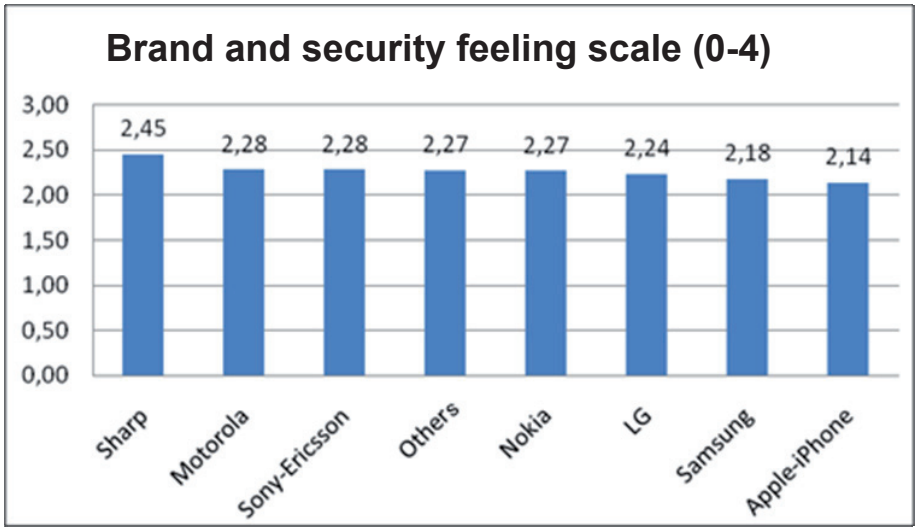


Figure 10. Brand and security feeling

CONCLUSIONS

In this paper, 38 different mobile phone models from 13 different manufacturers were intercepted in order to investigate the implementation of the Ciphering Indicator feature, which aims to alert users of switched-off encryption and possible interceptions. The documentation of the feature (or the absence of it) in various phones' manuals was also examined. In addition, user awareness regarding this feature and other security issues was surveyed by using the results of an empirical study employing a sample of 7,172 university students from 10 European countries.

Four approaches were followed by the manufacturers in the considered group regarding the Ciphering Indicator feature, ranging from no feature implementation to use of icons and an explanatory message. In the case of smartphones, phones with four different operating systems were tested and only one of them was found to implement the Ciphering Indicator feature. In general, the approach of each manufacturer seemed not to change over time, although one minor inconsistency was reported. The documentation of Ciphering Indicator in the considered phones' manuals was also investigated. It was found that the presence or the meaning of the icons used as a Ciphering Indicator was rarely documented. Finally, when examining the user element, awareness of the issue was found to be significantly low in the considered sample.

Although the Ciphering Indicator is a simple and efficient tool to alert users of possible communication interceptions, it seems to be neglected by both the industry and users. The results described in this paper emphasise the issue and can be employed to enhance awareness in both parties, considering the fact that security in mobile communications is an issue of growing concern.

REFERENCES

1. N. Jefferies, "Security in third-generation mobile systems", IEE Colloquium on Security in Networks (Digest No. 1995/024), **1995,** London, UK, pp.8/1-8/5.
2. R. J. Anderson, "Security Engineering: A Guide to Building Dependable Distributed Systems", 1st Edn., John Wiley and Sons, New York, **2001,** pp.354-358

3. L. Harte, R. Levine and G. Livingston, "GSM Superphones: Technologies and Services", 1st Edn., McGraw-Hill Professional, New York, **1998**, pp.121-152
4. E. Biham and O. Dunkelman, "Cryptanalysis of the A5/1 GSM stream cipher", *Lect. Notes Comput. Sci.*, **2000**, *1977*, 43-51.
5. E. Barkan, E. Biham and N. Keller, "Instant ciphertext-only cryptanalysis of GSM encrypted communication", *Lect. Notes Comput. Sci*., **2003**, *2729*, 600-616.
6. Technical Specification: "Digital cellular telecommunications system (Phase 2+), Security related network functions (GSM 03.20)", European Telecommunications Standards Institute, France, **2000**.
7. I. Androulidakis, "Intercepting mobile phone calls and short messages using a GSM tester", *Commun. Comput. Inform. Sci.*, **2011**, *160*, 281-288.
8. Technical Specification: "3rd Generation partnership project, technical specification group services and system aspects, 3G security, security architecture, V11.0.0", 3GPP Organizational Partners, Sophia Antipolis, Valbonne, France, **2011**, p.63.
9. K. Pentikoysis, "In search of energy-efficient mobile networking", *IEEE Commun. Magaz.*, **2010**, *48*, 95-103.
10. P. Baudisch, "My new PC is a mobile phone – techniques and technology for the new smallness", Proceedings of 12th International Conference on Human Computer Interaction with Mobile Devices and Services, **2010**, Lisbon, Portugal, pp.1-2.
11. K. A. Li, P. Baudisch and K. Hinckley, "Blindsight: Eyes-free access to mobile phones", Proceedings of 26th Annual SIGCHI Conference on Human Factors in Computing Systems, **2008**, Florence, Italy, pp.1389-1398.
12. V. Balakrishnan, S. F. Guan and R. G. Raj, "A one-mode-for-all predictor for text messaging", *Maejo Int. J. Sci. Technol.*, **2011**, *5*, 266-278.
13. D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond and D. Schmalstieg, "Real-time detection and tracking for augmented reality on mobile phones," *IEEE Trans. Visualiz. Comput. Graphics*, **2010**, *16*, 355-368.
14. Technical Specification: "Digital cellular telecommunications system (Phase2+), Specification of the subscriber identity module - mobile equipment (SIM - ME) interface, (GSM 11.11 V.5.5.0)", European Telecommunications Standards Institute, Sophia Antipolis, Valbonne, France, **1998**, p. 61.
15. Technical Specification: "European digital cellular telecommunications system (Phase 2), Security aspects, (GSM 02.09 V.4.4.0)", European Telecommunications Standards Institute, Sophia Antipolis, Valbonne, France, **1994**, p.12.
16. Technical Specification: "Digital cellular telecommunications system (Phase 2), Mobile Stations (MS) features, (GSM 02.07 V.4.8.2)", European Telecommunications Standards Institute, Sophia Antipolis, Valbonne, France**, 1998**, p.20.
17. Technical Specification, "Digital cellular telecommunications system (Phase 2+), Specification of the subscriber identity module - mobile equipment (SIM-ME) interface", European Telecommunications Standards Institute, Sophia Antipolis Cedex, France, **2005**, p.175.

18. Technical Specification: "3rd Generation partnership project, Technical specification group terminals, Characteristics of the USIM Application (V6.5.0)", 3GPP Organizational Partners, Sophia Antipolis, Valbonne, France, **2004**, p.94.
19. Technical Specification: "3rd Generation partnership project, Technical specification group services and system aspects service aspects, Service principles (Release 8, V8.11.0)", European Telecommunications Standards Institute, Sophia Antipolis, Valbonne, France**, 2008**, p.54.
20. User Guide: "8922M/S GSM Test Set", Agilent Part No. 08922-90211, Agilent Technologies Inc., South Queensferry (UK), **1998**, p.293.
21. Technical Specification: "The international identification plan for mobile terminals and mobile users", International Telecommunication Union, Geneva, Switzerland, **1998**, p.16.
22. I. Androulidakis, V. Christou, N. G. Bardis and I. Stilios, "Surveying users' practices regarding mobile phones' security features", Proceedings of 3rd International Conference on European Computing, **2009**, Tbilisi, Georgia, pp.25-30.
23. I. Androulidakis and G. Kandus, "A survey on saving personal data in the mobile phone", Proceedings of 6th International Conference on Availability, Reliability and Security, **2011**, Vienna, Austria, pp.633-638.
24. I. Androulidakis and G. Kandus, "Mobile phone brand categorization vs. users' security practices", *Eng. Technol. Appl. Sci. Res.*, **2011**, *1*, 30-35.
25. I. Androulidakis, "On a versatile and costless OMR system", *WSEAS Trans. Comput.*, **2005**, *2*, 160-165.
26. SPSS Technical, "SPSS for Windows, Rel. 16.0.2", SPSS Inc, Chicago, **2008**.