MAEJO INTERNATIONAL JOURNAL
OF SCIENCE AND TECHNOLOGY

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

CONTENTS

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.bioedit.co.uk (bioscience and medical papers), www.bioscienceeditingsolutions.com, www.scribendi.com, www.letpub.com, www.papersconsulting.com, www.stickerediting.com, Cambridge Proofreading (http://proofreading.org/), www.proofreadingservices.com, www.horizonproofreaders.org or www.manuscript-proofreding.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.

Websites :

7. S. Simon, "What is an odds ratio?", **2008**, http://www.childrensmercy.org/stats/definitions/or.htm (Accessed: October 2011).

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.

Manuscript Processing Time

As a result of a large number of submissions, there may be a long delay in the evaluation or publication of a paper. A duration of at least 6-8 mouths between submission and acceptance (or rejection) can normally be expected.

Maejo International Journal of Science and Technology

Communication

Antioxidant and anticancer activities of freshwater green algae, *Cladophora glomerata* and *Microspora floccosa*, from Nan River in northern Thailand

Ratiphan Laungsuwon[1, 2] and Warawut Chulalaksananukul[1-4,]*

[1] Biotechnology Program, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

[2] Department of Botany, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

[3] Aquatic Resources Research Institute, Chulalongkorn University, Bangkok 10330, Thailand

[4] Biofuels by Biocatalysts Research Unit, Chulalongkorn University, Bangkok 10330, Thailand

* Corresponding author, e-mail: warawut.c@chula.ac.th; tel/fax: +662 218 5482

Abstract: Organic solvent and hot water extracts of freshwater macroalgae, *Cladophora glomerata* and *Microspora floccosa,* harvested from Nan River in northern Thailand were screened for antioxidant and anticancer activities using DPPH free radical scavenging assay and inhibition of proliferation of the KB human oral cancer cell lines respectively. The ethyl acetate extract of *C. glomerata* showed the highest total phenol content (18.1±2.3 mg GAE/g), radical scavenging activity (49.8±2.7% DPPH scavenging at 100 μg/ml) and in vitro growth inhibition (IC$_{50}$=1420.0±66 μg/g) of the KB cell lines. These results indicate that *C. glomerata* could be a source of valuable bioactive materials.

INTRODUCTION

Algae are an important source of various bioactive compounds such as antioxidants, antimicrobials and antivirals [1]. These compounds are also important for protecting the algal cells against stressful conditions, e.g. ultraviolet radiation, temperature change and fluctuation in nutrient and salinity level. To enable rapid adaptation to new environmental conditions, algae produce a great variety of secondary metabolites that cannot be found in other organisms [2].

The biomass of macroalgae, represented mainly by a few species of Rhodophyta and Phaeophyta, is traditionally used to produce phycocolloids such as agar-agar, alginates and carrageenan. However, there are few reports on bioactive compounds from macroalgae. Some of the compounds are fucoidans and phlorotannins (algal polyphenols) produced by *Sargassum horneri* and *Ecklonia kurome*, which have been shown to have antiviral [3], antibacterial [4] and antioxidative properties [5]. Strong antioxidant activities of marine macroalgae have been reported [6-10], whereas those of freshwater macroalgae have been rarely investigated.

In northern Thailand, especially in Nan province, two species of freshwater macroalgae, *Cladophora glomerata* and *Microspora floccosa*, which belong to the Division Chlorophyta, are abundant in Nan River during the dry season [11]. The common names of these algae are "Kai" in Thai and "Mekong weed" in English. They have been used as a food source for many centuries by traditional culture. They are used in the manufacture of local food products such as crisps, baked goods, pasta and noodles. Besides being a popular food source, they are also believed to have many important health benefits such as rejuvenation, induction of appetite and expediting of recovery from many common maladies [12]. Some villagers also consume Kai to soothe stomach ulcers [13], but despite the widespread uses and claimed advantages of these algae, only few investigations on the chemical composition of *Cladophora* and *Microspora* species have been reported. Sterols, triterpenoids and volatile oils have been identified from some *Cladophora* species [14-16] which are distributed worldwide and often dominate in both fresh and marine waters [17]. While there have been several reports regarding antioxidant activities of green algae [6, 10, 18, 19], their anticancer activities have been rarely studied. As far as we know, the two species of freshwater green algae, *C. glomerata* and *M. floccose*, common in the Nan River, have not been evaluated for such bioactivities. The objective of this study is to determine the antioxidant and anticancer activities of various extracts of these two algae.

MATERIALS AND METHODS

Sampling and Identification

Cladophora glomerata Kiitzing and *Microspora floccose* (Vaucher) Thuret (Figure 1) were collected from three areas of Nan River, all located in Nan province, during the dry season (December 2009) when the algae were at their peak biomass (0-2 m in depth). Freshly collected algae were washed thoroughly in water to remove epiphytes, small invertebrates and extraneous matter. The samples were separated into two portions: one was used for morphological identification and the other was freeze-dried.

Preparation of Extracts and Preliminary Analysis

A 100-g portion of each freeze-dried macroalgal sample was extracted successively with 600 mL each of methanol, hexane and ethyl acetate at room temperature. Each extract was then clarified by centrifugation and the pellet was re-extracted twice with the same solvent. The supernatants were then pooled and filtered. The solvent was then removed from the filtrate by rotary evaporation and the dry crude extract was kept at 25°C and protected from light in a dessicator under an atmosphere of nitrogen gas until use.

Another 100-g sample of each freeze-dried alga was extracted with boiling deionised water for 1 hr and the water was removed by lyophilisation. The resulting crude extract was kept at 25°C and protected from light in a dessicator under an atmosphere of nitrogen gas until use.

All the extracts obtained were analysed by TLC on Kieselgel 60 F_{254} aluminum support plates. The spots were detected by UV irradiation (254 and 365 nm) and by heating.
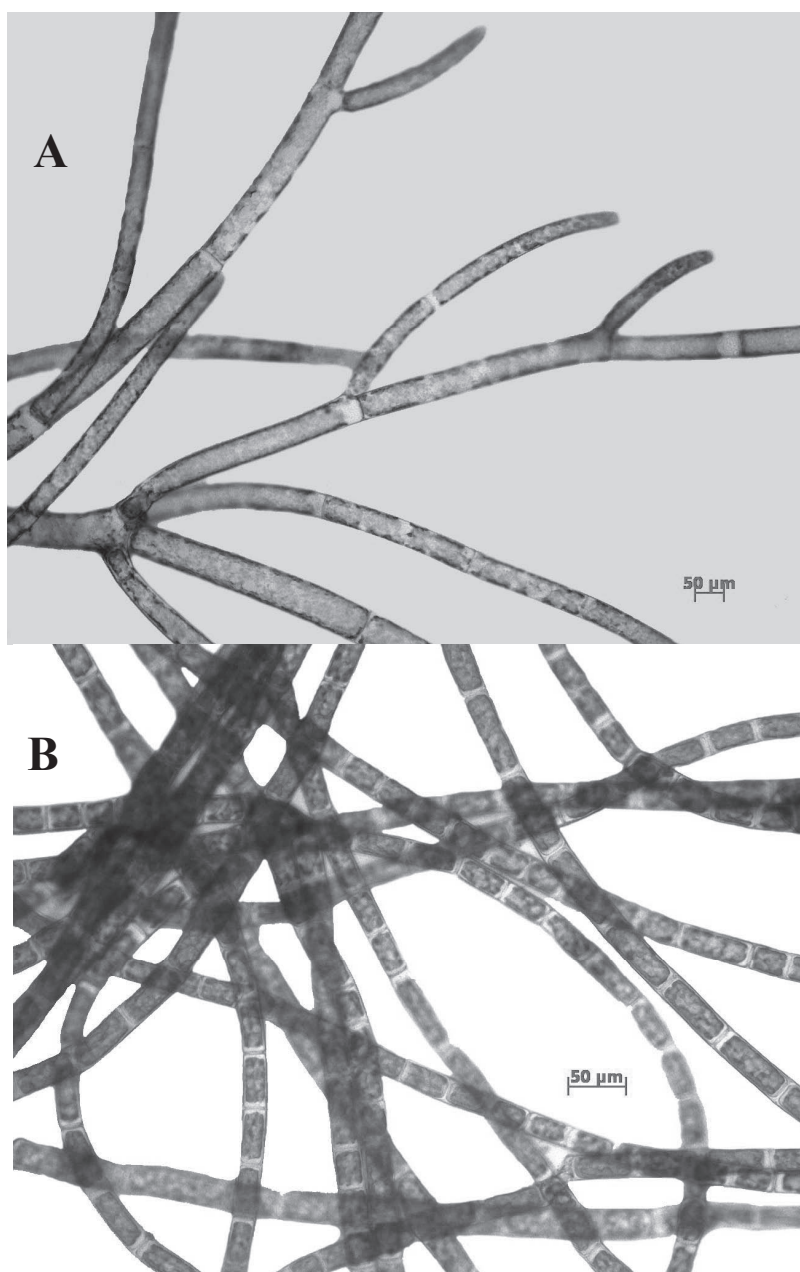
Figure 1. *Cladophora glomerata* Kützing (A) and *Microspora floccosa* (Vaucher) Thuret (B)

Determination of Total Phenolic Content (TPC)

The TPC was evaluated using Folin-Ciocalteu method as described previously [18, 19, 20] with modifications. Each extract (1.0 mL) was mixed with 1.0 mL of 2% Na_2CO_3 and 0.2 mL of 50% (v/v) Folin-Ciocalteu reagent was then added, mixed, allowed to stand at room temperature for 30 min. and then centrifuged. The absorbance of the supernatant was measured with a spectrophotometer at 750 nm. A calibration curve of gallic acid was prepared and the TPC were expressed as mg gallic acid equivalents (GAE)/g dry weight.

Free Radical Scavenging Activity by DPPH

The DPPH free-radical scavenging assay was performed according to established methods [20, 21] with some modifications. One mL of each extract in methanol at 100 μg/mL

was added to 2 mL of a solution of 0.004% DPPH in methanol. The mixture was shaken vigorously and allowed to stand for 30 min. at room temperature in the dark. The absorbance of the resulting solution was measured at 517 nm and converted into per cent DPPH consumed using the following formula: %DPPH= {[$Abs_{control}$ −Abs_{sample} −Abs_{blank})/$Abs_{control}$} x 100, where DPPH (2.0 mL) + methanol (1.0 mL) was used for $Abs_{control}$; DPPH (2.0 mL) + extract (1.0 mL) was used for Abs_{sample}; and methanol (2.0 mL) and extract (1.0 mL) was used for Abs_{blank}. Ascorbic acid was used as positive control.

Anticancer Assay against Human Oral Cavity Cell Lines (KB)

These experiments were based on the resazurin microplate assay (REMA) as described by Brien et al [22]. In brief, three KB cell lines (epidermoid carcinoma of oral cavity, ATCC CCL-17) at logarithmic growth phase were harvested and diluted to 7×10^4 cells/mL in fresh medium. Five μL of test sample diluted in 5% DMSO and 45 μL of cell suspension were successively added to a 384-well plate and incubated at 37°C in a 5% CO_2 incubator. After 3 days of incubation, 12.5 μL of 62.5 μg/mL resazurin solution was added to each well and the plate was then incubated at 37°C for 4 hr. Fluorescence signals were measured using a SpectraMax M5 multi-detection microplate reader (Molecular Devices, USA) at excitation and emission wavelengths of 530 nm and 590 nm respectively. Per cent inhibition of cell growth was calculated with the following equation: % Inhibition = [1- (FU_T/ FU_C)] x 100, where FU_T and FU_C are the mean fluorescent intensity from treated and untreated conditions respectively. Dose response curves were plotted from 6 concentrations of twofold serially diluted test compounds and the sample concentrations that inhibit cell growth by 50% (IC_{50}) were derived using SOFTMax Pro software (Molecular Devices, USA). Ellipticine and doxorubicin were used as positive controls and 0.5% DMSO as negative control.

Cytotoxicity against Normal Cell Lines (Vero)

The cytotoxicity experiments were based on the green fluorescent protein (GFP)-expressing Vero cell lines [23]. The method was generated in-house by stably transfecting the African green monkey kidney cell lines (Vero, ATCC CCL-81) with the pEGFP-N1 plasmid (Clontech). The cell line was maintained in minimal essential medium supplemented with 10% heat-inactivated fetal bovine serum, 2 mM L-glutamine, 1 mM sodium pyruvate, 1.5 g/L sodium bicarbonate and 0.8 mg/mL geneticin at 37°C in a humidified incubator with 5% CO_2.

The assay was carried out by adding 45 μL of cell suspension at 3.3×10^4 cells/mL to each well of a 384-well plate containing 5 μL of test compounds previously diluted in 0.5% DMSO. The plate was then incubated in a 37°C incubator with 5% CO_2 for 4 days. Fluorescence signals were measured with the SpectraMax M5 microplate reader in the bottom reading mode with excitation and emission wavelengths of 485 and 535 nm respectively. The fluorescence signal at day 4 was subtracted from the background fluorescence at day 0. The per cent inhibition was calculated with the following equation: % inhibition = [1-(FU_T / FU_C)] × 100, where FU_T and FU_C represent the fluorescence units of cells treated with test compound and untreated cells respectively.

The IC_{50} values (extract concentrations resulting in a 50% inhibition) were derived from dose-response curves using 6 concentrations of twofold serially diluted samples with the SOFTMax Pro software (Molecular Device, USA). Ellipticine and 0.5% DMSO were used as positive and negative controls respectively.

RESULTS AND DISCUSSION

TLC analysis of the extracts obtained from all locations revealed distinct different chemical profiles between the two algae. However, the same algal extracts from different locations showed similar TLC profiles and the presence of the same major secondary metabolites.

Phenolic compounds are known to act as antioxidants not only because of their ability to donate hydrogen atoms or electrons but also because of their stable radical intermediates [19], which prevent the oxidation of various ingredients, particularly fatty acids and oils. For example, minerals, polysaccharides and antioxidant properties were reported for macroalgae in the Noto Peninsula, Ishikawa, Japan [24]. The TPC of *C. glomerata* and *M. floccosa* extracts, when assayed at a concentration of 100 μg/mL, was found to vary with the extraction method and type of macroalga (Figure 2). The order of extraction efficiency varied between the two algae (ethyl acetate > hexane > hot water > methanol for *C. glomerata*, compared to hexane > methanol > ethyl acetate > hot water for *M. floccosa*). However, for each solvent the TPC level was always higher (2 times for methanol and hexane, and 3.3 times for ethyl acetate) for *C. glomerata* than *M. floccosa*.
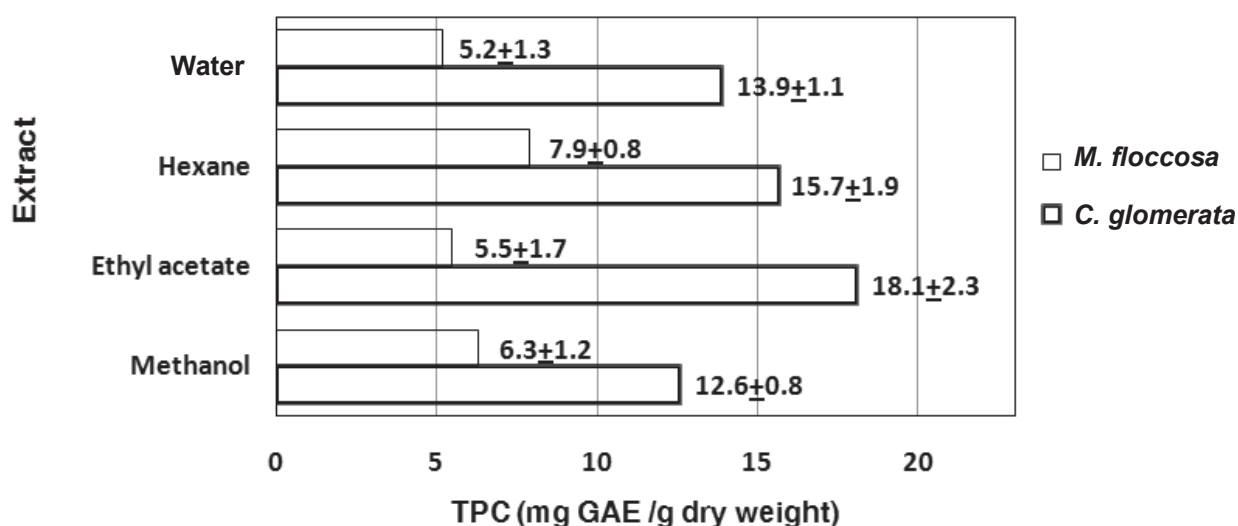
Figure 2. TPC of *C. glomerata* and *M. floccosa* extracts. The data are based on duplicates from three distinct areas of the Nan River. The results are expressed as means±SD (n=6).

Radical scavengers were evaluated in each of the *C. glomerata* and *M. floccosa* extracts by their reactivity towards the stable free radical DPPH (Figure 3). Similar to the TPC, the radical scavenger level varied with extraction method and type of macroalga. Indeed, the order of scavenging activity was observed to be the same as that of TPC, with %DPPH scavenging activity being 1.8 times (in methanol) to 5.6 times (in ethyl acetate) higher for *C. glomerata* than *M. floccosa*. However, all %DPPH scavenging activities observed were significantly lower than that of the ascorbic acid positive control at the same concentration.

From the determination of the effect of *C. glomerata* and *M. floccosa* extracts on in vitro inhibition of the growth (metabolism) of the KB cell lines in tissue culture, a significant decrease in the total cellular metabolic (reductase) activity (assumed number of viable cells) compared to the negative control was observed for the hexane and ethyl acetate extracts but not for the more polar methanol and hot water extracts of *C. glomerata*. In contrast, none of the four solvent extracts from *M. floccosa* elicited any significant cytotoxicity against the KB cell lines. The IC_{50}

of ethyl acetate and hexane extracts of *C. glomerata* were 1420.0±66 and 1662.0±48 μg/g respectively against the KB cell lines, and 2622.0±44 and 2574.0±30 μg/g respectively against the Vero cell lines (Table 1). These results suggest that *C. glomerata* contains certain useful biological compounds that have anticancer activity against KB cells and low cytotoxic against Vero cells. However, it is not yet known if this is a general cytotoxic activity towards any human cell lines or is indeed carcinoma-specific. Future research is needed to delineate the relative contribution of this pathway to cytotoxicity.
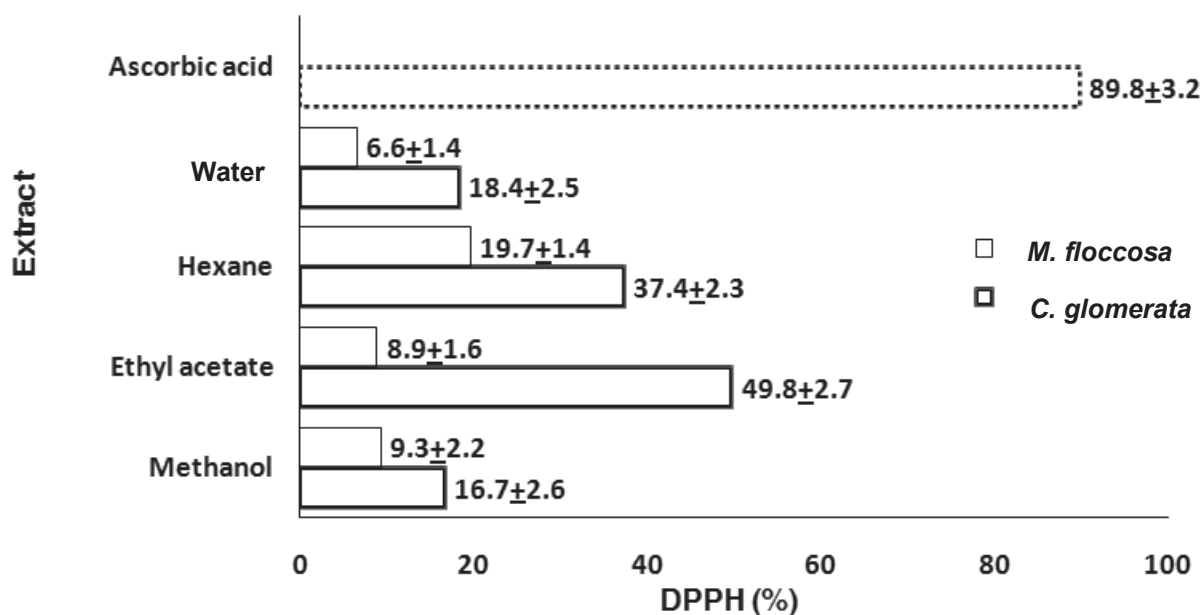
Figure 3. % DPPH radical scavenging activity of *C. glomerata* and *M. floccosa* extracts. The data are based on duplicates from three distinct areas of the Nan River. The results are expressed as means±SD (n=6).

Table 1. Cytotoxicity tests for *C. glomerata* and *M. floccosa* extracts against cell lines

Cell lines	IC$_{50}$ (μg/g)							
	C. glomerata				*M. floccosa*			
	Hot water	Methanol	Ethyl acetate	Hexane	Hot water	Methanol	Ethyl acetate	Hexane
KB	>10000	>10000	1420.0±66	1662.0±48	>10000	>10000	>10000	>10000
Vero	>5000	>5000	2622.0±44	2574.0±30	>5000	>5000	>5000	>5000

Note: The results are expressed as means±SD (n=6)

CONCLUSIONS

It has been shown that of the two freshwater green algae studied, i.e. *Cladophora glomerata* and *Microspora floccosa*, the former is a potential source of biologically active compounds that may be useful as therapeutic agents including an anticancer.

REFERENCES

1. O. Pulz and W. Gross, "Valuable products from biotechnology of microalgae", *Appl. Microbiol. Biotechnol.*, **2004**, *65*, 635-648.

2. I. Rodriguez-Meizoso, L. Jaime, S. Santoyo, F. J. Señoráns, A. Cifuentes and E. Ibáñez, "Subcritical water extraction and characterization of bioactive compounds from *Haematococcus pluvialis* microalga", *J. Pharm. Biomed. Anal.*, **2010**, *51*, 456-463.

3. T. Hoshino, T. Hayashi, K. Hayashi, J. Hamada, J. B. Lee and U. Sankawa, "An antivirally active sulfated polysaccharide from *Sargassum horneri* (Tunner) C. AGARDH", *Biol. Pharm. Bull.*, **1998**, *21*, 730-734.

4. T. Kuda, T. Kunii, H. Goto, T. Suzuki and T. Yano, "Varieties of antioxidant and antibacterial properties of *Ecklonia stolonifera* and *Ecklonia kurome* products harvested and processed in the Noto peninsula, Japan", *Food Chem.*, **2007**, *103*, 900-905.

5. S. J. Heo, E. J. Park, K. W. Lee and Y. J. Jeon, "Antioxidant activities of enzymatic extracts from brown seaweeds", *Bioresour. Technol.*, **2005**, *96*, 1613-1623.

6. Y. L. Chew, Y. Y. Lim, M. Omar and K. S. Khoo, "Antioxidant activity of three edible seaweeds from two areas in South East Asia", *LWT-Food Sci. Technol.*, **2008**, *41*, 1067-1072.

7. K. Iwai, "Antidiabetic and antioxidant effects of polyphenols in brown alga *Ecklonia stolonifera* in genetically diabetic KK-A(y) mice", *Plant Foods Hum. Nutr.*, **2008**, *63*, 163-169.

8. T. Kuda, M. Tsunekawa, H. Goto and Y. Araki, "Antioxidant properties of four edible algae harvested in the Noto Peninsula, Japan", *J. Food Comp. Anal.*, **2005**, *18*, 625-633.

9. Y. V. Yuan and N. A. Walsh, "Antioxidant and antiproliferative activities of extracts from a variety of edible seaweeds", *Food Chem. Toxicol.*, **2006**, *44*, 1144-1150.

10. M. Zubia, D. Robledo and Y. Freile-Pelegrin, "Antioxidant activities in tropical marine macroalgae from the Yucatan Peninsula, Mexico", *J. Appl. Phycol.*, **2007**, *19*, 449-458.

11. Y. Peerapornpisal, I. Pongsirikul and D. Kanjanapothi, "Potential of freshwater macroalgae as food and medicine", Final report submitted to Thailand Research Fund (TRF), **2005**.

12. P. Fahprathanchai, K. Saenphet, Y. Peerapornpisal and S. Aritajat, "Toxicological evaluation of *Cladophora glomerata* Kützing and *Microspora flocosa* Thuret in albino rats", *SE Asian J. Trop. Med. Pub. Health*, **2006**, *37*, 206-209.

13. Y. Peerapornpisal, D. Amornledpison, C. Rujjanawate, K. Ruangrita and D. Kanjanapothib, "Two endemic species of macroalgae in Nan River, northern Thailand, as therapeutic agents", *Sci. Asia*, **2006**, *32*, 71-76.

14. M. Kuniyoshi, K. Yamada and T. Higa, "A biologically active diphenyl ether from the green alga *Cladophora fascicularis*", *Experientia*, **1985**, *41*, 523-524.

15. S. V. Khotimchenko, "Distribution of glyceroglycolipids in marine algae and grasses", *Chem. Nat. Compd.*, **2002**, *38*, 223-229.

16. X. Huang, X. Zhu, L. Deng, Z. Deng and W. Lin, "Cycloartane triterpenes from marine green alga *Cladophora fascicularis*", *Chin. J. Oceanol. Limnol.*, **2006**, *24*, 443-448.

17. I. Elenkov, K. Stefanov, S. Dimitrova-Konaklieva and S. Popov, "Effect of salinity on lipid composition of *Cladophora vagabunda*", *Phytochem.*, **1996**, *42*, 39-44.

18. H-B. Li, K-W. Cheng, C-C. Wong, K-W. Fan, F. Chen and Y. Jiang, "Evaluation of antioxidant capacity and total phenolic content of different fractions of selected microalgae", *Food Chem.*, **2007**, *102*, 771-776.

19. M. Hajimahmoodi, M. A. Faramarzi, N. Mohammadi, N. Soltani, M. R. Oveisi and N. Nafissi-Varcheh, "Evaluation of antioxidant properties and total phenolic contents of some strains of microalgae", *J. Appl. Phycol.*, **2010**, *22*, 43-50.

20. G. Miliauskas, P. R. Venskutonis and T. A. van Beek, "Screening of radical scavenging activity of some medicinal and aromatic plant extracts", *Food Chem.*, **2004**, *85*, 231-237.

21. B-G. Wang, W-W. Zhang, X-J. Duan and X-M. Li, "*In vitro* antioxidative activities of extract and semi-purified fractions of the marine red alga, *Rhodomela confervoides* (Rhodomelaceae)", *Food Chem.*, **2009**, *113*, 1101-1105.

22. J. O′Brien, I. Wilson, T. Orton and F. Pognan, "Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity", *Eur. J. Biochem.*, **2000**, *267*, 5421-5426.

23. L. Hunt, M. Jordan, M. De Jesus and F. M. Wurm, "GFP-expressing mammalian cells for fast, sensitive, noninvasive cell growth assessment in a kinetic mode", *Biotechnol. Bioeng.*, **1999**, *65*, 201-205.

24. T. Kuda and T. Ikemori, "Minerals, polysaccharides and antioxidant properties of aqueous solutions obtained from macroalgal beach-casts in the Noto Peninsula, Ishikawa, Japan", *Food Chem.*, **2009**, *112*, 575-581.

Maejo International
Journal of Science and Technology

Full Paper

Effects of rejuvenator seal and fog seal on performance of open-graded friction course pavement

Nadeem A. Qureshi [1,*], Nam H. Tran [2], Donald Watson [2] and Syed M. Jamil [1]

[1] National Institute of Transportation (NIT), National University of Science and Technology, Islamabad, Pakistan
[2] National Centre for Asphalt Technology, 277 Technology Parkway, Auburn, AL, 36830

* Corresponding author, E-mail: nadeemqureshi612000@yahoo.com

Abstract: An open-graded friction course (OGFC) is a special-purpose surface layer of hot-mix asphalt (HMA) pavement that is increasingly being used around the world. Owing to its numerous benefits, OGFC is being regularly used as a final riding surface on interstate and high-traffic expressways by different highway agencies in the United States. However, some OGFC sections have experienced premature failure due to ravelling only after 6-8 years of service life. To maintain an effective, longer service life and enhanced performance of OGFC, preventive maintenance has been considered essential. There are several approaches to maintaining OGFC, one of which is the application of a fog seal and rejuvenator seal. A fog seal can reduce ravelling and extend the service life of OGFC while a rejuvenator seal can revitalise the existing aged asphalt binder in the top OGFC layer. This research focuses on optimising the fog and rejuvenator seal application rates by evaluating their effectiveness in terms of surface friction and durability. Three types of seal material were evaluated: Pavegaard (PG) and Pavepreserve (PP) asphalt rejuvenators and a cationic slow-setting asphalt emulsion (CSS-1H) as a fog seal. Improvement in abrasion resistance of OGFC pavement was observed on application of fog and rejuvenator seals but surface friction was reduced to some extent. Hamburg test clearly shows a trend that the medium application rate of 0.10 gallon/square yard is better in enhancing resistance to rutting/moisture susceptibility of OGFC.

Keywords: rejuvenator seal, fog seal, hot-mix asphalt pavement, open-graded friction course, surface friction

INTRODUCTION

An open-graded friction course (OGFC) is a thin and permeable surface layer of a hot-mix asphalt (HMA) mixture that incorporates a coarse aggregate skeleton with minimum fines. The load is supported through stone-to-stone contact and the asphalt binder keeps the skeleton intact. This inherent attribute of OGFC enhances resistance to rutting and its porous nature ensures immediate drainage of water from the pavement surface. It provides numerous benefits for the road users in terms of safety, environment and economy, including improved friction, minimised hydroplaning, reduction of splash and spray, improvement of night visibility and reduction of noise level [1]. However, the damaging actions of air, water, temperature and traffic over time raise durability issues in OGFC pavements. These issues lead to the problem of ravelling and reduced performance life [2, 3].

Advances in mixture design, construction techniques and practices, and use of durable materials, especially fibre- and polymer-modified asphalt binders have improved the performance and durability of OGFC pavements [4]. Regular preventive maintenance is one of the key measures to maintain good performance and durability of the OGFC pavement during its service life. Surface cleaning and application of fog seals or rejuvenator seals are two major methods used in preventive maintenance of OGFC. The surface cleaning technique is mainly adopted in Europe and Japan to maintain porosity and unclog OGFC pavements. Current maintenance activities in Denmark include cleaning of the voids by high-pressure water and air suction twice a year in order to maintain porosity during the pavement lifetime [5]. On the other hand, Japan is adopting frequent cleaning operations with only partial debris removal during each cleaning operation [6]. High-pressure washing is currently quite expensive and of questionable value. It is believed by local agencies in the United States that the OGFC functionality can be maintained by its self-cleaning capacity created in highways with relatively high-speed and high-volume traffic because of the suction generated by rolling tyres on the pavements [7].

Fog seal application is used as one of the preventive maintenance techniques in 5 out of 17 states in the US, where OGFC is routinely used. It provides a small film of unaged binder at the surface but friction and reduction in porosity can still be expected [8, 9]. The Federal Highway Administration recommends fog seal in two applications at a rate of 0.05 gallon/square yard (gal/sy) for each application, using a 50% dilution of asphalt emulsion without any rejuvenator agent [10]. Fog seal is utilised in New Mexico, Wyoming, South Carolina, Oregon and California to execute preventive maintenance [9, 11]. It is typically a light spray application of dilute asphalt emulsion used primarily to seal the existing asphalt surface to reduce ravelling and enrich a dry and weathered surface [12]. It enriches the top aged surface and its penetration to a certain depth is expected to extend the service life of OGFC [7, 13, 14]. Rejuvenator seals are maltene-based petroleum products that revitalise the existing aged asphalt binder in the top OGFC layer [15, 16]. The effectiveness of a rejuvenator is judged by the deep penetration to avoid a slippery surface. The main difference between a fog seal and a rejuvenator seal is in the chemical make-up. An asphalt rejuvenator is a petroleum product based on maltene (lighter component of asphalt), whereas a fog seal is an asphalt-based emulsion. Normally rejuvenator seals, rather than fog seals, are used for more aged and ravelled pavements. Fog or rejuvenator seals with or without sand application on OGFC have diverse effects depending on the application rate, age and condition of the pavement.

Estakhri and Agarwal [15] suggested that a fog seal application rate of 0.05 gal/sy was not effective in reducing the aging rate based on limited data of HMA. Another study indicated a 30% reduction of oxidation in HMA by an application of Gilsonite-Sealer-Binder (GSB) [17]. The stiffness and loss of surface fines in a high air-void-content mixture could be significantly reduced by the rejuvenators [15].

However, there are two trade-offs that must be considered when a fog or rejuvenator seal is applied on an OGFC surface. First, the seal will fill the surface voids and may reduce the drainage capacity of OGFC. Over time, the OGFC may become much like a dense-graded mix with little drainability and little reduction in splash and spray. Second, the seal may temporarily reduce the surface friction, or an inappropriate use of the seal may result in a slick surface [15]. If there is a significant loss in friction, the potential of reduced safety outweighs the benefits of applying the seal.

Owing to its numerous benefits, OGFC is being regularly used as the final riding surface on interstate and high-traffic expressways by different highway agencies in the United States, including the Alabama Department of Transportation. However, some OGFC sections have experienced premature failure due to ravelling only after 6-8 years of service life. Although surface distress is evident in all lanes, severe failure is more predominant in the outside lanes. An example of this surface distress is shown in Figure 1.

Figure 1. Ravelling in southbound outside lane at Milepost 177.9 (near Prattville)

The maintenance practices of OGFC pavements currently in vogue are not yielding effective results. The application rate of the fog and rejuvenator seals could be one of the probable causes of this issue. The objective of this study is to evaluate the effectiveness of the fog and rejuvenator seal application rates for the performance of OGFC in order to recommend an optimum application rate. The performance is assessed in terms of surface friction and durability.

MATERIALS AND METHODS

Methodology

The research methodology for evaluating the effects of using fog and rejuvenator seals on the performance of OGFC based on durability and surface friction is as illustrated in Figure 2.

Three types of seal material were evaluated, namely two asphalt rejuvenators: Pavegaard (PG) and Pavepreserve (PP), and a fog seal: a cationic slow setting (CSS-1H) asphalt emulsion [18-20]. These seal materials were obtained from Martin Company (Houston, Texas). Each material was applied at three different application rates. The evaluation was conducted through both the field work and laboratory testing. The former involved applying each fog seal material at a predetermined application rate on two OGFC test sections at the National Centre for Asphalt Technology (NCAT) pavement test track in Opelika, Alabama. Then the surface friction characteristics of treated and untreated OGFC surfaces were tested using a dynamic friction tester (DFT) and a circular texture meter (CTM). Cores were then cut from treated and untreated OGFC surfaces for laboratory testing, which included bulk specific gravity (Gmb) measurement, Hamburg wheel-tracking device (HWTD) test and Cantabro abrasion test.
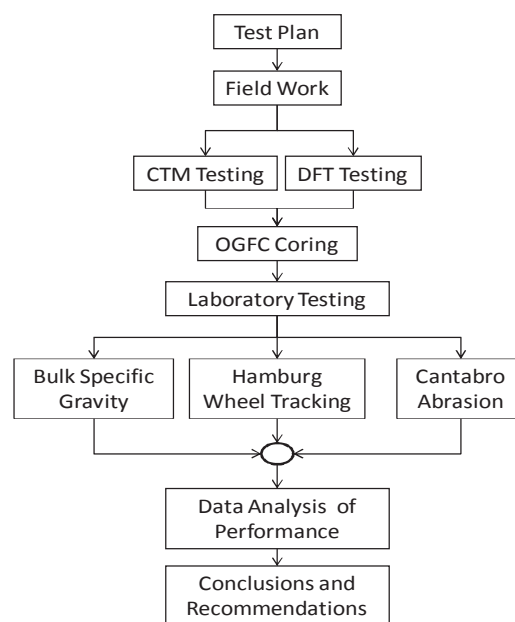
Figure 2. Steps for evaluating effects of fog and rejuvenator seals on performance of OGFC pavement

The test site, two sections (W4 and W5) on the inside lane at the NCAT pavement test track, was selected for field evaluation. The two sections were surfaced with two OGFC mixtures in 2000. The two OGFC mixtures used in sections W4 and W5 consisted of the same granite aggregate gradation (Table 1) and similar binder contents of 6.1% and 6.2% respectively. Modified binders of styrene butadiene rubber (SBR) and styrene butadiene styrene (SBS), which met the requirements of performance grade 76-22, were used in sections W4 and W5 respectively. The average thickness of the OGFC surfaces for sections W4 and W5 was 1.1 in. and 0.7 in. respectively, based on the thickness measurement of the cores extracted from the two sections. These sections were not trafficked regularly but were used for moving construction equipment during the construction, reconstruction and maintenance of the outside lane, which was trafficked during each research cycle. The site selection posed an inherent limitation as the effect of traffic on aging of asphalt binder was lacking in the real sense.

Table 1. Design gradation for OGFC mixture used in test sections W4 and W5

Sieve size, mm	Per cent passing by weight
19.00 (3/4 in.)	100.0
12.50 (1/2 in.)	95.0
9.50 (3/8 in.)	66.0
4.75 (#4)	23.0
2.36 (#8)	14.0
0.07 (#200)	8.6

Field Testing

Figure 3 shows the layout of test sections W4 and W5 containing squares (20x20 in. each) for evaluating the fog seal and rejuvenator seals as well as the type of fog/rejuvenator seal material and its application rate for each 20x20 in. square. In each square, a fog/rejuvenator seal material was evenly sprayed (Figure 4) at a predetermined application rate, and the surface was not sanded after the application because the fines from the sanding process might fill in the surface voids causing an adverse effect on the drainability of OGFC. To control the application rate, the sprayer was weighed before and during the spray application to determine the amount of fog/rejuvenator seal material applied in each square. In addition, a 2x2 in. geosynthetic pad whose weight had been predetermined was placed at the centre of each square during the spray application and then removed and oven-dried to a constant mass to determine the actual application rate of fog/ rejuvenator seal material for each square.

After the fog and rejuvenator seals had been cured for one week, the friction and macro texture characteristics of the OGFC surface in each square were measured by DFT in accordance with American Society for Testing and Materials (ASTM), Test no. E1911-09 AE01 [21] and by CTM in accordance with ASTM, Test no. E2157-09 [22]. After that, five 6-in. cores were taken from each square for laboratory testing.

Laboratory Testing

The five full depth cores (including the OGFC wearing course and the Superpave mixture in the underlying layer) taken from each square (except squares containing CSS-1H), shown in Figure 3, were used for further testing in the laboratory to determine the effects of the rejuvenators on the durability of OGFC. For each set of five cores, two cores were used for bulk specific gravity measurement and Cantabro abrasion test, two for HWTD test and one saved for future testing.

For bulk specific gravity measurement and Cantabro abrasion test, the OGFC surface layers were cut from two full-depth cores to prepare two test specimens. The bulk specific gravity of each OGFC specimen was determined using automatic vacuum sealing method in accordance with ASTM, Test no. D6752/D6752M-11 [23]. The specimens were then used for Cantabro abrasion test in accordance with Texas Department of Transportation standard [24].

For HWTD test, the top 1.5-in. layers including the OGFC wearing course and a part of the Superpave mix in the underlying layer were cut from two full-depth cores to prepare two test specimens. The two specimens were then used to run one HWTD test in compliance with American Association of State Highway and Transportation Officials (AASHTO) [25].

Square/Pad No.	Con	1	2	3	4	5	6	7	8	9
seal material										
Application rate gal/sy	0	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15
Section	**Section W4 (SBR Modified)**									

Square/Pad No.	Con	13	14	15	16	17	18	19	20	21
seal material										
Application rate gal/sy	0	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15
Section	**Section W5 (SBS Modified)**									

N ⟷ S

- Control
- Pavepreserve (PP), Rejuvenator Seal
- Pavegaard (PG), Rejuvenator Seal
- CSS - 1H, Fog Seal

Figure 3. Layout of test sections for evaluating fog and rejuvenator seals

20 × 20 in. Square

2 × 2 in. Square Pad

Figure 4. Spray application of seal material on OGFC pavement

194

RESULTS AND DISCUSSION

Micro and Macro Surface Friction Characteristics

Analyses of CTM and DFT testing results were conducted to assess the micro and macro surface friction characteristics on application of fog and rejuvenator seals. The macro and micro texture analyses were conducted for the mean profile depth (MPD) obtained from CTM and friction number measured at 20 km/h (DFT_{20}) and at 40 km/h (DFT_{40}) using DFT. The international friction index (IFI) parameter F_{60} was calculated and analysed to find the mutual effect of macrotexture and microtexture of the pavement surface. The IFI consists of two parameters: F_{60} and speed constant (Sp). F_{60} is the harmonised estimate of the friction at 60 km/hr computed from both the friction measurement and Sp, whereas Sp is linearly related to macrotexture measurements. The IFI parameter F_{60} can be estimated based on DFT and CTM results using Equation 1 as given in ASTM E1960-07(2011) [26]:

$$F_{60} = 0.081 + 0.732 \times DFT_{20} \times e^{\frac{-40}{Sp}}$$

(1)

where:

F_{60} = international friction index
DFT_{20} = friction number obtained at 20 km/h using DFT
Sp = speed constant = $14.2 + 89.7 \cdot$ MPD
MPD = mean profile depth obtained from CTM

Two analysis of variance (ANOVA) tests for all the measured MPD and DFT_{20} data were conducted. The results of these ANOVA tests at 95% confidence level ($P< 0.001$) show that the effect of the existing surfaces in sections W4 and W5 on the MPD and DFT_{20} measurements were statistically significant. Hence the effects of applying fog seal on the surface friction characteristics of OGFC were analysed separately for sections W4 and W5.

Figure 5 shows a graphical comparison of the effect of fog and rejuvenator seals on MPD for sections W4 and W5. ANOVA results indicate that the application of fog seal (CSS-1H) significantly affected MPD of the OGFC surfaces, especially at the medium and high application rates (0.1 and 0.15 gal/sy).

Figure 6 shows a graphical comparison of the effect of fog and rejuvenator seals on friction numbers for sections W4 and W5. ANOVA results show that the application of rejuvenator seals, especially PG for section W4, significantly affect DFT_{20} and DFT_{40}. For section W5, the application of fog seal and PG at low and medium application rates (0.05 and 0.1 gal/sy) significantly affected DFT_{20} and DFT_{40}. In the case of section W5, application of PP at high application rate significantly affected DFT_{20} and DFT_{40}.

Figure 7 shows the per cent decrease in IFI parameter F_{60} of OGFC due to the application of fog and rejuvenator seals for sections W4 and W5. The decrease in F_{60} depends on the type of modified binder used on the existing surface (W4 versus W5), the type of rejuvenator or fog seal material, and the application rate. The surface friction based on F_{60} was reduced 2-24%. Thus, fog and rejuvenator seals should be used with extreme caution on OGFC as they may cause a temporary loss of friction. Fog seal (CSS-1H) showed similar effects on the surface friction characteristics of the two sections. Rejuvenator seals (PG and PP) affected the surface friction characteristic of section W4 more than that of section W5.
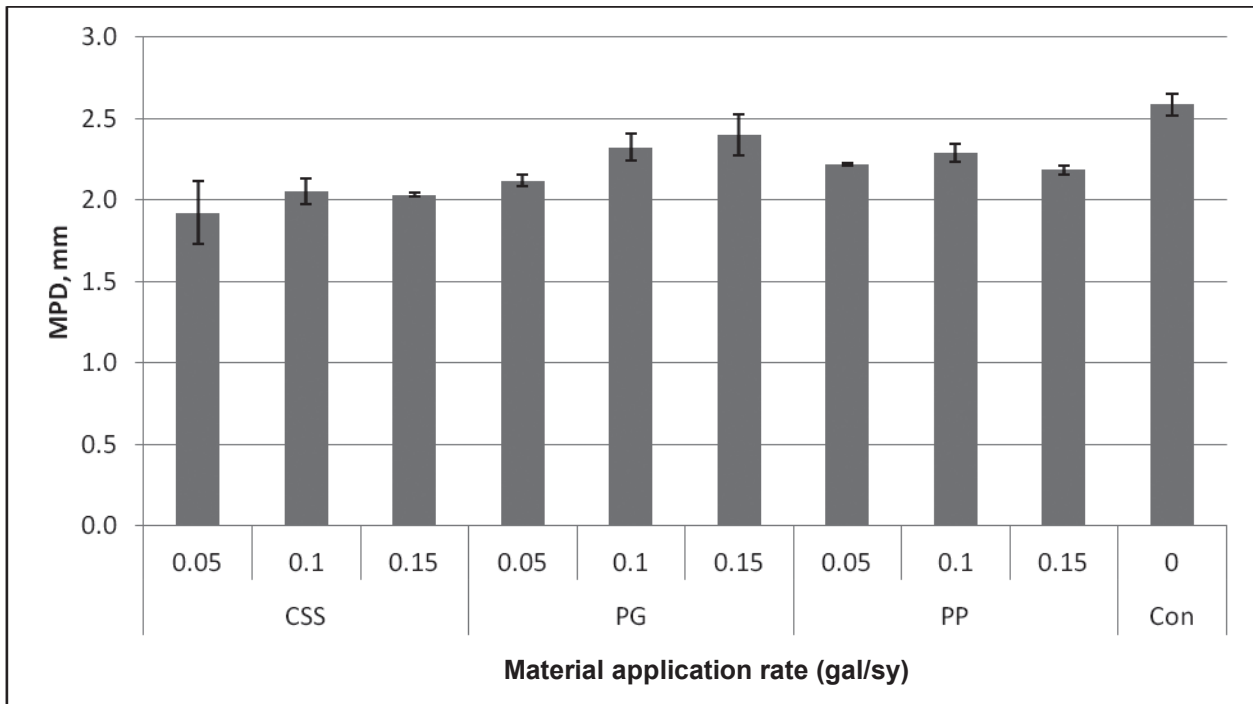
(a) Section W4
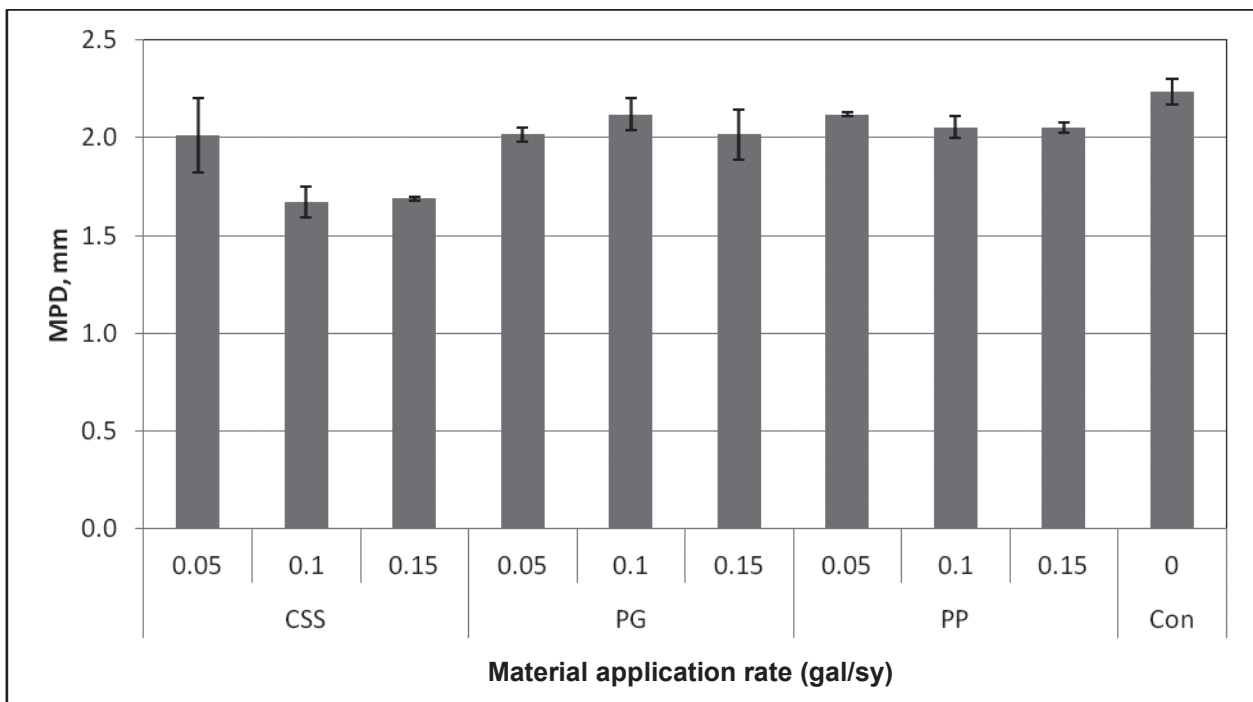
(b) Section 5

Figure 5. Effect of fog and rejuvenator seals on MPD for sections W4 and W5
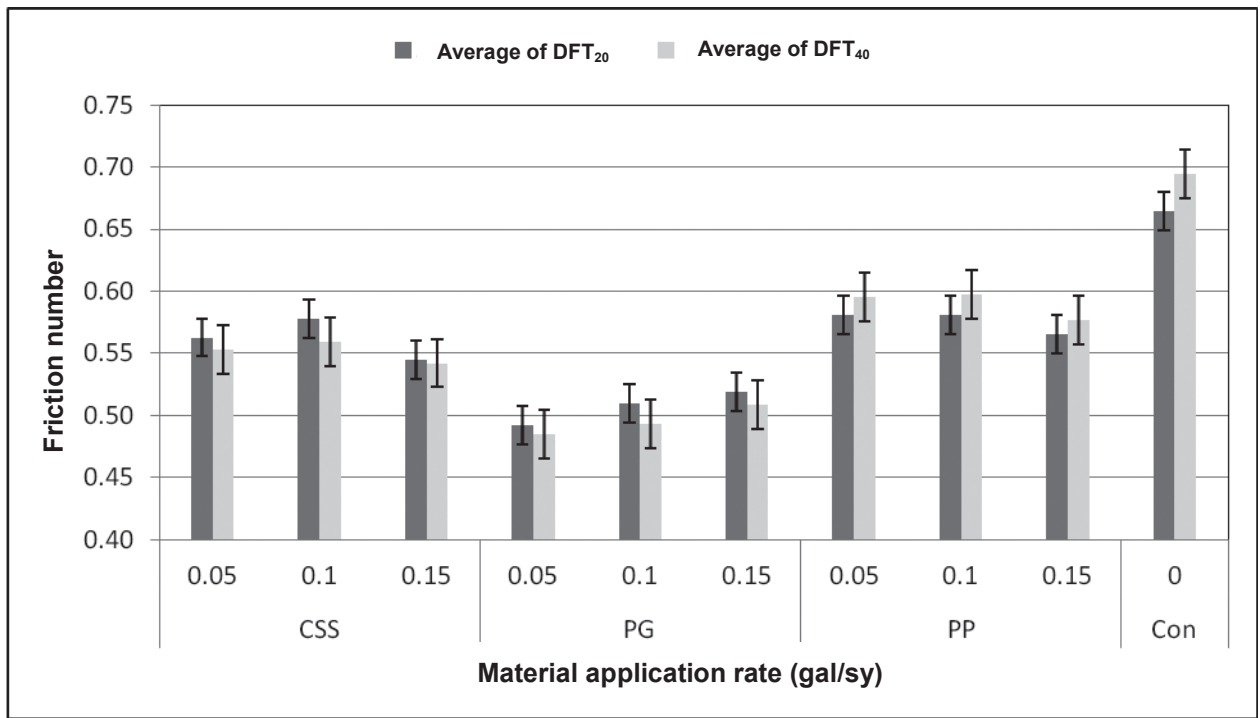
(a) Section W4
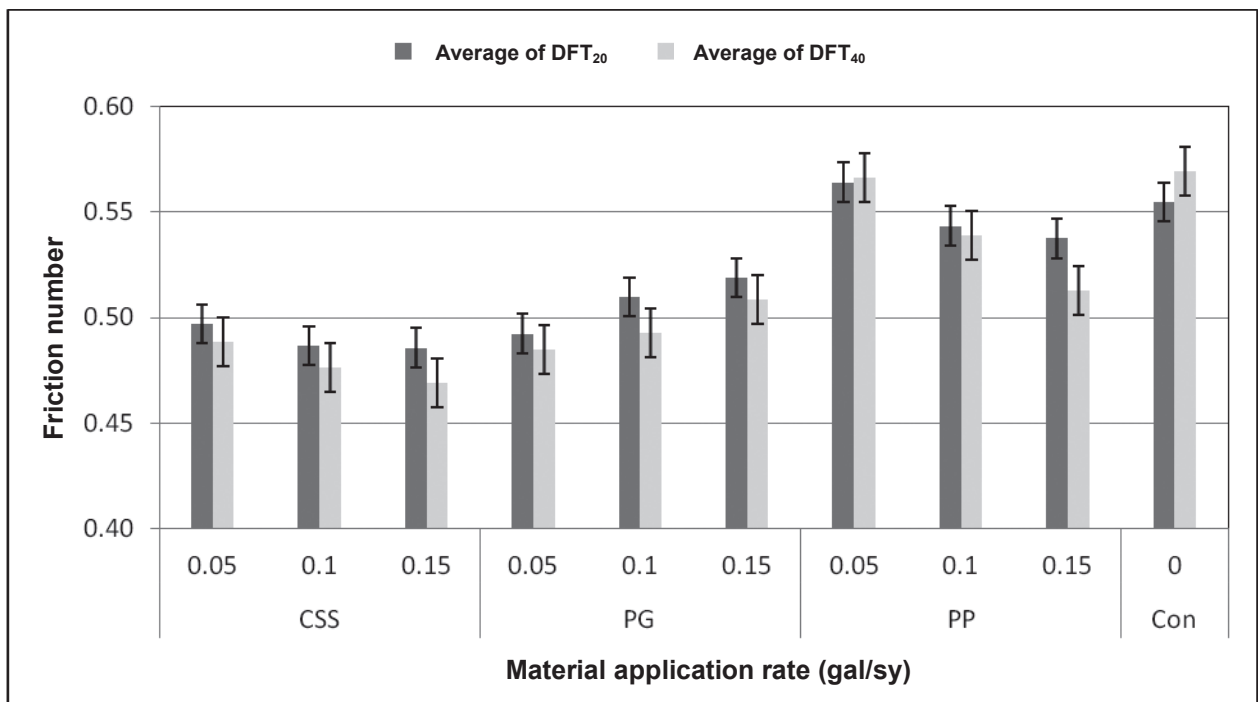
(b) Section W5

Figure 6. Effects of fog and rejuvenator seals on friction number for sections W4 and W5
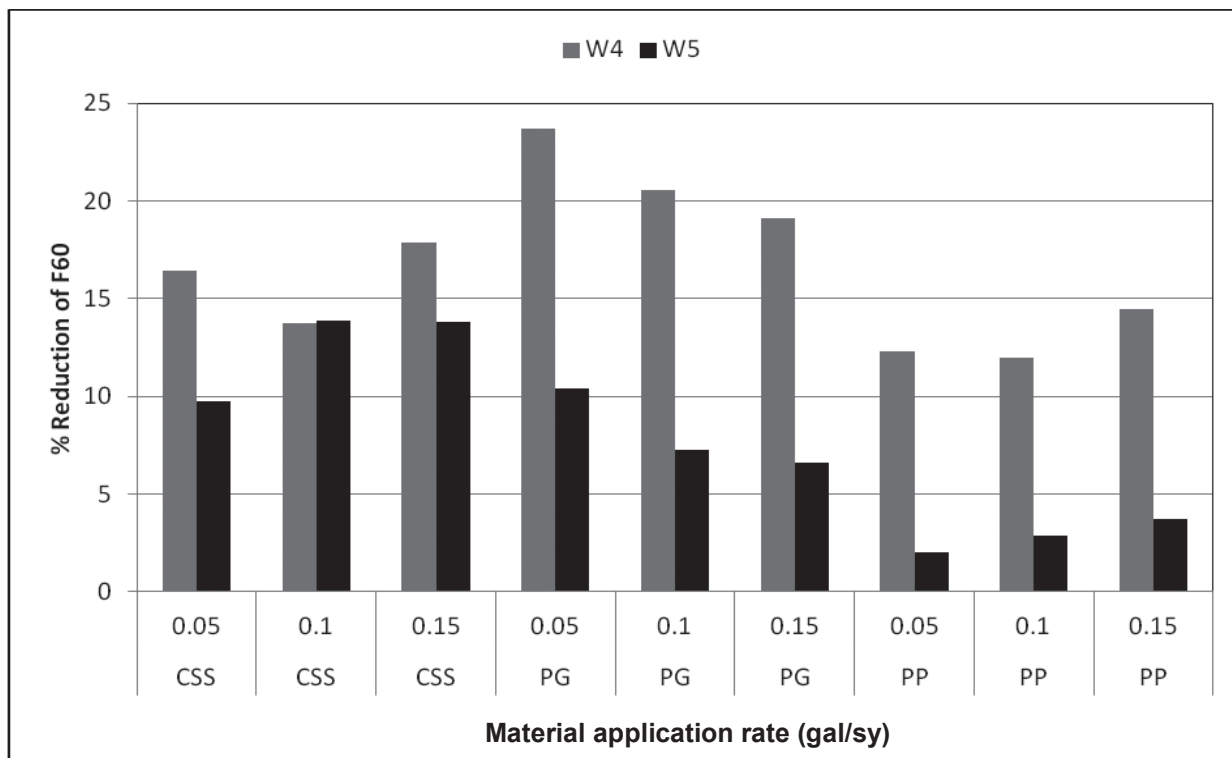
Figure 7. Per cent decrease in F_{60} due to fog and rejuvenator seals on OGFC

Air Void Measurement

Analysis of air voids was carried out to assess their impact on the functionality of OGFC on application of rejuvenator seal. ANOVA tests for all the measured air voids were conducted separately for sections W4 and W5. The results of the tests at 95% confidence level ($P = 0.25$ and $R^2 = 39.04\%$ for section W4, and $P = 0.033$ and $R^2 = 58.08\%$ for section W5) show that the differences in air voids were not statistically significant ($P > 0.05$) for section W4 while for

section W5 the statistical significance ($P < 0.05$) was at the borderline.

Figure 8 shows the air void measurements for the control squares with no treatment and for other squares with the rejuvenator seals sprayed at 0.05, 0.10 and 0.15 gal/sy. A trend of reduced air voids with increase in rejuvenator seal application rate, resulting in reduction in functionality especially for section W5, can be observed. This is a matter of concern as the capacity to drain water through OGFC pavements would be reduced and the functionality of the mix would be affected.

Cantabro Test Results

The Cantabro test indicates the mixture resistance to wear and ravelling [6, 27] and is recommended for use in a standard OGFC mix design procedure based on previous NCAT research [7, 14]. This test has been used to predict the durability of OGFC pavements during their service life.

Analysis of Cantabro per cent loss was conducted to evaluate the improvement in resistance to wear and ravelling on application of a rejuvenator seal. ANOVA tests were done for all values of Cantabro per cent loss. The results the tests at 95% confidence level ($P = 0.413$ and $R^2 = 50.21\%$ for section W4, and $P = 0.167$ and $R^2 = 65.03\%$ for section W5) show that the Cantabro per cent losses were not statistically significant ($P < 0.05$) for sections W4 and W5,

although they showed an intermediate existing relationship (R^2 = 50.21% for section W4 and R^2 = 65.03% for section W5).

Figure 9 shows the Cantabro loss results for OGFC specimens extracted from control squares and squares with the two rejuvenator seal products sprayed at the rates of 0.05, 0.10 and 0.15 gal/sy. The loss values were higher than what would be acceptable during mix design but they wee probably due to the age of the pavement and the thickness of the cores used for testing. Resistance to abrasion usually improves with an increase in binder content [3] and in this case each of the rejuvenator seal products appeared to improve the abrasion resistance. These results indicate that an application rate of 0.10 gal/sy or more should be suitable depending on the type of rejuvenator seal applied and the type of modified binder used on the existing surface (W4 versus W5) of OGFC pavement.
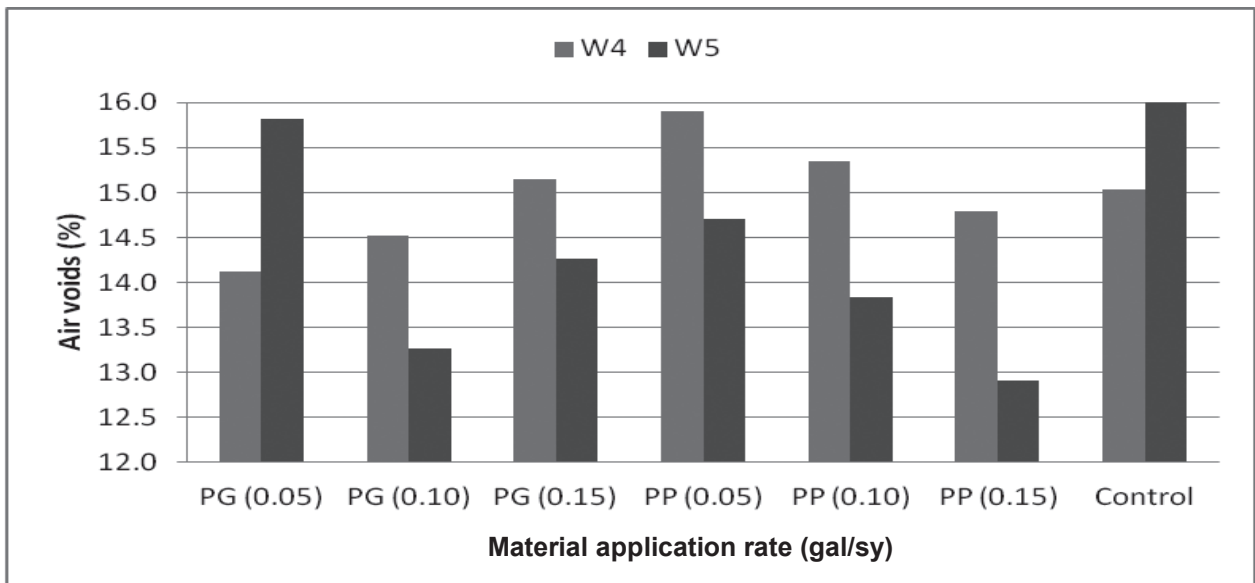
Figure 8. Per cent specimen air voids resulting from spraying with rejuvenator seals
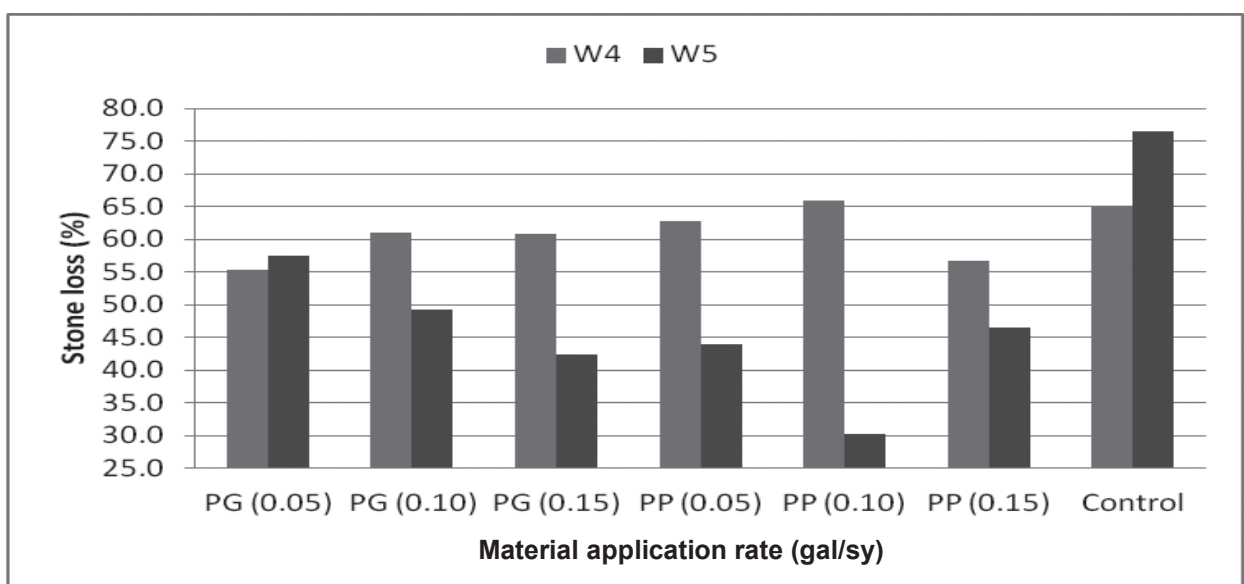
Figure 9. Effects of treatment type and application rate on Cantabro stone loss

Hamburg Test Results

The Hamburg wheel-tracking device (HWTD) was introduced in the United States as a result of the 1990 European asphalt study tour [6]. Analysis of Hamburg rut depth was conducted to evaluate the improvement in resistance to rutting on application of a rejuvenator seal. ANOVA tests comparing the rut depth to material application rate were conducted together for sections W4 and W5. The results of the tests at the 95% confidence level ($P = 0.047$ and $R^2 = 77.28\%$) show that the statistical significance ($P < 0.05$) of the rut depth for both sections was at the borderline, although there existed a strong relationship ($R^2 = 77.28\%$).

Figure 10 shows the average rut depth of cores (using the Hamburg rutting procedure) from control squares and squares with the two types of rejuvenator seal (PG and PP) applied with three different application rates (0.05, 0.10 and 0.15 gal/sy). Contrary to all other laboratory test data, the HWTD test results clearly show a trend that the medium application rate (0.10 gal/sy) was better in enhancing resistance to rutting and/or moisture. Based on these results, it appears that the typical application rate of 0.05 gal/sy used in the past may not be adequate to improve resistance to rutting damage.
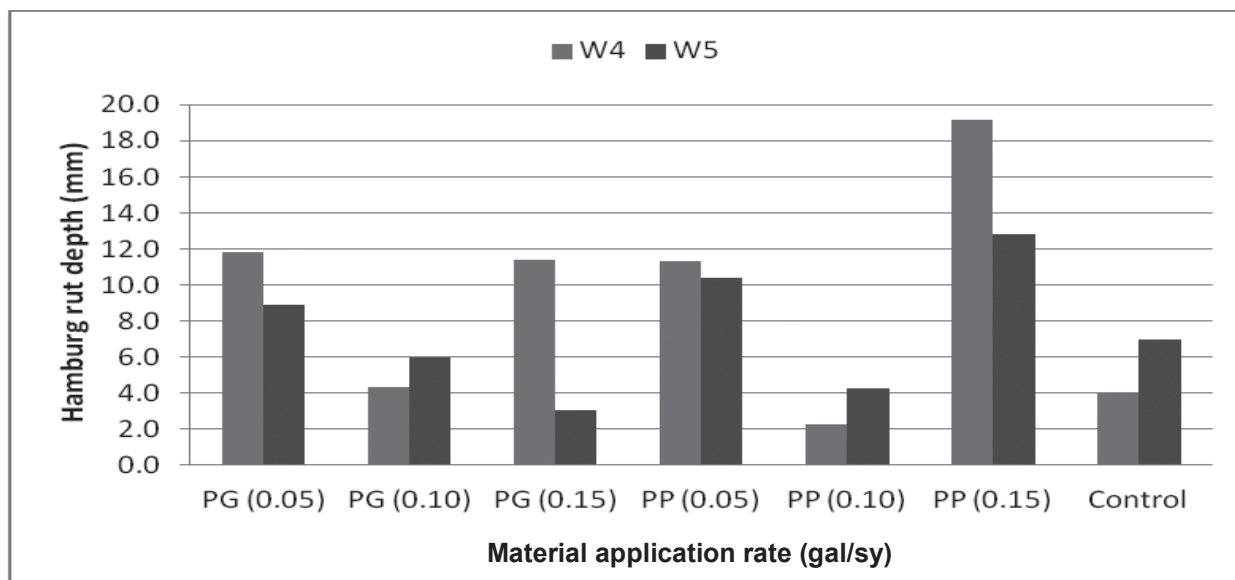
Figure 10. Average Hamburg rut depth after 10,000 cycles

CONCLUSIONS AND RECOMMENDATIONS

Fog and rejuvenator seals significantly affect the micro and macro texture of the OGFC surfaces in general. The surface friction may be reduced up to 24 per cent depending on the type of modified binder used in the pavement, the type of rejuvenator or fog seal material, and the application rate. Therefore, fog and rejuvenator seals should be used with caution on OGFC as they may cause a temporary loss of friction. A trend of reduced air voids was observed with increase in rejuvenator/fog seal application rates, which is a matter of concern as the functionality of OGFC mix would be affected. The Cantabro loss values were much higher than what would be acceptable for the OGFC mix design, but the probable reason for the high loss was the age and low thickness of the OGFC cores used for testing. The rejuvenator seals appear to improve the abrasion resistance. The Cantabro test indicates that an application rate of 0.10 gal/sy or more may be suitable depending on the type of rejuvenator being applied and the type of modified binder used in the pavement. Contrary to all other laboratory tests, the HWTD test

results clearly show a trend that the medium application rate (0.10 gal/sy) is better in improving resistance to rutting.

In order to expand and further validate this research, it is recommended that the OGFC sections from lightly and heavily trafficked interstate highways are selected for study and monitoring of the effect of rejuvenator and fog seals. Other rejuvenator and fog seal materials should also be selected to broaden the scope and efficacy of the research.

ACKNOWLEDGEMENT

REFERENCES

1. D. Watson, A. Johnson and D.J ared, "Georgia DOT's Progress in Open-Graded Friction Course Development", Transportation Research Board 1616, National Research Council, Washington, D.C., **1988**, pp.30-33.
2. E. R. Brown, P. S. Kandhal, F. L. Roberts, Y. R. Kim, D. Y. Lee and T. W. Kennedy, "Hot Mix Asphalt Materials, Mixture Design, and Construction", 2nd Edn., National Asphalt Pavement Association Research and Educational Foundation, Lanham (MD), **2009**, pp. 45-60.
3. "Pavement Interactive", http://pavementinteractive.org/index.php? title=Durability (Accessed: April 2010).
4. R. B. Mallick, P. S. Kandhal, L. A. Cooley and D. E. Watson, "Design, construction, and performance of new generation open-graded friction course mixes", NCAT report No. 2000-01, National Center for Asphalt Technology, Auburn University, Auburn, USA, **2000**, pp.18-19.
5. S. N. Thomsen, J. Kragh, E. Nielsen, H. Bendtsen and B. Andersen, "Noise Reducing Pavements−State of the Art in Denmark", Danish Road Institute, Copenhagen, **2005**, p.26.
6. T. D. Larson, "Report on the 1990 European asphalt study tour", American Association of State Highway and Transportation Officials, Washington, DC, **1991**, p.168.
7. P. S. Kandhal, "Design, Construction, and Maintenance of Open-Graded Asphalt Friction Courses", National Asphalt Pavement Association, Lanham (MD), **2002**, p.22.
8. D. F. Rogge, "Development of Maintenance Practices for Oregon F-mix", Oregon Department of Transportation, Washington, D.C., **2002**, pp.37-49.
9. A. E. Alvarez, A. E. Martin, C. K. Estakhri, J. W. Button, C. J. Glover and S. H. Jung, "Synthesis of Current Practice on the Design, Construction, and Maintenance of Porous Friction Courses", Texas Transportation Institute, Texas A&M University, College Station (TX), **2006**, pp.37-38.
10. "Technical advisory: Open graded friction courses, T 5040.31", Federal Highway Administration, U.S. Department of Transportation, Washington D.C., **1990**.
11. S. Shatnawi and B. D. Toepfer, "Fog Seal Guidelines", Office of Flexible Pavement Materials, Sacramento (CA), **2003**, p.20.
12. "A Basic Asphalt Emulsion Manual", 3rd Edn., Asphalt Institute, Lexington (KY), **1997**, p. 132.
13. E. H. S. Booth, R. Gaughan and G. Holleran, "Some uses of bitumen emulsions in SA and NSW", Proceedings of 14th Australian Road Research Board Conference, **1988**, Canberra, Australia, pp.387-401.

14. D. E. Watson, L. A. Cooley, K. A. Moore and K. Williams, "Laboratory Performance Testing of Open-Graded Friction Course Mixtures", Transportation Research Board, Washington, DC, **2004**, pp. 40-47.

15. C. K. Estakhri and H. Agarwal, "Effectiveness of Fog Seals and Rejuvenators for Bituminous Pavement Surfaces", Texas Transportation Institute, Texas A&M University, College Station (TX), **1991**, pp.3-4.

16. S. Shatnawi, "Maintenance Technical Advisory Guide, Volume 1: Flexible Pavement Preservation", 2nd Edn., California Department of Transportation, Sacramento (CA), **2008**, p.1.

17. Asphalt System, Inc. (The Solution), www.asi-roads.com/thesolution/thesolution.asp (Accessed: July 2011).

18. Martin Asphalt Company, "Paveguard product data sheet", http://www.themartincompanies. com/sites/themartincompanies.com/files/asphalt/pds/Product_Data_Sheet_CRF.pdf (Accessed: July 2011).

19. Martin Asphalt Company, "Pavepreserve product data sheet", http://www.themartin companies.com/sites/themartincompanies.com/files/asphalt/pds/Product_Data_Sheet_Recl amite.pdf (Accessed: July 2011).

20. Martin Asphalt Company, "CSS-1H product data sheet", http://www.themartincompanies. com/sites/themartincompanies.com/files/asphalt/pds/CSS-1h.pdf (Accessed: July 2011).

21. ASTM Standard E1911, 2009 AE01, "Standard test method for measuring paved surface frictional properties using the dynamic friction tester," ASTM International, West Conshohocken (PA), **2009**, DOI: 10.1520/E1911-09AE01 (www.astm.org).

22. ASTM Standard E2157, 2009, "Standard test method for measuring pavement macrotexture properties using the circular track meter", ASTM International, West Conshohocken (PA), **2009**, DOI: 10.1520/ E2157-09 (www.astm.org).

23. ASTM Standard D6752 / D6752M, 2011, "Standard test method for bulk specific gravity and density of compacted bituminous mixtures using automatic vacuum sealing method", ASTM International, West Conshohocken (PA), **2011**, DOI: 10.1520/D6752_D6752M-11 (www. astm.org).

24. Texas Department of Transportation Standard, "T-245-F: Cantabro loss", **2005**, http://ftp. dot.state.tx.us/pub/txdot-info/cst/TMS/200-F_series/pdfs/bit245.pdf (Accessed: July 2010).

25. AASHTO T 324-08, "Standard method of test for Hamburg wheel-track testing of compacted hot mix asphalt", American Association of State Highway and Transportation Officials, Washington, D.C. **2008**.

26. ASTM Standard 1960, 2007 (2011), "Standard practice for calculating international friction index of a pavement surface", ASTM International, West Conshohocken (PA), **2011**, DOI: 10.1520/E1960-07R11 (www.astm.org).

27. Sabita Manuals, "The Design and Use of Porous Asphalt Mixes: Manual 17", Southern African Bitumen and Tar Association, Roggebaai (South Africa),**1995**, pp.1-10.

Maejo International
Journal of Science and Technology

Report

Phytoplankton composition of Sazlidere Dam lake, Istanbul, Turkey

Nese Yilmaz

Department of Freshwater Biology, Fisheries Faculty, Istanbul University, Ordu St. No: 200 Laleli, Istanbul, Turkey

E-mail: nyilmaz@istanbul.edu.tr; tel: +90 212 455 57 00; fax: +90 212 514 03 79

Abstract: The phytoplankton composition of Sazlidere Dam lake was studied at 5 sampling sites between December 2003 - November 2005. A total of 67 taxa were recorded, representing Bacillariophyta (31), Chlorophyta (18), Cyanophyta (9), Chrysophyta (1), Cryptophyta (1), Dinophyta (3) and Euglenophyta (4). Bacillariophyta members constituted the dominant phytoplankton group in terms of species number. Nygaard's compound index value and composition of phytoplankton indicate that the trophic state of Sazlidere Dam lake was changing from oligotrophic to mesotrophic.

Keywords: phytoplankton, taxonomy, Sazlidere Dam lake, Turkey

INTRODUCTION

Water is essential for the survival of all organisms on the earth. About 97% of earth's water is found in the ocean, 2% is frozen as ice in the poles and the remaining 1% is available in the form of freshwater, which is used for daily needs, irrigation and industries [1]. Population growth, overurbanisation, integrated industry and increasing use of natural resources lead to water pollution problems in Turkey, as well as the rest of the world [2]. Access to clean and safe water has become more difficult because of the gradual decreases in available fresh water.

Phytoplankton are the primary producers in the food chain and they are very important organisms in aquatic systems. Phytoplankton species can be used as indicators of water quality, given their sensitivity and dynamic responses to changes in the surrounding environment [3]. According to the EU Water Framework Directive introduced in 2000, phytoplankton are one of the four biological elements suggested for assessing the ecological status and potential of surface waters in Europe [4, 5]. Taxonomic studies of phytoplankton are very effective in re-evaluation of the uses and stability of lakes. In particular, data on the algal flora of water sources are commonly used in the preservation, improvement and control of water quality. To date, few studies have been conducted to investigate the Sazlidere Dam lake and most of them are modelling studies. To our knowledge, this is the first report on the phytoplankton composition of Sazlidere Dam lake, one of the drinking water resources of the Istanbul metropolitan area. The

aim of the study is to determine the diversity of the phytoplanktonic algal flora of the Sazlidere reservoir.

STUDY AREA AND CLIMATE

Istanbul is located in the north-western part of Turkey and it extends both on European and Asian sides of the Bosphorus Strait. It is the only metropolis in the world that is situated on two continents. Its neighbours are the Black Sea in the north, Marmara Sea in the south, Kocaeli City in the east and Tekirdag City in the west. Istanbul has approximately 5100 km^2 of land area and a population of 13 million. In its long history, Istanbul served as a capital city of the Roman empire, the Byzantine empire, the Latin empire and the Ottoman empire. In addition to its rich history, high population and productive economy, Istanbul also has a wide variety of ecological features [6].

Istanbul is a kind of transition zone between the less rainy Mediterranean climate and the oceanic climate. In summer, less precipitation and high temperature are characteristic and the annual mean temperature is 14.5ºC for the last two decades. Between May-September the temperature is generally above 30ºC and between November-April it is rarely below 0ºC. The total precipitation for Istanbul averages 640 mm per year. The rain regime is winter-autumn-spring-summer and the rain type is Central Mediterranean. The relative humidity is between 73-77% in the city, which decreases to 65-68% in summer despite the effect of the seas. The dominant wind in the city is the northeast-orginated wind [7].

Sazlidere Dam is located in the Marmara region on the European peninsula of Istanbul, south-east of Lake Terkos (Durusu) and north-east of Lake Kucukcekmece (Figures 1-2). The dam is on the Sazlidere stream approximately 6 km from Lake Kucukcekmece. The reservoir aspect is 20 km east-west and 9 km north-south. The dam has a drainage area of 165 km^2 and average depth of 22.4 metres. It was completed in 1966 and is used to supply drinking and utility water, and as a sport fishing and recreational area. Sazlidere reservoir provides 55 million m^3 of water per year to Istanbul [8].

SAMPLING AND PHYTOPLANKTON IDENTIFICATION

This study was carried out between December 2003 - November 2005 at 5 sampling stations (Figure 1) on a monthly basis. The first station was at the stream input; station 2 was located in the littoral of Sazlibosna settlement area; station 3 was at the shore of the stone quarries; station 4 was situated in front of a road channel and station 5 was located on the deepest part of the dam, which is close to the dam embankment (Figures 3-7). The locations of the sampling sites are given in Table 1.

Samples were taken both on the surface and vertically in Nansen bottles and fixed with Lugol's iodine [9]. Phytoplankton species were identified using several comprehensive reports on the subject [10-19]. The proportional representation of phytoplankton groups is used to determine the primary production of lakes. Of those, Nygaard's compound index [20] is determined by dividing the total number of species of Cyanophyceae, Chlorococcales, Centrales and Euglenaceae by the number of species of Desmidiales. Nygaard's compound index was used to determine the trophic state of Sazlidere Dam.
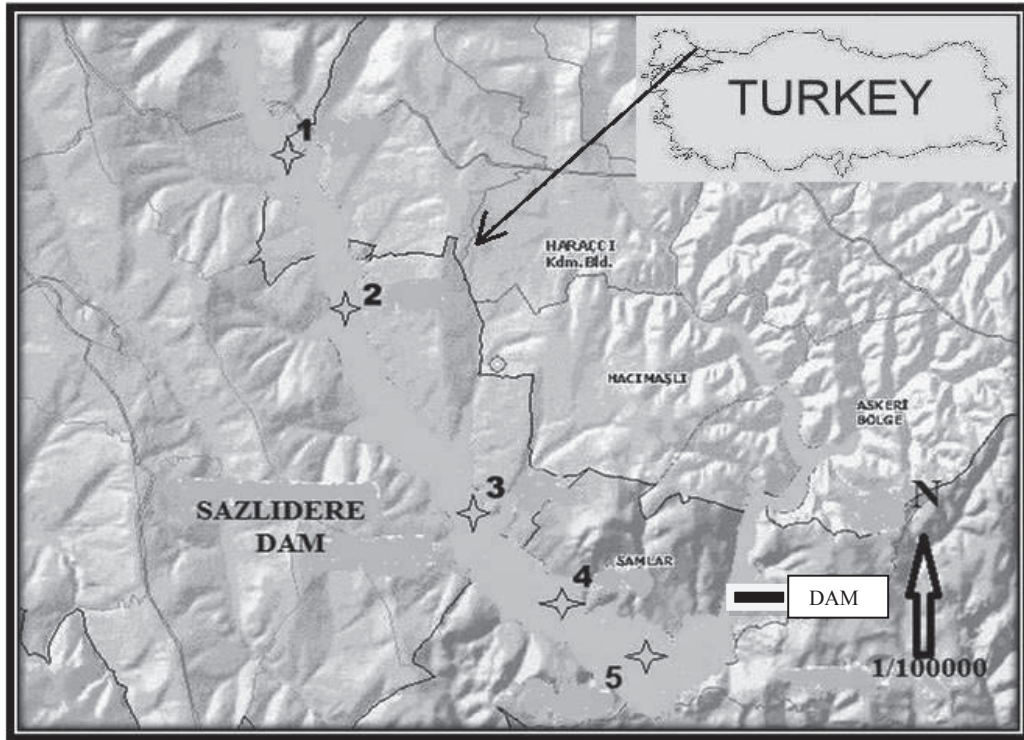
Figure 1. Map of Sazlidere Dam and sampling stations

Table 1. Locations of sampling stations

Station 1	N 41° 11 ' 31.4 "	E 028° 38 ' 37.9 "
Station 2	N 41° 09 ' 02.0 "	E 028° 40 ' 16.2 "
Station 3	N 41° 07 ' 57.9 "	E 028° 41 ' 10.02 "
Station 4	N 41° 07 ' 22.7 "	E 028° 42 ' 08.2 "
Station 5	N 41° 07 ' 08.09 "	E 028° 42 ' 34.5 "

Figure 2. Sazlidere Dam   **Figure 3.** Station 1

Figure 4. Station 2

Figure 5. Station 3

Figure 6. Station 4

Figure 7. Station 5

RESULTS AND DISCUSSION

A total of 67 phytoplankton taxa were identified. They belong to 7 divisions: Bacillariophyta (31 taxa), Chlorophyta (18 taxa), Cyanophyta (9 taxa), Euglenophyta (4 taxa), Dinophyta (3 taxa), Chrysophyta (1 taxon) and Cryptophyta (1 taxon). A list of the recorded taxa is given in Table 2 and the occurrence frequencies are given in Table 3. The major members of phytoplankton were Bacillariophyta, Chlorophyta and Cyanophyta and Bacillariophyta constituted the dominant phytoplankton group in terms of species numbers (48%). The distribution of phytoplankton groups by taxa (percentage) is shown in Figure 8.
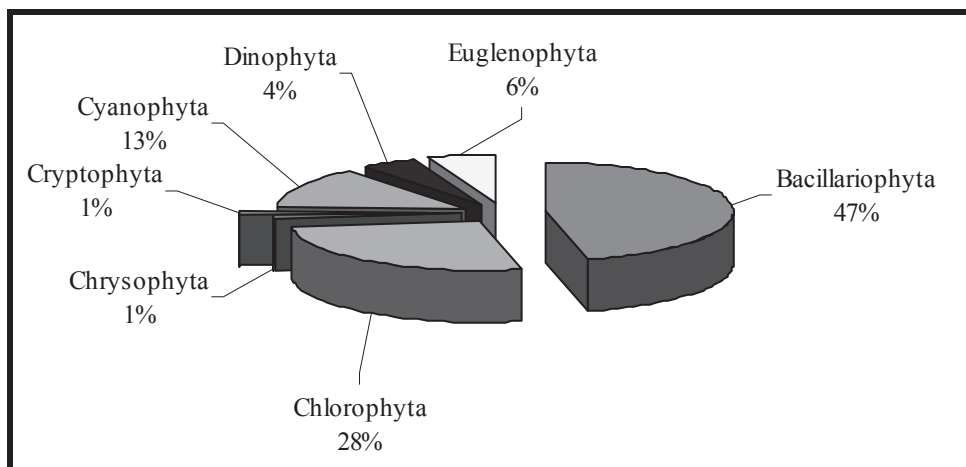
Figure 8. Sazlidere Dam phytoplankton diversity by division (percentage)

Table 2. List of recorded taxa

DIVISION BACILLARIOPHYTA
Order Centrales
Aulocoseira italica (Ehr.) Simonsen
Cyclotella atomus Hust.
Cyclotella ocellata Pant.
Melosira varians Ag.
Stephanodiscus astrea Grun.
Order Pennales
Achnanthes lanceolata (Breb.) Grun.
Amphora ovalis Kütz.
Cocconeis placentula Ehr.
Cymbelle affinis Kütz.
Cymbella tumida (Breb.) Van Heurck
Cymbella prostrata (Berkeley) Cl.
Diatoma vulgaris Bory
Diploneis elliptica (Kütz.) Cl.
Fragilaria crotonensis Kitton
Fragilaria pinnata Ehr.
Gomphonema clavatum Ehr.
Gomphonema truncatum Ehr.
Navicula ambigua Ehr.
Navicula cuspidata (Kütz.) Kütz.
Nitzschia acicularis (Kütz.) W. Sm.
Nitzshia linearis (Ag.) W. Sm.
Nitzschia palea (Kütz.) W. Sm.
Nitzschia panduriformis Gregory
Nitzschia sigmoidea (Nitzsch) W. Sm.
Pinnularia sp.
Rhapholodia gibba (Ehr.) O.Müll.
Rhoicosphenia sp.
Surirella ovata Breb.
Tabellaria sp.
Ulnaria acus (Kütz.) M. Aboal
Ulnaria ulna (Nitzsch) P. Compere

DIVISION CHLOROPHYTA
Order Chlorococcales
Coelastrum microporum (Naeg.) A. Braun
Kirchneriella sp.
Monoraphidium falcatus (Corda) Ralfs
Oocystis borgei J.Snow
Pediastrum dublex Meyen
Pediastrum simplex Meyen
Scenedesmus communis E.G.Hegewald
Scenedesmus dimorphus (Turp.) Kütz.
Scenedesmus quadricauda (Turp.) Breb.
Scenedesmus sp.
Sphaerocystis planctonica (Korshikov) Bourrelly

Tetraedron minumum (A. Braun) Hansgirg
Order Desmidiales
Closterium acutum (Breb.) Ralfs
Cosmarium depressum (Naeg.) P. Lundell
Cosmarium formosulum Hofmann
Staurastrum crenulatum (Naeg.) Delponte
Order Volvocales
Pandorina morum (O.F.Müller)Bory
Order Zygnematales
Mougeotia sp.

**DIVISION CYANOPHYTA
(CYANOBACTERIA)**
Order Chroococcales
Chroococcus limneticus Lemm.
Merismopedia glauca (Ehr.) Naeg.
Microcystis aeruginosa (Kütz.) Kütz.
Order Oscillatoriales
Oscillatoria tenuis C. Agarth Gomont
Oscillatoria sp.
Planktothrix rubescens D.C.
Spirulina major (Kütz.) Gomont
Order Nostocales
Anabaena flos-aquae [(Lyngb.) Breb.] Bornet *et*
Flauhault
Aphanizemenon ovalisporum Forti.

DIVISION EUGLENOPHYTA
Order Euglenales
Euglena gracilis G.A. Klebs
Phacus sp.
Trachelomonas hispida (Perty) F. Stein
Trachelomonas volvocina Ehr.

DIVISION DINOPHYTA
Order Peridiniales
Ceratium furca (Ehr.) Claparéde and Lachmann
Ceratium hirundinella (O.F. Müler) Dujardin
Peridinium bipes F. Stein

DIVISION CRYPTOPHYTA
Order Cryptomonadales
Cryptomonas ovata Ehr.

DIVISION CHRYSOPHYTA
Order Chrysomonadales
Dinobryon sertularia Ehr.

Table 3. Occurrence frequencies of phytoplankton in Sazlidere Dam (100-80% regularly present; 80-60% mostly present; 60-40% commonly present; 40-20% occasionally present; 20-1% rarely present)

OCCURRENCE (%)					
Total samples	20	20	20	18	20
Taxa	Station 1	Station 2	Station 3	Station 4	Station 5
BACILLARIOPHYTA					
Centrales					
Cyclotella atomus	61	60	90	90	70
Cyclotella ocellata	11	10	15	15	20
Pennales					
Cymbella prostrata	28	40	40	35	15
Fragilaria crotonensis	11	20	30	30	20
Navicula cuspidata	55	50	60	45	10
Ulnaria acus	39	15	30	20	15
Ulnaria ulna	33	55	55	55	40
CHLOROPHYTA					
Chlorococcales					
Scenedesmus quadricauda	28	35	2	30	35
CYANOPHYTA					
Oscillatoriales					
Oscillatoria tenuis	11	10	15	10	15
EUGLENOPHYTA					
Euglenales					
Trachelomonas hispida	17	10	15	20	10
DINOPHYTA					
Peridiniales					
Peridinium bipes	6	25	25	25	30
CRYPTOPHYTA					
Cryptomonadales					
Cryptomonas ovata	55	70	75	65	60
CHRYSOPHYTA					
Chrysomonadales					
Dinobryon sertularia	17	15	20	20	30

Bacillariophyta has been reported to be dominant in many algal studies in Turkish fresh waters [21-25]. The centric diatoms *Cyclotella atomus* and *Cyclotella ocellata*, which are typical components of oligotrophic lakes, were recorded in all seasons [26-27]. The benthic forms of the Order Pennales were based on wave-motion and wind exposure at the dam, as observed in previous studies at Lake Mogan, Lake Palandoken and Cakmak Dam lake [24, 28, 29]. It was reported that pennate diatoms were more abundant than centric diatoms in shallow lakes [28, 30]. *Ulnaria ulna*, a pennate diatom, is known to be characteristic of eutrophic lakes [31]. This species was also found to be dominant in Lake Hafik and Derbent Dam lake, which were oligotrophic [32, 33]. It was determined that the numbers of diatom taxa are higher in water samples taken in spring and summer. Generally, an increase in light and temperature leads to phytoplankton increase in spring. Diatoms, in particular, show greater abundance during this period [34].

Chlorophyta members are usually found widely and abundantly in mesotrophic and eutrophic lakes [27]. Furthermore, members of the Order Chlorococcales have been recorded in aquatic environments that are transitioning from oligotrophic to eutrophic character [27]. In this study *Sphaerocystis planctonica* and *Scenedesmus* species of the Order Chlorococcales were recorded mostly in autumn and winter samples.

Cyanophyta members usually prefer eutrophic environments [35]. *Oscillatoria tenuis* was found in all water samples. *Anabaena flos-aquae, Microcystis aeruginosa* and *Planktothrix rubescens* were recorded more in spring and summer samples than in other seasons.

Dinobryon sertularia (Chrysophyta) and *Euglena gracilis* (Euglenophyta) were recorded in all sampling stations. While *D. sertularia* is known to be characteristic of mesotrophic lakes [27], *E. gracilis* is found mostly in water with high levels of organic pollution [35].

When the vertical composition of the phytoplankton was examined, their individual numbers were found to decrease with increasing depth. As light penetration decreases with water depth, photosynthetic algae were found on and near the surface [34].

Nygaard's trophic state index can be a handy tool in determining the status of pollution in lakes. It involves only algal identification and the detailed analysis of physicochemical parameters can be omitted [20]. It was suggested by Rawson [36] that the numbers of species of certain groups present in phytoplankton would seem to have less ecological significance than the numbers of individuals of the dominant species [36]. A Nygaard compound index value less than 1 indicates oligotrophic conditions. Values of 1-2.5 indicate mesotrophic water; those of 3-5 indicate eutrophy and values of 5-20 show hyper-eutrophic conditions [20]. The compound index value was estimated to be 7.5 for Sazlidere Dam.

CONCLUSIONS

The algal flora of Sazlidere Dam did not have rich species variation except some dominant species. Phytoplankton of Sazlidere Dam consisted of centric and pennate diatoms, Chlorococcales members and blue-green algae. From Nygaard compound index value and the presence of eutrophic, mesotrophic and oligotrophic species, it was inferred that the trophic state of Sazlidere Dam lake was transitioning from oligotrophic to mesotrophic. Detailed ecological and limnological investigations should be continued at Sazlidere reservoir, which is an important source of drinking water for the Istanbul metropolitan area. The data obtained should provide a basis for determining the trophic state of the Sazlidere Dam water in order to manage and improve it.

ACKNOWLEDGMENTS

REFERENCES

1. A. K. De, "Environmental Chemistry", 5th Edn., New Age International Publishers, New Delhi, **2003**.
2. K. Fedra, "Water resources management in coastal zone: Issues of sustainability", *Eu. Water,* **2005**, *9/10*, 13-23.

3. C. S. Reynolds, "What factors influence the species composition of phytoplankton in lakes of different trophic status?", *Hydrobiol.*, **1998**, *369/370*, 11-26.

4. J. Padisak, G. Borics, I. Grigorszky and E. Soroczki-Pinter, "Use of phytoplankton assemblages for monitoring ecological status of lakes within the Water Framework Directive: The assemblage index", *Hydrobiol.*, **2006**, *553*, 1-14.

5. M. Katsiapi, M. Moustaka-Gouni, E. Michaloudi and K. A. Kormas, "Phytoplankton and water quality in a Mediterranean drinking-water reservoir (Marathonas reservoir, Greece) ", *Environ. Monit. Assess.*, **2011**, *181*, 563-575.

6. U. Yasar, I. I. Ozyigit and M.Serin, "Judas tree (Cercis siliquastrum L. subsp. siliquastrum) as a possible biomonitor for Cr, Fe and Ni in Istanbul (Turkey)", *Rom. Biotechnol. Lett.*, **2010**, *15*, 4983-4992.

7. V. Altay, I. I. Ozyigit and C. Yarici, "Urban ecological characteristics and vascular wall flora on the Anatolian side of Istanbul, Turkey", *Maejo Int. J. Sci. Technol.*, **2010**, *4*, 483-495.

8. A. Akca, "The wastewater disposal optimisation of water quality in Sazlidere basin and wetland costs evaluation" (in Turkish), *MSc Thesis*, **2005**, Yildiz Technical University, Turkey.

9. J. W. G. Lund, C. Kipling and E. D. le Cren, "The inverted microscope method of estimating algal numbers and the statiscal basis of estimations by counting", *Hydrobiol.*, **1958**, 11, 143-170.

10. T. V. Desikachary, "Cyanophyta", *PhD Thesis*, **1959**, University of Madras, India.

11. R. Patrick and C. W. Reimer, "The Diatoms of the United States: Exclusive of Alaska and Hawaii, Vol. 1", The Academy of Natural Sciences, Philadelphia, **1966**.

12. R. Patrick and C. W. Reimer, "The Diatoms of the United States: Exclusive of Alaska and Hawaii, Vol. 2 part 1", The Academy of Natural Sciences, Philadelphia, **1975**.

13. G. Huber-Pestalozzi, "Das Phytoplankton Des Süsswassers, Systematic und Biologie, Teil 2, Hälfte 2. Diatomeen", Schweizerbart Science Publishers, Stutgard, **1975**.

14. K. Krammer and H. Lange-Bertalot, "Bacillariophyceae: Teil 3: Centrales, Fragilariaceae, Eunotiaceae, Band 2/3", Gustav Fischer Pub., Jena, **1986**.

15. F. Hustedt, "Bacillariophyta (Diatomeae) heft 10", in "Die Süsswasser-flora Mitteleuropas", (Ed. A. Pascher), Gustav Fischer Pub., Jena, **1930**.

16. F. Hustedt, "The Pennate Diatoms", Koeltz Scientific Books, Koeningstein, **1985**.

17. G. W. Prescott, "Algae of the Western Great Lakes Area", W. C. Brown Co., Dubuque (IA), **1962**.

18. G. W. Prescott, "How to Know the Freshwater Algae", W. M. C. Brown Co., Dubuque (IA), **1954**.

19. D. M. John B. A. Whitton and A. J. Brook, "The Freshwater Algal Flora of the British Isles", Cambridge University Press, Cambridge, **2002**.

20. G. Nygaard, "Hydrobiological Studies on Some Danish Ponds and Lakes, Part II: The Quotient Hypothesis and Some New of Little Known Phytoplankton Organisms", I kommission hos Munksgaard, København (Denmark), **1949**, p.17.

21. Z. Altuner, "A qualitative and quantitative study on phytoplankton from one station in Lake Tortum" (in Turkish), *Doğa Bilim Dergisi*, **1984**, *8*, 162-181.

22. M. Temel, "Phytoplankton of Lake Sapanca" (in Turkish), *Istanbul Univ. J. Fish.*, **1992**, *1*, 25-40.

23. Z. Altuner and H. Gurbuz, "An investigation on phytoplankton community of Tercan Dam lake" (in Turkish), *Turk. J. Bot.*, **1994**, *18*, 443-450.

24. H. Gurbuz and Z. Altuner, "A qualitative and quantitative study on phytoplankton community in Palandoken (Tekederesi) pond" (in Turkish), *Turk. J. Biol.*, **2000**, *24*, 13-30.

25. N. Yilmaz and G. Aykulu, "An investigation on the seasonal variation of the phytoplankton density on the surface water of Sapanca Lake, Turkey", *Pak. J. Bot.* **2010**, *42*, 1213-1224.

26. G. E. Hutchinson, "A Treatise on Limnology: Introduction to Lake Biology and the Limnoplankton", John Wiley and Sons, New York, **1967**, p.385.

27. I. S. Trifonova, "Phytoplankton composition and biomass structure in relation to trophic gradient in some temperate and subarctic lakes of north-western Russia and the Prebaltic", *Hydrobiol.*, **1998**, *369/370*, 99-108.

28. O. Obali, "Seasonal variation of phytoplankton in Mogan Lake" (in Turkish), *Doğa Bilim Dergisi*, **1984**, *8*, 91-104.

29. E. Ersanli, "An investigation on phytoplankton in Cakmak Dam (Tekkeköy-Samsun)" (in Turkish), *PhD Thesis*, **2006**, Ondokuz Mayis University, Turkey.

30. A. Gonulol, "Studies on the phytoplankton of the Bayindir Dam lake", *Commun. Fac. Sci. Univ. Ankara*, **1985**, *3*, 21-38.

31. C. S. Reynolds, V. Huszar, C. Kruk, L. Naselli-Flores and S. Melo, "Towards a Functional Classification of the Freshwater Phytoplankton", *J. Plankton Res.,* **2002**, *24*, 417-428.

32. S. Kilinc and E. R. Sivaci, "A study on diatom flora in the past and present of two alkaline lakes" (in Turkish), *Turk. J. Bot.*, **2001**, *25*, 373-378.

33. B. Tas and A. Gonulol, "Planktonic algae of Derbent Dam lake (Samsun, Türkiye)" (in Turkish), *J. Fish. Sci. Commun.*, **2007**, *1*, 111-123.

34. S. Cirik and S. Gokpinar, "Plankton Knowledge and Culture" (in Turkish), Ege University Faculty of Fisheries Publications, Izmir (Turkey), **2006**, p.161.

35. J. Padisak and C. S. Reynolds, "Selection of phytoplankton associations in Lake Balaton, Hungary, in response to eutrophication and restoration measures, with special reference to the canoprokaryotes", *Hydrobiol.*, **1998**, *384*, 41-53.

36. D. S. Rawson, "Algal indicators of trophic lake types" *Limnol. Oceanogr.*, 1**956**, *1*, 18-25.

Review

Chemical constituents and biological activities of *Garcinia cowa* Roxb.

Thunwadee Ritthiwigrom [1,*], Surat Laphookhieo[2] and Stephen G. Pyne [3]

[1]Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand

[2]Natural Products Research Laboratory, School of Science, Mae Fah Luang University, Chiang Rai 57100, Thailand

[3]School of Chemistry, University of Wollongong, Wollongong, New South Wales 2522, Australia

* Corresponding author, e-mail: thunwadee.r@cmu.ac.th

Abstract: *Garcinia cowa* is an abundant source of bioactive phytochemicals. Phytochemical investigations of the plant parts indicated that the fruit, twig and stem are the best source of secondary metabolites, providing flavonoids, phloroglucinols and xanthones respectively. Seventy-eight of these compounds have been identified from the plant and several have interesting pharmacological activities.

Keywords: *Garcinia cowa*, flavonoids, xanthones, phloroglucinols

---

INTRODUCTION

Many pharmaceutical drug discoveries originated from traditional folk medicine and its associated plant materials and bioactive secondary metabolites. The Genus *Garcinia*, belonging to the Family Clusiaceae which comprises about 300 species, have been widely investigated in terms of their bioactive ingredients. Native to Asia, Africa, South America and Polynesia, the plants are small to medium sized evergreen trees which may grow up to 30 m in height and are widely distributed in the tropical and temperate regions of the world [1]. Twenty-nine species have been observed in Thailand, with 20, 13, 12, 7, 6 and 3 species found in the south, middle, north, east, north-east and west of the country respectively (Figure 1).

Garcinia is a rich source of secondary metabolites, especially triterpenes [1], flavonoids [5], xanthones [6] and phloroglucinols [7]. The latter two groups are well recognised as cheomotaxonomic markers for this genus [8a-e]. Many of the isolated compounds have a wide

range of pharmacological activities including anticancer, anti-inflammatory, antibacterial, antiviral, antifungal, anti-HIV, antidepressant and antioxidant [1, 9-13].

Figure 1. Species of *Garcinia* found in different parts of Thailand [2-4]

Garcinia cowa (Figure 2a), commonly known as Cha-muang in Thai, is widely distributed throughout Malaysia, Thailand and Myanmar. The fruits and young leaves are edible with a sour taste. The bark is dark brown with a yellow latex (Figure 2b). The plant has unisex flowers: yellow orange female flowers found at the end of branches and male flowers found along the branches as clusters (Figure 2c). The leaves are glossy, deep green, oblong and up to 6-15 cm in length and 2.5-6.0 cm in width (Figure 2d). The fruits are globose (2.5-6.0 cm in size),

green when young and dull orange or yellow at maturity with 5-8 shallow grooves, at least near the top, and contain 6-8 large 3-angled seeds (Figure 2e) [14].
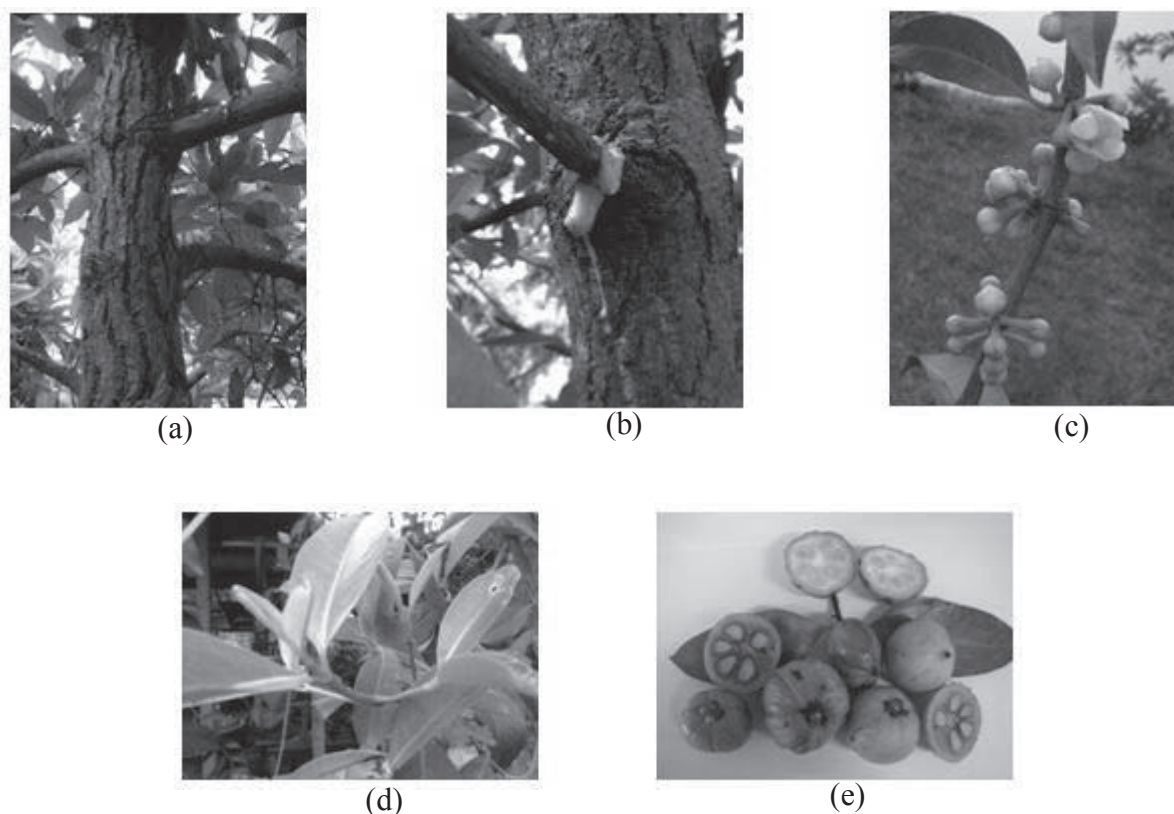
Figure 2. Parts of *G. cowa*: (a) branch, (b) bark and latex, (c) inflorescences, (d) leaves and (e) ripe and immature fruits (Photos taken by S. Laphookhieo, 2011)

Many parts of *G. cowa* have been used in traditional folk medicine. For example, the bark, latex and root have been used as an antifever agent [15, 16] while the fruit and leaves have been used for indigestion and improvement of blood circulation, and as an expectorant [16]. The chemical composition and biological activities of various parts of *G. cowa* have been investigated. The major compounds found were xanthones and phloroglucinols. However, minor compounds, including depsidones, terpenoids, steroids and flavonoids, were also observed. Currently, 78 compounds have been isolated from the twig [17], stem [18], fruit [19, 20] and latex [15]. This review mainly focuses on the chemical structures and biological activities of the phytochemicals isolated from *G. cowa* and covers the literature up to April 2012.

DISTRIBUTION AND BIOLOGICAL ACTIVITY

The biological activities of the extracts from various parts of *G. cowa* have been investigated, including the hexane and chloroform extracts of the fruit rind and methanol extract of the leaves and twigs [21-23].The hexane and chloroform extracts from the fruit rind of *G. cowa* were tested against four Gram-positive bacteria (*Bacillus cereus, B. coagulans, B. subtilis* and *Staphylococcus aureus*) and one Gram-negative bacterium (*Escherichia coli*). Both extracts significantly inhibited bacterial growth of the Gram-positive bacteria (IC$_{50}$s 15-30 µg/mL) but not *E. coli* (IC$_{50}$s 250-500 µg/mL) [21]. The extracts were also found to inhibit the growth of

Aspergillus flavus ATCC 46283, a common fungal food contaminant which produces aflatoxin B_1. The degree of inhibition of aflatoxin B_1 production (100% at a concentration of 2000 ppm) was found to be much higher than the inhibition of fungal growth (*ca* 40-60% at the same concentration) [22]. The methanol extracts of the leaves and twigs of *G. cowa* were evaluated for their ability to inhibit low-density lipoprotein peroxidation induced by copper ions. The twig extract had an IC_{50} value of 20.5 μg/mL and was more potent (higher % inhibition at 1000 μg/mL) than the leaf extract (IC_{50} not measured). The twig extract was more potent than the leaf extract on platelet aggregation of human whole blood induced by arachidonic acid, adenosine diphosphate and collagen. These activities may be due to the total phenolic content of these extracts, which were 19 and 61 mg of gallic acid equivalent per g of extract for the leaf and twig extracts respectively [23]. The structural types, chemical structures and biological activities of the natural products isolated from different parts of *G. cowa* were summarised in Table 1. A summary of the number of natural product compounds first discovered in each structural class is shown in Figure 3 and from different parts of the plant is shown in Figure 4.

Table 1. Structure, distribution and biological activity of phytochemicals isolated from different parts of *G. cowa*
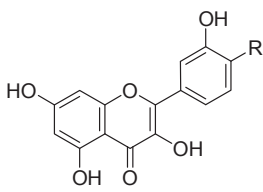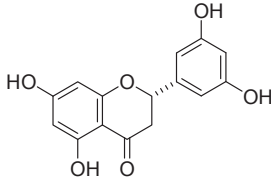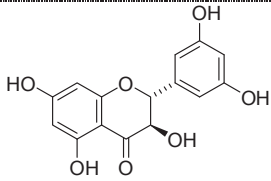
Category and structure	Name (code), source and reference	Biological activity
Depsidone		
	Cowadepsidone (**1**); twig [17]*	Cytotoxicity [17] NCI-H187 IC_{50} 31.47 μg/mL MFC-7 IC_{50} 36.03 μg/mL
Flavonoid		
	R = H Kaempferol (**2**); branch [24] R = OH Quercetin (**3**); stem [18]	
	2-(3,5-Dihydroxyphenyl)-2,3-dihy-dro-5,7-dihydroxy-3′,5,5′,7-tetrahydroxyflavanone (**4**); stem [18]	
	2-(3,5-Dihydroxy-phenyl)-2,3-dihy-dro-3,5,7-trihydroxy-(2*R*,3*R*)-3,3′,5,5′,7-pentahydroxy flavanone (**5**); stem [18]	

Table 1. (continued)
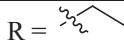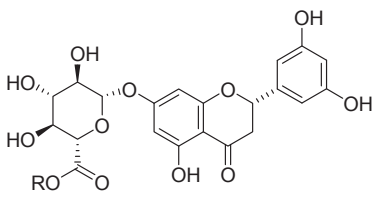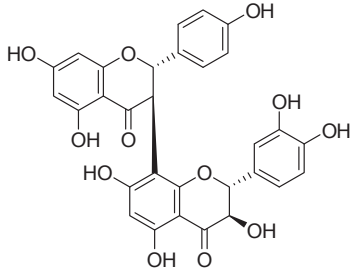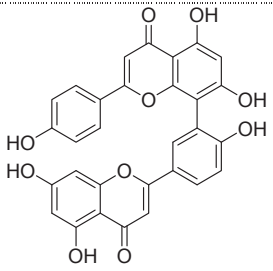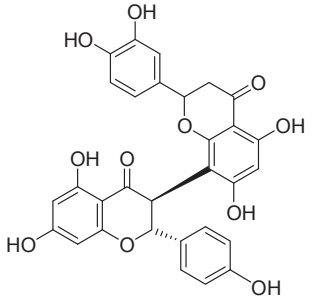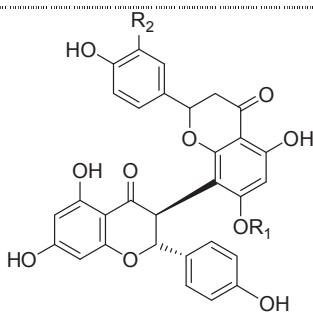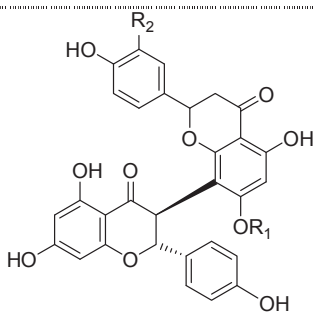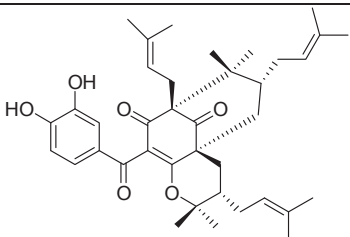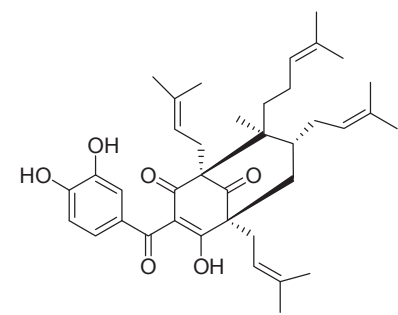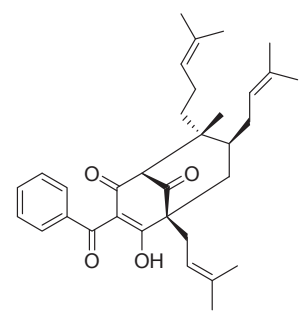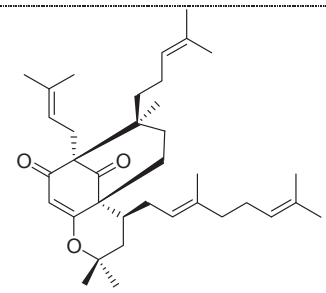
Category and structure	Name (code), source and reference	Biological activity
Flavonoid		

R = 〰

Garccowaside A (**6**); stem [18]*

R = 〰

Garccowaside B (**7**); stem [18]*

R = CH₃

Garccowaside C (**8**); stem [18]*

GB-2 (**9**); branch [24]

Amentoflavone (**10**); fruit [19]

Morelloflavone (**11**); fruit [19] twig [16]

Antioxidant activity [16]
DPPH assay   IC$_{50}$ 10.01 μg/mL
Hydroxyl radical scavenging
assay   IC$_{50}$ 3.11x10^{-4} μg/mL
Superoxide anion scavenging
assay   IC$_{50}$ 1.50x10^{-4} μg/mL

R₁ = R₂ = H

Volkensiflavone (**12**); twig [16]

R₁ = β-glucoside; R₂ = OH

Morelloflavone-7′-*O*-glucoside

or fukugiside (**13**); twig [16]

Antioxidant activity [16]
DPPH assay   IC$_{50}$ 12.92 μg/mL
Hydroxyl radical scavenging
assay   IC$_{50}$ 5.31x10^{-4} μg/mL
Superoxide anion scavenging
assay   IC$_{50}$ 6.39x10^{-4}μg/mL

Table 1. (continued)

Category and structure	Name (code), source and reference	Biological activity
	Phloroglucinol	
	Cambogin (**14**); fruit [25, 26], twig [7], stem [26]	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co
	Guttiferone K (**15a**)[#]; fruit [26]*, stem [26]	
	Guttiferone K (**15b**)[#]; twig [7]	Cytotoxicity [7] against HT-29 IC$_{50}$ 3.25$\pm$0.12 µg/mL HCT-116 CCD-18Co
	Chamuangone (**16**); leaf [27]*	Antibacterial activity [27] *S. pyogenes* MIC 7.8 µg/mL *S. viridans* MIC 15.6 µg/mL *H. pylori* MIC 15.6 µg/mL *B. subtilis, Enterococcus sp.* and *S. aureus* MIC 31.2µg/mL
	Garcicowin A (**17**); twig [7]*	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co

Table 1. (continued)
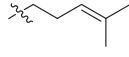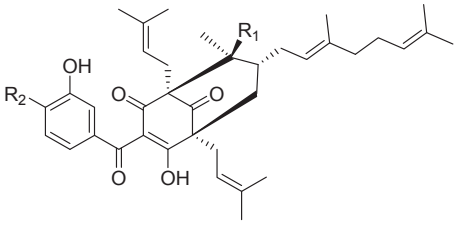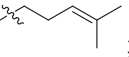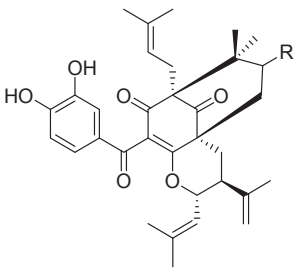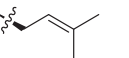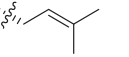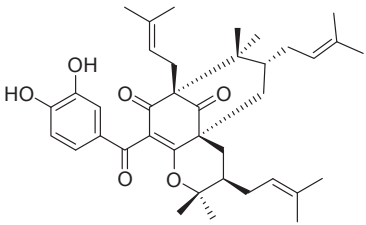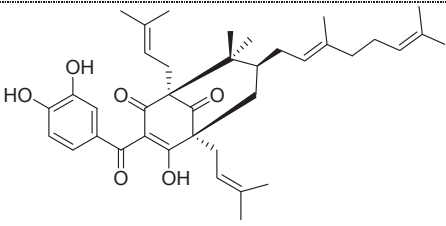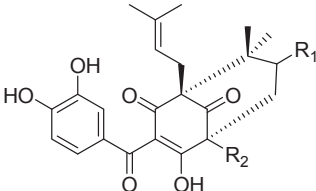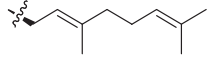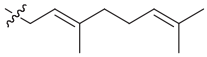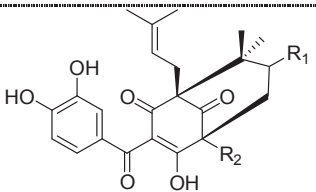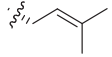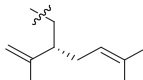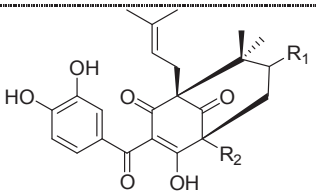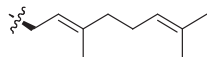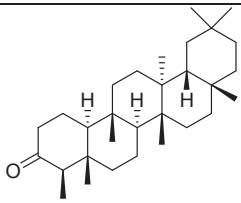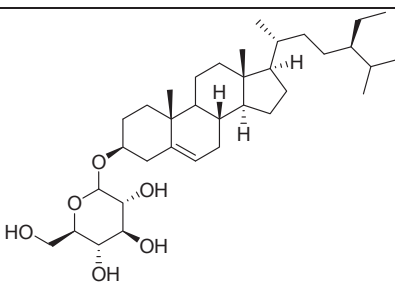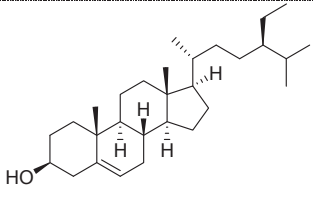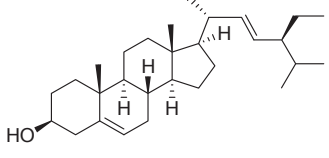
Category and structure	Name (code), source and reference	Biological activity
Phloroglucinol		
	$R_1 =$![]; $R_2 = H$ Garcicowin B (**18**); twig [7]*	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co
![structure]	$R_1 = CH_3$; $R_2 = OH$ Oblongifolin B (**19**); twig [7]	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co
	$R_1 =$![]; $R_2 = OH$ Oblongifolin C (**20**); twig [7]	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co
![structure]	$R =$![] Garcicowin C (**21**); twig [7]*	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co
	$R =$![] Garcicowin D (**22**); twig [7]*	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co
![structure]	30-Epicambogin (**23**); twig [7]	Cytotoxicity [7] against HT-29 IC_{50} 3.07$\pm$0.06 µg/mL HCT-116 CCD-18Co
![structure]	Oblongifolin A (**24**); twig [7]	

Table 1. (continued)

Category and structure	Name (code), source and reference	Biological activity
Phloroglucinol		
	$R_1 =$ $R_2 =$ Guttiferone B (**25**); twig [7]	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co
	$R_1 =$ $R_2 =$ Guttiferone F (**26**); twig [7]	
	$R_1 = R_2 =$ Oblongifolin D (**27**); twig [7]	Cytotoxicity [7] against HT-29 HCT-116 CCD-18Co
Terpene and Steroid		
	Friedelin (**28**); branch [24	
	Daucosterol (**29**); branch [24], fruit [19]	
	β-Sitosterol (**30**); branch [24], fruit [19]	
	Stigmasterol (**31**); branch [24]	

Table 1. (continued)
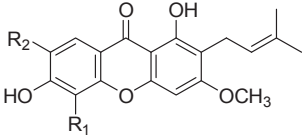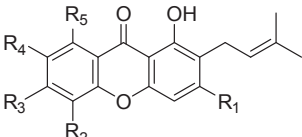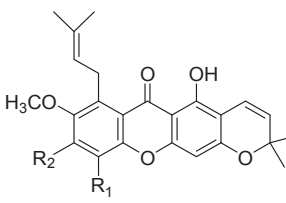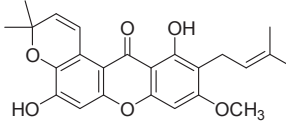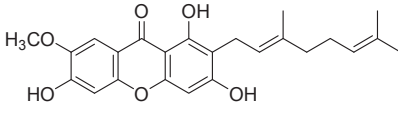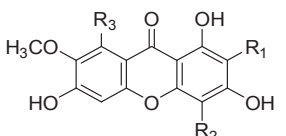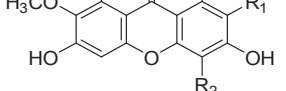
Category and structure	Name (code), source and reference	Biological activity
Xanthone		

Category and structure	Name (code), source and reference	Biological activity
(structure with R1 = OCH3 top right, OH, O, HO, R2)	R_1 = OCH$_3$; R_2 = H Cowaxanthone A (**32**); fruit [20] or cowagarcinone C; latex [15]*	Anti-inflammatory activity [16]
	R_1 = H; R_2 = OCH$_3$ 1,6-Dihydroxy-3,7-dimethoxy-2-(3-methyl-2-butenyl)xanthone (**33**); fruit [20]	
(structure with R5, O, OH, R4, R3, O, R2, R1)	R_1 = OH; R_2 = H; R_3 = R_4 = OCH$_3$; R_5 = prenyl Cowaxanthone B (**34**); fruit [20]*	Antibacterial activity [20] *S. aureus* MIC 128 µg/mL MRSA MIC 128 µg/mL Anti-inflammatory activity [16]
	R_1 = R_3 = R_4 = OCH$_3$; R_2 = H; R_5 = prenyl Fuscaxanthone C (**35**); fruit [20]	
	R_1 = R_3= OH; R_2 = R_5 = prenyl; R_4 = OCH$_3$ 7-*O*-Methylgarcinone E (**36**); stem [28]*, bark [29], fruit [20]	Antibacterial activity [20] *S. aureus* MIC 128 µg/mL MRSA MIC 64 µg/mL Antimalarial activity [29] *Plasmodium falciparum* IC$_{50}$ 1.5-3.0 µg/mL Antimalarial activity [29] *Plasmodium falciparum* IC$_{50}$ 1.5-3.0 µg/mL
	R_1 = R_3 = OH; R_2 = H; R_4 = OCH$_3$; R_5 = prenyl α-Mangostin (**37**); fruits [20], bark [29]	Antibacterial activity [20] *S.aureus* MIC 8 µg/mL MRSA MIC 8 µg/mL Anti-inflammatory activity [16] Antimalarial activity [29] *Plasmodium falciparum* IC$_{50}$ 1.5-3.0 µg/mL
	R_1 = R_4 = OCH$_3$; R_2 = H; R_3 = OH; R_5 = prenyl β-Mangostin (**38**); fruit [20], twig [17]	Antibacterial activity [20] *S. aureus* MIC 128 µg/mL MRSA MIC 64 µg/mL

Table 1. (continued)

Category and structure	Name (code), source and reference	Biological activity
Xanthone		

	R₁ = prenyl; R₂ = OH Cowaxanthone C (**39**); fruit [20]*	Antibacterial activity [20]
		S. aureus MIC 128 µg/mL
		MRSA MIC 128 µg/mL
		Anti-inflammatory activity [16]
	R₁ = H; R₂ = OH Mangostanin (**40**); fruit [20]	Antibacterial activity [20]
		S. aureus MIC 4 µg/mL
		MRSA MIC 4 µg/mL
		Anti-inflammatory activity [16]
	R₁ = H; R₂ = OCH₃ 6-*O*-Methylmangostanin (**41**); fruit [20]	

Cowaxanthone D (**42**); fruit [20]*	Anti-inflammatory activity [16]

Cowaxanthone (**43**); fruit [20], latex [15, 30*], twig [17], bark [29]	Antimalarial activity [29]
	Plasmodium falciparum
	IC₅₀ 1.5-3.0 µg/mL
	Cytotoxicity [17]
	NCI-H187 IC₅₀ 3.87 µg/mL
	KB IC₅₀ 15.43 µg/mL
	MFC-7 IC₅₀ 15.45 µg/mL

R₁ = R₃ = prenyl; R₂ = CHO Cowaxanthone E (**44**); fruit [20]*	

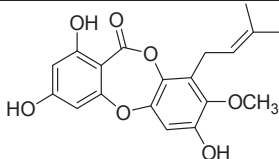| $R_1 \diagup\diagdown\diagup$ CH₂OH ; R₂ = H; R₃ = geranyl Cowanol (**45**); fruit [20], latex [15, 30*], twig [17], bark [29]	Antibacterial activity [20]
	S. aureus MIC >128 µg/mL
	MRSA MIC >128 µg/mL
	Anti-inflammatory activity [16]
	Antimalarial activity [29]
	Plasmodium falciparum
	IC₅₀ 1.5-3.0 µg/mL
	Cytotoxicity [17]
	NCI-H187 IC₅₀ 37.26 µg/mL
	KB IC₅₀ 32.34 µg/mL
	MFC-7 IC₅₀ 34.62 µg/mL
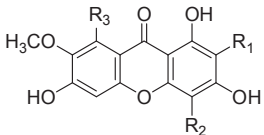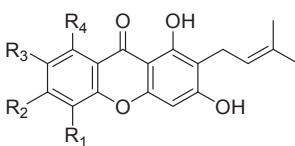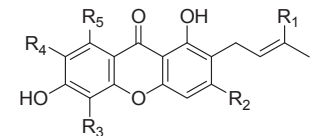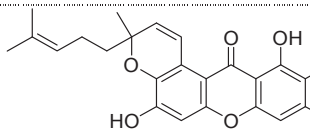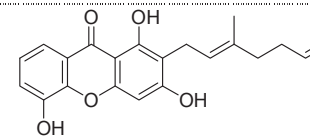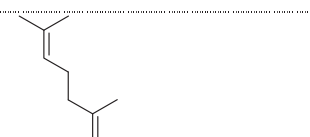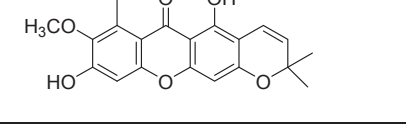
Table 1. (continued)

Category and structure	Name (code), source and reference	Biological activity
Xanthone		

	R_1 = prenyl; R_2 = H; R_3 = geranyl Cowanin (**46**); fruit [20], latex [15, 30*], twig [17], bark [29]	Antibacterial activity [20] *S .aureus* MIC >128 μg/mL MRSA MIC >128 μg/mL Anti-inflammatory activity [16] Antimalarial activity [29] *Plasmodium falciparum* IC_{50} 1.5-3.0 μg/mL
	R_1 = prenyl; R_2 = OH; R_3 = OCH_3; R_4= H 1,3,6-Trihydroxy-7-methoxy-2,5- bis(3-methyl-2-butenyl)xanthone (**47**); latex [15, 30*]	
	R_1 = H; R_2 = R_3 = OH; R_4 = geranyl Norcowanin (**48**); latex [30]*, twig [17]	Cytotoxicity [17] NCI-H187 IC_{50} 5.92 μg/mL KB IC_{50} 6.43 μg/mL MFC-7 IC_{50} 18.85 μg/mL
	R_1 = CH_3; R_2 = OH; R_3 = prenyl; R_4 = OCH_3; R_5 = geranyl Cowagarcinone A (**49**); latex [15]*	
	R_1 = CH_3; R_2 = R_3 = OCH_3; R_4 = R_5 = H Cowagarcinone B (**50**); latex [15]*	
	R_1 = CH_2OAc; R_2 = OH; R_3 = H; R_4 = OCH_3; R_5 = geranyl Cowagarcinone E (**51**); latex [15]*	
	Cowagarcinone D (**52**); latex [15]*	
	Mangostinone (**53**); latex [15]	
	Fuscaxanthone A (**54**); latex [15]	

Table 1. (continued)
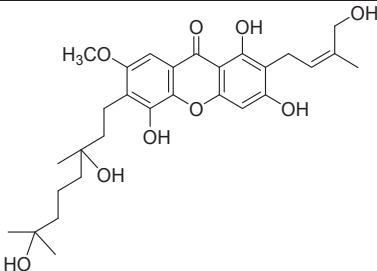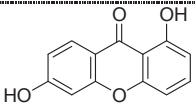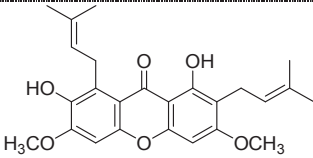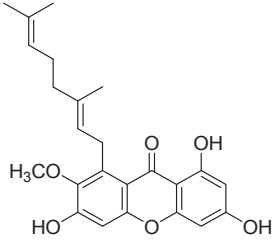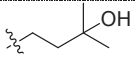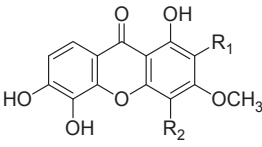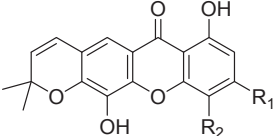
Category and structure	Name (code), source and reference	Biological activity
Xanthone		

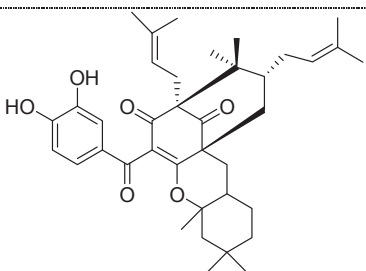|		
	Cowaxanthone F (**55**); twig [16]*	
	1,6-Dihydroxyxanthone (**56**); twig [16]	
	3,6-Di-*O*-methyl-γ-mangostin (**57**); twig [17]	Cytotoxicity [17] NCI-H187 IC$_{50}$ 8.58 μg/mL KB IC$_{50}$ 6.64 μg/mL MFC-7 IC$_{50}$ 10.59 μg/mL
	Rubraxanthone (**58**); stem [31, 32]	
	R$_1$ = H; R$_2$ = 1,5,6-Trihydroxy-3-methoxy-4-(3-hydroxyl-3-methylbutyl)xanthone (**59**) or nigrolineaxanthone T; stem [18, 33*]	
	R$_1$ = H; R$_2$ = prenyl Dulxanthone A (**60**); stem [18, 33, 34]	Cytotoxicity [34] HepG2
	R$_1$ = prenyl; R$_2$ = 4-(1,1-Dimethyl-prop-2-enyl)-1,5,6-trihydroxy-3-methoxy-2-(3-methylbut-2-enyl)xanthen-9-(9*H*)-one (**61**); stem [32]*	
	R$_1$ = OCH$_3$; R$_2$ = prenyl 1,5-Dihydroxy-3-methoxy-4-(3-methylbut-2-enyl)-6′,6′-dimethyl-2*H*-pyrano(2′,3′:6,7) xanthone (**62**); stem [18, 33*]	

Table 1. (continued)
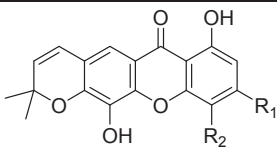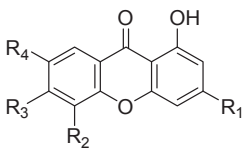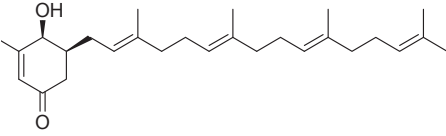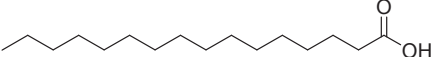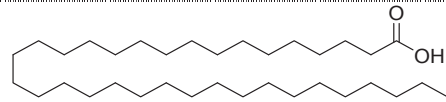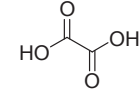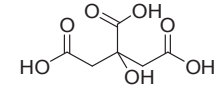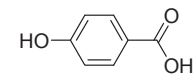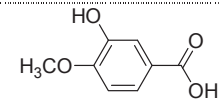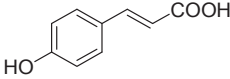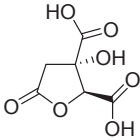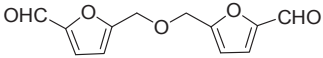
Category and structure	Name (code), source and reference	Biological activity
Xanthone		
	R_1 = OH; R_2 = H 1,3,5-Trihydroxy-6′,6′-dimethyl-2*H*-pyrano(2′,3′:6,7)xanthone (**63**); stem [18, 33]	
	R_1 = R_3 = R_4 = OH; R_2 = H 1,3,6,7-Tetrahydroxyxanthone (**64**) or norathyriol; stem [18, 33]	
	R_1 = R_2 = OH; R_3 = OCH_3; R_4 = H 1,3,5-Trihydroxy-6-methoxy-xanthone (**65**); stem [33]	
	R_1 = R_4 = OCH_3; R_2 = R_3 = OH 1,5,6-Trihydroxy-3,7-dimethoxyxanthone (**66**); stem [18, 33]	
	R_1 = R_2 = R_3 = H; R_4 = OH 1,7-Dihydroxyxanthone (**67**); stem [33]	
Miscellaneous compound		
	(2*E*,6*E*,10*E*)-(+)-4β-Hydroxy-3-methyl-5β-(3,7,11,15-tetramethyl-hexadeca-2,6,10,14-tetraenyl)cyclo-hex-2-en-1-one (**68**); stem [32]*	
	Palmitic acid (**69**); branch [24]	
	Tetratriacontanoic acid (**70**); branch [24]	
	Oxalic acid (**71**); fruit [35], fresh leaf [35], bark [35]	
	Citric acid (**72**); fruit [35], fresh leaf [35], bark [35]	
	4-Hydroxybenzoic acid (**73**); branch [24]	
	Isovanillic acid (**74**); branch [24]	

Table 1. (continued)

Category and structure	Name (code), source and reference	Biological activity
Miscellaneous compound		
	p-Coumaric acid (**75**); fruit [19]	
	(-)-Hydroxycitric acid lactone (**76**); fruit [35], fresh leaf [35], bark [35]	
	Cirsiumaldehyde (**77**); fruit [19]	

Note: Biological activity: MRSA (Methicillin resistant *Staphylococcus aureus*), NCI-H187 (Human small cell lung cancer), KB (Oral cavity cancer cell), MCF-7 (Breast cancer cell), HT-29 and HCT-116 (Human colon cancer cell), CCD-18Co (Normal human colon cell), HepG2 (Human hepatocellular liver carcinoma cell)
* Firstly discovered compound (new compound)
# Guttiferone K (**15a, 15b**) has been given two different structures in the literature [7, 26].
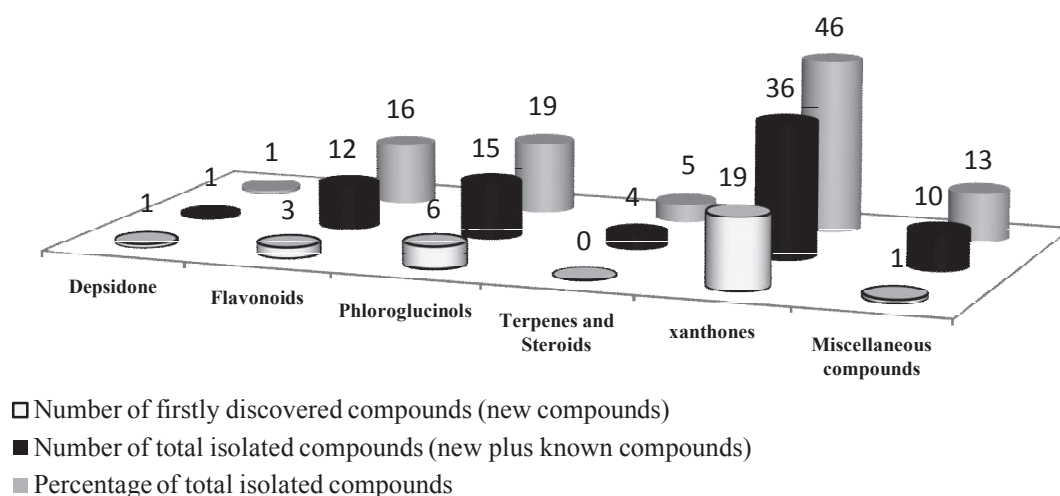
Figure 3. Classes and numbers of natural products isolated from *G. cowa*

CLASSESS OF COMPOUNDS ISOLATED FROM *G. COWA*

Depsidone

　　Depsidones comprise benzoic acid and phenol skeletons condensed at the *ortho*-positions through ester and ether linkages. This class of natural products is well known in the *Garcinia* species [36, 37]. However, cowadepsidone (**1**) was the first and only known depsidone from *G. cowa*. It was isolated from the twig extract and showed cytotoxicity against NCI-H187 and MFC-7 cancer cell lines [17].
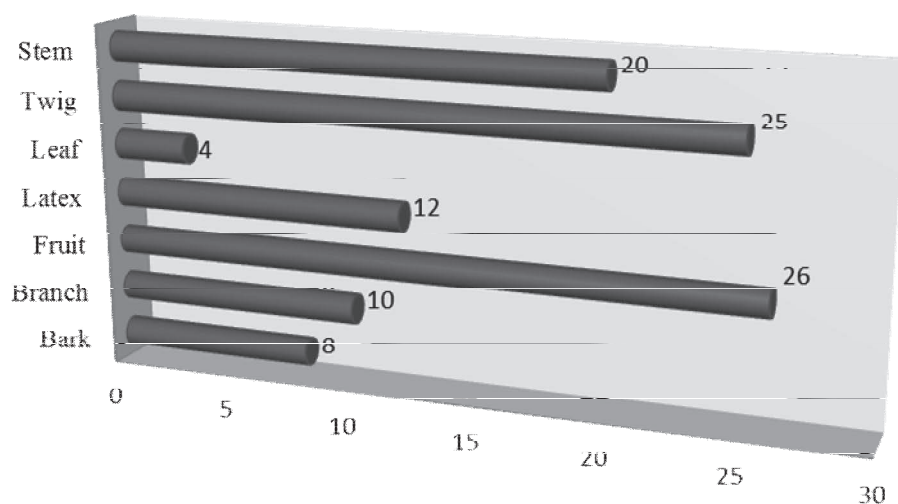
Figure 4. Numbers of natural products from different parts of *G. cowa*

Flavonoids

Twelve flavonoids (compounds **2-13** in Table 1) were isolated from *G. cowa* with garccowasides A (**6**), B (**7**) and C (**8**) being first reported as new compounds [18]. Of these compounds, only morelloflavone (**11**) and morelloflavone-7''-*O*-glucoside (**13**) showed strong antioxidant activities [16].

Phloroglucinols

Phloroglucinols are based on a phloroglucinol or 1,3,5-benzenetriol core skeleton or its 1,3,5-cyclohexanetrione (phloroglucin) tautomer. The phloroglucinols found in *G. cowa* have a benzoyl group and geranyl and polyprenyl units as substituent groups. So far, fifteen phloroglucinols (compounds **14-27** in Table 1) have been obtained from the twig including six new compounds: guttiferone K (**15a**), chamuangone (**16**), garcicowins A (**17**), B (**18**), C (**21**) and D (**22**) [7, 26, 27], and nine known phloroglucinols: cambogin (**14**), guttiferones K (**15b**), B (**25**) and F(**26**), oblongifolins B (**19**), C (**20**), A (**24**) and D (**27**), and 30-epicambogin (**23**). Some of them showed selective cytotoxicity against two cancer cell lines (HT-29 and HCT-116) and normal colon cells (CCD-18Co). Guttiferone K (**15**) and 30-epicambogin (**23**) exhibited highest cytotoxicity against cancer cell line HT-29 [7]. The name guttiferone K has been given to two different structures in the literature [7, 26] as shown in Table 1. Only one compound, chamuangone (**16**), was tested for its antibacterial activity and was found to be active against *S. pyogenes* (MIC = 7.8 μg/mL), *S. viridans* and *H. pylori* (MICs = 15.6 μg/mL), and *S. aureus, B. subtilis* and *Enterococcus* sp. (all of this bacteria shown MICs = 31.2 μg/mL) [27].

Terpenes and Steroids

Terpenes and steroids represent two large classes of natural products, although they are rare in *G. cowa*. Only four of these types of compounds (5% of the total compounds isolated) were present in *G. cowa*, viz. friedelin (**28**), daucosterol (**29**), *β*-sitosterol (**30**) and stigmasterol (**31**) [24]. None of these compounds were further studied for their biological activities. However, these compounds which were isolated from other plants had been investigated for their biological activities. Friedelin (**28**) from the root bark of *Terminalia avicennioides* exhibited antibacterial

226

activity against Bacillus Calmette-Guerin (BCG) with an MIC of 4.9 μg/mL [38]. Friedelin (**28**) and stigmasterol (**31**) isolated from the leaf of *Jatropha tanjorensis* were tested against human pathogenic microorganisms, i.e. Gram-positive bacteria: *Bacillus cereus, B. subtilis, S. aureus* and *S. epidermis*; Gram-negative bacteria: *Aeromonas hydrophila, Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, Proteus mirabilis, P. vulgaris, Salmonella paratyphi, S. paratyphi A, Vibrio alcaligenes* and *V. cholera*; and fungi: *Aspergillus fumigatus, Candida albicans, Microsporum gypseum* and *Trichophyton rubrum* using the agar-well diffusion and disk diffusion methods [39]. Friedelin (**28**), at the concentration of 2 μg/mL, showed maximum activity with 37-40, 17-40 and 31-33 mm of clear zone diameter against these three types of microorganisms respectively [39], while stigmasterol (**31**) at the same concentration exhibited maximum activity with 13-15, 8-17 and 7-8 mm of clear zone diameter respectively [39]. Daucosterol (**29**) from the roots of *Astragalus membranaceus* had no growth-inhibitory effect by direct contact but possessed immunomodulatory effect against disseminated candidiasis caused by *Candida albicans* [40]. β-Sitosterol (**30**) and stigmasterol (**31**), isolated from the bark of *Grewia tiliaefolia,* at the same concentration of 1 μg/mL showed antibacterial activity against the Gram-negative bacterium *P. aeruginosa* (ATCC-20852) with 18 and 20 mm of clear zones respectively and against *Klebsiella pneumonia* (MTCC-618) with 15 and 15 mm of clear zones respectively as determined by the agar diffusion method [41].

Xanthones

Xanthones, with two aromatic rings linked via carbonyl and ether linkages, are the major components of the *Garcinia* genus [8c-e]. They are commonly found in several parts of *G. cowa*, especially in the stem, fruit and latex. Thirty six xanthones (46% of the total isolated compounds) have been isolated and nineteen of them were first isolated from *G. cowa*. They are cowagarcinone C (**32**), cowaxanthone (**43**), cowanol (**45**), cowanin (**46**), 1,3,6-trihydroxy-7-methoxy-2,5-bis(3-methyl-2-butenyl)xanthone (**47**), norcowanin (**48**), cowagarcinones A (**49**), B (**50**), E (**51**) and D (**52**) from the latex [15, 30]; cowaxanthones B (**34**), C (**39**), D (**42**) and E (**44**) from the fruit [20]; 7-*O*-methylgarcinone E (**36**), 1,5,6-trihydroxy-3-methoxy-4-(3-hydroxyl-3-methylbutyl)xanthone (**59**), 4-(1,1-dimethyl-prop-2-enyl)-1,5,6-trihydroxy-3-methoxy-2-(3-methylbut-2-enyl)xanthen-9(9*H*)-one (**61**) and 1,5-dihydroxy-3methoxy-6′,6′-dimethyl-2*H*-pyrano(2′,3′:6,7)-4-(3-methylbut-2-enyl) xanthone (**62**) from the stem [18, 33]; and cowaxanthone F (**55**) from the twig [16]. Most of these xanthones showed interesting biological activities.

Antibacterial activity

Eight xanthones from the fruit: cowaxanthones B (**34**) and C (**39**), 7-*O*-methylgarcinone E (**36**), *α*-mangostin (**37**), *β*-mangostin (**38**), mangostanin (**40**), cowanol (**45**) and cowanin (**46**) were investigated for their antibacterial activity against *S. aureus* and MRSA. *α*-Mangostin (**37**) and mangostanin (**40**) showed significant activity against these bacteria. *α*-Mangostin (**37**) had a MIC value of 8 μg/mL against both *S. aureus* and MRSA while mangostanin (**40**) had an MIC value of 4 μg/mL against both bacteria [20].

Anti-inflammatory activity

Eight xanthones: cowaxanthones A (**32**), B (**34**), C (**39**) and D (**42**), α-mangostin (**37**), mangostanin (**40**), cowanol (**45**) and cowanin (**46**) were tested for their anti-inflammatory activity using the ethyl phenylpropiolate induced ear edema assay. All xanthones except cowanol were more active than the standard drug, phenylbutazone [16].

Antimalarial activity

Five xanthones isolated from the stem bark: 7-*O*-methylgarcinone E (**36**), α-mangostin (**37**), cowaxanthone (**43**), cowanol (**45**) and cowanin (**46**) had significant in vitro antimalarial activity against *Plasmodium falciparum* with IC_{50} values ranging between 1.5-3.0 µg/mL [29].

Anticancer activity

Six xanthones: cowaxanthone (**43**), cowanol (**45**), cowanin (**46**), norcowanin (**48**), 3,6-di-*O*-methyl-γ-mangostin (**57**) and dulxanthone A (**60**) isolated from twig were evaluated for their cytotoxicity against NCI-H187, KB, MFC-7 and/or HepG2 cell lines. Cowaxanthone (**43**), cowanin (**46**), norcowanin (**48**) and 3,6-di-*O*-methyl-γ-mangostin (**57**) exhibited significant cytotoxicity against the NCI-H187 cell line with IC_{50} values ranging between 3.87-8.58 µg/mL, and moderately inhibited KB and MCF-7 cancer cell lines with IC_{50} values ranging between 6.43-15.43 and 10.59-21.38 µg/mL respectively [17]. Dulxanthone A (**60**) was found to be cytotoxic against the HepG2 cell line [34].

Miscellaneous Compounds

Ten (13% of the total isolated compounds) of the miscellaneous class of compounds have been isolated, including a new discovery: (2*E*,6*E*,10*E*)-(+)-4β-hydroxy-3-methyl-5β-(3,7,11,15-tetramethyl-hexadeca-2,6,10,14-tetraenyl)cyclohex-2-en-1-one (**68**) [32]. None of the isolated compounds from this class were tested for their biological activities.

CONCLUSIONS

G. cowa is an important source of bioactive compounds. Among the parts of this tree, the fruit, twig and stem are the best source of metabolites, thirty of which have been isolated, i.e. one depsidone, one *α,β*-unsaturated cyclohexenone, three flavonoids, six phloroglucinols and nineteen xanthones. Some of these compounds show interesting pharmacological activities. α-Mangostin (**37**), cowanol (**45**) and cowanin (**46**) are commonly found in all parts of *G. cowa* and they can be used as chemotaxonomic markers of this species. The plant is still under investigation by our research group with the prospect of identifying new bioactive compounds in the near future.

REFERENCES

1. A. Kijjoa and L. M. M. Vieira, "Triterpenes from the plants of the Family Clusiaceae (Guttiferae): Chemistry and biological activities", in "Natural Products: Chemistry, Biochemistry and Pharmacology", (Ed. G. Brahmachari), Alpha Science International Ltd., Oxford, **2009**, Ch. 13.

2. T. Smitinand, "Thai Plant Names", The Forest Herbarium Royal Forest Department, Bangkok, **2001**, pp.246-248.

3. (a) J. F. Maxwell, "Vegetation of Doi Tung, Chiang Rai province, northern Thailand", *Maejo Int. J. Sci. Technol.*, **2007**, *1*, 10-63; (b) J. F. Maxwell, "Addendum to vegetation of Doi Tung, Chiang Rai province, northern Thailand", *Maejo Int. J. Sci. Technol.*, **2008**, *2*, 37-139.

4. S. Klaiklay, Y. Sukpondma, V. Rukachaisirikul and S. Phongpaichit, "Friedolanostanes and xanthones from the twigs of *Garcinia hombroniana*", *Phytochem.*, **2013**, *85*, 161-166.

5. W. H. Ansari, W. Rahman, D. Barraclough, R. Maynard and F. Scheinmann, "Biflavanoids and a flavanone-chromone from the leaves of *Garcinia dulcis* (Roxb.) Kurz", *J. Chem. Soc., Perkin Trans.1*, **1976**, *13*, 1458-1463.

6. V. Rukachaisirikul, T. Ritthiwigrom, A. Pinsa, P. Sawangchote and W. C. Taylor, "Xanthones from the stem bark of *Garcinia nigrolineata*", *Phytochem.*, **2003**, *64*, 1149-1156.

7. G. Xu, W. L. T. Kan, Y. Zhou, J.-Z. Song, Q.-B. Han, C.-F. Qiao, C.-H. Cho, J. A. Rudd, G. Lin and H.-X. Xu, "Cytotoxic acylphloroglucinol derivatives from the twigs of *Garcinia cowa*", *J. Nat. Prod.*, **2010**, *73*, 104-108.

8. (a) H.-D. Nguyen, B. T.-D. Trinh and L.-H. D. Nguyen, "Guttiferones Q-S, cytotoxic polyisoprenylated benzophenones from the pericarp of *Garcinia cochinchinensis*", *Phytochem. Lett.*, **2011**, *4*, 129-133; (b) S.-C. Chien, C.-F. Chyu, I.-S. Chang, H.-L. Chiu and Y.-H. Kuo, "A novel polyprenylated phloroglucinol, garcinialone, from the roots of *Garcinia multiflora*", *Tetrahedron Lett.*, **2008**, *49*, 5276-5278; (c) Q.-B. Han, N.-Y. Yang, H.-L. Tian, C.-F. Qiao, J.-Z. Song, D. C. Chang, S.-L. Chen, K. Q. Luo and H.-X. Xu, "Xanthones with growth inhibition against HeLa cells from *Garcinia xipshuanbannaensis*", *Phytochem.*, **2008**, *69,* 2187-2192; (d) L.-H. D. Nguyen, H. T. Vo, H. D. Pham, J. D. Connolly and L. J. Harrison, "Xanthones from the bark of *Garcinia merguensis*", *Phytochem.*, **2003**, *63,* 467-470; (e) L.-H. D. Nguyen and L. J. Harrison, "Xanthones and triterpenoids from the bark of *Garcinia vilersiana*", *Phytochem.*, **2000**, *53,* 111-114.

9. M. Hemshekhar, K. Sunitha, M. S. Santhosh, S. Devaraja, K. Kemparaju, B. S. Vishwanath, S. R. Niranjana and K. S. Girish, "An overview on genus *Garcinia*: Phytochemical and therapeutical aspects", *Phytochem. Rev.*, **2011**, *10*, 325-351.

10. Y. Ren, D. D. Lantvit, E. J. Carcache de Blanco, L. B. S. Kardono, S. Riswan, H. Chai, C. E. Cottrell, N. R. Farnsworth, S. M. Swanson, Y. Ding, X.-C. Li, J. P. J. Marais, D. Ferreira and A. D. Kinghorn, "Proteasome-inhibitory and cytotoxic constituents of *Garcinia lateriflora*: Absolute configuration of caged xanthones", *Tetrahedron*, **2010**, *66*, 5311-5320.

11. Y. Sukpondma, V. Rukachaisirikul and S. Phongpaichit, "Xanthone and sesquiterpene derivatives from the fruits of *Garcinia scortechinii*", *J. Nat. Prod.*, **2005**, *68*, 1010-1017.

12. H. W. Ryu, M. J. Curtis-Long, S. Jung, Y. M. Jin, J. K. Cho, Y. B. Ryu, W. S. Lee and K. H. Park, "Xanthones with neuraminidase inhibitory activity from the seedcases of *Garcinia mangostana*", *Bioorg. Med. Chem.*, **2010**, *18*, 6258-6264.

13. I. O. Pereira, M. J. Marques, A. L. R. Pavan, B. S. Codonho, C. L. Barbieri, L. A. Beijo, A. C. Doriguetto, E. C. D′Martin and M. H. dos Santos, "Leishmanicidal activity of benzophenones and extracts from *Garcinia brasiliensis* Mart. fruits", *Phytomed.*, **2010**, *17*, 339-345.

14. T. K. Lim, "Edible Medicinal and Non-Medicinal Plants, Vol. 2: Fruits", Springer Science and Business Media B.V., London, **2012**, pp.29-34.

15. W. Mahabusarakam, P. Chairerk and W. C. Taylor, "Xanthones from *Garcinia cowa* Roxb. latex", *Phytochem.*, **2005**, *66*, 1148-1153.

16. K. Panthong, N. Hutadilok-Towatana and A. Panthong, "Cowaxanthone F, a new tetraoxygenated xanthone, and other anti-inflammatory and antioxidant compounds from *Garcinia cowa*", *Can. J. Chem.*, **2009**, *87*, 1636-1640.

17. S. Cheenpracha, W. Phakhodee, T. Ritthiwigrom, U. Prawat and S. Laphookhieo, "A new depsidone from the twigs of *Garcinia cowa*", *Heterocycles*, **2011**, *83*, 1139-1144.

18. J. Shen, Z. Tian and J-s. Yang, "The constituents from the stems of *Garcinia cowa* Roxb. and their cytotoxic activities", *Pharmazie*, **2007**, *62*, 549-551.

19. J. Shen and J.-s. Yang, "Chemical constituents from fruit of *Garcinia cowa*", *Zhongguo Yaoxue Zazhi*, **2006**, *41*, 660-661.

20. K. Panthong, W. Pongcharoen, S. Phongpaichit and W. C. Taylor, "Tetraoxygenated xanthones from the fruits of *Garcinia cowa*", *Phytochem.*, **2006**, *67*, 999-1004.

21. P. S. Negi, G. K. Jayaprakasha and B. S. Jena, "Antibacterial activity of the extracts from the fruit rinds of *Garcinia cowa* and *Garcinia pedunculata* against food borne pathogens and spoilage bacteria", *LWT-Food Sci. Technol.*, **2008**, *41*, 1857-1861.

22. G. S. Joseph, G. K. Jayaprakasha, A. T. Selvi, B. S. Jena and K. K. Sakariah, "Antiaflatoxigenic and antioxidant activities of *Garcinia* extracts", *Int. J. Food Microbiol.*, **2005**, *101*, 153-160.

23. I. Jantan, F. A. Jumuddin, F. C. Saputri and K. Rahman, "Inhibitory effects of the extracts of *Garcinia* species on human low-density lipoprotein peroxidation and platelet aggregation in relation to their total phenolic contents", *J. Med. Plants Res.*, **2011**, *5*, 2699-2709.

24. J. Shen and J. Yang, "Chemical constituents of branch of *Garcinia cowa* Roxb", *Zhongcaoyao*, **2007**, *38*, 993-994.

25. G. K. Jayaprakasha, B. S. Jena, L. J. M. Rao and M. C. Varadaraj, "A process for the isolation of cambogin from *Garcinia cowa*", *Indian Pat.*, *No. 242799* (**2010**).

26. J. Shen and J.-S. Yang, "A novel benzophenone from *Garcinia cowa*", *Acta Chim. Sinica*, **2007**, *65*, 1675-1678.

27. A. Sakunpak and P. Panichayupakaranant, "Antibacterial activity of Thai edible plants against gastrointestinal pathogenic bacteria and isolation of a new broad spectrum antibacterial polyisoprenylated benzophenone, chamuangone", *Food Chem.*, **2012**, *130*, 826-831.

28. K. Lihitwitayawuid, T. Phadungcharoen, C. Mahidol and S. Ruchirawat, "7-*O*-Methylgarcinone e from *Garcinia cowa*", *Phytochem.*, **1997**, *45*, 1299-1301.

29. K. Likhitwitayawuid, T. Phadungcharoen and J. Krungkrai, "Antimalarial xanthones from *Garcinia cowa*", *Planta Med.*, **1998**, *64*, 70-72.

30. P. na Pattalung, W. Thongtheeraparp, P. Wiriyachitra and W. C. Taylor, "Xanthones of *Garcinia cowa*", *Planta Med.*, **1994**, *60*, 365-368.

31. H.-H. Lee and H.-K. Chan, "1,3,6-Trihydroxy-7-methoxy-8-(3,7-dimethyl-2,6-octadienyl) xanthone from *Garcinia cowa*", *Phytochem.*, **1977**, *16*, 2038-2040.

32. F. S. Wahyuni, L. T. Byrne, Dachriyanus, R. Dianita, J. Jubahar, N. H. Lajis and M. V. Sargent, "A new ring-reduced tetraprenyltoluquinone and a prenylated xanthone from *Garcinia cowa*", *Aust. J. Chem.*, **2004**, *57*, 223-226.

33. J. Shen and J.-S. Yang, "Two new xanthones from the stems of *Garcinia cowa*", *Chem. Pharm. Bull.*, **2006**, *54*, 126-128.

34. Z. Tian, J. Shen, A. P. Moseman, Q. Yang, J. Yang, P. Xiao, E. Wu and I. S. Kohane, "Dulxanthone A induces cell cycle arrest and apoptosis *via* up-regulation of p53 through mitochondrial pathway in HepG2 cells", *Int. J. Cancer*, **2008**, *122*, 31-38.

35. B. S. Jena, G. K. Jayaprakasha and K. K. Sakariah, "Organic acids from leaves, fruits, and rinds of *Garcinia cowa*", *J. Agric. Food Chem.*, **2002**, *50*, 3431-3434.

36. V. Rukachaisirikul, W. Naklue, S. Phongpaichit, N. H. Towatana and K. Maneenoon, "Phloroglucinols, depsidones and xanthones from the twigs of *Garcinia parvifolia*", *Tetrahedron*, **2006**, *62*, 8578-8585.

37. L. D. Ha, P. E. Hansen, F. Duus, H. D. Pham and L.-H. D. Nguyen, "Oliveridepsidones A-D, antioxidant depsidones from *Garcinia oliveri*", *Magn. Reson. Chem.*, **2012**, *50*, 242-245.

38. A. Mann, K. Ibrahim, A. O. Oyewale, J. O. Amupitan, M. O. Fatope and J. I. Okogun, "Antimycobacterial friedelane-terpenoid from the root bark of *Terminalia avicennioides*", *Amer. J. Chem.*, **2011**, *1*, 52-55.

39. M. B. Viswanathan, J. D. J. Ananthi and P. S. Kumar, "Antimicrobial activity of bioactive compounds and leaf extracts in *Jatropha tanjorensis*", *Fitoterapia*, **2012**, *83*, 1153-1159.

40. J.-H. Lee, J. Y. Lee, J. H. Park, H. S. Jung, J. S. Kim, S. S. Kang, Y. S. Kim and Y. Han, "Immunoregulatory activity by daucosterol, a β-sitosterol glycoside, induces protective Th1 immune response against disseminated Candidiasis in mice", *Vaccine*, **2007**, *25*, 3834-3840.

41. B. M. K. Ahamed, V. Krishna, H. B. Gowdru, H. Rajanaika, H. M. Kumaraswamy, S. Rajshekarappa, C. J. Dandin and K. M. Mahadevan, "Isolation of bactericidal constituents from the stem bark extract of *Grewia tiliaefolia* Vahl", *Res. J. Med. Plant*, **2007**, *1*, 72-82.

Full Paper

Integration of recommender system for Web cache management

Supawadee Hiranpongsin and Pattarasinee Bhattarakosol *

Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

* Corresponding author, e-mail: pattarasinee.b@chula.ac.th

_____

Abstract: Web caching is widely recognised as an effective technique that improves the quality of service over the Internet, such as reduction of user latency and network bandwidth usage. However, this method has limitations due to hardware and management policies of caches. The Behaviour-Based Cache Management Model (BBCMM) is therefore proposed as an alternative caching architecture model with the integration of a recommender system. This architecture is a cache grouping mechanism where browsing characteristics are applied to improve the performance of the Internet services. The results indicate that the byte hit rate of the new architecture increases by more than 18% and the delay measurement drops by more than 56%. In addition, a theoretical comparison between the proposed model and the traditional cooperative caching models shows a performance improvement of the proposed model in the cache system.

_____

INTRODUCTION

Due to the ease of use and the high volume of information that is available over the Internet via Web-based systems, the growth rate of Web usage is rapidly enlarging, seemingly with no end in sight. As a result, the quality of service (QoS) of every Internet service provider (ISP) is highly affected and various strategies have been employed to maintain their service levels. Unfortunately, none of these strategies can completely satisfy their customers. Thus, long delays and/or low throughput rates may occur.

One important metric of the QoS is performance. Different variables are applied to measure this indicator, such as average delay, average hit rate and average byte hit rate. The

values of these metrics are influenced by different factors such as limitation of hardware and unqualified service management systems. As a consequence, termination of the browsing transactions may occur [1-2] and performance drops.

In order to avoid this problem, many techniques have been proposed and applied to the management of the cache. Originally, the cache replacement algorithm was proposed for managing a single cache [3-5]. Currently, cache management is based on the cache farming architecture wherein many caches are combined to work together. As a rule, there are two different management models for cache farming: a hierarchical model [6] and a distributed model [7]. The hierarchical caching model consists of intermediate caches located in different levels of the network. Thus, each retrieved Web is copied to the intermediate caches along the traversal path. On the other hand, the distributed caching model allows the search to be performed on the distributed caches over the network.

As a consequence of the above methods, high duplication of objects in hierarchical caching and high bandwidth usage in distributed caching occur [8]. Moreover, when the ISP needs to increase their performance to satisfy their customers, the expanding of hardware and an adjustment of cache management policy always occur. These modifications cause system complexity for administrators and lead to the increasing of management expenses that is not appreciated by organisational managers.

Based on the problems mentioned above, a cache management mechanism that considers types of transaction has been proposed, including the user behaviour being counted as a management factor when managing the cache [9]. This paper proposes a new architecture of cache farming that integrates the concept of a recommender system where all arising behaviours are considered in managing the incoming transactions so the desired retrieval time and bandwidth can be maintained. Consequently, the proposed solution can be seen as an economical system because the number of caches need not expand when the enhancement of performance is desired [10-11]. Moreover, system complexity can be avoided and the cache management policy does not need to be altered.

RELATED WORK

Proxy server is the common solution to serving Web users for an organisation. The number of users and the resultant Web usage over the Internet has rapidly grown over the years. Thus, a long delay may occur for every browse since the amount of browsed content cannot be served or stored in the proxy's cache. To work around this, various researchers have proposed new caching algorithms to manage contents in the caches. In late 2005, Frequency Recency and Size Cache Replacement (FRES-CAR) [12], which considered the document reference recency, frequency and size was proposed. In addition, Bian and Chen introduced the Least Grade Replacement (LGR) [13] that brought the perfect history into account for Web cache optimisation. Other replacement policies exist, such as Semantic and Least Recently Used (SEMALRU) [14], which ignores the objects that are less related to an incoming object or least recently used; and Dump and Clear [15], which was developed to solve the locality problem based on Petersen Graph topology. In addition, many new cache management policies developed to perform in the mobile environment have been recently proposed, for example Distributed Alternative Binding Cache mechanism (DABC) [16] and Cache Replacement Policy Based on Neighbour Nodes' Condition (CRBNC) [17]. The DABC was presented to reduce the problem of limited cache size and also to eliminate the rate of registration signalling. The CRBNC, focusing

on increasing neighbour hits, used the cached data of neighbour nodes as an effective criterion in replacement.

Due to the fact that there are multiple caches with different sizes to be managed, two fundamental cache models have been proposed and implemented; these are the hierarchical and the distributed structure of a Web cache farm. With respect to the hierarchical structure, Foygel and Strelow [18] applied a prefetching algorithm to the hierarchical Web caches. Moreover, the Internet Small Computer System Interface (iSCSI) protocol [19] to communicate between a lower-level and its higher-level proxy server was proposed with a new Web caching scheme. In 2005, the Hierarchical Web Caching Placement and Replacement (HCPR) algorithm was drawn up [20]. This algorithm placed the most frequently referenced documents close to users in the leaf nodes of the hierarchy. Then the Content Distribution Network (CDN) architecture was applied to the hierarchical Web caching technology by Yang and Chi [21]. Additionally, the evaluation of cooperative caching was studied and used as background information for a better understanding of the current CDNs [22].

The cache content placement in emerging scenarios, such as Internet Protocol Television (IPTV) services, has drawn attention in several studies. For example, Borst et al. [23] developed the cooperative cache management algorithms that aimed to minimise the bandwidth costs. In the same year, Applegate et al. [24] created collaborative caching as a global optimisation solution. Later, Dai et al. [25] investigated the capacity provisioning problem in hierarchical caching networks based on a real-world IPTV system. In order to address this problem, an efficient collaborative caching mechanism with dynamic request routing for massive content distribution was proposed. This proposed mechanism achieved better results compared to conventional cache cooperation with static routing schemes.

Originally, the searches on the hierarchical caching model were usually reserved for depth-first search (DFS) exploration. Thus, a general framework of hierarchical caches that can be used by breadth-first search (BFS) was proposed [26]. In order to evaluate the performance of web caching, performance metrics were captured using Hit Rate (HR) and Byte Hit Rate (BHR) [27-28]. Additionally, a cost function model was implemented based on both performance metrics to investigate the suitability of applied policies over the two-level hierarchical cache model [29]. The results indicated that the performance obtained was higher when the lower-level cache used the Least Frequently Used (LFU) or Least Recently Used (LRU) and the upper-level cache used the Greedy Dual-Size (GDS).

In contrast to the hierarchical caching model, in the distributed caching architecture all browsing requests must be distributed over the network. Thus, Summary Cache protocol [30] was proposed to exchange the messages among caches in order to find documents in other caches. Unfortunately, the investigation [31] pointed out that this model caused congestion over the network. Therefore, the techniques of web caching and web prefetching were implemented in the proxy system to solve the problems of server load and congestion control [32].

After the implementation of the hierarchical caching and distributed caching architectures, a hybrid architecture that is the combination of both models has been proposed. One interesting hybrid model was proposed by Baek et al. [33], in which the reference table was employed at each level in the hierarchy except at the lowest level. Another solution in this model focused on some significant factors that were related to the QoS, such as communication between the caches and cache contents [34].

Since the development of Recommender System (RS) in the area of information retrieval is widely implemented in the information search algorithm to shorten the search time, the METIOREW system has been implemented for the Web search mechanism [35]. This system applies the document evaluation results from users as the 'intelligent bookmark' to finding the most relevant Web pages under the given search topics. In contrast, the Amazon.com system uses Item-to-Item Collaborative Filtering [36] as its recommender system. Similar to Amazon.com, the RS is widely implemented in many Internet activities and services such as course management systems [37] and e-learning systems [38-39]. Furthermore, a modification of the RS has been applied on some product or service systems to improve the efficiency of the RS, for example the RS implementation with fuzzy scatter difference [40]. Additionally, time context and group preferences were used to improve the customer profile in collaborative systems [41]. In 2011, Pukkhem and Vatanawood [42] employed learning styles and a word analysis technique to create a learning object recommendation model which provided learners with personalised learning object selection service. The results of experiments demonstrated that the accuracy resulted in high student satisfaction.

Although various caching management mechanisms are continuously proposed, the limitation of the QoS has not been eliminated because the retrieval process must rely on hardware and cache management policies. Thus, every ISP always extends the volume of its service caches for certain periods of time. This management method causes system complexity and increases the cost of management. In this paper the Behaviour-Based Cache Management Model (BBCMM) is therefore implemented to manipulate the cache farm in an economical way. In the following section, details of this mechanism are discussed.

ANALYSIS OF BROWSING BEHAVIOUR

Since the BBCMM is implemented based on the browsing behaviour of users under the service organisation, the first step in applying this model is to identify types of browsing characteristics of the organisation. The browsing methods can be identified by considering all existing transactions that pass through the cache farm in which a record is stored in the proxy log.

In this research samples of retrieval behaviour were collected from Chulalongkorn University with its many faculties and research activities, both in pure sciences and social sciences, and national and international subjects. As a consequence, there are a large number of categories of transactions during Internet access. In addition, under the Internet management process of the university, every browsing query must run through the proxy system and some part of this query is stored in the squid log file referred to as access.log. The data in the access.log file, consisting of 1.5 TB of compressed files, was sampled from January to August 2009 for this study. After uncompressing the files, the number of transactions was approximately 50 million per day.

The format of the access.log allows many fields to be stored, such as the retrieval time or the timestamp, the file size, the destination URL and the access situation (hit/miss). However, in this research only two fields are used: the destination URL and the file size. Since the types of website can be classified as static and dynamic [43], the URLs of all dynamic websites consider only the first part of the URL's name. For example, there is a CNN URL that presents the news of girl brides being abducted as a fabled HIV cure (http://thecnnfreedomproject.blogs.cnn.com/ 2012/05/27/girl-brides-abducted-as-fabled-hiv-cure/?hpt=hp_c2). This address will be truncated and considered only as thecnnfreedomproject.blogs.cnn.com. All these URLs will be counted for

frequency usage during the time of the study. This method is called frequency-based analysis. The result of this method is the 'Frequency Pattern (FP)', which shows the popularity of websites in a certain period of time.

It is a fact that congestion over the communication channel occurs because of a large volume of transferred information. Thus, values in the file size field of the access.log are analysed to find the load pattern of the retrieved URLs that is the real load of the cache. This load pattern is called 'Loading Size Pattern (LSP)' and this value can be applied as a weighting value in the architecture designing process of the cache farm.

After considering the existing values of the FP, three different groups can be identified. The first group, the low frequency (LF), is the group of URLs that are browsed less than three times per week. This group represents approximately 58% of all URLs. Most of these websites are sites where their pages were obtained from various server locations such as facebook.com or live.com. In addition, websites with frequency of only one or two times per week are authorised websites. The second group, the medium frequency (MF), accounts for 17% of the total URLs and are retrieved from four to eight times per week. The last group, the high frequency (HF), is the remaining 25% of the total URLs and they are retrieved more than eight times per week.

As a consequence of the LSP-based analysis, the Web size can be grouped into three specific ranges. The first range, the normal size (NS), accounts for 69% of all websites with the file size of 0-4,000 bytes. This NS can be further categorised into two sub-ranges as small size (SS) and medium size (MS). The SS is the group for the browsed files with sizes less than 1,100 bytes while the MS is the group for the browsed files with sizes of 1,100-4,000 bytes. The second range, the large size (LS), includes over 30% of all websites with the file size of 4,000 bytes - 10 Mbytes. The third range, the extra size (ES), is less than 1% of all websites and includes all files that are larger than 10 Mbytes. The ES range is independently grouped as another group due to the large size of the websites even though the number of sites browsed is small. Table 1 shows the grouping for Web cache based on the FP and the LSP.

Table 1. Grouping for Web cache

Type	Grouping criteria	Loading ratio (%)
1	LF with SS, MF with LS	25.11
2	LF with MS, MF with NS	23.37
3	HF with all size ranges except ES	25.62
4	All frequency groups with ES	25.89

Referring to Table 1, the probability that a retrieved URL will be a member of a group is approximately equal to 0.25 for every group. These criteria are employed to set up the proxy caches in the BBCMM. Nevertheless, groups of websites are altered when the pattern of browsed websites is changed. Consequently, the system configuration will occur in the classification process. This group classification must be performed on a fixed schedule, e.g. weekly or monthly, depending on the business objective and the policy of each organisation. Details of the BBCMM are presented in the following section.

PROPOSED ARCHITECTURE

Owing to the limitation of proxy caches of the ISPs and the existing policy of cache management, the number of caches must be increased when the number of users is expanded. Consequently, the complexity of cache management increases. Thus, in this investigation a new

architecture, BBCMM, is proposed to eliminate this weakness. As can be seen from Figure 1, the BBCMM consists of different groups of specific proxy server (SPS) whose main part is the proxy manager (PM) that works with the data in the Web profile database system (WPDB). Following is the description of each part of the BBCMM.
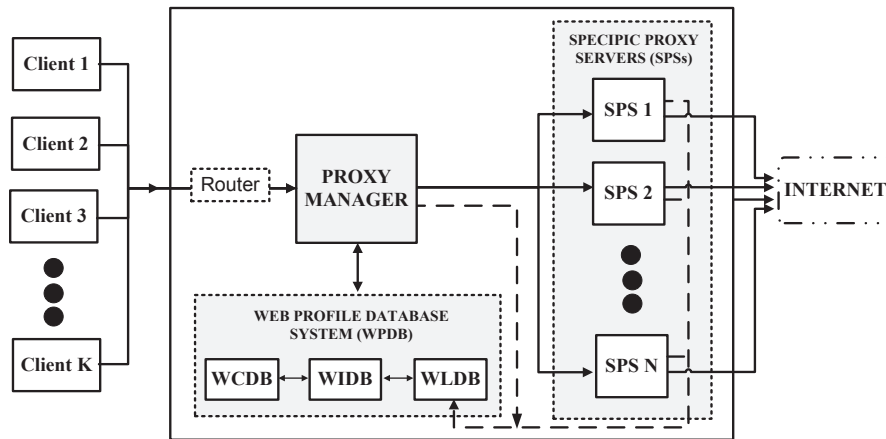
Figure 1. Behaviour-Based Cache Management Model (BBCMM)

Web Profile Database System

The WPDB consists of three sub-databases: a Web classification database (WCDB), a web identification database (WIDB) and a log database (WLDB), as illustrated in Figure 1. These sub-databases are implemented for Web classification when users call for websites. Of these three databases, the WCDB and the WIDB are used in the Web classification process when retrieving URLs. The WCDB contains a boundary value of both the browse frequency (bv_F) and the Web size (bv_S). These data are calculated and used to identify the group whenever a request by a user arrives at the PM. The WIDB is the database that stores the information for each Web, e.g. host name of web, browse frequency and size of web. The information stored in WIDB is used with both boundary values stored in the WCDB to identify the pattern of browsed websites by an automatic classification module (ACM) in the PM. Figure 2 shows the classification of the requested Web that is sent to the corresponding group or the SPS using the two boundary values, bv_F and bv_S, as indicators. Each group has bv_F and bv_S, which are denoted by $bv_F(i)$ and $bv_S(i)$ respectively, for the ith classified group. The $bv_F(i)$ and the $bv_S(i)$ are defined in equations 1 and 2 respectively.
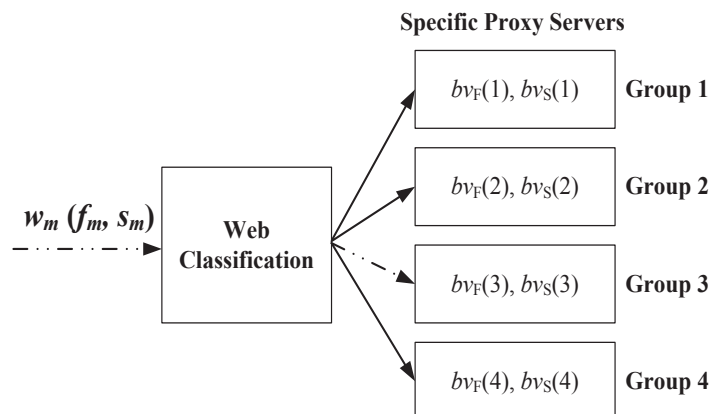
Figure 2. Web classification with boundary values

Given $F_l(i)$, $F_u(i)$, $S_l(i)$, $S_u(i) \in \Re$, and $F_l(i) < F_u(i)$ and $S_l(i) < S_u(i)$, let $F_l(i) = avg_F(i) -$ $std_F(i)$ and $F_u(i) = avg_F(i) + std_F(i)$, where $avg_F(i) \in \Re$ is the average of the browse frequency of group (i) and $std_F(i) \in \Re$ is the standard deviation of the browse frequency of group (i). In addition, $S_l(i) = avg_S(i) - std_S(i)$ and $S_u(i) = avg_S(i) + std_S(i)$, where $avg_S(i) \in \Re$ is the average of the Web size of group (i) and $std_S(i) \in \Re$ is the standard deviation of the Web size of group (i). Then,

$$bv_F(i) = [F_l(i), F_u(i)] = \{x \in \Re \mid F_l(i) \leq x \leq F_u(i)\} \qquad (1)$$
$$bv_S(i) = [S_l(i), S_u(i)] = \{y \in \Re \mid S_l(i) \leq y \leq S_u(i)\} \qquad (2)$$

Let m be the set of different requested websites, named $\{w_1, w_2,\ldots,w_m\}$; each w_m has the browse frequency denoted by f_m and the Web size denoted by s_m.

Since f_m, $s_m \in \Re$, and $f_m \in bv_F(i) \leftrightarrow F_l(i) \leq f_m \leq F_u(i)$ and $s_m \in bv_S(i) \leftrightarrow S_l(i) \leq s_m \leq S_u(i)$, thus, if $f_m \in bv_F(i)$ and $s_m \in bv_S(i)$, then the w_m is classified to the group (i). However, the data management and maintenance are recognised to handle data overloading in the database. For this reason, if the classified w_m is inactive for a specific time it will be removed from the database.

Consider the situation when a request is issued from a client. This request will be recorded into the WLDB so the system can be recovered when a failure or an unexpected event arises. The records in the WLDB are not only obtained from the retrieved command from the ACM, but also from the retrieved commands recorded in the local log database of each SPS, or caches. The transferring of data from all local log databases from each SPS is in the offline mode and is performed only before the RAM starts its evaluation. Nevertheless, all data of the WPDB system are necessary for the PM processes that are described in the following section.

Proxy Manager

The PM is an assigned proxy server in the cache farm, responsible for classifying websites into groups depending on the browsing patterns mentioned previously. This system consists of six modules: Record Analyser Module (RAM), Automatic Classification Module (ACM), Gateway-like Module (GM), Internet Communication Module (ICM), Squid Cache Module (SCM) and WPDB Communication Module (WPDB CM). One significant function of the system is to adapt the recommender system in the transaction identification when the transaction is indefinable in the normal classification process. Figure 3 demonstrates the PM's architecture in performing its tasks. The details of these modules are explained below.
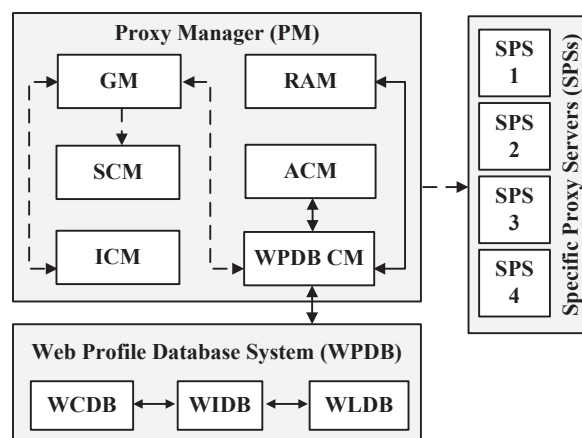
Figure 3. Proxy manager's architecture

Record Analyser Module

The RAM is the only module that works individually for pattern classification. This module deals with all URLs stored in the WLDB. The outcomes of the RAM computing are new values of browse frequencies and Web sizes that must be updated to the WIDB at the end of every day. These values are used to justify all unknown websites received from the ACM. Additionally, the re-analysing of all boundary values, $bv_F(i)$ and $bv_S(i)$, for every group (i) will be performed every weekend because websites are constantly inserted, updated and deleted. These new values will overwrite the previous values in the WCDB. However, these new boundaries may not be suitable since the number of websites in the existing groups under the newly defined indexes might break the load balancing policy of the network. Thus, websites with their frequencies and sizes that are close to the upper or lower bounds of their neighbours are changed to their neighbours' value until the number of websites in each group is significantly equivalent. Then the new boundary values, $bv_F(i)$ and $bv_S(i)$, for every group (i) will be recalculated and stored in the WCDB. As a result of the RAM process, the size of available caches for each group when passing the process of the ACM may change to obtain high performance and maintain the load balance of each group properly.

Automatic Classification Module

After all boundaries are calculated by the RAM, the ACM will use them for Web grouping. As mentioned in the RAM section, the pattern analysis of Web browsing will be performed every weekend. Thus, the process of the ACM, like the RAM, will automatically run every weekend. The group categorisation of the ACM is performed based on the popularity system called the recommender system. Since this system has two different approaches, the content-based (CB) approach and the collaborative-filtering (CF) approach [44], the ACM integrates both of them to gain the highest efficiency in the Web categorisation process.

Based on the calculation values obtained from the RAM process, the CF and the CB are applied to sort websites into suitable groups. Since there are various methods of the CF classification process, the user-based collaborative filtering technique is applied to identify websites in the WIDB. The real content of each website is unidentified; thus, in this investigation the browsing frequency is used as the indicator for users' preference based on the CF rule. Moreover, the file size is used as the Web's characteristic based on the CB concept.

Since the values in the WIDB contain the real browse frequency and average file size of retrieved websites during a week, every browse frequency and the average file size of each site will be compared with $bv_F(i)$ and $bv_S(i)$ in the WCDB for grouping. There are two possible situations in this comparison. The first is the normal case in which websites satisfy the condition of group (i) in the WCDB, $f_m \in bv_F(i)$ and $s_m \in bv_S(i)$; those websites are classified as websites of group (i). The second situation occurs when there is no suitable condition in the WCDB for websites in the WIDB, either $f_m \notin bv_F(i)$ or $s_m \notin bv_S(i)$ or neither of them. These websites will be assigned to the group with highest frequency under the assumption that the more popular the request is, the more chance occurs to meet the requirement.

Gateway-like Module

The GM classifies all arriving transactions from the Internet so they will be sent to suitable proxy servers. The situation of a transaction can either be classifiable or unclassifiable. If the requested transaction is classifiable, then the request will be sent via the ICM to the classified SPS that is linked to the PM system as shown in Figure 1. Otherwise, the request will

be sent to the SCM to retrieve the required information. The GM can classify the request using information in the read-only mode of the WIDB.

Internet Communication Module

The ICM sends a message from the PM to the SPSs according to the classification result of the ACM. The connection between this module and an SPS is connection-oriented, using Transmission Control Protocol/Internet Protocol (TCP/IP) to guarantee the delivery of data. The format of the sent message is the same as that of the packet received from the GM. The communication type of the ICM to any SPS is the simplex method because every SPS will return the requested page directly to the users.

Squid Cache Module

A cache module in the PM, this module is to serve all unidentified websites. The load of this module therefore depends on the existence of the number of undefined URLs.

WPDB Communication Module

Every process of the PM must use data from the WPDB system. This system consists of three sub-databases: WLDB, WIDB and WCDB. Thus, the WPDB CM creates a connection to the WPDB system to retrieve the required data. The connection performed by this module is connection-oriented using TCP/IP to prevent the occurrence of network congestion and reduce maintenance costs due to transmission loss.

Specific Proxy Server

The SPS is a group of caches defined based on the WIDB contents. Since there are many types of defined boundaries in the WCDB that are related to websites in the WIDB, there are many installed SPSs in the PM. However, the size of each SPS depends on the weekly calculation of the RAM. As a consequence, the size of each SPS is dynamic in order to maintain the performance of each server which affects the performance of the entire system.

EVALUATION METHOD

In this section, the implementation of the evaluation system environment is performed on HP-2 Quad Cores with XEON Processors and 16-GB main memory running Ubuntu 10.04 desktop. The database management system employs MySQL server version 5.1, and the phpMyAdmin running on Apache2 Web server is used to deal with MySQL server. All mechanisms maintaining the system are developed using Java.

In order to demonstrate that the proposed caching mechanism, the BBCMM, performs better than the single caching mechanism, a comparison between both systems is performed. Although the comparison is based on single caching mechanism, this mechanism is a general method that is implemented for the actual use in every organisation. The performance estimation of the BBCMM and the existing caching mechanism (EM) is based on the results of the trace-driven simulation implemented in a virtual machine environment. The outcomes of performance evaluation between the BBCMM and the EM are therefore trustworthy and acceptable. The details of simulation experiments are explained as follows.

Based on the defined groups in the previous section, the WCDB contains four pairs of boundaries in accordance with the four groups of websites in the WIDB. Thus, there are four SPSs in the simulation system and one SCM for undefined websites as mentioned in the GM section. Each SPS and the EM employ the Least Recently Used (LRU) algorithm as their cache

replacement policy. As a result, the similarity in individual cache management can be directly compared without bias or any external influence. The data for this evaluation is based on the real Internet usage of students and staff of a small-size government university in Thailand during October 2010. The simulation data are approximately 18 million records. Figure 4 illustrates the testing environment for the BBCMM.
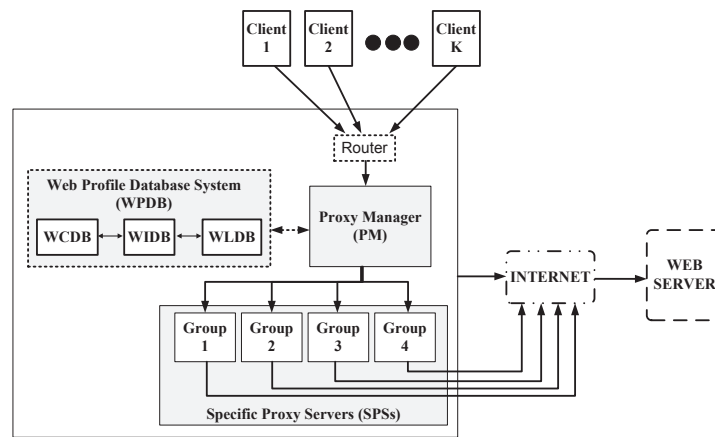
Figure 4. Simulation system environment for the BBCMM

SIMULATION RESULTS

To indicate the performance of the BBCMM and the EM, the performance metrics in this research constitute hit rates, miss rates, byte hit rates, byte miss rates and response times of hits and misses. These metrics are measured from running data in the simulation system obtained from the university. The results of this run are presented in Table 2.

Table 2. Result overview of simulation experiments

Metric	Implemented model		Difference (%)
	BBCMM	**EM**	
Hit rate (%)	41.82	39.53	+5.80
Miss rate (%)	58.18	60.47	-3.79
Byte hit rate (%)	43.39	36.76	+18.05
Byte miss rate (%)	56.61	63.24	-10.49
Average response time of hit (ms.)	85.44	196.20	-56.45
Average response time of miss (ms.)	2842.12	2431.25	+16.90

Referring to the results in Table 2, the hit rate of the BBCMM is 5.8% higher than the EM while the miss rate is also reduced. This can primarily be interpreted that the new architecture and caching mechanism are serving users better than the original concept. In addition, the byte hit rate shows an 18% increase from the EM and the byte miss rate is also reduced by approximately 10.5%. The average response time of the hit mode is more than 56% lower than the EM. Unfortunately, the simulation result shows that if websites are unidentified, users must wait for their requests longer than in the normal situation. The daily performance of the BBCMM can be drawn as in Figure 5, where every significant metric is measured within 30 days.
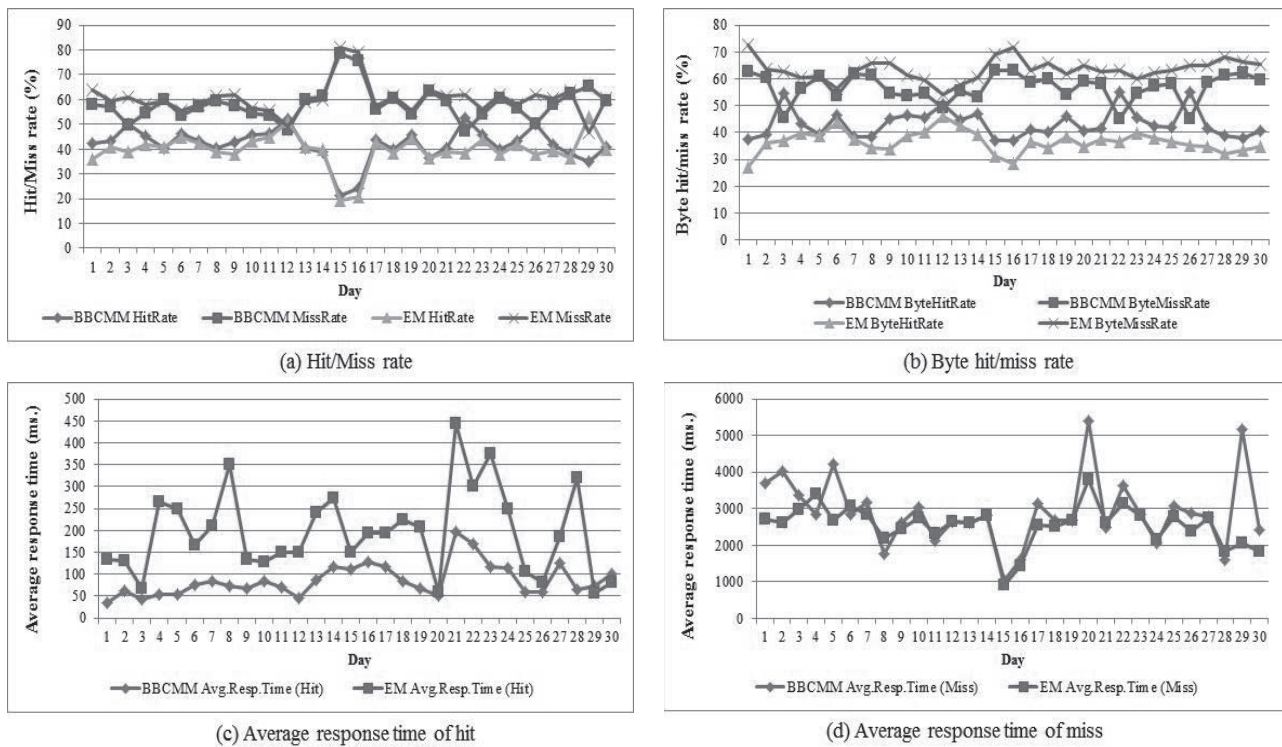
Figure 5. Daily performance comparison between BBCMM and EM

Referring to Figure 5(a), in the beginning of the measurement process the hit rate of the BBCMM is slightly higher than that of the EM. Nevertheless, after 15 days of this experiment the number of hits based on the BBCMM has the potential to be much better. This is because the variety of websites in each cache is small when starting the BBCMM, but all requested websites based on the defined patterns will be accrued to the system daily until most of the requests are available in the cache. However, if retrieved URLs have no pattern as expected, the daily hit rates of both the BBCMM and the EM will be dropped.

Since the hit rate is increased under BBCMM system, the simulation also illustrates that the number of byte hits is increased as shown in Figure 5(b). This is the result of the grouping mechanism of the cache because websites with similar characteristics are stored and manipulated in the same area. Thus, the possibility of the required website being found in the managed caches is higher compared to when different groups of websites are stored in the same area as the EM.

With respect to finding the requested web in the caches, the BBCMM is faster than the EM and the response time to obtain the required information is shorter for the BBCMM as shown in Figure 5(c). However, the retrieving information from the external sources of the BBCMM architecture is greater than the retrieval time of the EM. This is a result of the overhead required to search the non-existent content in each cache as shown in Figure 5(d).

The test results confirm that the cache management policy of BBCMM is highly efficient and QoS can be achieved. This is the result of categorising and grouping websites based on the available retrieval pattern; the classified websites are then stored in the individual caches based on the defined groups. In order to provide more insights into the proposed BBCMM, a comparison with models other than EM is performed. A theoretical comparison with two

common cooperative caching models, namely hierarchical model (HM) and distributed model (DM), is discussed in the following section.

THEORETICAL COMPARISON

The World Wide Web can be considered as a universal warehouse of web pages and links that offer large amounts of data for the Internet users. Thus, no one can imagine the number of users accessing that data. As a result, it is a major challenge for the proposed BBCMM, the HM and the DM to manipulate retrieved transactions. In fact, the main challenge and design principle in the cooperative caching architecture is to rapidly locate requested cached websites. To this end, this research focuses on two significant issues: the number of hops and the processing time for both hit and miss cases. The number of hops for traversing multiple levels in the cache system should be optimised by efficient locating and accessing of data, while the processing time handling each request may be reduced to shorten the response time.

The cache system is assumed to contain k clients and m web pages. A request for a specific web page originates from one of the clients. In addition, the client will not send another request before obtaining the previously requested page. Therefore, the maximum number of requests in the system is k. For a fair comparison, the HM, DM and BBCMM contain n nodes or caches and each web-caching model uses the LRU for page replacement.

Number of Hops

In all types of web caching model when the cache receives a requested transaction, it first checks the local cache for the requested page. If the requested page has a current copy of the object in the cache—the hit state, it simply returns the page to the user. Thus, the time complexity of the number of hops for both the HM and the DM is obviously $O(1)$ since they can return the page obtained from the local cache. With the results of classification and grouping mechanism of the BBCMM, after GM of the PM classifies the retrieved page to a suitable cache, it must send the transaction to the SPS or the SCM as mentioned previously. Therefore, the time complexity of the number of hops for the BBCMM is $O(1)$. As a result, the time complexity of the number of hops for all three caching models for the hit case is a constant value.

On the other hand, if the search in the cache returns a miss state, the required object will be sent to other cooperative caches and also to the original web server if the request cannot be found in the cache system. Since all caches that are arranged in a tree-like structure are compulsory objects for the HM, the time complexity of the number of hops in the miss state is $O(\log n)$, where n is the number of nodes. In contrast to the HM, the DM has no intermediate caches and its connection is performed among n peer cooperative caches. So when the miss state occurs, the content search will be performed on other *n-1* distributed caches. For this reason, the time complexity of the number of hops for the DM is $O(n)$. For the BBCMM, each SPS deals with the specific contents that are cached. If the request transaction is not satisfied by the SPS, this transaction will be fetched from the original web server without accessing other SPSs. Thus, the time complexity of the number of hops for the BBCMM when a miss occurs is $O(1)$.

Therefore, the proposed BBCMM can minimise the number of hops in the cache system on a miss case under the defined assumptions. Moreover, the number of hops can be maintained on a hit case when the BBCMM is implemented.

Processing Time

The processing time is measured from the time that the request is sent until the time when the user receives the response. Given that t_u is the processing time between the user and the local cache for the request, t_c the processing time between cooperative caches for the request, and t_o the processing time between the cache and the original web server for the request. Let t_{PM} be the processing time between the PM and the SPSs or the SCM inside the PM. There are two situations to consider: hit and miss.

Under the hit state, both the HM and the DM contain the requested page in their own caches. Thus, it is clear that the time complexity of the processing time for both of them is $O(t_u)$, where t_u is a constant value. In the case of BBCMM, the time complexity of the processing time is $O(t_u + t_{PM})$, where t_u and t_{PM} are constant values. As a consequence, with BBCMM, the transaction issued from the user will be sent first to the PM for the web page classification. Then the classified web page will be forwarded to the identified SPS or handled by the SCM. So two processing times, t_u and t_{PM}, are included in the elapsed time of the BBCMM.

The other situation, the miss state, occurs when the issued request cannot be satisfied by the local cache. Consequently, the search will be performed over cooperative caches for the HM and the DM. In order to search for the object among siblings, the total processing time t_c for the HM is $t_c \log(n)$ and that for the DM is $t_c n$. However, there is no processing time t_c for the BBCMM. Thus, the processing times for handling each transaction of the HM, the DM and the BBCMM are $t_u + t_c \log(n) + t_o$, $t_u + t_c n + t_o$ and $t_u + t_{PM} + t_o$ respectively, t_u, t_c, t_{PM} and t_o being constant values and n the number of caches. Consequently, the time complexities of the processing time for the request that is served by the caches in HM, DM and BBCMM are $O(\log n)$, $O(n)$ and $O(T)$ respectively, where n is a number of caches and T is a constant value.

In summary, the proposed BBCMM reduces the response time with shorter processing time on both of hit and miss states when comparing with the traditional caching models, the HM and the DM.

CONCLUSIONS

Retrieving information from any website within an organisation requires proxy and cache management systems to filter suitable information flowing in and out the organisation's networks. Thus, accessing performance depends on the cache management mechanism of the proxy server and the cache farm. Even though various techniques have been proposed and implemented to increase the service quality, there are still difficulties in maintaining the service quality desired as the number of transactions expands. The proposed BBCMM integrates the concepts of the recommender system, using frequencies and browsed file sizes, in order to maintain the quality of service, which is the desired response time for the Internet users. The components of the BBCMM are PM, WPDB and SPSs. Within the PM, there is an important module known as the ACM, which deals with the classifying of all retrieved websites.

This experiment has demonstrated that under a systematic organisation with a clear objective of browses, most of the important metrics have been changed positively. This can be accomplished while neither adding nor adjusting hardware or implementing changes in the cache management policy. These positive changes are, among others, the reduction of the response time of the BBCMM by more than 56% while the hit rate and the byte hit rate of the BBCMM are increased by more than 5% and 18% respectively. The comparison results measured by the time complexity of the number of hops and the processing time have shown that the proposed

BBCMM can improve the caching performance in the value of the time constant. The BBCMM can thus be an eco-proxy system that provides high service performance for an organisation.

REFERENCES

1. R. P. Wooster and M. Abrams, "Proxy caching that estimates page load delays", *Comp. Netw. ISDN Syst.*, **1997**, *29*, 977-986.

2. J. Mardesich, "The Web is no shopper's paradise", *Fortune*, **1999**, *140*, 188-198.

3. P. Cao and S. Irani, "Cost-aware www proxy caching algorithms", Proceedings of USENIX Symposium on Internet Technology and Systems, **1997**, Monterey, USA, pp.193-206.

4. S. Jin and A. Bestavros, "GreedyDual* Web caching algorithm: Exploiting the two sources of temporal locality in Web request streams", *Comp. Commun.*, **2001**, *24*, 174-183.

5. B. R. Haverkort, R. E. A. Khayari and R. Sadre, "A class-based least-recently used caching algorithm for world-wide web proxies", *Lect. Notes Comp. Sci.*, **2003**, *2794*, 273-290.

6. A. Chankhunthod, P. B. Danzig, C. Neerdaels, M. F. Schwartz and K. J. Worrell, "A hierarchical internet object cache", Proceedings of USENIX Annual Technical Conference, **1996**, San Diego, USA, pp. 153-164.

7. R. Tewari, M. Dahlin, H. M. Vin and J. S. Kay, "Beyond hierarchies: Design considerations for distributed caching on the Internet", Proceedings of the 19th IEEE International Conference on Distributed Computing Systems, **1999**, Austin, USA.

8. P. Rodriguez, C. Spanner and E. W. Biersack, "Analysis of Web caching architectures: Hierarchical and distributed caching", *IEEE/ACM Trans. Netw.*, **2001**, *9*, 404-418.

9. P. Bhattarakosol and V. Ngamaramvaranggul, "An Internet web management policy for government organization", Proceedings of Network Research Workshop in 18th APAN Meetings, **2004**, Cairns, Australia, pp.249-255.

10. W. Srisujjalertwaja and P. Bhattarakosol, "Customer-oriented policy for proxy management system", Proceedings of International Computer Symposium, **2004**, Taipei, Taiwan, pp.1168-1173.

11. S. Hiranpongsin and P. Bhattarakosol, "Intelligent caching algorithm for Web cache farming system", Proceedings of International Conference on Wireless Information Networks and Business Information System, **2009**, Kathmandu, Nepal.

12. G. Pallis, A. Vakali and E. Sidiropoulos, "FRES-CAR: An adaptive cache replacement policy", Proceedings of International Workshop on Challenges in Web Information Retrieval and Integration, **2005,** Tokyo, Japan, pp.74-81.

13. N. Bian and H. Chen, "A least grade page replacement algorithm for Web cache optimization", Proceedings of the First International Workshop on Knowledge Discovery and Data Mining, **2008**, Adelaide, Australia, pp.469-472.

14. K. Geetha, N. A. Gounden and S. Monikandan, "SEMALRU: An implementation of modified Web cache replacement algorithm", Proceedings of World Congress on Nature and Biologically Inspired Computing, **2009**, Coimbatore, India, pp.1406-1410.

15. B. Zhang and H. Wu, "A new distributed caching replacement strategy", Proceedings of 3rd IEEE International Conference on Communication Software and Networks, **2011**, Xi'an, China, pp.167-170.

16. S. S. Hasan, R. Hassan and F. E. Abdalla, "A new binding cache management policy for NEMO and MIPV6", *J. Theoret. Appl. Inform. Technol.*, **2012**, *36*, 113-117.

17. A. S. Ghasemi and A. M. Rahmani, "A new cache replacement policy based on neighbor nodes' condition in mobile environments", *J. Comp.*, **2012**, *4*, 164-169.

18. D. Foygel and D. Strelow, "Reducing Web latency with hierarchical cache-based prefetching", Proceedings of International Workshop on Scalable Web Services, **2000**, Washington, DC, USA, pp.103-108.

19. H. Lim and D. H. C. Du, "Design considerations for hierarchical Web proxy server using iSCSI", Proceedings of Symposium on Applications and the Internet, **2003**, Orlando, USA, pp.414-417.

20. W. Li, K. Wu, X. Ping, Y. Tao, S. Lu and D. Chen, "Coordinated placement and replacement for grid-based hierarchical Web caches", *Lect. Notes Comp. Sci.*, **2005**, *3795*, 430-435.

21. F. H. Yang and C. H. Chi, "Using hierarchical scheme and caching techniques for content distribution networks", Proceedings of 3rd International Conference on Semantics, Knowledge and Grid, **2007**, Xi'an, China, pp.535-538.

22. J. Zhang, "A literature survey of cooperative caching in content distribution networks", **2012**, arxiv.org/pdf/1210.0071 (Accessed: October, 2012).

23. S. Borst, V. Gupta and A. Walid, "Distributed caching algorithms for content distribution networks", Proceedings of IEEE INFOCOM, **2010**, San Diego, USA, pp.1-9.

24. D. Applegate, A. Archer, V. Gopalakrishnan, S. Lee and K. K. Ramakrishnan, "Optimal content placement for a large-scale VoD system", Proceedings of 6th International Conference on Emerging Networking Experiments and Technologies, **2010**, Philadelphia, USA.

25. J. Dai, Z. Hu, B. Li, J. Liu and B. Li, "Collaborative hierarchical caching with dynamic request routing for massive content distribution", Proceedings of IEEE INFOCOM, **2012**, Orlando, USA, pp.2444-2452.

26. R. Mateescu and A. Wijs, "Hierarchical adaptive state space caching based on level sampling", *Lect. Notes Comp. Sci.*, **2009**, *5505*, 215-229.

27. M. Busari and C. Williamson, "ProWGen: A synthetic workload generation tool for simulation evaluation of web proxy caches", *Comp. Netw.*, **2002**, *38*, 779-794.

28. L. Shi and Y. Zhang, "Optimal model of Web caching", Proceedings of 4th International Conference on Natural Computation, **2008**, Jinan, China, pp.362-366.

29. L. Shi, P. Yao, L. Wei and Y. Tao, "Cost-benefit analysis of the Web hierarchy caching model", *Inform. Technol. J.*, **2012**, *11*, 364-367.

30. L. Fan, P. Cao, J. Almeida and A. Z. Broder, "Summary cache: A scalable wide-area Web cache sharing protocol", *IEEE/ACM Trans. Netw.*, **2000**, *8*, 281-293.

31. M. Piatek, "Distributed Web proxy caching in a local network environment", **2004**, www.acm.org/src/subpages/papers/piatek.src.2004.pdf (Accessed: March, 2005).

32. C. V. Manikandan, P. Manimozhi, B. Suganyadevi, K. Radhika and M. Asha, "Efficient load reduction and congession control in Internet through multilevel border gateway proxy caching", Proceedings of IEEE International Conference on Computational Intelligence and Computing Research, **2010**, Coimbatore, India, pp.1-4.

33. J. Baek, G. Kaur and J. Yang, "A new hybrid architecture for cooperative Web caching", *J. Ubiq. Converg. Technol.*, **2008**, *2*, 1-11.

34. M. S. E. Oneis, H. Barada and M. J. Zemerly, "Towards an efficient Web caching hybrid architecture", Proceedings of 4th International Conference on Information Technology, **2009**, Amman, Jordan.

35. D. Bueno, R. Conejo and A. A. David, "METIOREW: An objective oriented content based and collaborative recommending system", *Lect. Notes Comp. Sci.*, **2002**, *2266*, 310-314.

36. G. Linden, B. Smith and J. York, "Amazon.com recommendations: Item-to-item collaborative Filtering", *IEEE Internet Comp.*, **2003**, *7*, 76-80.

37. J. Itmazi and M. Megías, "Using recommendation systems in course management systems to recommend learning objects", *Int. Arab J. Inform. Technol.*, **2008**, *5*, 234-240.

38. J. Bobadilla, F. Serradilla, A. Hernando and MovieLens, "Collaborative filtering adapted to recommender systems of e-learning", *Knowl.-Based Syst.*, **2009**, *22*, 261-265.

39. K. Souali, A. E. Afia, R. Faizi and R. Chiheb, "A new recommender system for e-learning environments", Proceedings of International Conference on Multimedia Computing and Systems, **2011**, Ouarzazate, Morocco, pp.1-4.

40. Y. Gong and Q. Xue, "Study on internet recommendation system of collaborative filtering based on scatter difference", Proceedings of International Conference on Computer, Mechatronics, Control and Electronic Engineering, **2010**, Changchun, China, pp.160-163.

41. M. Julashokri, M. Fathian, M. R. Gholamian and A. Mehrbod, "Improving recommender system's efficiency using time context and group preferences", *Adv. Inform. Sci. Service Sci.*, **2011**, *3*, 162-168.

42. N. Pukkhem and W. Vatanawood, "Personalised learning object based on multi-agent model and learners' learning styles", *Maejo Int. J. Sci. Technol.*, **2011**, *5*, 292-311.

43. J. Ravi, Z. Yu and W. Shi, "A survey on dynamic Web content generation and delivery techniques", *J. Netw. Comput. Appl.*, **2009**, *32*, 943-960.

44. M. Balabanovic and Y. Shoham, "Fab: Content-based, collaborative recommendation", *Commun. ACM*, **1997**, *40*, 66-72.

Full Paper

Simultaneous determination of plasma lopinavir and ritonavir by chemometrics-assisted spectrophotometry and comparison with HPLC method

Salinthip Jarusintanakorn[1], Kittisak Sripha[1], Chutima Matayatsuk Phechkrajang[1,*] and Prapin Wilairat[2]

[1] Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Mahidol University,
447 Sri-Ayuthaya Rd., Payathai, Rachathevi, Bangkok 10400, Thailand
[2] National Doping Control Centre, Mahidol University, 272 Rama VI Rd., Bangkok, Payathai,
Rachathevi, Bangkok 10400, Thailand

* Corresponding author, chutima.mat@mahidol.ac.th

Abstract: Chemometrics-assisted spectrophotometry for the determination of two protease inhibitors, lopinavir and ritonavir, in plasma was evaluated. A set of calibration mixtures (calibration set) was designed according to central composite design. The UV spectra obtained from the calibration set were subjected to partial least square regression (PLS-1) to construct the prediction models for lopinavir and ritonavir in unknown samples, which were then validated in a randomly selected set of synthetic mixtures of the drugs. An optimum model was obtained in the wavelength ranges of 215–249 nm and 240–279 nm with principal components 8 for both lopinavir and ritonavir respectively. The prediction models were used to analyse the two drugs in plasma and the results were compared with those obtained by high-performance liquid chromatography (HPLC). The PLS-1 model and the HPLC method were found to be comparable.

Keywords: protease inhibitor, lopinavir, ritonavir, chemometrics, PLS-1

INTRODUCTION

Lopinavir (LPV) and ritonavir (RTV) are important HIV-1 protease inhibitors, alone or in combination, for highly active antiretroviral therapy (HAART) [1]. However, there is a relationship between protease inhibitor concentration and pharmacological activity as well as toxicity or side effects. In addition, wide variability in pharmacokinetics of protease inhibitors was also found [2]. To avoid drug toxicity and improve efficacy, therapeutic drug monitoring is required for drugs with a narrow therapeutic range including protease inhibitors [1, 3-5]. Several

analytical methods have been described for the determination of protease inhibitors and other antiretroviral agents in biological samples [1, 3-17]. In the past decade, a combination of antiretroviral drugs including LPV and RTV in human plasma was usually determined by high-performance liquid chromatography – ultraviolet spectroscopy (HPLC-UV) or HPLC-UV/fluorescence [3-9]. Nowadays, liquid chromatography – mass spectrometry (LC-MS) and LC-MS/MS are the most widely adopted methods for quantitating protease inhibitors including LPV and RTV in several types of samples, e.g. human plasma [10-12], dried blood spot [13], dried plasma spot [14] and peripheral blood mononuclear cell [1, 15-17]. Although LC-MS/MS seems to be the most effective instrument due to its sensitivity and specificity when compared with conventional HPLC-UV, the price of the instrument and professional operation has been the limitation on routine therapeutic drug monitoring. On the other hand, a spectrophotometric measurement without prior chromatographic separation is interesting but the overlap of the absorbance spectra remains problematic. However, it is possible to solve this problem with chemometric approaches. Among these, multivariate calibration methods such as multiple linear regression (MLR), principle component regression (PCR) and partial least square (PLS) utilising spectrophotometric data are important for the determination of drugs in combination [18-19]. In our previous study [20], multivariate calibration methods, i.e. PCR and PLS-1, were developed for determination of LPV and RTV in syrup. Although those results were comparable with the developed HPLC method, the previous PCR and PLS-1 models cannot be applied to plasma samples owing to the difference between the syrup matrix and the human plasma matrix. A chemometrics-assisted spectrophoto- metric method, PLS-1, is therefore developed for the determination of LPV and RTV in plasma samples in this study. The same sample set was also analysed by HPLC. Results of the two methods are then compared to determine the reliability of the chemometric approach. To our knowledge, the determination of LPV and RTV in human plasma by PLS-1 has not been reported.

MATERIALS AND METHODS

Reagents, Apparatus and Software

All standard drugs of working standard grade were kindly donated by the Government Pharmaceutical Organisation (GPO), Thailand. Reagent-grade hexane and ethyl acetate were obtained from J. T. Baker. Acetonitrile and methanol (HPLC-grade) were purchased from Labscan. Sodium dihydrogen phosphate was obtained from Merck. The drug-free plasma samples were expired plasma kindly donated by Ramathibodi Hospital, Bangkok. These plasma samples were not in a category which needs ethics committee approval.

UV absorbance was recorded in a 10-mm quartz cell using a PerkinElmer (Lambda 35) UV-Visible spectrophotometer. Manipulation of the spectral data was carried out with UV Winlab software. The chromatographic system consisted of a Shimadzu LC-10AD, a CBM-10A (system controller) and a SPD-10V UV-Vis detector. Hypersil Gold C18 (250 × 4.6 mm i.d., particle size 5 μm) was supplied by Thermo Scientific (USA). A 1.73-GHz Intel personal computer with Windows XP operating system, together with the Unscrambler$^®$ (version 9.8) programme and other statistical analyses, was used to calculate the PLS model.

Preparation of Standard Solutions and Plasma Samples

Stock solutions of LPV (1 mg L^{-1}) and RTV (0.5 mg L^{-1}) were prepared by dissolving the required amount of each drug in methanol. Proper dilution of each stock solution was prepared with 50% methanol to obtain standard solutions for further use in the experiments. The linear concentration range of each drug was determined by one-component calibration. The absorbance spectra were recorded over a range of 200–400 nm using 50% methanol as blank.

Each plasma sample was prepared by spiking 400 µL of drug-free blank plasma with 100 µL of the desired standard mixture solution. Prior to the analysis, interference from the plasma samples was removed by liquid-liquid extraction. The plasma samples and organic solvent were mixed and then centrifuged at 3000 rpm for 15 min. The organic layer was transferred to a clean tube and evaporated under nitrogen at 50°C in a water bath. The extract was reconstituted by adding 2 mL of 50% methanol and centrifuged at 3000 rpm for 5 min., and 1.5 mL of the supernatant was then removed for analysis. For the HPLC method, the samples were cleaned and then filtered through a nylon membrane filter.

Chemometric Method

Two sets of standard LPV- and RTV-spiked plasma samples were prepared. The concentration of each drug in the two sets was within its respective linearity range. One set served as the calibration set, which consisted of 16 drug mixtures in plasma and four blank plasma samples. The second set was the validation set, which contained ten drug mixtures in plasma. The composition of the calibration set was selected based on central composite design (CCD) [21] while that of the validation set was randomly selected. Absorbance data of the two sample sets were collected at wavelengths between 200-400 nm. These data were applied to construct a model using Unscrambler[®] programme. The resulting model was then used to determine the LPV and RTV concentrations in the validation set (test set).

HPLC Method

HPLC conditions, comprising acetonitrile and 25 mM phosphate buffer pH 6 (50:50 v/v) and a C18 column (Hypersil Gold C18, 250 × 4.6 mm i.d., particle size 5 µm), were developed and validated. The flow rate of the system was 1 mL min^{-1}. The presence of the desired drugs was monitored by UV detection at 254 nm. Drug-spiked plasma samples were used for method development. To prove the reliability of the developed HPLC method, important validation parameters, i.e. linearity, accuracy, precision, specificity and system suitability, were investigated. Linearity was established by using the drug-mixture-spiked plasma. Six concentrations of LPV (2–20 µg mL^{-1}) and RTV (0.5–8 µg mL^{-1}) were assessed; indinavir (IDV) at a concentration of 2 µg mL^{-1} was used as internal standard. Linearity was demonstrated by plotting the peak area ratio of the drugs to the internal standard versus concentration of the drugs. Regression parameters were then computed. For accuracy and precision, drug-free plasma was spiked with three concentrations of LPV (3–18 µg mL^{-1}) and RTV (0.75–7 µg mL^{-1}) with five determinations for each concentration. The accuracy of the developed method was expressed as per cent recovery of the amount of added drug vs. amount detected, while precision was expressed as the relative standard deviation (% RSD) of per cent recovery at each concentration. The specificity of the method was confirmed by comparing chromatograms of the drug-spiked plasma samples with those of the drug-free plasma samples.

Three parameters, namely repeatability, resolution and tailing factor, provided the basis for the system suitability testing. Repeatability was based on five replicates of injected-drug mixtures (10 µg LPV mL^{-1} and 3 µg RTV mL^{-1}). Injection repeatability was expressed as % RSD of the peak area ratio of the drug to internal standard obtained from five-replicate injections. The resolution (R$_s$) of each drug and its adjacent peak was calculated and the tailing factor of the peak of the desired drug was also investigated.

RESULTS AND DISCUSSION

Since the spectra of LPV and RTV strongly overlap in the wavelength region of 200–270 nm [20], determination of one drug in the presence of the other by univariate calibration is not possible. Instead, a multivariate calibration technique, i.e. partial least square regression (PLS-1), is chosen for this study. Due to the complexity of plasma matrix and the absence of collinearity data, the PCR is not chosen.

PLS-1 Modelling

To construct the PLS-1 model, the one-component calibration (univariate calibration) of each drug was initially executed. A linearity range was determined in the concentration ranges of 2–20 µg mL^{-1} and 1–6 µg mL^{-1} for LPV and RTV respectively. The concentrations of LPV and RTV in the calibration set assigned for the CCD and the randomly selected concentrations of these drugs in the validation set were within the linearity range (2–12 µg mL^{-1} for LPV and 1–4 µg mL^{-1} for RTV). The composition of the calibration set corresponding to CCD is illustrated in Table 1. Individual analytes were independently modelled by PLS-1 with the optimum wavelength region and number of principal components (PCs) or factors. The optimum wavelength region was selected by visual observation; a region which was less interfered by another drug was chosen for testing.

Owing to the above-mentioned strongly overlapping drug spectra, several wavelength regions and PCs were studied to obtain the fitted models. Leave-one-out cross-validation (LOO-CV) was used to validate the PLS-1 model in the model development. The optimum number of factors (PCs) with minimum prediction error of sum squares for an optimum model was determined from the model development. The optimum prediction models were achieved in the wavelength regions of 215–249 nm with 8 PCs for LPV and 240–279 nm with 8 PCs for RTV. The resulting PLS-1 model for plasma samples contained more PCs than those in the PLS-1 model for syrup [20]. In plasma, the PCs of 8 were found to be the optimum components for both LPV and RTV models while the optimum PCs for LPV and RTV in syrup were 2 and 4 respectively. These may be due to a more complicate matrix in plasma compared to that in syrup.

In the calibration step (model development), statistical parameters such as the root mean square error of calibration, the correlation coefficient (r^2) and the relative error of prediction were also computed. In addition, the independent set of plasma samples containing different compositions of LPV and RTV (test set sample) that did not contribute to the calibration step was used to evaluate the proposed calibration models. The statistical parameter expressing the predictive applicability of a regression model was the standard error of prediction. All relevant statistical parameters of the optimal PLS-1 model for LPV and RTV determination were acceptable for all proposed calibration models as shown in Table 2.

Table 1. Composition of calibration set corresponding to CCD

Sample	Concentration (µg mL^{-1})		Coded level [22]	
	LPV	RTV	LPV	RTV
1	0	2.0	- α	0
2	16.0	2.0	+ α	0
3	8.0	0	0	- α
4	8.0	4.0	0	+ α
5	2.3	0.6	- 1	- 1
6	2.3	3.4	- 1	+ 1
7	13.7	3.4	+ 1	+ 1
8	13.7	0.6	+ 1	- 1
9	8.0	2.0	0	0
10	8.0	2.0	0	0
11	8.0	2.0	0	0
12	8.0	2.0	0	0
13	8.0	2.0	0	0
14	8.0	2.0	0	0
15	8.0	2.0	0	0
16	8.0	2.0	0	0

Table 2. Statistical parameters of the optimum PLS-1 model for LPV and RTV determination

Parameter	LPV	RTV
Spectral range (nm)	215 – 249	240 – 279
PCs	8	8
PRESS	0.6371	0.0213
RMSEC	0.1675	0.0365
r^2	0.9976	0.9986
REP (%)	2.50	1.82
SEP	0.3706	0.2144
SEN	1.02×10^{-3}	7.96×10^{-4}
SEL	9.60×10^{-3}	6.16×10^{-3}

Note: PRESS = prediction error of sum squares; PCs = number of principle component; RMSEC = root mean square error of calibration; r^2 = correlation coefficient; REP = relative error of the prediction; SEP = standard error of prediction; SEN = sensitivity; SEL = selectivity.

The proposed models were also validated in terms of linearity, accuracy, precision and limit of quantitation in order to show that the chemometric method is appropriate to the investigation of LPV and RTV in plasma samples by spiking plasma with different concentrations of the two drugs. Table 3 shows that all parameters meet the requirements of the bioanalytical method validation guidelines [23].

Table 3. Results of validation of PLS-1 model for LPV and RTV

Parameter[a]	LPV	RTV
Spectral range (nm)	215 – 249	240 – 279
PCs	8	8
Linearity	0.9988	0.9975
Accuracy	104	96
Repeatability	5.07	5.85
Inter-day precision	5.00	4.73
LOQ	1.77	0.51

[a] Linearity is expressed as the correlation coefficient (r^2) and accuracy as the average % recovery (n=15). Precision is expressed as relative standard deviation (% RSD) and consists of repeatability (n=15) and inter-day precision (n = 5). Limit of quantitation (LOQ) was calculated by the formula: LOQ = 10(SD/S), where SD is the standard deviation of the response and S is the slope of calibration curve.

HPLC Method Development and Method Validation

LPV and RTV can be well separated from each other and from other interferences present in plasma. The retention times of LPV, RTV and the internal standard IDV are 12.23 min., 10.55 min. and 5.16 min. respectively. Both LPV and RTV exhibit a linear relationship in a graph of the peak area ratio vs. concentration over the entire concentration range (2–20 µg mL^{-1} for LPV and 0.5–8 µg mL^{-1} for RTV) with acceptable linear regression parameters (Table 4). The stability of the proposed method was evaluated by studying the variation in the peak area ratio of LPV and RTV in drugs-spiked plasma.

Table 4. Summary of validation results and system suitability parameters for HPLC method

Parameter (Unit)	LPV	RTV
Linearity range (µg mL^{-1})	2-20	0.5-8
r^2	0.9917	0.9983
Slope	0.0913	0.7261
Intercept	0.0183	0.1118
Repeatability (% RSD, n=5)	1.43-2.62	0.73-3.37
Intermediate precision (% RSD, n=5)	2.10-3.89	2.03-7.86
Recovery (%)	100-104	86-101
Repeatability of peak area ratio (% RSD, n=5)	1.92	0.57
Limit of quantitation (µg mL^{-1})	2.96	0.96
Resolution	4.91	5.18
Tailing factor	1.08	1.09

As summarised in Table 4, the repeatability (intra-day precision) of the determination is indicated by the % RSD value for plasma analyses, which range between 1.43–2.62% for LPV and 0.73–3.37% for RTV. The intermediate precision (inter-day precision) RSD values obtained for the drugs-spiked plasma samples are below 3.89% for LPV and 7.86% for RTV over the entire concentration range of 3–18 µg mL^{-1} for LPV and 0.75–7 µg mL^{-1} for RTV. The results

show that the precision of the proposed HPLC method is satisfactory for the analysis of LPV and RTV in plasma samples. The accuracy of the method is expressed as per cent recovery of the added drug. The mean recoveries are 100–105% for LPV over the concentration range of 3–18 μg mL^{-1} and 86–101% for RTV over the concentration range of 0.75–7 μg mL^{-1}. All figures of merit for method validation meet the requirements of the bioanalytical guidelines [23].

The specificity of the method is illustrated in Figure 1. A comparison of the chromatogram of the drug-spiked plasma and that of the drug-free plasma shows that the endogenous substances in plasma do not elute at the retention times of LPV and RTV, indicating that LPV and RTV can be separated from the sources of interference present in plasma. Additionally, the precision of retention time, resolution and tailing factor (asymmetric peak), the three criteria used to assess the system suitability of the HPLC conditions, are all acceptable (Table 4).
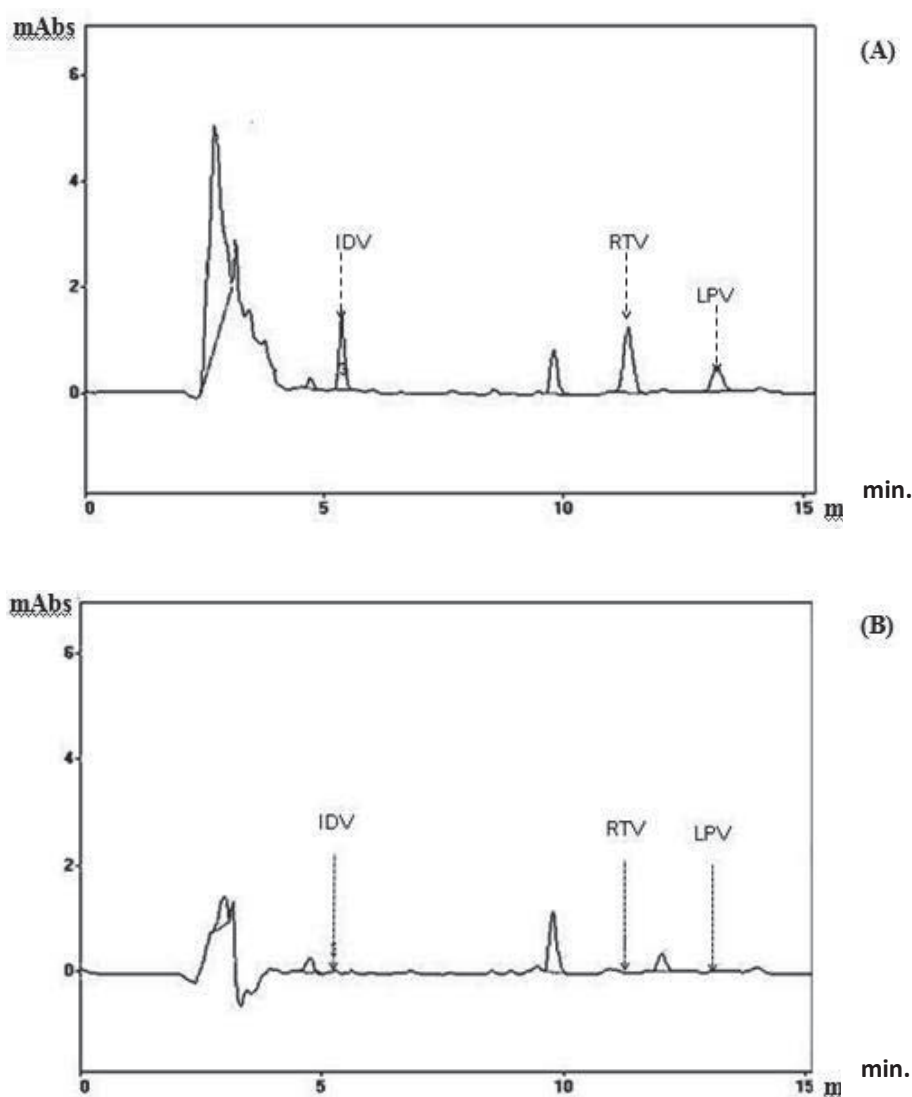
Figure 1. (A) Chromatogram of LPV (10 μg mL^{-1}), RTV (3 μg mL^{-1}) and IDV (2 μg mL^{-1}) spiked in plasma; (B) Chromatogram of drug-free plasma sample with expected retention times of IDV, RTV and LTV

Comparison of Chemometric and HPLC Methods

A set of 10 mixture samples, apart from those used in calibration step, were prepared and analysed using the developed HPLC method and PLS-1 model. The mean recoveries of LPV and RTV from human plasma obtained from the two methods were then compared by using t-test statistic. As shown in Table 5, the mean recovery results for the LPV- and RTV-spiked in plasma are not significantly different at the 95% confidence limit. These results indicate that the simple and rapid chemometric method can be used as an alternative to the HPLC method for the determination of LPV and RTV in plasma.

Table 5. Mean recoveries of LPV and RTV obtained using PLS-1 optimum model compared with HPLC method

LPV			RTV		
Parameter	**HPLC**	**PLS-1**	**Parameter**	**HPLC**	**PLS-1**
Mean % recovery (n=10)	106	103	Mean % recovery (n=10)	96	98
% RSD	3.37	4.29	% RSD	3.15	6.63

CONCLUSIONS

The developed PLS-1 model, employing UV absorbance data successfully determined LPV and RTV concentrations in plasma. Validation of the reliability of the PLS-1 prediction model gave acceptable validation results and was not significantly different from the reference HPLC method in the determination of LPV and RTV, which shows that this chemometrics-assisted spectro-photometric method can be used as an alternative to HPLC in the determination of the drug mixtures.

ACKNOWLEDGEMENTS

REFERENCES

1. A. D′Avolio, M. Simiele, M. Siccardi, L. Baietto, M. Sciandra, V. Oddone, F. R. Stefani, S. Agati, J. Cusato, S. Bonora and G. Di Perri, "A HPLC-MS method for the simultaneous quantification of fourteen antiretroviral agents in peripheral blood mononuclear cell of HIV infected patients optimized using medium corpuscular volume evaluation", *J. Pharm. Biomed. Anal.*, **2011**, *54*, 779-788.

2. L. John, F. Marra and M. H. Ensom, "Role of therapeutic drug monitoring for protease inhibitors", *Ann. Pharmacother.*, **2001**, *35*, 745-754.

3. J. E. Adaway and B. G. Keevil, "Therapeutic drug monitoring and LC-MS/MS", *J. Chromatogr. B*, **2012**, *883-884*, 33-49.

4. S. Notari, A. Bocedi, G. Ippilito, P. Narciso, L. P. Pucillo, G. Tossini, R. P. Donnorso, F. Gasparrini and P. Ascenzi, "Simultaneous determination of 16 anti-HIV drugs in human plasma by high-performance liquid chromatography", *J. Chromatogr. B*, **2006**, *831*, 258-266.

5. S. O. Choi, N. L. Rezk and A. D. Kashuba, "High-performance liquid chromatography assay for the determination of the HIV-protease inhibitor tipranavir in human plasma in combination with nine other antiretroviral medications", *J. Pharm. Biomed. Anal.*, **2007**, *43*, 1562-1567.

6. V. Albert, P. Modamino, C. F. Lastra and E. L. Mariño, "Determination of saquinavir and ritonavir in human plasma by reversed-phase high-performance liquid chromatography and the analytical error function", *J. Pharm. Biomed. Anal.*, **2004**, *36*, 835-840.

7. H. Rebiere, B. Mazel, C. Civade and P. A. Bonnet, "Determination of 19 antiretroviral agents in pharmaceuticals or suspected products with two methods using high-performance liquid chromatography", *J. Chromatogr. B*, **2007**, *850*, 376-383.

8. D. R. Weller, R. C. Brundage, H. H. Balfour Jr and H. E. Vezina, "An isocratic liquid chromatography method for determining HIV non-nucleoside reverse transcriptase inhibitor and protease inhibitor concentrations in human plasma", *J. Chromatogr. B*, **2007**, *848*, 369-373.

9. R. Verbesselt, E. V. Wijngaerden and J. de Hoon, "Simultaneous determination of 8 HIV protease inhibitors in human plasma by isocratic high-performance liquid chromatography with combined use of UV and fluorescence detection: Amprenavir, indinavir, atazanavir, ritonavir, lopinavir, saquinavir, nelfinavir and M8-nelfinavir metabolite", *J. Chromatogr. B*, **2007**, *845*, 51-60.

10. L. Dickinson, L. Robinson, J. Tjia, S. Khoo and D. Back, "Simultaneous determination of HIV protease inhibitors amprenavir, atazanavir, indinavir, lopinavir, nelfinavir, ritonavir and saquinavir in human plasma by high-performance liquid chromatography-tandem mass spectrometry", *J. Chromatogr. B*, **2005**, *829*, 82-90.

11. A. D´Avolio, M. Siccardi, M. Sciandra, L. Baietto, S. Bonora, L. Trentini and G. D. Perri, "HPLC-MS method for the simultaneous quantification of the new HIV protease inhibitor darunavir, and 11 other antiretroviral agents in plasma of HIV-infected patients", *J. Chromatogr. B*, **2007**, *859*, 234-240.

12. G. A. Temghare, S. S. Shetye and S. S. Joshi, "Rapid and sensitive method for quantitative determination of lopinavir and ritonavir in human plasma by liquid chromatography-tandem mass spectrometry", *E-J. Chem.*, **2009**, *6*, 223-230.

13. R. ter Heine, H. Rosing, E. C. van Gorp, J. W. Mulder, W. A. van der Steeg, J. H. Beijnen and A. D. Huitema,"Quantification of protease inhibitors and non-nucleoside reverse transcriptase inhibitors in dried blood spots by liquid chromatography-triple quadrupole mass spectrometry", *J. Chromatogr. B*, **2008**, *867*, 205-212.

14. A. D′Avolio, M. Simiele, M. Siccardi, L. Baietto, M. Sciandra, S. Bonora and G. D. Perri, "HPLC-MS method for the quantification of nine anti-HIV drugs from dry plasma spot on glass filter and their long term stability in different conditions", *J. Pharm. Biomed. Anal.*, **2010**, *52*, 774-780.

15. S. Colombo, A. Beguin, A. Telenti, J. Biollaz, T. Buclin, B. Rochat and L. A. Decosterd, "Intracellular measurements of anti-HIV drugs indinavir, amprenavir, saquinavir, ritonavir, nelfinavir, lopinavir, atazanavir, efavirenz and nevirapine in peripheral blood mononuclear cells by liquid chromatography coupled to tandem mass spectrometry", *J. Chromatogr. B*, **2005**, *819*, 256-276.

16. M. Ehrhartd, M. Möck, W. E. Haefeli, G. Mikus and J. Burhenne, "Monitoring of lopinavir and ritonavir in peripheral blood mononuclear cells, plasma, and ultrafiltrate using a selective and highly sensitive LC/MS/MS assay", *J. Chromatogr. B*, **2007**, *850*, 249-258.

17. R. ter Heine, M. Davids, H. Rosing. E. C. van Gorp, J. W. Mulder, Y. T. van der Heide, J. H. Beijnen and A. D. Huitema, "Quantification of HIV protease inhibitors and non-nucleoside reverse transcriptase inhibitors in peripheral blood mononuclear cell lysate using liquid chromatography coupled with tandem mass spectrometry", *J. Chromatogr. B*, **2009**, *877*, 575-580.

18. E. Dinte, I. Tomuta, E. M. Mut, R. I. Iovanov and S. E. Leucuta, "Chemometric methods for the simultaneous assay of chloramphenicol, chlorhexidine and metronidazole during *in vitro* dissolution of drugs from mucoadhesive buccal gels", *Farmacia*, **2010**, *58*, 572-582.

19. N. W. Ali, S. S. Abbas, H. E. Zaazaa, M. M. Abdelrahman and M. Abdelkawy, "Validated stability indicating methods for determination of nitazoxanide in presence of its degradation products", *J. Pharm. Anal.*, **2012**, *2*, 105-116.

20. C. M. Phechkrajang, E. E. Thin, L. Sratthaphut, D. Nacapricha and P. Wilairat, "Chemometrics-assisted UV spectrophotometric method for determination of lopinavir and ritonavir in syrup", *Int. J. Pharm. Pharm. Sci.*, **2012**, *4(suppl. 1)*, 492-496.

21. R. G. Brereton, "Chemometrics: Data Analysis for the Laboratory and Chemical Plant", John Wiley & Sons, Chichester, **2003**, pp.76-84.

22. Morgan, "Chemometrics: Experimental Design", John Wiley & Sons, Chichester, **2003**, pp.238-245.

23. U.S. Department of Health and Human Services, Food and Drug Administration, "Guidance for industry: Bioanalytical method validation", http://www.fda.gov/downloads/ Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ UCM070107.pdf (Accessed: July 2012).

Maejo International Journal of Science and Technology

Technical Note

Reduction in energy consumption and operating cost in a dried corn warehouse using logistics techniques

Korrakot Y. Tippayawong*, Peerat Piriyageera-anan and Teeraphat Chaichak

Department of Industrial Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai 50200, Thailand

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

Abstract: Corn is one of the major economic crops in Thailand. Corn postharvest operation involves various practices that consume a large amount of energy. Different energy conservation measures have been implemented but logistics consideration is not normally employed. In this work, attempt has been made to demonstrate that logistics techniques can offer a significant reduction in energy and cost. The main objective of this work is to identify and demonstrate possible approaches to improving energy efficiency and reducing operating cost for a dried corn warehouse operator. Three main problems are identified: (i) relatively high fuel consumption for internal transfer process, (ii) low quality of dried corn, and (iii) excess expenditure on outbound transportation. Solutions are proposed and implemented using logistics operations. Improvement is achieved using plant layout and shortest path techniques, resulting in a reduction of almost 50% in energy consumption for the internal transfer process. Installation of an air distributor in the grain storage unit results in a decrease in loss due to poor-quality dried corn from 17% to 10%. Excess expenditure on dried corn distribution is reduced by 6% with application of a global positioning system.

Keywords: agro-industry, logistics, plant layout design, dried corn warehouse

INTRODUCTION

Corn is one of the major upland crops in Thailand with large plantation areas throughout the whole country. Thailand has achieved considerable success in growing corn as an alternative to rice. Since the past few decades, corn cultivation has proved a remarkable success for farmers and the nation as a whole. Thailand's livestock industry consumed about 3-4 million tons of corn a year while 400,000 tons were exported. In 2009, however, the export amounted to over 840,000 tons and was worth more than 5,300 million Thai Baht (THB) [1]. Corn for domestic

and foreign consumption has become a significant contributor to the Thai economy and the demand and production continue to grow.

Typically, postharvest processes for corn involve: (i) field drying and harvesting, (ii) threshing and drying, (iii) shelling and cleaning, (iv) storage, and (v) transport [2]. In a typical warehouse, trucks, loaders, several types of dryers, storage facilities and silos can be found for drying and handling the grain. Drying is normally done before storing to ensure high quality and value. Storage is done by stacking the dried corn on a cement floor or in a silo. At this stage, deterioration of corn grain during storage due to heat liberated from the respiration process is normally controlled by passing cool air or ambient air through the bed to reduce the bed temperature. Energy is expended to power vehicles and various postharvest processing technologies are used. It is clear that a typical corn grain handling facility consumes a large amount of energy. There are a number of energy-saving techniques and many different energy conservation measures have been implemented in agro-product industries [3-6]. However, logistics techniques are not normally considered. In other industries, logistics considerations such as plant layout modification, warehouse operation, improved transportation network, and implementation of information technology have been successfully demonstrated [7-10].

Logistics activities incorporate three main areas; these are inbound, internal and outbound logistics. Inbound logistics involves activities in procurement process and supplier integration. Enhanced integration with suppliers can influence several dimensions of firm performance including cost, quality, technology, delivery, flexibility and profits [11]. Internal logistics can be viewed as operations in manufacturing, material transfer and handling, and inventory and warehouse control [12]. Internal logistics starts when raw materials enter the company until the product is ready for distribution. The activities of internal logistics affect numerous aspects of the products including cost, quality and performance. Consequently, they must be constantly monitored and evaluated from a continuous improvement perspective [13]. Outbound logistics encompasses such processes as distribution, marketing, sales and service. A system based on information technology is found to enhance such processes in many ways including accurate order delivery information used to provide appropriate capacity at expertise level regarding customer support. Moreover, the tracking of product delivery allows for increased customisation, fulfillment of customers' requirements and reduction in long-term transportation cost. This results in a higher inventory turnover with a higher market share driven by reliable and responsive availability of those products/services most desired by customers [14].

In this investigation an attempt is made to apply logistics techniques to improving energy efficiency and reducing the cost of operating a dried corn export business. The main focus is on the activities of internal and outbound logistics, where the main problems usually take place.

RESEARCH METHODOLOGY
Case Study

In this study a dried corn processing company (Nipen Kaset Paisan Co., Ltd.) is used as a case study. It is situated in Payao province, about 700 km north of Bangkok. Corn grain of 40% moisture or less is purchased from farmers, dried to below 14.5% moisture and kept in storage prior to transporting to customers. The monthly capacity is about 3,000 tons.

Data Collection and Preliminary Analysis

A preliminary interview with the company owner was conducted to identify the most important problems regarding energy and operating cost, which were: a high operational cost, large amounts of poor quality product and uncontrolled transportation costs. Data on activities were then collected based on internal and external operations. In the case of the former the details of plant layout and storage system of dried corn were collected. In the case of the latter the data on transportation cost were gathered, together with the quality of product in terms of per cent humidity. The data on current practices of the company were analysed to identify the activities with the highest costs using logistics cost analysis with activity-based costing (ABC) technique. The technique is defined by the Chartered Institute of Management Accountants [15] as 'an approach to the costing and monitoring of activities which involves tracing resource consumption and costing final outputs. Resources are assigned to activities, and activities to costing objects based on consumption estimates. The latter utilise cost drivers to attach activity costs to outputs.'

Four related costs: area utilisation, personnel, machines and consumables were used in the ABC calculations. The approach to the logistics ABC calculations is summarised in Figure 1 and the cost analysis results of existing activities are presented in Table 1. Problematic areas can be identified, which leads to an alternation of the existing warehouse, which could considerably reduce the cost on internal transfer. The costs on activity number 2, 3 and 6 have been reduced accordingly. Since high expenditure was spent on activity number 5 and 7, the actual root causes of the problem are further evaluated by cause and effect analysis using fish bone diagrams. The problem analysis on dried corn storage activity and dried corn delivery are shown in Figures 2 and 3 respectively. It is clearly seen that the problem of inappropriate dried corn storage can lead to the deterioration of the product quality. Additionally, high transport cost stemmed from the driving behaviour of the employees. The solutions to these problems are proposed in the next section.

Define activities

↓

Analyse 4 overall related costs

↓

Define criteria of cost allocation

↓

Calculate cost of each activity

↓

Define amount of each activity in a monthly period

↓

Calculate unit cost of each activity
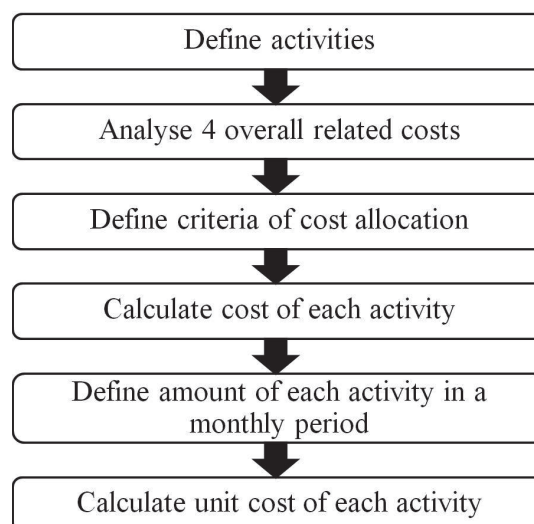
Figure 1. Approach to logistics cost calculation

Table 1. Existing monthly costs calculated by ABC technique

Activity	Cost (THB)	Number of activities (tons)	Unit cost (THB)
1. Pre-dry corn weighing	3,074	3,000	1.025
2. Transfer of pre-dry corn to dryer	37,000	3,000	12.33
3. Transfer of firewood to dryer	121,443	230	528.0
4. Corn drying process	59,655	3,000	19.88
5. Dried corn storage	94,656	2,490	38.01
6. Transfer of dried corn from dryer to truck	37,903	2,490	15.22
7. Dried corn delivery	651,459	2,490	261.6
8. Return of poor-quality dried corn	77,080	125	616.6
Total	1,081,470		
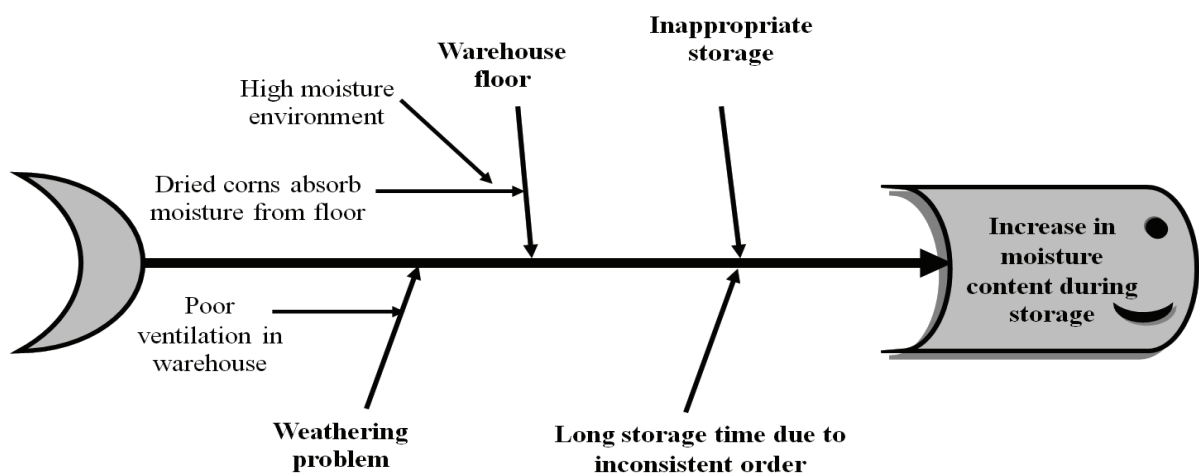
Figure 2. Cause and effect analysis for corn storage problem
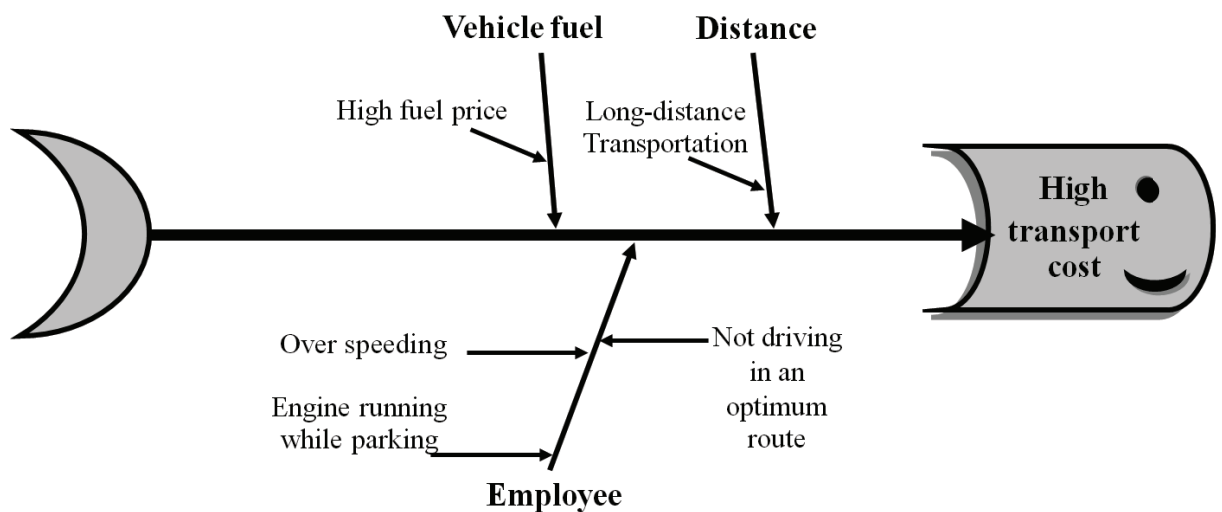
Figure 3. Cause and effect analysis for corn delivery problem

Improvement Approaches

Appropriate solutions were proposed and implemented to the company's existing practice. Economic feasibility of the proposed techniques was also evaluated using break-even analysis. These improvement approaches are:

Alteration of existing plant layout

Existing plant layout led to high energy consumption and long operation time. Modification should be made to the current layout. Unnecessary movements should be eliminated from the current transfer processes. A plant layout alteration should be adopted to reduce the overall transfer distance of internal logistics for higher efficiency in terms of energy consumption and time taken.

Installation of grain aeration system

Previously, dried grains were piled on a cement floor, exposed directly to ambient condition. Some fractions of the grains were prone to moisture absorption and recondensation, leading to damage. This was due mainly to poor storage condition. Improvement of the current storage should be implemented to preserve the quality of dried corn. This could be done by installing a proper aeration system to control the relative humidity and temperature of the dried grain bed. The aeration system consists of an air pump and a network of air distribution pipes laid on the base of the storage floor. There are holes along the length of these distribution pipes to supply air under the piles of dried grains. In this way the moisture content and temperature of the grains could be regulated.

Application of GPS in transportation

Fuel consumption of trucks plying the route from and to the company warehouses amounted to a high operating cost. The major problems involved overspeeding, driving out of the optimum route and keeping the engine running while parking. Installation of a global positioning system (GPS) in trucks would enable the company to track the product movement as well as the driving behaviour of drivers. This technique could increase reliability of product delivery and decrease vehicle fuel consumption during transportation.

RESULTS AND DISCUSSION

Change in Plant Layout Design

The plant layout in the case study with three internal transfer activities which appeared to have excess movements was modified. These activities were:

- Transfer of pre-dry corn to the dryer
- Transfer of firewood to the dryer
- Transfer of dried corn from the dryer to trucks for distributionx c

A wheel loader was used for these internal transfer processes. Improvements made were: (i) moving the location of pre-dry corn closer to the dryer, (ii) moving firewood station closer to the dryer, and (iii) shortening the path between the dryer and the distribution area. Results are schematically illustrated in Figures 4-6 respectively. Unnecessary movements were eliminated and the overall layout were changed, leading to a significant reduction (greater than 170 m) in the transfer distance. Table 2 summarises the percentage reduction in distance after the plant layout improvement on the three internal transfer activities. Table 3 shows savings in fuel consumption after the layout improvement. The cost assessment for overall logistic activities within the company is presented in Table 4. Only the vehicle fuel cost was considered. Three

internal activities are highlighted. The overall fuel cost was found to be reduced from about 135,000 THB to around 72,000 THB per month. This is almost a 50% saving on the internal transfer cost. It is clear that the change in plant layout can significantly lower energy usage.

(before)   (after)
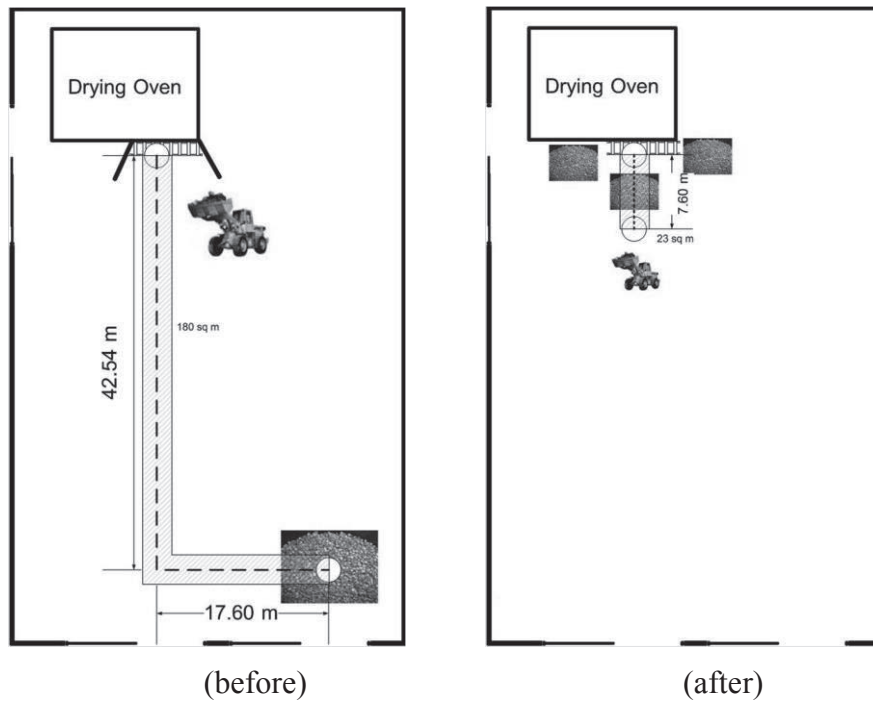
Figure 4. Layout modification of the transferring of corn to the dryer

(before)   (after)
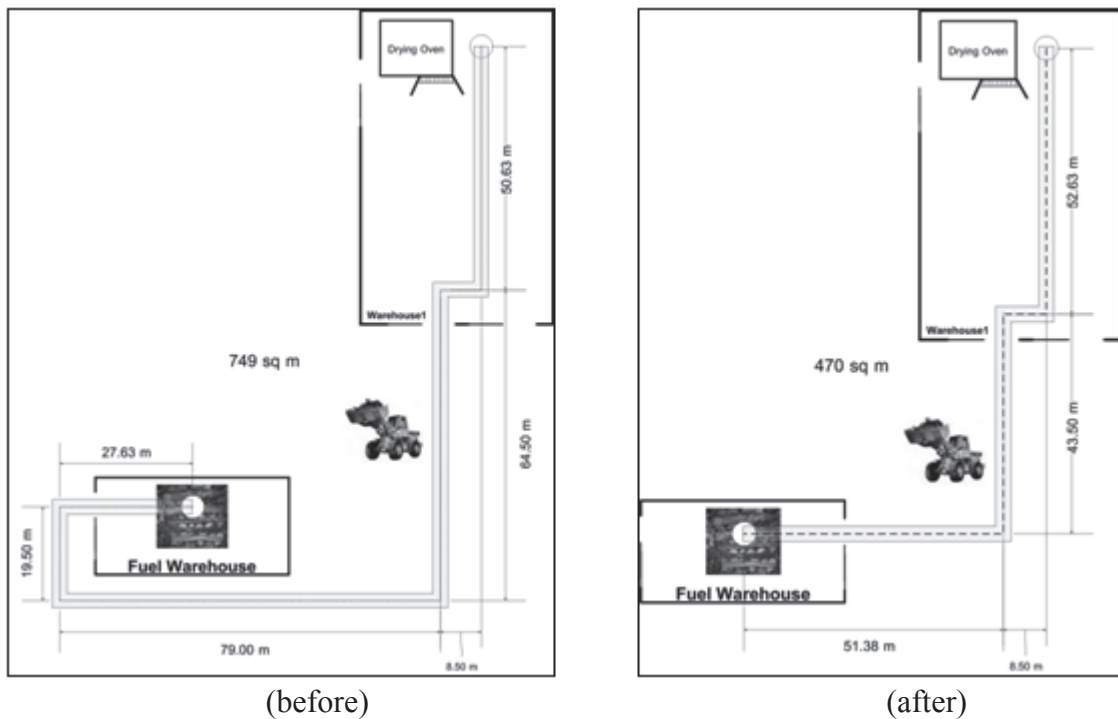
Figure 5. Layout modification of the transferring of firewood to the dryer

263
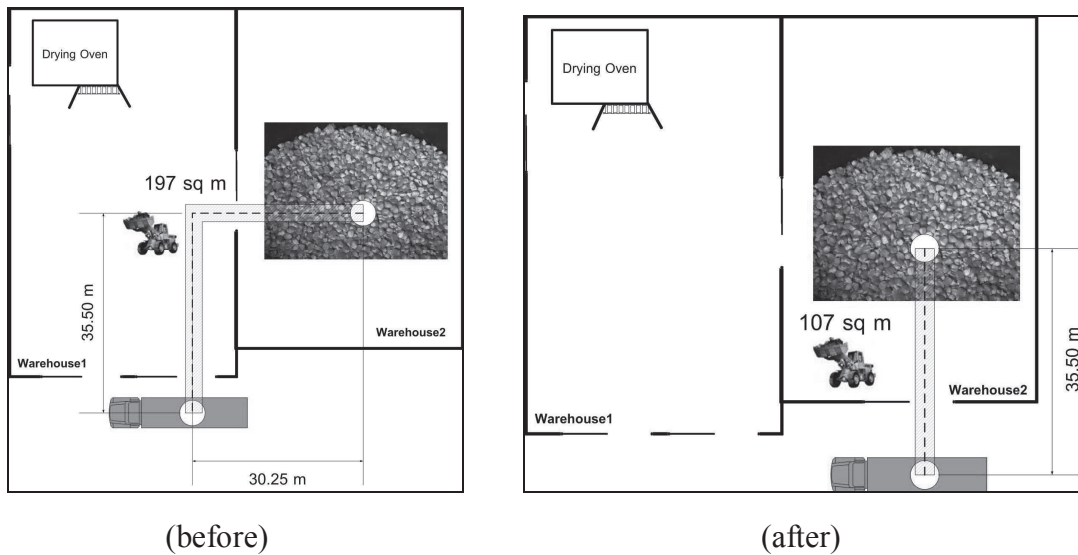
(before)   (after)

Figure 6. Layout modification of the transferring of dried corn to truck for distribution

Table 2. Distance reduction resulting from plant layout improvement

Activity	Distance before improvement (m)	Distance after improvement (m)	% Reduction
1. Transfer of pre-dry corn to dryer	60.1	7.6	87.0
2. Transfer of firewood to dryer	249.7	156.0	38.0
3. Transfer of dried corn from dryer to trucks	65.8	35.5	46.0

Table 3. Vehicle fuel consumption (monthly) before and after plant layout improvement

Activity	Fuel consumption (before)		Fuel consumption (after)	
	% Cost allocation	Cost (THB)	% Cost allocation	Cost (THB)
1. Pre-dry corn weighing	-	-	-	-
2. Transfer of pre-dry corn to dryer	5.05	21,665	0.75	2,740
3. Transfer of firewood to dryer	20.96	89,917	15.37	56,166
4. Corn drying process	-	-	-	-
5. Dried corn storage	-	-	-	-
6. Transfer of dried corn from dryer to truck	5.52	23,680	3.50	12,790
7. Dried corn delivery	61.23	262,678	71.88	262,672
8. Return of poor-quality dried corn	7.24	31,060	8.50	31,062
Total	100.00	429,000	100.00	365,430

Table 4. The comparison of costs before and after plant layout modification

Activity	Number of activities	Cost in THB (before)		Cost in THB (after)	
		Process cost	Unit cost	Process cost	Unit cost
2. Transfer of corn to dryer	3,000 tons	37,000	12.07	18,075	6.030
3. Transfer of firewood to dryer	230 tons	121,443	528.0	87,692	381.3
6. Transfer of dried corn to truck	2,490 tons	37,903	15.22	27,013	10.85
Total		196,346		132,780	

Installation of Grain Aeration System and GPS

When the grain aeration system was installed, it was found to operate well, with temperature and humidity being controlled at required values. With the aeration system, the corn grains were dried to 14.5% moisture content rather than the previous 13.8%. With a better storage system, over-drying of the product was no longer necessary. The installation also reduced firewood consumption by about 10% during the drying process (from 230 tons to 210 tons a month). Furthermore, the waste derived from poor quality product after storage was cut down by over 40% (from 510 tons to 300 tons).

Equipping trucks with the GPS increased monitoring and control of driving pattern as well as reliability of transportation. In this way the vehicle fuel was reduced; for a typical round-trip route of about 600 km, almost 20 litres or 6% of diesel fuel were saved.

Economic Analysis

An economic analysis using break-even point was employed and the results are shown in Table 5. Plant layout improvement was found to deliver the best return on energy cost; the transfer cost was reduced by almost 50%. The payback period of this investment was only 3 months, which was very attractive for the investor. Installation of the aeration system appeared to involve the highest initial investment cost. The substandard product was reduced by 40% per month, leading to an increased revenue by almost 50,000 THB per month. Moreover, the firewood utilised during the drying process was reduced by nearly 10%. Saving in energy was an additional benefit. However, the company had to invest around 1,100,000 THB with a payback period of almost two years. High initial investment could deter the owner from adopting this method. Installing of the GPS provided the shortest break-even time for investment with only 17 days to recover, although the cost saving was only 6% compared with operations without GPS.

Table 5. Break-even analysis of improvement approaches

Approch 1 : Plant layout improvement	
1. Plant layout improvement cost	191,240 THB
Total cost for improvement	191,240
Cost reduction (fuel consumption based)	
Internal transfer cost (before improvement)	135,262 THB/month
Internal transfer cost (after improvement)	71,696 THB/month
Cost saving (per month)	63,566
Break-even point	3 months

Approach 2: Aeration system installation	
1. Aeration system installation cost	851,676 THB
2. Yearly maintenance cost	129,600 THB
3. Electricity cost per year	144,000 THB
Total cost	1,125,276
Improvement detail	
Value of damage (before improvement)	119,340 THB/month
Value of damage (after improvement)	70,200 THB/month
Cost saving (per month)	49,140
Break-even point	23 months

Table 5. (continued)

Approach 3: GPS installation	
1. GPS equipment cost	14,500 THB
2. Yearly system maintenance cost	4,200 THB
Total cost for GPS	18,700
Improvement detail	
Fuel consumption (before GPS installation)	316.84 litres/round
Fuel consumption (after GPS installation)	297.72 litres/round
Fuel saving	19.12 litres/round
Fuel cost	28.6 THB/ litre
Saving per round	546.83 THB
Break-even point	34 rounds = 17 days (2 rounds/day)

CONCLUSIONS

It was clear that logistics considerations can offer options in the improvement on energy efficiency and cost reduction in dried corn production, hence significant savings for dried corn warehouse operators. Similar analysis may benefit other agro-product manufacturers who wish to adopt similar improvement approaches.

ACKNOWLEDGMENTS

REFERENCES

1. Office of Agricultural Economics, Ministry of Agriculture, Thailand, "Import-export of major agricultural crops", **2010**, www.oae.go.th (Accessed: Jan. 2010).

2. Food and Agriculture Organization (FAO), "Maize in human nutrition", **1993**, www.fao.org/docrep/ T0395E/T0395E00.htm (Accessed: Jan. 2010).

3. N. Dussadee and T. Kiatsiriroat, "Performance analysis and economic evaluation of thermosyphon paddy bulk storage", *Appl. Thermal Eng.*, **2004**, *24*, 401-414.

4. N. Tippayawong, C. Tantakitti and S. Thavornun, "Energy efficiency improvements in longan drying practice", *Energy*, **2008**, *33*, 1137-1143.

5. N. Tippayawong, C. Tantakitti, S. Thavornun and V. Peerawanitkul, "Energy conservation in drying of peeled longan by forced convection and hot air recirculation", *Biosyst. Eng.*, **2009**, *104*, 199-204.

6. V. Smil, P. Nachman and T. V. Long, "Technological changes and the energy cost of US grain corn", *Energy Agric.*, **1983**, *2*, 177-192.

7. L. C. Lin and G. P. Sharp, "Quantitative and qualitative indices for the plant layout evaluation problem", *Eur. J. Oper. Res.*, **1999**, *116*, 100-117.

8. M. T. Torres, M. C. Barros, P. M. Bello, J. J. Casares and J. M. Rodriguez-Blas, "Energy and material flow analysis: Application to the storage stage of clay in the roof-tile manufacture", *Energy*, **2008**, *33*, 963-973.

9. S. Jia, H. Peng, S. Liu and X. Zhang, "Review of transportation and energy consumption related research", *J. Transport. Sys. Eng. Inform. Technol.*, **2009**, *9*, 6-16.

10. S. McCabe, M. Roorda and H. Kwan, "Comparing GPS and non-GPS survey methods for collecting urban goods and service movements", Proceedings of 8th International Conference on Survey Methods in Transport, **2008**, Annecy, France.

11. B. Dehning, V. J. Richardson and R. W. Zmud, "The financial performance effects of IT-based supply chain management systems in manufacturing firms", *J. Oper. Manage.*, **2007**, *25*, 806-824.

12. B. S. Blanchard, "Logistics Engineering and Management", Pearson Prentice Hall, Upper Saddle River (N.J.), **2004**.

13. A. Gunasekaran, C. Patel and R. E. McGaughey, "A framework for supply chain performance measurement", *Int. J. Product. Econ.*, **2004**, *87*, 333-347.

14. S. K. Vickery, J. Jayaram, C. Droge and R. Calantone, "The effects of an integrative supply chain strategy on customer service and financial performance: An analysis of direct versus indirect relationships", *J. Oper. Manage.*, **2003**, *21*, 523-539.

15. Chartered Institute of Management Accountants, UK, **2013**, www.cimaglobal.com/ (Accessed: May 2013).

Maejo International Journal of Science and Technology

Full Paper

Some properties of a subclass of non-Bazilevic functions

Mohsan Raza [1, *], **Khalida Inayat Noor** [2] **and Kamran Yousaf** [2]

[1] Department of Mathematics, GC University Faisalabad, Pakistan
[2] Department of Mathematics, COMSATS Institute of Information Technology, Islamabad
  Pakistan

* Corresponding author, e-mail: mohsan976@yahoo.com

Abstract: The aim of this paper is to generalise the class of non-Bazilevic functions by using the concept of differential subordinations. The inclusion relations, the coefficient bound, the covering theorem and the famous Fekete-Szego inequality related with this subclass of analytic functions are studied.

Keywords: non-Bazilevic functions, differential subordination, Fekete-Szego inequality

---

INTRODUCTION

Let A denote a class of analytic functions of the form

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n, \tag{1}$$

which is defined in the open unit disc $E = \{z : |z| < 1\}$. A function f in A is said to be a starlike function of order ρ if and only if

$$\operatorname{Re} \frac{zf'(z)}{f(z)} > \rho, \ 0 \le \rho < 1, z \in E.$$

This class of functions is denoted by $S^*(\rho)$. It is noted that $S^*(0) = S^*$. Let f_1 and f_2 be two functions which are analytic in E. We say that the function f_1 is subordinate to the function f_2 in E (write $f_1 \prec f_2$ or $f_1(z) \prec f_2(z)$) if there exists a function w analytic in E with $w(0) = 0$ and $|w(z)| < 1$ in E such that $f_1(z) = f_2(w(z))$. A function f in A is said to be a Janowski starlike function denoted by $S^*[A, B]$ if and only if

$$\frac{zf'(z)}{f(z)} \prec \frac{1 + Az}{1 + Bz}, \ -1 \le B < A \le 1, z \in E.$$

Obradovic [1] introduced a class of functions $f \in A$ such that

$$\operatorname{Re}\left(f'(z)\left(\frac{z}{f(z)} \right)^{\alpha} \right) > 0, \ 0 < \alpha < 1, \ z \in E.$$

This class of functions was then called a non-Bazilevic type. Tuneski and Darus [2] obtained the Fekete Szego inequality for the non-Bazilevic class of functions. Using the concept of non-Bazilevic class of functions, Wang et al. [3] studied many subordination results for the class $N(\mu, \lambda, A, B)$ defined as:

$$N(\mu, \lambda, A, B) = \left\{ f \in A : (1+\lambda)\left(\frac{z}{f(z)} \right)^{\alpha} - \lambda f'(z)\left(\frac{z}{f(z)} \right)^{\alpha+1} \prec \frac{1+Az}{1+Bz} \right\},$$

where $\lambda \in \mathbb{C}, \ -1 \leq B < 1, \ A \neq B$ and $0 < \alpha < 1$.

Using the concept of subordination and Non-Bazilevicness, we generalize and define a subclass of non-Bazilevic functions as follows.

Definition 1. A function $f \in N_{\alpha,\mu}(A, B)$ if it satisfies the condition:

$$f'(z)\left(\frac{z}{f(z)} \right)^{\alpha} + \mu\left\{ 1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1 \right\} \prec \frac{1+Az}{1+Bz}, \ z \in E, \qquad (2)$$

where $\mu > 0$, $-1 \leq B < A \leq 1$ and $0 < \alpha < 1$.

For $A = 1 - 2\rho$, $B = -1$, we have the class $N_{\alpha,\mu}(\rho)$ defined as follows:

$$\operatorname{Re}\left\{ f'(z)\left(\frac{z}{f(z)} \right)^{\alpha} + \mu\left(1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1 \right) \right\} > \rho, \ z \in E.$$

Throughout in this paper we assume that $\mu > 0$, $-1 \leq B < A \leq 1$ and $0 < \alpha < 1$ unless otherwise specified.

PRELIMINARY RESULTS

We need the following lemmas, which will be used in our main results.

Lemma 1 [4]**.** If $-1 \leq B < A \leq 1, \beta > 0$ and the complex number γ satisfies $\operatorname{Re}\gamma \geq -\beta(1-A)/(1-B)$, then the differential equation,

$$q(z) + \frac{zq'(z)}{\beta q(z) + \gamma} = \frac{1+Az}{1+Bz}, \ z \in E,$$

has the univalent solution in E given by

$$
q(z) = \begin{cases} \dfrac{z^{\beta+\gamma}\left(1+Bz\right)^{\beta(A-B)/B}}{\beta\displaystyle\int_0^z t^{\beta+\gamma-1}\left(1+Bt\right)^{\beta(A/B-1)}dt} - \dfrac{\gamma}{\beta}, & B \neq 0, \\[4ex] \dfrac{z^{\beta+\gamma}e^{\beta Az}}{\beta\displaystyle\int_0^z t^{\beta+\gamma-1}e^{\beta At}dt} - \dfrac{\gamma}{\beta}, & B = 0. \end{cases}
$$

If $h(z) = 1 + c_1 z + c_2 z^2 + \dots$ satisfies

$$
h(z) + \frac{zh'(z)}{\beta h(z) + \gamma} \prec \frac{1+Az}{1+Bz}, \quad z \in E,
$$

then

$$
h(z) \prec q(z) \prec \frac{1+Az}{1+Bz},
$$

and $q(z)$ is the best dominant.

Lemma 2 [5]. Let ε be a positive measure on $[0,1]$ and let g be a complex-valued function defined on $E \times [0,1]$ such that $g(.,t)$ is analytic in E for each $t \in [0,1]$ and that $g(z,.)$ is ε-integrable on $[0,1]$ for all $z \in E$. In addition, suppose that $\operatorname{Re}\{g(z,t)\} > 0$, $g(-r,t)$ is real and $\operatorname{Re}\{1/g(z,t)\} \geq 1/(g(-r,t)$ for $|z| \leq r < 1$ and $t \in [0,1]$.

If $g(z) = \displaystyle\int_0^1 g(z,t)d_\varepsilon(t)$, then $\operatorname{Re}\{1/g(z)\} \geq 1/g(-r)$.

Lemma 3 [6]. Let a, b and $c \neq 0, -1, -2\dots$ be complex numbers. Then, for $\operatorname{Re}c > \operatorname{Re}b > 0$

(i) $\quad F(a,b,c;z) = \dfrac{\Gamma(c)}{\Gamma(c-b)\Gamma(b)}\displaystyle\int_0^1 t^{b-1}\left(1-t\right)^{c-b-1}\left(1-tz\right)^{-a}dt,$

(ii) $\quad {}_2F_1(a,b,c;z) = {}_2F_1(b,a,c;z),$

(iii) $\quad {}_2F_1\left(a,b,c;z\right) = \left(1-z\right)^{-a}\,{}_2F_1\left(a,c-b,c;\dfrac{z}{z-1}\right).$

Lemma 4 [7]. Let $-1 \leq B_1 \leq B_2 < A_2 \leq A_1 \leq 1$. Then

$$
\frac{1+A_2 z}{1+B_2 z} \prec \frac{1+A_1 z}{1+B_1 z}.
$$

Lemma 5 [8]. Let F be analytic and convex in E. If f, $g \in A$ and f, $g \prec F$. Then

$$
\lambda f + \left(1-\lambda\right)g \prec F, \quad 0 \leq \lambda \leq 1.
$$

Lemma 6 [9]. Let $f(z) = \sum_{k=0}^{\infty}a_k z^k$ be analytic in E and $F(z) = \sum_{k=0}^{\infty}b_k z^k$ be analytic and convex in E. If $f \prec F$, then

$$
|a_k| \leq |b_1| \quad (k \in \mathbb{N}).
$$

Lemma 7 [10]. If $p(z) = 1 + p_1 z + p_2 z^2 + \dots$ is a function with positive real part in E, then

$$| p_2 - vp_1^2 | \leq \begin{cases} -4v+2, & v \leq 0, \\ 2, & 0 \leq v \leq 1, \\ 4v-2, & v \geq 1. \end{cases}$$

When $v < 0$ or $v > 1$, equality holds if and only if $p(z)$ is $\dfrac{1+z}{1-z}$ or one of its rotations. If

$0 < v < 1$, then equality holds if and and only if $p(z) = \dfrac{1+z^2}{1-z^2}$ or one of its rotations. If $v = 0$,

equality holds if and only if $p(z) = \left(\dfrac{1}{2} + \dfrac{\eta}{2}\right)\dfrac{1+z}{1-z} + \left(\dfrac{1}{2} - \dfrac{\eta}{2}\right)\dfrac{1-z}{1+z}$, $(0 \leq \eta \leq 1)$ or one of its

rotations. If $v = 1$, equality holds if and only if p is the reciprocal of one of the functions such that equality holds in the case of $v = 0$. Although the above upper bound is sharp, when $0 < v < 1$, it can be improved as follows:

$$| p_2 - vp_1^2 | + v | p_1 |^2 \leq 2, \quad (0 < v \leq 1/2),$$

and

$$| p_2 - vp_1^2 | + (1-v) | p_1 |^2 \leq 2, \quad (1/2 < v \leq 1).$$

Lemma 8 [11]. If $p(z) = 1 + p_1 z + p_2 z^2 + \dots$ is a function with positive real part in E, then for a complex number v,

$$| p_2 - vp_1^2 | \leq 2 \max\{1, |2v-1|\}.$$

This result is sharp for the functions

$$p(z) = \frac{1+z^2}{1-z^2}, \quad p(z) = \frac{1+z}{1-z}.$$

MAIN RESULTS

Theorem 1. If $f \in N_{\alpha,\mu}(A, B)$, then

$$f'(z)\left(\frac{z}{f(z)}\right)^{\alpha} \prec \frac{1}{1/\mu Q(z)} = q(z) \prec \frac{1+Az}{1+Bz}, \tag{3}$$

where

$$Q(z) = \begin{cases} \displaystyle\int_0^1 t^{\frac{1}{\mu}-1}\left(\frac{1+Btz}{1+Bz}\right)^{\frac{1}{\mu}(A-B)/B} dt, & B \neq 0, \\ \displaystyle\int_0^1 t^{\frac{1}{\mu}-1} e^{\frac{1}{\mu}Az(t-1)} dt, & B = 0. \end{cases} \tag{4}$$

and $q(z)$ is the best dominant. In addition if $A < -\mu B$, $-1 \leq B < 0$, then $N_{\alpha,\mu}(A, B) \subset N(\alpha, \rho)$, where

$$\rho = \left\{ {}_2F_1\left(1, \frac{1}{\mu}\left(1 - \frac{A}{B}\right); \frac{1}{\mu} + 1; \frac{B}{B-1}\right) \right\}^{-1}. \tag{5}$$

This result is best possible.

Proof. Let

$$h(z) = f'(z)\left(\frac{z}{f(z)}\right)^{\alpha},$$

where $h(z)$ is analytic in E with $h(0) = 1$. Differentiating logarithmically, we obtain:

$$\frac{h'(z)}{h(z)} = \frac{f''(z)}{f'(z)} - \alpha\frac{f'(z)}{f(z)} + \frac{\alpha}{z}.$$

It follows easily that:

$$h(z) + \mu\frac{h'(z)}{h(z)} = f'(z)\left(\frac{z}{f(z)}\right)^{\alpha} + \mu\left\{1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1\right\}.$$

Since $f \in N_{\alpha,\mu}(A, B)$, therefore,

$$h(z) + \frac{h'(z)}{(1/\mu)h(z)} \prec \frac{1 + Az}{1 + Bz}.$$

Using Lemma 1 for $\lambda = \frac{1}{\mu}$ and $\gamma = 0$, we obtain:

$$h(z) \prec \frac{1}{(1/\mu)Q(z)} = q(z) \prec \frac{1 + Az}{1 + Bz},$$

where $q(z)$ is the best dominant of (3) and is given by (4). Next we show that $\inf_{|z|<1}\{\operatorname{Re}q(z)\} = q(-1)$. Now if we set $a = \frac{1}{\mu}(B - A)/B$, $b = \frac{1}{\mu}$ and $c = \frac{1}{\mu} + 1$, then it is clear that $c > b > 0$. It follows from (4) for $B \neq 0$ that:

$$Q(z) = (1 + Bz)^{a}\int_{0}^{1} t^{b-1}(1 + Btz)^{-a}\,dt.$$

By using Lemma 3, we get:

$$Q(z) = \frac{\Gamma(b)}{\Gamma(c)}\,_2F_1\left(1, a, c; \frac{Bz}{Bz + 1}\right). \tag{6}$$

To prove that $\inf_{|z|<1}\{\operatorname{Re}q(z)\} = q(-1)$, we need to show that:

$$\operatorname{Re}\{1/Q(z)\} \geq 1/Q(-1).$$

Since $A < -\mu B$ with $-1 \leq B < 0$, this implies that $c > a > 0$ and it follows that:

$$Q(z) = \int_{0}^{1} g(z,t)\,d_{\varepsilon}t,$$

where

$$g(z,t) = \frac{1 + Bz}{1 + (1-t)Bz},$$

$$d_{\varepsilon}t = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}t^{a-1}(1-t)^{c-a-1},$$

which is a positive measure on $[0,1]$. For $-1 \le B < 0$ it is clear that $\operatorname{Re} g(z,t) > 0$ and $g(-r,t)$ is real for $0 \le |z| \le r < 1$ and $t \in [0,1]$. Now using Lemma 2, we obtain:

$$\operatorname{Re}\{1/Q(z)\} \ge 1/Q(-r).$$

Now letting $r \to 1^-$, it follows:

$$\operatorname{Re}\{1/Q(z)\} \ge 1/Q(-1).$$

Further by taking $A \to -\mu B$ for the case $A = -\mu B$ and using (3), we get $N_{\alpha,\mu}(A,B) \subset N(\alpha,\rho)$.

Theorem 2. For $\mu_2 \ge \mu_1 \ge 0$ and $-1 \le B_1 \le B_2 < A_2 \le A_1 \le 1$,

$$N_{\alpha,\mu_2}(A_2,B_2) \subset N_{\alpha,\mu_1}(A_1,B_1).$$

Proof. Let $f \in N_{\alpha,\mu_2}(A_2,B_2)$. Then

$$f'(z)\left(\frac{z}{f(z)}\right)^\alpha + \mu_2\left\{1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1\right\} \prec \frac{1+A_2 z}{1+B_2 z}, \quad z \in E.$$

Since $-1 \le B_1 \le B_2 < A_2 \le A_1 \le 1$, therefore by Lemma 4, we have:

$$f'(z)\left(\frac{z}{f(z)}\right)^\alpha + \mu_2\left\{1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1\right\} \prec \frac{1+A_1 z}{1+B_1 z}, \quad z \in E.$$

Hence we have $f \in N_{\alpha,\mu_2}(A_1,B_1)$. For $\mu_2 = \mu_1 \ge 0$, we have the required result. When $\mu_2 > \mu_1 \ge 0$, Theorem 1 implies that:

$$f'(z)\left(\frac{z}{f(z)}\right)^\alpha \prec \frac{1+A_1 z}{1+B_1 z},$$

Now

$$f'(z)\left(\frac{z}{f(z)}\right)^\alpha + \mu_1\left\{1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1\right\}$$

$$= \left(1 - \frac{\mu_1}{\mu_2}\right)f'(z)\left(\frac{z}{f(z)}\right)^\alpha + \frac{\mu_1}{\mu_2}\left[f'(z)\left(\frac{z}{f(z)}\right)^\alpha + \mu_1\left\{1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1\right\}\right].$$

Using Lemma 5, we get the required result.

Theorem 3. Let $f \in N_{\alpha,\mu}(A,B)$, with $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$. Then

$$|a_2| \le \frac{A-B}{(2-\alpha)(1+\mu)}. \tag{7}$$

Proof. Since $f \in N_{\alpha,\mu}(A,B)$, therefore,

$$f'(z)\left(\frac{z}{f(z)}\right)^\alpha + \mu\left\{1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1\right\} \prec \frac{1+Az}{1+Bz}, \quad z \in E.$$

Now using the fact that $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ we obtain:

$$f'(z)\left(\frac{z}{f(z)}\right)^\alpha + \mu\left\{1 + \frac{zf''(z)}{f'(z)} - \alpha\frac{zf'(z)}{f(z)} + \alpha - 1\right\} = 1 + \{a_2(2-\alpha)(1+\mu)\}z + \dots.$$

By a well-known result due to Janowski and Lemma 6, we get:

$$|a_2(2-\alpha)(1+\mu)| \le A - B.$$

$$|a_2| \le \frac{A-B}{(2-\alpha)(1+\mu)}.$$

Therefore, we have the required result.

Theorem 4. Let $f \in N_{\alpha,\mu}(A,B)$. Then E is mapped by f on a domain that contains the disc $|w| < R_{\alpha,\mu}$, where

$$R_{\alpha,\mu} = \frac{(2-\alpha)(1-\mu)}{2(2-\alpha)(1-\mu)+(A-B)}. \tag{8}$$

Proof. Let w_0 be any complex number such that $f(z) \ne w_0$. Then

$$\frac{w_0 f(z)}{w_0 - f(z)} = z + \left(a_2 + \frac{1}{w_0}\right)z^2 + \dots,$$

is univalent in E, so that

$$\left|a_2 + \frac{1}{w_0}\right| \le 2$$

Therefore,

$$\left|\frac{1}{w_0}\right| - |a_2| \le 2$$

Hence,

$$|w_0| \ge \frac{(2-\alpha)(1-\mu)}{2(2-\alpha)(1-\mu)+(A-B)} = R_{\alpha,\mu}.$$

Theorem 5. Let $f \in N_{\alpha,\mu}(A,B)$ and of the form (1). Then

$$|a_3 - ta_2^2| \le \begin{cases} \dfrac{B_2}{2\beta}\left(2 - \dfrac{2\delta B_1 + \gamma B_2 + 2t\beta B_2}{\delta}\right), & t \le \sigma_1, \\[3mm] \dfrac{B_2}{\beta}, & \sigma_1 \le t \le \sigma_2, \\[3mm] \dfrac{B_2}{2\beta}\left(-2 + \dfrac{2\delta B_1 + \gamma B_2 + 2t\beta B_2}{\delta}\right), & \sigma_2 \le t, \end{cases}$$

where

$$\sigma_1 = -\frac{2\delta B_1 + \gamma B_2}{2\beta B_2},$$

$$\sigma_2 = \frac{4\delta - 2\delta B_1 - \gamma B_2}{2\beta B_2},$$

$$\sigma_3 = \frac{2\delta - 2\delta B_1 - \gamma B_2}{2\beta B_2},$$

$$B_1 = 1 + B,$$
$$B_2 = A - B,$$
$$\beta = (1 + 2\mu)(3 - \alpha),$$
$$\gamma = \alpha(\alpha - 3) + 2\mu(\alpha - 4),$$
$$\delta = (2 - \alpha)^2 (1 + \mu)^2.$$

Further, if $\sigma_1 \leq t \leq \sigma_3$, then

$$| a_3 - t a_2^2 | + \frac{1}{\beta B_2} \left(\frac{2\delta B_1 + \gamma B_2 + 2t\beta B_2}{2} \right) | a_2^2 | \leq \frac{B_2}{\beta}.$$

If $\sigma_3 \leq t \leq \sigma_2$, then

$$| a_3 - t a_2^2 | + \frac{1}{\beta B_2} \left(\frac{4\delta - 2\delta B_1 - \gamma B_2 - 2t\beta B_2}{2} \right) | a_2^2 | \leq \frac{B_2}{\beta}.$$

These results are sharp.

Proof. Since $f \in N_{\alpha,\mu}(A, B)$, therefore, we have:

$$f'(z)\left(\frac{z}{f(z)} \right)^\alpha + \mu \left\{ 1 + \frac{zf''(z)}{f'(z)} - \alpha \frac{zf'(z)}{f(z)} + \alpha - 1 \right\} \prec \frac{1 + Az}{1 + Bz}.$$

Now we can get after simple calculations:

$$f'(z)\left(\frac{z}{f(z)} \right)^\alpha + \mu \left\{ 1 + \frac{zf''(z)}{f'(z)} - \alpha \frac{zf'(z)}{f(z)} + \alpha - 1 \right\}$$
$$= 1 + (2 - \alpha)(1 + \mu) a_2 z + \qquad\qquad (9)$$
$$\left\{ \alpha(\alpha - 3) + 2\mu(\alpha - 4)\frac{a_2^2}{2} + (3 - \alpha)(1 + 2\mu) a_3 \right\} z^2 + \dots$$

Let $p(z) \prec \dfrac{1 + Az}{1 + Bz}$. Then

$$p(z) = \frac{(1 - A) + (1 + A) p_0(z)}{(1 - B) + (1 + B) p_0(z)}, \quad p_0 \in P,$$

where P is the well-known class of functions with positive real part. This implies that:

$$p(z) = 1 + \left(\frac{A - B}{2} \right) p_1 z + \left(\frac{A - B}{2} \right) \left\{ p_2 - \frac{1 + B}{2} p_1^2 \right\} z^2 + \dots, \qquad (10)$$

where

$$p_0(z) = 1 + \sum_{n=1}^{\infty} p_n z^n.$$

From (9) and (10) after comparing the coefficients of z and z^2, we obtain:

$$a_2 = \frac{(A - B) p_1}{2(2 - \alpha)(1 + \mu)},$$

$$a_3 = \frac{1}{(1 + 2\mu)(3 - \alpha)} \left[\left(\frac{A - B}{2} \right) \left\{ p_2 - \frac{1 + B}{2} p_1^2 \right\} - \left\{ \alpha(\alpha - 3) + 2\mu(\alpha - 4) \right\} \frac{a_2^2}{2} \right].$$

This implies that:

$$| a_3 - ta_2^2 |= \frac{A-B}{2(1+2\mu)(3-\alpha)} | p_2 - vp_2^2 |,$$

where

$$v = \frac{2(2-\alpha)^2(1+\mu)^2(1+B)+(A+B)\{\alpha(\alpha-3)+2\mu(\alpha-4)\}+2t(1+2\mu)(3-\alpha)(A-B)}{4(2-\alpha)^2(1+\mu)^2}.$$

Now using Lemma 7, we obtain the required result. Equality can be attained by the functions $F(z)$, defined as follows:

$$F'(z)\left(\frac{z}{F(z)}\right)^\alpha + \mu\left\{1+\frac{zF''(z)}{F'(z)}-\alpha\frac{zF'(z)}{F(z)}+\alpha-1\right\}=\frac{1+Az}{1+Bz}, \text{ if } t<\sigma_1, \text{ or } t>\sigma_2,$$

$$F'(z)\left(\frac{z}{F(z)}\right)^\alpha + \mu\left\{1+\frac{zF''(z)}{F'(z)}-\alpha\frac{zF'(z)}{F(z)}+\alpha-1\right\}=\frac{1+Az^2}{1+Bz^2}, \text{ if } \sigma_1<t<\sigma_2,$$

$$F'(z)\left(\frac{z}{F(z)}\right)^\alpha + \mu\left\{1+\frac{zF''(z)}{F'(z)}-\alpha\frac{zF'(z)}{F(z)}+\alpha-1\right\}=\frac{1+A\phi(z)}{1+B\phi(z)}, \text{ if } t=\sigma_1,$$

and

$$F'(z)\left(\frac{z}{F(z)}\right)^\alpha + \mu\left\{1+\frac{zF''(z)}{F'(z)}-\alpha\frac{zF'(z)}{F(z)}+\alpha-1\right\}=\frac{1-A\phi(z)}{1-B\phi(z)}, \text{ if } t=\sigma_2,$$

where $\phi(z)=\frac{z(z+\eta)}{1+\eta z}$ with $0\le\eta\le1$.

Theorem 6. Let $f\in N_{\alpha,\mu}(A,B)$ and of the form (1). Then for a complex number t,

$$| a_3 - ta_2^2 |\le \frac{B_2}{\beta}\max\left\{1,\left|-\frac{2\delta B_1+\gamma B_2+2t\beta B_2}{2\delta}+1\right|\right\}.$$

By using Lemma 8, we have the required result. Equality can be attained by the function:

$$F'(z)\left(\frac{z}{F(z)}\right)^\alpha + \mu\left\{1+\frac{zF''(z)}{F'(z)}-\alpha\frac{zF'(z)}{F(z)}+\alpha-1\right\}=\frac{1+Az}{1+Bz},$$

or

$$F'(z)\left(\frac{z}{F(z)}\right)^\alpha + \mu\left\{1+\frac{zF''(z)}{F'(z)}-\alpha\frac{zF'(z)}{F(z)}+\alpha-1\right\}=\frac{1+Az^2}{1+Bz^2}.$$

REFERENCES

1. M. Obradovic, "A class of univalent functions", *Hokkaido Math. J.*, **1998**, *27*, 329-335.

2. N. Tuneski and M. Darus, "Feketo-szego functional for non-Bazilevic functions", *Acta Math. Acad. Paedagog. Nyiregyhaz.*, **2002**, *18*, 63-65.

3. Z. Wang, C. Gao and M. Liao, "On certain generalized class of non-Bazilevic functions", *Acta Math. Acad Paedagog. Nyiregyhaz.*, **2005**, 21, 147-154.

4. S. S. Miller and P. T. Mocanu, "Univalent solutions of Briot-Bouquet differential equations", *J. Diff. Eqns.*, **1985**, *56*, 297-309.

5. D. R. Wilken and J. Feng, "A remark on convex and starlike functions", *J. London. Math. Soc.*, **1980**, *21*, 287-290.

6. E. T. Whittaker and G. N. Watson, "A Course of Modern Analysis", 4th Edn. Cambridge University Press, **1927**, pp.281-301.

7. M.-S. Liu, "On a subclass of p-valent close-to-convex functions of order β and type α", *J. Math. Study*, **1997**, *30*, 102-104 (in Chinese).

8. M.-S. Liu, "On certain subclass of analytic functions", *J. South China Normal Univ.*, **2002**, *4*, 15-20 (in Chinese).

9. W. Rogosinski, "On the coefficients of subordinate functions", *Proc. London. Math. Soc.*, **1945**, *48*, 48-82.

10. W. Ma and D. Minda, "A unified treatment of some special classes of univalent functions", Proceedings of the Conference on Complex Analysis, **1992**, Tianjin, China, pp.157-169.

11. V. Ravichandran, A. Gangadharan and M. Darus, "Fekete Szego inequality for certain class of Bazilevic functions", *Far East J. Math. Sci.*, **2004**, *15*, 171-180.

Full Paper

Reachability analysis of a class of Petri nets using place invariants and siphons

Xiu Yan Zhang [1], Zhi Wu Li [1,*], Chun Fu Zhong[1] and Abdulrahman M. Al-Ahmari [2]

[1] School of Electro-Mechanical Engineering, Xidian University, No.2 South Taibai Road, Xi'an 710071, China

[2] Department of Industrial Engineering, College of Engineering, King Saud University, P. O. Box 800, Riyadh 11421, Saudi Arabia

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

Abstract: This paper proposes a novel and computationally efficient approach to deal with the reachability problem by using place invariants and strict minimal siphons for a class of Petri nets called pipe-line nets (PLNs). First, in a PLN with an appropriate initial marking, the set of invariant markings and the set of strict minimal siphons are enumerated. Then a sufficient and necessary condition is developed to decide whether a marking is spurious by analysing the number of tokens in operation places of any strict minimal siphon and their bounds. Furthermore, an algorithm that generates the reachable markings by removing all the spurious markings from the set of invariant markings is proposed. Finally, experimental results show the efficiency of the proposed method.

Keywords: Petri nets, strict minimal siphons, place invariants, reachability analysis, flexible manufacturing system

---

INTRODUCTION

A flexible manufacturing system (FMS) is an automatically running system that consists of resources such as machines, robots, buffers, and conveyors. In an FMS, part processing sequences are executed concurrently, which have to compete for the limited system resources. This competition can cause deadlocks when some processes keep waiting indefinitely for other processes to release resources [1]. Deadlocks must be considered in FMSs since they may offset the advantages of these systems and even lead to catastrophic results such as long downtime and low use of some critical and expensive resources. Therefore, it is necessary to ensure that deadlocks will never occur in such a system.

To deal with deadlock problems in FMSs, Petri nets [2-6], automata [7-8], and graph theory [1] are major mathematical tools. Many researchers use Petri nets as a formalism to deal with deadlock problems [9-14]. There are mainly three approaches: deadlock detection and recovery [15-16], deadlock avoidance [17-19] and deadlock prevention [1, 9, 20, 21].

For Petri nets, there are two widely used analysis techniques for deadlock prevention in FMSs: structure analysis [4, 10, 11, 15, 22, 23] and reachability graph analysis [24-27]. The former always derives a deadlock prevention policy by structural objects of Petri nets, such as siphons and resource-transition circuits. The policy is often simple but always restricts the behaviour of a system in the sense that a part of permissive behaviour is excluded. Therefore, it is suboptimal in general. The latter, the reachability graph analysis, can obtain a liveness-enforcing supervisor with highly permissive or even maximally permissive behaviour. However, its computation is always expensive, which always suffers from a state explosion problem since it requires an enumeration of all or a part of reachable markings. Thus, to tackle this problem, it is necessary to explore more efficient approaches to compute reachable markings.

This paper proposes a novel approach to compute the set of reachable markings using P-invariants and strict minimal siphons in a class of Petri nets called PLNs. First, the set of invariant markings and the set of strict minimal siphons of a PLN are enumerated. As known, the set of invariant markings include spurious markings. Then a sufficient and necessary condition to identify the spurious markings is established. Finally the reachability set of the net is generated by removing all the spurious markings from the set of invariant markings.

PRELIMINARIES

Basics of Petri nets

A Petri net [2] is a four-tuple $N = (P, T, F, W)$ where P and T are finite and non-empty sets. P is a set of places and T is a set of transitions with $P \cap T = \phi$. $F \subseteq (P \times T) \cup (T \times P)$ is called the flow relation of the net, represented by arcs with arrows from places to transitions or from transitions to places. $W : (P \times T) \cup (T \times P) \to \text{IN}$ is a mapping that assigns a weight to an arc: $W(x, y) > 0$ if $(x, y) \in F$, and $W(x, y) = 0$ otherwise, where $x, y \in P \cup T$ and $\text{IN} = \{0, 1, 2, ...\}$ is the set of non-negative integers. $N = (P, T, F, W)$ is said to be ordinary if $\forall (x, y) \in F$, $W(x, y) = 1$. A net $N = (P, T, F, W)$ is pure (self-loop free) if $\forall x, y \in P \cup T$, $W(x, y) > 0$ implies $W(y, x) = 0$. A pure net $N = (P, T, F, W)$ can be represented by its incidence matrix $[N]$, where $[N]$ is a $|P| \times |T|$ integer matrix with $[N](p, t) = W(t, p) - W(p, t)$.

A marking M of a Petri net N is a mapping from P to IN. $M(p)$ denotes the number of tokens in place p. A place p is marked at M if $M(p) > 0$. A subset $S \subseteq P$ is marked at M if at least one place in S is marked at M. The sum of tokens in all places in S is denoted by $M(S)$, i.e. $M(S) = \sum_{p \in S} M(p)$. S is said to be empty or unmarked at M if $M(S) = 0$. M_0 is called an initial marking of N and (N, M_0) is called a net system or marked net.

Let $x \in P \cup T$ be a node of net N. The preset of node x is defined as ${}^\bullet x = \{y \in P \cup T \mid (y, x) \in F\}$, while the postset of x is defined as $x^\bullet = \{y \in P \cup T \mid (x, y) \in F\}$. These notations can be extended to a set of nodes as follows: given $X \subseteq P \cup T$, ${}^\bullet X = \bigcup_{x \in X} {}^\bullet x$ and $X^\bullet = \bigcup_{x \in X} x^\bullet$. For $t \in T$, $p \in {}^\bullet t$ is called an input place of t and $p \in t^\bullet$ is called an output place of t. For $p \in P$, $t \in {}^\bullet p$ is called an input transition of p and $t \in p^\bullet$ is called an output

transition of p. A state machine is an ordinary Petri net satisfying $|{}^{\bullet}t| = |t^{\bullet}| = 1$, $\forall t \in T$. A marked graph is an ordinary Petri net satisfying $|{}^{\bullet}p| = |p^{\bullet}| = 1$, $\forall p \in P$. A sequence of nodes $x_1 \ldots x_i \ldots x_n$ is called a path of N if $\forall i \in \{1, 2, \ldots, n-1\}$, $x_{i+1} \in x_i^{\bullet}$, where $x_i \in P \cup T$. An elementary path from x_1 to x_n is a path whose nodes are all different (perhaps, except for x_1 and x_n). It is called a circuit if $x_1 = x_n$. A Petri net N is said to be strongly connected if there is a sequence of nodes $x, a, b, \ldots, c, y$ in N for $\forall x, y \in P \cup T$ such that (x, a), (a, b), $\ldots$, $(c, y) \in F$, where $\{a, b, \ldots, c\} \subseteq P \cup T$.

A transition $t \in T$ is enabled at M if $\forall p \in {}^{\bullet}t$, $M(p) \geq W(p, t)$. This fact is denoted as $M[t\rangle$. Firing it yields a new marking M' such that $\forall p \in P$, $M'(p) = M(p) - W(p, t) + W(t, p)$, denoted as $M[t\rangle M'$. M' is called an immediately reachable marking from M. The marking M'' is said to be reachable from the marking M if there exists a sequence of enabled transitions $\sigma = t_0 t_1 \ldots t_n$ and markings $M_1, M_2, \ldots$, and M_n such that $M[t_0\rangle M_1[t_1\rangle M_2 \ldots M_n[t_n\rangle M''$ holds, which is denoted as $M[\sigma\rangle M''$. The set of markings reachable from M_0 by firing any possible sequence of transitions in N is called the reachability set of Petri net (N, M_0) and is denoted by $R(N, M_0)$. A reachability graph is a directed graph whose nodes are markings in $R(N, M_0)$ and arcs are labelled by the transitions of N. An arc from M_1 to M_2 is labelled by t if $M_1[t\rangle M_2$. N' is the reverse net of N obtained by reversing the direction of all arcs in N with the initial marking unchanged.

A P-vector is a column vector $I : P \to \mathbf{Z}$ indexed by P and a T-vector is a column vector $J : T \to \mathbf{Z}$ indexed by T, where $\mathbf{Z}$ is the set of integers. P-vector I is called a P-invariant (place invariant) if $I \neq \mathbf{0}$ and $I^T[N] = \mathbf{0}^T$. T-vector J is called a T-invariant (transition invariant) if $J \neq \mathbf{0}$ and $[N]J = \mathbf{0}$. A P-invariant I is said to be a P-semiflow if every element of I is non-negative. $\|I\| = \{p \mid I(p) \neq 0\}$ is called the support of I. I is called a minimal P-invariant if $\|I\|$ is not a superset of the support of any other one and its components are mutually prime. Let I be a P-invariant of (N, M_0) and M be a reachable marking from M_0. Then $I^T M = I^T M_0$.

Let X be a matrix where each column is a P-semiflow of the net (N, M_0) and $I_X(N, M_0) = \{M \in \mathrm{IN}^{|P|} \mid X^T M = X^T M_0\}$ denotes the set of invariant markings, where $\mathrm{IN}^{|P|}$ is a set of non-negative vectors, each of which has a length of $|P|$. It can be noted that $R(N, M_0) \subseteq I_X(N, M_0)$.

A non-empty set $S \subseteq P$ is a siphon if ${}^{\bullet}S \subseteq S^{\bullet}$. $S \subseteq P$ is a trap if $S^{\bullet} \subseteq {}^{\bullet}S$. A siphon is minimal if there is no siphon contained in it as a proper subset. A minimal siphon is said to be strict if ${}^{\bullet}S \subset S^{\bullet}$.

S³PR Nets

Definition 1 [9]. A simple sequential process (S²P) is a Petri net $N = (P_A \cup \{p^0\}, T, F)$, where the following statements are true: (1) $P_A \neq \phi$ is called a set of operation places; (2) $p^0 \notin P_A$ is called the process idle place; (3) N is a strongly connected state machine; and (4) every circuit of N contains place p^0.

Let $N = (P,T,F)$ be an S^{2}P with idle process place p^0. Let C be a circuit of N, and x and y be two nodes of C. Node x is said to be previous to y if there exists a path in C from x to y, the length of which is greater than one and does not pass over the idle place p^0. This fact is denoted by $x <_C y$. Let x and y be two nodes in N. Node x is said to be previous to y in N if there exists a circuit C such that $x <_C y$. This fact is denoted by $x <_N y$.

Definition 2 [9]. A system of simple sequential processes with resources (S^{3}PR) $N = (P^0 \cup P_A \cup P_R, T, F)$ is defined as the union of a set of nets $N_i = (\{p_i^0\} \cup P_{A_i} \cup P_{R_i}, T_i, F_i)$ sharing common places, where the following statements are true:

(1) p_i^0 is called the process idle places of N_i. Elements in P_{A_i} and P_{R_i} are called operation places and resource places respectively;

(2) $P_{R_i} \neq \phi$; $P_{A_i} \neq \phi$; $p_i^0 \notin P_{A_i}$; $(P_{A_i} \cup \{p_i^0\}) \cap P_{R_i} = \phi$;

(3) $\forall p \in P_{A_i}$, $\forall t \in {}^\bullet p$, $\forall t' \in p^\bullet$, $\exists r_p \in P_{R_i}$, ${}^\bullet t \cap P_{R_i} = t'^\bullet \cap P_{R_i} = \{r_p\}$;

(4) $\forall r \in P_{R_i}$, ${}^{\bullet\bullet}r \cap P_{A_i} = r^{\bullet\bullet} \cap P_{A_i} \neq \phi$ and ${}^\bullet r \cap r^\bullet = \phi$;

(5) ${}^{\bullet\bullet}(p_i^0) \cap P_{R_i} = (p_i^0)^{\bullet\bullet} \cap P_{R_i} = \phi$;

(6) N_i' is a strongly connected state machine, where $N_i' = (P_{A_i} \cup \{p_i^0\}, T_i, F_i)$ is the resulting net after the places in P_{R_i} and related arcs are removed from N_i. Every circuit of N_i' contains place p_i^0;

(7) any two N_i's are composable when they share a set of common places. Every shared place must be a resource place; and

(8) transitions in $(p_i^0)^\bullet$ and ${}^\bullet(p_i^0)$ are called source and sink transitions of an S^{3}PR respectively.

In an S^{3}PR, P^0 is called the set of process idle places, P_A is called the set of operation places and P_R is called the set of resource places.

LS^{3}PR Nets

Definition 3 [28]. An S^{3}PR $N = (P,T,F)$ is called a linear S^{3}PR (LS^{3}PR) if

(1) $P = P^0 \cup P_A \cup P_R$ is a partition of places, where

(1.a) $P^0 = \{p_1^0, p_2^0, \ldots, p_k^0\}$, $k > 0$,

(1.b) $P_A = \bigcup_{i=1}^{k} P_{A_i}$, where $P_{A_i} \cap P_{A_j} = \phi$, for all $i \neq j$,

(1.c) $P_R = \{r_1, r_2, \ldots, r_n\}$, $n > 0$;

(2) $T = \bigcup_{i=1}^{k} T_i$, where $T_{A_i} \cap T_{A_j} = \phi$, for all $i \neq j$;

(3) $\forall i \in \{1, 2, \ldots, k\}$, the subnet N_i generated by $\{p_i^0\} \cup P_{A_i} \cup T_i$, is a strongly connected state machine such that every cycle of N_i contains place $\{p_i^0\}$ and $\forall p \in P_{A_i}$, $|p^\bullet| = 1$;

(4) $\forall i \in \{1, 2, \ldots, k\}$, $\forall p \in P_{A_i}$, ${}^{\bullet\bullet}p \cap P_R = p^{\bullet\bullet} \cap P_R$ and $|{}^{\bullet\bullet}p \cap P_R| = 1$; and

(5) N is strongly connected.

Definition 4 [28]. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be an LS^{3}PR. Given $p \in P_A$, if ${}^{\bullet\bullet}p \cap P_R = \{r_p\}$, r_p is called the resource used by p. For $r \in P_R$, $H(r) = {}^{\bullet\bullet}r \cap P_A$ is called the set of holders of r.

Definition 5 [28]. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be an LS^{3}PR. An initial marking M_0 is called an admissible initial marking for N if

(1) $M_0(p^0) \geq 1, \forall p^0 \in P^0$;

(2) $M_0(p) = 0, \forall p \in P_A$; and

(3) $M_0(r) \geq 1, \forall r \in P_R$.

REACHABILITY ANALYSIS FOR A PIPE-LINE NET (PLN)

Definition of a PLN

Definition 6. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be an LS^3PR and N_i be the subnet generated by $\{p_i^0\} \cup P_{A_i} \cup T_i$. The two subnets N_i and N_j are said to be mutually reversed if $\exists r, r_1 \in P_R$ ($r \neq r_1$) such that one of the two following statements holds: 1) $p_1 <_{N_i} p$ and $q <_{N_j} q_1$; and 2) $p <_{N_i} p_1$ and $q_1 <_{N_j} q$, where $p, p_1 \in P_{A_i}$, $q, q_1 \in P_{A_j}$, $i \neq j$, $p, q \in H(r)$, and $p_1, q_1 \in H(r_1)$.

The net shown in Figure 1 is an LS^3PR with $P_0 = \{p_1, p_7, p_{11}\}$, $P_A = \{p_2, p_3, p_4, p_5, p_6, p_8, p_9, p_{10}, p_{11}, p_{12}, p_{13}\}$ and $P_R = \{p_{14}, p_{15}, p_{16}, p_{17}, p_{18}\}$. It consists of three subnets: N_1 generated by $\{p_1\} \cup \{p_2, p_3, p_4, p_5, p_6\} \cup \{t_1, t_2, t_3, t_4, t_5, t_6\}$ N_2 generated by $\{p_7\} \cup \{p_8, p_9, p_{10}\} \cup \{t_7, t_8, t_9, t_{10}\}$ and N_3 generated by $\{p_{11}\} \cup \{p_{12}, p_{13}\} \cup \{t_{11}, t_{12}, t_{13}\}$. In the net, $p_2, p_3, p_4, p_5, p_6 \in P_{A_1}$, $p_8, p_9, p_{10} \in P_{A_2}$, $p_{12}, p_{13} \in P_{A_3}$, $p_2, p_8 \in H(p_{14})$, $p_3, p_9 \in H(p_{15})$, $p_4, p_{10} \in H(p_{16})$, $p_5, p_{12} \in H(p_{17})$ and $p_6, p_{13} \in H(p_{18})$. Since $p_2 <_{N_1} p_3$ and $p_9 <_{N_2} p_8$; $p_3 <_{N_1} p_4$ and $p_{10} <_{N_2} p_9$; and $p_2 <_{N_1} p_4$ and $p_{10} <_{N_2} p_8$, the two subnets N_1 and N_2 are mutually reversed. Since $p_5 <_{N_1} p_6$ and $p_{12} <_{N_3} p_{13}$, the two subnets N_1 and N_3 are not mutually reversed.
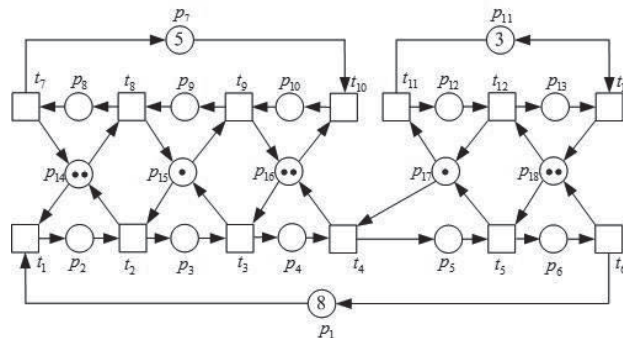
Figure 1. A Petri net model [28]

Definition 7. An LS^3PR $N = (P^0 \cup P_A \cup P_R, T, F)$ is called a PLN if

(1) $\forall r \in P_R$, $\forall i \in \{1, 2, \ldots, k\}$, $|H(r)| = 2$ and $|H(r) \cap P_{A_i}| \leq 1$; and

(2) $q_1 \in q^{\bullet\bullet} \cap P_{A_j}$ holds if the two subnets N_i and N_j are mutually reversed with $p, p_1 \in P_{A_i}$, $q, q_1 \in P_{A_j}$, $p_1 \in^{\bullet\bullet} p \cap P_{A_i}$ $p, q \in H(r)$, $p_1, q_1 \in H(r_1)$, $r, r_1 \in P_R$ and $r \neq r_1$.

Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be a PLN with $t \in T$. $^{(p)}t$ and $t^{(p)}$ denote the sets of input and output operation places of t respectively, and $^{(r)}t$ and $t^{(r)}$ denote the sets of input and output resources of t respectively. Hence $^{\bullet}t =^{(p)} t \cup^{(r)} t$ and $t^{\bullet} = t^{(p)} \cup t^{(r)}$.

As shown in Figure 1, $|H(p_{14})| = 2$, $|H(p_{15})| = 2$, $|H(p_{16})| = 2$, $|H(p_{17})| = 2$ and $|H(p_{18})| = 2$. Also we have $|H(p_{14}) \cap P_{A_1}| = 1$, $|H(p_{14}) \cap P_{A_2}| = 1$, $|H(p_{14}) \cap P_{A_3}| = 0$, $|H(p_{15}) \cap P_{A_1}| = 1$, $|H(p_{15}) \cap P_{A_2}| = 1$, $|H(p_{15}) \cap P_{A_3}| = 0$, $|H(p_{16}) \cap P_{A_1}| = 1$, $|H(p_{16}) \cap P_{A_2}| = 1$, $|H(p_{16}) \cap P_{A_3}| = 0$, $|H(p_{17}) \cap P_{A_1}| = 1$, $|H(p_{17}) \cap P_{A_2}| = 0$, $|H(p_{17}) \cap P_{A_3}| = 1$, $|H(p_{18}) \cap P_{A_1}| = 1$, $|H(p_{18}) \cap P_{A_2}| = 0$ and $|H(p_{18}) \cap P_{A_3}| = 1$. Moreover, in the two subnets N_1 and N_2,

$p_2, p_8 \in H(p_{14})$, $\qquad$ $p_3, p_9 \in H(p_{15})$, $\qquad$ $p_2 \in {}^{\bullet\bullet} p_3 \cap P_{A_1}$, $\qquad$ $p_8 \in p_9{}^{\bullet\bullet} \cap P_{A_2}$, $\qquad$ $p_3, p_9 \in H(p_{15})$,

$p_4, p_{10} \in H(p_{16})$, $p_3 \in {}^{\bullet\bullet} p_4 \cap P_{A_1}$ and $p_9 \in p_{10}{}^{\bullet\bullet} \cap P_{A_2}$. Therefore, the net in Figure 1 is a PLN.

A Sufficient and Necessary Condition for Reachability in a PLN

Definition 8. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be a PLN with an admissible initial marking M_0. The maximum number of tokens in place p is called the bound of place p, denoted by b_p. That is to say, $b_p = \max\{M(p) \mid M \in R(N, M_0)\}$.

Lemma 1 [28]. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be an LS^{3}PR. The set of minimal P-semiflows of N is $I = I_R \cup I_{SM}$, where $I_R = \bigcup_{r \in P_R} H(r) \cup \{r\}$ and $I_{SM} = \bigcup_{i \in \{1,\dots,k\}} P_{A_i} \cup \{p_i^0\}$.

Lemma 2. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be a PLN with an admissible initial marking M_0. Then $\forall p^0 \in P^0$, $b_{p^0} = M_0(p^0)$; $\forall r \in P_R$, $b_r = M_0(r)$; and $\forall p \in P_A$, $b_p = M_0(r_p)$.

Proof. From Lemma 1, the set of minimal P-semiflows of a PLN consists of two subsets. The first corresponds to the token conservation law associated with resources. $\forall r \in P_R$, the P-semiflow $I_r = H(r) \cup \{r\}$, states that for each reachable marking M, the token conservation law $\sum_{p \in H(r)} M(p) + M(r) = M_0(r)$ is true. The second subset is associated with the token conservation law for each state machine (in the sense of processes). $\forall i \in \{1,\dots,k\}$, $p_i^0 \in P^0$, the P-semiflow $I_{SM_i} = P_{A_i} \cup \{p_i^0\}$, establishes the invariant relation $\sum_{p \in P_{A_i}} M(p) + M(p_i^0) = M_0(p_i^0)$ for each reachable marking M. Taking into account of $M \geq 0$, it is easy to see that $\forall p^0 \in P^0$, $b_{p^0} = M_0(p^0)$; $\forall r \in P_R$, $b_r = M_0(r)$; and $\forall p \in P_A$, $b_p = M_0(r_p)$.

Definition 9. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be a PLN. A circuit C that contains resources and transitions only is called a resource-transition circuit if ${}^{(r)}C_T = C_T{}^{(r)} = C_R$, where C_R and C_T denote the sets of all resources and transitions of C respectively.

Definition 10. Let $C(r_1, t_1, r_2, t_2, \dots, r_m, t_m)$ be a resource-transition circuit in a PLN, where 1) $m \geq 2$; 2) $\forall i \in \{1, 2, \dots, m\}$, $r_i \in {}^\bullet t_i$; 3) $\forall i \in \{2, \dots, m\}$, $r_i \in t_{i-1}{}^\bullet$; and 4) $r_1 \in t_m{}^\bullet$. r_i is called a connected resource if there exists $r_i = r_j$ in $C(r_1, t_1, r_2, t_2, \dots, r_m, t_m)$, where $i, j \in \{1, 2, \dots, m\}$ and $i \neq j$.

As shown in Figure 1, there are three resource-transition circuits in the net: $C_1 = C(p_{14}, t_8, p_{15}, t_2)$, $C_2 = C(p_{15}, t_9, p_{16}, t_3)$ and $C_3 = C(p_{14}, t_8, p_{15}, t_9, p_{16}, t_3, p_{15}, t_2)$ with ${}^{(r)}C_{1_T} = C_{1_T}{}^{(r)} = \{p_{14}, p_{15}\} = C_{1_R}$, ${}^{(r)}C_{2_T} = C_{2_T}{}^{(r)} = \{p_{15}, p_{16}\} = C_{2_R}$ and ${}^{(r)}C_{3_T} = C_{3_T}{}^{(r)} = \{p_{14}, p_{15}, p_{16}\} = C_{3_R}$. p_{15} appears twice in C_3. Hence, p_{15} is a connected resource.

Theorem 1 [11]. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be a PLN with an admissible initial marking M_0. $S = S_A \cup S_R$ is a strict minimal siphon of N if

1) $S_A \neq \phi$, $S_R \neq \phi$;

2) $S_R = C_R$, where C_R is the set of all resources of a resource-transition circuit C in N; and

3) $S_A = \{p \mid p \in \bigcup_{r \in C_R} H(r) \wedge (p^{\bullet\bullet} \cap (P_A \cup P^0)) \varnothing \bigcup_{r \in C_R} H(r)\}$.

Corollary 1. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be a PLN and $S = S_A \cup S_R$ be a strict minimal siphon of N. Then $|S_A| = 2$.

Proof. From Theorem 1, we have $S_R = C_R$ and $S_A = \{p \mid p \in \bigcup_{r \in C_R} H(r) \wedge (p^{\bullet\bullet} \cap (P_A \cup P^0)) \emptyset \bigcup_{r \in C_R} H(r)\}$. We accordingly have the following two cases:

(1) $|S_R| = 2$. By the definition of resource-transition circuits, there necessarily exist two transitions in C_T that is the set of all the transitions of resource-transition circuit C associated with S. By the definition of a PLN, the two transitions in C_T necessarily belong to two subnets that are mutually reversed. Since $S_A = \{p \mid p \in \bigcup_{r \in C_R} H(r) \wedge (p^{\bullet\bullet} \cap (P_A \cup P^0)) \emptyset \bigcup_{r \in C_R} H(r)\}$, $\forall p \in S_A$, $p \in C_T^{\bullet}$ holds. Therefore, $|S_A| = 2$ is true.

(2) $|S_R| > 2$. By the definition of resource-transition circuits, there necessarily exist more than two transitions in C_T. By the definition of a PLN, the transitions in C_T necessarily belong to two subnets that are mutually reversed. Since $S_A = \{p \mid p \in \bigcup_{r \in C_R} H(r) \wedge (p^{\bullet\bullet} \cap (P_A \cup P^0)) \emptyset \bigcup_{r \in C_R} H(r)\}$, $\forall p \in S_A$, $p \in C_T^{\bullet}$ and $p \notin H(r_c)$ hold, where r_c is a connected resource in S_R. Therefore, $|S_A| = 2$ is true.

Definition 11. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be a PLN. An initial marking M_0 is called an appropriate initial marking of N if

(1) $M_0(p^0) \geq 1, \forall p^0 \in P^0$;

(2) $M_0(p) = 0, \forall p \in P_A$; and

(3) $\forall r \in P_R$; if r is a connected resource, $M_0(r) = 1$, otherwise $M_0(r) \geq 1$.

Definition 12. The markings in the set of invariant markings $I_X(N, M_0)$ that are not in the reachability set $R(N, M_0)$ are called spurious markings.

A backward firing in N is equivalent to a forward firing in the reverse net N'[30]. This implies that the directed path in the reachability graph of N' from M' to M is just the reverse path in the reachability graph of N from M to M'. Similarly, a spurious marking in N does not have directed paths from reachable markings and the corresponding marking in N' does not have directed paths to reachable markings.

Theorem 2. Let $N = (P^0 \cup P_A \cup P_R, T, F)$ be a PLN with an appropriate initial marking M_0, $\prod$ be the set of strict minimal siphons, and S be a strict minimal siphon in N. A marking M is spurious in the set of invariant markings of N if $\forall S \subseteq \prod$, $M(S_A) = \sum_{p \in S_A} b_p$, where S_A is the set of operation places of S and b_p is the bound of place p.

Proof. 1. We first prove the sufficiency. By Definition 8, $\forall p \in S_A$, $b_p = \max\{M(p) \mid M \in R(N, M_0)\}$ holds. $\forall S \subseteq \prod$, $M(S_A) = \sum_{p \in S_A} b_p$ means that the number of tokens in each operation place of any strict minimal siphon S reaches its bound at marking M. We have to prove that M is a spurious marking. That is to say, we need to show that there exist no directed paths from initial marking M_0 to M in N. By Corollary 1, $|S_A| = 2$. Without loss of

generality, let $S_A = \{p_1, p_2\}$. From Theorem 1, $|S_R| \geq 2$ holds. We accordingly have the following two cases.

(1) $|S_R| = 2$. From Theorem 1, $S_R = \{r_{p_1}, r_{p_2}\}$ holds. From the proof of Corollary 1, there necessarily exist two transitions in C_T that is the set of all transitions of resource-transition circuit C associated with S, and $\forall p \in S_A$, $p \in C_T^\bullet$ holds. Without loss of generality, let $C_T = \{t_1, t_2\}$, $t_1 \in {}^\bullet p_1$, and $t_2 \in {}^\bullet p_2$.

By contradiction, suppose that M is reachable from M_0 in N with $M(p_1) = M_0(r_{p_1})$ and $M(p_2) = M_0(r_{p_2})$. By Definition 11, $M_0(p_1) = 0$ and $M_0(p_2) = 0$ hold. According to the token conservation law, $M(r_{p_1}) = 0$ and $M(r_{p_2}) = 0$ hold. Since M is reachable from M_0 in N, there necessarily exists a reachable marking M' in the reachability graph of N such that $M'[t_1\rangle M$ or $M'[t_2\rangle M$ holds. This implies that t_1 or t_2 must be enabled at M in the reverse net N' such that $M[t_1\rangle M'$ or $M[t_2\rangle M'$. By the definition of a PLN and Definition 10, $r_{p_1} \in t_2^\bullet$ and $r_{p_2} \in t_1^\bullet$ hold in N. This implies that $r_{p_1} \in {}^\bullet t_2$ and $r_{p_2} \in {}^\bullet t_1$ hold in the reverse net N'. Therefore, both t_1 and t_2 are disabled at M in the reverse net N', which contradicts that t_1 or t_2 is enabled at M in the reverse net N'. Thus, M is a spurious marking in N.

(2) $|S_R| > 2$. From Theorem 1, there necessarily exist connected resources in S_R. We denote R_C as the set of connected resources in S_R. From the proof of Corollary 1, there necessarily exist more than two transitions in C_T, and $\forall p \in S_A$, $p \in C_T^\bullet$ and $p \notin H(r_c)$ hold, where $r_c \in R_C$. Let $t_1 \in {}^\bullet p_1$ and $t_2 \in {}^\bullet p_2$. From Theorem 1, $r_{p_1} \notin R_C$ and $r_{p_2} \notin R_C$ hold. Since M_0 is an appropriate initial marking, $\forall r_c \in R_C$, $M_0(r_c) = 1$ holds.

By contradiction, suppose that M is reachable from M_0 in N with $M(p_1) = M_0(r_{p_1})$ and $M(p_2) = M_0(r_{p_2})$. From the proof of case (1), since there exist connected resources in S_R, $\forall r_c \in R_C$, $M(r_c) = 1$ may hold according to the token conservation law. Therefore, t_1 or t_2 may be enabled at M in the reverse net N'. Similarly, there exist a sequence of transitions σ in $R_C^\bullet$, which may be enabled from M in N' such that $M[\sigma\rangle M''$ with $M''(r_{p_2}) = 0$. Therefore, the transition $t \in r_{p_2}^\bullet$ must be disabled at M''. That is to say, M does not have directed path to M_0 in N'. This implies M_0 does not have directed path to M in N, which contradicts that M is reachable in N. Therefore, we can conclude that M is a spurious marking.

2. We prove the necessity. By contradiction, suppose that $\forall S \subseteq \Pi$, $M(S_A) \neq \sum_{p \in S_A} b_p$. Since M is a marking in the set of invariant markings of N, the token conservation law $\sum_{p \in H(r)} M(p) + M(r) = M_0(r)$ is true. From Lemma 2, $\forall p \in S_A$, $b_p = M_0(r_p)$ holds. Note that $M \geq \mathbf{0}$. We can conclude that $\forall p \in S_A$, $M(p) \leq b_p$. Therefore, $M(S_A) < \sum_{p \in S_A} b_p$ holds.

$M(S_A) < \sum_{p \in S_A} b_p$ means that at marking M the number of tokens in each operation place of S does not reach its bound at the same time. From the proof of the sufficiency, we can similarly prove that M is reachable from M_0, which contradicts that M is a spurious marking. Therefore, if M is a spurious marking in the set of invariant markings of N, $\forall S \subseteq \Pi$, $M(S_A) = \sum_{p \in S_A} b_p$ holds.

An Algorithm Computing the Set of Reachable Markings

From Theorem 2, given a PLN with an appropriate initial marking M_0, all the spurious markings can be identified from the set of invariant markings. Then the set of reachable markings can be calculated by removing all the spurious markings from the set of invariant markings. An algorithm to compute the set of markings reachable from M_0 is presented as follows:

Algorithm 1. Computation of reachable markings

Input: a PLN model (N, M_0).

Output: The reachability set $R(N, M_0)$.

1) Check if N is a PLN and M_0 is an appropriate initial marking. If not, exit.

2) Compute the set of minimal P-semiflows by Lemma 1.

3) Compute the bounds of all the places by Lemma 2. According to the token conservation law, enumerate the set of invariant markings $I_X(N, M_0) = \{M \in \mathrm{IN}^{|P|} \mid X^T M = X^T M_0\}$, where X is a matrix whose column each is a P-semiflow of the net (N, M_0), and $\mathrm{IN}^{|P|}$ is a set of non-negative vectors with a length of $|P|$.

4) Compute the set C of resource-transition circuits by Definition 9.

5) Compute the set of strict minimal siphons $\prod$ due to Theorem 1.

6) **if** $\{\prod = \phi\}$ **then** $R(N, M_0) = I_X(N, M_0)$.

   else $R(N, M_0) = I_X(N, M_0) \setminus \{M \mid \forall S \subseteq \prod, M(S_A) = \sum_{p \in S_A} b_p\}$.

7) Output $R(N, M_0)$.

8) End.

AN EXAMPLE

To practically test the efficiency of the proposed method, a C program has been developed, which implements the algorithm and runs on a Windows XP operating system with Intel CPU Core 2.60 GHz and 3 GB memory.

Take the net in Figure 1 as an example. There are 18 places and 13 transitions. It is a PLN and M_0 is an appropriate initial marking. By Lemma 1, the net has eight minimal P-semiflows as follows: $I_{SM_1} = \{p_1, p_2, p_3, p_4, p_5, p_6\}$, $I_{SM_2} = \{p_7, p_8, p_9, p_{10}\}$, $I_{SM_3} = \{p_{11}, p_{12}, p_{13}\}$, $I_{r_1} = \{p_2, p_8, p_{14}\}$, $I_{r_2} = \{p_3, p_9, p_{15}\}$, $I_{r_3} = \{p_4, p_{10}, p_{16}\}$, $I_{r_4} = \{p_5, p_{12}, p_{17}\}$ and $I_{r_5} = \{p_6, p_{13}, p_{18}\}$.

By Lemma 2, the bounds of all the places are as follows: $b_{p_1} = 8$, $b_{p_2} = 2$, $b_{p_3} = 1$, $b_{p_4} = 2$, $b_{p_5} = 1$, $b_{p_6} = 2$, $b_{p_7} = 5$, $b_{p_8} = 2$, $b_{p_9} = 1$, $b_{p_{10}} = 2$, $b_{p_{11}} = 3$, $b_{p_{12}} = 1$, $b_{p_{13}} = 2$, $b_{p_{14}} = 2$, $b_{p_{15}} = 1$, $b_{p_{16}} = 2$, $b_{p_{17}} = 1$ and $b_{p_{18}} = 2$. By the definition of the set invariant markings $I_X(N, M_0)$, we can obtain $I_X(N, M_0)$ that has 1944 markings.

By Definition 8, the net has three resource-transition circuits: $C_1 = C(p_{14}, t_8, p_{15}, t_2)$, $C_2 = C(p_{15}, t_9, p_{16}, t_3)$ and $C_3 = C(p_{14}, t_8, p_{15}, t_9, p_{16}, t_3, p_{15}, t_2)$. By Theorem 1 we can find that the net correspondingly has three strict minimal siphons: $S_1 = \{p_3, p_8, p_{14}, p_{15}\}$, $S_2 = \{p_4, p_9, p_{15}, p_{16}\}$ and $S_3 = \{p_4, p_8, p_{14}, p_{15}, p_{16}\}$.

By Theorem 2, a marking M with $M(p_3) = 2$ and $M(p_8) = 1$ is spurious in $I_X(N, M_0)$, a marking M' with $M'(p_4) = 2$ and $M'(p_9) = 1$ is spurious, and a marking M'' with $M''(p_4) = 2$ and $M''(p_8) = 2$ is also spurious. We impose the constraints on $I_X(N, M_0)$ as

follows: $\quad M(p_4)+M(p_9)<3$, $\quad M(p_4)+M(p_9)<3 \quad$ and $\quad M(p_4)+M(p_8)<4$, $\quad$ where $M \in I_X(N,M_0)$. Then the reachability set $R(N,M_0)$ that has 1710 markings is generated by removing 234 spurious ones from $I_X(N,M_0)$.

The software package INA2003 [30] can also compute the reachability set. For comparison, reachability analysis of the Petri net in Figure 1 is conducted by the use of INA. The tool generates the reachability set consisting of 1710 markings, which are in agreement with the markings in the reachability set $R(N,M_0)$ generated by the proposed method, validating the correctness of the proposed algorithm.

TINA [31] is a toolbox for editing and analyzing Petri nets, which can also compute the reachability set. For comparison, reachability analysis of the Petri net in Figure 1 is conducted by the use of TINA. The toolbox generates a reachability set consisting of 1710 markings, which are also in agreement with the markings in the reachability set $R(N,M_0)$ generated by the proposed method.

EXPERIMENTAL RESULTS

The net structure in Figure 1 is selected for experimental studies. We vary the initial markings of resource places p_{14}, p_{15}, p_{16}, p_{17} and p_{18}, and idle places p_1, p_7 and p_{12}. Table 1 shows various parameters in the net, where the first column represents the initial tokens in places p_1, p_7, p_{12}, p_{14}, p_{15}, p_{16}, p_{17} and p_{18}. N_I, N_S and N_R indicate the numbers of invariant markings, spurious markings and reachable markings respectively. The fifth column shows the total CPU time for computing $R(N,M_0)$ by using the proposed method. The sixth and the last columns show the total CPU time for computing $R(N,M_0)$ by using INA and the total CPU time for computing $R(N,M_0)$ by using TINA for comparison purpose respectively.

Table 1. Parameters in the model depicted in Figure 1 with varying markings

p_1, p_7, p_{12}, p_{14}, p_{15}, p_{16}, p_{17}, p_{18}	N_I	N_S	N_R	CPU time (s)	INA time (s)	TINA time (s)
8, 5, 3, 2, 1, 2, 1,2	1,944	234	1,710	<1	<1	<1
15, 9, 6, 4, 1, 4, 2, 4	60,750	2,790	57,960	<1	216	3
29, 17, 12, 8, 1, 8, 4, 8	4,100,625	61,425	4,039,200	22	>7200	—
36, 21, 15, 10, 1, 10, 5, 10	18,112,248	184,338	17,927,910	398	—	—
43, 25, 18, 12, 1, 12, 6, 12	63,299,964	466,284	62,833,680	1432	—	—
50, 29, 21, 14, 1, 14, 7, 14	186,624,000	1,041,120	185,582,880	5365	—	—

As shown in Table 1, we can see that the proposed method becomes more efficient with the increase of the initial markings. Note that "—" in Table 1 means that the computation cannot be finished with a reasonable time or memory is overflowed.

CONCLUSIONS

The set of reachable markings play an important role in the deadlock control in Petri nets. This paper presents a novel approach in computing the set of reachable markings using P-invariants and strict minimal siphons without the construction of reachability graph that often makes the analysis intractable. The method is applied to a small class of Petri nets called PLNs that are a subclass of LS^{3}PR. Experimental results show its efficiency via studying a number of examples. Future work should extend the method in this paper to more general classes of Petri nets.

REFERENCES

1. M. P. Fanti and M. C. Zhou, "Deadlock control methods in automated manufacturing systems", *IEEE Trans. Syst. Man Cybern. A* , **2004**, *34(1)*, 5-22.
2. T. Murata, "Petri nets: Properties, analysis and applications", *Proc. IEEE*, **1989**, *77*, 541-580.
3. Z. W. Li and M. C. Zhou, "Deadlock Resolution in Automated Manufacturing Systems: A Novel Petri Net Approach", Springer, London, **2009**.
4. D. Y. Chao, "Improvement of suboptimal siphon and FBM-based control model of a well-known S^{3}PR", *IEEE Trans. Autom. Sci. Eng.*, **2011**, *8*, 404-411.
5. N. Q. Wu and M. C. Zhou, "System Modeling and Control with Resource-Oriented Petri Nets", CRC Press, New York, **2010**.
6. K. Xing, M. C. Zhou, H. Liu and F. Tian, "Optimal Petri-net-based polynomial-complexity deadlock-avoidance policies for automated manufacturing systems", *IEEE Trans. Syst. Man Cybern. A* , **2009**, *39*, 188-199.
7. P. J. Ramadge and W. M. Wonham, "Supervisory control of a class of discrete event processes", *SIAM J. Control Optim.*, **1987**, *25*, 206-230.
8. P. J. G. Ramadge and W. M. Wonham, "The control of discrete event systems", *Proc. IEEE*, **1989**, *77*, 81-98.
9. J. Ezpeleta, J. M. Colom and J. Martinez, "A Petri net based deadlock prevention policy for flexible manufacturing systems", *IEEE Trans. Robot. Autom.*, **1995**, *11*, 173-184.
10. Z. W. Li and M. C. Zhou, "Elementary siphons of Petri nets and their application to deadlock prevention in flexible manufacturing systems", *IEEE Trans. Syst. Man Cybern. A*, **2004**, *34*, 38-51.
11. Z. W. Li and M. C. Zhou, "Two-stage method for synthesizing liveness-enforcing supervisors for flexible manufacturing systems using Petri nets", *IEEE Trans. Ind. Inform.*, **2006**, *2*, 313-325.
12. Z. W. Li, M. C. Zhou and M. D. Jeng, "A maximally permissive deadlock prevention policy for FMS based on Petri net siphon control and the theory of regions", *IEEE Trans. Autom. Sci. Eng.*, **2008**, *5*, 182-188.

13. Y. F. Chen, Z. W. Li, M. Khalgui and O. Mosbahi, "Design of a maximally permissive liveness-enforcing Petri net supervisor for flexible manufacturing systems", *IEEE Trans. Autom. Sci. Eng.*, **2011**, *8*, 374-393.

14. K. Xing, M. C. Zhou, F. Wang, H. Liu and F. Tian, "Resource-transition circuits and siphons for deadlock control of automated manufacturing systems", *IEEE Trans. Syst. Man Cybern. A*, **2011**, *41*, 74-84.

15. T. K. Kumaran, W. Chang, H. Cho and R. A. Wysk, "A structured approach to deadlock detection, avoidance and resolution in flexible manufacturing systems", *Int. J. Prod. Res.*, **1994**, *32*, 2361-2379.

16. R. A. Wysk, N. S. Yang and S. Joshi, "Resolution of deadlocks in flexible manufacturing systems: Avoidance and recovery approaches", *J. Manuf. Syst.*, **1994**, *13*, 128-138.

17. I. B. Abdallah and H. A. ElMaraghy, "Deadlock prevention and avoidance in FMS: A Petri net based approach", *Int. J. Adv. Manuf. Technol.*, **1998**, *14*, 704-715.

18. J. Ezpeleta and L. Recalde, "A deadlock avoidance approach for nonsequential resource allocation systems", *IEEE Trans. Syst. Man Cybern. A*, **2004**, *34*, 93-101.

19. N. Q. Wu and M. C. Zhou, "Modeling and deadlock avoidance of automated manufacturing systems with multiple automated guided vehicles", *IEEE Trans. Syst. Man Cybern. B*, **2005**, *35*, 1193-1202.

20. K. Y. Xing, B. S. Hu and H. X. Chen, "Deadlock avoidance policy for Petri-net modeling of flexible manufacturing systems with shared resources", *IEEE Trans. Autom. Control*, **1996**, *41*, 289-295.

21. D. Liu, Z. W. Li and M. C. Zhou, "Liveness of an extended S^{3}PR", *Automatica*, **2010**, *46*, 1008-1018.

22. Y. S. Huang, M. D. Jeng, X. L. Xie and S. L. Chung, "Deadlock prevention policy based on Petri nets and siphons", *Int. J. Prod. Res.*, **2001**, *39*, 283-305.

23. H. S. Hu, M. C. Zhou and Z.W. Li, "Liveness enforcing supervision of video streaming systems using nonsequential Petri nets", *IEEE Trans. Multimed.*, **2009**, *11*, 1457-1465.

24. A. Ghaffari, N. Rezg and X. L. Xie, "Design of a live and maximally permissive Petri net controller using the theory of regions", *IEEE Trans. Robot. Autom.*, **2003**, *19*, 137-141.

25. M. Uzam and M. C. Zhou, "An improved iterative synthesis method for liveness enforcing supervisors of flexible manufacturing systems", *Int. J. Prod. Res.*, **2006**, *44*, 1987-2030.

26. M. Uzam and M. C. Zhou, "An iterative synthesis approach to Petri net-based deadlock prevention policy for flexible manufacturing systems", *IEEE Trans. Syst. Man Cybern. A*, **2007**, *37*, 362-371.

27. Y. F. Chen and Z. W. Li, "Design of a maximally permissive liveness-enforcing supervisor with a compressed supervisory structure for flexible manufacturing systems", *Automatica*, **2011**, *47*, 1028-1034.

28. J. Ezpeleta, F. García-Vallés and J. M. Colom, "A class of well structured Petri nets for flexible manufacturing systems", *Lect. Notes Comput. Sci.*, **1998**, *1420*, 64-83.

29. J. Júlvez, L. Recalde and M. Silva, "On reachability in autonomous continuous Petri net systems", *Lect. Notes Comput. Sci.,* **2003**, *2679*, 221-240.

30. P. H. Starke, "INA: Integrated Net analyzer", http://www2.informatik.hu-berlin.de/~starke/ina.html (Accessed: July 2013).

31. B. Berthomieu, F. Vernadat, and S. D. Zilio, "TINA: Time Petri Net Analyzer", http://www.laas.fr/tina (Accessed: April 2012).

Maejo International
Journal of Science and Technology

Communication

On the Diophantine equation $2^x + 11^y = z^2$

Somchit Chotchaisthit

Department of Mathematics, Faculty of Science, Khon Kaen University, Khon Kaen, 40002, Thailand

E-mail: somchit@kku.ac.th

Abstract: In this paper it is shown that (3,0,3) is the only non-negative integer solution of the Diophantine equation $2^x + 11^y = z^2$.

Keywords: exponential Diophantine equation, Catalan's conjecture

INTRODUCTION

Solving Diophantine equations of the form $2^x + p^y = z^2$, where p is prime, has been widely studied by many mathematicians. For example, Acu [1] proved in 2007 that $(x, y, z) = (3,0,3), (2,1,3)$ are the only two non-negative solutions for the case $p = 5$. On the other hand, it was shown in 2010 by Suvarnamani et al. [2] that there are no non-negative solution (x, y, z), with x even, for the case $p = 7, 11$. The complete sets of non-negative solutions to some Diophantine equations of similar forms were also studied by Sándor in 2002 [3].

Later in 2011, Suvarnamani [4] published a paper on finding non-negative solutions to the Diophantine equations of the form $2^x + p^y = z^2$ for every prime p. Having this result, it might seem at first glance that the problem of solving the Diophantine equations of such form did come to an end. Nevertheless, for the case where x is even (i.e. $x = 2m$ for some integer m), these equations can be rewritten as $4^m + p^y = z^2$, whose complete set of non-negative solutions for each prime p was readily found by the author [5]. In addition, for the case where x is odd, Suvarnamani's proof [4] unfortunately contains a misleading argument which significantly affects its correctness: it was stated (p.1417, line 14-15) that

$$(z - 2^{k+\frac{1}{2}})(z + 2^{k+\frac{1}{2}}) = p^y,$$

for some integer k, p, y, z, and thus,

$$z - 2^{k+\frac{1}{2}} = p^u \text{ and } z + 2^{k+\frac{1}{2}} = p^{y-u},$$

for some integer $u \geq 0$. This is clearly absurd since $z \pm 2^{k+\frac{1}{2}}$ is irrational while p^u, p^{y-u} are integers. Since no extra information can be obtained from Suvarnamani's proof, solving the Diophantine equation of the form $2^x + p^y = z^2$, where p is prime, now remains an open problem.

Inspired by all the aforementioned results, this paper therefore aims to study the Diophantine equation $2^x + p^y = z^2$, particularly where $p = 11$, in detail. To be precise, the objective is to show that $(x, y, z) = (3,0,3)$, which clearly satisfies such equation, is its only non-negative solution.

MAIN RESULTS

In this study, Catalan's conjecture [6], which states that the only solution in integers $a > 1$, $b > 1, x > 1, y > 1$ of the equation $a^x + b^y = 1$ is $(a, b, x, y) = (3,2,2,3)$, is used.

In the main theorem, the Diophantine equation $2^x + 11^y = z^2$ is considered.

Theorem. *The Diophantine equation* $2^x + 11^y = z^2$ *has only one solution in non-negative integer, namely* $(x, y, z) = (3,0,3)$.

Proof. Consider the following cases:

Case 1: $x = 0$. It can be easily checked that if the Diophantine equation $1 + 11^y = z^2$ has a solution, then $y \geq 2$ and z is an even integer greater than 3. Thus,

$$11^y = z^2 - 1 = (z+1)(z-1).$$

Then there are non-negative integers α, β such that $11^\alpha = z+1$, $11^\beta = z-1$, $\alpha > \beta$ and $\alpha + \beta = y$. Therefore,

$$11^\beta(11^{\alpha-\beta} - 1) = 11^\alpha - 11^\beta = (z+1) - (z-1) = 2,$$

which implies that $\beta = 0$ and $11^\alpha - 1 = 2$. This contradicts the fact that α is a non-negative integer. Therefore, the Diophantine equation $1 + 11^y = z^2$ has no solution.

Case 2: $x = 1$. If the Diophantine equation $2 + 11^y = z^2$ has a solution, then this implies that $z^2 \equiv 2 \pmod{11}$ has a solution. However, it is easy to check that $z^2 \equiv 2 \pmod{11}$ does not have a solution, a contradiction. Thus, the Diophantine equation $2 + 11^y = z^2$ also has no solution.

Case 3: $x \geq 2$. Therefore $2^x \equiv 0 \pmod 4$ and $z^2 \equiv 0, 1 \pmod 4$, which implies that $11^y \equiv 0, 1 \pmod 4$. It can be observed that if y is an odd non-negative integer, then $11^y \equiv 3 \pmod 4$. Therefore, y is an even non-negative integer. Let $y = 2k$ for some non-negative integer k. Thus,

$$2^x = z^2 - 11^{2k} = (z + 11^k)(z - 11^k).$$

Then there are non-negative integers α, β such that $2^\alpha = z + 11^k$, $2^\beta = z - 11^k$ with $\alpha > \beta$ and $\alpha + \beta = x$. Therefore,

$$2^\beta(2^{\alpha-\beta} - 1) = 2^\alpha - 2^\beta = (z + 11^k) - (z - 11^k) = 2(11^k).$$

This implies that $\beta = 1$ and

$$2^{\alpha-1} - 1 = 11^k. \tag{*}$$

So $z = 11^k + 2$ and $\alpha > \beta = 1$. By Catalan's conjecture, $2^{\alpha-1} - 1 = 11^k$ has no solution only when $\alpha - 1 > 1$ and $k > 1$. Thus, it suffices to consider only the case when $\alpha - 1 \leq 1$ or $k \leq 1$, i.e. $\alpha = 2$ or $0 \leq k \leq 1$. From (*), it is easy to see that $\alpha = 2$ if and only if $k = 0$. This implies

that $(x, y, z) = (3, 0, 3)$. Finally, one can easily check that $2^{\alpha-1} - 1 = 11^k$ dose not have a solution when $k = 1$.

It is easy to check that $(x, y, z) = (3, 0, 3)$ is the only non-negative integer solution of $2^x + 11^y = z^2$. This finishes the proof. □

OPEN PROBLEM

It is to be noted that all finding of the solutions of Diophantine equation in the case of $2^x + p^y = z^2$ where p is prime in general is still an open problem. For example, it is not known how to find all non-negative integer solutions of $2^x + p^y = z^2$ where p = 7, 13, 29, 37 or 257.

ACKNOWLEDGEMENTS

REFERENCES

1.   D. Acu, "On a diophantine equation $2^x + 5^y = z^2$", *Gen. Math.*, **2007**, *15*, 145-148.
2.   A. Suvarnamani, A. Singta and S. Chotchaisthit, "On two Diophantine equations $4^x + 7^y = z^2$ and $4^x + 11^y = z^2$", *Sci. Techno. RMUTT J.*, **2011**, *1*, 25-28.
3.   J. Sandor, "Geometric Theorems, Diophantine Equations, and Arithmetic Functions", American Research Press, Rehoboth, **2002**, pp.89-92.
4.   A. Suvarnamani, "Solutions of the diophantine equations $2^x + p^y = z^2$", *Int. J. Math. Sci. Appl.*, **2011**, *1*, 1415-1419.
5.   S. Chotchaisthit, "On the diophantine equation $4^x + p^y = z^2$ where p is a prime number", *Amer. J. Math. Sci.*, **2012,** *1*, 191-193.
6.   P. Mihăilescu, "Primary cyclotomic units and a proof of Catalan's conjecture". *J. Reine Angew. Math.* **2004**, *572*, 167-195.

Maejo International
Journal of Science and Technology

Review

Automated microaneurysm detection algorithms applied to diabetic retinopathy retinal images

Akara Sopharak [1],*, Bunyarit Uyyanonvara [2] and Sarah Barman [3]

[1] Faculty of Science and Arts, Burapha Univeristy, Chanthaburi Campus, 57 Moo1 Kamong, Thamai, Chanthaburi 22170, Thailand

[2] Department of Information Technology, Sirindhorn International Institute of Technology (SIIT), Thammasat University, 131 Moo 5, Tiwanont Road, Bangkadi, Muang, Pathumthani 12000, Thailand

[3] School of Computing and Information Systems, Kingston University, Penrhyn Road, Kingston upon Thames, Surrey KT1 2EE, UK

* Corresponding author, e-mail: akara@buu.ac.th

Abstract: Diabetic retinopathy is the commonest cause of blindness in working age people. It is characterised and graded by the development of retinal microaneurysms, haemorrhages and exudates. The damage caused by diabetic retinopathy can be prevented if it is treated in its early stages. Therefore, automated early detection can limit the severity of the disease, improve the follow-up management of diabetic patients and assist ophthalmologists in investigating and treating the disease more efficiently. This review focuses on microaneurysm detection as the earliest clinically localised characteristic of diabetic retinopathy, a frequently observed complication in both Type 1 and Type 2 diabetes. Algorithms used for microaneurysm detection from retinal images are reviewed. A number of features used to extract microaneurysm are summarised. Furthermore, a comparative analysis of reported methods used to automatically detect microaneurysms is presented and discussed. The performance of methods and their complexity are also discussed.

Keywords: microaneurysm, diabetic retinopathy, retinal image, automated detection

INTRODUCTION

Diabetic retinopathy (DR) is a severe and widespread eye disease which can be regarded as a manifestation of diabetes on the retina. It is a major public health problem and it remains the leading cause of blindness in people of working age (20-65 years). After duration of 10 years,

around 7% of people with diabetes will have developed retinopathy, rising to 90% after 25 years. DR is a leading cause of blindness both in the United States and Asia. The global prevalence of diabetes among adults aged 20 years or more in 2000 was around 171 million (2.8 % of the world population) and is expected to rise to 366 million (4.4% of the estimated world population) by the year 2030 [1, 2]. The increasing number of individuals with diabetes worldwide suggests that diabetic retinopathy will continue to be a major contributor to vision loss and associated functional impairment. People with untreated diabetes are said to be 25 times more at risk for blindness than the general population and 2% of diabetic patients will become blind. The growing numbers of diabetic patients will increase the pressure on available infrastructure and resources [3].

Diabetic retinopathy occurs when the increased glucose level in the blood damages the capillaries. It is characterised and graded by the development of retinal microaneurysms (MAs), haemorrhages and exudates. MAs are focal dilations of retinal capillaries and appear as small, round, dark red dots. MAs are swellings of the capillaries caused by a weakening of the vessel wall. In retinal photographs, although the capillaries are not visible, MAs appear as dark red isolated dots. Haemorrhages occur when blood leaks from the retinal vessels and appear as round small red dots or blots indistinguishable from MAs. Exudates occur when proteins or lipids leak from blood vessels and appear yellowish in colour. It is difficult to detect MAs as their pixels are similar to that of blood vessels. MAs are hard to distinguish from noise or background variations because of a typically low contrast.

With a large number of patients, the number of ophthalmologists is often not sufficient to cope, especially in rural areas or if the workload of local ophthalmologists is substantial. Early screening for diabetic retinopathy could improve the prognosis of proliferative retinopathy and reduce risk factors to lower the rate of blindness. The two treatments for diabetic retinopathy are laser and vitrectomy surgery. Even though there are treatments for diabetic retinopathy, they cannot restore lost vision. Thus, early screening for diabetic retinopathy is the best way to prevent further vision loss [4-7].

In this paper we are concerned with methods that will automatically detect MAs as the earliest clinically localised characteristic of DR [8]. Automatic MA detection can assist ophthalmologists in preventing and treating the disease more efficiently. The MA detection can be used to grade the progression of DR into four stages: no DR, mild DR, moderate DR and severe DR, as shown in Table 1. This paper reviews automated MA detection methods. The overall procedure described in the following sections is shown in Figure 1.

Table 1. Criteria used for grading diabetic retinopathy [9]

DR stage	
Grade 0 (no DR)	$MA = 0$ and $H = 0$
Grade 1 (mild)	$1 \leq MA \leq 5$ and $H = 0$
Grade 2 (moderate)	$5 < MA < 15$ or $0 < H \leq 5$
Grade 3 (severe)	$MA \geq 15$ or $H > 5$

Note: MA = microaneurysm, H = haemorrhage

```
┌─────────────────────────┐          ┌─────────────────────────┐
│    Image Acquisition    │ ──────▶  │  Candidate MA Detection │ ──┐
└─────────────────────────┘          └─────────────────────────┘   │
            │                                    │                  │
            ▼                                    ▼                  │
┌─────────────────────────┐          ┌─────────────────────────┐   │
│   Image Pre-processing  │          │   Feature Extraction    │   │
└─────────────────────────┘          └─────────────────────────┘   │
            │                                    │                  │
            ▼                                    ▼                  │
┌─────────────────────────────────┐  ┌─────────────────────────┐   │
│ Vessel/ Optic Disc Detection    │  │    Fine MA Detection    │   │
│        and Elimination          │  └─────────────────────────┘   │
└─────────────────────────────────┘              │                 │
            │                                     ▼                 │
            └──────────────────────▶  ┌─────────────────────────┐   │
                                      │  Performance Evaluation │ ◀─┘
                                      └─────────────────────────┘
```
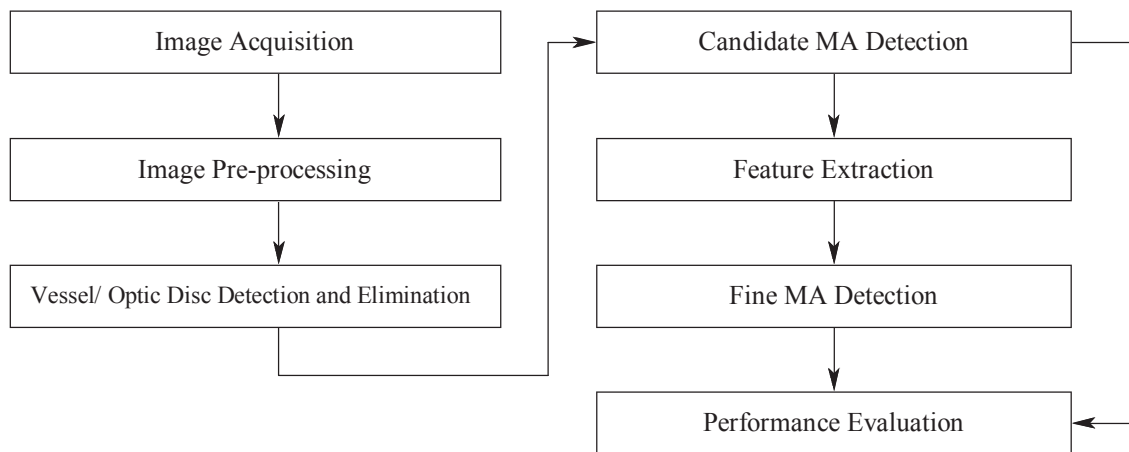
Figure 1. Procedure of MA detection

RETINAL IMAGE ACQUISITION

Digital retinal images are taken by a fundus camera (mydriatic or non-mydriatic). To maximise the view of the fundus, a pupil dilation process must be taken before photographing. Images can be stored in JPEG, GIF or PNG image format files. Digital image acquisition is usually over a large field of view: seven 30°, three wide-angle 60° or nine overlapping 45° fields. To have a great accuracy for the screening of diabetic retinopathy, 45° field fundus photography is recommended. Images are usually rectangular and the largest dimension of the rectangle ranges from 600 to 1024 pixels [3-9]. The angular field of view and the image size can be different depending on the camera setting.

Abnormal diabetic retinopathy on fundus images is shown in Figure 2. A mydriatic fundus image where retinal features are clearly visible is shown in Figure 2(a) while a poor non-mydriatic fundus image is shown in Figure 2(b). Most of the published research in the field used images originating from their own sources while a few researchers used images from existing databases. The available retinal image databases are DRIVE (Digital Retinal Image for Vessel Extraction), STARE (Structured Analysis of Retina) and DIARETDB1. Table 2 shows the databases that provide images containing MAs.

The STARE database [10] provides images with a variety of diagnoses. The images are captured by a TopCon TRV-50 fundus camera with 35 degree field of view. The image size is 605x700 pixels, with 24 bits per pixel. Microaneurysms are categorised into 4 states; these are 'Many', 'Few', 'Absent' and 'Unknown'. The DRIVE database [11] provides retinal images for vessel segmentation. The images are captured by a Canon CR5 non-mydriatic 3CCD camera with a 45-degree field of view. The image size is 565x584 pixels. The DIARETDB1 database [12] provides images for DR detection. Images are classified into 2 types, abnormal (at least mild non-proliferative signs of DR) and normal. Images are captured with a 50-degree field of view.

(a)   (b)
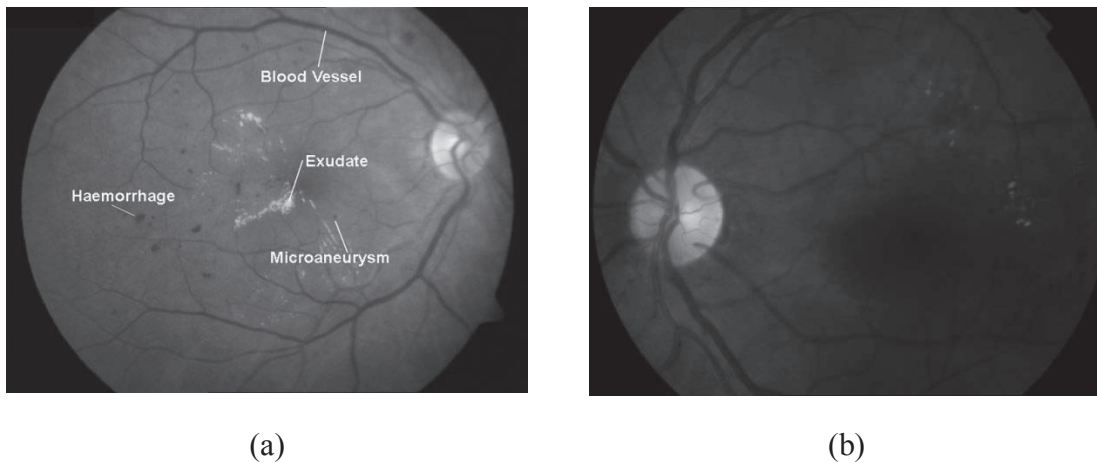
Figure 2. Abnormal diabetic retinopathy images: (a) mydriatic fundus image showing exudates, microaneurysms and haemorrhages, (b) non-mydriatic fundus image

Table 2. Retinal image databases that provide images containing MAs [10, 12]

Database	State	No. of images
STARE	1. Many	32
	2. Few	40
	3. Absent	298
	4. Unknown	27
		397
DIARETDB1	1. Abnormal	84
	2. Normal	5
		89

IMAGE PRE-PROCESSING

The quality of a retinal image has an impact on the performance of lesion detection algorithms. There are many factors that can cause an image to be of poor quality such as low contrast, noise, non-uniform illumination, variation in light reflection and diffusion, difference in retinal pigmentation and differences in cameras. Pre-processing is an important step in order to attenuate such image variations and improve image quality. There are two main categories of pre-processing: correction of non-uniform illumination and enhancement (noise removal, contrast improvement).

A Red-Green-Blue (RGB) image consists of three planes of colour. The green plane is usually used for further analysis since MAs have the highest contrast with the background in this colour plane [13-16].

Shade Correction

To remove non-uniform illumination, a shade correction is applied [9, 14, 15, 17-21]. A shade-corrected image is accomplished by subtracting the background image from the green image. The background image is produced by smoothing the original image with a low-pass filter—a mean or median filter whose size is greater than the largest retinal feature. Spencer et al. [17] and Frame et al. [19] produced the background image by median filtering the green band

image with a 25x25 pixel kernel. The size of filter is wider than the widest blood vessel in their test image set. Lee et al. [15] applied shade-correction using a 56x56 median filter.

Median Filter

To remove noise, the median filter is applied. The median filter replaces the value of a pixel by the median of the gray levels in the neighbourhood pixels. The median filter is much less sensitive than the mean of outliers. Median filtering is better able to remove these outliers without reducing the sharpness of the image. The median filter has a benefit of simultaneously reducing noise and preserving edges. Fleming et al. [16] used a 3x3 median filter to remove salt-and-pepper noise.

Adaptive Contrast Enhancement

Adaptive contrast enhancement was first proposed by Sinthanayothin et al. [5] in order to emphasise features in the retinal image. The mean and variance of the intensity within a sublocal region were considered and the transformation function was applied. Usher et al. [22] applied the local adaptive contrast enhancement on the intensity component of the Hue-Saturation-Intensity (HSI) colour model to enhance contrast and normalise intensity. The limitation of this technique is that it not only adjusts the contrast but also increases noise.

SEGMENTATION OF OTHER RETINAL LANDMARKS

MAs do not appear in a vessel but many MA candidates can be red/dark spots within a retinal vessel. In order to reduce false MA detection, prominent structures within the retinal images such as blood vessels and the optic disc have to be removed to prevent a misclassification.

Vessel Detection

There are a number of proposed vessel detection methods. A top-hat transformation technique was used by Spencer et al. [17], Hipwell et al. [20] and Niemeijer et al. [21]. It is based on morphologically opening with a structuring element at a different orientation. Niemeijer et al. [21] used a transform with a linear structuring element size of 9 pixels. A vasculature map was obtained by taking the maximum pixel value at each pixel location in all 12 opened images. Adaptive thresholding was used by Zhang et al. [14] to locate vessels as shown in Figure 3. Sinthanayothin et al. [5] used the mean of a multilayer perceptron neural network to identify blood vessels. Raman et al. [23] used a Gaussian matched filter to detect and remove vessels from the image.

Optic Disc Detection

The optic disc is one of the landmark features in a retinal image. It is important to detect the location of the optic disc in retinal image analysis. To prevent the optic disc from interfering with MA detection, the optic disc is detected and eliminated from the image and any further consideration. Principle component analysis was used by Li and Chutatape [24] and Osareh et al. [25]. Lalonde et al. [26] used pyramidal decomposition and Hausdorff-based template matching guided by scale tracking of large objects using multi-resolution image decomposition. This method is effective but rather complex. The Hough transform was used by Ege et al. [4], Zheng et al. [7] and Lowell et al. [27] for localising the optic disc. The method presents an optic disc

detection using an active contour which is a model-based method for localising and tracking image structures.

(a)   (b)
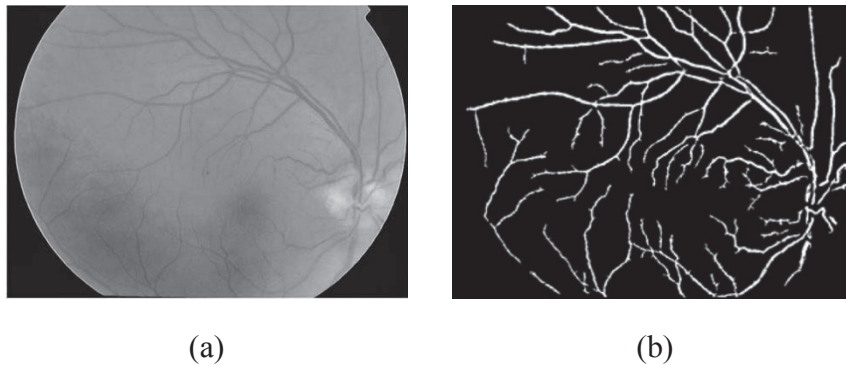
Figure 3. Results of vessel detection: (a) retinal image (b) detected vessels [14]

ALGORITHM FOR MICROANEURYSM DETECTION

In the literature, several algorithms have been employed for MA detection. Previously published methods for MA detection have been shown to work on fluorescein angiograms [17-19, 28] or colour images [4-6, 9, 14, 20-23, 29, 30-32], in which the MAs and other retinal features are clearly visible. In fluorescein angiogram images the contrast between the microaneurysms and background is greater than in digital colour photographs. However, a frequency rate for mortality of 1:222,000 from 221,781 fluorescein angiograms associated with the intravenous use of fluorescein prohibits the application of this technique for large-scale screening purposes [20].

In this paper several techniques are introduced. These are recursive region growing (RRG), watershed transform, mathematic morphology, multi-scale correlation coefficients, matched filtering, discrimination function, k-NN (nearest neighbour) classifier and neural networks. The algorithms reviewed can be divided into two main approaches: the first approach extracts MAs using a single method while the second approach first roughly detects MAs as candidate MAs. The candidate detection is to identify all possible MA candidates in a retinal image. Feature extraction is then applied and fine MA detection is applied in the last step. A quantitative analysis groups articles according to the methodological approach used for each step in the process. We categorise these methods as follows.

Recursive Region Growing (RRG)

RRG [17] is a straightforward method; one pixel is selected as a seed point and then iteratively grown by comparing pixels in the neighborhood of the region. The difficulty of this technique is the selecting of threshold values, region seed points and stopping criteria. The region growing method cannot segment MAs with dramatic background intensity changes. In addition, the performance of the method depends on the selection of the seed pixels (starting points of region growing).

Spencer et al. [17] applied RRG algorithm on the Gaussian matched filter image. The size, shape and energy features were used to finalise the MA detection. Based on fluorescein angiogram images, the sensitivity and specificity were 82% and 86% respectively, with 100 false positives per image. Only four images were tested. Cree et al. [18] developed the technique of Spencer et al. [17] by including a process for removing the need for operator intervention in selecting regions of interest and another one for registering images to allow sequential

comparison of MAs based on a cross-correlation algorithm on angiogram images. The sensitivity was 82% with 5.7 false positives per image. Sinthanayothin et al. [5] combined RRG with a 'moat operator' to detect MAs. The operator was used to enhance the edges of the red lesions by creating a trough around them. Thresholding was then used to binarise the image. A dataset of 30 images was used. The sensitivity and specificity were 77.5% and 88.7% respectively. Usher et al. [22] extracted candidate MAs using a combination of RRG and adaptive intensity threshold.

Watershed Transform

The watershed transform [33] is a technique that operates on the gray level image to segment regions of the image. It is a method based on the morphological technique. The watershed transform is computed from the gradient of the original image so that the catchment basin (lakes) boundaries are located at high gradient points. The limitations of this technique are over-segmentation, sensitivity to noise and poor detection of thin or low signal-to-noise ratio structures.

Luo et al. [34] performed the watershed transform on the colour difference image to extract MAs. Firstly, the Luv-u plane colour model was considered instead of the RGB colour model as the former showed a high sensitivity to dark objects on retinal images. The 2D histogram distribution on the Luv-u plane was then used to obtain the object-based colour difference image and the watershed was applied in the last step. The performance of the system was not reported.

Mathematic Morphology

Mathematic morphology [17-19] is a technique to extract image components. There are two basic terms: the image and the structuring element. Morphological image processing is performed by sliding a structuring element over an image. Morphology is a technique of image processing based on shapes. The value of each pixel in the output image is based on a comparison of the corresponding pixel in the input image with its neighbours. On the binary image, the basic operations of mathematic morphology are dilation (dilating the boundary), erosion (eroding the boundary), opening (erosion followed by dilation) and closing (dilation followed by erosion). On a grey level image, dilation brightens small dark areas. Erosion darkens small bright areas like noise or small spur. Opening darkens small bright areas and removes small bright spots like noise. Closing brightens small dark areas and removes small dark holes. Niemeijer et al. [21] applied the top-hat transformation by subtracting the image opened by the structuring element from the original image to extract the vasculature map. The vasculature regions were then removed from a shade-corrected image. The matched filter was applied in order to enhance the contrast. The resultant image was then thresholded. The binary pixels were set as a starting point. A region growing algorithm was then applied in the final step for candidate MA regions. The limitation of this technique is on the vessel detection step, in which red lesions which are larger than the linear structuring element cannot be detected. The transformation and threshold was also used by Dupas et al. [9] and Raman et al. [23] to detect candidate MAs.

Multi-Scale Correlation Coefficients (MSCF)

MSCF is used to detect a bright spot in the image [14, 35]. It involves applying a sliding neighbourhood filter with multi-scale Gaussian kernels to the fundus image in order to calculate a correlation coefficient for each pixel. Zhang et al. [14] applied MSCF to detect candidate MAs.

Gaussian kernels with five different scales were applied to the fundus image. The sigmas of the Gaussian function were 1.1, 1.2, 1.3, 1.4 and 1.5. Thresholding was then applied in order to determine the number of MA candidates and region-growing-based algorithms were applied to allow a fit to the true MA size. The result of MA detection using multi-scale correlation coefficients by Zhang et al. [14] is shown in Figure 4.

(a)   (b)   (c)
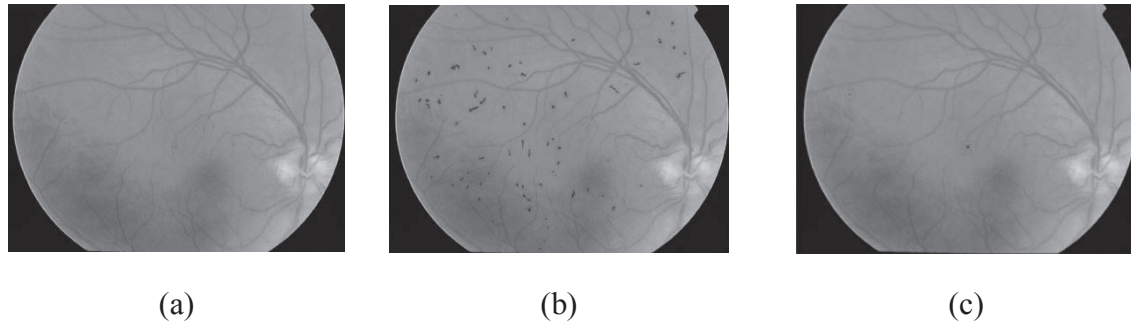
Figure 4. Results of MA detection by Zhang et al. [14]: (a) original retinal image, (b) candidate MAs detected, (c) fine MAs detected

Matched Filtering

The matched filtering [28, 36] is one of the template matching techniques. It is based on the spatial properties of the object to be recognised. Matched filtering convolves a kernel with the retinal image. The kernel is designed to model properties in the image at some unknown position and orientation, and the matched filter response indicates the presence of the feature. As the MA intensity profile has a Gaussian shape, a Gaussian-matched filter is usually used.

Hipwell et al. [20] applied a Gaussian filter to retain candidate MAs. Kahai et al. [37] used the response of the matched filter tuned to detect microaneurysms of size 2x2. A raw image was convolved with the filter and a threshold was set in order to determine the location of the microaneurysms. The size of the filter provided an estimate for the size of the microaneurysms. Two circular symmetric Gaussian kernels (widths = 17 and 13) were applied to the top-hat image. The result was comprised of a candidate MA seed image. A region growing algorithm was applied in the last step [29]. The RRG and adaptive intensity thresholding with addition of a 'moat operator' was used to identify candidate MAs by Sinthanayothin et al. [5].

Discrimination Function

The application of a discrimination function is a statistical analysis. It can be used to determine which variables discriminate between two or more groups. A fine MA detection step was employed by Zhang et al. [14], who used a discrimination function. The candidate MAs were classified into 2 groups: true MA and false MA. The maximum and minimum values of each feature of a true MA was taken and stored in a discrimination table. This table was used to eliminate any candidates whose feature values were greater than the maximum or less than the minimum. The candidate MAs whose feature values were between minimum and maximum were detected as true MAs. The system used 100 images (50 training images and 50 testing images). The performance was shown in terms of false positives per image, which is 0.357. The values of sensitivity and specificity were not presented.

Hipwell et al. [20] used the discrimination function to grade candidate MAs as present or absent. The programme was trained on a set of 102 images and tested on a set of 62 images. The results were compared with two standard photographs according to the EURODIAB IDDM

complication study protocol [38]. The sensitivity and specificity were 81% and 93% respectively. The limitation of this work is that the MA is classified only after haemorrhages are removed.

Raman et al. [23] used the different intensity and shape features that characterise the MAs to distinguish MAs from other objects. The sensitivity and specificity were both 70%. Streeter and Cree [29] located candidate microaneurysms by a standard method that enhances small round features after background intensity correction. Each candidate was then classified based on colour and standard morphological features. Linear discriminant analysis was used and a microaneurysm detection sensitivity of 56% at 5.7 false positives per image was reported.

k-NN Classifier

The k-NN classifier has been broadly used in image analysis applications due to its conceptual simplicity and general applicability. The NN classifier simply classifies a test instance with the class of the k-nearest training instance according to some distance measure. The Mahalanobis and Euclidean distance are usually used. Niemeijer et al. [21] performed a Linear Discriminant classifier, a Quadratic Discriminant classifier and k-NN classifier on candidate MA objects. The k-NN classifier showed the best performance with k=55. A set of 40 images was used to train the classifier in the first candidate object extraction. A set of 100 images was then used to train and test the system. The sensitivity and specificity were 100% and 87% respectively. Dupas et al. [9] used a total of 94 images, of which 68 contained at least 1 MA and were used to train the classifier. They reported 88.47% sensitivity with 2.13 false positives per image.

Neural Network

Application of a neural network is an unsupervised method [30]. A training process is needed. The network is comprised of an input layer, a hidden layer and an output layer. A weight is used to decide the probability of input data belonging to a particular output. The weights between the layers have to be randomly initialised. They are usually set to small values often in the range of [-0.5, 0.5]. Once an input is presented to the neural network and a corresponding desired or target response is set as the output, an error is composed from the difference between the desired response and the real system output. The error information is fed back to the system, which makes all adjustments. The weight is adjusted by training the network with a known output. This process is repeated until the desired output is acceptable. The network is then tested with the data from seen training data and unseen test data.

Usher et al. [22] applied a neural network to candidate MA images with a 500-image training set and a 773-image testing set. The images were classified as having the presence or absence of lesions. The sensitivity and specificity were 95.1% and 46.3% respectively for any exudate and/or haemorrhage/MA detection. A back propagation neural network was applied by Gardner et al. [6] to the fundus images to detect MAs. The weight update method, delta rules, was used in back propagation. The training set comprised 147 lesion images and 32 normal images. The image was divided into a region of 20x20 pixels or 30x30 pixels before training. The sub-images were classified as 'normal without vessel', 'normal vessel', 'exudate' and 'haemorrhage/ microaneurysm.' The sensitivity and specificity were both 73.8 %. Kamel et al. [30] used the learning vector quantisation neural network to detect the MA location in retinal angiograms. A small window of 32x32 pixels was used to train the network. To enhance the performance, a multi-stage training procedure was applied. The error value of 2.67% was

presented instead of accuracy value. The limitation is that the network requires longer training to achieve the desired output.

Miscellaneous

Streeter and Cree [29] located candidate MAs by a standard method that enhances small round features after background intensity correction. Each candidate was then classified based on colour and standard morphological features. An MA detection sensitivity of 56% at 5.7 false positives per image was reported.

Kose et al. [31] proposed the inverse segmentation method. The image along with inverse segmentation was employed to measure and follow up the degeneration in retinal images. The pixel values which have lower intensity than the background image intensity values were segmented as dark lesions. The sensitivity and specificity were 95.1% and 99.3% respectively.

Ege et al. [4] tested the system on 134 retinal images. Naive Bayes, Mahalanobis distance and k-NN classifier were employed on candidate images containing MAs to compare the performance. The Mahalanobis distance performed the best with a sensitivity of 69%.

Walter and Klein [32] proposed a method based on diameter closing and kernel density estimation for automatic classification. The mean of the diameter closing was applied for candidate MA detection. A set of 94 images was tested. The sensitivity was 88.47% with 2.13 false positives per image.

FEATURE EXTRACTION

Features are extracted based on shape, greyscale, pixel intensity, colour intensity, responses of Gaussian filter banks and correlation coefficient values [4, 9, 14, 17, 18, 20-22, 28, 29, 32]. Ege et al. [4] proposed 6 features based on shape and colour. Dupas et al. [9] used a feature based on size, contrast, circularity, grey-scale level and colour. Zhang et al. [14] applied 31 features on true MA classification. Spencer et al. [17] and Frame et al. [19] proposed 13 features. Hipwell et al. [20] applied 13 features based on shape, intensity and size, e.g. circularity, perimeter length and length ratio, but did not clearly clarify the full features list. Niemeijer et al. [21] combined Spencer-Frame features with new proposed features. In total, 68 features were used. Usher et al. [22] extracted features from candidate MAs based on size, shape, hue and intensity. Hafez and Azeem [28] proposed 6 features to extract MAs from fluorescein angiograms. The 19 features based on colour, standard morphology and Fourier descriptors were used by Streeter and Cree [29]. Walter and Klein [32] proposed 15 features based on size, shape, colour and grey level.

Streeter and Cree [29] proposed a feature selection step to select the optimal feature set. The forward-backward feature selection was applied to the full feature set. The forward feature selection starts with an empty data set and proceeds by expanding the data set with the feature, of which addition to the data set increases the performance most. In contrast to the forward feature selection, the backward feature selection starts with an original full data set and proceeds by removing a single feature. Finally, a subset of sixteen features is selected by the forward-backward feature selection which gives the same discrimination of the full set.

All the methods discussed confirm that features can be grouped into shape-based features, pixel-intensity-based features, Fourier-descriptor-based features and colour-based features. They are listed in Tables 3-6.

Table 3. Features based on MA shape

No.	Shape feature
1	Area of pixels in the candidate [4, 14, 17, 18, 32]
2	Perimeter of the object [14, 17, 18, 28]
3	Aspect ratio between length of largest and second largest eigenvector of covariance matrix of the object [14, 17, 18, 28]
4	Circularity [4, 14, 17, 18, 32]
5	Average Gaussian filter response of the green image with σ=1,2,4,8 (all 4 features) [14]
6	Standard deviation response of the green image after Gaussian filtering with σ=1,2,4,8 [14]
7	Maximum, minimum and average correlation coefficients of the candidate. Candidates with higher coefficients are more likely to be true microaneurysms [14].
8	Major axis length of the candidate [14]
9	Minor axis length of the candidate [14]
10	Compactness [14, 21]
11	Mean and standard deviation of filter outputs under the object. Filters consist of the Gaussian and its derivatives up to second order at scales σ= 1, 2, 4, 8 pixels [21].
12	Average value under the object of absolute difference between two largest eigenvalues of the Hessian tensor. The scale σ= 2 is used for the Gaussian partial derivatives that make up the Hessian [21].
13	Average output under the object of iris filter used on shade-corrected images with minimum circle radius of 4, maximum radius of 12 and 8 directions [21]
14	Total energy, calculated by summing grey levels of those pixels in original image, which coincides with overlaying binary object [28]
15	Mean energy. Mean energy of the object is total energy divided by area of that object [28].
16	Energy difference. A measure of object energy independent of background fluorescence is obtained by calculating the difference in the energies of the object and the background retina at that location [28].
17	Mean energy difference. Mean energy difference of the object is total energy difference divided by area of that object [28].
18	Minor axis variance [4]
19	Major axis variance [4]
20	Ratio of major axis variance and minor axis variance [4, 32]

Table 4. Features based on MA pixel intensity

No.	Pixel intensity feature
1	Total intensity of the object in original green plane image [14, 17, 18]
2	Total intensity of the object in shade-corrected image [14, 17, 18]
3	Mean intensity under the object in original green plane image [14, 17, 18]
4	Mean intensity under the object in shade-corrected image [14, 17, 18]
5	Normalised intensity in original green plane image [14]
6	Normalised intensity in shade-corrected image [14]
7	Normalised mean intensity in original green image [14]
8	Normalised mean intensity in shade-corrected image [14]
9	Intensity of the region growing seed in match filtered image [14]
10	Mean of opening image from top-hat step [29]
11	Standard deviation of opening image from top-hat step [29]
12	Maximum value of top-hat by diameter [32]
13	Mean value of top-hat [32]
14	Dynamic, morphological contrast measure [32]
15	Outer mean value of preprocessed image adds information about the surroundings of the candidate [32].
16	Outer standard deviation of preprocessed image gives information about grey level variation within the surroundings of the candidate [32].
17	Inner standard deviation of preprocessed image [32]
18	Inner mean value of top-hat transform helps identifying candidates situated on vessels [32].
19	Inner range is simply the dynamic range of preprocessed image in the candidate region and exploits information about smoothness of transition between candidate and environment [32].
20	Outer range is dynamic range of preprocessed image on the surroundings of candidates [32].
21	Grey level contrast between inner and outer region [32]
22	Half-range area is number of pixels inside candidate region with a grey level greater than arithmetical mean of minimum and maximum in candidate region [32].

Table 5. Features based on MA Fourier descriptor

No.	Fourier descriptor feature
1	(x,y) Perimeter indices [29]
2	Perimeter-area-based non-linear feature [29]
3	Second-moment-based non-linear feature [29]

Table 6. Features based on MA colour

No.	Colour feature
1	Difference between mean pixel values inside the object and mean values in circular region centred on object in the red plane from RGB colour space [14, 21]
2	Difference between mean pixel values inside the object and mean values in circular region centred on object in the green plane from RGB colour space [14]
3	Difference between mean pixel values inside the object and mean values in circular region centred on object in the blue plane from RGB colour space [14]
4	Difference between mean pixel values inside the object and mean values in circular region centred on object in the hue plane from HSI colour space [14]
5	Mean of pixel value within candidate regions of each component of RGB and HSI colour spaces [29]
6	Standard deviation of pixel value within candidate regions of each component of RGB and HSI colour spaces [29]
7	Second moments (in x,y and radial directions) of pixel value within candidate regions of each component of RGB and HSI colour spaces [29]
8	Complexity [29]
9	Aspect ratio [29]
10	Object mean colour, green channel – background colour, green channel [4]
11	Colour contrast [32]

PERFORMANCE MEASUREMENT

In the literature, to evaluate classifier performance, the sensitivity and specificity on a per-pixel basis are used. All measures can be calculated based on four values, namely the true positive rate, the false positive rate, the false negative rate and the true negative rate. Sensitivity is the percentage of the actual MA pixels that are detected, and specificity is the percentage of non-MA pixels that are correctly classified as non-MA pixels. The number of true positives represents the number of MA pixels correctly detected, whereas the number of false positives represents the number of MA pixels wrongly detected as MA pixels. There is some work in the literature where the average number of false positives per image is used as one of the performance evaluation values [9, 14, 18, 29, 31]. There are other useful classifier performances used to evaluate the system but not reported in the literature, such as precision, accuracy, precision and recall and positive likelihood ratio (PLR). Precision is the percentage of detected pixels that are actually MAs. Accuracy is the overall per-pixel success rate of the classifier. Precision and recall is the average of the precision and recall (also known as sensitivity). A PLR is the proportion of the probability of MA pixels which are positively detected (true positive rate) and that of non-MA pixels which are positively detected (false positive rate). PLR is based on the ratio of sensitivity to specificity.

COMPARING ALGORITHM PERFORMANCE

In this section the experimental results of MA detection using RRG, mathematical morphology, multi-scale correlation coefficients, discrimination function, matched filter, k-NN classifier, neural network and miscellaneous classifiers are presented. The performance on the test set is evaluated by comparing the classifier's result to a ground truth. The sensitivity of detection of images with one or more MAs gives the success rate for detection of this early sign of diabetic retinopathy. The performance of each classifier in terms of sensitivity, specificity and false positive rate per image is summarised in Table 7 and a graphical representation is shown in Figure 5.

Because all the algorithms run on different platforms, the performance of each algorithm cannot be measured and compared using run-time. The computation complexity of each algorithm is compared instead as in Table 8. A classifier selection factor is also presented in Table 9.

The weakness of the RRG, mathematical morphology, multi-scale correlation coefficients, discrimination function and matched filter is that they require many predetermined features, while the k-NN classifier and neural network require a learning phase which is time-consuming.

Table 7. Summary of reported sensitivity, specificity and number of false positives per image of MA detection

Classifier	Author	Se (%)	Sp (%)	No. of FP per image
Recursive region growing (RRG)	Spencer et al. [17]	82.00	86.00	100
	Cree et al. [18]	82.00	NR	5.7
	Sinthanayothin et al. [5]	77.50	88.70	NR
Discrimination function +	Hipwell et al. [20]	81.00	93.00	NR
matched filter	Raman et al. [23]	70.00	70.00	NR
	Streeter et al. [29]	56.00	NR	5.7
Discrimination function + MSCF	Zhang et al. [14]	NR	NR	0.36
k-NN classifier	Ege et al. [4]	69.00	NR	NR
k-NN classifiter +	Niemeijer et al. [21]	100.00	87.00	NR
Mathemetic morphology	Dupas et al. [9]	88.47	NR	2.13
Neural network	Gardner et al. [6]	73.80	73.80	NR
	Kamel et al. [30]	NR	NR	2.67
Neural network + RRG + moat opeartor	Usher et al. [22]	95.10	46.30	NR
Miscellaneous	Kose et al. [31]	95.10	99.30	NR
	Walter et al. [32]	88.47	NR	2.13

Note: Se = sensitivity, Sp = specificity, FP = false positives, MSCF = multi-scale correlation coefficients, NR = not reported
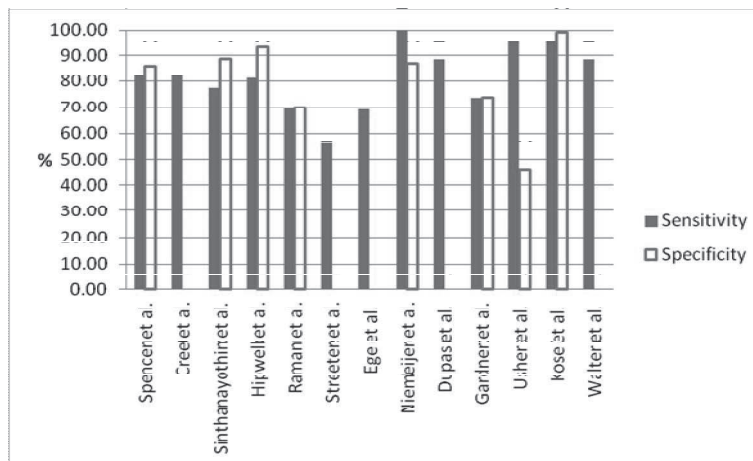
Figure 5. Graphical representation of sensitivity and specificity

Table 8. Time complexity (for one image)

Classifier	Training time complexity	Testing time complexity
Recursive region growing (RRG)	-	O (n^2i)
Mathematic morphology (watershed transform, top-hat)	-	O (n^2i)
Multi-scale correlation coefficients (MSCF)	-	O (n^2i)
Matched filter	-	O (n^2i)
Discrimination function	-	O (n^2if)
Nearest neighbour	O (1)	O (nft)
Neural network	O (m^2f^2)	O (nft)

Note: m is number of training data (number of training pixels); n is number of testing data (number of testing pixels); i is number of iteration; f is number of features; t is number of training points.

Table 9. Classifier selection factor

Classifier	Parameter sensitive	Require learning phase	High computation cost	Require high computer system
Recursive region growing (RRG)	Yes			
Mathematic morphology (watershed transform, top-hat)	Yes			
Multi-scale correlation coefficients (MSCF)	Yes		Yes	
Matched filter	Yes			
Discrimination function	Yes		Yes	
Nearest neighbour		Yes	Yes	Yes
Neural network		Yes	Yes	Yes

DISCUSSION AND CONCLUSIONS

Ophthalmologists use retinal photographs to follow up, diagnose and treat eye diseases. The appearance of the fundus image can provide a great deal of pathological information about eye diseases. Another advantage of digital fundus photography is the capacity to transmit images to a centralised reading centre for grading. It can also indicate early signs of diabetic retinopathy. This is because retinal details may be easier to visualise in fundus photographs as opposed to direct examination [39].

Most of the algorithm reviewed in earlier sections work on fluorescein angiography or colour images taken on patients with dilated pupils, in which the MA and other retinal features are clearly visible. The examination time and effect on the patient could be reduced if the detection system could succeed on images taken from patients with non-dilated pupils. However, the quality of these images will be poor and it greatly affects the performance of those algorithms. Therefore, methods for MA detection on non-dilated fundus images still need to be developed. On poor contrast images, small MAs are removed during the pre-processing step while thresholding is not effective given the variations in background intensity. Some small MAs could also be missed during the candidate MA detection step. It is difficult to determine which is the best algorithm at each stage due to the variability of factors such as the source of the test image set, the size of sample, which ranges from four images to over 3700 images, and the way in which the performance of each method is evaluated. Abramoff et al. [40] suggested that the performance of algorithms should be evaluated on a published digital image data set. The other problem in evaluation of the systems is that there are only a few gold standard images read by experts to compare with.

For feature extraction, in early research few features were used. Then subsequently, more features were introduced to improve the classification efficiency. Due to a large number of features, feature selection is needed to reduce irrelevant features. There are interesting methods such as the entropy-based [41], the correlation-based feature selection [42], Chi-squared method [43], signal-to-noise statistic [44], microarray data [45], relief [46], random forest feature selection [47], Gini index [48], sequential forward selection and sequential backward feature selection [49].

The evaluation of a system for automatic detection of DR from colour fundus photographs using published algorithms [40] cannot yet be recommended for clinical practice because the false negative rate is still not acceptable and the processing time is still too slow and large computational efforts are required. Therefore, there is a compelling need for algorithm improvement including better detection of early sign of DR such as MAs. Large scale retinal image sets from patients with diabetes, which are taken with different camera types are needed for testing before being put into direct clinical use.

Further investigations should improve the efficiency of the detection systems. To increase the accuracy rate, more advanced classifiers such as Support Vector Machines (SVMs) [50, 51] and Adaptive Boosting (AdaBoost) [52-55] could be used. Both of them are machine learning algorithm. SVMs map training data into a high-dimensional feature space in which we can construct separated hyper-planes, maximising the margin or distance from the hyper-plane to the nearest training data points. AdaBoost is a meta-algorithm that sequentially selects weak classifiers from a candidate pool and weights each of them based on their error. Similar to SVMs, AdaBoost works by combining several votes. Instead of using support vectors, AdaBoost uses weak learners.

With improved technology, telemedicine is a very useful application. A hospital from a rural area with limited ophthalmologist resource could transmit a patient's fundus image to a central clinic for diagnosis by an experienced clinician. An automatic DR detection system could help in the screening process or act like a decision support system [56-60]. Only the patients with a risk of DR will be further examined by the ophthalmologist. Therefore, the workload of the ophthalmologist is reduced and the patient will be examined faster, which will lead to early treatment and reduction of blindness.

REFERENCES

1. D. S. Fong, L. P. Aiello, F. L. Ferris and R. Klein, "Diabetic retinopathy", *Diabetes Care*, **2002**, *25*, 590-593.

2. The University of Michigan Kellogg Eye Center, "Diabetic retinopathy", **2009**, http://www.kellogg.umich.edu/patientcare/conditions/diabetic.retinopathy.html (Accessed: April 2012).

3. P. Ruamviboonsuk, N. Wongcumchang, P. Surawongsin, E. Panyawatananukul and M. Tiensuwan, "Screening for diabetic retinopathy in rural area using single-field, digital fundus images", *J. Med. Assoc. Thai.,* **2005**, *88*, 176-180.

4. B. M. Ege, O. K. Hejlesen, O. V. Larsen, K. Moller, B. Jennings, D. Kerr and D. A. Cavan, "Screening for diabetic retinopathy using computer based image analysis and statistical classification", *Comput. Meth. Programs Biomed.*, **2000**, *62*, 165-175.

5. C. Sinthanayothin, J. F. Boyce, T. H. Williamson, H. L. Cook, E. Mensah, S. Lal and D. Usher, "Automated detection of diabetic retinopathy on digital fundus images", *Diabet. Med.*, **2002**, *19*, 105-112.

6. 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.

7. Z. Liu, C. Opas and S. M. Krishnan, "Automatic image analysis of fundus photograph", Proceedings of 19th Annual International Conference of IEEE on Engineering in Medicine and Biology, **1997**, Chicago, USA, pp.524-525.

8. P. Massin, A. Erginay A. B. Mehidi, E. Vicaut, G. Quentel, Z. Victor, M. Marre, P. J. Guillausseau and A. Gaudric, "Evaluation of a new non-mydriatic digital camera for detection of diabetic retinopathy", *Diabet. Med.*, **2003**, *20*, 635-641.

9. B. Dupas, T. Walter, A. Erginay, R. Ordonez, N. Deb-Joardar, P. Gain, J. C. Klein and P. Massin, "Evaluation of automated fundus photograph analysis algorithms for detecting microaneurysms, haemorrhages and exudates, and of a computer-assisted diagnostic system for grading diabetic retinopathy", *Diabet. Metab.*, **2010**, *36*, 213-220.

10. National Institutes of Health USA, "Structured analysis of the retina", **2004**, http://www.ces.clemson.edu/~ahoover/stare/ (Accessed: March 2012).

11. Image Sciences Institute, "DRIVE: Digital retinal images for vessel extraction", **2001**, http://www.isi.uu.nl/Research/Databases/DRIVE (Accessed: March 2012).

12. T. Kauppi, V. Kalesnykiene, J. K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, J. Pietila, H. Kalvianen and H. Uusitalo, "DIARETDB1- Standard diabetic retinopathy database calibration level 1", **2007**, http://www2.it.lut.fi/project/imageret/diaretdb1/ (Accessed: March 2012).

13. T. Teng, M. Lefley and D. Claremont, "Progress towards automated diabetic ocular screening: A review of image analysis and intelligent systems for diabetic retinopathy", *Med. Biol. Eng. Comput.*, **2002**, *40*, 2-13.

14. B. Zhang, X. Wu, J. You, Q. Li and F. Karray, "Detection of microaneurysms using multi-scale correlation coefficients", *Pattern Recogn.*, **2010**, *43*, 2237-2248.

15. S. C. Lee, Y. Wang and E. T. Lee, "Computer algorithm for automated detection and quantification of microaneurysms and hemorrhages (hma's) in colour retinal images", Proceedings of International Conference on Image Perception and Performance, **1999**, San Diego, USA, pp.61-71.

16. A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson and P. F. Sharp, "Automated microaneurysm detection using local contrast normalization and local vessel detection", *IEEE Trans. Med. Imag.*, **2006**, *25*, 1223-1232.

17. T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp and J. V. Forrester, "An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus", *Comput. Biomed. Res.*, **1996**, *29*, 284-302.

18. M. J. Cree, J. A. Olson, K. C. McHardy, P. F. Sharp and J. V. Forrester, "A fully automated comparative microaneurysm digital detection system", *Eye*, **1997**, *11*, 622-628.

19. 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.

20. J. H. Hipwell, F. Strachan, J. A. Olson, K. C. McHardy, P. F. Sharp and J. V. Forrester, "Automated detection of microaneurysms in digital red-free photographs: A diabetic retinopathy screening tool", *Diabet. Med.*, **2000**, *17*, 588-594.

21. M. Niemeijer, B. van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sanchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu, X. Wu, G. Cazuquel, J. You, A. Mayo, Q. Li, Y. Hatanaka, B. Cochener, C. Roux, F. Karray, M. Garcia, H. Fujita and M. D. Abramoff, "Retinopathy online challenge: Automatic detection of microaneurysms in digital colour fundus photographs", *IEEE Trans. Med. Imag.*, **2010**, *29*, 185-195.

22. D. Usher, M. Dumskyj, M. Himaga, T. H. Williamson, S. Nussey and J. Boyce, "Automated detection of diabetic retinopathy in digital retinal images: A tool for diabetic retinopathy screening", *Diabet. Med.*, **2004**, *21*, 84-90.

23. B. Raman, E. S. Bursell, M. Wilson, G. Zamora, I. Benche, S. C. Nemeth and P. Soliz, "The effects of spatial resolution on an automated diabetic retinopathy screening system's performance in detecting microaneurysms for diabetic retinopathy", Proceedings of 17th IEEE Symposium on Computer-Based Medical Systems, **2004**, Washington, DC, USA, pp.128-133.

24. H. Li and O. Chutatape, "A model-based approach for automated feature extraction in fundus images", Proceedings of 9th IEEE International Conference on Computer Vision, **2003**, Washington, DC, USA, pp.394-399.

25. A. Osareh, M. Mirmehdi, B. Thomas and R. Markham, "Automated identification of diabetic retinal exudates in digital colour images", *Br. J. Ophthalmol.*, **2003**, *87*, 1220-1223.

26. M. Lalonde, M. Beaulieu and L. Gagnon, "Fast and robust optic disc detection using pyramidal decomposition and Hausdorff-based template matching", *IEEE Trans. Med. Imag.*, **2001**, *20*, 1193-1200.

27. J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher and L. Kennedy, "Optic nerve head segmentation", *IEEE Trans. Med. Imag.*, **2004**, *23*, 256-264.

28. M. Hafez and S. A. Azeem, "Using adaptive edge technique for detecting microaneurysms in fluorescein angiograms of the ocular fundus", Proceedings of 11th Mediterranean Electrotechnical Conference, **2002**, Rochester (NY), USA, pp.479-483.

29. L. Streeter and M. J. Cree, "Microaneurysm detection in colour fundus images", Proceedings of Image Vision Computing New Zealand, **2003**, Palmerston North, New Zealand, pp.280-285.

30. M. Kamel, S. Belkassim, A. M. Mendonca and A. Campilho, "A neural network approach for the automatic detection of microaneurysms in retinal angiograms", Proceedings of 1st International Joint Conference on Neural Networks, **2001**, Washington, DC, USA, pp.2695-2699.

31. C. Kose, U. Sevik, C. Ikibas and H. Erdol, "Simple methods for segmentation and measurement of diabetic retinopathy lesions in retinal fundus images", *Comput. Meth. Programs Biomed.*, **2012**, *107*, 274-293.

32. T. Walter and J. C. Klein, "Automatic detection of microaneurysms in colour fundus images of the human retina by means of the bounding box closing", Proceedings of 3rd International Symposium on Medical Data Analysis, **2002**, Rome, Italy, pp.210-220.

33. S. Derivaux, G. Forestier, C. Wemmert and S. Lefevre, "Supervised image segmentation using watershed transform, fuzzy classification and evolutionary computation", *Pattern Recogn. Lett.*, **2010**, *31*, 2364-2374.

34. G. Luo, O. Chutatape, H. Li and S. M. Krishnan, "Abnormality detection in automated mass screening system of diabetic retinopathy", Proceedings of 14th IEEE Symposium on Computer-Based Medical Systems, **2001**, Bethesda (MD), USA, pp.132-137.

35. J. C. Olivo-Marin, "Extraction of spots in biological images using multiscale products", *Pattern Recogn.*, **2002**, *35*, 1989-1996.

36. R. J. Winder, P. J. Morrow, I. N. McRitchie, J. R. Bailie and P. M. Hart, "Algorithms for digital image processing in diabetic retinopathy", *Comput. Med. Imag. Graph.*, **2009**, *33*, 608-622.

37. P. Kahai, K. R. Namuduri and H. Thompson, "A decision support framework for automated screening of diabetic retinopathy", *Int. J. Biomed. Imag.*, **2006**, *2006*, 1-8.

38. S. J. Aldington, E. M. Kohner, S. Meuer, R. Klein and A. K. Sjolie, "Methodology for retinal photography and assessment of diabetic retinopathy: The EURODIAB IDDM complications study", *Diabetologia*, **1995**, *38*, 437-444.

39. J. Hayashi, T. Kunieda, J. Cole, R. Soga, Y. Hatanaka, M. Lu, T. Hara and H. Fujita, "A development of computer-aided diagnosis system using fundus images", Proceedings of 7th International Conference on Virtual Systems and Multimedia, **2001**, Berkeley (CA), USA, pp. 429-438.

40. M. D. Abramoff, M. Niemeijer, M. S. Suttorp-Schulten, M. A. Viergever, S. R. Russell and B. van Ginneken, "Evaluation of a system for automatic detection of diabetic retinopathy from colour fundus photographs in a large population of patients with diabetes", *Diabet. Care.*, **2008**, *31*, 193-198.

41. U. Fayyad and K. Irani, "Multi-interval discretization of continuous-valued attributes for classification learning", Proceedings of 13th International Joint Conference on Artificial Intelligence, **1993**, Chambery, France, pp.1022-1029.

42. M. A. Hall, *"*Correlation-based feature selection machine learning"*, PhD Thesis*, **1999**, University of Waikato, New Zealand.

43. H. Liu and R. Setiono, "Chi2: Feature selection and discretization of numeric attributes", Proceedings of 7th International Conference on Tools with Artificial Intelligence, **1995**, Herndon (VA), USA, pp. 338-391.

44. T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gassenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield and E. S. Lander, "Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring", *Science*, **1999**, *286*, 531-537.

45. V. Vinaya, N. Bulsara, C. J. Gadgil and M. Gadgil, "Comparison of feature selection and classification combinations for cancer classification using microarray data", *Int. J. Bioinform. Res. Appl.*, **2009**, *5*, 417-431.

46. I. Kononenko, "Estimating attributes: Analysis and extensions of RELIEF", Proceedings of the European Conference on Machine Learning, **1994,** Catania, Italy, pp.171-182.

47. L. Breiman, "Random forest**s"**, *Mach. Learn.*, **2001**, *45*, 5-32.

48. L. Breiman, J. H. Friedman, C. J. Stone and R. A. Olshen, "Classification and Regression Trees", Chapman and Hall/CRC, Boca Raton, **1984**.

49. G. H. John, R. Kohavi and K. Pfleger, "Irrelevant features and the subset selection problem", Proceedings of 11th International Conference on Machine Learning, **1994**, New Brunswick (NJ), USA, pp.121-129.

50. C. J. C. Burges, "A tutorial on support vector machines for pattern recognition", *Data Mining Knowl. Discov.*, **1998**, *2*, 121-167.

51. A. Osareh, M. Mirmehdi, B. Thomas and R. Markham, "Comparative exudate classification using support vector machines and neural networks", Proceedings of 5th International Conference on Medical Image Computing and Computer-Assisted Intervention--Part II, **2002**, Tokyo, Japan, pp.413-420.

52. Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting**"**, *J. Comput. Syst. Sci.*, **1997**, *55*, 119-139.

53. R. Nishii and S. Eguchi, "Supervised image classification based on adaboost with contextual weak classifiers", Proceedings of 4th IEEE International Geoscience and Remote Sensing Symposium, **2004**, Anchorage (AK), USA, pp.1467-1470.

54. Y. W. Wu and X. Y. Ai, "Face detection in color images using AdaBoost algorithm based on skin color information", Proceedings of 1st International Workshop on Knowledge Discovery and Data Mining, **2008**, Adelaide, Australia, pp.339-342.

55. J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, A. W. Toga and P. M. Thompson, "Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation", *IEEE Trans. Med. Imag.*, **2010**, *29*, 30-43.

56. O. Faust, U. R. Acharya, E. Y. K. Ng, K. H. Ng and J. S. Suri, "Algorithms for the automated detection of diabetic retinopathy using digital fundus images: A review", *J. Med. Syst.*, **2012**, *36*, 145-157.

57. N. Patton, T. M. Aslam, T. MacGillivray, I. J. Deary, B. Dhillon, R. H. Eikelboom, K. Yogesan and I. J. Constable, "Retinal image analysis: Concepts, applications and potential", *Prog. Retin. Eye Res.*, **2006**, *25*, 99-127.

58. W. L. Yun, U. R. Acharya, Y. V. Venkatesh, C. Chee, L. C. Min and E. Y. K. Ng, "Identification of different stages of diabetic retinopathy using retinal optical images", *Inform. Sci.,* **2008**, *178*, 106-121.

59. A. Osareh, B. Shadgar and R. Markham, "A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images", *IEEE Trans. Inform. Technol. Biomed.*, **2009,** *13*, 535-545.

60. N. Singh and R. C. Tripathi, "Automated early detection of diabetic retinopathy using image analysis techniques", *Int. J. Comput. Appl.*, **2010**, *8*, 18-23.

Maejo International Journal of Science and Technology

Full Paper

Patent quality determinants based on technology life cycle with special reference to solar-cell technology field

Jungkyu Park[1] and Eunnyeong Heo [2,*]

[1] R&D Strategy Office, Korea Institute of Geoscience and Mineral resources, 124 Gwahang-no, Daejeon, South Korea

[2] Department of Energy Systems engineering, Seoul National University, 599 Gwanang-no, Seoul, South Korea

* Corresponding author, e-mail: heoe@snu.ac.kr

Abstract: The purpose of this study is to illustrate the necessity of considering the technology life cycle when creating a distinct R&D strategy planning, if the aim is to enhance patent quality. The study also suggests an effective R&D strategy for the solar- cell technology field. It uses count data models to introduce the concept of technology life cycle and analyses the determinants of patent quality that depend upon the technology life cycle. Empirical results show that three variables influence patent quality in contrary manners, depending on the stage of the technology life cycle. This means that an R&D strategy for each of the three variables should be established while considering the technology life cycle. As a result, this study clarifies that the technology life cycle needs to be considered when establishing an R&D strategy that will enhance patent quality, and it suggests that distinct R&D strategy planning should be done for the solar-cell technology field in particular while bearing in mind the technology life cycle.

Keywords: technology life cycle, patent quality, count data model, solar-cell technology

---

INTRODUCTION

Today's crises involving patent trolls, increased patent litigation and technology standards competition point to the importance of intellectual property rights, the possession of which can create economic value in a knowledge-based economy. There is a positive relationship between a patent's economic value and its citation counts, and the latter is used as a

proxy to determine the quality of the patent [1–3]. In other words, a patent with high economic value will be cited more frequently [4–5].

Hence, R&D policymakers need to pay attention to the determinants of patent citation counts to establish R&D strategies that give rise to stronger R&D economic performance. On the other hand, technology is created so it can be further developed and introduced into a variety of R&D environments, until it eventually reaches a stage of decline. Therefore, to achieve stronger R&D economic performance, any R&D strategy should be differentiated by virtue of the technology life cycle. However, R&D strategies have generally been established through some experts' peer views, without due consideration for either the determinants of patent citation counts or the technology life cycle.

The current study offers an R&D strategy that bears in mind these considerations. In particular, the concept of the technology life cycle is initially introduced to identify the stage of technology development. Empirical analyses are used to emphasise differences in R&D strategies while taking into consideration the determinants of patent citation counts in terms of the technology life cycle. Moreover, with regard to technology development undertaken to improve patent quality, this study suggests an R&D strategy for the solar-cell technology field.

LITERATURE REVIEW

Patent data have been considered useful in analysing various trends including technological change, technological development, and economic growth [6–9]. In particular, patent citation counts data have been used in applied research fields to study patent quality, knowledge spillover, economic value and so on. Trajtenberg [10], Narin et al. [11] and Carpenter et al. [12] each attempted to put forward patent counts that are weighted by citations as indicators of the value of innovations and hence overcome the limitations of simple counts—limitations that have hindered assessments of technology importance or value in economic research. Moreover, Harhoff et al. [3] and Sampat and Ziedonis [13] each suggested that there is a positive relationship between patent citation counts and economic value, and the studies of Fung and Chow [14], Hu and Jaffe [15], and Jaffe et al. [16] each showed that patent citation data are a good proxy for knowledge flow into an industry.

The studies of Sampat [17] and Lee et al. [5] are representative pieces of work that discuss patent quality in relation to patent citation counts. Lee et al. [5] identified the factors that affect patent citation counts using US patents that belong to particular government-funded research institutes in South Korea. Sampat [17] meanwhile collected a large quantity of patent data issued between 2002 and 2003 and analysed patent quality and any patents (or published articles) related to an invention.

As shown above, most studies to date verify the relationship between patent citation counts and patent quality (or the economic value of a patent). However, few studies discuss R&D strategies that can be used to improve patent quality while using determinants of patent citation counts. Moreover, prior studies did not apply the concept of the technology life cycle in spite of changes in the technology development environment. To fill this research gap and address the importance of R&D strategies, this study addresses the concept of technology life cycle and analyses the determinants of patent citation counts.

MODELS FOR COUNT DATA

The most useful model for use with the count data is the Poisson distribution. It can be used to model the number of occurrences of a type of event such as the numbers of patent applications, patent citations or car accidents [18–19]. If the discrete random variable Y is Poisson distributed with an intensity or rate parameter μ ($\mu > 0$), then Y has the density [20–21]:

$$Prob(Y_i = y_i) = \frac{e^{-\mu}\mu^{y_i}}{y_i!} \qquad y = 1, 2, 3, \dots \tag{1}$$

where μ is a positive real number equal to the expected number of occurrences during the given interval; e is the base of the natural logarithm; and y is the number of occurrences of an event. The mean of the Poisson distribution is equal to its variance, i.e. E(Y) = Var(Y) = μ, which is a unique feature of this distribution.

A regression model specifies the parameter μ as varying across individuals according to a specific function of regressor vector x and parameter vector β. The typical Poisson specification is $\mu = \exp(x'\beta)$. The method of maximum likelihood is widely used to estimate the parameter. The log-likelihood function of the Poisson estimation models is as follows [20–21]:

$$\ln L_{Poisson} = \sum_{y_i} \left[-\mu_i + y_i \ln(\mu_i) - \ln(y_i!) \right] \tag{2}$$

The equidispersion property of Poisson distribution, *E(Y) = Var(Y) = μ*), is violated in many research studies because the overdispersion problem is common. Researchers have proposed many extensions for the count data model to improve the validity of the equidispersion assumption inherent in the Poisson model [22]. One of these is the negative binomial model—a model more general than the Poisson model because it accommodates overdispersion. The negative binomial distribution arises as a continuous mixture of the Poisson distribution, where the mixing distribution of the Poisson rate is a gamma distribution. In equation (1), by replacing μ with μv, where v is a random variable, $y \sim Poisson(y|\mu v)$. If v is specified, i.e. $E(v) = 1$, $Var(v) = \sigma^2$, v preserves the mean but increases the dispersion. In $v \sim Gamma(1, \alpha)$, α is the variance parameter of the gamma distribution. The gamma function, $\Gamma(\alpha)$, is defined by $\Gamma(\alpha) = \int_0^\infty e^{-t} t^{a-1} dt$, $\alpha > 0$. Thus, a negative binomial distribution is denoted by $NB(\mu, \alpha)$, and its probability mass function is a mixture density [20–21]:

$$Prob(Y_i = y_i) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1})\Gamma(y+1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu} \right)^{\alpha^{-1}} \left(\frac{\mu}{\alpha^{-1} + \mu} \right)^{y} \tag{3}$$

The moments of negative binomial distributions are $E(y) = \mu$, $Var(y) = \mu(1 + \alpha\mu)$. The log-likelihood function of the negative binomial estimation model is as follows [20–21]:

$$\ln L_{NB} = \sum [y \ln(\frac{\alpha\mu_i}{1 + \alpha\mu_i}) - \frac{1}{\alpha} \ln(1 + \alpha\mu_i) + \ln\Gamma(y_i + \frac{1}{\alpha}) - \ln\Gamma(y_i + 1) - \ln\Gamma(\frac{1}{\alpha})] \tag{4}$$

TECHNOLOGY LIFE CYCLE

The technology life cycle comprises a pattern of dynamic characteristics pertaining to technology, in which its innovative and economic outcomes change over time. Ford and Ryan [23] developed the idea of technology life cycle and broke it into six distinct periods─technology development, application, application launch, application growth, technology maturity and degraded technology─from the viewpoint of technology selling, based on its position within the product life cycle [24].

Haupt et al. [25] pointed out a good reason for using the patent approach to measure the technology life cycle: patents inform the public about technological developments since they contain technological expertise and inform the public about the commercial potential of certain technologies. In its patent portfolio, the Japan Intellectual Property Association (JIPA) breaks the technology life cycle into five distinct periods using the trends pertaining to patent applications and growth rates: technology seed, growth, development, maturity and technology declining period (Figure 1) [26].
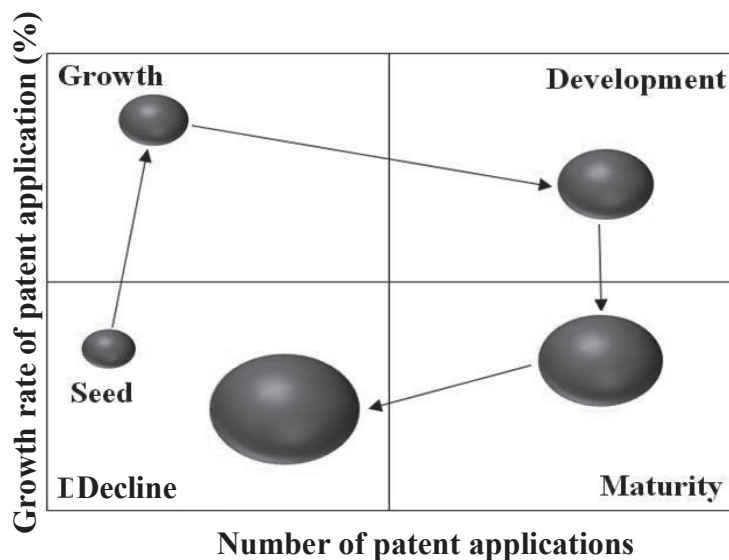
Figure 1. Characteristics of patent application in terms of technology life cycle period

A technology growth period is defined by an increase in its number of applications and in its growth rate. Similarly, an increase in the application number and in the static growth rate of applications are observed during the technology development period. In the technology maturity period the number of applications stagnates and the application growth rate decreases. A summary of the technology life cycle's characteristics vis-à-vis the number of applications and the growth rate of applications is provided in Table 1.

Table 1.   Characteristics of the technology life cycle period

Life cycle period	Description	
Seed	No. of applications (↓),	Growth rate of applications (↓)
Growth	No. of applications (↑),	Growth rate of applications (↑)
Development	No. of applications (↑),	Growth rate of applications (→)
Maturity	No. of applications (→),	Growth rate of applications (↓)
Declining	No. of applications (↓),	Growth rate of applications (↓)

DATASET AND VARIABLES

Dataset

Solar-cell technology is a significant field that promises a new form of renewable energy. It has the characteristic of being widely distributed in other technology areas such as liquid-crystal displays, semiconductor and lighting. Its development is being promoted in various countries through government support, along with the commercialisation of the technology and rapid expansion into various markets. Therefore, a large quantity of patent applications in the last 20 years is available for analysis. Furthermore, one should bear in mind that in terms of the technology life cycle, solar-cell technology has already passed its maturity period [27].

The technology classes and technology ranges for patent searching in the solar-cell technology field are shown in Table 2, following the precedent set by the quasi-government institute—the Korea Institute for Advancement of Technology (KIAT) [28].

Table 2.   Solar cell technology classes and ranges

Technology class	Technology range
Material	Silicon cell, inorganic compound cell, organic compound cell, dye cell
Manufacture	Manufacturing process (wafer, ingot, module, array)
Module	Thin-film type (CIGS), wafer type, flat-bed type, array
Electrode	Electrode for solar-cell structure, electrode for solar-cell manufacturing process

The patent search results show that the number of US patent applications has been in decline since 2001. However, there has been a rapid increase in the number of patent applications in South Korea and Europe, where patent applications were not very active in the 1990s. On the other hand, the number of patent applications in Japan did not change substantially, as shown in Table 3. Although patent application information related to the patent offices of South Korea (KIPO), Japan (JPO), and the European Union (EPO) would have various implications, this study was compelled to use US patent data because only the US Patent and Trademark Office (USPTO) provides patent citation information. Thus, after removing data noise, the 4,447 granted US patents found to be related to the solar-cell technology field yielded

1,466 valid data records, and then this valid data should be grouped by technology life cycle for empirical analysis in this study.

Table 3. Number of patent applications by year and patent office for solar-cell technology

Year	KIPO(KR)	USPTO(US)	JPO(JP)	EPO(EU)
1990	3	65	-	20
1991	1	80	29	33
1992	3	90	85	49
1993	2	58	75	43
1994	11	88	71	32
1995	16	81	86	36
1996	17	90	80	36
1997	29	91	110	44
1998	20	100	130	58
1999	33	116	110	86
2000	40	128	102	77
2001	58	143	98	79
2002	45	116	98	67
2003	75	87	106	56
2004	93	54	107	86
2005	126	37	105	92
2006	204	26	76	174
2007	347	14	113	218
2008	240	2	69	116
2009	44	-	22	21

However, when using patent data, it is not possible to classify all the life cycle periods therein in detail. Thus, the current study considers only the period of technology maturity and compares differences in patent quality determinants between the pre-technology and post-technology maturity periods. To classify the technology maturity period, a patent portfolio along with its number of patent applications and growth rate was constructed following JIPA's technology life cycle model (Figure 2). This patent portfolio data show that solar-cell technology matured in 2001. For empirical analysis, patent data are separated into pre-technology maturity and post-technology maturity periods, based on the year 2001. In its analysis this study uses 1,130 patents granted before 2001 and 336 patents granted after that year.
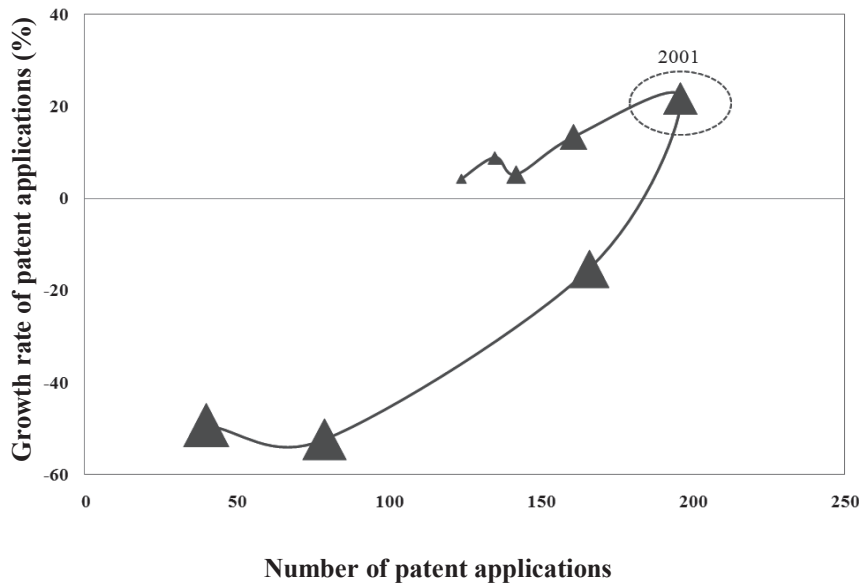
Figure 2. Patent portfolio of solar-cell technology (The size of the symbols ▲ represents the accumulated number of patent applications.)

Variables

The dependent variable for this study is the number of patent citation counts; this figure is a proxy of the patent quality as discussed in previous research [2, 5, 10]. The independent variables used in this study are listed in Table 4. A variety of information from the front pages of the patent documents were used to consider the strategies for R&D planning; in this sense the current study follows the lead of Lee et al. [5] and Daines [29].

Table 4. Description of independent variables

Variable	Measurement	Description
NA	No. of assignees	Size of research
DC	Domestic collaboration	Domestic joint research
INTC	International collaboration	International joint research
NINV	No. of inventors	Size of research team
2INV	Two or more nationalities of inventors	Linguistic problem
BCitation	No. of backward citations	Size of knowledge from outside
NNONP	No. of non-patent citations	Scientific linkage
NSELF	No. of self citations	Technological(knowledge) cumulativeness
USIC	No. of citations of US-invented patent	Degree of dependence on US technology
JPIC	No. of citations of JP-invented patent	Degree of dependence on JP technology
EUIC	No. of citations of EU-invented patent	Degree of dependence on EU technology
NCLAIM	No. of claims	Size of patent right
NFAM	No. of family patents	Size of potential market
NIPC	No. of international patent classification	Size of application range

To examine how the size of the research project and joint research (i.e. collaboration) affects patent citation counts, the number of assignees (NA) was used as an independent variable and dummy variables were set up for domestic collaboration (DC) and international collaboration (INTC). In addition, this study used the number of inventors (NINV) as an independent variable to observe the effect of the size of the research team. A dummy variable of two or more nationalities of inventors (2INV) was set up to investigate any language-related effects—a variable discussed by Maurseth and Verspagen [30]. In addition, the numbers of cases that cited other patents (BCitation) and non-patent documents (NNONP) as well as the number of self-citations (NSELF) were set up as independent variables to examine the influence of the degree of technology dependence type and the accumulation of knowledge.

The variables USIC, JPIC and EUIC were set up to accurately analyse the degree of technology dependence on the US, JP and EU respectively, and the nationality of the referenced patent's assignee was investigated rather than the nationality of the patent itself. Besides these variables, NCLAIM, NFAM and NIPC were considered independent variables and they were used to measure the size of the patent rights, the size of the potential market and the possibility of application to other fields of technology respectively.

The basic statistics with respect to the independent variables are summarised in Tables 5 and 6. In terms of the mean statistics of NA, DC, INTC, NINV, 2INV and NIPC, no statistically significant difference was found between the pre-technology and post-technology maturity periods. On the other hand, in the post-technology maturity period the mean statistics for BCitation, NNONP, NSELF, USIC, JPIC and EUIC related to technology dependence were about   twice   those in the pre-technology maturity period.   It is assumed that basic technologies

Table 5.   Descriptive statistics of variables (pre-technology maturity period)
for solar-cell technology

Variable	Mean	Std. dev.	Min.	Max.
NA	1.0	0.2	1	4
DC	0.0	0.2	0	1
INTC	0.0	0.1	0	1
NINV	2.8	1.7	1	10
2INV	0.0	0.1	0	1
BCitation	10.3	9.3	0	109
NNONP	2.6	4.4	0	43
NSELF	1.0	2.9	0	71
USIC	5.0	6.4	0	78
JPIC	4.0	5.0	0	100
EUIC	1.2	1.8	0	16
NCLAIM	19.3	16.8	1	236
NFAM	7.4	8.1	1	167
NIPC	1.8	0.9	1	5

Table 6. Descriptive statistics of variables (post-technology maturity period) for solar-cell technology

Variable	Mean	Std. dev.	Min.	Max.
NA	1.1	0.3	1	5
DC	0.1	0.2	0	1
INTC	0.0	0.1	0	1
NINV	2.7	1.8	1	10
2INV	0.1	0.2	0	1
BCitation	22.9	31.6	0	210
NNONP	7.2	14.4	0	94
NSELF	2.1	5.4	0	63
USIC	11.5	17.9	0	125
JPIC	7.2	9.4	0	59
EUIC	3.2	7.7	0	77
NCLAIM	20.4	15.6	1	119
NFAM	10.3	13.0	1	78
NIPC	1.6	1.0	1	6

related to solar cells were validated on the basis of prior research during the pre-technology maturity period and that these technologies were put into practical use in the post-technology maturity period. Meanwhile, Tables 5 and 6 show that in the solar-cell technology field, more US-invented and JP-invented patents were used in technology development than were EU-invented patents, irrespective of the technology life cycle period. The mean statistics of NCLAIM and NFAM in the post-technology maturity period tended to be relatively higher than those in the pre-technology maturity period. It is reasonable to assume that these results stem from a consideration of technology commercialisation in the post-technology maturity period.

EMPIRICAL RESULTS AND DISCUSSION

The current study estimated parameters by using STATA statistical software version 10.0. According to the estimation results, the effect of the number of assignees (NA), which indicates the size of the research project, on the patent citation count differs by technology life cycle period. With a one-unit increase in NA comes a patent citation count decrease of 0.436 unit in the pre-technology maturity period, while with a one-unit increase in NA comes a patent citation count increase of 0.485 unit in the post-technology maturity period. Similarly, if domestic joint research (DC) takes place in the pre-technology maturity period, the patent citation count increases by 0.614 unit, whereas in the post-technology maturity period, it decreases by 1.006 unit. Likewise, the effects of NA and DC on the patent quality are sensitive to the technology life cycle period. In other words, these results show that the technology life cycle period is a crucial factor in establishing an R&D strategy that enhances patent quality. On the other hand, Table 7 shows that undertaking international joint research (INTC) enhances patent citation count, irrespective of the technology life cycle period involved.

Regarding the effect of the number of inventors (NINV)—a variable that indicates the size of the research team—on the patent citation count, it was found that a one-unit increase in NINV leads to an increase of 0.018 unit in patent citation count in the pre-technology maturity period. Especially, this study considers the effect of a multinational research team (2INV) and uses it as a variable. As a result, irrespective of the technology life cycle period involved, and all else being equal, if a research team comprises researchers from a number of different countries, the patent citation count decreases by 0.165 unit. This result is similar to that of the study of Maurseth and Verspagen [30], who suggested that patent citation count is higher if the citing region belongs to the same linguistic group. Therefore, an R&D planner needs to consider the size of the research team as well as linguistic issues therein as the R&D strategy is formulated.

Typically, there are two paths of technology knowledge inflow. One involves knowledge inflow from a patent (BCitation); the other involves knowledge inflow from a nonpatent (NNONP) such as a journal article or technology magazine. As described in Table 7, the effects of the two paths on patent citation count differ. With a one-unit increase in BCitation comes a patent citation count decrease of 0.196 units, while with such an increase in NNONP, the patent citation count increases by 0.013 unit. The citation of one's prior own patent (NSELF), a special path of technology knowledge inflow, affects patent citation count differently. Among studies that use a substantial amount of NSELF in the pre-technology maturity period, patent citation count increases while it decreases in the post-technology maturity period. Technology knowledge accumulation and the referencing of non-patent documents can serve as significant components of an R&D strategy in the early stages of technology development. These results point to the importance of the technology knowledge inflow path in enhancing patent quality.

Results pertaining to each of the models indicate that the use of US-invented patents (USIC), JP-invented patents (JPIC) and EU-invented patent (EUIC) as references is important to the patent quality in the pre-technology maturity period. In the post-technology maturity period, only the Poisson model shows any significant positive impact on patent quality. This result indicates that the US, JP and the EU are the leading countries in solar-cell technology, and we need to be mindful of their prior patents and research trends.

Table 7 shows that high numbers of claims (NCLAIM), family patents (NFAM) and international patent classifications (NIPC) can increase patent citation count, which is consistent with the viewpoint of Lee et al. [5]. Moreover, the coefficient values for NCLAIM, NFAM and NIPC are much higher in the post-technology maturity period than in the pre-technology one. This could be a natural result because each of these variables represents the size of the patent rights, the potential market and the application field, all of which are related to economic performance (e.g. technology licensing and commercialisation) in the post-technology maturity period.

Thus far, we have investigated the effects of the aforementioned variables on patent quality. In terms of the technology life cycle, three variables—NA, DC and NSELF—are especially noteworthy. Unlike other variables, these variables have differential effects on patent quality as per the technology life cycle period (Table 8). This means that small-sized research projects or those featuring low technology cumulativeness or domestic joint research values need to be encouraged to acquire a high patent quality in the pre-technology maturity period. While a

Table 7. Estimation results by technology life cycle period and model for solar-cell technology

Variable	Pre-technology maturity period		Post-technology maturity period	
	Poisson model	Negative Binomial model	Poisson model	Negative binomial model
NA	-0.436***	-0.389	0.485*	0.574
	(0.167)	(0.398)	(0.257)	(0.664)
DC	0.614***	0.551	-1.006**	-1.291
	(0.194)	(0.493)	(0.471)	(0.978)
INTC	0.721***	0.705	0.871*	1.132
	(0.218)	(0.588)	(0.457)	(1.367)
NINV	0.018***	0.017	0.048	0.055
	(0.006)	(0.018)	(0.031)	(0.067)
2INV	-0.165**	-0.126	-0.077	0.218
	(0.077)	(0.237)	(0.267)	(0.544)
BCitation	-0.196***	-0.167***	-0.077**	-0.037
	(0.022)	(0.055)	(0.035)	(0.066)
NNONP	0.013***	0.010	0.009**	-0.012
	(0.002)	(0.007)	(0.004)	(0.013)
NSELF	0.011**	0.008	-0.052***	-0.007
	(0.005)	(0.014)	(0.019)	(0.032)
USIC	0.202***	0.177***	0.067*	0.031
	(0.022)	(0.056)	(0.038)	(0.072)
JPIC	0.170***	0.141**	0.077**	0.040
	(0.023)	(0.056)	(0.036)	(0.070)
EUIC	0.211***	0.181***	0.074**	0.015
	(0.023)	(0.058)	(0.037)	(0.073)
NCLAIM	0.006***	0.008***	0.021***	0.026***
	(0.000)	(0.002)	(0.002)	(0.010)
NFAM	0.008***	0.013***	0.034***	0.038**
	(0.001)	(0.004)	(0.004)	(0.015)
NIPC	0.122***	0.117***	0.175***	0.241*
	(0.010)	(0.033)	(0.052)	(0.144)
constant	2.231***	2.126***	-1.498***	-1.862**
	(0.170)	(0.407)	(0.291)	(0.760)
Log-likelihood	-6884.6	-3673.9	-633.9	-431.3
lnalpha		-0.153***		1.149***
		(0.047)		(0.149)

Notes: 1) *** p<0.01, ** p<0.05, * p<0.1
      2) Standard errors are in parentheses.
      3) lnalpha is the log-transformed overdispersion parameter.

large research project and the use of others' knowledge are both essential characteristics, domestic joint research needs to be avoided in the post-technology maturity period if patent quality is to be enhanced. In other words, this result highlights the importance of considering technology life cycle when developing an R&D strategy: an R&D planner within the solar-cell technology field needs to consider the size of a research project, the nature of its collaboration and accumulated knowledge in relation to the technology life cycle.

Table 8. Noteworthy effects of variables on patent quality by technology life cycle period

Variable	Pre-technology maturity period	Post-technology maturity period
NA	Increased NA → Decreased patent quality	Increased NA → Increased patent quality
DC	Adopted DC → Increased patent quality	Adopted DC → Decreased patent quality
NSELF	Increased NSELF → Increased patent quality	Increased NSELF → Decreased patent quality

CONCLUSIONS

It is important to establish an appropriate plan for an R&D strategy in order to obtain strong R&D performance in relation to a technology's position within the technology life cycle. This is because the R&D environment is rapidly changing in terms of the technology life cycle. Such a change is caused by rapid fluctuations in market needs and the fact that technology is advancing at an ever-growing pace.

The primary aim of this study was to investigate differences in the determinants of patent citation counts. It also asserts that R&D strategies for solar-cell technology should be prepared while bearing in mind the technology life cycle. The empirical results of this study are in agreement with our intuitive expectations. When a technology develops in some R&D field, the accumulation of the technological knowledge generally reaches a certain level through development from independent basic studies; thereafter, the outcomes are grafted into various fields and disperse outwards.

This empirical analysis has confirmed the veracity of intuitive knowledge that is related to the technology life cycle. Thus, to enhance the R&D performance, every R&D programme should adopt a distinct R&D strategy based on the technology period within the technology life cycle.

Since this study focused solely on the solar-cell technology field, it does have some limitations. Thus, to improve the generalisability of the results, it is necessary to undertake similar analyses in other fields and compare and analyse the results obtained with those presented here. In addition, for a more specific R&D planning, the dynamic determinants of patent citation counts should be distinguished in greater detail by further refining the technology life cycle.

REFERENCES

1. M. Hirschey and V. J. Richardson, "Valuation effects of patent quality: A comparison for Japanese and US firms", *Pacif.-Basin Finan. J.*, **2001**, *9*, 65-82

2. M. Hirschey and V. J. Richardson, "Are scientific indicators of patent quality useful to investors?", *J. Empiric. Finan.*, **2004**, *11*, 91-107.

3. D. Harhoff, F. Narin, F. M. Scherer and K. Vopel, "Citation frequency and the value of patented innovations", *Rev. Econ. Statist.*, **1999**, *81*, 511-515.

4. M. B. Albert, D. Avery, F. Narin and P. McAllister, "Direct validation of citation counts as indicators of industrially important patents", *Res. Policy*, **1991**, *20*, 251-259.

5. Y. G. Lee, J. D. Lee, Y. I. Song and S. J. Lee, "An in-depth empirical analysis of patent citation counts using zero-inflated count data model: The case of KIST", *Scientometrics*, **2007**, *70*, 27-39.

6. Z. Griliches, "Patent statistics as economic indicators: A survey", *J. Econ. Lit.*, **1990**, *28*, 1661-1707.

7. W. S. Comanor and F. M. Scherer, "Patent statistics as a measure of technical change", *J. Polit. Econ.*, **1969**, *77*, 392-398.

8. S. J. Liu and J. Shyu, "Strategic planning for technology development with patent analysis", *Int. J. Technol. Manage.*, **1997**, *13*, 661-680.

9. A. B. Jaffe, M. Trajtenberg and R. Henderson, "Geographic localization of knowledge spillovers as evidenced by patent citations", *Quart. J. Econ.*, **1993**, *108*, 577-598.

10. M. Trajtenberg, "A penny for your quotes: Patent citations and the value of innovations", *RAND J. Econ.*, **1990**, *21*, 172-187.

11. F. Narin, E. Noma and R. Perry, "Patents as indicators of corporate technological strength", *Res. Policy*, **1987**, *16*, 143-155.

12. M. P. Carpenter, F. Narin and P. Woolf, "Citation rates to technologically important patents", *World Pat. Inform.*, **1981**, *3*, 160-163.

13. B. N. Sampat and A. A. Ziedonis, "Patent citations and the economic value of patents", in "Handbook of Quantitative Science and Technology Research" (Ed. H. F. Moed, W. Glanzel and U. Schmoch), Kluwer Academic, Dordrecht, **2004**, Ch.12.

14. M. K. Fung and W. W. Chow, "Measuring the intensity of knowledge flow with patent statistics", *Econ. Lett.*, **2002**, *74*, 353-358.

15. A. G. Z. Hu and A. B. Jaffe, "Patent citations and international knowledge flow: The cases of Korea and Taiwan", *Int. J. Ind. Org.*, **2003**, *21*, 849-880.

16. A. B. Jaffe, M. Trajtenberg and M. S. Fogarty, "Knowledge spillovers and patent citations: Evidence from a survey of inventors", *Amer. Econ. Rev.*, **2000**, *90*, 215-218.

17. B. N. Sampat, "Determinants of patent quality: An empirical analysis", Working paper, Columbia University, New York, **2005**.

18. J. Mullahy, "Specification and testing of some modified count data models", *J. Econometrics*, **1986**, *33*, 341-365.

19. D. Lambert, "Zero-inflated Poisson regression, with an application to defects in manufacturing", *Technometrics*, **1992**, *34*, 1-14.

20. A. C. Cameron and P. K. Trivedi, "Regression Analysis of Count Data", Cambridge University Press, Cambridge, **1998**, pp.59-85.

21. A. C. Cameron and P. K. Trivedi, "Econometric models based on count data: Comparisons and applications of some estimators and tests", *J. Appl. Econometrics*, **1986**, *1*, 29-53.

22. J. A. Hausman, B.H. Hall and Z. Griliches, "Econometric models for count data with an application to the patents-R&D relationship", *Econometrica*, **1984**, *52*, 909-938.

23. D. Ford and C. Ryan, "Taking technology to market", *Harvard Bus. Rev.*, **1981**, *59*, 117-126.

24. T. Levitt, "Exploit the product life cycle", *Harvard Bus. Rev.*, **1965**, *43*, 81-94.

25. R. Haupt, M. Kloyer and M. Lange, "Patent indicators for the technology life cycle development", *Res. Policy*, **2007**, *36*, 387-398.

26. J. K. Park, "A study on the scope of government R&D planning: focused on the energy and resources production technology", *Econ. Environ. Geol.*, **2012**, *45*, 579-587.

27. D. M. Bagnall and M. Boreland, "Photovoltaic technologies", *Energ. Policy*, **2008**, *36*, 4390-4396

28. B. Y. Jang, Y. H. Lee, M. H. Son and J. K. Park, "Searching Convergence Technologies and Establishing Countermeasures by Using Patent Analysis", Korea Institute for Advancement of Technology, Seoul, **2010**, pp. 317-330.

29. G. P. Daines, "Patent citations and licensing value", *MBA Thesis*, **2007**, Massachusetts Institute of Technology, USA.

30. P. B. Maurseth and B. Verspagen, "Knowledge spillovers in Europe: A patent citations analysis", *Scand. J. Econ.*, **2002**, *104*, 531-545.

Maejo International Journal of Science and Technology

Communication

Geotechnical maps for recommendation on bored pile capacity in Nakhon Ratchasima municipality, Thailand

Suksun Horpibulsuk*, Nunthapon Rathanamanee, Nutthachai Prongmanee, Arnon Cholphatsorn and Avirut Chinkulkijniwat*

School of Civil Engineering, Suranaree University of Technology, 111 University Avenue, Muang District, Nakhon Ratchasima 30000, Thailand

* Corresponding authors, e-mail: suksun@g.sut.ac.th and avirut@sut.ac.th

---

Abstract: This paper presents the development of geotechnical maps in Nakhon Ratchasima municipality, Nakhon Ratchasima province, Thailand based on the boring logs and in situ test results collected from public and private sector sources. The standard penetration number, N was used to identify the soil type. The soil deposits in Nakhon Ratchasima municipality are divided into three layers: medium-to-stiff silty clay with $N < 30$, first hard silty clay with $30 < N < 50$, and second hard silty clay with $N > 50$. The medium-to-stiff silty clay layer has a thickness varying from 1.8 to 7.5 metres and an average N value of 14 with a relatively low standard deviation of 1.08. The first hard silty clay layer has a thickness varying from 1.2 to 3.0 metres and an average N value of 42 with a standard deviation of 1.37. For a practical application in foundation engineering in which the pile tips of the bored piles are located in the second hard stratum with $N > 50$, eight pile tip zones with approximated load capacity are recommended for pile lengths of 3-10 metres.

Keywords: geotechnical map, Nakhon Ratchasima municipality, standard penetration number, bored pile.

INTRODUCTION

The infrastructure in the city of Nakhon Ratchasima, Nakhon Ratchasima province, Thailand has grown rapidly in recent years. Soil deposits in this area are generally silty clay, which is wind-blown and deposited over several decades. The top soil is problematic clay sensitive to changes in water content. Laboratory and field investigations on its collapsing behaviour due to wetting were conducted by Kohgo et al. [1] and Kohgo and Horpibulsuk [2]. Due to the moderate strength of the soil when dry, many low-rise and medium-rise buildings

were constructed on shallow foundations. When soil moisture changed due to rain and waste water from the buildings, building movements occurred [3]. This movement may be either a settlement or a heave due to changes in effective stress [4-6].

To avoid this problem, piled foundations have been used to transfer the load of the super-structure to the hard soil stratum. The driven pile is not allowed in this area due to vibration problems. Dry-process bored piles have been commonly used instead because the ground water level is very deep and the soil deposit is cohesive, which can protect the borehole from collapse. A casing is therefore not required for construction. The pile tip and pile diameter are generally determined from shear-strength parameters, which are obtained from either laboratory or in situ tests. The standard penetration test (SPT) is suitable for stiff-to-hard clay. Soil samples can be obtained during the test to determine soil index properties. The standard penetration number, N, which is the number of blows per foot, is practically used to approximate the strength parameters and the pile load capacity. For clayey soils, the N value is directly related to the undrained shear strength.

Available empirical relationships between undrained shear strength, S_u, and standard penetration number, N, are limited to $N < 30$ and dependent upon soil characteristics [7,8]. Soil with higher plasticity exhibits higher undrained shear strength for the same N value. For a known clay characteristic, the available relationships are sufficient for designing a driven pile with a large section that cannot penetrate the hard clay ($N > 30$). The same is not true for a bored pile for which a borehole can be advanced by a boring machine. Horpibulsuk et al. [3] back-calculated the pile load test results for micro-piles in stiff-to-hard silty clay in the campus of Suranaree University of Technology, where the clay is of medium plasticity. They concluded that for a low N (< 30), a linear relationship between S_u and N exists and is close to that proposed by Terzaghi and Peck [7]. The same relationship was also observed for $N > 30$. The relationship between S_u and N for $N < 68$ was proposed as follows [3]:

$$S_u = \frac{2}{3} N \tag{1}$$

where S_u is expressed as ton/m^2 and N is expressed as number of blows/ft.

The traditional method for approximating pile load capacity in Nakhon Ratchasima province is the use of a static formula. The designed load capacity of a single pile is then proved by the pile load test. The designed strength parameters are obtained from a standard penetration test. The soil profile in Nakhon Ratchaisma province varies significantly because the province is located in a mountainous area. Consequently, significant boring logs are generally required for appropriate design. Geological maps showing an overview of the soil profile in Nakhon Ratchasima province are considered as a necessary tool for economical and engineering purposes. This paper aims to collect boring logs and in situ test results in Nakhon Rachasima municipality to develop geotechnical maps. The recommended pile tip zones and approximated load capacity of bored piles for different locations are finally presented based on the developed contour maps. The pile tip zones are classified according to the depth of the hard clay stratum where the N values are greater than 50. The ultimate loads for each zone are approximated using a static formula. The recommended pile tip zones and load capacity are useful for a rapid determination of the pile diameter required to attain the allowable load, and hence a cost estimate of the pile installation for contractors, designers and owners.

METHODOLOGY

Geotechnical maps for bored pile designs in Nakhon Ratchasima municipality were developed using 139 boring logs collected from public and private sector sources. Most of the data were from Mr. Taveesak Wintachai, a professional engineer, and his team. They have performed several boring and in situ tests in Nakhon Ratchasima province and the pile foundations in many construction projects were designed based on their data. Soil layers were generally classified based on the undrained shear strength approximated from the N value. Standard penetration tests were performed according to the American Society for Testing and Materials (ASTM) standards. The groundwater table was measured in all boreholes after one day of boring. Generally, the groundwater table is not detected because the groundwater is very deep in this area. Figure 1 shows a boring log in Muang district, Nakhon Ratchasima. A typical soil profile in the studied area is summarised and shown in Figure 2. The soil profile consists of a medium-to-stiff silty clay layer and the first and second hard-clay layers. The raw data include the ground elevation above mean sea level and the thickness and soil properties of the medium-to-stiff clay layer and the first hard-clay layer. The thickness of the second hard-clay layer, with N values greater than 50, is not within the scope of this study because it is impossible to advance the borehole through this layer using dry boring technique (the pile tip being located in the second hard-clay layer). The Land Desktop 2006 was used to create two types of contour maps, which are maps of ground elevation and soil thickness. For the map of ground elevation, the input data are the northern (N) and eastern (E) coordinates and ground elevation (Z). The input data are (N, E) coordinates and thickness of each soil layer for the map of soil thickness.

Based on the geotechnical maps, the pile tip zones are recommended for a rapid estimation of the pile length and diameter to attain the required load capacity. The ultimate loads for each zone were approximated for 4 pile diameters (0.35, 0.40, 0.50 and 0.6 metres) based on static formulae. These four pile diameters correspond to the typical drilling head in Thailand. The ultimate load, Q_u, consists of skin friction, Q_{su}, and end bearing resistance, Q_{bu}, which can be estimated from the following equations [9,10].

$$Q_{su} = \alpha S_u p L \tag{2}$$
$$Q_{bu} = w N_c S_u A \tag{3}$$

where α is the adhesion factor, which is taken as 0.45 for a bored pile [11]; p is the pile perimeter; L is the length of the pile shaft; N_c is the bearing capacity factor equal to 9.0 [12]; A is the cross sectional area of the pile toe and w is the reduction factor. The end bearing resistance is mainly controlled by the soil shear strength at the pile tip, where the disturbance during soil boring is commonly significant for a large pile diameter. It is thus recommended that the w value decreases as the pile diameter increases. It is equal to 0.8 or 0.75 for pile diameters less or greater than 1.0 metre respectively [11]. The allowable load for the bored pile can be obtained using a safety factor of 2.5 to 3.0.

Figure 1. Boring log in Muang district, Nakhon Ratchasima (data from Department of Public Works and Town & Country Planning)
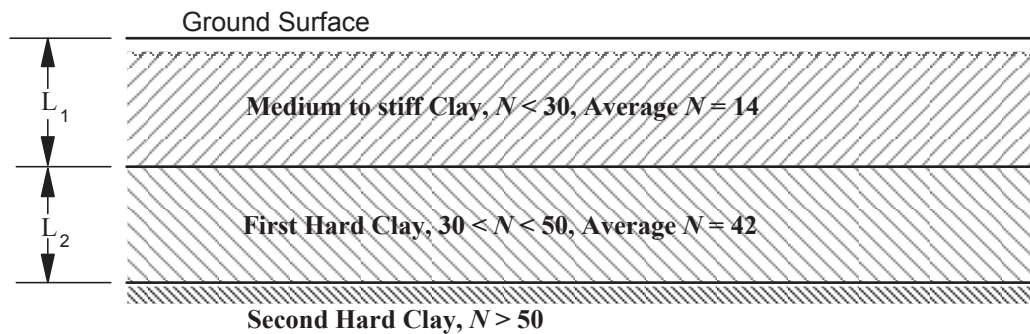
Figure 2. Typical soil profile in Nakhon Ratchasima municipality

RESULTS AND DISCUSSION

From the boring logs, the soil deposit was determined to be silty clay (Figure 2), i.e. the medium-to-stiff clay layer with N values less than 30, the first hard-clay layer with $30 < N < 50$ and the second hard-clay layer with N values greater than 50. The average N value of the medium-to- stiff silty clay layer is 14 with a very low standard deviation of 1.08, while that of the first hard-clay layer is 42 with a very low standard deviation of 1.37. Figure 3 shows the contour map of the ground elevation in Nakhon Ratchasima municipality, which varies from 169 to 216 m above mean sea level. This significant difference in ground elevation leads to a large difference in the thicknesses of the medium-to-stiff clay and hard-clay layers. The thickness of

the medium-to-stiff clay layer, L_1, varies from 1.80 to 7.00 metres (Figure 4) and the thickness of the first hard-clay layer, L_2, is 1.20 to 3.00 metres (Figure 5). The thickness of the medium-to-stiff clay layer, L_1, in Figure 4 is the difference between the depth of medium-to-stiff clay layer and the ground level obtained from Figure 3. The thickness of the first hard-clay layer, L_2, in Figure 5 is the difference between the depth of first hard-clay layer and the depth of medium-to-stiff clay layer.
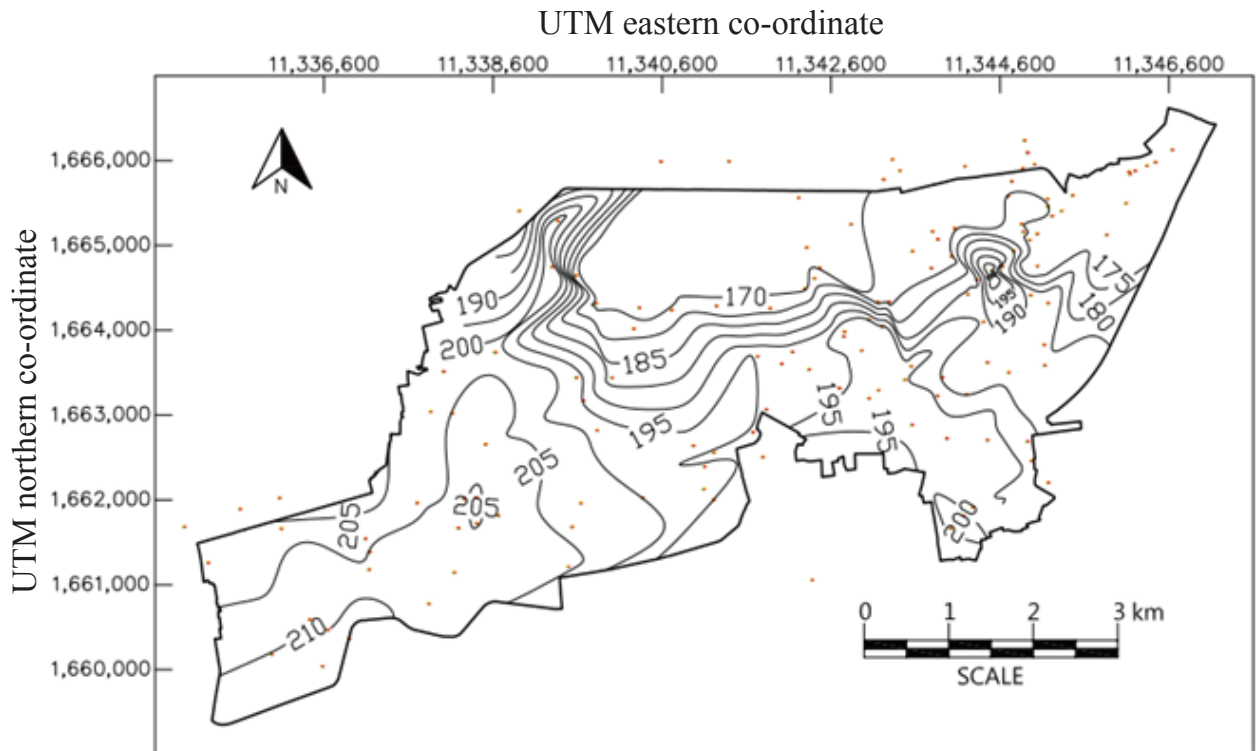
Figure 3. Contour map of ground elevation in Nakhon Ratchasima municipality. Red dots represent position of boring logs.

The minimum pile length must be greater than the sum of L_1 and L_2 to ensure that the pile tip of the bored pile is located in the second hard-clay layer. Due to a significant difference in the geotechnical profile in this area, the required pile length varies from 3 to 10 metres. Having obtained Figures 4 and 5, the average values of L_1 and L_2 were calculated and the area was zoned. For each zone, the mean standard deviation, SD, was less than 10% of the average value, showing that the value can be representative. The recommended pile tips in Figure 6 are the sum of L_1 and L_2. The approximated ultimate load for each zone is presented in Table 1. In the approximation, the w value was 0.8 for all pile diameters and the average N values of 14 and 42 were used for the medium-to-stiff clay and first hard-clay layers respectively. These N values correspond to the undrained shear strength of 9.3 and 28 ton/m^2 respectively. The average thicknesses of the medium- to-stiff clay and first hard-clay layers used for the approximation of pile load capacity are $L_1 = 1.80 \pm 0.01$ m and $L_2 = 1.20 \pm 0.14$ m for zone 1, $L_1 = 2.5 \pm 0.20$ m and $L_2 = 1.5 \pm 0.06$ m for zone 2, $L_1 = 3.90 \pm 0.15$ m and $L_2 = 1.10 \pm 0.22$ m for zone 3, $L_1 = 4.40 \pm 0.23$ m and $L_2 = 1.60 \pm 0.32$ m for zone 4, $L_1 = 4.80 \pm 0.37$ m and $L_2 = 2.20 \pm 0.27$ m for zone 5, $L_1 = 5.50 \pm 0.45$ m and $L_2 = 2.50 \pm 0.25$ m for zone 6, $L_1 = 6.30 \pm 0.75$ m and $L_2 = 2.70 \pm 0.98$ m for zone 7, and $L_1 = 7.00 \pm 0.14$ m and $L_2 = 3.00 \pm 0.28$ m for zone 8.
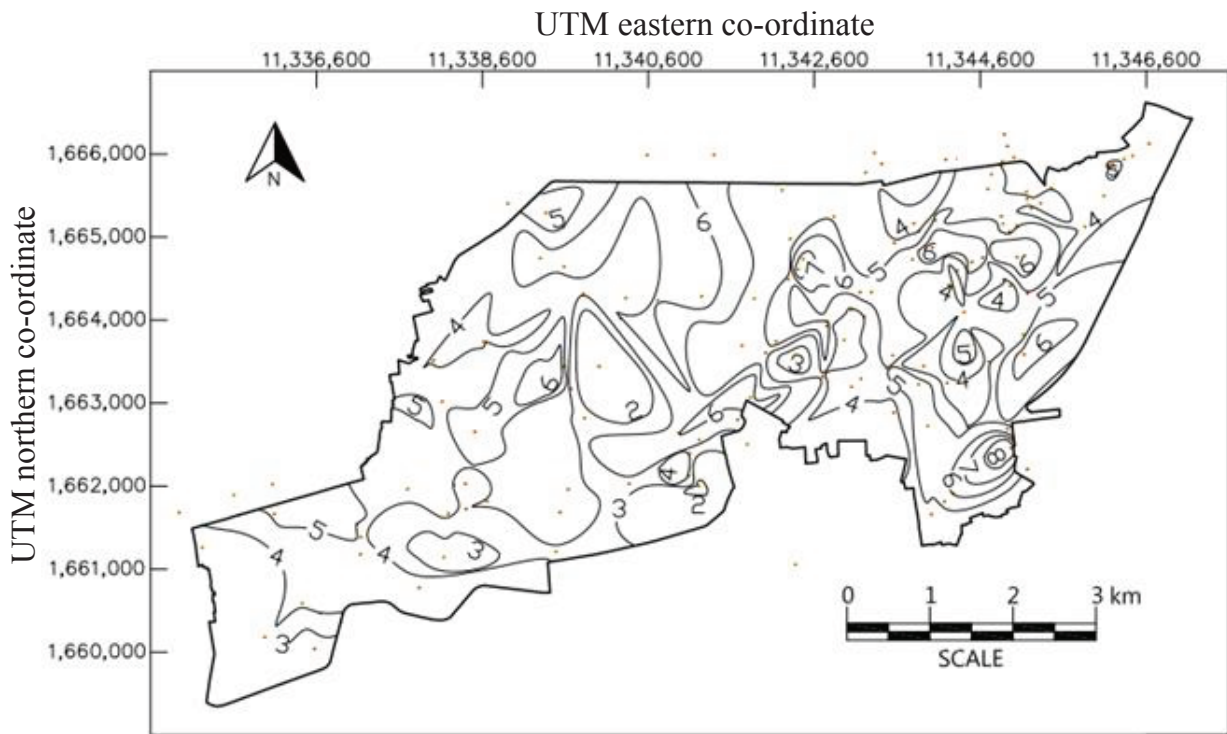
Figure 4. Contour map of the thickness of the medium-to-stiff clay layer ($N < 30$). Red dots represent position of boring logs.
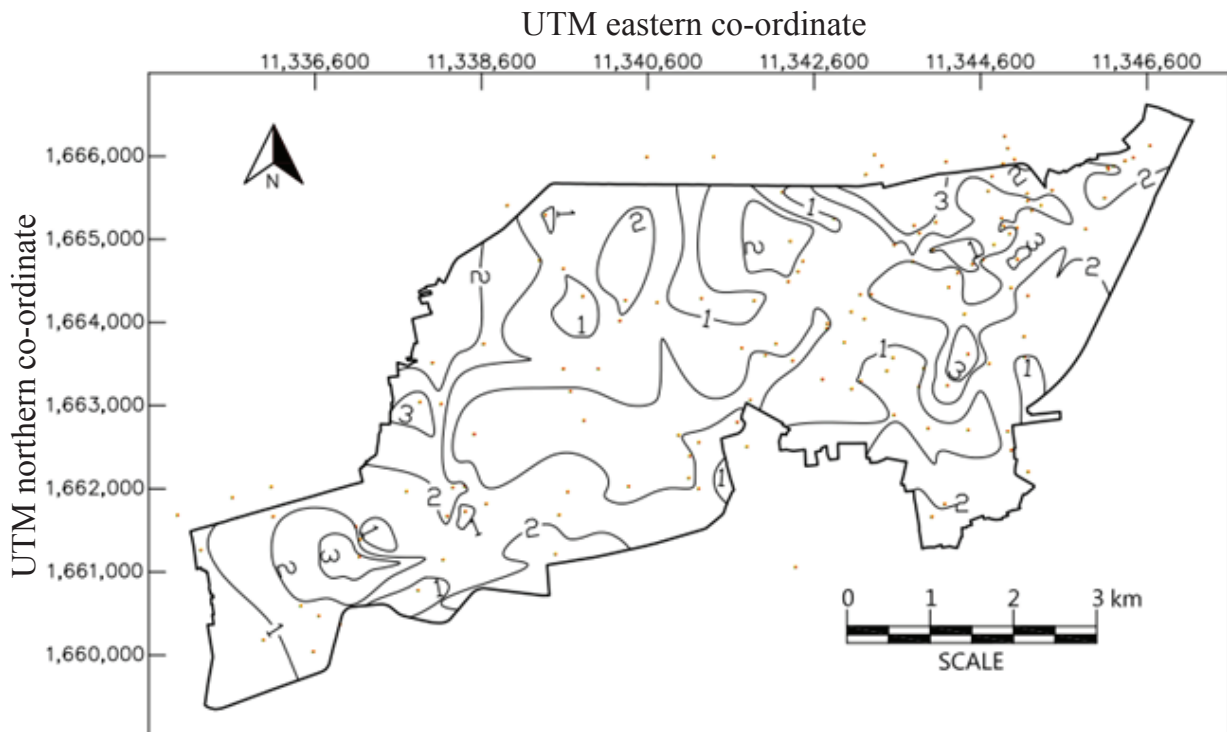
Figure 5. Contour map of the thickness of the first hard-clay layer ($30 < N < 50$). Red dots represent position of boring logs.

UTM eastern co-ordinate
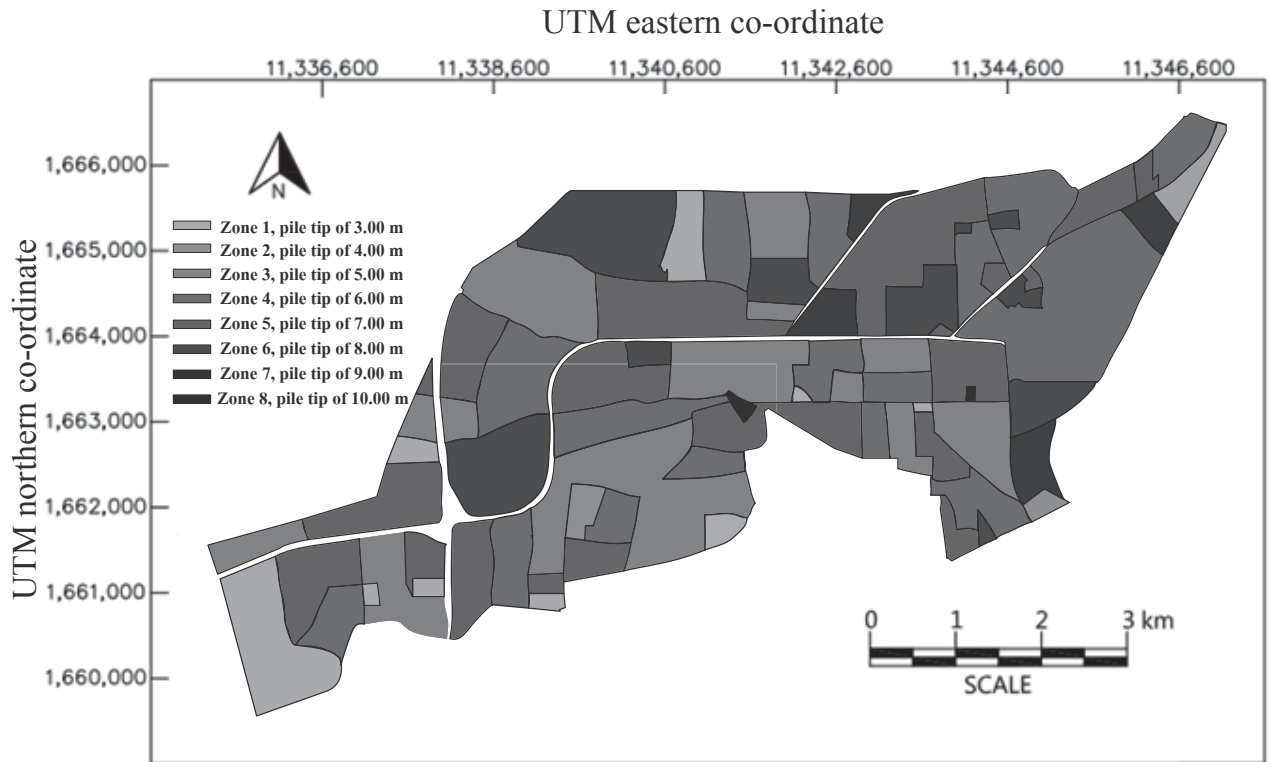
Figure 6. Recommended pile tips for eight zones in Nakhon Ratchasima municipality

Table 1. Approximated ultimate load of single pile for eight zones

Pile diameter (m)	Approximated ultimate load (ton)							
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7	Zone 8
0.35	49.4	56.8	57.7	67.0	77.1	84.5	91.0	98.4
0.4	60.5	69.0	70.0	80.6	92.2	100.6	108.0	116.5
0.5	85.7	96.2	97.6	110.7	125.3	135.8	145.1	155.6
0.6	114.8	127.5	129.1	144.9	162.3	175.0	186.1	198.7

The following is an example of the estimation of bored pile capacity in Zone 1 for a pile diameter of 0.35 m. The average L_1 and L_2 are 1.8 and 1.2 metres and the corresponding undrained shear strength values are 9.3 and 28.0 ton/m^2 respectively. Using Eq. (2), Q_{su} can be determined:

$$Q_{su1} = 0.45 \times 9.3 \times 1.10 \times 1.80 = 8.3 \ \text{ton/m}^2$$

$$Q_{su2} = 0.45 \times 28 \times 1.10 \times 1.20 = 16.6 \ \text{ton/m}^2$$

$$Q_{su} = Q_{su1} + Q_{su2} = 8.29 + 16.63 = 24.9 \ \text{ton/m}^2$$

The end bearing resistance is approximated from Eq. (3) as follows:

$$Q_{sb} = 0.85 \times 9 \times 33.33 \times 0.0962 = 24.5 \ \text{ton/m}^2.$$

Thus, the approximated load capacity is equal to 24.9 + 24.5 = 49.4 ton/m^2.

The proposed geotechnical maps and the approximated ultimate loads were determined from the available data and the empirical parameters (α and w). They are useful as a tool for a rapid estimation of the pile diameter and length as well as the installation cost. It must be stated that site investigation and pile load tests are still required prior to the installation of bored piles.

For some areas where the adjacent areas require an abrupt change of pile length, a long pile is conservatively recommended at the boundary between the two areas.

CONCLUSIONS

This paper deals with the development of geotechnical maps and the approximated ultimate load capacity in Nakhon Ratchasima municipality. The development is based on 139 boreholes which cover the studied area. The ground elevation in this area varies significantly with a maximum difference of 47 metres. The soil profile consists of 3 layers: the medium-to-stiff clay layer with average N value of 14, the first hard-clay layer with average N value of 42, and the second hard-clay layer with N value of greater than 50. The variation of the N value for each layer is considered as low with a standard deviation of less than 1.37. The contour of ground elevation and the thickness of the medium-to-stiff clay and the first hard-clay layers are introduced. The length of the bored pile varies between 3-10 metres in eight different zones. The zone map and the approximated load capacity are very useful as a practical tool for a rapid estimation of the pile diameter and length as well as the installation cost. However, the site investigation and pile load tests are still required prior to the installation of the bored piles.

ACKNOWLEDGEMENTS

REFERENCES

1. Y. Kohgo, S. B. Tamrakar and H. G. Tang, *"Investigations on the mechanical properties of typical soils distributed in northeast Thailand for the construction of irrigation facility"*, Technical Report, Japan International Research Centre for Agricultural Sciences, Tsukuba, **1997**.

2. Y. Kohgo and S. Horpibulsuk, "Simulations of volume change behavior of yellow soil sampled from Khon Kaen city in northeast Thailand", Proceedings of 11th Asian Regional Conference on Soil Mechanics and Geotechnical Engineering, **1999**, Soul, Korea, pp.141-144.

3. S. Horpibulsuk, A. Kumpala and W. Katkan, "A case history on underpinning for a distressed building on hard residual soil underneath non-uniform loose sand", *Soils Found.*, **2008**, *48*, 267-285.

4. Y. Kohgo, M. Nakano and T. Miyazaki, "Theoretical aspects of constitutive modeling for unsaturated soils", *Soils Found.*, **1993**, *33*, 49-63.

5. Y. Kohgo, M. Nakano and T. Miyazaki, "Verification of the generalized elastoplastic model for unsaturated soils", *Soils Found.*, **1993**, *33*, 64-73.

6. D. G. Fredlund and H. Rahardjo, "Soil Mechanics for Unsaturated Soils", John Wiley and Sons, New York, **1993**.

7. K. Terzaghi and R. B. Peck, "Soil Mechanics in Engineering Practice", 2nd Edn., John Wiley, New York, **1967**.

8. W. M. Zobel, "Soil Mechanics, Design Manual 7.1", Department of the Navy, Naval Facility Engineering Command, Alexandria (VA), **1982**.

9. H. G. Poulos and E. H. Davis, "Pile Foundation Analysis and Design", John Wiley and Sons, New York, **1980**.

10. T. R. Whitaker, "The Design of Piled Foundations", Pergamon Press, Oxford, **1970**.

11. A. W. Skempton, "Summing-up", Proceedings of Symposium on Large Bored Piles, **1996**, London, UK, pp. 155-157.

12. A. W. Skempton, "The bearing capacity of clay", Proceedings of Building Research Congress, **1951**, London, UK, pp.180-189.

Full Paper

Cloning and sequence of cDNA encoding 1-aminocyclo-propane-1-carboxylate oxidase in *Vanda* flowers

Noppamart Lokkamlue [1] and Pattana Srifah Huehne [1, 2, *]

[1] Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand
[2] Laboratory of Biotechnology, Chulabhorn Research Institute, Bangkok 10210, Thailand

* Corresponding author, e-mail: fscipns@ku.ac.th

Abstract: The 1-aminocyclopropane-1-carboxylate oxidase (ACO) gene in the final step of ethylene biosynthesis was isolated from ethylene-sensitive *Vanda* Miss Joaquim flowers. This consists of 1,242 base pairs (bp) encoding for 326 amino acid residues. To investigate the specific divergence in orchid ACO sequences, the deduced *Vanda* ACO was aligned with five other orchid ACOs. The results reveal that the ACO sequences within *Doritaenopsis*, *Phalaenopsis* and *Vanda* show highly conserved and almost 95% identical homology, while the ACOs isolated from *Cymbidium*, *Dendrobium* and *Cattleya* are 87–88% identical to *Vanda* ACO. In addition, the 2-oxoglutarate- Fe(II)_oxygenase (Oxy) domain of orchid ACOs consists of a higher degree of amino acid conservation than that of the non-haem dioxygenase (DIOX_N) domain. The overall homology regions of *Vanda* ACO are commonly folded into 12 α-helices and 12 β-sheets similar to the three dimensional template-structure of *Petunia* ACO. This *Vanda ACO* cloned gene is highly expressed in flower tissue compared with root and leaf tissues. In particular, there is an abundance of *ACO* transcript accumulation in the column followed by the lip and the perianth of *Vanda* Miss Joaquim flowers at the fully-open stage.

Keywords: 1-aminocyclopropane-1-carboxylate oxidase, Miss Joaquim vanda, vanda orchid, ethylene, flower senescence

_____

INTRODUCTION

The flower of *Vanda* Miss Joaquim (*Papilionanthe* Miss Joaquim, *Papilionanthe hookeriana* × *Papilionanthe teres*), the national flower of Singapore, is particularly striking in shape and style. However, it is difficult to use them as cut flowers because the removal of the pollinia (emasculation) induces severe senescence-related phenomena such as the destruction of

anthocyanin (colour fading) within 24 hours due to the release of endogenous ethylene. The flower of *Vanda* Miss Joaquim produces ethylene at a high rate and is very sensitive to ethylene [1]. Fading of the petal, also found in *Vanda* Rose Marie, first becomes evident after 8–12 hours in self-pollinated flowers, and after approximately 24–30 hours in emasculated flowers [2, 3]. In contrast, control flowers with no self-pollination or emasculation stay fresh for about 1 week and produce no ethylene. The production of ethylene in *Vanda* Miss Joaquim flower can reach the peak level of 3.442 nL/gf/hr at 32 hours after emasculation [4]. At the same ethylene concentration, *Dendrobium* Pompadour flowers, which are less sensitive to ethylene, can produce ethylene at a peak of 4.5 nL/gf/hr on the 25th day after anthesis [5] and do not display obvious senescent symptoms in the early stage of flower senescence [3, 4]. The endogenous ethylene production in ethylene-sensitive ornamental species such as the *Phalaenopsis* orchid, rose and petunia is often induced by pollination [6], and the treatment of these flowers with exogenous ethylene accelerates flower senescence or programmed cell death [7].

The ethylene biosynthetic pathway has been well documented by Yang and Hoffman [8]. Ethylene is basically synthesised by two major enzymes, namely 1-aminocyclopropane-1-carboxylic acid synthase (ACC synthase, ACS) and 1-aminocyclopropane-1-carboxylate oxidase (ACO). ACO is an ethylene-forming enzyme that oxidises the intermediate ACC to ethylene in the final step. During fruit ripening or flower senescence and wounding or elicitor treatment, the endogenous ethylene biosynthesis is highly regulated by an increases in ACS and ACO activity [9-11]. Although ethylene biosynthesis is mainly regulated by ACS, in some cases it has been suggested that the ethylene biosynthetic pathway is highly regulated by increased ACO activity during flower senescence [11]. Unlike ACS gene expression patterns, the level of ACO gene expression is suggested in terms of a developmentally regulated and tissue-specific manner [12]. Consequently, ACO has been one of the molecular markers for both ethylene formation and ethylene responsiveness [13].

ACO is classified as a member of the large 2-oxoglutarate (OG) Fe(II)–dependent dioxygenase superfamily which require ferrous ions and ascorbate for enzyme activity [14]. To date, homologous *ACO* genes have been isolated from flowers of a range of orchid species such as the self-pollinated flower of *Phalaenopsis* [15], the senescing perianth of *Doritaenopsis* [16], the necrotic flower of *Cymbidium* [17] and the fully-open flowers of *Dendrobium* [18, 19] and *Cattleya* [20]. The increasing abundance of ACO transcripts in the gynoecium and labellum of *Phalaenopsis* flower is coordinately regulated by emasculation, auxin and ethylene [12]. In addition, the dramatically increased level of ACO transcripts in half-open and fully-open *Dendrobium* flowers suggests that ACO may play a crucial role in flower opening and the senescence process [19, 21].

However, there is no information available on the *ACO* gene from *Vanda* Miss Joaquim, whose flowers are highly sensitive to endogenous ethylene, leading to the fading of flower colour within 24 hours after removal of the pollinia [1]. The *Vanda* flower has been considered to be the highest ethylene-producing tissue [22]. Thus, in the paper the *ACO* gene from *Vanda* Miss Joaquim flowers is cloned and identified as representing the mRNA transcript that is produced in ethylene-sensitive flowers. The differences in sequence homology of the *Vanda* ACO and five orchid ACOs reported in GenBank database are also analysed to determine the relationships among orchid ACOs.

MATERIALS AND METHODS

Total RNA Extraction

Samples of the flower of *Vanda* Miss Joaquim were purchased from an orchid nursery, Pathum Thani, Thailand. Four hours after removal of pollinia, the fully-open flowers were used for total RNA extraction using lithium chloride precipitation method [23]. A sample of 200 mg of flower tissue was ground to a fine powder in liquid nitrogen and then extracted with 50 μL of an extraction buffer (0.2 M Tris-HCl (pH 7.5), 0.1 M LiCl, 5 mM EDTA and 1% (w/v) sodium dodecyl sulphate). After vortexing for 30 sec., 50 μL of phenol:chloroform:isoamyl (25:24:1) was added and the mixture was incubated at room temperature. Cellular debris and denatured proteins were collected by centrifugation at 3,000×g for 20 min. at 4°C and then the supernatant was adjusted to a final concentration of 3M LiCl by adding an equal volume of 6M LiCl, whereupon the RNA precipitated overnight at 4°C. The RNA was collected by centrifugation as described above. The pellet obtained was suspended in 3M LiCl after it was collected by centrifugation; then, the pellet was resuspended in diethylpyrocarbonate-treated water. The total RNA was precipitated with ethanol and resuspended in a small volume of diethylpyrocarbonate-treated water.

Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)

The total RNA (1 μg) was reverse transcribed to cDNA by Ready-To-Go You-Prime First-Strand Beads (GE Healthcare, USA) with two primers, namely RACE-T (5'-GAC TCG AGT CGA CAT CG (T)$_{17}$-3') and 5'ACO-R (5'-AAG GAG CCG AGG TTT GAG GCC-3'). The reaction mixture (25 μL) consisted of 2.5 μL of 10× PCR buffer, 7.5 μM of each primer (forward and reverse primers), 0.5 μL of i-Taq^{TM} DNA polymerase (iNtRON Biotechnology, Korea) and 50 ng of RNA template. By amplifying the 927 bp-5' region and 579 bp-3' region of the *Vanda* ACO gene, two pairs of primers, viz. 5'ACO-F (5'-AGC CAT GGA GAG CGG AAG C-3') and 5'ACO-R; and 3'ACO-F (5'-GGG GAT CAG CTC GAG GTT ATA ACA-3') and RACE-T, were respectively generated using the following conditions: denaturising for 5 min. at 95°C followed by 10 cycles of amplification with 30 sec. of denaturising at 95°C, 1 min. of annealing at 50°C, 2 min. of extension at 72°C followed by 20 cycles of amplification with 30 sec. of denaturising at 95°C, 1 min. of annealing at 55°C, 2 min. of extension at 72°C and a final extra extension step of 10 min. at 72°C for cycle completion. The PCR products were cloned into pGEM$^{®}$-T Easy (Promega Co., USA) using the manufacturer's instructions and then were transformed into *Escherichia coli* XL1-blue strain and confirmed to be *ACO* by DNA sequencing. The full-length cDNA sequence data for genes in this article have been deposited at GenBank [24].

Primary Structure Analysis

An ACO monomer is known to have activity, although this ACO protein generally functions as a dimer. Hence, six orchid ACO sequences of *Cattleya bicolor* (GenBank ID: AAT02192), *Cymbidium* hybrid cultivar (GenBank ID: BAF36562), *Dendrobium* hybrid cultivar (GenBank ID: ABK32881), *Doritaenopsis* sp. (GenBank ID: AAA21611), *Phalaenopsis* hybrid cultivar (GenBank ID: AAR00506) and *Vanda* Miss Joaquim ACO (GenBank ID: ACS34759), and an out-group *Petunia* ACO (GenBank ID: Q08506) were retrieved from the NCBI protein database [24]. The protein sequences were aligned using Muscle alignment [25], and the phylogenetic tree of plant ACO amino acid sequences was performed with the MEGA package version 5.1 using the neighbour-joining method. The bootstrap consensus tree inferred from

1,000 replicates and branches corresponding to partitions was reproduced in less than 50% bootstrap collapsed replicates. The evolutionary distances were computed using the Dayhoff matrix–based method and were in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option) [26].

3D Structure Homology Modelling

To identify the homologous regions of the deduced protein structure, the *Vanda* ACO sequence was transformed into a 3D structure using the SWISS-MODEL repository of homology models of annotated 3D protein structures by employing the most homologous of the amino acid sequences resulting from the former 1W9Y_A sequence as standard model [27]. Figure construction and Deepview/Swiss-PdbViewer 2.SP5 were performed for superimposing [28]. The 3D structures of the *Vanda* and *Petunia* ACOs were constructed by the Accelrys DS Visualiser [29].

Gene Expression Analysis by Qualitative Real-Time PCR (qPCR)

To evaluate the *ACO* transcripts in the various tissues of *Vanda* Miss Joaquim, qPCR was performed and *5.8S rRNA* was used as the reference gene. One microgram of total RNA from each of the root, leaf, fully-open (senescing) flower, lip, perianth and column of *Vanda* Miss Joaquim was reverse transcribed to cDNA with two primers, specifically RT-ACO(R) (5′-ATG GCG GAG GAA GAA GGT GCT-3′) and 5.8S rRNA(R) (5′-GCT TGA AGC CCA GGC AGA CG-3′), using Ready-To-Go You-Prime First-Strand Beads (GE Healthcare, USA). Using the SuperScript III One-Step RT-PCR System with a Platinum® *Taq* DNA Polymerase kit (Invitrogen, USA), the 207 bp of *ACO* and 198 bp of *5.8S rRNA* were generated together using two pairs of primers: RT-ACO(F) (5′-GAC GCC TGT GAG AAC TGG GG-3′) and RT-ACO(R) for *ACO*; and 5.8S rRNA(F) (5′-ATG ACT CTC GAC AAT GGA TTT-3′) and 5.8S rRNA(R). Meanwhile, the standard curve of *ACO* gene copy numbers was constructed from six serial dilutions with final concentrations of 6.9×10^4, 4.4×10^4, 1.4×10^4, 9.2×10^3, 4.2×10^3 and 2.2×10^3 copies of the cloned *ACO* gene in pACOJ9 of pGEM-T Easy vector (Promega, USA). Each orchid sample reaction containing 0.6 μL of cDNA template along with 7.5 μM primers in a final reaction volume of 10 μL was set up in triplicate to ensure the reproducibility of the results. The real-time PCRs were generated using the following conditions: denaturising for 5 min. at 94°C followed by 35 cycles of amplification, 20 sec. of denaturising at 95°C, 15 sec. of annealing at 60°C and 30 sec. of extension at 72°C by Eppendorf Mastercycle® ep realplex real-time PCR (Eppendorf, USA). At the end of the PCR run, a melting curve was generated and analysed using the following conditions: denaturising for 15 sec. each at 95°C, 60°C and 95°C. The mean and standard deviation of the copy number of the genes were calculated. Statistical analysis at 95% significance level was performed using one-way analysis of variance, and multiple comparisons were done using Duncan's multiple range test [30].

RESULTS AND DISCUSSION

Vanda ACO cDNA and Its Organisation

cDNAs corresponding to the induced *ACO* gene were isolated from the fully-open flower of *Vanda* Miss Joaquim 4 hr after emasculation, whereupon the flowers developed colour fading within 24 hr (Figure 1). A full-length clone of *Vanda ACO* was constructed from two RT-PCR fragments of the 5′-end and the 3′-end cDNA and further subjected to sequence analysis. The obtained full-length *Vanda ACO* cDNA sequence is composed of 1,242 bp, encoding a

polypeptide of 326 amino acids (Figure 2) with a predicted molecular weight of 37.2 kDa and a calculated isoelectric point of 5.16. The *Vanda ACO* accession number submitted to the GenBank database is GQ140315 for the ACO nucleotide (*Papilionanthe hookeriana* × *Papilionanthe teres*) and ACS34759 for its deduced amino acids. Apart from the highly conserved sequence among orchid ACOs (88–95% amino acid identity), *Vanda* ACO has sequence similarity to other monocotyledon species according to a Blastp search: it shares 77, 76 and 76% amino acid identity with the ACO of Moso bamboo (*Phyllostachys pubescens*) (GenBank ID: BAB32502), Hardy sugar cane (*Saccharum arundinaceum*) (GenBank ID: ABM74187) and rice (*Oryza sativa*) (GenBank ID: EEC84681) respectively.
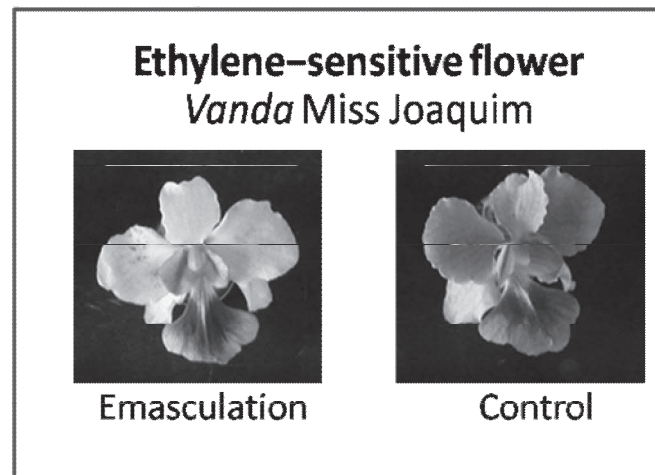
Figure 1. Flower of *Vanda* Miss Joaquim displays colour fading 24 hr after emasculation.

From the descriptions of the multi-domains in the 2OG-Fe(II) oxygenase superfamily, two highly conserved domains are disclosed in *Vanda* ACO by Blastp search (Figure 3). The DIOX_N domain of 2OG corresponding to morphine synthesis in opium poppy [31, 32] is located at the N terminal (residues 2–98). Furthermore, two potential leucine zipper domains for dimerisation of ACO proteins binding to the cell membrane [32] are predicted for leucine zipper I located at amino acid position 28–44 and for leucine zipper II located at residues 103–147 as shown in Figure 2. The second conserved domain, the Fe(II)-dependent oxygenase of the Fe(II) oxygenase superfamily (2OG-FeII_Oxy domain), is located at the residues 157–257, where an active site for a ferrous ion binding site (His181, Asp183 and His238) exists as the first template proposed in isopenicillin N synthase [32] and *Petunia* ACO crystal structure analysis [33].

The optimal orchid ACO sequence alignment shows a well conserved motif of the superfamily of Fe(II) ascorbate enzymes with a consensus His-Thr-Asp-Xaa-His-Arg sequence being the common Fe(II)-binding motif and shares a motif which is the putative co-substrate hydrogen binding site, Arg-Met-Ser, with an identical motif of ACO isolated from petunia [34]. The *Vanda* ACO amino acid sequences also contain 12 conserved residues (Pro7, Ala29, Gly34, His41, His181, Asp183, Leu199, Glu200, Gly222, His238, Arg248 and Ser250) of the ferrous ion and the ascorbate-requiring enzyme member of the Fe(II) oxygenase superfamily (Figures 2-4) corresponding to the ACO1 and ACO2 of papaya [35].

```
1    atggagagcggaagcttccctgtgattaacatggagcttctccagggtcctcagcggccg    60
     M   E   S   G   S   F   P   V   I   N   M   E   L   L   Q   G   P   Q   R   P
                             β1                α1                        α2
61   gccgccatggctcttctccgagacgcctgtgagaactggggtttcttcgagttactgaac   120
     A   A   M   A   L   L   R   D   A   C   E   N   w   G   F   F   E   L   L   N
                 α2                                              β2
121  catggaatcacccacgagctaatggaccgggtggaggctgtaaacaaggaacattaccgg   180
     H   G   I   T   H   E   L   M   D   R   V   E   A   V   N   K   E   H   Y   R
                         α3
181  cggttccgagaacagcgcttcaaagaattcgcgtccaaaaccctagattccgtggagaag   240
     R   F   R   E   Q   R   F   K   E   F   A   S   K   T   L   D   S   V   E   K
                 α3                                  α4
241  gtggactccgataatctggattgggagagcacttcttcctccgccatctccccacctcc   300
     V   D   S   D   N   L   D   W   E   S   T   F   F   L   R   H   L   P   T   S
                                 β3
301  aacatctcccaaatccccgatctggatgacgattgccggaaaacgatgaaggaattcgca   360
     N   I   S   Q   I   P   D   L   D   D   D   C   R   K   T   M   K   E   F   A
                                         α5
361  gtggagctggagaagctcgcggagagattgctggatctgttgtgcgaagatttgggtctg   420
     V   E   L   E   K   L   A   E   R   L   L   D   L   L   C   E   D   L   G   L
                         α5
421  gagaaagggtatttgaagagggttttctgcggggatcggatggcctgccgacgttcggg   480
     E   K   G   Y   L   K   R   V   F   C   G   G   S   D   G   L   P   T   F   G
                     α6                                          β4
481  acgaaggtgagtaattatccgccgtgtcccaagccggagctgataaagggattgagggct   540
     T   K   V   S   N   Y   P   P   C   P   K   P   E   L   I   K   G   L   R   A
             β4                              α7
541  catacggatgcaggagggatcattctattgtttcaggatgataaggttagcgggcttcag   600
     H*  T   D*  A   G   G   I   I   L   L   F   Q   D   D   K   V   S   G   L   Q
                 β5                                              β6
601  ctgcttaaggacggggaatggattgatgttcctccgctgcgctattccattgttgtcaat   660
     L   L   K   D   G   E   W   I   D   V   P   P   L   R   Y   S   I   V   V   N
     β6                  β7                              β8
661  attggggatcagctcgaggttataacaaatggaaaatacaagagtgtgatgcacagagtg   720
     I   G   D   Q   L   E   V   I   T   N   G   K   Y   K   S   V   M   H*  R   V
     β8          α8                                          β9
721  gtggcccaaaccaatggaaaccgcatgtcgatcgcgtcgttctacaaccccggcagcgac   780
     V   A   Q   T   N   G   N   R   M   S   I   A   S   F   Y   N   P   G   S   D
                             β10
781  gccgtgatctttccggcgccggcgctggtagagaaagaggcggaggaggaggagaagaga   840
     A   V   I   F   P   A   P   A   L   V   E   K   E   A   E   E   E   E   K   R
         β11                  α9                      α10
841  gaaacttatccaaagttcgtgttccaagactatatgaatctgtacattcgcaaaaagttc   900
     E   T   Y   P   K   F   V   F   Q   D   Y   M   N   L   Y   I   R   K   K   F
                     β12                  α11
901  gaggcgaaggagccgaggtttgaggccatgaagacaatggaagaagttgttatcagctct   960
     E   A   K   E   P   R   F   E   A   M   K   T   M   E   E   V   V   I   S   S
                         α12
961  caacctattcccactgcttgattgtgattaattggcttattatttgatgacaatttagtg  1020
     Q   P   I   P   T   A   *
1021 gattagggtttaattttatgtgttggtgttggtttgagtctagtggtttgatgtgaagat  1080
1081 ttatgtttattgtttggagtctatcatttatatatatatatatatatatatctatata  1140
1141 gttgttatctatgggtagtttggctattaaatttgttccttgtggtttgggtgtaatgaa  1200
1201 gagtgtgagtgagttaatgtaaaaaaaaaaaaaaaaaaaaaaa   1242
```

Figure 2. Nucleotide and deduced amino acid sequence of the ACO of *Vanda* Miss Joaquim. Twelve α-helices (solid line boxes) and β-sheets (dashed line boxes) are generated by Accelrys DS Visualize version 2.0.1.7347. The 12 amino acid residues conserved across all members of this enzyme which require ascorbate are indicated by bold letters together with an asterisk marking the Fe^{+2} binding residues.
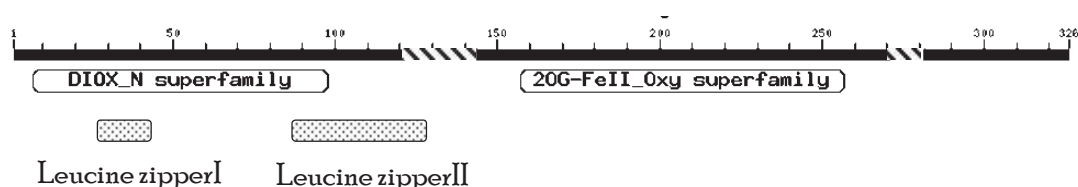
Figure 3. Diagram of three predicted conserved domains. The black bar represents a polypeptide chain with two hatched areas for spacers. The boxes indicate four conserved domains of the non-haem dioxygenase (DIOX_N) superfamily, Fe(II)-dependent oxygenase (2OG-FeII_Oxy) superfamily and two leucine zippers that are placed along the polypeptide.

Conservation of Sequence Similarity of Six Orchid ACOs

The full lengths of six orchid ACOs isolated from *Dendrobium* hybrid cultivar, *Doritaenopsis* sp., *Cattleya bicolor*, *Cymbidium* hybrid cultivar, *Phalaenopsis* hybrid cultivar and *Vanda* Miss Joaquim are 324–326 residues. These are 5–7 residues longer than the *Petunia* ACO (GenBank ID: Q08506; PDB ID: 1W9Y_A), an out-group candidate. Multiple alignments of the *Vanda* ACO–amino acid sequence with six published ACOs reveal highly conserved sequences with 91–95% identity among three orchids (*Doritaenopsis, Phalaenopsis* and *Vanda*) (Figure 4). On the other hand, the *Dendrobium, Cattleya* and *Cymbidium* ACOs show only 87–88% identity with the *Vanda* ACO sequence. However, the *Vanda* ACO shares only 69% identity with the *Petunia* ACO. The alignments of these seven ACOs demonstrate that the identities of the amino acid sequences are 94.5, 94 and 96% on average when the whole proteins, DIOX_N domain (113 amino acids) and 2OG-Fe(II)_Oxy domain (101 amino acids) of three orchid ACOs of *Doritaenopsis, Phalaenopsis* and *Vanda*, respectively, are clustered in the same subgroup.

Surprisingly, the flowers of these three orchid species are highly sensitive to endogenous ethylene after the pollinia have been dislodged. The ACOs isolated from less ethylene-sensitive species such as *Dendrobium, Cattleya* and *Cymbidium* are appropriately clustered in the second subgroup with less homology and slight divergence at identity levels of 87.33% (whole protein), 85.33% (DIOX_N domain) and 95% (2OG-Fe(II)_Oxy domain) on average. However, the amino acids of *Vanda* ACO share only 68%, 50% and 89% identity with *Petunia* ACO, DIOX_N domain and 2OG-Fe(II)_Oxy domain respectively.

The highly conserved sequences in the 2OG-Fe(II)_Oxy domain consist of the active site for Fe^{2+} binding (His181, Asp183 and His238) and the corresponding site for ascorbate binding (Arg248 and Ser250) as described by Tang and colleagues [34]. These sequences are clearly evident in the six orchid ACOs with the phylogenetic trees generated by either the complete or partial (domain) protein sequences displaying the six distinctive orchid ACOs in a similar topology to the two major subgroups of the ACOs isolated from ethylene-sensitive species (*Vanda, Phalaenopsis*, and *Doritaenopsis*) and less ethylene-sensitive species (*Cymbidium, Dendrobium* and *Cattleya*) (Figure 5). Interestingly, in this study the clustering groups of orchid ACOs classified by amino acid similarity correlate well with the characteristics of orchid flowers in terms of ethylene sensitivity, senescence phenomenon and ethylene production.

```
                                           DIOX_N domain
Phalaenopsis_ACO   MESGSFPVINMELLQGSQRPAAMALLRDACENWGFFELLNHGISHELMNRVEAVNKEHYR 60
Doritaenopsis_ACO  MESGSFPVINMELLQGSQRPAAMALLRDACENWGFFELLNHGISHELMNRVEAVNKEHYR 60
Vanda_ACO          MESGSFPVINMELLQGPQRPAAMALLRDACENWGFFELLNHGITHELMDRVEAVNKEHYR 60
Cattleya_ACO       MESGSFPVINMELLESSQRPEAMALLREACENWGFFELLNHGISTELMNRVETVNKENYR 60
Dendrobium_ACO     MESRSFPVINMELLEGSQRSDAMAVLRDACENWGFFELLNHGISHDLMNRVETVNKEHYR 60
Cymbidium_ACO      MESGSFPVINMELLEGSSRPEAMAVLRDACQNWGFFELLNHGISHELMNRVEAVNKEHYR 60
Petunia_ACO        MEN--FPIISLDKVNGVERAATMEMIKDACENWGFFELVNHGIPREVMDTVEKMTKGHYK 58
                   **..  :**.*..:.. :: :..*.. :*.: *:**.*******:****. ::*: ** :.* :*:
                                           DIOX_N domain
Phalaenopsis_ACO   RFREQRFKEFASKTLDSVENVDPDNLDWESTFFLRHLPTSNISQIPDLDDDCRATMKEFA 120
Doritaenopsis_ACO  RFREQRFKEFASKTLDSVENVDPDNLDWESTFFLRHLPTSNISQIPDLDDDCRATMKEFA 120
Vanda_ACO          RFREQRFKEFASKTLDSVEKVDSDNLDWESTFFLRHLPTSNISQIPDLDDDCRKTMKEFA 120
Cattleya_ACO       RFREQRFKEFAAKTLDSGEKVDGDNLDWESTFFLRHLPTSNISQIPDLDEDCRSTMKEFA 120
Dendrobium_ACO     RFREQRFKEFAAKTLDSGENVADNLDWESTFFLRHLPTSNISQIPDLDEDCRSTMKEFA 120
Cymbidium_ACO      RFREQRFKEFAAKTLASGEEVDSDNLDWESTFFLRHLPTSNISEVPDLDDDCRSTMKEFA 120
Petunia_ACO        KCMEQRFKELVASKALEGVQAEVTDMDWESTFFLKHLPTSNISEVPDLDEEYREVMRDFA 118
                    . ******:..:.... ...:.  :.*********:***: ****::****:: *  .*::**
                                              2OG-Fe(II)-Oxy domain
Phalaenopsis_ACO   RELEKLAERLLDLLCEDLGLEKGYLKRVFCGGSDGLPTFGTKVSNYPPCPKPDLIKGLRA 180
Doritaenopsis_ACO  RELEKLAERLLDLLCEDLGLEKGYLKRVFCGGSDGLPTFGTKVSNYPPCPKPDLIKGLRA 180
Vanda_ACO          VELEKLAERLLDLLCEDLGLEKGYLKRVFCGGSDGLPTFGTKVSNYPPCPKPELIKGLRA 180
Cattleya_ACO       RELEKLAERLLDLLCEDLGLEKGYLKRVFCGGSDGLPTFGTKVSNYPPCPKPELIKGLKA 180
Dendrobium_ACO     LELEKLAERLLDLLCEDLGLEKGYLKRVFCGGSDGLPTFGTKVSNYPPCPKPELIKGLRA 180
Cymbidium_ACO      REVEKLAESLLDLLCEDLGLEKGYLKRVFCGGSDGLPTFGTKVSNYPPCPKPELIKGLRA 180
Petunia_ACO        KRLEKLAEELLDLLCENLGLEKGYLKNAFYGSKG--PNFGTKVSNYPPCPKPDLIKGLRA 176
                    . ***** *******.********** * *.  * ***************.******.*
                                       2OG-Fe(II)-Oxy domain
Phalaenopsis_ACO   HTDAGGIILLFQDDKVSGLQLLKDGEWIDVPPLRYSIVVNIGDQLEVITNGKYKSVLHRV 240
Doritaenopsis_ACO  HTDAGGIILLFQDDKVSGLQLLKDREWIEVPPLRYSIVVNIGDQLEVITNGKYKSVLHRV 240
Vanda_ACO          HTDAGGIILLFQDDKVSGLQLLKDGEWIDVPPLRYSIVVNIGDQLEVITNGKYKSVMHRV 240
Cattleya_ACO       HTDAGGIILLFQDDKVSGLQLLKDGQWIDVPPLRHSIVVNIGDQLEVITNGKYKSVMHRV 240
Dendrobium_ACO     HTDAGGIILLFQDDKVSGLQLLKDGEWVDVPPMRHSIVVNIGDQLEVITNGRYKSVMHRV 240
Cymbidium_ACO      HTDAGGIILLFQDDKVSGLQLLKDGEWVNVPPMRHSIVINIGDQLEVITNGKYKSVLHRV 240
Petunia_ACO        HTDAGGIILLFQDDKVSGLQLLKDGQWIDVPPMRHSIVVNLGDQLEVITNGKYKSVMHRV 236
                   *****************.******* :*::***:*:***:*:**********:****:***
                   2OG-Fe(II)-Oxy domain
Phalaenopsis_ACO   VAQTDGNRMSIASFYNPGSDAVIFPAPALVEKEAEEEEKKEIYPKFVFQDYMNLYIRKKF 300
Doritaenopsis_ACO  VAQTDGNRMSIASFYNPGSDAVIFPAPALVEKEAEEEEKKEIYPKFVFQDYMNLYIRKKF 300
Vanda_ACO          VAQTNGNRMSIASFYNPGSDAVIFPAPALVEKEAEEEEKRETYPKFVFQDYMNLYIRKKF 300
Cattleya_ACO       VAQTDGNRMSIASFYNPGSDAVISPAPELVEKEAEEEEKKEATYPKFVFEDYMNLYIRKKF 300
Dendrobium_ACO     VAQTDGNRMSIASFYNPGSDAVIFPAPELVEKEAEEKKE-VYPKFVFEDYMKLYVRQKF 299
Cymbidium_ACO      VAQTDGNRMSIASFYNPGSDAVIFPAPALVEREAEEKKIE-TYPKFVFQDYMNLYVRQKF 299
Petunia_ACO        IAQKDGARMSLASFYNPGSDAVIYPAPALVEKEAE--ENKQVYPKFVFDDYMKLYAGLKF 294
                   :**.:* ***.****** ****** *** ***:***   :  ******:***:**   **
Phalaenopsis_ACO   EAKEPRFEAMKSM-EIVMSSQPIPTA 325
Doritaenopsis_ACO  EAKEPRFEAMKSM-EIVMSSQPIPTA 325
Vanda_ACO          EAKEPRFEAMKTMEEVVISSQPIPTA 326
Cattleya_ACO       EAKEPRFEAMKTM-ETVSGSQLIPTA 325
Dendrobium_ACO     EAKEPRFEAMKTM-DTVISSQLIPTA 324
Cymbidium_ACO      EAKEPRFEAMKSM-ESVSSSQLIPTA 324
Petunia_ACO        QAKEPRFEAMKAM-ETDVKMDPIATV 319
                   :**********:*  :   : *.*.
```

Figure 4. Amino acid alignment of six orchid ACO sequences assessed from the GenBank ID: ABK32881 (*Dendrobium*), AAA21611 (*Doritaenopsis*), AAT02192 (*Cattleya*), BAF36562 (*Cymbidium*), AAR00506 (*Phalaenopsis*) and ACS34759 (*Vanda*). Two conserved domains of DIOX_N and three conserved domains of 2OG- Fe(II)_Oxy are boxed by dotted lines and solid lines respectively.
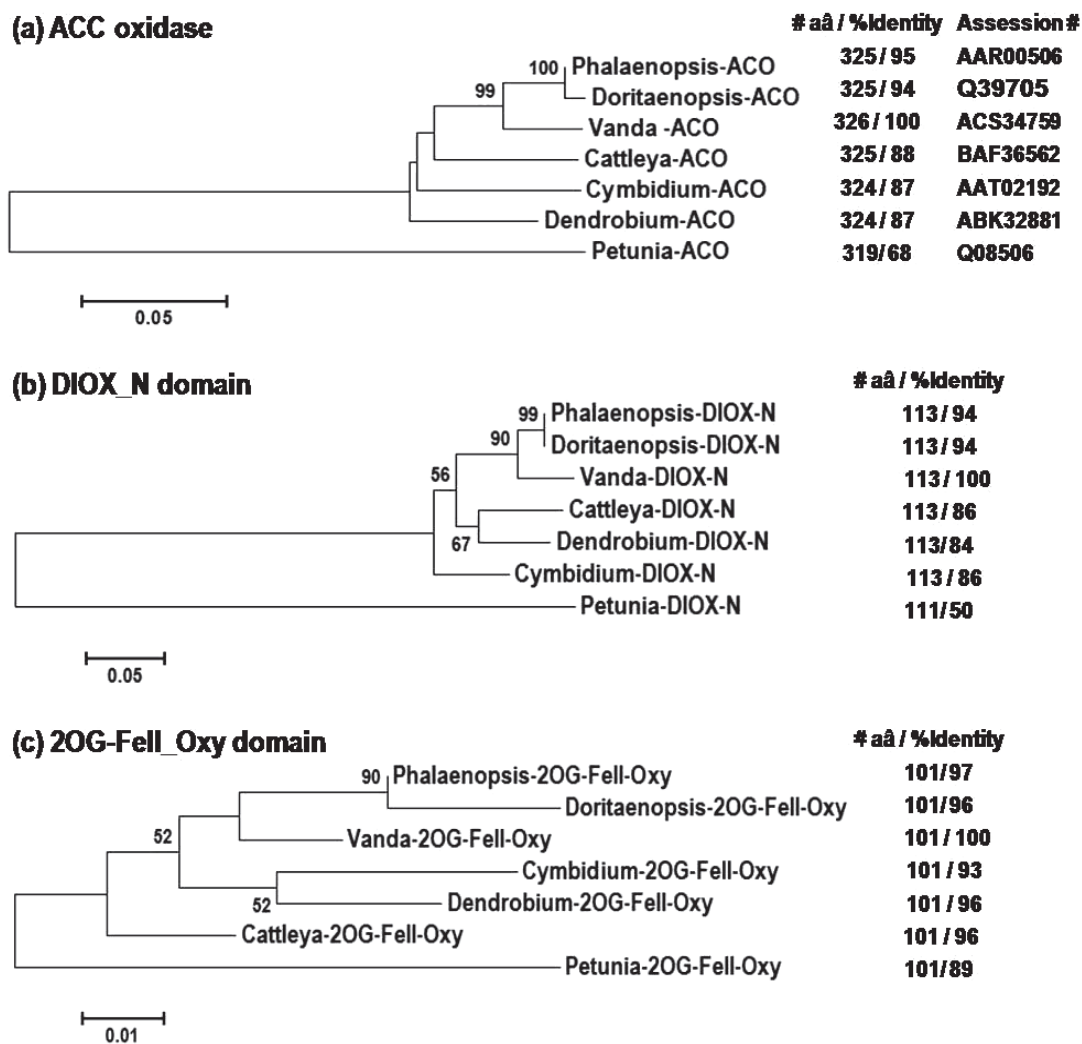
Figure 5. Neighbour-joining phylogenetic relationship of *Vanda* ACO to other orchid ACOs. The ACO proteins used in the construction are retrieved from the database as complete proteins (a), DIOX_N domain (b), and 2OG-Fe(II)_Oxy domain (c). The clustering is performed with 1,000 replicates for bootstrapping analysis using MEGA5.1 computer software, and only bootstrap values higher than 50% are indicated. The term # aa is defined as the number of amino acids.

Homologous Regions of *Vanda* and *Petunia* ACO 3D Structures

To identify the homologous regions of *Vanda* ACO structure in 3D-structure template, the initial 3D structure of *Vanda* ACO was analysed by SWISS-MODEL using *Petunia* ACO structure (PDB ID: 1W9Y_A) [33] as the template model. Both *Vanda* and *Petunia* ACO 3D structures constructed by an Accelrys DS Visualiser are comparable in folding structures with 12 α-helices complexes with 12 β-sheets, where 2 α-helices (α7 and α8) and 6 β-sheets (β4–β10) contribute to most of the 2OG-FeII_Oxy domain. By using a pairwise amino acid sequence comparison of these two proteins, highly conserved sequences are found in all α-helix and β-sheet regions. In particular, the most conserved sequence regions are located in the 2OG-Fe(II)_Oxy domain between 157–257 residues (according to *Vanda* sequences), where the active site of the Fe^{2+} binding site (His181, Asp183 and His238) is located. However, the regions in some linkage sequences between α3–α4, α4–β3, α6–β4 and α10–β12 all contain gaps which correspond to the different conformations of the 3D structures of these sequence-variable regions (Figures 6a-6b).

**(a)**   *Vanda*   *Petunia*

(b)
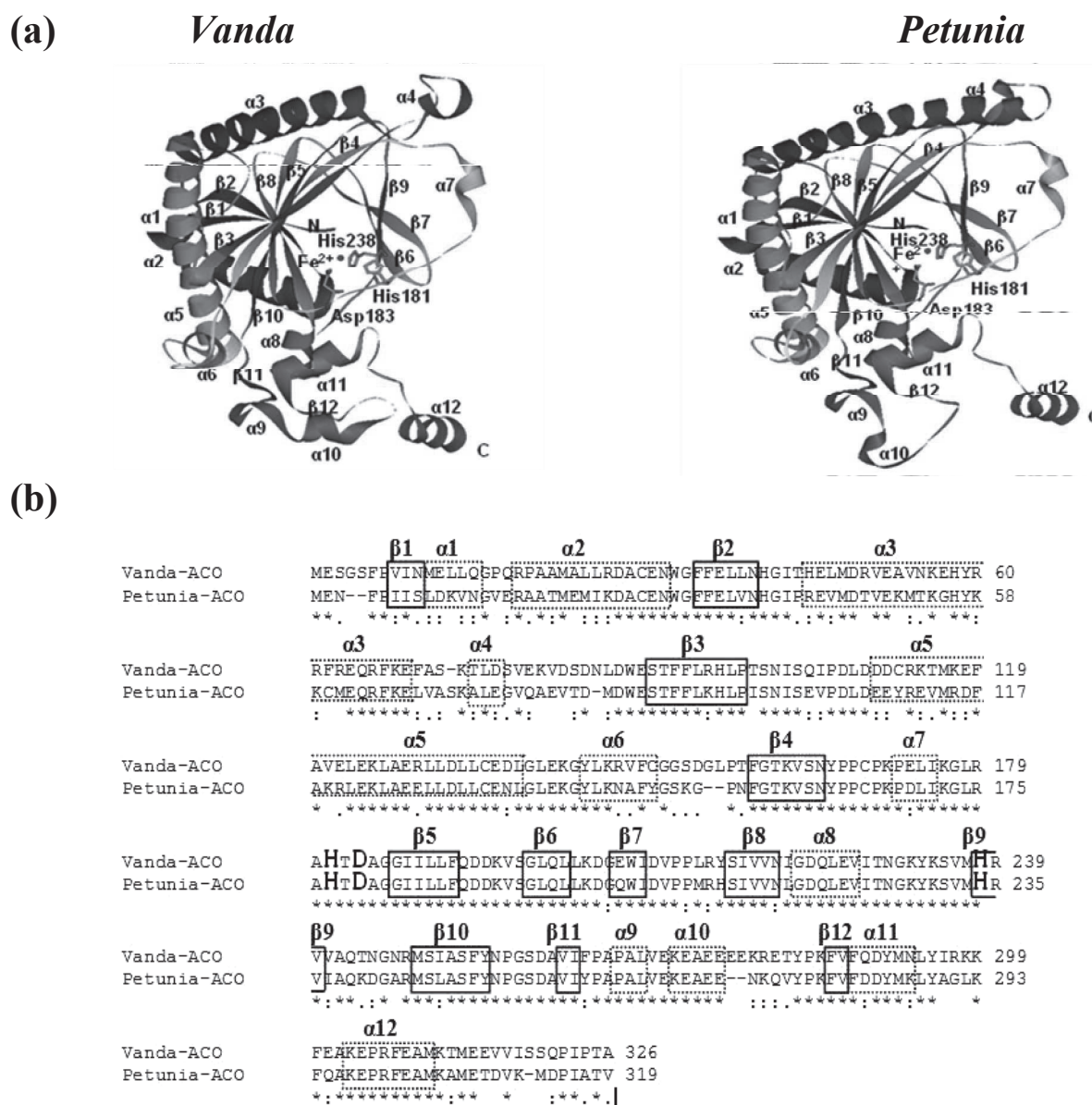
Figure 6. (a) Comparison of 3D structures of *Vanda* ACO and *Petunia* ACO (1W9Y_A) as constructed by the Accelrys DS Visualiser programme. Both the β-sheets and α-helices are numbered. The iron binding site of the 2OG-Fe(II) oxygenase superfamily consisting of three amino acid residues (His181, Asp183 and His238) together with the Fe^{2+} ion is shown. (b) Twelve α-helices (solid boxes) and β-sheets (dotted boxes) are marked and three Fe^{+2} binding residues, His181 (H), Asp183 (D) and His238 (H), are indicated by bold letters.

Vanda ACO has a 3D structural character similar to the *Petunia* template reported by Zhang and colleagues [33], regardless of the 68% sequence identity over 326 residues. The predicted folding in the core conformation region of the iron binding site is conserved and protected by two enclosing chains of α-helices, which serve as the backbone (with the N-terminal region consisting of α-1 to α-5) and the lid (with the C-terminal region consisting of α-7 to α-11). However, the variable regions between the α3 and α4 helices only create longer extended α3 helices of the *Petunia* backbone. In addition, *Petunia* flowers are more likely to be classified as less sensitive to ethylene, which is contrary to the flowers of *Vanda* Miss Joaquim [4, 36].

Expression Pattern of ACO Gene in *Vanda* Miss Joaquim

In evaluating the *ACO* mRNA transcript levels in *Vanda* orchid tissues by qPCR, the *ACO* gene (clone pACOJ9) isolated from emasculated fully-open flowers was determined to be more highly expressed in the reproductive tissues (perianth, lip and column) than in the vegetative tissues (root and leaf). In particular, the abundance of *ACO* transcripts found in the column tissue of the fully-open flower (67.7×10^3 copies/µg total RNA) was significantly greater (at 7-, 4-, 3.5- and 2.5-fold) compared to that in the root (9.8×10^3 copies/µg total RNA), perianth (16.4×10^3 copies/µg total RNA), leaf (17.6×10^3 copies/µg total RNA) and lip tissues (26.3×10^3 copies/µg total RNA) (Figure 7). However, comparable expression levels of the *ACO* transcripts were found in both leaf and perianth samples. Similar results were obtained from triplicate measurements. The mean and standard deviation of the copy number of genes were calculated throughout and the differences were considered significant at the *P* <0.05 level. The results suggest that the *Vanda* ACO clone gene obtained in this study corresponds to the flower emasculation of senescing flowers.
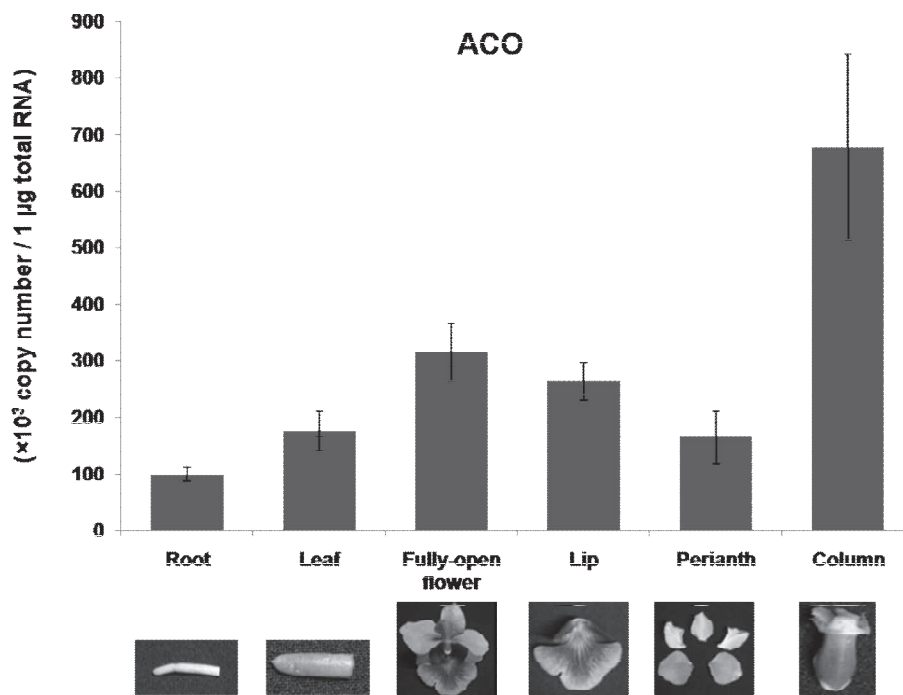
Figure 7. Expression analysis of *ACO* transcripts in various tissues of *Vanda* Miss Joaquim evaluated by quantitative real-time PCR with 5.8 *rRNA* used as internal control

Although found in most tissues of *Vanda* Miss Joaquim, the most abundant *ACO* transcript was detected early in the column tissue of the fully-open flower, followed by the lip and then much less in the perianth, leaf and root. This ACO gene expression pattern is commonly found in *Phalaenopsis* cv SM9180 flowers, which are ethylene sensitive [12]. After pollination, the stigma of the orchid column complex tissue is an initial source of ethylene production in flowers, after which it becomes a senescing organ and plays a significant role in the induction of the senescence programme and ethylene production in these organs [12, 15]. The *ACO* transcripts and activity in the column of *Phalaenopsis* cv SM9180 rapidly increase to very high

levels 6 hr after pollination but to a slower extent in the perianth, where their presence is first observed at 12 hr [15]. Ethylene is mainly produced in the orchid column, which acts as the central organ and then translocates the signal, either ethylene or ACC, to other parts of the floral organs to induce ethylene biosynthesis, in particular the sepal and petal where ACC is oxidised to ethylene [15, 37].

Meanwhile, after removal of the pollinia from the column without pollination, the endogenous ethylene biosynthesis is rapidly produced, resulting in colour fading of *Vanda* Miss Joaquim (Figure 1). It can be suggested that the *Vanda* ACO encoding gene in ethylene biosynthesis is regulated by pollinia in *Vanda* Miss Joaquim. In addition, the high level of *ACO* in the column of senescing *Vanda* Miss Joaquim is likely to be ready to convert ACC to ethylene after emasculation. The *Vanda ACO* gene therefore has a potential to be used as a molecular key for ethylene responsiveness by emasculation induced in the ethylene-sensitive flowers, which is consistent with *Phalaenopsis ACO* but contrary to the constitutive *ACO* expression in carnation and petunia styles as well as *Dendrobium* column tips [38, 39]. The dramatically increasing abundance of *ACO* transcripts in the gynoecium and labellum of *Phalaenopsis* flower is coordinately regulated by emasculation, auxin and ethylene after self-pollination [15, 37].

CONCLUSIONS

The cDNA *ACO* sequence isolated from highly ethylene-sensitive flowers of *Vanda* Miss Joaquim shares 94.5% on average of the sequences of *Phalaenopsis* and *Doritaenopsis* ACOs, whose flowers are sensitive to the induction of endogenous ethylene through the removal of pollinia from the stigma of column tissue, where most of the *ACO* transcript accumulates. The most variable regions (85.33% on average) are located in the DIOX_N domain while most sequences in the 2OG-Fe(II)_Oxy domain are conserved (95.14% on average). The variable regions verified by the 3D-structure comparison between the structures of *Vanda* ACO and *Petunia* ACO (1W9Y_A) are located in some linkage sequences. The ACO sequence analysis of the orchid *Vanda* Miss Joaquim has provided much useful information for the identification of a potential target sequence to suppress *ACO* transcripts and thus reduce ethylene production in *Vanda* orchid flowers.

ACKNOWLEDGEMENTS

REFERENCES

1. E. K. Akamine, "Ethylene production in fading *Vanda* orchid blossoms", *Science,* **1963**, *140*, 1217-1218.
2. S. P. Burg and M. J. Dijkman, "Ethylene and auxin participation in pollen induced fading of *Vanda* orchid blossoms", *Plant Physiol.*, **1967**, *42*, 1648-1650.
3. J. Arditti, "Aspects of the physiology of orchids", in "Advances in Botanical Research" (Ed. H. W. Woolhouse), Academic Press, London, **1979**, pp.421-638.
4. C. J. Goh, A. H. Halevy, R. Engel and A. M. Kofranek, "Ethylene evolution and sensitivity in cut orchid flowers", *Sci. Hortic.*, **1985**, *26*, 57-67.

5. H. Nair and T. H. Fong, "Ethylene production and 1-aminocyclopropane-1-carboxylic acid levels in detached orchid flowers of *Dendrobium* 'Pompadour'", *Sci. Hortic.,* **1987**, *32*, 145-151.

6. H. J. Rogers, "Programmed cell death in floral organs: How and why do flowers die?", *Ann. Bot.*, **2006**, *97*, 309-315.

7. H. Thomas, H. J. Ougham, C. Wagstaff and A. D. Stead, "Defining senescence and death", *J. Exp. Bot.*, **2003**, *54*, 1127-1132.

8. S. F. Yang and N. E. Hoffman, "Ethylene biosynthesis and its regulation in higher plants", *Ann. Rev. Plant Physiol.*, **1984**, *35*, 155-189.

9. C. J. Brady and J. Speirs, "Ethylene in fruit ontogeny and abscission", in "The Plant Hormone Ethylene" (Eds. A. K. Mattoo and J. C. Suttle), CRC Press, Boca Raton (FL), **1991**, pp.235-258.

10. G. Felix, D. G. Grosskopf, M. Regenass and T. Boller, "Rapid changes of protein phosphorylation are involved in transduction of the elicitor signal in plant cells", *Proc. Natl. Acad. Sci. USA*, **1991**, *88*, 8831-8834.

11. W. R. Woodson, K. Y. Park, A. Drory, P. B. Larsen and H. Wang, "Expression of ethylene biosynthetic pathway transcripts in senescing carnation flowers", *Plant Physiol.*, **1992**, *99*, 526-532.

12. S. D. O'Neill, J. A. Nadeau, X. S. Zhang, A. Q. Bui and A. H. Halevy, "Interorgan regulation of ethylene biosynthetic genes by pollination", *Plant Cell*, **1993**, *5*, 419-432.

13. S. C. Peck, K. Pawlowski and H. Kende, "Asymmetric responsiveness to ethylene mediates cell elongation in the apical hook of peas", *Plant Cell*, **1998**, *10*, 713-720.

14. A. J. Hamilton, M. Bouzayen and D. Grierson, "Identification of a tomato gene for the ethylene-forming enzyme by expression in yeast", *Proc. Natl. Acad. Sci. USA*, **1991**, *88*, 7434-7437.

15. J. A. Nadeau, X. S. Zhang, H. Nair and S. D. O'Neill, "Temporal and spatial regulation of 1-aminocyclopropane-1-carboxylate oxidase in the pollination-induced senescence of orchid flowers", *Plant Physiol.*, **1993**, *103*, 31-39.

16. J. A. Nadeau and S. D. O'Neill, "Nucleotide sequence of a cDNA encoding 1-aminocyclopropane-1-carboxylate oxidase from senescing orchid petals", *Plant Physiol.*, **1995**, *108*, 833-834.

17. S. Mita, R. Henmi and H. Ohno, "Enhanced expression of genes for ACC synthase, ACC oxidase, and NAC protein during high-temperature-induced necrosis of young inflorescences of *Cymbidium*", *Physiol. Plant.*, **2006**, *128*, 476-486.

18. Z. Razali, S. Chandran, L. A. Ling, A. N. Boyce and H. Nair, "Isolation and characterization of senescence-associated ethylene genes from *Dendrobium* orchids", in "Biotechnology and Sustainable Agriculture 2006 and Beyond" (Ed. Z. Xu, J. Li, Y. Xue and W. Yang), Springer, Dordrecht, **2007**, pp.327-332.

19. T. Nagtong, S. Thanonkeo, P. Klanrit and P. Thanonkeo, "Cloning and characterization of 1-aminocyclopropane-1-carboxylate oxidase gene from orchid (*Dendrobium* spp.)", *World Appl. Sci. J.*, **2009**, *7*, 11-18.

20. B. Q. Zheng, Y. Wang, Z. H. Peng and X. H. Li, "Cloning of ACC oxidase gene from *Cattleya* flower and construction of its plant antisense expression vector", *J. Nucl. Agric. Sci.*, **2009**, *23*, 422-446.

21. L. Lerslerwong and S. Ketsa, "Autocatalytic ethylene production by *Dendrobium* flowers during senescence induced by exogenous ethylene", *Thai J. Agric. Sci.,* **2008**, *41*, 91-99.

22. J. Arditti and C. R. Harrison, "Postpollination phenomena in orchid flowers VIII: Water and dry weight relations", *Bot. Gaz.,* **1979**, *140*, 133-137.

23. S. Lievens, S. Goormachtig and M. Holsters, "Identification of differentially expressed mRNAs using the differential display technique", Workshop on Genome Diversity and Genome Expression in Plants EMBO-Course, **1997**, Ghent, Belgium, pp.1-17.

24. National Center for Biotechnology Information, "Genomics and proteomics information", **2009**, http://www.ncbi.nlm.nih.gov (Accessed: October 2012).

25. R. C. Edgar, "Muscle: Multiple sequence alignment with high accuracy and high throughput", *Nucleic Acids Res.*, **2004,** *32*, 1792-1797.

26. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar, "MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods", *Mol. Biol. Evol.*, **2011**, *28*, 2731-2739.

27. K. Arnold, L. Bordoli, J. Kopp and T. Schwede, "The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling", *Bioinformatics*, **2006**, *22*, 195-201.

28. N. Guex and M. C. Peitsch, "SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling", *Electrophoresis*, **1997**, *18*, 2714-2723.

29. Accelrys, "DS visualizer and activeX control 3.5", **2007**, http://www.accelrys.com/products/discovery-studio/visualization-download.php (Accessed: October 2012).

30. R. J. Freund and W. J. Wilson, "Statistical Methods", 2nd Edn., Academic Press, London, **2003**.

31. J. M. Hagel and P. J. Facchini, "Dioxygenases catalyze the o-demethylation steps of morphine biosynthesis in opium poppy", *Nat. Chem. Biol.*, **2010**, *6*, 273-275.

32. P. L. Roach, I. J. Clifton, V. Fülöp, K. Harlos, G. J. Barton, J. Hajdu, I. Andersson, C. J. Schofield and J. E. Baldwin, "Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes", *Nature*, **1995**, *375*, 700-704.

33. Z. Zhang, J. S. Ren, I. J. Clifton and C. J. Schofield, "Crystal structure and mechanistic implications of 1-aminocyclopropane-1-carboxylic acid oxidase-the ethylene-forming enzyme", *Chem. Biol.*, **2004**, *11*, 1383-1394.

34. X. Tang, H. Wang, A. S. Brandt and W. R. Woodson, "Organization and structure of the 1-aminocyclopropane-1-carboxylate oxidase gene family from *Petunia* hybrid", *Plant Mol. Biol.*, **1993**, *23*, 1151-1164.

35. Y. T. Chen, Y. R. Lee, C. Y. Yang, Y. T. Wang, S. F. Yang and J. F. Shaw, "A novel papaya ACC oxidase gene (*CP-ACO2)* associated with late stage fruit ripening and leaf senescence", *Plant Sci.*, **2003**, *164*, 531-540.

36. R. Porat, Y. Reuveny, A. Borochov and A. H. Halevy, "Petunia flower longevity: The role of sensitivity to ethylene", *Physiol. Plant.*, **1993**, *89*, 291-294.

37. E. J. Woltering, "Interorgan translocation of 1-aminocyclopropane-1-carboxylic acid and ethylene coordinates senescence in emasculated *Cymbidium* flowers", *Plant Physiol.*, **1990**, *92*, 837-845.

38. J. C. Pech, A. Latché, C. Larrigaudihe and M. S. Reid, "Control of early ethylene synthesis in pollinate petunia flowers", *Plant Physiol. Biochem.*, **1987**, *25*, 431-437.

39. H. Nair, Z. Idris and J. Arditti, "Effects of 1-aminocyclopropane-1-carboxylic acid on ethylene evolution and senescence of *Dendrobium* Orchidaceae flowers", *Lindleyana*, **1991**, *6*, 49-58.