Journal Information
Maejo International Journal of Science and Technology (ISSN 1905-7873 © 2010), the international journal for preliminary communications in Science and Technology is the first peer-refereed scientific journal of Maejo University (www.mju.ac.th). Intended as a medium for communication, discussion, and rapid dissemination of important issues in Science and Technology, articles are initially published online in an open access format, which thereby gives authors the chance to communicate with a wide range of readers in an international community.

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

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

Contact Information
Editorial office: Maejo International Journal of Science and Technology (MIJST), 1st floor, Orchid Building, Maejo University, San Sai, Chiang Mai 50290, Thailand
Tel: +66-53-87-3880, Fax: +66-53-87-3880
E-mails: duang@mju.ac.th

MAEJO INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY

MAEJO INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY

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

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

MAEJO INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY

Volume 4, Issue 1 (January - April 2010)

CONTENTS

---

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

MAEJO INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY

Volume 4, Issue 1 (January - April 2010)

Author Index

Instructions for Authors

Manuscript Preparation

Manuscripts should 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 . The printing area is 17.5 cm x 24.7cm (6.9" x 9.73"). For A4 paper, margins must be 1.5 cm on top, 2.0 cm at the bottom and 1.75 cm on both left and right sides of the paper. For US letter-sized paper, margins must be 1.5 cm (0.59") on top, 1.74 cm (0.68") at the bottom and 2.05 cm (0.8") on both left and right sides. 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.

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.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), i

Maejo International Journal of Science and Technology

ISSN 1905-7873

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

Editor's Note

With the start of this volume, this journal will enter its fourth year of activity. Somewhat insecurely and in an almost frowning atmosphere, it was launched almost three years ago by the bold initiative of Maejo University's research administrator, Associate Professor Chalermchai Panyadee and also as a result of the far-sighted vision and the venturing cowboy-like spirit of the University President, Associate Professor Thep Pongpanich. As the real founders of this journal, they both have made a fair share of upgrading the journal standard for the country's scientific community.

Apparently, the journal is one of the newest ever published in this country although it seems to have gained overwhelming support from contributors both local and abroad. Unfortunately, this has forced us to reluctantly narrow down the scope of the journal in order to cope with the number of contributions. It now does not cover the field of engineering science except that of agro-engineering.[*]

We sincerely thank all the helpful members of the editorial board and all the enthusiastic contributors, both new and old, regular and casual, for choosing our journal to publish their research work. They are the real supporters and a key factor in the steady progress of this journal.

Duang Buddhasukh

[*] **Corrigendum:** The field of engineering is still covered in this volume.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 1-7

Maejo International Journal of Science and Technology

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

Full Paper

Asymptotic confidence interval for the coefficient of variation of Poisson distribution: a simulation study

Wararit Panichkitkosolkul

Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Phathum Thani, 12121, Thailand
E-mail: wararit@mathstat.sci.tu.ac.th

Received: 5 October 2009 / Accepted: 13 January 2010 / Published: 15 January 2010

Abstract: A new asymptotic confidence interval constructed by using a confidence interval for the Poisson mean is proposed for the coefficient of variation of the Poisson distribution. The following confidence intervals are considered: McKay's confidence interval, Vangel's confidence interval and the proposed confidence interval. Using Monte Carlo simulations, the coverage probabilities and expected lengths of these confidence intervals are compared. Simulation results show that all scenarios of the new asymptotic confidence interval have desired minimum coverage probabilities of 0.95 and 0.90. In addition, the newly proposed confidence interval is better than the existing ones in terms of coverage probability and expected length for all sample sizes and parameter values considered in this paper.

Keywords: coefficient of variation, confidence interval, coverage probability, expected length, Poisson distribution

Introduction

The coefficient of variation is a dimensionless number that quantifies the degree of variability relative to the mean [1]. The population coefficient of variation is defined as:

$$\kappa = \frac{\sigma}{\mu}, \quad (1)$$

where σ is the population standard deviation and μ is the population mean. The typical sample estimate of κ is given as:

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 1-7

$$\hat{\kappa} = \frac{S}{\bar{X}}, \tag{2}$$

where S is the sample standard deviation, the square root of the unbiased estimator of the variance, and $\bar{X}$ is the sample mean.

The coefficient of variation has long been widely used as a descriptive and inferential quantity in many applications of science, economics and other areas. In chemical experiments, it is often used as a yardstick of precision of measurements; two measurement methods, for example, may be compared on the basis of their respective coefficients of variation. Relative risks in finance and actuarial science can be measured using the coefficient of variation [2]. The test for equality of the coefficients of variation of two stocks can also help determine whether the two stocks possess the same risk. In physiological science, the coefficient of variation can be applied to assess the homogeneity of bone samples [3]. It has been used as a tool in uncertainty analysis of fault trees [4] and in assessing the strength of ceramics [5].

Though useful as a point estimate, perhaps for the best use of the estimated coefficient of variation it is necessary to construct a confidence interval for the population quantity [1]. A confidence interval provides much more information about the population value of the quantity of interest than does a point estimate [e.g. 6-8].

An approximate $(1-\alpha)100\%$ confidence interval for the coefficient of variation for a normal distribution is given [e.g. 9] by:

$$CI = \left\{ \frac{\hat{\kappa}}{\sqrt{t_1(\theta_1\hat{\kappa}^2+1)-\hat{\kappa}^2}}, \frac{\hat{\kappa}}{\sqrt{t_2(\theta_2\hat{\kappa}^2+1)-\hat{\kappa}^2}} \right\} \tag{3}$$

where $\nu = n-1$, $t_1 \equiv \chi^2_{\nu,1-\alpha/2}/\nu$, $t_2 \equiv \chi^2_{\nu,\alpha/2}/\nu$, and $\theta = \theta(\nu,\alpha)$ being a known function selected so that a random variable $W_\nu \equiv Y_\nu/\nu$, where Y_ν has a χ^2_ν distribution and has approximately the same distribution as a pivotal quantity $Q \equiv \frac{K^2(1+\kappa^2)}{(1+\theta K^2)\kappa^2}$. This pivotal quantity can be used to construct hypothesis tests and confidence interval for κ.

McKay [10] has proposed using $\theta = \frac{\nu}{\nu+1}$ as a good approximation for the confidence interval in equation (3) but was unable to investigate the small-sample distribution of Q. McKay's approximate confidence interval is

$$CI_1 = \left\{ \hat{\kappa}\left[\left(\frac{\chi^2_{\nu,1-\alpha/2}}{\nu+1}-1\right)\hat{\kappa}^2+\frac{\chi^2_{\nu,1-\alpha/2}}{\nu}\right]^{-1/2}, \hat{\kappa}\left[\left(\frac{\chi^2_{\nu,\alpha/2}}{\nu+1}-1\right)\hat{\kappa}^2+\frac{\chi^2_{\nu,\alpha/2}}{\nu}\right]^{-1/2} \right\} \tag{4}$$

where $\nu = n-1$, the degrees of freedom of a χ^2 distribution. Several authors have carried out numerical investigations of the accuracy of McKay's confidence interval. For instance, Iglewicz and Myers [11] compared McKay's confidence interval with the exact one based on the noncentral t distribution and found that McKay's confidence interval is efficient for $n \geq 10$ and $0 < \kappa < 0.3$.

Vangel [9] proposed a new confidence interval called the modified McKay's confidence interval for the coefficient of variation and also proposed the use of the function for θ as $\theta = \frac{\nu}{\nu+1}\left[\frac{2}{\chi^2_{\nu,\alpha}}+1\right]$, claiming that the modified McKay's method gives for the coefficient of variation

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 1-7

the confidence intervals that are closely related to McKay's confidence interval, albeit usually more accurate. The modified McKay's confidence interval for the coefficient of variation is given by:

$$CI_2 = \left\{ \hat{\kappa}\left[\left(\frac{\chi^2_{\nu,1-\alpha/2}+2}{\nu+1}-1\right)\hat{\kappa}^2+\frac{\chi^2_{\nu,1-\alpha/2}}{\nu}\right]^{-1/2}, \hat{\kappa}\left[\left(\frac{\chi^2_{\nu,\alpha/2}+2}{\nu+1}-1\right)\hat{\kappa}^2+\frac{\chi^2_{\nu,\alpha/2}}{\nu}\right]^{-1/2} \right\}. \quad (5)$$

When data are normally distributed, McKay's confidence interval (CI_1) and modified McKay's confidence interval (CI_2) can be used very well in terms of coverage probability and expected length. However, for non-normal data, these confidence intervals cannot be used in practice. The aim of this paper is to construct a new asymptotic confidence interval for the coefficient of variation of the Poisson distribution. The modified asymptotic confidence interval is obtained by applying a confidence interval for the Poisson mean. Additionally, the coverage probabilities and the expected lengths of the new and existing confidence intervals for the coefficient of variation are compared through a Monte Carlo simulation study.

New Asymptotic Confidence Interval for the Coefficient of Variation of the Poisson Distribution

In this section, a new asymptotic confidence interval for the coefficient of variation of the Poisson distribution is presented. The newly-proposed confidence interval is based on a confidence interval for the Poisson mean. The population coefficient of variation for a Poisson distribution is given by:

$$\kappa = \frac{\sigma}{\mu} = \frac{\sqrt{\lambda}}{\lambda} = \frac{1}{\sqrt{\lambda}}$$

In order to construct a new asymptotic confidence interval, first, a confidence interval for the Poisson mean is used. One $(1-\alpha)100\%$ confidence interval for the Poisson mean with continuity correction [12] is defined as:

$$\left(\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}} \ , \ \bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}} \right), \quad (6)$$

where $X_i \sim Poi(\lambda), i = 1,2,...,n$, $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, and $Z_{1-\frac{\alpha}{2}}$ is a $\left(1-\frac{\alpha}{2}\right)$th quantile of the standard normal distribution. From Equation (6), a confidence interval for the coefficient of variation of the Poisson distribution can be derived as follows:

$$P\left(\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}} < \lambda < \bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}} \right) = 1-\alpha$$

$$1-\alpha = P\left(\sqrt{\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}} < \sqrt{\lambda} < \sqrt{\bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}} \right)$$

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 1-7

$$= P\left(\frac{1}{\sqrt{\bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}}} < \frac{1}{\sqrt{\lambda}} < \frac{1}{\sqrt{\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}}} \right)$$

$$= P\left(\frac{1}{\sqrt{\bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}}} < \kappa < \frac{1}{\sqrt{\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}}} \right). \quad (7)$$

Thence, the new $(1-\alpha)100\%$ asymptotic confidence interval for the coefficient of variation of the Poisson distribution is obtained, which is:

$$CI_3 = \left\{ \left(\sqrt{\bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}} \right)^{-1}, \left(\sqrt{\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}+0.5}{n}}} \right)^{-1} \right\}. \quad (8)$$

Monte Carlo Simulation Results

The performance of the estimated coverage probabilities of the confidence intervals (4), (5) and (8) and their expected lengths were examined via Monte Carlo simulations, with particular emphasis on comparison between the new and existing approaches. Data were generated from Poisson distribution with κ = 0.1, 0.2 and 0.3, and sample size n = 10, 15, 25, 50 and 100. All simulations were performed using programs written in the R statistical software [13-15] with the number of simulation runs M = 50,000 at level of significance $\alpha = 0.05$ and 0.10. The simulation results are shown in Tables 1-2, in which the following information, viz. the estimated coverage probabilities of the confidence intervals in (4), (5) and (8) (CI_1, CI_2 and CI_3) and their expected lengths for the coefficient of variation of a Poisson distribution at α = 0.05 and 0.10, is presented respectively. As can be seen the Tables, all of the confidence intervals, CI_1, CI_2 and CI_3, have minimum coverage probability of $1-\alpha$ for all sample sizes and values of κ. The estimated coverage probability increases as the value of κ gets larger (e.g. for CI_3, n=10 and $\alpha = 0.05$; 0.9506 for κ = 0.1; 0.9511 for κ = 0.2; and 0.9559 for κ = 0.3). In addition, the empirical coverage probabilities of the proposed confidence interval, CI_3, are closer to the nominal value of $1-\alpha$ than those of CI_1 and CI_2. Furthermore, the expected lengths of the proposed confidence interval, CI_3, are much shorter than those of CI_1 and CI_2 in all conditions. The expected length increases as the value κ gets larger (e.g. for CI_3, n=10 and $\alpha = 0.05$; 0.0062 for κ = 0.1; 0.0254 for κ = 0.2; and 0.0590 for κ = 0.3). Moreover, when the sample size increases, the expected length is shorter (e.g. for CI_3, κ=0.1 and $\alpha = 0.05$; 0.0062 for n=10; 0.0051 for n=15; 0.0039 for n=25; 0.0028 for n=50; and 0.0020 for n=100).

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 1-7

Table1. Estimated coverage probabilities and expected lengths of 95% confidence intervals in (4), (5) and (8) for a Poisson distribution

n	κ	Coverage probability			Expected length		
		CI_1	CI_2	CI_3	CI_1	CI_2	CI_3
10	0.1	0.9506	0.9508	0.9506	0.1133	0.1126	0.0062
	0.2	0.9556	0.9565	0.9511	0.2449	0.2387	0.0254
	0.3	0.9618	0.9631	0.9559	0.4248	0.3958	0.0590
15	0.1	0.9517	0.9518	0.9514	0.0845	0.0843	0.0051
	0.2	0.9556	0.9561	0.9520	0.1786	0.1764	0.0206
	0.3	0.9625	0.9628	0.9540	0.2940	0.2850	0.0476
25	0.1	0.9522	0.9524	0.9521	0.0612	0.0611	0.0039
	0.2	0.9575	0.9577	0.9520	0.1278	0.1270	0.0159
	0.3	0.9632	0.9635	0.9539	0.2057	0.2027	0.0365
50	0.1	0.9519	0.9518	0.9515	0.0414	0.0413	0.0028
	0.2	0.9568	0.9569	0.9514	0.0857	0.0855	0.0112
	0.3	0.9632	0.9634	0.9536	0.1362	0.1354	0.0257
100	0.1	0.9516	0.9516	0.9512	0.0286	0.0286	0.0020
	0.2	0.9560	0.9560	0.9517	0.0591	0.0590	0.0079
	0.3	0.9632	0.9635	0.9541	0.0933	0.0930	0.0181

Table2. Estimated coverage probabilities and expected lengths of 90% confidence intervals in (4), (5) and (8) for a Poisson distribution

n	κ	Coverage probability			Expected length		
		CI_1	CI_2	CI_3	CI_1	CI_2	CI_3
10	0.1	0.9022	0.9025	0.9021	0.0909	0.0905	0.0052
	0.2	0.9045	0.9053	0.9023	0.1936	0.1895	0.0212
	0.3	0.9193	0.9206	0.9100	0.3255	0.3081	0.0491
15	0.1	0.9031	0.9033	0.9020	0.0690	0.0688	0.0043
	0.2	0.9114	0.9118	0.9009	0.1449	0.1432	0.0173
	0.3	0.9188	0.9197	0.9014	0.2354	0.2291	0.0398
25	0.1	0.9034	0.9035	0.9005	0.0506	0.0505	0.0033
	0.2	0.9084	0.9094	0.9023	0.1053	0.1047	0.0134
	0.3	0.9209	0.9215	0.9059	0.1686	0.1664	0.0306
50	0.1	0.9029	0.9030	0.9006	0.0344	0.0344	0.0023
	0.2	0.9106	0.9109	0.9045	0.0713	0.0711	0.0094
	0.3	0.9179	0.9188	0.9035	0.1130	0.1123	0.0215
100	0.1	0.9016	0.9017	0.9004	0.0239	0.0239	0.0016
	0.2	0.9093	0.9096	0.9044	0.0494	0.0493	0.0067
	0.3	0.9206	0.9208	0.9072	0.0779	0.0777	0.0152

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 1-7

Conclusions

A new asymptotic confidence interval which is based on a confidence interval for the Poisson mean has been developed for the coefficient of variation of the Poisson distribution. The developed confidence interval was compared with those of McKay and Vangel through a Monte Carlo simulation study. All confidence intervals have minimum coverage probabilities $1-\alpha$. The estimated coverage probabilities and the expected lengths of both McKay's and Vangel's confidence intervals are slightly different when sample sizes are small while they are virtually the same when sample sizes are large. The new asymptotic confidence interval proposed in this paper has several advantages over the existing confidence intervals of McKay and Vangel. Firstly, it has the estimated coverage probabilities which are closer to the nominal value of $1-\alpha$. Secondly, it has shorter expected lengths than those of the other two confidence intervals while having reasonable coverage probabilities. Thus, if a confidence interval with minimum coverage probability equal to a pre-specified value and with a shorter expected length is preferred, the newly proposed confidence interval may be the one of choice.

Acknowledgements

The author would like to thank the editor and the three anonymous referees for their suggestions and helpful comments in improving this paper. The author also acknowledges the excellent comments provided by Dr. Gareth Clayton on earlier drafts of this paper.

References

1. K. Kelley, "Sample size planning for the coefficient of variation from the accuracy in parameter estimation approach", *Behav. Res. Meth.*, **2007**, *39*, 755-766.
2. E. G. Miller and M. J. Karson, "Testing the equality of two coefficients of variation", American Statistical Association: Proceedings of the Business and Economics Section, Part I, **1977**, pp. 278-283.
3. A. J. Hamer, J. R. Strachan, M. M. Black, C. Ibbotson and R. A. Elson, "A new method of comparative bone strength measurement", *J. Med. Eng. Technol.*, **1995**, *19*, 1-5.
4. K. Ahn, "On the use of coefficient of variation for uncertainty analysis in fault tree analysis", *Reliab. Eng. Syst. Safe.*, **1995**, *47*, 229-230.
5. J. Gong and Y. Li, "Relationship between the estimated Weibull modulus and the coefficient of variation of the measured strength for ceramics", *J. Am. Ceram. Soc.*, **1999**, *82*, 449-452.
6. M. Smithson, "Correct confidence intervals for various regression effect sizes and parameters: The importance of noncentral distributions in computing intervals", *Edu. Psychol. Meas.*, **2001**, *61*, 605-632.
7. B. Thompson, "What future quantitative social science research could look like: Confidence intervals for effect sizes", *Edu. Researcher*, **2002**, *31*, 25-32.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 1-7

8. J. H. Steiger, "Beyond the F test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis", *Psychol. Meth.*, **2004**, *9*, 164-182.
9. M. G. Vangel, "Confidence intervals for a normal coefficient of variation" *Amer. Statist.*, **1996**, *50*, 21-26.
10. A. T. McKay, "Distribution of the coefficient of variation and the extended t distribution", *J. Roy. Statist. Soc. Ser. B*, **1932**, *95*, 695-698.
11. B. Iglewicz and R. H. Myers, "Comparisons of approximations to the percentage points of the sample coefficient of variation", *Technometrics*, **1970**, *12*, 166-169.
12. L. Barker, "A comparison of nine confidence intervals for a Poisson parameter when the expected number of events is ≤ 5", *Amer. Statist.*, **2002**, *56*, 85-89.
13. The R Development Core Team, "An Introduction to R", R Foundation for Statistical Computing, Vienna, **2008a**.
14. The R Development Core Team, "R: A Language and Environment for Statistical Computing", R Foundation for Statistical Computing, Vienna, **2008b**.
15. R. Ihaka and R. Gentleman, "R: A language for data analysis and graphics", *J. Comput. Graph. Statist.*, **1996**, *5*, 299-314.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

Maejo International Journal of Science and Technology

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

Full Paper

Development of carvedilol assay in tablet dosage form using HPLC with fluorescence detection

Pattana Sripalakit [1,*], Somsak Kaewnok [2] and Sakawrat Tubtonglang [2]

[1] Department of Pharmaceutical Chemistry and Pharmacognosy, Faculty of Pharmaceutical Sciences, Naresuan University, Phitsanulok 65000, Thailand
[2] Bioequivalence Test Centre, Faculty of Pharmaceutical Sciences, Naresuan University, Phitsanulok 65000, Thailand
[*] Corresponding author, e-mail: pattanas@nu.ac.th

Received: 23 July 2009 / Accepted: 22 January 2010 / Published: 2 February 2010

Abstract: A simple HPLC method was developed and validated for quantitation of carvedilol in dissolution medium and tablet dosage form. Chromatographic separation was achieved on a Alltima® C18 (250 mm×4.6 mm) column using a mobile phase containing 0.01 M Na_2HPO_4 in water and acetonitrile (30:70 v/v) adjusted to pH 3.0 by ortho-phosphoric acid at a flow rate of 1.0 ml/min and employing fluorescence detection with 300- nm excitation and 343-nm emission wavelengths. The method was validated for specificity, linearity, accuracy, precision and stability. Dissolution test parameters were also investigated. Moreover, the proposed analytical method was applied to monitor the formulation content uniformity and labelled amount of commercially available carvedilol drugs.

Keywords: carvedilol, HPLC, validation, dissolution test, quality control

Introduction

Carvedilol, or (±)-1-9H-(carbazol-4-yloxy)-3-[[2-(2-methoxyphenoxy)ethyl]amino]-2-propanol (Figure 1), is an antihypertensive agent with β- and α_1-adrenergic receptor blocking activities [1-3]. Carvedilol has much greater antioxidant activity than other commonly-used β-blockers [4-5]. It has been prescribed as an antihypertensive agent and an angina agent [6-7] and for treatment of congestive heart failure [8].

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

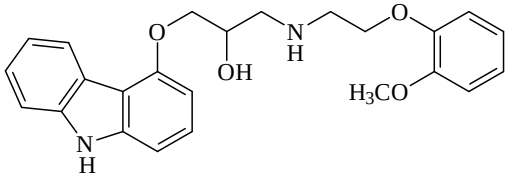


Figure 1. Chemical structure of carvedilol

High-performance liquid chromatography (HPLC) with fluorescence detector [9-14], mass spectrometer [15-16] or electrochemical detection [17] has been used for the analysis of carvedilol and its enantiomers in biological samples. Determination of cavedilol by capillary electrophoresis has also been reported [14,18]. There have been few published articles on the evaluation of carvedilol in pharmaceutical formulations. That using HPLC with UV detector [19-21] and differential pulse voltammetric determination [22] have been presented.

The dissolution test and quantitative assay are very important features of the quality control of drugs in the pharmaceutical industry. The dissolution test is currently used as an in vitro bioequivalence test and generally for obtaining dissolution profile and profile comparison to establish the similarity of pharmaceutical dosage forms [23-24]. To the best of our knowledge, there is no official assay guideline for carvedilol in dosage forms and dissolution samples in any pharmacopoeia, nor any dissolution test for this pharmaceutical in dosage forms reported in the literature. Thus, in this paper an attempt is made to develop and validate a simple, efficient and reliable method for the determination of carvedilol intended for pharmaceutical applications by HPLC using fluorescence detection. Carvedilol assay in tablet formulation and dissolution samples is described and the optimisation of a dissolution protocol for carvedilol-containing tablets is presented. Evaluation of the dissolution profiles of two marketed carvedilol products by the optimised method is also reported.

Materials and Methods

Chemicals and reagents

Standard carvedilol (99.91%) was obtained from Salutas Pharma GmbH (Barleben, Germany) and was used as certified reference compound for quantitative analysis. Other chemicals were of analytical reagent grade purchased from various sources. All solvents were of HPLC grade obtained from VWR Prolabo (Leuven, Belgium). All experiments were performed with purified water obtained from TKA ROS 300 (Niederelbert, Germany).

Chromatographic conditions

The HPLC system used for the assay consisted of a dual plunger pump (LC-10ATVP, Shimadzu, Kyoto, Japan), a fluorescence detector (RF-10AXL, Shimadzu), a system controller (SCL-10AVP, Shimadzu) and a Rheodyne (7725) sample injector (Rohnert Park, CA) fitted with a 20-µl sample loop. The separation was performed at ambient temperature on an Alltima® C18 (250 mm×4.6 mm i.d., 5 µm, 250 Å) column purchased from Alltech (Deerfield, IL). The column was fitted with a guard column packed with C18 (4.0 mm×3.0 mm i.d.; Phenomenex Torrance, CA). The mobile phase was a

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

mixture of 0.01 M Na_2HPO_4 in water and acetonitrile (30:70 v/v) adjusted to pH 3.0 by ortho-phosphoric acid and had a flow rate of 1.0 ml/min. The mobile phase was degassed by ultrasonication prior to use and was allowed to recirculate during the analysis. The peak areas were determined using a fluorescence detector with excitation wavelength and emission wavelength set at 300 nm and 343 nm respectively [25].

Preparation of standard solutions

A stock solution of carvedilol (1 mg/ml) was prepared with the mobile phase as solvent. Calibration standard solutions were prepared by diluting the stock solution to 1, 5, 10, 20, 30 and 40 μg/ml with the dissolution medium.

Analytical method validation

Specificity: Specificity was assessed by examining peak interferences from dissolution medium. This was done by inspecting chromatograms of blank and spiked medium samples.

Linearity: Six-point standard calibration curves were prepared over a concentration range of 1-40 μg/ml for carvedilol. The data of peak area versus drug concentration were constructed by unweighted least-square linear regression analysis.

Accuracy and precision: Accuracy and precision were determined from six replicates of each carvedilol concentration (1, 5, 10, 20, 30 and 40 μg/ml) within the range of the calibration curve. Accuracy and precision were expressed as % accuracy and % coefficient of variation (CV) respectively.

Stability: The dissolution medium containing standard carvedilol was kept at 37±0.5°C for 2 hr under light shaking and then left at room temperature for 24 hr. The response of the 24-hr aged solutions was evaluated against a freshly-prepared standard solution.

Dissolution

Dissolution of carvedilol tablets was optimised using Dilatrend® (carvedilol tablets: 6.25, 12.5 and 25 mg, manufactured by Roche S.p.A., Segrate, Italy). In each experiment, twelve tablets were randomly selected. Dissolution testing was performed in accordance with the USP <711> [26] using apparatus II (VK 10-1500, Vankel Industries Inc., Cary, NC). The dissolution apparatus was used with paddles at 50 rpm and a bath temperature of 37±0.5°C. The dissolution media were evaluated using 0.1 N HCl solution (pH 1.2), acetate buffer (pH 4.5) and phosphate buffer (pH 6.8) [26]. Dissolution was carried out according to the drug release guidelines [27]; 900 ml of the freshly prepared medium was used in a rotating vessel. At each sampling time point, the dissolution sample (5 ml) was collected from each vessel and filtered through a 0.45-μm porosity nitrocellulose membrane (Millipore, Bedford, MA). Fresh medium (5 ml) weas replaced in each vessel after sampling. A 20-μl aliquot of each sample was injected into the HPLC system for analysis. The quantity of carvedilol in the dissolution medium was calculated from a calibration curve. The results were estimated as % labelled amount of the dissolved active ingredient.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

Application to drug quality controls: dissolution profile comparison

Dilatrend® (6.25, 12.5 and 25 mg) as reference product and the same does of Brand A (generic carvedilol tablets) as test product were studied. The procedure for dissolution as previously described above was followed. According to the US FDA guideline [28] for dissolution profile comparisons, the difference factor (f_1) and similarity factor (f_2) were calculated as follows:

$$f_1 = \{(\Sigma^n_{t=1} | R_t - T_t |) / (\Sigma^n_{t=1} R_t)\} \times 100$$

$$f_2 = 50 \times \log \{(1 + (1/n) \Sigma^n_{t=1} (R_t - T_t)^2)^{-0.5} \times 100\}$$

in which R_t and T_t are the percentages of Dilatrend® and Brand A respectively that were dissolved at each time point, and n is the number of sampling time points.

Assay in tablet formulation

Standard preparation: An accurately weighed quantity of carvedilol working standard was dissolved in the mobile phase to afford a solution having a concentration of 0.025 mg/ml.

Assay preparation: Twenty tablets of the test or reference product were weighed and then finely powdered. An accurately weighed portion of the powder, equivalent to about 12.5 mg of carvedilol, was transferred to a 50-ml volumetric flask and 20 ml of mobile phase was added. The volumetric flask was shaken mechanically for 5 min, sonicated for 10 min and diluted to volume. One ml of this solution was transferred to a 10-ml volumetric flask and diluted with mobile phase to volume. A portion of this solution was filtered through a 0.45-μm-porosity nylon filter membrane (Millipore).

Procedure: A 20-μl aliquot of standard or sample preparation (test and reference products) was injected into the HPLC system described above. The quantity (in mg) of carvedilol in the portion of tablets was obtained by the formula: 500 C (r_U / r_S), in which 500 is the dilution factor, C is the concentration (in mg/ml) of carvedilol in the standard preparation, and r_U and r_S are the carvedilol peak areas obtained from the assay and standard preparations respectively. The results were then estimated as % labelled amount.

For the determination of dosage-unit uniformity by assay of individual units [26], 10 units each of the test and reference products were selected. Each tablet was finely powdered, transferred to a 50-ml volumetric flask and diluted to volume. A portion of this solution was transferred to a 10-ml volumetric flask and diluted with mobile phase to volume, to afford a final concentration of 0.025 μg/ml. A portion of this solution was filtered through a 0.45-μm-porosity nylon filter membrane and the filtrate (20μl) was analysed by HPLC in the same manner as above. The content of carvedilol in each tablet was calculated by comparison with the standard solution at 0.025 μg/ml.

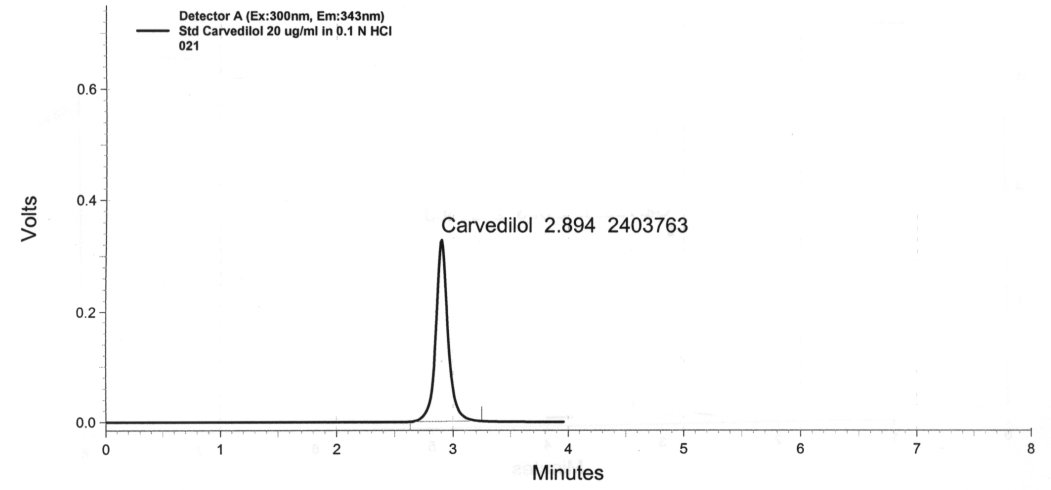
Results and Discussion

Specificity and optimisation of chromatographic conditions

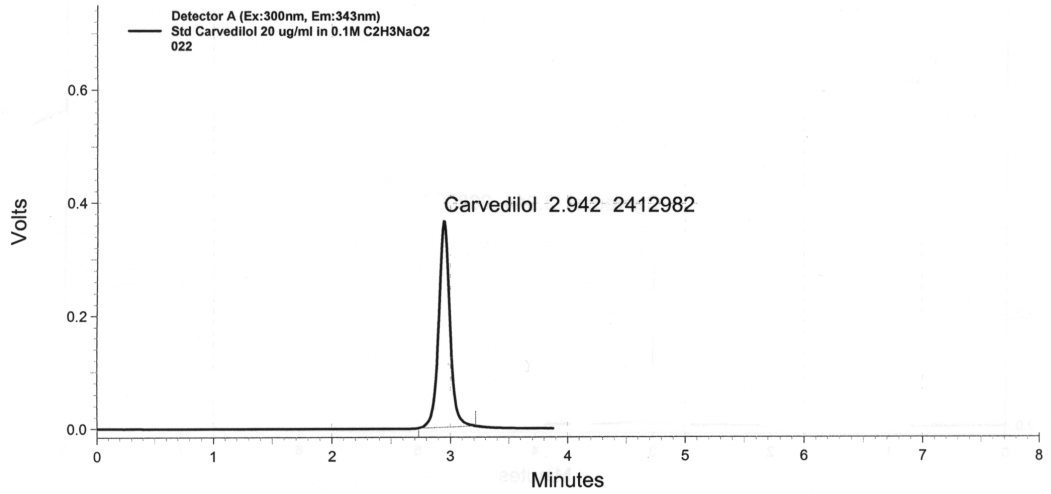
The method demonstrates excellent chromatographic specificity with no interference from tablet excipients, mobile phase or dissolution medium at the retention time of carvedilol. Representative chromatograms of carvedilol in the three dissolution media are shown in Figure 2. At equal concentration, a smaller peak area of carvedilol is observed in the phosphate buffer pH 6.8 as compared with those in the HCl solution pH 1.2 and the acetate buffer pH 4.5. Apparently, the pH of

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

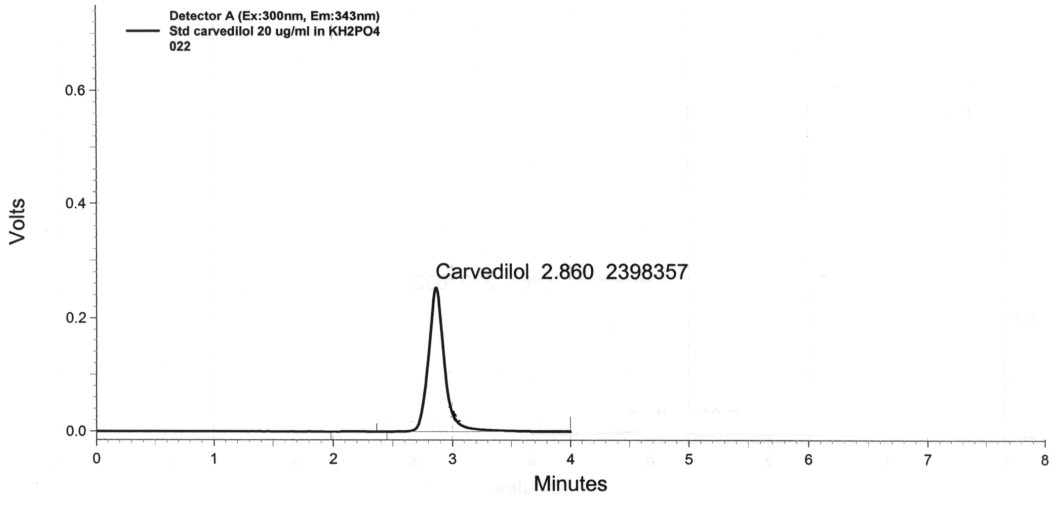
the medium has certain effect on the fluorescence intensity of carvedilol. The retention time of carvedilol is 2.8 min and each analysis can be done within 4 min under specified conditions.



(A)



(B)



(C)

Figure 2. Representative HPLC chromatograms of carvedilol (equal concentration) in: (A) 0.1 N HCl solution pH 1.2; (B) acetate buffer pH 4.5; (C) phosphate buffer pH 6.8

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

Method validation

According to Category III of the compendial assay procedures [26], a minimal assessment is required in terms of linearity, accuracy, precision and stability in three different media. The calibration curves for carvedilol in all dissolution media show good linearity with regression coefficient greater than 0.99 in the concentration range of 1-40 μg/ml. This means that there is a good correlation between peak area and drug concentration. The equation of linear regression and regression coefficient of the calibration curve for each medium is presented in Table 1. The results of the accuracy and precision determinations are shown in Table 2. The accuracy is between 95-105% and the intra-day precision expressed as % CV is less than 6.67% for the three dissolution media. The solutions remained stable in all dissolution media tested for the time period specified and no degradation products were observed in any chromatogram.

Table 1. Slope, intercept and regression coefficient of calibration curves obtained from three different dissolution media (*n*=6)

Medium	Slope	Intercept	Regression coefficient
0.1 N HCl solution pH 1.2	124367.00	-25017.39	0.9994
Acetate buffer pH 4.5	117218.42	4996.24	0.9989
Phosphate buffer pH 6.8	122211.38	-22638.78	0.9981

In vitro dissolution study

A dissolution test is normally employed for lot-to-lot quality control of pharmaceuticals in solid dosage form. Since carvedilol is not officially available in the pharmacopoeia, we have developed the dissolution testing condition for this drug. Drug release was carried out in accordance with the US pharmacopoeia general methods (Apparatus II) [26]. The temperature was kept constant at 37±0.5°C and the volume in each vessel kept at 900 ml. The dissolution parameters such as pH of medium, stirring speed and sampling time interval were optimised in terms of dissolution rate and precision. The most suitable dissolution method is shown in Table 3. In all three dosage strengths, Dilatrend® exhibits delayed dissolution in phosphate buffer pH 6.8 (Figure 3C) compared with dissolution in HCl solution pH 1.2 (Figure 3A) and that in acetate buffer pH 4.5 (Figure 3B). From the dissolution profile (Figure 3), 12.5 mg of Dilatrend® in HCl solution pH 1.2 and acetate buffer pH 4.5 show fastest dissolution. The different dissolution rates might stem from the difference in solubility in different pH media. For routine quality control test using a single-point specification [26], the acceptance criteria of tolerance should be at least 80% (Q) dissolution within 30 min in acetate buffer pH 4.5.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

Table 2. Accuracy and precision of the method for determining the concentration of carvedilol in three dissolution media (*n*=6)

Medium	Actual concentration (μg/ml)	Detected concentration (Mean ± SD; μg/ml)	Accuracy (%Accuracy)	Precision (%CV)
0.1 N HCl solution pH 1.2	1	1.04 ± 0.03	104.49	2.42
	5	4.82 ± 0.16	96.46	3.29
	10	9.89 ± 0.26	98.89	2.64
	20	20.19 ± 0.33	100.96	1.62
	30	30.39 ± 0.09	101.31	0.29
	40	39.66 ± 0.36	99.15	0.92
Acetate buffer pH 4.5	1	0.98 ± 0.07	98.42	6.67
	5	4.77 ± 0.17	95.40	3.58
	10	10.00 ± 0.53	99.95	5.32
	20	20.42 ± 0.31	102.10	1.53
	30	30.04 ± 1.96	100.14	0.59
	40	39.79 ± 0.71	99.47	1.78
Phosphate buffer pH 6.8	1	1.04 ± 0.05	103.91	5.02
	5	4.97 ± 0.22	99.37	4.42
	10	10.08 ± 0.14	100.85	1.39
	20	19.78 ± 0.66	98.91	3.31
	30	30.14 ± 0.59	100.46	1.95
	40	39.99 ± 1.31	99.97	3.28

Table 3. Optimal conditions for dissolution test of carvedilol tablets

Condition	Data / Unit
Apparatus	Apparatus II (Paddle) [26]
Dissolution medium	Acetate buffer pH 4.5
Volume of dissolution medium	900 ml
Temperature of dissolution medium	37±0.5°C
Revolution of stirrer	50 rpm
Number of tablet in vessel	1 tablet
Sampling time	5, 10, 15, 20, 30 and 45 min
Sampling volume	5.0 ml
Medium replacement	Yes

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

Application to drug quality controls: comparison of dissolution profiles

Approval of multi-source formulations using comparative in vitro dissolution studies should be based on generation of comparative dissolution profiles rather than a single-point dissolution test [27]. When comparing the test and reference products, dissolution profile can be compared using f_1 and f_2. Two dissolution product profiles are declared similar if f_1 is between 0-15 and f_2 is between 50-100 [28]. The results of dissolution efficiency in the three dissolution media, with Dilatrend® and Brand A as reference and test products respectively, are presented in Table 4. The dissolution profiles of 6.25-mg and 12.5-mg formulations in all dissolution media show f_1 and f_2 within acceptable ranges. The results of the two formulations therefore reflect sameness of the two curves and thus equivalence of the in vitro performance of the two products. On the contrary, the 25-mg formulation shows disagreement with the above guideline. This difference might be due to the excipient in the formula and the size of tablet, both of which can vary among brands. It is then necessary to carry out an in vivo study to guarantee the bioequivalence between the products.

Table 4. The difference and similarity factors between Dilatrend® (reference product) and Brand A (test product) in three different dissolution media

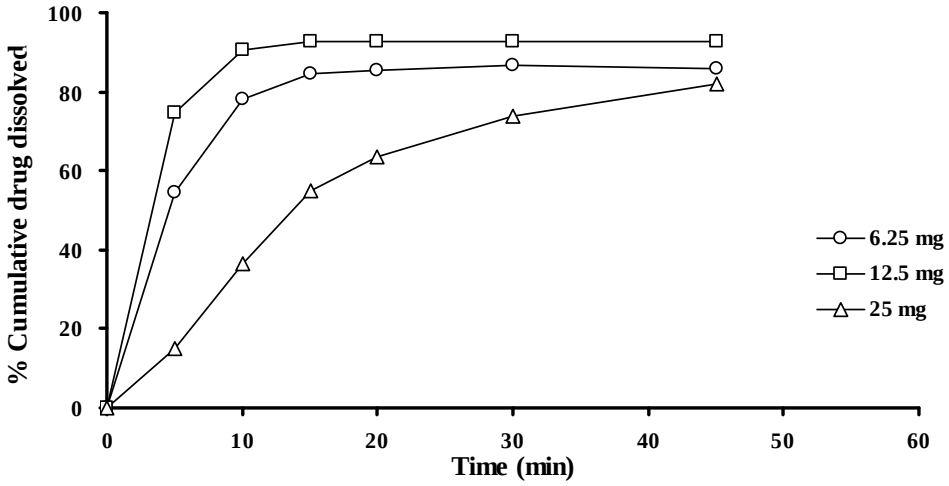
Dissolution medium	6.25 mg		12.5 mg		25 mg	
	f_1	f_2	f_1	f_2	f_1	f_2
pH 1.2 0.1 N HCl solution	4.20	69.41	7.71	56.36	41.30	28.52
pH 4.5 acetate buffer	5.75	61.59	0.89	88.24	10.36	38.46
pH 6.8 phosphate buffer	6.15	68.96	7.56	68.02	17.59	55.74

f_1 = difference factor (0-15), f_2 = similarity factor (50-100)

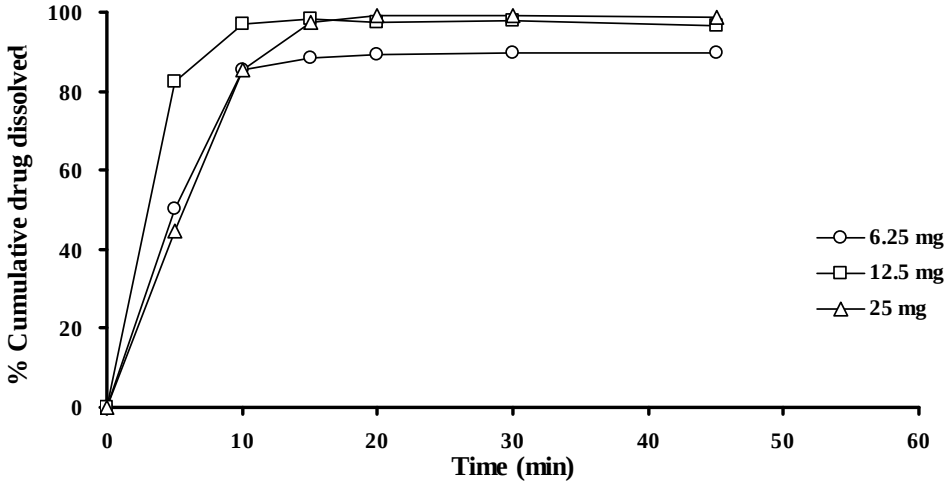
Assay in tablet formulation

The validated HPLC assay was applied to the quality control of two products. The % labelled amount and content uniformity are presented in Table 5. None of the formulation tested contains less than 95% of the labelled amount. Results of content uniformity experiment show that carvedilol content in each tablet from every product examined is in the range of 85.0-115.0 % and the RSD values are less than 6%. According to the acceptance limit of pharmacopoeia [24], this indicates a uniform distribution of drug in the tablets without any significant variation.

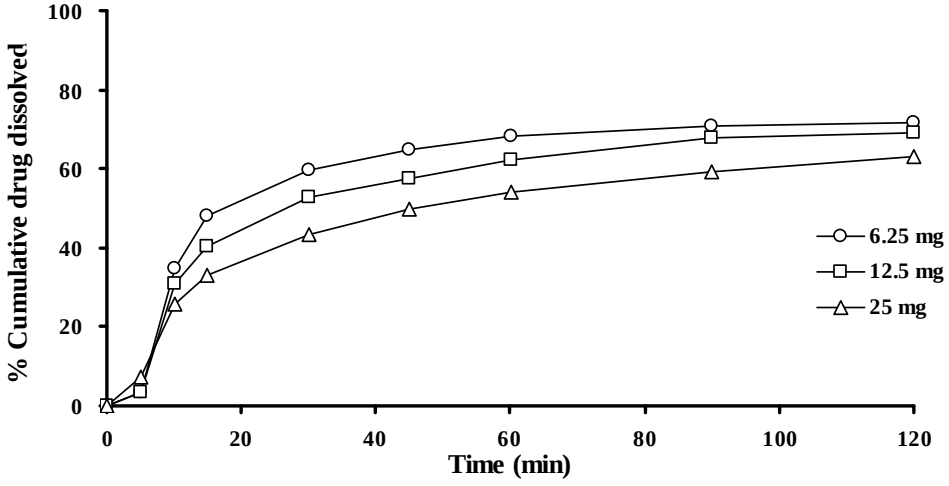
Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19



(A)



(B)



(C)

Figure 3. Dissolution profiles of carvedilol tablets (Dilatrend®) in: (A) 0.1 N HCl solution pH 1.2; (B) acetate buffer pH 4.5; (C) phosphate buffer pH 6.8

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

Table 5. Content of carvedilol in Dilatrend® and Brand A tablets

	Dose	Dilatrend®	Brand A
% Labelled	6.25 mg	95.82±0.45	97.74±1.27
amount	12.5 mg	98.07±0.93	95.71±0.38
(Mean ± SD)	25 mg	97.78±0.80	96.76±0.90
Content	6.25 mg	95.76-99.18 (1.38)	96.02-102.73 (1.89)
uniformity	12.5 mg	94.44-97.51 (1.02)	93.55-97.60 (1.67)
(Range in %)	25 mg	95.01-99.30 (1.36)	97.01-100.65 (1.45)

Note: Numbers in parentheses represent % RSD.

Conclusions

A method of quantitative determination of carvedilol using HPLC with fluorescence detector has been developed for the dissolution test and the quality control of the tablet formulation. The validation results have demonstrated that this method is accurate, precise, linear and specific. The dissolution test developed for carvedilol tablets is considered satisfactory. The optimal conditions for the dissolution profile determination are: 900 ml of acetate buffer (pH 4.5) medium at 37±0.5°C and paddle apparatus with 50-rpm stirring speed. The drug delivery requirement should be at least 80% dissolved in 30 min. The method can also be applied for quality control of drug content in pharmaceutical preparations.

Acknowledgements

The authors would like to thank Bioequivalence Test Centre, Faculty of Pharmaceutical Sciences, Naresuan University (Phitsanulok, Thailand) for the financial support, and all staff of the Bioequivalence Test Centre for their kind technical assistance.

References

1. R. R. Ruffolo Jr., M. Gellai, J. P. Hieble, R. N. Willette and A. J. Nichols, “The pharmacology of carvedilol” *Eur. J. Clin. Pharmacol.*, **1990**, *38*, S82-S88.
2. A. J. Nichols, M. Gellai and R. R. Ruffolo Jr., “Studies on the mechanism of arterial vasodilation produced by the novel antihypertensive agent, carvedilol”. *Fundam. Clin. Pharmacol.*, **1991**, *5*, 25-38.
3. C. de Mey, K. Breithaupt, J. Schloos, G. Neugebauer, D. Palm and G. G. Belz, “Dose-effect and pharmacokinetic-pharmacodynamic relationships of the beta 1-adrenergic receptor blocking properties of various doses of carvedilol in healthy humans”, *Clin. Pharmacol. Ther.*, **1994**, *55*, 329-337.
4. K. Nakamura, K. Kusano, Y. Nakamura, M. Kakishita, K. Ohta, S. Nagase, M. Yamamoto, K. Miyaji, H. Saito, H. Morita, T. Emori, H. Matsubara, S. Toyokuni and T. Ohe, “Carvedilol

decreases elevated oxidative stress in human failing myocardium", *Circulation*, **2002**, *105*, 2867-2871.

5. L. M. Kukin, J. Kalman, H. R. Charney, K. D. Levy, C. Buchholz-Varley, N. O. Ocampo and C. Eng, "Prospective, randomized comparison of effect of long-term treatment with metoprolol or carvedilol on symptoms, exercise, ejection fraction, and oxidative stress in heart failure", Circulation, **1999**, *99*, 2645-2651.
6. M. Packer, B. M. Fowler, B. E. Roecker, J. S. A. Coats, A. H. Katus, H. Krum, P. Mohacsi, L. J. Rouleau, M. Tendera, C. Staiger, L. T. Holcslaw, I. Amann-Zalan and L. D. DeMets, "Effect of carvedilol on the morbidity of patients with severe chronic heart failure: results of the carvedilol prospective randomized cumulative survival (COPERNICUS) study", *Circulation*, **2002**, *106*, 2194-2199.
7. R. R. Ruffolo, D. A. Boyle, D. P. Brooks, G. Z. Feuerstein, R. P. Venuti, M. A. Lukas and G. Poste, "Carvedilol: a novel cardiovascular drug with multiple actions", *Cardiovasc. Drug Rev.*, **1992**, *10*, 127-157.
8. P. A. Poole-Wilson, K. Swedberg, J. G. Cleland, A. Di Lenarda, P. Hanrath, M. Komajda, J. Lubsen, B. Lutiger, M. Metra W. J. Remme, C. Torp-Pedersen, A. Scherhag and A. Skene, "Rationale and design of the carvedilol or metoprolol European trial in patients with chronic heart failure: COMET", *Eur. J. Heart Fail.*, **2002**, *4*, 321-329.
9. N. Hokama, N. Hobara, H. Kameya, S. Ohshiro and M. Sakanashi, "Rapid and simple micro-determination of carvedilol in rat plasma by high-performance liquid chromatography" *J. Chromatogr. B*, **1999**, *732*, 233-238.
10. G. Lamprecht and K. Stoschitzky, "Determination of carvedilol in human plasma by high-performance liquid chromatography applying on-line deproteination and column switching", *Chromatographia*, **2004**, *59*, 551-554.
11. G. Lamprecht, L. Gruber, K. Stoschitzsky and W. Lindner, "Enantioselective analysis of (R)- and (S)-carvedilol in human plasma by high-performance liquid chromatography", *Chromatographia*, **2002**, *56*, S25-S29.
12. M. Saito, J. Kawana, T. Ohno, M. Kaneko, K. Mihara, K. Hanada, R. Sugita, N. Okada, S. Oosata, M. Nagayama, T. Sumiyoshi and H. Ogata, "Enantioselective and highly sensitive determination of carvedilol in human plasma and whole blood after administration of the racemate using normal-phase high-performance liquid chromatography", *J. Chromatogr. B*, **2006**, *843*, 73-77.
13. A. Zarghi, S. M. Foroutan, A. Shafaati and A. Khoddam, "Quantification of carvedilol in human plasma by liquid chromatography using fluorescence detection: application in pharmacokinetic studies", *J. Pharm. Biomed. Anal.*, **2007**, *44*, 250-253.
14. L. Clohs and K. M. McErlane, "Comparison between capillary electrophoresis and high-performance liquid chromatography for the stereoselective analysis of carvedilol in serum", *J. Pharm. Biomed. Anal.*, **2003**, *31*, 407-412.
15. E. Yang, S. Wang, J. Kratz and M. J. Cyronak, "Stereoselective analysis of carvedilol in human plasma using HPLC/MS/MS after chiral derivatization", *J. Pharm. Biomed. Anal.*, **2004**, *36*, 609-615.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 8-19

16. N. C. do Carmo Borges, G. D. Mendes, D. de Oliveira Silva, V. Marcondes Rezende, R. E. Barrientos-Astigarraga and G. de Nucci, "Quantification of carvedilol in human plasma by high-performance liquid chromatography coupled to electrospray tandem mass spectrometry: application to bioequivalence study", *J. Chromatogr. B*, **2005**, *822*, 253-262.
17. M. Machida, M. Watanabe, S. Takechi, S. Kakinoki and A. Nomura. "Measurement of carvedilol in plasma by high-performance liquid chromatography with electrochemical detection", *J. Chromatogr. B*, **2003**, *798*, 187-191.
18. J. Oravcova, D. Sojkova and W. Lindner, "Comparison of the Hummel-Dreyer method in high-performance liquid chromatography and capillary electrophoresis conditions for study of the interaction of (RS)-, (R)- and (S)-carvedilol with isolated plasma proteins", *J. Chromatogr. B*, **1996**, *682*, 349-357.
19. L. J. Patel, B. N. Suhagia, P. B. Shah and R. R Shah, "RP-HPLC and HPTLC methods for the estimation of carvedilol in bulk drug and pharmaceutical formulations", *Indian J. Pharm. Sci.*, **2006**, *68*, 790-793.
20. J. Stojanovic, V. Marinkovic, S. Vladimirov, D. Velickovic and P. Sibinovic, "Determination of carvedilol and its impurities in pharmaceuticals", *Chromatographia*, **2005**, *62*, 539-542.
21. J. Stojanovic, S. Vladimirov, V. Marinkovic, D. Velickovic and P. Sibinovic, "Monitoring of the photochemical stability of carvedilol and its degradation products by the RP-HPLC method", *J. Serb. Chem. Soc.*, **2007**, *72*, 37-44.
22. A. Radi and T. Elmogy, "Differential pulse voltammetric determination of carvedilol in tablets dosage form using glassy carbon electrode", *Il Farmaco*, **2005**, *60*, 43-46.
23. T. O'Hara, A. Dunne, J. Butler and J. Devane, "A review of methods used to compare dissolution profile data", *Pharm. Sci. Technol. Today*, **1998**, *5*, 214-223.
24. V. P. Shah, "Dissolution: a quality control test vs. bioequivalence test", *Dissol. Tech.*, **2001**, *8*, 1-2.
25. L. X. Xu, N. Hui, L. Y. Ma and H. Y. Wang, "Study on fluorescence property of carvedilol and determination of carvedilol by fluorimetry", *Spectrochim. Acta Part A*, **2005**, *61*, 855-859.
26. "The United States Pharmacopoeia", 28th Edn., United States Pharmacopoeial Convention, Rockville, MD, **2005**.
27. "Multisource (Generic) Pharmaceutical Products: Guidelines on Registration Requirements to Establish Interchangability", World Health Organisation., Geneva, **2005**.
28. "Guidance for Industry: Dissolution Testing of Intermediate Release Solid Oral Dosage Forms", U.S. Department of Health and Human Services, Food and Drug Administration, Rockville, MD, **1997**.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

Maejo International Journal of Science and Technology

ISSN 1905-7873

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

Full Paper

The value of h/e^2 from quantum Hall effect

Keshav N. Shrivastava

Department of Physics, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia

E-mail: keshav1001@yahoo.com

Received: 25 September 2009 / Accepted: 1 February 2010 / Published: 4 February 2010

Abstract: The quantum Hall effect and the emergence of the value of h/e^2 is found to be understood within five steps. Here h is the Planck's constant and e is the charge of the electron. The Hall resistivity is found to become a function of spin. For positive spin, one value is found but for negative sign in the spin, another value occurs. In this way, there is never only one value of the resistivity but doubling of values. The value of h/e^2 is a special case of the more general dependence of resistivity on the spin. We investigate the effect of Landau levels. For extreme quantum limit, n=0, the effective charge of the electron becomes (1/2)ge. The fractional charge arises for a finite value of the angular momentum. There is a formation of spin clusters. As the field increases, there is a phase transition from spin ½ to spin 3/2 so that g value becomes 4 and various values of n in Landau levels, g(n+1/2), form plateaus in the Hall resistivity. For finite values of the orbital angular momenta, many fractional charges emerge. The fractional as well as the integral values of the charge are in full agreement with the experimental data. The generalised constant is h/[(1/2)ge]e which under special conditions becomes h/e^2, the ratio of Planck's constant to the square of the electron charge. The flux is usually quantised in units of φ_o =hc/e. When the angular momentum is properly taken into account, hc/e is replaced by hc/(1/2)ge. Thus, we predict a new superfluid which has (1/2)ge in place of the charge, e.

Keywords: Hall effect, constant h/e^2, spin, charge

Introduction

Recently, we have shown that fractional charges occur in the quantum Hall effect and it can be explained by a few steps [1]. The quantum Hall effect is an experimental observation of plateaus in the Hall current which are explained by means of a wave function so that there is a concept of quasiparticles. These quasiparticles may be bosons, fermions or anyons. That is for the theorists to

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

resolve with or without the use of experimental data. In the constrictions of wires, electron clusters are formed which have spin. It is possible to suggest that repulsive Coulomb interactions give rise to fractional charges as compared with the charge-density waves. Laughlin [2] has suggested a possible wave function which might explain a few fractional charges. Wilczek [3] has the ideas of anyons which obey fractional statistics. It was also suggested that flux quanta may be attached to the electrons, which might explain the symmetries found in the plateaus. Anderson [4] has suggested an alternative to the Laughlin's wave function. Alex Mueller [5] realised the importance of doping in understanding high temperature superconductors, which in turn are important for the understanding of pairing of electrons. We explained the quantum Hall effect with more than 101 plateaus by using the spin [6]. There are six steps in this mesa:

Step 1: Magnetic field. The energy of an electron in a magnetic field is given by,

$$g\mu_B H = \eta\omega_c \quad (1)$$

where g is the Lande factor, μ_B is the Bohr magneton and c is the velocity of light,

$$\mu_B = \frac{e\eta}{2mc} \quad (2)$$

Substituting (2) in (1),

$$\frac{1}{2} g \frac{eH}{mc} = \omega_c \quad (3)$$

Thus, the effective charge of an electron can be written as,

$$e^* = \frac{1}{2} ge . \quad (4)$$

This is an important step because it gives the effective charge of a quasiparticle. When g=2, which is the spin-only value, the effective charge becomes e*= e.

Step 2: Landau levels. The energy levels of an electron in two dimensions look like that of a harmonic oscillator,

$$(n + \frac{1}{2})\eta\omega_c = (n + \frac{1}{2}) g\mu_B H \ . \quad (5)$$

For n = 0,

$$E_o = \frac{1}{2} g\mu_B H = \frac{1}{2}\eta\omega_c \quad (6)$$

We can remember, just in case we need this energy term with a factor of ½.

Step 3: Hall effect. The classical Hall resistivity is linearly proportional to the magnetic induction. It is used to determine the concentration, *n*, the number of electrons per unit area.

$$\rho = \frac{B}{nec} \quad (7)$$

The flux within the area, A, is quantised in the units of $\phi_o = hc/e$,

$$B.A = n'\phi_o \quad (8)$$

Substituting (8) in (7),

$$\rho = \frac{n'\phi_o}{nAec} = \frac{h}{ie^2} \text{ (i = integer)} \quad (9)$$

Substituting (4) in (9),

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

$$\rho = \frac{h}{(\frac{1}{2}ge)e} \tag{10}$$

Note that one e comes from the Hall effect and the other comes from flux quantisation.

Step 4: Fractional values. We define the g values linear in the angular momenta and allow both signs of spin in the total angular momentum, j = $l \pm s$. Then,

$$g = \frac{2j+1}{2l+1} = \frac{2(l \pm s)+1}{2l+1} \tag{11}$$

$$\frac{1}{2}g = \frac{l+\frac{1}{2} \pm s}{2l+1}. \tag{12}$$

For s=1/2, $\frac{1}{2}g_{+} = \frac{l+1}{2l+1}, \frac{1}{2}g_{-} = \frac{l}{2l+1}$, which we tabulate (Table 1):

Table 1. The fractional charge for two signs of s

l	$\frac{1}{2}g_{+}$	$\frac{1}{2}g_{-}$
0	1	0
1	2/3	1/3
2	3/5	2/5

The (1/2)g=1 is the correct spin-only value and all of the fractions are correct including the 1/3 charge. These fractional values agree with the experimental data.

Step 5: Spin 3/2. For l = 0, s = 3/2, $\frac{1}{2}g_{+} = \frac{l+\frac{1}{2}+\frac{3}{2}}{2l+1} = 2$, $g_{+} = 4$. The values of $n+\frac{1}{2}$ are,

½, 3/2, 5/2, 7/2, 9/2, …

and the values of $g_{+}(n+\frac{1}{2})$ are,

2, 6, 10, 14, 18, …

This series has been observed in the experimental data. Note that if one-particle states are at g/2, then two-particle states occur at g.

Step 6: Comparison.

(i) Equal spin pairing: Balian and Werthamer [7]. Of course these days a better calculation with proper treatment of singlets and triplets is available.

(ii) Zero momentum, spin singlet pairs, k↑ and -k↓: B.C.S. pairing in the conduction band [8].

(iii) Proton spin triplets, ↑↑: Leggett [9].

(iv) Our result: Shrivastava [10],

Spin up, ↑, charge e*=2/3;

Spin down, ↓, charge e*=1/3.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

$$\rho = \frac{h}{(\frac{1}{2} g_{\pm} e)e} . \qquad (13)$$

The theory of fractional charges compares in quality with those of Balian and Werthamer [7], Bardeen et al. [8] and Leggett [9], and explains 101 fractional charges. Hence the flux quantises as B×area=n'hc/[(1/2)$g_{\pm}$e].

The Value of h/e^2

The resistivity at the plateaus is quantised in the units of h/e^2. Usually, the electron is associated with the electromagnetic field, the same way as the charge density is, in the Maxwell equations. The electric and magnetic field vectors are linked to the charge density. However, the charge is defined in such a way that the effect of self electromagnetic fields is already included in the value of the charge,

$$e = 1.602\ 176\ 487(40)\ \times 10^{-19}\ \text{Coulomb} \qquad (14)$$

The Planck's constant is associated with the frequency or the wave length of a particle,

$$h = 6.626\ 068\ \ 96(33)\ \times 10^{-34}\ \text{Js} . \qquad (15)$$

It is a matter of pencil calculation to show that,

$$h/e^2 = 258\ 12.807\ 5651\ \ \text{Ohm} . \qquad (16)$$

This constant was measured by von Klitzing et al [11]. In their paper, the value given is 25813 Ω. The calculation of h/e^2 does not require that there should be two dimensionality or there should be Landau levels. However, the experimental value requires the Hall geometry. The value of h/e^2 does not require any electrodynamic correction. The fine structure constant is defined in such a way that,

$$h/e^2 = \mu_o\, c/(2\alpha) \qquad (17)$$

where $\mu_o = 4\pi \times 10^{-7}$ H/m and c is the velocity of light. The above expression is actually an identity because,

$$\alpha = \frac{e^2}{4\pi\varepsilon_o \eta c} \qquad (18)$$

where $\varepsilon_o = 1/(\mu_o c^2)$. At the present time, the value of the inverse fine structure constant [12] is $1/\alpha$ = 137.035 999 084(51), which is another way of writing the value of h/e^2. These are one and the same and not two different quantities. How the accuracy has become so high is another question but in 1965, the value was 137.0388(6). The gyromagnetic ratio of the electron is given by Mohr et al. [12],

$$g/2 = 1.001\ 159\ 652\ 180\ 73(28) \qquad (19)$$

This value is related to the fine structure constant,

$$g/2 = 1 + C_2(\alpha/\pi) . \qquad (20)$$

In this way, g is related to α and α determines h/e^2. However, g is subject to the electrodynamical corrections whereas h/e^2 is not. The electron is associated with the electromagnetic field because of the charge. The electromagnetic field is quantised in terms of photons. Therefore, there are many Feynmann

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

diagrams which describe the electron-photon interaction so that many more terms arise in (20) which have to be carefully added. The Lande's formula gives,

$$g/2 = 1 \tag{21}$$

for $l = 0$ and the electrodynamic correction is,

$$\frac{g}{2}|_{electrodynamic} = 0.00115965218073(28) \tag{22}$$

so that,

$$\frac{g}{2} = \frac{g}{2}|_{Lande} + \frac{g}{2}|_{electrodynamic} \ . \tag{23}$$

The g value can be separated into electrodynamic part and Lande's part but in the case of the value of the charge such a separation is not available. The Lande's formula [13] does not contain the electrodynamics but it contains the angular momenta, L, S and J. If there is any correction to the value of the charge due to the electrodynamics, it is already included in the tabulated value of e. There is a problem of gauge invariance as to which h/e is fixed and only one e in h/e^2 is subject to measurement. If both values of e are equal we get the h/e^2. In our theory [14-19] the resistivity is,

$$\rho = \frac{h}{\frac{1}{2}ge^2} \tag{24}$$

where (1/2)g does not include the electrodynamic correction. In fact, such electrodynamic corrections are already included in h/e^2. We use the definition $g=(2j+1)/(2l+1)$ so that for $j = l \pm s$, there are two values of g which we call $g_{\pm}$,

$$g_{\pm} = \frac{2(l \pm s)+1}{2l+1}. \tag{25}$$

Note that this value of $g_{\pm}$ does not have the electrodynamic correction. The expression (17) suggests that h/e^2 is equivalent to α and (20) relates α to g value. When $l = 0$,

$$g_{\pm} = 2(\pm s)+1. \tag{26}$$

For s=1/2 for + sign, $g_+=2$ so that $(1/2)g_+=1$ and the result (24) gives h/e^2. For s=1/2 and negative sign, $g_-=0$ and we get $\rho \to \infty$, or the conductivity, $\sigma \to 0$. We call these values von Klitzing constants, which now have two values,

$$R_K = h/e^2 \tag{27}$$

and

$$R_K = \infty \ . \tag{28}$$

For l =1, s=1/2 for positive sign, (25) gives,

$$g_+ = \frac{2(1+\frac{1}{2})+1}{3} = \frac{4}{3} \tag{29}$$

or $(1/2)g_+=2/3$, which makes von Klitzing value,

$$R_K = \frac{h}{\frac{2}{3}e^2}. \tag{30}$$

For l =1, s=1/2 and negative sign in (25),

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

$$g_{-}=\frac{2(1-\frac{1}{2})+1}{3}=\frac{2}{3} \quad (31)$$

or $(1/2)g_{-}=1/3$ so that the von Klitzing resistivity becomes,

$$R_K=\frac{h}{\frac{1}{3}e^2} \quad (32)$$

In this way many values of the Klitzing constant can be predicted. The fractional values calculated here agree with the measured values of Tsui et al [20].

The Harmonic Oscillator

The eigen values of the harmonic oscillator are given by,

$$E_n=(n+\frac{1}{2})\eta\omega \quad (33)$$

where,

$$\eta\omega = g\frac{e\eta}{2mc}B. \quad (34)$$

For n = 0, $E_o=(1/2)\eta\omega$ so that the frequency becomes,

$$E_o=\frac{1}{2}\eta\omega=\frac{1}{2}g\frac{e\eta}{mc}B \quad (35)$$

This means that we can replace e by (1/2)ge or e*=(1/2)ge. The von Klitzing resistivity now becomes,

$$R_K=\frac{h}{\frac{1}{2}g_{\pm}e^2} \quad (36)$$

where we can generate a lot of values by changing *l* and s but it is clear that there are pairwise values due to ± and not single value, i.e. there is a doubling of values. From (25) we can calculate the values of $g_{\pm}$ for various values of *l* and s, which gives values of the resistivity. We use the harmonic oscillator-type expression, so that (24) becomes,

$$R_K=\frac{h}{(n+\frac{1}{2})g_{\pm}e^2}\ . \quad (37)$$

For,

n = 0, 1, 2, 3, 4, 5, 6,

the values of n+(1/2) are,

0, 3/2, 5/2, 7/2, 9/2, …

For S=3/2, *l* =0 we have for the positive sign,

$g_{\pm}=2(\pm 3/2)+1=4$(for + sign). The values of $g_{+}(n+1/2)$ are now,

0, 6, 10, 14, 18, … (38)

This series is actually observed in the experimental data. As we can see, there is no need of random topological numbers, nor of Chern numbers or Hofstadter butterfly [21]. The growth of the series such as that in (25) is not a fractal growth and it does not have a constant chemical length.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

The g Value and α

The electron produces its own electromagnetic field which changes the g value. This is a small field but quite noticeable in ordinary electron-spin resonance experiments. The magnetic moment of the electron is,

$$\mu = -\frac{1}{2} g\mu_B \frac{S}{\eta/2} \qquad (39)$$

where S is the spin. Usually S=1/2, but in solid state electron clusters are formed so that it is not limited to ½ and it may be 1, 3/2, 5/2, etc. The accurate value of g/2 is needed to obtain the magnetic moment of the electron. Therefore, it is important to calculate the energy contributions of the electron-photon interaction which can be used to redefine the g value. Thus, an expansion has been considered,

$$\frac{1}{2} g = 1 + c_2(\alpha/\pi) + c_4(\alpha/\pi)^2 + c_6(\alpha/\pi)^3 + c_8(\alpha/\pi)^4 + c_{10}(\alpha/\pi)^5 + \ldots \qquad (40)$$

in which all of the coefficients have been carefully calculated to find,

$$\alpha^{-1} = 137.035\ 999\ 084(33) \qquad (41)$$

These calculations are limited to l =0, s=1/2 only. Therefore, two values of g are not obtained. Even then there are two values due to the ± in (25). One of these values is zero and the other is 2 besides the electrodynamic correction which is known for l =0. Let us take only 2 terms and substitute 0 and 2 for the g value. Then we obtain two equations,

$$\frac{1}{2}(2 + g_{ed}) = 1 + c_2^{(+)}(\alpha^{(+)}/\pi) \qquad (42a)$$

$$\frac{1}{2}(0 + g_{ed}) = 1 + c_2^{(-)}(\alpha^{(-)}/\pi) \qquad (42b)$$

leaving out small terms. The solution of the second of these gives negative value for $c_2^{(-)}\alpha^{(-)}$, which means that $c_2^{(+)}$ is not equal to $c_2^{(-)}$. Therefore, the values of the coefficients depend on the g values. The sign of the spin is contained in the g value so that both the positive and negative spin values are important.

Two Constants h/e^2 and $h/(g_\pm/2)e^2$

The resistivity at n = 0 in (24) for positive sign of the spin is,

$$\rho = \frac{h}{\frac{1}{2} g_\pm e^2} \qquad (43)$$

where $g_\pm$ must be taken from (25) and is free from the electrodynamic effects. We list some of the values which give the quantisation of the resistivity:

h= 6.626068 960 (330)× 10^{-34} Js,

e=1.602 176 487(40) × 10^{-19} Coulomb,

h/e^2= 25812.807 5651 Ohm {pencil calculation},

(1/2)g=1.001 159 652 180 73(28) [22],

h/e^2= 25812.807557(18) Ohm [12].

By taking only two terms from the right hand side of (40), we find that the charge can be completely eliminated,

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

$$\rho = \frac{h}{e^2} = \frac{hc_2'}{\pi(4\pi\varepsilon_o \eta c)(\frac{g}{2} - 1)} \qquad (44)$$

but the two values of the resistivity are exactly equal. The error in the experimental value of 25812.8 Ω is perhaps not more than ±0.20 Ω. The expression (25) gives the doubling of values due to ± signs and gives the correct fractional values of the charges which agree with the measured values.

Spin and Resistivity

There is a special case when (1/2)g=1,

$$\frac{1}{2} g_{\pm} = \frac{l + \frac{1}{2} \pm s}{2l + 1} \qquad (45)$$

which occurs for l =0, s=+1/2. For this case the resistivity (24) is the same as von Klitzing's value. In cases of finite l and s, the physics of the problem is different from that of von Klitzing et al. [11], so that von Klitzing's constant becomes a special case of "spin-dependent" phenomenon [6]. The values of $\rho_{K(+)}$ and $\rho_{K(-)}$ from the expression,

$$\rho_{K(\pm)} = \frac{h}{\frac{1}{2} g_{\pm} e^2} \qquad (46)$$

are given in Table 2 along with the values of $g_{\pm}$. A plot of $\rho_{K(\pm)}$ as a function of l is given in Figure 1.

Table 2. The values of various constants by varying the value of l

l	$\rho_{K(+)}$	$\rho_{K(-)}$	$(1/2)g_+$	$(1/2)g_-$
0	1	∞	1	0
1	3/2	3/1	2/3	1/3
2	5/3	5/2	3/5	2/5
3	7/4	7/3	4/7	3/7
4	9/5	9/4	5/9	4/9
5	11/6	11/5	6/11	5/11
6	13/7	13/6	7/13	6/13
7	15/8	15/7	8/15	7/15
8	17/9	17/8	9/17	8/17
9	19/10	19/9	10/19	9/19
∞	2	2	1/2	1/2

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

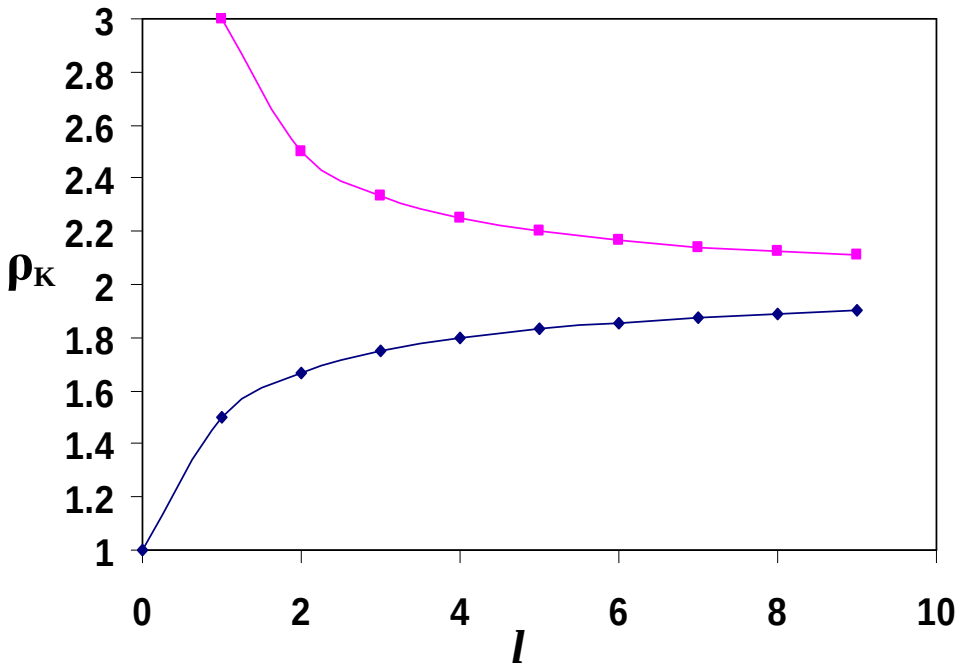


Figure 1. The variation of resistivity as a function of *l*: the upper curve is (-) spin and lower curve is (+) spin.

At l =0, ρ_K=1, we obtain the von Klitzing's constant. Otherwise, there are many values and the von Klitzing constant is a special case of more general constants:

$\rho_{K(+)}$ (l =0, s=1/2) = Von Klitzing's constant (47)

$\rho_{K(\pm)}$ ($l \neq 0$, s=±1/2, ± 1, ± 3/2, ±,2, ...) = General constants. (48)

A plot of $\rho_{K(\pm)}$ as a function of (1/2)g from Table 2 is given in Figure 2. When (1/2)g=1, we obtain the von Klitzing's constant, otherwise the more general constants exist.

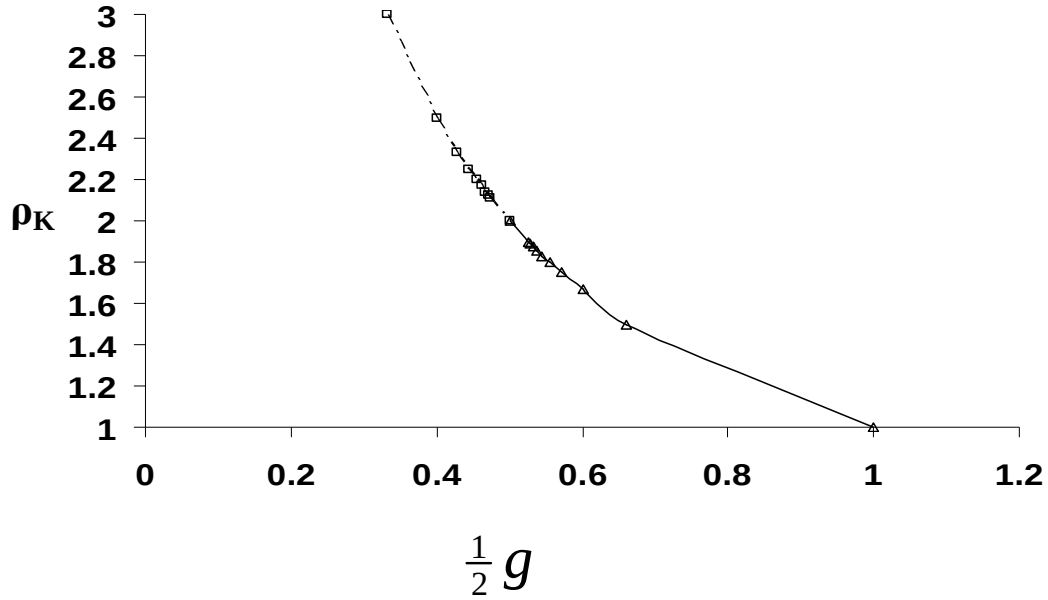


Figure 2. Plot of resistivity as a function of (1/2)g. The continuous line on the right hand side of 0.5 has (+) spin and the broken line on the left hand side of 0.5 has (-) spin.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

Turning Points

As the gate voltage is increased, the resistivity starts turning towards the plateau. This phenomenon can occur when spins start turning. When the resistivity is at the Hall effect value away from the plateau region, the electron spin starts turning until the area is so adjusted as to satisfy the flux quantisation, which means that the vortex area becomes an integral multiple of the flux quanta divided by the field area, i.e. $n\varphi_o/B$. The area in the Hall region is infinite. As the spins turn, the area starts reducing from the infinite value to the quantised value.

The change in resistivity from the turning point to the plateau is about 72.7 Ohm, compared with $h/4e^2$=6453.201 Ohm. A plot of the resistivity as a function of gate voltage is given in Figure 3. At the turning point the resistivity is,

$$\rho_{turn} = 6471.21\ \Omega \qquad (49)$$

compared with the pencil calculation of $h/4e^2$= 6453.03 Ω . These two values are off by 18.18 Ω. In order to compare the turning point value with the plateau value, we define,

$$\delta\rho = \rho_{turn} - \rho_{plateau.} \qquad (50)$$

Then the value of $\rho_{plateau}$ is 25812.8075 Ω whereas $\rho_{turn}(i=1)$ is 25884.84 Ω so that,

$$\delta\rho = 72.0\ \Omega. \qquad (51)$$

This, in principle, makes the value of h/ie^2 (i=integer) quite uncertain. The experimental uncertainty in 25812.8 is only 0.2 Ω but the in-principle uncertainty is 2.8×10^{-3}, which is a few parts per thousand. The plateau measurement is obviously much more accurate than the difference between the plateau and the turning point values. In such a case the "in principle" value will play a dominant role. The plateau

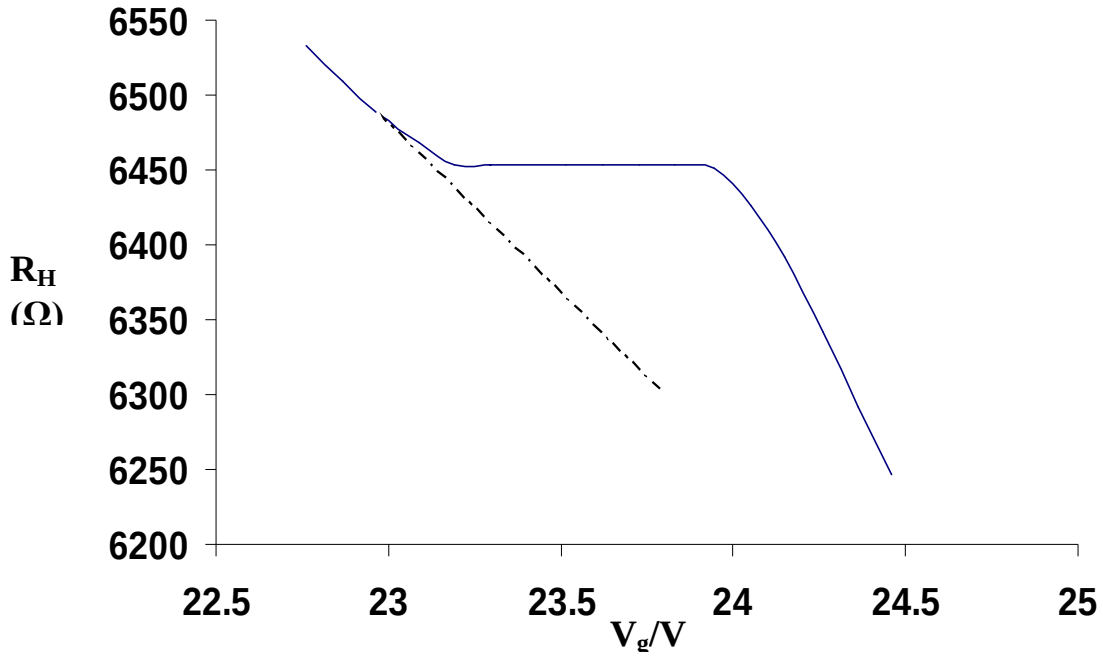


Figure 3. Plot of resistivity as a function of gate voltage. As the gate voltage is increased, the data shows "turning point" before reaching the plateau.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 20-32

value can be measured up to 8 digits, which means that the accuracy is 1 part in 10^8. If that is the case, the plateau is sharply peaked but the distribution may be extended up to the turning point. It is said that the centre of a line can be located to a large accuracy. That does not mean that there is no line width. The line is an envelope of a large number of events so that there is a finite width. The accuracy of measurement is thus not the accuracy of locating the plateau but the location of the turning point. In Laughlin's work [2] an effort is made to obtain the fractions by correlations. In the present work the fractions arise from the spin. In Laughlin's theory, incompressibility is needed, otherwise the area, A, in the flux quantisation will make the charge flow. The charge can be fractionalised only when A=constant. The flux quantisation condition, B.A=n'hc/e, demands that if the charge has to change, the area A must be a constant. We can check this constancy of the area by applying pressure at the point of the plateaus. In fact, such an experiment has been done, for example by Leadley et al. [23], who recorded the resistivity of GaAs/$Ga_{0.7}Al_{0.3}As$ as a function of pressure and applied magnetic field. It was reported that the dip in the xx-component of the resistivity varies as a function of the applied field. For a pressure of 18.7 kbar the dip at the fractional charge of 1/3 (ν =3) is almost completely wiped off but appears again when the pressure is increased to 20 kbar. In any case, there is some dependence on the pressure so that the area cannot be held constant. In Figure 4 we show the resistivity as a function of pressure, which shows that the system is not incompressible so that the incompressible model of the fractional charge is not necessarily applicable to the present experimental situation. Hence the fractionalisation of the charge is due to spin and the flux is quantised as, B.A= n'φ_o, where φ_o is hc/e which is changed to hc/(1/2)ge, where g=(2j+1)/(2 *l* + 1). Thus, there is a quantum superfluid in which

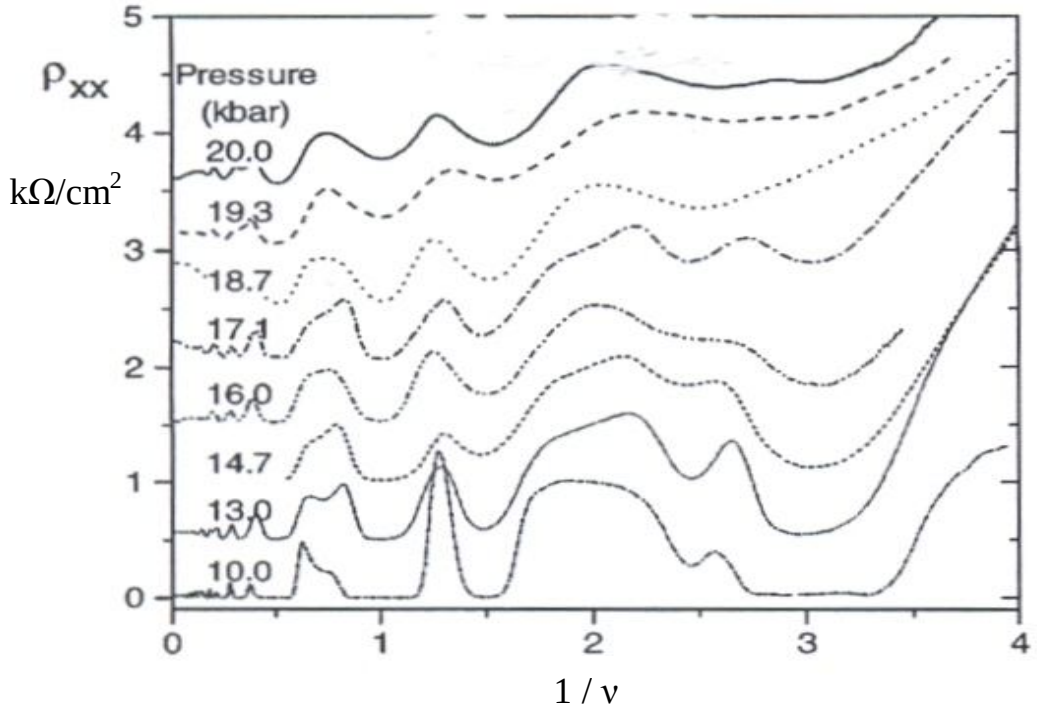


Figure 4. The resistivity of GaAs/AlGaAs as a function of applied pressure. As the pressure is changed from 10.0 kbar to 20.0 kbar, there is considerable change in the behaviour near ν=3, showing that the incompressibility condition is not found [24].

finite angular momenta occur. We have also performed the calculations with Pfaffian determinant from which we find that the non-Abelian determinant is unlikely to correspond to the real material. It is found that the time runs faster in the non-Abelian than in the Abelian wave function [25]. For small matrices such as 2×2, the positive spin works just as good as the negative spin. However, it is interesting to learn that the negative spin plays an important role and it is quite feasible to use it to define the dimensions [26].

Conclusions

The fractional charges occurring in the quantum Hall effect can be explained by the spin properties. It also means that, without spin and by orbital correlations alone, the fractional charge does not occur. If spin is ignored, the correct charge is proportional to the orbital angular momentum as, $E_{effective(spinless)} \propto \frac{1}{2L+1}$. The von Klitzing's constant, h/e^2, is related to the g value for $l = 0$ and s =1/2. There is a small correction to this g value due to electrodynamics. Such a correction is already included in the value of the charge of the electron. Another modification to the von Klitzing's constant arises from $g_{\pm}$ which is due to (2j+1)/(2l+1). This effect produces fractional values which agree with the experimental observations. In addition to the von Klitzing's constant, there are more general constants which depend on spin. It is impossible for the von Klitzing constant to occur alone. The constants occur in pairs. For finite values of *l* and s including the values other than ½, a whole series of constants arise. There is a "turning point" phenomenon. Although the value of h/e^2 is very sharply peaked, it cannot be ignored that there is a distribution. We are able to construct the basic theory which correctly gives the fractional charges in agreement with the data. We find that it is necessary to introduce negative spin and that the charge gets coupled to spin. There is a new condition on the flux quantisation which depends on spin.

References

1. K. N. Shrivastava, "Elementary theory of quantum Hall effect", *Maejo Int. J. Sci. Technol.*, **2008**, *2*, 285-297.
2. R. B. Laughlin, "Anomalous quantum Hall effect: An incompressible liquid with fractionally charged excitations", *Phys. Rev. Lett.*, **1983**, *50*, 1395-1398.
3. F. Wilczek, "Magnetic flux, angular momentum and statictics", *Phys. Rev. Lett.*, **1982**, *48*, 1144-1146.
4. P. W. Anderson, "Remarks on the Laughlin theory of fractionally quantized Hall effect", *Phys. Rev. B*, **1983**, *28*, 2264-2265.
5. K. A. Mueller, "On the superconductivity in hole doped cuprates", *J. Phys. Condens. Matter*, **2007**, *19*, 251002-1-13.
6. K. N. Shrivastava, "Interpretation of fractions in quantum Hall effect: Wei Pan's data", *Amer. Inst. Phys. Conf. Proc. (USA)*, **2009**, *1150*, 59-67.
7. R. Balian and N. R. Werthamer, "Superconductivity with pairs in a relative p wave", *Phys. Rev.*, **1963**, *131*, 1553-1564.

8. J. Bardeen, L. Cooper and J. R. Schrieffer, "Theory of superconductivity", *Phys. Rev.*, **1957**, *108*, 1175-1204.
9. A. J. Leggett, "Interpretation of recent results on He3 below 3mK: A new liquid phase", *Phys. Rev. Lett.*, **1972**, *29*, 1227-1230.
10. K. N. Shrivastava, "Rational numbers of the fractionally quantized new states and their observation in fractionally quantized Hall effect", *Phys. Lett. A*, **1986,** *113*, 435-436; *115*, 459(E).
11. K. von Klitzing , G. Dorda and M. Pepper , "New method for high accuracy determination of the fine structure constant based on quantized Hall resistance", *Phys. Rev. Lett.*, **1980,** *45*, 494-497.
12. P. J. Mohr, B. N. Taylor and D. B. Newell, "The fundamental physical constants", *Phys. Today*, **2007**, *60*, 52-54.
13. A. Lande, "Quantum Mechanics", Cambridge University Press, Cambridge, **1951**.
14. K. N. Shrivastava, "Additional Dirac matrix in quantum Hall effect", *Amer. Inst. Phys. Conf. Proc. (USA)*, **2008**, *1017*, 47-56.
15. K. N. Shrivastava, "Energy band crossing points in multilayers of graphene", *Amer. Inst. Phys. Conf. Proc. (USA)*, **2008**, *1017*, 326-330.
16. K. N. Shrivastava, "The theory of the quantum Hall effect", *Amer. Inst. Phys. Conf. Proc (USA)*, **2008**, *1017*, 422-428.
17. K. N. Shrivastava, "Infrared of thin film graphene in a magnetic field and the Hall effect", *Proc. SPIE (USA)*, **2008**, *7155*, 71552F-1-8.
18. A. N. Rosli and K. N. Shrivastava, "Von Klitzing's constant as a special case of generalized constants", *Amer. Inst. Phys. Conf. Proc. (USA)*, **2009**, *1136*, 469-473.
19. K. N. Shrivastava, "The fractional quantum Hall effect: The cases of 5/2 and 12/5", *Amer. Inst. Phys. Conf. Proc. (USA)*, **2009**, *1169*, 48-54.
20. D. C. Tsui, H. L. Stormer and A. C. Gossard, "Two dimensional magnetotransport in the extreme quantum limit", *Phys. Rev. Lett.*, **1982**, *48*, 1559-1562.
21. D. R. Hofstadter, "Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields", *Phys. Rev. B*, **1976**, *14*, 2239- 2249.
22. D. Hanneke, S. Fogwell and G. Gabrielse , "New measurement of the electron magnetic moment and the fine structure constant", *Phys. Rev. Lett.*, **2008**, *100*, 120801-1-4.
23. D. R. Leadley, R. J. Nicholas, D. K. Maude, A. N. Utjuzh, J. C. Portal and C. T. Foxon,"Fractional quantum Hall effect measurements at zero g factor", *Phys. Rev. Lett.*, **1997**, *79*, 4246-4249.
24. K. N. Shrivastava, "The derivation of the von Klitzing's constant", *Natl. Acad. Sci. Lett. (India)*, **2009**, *32*, 243-249.
25. K. N. Shrivastava, "Pfaffian states: Quantum computation", *Amer. Inst. Phys. Conf. Proc. (USA)*, **2009**, *1169*, 241-246.
26. K. N. Shrivastava, "Negative-spin quasiparticles in quantum Hall effect", *Phys. Lett. A*, **2004**, *326*, 469-472.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

Maejo International Journal of Science and Technology

ISSN 1905-7873

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

Full Paper

Effect of heat on laterised concrete

Efe E. Ikponmwosa* and Musbau A. Salau

Department of Civil and Environmental Engineering, University of Lagos, Lagos, Nigeria

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

Received: 19 March 2009/ Accepted: 3 February 2010 / Published: 11 February 2010

Abstract: This study presents the results of investigation of the effects of temperature variation on the compressive strength of laterised concrete. Cube specimens were cast, cured in water at ambient laboratory temperature and subjected to different temperature regimes before testing. A concrete mix ratio of 2:3:6 (cement: laterite/sand: granite) with water/cement ratio of 0.65 was adopted for this investigation. The laterite content in the fine aggregate was varied from 0 to 100% at 25% interval. Specimens cured for 7 and 28 days were subjected to uniaxial compressive loading tests at room and elevated temperatures of 250, 500 and 750^{o}C. The results show that normal concrete cannot withstand appreciable load above 250^{o}C while laterised concrete with 25% laterite in the fine aggregate is able to resist higher load with increase in age and at temperature up to 500^{o}C. It is also observed that there is no appreciable increase in strength at higher temperatures. The peak compressive strength value of 30.44 N/mm^{2} is recorded for the mix with 25% laterite-75% sand at 500^{o}C. This is an indication that the strength of laterised concrete is generally sufficient for use at elevated temperature not exceeding 500^{o}C.

Keywords: laterite, laterised concrete, compressive strength

Introduction

Laterised concrete is one in which the fine aggregate has been partially or wholly substituted with laterite soil in its natural form [1-2]. This soil in readily available in Nigeria and constitutes one of the locally and readily available but underutilised building materials. Its neglect as a structural engineering material is associated with the uncertainty of its strength and other structural

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

characteristics. Osunade et al. [3] investigated the shear and tensile strength properties of laterised concrete under laboratory temperature of 20±1^{o}C. They observed that, as in normal concrete, the strength development of test specimens was more rapid at an early curing age than at later age. A higher percentage of the long-term shear and tensile strength of laterised concrete was significantly acquired at an early curing age. Salau [2] investigated long-term deformation of short columns of laterised concrete without taking into consideration the change in temperature and concluded that laterised concrete specimens experience more creep and shrinkage deformation when compared to their corresponding normal concrete specimens. A consistent pattern of creep-time curves in all cases of laterite content was obtained. The shrinkage-time curves were also observed to be consistent but different from the creep-time curves.

Ikponmwosa and Falade [4] reported on the study of strength properties of fibre-reinforced laterised concrete under normal laboratory temperature. A consistent trend of increase in strength with age was observed in the specimens. A proportion of 45% laterite content as replacement of sharp sand in concrete produced the highest compressive strength. At this laterite content, a reduction of 18% in the cost of fine aggregate in concrete was obtained at the prevailing market price. Although the strength characteristic of laterised concrete was found to be generally lower than that of normal concrete, it was sufficient for use in general concrete work. Concrete with 25% laterite content in the fine aggregate compared favourably with those of normal concrete of similar mix proportion by weight and water/cement ratio, and thus was evidently desirable for use in the determination of the effect of heat on laterised concrete.

Oluwaseyi and Mnse [5] investigated the weathering characteristics of laterised concrete with laterite-granite fines ratio as a factor in ascertaining its suitability as a substitute for the conventional fine aggregate. They found that the compressive strength of laterised concrete with laterite-granite fines decreased when subjected to alternate wetting and drying. It was also observed that laterised concrete with 40-60% laterite-granite fines subjected to a temperature variation range of 75-125^{o}C attained compressive strength of 22.52 N/mm^{2}. However, the critical failure temperature of the laterised concrete was not ascertained.

Nijland and Larbi [6] reported that heating of concrete by fire might result in a variety of structural failure such as cracking, spalling, debonding of aggregate and rebars, expansion and mineralogical/chemical changes such as discolouration, dehydration and dissociation. They stated that when a concrete structure is exposed to fire, differential expansion and contraction of various components and constituents within the concrete take place. Also, depending on the period of exposure to the fire, considerable changes may take place in a structure ranging from surface cracking and spalling, discolouration of the concrete, excessive expansion, buckling, warping and loss of strength of reinforcement if present, to cracking around and across aggregate, loss of bond to aggregate and reinforcement, and internal cracking within the cement paste. The failure temperature, however, was not ascertained. St. John et al. [7] stated that with regard to cement paste, evaporation, dissolution, dehydration and dissociation of ettringite, gypsum, calcium hydroxide, calcium carbonate

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

and other phases such as calcium silicate hydrates in the cement paste may occur. The causes of this phenomena need to be ascertained.

Hansen and Ericsson [8] studied the effects of temperature change between room temperature and 100oC on the behaviour of cement paste, mortar and normal concrete under load. Results of their investigation show that cement paste and mortal beams deflect excessively when heated after application of load. Their findings also indicate that deflection occasionally leads to failure at low stresses and after moderate heating. It was also observed that deflection increases with higher rate of heating and the temperature at failure is lower for cement paste than for cement mortar. Likewise, deflection was observed to be large and the temperature at failure lower for saturated beams than for dry beams. Lastly, the researchers opined that rapid rates of heating permanently reduce the modulus of elasticity of cement mortar: an indication of internal destruction of the material structure. The study further concluded that thermal cycling leads to excessive deflection and occasionally to failure. It is, however, important to know the behaviour of normal concrete at higher temperature up to 1000^{o}C.

Salau [9] presented a report of comparative study of heat-resistant characteristics of normal concrete and concrete containing basalt as aggregate. The study revealed that the water adsorption and porosity of basalt concrete falls within an acceptable range for normal refractory material, and that while basalt concrete specimens display some coherence and stability in volume expansion, the normal concrete starts chipping, cracking and spalling at temperature above 450^{o}C. Furthermore, at the end of the first heating-cooling cycle, the subsequent cycles have no significant effect on the weight loss of both types of concrete, and at the end of the third heating-cooling cycle, the normal concrete has lost its entire strength while the basalt concrete still shows more than 17% residual compressive strength which it retains even after the sixth cycle.

Castillo [10] reported on the effect of transient high temperature on the uniaxial compressive strength of high-strength normal concrete. The temperatures studied varied from 100 to 800oC. The presence of loads in the structure was simulated by preloading the test specimens during the heating period. It was found that exposure to temperatures between 100-300^{o}C reduces the compressive strength of high-strength concrete by 15-20%. For temperatures between 400-800^{o}C, the compressive strength decreases to 30% of that at room temperature. A third of the preloaded specimens failed explosively during the heating period. In the remaining specimens the presence of a preload had a beneficial effect and a smaller loss of strength was observed compared to unstressed specimens. The author concluded that exposure to high temperatures causes the modulus of elasticity to decrease in all specimens irrespective of the preload condition and the strength of concrete.

Laterised concrete has been studied for its suitability for use in the construction of structural members. However, the effect of heat on this type of concrete is not well documented. Due to the plasticity and fineness of laterite fines compared to sharp sand, the effect of temperature variation will definitely influence the strength characteristics of laterised concrete. This work therefore aims to study its performance when exposed to different temperature regimes.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

Materials and Methods

The fine aggregate (i.e. sharp sand) used in this study was obtained from an upland river bed and was passed through a 2.36-mm sieve and retained on 63-µm sieve while the coarse aggregate was crushed granite with 12-mm maximum grain size. Laterite fines, reddish in colour, absorbent and non-granular, was obtained from Arepo area of Ogun State. The cement was ordinary Portland cement with properties conforming to British standard BS 12 [11] and with average bulk density in the range of 3050-3150 kg/m^3. The water used was clean, potable water, free from impurities. The dry densities of the constituent materials were determined in the laboratory. The test samples of the sharp sand, laterite and granite chips were obtained in accordance with BS1377-1 [12] and tests whose results are reported in this investigation were carried out in accordance with BS1377-2 [13]. The test samples were air-dried and particle size distribution analyses of the aggregates carried out. The test sieve openings used for the particle size distribution analysis for sand and laterite samples ranged from 63 micron to 3.2 mm while for the granite chippings they ranged between 2.0-19.0 mm.

The coefficient of uniformity (C_u) and coefficient of curvature (C_c), which are used to standardise gradation criteria for the sand, laterite and granite, are obtained from the relationships $C_u = (D_{60}/D_{10})$ and $C_c = ((D_{30})^2/(D_{60} * D_{10}))$ where D_{60}, D_{30} and D_{10} = diameter (mm) of the 60%, 30% and 10% passing size respectively [14].

The liquid limit (LL) of a soil sample, which is a measure of the level of water content at which the soil changes from plastic to liquid, was determined by the cone penetrometre method [15] based on the measurement of penetration of a standardised cone of specific mass into the soil from which material retained on a 425-µm test sieve has been removed. Also, the plastic limit (PL), which shows the level of water content at which the soil whose material retained on a 425-µm test sieve has been removed starts to exhibit plastic behaviour, was determined by the 'soil snake test' [15].

The plasticity index (PI), a measure of the plasticity of the soil samples, was determined for both the sand and laterite. This index indicates the water content at which the soil specimens exhibit plastic properties. The plasticity index is the difference between the liquid limit and the plastic limit, i.e. $PI = LL - PL$.

The slump test [16] was carried out on the fresh concrete test specimens to determine the consistency. The mould for the slump test measures 305 mm in height. The base diameter is 203 mm while the smaller opening at the top is 102 mm. The slump cone is filled in three layers with tamping between each filling to remove voids. The concrete is levelled off at the top of the cone. With the cone removed, the height of the slump is then measured.

The concrete mix proportion of 2:3:6 (cement: fine aggregate: coarse aggregate) by weight with a water/cement ratio of 0.65 was adopted throughout the experiment with the fine aggregate being a mixture of laterite fines and/or sharp river sand. The percentage of laterite in the fine aggregate was varied between 0-100% at 25% interval. The normal concrete specimens, i.e. without laterite, served as control for the experiment. A total of 120 150-mm-cube specimens were cast, cured and tested. The preparation and curing of the specimens were in accordance with British Standard BS1881 [16-17]. The specimens were demoulded $24 \pm ½$ hours after casting and stored in a curing tank

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

containing clean water until testing. After 7 and 28 days of curing, the cube strength was determined on a 600 kN Avery Denison universal testing machine using a loading rate of 120 kN/min.

At each curing age, the cubes were removed from the curing tank. The average weight and density of three cubes for each age were initially determined and the specimens were left for 2 hours before testing. This was followed by the crushing test of the specimens without heating for compressive strength determination. For each mix proportion, 3 cubes were subsequently subjected to heating for one hour at 250, 500 or 750^{o}C in a carbolite furnace with regulated temperature up to 1000^{o}C, and allowed to cool in air for 24 hours before testing.

Results and Discussion

Preliminary tests

The preliminary testing of the concrete constituent materials showed that the coarse aggregate ranged in particle sizes between 3.12-12.7 mm, had a dry density of 2930 kg/m^3, a uniformity coefficient of 2.18 and a coefficient of curvature of 1.23. The dry density of the sieved sand was 2683 kg/m^3 with a uniformity coefficient of 1.81. The laterite fines also had the same particle sizes as those of the sand but with a dry density of 2550 kg/m^3 and a uniformity coefficient of 2.0. The coefficients of curvature were 1.09 and 0.90 for sand and laterite respectively. The grading curves of the aggregates are shown in Figure 1. The moisture content of the lateritic soil, sharp sand and granite was 13.61, 6.14 and 0.6% respectively. The liquid limit of the laterite was 54.0% while the plastic limit was 21.8% and the plasticity index was 32.2%.

The average density of the plain concrete specimens was 2473 kg/m^3 while for the laterised concrete the value varied in the range of 1896-2459 kg/m^3 depending on the laterite content and the elevated temperature. Generally, a trend of significant decrease in weight and density of the test specimens with higher laterite content and temperature was observed. This confirms the findings of Balogun [18] who reported the average densities of laterised concrete specimens in the range of 1980-2540 kg/m^3. However, the reduction in density of the test specimens as the temperature increases which averages above 15% cannot be classified as insignificant. From the slump test results the mix without laterite (i.e. normal concrete) had the highest workability value of 65 mm. With the incorporation of laterite fines in the mix, the value decreased. This might be due to the finer particle sizes associated with laterite fines.

Colouration and deterioration of concrete specimens

At room temperature, the initial colour of the normal concrete cubes was grey while that of the laterised concrete cubes was reddish due to the colour of the incorporated laterite fines. Up to 250^{o}C, the colour of the test specimens did not change. However, as the temperature increased from 250 to 500^{o}C, the colour of the normal specimens changed from grey to pinkish brown while that of the laterised specimens changed from reddish to dark brown. In the 500-750^{o}C temperature range, the colour of the normal specimens changed from pinkish brown to dark brown while that of the laterised

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

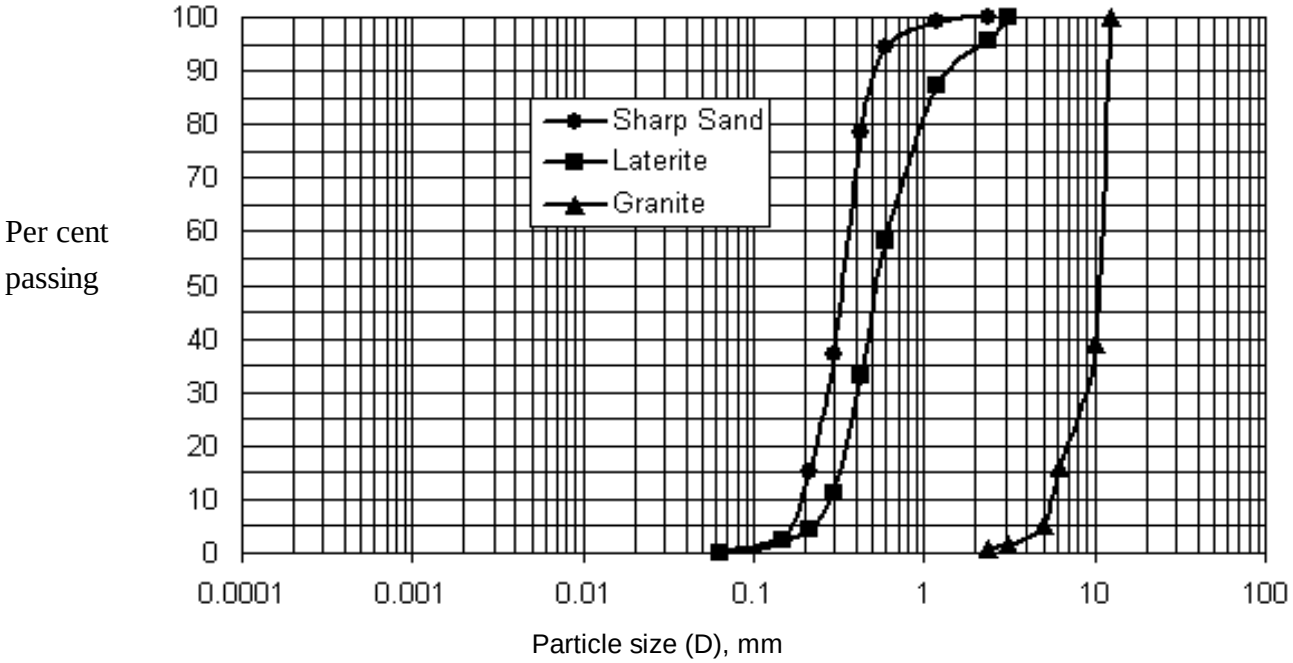


Figure 1. Particle size distribution of the aggregates

specimens changed from dark brown to a very dark brown. This might be due to the effect of heat on the iron oxide content in the laterite. The normal cubes started showing minor signs of cracking and spalling on the exposed surface at temperature between 400 and 600^{o}C. This is in agreement with the observation made by Salau [9]. In the case of laterised concrete, minor cracking and spalling occurred above 750^{o}C.

Effect of temperature on weight and density of test specimens

The variation of specimen weight with temperature for various values of laterite content is shown in Table 1 and Figure 2. For all specimens (i.e. normal and laterised) the weight decreases with increase in temperature. At 7 days of curing, the weight generally decreased with higher laterite content except for the specimens subjected to 750^{o}C heat. At ambient temperature (21^{o}±0.5^{o}C), the weight of normal concrete decreased from 8.3-8.4 kg to 7.8-7.9 kg for concrete with 100% laterite fines as aggregate indicating a weight loss of about 6 % due mainly to the density effect..

Maximum loss in weight of 16.7-18.1% was recorded for normal concrete subjected to 750 ^{o}C heat at 28- and 7-day curing ages respectively when compared with weight of concrete at room temperature. Specimens with 25-100% laterite fines recorded a reduction in weight varying between 14.4-5.1% and 14.6-5.1% and for 28- and 7-day curing ages respectively. The greater reduction in weight for normal concrete should be due to the near-total loss of water at higher temperature from all the voids while lateritic clay, having finer particles and lower density than sand, seems to exhibit the capacity to retain more moisture in its structure under the same condition [2].

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

Table 1. Average weights of specimens (kg) after heating at different temperatures

% Laterite in fine aggregate	21 ± 0.5°C		250°C		500°C		750°C	
	7-Day curing	28-Day curing	7-Day curing	28-Day curing	7-Day curing	28-Day curing	7-Day curing	28-Day during
0	8.3	8.4	8.0	7.7	7.6	7.3	6.8	7.0
25	8.2	8.3	7.8	8.0	7.3	7.7	7.0	7.1
50	8.1	8.1	7.8	8.0	7.3	7.7	7.0	7.1
75	7.9	8.0	7.6	7.8	7.4	7.4	6.4	7.2
100	7.8	7.9	7.7	7.8	7.5	7.7	7.4	7.5

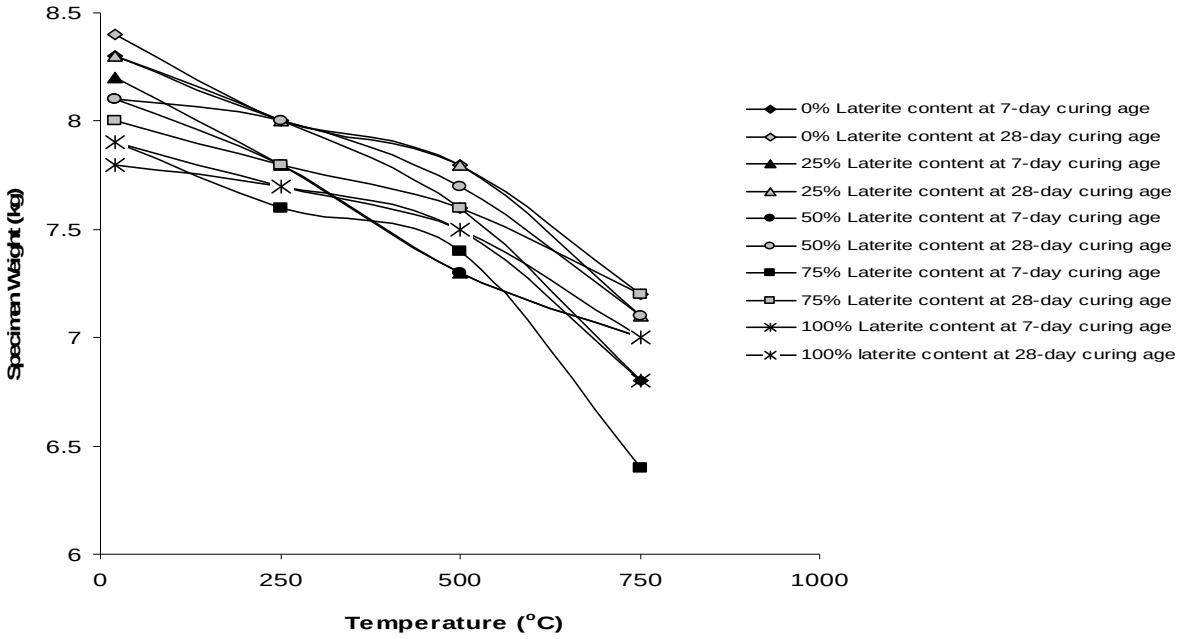


Figure 2. Variation of specimen weight with temperature for various values of laterite content at 7 and 28 days

From the above results, it can also be seen that beyond 500°C the weight loss is highest in normal concrete cubes. This could stem from the fact that normal concrete continues to chip with cracking and spalling at an average temperature of 450°C as reported by Salau [9]. This shows that the inclusion of laterite fines seems to reduce the negative effect of temperature on concrete.

Compressive strength at varying temperature

Table 2 and Figure 3 show the variation of compressive strength with temperature for the test concrete cubes with various levels of laterite content. It can be seen that for all concrete specimens (normal and laterised) the compressive strength increases with temperature up to 250°C. However, beyond 250°C compressive strength of normal concrete (at 750°C) decreases to about 50% of the room temperature values. For concrete with 25-75% laterite content the strength continues to increase

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

appreciably up to 500°C, then decreases. However, the decrease is not appreciable for specimens containing 50% or more of laterite. It is also apparent that concrete with 25% laterite fines can still function well with increased load-bearing capacity when subjected to heat not exceeding 500°C for one hour. The maximum compressive strength of the test specimens is also achieved with 25% laterite content at about 500°C test temperature for both 7- and 28-day-cured specimens. Generally, the improved crushing load and compressive strength characteristics of test specimens when heated up to 500°C may be attributed to improved bonding between the constituent materials as a result of the heating process.

Table 2. Average compressive strength (N/mm^2) of concrete specimens with different percentages of laterite

% Laterite in fine aggregate	21° ± 0.5°C		250°C		500°C		750°C	
	7-Day curing	28-Day curing	7-Day curing	28-Day curing	7-Day curing	28-Day curing	7-Day curing	28-Day curing
0	19.15	20.44	24.71	26.76	19.64	20.27	9.11	10.58
25	21.87	23.02	25.42	28.36	28.67	30.44	11.64	13.69
50	20.53	21.96	23.02	24.13	24.00	25.11	12.00	18.76
75	19.20	20.36	21.16	21.87	22.04	22.27	16.49	19.29
100	14.13	16.53	18.98	19.20	19.56	20.80	20.53	21.96

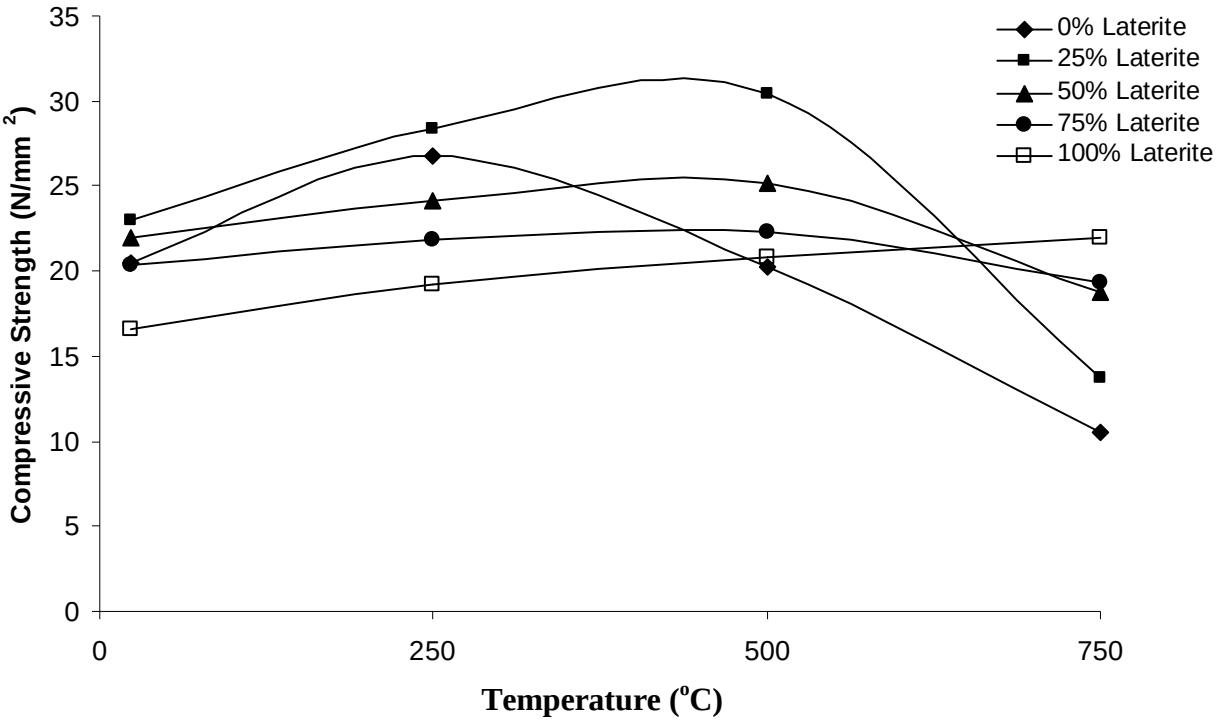


Figure 3. Variation of compressive strength with temperature for concrete of 28-day curing age with different levels of laterite content

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

Conclusions

Laterised concrete can be classified as normal-weight concrete as the density of all test specimens of 28-day curing age exceeds 2000 kg/m^3 based on the cube weight of not less than 7.0 kg even at 750^{o}C.

The compressive strength of normal concrete increases with temperature up to 250^{o}C. However, concrete with 25% laterite in the fine aggregate can withstand heat up to 500^{o}C with increased compressive strength, which is about 30 N/mm2. There is no appreciable difference in strength of specimens of the two different ages when subjected to high temperatures.

From the results of the investigation, it is recommended that concrete with 25% laterite fines can be used in structural components to withstand temperature variation up to 500^{o}C. There should be a considerable economic saving if laterised concrete is used in areas of high temperature up to 500^{o}C. In such situation the application of abundant and locally available laterite is ideal as a construction material.

References

1. F. Falade, "Structural design and economy of normal and lightweight concrete—a comparative study", Proceedings of 3rd International Conference on Quality, Reliability and Maintenance (QRM 2000), **2000**, University of Oxford, UK, pp.219-222.
2. M. A. Salau, "Long-term deformations of laterised concrete short columns", *Build. Environ. J.*, **2003**, *38*, 469-477.
3. J. A. Osunade, P. E. Adeyefa and F. Lasisi, "Shear and tensile strengths of unreinforced laterised concrete", *Ife J. Technol.*, **1990**, 2, 71-84.
4. E. E. Ikponmwosa and F. A. Falade, "A study on the properties of fibre reinforced laterised concretes", *J. Raw Mater. Res.*, **2006**, *3*, 46-55.
5. L. Oluwaseyi and M. A. Mnse, "The Influence of weather on the performance of laterised concrete". *J. Eng. Appl. Sci.*, **2007**, *2*, 129-135.
6. T. G. Nijland and J. A. Larbi, "Unraveling the temperature distribution in fire-damaged concrete by means of PFM microscopy: Outline of the approach and review of potentially useful reactions", *Heron*, **2001**, *46*, 253-264.
7. D. A. St. John, A. W. Poole and I. Sims, "Concrete Petrography: A Handbook of Investigative Techniques", Arnold, London, **1998**.
8. T. C. Hansen and L. Ericsson, "Temperature change effect on behaviour of cement paste, mortar, and concrete under load", *J. Am. Concrete Insti.*, **1966**, *63*, 489-504.
9. M. A. Salau, "A comparative study of heat resistant characteristics of normal concrete with concrete containing basalt as aggregates", *J. Eng. Res.*, **1990**, *JER-1*, 1-11.
10. C. Castillo, "Effect of transient high temperature on high-strength concrete", *MS. Thesis*, **1987**, Rice University, USA.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 33-42

11. British Standard BS12, "Portland Cement (Ordinary and Rapid Hardening), Part 2", British Standard Institution, London, **1991**.
12. British Standard BS1377-1, "Methods of Test for Soils for Civil Engineering Purposes. Classification Test", British Standard Institution, London, **1990**.
13. British Standard BS1377-2, "Methods of Test for Soils for Civil Engineering Purposes. General Requirements and Sample Preparation", British Standard Institution, London, **1990**.
14. B. M. Das, "Fundamentals of Geotechnical Engineering", International SI Edn., Thompson Learning, Toronto, **2008**.
15. S. K. Garg, "Geotech Engineering: Soil Mechanics and Foundation Engineering", Khanna Publishers, Delhi, **2008**.
16. British Standard BS 1881, "Methods of Testing Concrete, Part 2", British Standard Institution, London, **1983**.
17. British Standard 1881, "Methods of Making and Curing Test Specimens, Part 3", British Standard Institution, London, **1983**.
18. L. A. Balogun, "Effect of temperature on the residual compressive strength of laterized concrete", *Build. Environ.*, **1986**, *21*, 221-226.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

Maejo International Journal of Science and Technology

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

Communication

Modelling of natural convection along isothermal plates and in channels using diffusion velocity method

Ademola A. Dare* **and Moses O. Petinrin**

Department of Mechanical Engineering,University of Ibadan, Nigeria

* Corresponding author e-mail: ademola.dare@mail.ui.edu.ng

Received: 19 January 2009/ Accepted: 5 February 2010 / Published: 17 February 2010

Abstract: The diffusion velocity method (DVM), a version of vortex element method (VEM), was used to model the steady state, laminar natural convection flows along isothermal vertical plates and in isothermal vertical channels. For each case, numerical models were developed using DVM from the vorticity transport equation and the energy equation. This study shows that the diffusion velocity method is a viable numerical tool at modelling not only fluid flow problems but also the heat transfer problems.

Keywords: natural convection, isothermal plates, isothermal channels, diffusion velocity method

Introduction

Many engineering problems are represented by non-linear, partial differential equations. In order to simplify these problems which may be difficult to solve analytically, different numerical methods are used to analyse the problems. Some of these numerical methods are: the finite difference method, the finite element method and the Monte Carlo method.

Another numerical method that has been used successfully to study and analyse very complex, unsteady fluid flows and thermal engineering problems is the vortex element method (VEM). This method is simple to implement as it requires simple mathematics compared to other numerical methods that may require many mathematical operations such as variational calculus and matrix inversions [1].

An engineering problem involves flows of gases or liquids over solid bodies. Examples include: air flows over cars and aeroplanes; wind blowing over bridges and buildings; and sea waves

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

slashing against the supporting columns of an off-shore oil rig. Often these flows do not follow the contour of the solid surface completely, but may be formed separate from it, e.g. a wake behind a ship. Such separated flows are difficult to handle by conventional numerical schemes. The main reason for basing the numerical method on the vorticity is that typically only a small portion of the flow contains vorticity. This can lead to significant savings in storage and computational effort [2]. A number of numerical schemes model diffusion by vortex methods without using a mesh, which is based on the Lagrangian approach. However, the VEM is grid-free in implementation [2-3].

Recent applications of the vortex methods based on the Biot-Savart law have been extended to numerical prediction of unsteady and complex characteristics of various flows related to difficult engineering problems concerning flow-induced vibration, off-design operation of fluid machinery, automobile aerodynamics, biological fluid dynamics, etc. Kamemoto [4] described how VEM, with its simple algorithms based on physics of flow, has been used to find the virtual operation of fluid machinery (pumps and water turbines), and the calculated flows around bluff bodies, an oscillating airfoil and a swimming fish.

Cheng et al. [5] developed a hybrid vortex method to simulate two-dimensional viscous incompressible flows over a bluff body. It was based on a combination of the diffusion-vortex method and the vortex-in-cell method whereby the flow field is divided into two regions. In the region near the body surface the diffusion-vortex method is used to solve the Navier-Stokes equations while the vortex-in-cell method is used in the exterior domain. They compared the results obtained with those from the finite difference method and other vortex methods and experiments. Which showed that the method is well adapted to calculate two-dimensional external flows at high Reynolds number.

Ghoniem and Oppenheim [6] applied random-walk vortex method to an assortment of problems of diffusion of momentum and energy in one dimension as well as heat conduction in two dimensions in order to assess its validity and accuracy. The numerical solutions obtained were found to be in good agreement with the exact solution except for a statistical error introduced by using a finite number of elements. They claimed further that the error could be reduced by increasing the number of elements or by using ensemble averaging over a number of solutions.

The concept of the diffusion velocity method (DVM), a version of the VEM, was extensively discussed by Ogami [7]. The velocity is defined in order that the vorticity is conserved in the transfer of diffusion process as it is so in the convection process. Unlike the other vortex element methods, this technique handles vorticity equation in a deterministic manner by calculating the diffusion velocity to account for diffusion in the flow [8]. Ogami [8] used the DVM to simulate the diffusion of vorticity and temperature from one-dimensional vorticity transport equation and compared the results with the analytical solutions. The method was successfully used to treat the diffusion equation (Re = 0), the boundary layer and two-dimensional flows around a circular cylinder (Re = 0.1 ~ 10^7), aerofoil, the Burger equation, and the equations of incompressible fluid.

This study employs the diffusion velocity technique to model the natural convection along isothermal plates and in isothermal channels. The channels, consisting of two parallel vertical plates, are asymmetrically heated at uniform wall temperature. The Nusselt numbers are obtained with the velocity and the temperature distributions.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

The Governing Equations

The continuity equation and the Navier-Stokes equations for Newtonian, two-dimensional, incompressible flow are presented as follows:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \quad (1)$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=-\frac{1}{\rho}\frac{\partial P}{\partial x}+f_x+\nu\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right) \quad (2)$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}=-\frac{1}{\rho}\frac{\partial P}{\partial y}+f_y+\nu\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right) \quad (3)$$

Also, the energy equation, assuming no viscous dissipation or thermal generation, is

$$\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}=\alpha\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}\right) \quad (4)$$

Variables u and v are the velocity components in x and y directions respectively, f_x and f_y are the components of the gravitational acceleration in the x and y directions respectively, T is the fluid temperature, P is the fluid pressure, t is the time, α is the fluid thermal diffusivity, ν is the kinematic viscosity, and ρ is the fluid density [9-10].

Following the Lagrangian scheme, the alternative expression of the governing equations of viscous and incompressible flow gives the vorticity transport equation as

$$\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}=\nu\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}\right)+g\beta\frac{\partial T}{\partial y} \quad (5)$$

where vorticity, $w=\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}$ (6)

Considering the diffusive term only in the vorticity equation, then

$$\frac{\partial w}{\partial t}=\nu\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}\right) \quad (7)$$

The solution to equation (7) was given by Batchelor [11] as

$$w(r,t)=\frac{\Gamma}{4\pi\nu t}e^{(-r^2/4\nu t)} \quad (8)$$

where $r=\sqrt{(x^2+y^2)}$ and Γ is the vortex strength or circulation [12]

Numerical Formulations

The initial velocity, γ, is induced by a buoyancy effect created by the temperature difference between the plate and the fluid in contact with its surface, and is given as

$$\gamma=-g\beta\frac{\partial T}{\partial y}\Delta t\Delta s \quad (9)$$

where Δs is the elemental length, Δt is the time step, β is the volumetric thermal expansion coefficient, and g is acceleration due to gravity.

Velocity and temperature vortices with strengths Γ and Θ respectively, which are created on one elemental surface of m elements into which a plate is divided are given by

$$\Gamma_q=\gamma\Delta s \quad (10)$$

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

$$\Theta_q = Tu_{Tq}\Delta t = -\alpha\Delta t\left.\frac{\partial T}{\partial y}\right|_{y=\delta} \quad (11)$$

where $m = \frac{L}{\Delta s}$; $\delta = \sqrt{\alpha\Delta t}$; q = 1, 2, 3,...m; and L is the plate length.

The elemental length has a great influence on the strength of the vortices. The vortices are initially distributed and separated at distance Δs before diffusing. The vorticity of each velocity vortex and the temperature of each temperature particle are respectively given by

$$w(r,\Delta t) = \frac{\Gamma_i}{4\pi\nu\Delta t}\left[e^{(-r_1^2/4\nu\Delta t)} - e^{(-r_2^2/4\nu\Delta t)}\right] \quad (12)$$

$$T(r,\Delta t) = \frac{\Theta_i}{4\pi\alpha\Delta t}\left[e^{(-r_1^2/4\alpha\Delta t)} - e^{(-r_2^2/4\alpha\Delta t)}\right] \quad (13)$$

For flow on a flat plate, each velocity vortex or temperature particle has a corresponding negative image. The distances between a vortex and the inducing vortex and its corresponding image can be deduced respectively by

$r_1(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$ and $r_2(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (y_j + y_i)^2}$

Here $i = j$ = 1, 2, 3,…N, and N is the number of vortices on the plate surface and in space. However, i is not equal to j most of the times. Figure 1 shows an inducing vortex with its image and how r_1 and r_2 are determined from a vortex of interest.

The divergences of equations (12) and (13) are respectively $\nabla.w$ and $\nabla.T$, which can be written in the forms:

$$\nabla.w = \frac{\partial w}{\partial xj} + \frac{\partial w}{\partial y_j} \quad (14)$$

$$\text{and } \nabla.T = \frac{\partial T}{\partial xj} + \frac{\partial T}{\partial y_j} \quad (15)$$

The diffusion velocity of each velocity vortex is induced by all the velocity vortices and their corresponding images. Also, the diffusion velocity of temperature particle is induced by all the temperature particles and their corresponding images. As discussed by Ogami [8], the diffusion velocities of each velocity vortex and temperature particle are now respectively given as

$$u_{wj} = \sum_{i=1}^{N}[-\frac{\nu}{w}\nabla.w] \quad (16)$$

$$u_{Tj} = \sum_{i=1}^{N}[-\frac{\alpha}{T}\nabla.T] \quad (17)$$

The diffusion distances of each velocity vortex and temperature particle can be expressed respectively as

$$d_{vj} = u_{wj} \times \Delta t \quad (18)$$

$$\text{and } d_{Tj} = u_{Tj} \times \Delta t \quad (19)$$

These distances are added to the original positions to move the vortex and particle to other positions. For convection, each vortex is repositioned to a new location using the induced velocities (u, v) by neighbouring velocity vortices.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

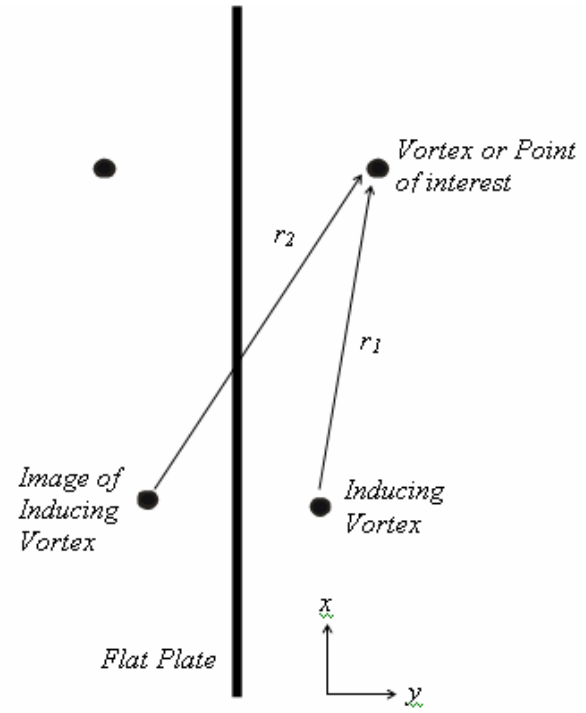


Figure 1. Interaction of vortices on a vertical flat plate

The slip flow condition is met when the velocity on each elemental surface is approximately zero and the numerically calculated temperature is approximately equal to the initial temperature of the plate. To check for the slip flow condition, the velocity on the elemental surface can be determined as

$$v_e = \sum_{i=1}^{N}\left[\frac{\Gamma_i}{2\pi r_1} - \frac{\Gamma_i}{2\pi r_2}\right] \tag{20}$$

The corresponding temperature on the surface is determined as

$$T_e = \sum_{i=1}^{N}\frac{\Theta_i}{4\pi\alpha\Delta t}\left(e^{(-r_1^2/4\alpha\Delta t)} - e^{(-r_2^2/4\alpha\Delta t)}\right) + T_B \tag{21}$$

The practical substitution for the given boundary condition is

$$T_B = T_w erfc(\frac{y_j}{2\sqrt{\alpha\Delta t}}) \tag{22}$$

where T_w is the wall temperature. At each time step, the strength of a vortex is increased by adding $-g\beta\frac{\partial T}{\partial y}\Delta t\Delta s$ to equation (9) and T_e is substituted for T in equation (11) until the slip flow condition is met.

When a vortex is very close to the wall or an inducing vortex, i.e. when $r < \frac{\Delta s}{\pi}$ (r can be r_1 or r_2), then $\frac{\Gamma_i r}{2\pi r_0^2}$ replaces $\frac{\Gamma_i}{2\pi r}$ where it is appropriate. The minimal distance, r_0, is equal to $\frac{\Delta s}{\pi}$. Likewise for temperature particles, $\frac{\Theta_i r}{2\pi r_0^2}$ replaces $\frac{\Theta_i}{2\pi r}$.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

The velocity and temperature distributions at specific locations in the hydrodynamic and thermal boundary layers are determined when the slip flow condition is meant for the velocities and temperatures on the plate surfaces. The two components of the velocity distribution are obtained as

$$v_{mnx}(r) = \sum_{i=1}^{N}\left[\frac{\Gamma_i}{2\pi r_1}\left(\frac{y_j - y_i}{r_1}\right) - \frac{\Gamma_i}{2\pi r_2}\left(\frac{y_j + y_i}{r_2}\right)\right] \tag{23}$$

$$v_{mny}(r) = \sum_{i=1}^{N}\left[\frac{\Gamma_i}{2\pi r_2}\left(\frac{x_j - x_i}{r_2}\right) - \frac{\Gamma_i}{2\pi r_1}\left(\frac{x_j - x_i}{r_1}\right)\right] \tag{24}$$

The temperature at any point can be obtained from

$$T_{mn}(r) = \sum_{i=1}^{N}\left[\frac{\Theta_i}{2\pi r_1} - \frac{\Theta_i}{2\pi r_2}\right] + T_B \tag{25}$$

where *m* and *n* are the positions of the velocity or temperature along and normal to the plate respectively.

Natural Convection in Channels

The type of channels described here are two parallel plates placed vertically. The number of images of a vortex or a particle is infinite, but for easy computation the number of the images is reduced to eight; therefore, we have one positive vortex with four positive images and four negative images [1].

As discussed by Petinrin [13], the vorticity of each velocity vortex and the temperature of each temperature particle can then be respectively given as

$$w(r,\Delta t) = \frac{\Gamma_i}{4\pi\nu\Delta t}\left[e^{(-r_1^2/4\nu\Delta t)} - e^{(-r_2^2/4\nu\Delta t)} + e^{(-r_3^2/4\nu\Delta t)} + e^{(-r_4^2/4\nu\Delta t)} + e^{(-r_5^2/4\nu\Delta t)} + e^{(-r_6^2/4\nu\Delta t)} - e^{(-r_7^2/4\nu\Delta t)} - e^{(-r_8^2/4\nu\Delta t)} - e^{(-r_9^2/4\nu\Delta t)}\right] \tag{26}$$

$$T(r,\Delta t) = \frac{\Theta_i}{4\pi\alpha\Delta t}\left[e^{(-r_1^2/4\alpha\Delta t)} - e^{(-r_2^2/4\alpha\Delta t)} + e^{(-r_3^2/4\alpha\Delta t)} + e^{(-r_4^2/4\alpha\Delta t)} + e^{(-r_5^2/4\alpha\Delta t)} + e^{(-r_6^2/4\alpha\Delta t)} - e^{(-r_7^2/4\alpha\Delta t)} - e^{(-r_8^2/4\alpha\Delta t)} - e^{(-r_9^2/4\alpha\Delta t)}\right] \tag{27}$$

Also, the distances of each vortex from the inducing vortex and its images can be determined as r_1 to r_9, where h is the distance of the plates apart, i.e.

$$r_1(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2},$$

$$r_2(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (y_j + y_i)^2},$$

$$r_3(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (y_j - (y_i + 2h))^2},$$

$$r_4(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (y_j - (y_i + 4h))^2},$$

$$r_5(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i + 2h)^2},$$

$$r_6(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i + 4h)^2},$$

$$r_7(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (2h - y_j - y_i)^2},$$

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

$$r_8(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (4h - y_j - y_i)^2}\ ,$$
$$r_9(x_j, y_j) = \sqrt{(x_j - x_i)^2 + (2h + y_j + y_i)^2} \qquad (28)$$

Results and Discussion

The results of the numerical analysis, which was solved with Visual Basic programming language, are automatically displayed on Microsoft Excel Workbook. The velocity and temperature distributions are displayed on the Sheet 1 and Sheet 2 of the workbook respectively.

The input parameters used to simulate natural convection on a single plate lying vertically are listed in Table 1. The fluid (air) thermophysical properties of the fliud are taken at film temperature.

Table 1. The input parameters for natural convection over a single plate

Length of the plates	0.5 m
Fluid (air) temperature	10^0C
Plate wall temperature	30^0C
Coefficient of thermal expansion	0.00341 (K^{-1})
Kinematic viscosity of air at 20^0C	0.0000157 m^2/s
Thermal diffusivity of air at 20^0C	0.000022 m^2/s

The logarithmic plot of Nusselt number against Rayleigh number is presented in Figure 2 by varying the plate wall temperature, *Tw*, from 40^0C to 120^0C while keeping the fluid (air) temperature constant at 10^0C. The fluid properties are taken at film temperature. It can be deduced that the Nusselt number increases with the Rayleigh number for a fixed plate length as in Table 2. The slope of the plot is 3.47 while the intercept is -28.22 on the log scale and 6.08E-29 on the normal scale.

Therefore, the relationship between Nusselt number and Rayleigh number is $Nu = (6.08E\text{-}29)Ra^{3.47}$. Comparing this relationship with $Nu = 0.59Ra^{0.25}$ as reported by Incropera and Dewitt [10], there is much deviation in the two correlations which may be due to convergence difficulties at the plate surface. The correlation coefficient of the plot is 0.9065.

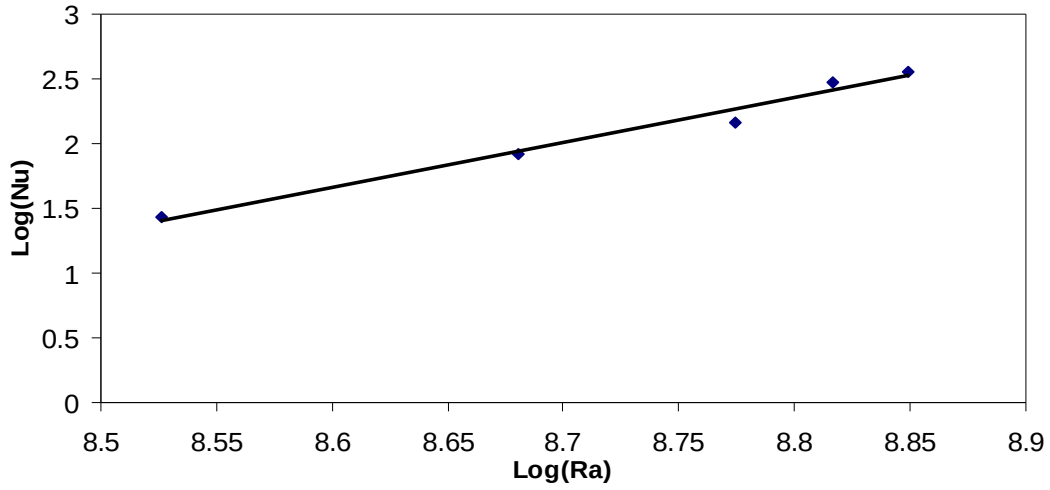


Figure 2. Logarithmic plot of Nusselt number (Nu) against Rayleigh number (Ra) for a vertical plate

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

Table 2. Distribution of Rayleigh number with Nusselt number for natural convection over a single plate

Tw (^0C)	*Ra*	*Nu*	*Log(Ra)*	*Log(Nu)*
40	336123348	27.0125	8.5265	1.4316
60	479235019	84.0570	8.6805	1.9246
80	594247932	144.6929	8.7740	2.1604
100	655562931	297.4379	8.8166	2.4734
120	706467195	362.2841	8.8491	2.5590

The Nusselt number increases with the wall temperature as presented in Figure 3. However, the relationship between them is not linear owing to the Nusselt number being also dependent on some of the fluid properties, e.g. the thermal conductivity, which keeps changing as the wall temperature changes.

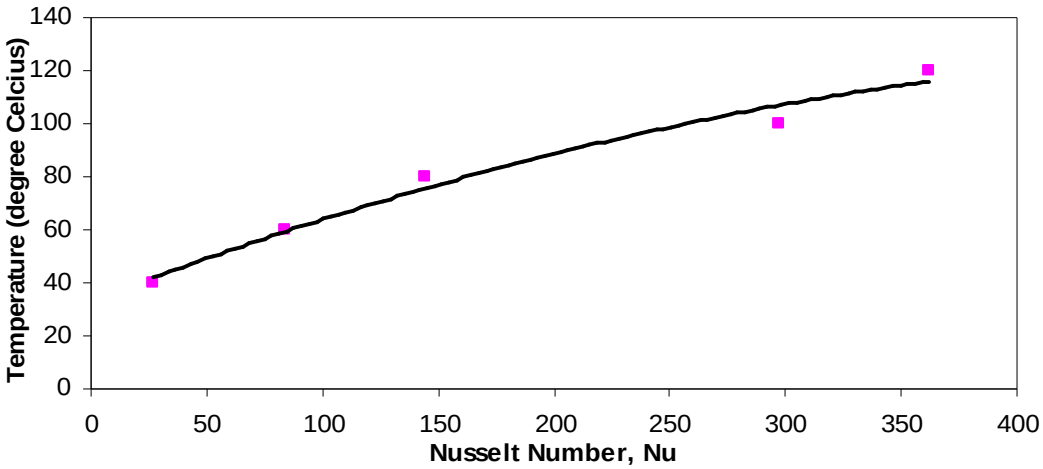


Figure 3. Effect of wall temperature on Nusselt number

The input parameters for the symmetrically heated, isothermal plates listed in Table 3 were used to simulate natural convection in channels. The channels are two parallel plates placed vertically.

The graph of logarithm of Nusselt number against logarithm of Rayleigh number in Figure 4 is obtained by varying the magnitude of the space, *S*, between the channel plates from *S* = 0.1m to 0.3m at a step of 0.05m. Input parameters are fed from Table 3. The slope of the plot is 0.51 while the intercept is -0.75 on the log scale and 0.18 on the normal scale.

Therefore, the relationship between Nusselt number and Rayleigh number is $Nu = 0.18Ra^{0.51}$ (Figure 4). The Figure also shows a deviation in the Nusselt-Rayleigh relationship from other workers. Ofi and Hetherington [14] with finite element method obtained a relationship $Nu = 0.699Ra^{0.26}$; Bodoia and Osterle[15] obtained $Nu = 0.56Ra^{0.25}$; and Ogundare[1] obtained $Nu = 0.43Ra^{0.21}$.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

Table 3. Input parameters for natural convection in vertical channel

Length of plates	0.5 m
Gap between plates	0.1m
Fluid (air) temperature	10^0C
Wall temperature of first plate	60^0C
Wall temperature of second plate	60^0C
Coefficient of thermal expansion	0.00325
Kinematic viscosity of air at 35^0C	0.0000171 m^2/s
Thermal diffusivity of air at 35^0C	0.0000241 m^2/s

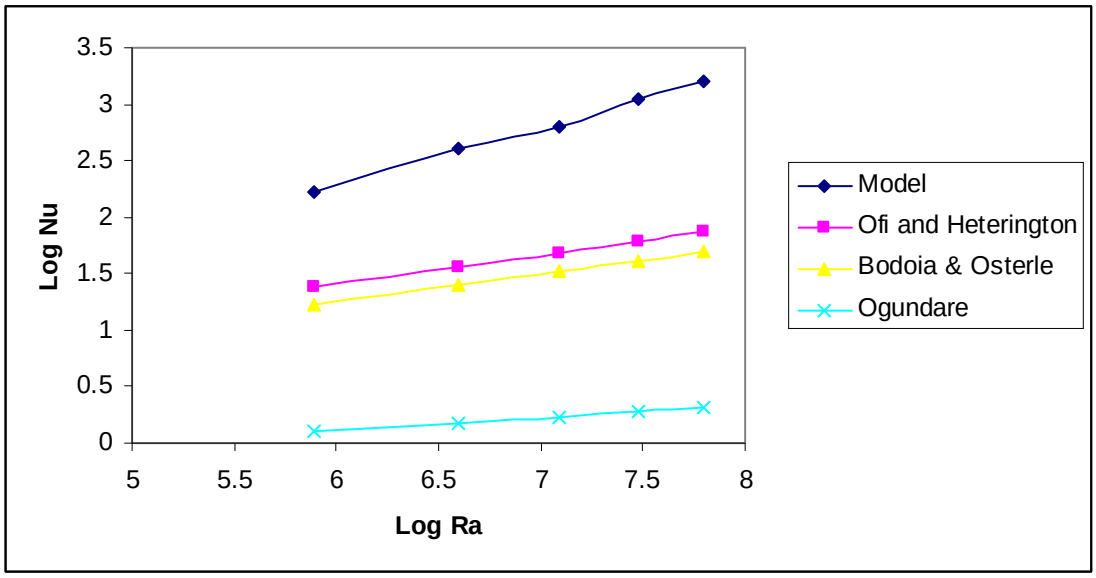


Figure 4. Logarithm plot of Nusselt number (Nu) against Rayleigh number (Ra) for vertical channel

Conclusions

Modelling of natural convection in isothermal vertical plates and channels has been successfully carried out with the diffusion velocity method, a version of the vortex element method. However, a large deviation recorded for correlation of Nusselt number and Rayleigh number for both the plate and the channel with existing correlations may stem from convergence difficulties encountered at the plate surface.

From the results obtained, it is established that as the wall temperature increases while keeping the mainstream fluid temperature constant, the thermal boundary layer thickness increases. The study has also established that the diffusion velocity method is a viable numerical tool capable of modelling fluid and heat transfer problems.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 43-52

References

1. M. A. Ogundare, "Vortex element modelling of fluid and heat flow in vertical channels and ducts", *PhD. Thesis*, **2006**, University of Ibadan, Nigeria.
2. S. A. Subramanian, "New mesh-free vortex method", **1996**, Retrieved 20th Feb, 2008 from *http://www.eng.fsu.edu/~dommelen/papers/subram/style_a/node78.html*
3. L. Liu, J. Feng, J. Fan and K. Cen, "Recent development of vortex method in incompressible viscous bluff body flows", *J. Zhejiang Univ. SCI*, **2005**, *6A*, 283-288.
4. K. Kamemoto, "On contribution of advanced vortex element methods toward virtual reality of unsteady vortical flows in the new generation of CFD", *J. Braz. Soc. Mech. Sci. & Eng*, **2004**, *26*, 368-379.
5. M. Cheng, Y. T. Chew and S. C. Luo, "A hybrid vortex method for flows over a bluff body", *Int. J. Numer. Meth. Fluids*, **1997**, *24*, 253-274.
6. A. F. Ghoniem and A. K. Oppenhim, "Random element method for numerical modeling of diffusional processes", Proceedings of 8th International Conference on Numerical Methods in Fluid Dynamics, **1982**, Aachen, Germany, pp. 224-232.
7. Y. Ogami, "Simulation of heat-fluid motion by the vortex method", *JSME Int. J. Ser. B*, **2001**, *44*, 513-519.
8. Y. Ogami, "Simulation of heat-vortex interaction by the diffusion velocity method", ESAIM Proceedings of 3rd International Workshop on Vortex Flows and Related Numerical Methods, **1999**, Toulouse, France, Vol. *7*, pp 314-324**.**
9. F. Kreith and M. S. Bohn, "Principles of Heat Transfer", 5th Edn., West Publishing Co., St. Paul, **1993**.
10. F. P. Incropera and D. P. Dewitt, "Fundamentals of Heat and Mass Tranfer", 5th Edn., John Wiley & Sons, New York, **2005**.
11. G. K. Batchelor, "An Introduction to Fluid Dynamics", Cambridge University Press, Cambridge, **1967**.
12. R. I. Lewis, "Vortex Element Methods for Fluid Dynamic Analysis of Engineering Systems", University Press, Cambridge, **1991**.
13. M. O. Petinrin, "Diffusion velocity modelling of forced and natural convection in isothermal plates and channels", *MSc. Thesis*, **2008**, University of Ibadan, Nigeria.
14. O. Ofi and A. J. Hetherington, "Application of the finite element method to natural convection heat transfer from the open vertical channel", *Int. J. Heat Mass Transf.*, **1977**, *20*, 1195-1204.
15. J. R. Bodoia and J. F. Osterle, "The development of free convection between heated vertical plates", *J. Heat Transf.*, **1962**, *84*, 40-44.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

Maejo International Journal of Science and Technology

ISSN 1905-7873

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

Full Paper

Crossing points in the electronic band structure of vanadium oxide

Noriza A. Zabidi, Hasan A. Kassim and Keshav N. Shrivastava*

Department of Physics, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia

* Corresponding author, e-mail: keshav1001@yahoo.com

Received: 5 December 2009 / Accepted: 1 March 2010 / Published: 3 March 2010

Abstract: The electronic band structures of several models of vanadium oxide are calculated. In the models 1-3, every vanadium atom is connected to 4 oxygen atoms and every oxygen atom is connected to 4 vanadium atoms. In model 1, a=b=c 2.3574 Å; in model 2, a= 4.7148 Å, b= 2.3574 Å and c= 2.3574 Å; and in model 3, a= 4.7148 Å, b= 2.3574 Å and c= 4.7148 Å. In the models 4-6, every vanadium atom is connected to 4 oxygen atoms and every oxygen atom is connected to 2 vanadium atoms. In model 4, a=b= 4.551 Å and c= 2.851 Å; in model 5, a=b=c= 3.468 Å; and in model 6, a=b=c= 3.171 Å. We have searched for a crossing point in the band structure of all the models. In model 1 there is a point at which five bands appear to meet but the gap is 7.3 meV. In model 2 there is a crossing point between G and F points and there is a point between F and Q with the gap ≈ 3.6608 meV. In model 3, the gap is very small, ~ 10^{-5} eV. In model 4, the gap is 5.25 meV. In model 5, the gap between Z and G points is 2.035 meV, and in model 6 the gap at Z point is 4.3175 meV. The crossing point in model 2 looks like one line is bent so that the supersymmetry is broken. When pseudopotentials are replaced by a full band calculation, the crossing point changes into a gap of 2.72 x 10^{-4} eV.

Keywords: vanadium oxide, band crossing points, supersymmetry

Introduction

Recently, it has been observed that there are crossing points in the band structure. The valence band appears as a cone with apex on top while the conduction band does with apex at the bottom. The apex point of the conduction band sits on top of the apex point of the valence band. If we take a point

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

on the cone of the conduction band with energy E, then for the same wave vector, in the valence band there is a point with energy –E. When the magnitude of the energy E in the conduction band is equal to the magnitude of the energy –E in the valence band, there is supersymmetry; otherwise the supersymmetry is broken. In the theory of relativity, the energy of a particle of momentum p and mass m is given by $\pm (c^2p^2 + m^2c^4)^{1/2}$ where the magnitude of the energy with positive sign is exactly equal to that with negative sign. This is called the supersymmetry. In the Dirac equation, the negative energy states are associated with a particle of positive charge called the positron and the positive energy solutions are associated with the electron. The electron energy solutions found in the non-relativistic Schrödinger equation may be mapped to that of the Dirac equation. The two-energy solutions in $\pm (c^2p^2 + m^2c^4)^{1/2}$ never cross but it is possible to select a point in the middle of the two solutions which is named as the "Dirac point" [1, 2]. The geometric average of the two energies in which one is a constant and the other varies as n will vary as the square root of n. For zero mass, mc^2 term is zero and the c^2p^2 term varies linearly as a function of momentum. This linear dependence with both ± signs forms a crossing point called the "Dirac point" as shown in Figure 1.

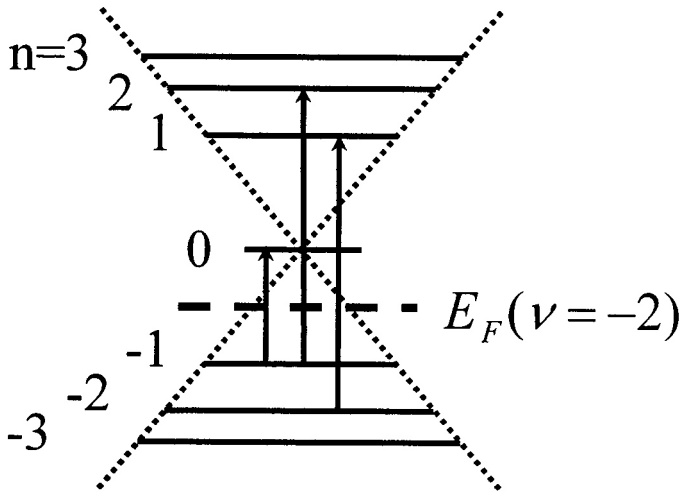


Figure 1. The crossing point of the two dash lines becomes a cone when visualised in three dimensions. The crossing point can be at the top of one cone and at the bottom of another cone and hence matches with the apex point. The position of Fermi energy can be fixed from an independent calculation [2].

The energy mc^2 in the Dirac equation is of the order of 0.5×10^6 eV, which is very large compared with the energies $\approx 10^{-3}$ eV found in the semiconductors. If there is a point in the centre of the conduction and valence bands it should be called the Schrödinger point. The object of the present study is to look for crossing points in the electron energy bands obtained from the solutions of the density-functional theory. In a calculation of band structure of a monolayer of carbon atoms, the crossing point with the supersymmetry was searched [3]. In a recent study of the crossing points and band bending in TiO_2/VO_2 nanostructures, semi-Dirac points have been discussed [4]. The compounds of vanadium show several different valencies. In VO_2, the valency of vanadium is 4. The VO_2 shows several phases as a function of temperature and pressure. For T > 340 K it has a pseudo R phase. For the pressure larger than 200 bar and temperature less than 340 K, the phase is called M_1, and for T < 340 K but larger than -0.0099 P + 340 K, where P is the pressure measured in units of bar, the phase is

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

called M_2 [5]. There is a Peierls distortion so that the vanadium atoms occur in pairs [6]. The electronic configuration of vanadium is $3d^3 4s^2$. In the tetravalent position, it is $3d^1$. Hence, two atoms form a spin triplet. The experimental optical gap is 0.6 eV [7] but the calculated value is -0.04 eV [8]. Hence, it will be of interest to calculate the gap energy. The compounds $V_{2n}O_{5n-2}$ are known as Wadsley phase and V_nO_{2n-1} are called Magneli phase. It is clear that many phases with different valencies occur and there are phase transitions as a function of pressure. The Raman spectra of VO_2 show a phase transition at a pressure of 12 GPa. It has been shown that the metal-insulator transition and the electronic structure are correlated [9]. The M_1 phase of VO_2 is charge ordered [10]. There are a series of avalanches which depend on the particle size before the transition is completed [11]. A point-ion model has been described by Nakatsugawa and Iguchi [12]. The early ideas of the metal-insulator transition were given by Zylbersztejn and Mott [13].

In the present work, we have made six models of vanadium oxide. Out of these six models, three have vanadium valency equal to 2 and the remaining 3 have vanadium valency equal to 4. In all of the cases we optimise the structures and calculate the band stuctures to look for crossing points. We find a crossing point only when pseudopotentials are used in Castep programme. In the full calculation, taking into account all of the electrons, crossing disappears and gaps appear. We report all of the gaps found. The models are calculated using spin unpolarised orbital for both Castep and DMol3 programmes. We take the P1 space group for all of the models.

Materials and Methods

The electronic band structure is calculated by using the density-functional theory which is an approximation of the Schrödinger equation. The potential energy is expressed in the form of electron density and then the energy is minimised to obtain the Kohn-Sham equation [14, 15], which can be solved in local density approximation (LDA) as well as in the generalised gradient approximation (GGA). The actual atoms connect by exchange and correlation integrals so the linear GGA is quite approximate but may be compared with the LDA. We calculate the band structure by using the DMol3 (all electrons) as well as the Castep (pseudopotentials) programmes [16]. We present the results of the calculation of the gap energy in both approximations. All of the results are obtained in LDA and unpolarised orbitals are used. Both the Castep and the DMol3 were kindly provided by Accelrys Software Inc, San Diego, CA. The geometry and the unit cell sizes (a, b and c) are calculated by minimising the energy of the Schrödinger equation for each model.

Results and Discussion

Models of vanadium oxide

Model 1: VO. In this model, V atom is coordinated with 4 oxygen atoms and every oxygen atom is coordinated with 4 vanadium atoms. The V-O distance is 1.667Å and a=b=c=2.3574Å. The coordinates of k points are G (0, 0, 0), F (0, 0.5, 0), Q (0, 0.5, 0.5) and Z (0, 0, 0.5). The model is shown in Figure 2 and the calculated band structure along with the density of states (DOS) is shown in Figure 3, which

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

also identifies the various zone points. The Fermi energy, the binding energy and the gaps at the k points are given in Table 1.

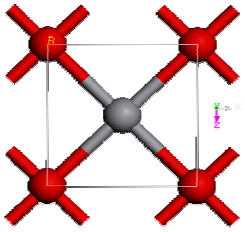


Figure 2. Model 1: The central V atom is connected to 4 oxygen atoms.

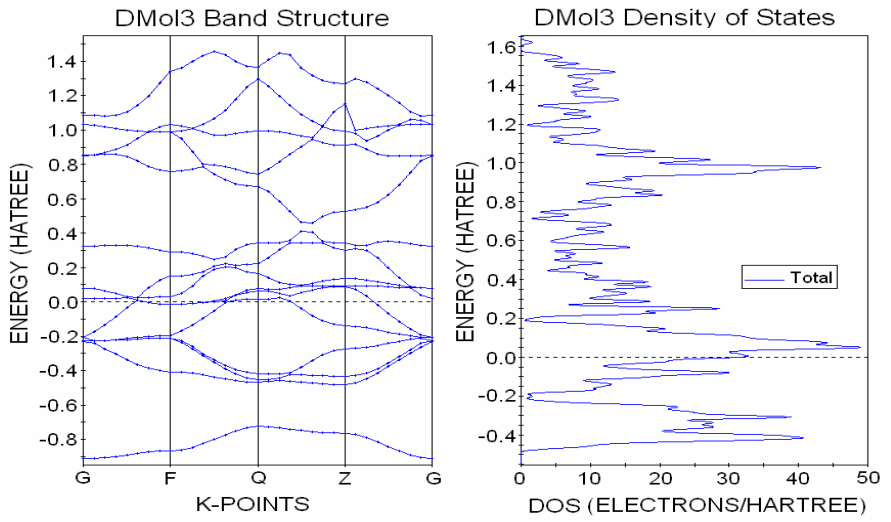


Figure 3. The band structure and the DOS of model 1

Table 1. The Fermi energy, the binding energy and the energy gap at various k points

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fermi energy (eV)		7.2234	7.0312	7.1722	6.6447	6.9399	6.8908
Bind. energy (eV)		10.931	21.8223	43.6524	19.8832	20.4224	17.4416
Energy gap (eV)	G	6.1907	4.2791	0.5851	0.9252	0.8844	0.1361
	F	1.1565	0.8028	0.7891	0.7551	1.2586	1.5579
	Q	11.7012	5.4152	0.2177, 1.1905	1.83	2.5035	1.4558
	Z	6.1907	2.4899	2.9729	0.1905, 2.0953	0.9456	1.1837

Model 2: V_2O_2. In this model every vanadium atom is coordinated to 4 oxygen atoms and every oxygen atom is also coordinated to 4 vanadium atoms. The layering of atoms is such that b=c=2.3574Å

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

and a=4.7148Å. Thus, the cell is much bigger than in model 1. The V-O distance is 1.667Å. The model is shown in Figure 4 and the calculated band structure and the DOS are shown in Figure 5. The gap energies at the k-points are given in Table 1.

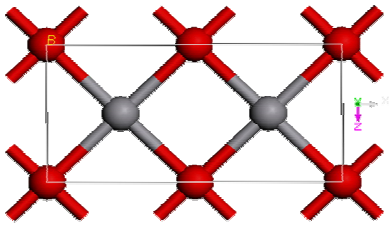


Figure 4. Model 2: V_2O_2 in a large unit cell

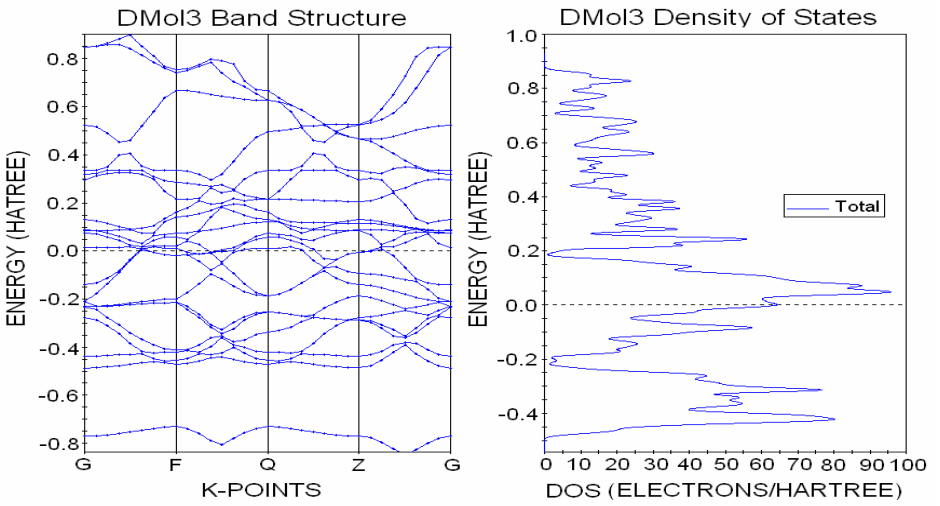


Figure 5. The band structure and the DOS of model 2

Model 3: V_2O_2. In this model a=c=4.7148Å and b=2.3574Å. The binding energy in this model is twice that of model 2. The model is shown in Figure 6 and the band structure along with the DOS is given in Figure 7. The gap energies are given in Table 1. The gap at the G point is in agreement with the experimental value [6]. The calculated value of 0.58 eV is to be compared with the experimental value of 0.6 eV.

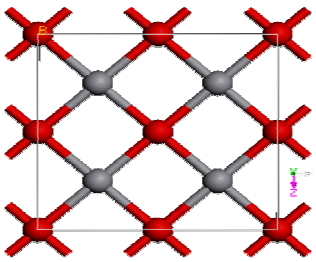


Figure 6. Model 3: Another model of V_2O_2

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

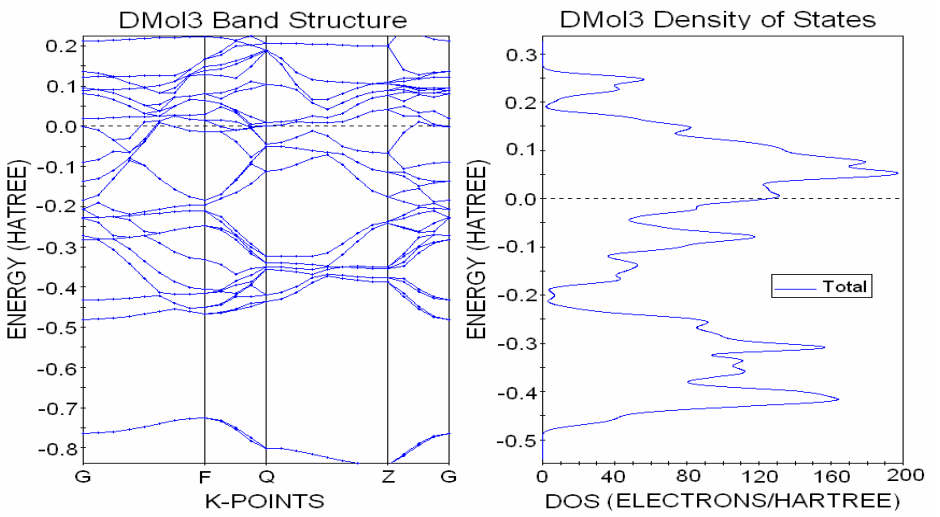


Figure 7. The band structure and the DOS of model 3

Model 4: VO_2. In this model shown in Figure 8, every vanadium atom is connected to 4 oxygen atoms but every oxygen atom is connected to only 2 vanadium atoms with the V-O distance of 1.895Å. The unit cell parameters are a=b=4.551Å and c=2.851Å. The calculated band structure and the DOS are given in Figure 9, and the gap energies are given in Table 1.

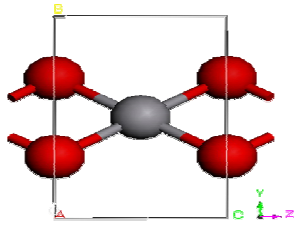


Figure 8. Model 4: One vanadium atom is connected to 4 oxygen atoms and one oxygen atom is connected to 2 vanadium atoms.

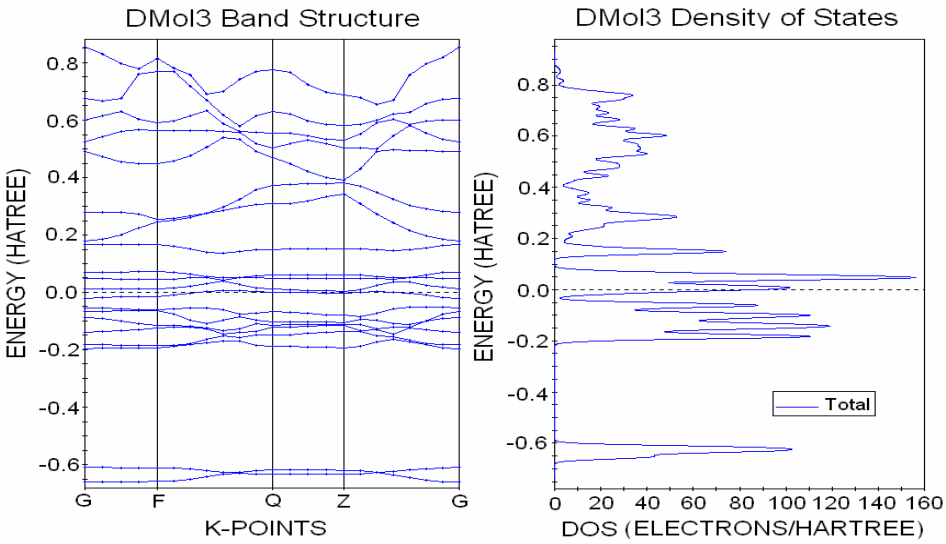


Figure 9. The band structure and the DOS of model 4

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

Model 5: VO_2. This model is shown in Figure 10. It has the unit cell of a=b=c=3.468Å. The V-O distance is 1.734Å and the band structure along with the DOS is shown in Figure 11. The calculated gap energies are given in Table 1.

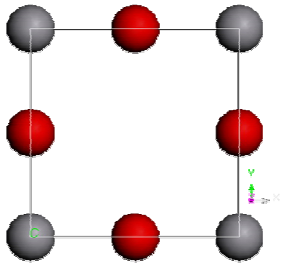


Figure 10. Model 5: The cubic model of vanadium oxide

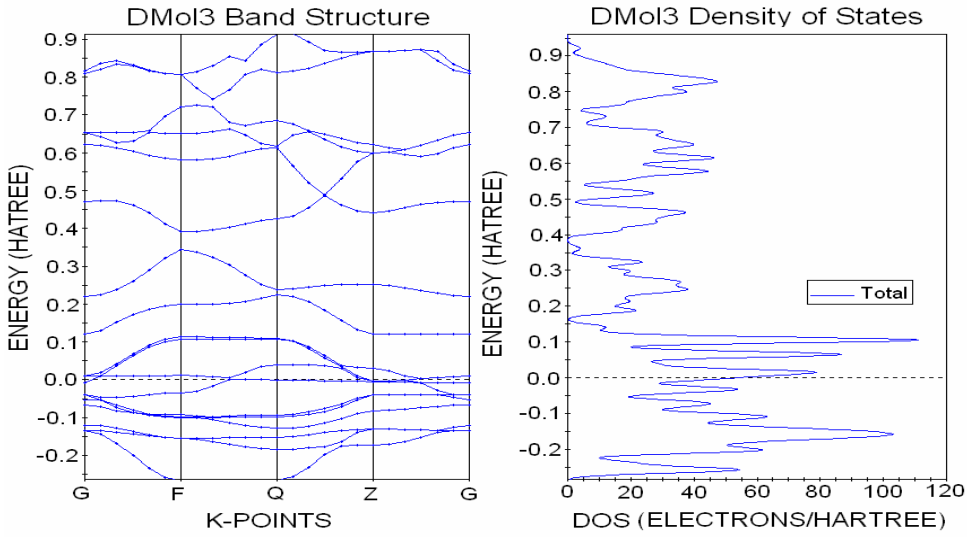


Figure 11. The band structure and the DOS of model 5

Model 6: VO_2. In this model the atoms are rearranged similar to those in model 5. The V-O distance is slightly elongated to 1.831Å and a=b=c=3.171Å. The model is shown in Figure 12 and the band structure along with the DOS is given in Figure 13. The calculated gap energies are given in Table 1.

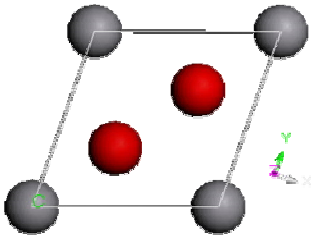


Figure 12. Model 6 of vanadium oxide (VO_2)

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

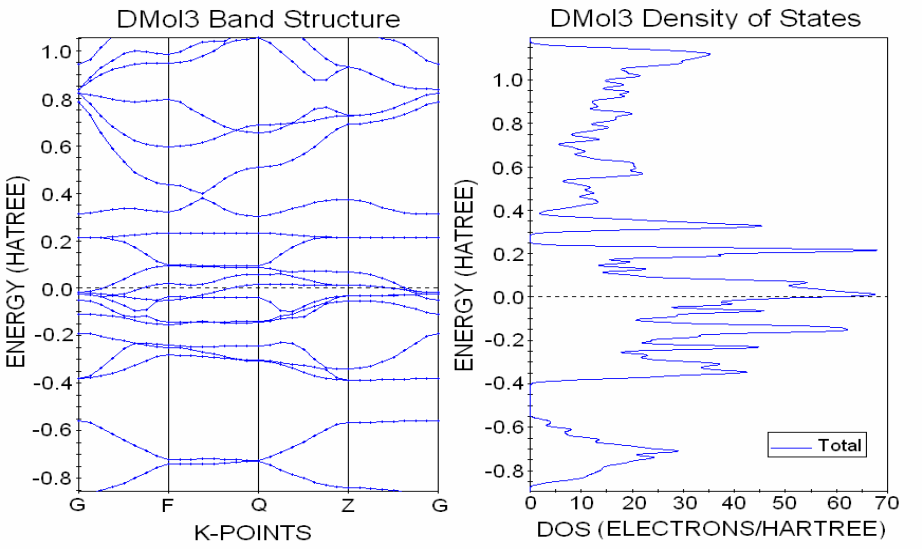


Figure 13. The band structure and the DOS of model 6

Crossing points

When we magnify all the band structures to search for crossing points, we find that all crossing points actually open a gap except in model 2 in Castep (pseudopotentials). When the program takes into account all the orbitals, these crossings also change into a small gap. These small gaps are shown in Table 2. It can be observed that Castep gives zero value for the gap, whereas DMol3 gives a small value, 2.72 x 10^{-4} eV, in model 2 in between G and F points. This is due to the interactions between electrons. Thus, the pseudopotential method does not see the small gaps at some places which can be seen when all of the electrons are considered properly.

Supersymmetry

In Figure 14, the crossing point obtained from the Castep calculation of model 2 is shown. In fact when all of the electrons are taken into account, this crossing point opens up a gap of about 2.72 x 10^{-4} eV. If we look at the cross carefully we find that it is not a cross of two lines but a confluence point of four lines.

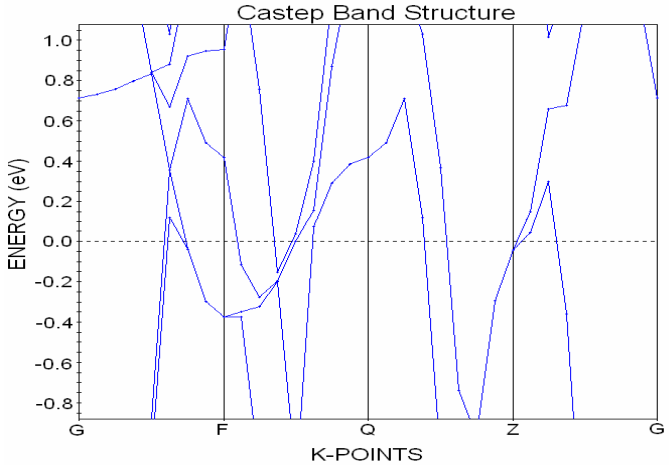


Figure 14. Crossing point in the space between G and F points at about 0.3 eV

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

Table 2. Small energy gaps between two k points or at a k point in eV

Model	**Castep (pseudopotentials)**	**DMol3 (all electrons)**
Model 1	7.3×10^{-3}(G-F)	17.2×10^{-2} (G-F)
Model 2	0.00 (G-F) 3.66×10^{-3} (F-Q)	2.72×10^{-4} (G-F) 18.12×10^{-3} (F-Q) 1.53×10^{-3} (F-Q)
Model 3	7.35×10^{-3} (G-F) 2.88×10^{-5} (Q) 6.25×10^{-5} (Q)	10.3×10^{-3} (G-F) 34.5×10^{-3} (F-Q)
Model 4	5.25×10^{-3}(F)	32.5×10^{-3} (F-Q)
Model 5	9.48×10^{-3} (G-F) 2.04×10^{-3} (Z-G)	13.03×10^{-3} (G-F) 11.05×10^{-3} (G-F)
Model 6	12.43×10^{-3} (Z) 4.32×10^{-3} (Z) 10.43×10^{-3} (Z-G) 10.23×10^{-3} (Z-G) 4.38×10^{-3} (Z-G)	32.27×10^{-3} (G-F) 3.62×10^{-4} (Z-G) 9.36×10^{-3} (Z-G) 11.43×10^{-3} (Z)

When we draw a horizontal line and look for the positive solutions above the line and compare them with those below the horizontal line, the supersymmetry is not found. The expanded version of the Castep band structure of model 3 shown in Figure 15 also shows gaps between G and F points which in DMol3 show gaps.

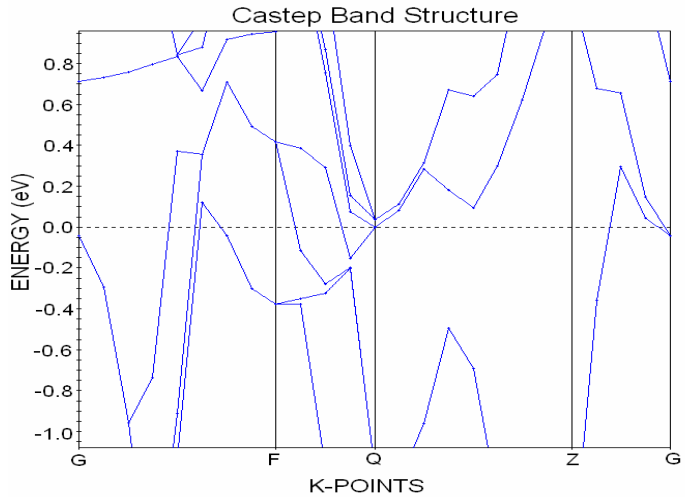


Figure 15. The expanded version of band structure of model 3

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

Conclusions

We have studied 6 different models of vanadium oxide and have calculated the gaps in all of the cases. In models 1-3, the valency of vanadium is two so that we see the alignment of V^{2+} ions, whereas in models 4-6, vanadium has a valency of 4, thus showing the ordering of V^{4+} ions. We have searched for the crossing points by using two valencies and six models but discovered "confluence points." The crossing may be called the "Dirac point." However, our calculation is non-relativistic and hence "Schrödinger point" is a more appropriate terminology for the crossing point.

Acknowledgements

We are grateful to the University of Malaya for supporting this work. The Fundamental Research Grants Scheme (FRGS) of the Ministry of Higher Education has kindly provided financial support. The Malaysian Academy of Sciences has supported the initial stages of the work through the Scientific Advancement Grants Allocation (SAGA). We are also grateful to the university authorities for the University of Malaya Research Grant (UMRG) which includes travel grants.

References

1. Y. Zhang, Y-W. Tan, H. L. Stormer and P. Kim, "Experimental observation of the quantum Hall effect and Berry's phase in graphene", *Nature*, **2005**, *438*, 201-204.
2. K. N. Shrivastava, "Infrared of thin film graphene in a magnetic field and the Hall effect", *Proc. SPIE*, **2008**, *7155*, 71552F1-71552F8.
3. N. A. Zabidi, H. A. Kassim and K. N. Shrivastava, "Energy band crossing points in multilayer of graphene", *AIP Conf. Proc.*, **2008**, *1017*, 326-330.
4. V. Pardo and W. E. Pickett, "Half-metallic semi-Dirac-point generated by quantum confinement", *Phys. Rev. Lett.*, **2009,** *102*, 166803_1-166803_4.
5. J. P. Pouget, H. Laurois, J. P. D'Haenens, P. Merenda and T. M. Rice, "Electrons localization induced by unaxial stress in pure VO_2", *Phys. Rev. Lett.*, **1975**, *35*, 873-875.
6. M. Marezio, D. B. McWhan, J. P. Remeika and P. D. Dernier, "Structural aspects of the metal-insulator transitions in Cr-doped VO_2", *Phys. Rev. B*, **1972,** *5*, 2541-2551.
7. S. Shin, S. Suga, M. Taniguchi, M. Fujisawa, H. Kanzaki, A. Fujimori, H. Daimon, Y. Veda, K. Kosuge and S. Kachi, "Vacuum ultraviolet reflectance and photoemission study of the metal-insulator phase transitions in VO_2, V_6O_{13} and V_2O_3", *Phys. Rev. B*, **1990**, *41*, 4993-5009.
8. R. M. Wentzcovitch, W. W. Schulz and P. B. Allen, "VO_2: Peierls or Mott-Hubbard? A view from the band theory", *Phys. Rev. Lett.*, **1994**, *72*, 3389-3392 .
9. D. Ruzmetov, S. D. Senanayake, V. Narayanmurti and S. Ramanathan, "Correlation between metal -insulator transition characteristis and electronic structure changes in vanadium oxide thin films", *Phys. Rev. B*, **2008**, *77*, 195442_1-195442_5.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 53-63

10. M. M. Qazilbash, A. A. Schafgans, K. S. Burch, S. J. Yun, B. G. Chae, B. J. Kim, H. T. Kim and D. N. Basov, "Electrodynamics of VO_2 and V_2O_3", *Phys. Rev. B*, **2008**, *77*, 115121_1-115121_10.
11. A. Sharoni, J. B. Ramirez and I. K. Schuller, "Multiple avalanches across the metal-insulator transitions of vanadium oxide nanoscaled junction", *Phys. Rev. Lett.*, **2008**, *101*, 026404_1-026404_4.
12. H. Nakatsugawa and E. Iguchi, "Electronic structures in VO_2 using the periodic polarizable point ion shell model and DV-Xα method", *Phys. Rev. B*, **1997**, *55*, 2157-2163.
13. A. Zylbersztejn and N. F. Mott, "Metal-insulator transition in vanadium dioxide", *Phys. Rev. B*, **1975**, *11*, 4383-4395.
14. W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlations effects", *Phys. Rev.*, **1965***, 140,* A1133-A1138.
15. P. Hohenberg and W. Kohn, "Inhomogeneous electtron gas", *Phys. Rev.*, **1964**, *136*, B864- B871.
16. M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, "Iterative minimization techniques for ab anitio total energy calculations: Molecular dynamics and conjugate gradients", *Rev. Mod. Phys.*, **1992**, *64*, 1045-1097.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Maejo International Journal of Science and Technology

ISSN 1905-7873

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

Full Paper

Urban vascular flora and ecological characteristics of Kadıköy district, Istanbul, Turkey

Etem Osma[1], İbrahim İ. Özyiğit[1,*], Volkan Altay[2] and Memduh Serin[1]

[1] Department of Biology, Science and Arts Faculty, Marmara University, 34722, Göztepe, Kadıköy, Istanbul, Turkey

[2] Department of Biology, Science and Arts Faculty, Mustafa Kemal University, 31040, Hatay, Turkey

[*]Corresponding author, e-mail: ilkozyigit@marmara.edu.tr ; phone: +90 216 346 45 33; fax: +90 0 216 347 87 83

Received: 28 October 2009 / Accepted: 11 January 2010 / Published: 10 March 2010

Abstract: Cities are usually considered unnatural places lacking ecological benefits. Many ecological studies have been done in suburban or rural areas and have excluded the city. However, urban ecology has become increasingly important. Today, it is well known that rapid and uncontrolled urbanisation accompanied by insufficient infrastructure has resulted in the degradation of many forests and grasslands in metropolitan areas. Land use changes due to urbanisation during the second half of the 20th century have caused widespread decline of biodiversity of many animal and plant taxa, especially in densely populated regions such as Kadıköy district, Istanbul, Turkey. In this study, different kinds of urban habitats within the boundaries of Kadıköy are described. Plant samples were collected, dried, labelled and identified according to standard herbarium procedures and all the greenery in the district was explored during vegetation seasons. A total of 561 vascular plant taxa were determined, wherein 412 (337 species, 44 subspecies and 31 varieties) were native and 149 (143 species and 6 varieties) were exotic and cultivated. The major native taxa were Asteraceae (46 species) while Rosaceae had the most exotics (22 species). Archaeophytes and neophytes, endemics, rare, endangered, medicinal, and poisonous species are also reported. In addition, the requirements of making ecological studies in other urban areas are mentioned.

Keywords: urban flora, urban habitats, urban ecology, Kadıköy, Istanbul

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Introduction

Over the last 50 years, the world has witnessed a dramatic growth in urban populations and today approximately 50% of the world's people live in urban areas. A recent United Nation's report notes that the present world urban population of 3 billion is expected to reach 5 billion by 2030. Although the rate of urbanisation is declining in many regions of the world, most cities, metropolitan regions and urban areas continue to grow. Monitoring these developments and creating sustainable urban environments remain as crucial issues on the international development agenda [1].

Cities are the most important naturally and/or anthropogenically influenced ecosystems in which millions of people live [2-3]. Natural disturbances include storms, floods, drought, and diseases. Anthropogenic disturbances include ploughing, mowing, burning, grazing and soil compaction by livestock [2]. In addition, the construction and use of tracks, roads, canals, railways and airports have involved many direct and indirect changes on ecosystems [2, 4]. Because of these anthropogenic activities, plant and animal species are estimated to go extinct worldwide at a rate of 0.5% per year, mainly due to habitat loss and fragmentation [5-6]. Land use changes due to urbanisation and agricultural intensification have caused widespread decline of biodiversity for many organisms such as butterflies, birds and plants in populated regions of Western Europe during the second half of the 20th century [6-9].

The improvement of the natural environment is a challenge for cities with adverse natural conditions. Flora and vegetation are very important components of urban ecosystems. The urban flora consists of plants that naturally grow there [3, 10]. This also reflects different historical phases of urban development and land use. Species occur which are directly linked to the economic and cultural life of cities. In addition, the value of the flora and vegetation for nature conservation, urban recreation, and nature enjoyment in the city is enhanced [11].

Istanbul, which is located in the north-west part of Turkey, is one of the biggest metropolitan areas of the world. It has an area of around 5,750 km^2, a population of 11,008,790 (2007), and is Turkey's cultural and financial centre [12-14]. The city provides around 40 % of the country's tax revenue and is home to around 38% of the country's industrial companies and around 55% of the country's commercial companies [12]. The city has experienced more pressure from high increase in population, mostly from eastern Turkey, and rapid increase in urban areas with consequent decrease in green areas [15]. In Istanbul, 99% of the population live in the city centre or in suburban zones [14]. This rapid, uncontrolled and illegal urbanisation has caused degradation of forest and open land in the city, especially during the last two decades [15].

In this study, the urban flora and ecological characteristics of Kadıköy district, which is an old settlement on the Anatolian side of Istanbul, is presented. In addition, native, exotic, and cultivated plants are mentioned. Endemic, rare, and endangered plant species and their habitats are listed in the appendix. The negative effects of unplanned urbanisation and industrialisation which were made without ecological consideration are also pointed out.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

General Information

Location: Kadıköy district is located on the Asian (Anatolia) side of Istanbul (N 40° 59' 08" , E 29° 01' 45") on the north coast of the Marmara Sea. The district covers 34 km^2 and has the typical physical characteristics of coastal cities in the region (Figure 1). Neighbouring districts are Üsküdar and Ümraniye to the north, Maltepe to the south-east, and Bostancı to the east. The coast extends through the Bosphorus Strait, which extends to Europe and Asia. Thus, Kadıköy has the most important ecological aspects compared to any other district in Istanbul [16].

Figure 1. Satellite view of Istanbul (upper left corner) and Kadıköy district. The study area is circled. (This picture was prepared by using The Google Earth Programme.)

History: The establishment date of Kadıköy (Chalcedon) is accepted as 675 BC and it was the first settlement which the Greeks from Megara established on the Bosphorus. Byzantium was later established on the other side of the strait in 667 BC (Figure 2). In its history, Chalcedon had changed hands many times as Persians, Bithynians, Romans, Byzantines, Arabs, Crusaders and Turks passed through the area, and was finally under the Ottomans in 1353. In 1453, Istanbul was conquered by Fatih Sultan Mehmet and Kadıköy was given to Hıdır Bey, who was the first Otoman judge (Kadı) for the Sultan of Istanbul. The name Kadıköy, which means "village of the judges", comes from Kadı Hıdır Bey. After that, Kadıköy became a popular market for many goods and in time, developed into a residential area for people commuting to the city [16].

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

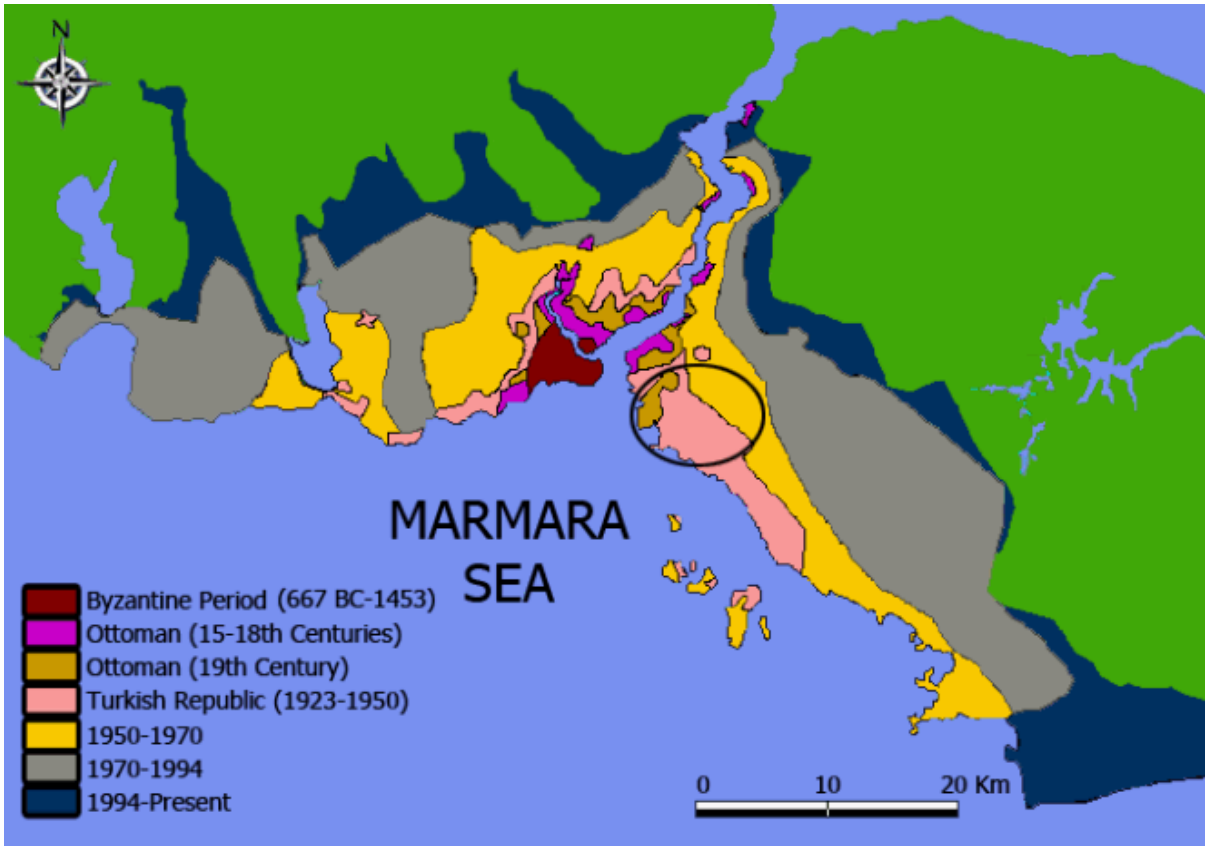


Figure 2. The historical growth of Istanbul (modified from Karakuyu [13])**.** The study area is circled.

Population: The population of Kadıköy has increased rapidly in the last 35 years. While the population was 57,901 in 1940, it reached 241,593 in 1970 and 744,670 in 2007 [14, 16].

Topography: The topography of Kadıköy is similar to those of many other coastal districts on the Anatolian side of Istanbul. It has an elevation range of 1-600 m above sea level. Kadıköy District, which has lower elevation on the shore of the Marmara Sea, is also the west end of the Kocaeli Plateau where it is bounded by the Bosphorus Strait [16-17]. There are also seven hills in Kadıköy, namely (from west to east) Kayışdağı (partially), Göztepe, Fikirtepe, Acıbadem, Altıyol, Cevizlik and Koşuyolu. The ground between these hills is slightly slanted. The coastal areas between Kadıköy and Bostancı are flat. Kurbağalıdere is the only stream, which has its source in Kayışdağı and flows into the Marmara Sea [16].

Climate: Kadıköy has a slightly cooler Mediterranean climate, which is associated with the climate of the Black Sea [17-18]. January and February are the coldest months while July and August are the warmest. Annual precipitation is about 690.7 mm, most occurring in winter. In the summer, low precipitation and high temperatures prevail with an annual mean temperature of 14.5^{o}C for the past two decades. Between May and September, the temperature is generally above 30^{o}C and between November and April it is rarely below 0^{o}C. In the growing period, the daily mean temperature is 8^{o}C with about 280 days (between 15 March-20 December). The relative humidity is about 75% and this value decreases to 65% in summer despite of the effect of the sea. Information about the climate of the study area was derived from the meteorological station at Göztepe-Kadıköy [18-19].

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Urban habitats: In the study area, vascular plants were found in parks, gardens, meadows, forests, refuges, civic squares, abandoned land and cemeteries.

Materials and Methods

Collection of the flora was conducted from March to July during 2002-2008. Each year the grasslands, rocky sites, shrub lands and forests were inventoried between March and July. The identification of plant specimens was made by using Davis [20]. The collected plant samples were deposited in MÜFE Herbarium (Faculty of Science and Arts Herbarium, Marmara University). The flora is listed in Appendix 1 (Native plant list) and Appendix 2 (Exotic and cultivated plant list). The family order is according to Davis. The genera and species are arranged alphabetically. Life forms [phanerophytes (Ph), chamaephytes (Ch), hemicryptophytes (H), therophytes (Th), geophytes (G), helophytes (He)] and phytogeographical origins [Euro-Siberian (Euro.-Sib), Irano-Turanian (Ir.-Tur.), Mediterranean (Medit.), East Mediterranean (E. Medit.)] are noted with the scientific names of collected species and they were determined according to Raunkier system [21].

Archaeophyte and neophyte plants are also emphasised. Exotic and cultivated plants were alphabetically listed by using Bailey [22] (Appendix 2). Endemic, rare, endangered, medicinal, and poisonous species are noted in the lists. The categories and criteria of rare and endangered species are according to Ekim et al.[23] and "Red Data List" of International Union for the Conservation of Nature and Natural Resources (IUCN) [24]. Poisonous plants were noted according to Seçmen and Leblebici [25] and Baytop [26] while medicinal plants are according to Baytop [27] and Baytop and Kadıoğlu [28].

Results and Discussion

A total of 561 vascular plant taxa were found with 412 native (337 species, 44 subspecies and 31 varieties) belonging to 68 families and 239 genera, and 149 exotic and cultivated (143 species) belonging to 61 families and 112 genera (Appendices 1 and 2). While 268 native species (70.20%) were dicots, 62 native species (18.40 %) were monocots. Only one native species (0.30%) was a pteridophyte and 6 native species (1.78 %) were gymnosperms. There were 173 perennial, 152 annual and 12 biennial species in the native flora.

The following families have the most native species: Asteraceae (46 species, 13.65%), Poaceae (39 species, 11.57%), Fabaceae (33 species, 9.79%), Lamiaceae (16 species, 4.75%) and Liliaceae (14 species, 4.15%). Asteraceae, as it is known, has the most species in the whole flora of Turkey [20]. In fact, it is also a very large cosmopolitan family represented by 13 tribes, 84 genera and over 240 species distributed all over the world [2-4, 11, 29-33]. The most common genera were *Trifolium* (11 species, 3.26%), *Ranunculus* (5 species, 1.48%), *Quercus* (5 species, 1.48 %), *Rumex*, *Geranium*, *Medicago*, *Plantago* and *Veronica* (4 species, 1.19%).

For native plants, the largest groups of life forms were therophytes (44.81%) and hemicryptophytes (30.56%). The percentages of other life forms were as follows: phanerophytes (13.65%), geophytes (8.31%), chamaephytes (1.48%) and helophytes (1.19%). The significant therophytic and hemicryptophytic components of the flora correlate with the submediterranean climate

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

of Kadıköy district and demonstrate an intermediate position between Mediterranean and temperate ecosystems [33-34].

The most common phytogeographical elements found were Medit. El. (44 species, 13.06%), Euro-Sib. El. (31 species 9.20%) and E. Medit. El. (7 species, 2.08%) for native plants. This is because the research area is mostly affected by a Mediterranean climate. In addition, the northern side of Istanbul is partly affected by the oceanic climate and this situation can result in plants belonging to Euro-Sib. El. in Kadıköy district [17, 34].

In the research area, 8 taxa (2.37%) were cosmopolitan and 59 taxa (18.00%) were widespread. The most common native plants are *Sonchus asper* (L.) Hill ssp. *glaucescens* (Jordan) Ball (Asteraceae), *Rapistrum rugosum* (L.) All. (Brassicaceae), *Chenopodium album* L. (Chenopodiaceae), *Lolium perenne* L. (Poaceae), *Cynodon dactylon* (L.) Pers. var. *dactylon* (Poaceae), *Stellaria media* (L.) Vill. subsp. *media* (Caryophyllaceae), *Lamium purpureum* L. var. *purpureum* (Lamiaceae) and *Parietaria judaica* L. (Urticaceae)-especially on walls.

Because Kadıköy district has mostly completed its urbanisation process, perennial habitats have decreased. Present habitats are represented by annual taxa, e.g. *Stellaria media* (L.) Vill. ssp. *media* (Caryophyllaceae) and *Lamium purpureum* L. var. *purpureum* (Lamiaceae), *Euphorbia helioscopia* (Euphorbiaceae), *Capsella bursapastoris* L. (Brassicaceae), *Veronica persica* Poir. (Scrophulariaceae).

Archaeophyte (28 species) and neophyte (15 species) species were found (Table 1) [34-36]. In addition, 2 endemic (0.48 %) and 2 rare taxa (0.71 %) were collected. The endemics were *Ballota nigra* L. subsp. *anatolica* P. H. Davis (0.24 %) and *Cirsium polycephalum* DC. (0.24 %). The rare species were *Cymbalaria muralis* Gaertner and *Albizzia julibrissin* (Willd.) Durazz. [23]. We also determined the poisonous (83 species) and medicinal (119 species) plants for both native and exotic plants.

Kadıköy still has a rich flora even with millennia of urbanisation. It reflects a Mediterranean ecosystem and shows increased floristic diversity with temperature, water availability, and topography. These relate to the richness of the area along with the human influence which has created a variety of habitat types through traditional land use practices [17-18, 34].

It is obvious that rapid urbanisation has created problems for the survival of some species [6, 37]. In urbanised areas, therophytes increased while rare species and wasteland plants decreased from the suburbs to the centre [11]. *Lemna minor* L. (Lemnaceae), *Iris sintenisii* Janka (Iridaceae), *Cephalanthera longifolia* (L.) Fritsch., *Neotinea masculata* (Desf.) Steam, *Orchis papilionacea* L., *Orchis laxiflora* Lam., *Orchis tridentate* Scop. subsp. *lactea* (Poir.) Rouy (Orchidaceae) were present in the 1960s. Unfortunately, they have now been extirpated [38]. In addition, populations of *Phragmites australis* (Cav.) Trin. *ex* Steudel (Poaceae), *Typha latifolia* L. (Typhaceae), *Juncus heldreichianus* Marsson *ex* Parl. and *J. conglomeratus* L. (Juncaceae) have been greatly reduced. Additionally, an increase of non-native plants, ruderals, cultivated species and annuals has occurred.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Table 1. List of archaeophyte and neophyte plants

Archaeophytes (before 1500 AD)	Neophytes (after1500 AD)
Lamium purpureum L. (Lamiaceae)	*Conyza canadensis* (L.) Cronquist (Asteraceae)
Lamium amplexicaule L.	*Datura stromanium* L. (Solanaceae)
Ballota nigra L.	*Veronica persica* Poiret (Scrophulariaceae)
Euphorbia helioscopia L. (Euphorbiaceae)	*Cymbalaria muralis* Gaertner
Euphorbia peplus L.	*Ailanthus altissima* (Miller) Swingle (Simaroubaceae)
Sinapis arvensis L. (Brassicaceae)	*Aesculus hippocastanum* L. (Hippocastaneae)
Capsella bursa pastoris (L.) Medik.	*Oxalis corniculata* L. (Oxalidaceae)
Sisymbrium officinale (L.) Shop.	*Diplotaxis tenuifolia* (L.) DC. (Brassicaceae)
Fumaria officinalis L. (Papaveraceae)	*Cardaria draba* (L.) Desv.
Papaver rhoeas L.	*Sisymbrium altissimum*L.
Stellaria media (L.) Vill. (Caryophyllaceae)	*Raphanus raphanistrum* L.
Cerastium glomeratum Thuill	*Robinia pseudoacacia* L. (Fabaceae)
Cichorium intybus L. (Asteraceae)	*Amaranthus retroflexus* L. (Amaranthaceae)
Bellis perennis L.	*Syringia vulgaris* L. (Oleaceae)
Solanum nigrum L. (Solanaceae)	*Lolium multiflorum* Lam. (Poaceae)
Chenopodium album L. (Chenopodiaceae)	
Setaria viridis (L.) P. Beauv. (Poaceae)	
Bromus sterilis L.	
Echinochloa crus - galli (L.) P. Beauv.	
Plantago lanceolata L. (Plantaginaceae)	
Melilotus alba Desr. (Fabaceae)	
Melilotus officinalis (L.) Desr.	
Geranium pusillum Burm. fil. (Geraniaceae)	
Geranium dissectum L.	
Geranium molle L.	
Echium vulgare L. (Boraginaceae)	
Anagallis arvensis L. (Primulaceae)	
Malva neglecta Wallr. (Malvaceae)	

The changes in environmental conditions in human-altered sites provide specific niches, which are often colonised by aliens rather than by native species. Higher temperatures and limited soil moisture are characteristic factors in rural to urban gradients. Many alien species in temperate zones originating from warmer areas easily adapt to disturbed urban conditions [39]. These species, which are imported to an area due to anthropogenic activities, sometimes cover the area more frequently than native plants [40-41]. Some of those invasive alien species are: *Hyacinthus orientalis* L. and *Tulipa species* L. (Liliaceae), *Acer negundo* L. (Aceraceae), *Robinia pseudoacacia* L. (Fabaceae), *Ailanthus altissima* (Miller) Swingle (Simaroubaceae), *Celtis australis* L. (Ulmaceae), *Platanus orientalis* L. (Platanaceae), and *Viola x wittrockiana* Gams. (Violaceae). *Robinia pseudoacacia* L. and *Acer negundo* L. are the most successful non-native tree species which also appear spontaneously in different habitats in Berlin, Germany [34, 40-41].

The effects of urbanisation are more intense in the inner city and this can support specialised urban plant communities. This study has demonstrated the decrease in indigenous species, the immigration of alien species, and the establishment of new ecotypes in an old urban area.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Acknowledgements

We express our gratitude to Professor Dr. Munir Ozturk from Ege University (Biology Department) for his valuable suggestions in the completion of this study and Ms Sedef Çakır for rechecking the English.

References

1. http://www.unis.unvienna.org (The website of United Nations Information Service), last access: September **2009**.
2. K. H. Shaltout and M. A. El-Sheikh, "Vegetation of the urban habitats in the Nile Delta region, Egypt", *Urban Ecosyst.*, **2002**, *6*, 205-221.
3. G. S. Antipina, **"**Urban flora as a component of the urban ecosystems in the Taiga Zone: An example of Karelian Cities", *Russian J. Ecol.*, **2003**, *34*, 215-218.
4. D. Lausi and P. L. Nimis, "Roadside vegetation in boreal South Yukon and adjacent Alaska", *Phytocoenologia*, **1985**, *13*, 103-138.
5. P. M. Vitousek, H. A Mooney, J. Lubchenco and J. M. Melillo, "Human domination of earth's ecosystems", *Science*, **1997**, *277***,** 494-499.
6. S. Van der Veken, K. Verheyen and M. Hermy, "Plant species loss in an urban area (Turnhout, Belgium) from 1880 to 1999 and its environmental determinants", *Flora*, **2004**, *199*, 516-523.
7. D. Maes and H. Van Dyck, "Butterfly diversity loss in Flanders (north Belgium): Europe's worst case scenario?", *Biol. Conserv.*, **2001**, *99***,** 263-276.
8. P. F. Donald, R. E. Green and M. F. Heath, "Agricultural intensification and the collapse of Europe's farmland bird populations", *Proc. R. Soc. London Ser.B*, **2001**, *268*, 25-29.
9. K. Thompson and A. Jones, "Human population density and prediction of local plant extinction in Britain", *Conserv. Biol.*, **1999**, *13*, 185-189.
10. N. G. Il'minskikh, "Ecotopological structure of the urban flora", Proceedings of Third Workshop on Comparative Floristics, **1998**, Kunkur, Russia, pp. 269-276.
11. U. Maurer, T. Peschel and S. Schmitz "The flora of selected urban land-use types in Berlin and Potsdam with regard to nature conservation in cities", *Landscape Urban Plan.*, **2000**, *46,* 209-215.
12. S. Kaya and P. J. Curran "Monitoring urban growth on the European side of the Istanbul metropolitan area: A case study", *Int. J. Appl. Earth Observ. Geoinform.*, **2006**, *8*, 18-25.
13. M. Karakuyu, "Istanbul'un mekansal gelişiminin analizi (The urban development analysis of Istanbul)", 4th GIS Days in Türkiye, September 13-16, **2006**, Fatih University, Istanbul,Turkey, pp.1-8.
14. http://tuikapp.gov.tr (The official website of Turkish Republic Office of Prime Ministry Statistics Institution), last access: September **2009**.
15. A. Geymen and I. Baz, "Monitoring urban growth and detecting land-cover changes on the Istanbul metropolitan area", *Environ. Monit. Assess.*, **2008**, *136*, 449-459.
16. http://www.Kadıköy-bld.gov.tr (The official website of Kadıköy district municipality), last access: September **2009**.

17. E. Osma, "Kadıköy İlçesi (Istanbul) Kentsel Ekolojisi (Urban ecology of Kadıköy District-Istanbul)", *MSc. Thesis*, **2003**, Marmara University, Turkey.
18. V. Altay, "Kartal (Istanbul)'nin Kentsel Ekolojisi (Urban ecology of Kartal- Istanbul)", *MSc. Thesis*, **2004**, Marmara University, Turkey.
19. http://www.dmi.gov.tr (The official website of Turkish State Meteorological Service), last access: September **2009**.
20. P. H. Davis, "Flora of Turkey and the East Aegean Islands", Vols: 1-11, Edinburgh University Press, Edinburgh, **1965-2001**.
21. C. Raunkier, "The Life Forms of Plants and Statistical Plant Geography", Clarendon Press, Oxrford, **1934**.
22. L. H. Bailey, "Manual of Cultivated Plants", McMillan, New York, **1949**.
23. T. Ekim, M. Koyuncu, M. Vural, H. Duman, Z. Aytaç and N. Adıgüzel, "Türkiye Bitkileri Kırmızı Kitabı (Red Data Book of Turkish Plants) Pteridophyta and Spermatophyta", Türkiye Tabiatını Koruma Derneği Yayınları**,** Ankara**, 2000**.
24. http://www.iucn.org (The website of International Union for Conservation of Nature), last access: September **2009**.
25. Ö. Seçmen and E. Leblebici, "Yurdumuzun Zehirli Bitkileri (The poisonous plants of our country)", Ege Universitesi Fen Fakültesi Kitaplar Serisi, No:103, Bornova, Izmir, **1987**.
26. T. Baytop, "Türkiye'de Zehirli Bitkiler, Bitki Zehirlenmeleri ve Tedavi Yöntemleri (Poisonous plants, plant poisonings and treatment methods in Turkey)", Istanbul Universitesi Yayınları, No: 3560, Istanbul, **1989**.
27. T. Baytop, "Türkiye'de Bitkiler ile Tedavi (Treatment with plants in Turkey)", Istanbul Universitesi Eczacılık Fakültesi Yayınları, No: 40, Istanbul, **1984**.
28. T. Baytop and D. Kadıoğlu, "Istanbul Bölgesi'nin Tıbbi Bitkileri (Medicinal plants of Istanbul Province)", 14. Bitkisel İlaç Hammaddeleri Toplantısı, Bildiriler, 29-31 May **2002**, Eskişehir, Turkey.
29. O.Adedeji and O.A. Jewoola, "Importance of leaf epidermal characters in the Asteraceae family", *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, **2008**, *36*, 7-16.
30. C. Yarci, M. Serin and V. Altay, "The segetal vegetation of Kocaeli Province (Turkey)", *Ekoloji*, **2007**, *63*, 23-33.
31. J. F. Maxwell, "Vascular flora of the Emerald Pool area, Krabi province, southern Thailand", *Maejo Int. J. Sci. Technol.*, **2009**, *3*, 1-25.
32. J. F. Maxwell, "Vegetation and vascular flora of the Mekong River, Kratie and Steung Treng Provinces, Cambodia", *Maejo Int. J. Sci. Technol.*, **2009**, *3*, 143-211.
33. D. Amanatidou, "Analysis and evaluation of a traditional cultural landscape as a basis for its conservation management. A case study in Vikos-Aoos National Park-Greece", *PhD. Thesis*, **2005**, Albert-Ludwigs University, Germany.
34. V. Altay, İ. İ. Özyiğit and C. Yarci, "Urban flora and ecological characteristics of the Kartal District (Istanbul): A contribution to urban ecology in Turkey", *Sci. Res. Essay*, **2010**, *5*, 183-200.
35. H. Schepker and I. Kowarik, "Bekämpfung von Neophyten in Niedersachsen.: Ursachen. Umfang, Erfolg", *Neobiota*, **2002**, *1*, 343-354.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

36. I. Kowarik, "Human agency in biological invasions: Secondary releases foster naturalisation and population expansion of alien plant", *Biol. Invasions*, **2003**, *5*, 293-312.
37. Z. Chocholousková and P. Pysek, "Changes in composition and structure of urban flora over 120 years: A case study of the city of Plzen", *Flora*, **2003**, *198*, 366-376.
38. A. Baytop, "Aydos ve Kayışdağ havalisinin çiçekli bitkileri (The flowering plants of Aydos and Kayışdağ areas)", *Türk Biyolog Dergisi*, **1962**, *12*, 75-111.
39. R Wittig, "The origin and development of the urban flora of Central Europe", *Urban Ecosystems*, **2004**, *7*, 323-339.
40. S. Zerbe, U. Maurer, S. Schmitz and H. Sukopp, "Biodiversity in Berlin and its potential for nature conservation", *Landscape and Urban Planning*, **2003**, *62*, 139-148.
41. I. Kowarik, "Einführung und Ausbreitung nichteinheimischer Gehölzarten in Berlin und Brandenburg und ihre Folgen für Flora und Vegetation. Verhandlungen des Botanischen Vereins von Berlin und Brandenburg", *Beiheft*, **1992**, *3*, 1-188.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Appendix 1: Native plants of Kadıköy district

Abbreviations and symbols: * poisonous plant, ** medicinal plant, LR: lower risk, CR: critically endangered, VU: vulnerable, DD: data deficient. (The LR, CR, VU and DD values are according to IUCN standards) [23- 28].

PTERIDOPHYTA

HYPOLEPIDACEAE
Pteridium aquilinum (L.) Kuhn (*), G

SPERMATOPHYTA

GYMNOSPERMAE

PINACEAE
Cedrus libani A Rich. Ph
Pinus brutia Ten. Ph, E. Medit. El.
P. pinea L. (**), Ph
P. sylvestris L. Ph, Euro-Sib. El.

CUPRESSACEAE
Cupressus sempervirens L. (**), Ph
Juniperus oxycedrus L. subsp. *oxycedrus* (**), Ph, widespread

ANGIOSPERMAE

DICOTYLEDONEAE

RANUNCULACEAE
Anemone coronaria L. (*), G, Medit. El.
Consolida ambiqua (L.) P. Bass. & Heyw. Th
Ranunculus constantinopolitanus (DC.) d'Urv. (*), H, widespread
R. ficaria L. subsp. *calthifolius* (Reichb.) Arc. (*), (**), G
R. marginatus d'Urv. var. *trachycarpus* (Fisch. & Mey.) Azn. (*), Th
R. marginatus d'Urv. var. *marginatus* (*), Th
R. muricatus L. (*), Th

PAPAVERACEAE
Fumaria officinalis L. (*), (**), Th
Glaucium flavum Crantz H, widespread
Papaver dubium L. Th
P. rhoeas L. (**), Th, widespread

BRASSICACEAE
Calepina irregularis (Asso) Thellung Th
Capsella bursa-pastoris (L.) Medik. Th, cosmopolitan
Cardaria draba (L.) Desv. subsp. *draba* H, widespread
Diplotaxis tenuifolia (L.) DC. (*), H.
Erophila verna (L.) Chevall. Th
Lepidium graminifolium L. H
Neslia apiculata Fisch. Th, widespread
Raphanus raphanistrum L. (**), Th

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Rapistrum rugosum (L.) All. Th
Sinapis arvensis L. Th, widespread
S. alba L. Th
Sisymbrium altissimum L. Th, widespread
S. officinale (L.) Shop. (*), Th, widespread
Thlaspi perfoliatum L. Th, widespread

RESEDACEAE
Reseda lutea L. var. *lutea* Th, widespread

CISTACEAE
Cistus creticus L. (**), Ph, Medit. El.
C. salviifolius L. Ph

VIOLACEAE
Viola sieheana Becker Th

POLYGALACEAE
Polygala vulgaris L. (**), H, Euro-Sib. El.

PORTULACACEAE
Portulaca oleracea L. Th

ILLECEBRACEAE
Scleranthus perennis L. H

CARYOPHYLLACEAE
Cerastium glomeratum Thuill Th, cosmopolitan
Dianthus leptopetalus Willd. H
Moenchia mantica (L.) Bartl. subsp. *mantica* Th
Petrorhagia prolifera (L.) Ball. & Heywood Th
Sagina maritima Don Th
Silene vulgaris (Moench) Garcke var. *vulgaris* (**), H
S. nocturna L. Th, Medit. El.
S. dichotoma Ehrh. Th
Spergularia bocconii (Scheele) Aschers. & Graebn. Th, Medit. El.
Stellaria media (L.)Vill. subsp. *media* (**), Th
S. media (L.) Vill. subsp. *pallida* (Dumort.) Aschers. & Graebn. (**), Th
Telephium imperati L. subsp. *orientale* (Boiss.) Nyman H

POLYGONACEAE
Polygonum arenastrum Bor. Th
P. aviculare L. Th, cosmopolitan
P. lapathifolium L. Th
Rumex acetosella L. (**), H, cosmopolitan
R. conglomeratus Murray H
R. crispus L. (**), H
R. pulcher L. H

CHENOPODIACEAE
Atriplex hastate L. Th
Chenopodium album L. subsp. *album* var. *album* Th
Salsola ruthenica L. Th

AMARANTHACEAE
Amaranthus blitoides S. Wats. Th
A. deflexus L. H
A. retroflexus L. Th

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

HYPERICACEAE
Hypericum calycinum L. (*), Ch
H. cerastoides (Spach) Robson (*), Ch
H. perforatum L. (*), (**), H

MALVACEAE
Alcea pallida Waldst. & Kit. H
Lavatera punctata All. Th
Malva neglecta Wallr. Th
M. nicaeensis All. Th
M. sylvestris L. (**), H

TILIACEAE
Tilia argentea Desf. ex DC. (**), Ph

LINACEAE
Linum bienne Miller (*), Th, Medit. El.
L. trigynum L. (*), Th, Medit. El.

GERANIACEAE
Erodium cicutarium (L.) L'Hérit subsp. *cicutarium* Th
E. malacoides (L.) L'Herit. Th, Medit. El.
Geranium dissectum L. Th
G. molle L. subsp. *molle* Th
G. purpureum Vill. Th
G. pusillum Burm. fil. Th

OXALIDACEAE
Oxalis corniculata L. Th, cosmopolitan

ZYGOPHYLLACEAE
Tribulus terrestris L. (*), (**), Th

SIMAROUBACEAE
Ailanthus altissima (Miller) Swingle (*), H

RHAMNACEAE
Paliurus spina-christi Miller (**), Ph

ANACARDIACEAE
Pistacia terebinthus L. subsp. *terebinthus* (**), Ph

FABACEAE
Cercis siliquastrum L. var. *siliquastrum* Ph
Gleditsia triacanthos L. Ph
Dorycnium pentaphyllum Scop. subsp. *herbaceum* (Vill.) Rouy H
Hippocrepis unisiliquosa L. Th
Lathyrus digitatus (Bieb.) Fiori H, E. Medit. El.
Lotus corniculatus L. var. *corniculatus* (*), H, widespread
Medicago lupulina L. Th, Widespread
M. orbicularis (L.) Bart. Th
M. polymorpha L. var. *vulgaris* (Benth.) Shinners Th, widespread
M. sativa L. H
Melilotus alba Desr. (*), Th, Widespread
M. officinalis (L.) Desr. (*),Th, Widespread
Onobrychis caput–galli (L.) Lam. Th, Widespread, Medit. El.
O. oxydonta Boiss. H, Widespread
Ononis spinosa L. subsp. *leiosperma* (Boiss.) Sirj. H, widespread
Robinia pseudacacia L. (*), (**), Ph

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Psoralea bituminosa L. H, Medit. El.
Scorpiurus muricatus L. Th
Spartium junceum L. (*), (**), Ph, Medit. El.
Trifolium angustifolium L. var. *angustifolium* Th
T. campestre Schreb. Th, widespread
T. constantinopolitanum Ser. Th, widespread
T. nigrescens Viv. subsp. *petrisavii* (Clem.) Holmboe Th, widespread
T. pratense L. H
T. repens L. var. *repens* H
T. resupinatum L. var. *resupinatum* Th
T. scabrum L. Th, widespread
T. stellatum L. var. *stellatum* Th
T. subterraneum L. Th
T. tomentosum L. Th
Vicia cracca L. subsp. *cracca* H, Euro-Sib. El.
V. hybrida L. Th, Widespread
V. sativa L. subsp. *nigra* (L.) Ehrh. var. *segetalis* (Thuill) Ser Th

ROSACEAE
Crataegus monogyna Jacq. subsp. *monogyna* (**), Ph
Mespilus germanica L. (**), Ph
Potentilla reptans L. H, widespread
Rosa canina L. (**), Ph
Rubus canescens DC. var. *canescens* Ph, widespread
Sanguisorba minor Scop. H
Sarcopoterium spinosum (L.) Spach (**), Ph, E. Medit. El.

CUCURBITACEAE
Ecballium elaterium (L.) A. Rich. (*), (**), H, Medit. El.

CRASSULACEAE
Sedum hispanicum L. Ch

APIACEAE
Ammi visnaga (L.) Lam. (**), Th
Berula erecta (Huds.) Coville H
Conium maculatum L. (*), Th
Daucus guttatus Sm. Th
Eryngium campestre L. var. *virens* Link (**), H, widespread
Ferulago confusa Velen. H, Euro-Sib. El.
Foeniculum vulgare Miller (**), H
Scandix pecten-veneris L. Th, widespread
Tordylium apulum L. Th, Medit. El.
Torilis nodosa (L.) Gaertner Th

ARALIACEAE
Hedera helix L. (*), (**), Ph

CORNACEAE
Cornus mas L. (**), Ph

CAPRIFOLIACEAE
Sambucus ebulus L. (**), H, Euro-Sib. El.

VALERIANACEAE
Centranthus ruber (L.) DC. (**), G

DIPSACACEAE
Knautia orientalis L. Th, E. Medit. El.
Scabiosa columbaria L. subsp. *columbaria* var. *columbaria* H

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

ASTERACEAE
Anthemis cf. c*hia* L. Th
A. cretica L. subsp. *tenuiloba* H
A. tinctoria L. var. *tinctoria* H, widespread
Arctium minus (Hill.) Bernh. subsp. *minus* H
Artemisia absinthium L. (*), (**), H
Bellis perennis L. H, Euro-Sib. El.
Calendula arvensis L. Th
Carduus nutans L. H
Carlina corymbosa L. H, Medit. El.
Carthamus lanatus L. (**), Th, widespread
Centaurea diffusa Lam. Th, widespread, Medit. El.
C. iberica Trev. ex Sprengel Th, widespread
C. solstitialis L. subsp. *solstitialis* Th, widespread
Chondrilla juncea L. var. *juncea* (**), H, widespread
Chrysenthemum segetum L. Th
Cichorium intybus L. (**), H, widespread
Cirsium creticum (Lam.) d'Urv. subsp. *creticum* H, E. Medit. El.
C. polycephalum DC. Endemic (CR), H
Conyza canadensis (L.) Cronguist Th
Crepis foetida L. Th
C. sancta (L.) Babcock Th, widespread
C. zacintha (L.) Babcock Th, Medit. El.
Echinops microcephalus Sm. H, Medit. El.
Erigeron acer L. H
Filago vulgaris Lam. Th
Helminthotheca echioides (L.) Holub Th
Hypochoeris glabra L. Th
Inula oculus–christi L. (*), G
I. viscosa (L.) Aiton (*), (**), H, Medit. El.
Lactuca saligna L. (*), Th
Lapsana communis L. Th
Matricaria chamomilla L. (**), Th
Pallenis spinosa (L.) Cass. Th
Picnomon acarna (L.) Cass. Th, widespread, Medit. El.
Picris hieracioides L. Th, Euro-Sib. El.
Scariola viminea (L.) F. W. Schmiat H, widespread
Scolymus hispanicus L. (**), H, Medit. El.
Sonchus asper (L.) Hill subsp. *glaucescens* (Jordan) Ball Th, widespread
Senecio vulgaris L. (*), (**), Th
Silybum marianum (L.) Gaertner (**), H, Medit. El.
Taraxacum officinale Weber (**), Ch
Tragopogon longirostris Bisch. *ex* Schultz H
Tussilago farfara L. (**), G, widespread, Euro-Sib. El.
Urospermum picroides (L.) F.W. Schmidt Th, Medit. El.
Xanthium spinosum L. (**), Th
X. strumarium L. subsp. *cavanillesii* (Scouw) D. Löve & P. Dansereau (**), Th

ERICACEAE
Arbutus unedo L. (**), Ph
Erica arborea L. Ph
E. manupuliflora Salisb. Ph

PRIMULACEAE
Anagallis arvensis L. var. *arvensis* (*), Th
A. arvensis L. var. *caerulea* (L.) Gouan (*), Th
Primula vulgaris Huds. subsp. *sibthorpi* (Hoffmanns.) W. W. Sm. & Forrest (*), H, Euro-Sib. El.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

OLEACEAE
Jasminum fruticans L. (**), Ph, Medit.El.
Ligustrum vulgare L. (*), Ph, Euro-Sib. El.
Olea europaea L. var. *europaea* (**), Ph
Phillyrea latifolia L. Ph, Medit. El.

APOCYNACEAE
Nerium oleander L. (*), (**), Ph, Medit. El.

ASCLEPIADACEAE
Cionura erecta (L.) Griseb. (*), Th, widespread, E. Medit. El.

CONVOLVULACEAE
Calystegia sepium (L.) R. Br. G
Convolvulus arvensis L. (**), H
C. cantabrica L. H

BORAGINACEAE
Anchusa azurea Miller H
Borago officinalis L. (**), Th
Echium italicum L. (*), H
E. plantagineum L. (*), Th
E. vulgare L. (*), H
Heliotropium europaeum L. (*), Th
Myosotis ramosissima Rochel ex Schultes Th
Trachystemon orientalis (L.) G. Don (**), G, Euxine El.

SOLANACEAE
Datura stramonium L. (*), (**), Th, cosmopolitan
Solanum dulcamara L. (*), (**), H, widespread, Euro-Sib. El.
S. nigrum L. subsp. *nigrum* (*), (**), Th, cosmopolitan

GENTIANACEAE
Blackstonia perfoliata (L.) Hudson Th
Centaurium erythraea Rafn. H

SCROPHULARIACEAE
Antirrhinum majus L. subsp. *majus* H
Bellardia trixago (L.) All. Th
Cymbalaria muralis Gaertner Rare (VU), Th
Kickxia spuria (L.) Dumort subsp. *integrifolia* Th
Linaria genistifolia (L.) Miller H
Parentucellia latifolia (L.) Caruel subsp. *latifolia* Th
Verbascum sp. H
Veronica chamaedrys L. H, Euro-Sib. El.
V. cymbalaria Bodard Th, Medit. El.
V. hederifolia L. Th, widespread
V. persica Poiret Th

OROBANCHACEAE
Orobanche minor L. (*), G

LAMIACEAE
Ballota nigra L. subsp. *anatolica* P. H. Davis Endemic (LR), H, Ir.-Tur. El.
Calamintha nepeta (L.) Savi subsp. *glandulosa* (Req.) P.W. H
Lamium amplexicaule L. Th, widespread, Euro-Sib. El.
L. purpureum L. Th, Euro-Sib. El.
Lavandula stoechas L. subsp. *stoechas* (**), Ph, Medit. El.
Melissa officinalis L. (**), H

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Mentha longifolia (L.) Hudson subsp. *typhoides* (Brig.) Harley var. *typhoides* H, widespread
M. pulegium L. (**), H
Origanum vulgare L. (**), H
Prunella vulgaris L. H, widespread, Euro-Sib. El.
Rosmarinus officinalis L. Ph, Medit. El.
Salvia verbenaca L. Ch
Scutellaria albida L. subsp. *albida* H, Medit. El.
Sideritis montana L. subsp. *montana* Th, widespread, Medit. El.
Stachys byzantina C. Koch H, Euro-Sib. El.
Teucrium chamaedrys L. (**), G
Thymus longicaulis C. Presl subsp. *longicaulis* var. *longicaulis* (**), Ch

PLANTAGINACEAE
Plantago coronopus L. Th, Euro-Sib. El.
P. lagopus L. Th, Medit. El.
P. lanceolata L. (**), H
P. major L. subsp. *intermedia* (Gilib.) Lange (**), H, widespread

THYMELAEACEAE
Daphne pontica L. (*), (**), Ph, Euxine El.

ELAEAGNACEAE
Elaeagnus angustifolia L. (**), Ph, widespread

LAURACEAE
Laurus nobilis L. (**), Ph, Medit. El.

SANTALACEAE
Osyris alba L. H, Medit. El.

EUPHORBIACEAE
Euphorbia helioscopia L. (*), (**), Th
E. peplus L. var. *peplus* (*), Th
E. peplus L. var. *minima* DC. (*), Th
E. segueriana Necker subsp. *niciciana* (Borbas *ex* Novak) Rech. fil. (*), H
Mercurialis annua L. (*), Th

URTICACEAE
Parietaria judaica L. (**), H, widespread
Urtica dioica L. (**), H
U. pilulifera L. (*), Th, Medit. El.

ULMACEAE
Celtis australis L. (**), Ph, Medit. El.

PLATANACEAE
Platanus orientalis L. (**), Ph, widespread

FAGACEAE
Quercus coccifera L. (*), (**), Ph, Medit. El.
Q. frainetto Ten. Ph, Euro-Sib. El.
Q. infectoria Olivier Ph
Q. ilex L. Ph, Medit. El.
Q. pubescens Willd. (*), Ph

SALICACEAE
Populus alba L. Ph, Euro-Sib. El.
P. tremula L. (**), Ph, widespread, Euro-Sib. El.
Salix alba L. (**), Ph, widespread, Euro-Sib. El.
S. babylonica L. Ph

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

RUBIACEAE
Cruciata taurica (Palas ex Willd.) Ehrend H, widespread
Galium aparine L. Th
Rubia tinctorum L. (**), H, widespread, Ir.-Tur. El.
Sherardia arvensis L. Th

MONOCOTYLEDONEAE

LILIACEAE
Asparagus acutifolius L. (**), H, Medit. El.
Allium neapolitanum Cyr. G, Medit. El.
A. paniculatum L. subsp. *paniculatum* G, Medit. El.
A. scorodoprasum L. subsp. *rotundum* (L.) Stearn G, widespread, Medit. El.
Asphodelus aestivus Brot. (**), G, Medit. El.
A. fistulosus L. G, Medit. El.
Gagea bohemica (Zauschn) Schultes & Schultes fil. G
Muscari comosum (L.) Miller G, widespread
M. neglectum Guss. G, widespread
Ruscus aculeatus L. var. *angustifolius* Boiss. (*), (**), H
Ornithogalum sigmoideum Freyn & Sint. (*), G, Euro-Sib. El.
O. umbellatum L. (**), G
Scilla autumnalis L. G, Medit. El.
Smilax excelsa L. (**), H

IRIDACEAE
Iris germanica L. G
I. sintenisii Janka G, Euro-Sib. El.
I. suaveolens Boiss. & Reuter G, E. Medit. El.

DIOSCOREACEAE
Tamus communis L. (*), (**), H

TYPHACEAE
Typha latifolia L. He

JUNCACEAE
Juncus conglomeratus L. He
J. heldreichianus Marsson *ex* Parl. He
Luzula multiflora (Ehrh. *ex* Retz.) Lej. H, widespread, Euro-Sib. El.

CYPERACEAE
Carex flacca Schreber G

POACEAE
Aegilops geniculata Roth. Th, Medit. El.
Agrostis capillaris L. var. ca*pillaris* H
A. stolonifera L. H, widespread, Euro-Sib. El.
Aira caryophyllea L. Th, Euro-Sib. El.
Alopecurus myosuroides Hudson Th
Anthoxanthum odoratum L. H
Avena barbata Pott *ex* Link Th
A. sterilis L. subsp. *sterilis* Th
A. wiestii Steudel Th
Brachypodium sylvaticum (Hudson) P. Beauv. H, widespread, Euro-Sib. El.
Briza maxima L. Th
Bromus hordeaceus L. Th
B. japonicus Thunb. subsp. *japonicus*
B. sterilis L. Th, widespread

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Catabrosa aquatica (L.) P. Beauv. G, widespread
Catapodium rigidum (L.) C. E. Hubbard *ex* Dony Th
Cynodon dactylon (L.) Pers var. *dactylon* (**), H
Cynosurus cristatus L. H, Euro-Sib. El.
C. echinatus L. Th, Medit. El.
Dactylis glomerata L. subsp. *hispanica* (Roth) Nyman H
Dasypyrum villosum (L.) Cand. Th, Medit. El.
Digitaria sanguinalis (L.) Scop. Th
Echinochloa crus–galli (L.) P. Beauv. Th
Holcus lanatus L. H, Euro-Sib. El.
Hordeum bulbosum L. G, widespread
H. marinum Hudson. Th
H. murinum L. subsp. *leporinum* (Link) Arc. var. *leporinum* Th
Lolium multiflorum Lam. Th
Lolium perenne L. H, Euro-Sib. El.
Melica ciliata L. subsp. *ciliata* G, widespread
Paspalum paspalodes (Michx.) Scribner G
Phalaris aquatica L. H
Phragmites australis (Cav.) Trin. ex Steudel (**), He, widespread, Euro-Sib. El.
Poa annua L. Th, cosmopolitan
P. bulbosa L. G
Rostraria cristata (L.) Tzvelev Th
Setaria verticillata (L.) P. Beauv. Th
S. viridis (L.) P. Beauv. Th, widespread
Sorghum halepense (L.) Pers. var. *halepense* H

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

Appendix 2: Exotic and cultivated plants of Kadıköy district

Abbreviations and symbols: * poisonous plant, ** medicinal plant, LR: lower risk, CR: critically endangered, VU: vulnerable, DD: data deficient. (The LR, CR, VU and DD values are according to IUCN standards) [23- 28].

GINKGOACEAE
Ginkgo biloba L.

TAXACEAE
Taxus baccata L. (*)

PINACEAE
Cedrus libani A. Richard
C. deodora (Roxb.) Loud.
C. atlantica (Endl) Manetti *ex* Carrière "Glauca"
Picea orientalis (L.) Link
P. punges Endelm
Pinus pinaster Ait.
P. pinea L. **(**)**
P. sylvestris L.
P. mugo Turra
P. strobus L.

CUPRESSACEAE
Chamaecyparis lawsoniana (Murr.) Parl.
Cupressus arizonica Grene
C.macrocarpa Hartw. *ex* Gord.
Juniperus horizontalis Moench
Thuja occidentalis L. "Globosa"
T. orientalis L.

ARECACEAE
Chamaerops excelsa Thunb.
Phoenix canariensis Chabaud.

LILIACEAE
Allium orientale Boiss.
Fritillaria persica L.
Hemerocallis fulva L.
Hyacinthus orientalis L. (*)
Lilium candidum L.
Ornithogalum nutans L. (*)
Tulipa species L.

IRIDACEAE
Crocus sativus L. (*)
Iris germanica L.

AGAVACEAE
Agave americana L.
A. americana L. "Marginata"
Yucca flamentosa L.
Y. gloriosa L.

AMARYLLIDACEAE
Narcissus pseudonarcissus L. "Dutch Master" (*)

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

MORACEAE
Ficus carica L. (**)
Maclura pomifera (Rafin.) Schneider (*)
Morus alba L. 'Pendula'
M. nigra L.

BERBERIDACEAE
Berberis thunbergii (Koch) DC. var. *atropurpurea* Chenault

MAGNOLIACEAE
Liriodendron tulipifera L.
Magnolia grandiflora L.

SAXIFRAGACEAE
Hydrangea macrophylla (Thunb.) Ser. (*)
Philadelphus coronarius L.

ROSACEAE
Chaenomeles japonica (Thunb.) Lindl. ex Spach
Cotoneaster franchetii Boiss.
Cydonia oblonga Miller **(**)**
C. japonica Pers.
Eriobotrya japonica (Thunb.) Lindl. (**)
Fragaria vesca L. (**)
Kerria japonica (L.) DC. "Pleniflora"
Laurocerasus officinalis Roemer (*),(**)
Malus floribunda Sieb. ex Van Houtte
M. sylvestris Miller (**)
Persica vulgaris Miller (*),(**)
Prunus avium L**.** (*)
P. cerasus L. (*),(**)
P. domestica L. (*),(**)
P. mahalep L. (**)
P. serrulata Lindl.
Pyracantha coccinea Roemer (**)
Pyrus communis L. (**)
Rosa damascena Miller (**)
R. gallica L.
R. multiflora Thunb.
Spiraea x vanhouttei (Briot) Zab.

FABACEAE
Acacia cyanophylla Lindl.
A. dealbata Link
Albizzia julibrissin (Willd.) Durazz. Rare (VU)
Caesalpinia gilliesii (Wall. ex Hook.) Wall. ex D. (*)
Cassia acutifolia Del.
Robinia pseudoacacia L. "Umbraculifera"
Sophora japonica L. var. *pendula* Loud. (*)
Wisteria sinensis (Sim.) DC. (*)

GERANIACEAE
Pelargonium zonale (L.) L'Herit. ex Ait.

OXALIDACEAE
Oxalis floribunda Linn.

BUXACEAE
Buxus sempervirens L. (*)

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

CELASTRACEAE
Euonymus japonicus L. *"*Aureo-variegatus*"*

ACERACEAE
Acer campestre L.
A. negundo L.
A. palmatum Thunb.
A. platanoides L.
A. pseudoplatanus L.

HIPPOCASTANACEAE
Aesculus x carnea Briottii
A. hippocastaneum L. (*), (**)

TILIACEAE
Tilia tomentosa (DC.) Moench

MALVACEAE
Hibiscus syriacus L.

TAMARICACEAE
Tamarix tetrandra Pallas ex Bieb.

VIOLACEAE
Viola x wittrockiana Gams.

LYTHRACEAE
Lagerstroemia indica L.

PUNICACEAE
Punica granatum L. (**)

MYRTACEAE
Eucalyptus camaldulensis Dehnh.
Callistemon citrinus (Curtis) Skeels

ARALIACEAE
Fatsia japonica (Thunb.) Decne. & Planch.

PRIMULACEAE
Primula vulgaris Huds.

OLEACEAE
Forsythia x intermedia Zabel
Fraxinus excelsior L. (**)
Jasminum fruticans L.
Syringa vulgaris L.

APOCYNACEAE
Vinca major L. (*)

VERBENACEAE
Clerodendron trichotomum Thunb.
Lantana camara L. "Aulanche" (*)
Vitex agnus-castus L. (**)

BIGNONIACEAE
Catalpa bignonioides Walt.
Campsis radicans (L.) Seem.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

CAPRIFOLIACEAE
Lonicera japonica Thunb.
Viburnum opulus L. "Sterile"
V. tinus L.

ASTERACEAE
Bellis perennis L. "Pompenette Red"
Calendula officinalis L.
Chrysanthemum maximum Ramond
Cosmos bipinnatus Cav.
Santolina chamaecyparissus L.
Senecio cineraria DC.
Tagetes erecta L.

MELIACEAE
Melia azaderach L. (*)

NYCTAGINACEAE
Bougainvillea spectabilis Willd
Mirabilis jalaba L. (*)

AIZOACEAE
Carpobrotus acinaciformis Folia

CACTACEAE
Opuntia ficus-indica (L.) Mill.

PASSIFLORACEAE
Passiflora coerulea L.

SAPINDACEAE
Koelreuteria paniculata Laxm.

VITACEAE
Parthenocissus quinquefolia (L.) Planch. (*)
Vitis vinifera L. (**)

EBENACEAE
Diospyros kaki L.
D. lotus L. **(**)**

BUDDLEIACEAE
Buddleia davidii Franch.

PITTOSPORACEAE
Pittosporum tobira Ait.

BRASSICACEAE
Brassica oleracea L. var. *acephala* DC.

CONVOLVULACEAE
Ipomea tricolor Cav.

JUGLANDACEAE
Juglans regia L. (**)

RANUNCULACEAE
Eranthis hyemalis (L.) Salisb.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 64-87

PLATANACEAE
Platanus acerifolia (Ait.) Willd.

THEACEAE
Camelia japonica L.

POLYGONACEAE
Polygonum bistorta L.

ARALIACEAE
Hedera colchica (C. Koch) C. Koch "Sulphur Heart" (*)
H. helix L. "Aureovariegata*"* (**)

RUTACEAE
Citrus sinensis (L.) Osbeck

ARACEAE
Anthurium andreanum Lenny

CANNACEAE
Canna x generalis Hortus

AMARANTHACEAE
Amaranthus caudatus L.

ONAGRACEAE
Oenothera biennis L.

AQUIFOLIACEAE
Ilex aquifolium L. (*)

CORYLACEAE
Corylus avellana L. var. *avellana* (**)

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 88-92

Maejo International Journal of Science and Technology

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

Full Paper

Application of separation of variables in Green's function to typical half-strip problem for elastic material

Paul J. Udoh* **and Asuquo E. Eyo**

Department of Mathematics, Statistics and Computer Science, University of Uyo, Nigeria

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

Received: 30 July 2009 / Accepted: 10 March 2010 / Published: 15 March 2010

Abstract: This paper presents the method of separation of variables in Green's function and its application to the problem of half-strip of elastic material. Unlike that of the full strip, the half-strip problem differs in specification of the boundary conditions. Though similar to the full-strip problem in conjugality conditions, the half-strip problem is amenable to the application of Poisson's equation instead of Laplace's equation. The results show that this problem is one-dimensional and the shearing stresses are non-symmetric.

Keywords: Green's function, half-strip pboblem, elastic material

Introduction

The method of separation of variables is the most general one for constructing Green's functions. This method was used to solve a full-strip elastic problem [1]. The approach here differs from that of Humphrey and Rajagopol [2] where the basic idea is that of decomposing the motion into two parts: one due to traction-free uniform heating and the other to isothermal mechanical loading.

However, in the method of separation of variables, the approach is holistic without analogous representation into singular and regular parts [3]. The method of separation of variables enables us to construct the Green's functions both for Laplace's and Poisson's operators [4] for simple boundary-value problems. This method allows us to represent the functions in an infinite series as is done in regularity conditions [5] and for the strain energy **W** [2]. The method adopted here enables us to construct and take into account the property of orthogonality for trigonometric functions in terms of one-dimensional differential equations.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 88-92

Mathematical Formulations

The Green's function G(x,ξ) for Poisson's equation $\nabla_x^2 G(x,\xi) = -\delta(x-\xi)$, and for the half-strip we specify the intervals as $0 < x_1 < \infty$ and $0 \le x_2 \le a_2$ under the following conditions:

$$\frac{\partial G}{\partial x_1} = 0 \; ; x_1 = 0, \; 0 \le x_2 \le a_2 \tag{1}$$

$$\frac{\partial G}{\partial x_1} = 0 \; ; x_2 = 0, a_2 \; ; \; 0 \le x_1 \le \infty \tag{2}$$

Here, the function G(x,ξ) should be bounded at infinity, i.e. $G|_{x_1 = 0} < \infty$

Computational Procedure

Following our earlier results [6], the solution to this problem is sought by means of the method of separation of variables leading to the following general trigonometric series:

$$G = a_0 + \sum_{n=1}^{\infty} a_n \cos v_1 x_2 + \sum_{m=1}^{\infty} b_m \sin v_1 x_2 \tag{3}$$

where, a_0, a_n and b_m are the functions of the independent variable x_i. The boundary conditions of this problem with respect to the independent variables x_2 reduce equation (3) with the general series to the following form:

$$G = a_0 + \sum_{n=1}^{\infty} a_n \cos v_1 x_2, \; v_1 = \frac{n\pi}{a_2}, n = 1,2,3, \ldots\ldots\ldots \tag{4}$$

Substituting this expression in (4) for the Green's function G into the Poisson's $\nabla_x^2 G(x,\xi)$, we get

$$G = a_0'' + \sum_{n=1}^{\infty} (a_n'' - v_1^2 a_n) \cos v_1 x_2 = -\delta(x_1 - \xi_1)\,\delta(x_2 - \xi_2) \tag{5}$$

where, in the method of separation of variables, it takes the form:

$$\delta(x-\xi) = \delta(x_1 - \xi_1)\,\delta(x_2 - \xi_2) \tag{6}$$

Integrating equation (5) with respect to the independent variable x_2, we note that all the integrals with the exceptions of equations (7) and (8):

$$\int_0^{a_2} a_0'' \, dx_2 = a_0'' a_2 \tag{7}$$

$$\int_0^{a_2} \delta(x_1 - \xi_1)\,\delta(x_2 - \xi_2)\, dx_2 = \delta(x_1 - \xi_1) \tag{8}$$

are equal to zero. Therefore, to determine the function $a_0(x_1)$, we obtain the following differential equation of the second order:

$$a_0'' = -a_2^{-1}\,\delta(x_1 - \xi_1) \tag{9}$$

and the boundary condition and the condition at infinity:

$$a_0'(x_1 = 0); \qquad a_0(x_1 = \infty) < \infty \tag{10}$$

To construct the Green's function for this boundary-value problem, we use the standard technique [2]. The general solution of the 1D equation is sought in the form of

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 88-92

$$a_2(x_1) = \begin{cases} c_1 x_1 + c_2, & x_1 \le \xi_1 \\ k_1 x_1 + k_2, & x_1 \ge \xi_1 \end{cases} \quad (11)$$

From the boundary condition and the condition at infinity, we obtain

$$a_0'(x_1 = 0) = 0 \Rightarrow c_1 = 0; \qquad a_0(x_1 = \infty) < \infty \Rightarrow k_1 = 0 \quad (12)$$

So the function $a_0(x_1)$ can be rewritten in the form:

$$a_2(x_1) = \begin{cases} c_2, & x_1 \le \xi_1 \\ k_2, & x_1 \ge \xi_1 \end{cases} \quad (13)$$

which satisfies the ordinary differential equation of the first order. Next, from the condition of conjugality at the point $x_1 = \xi_1$,

$$a_0(x_1 = \xi_1 - 0) = a_0(x_1 = \xi_1 - +0) \quad (14)$$

We obtain the following result: $c_2 = k_2 = b =$ constant.
Finally, for the desired function, we find

$$a_0(x_1, \xi_1) = \begin{cases} b, & x_1 \le \xi_1 \\ b, & x_1 \ge \xi_1 \end{cases} \quad (15)$$

Now, to determine the function a_m, we multiply both parts of the equation (5) by

$$\cos v_2 x_2 \qquad (v_2 = s\pi / a_2, \quad s = 1,2,3.....) \quad (16)$$

and take the integral with respect to the variable x_2.
Taking into account the property of orthogonality for trigonometric functions, we get the respective 1D differential equation:

$$(a_n'' - v_1^2 a_m)\frac{a_2}{2} = -\delta(x_1 - \xi_1)\cos v_1 \xi_2 \quad (17)$$

and the boundary conditions:

$$a_n'(x_1 = 0); \qquad a_n'(x_1 = \infty) < \infty \quad (18)$$

In solving equation (17), the following integrals have been calculated:

$$\int_0^{a_2} \cos v_2 x_2 \cos v_1 x_2 d x_2 = \begin{cases} 0, & v_1 \ne v_2 \\ a_0/2, & v_1 = v_2 \end{cases} \quad (19)$$

$$\int_0^{a_2} \delta(x_1 - \xi_1)\,\delta(x_2 - \xi_2)\cos v_2 x_2 \, dx_2 = \delta(x_1 - \xi_1)\cos v_2 \xi_2 \quad (20)$$

By taking a notation, $a_m = 2a_2^{-1}\bar{a}_m \cos v_2 \xi_2$ (21)

with the conditions in (18), we reduce equation (19) to the boundary-value problem formulated to determine the function $\bar{a}_n$, i.e.

$$(a_n'' - v_1^2 a_m) = -\delta(x_1 - \xi_1) \quad (22)$$

$$a_n'(x_1 = 0) = 0 \quad (23)$$

$$a_n'(x_1 = \infty) < \infty \quad (24)$$

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 88-92

Using the standard technique [7], the general solution of equation (11) is written in the form:

$$\bar{a}_n = \begin{cases} c_1 e^{-v_1 x_1} + c_2 e^{v_1 x_1} \quad ; & x_1 \le \xi_1 \\ k_1 e^{-v_1 x_1} + k_2 e^{v_1 x_1} \quad ; & x_1 \ge \xi_1 \end{cases} \tag{25}$$

Then from the conditions of conjugality at the point $x_1 = \xi_1$,

$$\bar{a}_n (x_1 = \xi_1 - 0) = \bar{a}_n (x_1 = \xi_1 + 0) \tag{26}$$

$$\bar{a}_n (x_1 = \xi_1 - 0) - \bar{a}_n (x_1 = \xi_1 + 0) = 1 \tag{27}$$

we get a system of two simultaneous linear algebraic equations:

$$(c_1 - k_1)\, e^{-v_1 \xi_1} + (c_2 - k_2)\, e^{v_1 \xi_1} = 0 \tag{28}$$

$$v_1 \left[(c_1 - k_1)\, e^{-v_1 \xi_1} - (c_2 - k_2)\, e^{v_1 \xi_1} \right] = -1 \tag{29}$$

Then from the boundary conditions and conditions at infinity, we get

$$\bar{a}_m (x_1 = 0) = 0 \Rightarrow c_2 - c_1 = 0 \tag{30}$$

$$\bar{a}_n' (x_1 = \infty) < \infty \Rightarrow k_2 = 0 \tag{31}$$

For the coefficients, we finally obtain the following values:

$$k_1 = \frac{e^{-v_1 \xi_1} - e^{v_1 \xi_1}}{2 v_1}; \quad c_1 = c_2 = \frac{e^{-v_1 \xi_1}}{2 v_1}; \quad k_2 = 0 \tag{32}$$

Using standard transformations, the Green's functions in equation (25) can be represented as

$$\bar{a}_n (x_1, \xi_1) = \begin{cases} \dfrac{1}{2v_1} (e^{v_1 (x_1 - \xi_1)} + e^{-v_1 (x_1 - \xi_1)}); & x_1 \le \xi_1 \\ \dfrac{1}{2v_1} (e^{-v_1 (x_1 - \xi_1)} + e^{-v_1 (x_1 - \xi_1)}); & x_1 \ge \xi_1 \end{cases} \tag{33}$$

On account of equation (21), the Green's function G(x, ξ) for the initial boundary-value problem can be written in the form:

$$G(x, \xi) = \begin{cases} G_\ell (x, \xi) = b + \dfrac{2}{a_2} \displaystyle\sum_{n=1}^{\infty} \frac{1}{2v_1} (e^{v_1 (x_1 - \xi_1)} + e^{-v_1 (x_1 - \xi_1)}) \cos v_1 x_2 \cos v_1 \xi_2; & x_1 \le \xi_1 \\ G_\ell (x, \xi) = b + \dfrac{2}{a_2} \displaystyle\sum_{n=1}^{\infty} \frac{1}{2v_1} (e^{v_1 (x_1 - \xi_1)} + e^{-v_1 (x_1 - \xi_1)}) \cos v_1 x_2 \cos v_1 \xi_2; & x_1 \ge \xi_1 \end{cases} \tag{34}$$

However, by making use of the known sum [1, 3, 4],

$$\sum_{n=1}^{\infty} \frac{p^n}{n} \cos n\alpha = -\ln \sqrt{1 - 2p \cos \alpha + p^2}; \; p^2 < 1, \; 0 \le \alpha \le 2\pi; \quad \text{or} \; p^2 \le 1, \; 0 \le \alpha \le 2\pi \tag{35}$$

and trigonometric formula:

$$\cos v_1 x_2 \cos v_1 \xi_2 = \frac{1}{2} \left[\cos v_1 (x_1 - \xi_2) + \cos v_1 (x_2 + \xi_2) \right], \tag{36}$$

then one can take the sum of the ordinary infinite series. After some computation, we obtain the final expression for the Green's function of the initial boundary-value Neumann's problem for the half-strip as:

$$G = b - \frac{1}{2\pi} \ln E\, E_1\, E_2\, E_{12} \tag{37}$$

where the functions E, E_1, E_2, and E_{12} are determined by the expressions:

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 88-92

$$E = \sqrt{1 - 2e^{\frac{\pi}{a_2}(x_1 - \xi_1)} \cos\pi/a_2 (x_2 - \xi_2) + e^{\frac{2\pi}{a_2}(x_1 - \xi_1)}} \quad (38)$$

$$E_1 = \sqrt{1 - 2e^{-\frac{\pi}{a_2}(x_1 + \xi_2)} \cos\pi/a_2 (x_2 - \xi_2) + e^{\frac{-2\pi}{a_2}(x_1 + \xi_1)}} \quad (39)$$

$$E_2 = \sqrt{1 - 2e^{\frac{\pi}{a_2}(x_1 - \xi_1)} \cos\pi/a_2 (x_2 + \xi_2) + e^{\frac{2\pi}{a_2}(x_1 - \xi_1)}} \quad (40)$$

$$E_{12} = \sqrt{1 - 2e^{-\frac{\pi}{a_2}(x_1 + \xi_2)} \cos\pi/a_2 (x_2 + \xi_2) + e^{\frac{-2\pi}{a_2}(x_1 + \xi_1)}} \quad (41)$$

Conclusions

In this paper, we have made use of the method of separation of variables in Green's functions to solve the problem of the half-strip problem. A particular focus is given to the problem of the half-strip using Poisson's equation. For ease of formulation, we use trigonometric functions to construct the Green's functions for the Poisson's equation. The results differ from the full-strip problem by shearing strain as shown in equation (41).

References

1. B. A. Boley, "A method for the construction of Green's functions", *Q. J. Appl. Math.*, **1956**, *14*, 249-257.
2. J. D. Humphrey and K. R. Rajagopol, "Finite thermoealsticity of constrained elastomers subject to biaxial loading", *J. Elasticity*, **1998**, *40*, 189-200.
3. M. D. Greenberg, "Application of Green's Functions in Science and Engineering", Prentice-Hall Inc., Englewood Cliffs, N. J., **1971**.
4. V. D. Seremet, "Handbook of Green's Functions and Matrices", WIT Press, Southampton, **2003**.
5. S. P. Gavelia, "On calculation of Green's matrices for applied problems allowing separation of variables", *J. Boundary Value Problems Math. Phys. Kiev*, **1979**, *13*, 59-67.
6. P. J. Udoh, "On the separation of variables in Green's functions for elastic materials", *Kenya J. Sci.*, **2010** (in press).
7. Y. A. Melnikov, "Green's Functions in Applied Mechanics", Computational Mechanics Publications, Southampton, **1995**, p. 267.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 93-100

Maejo International Journal of Science and Technology

ISSN 1905-7873

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

Communication

Physicochemical characteristics of the rhizosphere soils of some cereal crops in Ambo Woreda, West Shoa, Ethiopia

Louis E. Attah

Department of Chemistry, Ambo University, P.O. Box 19, Ambo, Ethiopia

E-mail: louattah@yahoo.com

Received: 27 July 2009 / Accepted: 6 March 2010 / Published: 16 March 2010

Abstract: In this study, physicochemical properties of rhizosphere soils of some cereal crops in Ambo Woreda, West Shoa in Ethiopia have been investigated. Soil samples were collected from four different localities, viz. Awaro, Senkele, Meja and Guder, and their edaphic characteristics are determined. The soils are dominated by clay (40.4-45.8%) along with coarse particles of sand. Bulk density, organic carbon (1.52-1.81%) and electrical conductivity (1.3-1.9 dSm) are low in all the soil samples. The soils are acidic with pH varying from 6.2 to 6.7. There are similarities in the relatively low content of available phosphorus (1.4-2.4 mg kg^{-1}) and high available nitrogen content (480-986 mg kg^{-1}) in all the soil samples while available potassium content (240-496 mg kg^{-1}) is found to be medium in Awaro soil but high in the other three soil samples. Deficiencies are observed in the levels of available micro-nutrients (Cu: 1.2-1.8 µg g^{-1}, Zn: 1.2-1.8 µg g^{-1} and Mn: 3.2-3.8 µg g^{-1}) while the Fe content is sufficient in all the soil samples (340-496 µg g^{-1}). With proper soil management, the farmlands studied are recommended for the cultivation of cereal crops.

Keywords: rhizosphere soil, cereal crops, nutrients, edaphic characteristics

Introduction

The importance of soil fertility and plant nutrition to the health and survival of all life cannot be overstated. Understanding of the diversity, distribution, characteristics and processes of soil is important for agricultural development and productivity of agricultural systems. Rhizosphere, the region of soil in

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 93-100

the vicinity of the root and under its constant influence, is a dynamic soil environment. In general, there is a higher nutrient availability in rhizosphere soil than in non-rhizosphere soil [1-2]. Soils differ widely in the ability to meet plant requirements. Most soils have moderate natural soil fertility but can be considerably improved by soil amelioration. For successful farming, the natural fertility of the soil is often less important than its potential productivity after the removal of its inherent limiting factors. Soils with high natural fertility can produce substantial crop yields without added fertilisers and can achieve even higher yields with additional supply of critical nutrients. Good soil fertility provides the basis for all other measures for successful farming.

Cereal crops such as wheat, teff, maize and corn can grow in many soil types ranging from heavy clay to almost pure peat or sandy soils. These crops require well-aerated soil in which roots and water can move easily to a certain depth. One of the basic limiting factors for cereal crop yield including that of wheat in Ethiopia is poor soil fertility. Two macro-nutrients, i.e. nitrogen and phosphorus, are deficient in most of the soils of the highland plateau [3]. According to Aydemir [4], 50% of nitrogen in the fertiliser which is applied to soil is taken by the plants in the first year. Of this, 30% is fixed by microorganisms, 15% is lost by denitrification and 5% is lost by leaching. Although the role of micro-nutrients in plant nutrition is as vital as that of macro-nutrients, under subsistence farming system like the one in Ethiopia where farmers cannot afford to buy the recommended doses of primary macro-nutrients (N, P, K), application of micro-nutrients to crops are not usually considered as a priority in alleviating soil fertility problems. Considerable variation in micro-nutrient content of soils in Ethiopia has been reported by Desta [5]. The ability of soil to supply nutrients for crop growth and maintain soil physical conditions to optimise crop yield is known to be an important component of soil fertility that determines the productivity of an agricultural system [6]. This ability of soil decreases with increase in soil erosion, which involves detachment and transportation of soil particles, exposing relatively infertile subsoil which is poor in biological and chemical processes and resulting in decreased crop production [7-8].

Gunes et al. [9] have evaluated the soil fertility status in Ankara district of Turkey for the production of cereal crops. Physicochemical properties of soils under vegetable cultivation in the highland of Cameroon have been investigated by Salubin et al [10]. The variation in the physico-chemical properties of a series of soils in south-western Nigeria has been reported by Olatunji et al [11]. Udotong et al. [12] has reported the microbiological and physicochemical properties of wetland soils in Eket, Nigeria. This communication reports the evaluation of the physicochemical characteristics of rhizosphere soils in some areas of Ambo Woreda, West Showa, Ethiopia during the dry season and their suitability for the cultivation of some common cereal crops in this area.

Materials and Methods

Description of the study site

This study is conducted on 4 locations (Awaro, Senkele, Guder and Meja) in Ambo district of West Shoa, Ethiopia, which is about 115 km from Addis Ababa, the capital city (Figure 1). About 88% of its terrain having gradients between 0.5 and 15%, which favours mechanised farming, the West Shoa zone is endowed with a high potential for agriculture. The soil samples were collected from farmlands at

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 93-100

a depth of 0-15 cm during the dry-season months of October-December, 2008. The average maximum temperature in all the study areas was 29.6°C during the year of study.

Awaro is located 5 km in the eastern direction from Ambo town. It has a moderate slope and the land has been cultivated for vegetables with some grass-covered areas mainly used for the grazing of cattle. Senkele, which is about 6 km west of Ambo, has a gentle slope and the land has been used for only 3 years for the cultivation of maize. Guder, located within 7 km of the town, is a nearly flat slope and its farmland has been cultivated for 13 years primarily for maize, teff and wheat. Meja is 4 km in the southern direction of Ambo town. It has a moderate slope and has been cultivated continuously for 11 years for teff and wheat.

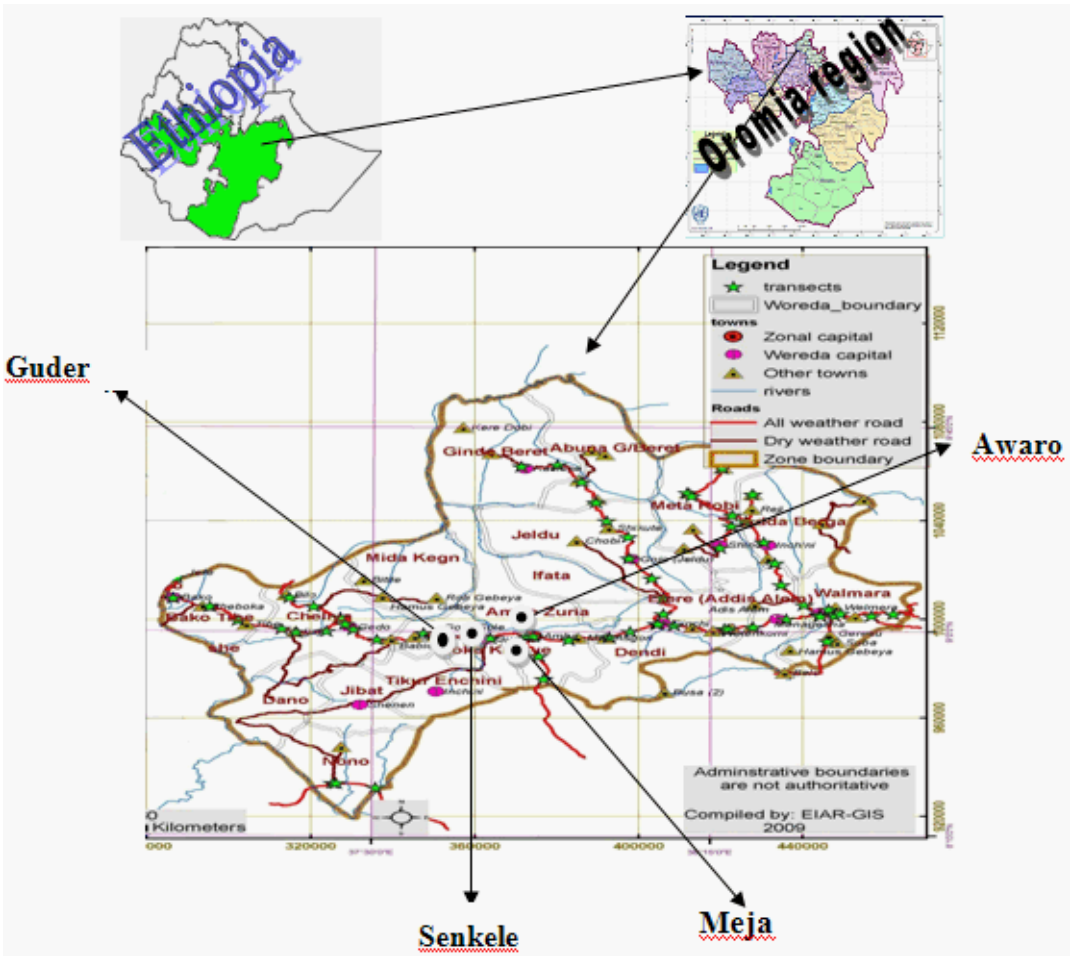


Figure 1. Location map of West Shoa zone in Oromia region of Ethiopia

Sample analysis

The soil samples were air-dried, crushed and passed through a 2-mm sieve and then mixed thoroughly to obtain a homogeneous mixture. Particle size analysis was performed using the Bouyoucous hydrometer method [13]. Soil pH and electrical conductivity were determined in a soil:water (1:2.5) mixture using a pH meter. Soil organic carbon was assessed according to Piper [14]. Moisture content, bulk density and nitrogen were determined according to the method of Jackson [15].

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 93-100

Available phosphorus was determined by the method of Kitson and Mellon [16] while potassium was determined by a flame photometer. Other nutrients, viz. copper, iron, manganese and zinc, were analysed by an atomic absorption spectrophotometer (Unicam 919). All results were presented as an average of five determinations (n=5) ± standard deviation (SD).

Results and Discussion

Physicochemical characteristics of the soil samples are presented in Table 1. The particle size reveals the texture of the soils as varying from clay loam to red and black clay loam with sand, silt and clay fractions ranging between 30.7-34.2%, 24.5-28.6% and 40.4-45.8% respectively. The characteristics of a soil largely determine its utilisation [17]. From this result, the texture of the soils under investigation can be classified as loam with excellent properties for crop cultivation. The moisture content of all soils is similar except for Meja soil, which has a slightly lower moisture content (8.5%). The soil pH, which varies from 6.2 to 6.7, lies within the preferred range for most crops.

The fertility of soil is intimately linked with its organic matter which has an influence on the physical, chemical and biological properties of the soil. It is well known that under continuous agricultural practice, the organic matter content in the top soil will decrease. The organic carbon content in all studied soil samples is found to be very low (< 4%). This may be attributed to intensive agricultural practices that aggravate organic carbon oxidation [18]. The bulk density of the studied soils is found to be equal or less than 1.44, which is common in cultivated soils. The low bulk density found in soils from Senkele, Meja and Guder indicates that the soils are not compacted and have more porosity. This is beneficial to root activity, water infiltration into soil, and overall growth of crops. The high bulk density of Awaro soil (uncultivated) is unexpected since the bulk density is generally inversely correlated with organic matter content. Soil with very high bulk density can limit root growth, air circulation and availability of less mobile essential plant nutrients such as P and K [19]. The electrical conductivity values of the saturated extracts of the four soil samples can be classified as Index 0 in the ADAS recommendations [20]. This means that they are within the normal range found for outdoor soils and pose no restriction for field-crop cultivation nor adversely affect crop yield.

The available major nutrients in the soil samples are presented in Table 2. Of all the major nutrients, phosphorus probably has the most complicated chemistry in the soil, at least as far as the assessment of its level and of the fertiliser requirement for it are concerned. Phosphorus occurs in soil in both organic and inorganic forms, the latter being more important for crop nutrition. A low available phosphorus is indicative of acute deficiency. There is a similarity among all the soil samples in a very low phosphorus level which falls between Index 0-1 category [20]. At this level, there is a possibility of failure of arable crops if there is no further application of P fertiliser. The cultivated soils of Senkele, Meja and Guder have a slightly higher phosphorus content than that in the uncultivated Awaro soil. The difference could be accounted for by the P fertiliser application over the years in the cultivated areas. The nitrogen content in all the soils is high, ranging between 460±6.5 - 986±5.2 mg Kg^{-1}. The main source of both N and P in the farmlands is organic matter.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 93-100

Table 1. Physicochemical properties of rhizosphere soils of Ambo Woreda

Soil property	Awaro	Senkele	Meja	Guder
Sand (%)	34.2 ± 0.05	32.4 ± 0.02	30.7 ± 0.06	32.8 ± 0.02
Silt (%)	26.5 ± 0.02	24.6 ± 0.04	28.6 ± 0.04	24.5 ± 0.06
Clay (%)	40.4 ± 0.05	44.5 ± 0.05	43.7 ± 0.02	45.8 ± 0.04
Moisture (%)	10.6 ± 0.05	9.5 ± 0.04	8.5 ± 0.04	10.8 ± 0.05
pH	6.2 ± 0.04	6.4 ± 0.05	6.7 ± 0.02	6.2 ± 0.05
Electrical conductivity (dSm^{-1})	1.3 ± 0.02	1.6 ± 0.02	1.9 ± 0.02	1.8 ± 0.02
Bulk density ($g\ cm^{-3}$)	1.44 ± 0.05	0.84 ± 0.02	1.24 ± 0.04	1.27 ± 0.06
Organic carbon (%)	1.81 ± 0.02	1.66 ± 0.02	1.52 ± 0.04	1.62 ± 0.05
Textural class	Clay loam	Clay loam	Red clay loam	Black clay loam

Table 2. Available macro- and micro-nutrients in the rhizosphere soils of Ambo Woreda

Location	Macro-nutrient ($mg\ kg^{-1}$)			Micro-nutrient ($\mu g\ g^{-1}$)			
	N	P	K	Cu	Zn	Mn	Fe
Awaro	480 ± 6.5	1.4 ± 0.2	240 ± 3.2	1.2 ± 0.02	1.3 ± 0.04	3.8 ± 0.2	60.2 ± 0.8
Senkele	645 ± 4.6	1.6 ± 0.4	340 ± 4.4	1.8 ± 0.02	1.4 ± 0.04	3.4 ± 0.4	96.5 ± 0.8
Meja	648 ± 5.2	1.8 ± 0.02	395 ± 4.8	1.6 ± 0.02	1.2 ± 0.02	3.2 ± 0.06	98.4 ± 1.6
Guder	986 ± 5.2	2.4 ± 0.08	496 ± 4.8	1.6 ± 0.04	1.8 ± 0.02	3.6 ± 0.04	108 ± 1.8

Available potassium content is medium (240±8.2 mg Kg^{-1}) in the Awaro soil but high in the other three soil samples (>281.6 mg Kg^{-1}) [21], which could be attributed to the continuous use of animal waste (farmyard manure) for several years in those three locations. According to the farmers, good yields are obtained for maize and corn crops in both Senkele and Meja farms. Similarly, wheat yield obtained from Guder is also good, although its maize yield is poor.

The availability of trace elements for plants is influenced by many soil and environmental factors as reported by Jones [22]. The concentrations of Cu, Zn, Mn and Fe in all the soil samples indicate deficiencies in Cu, Zn and Mn (Table 2). Available Fe content (60.2±0.8 - 108 ±1.8 $\mu g\ g^{-1}$) in all the soil samples is considered sufficient [22]. As reported by Sillanpaa [23], total Cu content in soil generally falls in the range of 2-100 ppm. The strong interaction which is generally held to occur between Cu and soil organic matter does not affect Cu availability to plants, although it does influence

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 93-100

the concentration of Cu in soil solutions [24]. From the above results, Cu level is low in all the soils as compared to the critical value given for Nigerian sandy soil [25-26]. This is in agreement with the conclusion made based on the assessment of Ethiopian soils [27]. Cereal plants such as wheat and maize are particularly sensitive to Cu deficiency. The low Cu content (<2ppm) in the soils may account for the poor yield in maize crops experienced in the Guder farmland.

A wide range of crops are sensitive to Mn deficiency, which is common in calcareous soils and soils of high pH. In soil, Mn originates primarily from the decomposition of ferromagnesium rocks and its content varies from 20 to 10,000 ppm with an approximate mean of 1,000 ppm as reported by Lindsay [28]. The deficiency in Mn observed in the studied soil samples may be attributed to the low organic carbon content, the mildly acidic soil pH and the potential adsorption of Mn on Fe or aluminum oxide present in the soils [29].

Plants vary in their zinc requirement as well as their ability to extract zinc from soil. Cereal crops such as maize are sensitive to Zn deficiency. The usual range of Zn in soil (1-900 ppm) with an approximate mean of 90 ppm has been reported by Davies [30] and Fairbridge and Finkl [31]. On the other hand, total zinc content in soil generally falls within the range of 10 to 300 ppm according to Sillanpaa [23]. From Table 2, a very low level of Zn is obtained for all the soil samples. Zn availability is mainly related to pH and complexation, and as pH increases organically-bound Zn decreases, which could be responsible for the apparent Zn deficiency in the soil samples [32].

Conclusions

The physicochemical characteristics of Ambo Woreda rhizosphere soils of some cereal crops have been determined. Results show that two macronutrients (N, K) are found optimal while available P is deficient for major crops such as maize, wheat, teff and corn grown in the area. In order to obtain optimal yield for the cereal crops in the farmlands studied, therefore, the deficiencies in the available P and micronutrients (Cu, Zn and Mn) along with the low organic carbon content of the soils should be remedied by appropriate soil management through improved drainage and application of minerals and organic fertilisers.

Acknowledgements

The author is grateful to Prof. T. Selvaraj of the Departmemt of Plant Sciences, Ambo University, and to D. E. Olana of Muthaiyah Research Foundation for Biological Science, Thanjavur, India for most of the analyses of the soil samples.

References

1. E. H. Curl and B. Truelove, “The Rhizosphere”, Springer-Verlag, New York, **1986**, p. 228.
2. C. H. Yong and D. E. Crowley, “Rhizosphere microbial community structure in relation to root location and plant iron nutritional status”, *Appl. Environ. Microbiol.*, **2000**, *66*, 345-351.

3. A. Woldead, M. Tekaligin, B. Mengesha and A. Tefera, "Soil fertility management studies on wheat in Ethiopia", in: "Wheat Research in Ethiopia : A Historical Perspective", (Ed. H. Gebre-marian, D. G. tanner and H. Mengistu), IAR/CIMMYT, Addis Ababa, **1991**, pp.137-172.
4. O. Aydemir, "Fertilizers and environmental pollution", *Atatuk Univ. J. Fac. Agric., Erzurum*, **1979**, *10*, 189-197.
5. B. Desta, "Micronutrient status of some Ethiopian soils", Soil Science Bulletin 4, Institute of Agricultural Research**,** Addis Ababa, Ethiopia, **1982**.
6. R. Lal, "Soil erosion from tropical arable lands and its control", *Adv. Agron.*, **1984**, *37*, 183-248.
7. L. J. Cihacek and J. B. Swan, "Effects of erosion on soil chemical properties in the north central region of the United States", *J. Soil Water Conserv.*, **1994**, *49*, 259-265.
8. G. K. Sims, "Biological degradation of soil", *Adv. Soil Sci.*, **1990**, *11*, 289-330.
9. A. Gunes, M. Alpaslan, M. R. Karaman, K. Yalvac and S. Yasar, "Evaluations of nitrogen, phosphorous, potassium, boron and zinc nutritional status of wheat grown in Ankara district with soil and plant samples", *J. Agric. Fac. Gaziosmanpassa Univ. Tokat*, **2003**, *20*, 149-155.
10. A. R. Salubin, A. R. Zulfahni, K. M. Lai, P. Errol and L. Tukimat, "Physico-chemical properties of soils under vegetables cultivation in Cameroon Highland", Proceedings of the Regional Symposium on Environment and Natural Resources, 10-11th April **2002**, Kuala Lumpur, Malaysia, pp. 668-674.
11. O. O. Olatunji, G. E. Akinbola, G. O. Oyediran, B. A. Lawal, W. B. Akanbi, J. C. Obi and F. M. Owoade, "Mycorrhiza fungi distribution in six different soil types of South Western Nigeria", *Res. J. Agron.*, **2008**, *2*, 52-55.
12. I. A. Udotong, U. M. J. Ofonime and J. I. R. Udotong, "Microbiological and physico-chemical studies of wetland soils in Eket, Nigeria", *Proc. World Acad. Sci., Eng. Technol.*, **2008**, *34*, 837-842.
13. G. T. Bouyoucous, "Improved hydrometer method for making particle size analysis of soils", *Agron. J.*, **1962**, *54*, 464-465.
14. C. S. Piper, "Soil and Plant Analysis", Interscience Publishers Inc., New York, **1950**, p. 368.
15. M. L. Jackson, "Soil Chemical Analysis", Prentice Hall of India, New Delhi, **1973**, p. 245.
16. R. E. Kitson and M. G. Mellon, "Colorimetric determination of phosphorus as molybdovanado phosphoric acid", *Ind. Eng. Chem. Anal. Ed.*, **1944**, *9*, 379-383.
17. J. Hassink, "Effect of soil texture on carbon and nitrogen mineralization in grassland soils", *Biol. Fertil. Soils*, **1992**, *14*, 126-134.
18. N. Wakene and G. Herf, "Forms of phosphorus and status of available micronutrients under different land use systems of in Bako area of Ethiopia", *Ethiopian J. Nat. Res.*, **2003**, *5*, 17-37.
19. M. S. Dolan, R. H. Dowdy, W. B. Voorhees, J. F. Johnson and A. M. Bidwell-Scrader, "Corn phosphorus and potassium uptake in response to soil compaction", *Agron. J.*, **1992**, *84*, 639-642.
20. ADAS (UK Agricultural Development and Advisory Service), "Fertilizer recommendations", Bulletin 209, HMSO (Her Majesty Stationary Office), London, **1973**.
21. Anonymous, "Physico-Chemical Properties of Soils", Agricultural Division Publications, Tamil Nadu Agricultural University, Tamil Nadu, India, **1988**, p. 218.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 93-100

22. J. B. Jones, "Plants tissue analysis for micronutrients", in: "Micronutrients in Agriculture" (Ed. J. J. Mortvedt, P. M. Giodano and W. L. Lindsay), Soil Science Society of America, Madison, Wisconsin, **1972**, p. 319.
23. M. Sillanpaa,"Trace elements in soil and agriculture", Soils Bulletin No.17, FAO, Rome, **1972**, p. 151.
24. E. W. Russell, "Soil Conditions and Plant Growth", 10th Edn., Longman, London, **1973**, p. 34.
25. A. Kabata-Pendias and H. Pendias, "Trace Elements in Soils and Plants", 3rd Edn., CRC Press, Boca Raton, **2001**.
26. J. R. Landon, "Booker Tropical Soil Manual: A Handbook for Soil Survey and Agricultural Land Evaluation in the Tropics and Subtropics", Longman Scientific and Technical, Essex, **1991**, p. 146.
27. M, Sillanpaa, "Micronutrient Assessment at Country Level: An International Study", FAO, Rome, **1990**, p. 208.
28. W. L. Lindsay, "Inorganic phase equilibrium of micronutrients in soils", in: "Micronutrients in Agriculture" (Ed. J. J. Mortvedt, P. M. Giordano and M. L. Lindsay), Soil Science Society of America, Madison, Wisconsin, **1972**, p. 41.
29. N. C. Brady and R. R. Well, "The Nature and Properties of Soil", 13th Edn., Pearson Education Inc., New York, **2002**, p. 959.
30. B. E. Davies, "Applied Soil Trace Elements", Wiley, Chichester, **1980**, p. 482.
31. R. W. Fairbridge and C. W. Finkl, "The Encyclopedia of Soil Science, Part 1", Dowden, Hutchinson and Ross Inc. Stroudsburg, Pennsylvania, **1979**, p. 98.
32. R. E. Lucas and B. D. Knezek, "Climatic and soil conditions promoting micronutrient deficiencies in plants", in: "Micronutrients in Agriculture", (Ed. J. J. Mortvedt, P. M. Giordano and W. L. Lindsay), Soil Science Society of America, Madison, Wisconsin, **1972**, pp. 265-288.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

Maejo International Journal of Science and Technology

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

Full Paper

Nutritional and fatty acid profiles of sun-dried edible black ants (*Polyrhachis vicina* Roger)

Subhachai Bhulaidok, Oranut Sihamala, Lirong Shen and Duo Li*

Department of Food Science and Nutrition, Zhejiang University, Hangzhou, China

*Corresponding author, e-mail: duoli@zju.edu.cn

Received: 2 September 2009 / Accepted: 21 March 2010 / Published: 24 March 2010

Abstract: Determination of the nutritional composition of sun-dried edible black ants (*Polyrhachis vicina* Roger) cultivated in Zhejiang and Guizhou Provinces, China, was carried out. The Zhejiang and Guizhou ants contained 31.5% and 41.5% protein, 15.7% and 15.9% lipid, and 25.4% and 26.4% fibre respectively. Monounsaturated fatty acids were the most predominant fatty acids (71.4–72.7% of total fatty acids) found in both ant samples, followed by saturated fatty acids (23.8–25.5%) and polyunsaturated fatty acids (3.1–3.7%). A significant amount of n-3 fatty acids was detected: 87.4 mg/100g and 145.6 mg/100g in Zhejiang and Guizhou ants respectively. Phosphorus, iron and calcium were the main minerals found in the ant samples. A small amount of selenium was also found.

Keywords: edible black ants, *Polyrhachis vicina*, nutritional composition

Introduction

Edible black ants (*Polyrhachis vicina* Roger) are widely distributed in China. However, they are cultivated mainly in Zhejiang and Guizhou Provinces in southern China. They have been made into various products such as wines, powders and capsules. Some of these products have been exported to South Korea, Japan, Thailand and other South-East Asian countries [1].

Globally, there are more than 500 edible insect species, the most popular being ants, grasshoppers, silk worm pupae, locusts, beetles, crickets and bamboo worms [2-3]. Edible black ants

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

have been used as human food and health food for hundreds of years in China [4]. They are an important cheap source of high-protein food in rural areas. In China, edible black ants have been extensively studied and shown to have many pharmacological properties such as anti-cancer, anti-fatigue and anti-inflammatory [5].

Sun drying is a common traditional method for drying fresh ants; it is still used in edible black ant processing industries in China. The fresh ants are put in water or in a freezer at -20°C and then were sun dried. While a few studies have investigated the nutrition composition of hot-air-dried black ants [6–7], for sun-dried ants which are commonly eaten, there is limited information on their nutritive value. In the present study, therefore, the determination of the nutritional composition of sun-dried edible black ants is carried out.

Materials and Methods

Sample preparation

The cultivated adult edible black ants (*Polyrhachis vicina*) were collected in Zhejiang and Guizhou Provinces, China, in June 2006. The fresh ants were stored at -20°C overnight, then sun dried for 3 days (10 h/day) and powdered with mortar and pestle until homogeneous.

Sampling

For each experiment, three samples were randomly collected from different batches of sun drying.

Proximate analysis

All solvents used were of analytical grade or HPLC grade and were purchased from Merck (Darmstadt, Germany) or Sigma-Aldrich (Steinheim, Germany). Crude protein, crude lipid, ash, crude fibre and moisture content were determined using the standard methods of the AOAC [8]. Crude protein was determined by the micro-Kjeldahl procedure; the factor N × 6.25 was used to convert nitrogen into crude protein. Results were calculated as per cent of dry weight (DW). Crude lipid content was obtained by the Soxhlet extraction method using petroleum ether as solvent. Ash content was obtained by heating the sample at 550°C to constant weight. Moisture content was determined by heating the sample to constant weight at 105°C. Crude fibre was determined gravimetrically after chemical digestion and solubilisation of other materials present. The fibre residual weight was then corrected for ash content after ignition. Carbohydrate content was estimated by subtracting the sum of the weights of crude protein, crude lipid, ash and crude fibre from the total dry matter.

Amino acid analysis

Determination of amino acid profile was conducted according to the standard AOAC methods [8]. The ground ant sample was hydrolysed in 6 N HCl at 110°C under nitrogen atmosphere for 24 h. The hydrolysed solution was used in the determination of free amino acids with a Hitachi L8800

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

(Hitachi, Japan) automatic amino acid analyser. The results were reported as milligrams of amino acid per 100 g of protein.

Fatty acid analysis

An oil sample for fatty acid analysis was obtained according to a previous study with modification [9]. Approximately 2 g of well-ground sun-dried ant sample were extracted with 20.0 mL of chloroform–methanol (2:1,v/v) containing 10 mg/L of butylated hydroxyl toluene (BHT) and 0.9 mg/mL of C19:0 (nonadecanoic acid) as internal standard. Then, the mixture was stored in the fume hood for 24 h, then filtered and transferred to a separating funnel and added with 0.9% sodium chloride (10 mL). After shaking, the phases were allowed to separate and the separated lower phase was then concentrated at 38°C and transferred to a 10-mL volumetric flask and made up to volume with chloroform containing 10 mg/L of BHT. The total oil content was determined by evaporating off the solvent at 38°C under N_2 to constant weight.

The fatty acid methyl esters (FAMEs) of the total lipid extract were prepared by transesterification using 0.9M H_2SO_4 in methanol [10]. Briefly, lipid solution (1 mL), 0.9M H_2SO_4 in methanol (3 mL) and toluene (1 mL) were added to a Teflon-capped tube and the mixture shaken strongly, then submerged in water bath at 70°C for 2 h. Then n-hexane (2 mL) and 0.9% sodium chloride (1 mL) were added to the tube and centrifuged at 1200 rpm for 15 min. The supernatant was then dropped into water (2 mL) and, after separating off the aqueous phase, dried with a little sodium sulphate anhydrous. The crude FAMEs solution obtained was filtered through a Sep-pak silica column (Alltech Associates, Inc., Deerfield, IL) before injection into the gas chromatograph, which is a Shimadzu GC-14C system equipped with a flame ionisation detector (Shimadzu Corp., Kyoto, Japan), a fused silica capillary column (DB-23, 60m×0.248mm×0.25μm: Agilent Technologies, Inc., Palo Alto, CA, USA) and the N2010 Chromatography Data System (Zhida Information Technologies, Inc., Hangzhou, China). Injection and detection temperature was 270°C and 270°C respectively. The column temperature was kept at 100°C for 3 min and programmed to 190°C at a rate of 20°C/min and kept at 190°C for 10 min. It was then increased to 205°C at a rate of 5 °C/min and kept at 205°C for 6 min. Finally, it was increased to 230°C at a rate of 10°C/min and kept at 230°C for 5 min. Individual fatty acids were identified by means of purified standards (Sigma-Aldrich, Deisenhofen, Germany) and quantified by means of the internal standard method [11].

Mineral analysis

Sun-dried ant samples were subjected to acid digestion and analysed according to the procedure described by Farías et al. [12] with modification. A sample of ant powder (2 g) was ashed in a muffle furnace for 4 h at 550°C. After cooling, HNO_3 (30 mL), H_2O_2 (2 mL) and $HClO_4$ (5 mL) were added. Then the mixture was boiled to near dryness. The residue was added with doubly distilled water (10 mL) and filtered. Then it was diluted with doubly distilled water to a volume of 100 mL and used for analysis. Mineralogical analysis was carried out using ICP emission spectroscopy (IRIS Intrepid II XSP,

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

Thermo Elemental Corporation, USA). All determinations were performed in triplicate and data represented on a dry weight basis as mean values ± standard deviation.

Statistical analysis

All experiments were performed in triplicate. The design was completely randomised [13]. The data were compared by one-way analysis of variance (ANOVA), and the mean values were compared by least squares difference method with SPSS 16.0 program. Significance was set as $p<0.05$.

Results and Discussion

Proximate composition

The proximate composition of the sun-dried edible black ants is shown in Table 1. Similar to previous report [7], both Zhejiang and Guizhou ants are low in moisture, moderate in lipid and fibre, and high in protein. However, the ants from the two locations are significantly different ($p<0.05$) with respect to the moisture, protein and carbohydrate content. Low moisture content means a good shelf life characteristic while the most abundant component is protein, indicating that these sun-dried ants are a good protein supplement. The protein content of the Guizhou ants is significantly higher than that of the Zhejiang ants ($p<0.05$), while the reverse is true for moisture content. It is possible that a slight difference in sun-drying conditions might have led to different dehydrating rates for the two ant samples, resulting in a difference in protein content [14].

High crude fibre in food is known to promote digestibility and enhance health benefits such as reduction of the risk of gastrointestinal cancers [15]. From Table 1, fibre is the second major component in both ant samples while lipid is the third most abundant component (about 15%). Hot-air-dried Hangzhou black ants contained 9.0% of lipid [7]. Thus, lipid loss may be more intensive in hot-air-dried ants than that in sun-dried ants. Lipid may be lost throngh moisture evaporation during hot-air drying and extensive heat treatment also appears to increase lipid loss phenomenon [16]. The quantity of ash in ant samples supposedly represents the amount of nutritionally important minerals [17]. From Table 1, the ash content is significantly higher in Guizhou ants than that in Zhejiang ants.

Table 1. Proximate composition of sun-dried edible black ants

Component (g/100g)	Zhejiang sun-dried ants	Guizhou sun-dried ants
Moisture	12.8 ± 0.1^{b}	8.6 ± 0.1^{a}
Crude protein	31.5 ± 0.5^{a}	41.5 ± 1.2^{b}
Crude fibre	25.4±0.9	26.4±1.4
Crude lipid	15.7±0.1	15.9±1.1
Crude carbohydrate	12.4 ± 1.1^{b}	3.8 ± 1.2^{a}
Ash	2.2 ± 0.1^{a}	3.8 ± 0.1^{b}

Note: Values are means ± standard deviations (n=3). Those having different superscript letters in a row differ significantly at $p<0.05$.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

Amino acid profile

The amino acid composition of the sun-dried ants is shown in Table 2. Eighteen amino acids were detected in both Zhejiang and Guizhou ants, similar to previous reports for hot-air-dried ants [6-7]. The major amino acids in both Zhejiang and Guizhou sun-dried ant proteins are glycine and glutamic acid, and the least abundant ones are cysteine and methionine. The proportion of a few amino acids (threonine, valine, lysine, phenylalanine, tryptophan, serine, glycine, arginine, proline and histidine) seems to be significantly influenced by the location of the ants. Ants need a nutritional source of arginine, histidine, leucine, isoleucine, lysine, methionine, phenylalanine, threonine, trythophan and valine, which is the same as for human young [18]. Total essential amino acid content in Guizhou ants is significantly higher than that in Zhejiang ants, which may be attributable to difference in amino acid content in the feed for the ants. Nevertheless, the amino acid profiles of the Zhejiang and Guizhou sun-dried ants are similar, their essential amino acid content being 34.6% and 36.1% respectively of the total protein, which is similar to that in egg (38.8% of total protein) [19] but substantially higher than that of Hangzhou hot-air-dried ants (18.1% of total protein) [7]. Also, the protein of the sun-dried ants has a higher proportion of methionine, valine, lysine, threonine, isoleucine, leucine, phenylalanine and tryptophan than that of the Hangzhou hot-air-dried ants. These results seem to imply that there is a protein and amino acid loss probably caused by the Maillard reactions during the hot-air drying process [20] whereas sun drying does not induce as much loss of amino acids in the ants.

Fatty acid composition

The sun-dried ants are rich in lipid. The composition and concentration of fatty acids in the sun-dried ant oil are presented in Tables 3-4. Twenty-two and 23 fatty acids are identified in Guizhou and Zhejiang ant oil respectively. Unsaturated fatty acids vary between 74.5-76.4% with monounsaturated fatty acids (MUFA) and polyunsaturated fatty acids (PUFA) ranging between 71.4-72.7% and 3.1-3.7% respectively. As a result, sun-dried ant samples could be a good source of unsaturated fatty acids. The 16:0 fatty acid is found as the major saturated fatty acid while the 18:1 acid is the main MUFA (60.5-63.0%).

The n-3 PUFA are known to prevent cardiovascular disease [21] and are anti-inflammatory [22]. Total PUFA content (n-3 + n-6) is higher in Guizhou than in Zhejiang ants. Since both groups of ants feed on similar diet (such as insects, larvae and algae), the difference, though small, in the proportions of their PUFA and total unsaturated fatty acids could possibly be attributed to the climate factor [23-25], the average annual temperature in Guizhou Province being lower than that in Zhejiang Province. The 18:3(n-3) is the main n-3 PUFA in both Zhejiang and Guizhou ants, amounting to 63.0 mg and 123.5 mg per 100 g of total fatty acids respectively. With the quantities of total n-3 PUFA being as they are (Tables 3-4), the Zhejiang and Guizhou sun-dried ants can be utilised as a source of these fatty acids. It has been recommended that human beings evolve on a diet with n-3/n-6 ratio of approximately 1 [26-27]. High ratio of n-3/n-6 is advantageous for reducing the risk of cardiovascular diseases [26-27]. In this study, the ratio of n-3/n-6 is found to be 0.41 and 0.42 for Zhejiang and Guizhou sun-dried

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

ants respectively. These ants are also a good source of unsaturated fatty acids which are beneficial to human health.

Table 2. Amino acid composition (mg/100g crude protein) of edible black ant samples

Amino acid	Zhejiang sun-dried ants	Guizhou sun-dried ants	Hangzhou hot-air-dried ants (Shen et al.[7])	Egg (Wang [19])
Essential amino acids				
Threonine	4016.7±1.3[a]	4263.3±1.3[b]	2260.0	4470.0
Valine	6883.3±1.5[a]	8273.3±1.4[b]	3430.0	5420.0
Methionine	1626.7±1.3	2326.7±1.1	1190.0	2810.0
Isoleucine	4650.0±2.1	4570.0±2.0	2260.0	4880.0
Leucine	6973.3±2.1	6980.0±2.0	3920.0	8110.0
Lysine	4370.0±1.7[b]	3736.7±1.4[a]	2200.0	6590.0
Phenylalanine	3340.0±1.9[a]	3496.7±1.1[b]	1760.0	4820.0
Tryptophan	2736.7±1.4[b]	2460.0±1.5[a]	1120.0	1720.0
Total essential amino acids	34596.7±1.6[a]	36106.7±2.2[b]	18140.0	38820.0
Non-essential amino acids				
Aspartic acid	7516.7±1.6	7756.7±1.8	5050.0	8920.0
Serine	4823.3±0.8[a]	5106.7±1.3[b]	2940.0	6720.0
Glutamic acid	11206.7±1.6	11570.0±1.4	7450.0	12130.0
Glycine	12860.0±1.3[b]	11596.7±2.1[a]	5690.0	3020.0
Alanine	9483.3±1.5	8843.3±1.0	4540.0	5030.0
Tyrosine	5340.0±1.7	5330.0±1.2	2820.0	3810.0
Arginine	3523.3±1.6[a]	3790.0±1.0[b]	2730.0	5700.0
Proline	7366.7±1.0[b]	6766.7±1.1[a]	2830.0	3380.0
Cysteine	696.7±1.5	730.0±1.5	3390.0	2090.0
Histidine	2580.0±1.2[b]	2396.7±2.0[a]	ND	1900.0
Total non-essential amino acids	65396.7±1.6[b]	63886.8±1.8[a]	37440.0	52700.0

Note: Values are means ± standard deviations (n=3). Values having different superscript letters in a row differ significantly at $p<0.05$.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

Table 3. Fatty acid composition (as % of total fatty acids) of sun-dried edible black ants

Fatty acid	Zhejiang ants	Guizhou ants
12:0	0.7±0.3	0.6±0.3
14:0	0.6±0.1	0.6±0.1
15:0	0.2±0.1	0.2±0.1
16:0	19.0±1.1	17.6±0.2
17:0	0.1±0.1	0.1±0.1
18:0	4.3±0.2	4.1±0.5
20:0	0.2±0.1	0.3±0.1
22:0	0.3±0.1	0.3±0.1
24:0	0.1±0.1	ND
Total SFA	25.5±1.0	23.8±1.8
14:1	0.1±0.1	0.1±0.1
15:1	0.2±0.1	0.1±0.1
16:1	8.9±0.2	8.2±0.2
17:1	0.5±0.4	0.3±0.1
18:1	60.5±0.8	63.0±0.3
20:1	0.1±0.1	0.1±0.1
22:1	1.1±0.2	0.9±0.1
Total MUFA	71.4±1.3	72.7±1.2
16:2 (n-6)	0.3±0.1	0.3±0.1
18:2 (n-6)	1.7±0.1	2.0±0.1
20:2 (n-6)	0.1±0.1	0.1±0.2
20:4 (n-6)	0.1±0.1	0.2±0.1
Total *n*-6	2.2±0.7	2.6±0.8
18:3 (n-3)	0.6±0.2	0.9±0.1
20:5 (n-3)	0.2±0.2	0.1±0.1
22:5 (n-3)	0.1±0.1	0.1±0.1
Total *n*-3	0.9±0.3	1.1±0.4
Total PUFA	3.1±0.6	3.7±0.7
PUFA:SFA	0.12	0.16
n-3/n-6	0.41	0.42

Note: Values are means ± standard deviations (n=3); ND = not detected, SFA = saturated fatty acids, MUFA = monounsaturated fatty acids, PUFA = polyunsaturated fatty acids

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

Table 4. Fatty acid concentration (mg/100 g) in sun-dried edible black ant oils

Fatty acid	Zhejiang sun-dried ant oil	Guizhou sun-dried ant oil
12:0	66.1±1.8	80.0±1.2
14:0	64.6±1.0	82.3±1.6
15:0	24.0±1.7	23.1±1.9
16:0	1935.7±1.4	2410.7±1.4
17:0	12.8±1.2	17.8±1.7
18:0	441.0±1.3	562.8±1.2
20:0	22.1±1.6	40.7±1.9
22:0	30.9±1.9	34.5±1.6
24:0	5.5±1.3	ND
Total SFA	2602.7±1.4	3251.9±1.2
14:1	11.8±1.8	8.2±1.1
15:1	16.5±1.1	7.3±0.6
16:1	898.9±1.6	1114.8±1.7
17:1	47.6±1.6	41.8±1.8
18:1	6118.9±1.6	8612.8±1.6
20:1	9.8±1.4	11.1±1.6
22:1	111.2±1.7	127.1±1.8
Total MUFA	7214.7±1.8	9923.1±1.3
16:2n-6	24.8±1.4	39.1±1.2
18:2n-6	169.0±1.3	276.8±1.4
20:2n-6	7.3±1.7	12.9±1.3
20:4n-6	14.3±1.4	21.2±1.5
Total *n*-6	215.4±1.9	350.0±1.2
18:3n-3	63.0±1.7	123.5±1.5
20:5n-3	17.6±1.5	13.9±1.7
22:5n-3	6.8±1.7	8.2±1.1
Total *n*-3	87.4±1.5	145.6±1.7
Total PUFA	302.8±1.2	495.6±1.7

Note: Values are means ± standard deviations (n=3); ND = not detected, SFA = saturated fatty acids, MUFA = monounsaturated fatty acids, PUFA = polyunsaturated fatty acids

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

Mineral composition

The mineral content of the sun-dried ants is shown in Table 5. The most abundant mineral in both sun-dried ants is phosphorus and the least abundant is selenium. Evidently, sun-dried ants are rich in phosphorus, iron and calcium. Iron content in Zhejiang ants is significantly higher than that in Guizhou ants while for calcium content the opposite is true. This may be ascribed partly to the difference in mineral content in their feeds. It is also noted that the relative proportion of the major minerals found in this investigation is substantially different from that obtained for oven-dried ants [7]. It can be concluded, however, that sun-dried ants are a good source of many important minerals such as phosphorus, calcium, iron and selenium. Iron can be antioxidants [28] while magnesium and zinc can prevent cardiomyopathy and growth retardation [29]. Selenium is an essential component in several major metabolic pathways including thyroid hormone metabolism, and also has other important biological functions such as antioxidative and immune regulating functions [30] and anti-cancer activities [31].

Table 5. Mineral content (mg/100g of sun-dried ants)

Mineral	Zhejiang ants	Guizhou ants	Hangzhou hot-air-dried ants [7]
Phosphorus	387.7±1.7	417.0±1.0	158.0
Iron	118.0±1.1^{b}	53.7±1.0^{a}	94.1
Calcium	49.1±0.7	108.0±0.4	175.4
Magnesium	65.3±1.3	67.6±1.0	103.1
Manganese	25.9±1.0^{a}	32.3±0.5^{b}	21.0
Zinc	17.6±1.1^{b}	11.9±0.1^{a}	22.7
Copper	2.4±0.5	1.9±0.1	2.4
Lead	0.4±0.1^{a}	0.6±0.1^{b}	0.1
Nickel	<2.0±0.1	<2.0±0.1	0.7
Chromium	<2.0±0.1	<2.0±0.1	1.7
Selenium	2.9x10^{-2}±0.1^{a}	3.8×10^{-2}±0.1^{b}	0.1

Note: Values are means ± standard deviations (n=3). Values having different superscript letters in a row differ significantly at $p<0.05$.

Conclusions

The present study has shown that Zhejiang and Guizhou sun-dried edible black ants are a rich source of protein, lipid, fibre and essential amino acids. In addition, they contain essential fatty acids such as 18:3 (n-3), 20:5 (n-3) and 22:5 (n-3) and are also rich in minerals such as phosphorus, iron, calcium, magnesium, zinc and selenium. Thus, they seem to have a potential in being used as a nutraceutical or alternative nutritional food source for humans.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

Acknowledgements

The authors acknowledge support from Rajamangala University of Technology Isan, Kalasin Campus, Thailand in the form of a research scholarship. The authors also thank the Zhejiang Shikang Biotech Co., TCD, Wenzhou and the Hangzhou Tianma Exploration Company of Animals and Plants, Hangzhou, Zhejiang Province, China for supplying the ant samples used in this study.

References

1. L. R. Shen and Y. C. Ren, "Advances on exploration and utilization of the edible and medical ants in China", *Entomol. Knowledge*, **1999**, *36*, 251–254.
2. United Nations, **"**Edible insects provide food for thought**"**, http://www.scoop.co.nz/stories/WO0802/S00211.htm (Accessed: 1 December 2008)
3. G. R. DeFoliart, "Insects as food: Why the western attitude is important", *Annu. Rev. Entomol.*, **1999**, *44*, 21–50.
4. Z. C. Wu, "Ant and ant therapy", Shanghai Publishing House of Scientific Popularizing, Shanghai (China), **1994**.
5. J. Kou, Y. Ni, N. Li, J. Wang, L. Liu and Z-H. Jiang, "Analgesic and anti-inflammatory activities of total extract and individual fractions of Chinese medicinal ants *Polyrhachis lamellidens*", *Biol. Pharm. Bull*., **2005**, *28*, 176–180.
6. S. L. Li, Y. J. Ren, X. Cui, Z. S. Qian and Q. Dong, "Analysis of nutritional components in *Formica sanguinea* Latr. and *Polyrhachis vicina* Roger", *J. Chinese Pharm*., **1994**, *29*, 521–522.
7. L. R. Shen, D. Li, F. Q. Feng and Y. Ren, "Nutritional composition of *Polyrhachis vicina* Roger (Edible Chinese black ant)", *Songklanakarin J. Sci. Technol*., **2006**, *28* (Suppl.1), 107–114.
8. P. Cunniff (Ed.), "Official Methods of Analysis of AOAC International", 16th Edn., AOAC International, Arlington V.A., 1**995**.
9. L-F. Yang, S. Siriamornpun and D. Li, "Polyunsaturated fatty acid content of edible insects in Thailand", *J. Food Lipids,* **2006,** *13*, 277–285.
10. H. Li, F-Q. Feng, L-R. Shen, Y. Xie and D. Li, "Nutritional evaluation of different bacterial douchi", *Asia Pac. J. Nutr.*, **2007**, *16 (Suppl 1)*, 215–221.
11. H. T. Slover and E. Lanza, "Quantitative analysis of food fatty acids by capillary gas chromatography". *J. Am. Oil Chem. Soc.*, **1979**, *56*, 933–943.
12. S. Farías, S. P. Arisnabarreata, C. Vodopivez and P. Smichowski, "Levels of essential and potential toxic metals in Antarctic macro algae", *Spectrochim. Acta Part B*, **2002**, *57*, 2133–2140.
13. W. G. Cochran and G. M. Cox, "Experimental Designs", John Wiley and Sons Inc, New York, **1975**, pp. 95–100.
14. L. Fachruddin, "A variety of beef jerky making", Kanisius, Yogyakarta, **1997**. (In Indonesian)

15. J. Salvin, D. R. Jacobs and L. Marquart, "Whole grain consumption and chronic disease: Protective mechanisms", *Nutr. Cancer*, **1997**, *27,* 14–21.
16. T. Wu and L. Mao, "Influences of hot air drying and microwave drying on nutritional and odorous properties of grass carp (*Ctenopharyngodon idellus*) fillets", *Food Chem*., **2008**, *110*, 647–653.
17. V. Vadivel and K. Janardhanan, "The nutritional and anti-nutritional attributes of sword bean [Canavalia gladiata (Jacq) DC.]: An underutilized tribal pulse from south India", *Int. J. Food Sci. Tech.*, **2004**, *39,* 917–926.
18. H. Genç, "General principles of insect nutritional ecology", *Trakya Univ. J. Sci.*, **2006**, *7*, 53–57.
19. G. Y. Wang, "Table of Food Ingredients (Representative Values in China)", People's Sanitation Press, Beijing, **1991**, pp.78–79.
20. M. T. García-Arias, E. Álvarez-Pontes, M. C. García-Linares, M. C. García-Fernández and F. J. Sánchez-Muniz, "Grilling of sardine fillets. Effects of frozen and thawed modality on their protein quality", *LWT-Food Sci. Technol.*, **2003**, *36*, 763–769.
21. F. B. Hu, L. Bronner, W. C. Willett, M. J. Stampfer, K. M. Rexrode, C. M. Albert, D. Hunter and J. E. Manson, "Fish and omega-3 fatty acid intake and risk of coronary heart disease in women", *J. Am. Med. Assoc.*, **2002**, *287*, 1815–1821.
22. A. P. Simopoulos, "Omega-3 fatty acids in inflammation and autoimmune diseases", *J. Am. Coll. Nutr.*, **2002**, *21*, 495–505.
23. K. Uysal, M. Bülbül, M. Dönmez and A. K. Seçkin, "Changes in some components of the muscle lipids of three fresh water fish species under natural extreme cold and temperate conditions", *Fish Physiol. Biochem.*, **2008**, *34*, 455–463.
24. S. I. Kiyashko, A. B. Imbs, T. Narita, V. I. Svetashev and E. Wada, "Fatty acid composition of aquatic insect larvae *Stictochironomus pictulus* (Diptera: Chironomidae): Evidence of feeding upon methanotrophic bacteria", *Comp. Biochem. Physiol. B Biochem. Mol. Biol.*, **2004,** *139*, 705–711.
25. G. V. Skuladottir, H. B. Schioth, E. Gudmundsdottir, B. Richards and L. Jonsson, "Fatty acid composition of muscle heard and liver lipids in Atlantic Salmon, *Salmo Salar*, at extremely low environmental temperature", *Aquaculture,* **1990**, *84*, 71–80.
26. P. M. Herold and J. E. Kinsella, "Fish oil consumption and decreased risk of cardiovascular disease: A comparison of findings from animal and human feeding trials", *Am. J. Clin. Nutr*., **1986**, *43*, 566–598.
27. A. P. Simopoulos, "The importance of the ratio of omega-6/omega-3 essential fatty acids", *Biomed. Pharmacother.*, **2002**, *56*, 356–379.
28. J. L. Buss, F. M. Torti and S. V. Torti, "The role of iron chelation in cancer therapy", *Curr. Med. Chem*., **2003**, *10,* 1021–1034.
29. V. C. Chaturvedi, R. Shrivastava and R. K. Upreti, "Viral infections and trace elements: A complex interaction", *Curr. Sci.*, **2004**, *87*, 1536–1554.
30. K. M. Brown and J. R. Arthur, "Selenium, selenoproteins and human health", *Publ. Health Nutr.*, **2001**, *4*, 593–599.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 101-112

31. M. P. Rayman, "Selenium in cancer prevention: A review of the evidence and mechanism of action", *Proc. Nutr. Soc.*, **2005**, *64*, 527–542.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

Maejo International Journal of Science and Technology

ISSN 1905-7873

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

Full Paper

Selection of efficient wavelengths in NIR spectrum for determination of dry matter in kiwi fruit

Lü Qiang[1], Tang Mingjie[1], Cai Jianrong[1,*], Lu Huazhu[1] and Sumpun Chaitep[2]

[1] School of Food and Biological Engineering, Jiangsu University, Zhenjiang, Jiangsu, 212013, China

[2] Department of Mechanical Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, 50200, Thailand

* Corresponding author, e-mail: lvqiang1111@gmail.com

Received: 19 November 2009 / Accepted: 9 March 2010 / Published: 2 April 2010

Abstract: The feasibility of using efficient wavelengths in the near-infrared (NIR) spectrum for the rapid determination of the dry matter (DM) in kiwi fruit was investigated. Partial least squares (PLS), synergy interval PLS (siPLS) and genetic algorithm siPLS (GA-siPLS) were comparatively performed to calibrate regression models. The number of wavelengths and the number of PLS components were optimised as per the root mean square error of cross-validation (RMSECV) in the calibration set. The performance of the final model was evaluated by the root mean square error of prediction (RMSEP) and the correlation coefficient (r) in the prediction set. Results indicate that the performance of GA-siPLS model is the best one compared to PLS and siPLS models. The optimal model was achieved with r = 0.9020 and RMSEP = 0.5315 in the prediction set. This work shows that it is feasible to determine DM in kiwi fruit using NIR spectroscopy and that GA-siPLS algorithm is most suitable in solving the problem of selection of efficient wavelengths.

Keywords: kiwi fruit, dry matter, NIR spectroscopy, partial least squares (PLS), synergy interval partial least squares (siPLS), genetic algorithm siPLS.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

Introduction

Kiwi fruit are harvested unripe though physiologically mature but must be left in natural storage to ripen before consumption [1]. Timing of the harvest has a decisive effect on the subsequent postharvest shelf life and fruit quality [2-3]. The dry matter (DM) in kiwi fruit has been proposed as a maturity indicator for the proper time of harvest and also as a predictor of the sensory quality of the fruit once it is ripe [4-6].

Near infrared (NIR) spectroscopy is a fast, accurate and non-destructive technique that can be deployed as a replacement of individuals' labour skills and time-consuming methods. The NIR spectroscopy has been used to grade fruits [7-8], predict fruit maturity [9] and indicate optimal harvesting time [10]. Kiwi fruit are a commodity the sorting of which, based on pre-selected NIR spectral features, can be used to grade them at harvest on the basis of DM. Recent research has established that NIR spectroscopic analysis can be used to assess kiwifruit DM and/or soluble-solid content of the ripe fruit [1, 3, 6, 11-12].

In addition to these, NIR spectral data calibrations have been made with the classical multivariate calibration analysis, e.g. partial least squares (PLS) regression [13-14]. Many spectral pretreatment methods have been developed to reduce the effects of variations in the spectral data that are not related to the chemical variations in the samples [15-16]. These methods generally improve the calibrations. However, they did not take into account that there might be spectral regions that do not contain any information about the chemical variations in the samples [17]. In fact, one of the major tasks in multivariate data analysis is to select appropriate spectral regions in order to achieve the best performance. A number of researchers have constructed PLS models in different spectral regions to quantify ingredient content in kiwi fruit. However, these regions were selected manually [2-3]. Without prior detailed knowledge about NIR spectroscopy, spectral regions selected manually might as well weaken the performance of the calibration model.

According to some other researchers, both theoretical and experimental evidence has been published to the effect that spectral region selection can significantly improve the performance of these calibration techniques [18-19]. It is important to select specific regions that contain much information based on which of the more stable models can be generated with superior interpretability and lower prediction error. Methods [e.g.19] have been recently described in the literature in implementing spectral region selection and PLS used for multivariate calibration in each subset.

A graphically oriented local modelling procedure called interval partial least squares (iPLS) has been presented for use on NIR spectral data. It has been shown that selective optimum interval in the spectral data can yield precision prediction models [19-20]. A method called synergy interval partial least squares (siPLS) has also been proposed to be used to select several interval spectral data which can split the data set into a number of intervals (variable-wise) and to calculate all possible PLS model combinations of two, three or four intervals [17]. Genetic algorithm (GA) has already been used in variable selection problem and seems to be a solution to the multivariate selection of variables [21-22].

This study investigates and compares the results provided by PLS, siPLS and GA-siPLS procedures for NIR quantitative analysis of DM in kiwi fruit. Two specific objectives of this research

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

are: (1) to establish relationships between the NIR measurements and the DM of kiwi fruit based on the new method, and (2) to compare the prediction performance of calibration models at different wavelengths and then find out the optimal wavelengths and develop the best calibration models.

Materials and Methods

Sample preparation

One hundred and twelve "*Zhonghua*" kiwi fruit samples, purchased from a farm in Zhouzhi, Shanxi Province, China, were used in this study. All sizes of the fruit from peewee to jumbo were used. However, the fruit with irregular shape were not incorporated in the data analysis. The fruit were sent to our laboratory in October 2008, then were stored for one month. Experiments were done under controlled condition (20^{o}C). Before being examined by NIR technique, the fruit were acclimatised to equilibrium for 12 h in the controlled condition.

Collection of spectra

The NIR spectra were measured in the reflectance mode using the FT-NIR spectrophotometer (Antaris^{TM} Analyser, Thermo Electron Co., USA) with an integrating sphere. Each spectrum was obtained from an average of 32 scans. The range of spectrum was 10,000-4,000 cm^{-1} and the data were collected in 1.928 cm^{-1} intervals, which resulted in 3,112 variables. Each kiwi fruit was measured three times around equatorial locations. The average of the three spectra, which were measured at the equator of each kiwifruit, was used in the sequence analysis.

Measurement of kiwi fruit reference DM

The fruit DM was determined by cutting two equatorial slices of approximately 3-mm thickness each, and drying them at 65^{o}C to constant weight (approximately 24 h). The fruit DM was calculated from the final dry weight and the initial wet weight of the slices, recorded as a percentage of fresh weight.

Software

All algorithms were implemented in Matlab V7.0 (Mathworks, USA) under Windows XP. Result Software (Antaris System, Thermo Electron Co., USA) was used in NIR spectral data acquisition. The iPLS, siPLS and GAPLS algorithms used in this work were downloaded from http://www. models.kvl.dk/.

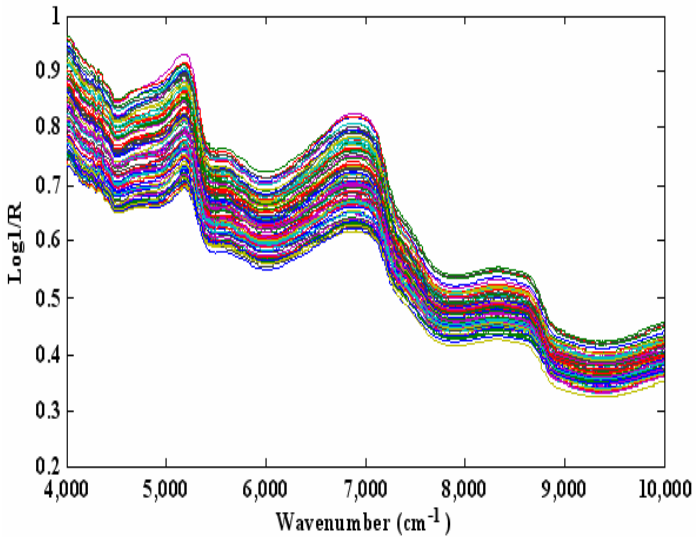
Results and Discussion

Spectral pre-processing

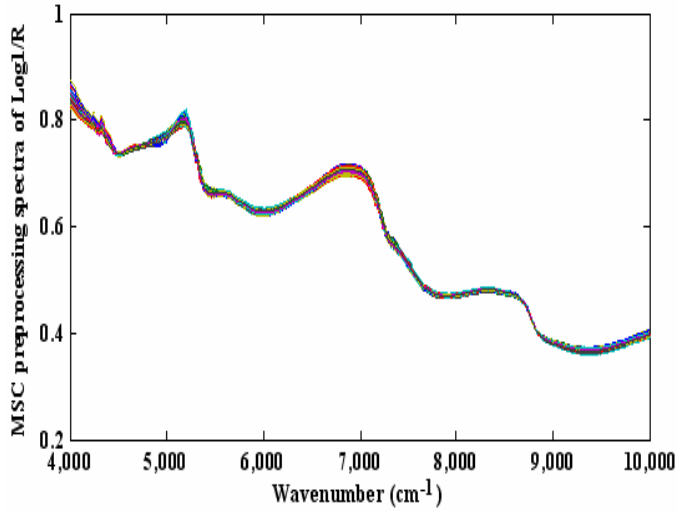
Figure 1(a) presents the raw spectral profile of the kiwi fruit, the raw spectral data being conducted on spectral pre-processing. Each mean spectrum was recorded as log(1/R), where R is the reflectance. In this research, the spectral data were analysed with multiplicative scatter correction (MSC) pre-processing technique because MSC is an important procedure for the correction of

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

scattered light, and the technique is often used to correct for additive and multiplicative effects in the spectra [23]. The spectra after MSC pre-processing are presented in Figure 1(b).



(a)



(b)

Figure 1. Spectra of kiwi fruit obtained from (a) raw data and (b) MSC pre-processed data

Calibration of models

All 112 samples were divided into two subsets. The first one was the calibration set, which was used to build the models, whereas the other was the prediction set, which was used to test the robustness of the established models. To avoid bias in the subset division, it was made by sorting all samples according to their respective y-value (viz. the reference measurement value of dry matter). In order to achieve a 2/1 division of calibration/prediction spectra, one spectrum of every three samples was assigned to the prediction set so that finally the calibration set contained 74 spectra and the remaining 38 spectra constituted the prediction set. Seen from Table 1 is the range of y-value in

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

the calibration set that covers the range in the prediction set. Therefore, the distribution of the samples was appropriate in both the calibration and prediction sets.

Table 1. Reference measurements of DM and sample numbers in calibration and prediction sets

Set	Unit	Number of samples	Mean value	Range	Standard deviation	CV /%
Calibration set	% (g/g)	74	16.1736	13.526-18.757	1.2554	7.7622
Prediction set	% (g/g)	38	16.2237	13.758-18.584	1.2202	7.5210

Note: CV = coefficient of variation

The performance of the final PLS model was evaluated in terms of the root mean square error of cross-validation (RMSECV), the root mean square error of prediction (RMSEP), and the correlation coefficient (r). For RMSECV, a leave-one-sample-out cross-validation was performed: the spectrum of one sample of the training set was deleted from this set and a PLS model was built with the remaining spectra of the calibration set. The left-out sample was predicted with this model and the procedure was repeated by leaving out each of the samples of the calibration set. The RMSECV was calculated by Eq. 1 [24]:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{I_c}(\hat{y}_i - y_i)^2}{I_c - 1}} \quad (1)$$

where $\hat{y}_i$ is the predicted value of the i th observation, y_i the measured value of i th observation and I_c the number of observation in the calibration set. The number of PLS factors included in the model was chosen according to the lowest RMSECV. This procedure was repeated for each of the pre-processed spectra.

For the prediction set, the RMSEP was calculated by Eq. 2 [24]:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{I_p}(y_i - \tilde{y}_i)^2}{I_p}} \quad (2)$$

where $\tilde{y}_i$ is the predicted value for sample i of the prediction set, y_i is the measured value for sample i of the prediction set, and I_p is the number of observation in the prediction set. The correlation coefficients (r) between the predicted and measured values were calculated by Eq. 3 [24] for both the calibration and prediction sets:

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

$$r = \sqrt{1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}} \quad (3)$$

where $\hat{y}_i$ and y_i are the predicted and measured values respectively of sample i in calibration or prediction set, $\bar{y}_i$ is the mean of the reference measurement results for all samples in the calibration or prediction set, and n is the number of observation in the calibration or prediction set.

To verify the superior capability of the PLS calibration models based on the selected region by different methods, each calibration model mentioned above was used to predict the calibration data set and the prediction data set. The RMSECV, RMSEP and correlation coefficients of each model for the calibration data set (r_c) and validation data set (r_p) were taken into account.

Results of PLS model

In the application of PLS algorithm, it is generally known that the number of PLS components is a critical parameter in calibrating the model. The optimum number of PLS components is determined by the lowest RMSECV, which is 0.5513 when 12 PLS components are included in the calibration model. Therefore the optimal number of PLS components is 12.

In the optimal model, RMSECV is 0.5513 and correlation coefficient (r) is 0.8913 in calibration set. When the performance of PLS model is evaluated by the samples in the prediction set, RMSEP is 0.5926 and correlation coefficient (r) is 0.8806. Figure 2 represents the scatter plot showing a correlation between reference and NIR-predicted DM in the prediction set by PLS model.

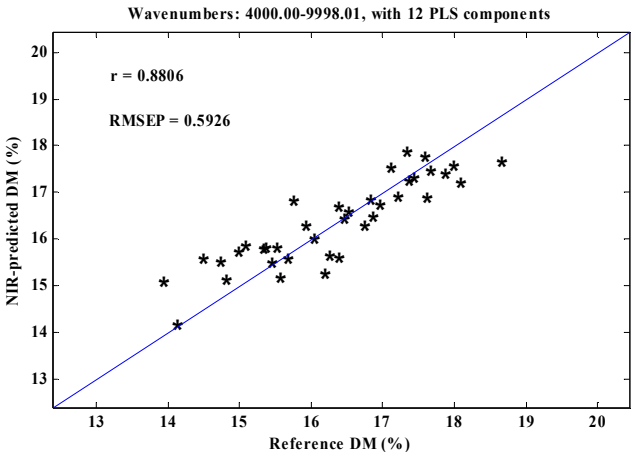


Figure 2. Reference versus NIR-predicted DM by PLS in prediction set

Results of siPLS model

The synergy interval PLS (siPLS) algorithm used here has been developed by Nørgaard et al. [19]. First, the data set is split into a number of intervals (variable-wise). Next, PLS regression models are established for all possible combinations of two, three or four intervals. Thereafter, RMSECV is calculated for every combination of intervals. The combination of intervals with the lowest RMSECV is then chosen.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

The number of intervals is also optimised according to RMSECV in siPLS model calibration. Table 2 shows the results of siPLS model calibration when splitting the spectra into different numbers of intervals. The optimal siPLS model is obtained with 15 intervals and 10 PLS components, the lowest RMSECV being 0.5139. The optimal combination of intervals selected is 3, 4, 8 and 12. It corresponds to 4,802.04-5,201.14, 5,203.07-5,602.16, 6,807.16-7,204.33 and 8,403.55-8,800.71 cm^{-1} in the spectral regions as shown in Figure 3.

For the optimal model, RMSECV is 0.5139, and correlation coefficient (r) is 0.9062 in the calibration set. When the performance of siPLS model is evaluated by the samples in the prediction set, RMSEP is 0.5710 and correlation coefficient (r) is 0.8903. Figure 4 represents the scatter plot showing a correlation between reference and NIR-predicted DM in the prediction set by siPLS model.

Table 2. Results of siPLS model calibration for different spectral regions

Number of intervals	No.of PLS components	Selected intervals	Calibration set		Prediction set	
			r	RMSECV	r	RMSEP
13	9	[3 7 10 12]	0.9115	0.4992	0.8829	0.5822
14	9	[3 7 11 13]	0.9247	0.4625	0.8862	0.5800
15	**10**	**[3 4 8 12]**	**0.9062**	**0.5139**	**0.8903**	**0.5710**
16	7	[1 9 13]	0.9014	0.5254	0.8348	0.6693
17	13	[2 4 7 13]	0.9056	0.5142	0.8592	0.6213
18	6	[3 9 14]	0.8896	0.5538	0.8352	0.6817
19	8	[3 4 7 15]	0.9283	0.4499	0.8695	0.6086
20	6	[1 11 16]	0.9032	0.5196	0.8548	0.6370

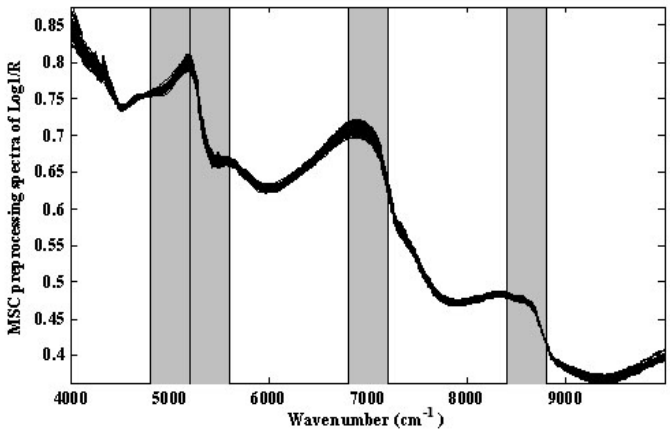


Figure 3. Optimal spectral regions selected by siPLS with wavenumbers of 4,802.04-5,201.14, 5,203.07-5,602.16, 6,807.16-7,204.33 and 8,403.55-8,800.71 cm^{-1}

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

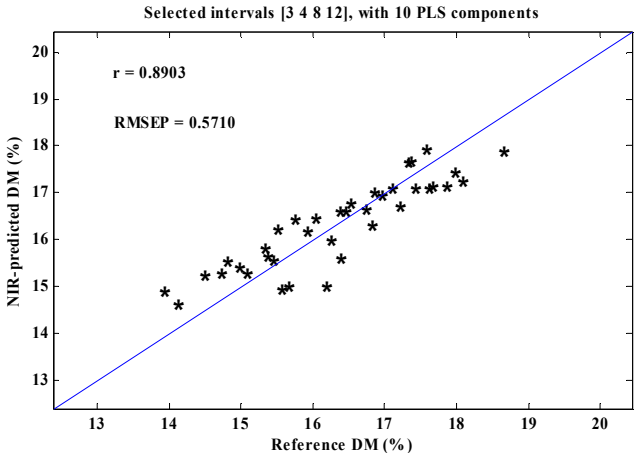


Figure 4. Reference versus NIR-predicted DM by siPLS in prediction set

Results of GA-siPLS model

GA is an optimisation method based on the principles of genetics and natural selection. This algorithm is inspired by the theory of evolution. In a living environment, the 'best' individuals have a greater chance to survive and a greater probability to spread their genomes by reproduction. The mating of two 'good' individuals causes the mixing of their genomes, which may result in a 'better' offspring. The terms 'good', 'better' and 'best' are related to the fitness of the individuals to their environment [21-22].

The number of wavelengths is also similarly optimised by RMSECV using GA in the optimal combination of intervals (4,802.04-5,201.14, 5,203.07-5,602.16, 6,807.16-7,204.33 and 8,403.55-8,800.71 cm^{-1}) selected by siPLS model. The optimal parameters are set as follows: number of generations = 100, population size = 30, mutation probability = 0.1, and recombination probability = 0.8. Figure 5 shows the selected frequency versus DM variable in the first spectral region (4,802.04-5,201.14 cm^{-1}) of the optimal combination of intervals. Wavelength variables are individually added to PLS model in accordance with the selected frequency. The best number of wavelength variables is then identified according to the RMSECV of the model. The optimal GA-siPLS model is obtained with 229 wavelengths and 9 PLS components when the lowest RMSECV is 0.4724.

In the optimal model, RMSECV is 0.4724 and correlation coefficient (r) is 0.9209 in the calibration set. When the performance of GA-siPLS model is evaluated by the samples in the prediction set, RMSEP is 0.5315 and correlation coefficient (r) is 0.9020 in the prediction set. Figure 6 represents the scatter plot showing a correlation between reference and NIR-predicted DM in the prediction set by GA-siPLS model.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

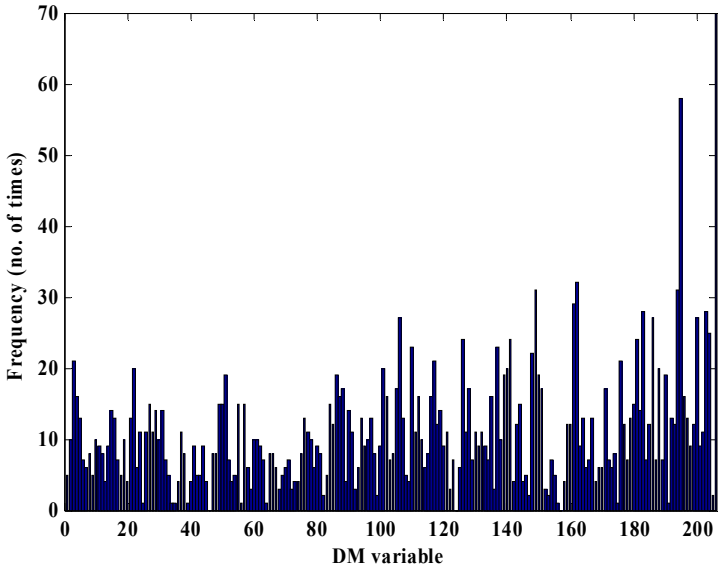


Figure 5. Selected frequency versus DM variable

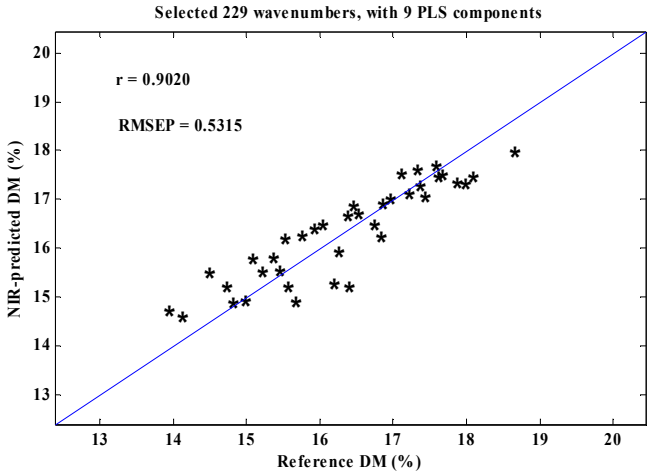


Figure 6. Reference versus NIR-predicted DM by GA-siPLS in prediction set

Table 3 shows results from different PLS models. Comparing among these models, one can see that GA-siPLS seems to be the best one followed by siPLS. Such phenomena can be explained by the following: (1) PLS is performed on full spectral range (4,000.00-10,000 cm^{-1}) to calibrate the global model. Thus, some noisy spectral information has inevitably weakened the performance of the model; (2) siPLS overcomes the disadvantages of PLS since siPLS combines with two, three or four intervals to calibrate the PLS model so as to remove some noisy regions and obtain useful information in the calibrated model; and (3) in contrast with siPLS, GA-siPLS selects interesting variable wavelengths and removes noisier spectral information based on siPLS.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

Table 3. Results from different PLS models

Model	Number of variables	PLS components	Calibration set		Prediction set	
			r	RMSECV	r	RMSEP
PLS	3112	12	0.8913	0.5513	0.8806	0.5926
siPLS	830	10	0.9062	0.5139	0.8903	0.5710
GA-siPLS	229	9	0.9209	0.4724	0.9020	0.5315

Conclusions

In the present study, it has been demonstrated that NIR spectroscopy is a suitable tool for quantification of dry matter in kiwi fruit with small prediction errors over the entire range studied. Three models were studied. The PLS model was performed on full spectral region (4,000.00-10,000 cm^{-1}, 3112 variables) to calibrate the model. It requires a large number of variables and some noisy spectral information has inevitably reduced the prediction accuracy of the model. The siPLS model was performed on four intervals of the spectral region (4,802.04-5,201.14, 5,203.07-5,602.16, 6,807.16-7,204.33 and 8,403.55-8,800.71 cm^{-1}, 830 variables) to calibrate the model. Some noisy regions were removed so as to reduce variables and improve prediction accuracy. The GA-siPLS model was performed on the most informative wavelengths (4,802.04-5,201.14 cm^{-1}, 229 variables) to calibrate the model. Compared with PLS and siPLS models, the GA-siPLS model requires fewer variables and improves prediction accuracy by removal of more spectral noises.

Acknowledgements

This work was financially supported by the National High Technology Research and Development Program of China (Project No. 2006AA10Z263) and the Key Natural Science Foundation of Jiangsu Province (Grant No. BK2006707-1). We are grateful to the website http://www.models.kvl.dk/ for their generous contribution in permitting the download of software for iPLS, siPLS and GAPLS free of charge.

References

1. V. A. McGlone, R. B. Jordan, R. Seelye and P. J. Martinsen, "Comparing density and NIR methods for measurement of kiwi fruit dry matter and soluble solids content", *Postharvest Biol. Technol.*, **2002**, *26*, 191-198.
2. D. C. Slaughter and C. H. Crisosto, "Nondestructive internal quality assessment of kiwi fruit using near-infrared spectroscopy", *Semin. Food Anal.*, **1998**, *3*, 131-140.
3. V. A. McGlone, C. J. Clark and R. B. Jordan, "Comparing density and VNIR methods for predicting quality parameters of yellow-fleshed kiwi fruit (Actinidia chinensis)", *Postharvest Biol. Technol.*, **2007**, *46*, 1-9.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

4. J. Burdon, D. McLeod, N. Lallu, J. Gamble, M. Petley and A. Gunson, "Consumer evaluation of "Hayward" kiwi fruit of different at-harvest dry matter contents", *Postharvest Biol. Technol.*, **2004**, *34*, 245-255.
5. R. B. Jordan, E. F. Walton, K. U. Klages and R. J. Seelye, "Postharvest fruit density as an indicator of dry matter and ripened soluble solids of kiwi fruit", *Postharvest Biol. Technol.*, **2000**, *20*, 163-173.
6. V. A. McGlone and S. Kawano, "Firmness, dry-matter and soluble-solids assessment of postharvest kiwi fruit by NIR spectroscopy", *Postharvest Biol. Technol.*, **1998**, *13*, 131-141.
7. I. W. Budiastra, Y. Ikeda and T. Nishizu, "Prediction of individual sugars and malic acid concentrations of apples and mangoes by the developed NIR reflectance system", *J. Jpn. Soc. Agric. Machine.*, **1998**, *60*, 117-127.
8. V. Steinmetz, J. M. Roger, E. Molto and J. Blasco, "On-line fusion of colour camera and spectrophotometer for sugar content prediction of apples", *J. Agric. Eng. Res.*, **1999**, *73*, 207-216.
9. A. Peirs, N. Scheerlinck and B. M. Nicolaï, "Temperature compensation for near infrared reflectance measurement of apple fruit soluble solids contents", *Postharvest Biol. Technol.*, **2003**, *30*, 233-248.
10. B. Park, J. A. Abbott, K. J. Lee, C. H. Choi and K. H. Choi, "Near-infrared diffuse reflectance for quantitative and qualitative measurement of soluble solids and firmness of delicious and dala apples", *Trans. Am. Soc. Agric. Eng.*, **2003**, *46*, 1721-1731.
11. C. J. Clark, V. A. McGlone, H. N. De Silva, M. A. Manning, J. Burdon and A. D. Mowat, "Prediction of storage disorders of kiwi fruit (Actinidia chinensis) based on visible-NIR spectral characteristics at harvest", *Postharvest Biol. Technol.*, **2004**, *32*, 147-158.
12. P. N. Schaare and D. G. Fraser, "Comparison of reflectance, interactance and transmission modes of visible-near infrared spectroscopy for measuring internal properties of kiwi fruit (Actinidia chinensis)", *Postharvest Biol. Technol.*, **2000**, *20*, 175-184.
13. J. Lammertyn, A. Peirs, J. D. Baerdemaeker and B. Nicolaï, "Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment", *Postharvest Biol. Technol.*, **2000**, *18*, 121-132.
14. A. Peirs, J. Lammertyn, K. Ooms and B. M. Nicolaï, "Prediction of the optimal picking date of different apple cultivars by means of VIS/NIR-spectroscopy", *Postharvest Biol. Technol.*, **2001**, *21*, 189-199.
15. Q. S. Chen, J. W. Zhao, X. Y. Huang, H. D. Zhang and M. H. Liu, "Simultaneous determination of total polyphenols and caffeine contents of green tea by near-infrared reflectance spectroscopy", *Microchem. J.*, **2006**, *83*, 42-47.
16. Q. S. Chen, J. W. Zhao, H. D. Zhang and X. Y. Wang, "Feasibility study on qualitative and quantitative analysis in tea by near infrared spectroscopy with multivariate calibration", *Anal. Chim. Acta.*, **2006**, *572*, 77-84.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 113-124

17. C. Abrahamsson, J. Johansson, A. Sparén and F. Lindgren, "Comparison of different variable selection methods conducted on NIR transmission measurements on intact tablets", *Chemom. Intell. Lab. Syst.*, **2003**, *69*, 3-12.
18. O. Kleynen, V. Leemans and M. F. Destain, "Selection of the most efficient wavelength bands for 'Jonagold' apple sorting", *Postharvest Biol. Technol.*, **2003**, *30*, 221-232.
19. L. Nørgaard, A. Saudland, J. Wagner, J. P. Nielsen, L. Munck and S. B. Engelsen, "Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy", *Appl. Spectrosc.*, **2000**, *54*, 413-419.
20. R. Leardi and L. Nørgaard, "Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions", *J. Chemom.*, **2004**, *18*, 486-497.
21. R. Leardi and A. L. González, "Genetic algorithms applied to feature selection in PLS regression: How and when to use them", *Chemom. Intell. Lab. Syst.*, **1998**, *41*, 195-207.
22. R. Leardi, "Application of genetic algorithm-PLS for feature selection in spectral data sets", *J. Chemom.*, **2000**, *14*, 643-655.
23. X. L. Chu, H. F. Yuan and W. Z. Lu, "Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique", *Prog. Chem.*, **2004**, *4*, 528-542.
24. X. B. Zou, J. W. Zhao and Y. X. Li, "Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of 'Fuji' apple based on BiPLS and FiPLS models", *Vib. Spectrosc.*, **2007**, *44*, 220-227.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 125-135

Maejo International Journal of Science and Technology

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

Full Paper

Field programmable gate array (FPGA) implementation of novel complex PN-code-generator- based data scrambler and descrambler

Ghulam M. Bhat [1], Muhammad Mustafa [2], Shabir A. Parah [2,*], Javaid Ahmad [2]

[1] University Science Instrumentation Centre (USIC), University of Kashmir, Srinagar-190006, India
[2] Department of Electronics and Instrumentation Technology, University of Kashmir Srinagar-190006, India

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

Received: 4 August 2009 / Accepted: 17 March 2010 / Published: 7 April 2010

Abstract: A novel technique for the generation of complex and lengthy code sequences using low-length linear feedback shift registers (LFSRs) for data scrambling and descrambling is proposed. The scheme has been implemented using VHSIC hardware description language (VHDL) approach which allows the reconfigurability of the proposed system such that the length of the generated sequences can be changed as per the security requirements. In the present design consideration the power consumption and chip area requirements are small and the operating speed is high compared to conventional discrete I.C. design, which is a pre-requisite for any system designer. The design has been synthesised on device EP2S15F484C3 of Straitx II FPGA family, using Quarts Altera version 8.1. The simulation results have been found satisfactory and are in conformity with the theoretical observations.

Keywords: complex code generator, VHDL, FPGA, scrambler, descrambler

Introduction

Since the birth of the cellular industry, security has been a major concern for both service providers and subscribers. Service providers are primarily concerned with preventing fraudulent operations such as cloning or subscription fraud while subscribers are mainly concerned with privacy issues. With the advent of second-generation digital technology platforms like TDMA/CDMA-IS-41,

operators were able to enhance their network security by using improved encryption algorithms and other means. The noise-like signature of a CDMA signal over the air interface makes eavesdropping very difficult. This is due to the CDMA "long code," a pseudo-random noise (PN) sequence of long length, which is used to scramble voice and data transmissions. The PN sequence is represented as a sequence of 1's and 0's with certain properties and the sequences are generally categorised into two classes: (1) periodic sequences and (2) aperiodic sequences. The class of sequences used in spread-spectrum communication is usually periodic. There are many types of periodic sequences, some popular ones being (i) maximal-length linear shift register sequences (m-sequences), (ii) quadratic residue sequences (q-r sequences), (iii) Hall sequences, and (iv) twin primes. Among all these sequences, the most commonly used ones are m-sequences.

The PN key generation is of paramount importance for any secure communication system. Maximal sequences are easily generated by linear feedback shift registers (LFSRs). A LFSR [1-2] consists of a shift register and a feedback network (or a parity generator) consisting of only modulo-2 adders (XOR gates). The output of the feedback network is applied to the input of the shift register. The feedback network provides output logic 0 when an even number of input is at logic 1 and generates logic 1 when an odd number of input is at logic 1 state. Maximal-sequence codes generated by using LFSRs are not adequately secure when smaller lengths of LFSRs are employed.

Many circuits using LFSR for generation of complex codes have been reported in the literature [3-5]. Feedback-with-carry shift registers (FCSRs) and schemes using various combinations of LFSRs have been proposed for complex key generation but seem to be expensive for commercial application [6]. The security of the encrypted data is a direct function of the number of stages of shift register used to generate the key. This means, to increase data security, the number of stages of the shift register is to be increased, which leads to an increase in the complexity of the system in terms of power, space and cost [7].

In this paper a relatively low-length shift register for the generation of highly secure m-sequence generator is presented, wherein the feedback tappings keep on changing in a pseudo-random manner, which makes the generated codes quite complex. The complex code generated is used to scramble incoming plain text. At the receiving end, the same code is generated and successfully used to decrypt the transmitted data. The simplicity of the circuit along with the complexity of the generated codes makes the circuit attractive for secure message communication applications. The proposed technique of complex code generation is modelled in VHSIC hardware description language (VHDL), synthesised and simulated for a field programmable gate array (FPGA) target device. This approach allows the reconfigurability [8-10] of the proposed system such that the key complexity can be further enhanced as per the security requirements. Further, this type of implementation offers many advantages over conventional IC design vis-à-vis dynamic power consumption, space occupied and stray capacitances [11].

Description of Proposed Complex-Code-Based Data Scrambler and Descrambler

The proposed scheme is shown in Figure 1. The heart of the system is a complex code generator, as shown in Figure 2, which generates a long and highly complex key stream. The various

available PN codes at the output of the complex code generator (Q1 to Qn) are applied as input to the AND gates of the scrambler that act as tap gains. The tap gain outputs are applied to the logical function that produces a complex key to scramble the plain text. The complex key generated is XORed with the plain text to produce cipher text, also known as cryptogram. At the receiving end the same key is generated and used to decrypt the incoming cipher text. The whole system has been modelled using VHDL and hence one can exploit the reconfigurability feature of Description Languages to enhance the system security and optimise the power consumption and speed of operation. The complex code generator used to scramble the incoming plain text uses an 8-bit LFSR which can generate 16 different sets of 255-bit code sequences, depending upon 16 valid sets of feedback tappings, viz. {8,4,3,2}, {8,6,5,4}, {8,6,5,3}, {8,5,3,2}, {8,6,5,2}, {8,6,3,2}, {8,5,3,1}, {8,7,4,3}, {8,6,5,1}, {8,7,3,2}, {8,7,6,1}, {8,7,2,1}, {8,7,6,5,2,1}, {8,7,6,3,2,1}, {8,6,4,3,2,1} or {8,7,6,5,3,2}. Any one of these sets of feedback tappings can be used at a time so that a particular combination of the output of the LFSR is connected back to its input through a modulo-2 adder. Thus, any one of the above sets of feedback connections can be selected at a time to generate the corresponding code-sequence, in part or in full, depending on the time for which the selected feedback remains connected. If these feedback connections are changed synchronously in a random manner, the output sequence (Q's in Figure 2) also changes correspondingly. For simplicity of demonstration of the scheme, only seven sets of feedback connections, viz. {8,4,3,2}, {8,6,5,4}, {8,6,5,3}, {8,5,3,2}, {8,6,5,2}, {8,6,3,2} and {8,6,5,1}, are chosen here. These sets of feed-back connections are obtained by XORing various output streams of the LFSR. One set of these connections can be selected at a time with the help of an 8-to-1 line multiplexer (MUX) controlled by a 3-bit word generated by another PN sequence as shown in Figure 2. Since '000' state is to be avoided as the control word of the MUX (because the PN sequence generates the required control word), only seven input signals of the MUX (one at a time) will be chosen depending upon the control word. Thus, feedback tappings are changed randomly, selecting one at a time out of the set of seven different sets. Hence, depending upon the value of N (the divide-by factor in the circuit), together with the initial state of the code generator and the initial state of the MUX, a complex code sequence is generated. The complex codes generated (Q1 to Q8) are used as input to AND gates to generate the tap gains. These tap gains are given to the logic function generator which manipulates the data and produces the complex key, the key being modulo-2-added with plain text and thus resulting in a cryptogram. The cryptogram is applied to the shifting logic so as to increase the complexity of the key and hence the security of data. At the receiving end, the same key is generated and used for the successful decryption of received data.

Simulation Results and Verification

The proposed scheme has been modelled in VHDL, synthesised and simulated for target device EP2S15F484C3 of Straitx II FPGA family using Quarts Altera version 8.1. It is important to mention that Stratix II devices can be used for implementing memory functions and complex logic functions such as digital signal processing, wide data-path manipulation, data transformation, and

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 125-135

microcontrollers. The high-pin-count Stratix II devices contain a two-dimensional row- and column-based architecture to implement custom logic. The devices provide adaptive look-up tables (ALUTs), memory bits, and adaptive logic modules (ALMs). They also include input and output low-voltage differential signalling (LVDS) channels and provide dedicated circuitry to support differential I/O standards at up to 1 Gbps when using dynamic phase alignment (DPA) and 840 Mbps when not using

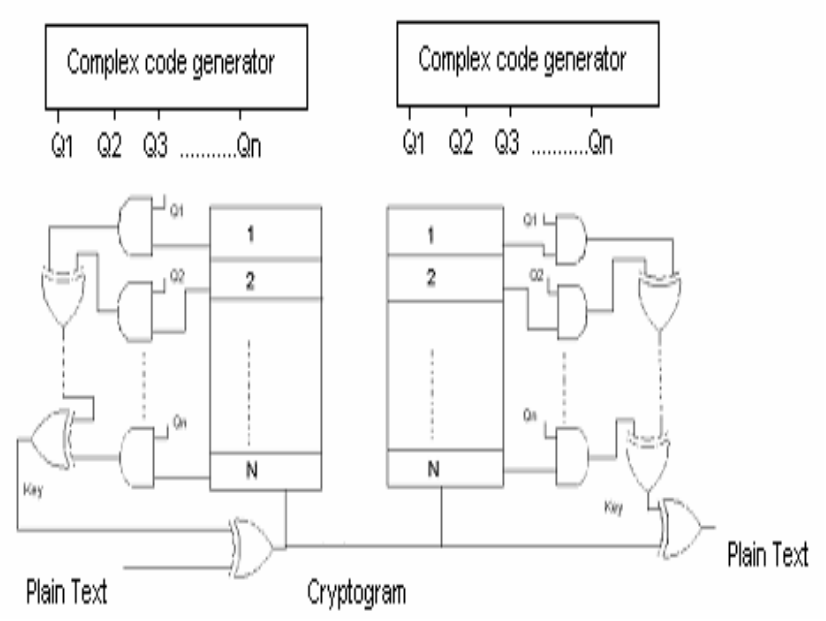


Figure 1. Complex-code-generator-based data scrambler and descrambler

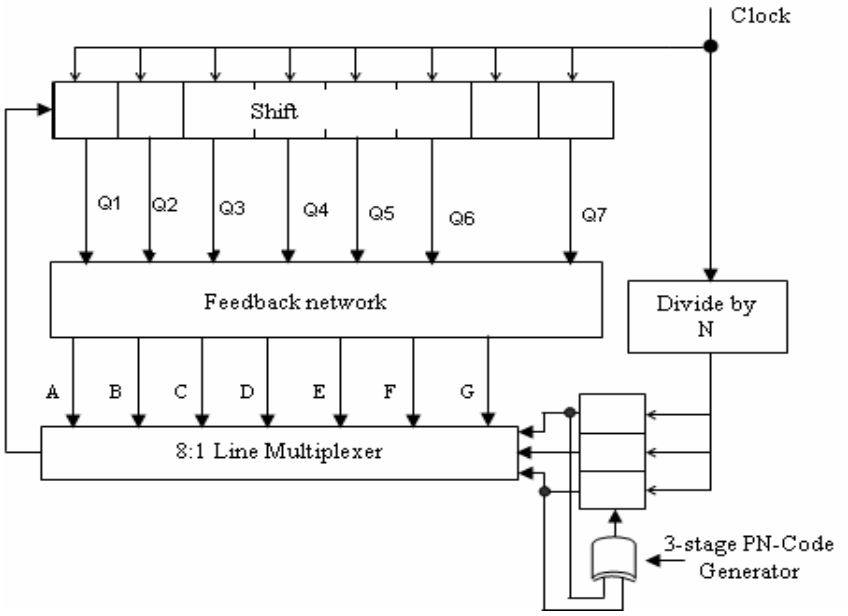


Figure 2. Complex code generator

DPA. Stratix II devices provide enhanced phase-locked loops (PLLs), fast PLLs, and regional clock networks to increase performance, and also provide advanced clock interfacing and clock-frequency synthesis. The devices also contain 16 dedicated clock pins for controlling signals with large fan-outs. In addition, all Stratix II devices include enhanced and fast (PLL) circuitry.

The wave forms obtained at various check nodes (output terminals at which output codes are available) of the complex code generator corresponding to all the 255 × 7 sequence combinations have been investigated and found in conformity with the theoretical observations. As shown in Figure 3, 'Clk' is an input signal which drives the PN sequence generator. PN (7) to PN (0) are the resultant output PN sequences. For seed value (initial value with which LFSR is loaded) of '11111111' and feedback tapping (8, 4, 3, 2), theoretically, the LFSR output sequence (with PN(7) as MSB and PN(0) as LSB) passes through the following states with subsequent clock cycles: '01111111', '00111111', '10011111', '01001111', '00100111', '000100111', '000010011', '10000100', and so on. (The pattern generated is a function of feedback tapping. Hence, the next pattern depends on the previous one and the feedback tapping.) The wave form obtained for LFSR seed value '11111111' and selected tap position (8, 4, 3, 2) for a few clock cycles is presented in Figure 3, which confirms that the results obtained are in agreement with the theoretical values (shown later.)

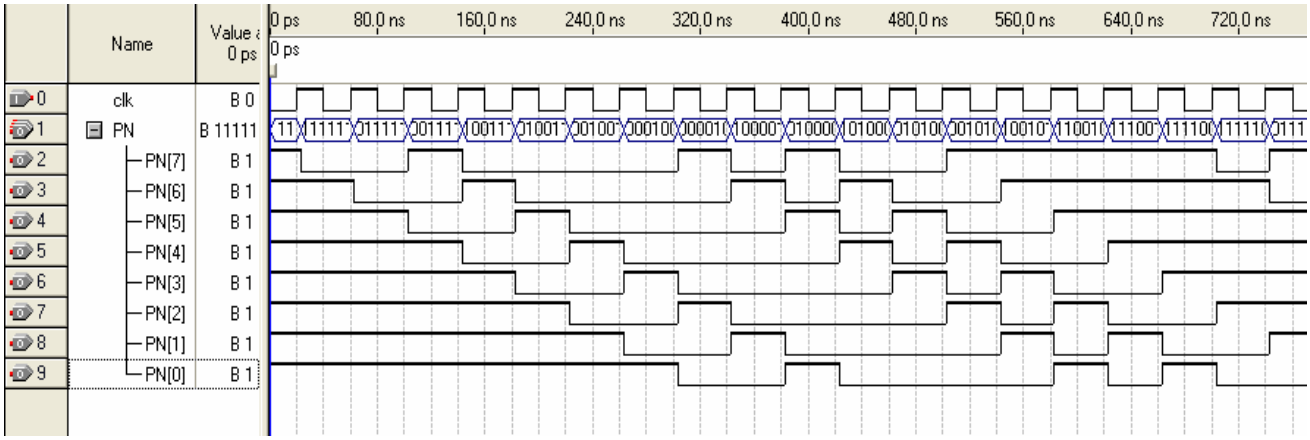


Figure 3. Wave forms obtained at various check points of complex code generator

As shown in Figure 3, prior to application of clock pulse, the value of output codes from PN(7) to PN(0) is '11111111'. With the arrival of positive-going edge of first clock pulse the output changes to '01111111'. The arrival of subsequent clock pulses then keeps on changing the output to '00111111', '10011111', '01001111', '00100111', '000100111', '000010011', '10000100' and so on in that sequence, which is in complete agreement with the theoretical results. The complex codes obtained from the complex code generator as shown in Figure 3 are used to scramble incoming data stream at the transmitter, as shown in Figure 4. At the receiving end, the same codes are generated for successful decryption of data.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 125-135

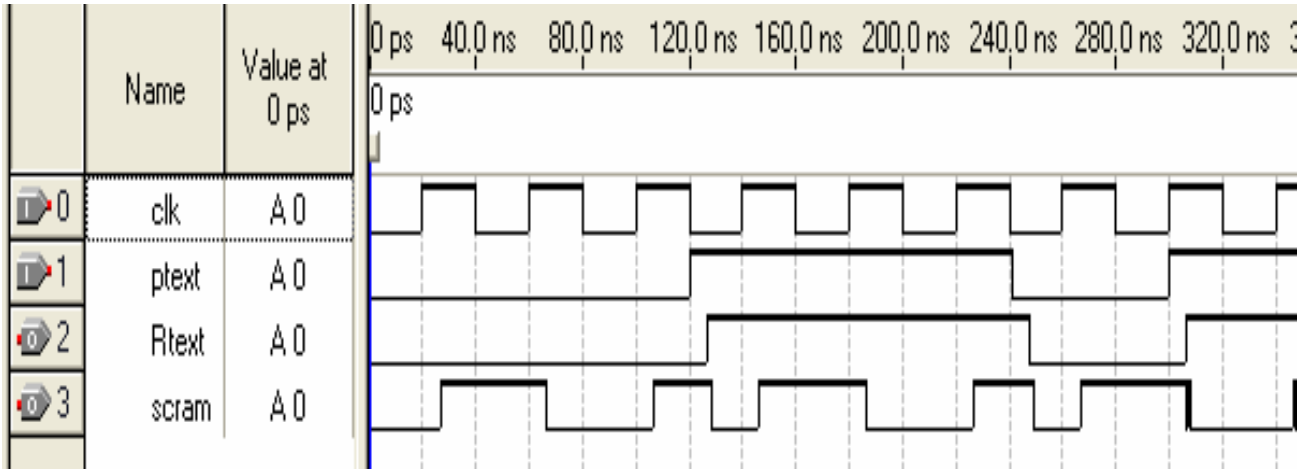


Figure 4. Wave forms obtained at various check points of data scrambler and descrambler

The simulation results are shown in Figure 4. 'Clk' represents the input signal clock used to drive the system, 'Ptext' is the input plain text at the transmitting end, and 'Rtext' represents the received data after decryption at the receiver. 'Scram' represents the scrambled signal corresponding to the input data. As shown in the figure, transmitted data and the received version of it are in complete agreement. However, the wave forms of the transmitted data and its scrambled version do not resemble at all.

The RTL viewers of the implemented designs, viz. scrambler and descrambler, are shown in Figures 5 and 6 respectively. Figure 7 shows RTL viewer of the implemented complex code generator. Finally, a comparison of the codes generated by the proposed scheme with those generated conventionally [3] is presented in Table 1. The comparison is depicted graphically in Figure 8. It is clear that the codes generated using the same length of LFSR are much more complex (more lengthy) than those generated using the conventional PN code generator.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 125-135

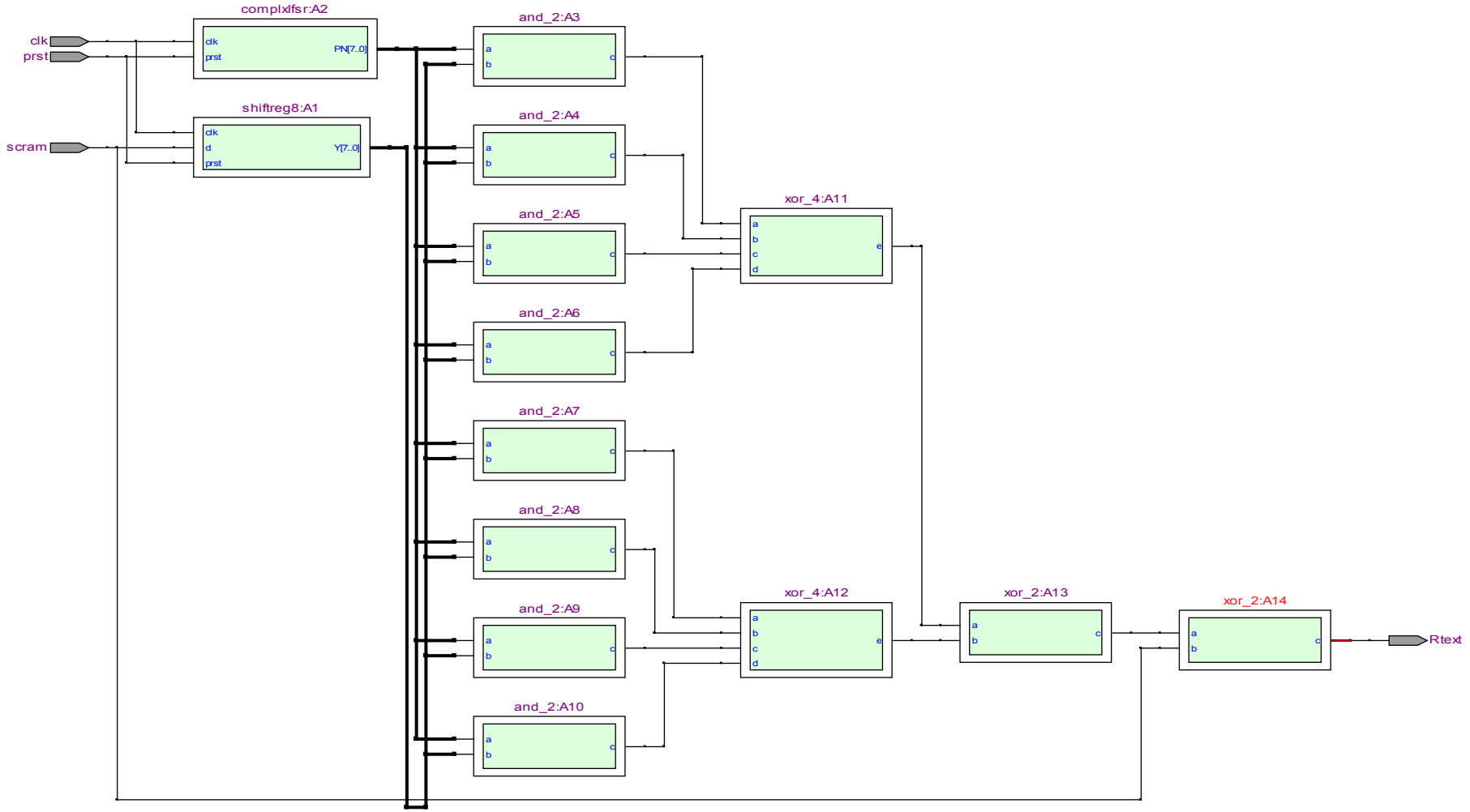


Figure 5. RTL viewer of complex-code-based data scrambler

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 125-135

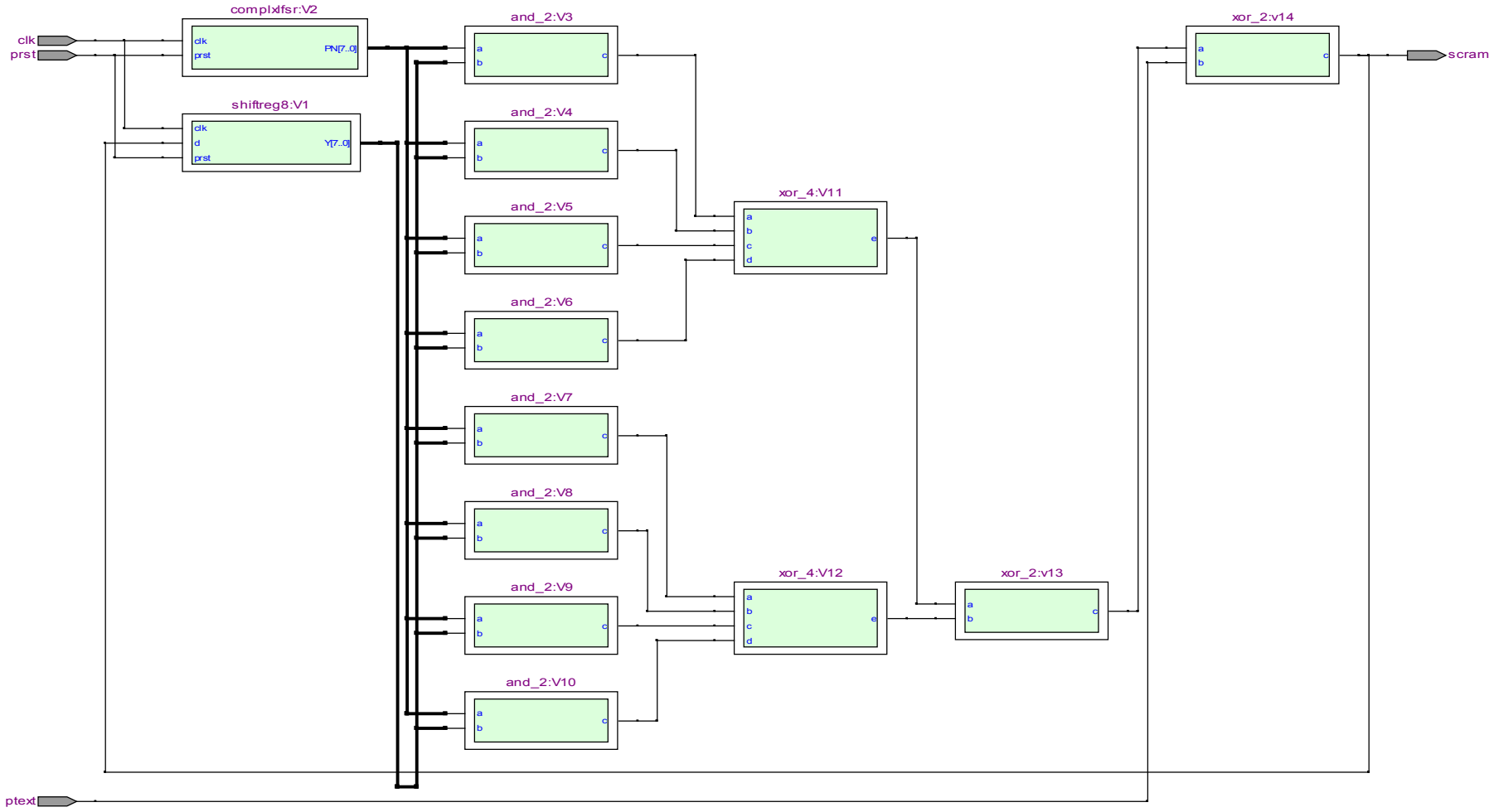


Figure 6. RTL viewer of complex-code-based data descrambler

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 125-135

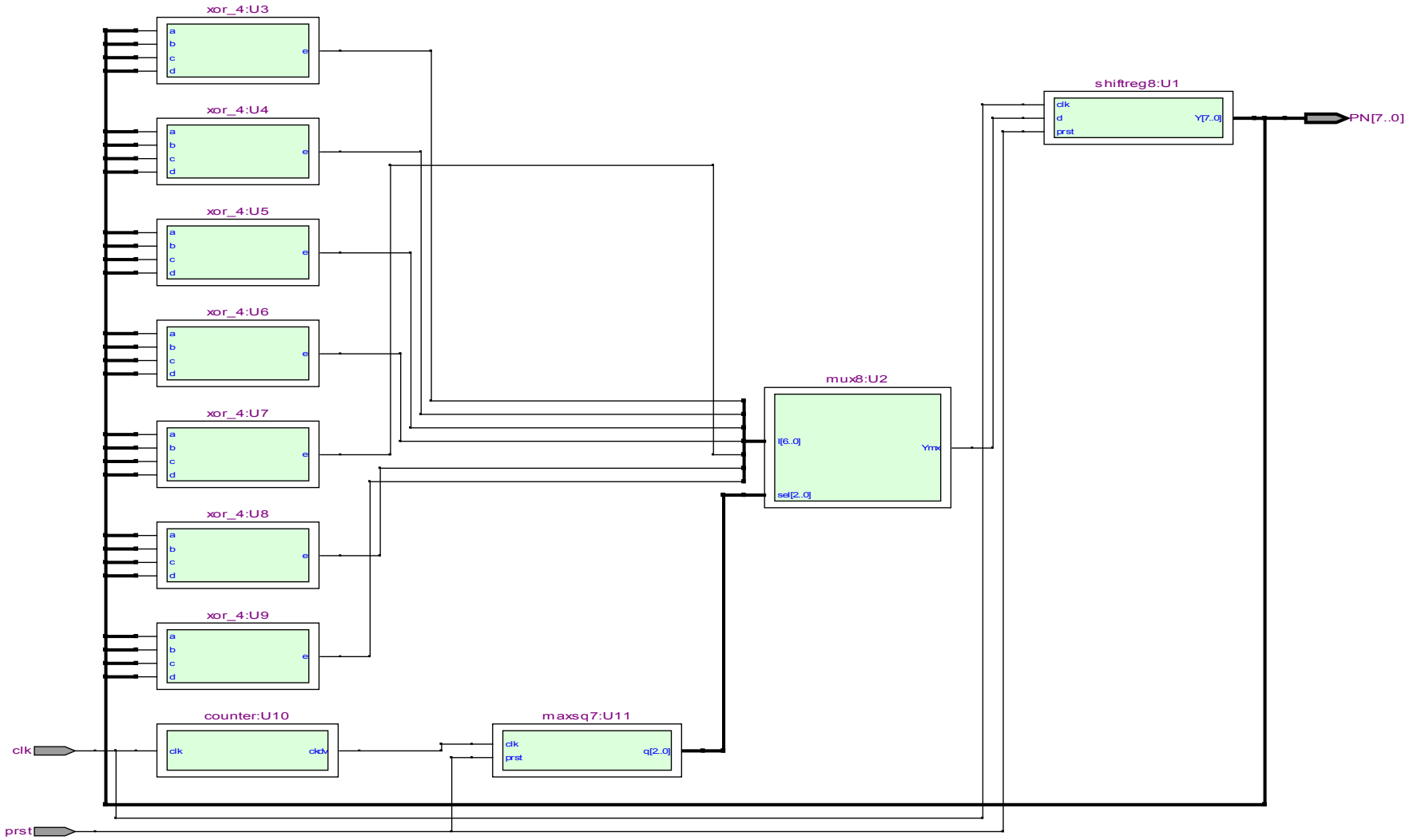


Figure 7. RTL viewer of complex code generator

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 125-135

Table 1. Resource utilisation summary

S.No	Length of shift register	No. of feedback tappings used for complex code generation	No. of feedback tappings used for code generation in conventional method	Length of code generated using		Length improvement factor
				Conventional method	Proposed method	
1	5	3	1	31	93	3
2	6	3	1	63	189	3
3	7	7	1	127	889	7
4	8	7	1	255	1785	7
5	9	10	1	511	5110	10
6	10	10	1	1023	10230	10
7	11	10	1	2047	20470	10
8	14	14	1	16383	229362	14
9	15	21	1	32767	688107	21

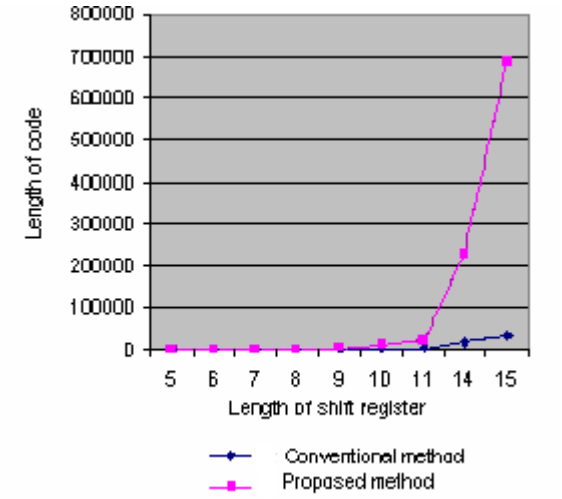


Figure 8. Code length comparison

Conclusions

A new modified scheme for complex PN-code-based data scrambler and descrambler has been presented. The proposed scheme uses a complex code generator, capable of generating lengthy codes using a LFSR with relatively less number of stages, for data encryption and decryption. The proposed scheme has been modelled in VHDL synthesised and simulated for target device EP2S15F484C3 of Straitx II FPGA family using Quarts Altera version 8.1 However, for defense and aerospace applications where design certainly carries special technical challenges, the choice for the target device

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 125-135

can be Xilinx Virtex-5Q family of FPGAs. This provides 65-nm, high-density and high- performance FPGA technology suitable for secure communication systems. The proposed scheme is capable of providing a range of applications in Spread Spectrum Modulation, Code Division Multiple Access and Global Positioning Systems. The scheme can be synthesised and implemented on any of the existing CPLD and FPGA systems as per the degree of optimisation required.

References

1. B. Schneier, "Applied Cryptography", 2nd Edn., John Wiley & Sons Inc., New York, **1996**.
2. P. Pedron, "Linear feedback shift registers", *US Patent 5105376* (**1992**).
3. W. Ahmad and G. M. Bhat, "Scrambler for data", *Electron. World*, **1997**, *3*, 227-228.
4. M. Izharudin, O. Farooq and A. Waseem, "FPGA-based Implementation of Stream Ciphering Circuit", Proceedings of International Conference on ROVSIP, **2005**, Penang, Malaysia, pp. 16-19
5. R.C. Dixon, "Spread-Spectrum Systems with Commercial Applications", John Wiley & Sons Inc., New York, **1994**.
6. A. Wasim, "Development of low-cost secure communication techniques", AICTE (R&D) Project (**2002**), Department of Electronics Engineering, Aligarh Muslim University, Aligarh, India.
7. K. Jarvinen, M. Tommiska and J. Skytta,"Comparative survey of high-performance cryptographic algorithm implementations on FPGAs", *IEE Proc. Inform. Secur.*, **2005**, *152*, 3-12.
8. F. X. Standaert, G. Piret, G. Rouvroy and J. J. Quisquater, "FPGA implementations of the ICEBERG block cipher", *Proc. ITCC*, **2005**, *1*, 556-561.
9. K. Gaj and P. Chodowiec, "Fast implementation and fair comparison of the final candidates for advanced encryption standard using field programmable gate arrays", in "Topics in Cryptology - CT-RSA 2001" (Ed. D. Naccache), Springer, San Francisco, **2001**, pp. 84-99.
10. R. Mita, G. Palumbo and M. Poli, "Pseudo random sequence generators with improved inviobility performance", *IEE Proc. Circuits Devices Syst.*, **2006**, *153*, 375-382.
11. M. Feldhofer, J. Wolkerstorfer and V. Rijmen, "AES implementation on a grain of sand", *IEE Proc. Inf. Secur.*, **2005**, *152*, 13-20.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

Maejo International Journal of Science and Technology

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

Full Paper

Enhancement of transparency and accuracy of credit scoring models through genetic fuzzy classifier

Adel Lahsasna, Raja N. Ainon* and Teh Y. Wah

Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia

*Corresponding author, e-mail: ainon@um.edu.my

Received: 1 October 2009 / Accepted: 7 April 2010 / Published: 19 April 2010

Abstract: Credit risk evaluation systems play an important role in the financial decision-making by enabling faster credit decisions, reducing the cost of credit analysis and diminishing possible risks. Credit scoring is the most commonly used technique for evaluating the creditworthiness of the credit applicants. The credit models built with this technique should satisfy two important criteria, namely accuracy, which measures the capability of predicting the behaviour of the customers, and transparency, which reflects the ability of the model to describe the input-output relation in an understandable way. In our paper, two credit scoring models are proposed using two types of fuzzy systems, namely Takagi-Sugeno (TS) and Mamdani types**.** The accuracy and transparency of these two models have been optimised. The TS fuzzy credit scoring model is generated using subtractive clustering method while the Mamdani fuzzy system is extracted using fuzzy C-means clustering algorithm. The accuracy and transparency of the two resulting fuzzy credit scoring models are optimised using two multi-objective evolutionary techniques. The potential of the proposed modelling approaches for enhancing the transparency of the credit scoring models while maintaining the classification accuracy is illustrated using two benchmark real world data sets. The TS fuzzy system is found to be highly accurate and computationally efficient while the Mamdani fuzzy system is highly transparent, intuitive and humanly understandable.

Keywords: credit scoring, fuzzy classifier, genetic algorithms, transparency

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

Introduction

The global financial crisis of 2008 reveals the importance of the credit risk evaluation decisions not only on the financial institutions and banks but also on both global and local economy. Many major banks collapsed and others suffered heavy losses as a result of mortgage payment default. Hence, decision support tools that aim to enhance the manager's decision may play a valuable role in decision-making by allowing faster and more accurate decisions.

Credit scoring is the most commonly used method for evaluating the creditworthiness of the applicants. Before this method came into use, judgmental method was the only way to differentiate between the good applicants who are likely to repay their debts and the bad ones who are denied because of the high potential of defaulting on their debts. This approach to credit assessment has been criticised for being inconsistent, costly and time consuming. In recent years, credit scoring which replaced the judgmental method, aims at classifying credit applicants into bad and good customers with respect to their features such as age, income, and marital status [1].

Accuracy and transparency are two important criteria that should be satisfied by any credit scoring model. A highly accurate credit model enables correct assessment, thus avoiding any heavy losses associated with wrong predictions while transparent credit model enables financial analysts to understand the decision process.

The literature on credit scoring shows that statistical methods such as linear discriminant analysis and logistic regression are the most commonly used methods in building credit scoring models [2]. However, artificial intelligence techniques such as neural networks and genetic algorithms provide a new alternative to statistical methods in optimising non-linear, complex and real world systems [1, 3-5].

The main reason artificial intelligence techniques are seldom used in credit risk evaluation industry is the lack of explanatory capabilities of these methods. Hence, the enhancement of the transparency of the artificial-intelligence-based credit scoring model is one of the key factors of their successful deployment [6]. The main advantage of the fuzzy system is its transparency. Through the hybridisation of the transparency of the fuzzy system with the excellent learning capacity of the artificial intelligence techniques, some limitations of single-methods transparency may be overcome. Using this approach, some credit scoring models have been proposed using neuro-fuzzy [7-8] and genetic fuzzy [9] techniques to solve the transparency problem. Hoffman et al. [9] proposed a genetic fuzzy system for credit scoring and compared it with Nefclass, a neuro-fuzzy algorithm. The results showed that the performance of the genetic fuzzy algorithm is better than Nefclass [10] while the latter is more transparent. In addition, the above stated study reveals the classical trade-off between accuracy and transparency in the fuzzy systems. Hence, such problem has to be carefully addressed and balanced based on the needs and the objective of the credit scoring user.

In a recent study [3], the main soft computing methods applied in credit scoring models were surveyed. However, the multi-objective genetic algorithm, which is an efficient technique to get a maximum trade-off between conflicting objectives, has not been investigated for its handling of the accuracy and transparency trade-off in the fuzzy-based credit scoring models. The multi-objective genetic algorithm has, however, been successfully applied in the design phase of the

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

fuzzy-rule-based system modelling [11]. Specifically, it has been used in this phase to find an appropriate balance between transparency and complexity of the fuzzy-rule-based system [12]. Moreover, the multi-objective Pareto optimal solutions, the adopted approach in this paper, give more realistic solutions to the problem by allowing the decision-maker to choose between different solutions based on his needs and conditions. This paper aims at investigating the significance of this approach for addressing the above stated problem using two real-world data sets.

In this paper, two credit scoring models are built using Takagi-Sugeno (TS) and Mamdani fuzzy systems. Particularly, the accuracy and transparency of the resulting credit fuzzy models are enhanced using two different multi-objective genetic algorithms. To illustrate the potentiality of the proposed methods, two benchmark data sets, namely German [13] and Australian [14] credit data sets, are used. An overview of the multi-objective genetic optimisation methods is first detailed, followed by the description of the adopted methodology and the data sets used in this study.

Multi-Objective Genetic Algorithms

Many real-world problems have multiple conflicting objectives that should be simultaneously considered, as the optimisation of a particular solution with respect to one objective can give unacceptable results with respect to other objectives. A reasonable approach to multi-objective optimisation problem is to find a set of solutions, each of which achieves the objectives in a balanced way without being dominated by any other solution. Genetic algorithms, the meta-heuristic techniques inspired by the evolutionary biology, are well suited to this class of problems [15].

There are two approaches in multi-objective genetic algorithms optimisation. The first is to combine the various objective functions into a single function in a linear fashion using weight factors. The drawback of this approach lies in the determination of the optimal weight values that characterise the user preferences. The second approach finds the non-dominated Pareto optimal set of solutions for all optimal compromises between the conflicting objectives. It is a practical approach as the decision-maker can find solutions with different trade-off levels. A number of algorithms have been proposed [16-17] and the elitist non-dominated sorting genetic algorithm II (NSGA-II) [18] is among the well-known and most commonly used multi-objective genetic algorithms in the literature.

The NSGA-II was introduced to overcome the following drawbacks of NSGA [19]: (i) computation complexity, (ii) non-elitism approach and (iii) the need for specifying a sharing parameter. This algorithm has two features which makes it an efficient algorithm. The first one is that the fitness function of the solution is based on non-dominated ranking and a crowding measure, and the second is the elitist-generation update procedure. A non-dominated rank is assigned to each individual using the relative fitness. The concept of non-dominated solution can be defined as follows: Individual or solution 'A' dominates 'B' if the two following conditions hold:

(i) 'A' is strictly better than 'B' in at least one objective and

(ii) 'A' is no worse than 'B' in all objectives.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

An outline of the elitism-preserving mechanism of NSGA-II is written as follows:
Step 1: Generate an initial population with *N* chromosomes.
Step 2: Generate an offspring population by iterating the following procedures *N* times:
(1) Select a pair of parent solutions from the current population.
(2) Generate an offspring from the selected parent solutions by genetic operations.
Step 3: Merge the offspring population and current population. Then select the best *N* solutions from the merged population to construct the next population.
Step 4: If a pre-specified stopping condition is satisfied, terminate the execution of the algorithm. Otherwise, return to Step 2. In the former case, we choose all the non-dominated solutions in the merged population in Step 3 as the final solutions.

Controlled elitist genetic algorithm, a variant of NSGA-II, was proposed by Deb and Goel [20] for controlling the extent of the elite members of the population to maintain the diversity of the population for convergence to an optimal Pareto front. The controlling mechanism is accomplished by allowing only a certain portion of the population to be included in the currently-best-non-dominated solutions. The controlled NSGA-II has a better convergence than the original NSGA-II [20], and since we apply the multi-objective genetic algorithm in different steps of optimisation, we choose to use the controlled NSGA-II in our study rather than the original NSGA-II in order to reduce the computational cost.

Methodology

The credit scoring models were implemented using MATLAB 7.5.0. The two proposed methods, which are based on Takagi-Sugero (TS) and Mamdani fuzzy systems are described in the respective order as follows.

First approach: Takagi-Sugeno-fuzzy-based system

In the first approach, the fuzzy systems were extracted from the data by a subtractive clustering method and then the resulting fuzzy rules were optimised to increase the accuracy using genetic algorithms. In the last two steps a multi-objective genetic algorithm was applied to preserve the accuracy of the fuzzy model to a given value while enhancing the transparency of the fuzzy model by reducing the customer input and fuzzy sets in the rule base. The steps are outlined below.

Step 1: Structure and parameter initialisation using subtractive clustering algorithm. In this step, a fuzzy system of TS type was generated using subtractive clustering method [21] which is an efficient and fast algorithm used for estimating the number of clusters and the location of cluster centres in a set of data. The linear least-square estimation was then used to determine each rule consequent equation. This algorithm has the advantage of describing the TS fuzzy model with few rules [21]. The TS fuzzy model [22] uses fuzzy rules with fuzzy antecedents and functional consequent parts. This model is represented by a series of fuzzy rules of the form:

$$R_k: IF\ x\ is\ A^k\ \ Then\ \ y\ is\ f(x) \qquad (1)$$

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

where R_k is the label of the k^{th} fuzzy rule, f represents the output variable y, and A^k is the fuzzy set that is defined over input x, where $x = (x_1, \ldots, x_n)$ is the n-dimensional pattern vector. A^k is represented by Gaussian membership functions of the form:

$$\mu_{ik}(x_i) = \exp\left(-\frac{(x_i - c_{ik})^2}{2a_{ik}^2}\right) \quad (2)$$

where c_{ik} and a_{ik} are the centre and the width of the Gaussian function respectively.

Step 2: Structure and parameter optimisation by genetic algorithm. In the second step, a genetic algorithm was applied to increase the accuracy of the initial fuzzy system by searching for the most suitable value of centre c_{ik} and width a_{ik} of each fuzzy set in the rule base. The fitness function of the genetic algorithm for an individual S_i is given by:

$$Fitns(S_i) = fitn_{acc}(S_i) \quad (3)$$

where $fitn_{acc}(S_i)$ is the fitness function of the accuracy measured by the percentage of correctly classified training patterns. The parameters of membership functions in the antecedents of each fuzzy rule were encoded into a chromosome. Thus, the i-th chromosome is a string of the form:

$$S_i = (\underbrace{c_{11}^{(i)}, a_{11}^{(i)}, \ldots, c_{n1}^{(i)}, a_{n1}^{(i)}}_{premise\ of\ rule\ 1}, \ldots, \underbrace{c_{1K}^{(i)}, a_{1k}^{(i)}, \ldots, c_{nk}^{(i)}, a_{nk}^{(i)}}_{premise\ of\ rule\ K}) \quad (4)$$

The first individual of the initial population was generated as a copy of the premise parameters of the initial fuzzy rules generated from step 1. The initialisation from a good population may speed up the convergence of the solution. The remaining individuals were initialised with random values. The best individuals in the population were always selected and kept unchanged in the next generations according to the elitist strategy. The simplest form of crossover, which is the single-point crossover, was adopted. At the end of this step, the highest-accuracy fuzzy model was obtained.

Step 3: Feature selection using multi-objective genetic algorithm. The objective of this step is to reduce input dimensions by choosing the relevant subset of features. To achieve this, we need first to keep the accuracy achieved in the previous GA optimisation step as high as possible while choosing the subset which contains the smallest number of features.

The modelling objectives of fuzzy system S in this step can be written as follows:

$$\text{Maximise } f_{acc}(S), \ \text{Minimise } f_{input}(S) \quad (5)$$

where $f_{acc}(S)$ is the fuzzy system accuracy measured by the percentage of correctly classified training patterns and $fitn_{input}(S)$ is the total number of selected features of a fuzzy system. To simultaneously achieve these two objectives, controlled elitist genetic algorithm (controlled NSGA II) was applied. The results of this step are Pareto-front solutions that represent a number of fuzzy models with different accuracy numbers of input values. The fuzzy model chosen in this case is based on the need of the user, that is, if the accuracy is more important than the transparency then a fuzzy model with high accuracy and high number of features will be chosen. As stated before, our objective is to enhance the transparency while keeping almost the same accuracy. So a fuzzy model which has accuracy value near to the initial fuzzy model was chosen.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

The final result of this step is a fuzzy model with relatively good classification accuracy and relatively fewer number of features.

The following is the design of the chromosome used for feature selection (step3) as well as the genetic operators applied for exploring the search space by testing every possible combination of candidate features and selecting the relevant ones.

Chromosome design

The chromosome S_i which represents the selected features is denoted by a concatenated binary bit string of length n (n is the total number of features in the data set), where each binary bit denotes whether a given input is selected during the feature selection process. In this implementation, the selected features were set to 1 while the non-selected features were set to 0. Figure 1 shows the structure of an example of one chromosome after the selected feature process in the Australian data set. The total number of features is 14 and the selected inputs which have the value 1 are: 1, 2, 5, 8, 9, 10, 13 and 14.

F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14
1	1	0	0	1	0	0	1	1	1	0	0	1	1

Figure 1. Chromosome of the genetic algorithm used in the feature selection

Genetic operators

A new population P of chromosomes (fuzzy systems) was generated using the genetic operations: selection, crossover and mutation. To generate a new fuzzy system S_i, first a pair of parent fuzzy systems was selected from the current population using tournament selection based on the Pareto ranking and the crowding distance. In order to maintain the diversity in the next population, the best non-dominated solutions were kept down to only 35% (which is the default value defined by the algorithm [23]) of the population. In addition, the crowding measure was used to calculate the crowding distance for each individual on a non-dominated front. After the selection step, the uniform crossover and uniform mutation with a range of 0.01 were applied. These genetic operations were applied for fuzzy sets selection step (next step) and also for fuzzy sets replacement process in the second approach because they involve similar problem.

Step 4: Fuzzy sets selection using multi-objective genetic algorithm. The final step in the optimisation process is the removal of the fuzzy sets whose effect on the fuzzy system is not important. In this case, we have two objectives; the first one is to maximise the classification accuracy of the fuzzy system while minimising the number of fuzzy sets in the fuzzy system is the second objective. The objectives of the fuzzy system S are written as follows:

$$\text{Maximise } f_{acc}(S) \text{ , Minimise } f_{sets}(S) \tag{6}$$

where $f_{acc}(S)$ is the accuracy of the fuzzy system measured by the percentage of correctly classified training patterns and $f_{sets}(S)$ is the number of selected fuzzy sets. To accomplish these two objectives, controlled elitist genetic algorithm was applied. Among the Pareto-front solutions

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

that represent fuzzy models with different accuracy numbers of fuzzy set trade-off, a fuzzy system with accuracy almost equal to the initial fuzzy system was chosen. After this final stage, a fuzzy model with relatively good classification accuracy and less fuzzy sets was obtained.

The following is a description of the chromosome design used for fuzzy set selection (step 4). In this step, we adopt the same genetic operators as in step 3.

Chromosome design

The chromosome S_i which represents the selected antecedents is denoted by a concatenated binary bit string of length $L_i = n' \times K$, where n' and K are the number of inputs after feature selection phase and the number of rules in the fuzzy system respectively. Each binary bit in the string denotes whether a given fuzzy antecedent is selected. In this case, the selected antecedents were set to 1 and non-selected antecedents were set to 0. Figure 2 shows the structure of the chromosome used in the antecedent fuzzy sets selection phase.

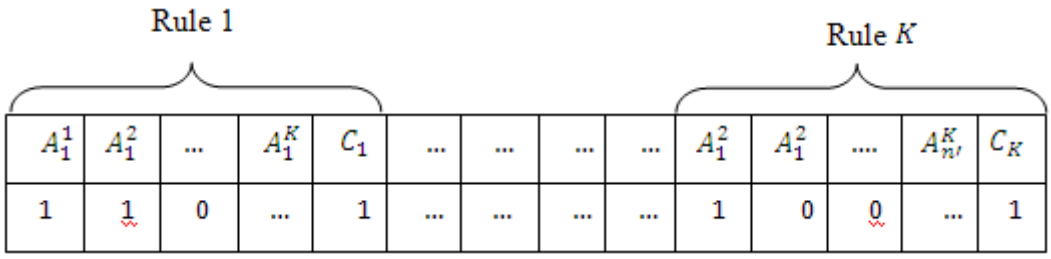


Figure 2. Chromosome of the genetic algorithm used in the antecedent fuzzy set selection phase

Second approach: Mamdani-fuzzy-based system

In the second approach, we used Mamdani fuzzy system [24] in place of Takagi-Sugeno one. First, the initial Mamdani fuzzy system was generated using fuzzy C-means clustering (FCM) method [25]. The generated Mamdani fuzzy rules are written as:

$$R_k: IF\ x\ is\ A^k\ \ Then\ \ y\ is\ B^k \qquad (7)$$

where R_k is the label of the k^{th} fuzzy rule, A^k is the fuzzy set defined over the input x where $x = (x_1, \ldots, x_n)$ is the n-dimensional pattern vector while B^k is a fuzzy set defined over the output variable y. All the fuzzy sets in the rule base are represented by Gaussian function.

The next step is to replace the fuzzy sets of the generated fuzzy system by new fuzzy sets. The reason behind this replacement is that the fuzzy sets resulting from clustering or learning method are usually not interpretable [10]. On the other hand, the new predefined fuzzy sets have clear linguistic interpretations such as low, average and high. The linguistic values of each attribute x_i have to be defined before starting the replacement process. In our case, we use five linguistic values: very low, low, average, high and very high, and each of the linguistic values is defined within a specific range of values. Figure 3 shows an example of five linguistic values for the credit amount attribute in German credit data. These new fuzzy sets replace the existing ones

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

of the credit amount attribute in the fuzzy-rule-based system. This idea is similar to that applied by Ishibuchi et al [26].

In replacing the existing fuzzy sets by the new ones the following must be considered:
- The replacement of an existing fuzzy set A^{ik} of the rule R_k and the x_i attributes by A'^{ik} where A'^{ik} is one of the linguistic values defined over the x_i attribute. (For example, A'^{ik} could be either low, average or high.)
- The replacement procedure has to improve the classification accuracy of the fuzzy system.
- In addition to the five linguistic values, 'don't care' is another linguistic value and it refers to unimportant fuzzy set that can be deleted without effecting the fuzzy system performance.

In the first subsection below more explanation on the problem of replacing the existing fuzzy sets with linguistic values is given and the proposed solution is described. In the second subsection, a description of the chromosome design used for Mamdani-based fuzzy system is given.

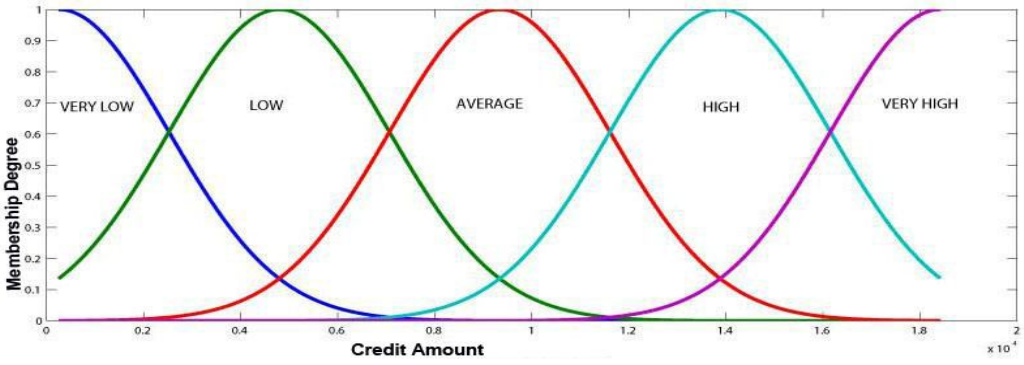


Figure 3. Linguistic fuzzy sets of the credit amount attribute – German data set

Problem formulation

Let K_i be the number of linguistic values in each attribute. So, for each attribute x_i , we have K_i possible antecedent fuzzy sets. In addition, 'don't care' is considered as another fuzzy set. In this case, we have $(K_i + 1)$ possible cases and each antecedent fuzzy set A^{ik} in the fuzzy rules is selected from the given K_i linguistic values and 'don't care'. The total number of possible combinations of the antecedent linguistic values in the fuzzy rules is $(K_{11} + 1) * (K_{22} + 1) * \ldots * (K_{in} + 1)$, where K_{in} is the number of linguistic values of x_i and n is the number of antecedent fuzzy sets in the fuzzy rules.

The task now is to search for the best combination of these antecedent linguistic values that achieves the two objectives, namely maximising the classification accuracy and maximising the transparency by increasing the number of 'don't care' fuzzy sets in the rule base. These two objectives of the fuzzy system S can be written as:

$$\text{Maximise } f_{acc}(s), \text{ Maximise } f_{transp}(s) \quad (8)$$

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

where f_{acc}(s) is the classification accuracy of the fuzzy system measured by the percentage of correctly classified training patterns and f_{transp} is the transparency measured by the number of 'don't care' fuzzy sets in the rule base. To solve this combinatorial problem with these two objectives, a controlled elitist genetic algorithm is applied. Since 'don't care' conditions can be omitted, fuzzy rules with many 'don't care' conditions are written as short fuzzy rules.

Chromosome design

The chromosome S_i is coded as follows:

$$S_i = (\underbrace{A_1^1\ A_2^1\ A_3^1\ \ldots\ A_n^1\ C_1}_{\text{Premise and consequent of rule 1}}\ ,\ldots,\ \underbrace{A_1^K\ A_2^K\ A_3^K\ \ldots\ A_n^K\ C_K}_{\text{Premise and consequent of rule } K}) \quad (9)$$

where n and K denote the number of features and fuzzy rules respectively. A_i^j is the linguistic antecedent value and C_K is the consequent class. The length of the chromosome is $(n+1)\times K$. We used six linguistic values: very low, low, average, high, very high and 'don't care'. Each of these linguistic values is defined by a number. In our case, we set the values 0, 1, 2, 3, 4 and 5 to denote 'don't care', very low, low, average, high and very high respectively. For the consequent class, we set 0 and 1 for negative and positive class respectively. In this case, each antecedent condition $A_i^j \in \{0,1,2,3,4,5\}$ and the consequent class $C_K \in \{0,1\}$.

The following is an example to further explain this idea. Assume that we generate a fuzzy system with 3 inputs and 2 rules in the clustering step and then we get the string 01204521 as one of the best Pareto solutions at the end of the multi-objective optimisation process. Figure 4 shows the decoding process of the 01204521 string. Since we have 3 inputs and one output, the length of string encoding one rule is four. Decoding process of the previous string results in the following rules:

Rule1: If input 1 is 'don't care', input 2 is very low and input 3 is low, then outcome is negative.

Rule2: If input 1 is high, input 2 is very high and input 3 is low, then outcome is positive.

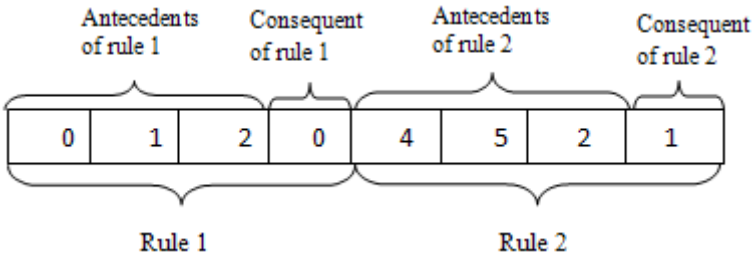


Figure 4. Chromosome coding with 3 inputs and 2 rules

Data Sets

We used two data sets, namely Germany credit data set [13] and Australian credit data set [14]. Both data sets are made publicly available to all users from the UCI Repository of Machine Learning Databases [13] and are mostly used to compare the performance of various classification models.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

To eliminate the skewness and bias in the training and test samples, a common validation method called random sub-sampling validation was applied [27]. Using this method, the whole of the two data sets (German and Australian) were randomly divided into two parts: one for training and the other for testing. As commonly done in similar studies [27], 70% of the data were used for training purpose and 30% for testing the performance of the model. This process was repeated five times to create five pairs of training and test samples for each data set. To calculate the fuzzy classification accuracy for example, five fuzzy systems were built using the training data sets of the five data partitions (S1, S2, S3, S4 and S5) and the accuracy result (validation result) was averaged over the classification accuracy of the corresponding five test sets. The advantage of this technique over *k*-fold and leave-one-cross validation methods is that the proportion of the training/testing split is independent of the number of iterations (folds). Table 1 presents the features of the data sets used in this study. (See Tables 9 and 10 in Appendix for more details on the data attributes.)

Table 1. German and Australian data sets

	No. of inputs	**Training set size**	**Testing set size**	**Data set size**
German data	20	700	300	1000
Australian data	14	383	207	690

Results and Discussion

The first approach: TS-fuzzy-based system

The results of the first step are summarised in Table 2. Since the transparency of the credit scoring model is one of the two modelling objectives, a compact fuzzy system with 3 fuzzy rules and relatively good accuracy was generated in the first step for both data sets.

Table 2. Classification accuracy of initial fuzzy systems with 3 rules using five randomly generated samples

		S1	**S2**	**S3**	**S4**	**S5**	**Average**
Australian data set	Training accuracy (%)	88.61	88.82	89.03	88.82	87.58	88.57
	Testing accuracy (%)	89.37	86.96	86.96	86.96	86.47	87.34
German data set	Training accuracy (%)	75.57	75.71	77.86	76.29	77.43	76.57
	Testing accuracy (%)	75.00	69.67	74.67	73.33	71.33	72.80

In the second step, a genetic algorithm was applied to improve the performance of the fuzzy system generated from the first step. The disadvantage of the genetic algorithm is its computational cost, but starting from a good point (initial fuzzy model) has speeded up the convergence of the genetic algorithm. For example, the genetic algorithm in the case of the Australian data-S1 took around 15 epochs to converge while it took around 16 epochs in the case

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

of the German data-S1. This reveals the complementary functioning between fuzzy clustering and genetic algorithm. The results of this step are shown in Table 3. As the table indicates, there is an increase in the prediction accuracy for both data sets. For the Australian data set, the classification accuracy for testing data increases from 87.34% to 88.89% while that for the German data set has a relatively significant increase from 72.80% to 77.07%. The difference in the enhancement may be due to the difference in the degree of the complexity in the two data sets. The German data is more complicated than the Australian one as the former has 20 inputs while the latter has 14 inputs. Thus, the genetic algorithm seems to be more efficient in dealing with this complexity than the subtractive clustering algorithm.

In the third step, a multi-objective genetic algorithm was applied to select the most relevant inputs in both data sets. Table 4 summarises the results of this step. The names and characteristics of the selected inputs for the two data sets are listed in Tables 9-10 (Appendix). For example, as Table 4 shows, the selected inputs for the German data-S1 are: 1, 2, 3, 5, 8, 12, 14 and 19 and the corresponding attributes are the following: (1) Status of existing checking account, (2) Duration in month, (3) Credit history, (5) Credit amount, (8) Installment rate, (12) Property, (14) Other installment plans, and (19) Telephone. The Australian data attributes have been changed to meaningless symbols to protect the confidentiality of the data. As can be seen from Table 4, there is an improvement in the transparency of the fuzzy systems for both data sets represented by the decrease in the number of the inputs while there is a slight decrease in the accuracy of the fuzzy systems. Particularly, in the Australian data set the accuracy is kept almost the same. Its average number of inputs of fuzzy models is 5.8 while it is 8.6 for the German data set. The new fuzzy system of the Australian data-S1 has only 6 inputs and the other inputs have been deleted without affecting the prediction accuracy of the model. Hence, there is a complexity (in this case unnecessary attributes) that should be removed without decreasing the model performance and that has no relation with the accuracy-transparency trade-off. Furthermore, there are some cases where removing some inputs may increase the accuracy (like in the case of Australian data-S5). On the other hand, results from the German data case reveal accuracy-transparency trade-off. For example, the performance of the fuzzy system of sample 3 that contains 11 inputs is better than that of the other fuzzy models with fewer numbers of inputs.

Table 3. Classification accuracy of fuzzy systems after applying genetic algorithm on initial fuzzy systems

		S1	S2	S3	S4	S5	Average
Australian data set	Training accuracy (%)	88.82	88.00	90.06	89.65	89.44	89.19
	Testing accuracy (%)	89.86	87.92	88.41	89.37	88.89	88.89
German data set	Training accuracy (%)	78.00	77.71	78.86	77.43	78.29	78.06
	Testing accuracy (%)	78.00	74.67	77.67	78.00	77.00	77.07

In the final step, the unnecessary fuzzy sets were removed. Table 5 shows the results of this step for both data sets. For the Australian data set, the average number of fuzzy sets per rule decreases from 5.8 to 3 while the prediction accuracy is maintained as it was before transparency

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

optimisation. In the case of the German data set, there is an improvement in the transparency of the fuzzy system from 8.6 to 4.9 fuzzy sets per rule with a slight decrease in the prediction accuracy. The fuzzy rules resulting from the last step of the Australian data-S1 are depicted in Figures 5-6. As Table 5 shows, 8 fuzzy sets are removed with only a slight decrease in the classification accuracy from 89.86% to 89.37%.

Table 4. Classification accuracy of fuzzy systems and their corresponding number of selected inputs after applying feature selection procedure

		S1	**S2**	**S3**	**S4**	**S5**	**Average**
Australian data set	No.of selected inputs	6	5	6	7	5	**5.8**
	Inputs selected	1, 4, 8, 9, 11, 13	1, 3, 4, 10, 11	1, 4, 7, 9, 11, 13	1, 3, 6, 8, 9, 10, 11	2, 7, 8, 11, 14	
	Training accuracy (%)	87.58	87.58	87.37	88.61	88.41	**87.91**
	Testing accuracy (%)	89.86	87.92	87.92	88.41	89.37	**88.70**
German data set	No.of selected inputs	8	9	11	6	9	**8.6**
	Inputs selected	1, 2, 3, 5, 8, 12, 14, 19	1, 2, 5, 8, 9, 10, 11, 15, 18	1, 3, 6, 7, 9, 10, 13, 14, 15, 18, 19	1, 5, 7, 10, 16, 17	1, 2, 4, 6, 9, 10, 13, 17, 19	
	Training accuracy (%)	75.00	75.42	76.57	75.43	75.57	**75.60**
	Testing accuracy (%)	75.33	73.00	77.00	75.00	75.67	**75.20**

Table 5. Classification accuracy of fuzzy systems and their corresponding number of selected fuzzy sets after applying fuzzy set selection procedure

		S1	**S2**	**S3**	**S4**	**S5**	**Average**
Australian data set	No.of selected fuzzy sets (sets/rule)	10	12	8	8	7	**3**
	Training accuracy (%)	88.00	87.58	88.20	88.82	88.41	**88.20**
	Testing accuracy (%)	89.37	88.41	87.92	88.89	89.41	**88.60**
German data set	No.of selected fuzzy sets (sets/rule)	12	12	14	17	18	**4.9**
	Training accuracy (%)	75.29	74.57	75.43	75.71	76.57	**75.51**
	Testing accuracy (%)	75.67	73.00	75.33	75.00	76.00	**75**

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

(1) IF Input4 is Gaussian(0.6656 2.042) AND IF In- put8 is Gaussian(0.3767 0.03808) AND IF Input11 is Gaussian (0.3998 0.1251) AND IF Input13 is Gaussian(305.7 199.7)Then The customer is GOOD

(2) IF Input9 is Gaussian(0.3874 0.9673) AND IF Input11 is Gaussian(0.4078 0.9664) Then The customer is BAD

(3) IF Input1 is Gaussian(0.3679 0.2458) AND IF Input4 is Gaussian(0.6675 2.005) AND IF Input9 is Gaussian (0.3693 0.8825) AND IF Input13 is Gaussian(312.1 0.06008) Then The customer is BAD

Figure 5. Approximate TS fuzzy rules generated from Australian data-S1

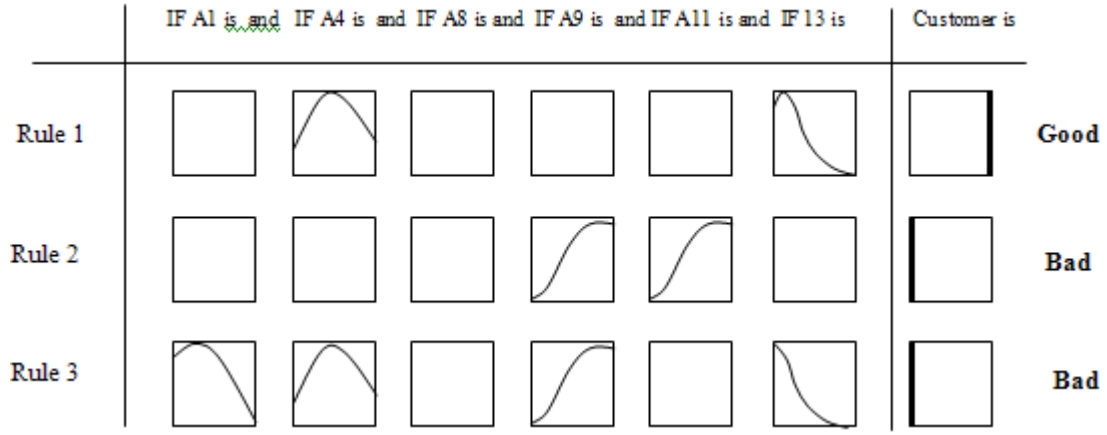


Figure 6. Fuzzy rules after the fuzzy set selection step (Australian data-S1)

Evaluation of the transparency of the antecedent fuzzy sets. The fuzzy IF-THEN rules depicted in Figure 5 are approximate fuzzy rules and they do not give an accurate description of the antecedent of the fuzzy rules and therefore this fuzzy system is not considered transparent despite its compactness, as it does not satisfy one of the comprehensibility measures which is the linguistic representation of the produced fuzzy rules. However, some useful information like the attributes that influence or contribute to the system decision can be extracted and some of the antecedent fuzzy sets can be easily understood using the graphical plot of fuzzy rules and even transformed into linguistic values. For example, the 11th attribute of the Australian data set is categorical and has two values; the first one is 0 and the second is 1. As Figure 5 shows, this attribute is only included in the first and second rules that represent the good and bad classes respectively. It can be seen that the value of the 11th attribute is 0 in the first rule while it takes the value 1 in the second rule and it can be concluded that the 11th attribute of the Australian data set has an effect in the discrimination between the good and bad customers. Furthermore, linguistic values can be assigned to both of the two fuzzy sets. Therefore, rather than saying IF Input 11 is Gaussian (0.3998 0.1251) in the first rule and IF Input 11 is Gaussian (0.4078 0.9664) in the second rule, we can say IF Input 11 belongs to category 1 (value 0) and IF Input 11

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

belongs to category 2 (value 1) in the first and the second rule respectively. Another issue for the continuous attribute is when there is too much overlapping between the antecedent fuzzy sets. This problem prevents the readability of the fuzzy sets and causes a lack of comprehensibility of the fuzzy system. To overcome this problem, a similarity-driven procedure [28] can be applied to merge similar antecedent fuzzy sets into a given attribute.

Evaluation of the accuracy of the TS fuzzy system. To evaluate the performance of our approach, the credit scoring model developed in this study is compared with the other benchmark methods [1, 29-30] applied on the same data sets. These methods are usually used to test the classification accuracy of the new algorithms applied for credit scoring models. More details about the main characteristics of these methods and their application in the credit scoring systems are described by Lahsasna et al [3]. As Table 6 shows, the first approach applied in this study (TS fuzzy system) compares favourably with the other methods such as genetic programming (GP), artificial neural networks (ANNs), radial basis function (RBF) and genetic algorithms-support vector machines ('GA+SVM') hybrid method while it is superior to some methods such as classification and regression tree (CART), rough sets, and the popular decision tree algorithm C4.5. Even though the machine learning methods are accurate classification methods, the lack of transparency of these methods is a major drawback especially when the end user needs to get some information about the credit system. Unlike these black-box methods like ANNs, SVM and genetic algorithm, TS-fuzzy-based method gives some useful information (such as defining the customer's attributes) that influences the system decision and the approximate values of these attributes.

Table 6. Classification accuracy of 'GA+SVM', GP, CART, C4.5, rough sets, ANNs, RBF and TS-fuzzy-based system

Author	Method	Classification accuracy for Australian data (%)	Classification accuracy for German data (%)
[29]	GA+SVM	86.9	77.92
[30]	GP	88.27	77.34
[30]	CART	85.81	70.59
[30]	C4.5	87.06	73.17
[30]	Rough sets	83.72	74.57
[30]	ANNs	87.93	75.51
[1]	RBF	87.78	75.63
This paper	TS-fuzzy-based system (accuracy only)	88.89	77.07

In the case where the end user is only interested in prediction (i.e. getting the best classification accuracy), the use of TS-fuzzy-based system which results from step 2 (structure and parameter optimisation by genetic algorithm) is recommended as it is more accurate than the

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

TS-fuzzy-based system resulting from the final step where both the accuracy and the transparency have been considered. Alternatively, other methods such as GP, ANNs, RBF and 'GA+SVM' can be used in such case.

The second approach: Mamdani-fuzzy-based system

The results obtained at the end of this step are summarised in Table 7, which shows the degree of performance and the level of compactness of each fuzzy system using the rate of correctly classified testing patterns and the number of antecedents per rule respectively. The number of antecedent fuzzy sets per rule has been reduced from 20 and 14 to 6.7 and 3.8 fuzzy sets per rule for the German and Australian data sets respectively. So this method gives relatively good results in enhancing the compactness of the initial fuzzy system resulting from the clustering step.

Table 7. Accuracy and transparency results for Mamdani-based fuzzy system

		S1	**S2**	**S3**	**S4**	**S5**	**Average**
German data set	Training accuracy (%)	78.28	78.57	76.57	77.14	78.86	**77.88**
	Testing accuracy (%)	74.33	71	73.33	72.33	72	**72.60**
	Total no.of fuzzy sets in fuzzy system	160	160	160	160	160	**160**
	Total no.of selected fuzzy sets in fuzzy system	56	61	53	49	50	**53.8**
	Average no.of selected fuzzy sets per rule	7	7.6	6.6	6.1	6.25	**6.71**
Australian data set	Training accuracy (%)	86.75	87.78	90.26	86.75	88	**87.91**
	Testing accuracy (%)	88.88	86	86	86	84.05	**86.19**
	Total no.of fuzzy sets in fuzzy system	98	98	98	98	98	**98**
	Total no.of selected fuzzy sets in fuzzy system	19	26	25	36	29	**27**
	Average no.of selected fuzzy sets per rule	2.71	3.71	3.57	5.14	3.6	**3.8**

Evaluating the antecedent fuzzy sets comprehensibility. The antecedent fuzzy sets of this system become well defined and distinguishable. In Figure 3, the antecedent fuzzy sets of the attribute credit amount are plotted. This attribute may be assigned five well defined linguistic values: very low, low, average, high and very high, and every linguistic fuzzy set has a specific range of values. The fuzzy rules are humanly understandable because of the natural language used. Hence, the descriptive Mamdani fuzzy rules generated have the capacity to represent the knowledge characterising the relations between the customer features and his creditworthiness in a series of linguistic fuzzy rules, thus rendering the decision process of the system understandable

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

and helping the manager in the financial institution to make useful financial analysis and then make the right decisions.

Comparison between TS and Mamdani fuzzy systems. Table 8 shows accuracy and transparency results for German and Australian data sets using TS and Mamdani fuzzy systems. Transparency results are shown using transparency 1, which indicates the number of rules, and transparency 2, which defines the number of fuzzy sets per rule. In addition, Figure 7 shows Mamdani fuzzy rules generated from Australian data set while Figure 8 shows the fuzzy rules extracted from German data set. The fuzzy system uses IF-THEN rules with linguistic values. By comparing this rule set with the rules extracted from TS fuzzy approach using the performance and comprehensibility criteria, the following results are noted.

Performance : TS fuzzy system performance is 88.60% and 75% for the Australian and German data sets respectively while for Mamdani fuzzy system it is 86.19% and 72.60% for the same data sets. These results indicate that TS fuzzy system is more powerful than Mamdani fuzzy system and thus the former is the better choice for predicting the customer's creditworthiness.

Comprehensibility : Compared to Mamdani fuzzy system TS fuzzy system is more compact as it uses only 3 rules for both data sets while the former uses 7 rules for Australian data and 8 rules for German data. Furthermore, TS fuzzy system generally uses slightly smaller number of antecedent fuzzy sets per rule than Mamdani system for both data sets. Despite these strong points of the TS fuzzy system, however, the Mamdani fuzzy system has a major advantage in the capacity to represent the fuzzy rules in an intuitive way using linguistic fuzzy rules. This capacity represents the true level of comprehensibility. The approximate fuzzy rules of the TS fuzzy system do not give a clear idea about the underlying relation between the customer features and their creditworthiness or generally between the input and the output of the data.

Table 8. Accuracy and transparency results for TS- and Mamdani-based fuzzy systems

Data	**TS-fuzzy-based system**			**Mamdani-fuzzy-based system**		
	Accuracy	**Transparency 1**	**Transparency 2**	**Accuracy**	**Transparency 1**	**Transparency 2**
German	75%	3 rules	4.9 sets/rule	72.60%	8 rules	3.8 sets/rule
Australian	88.60%	3 rules	3 sets/rule	86.19%	7 rules	6.71 sets/rule

Despite the enhancement of the comprehensibility of approximate of TS fuzzy sets using similarity-driven method, the problem is not definitely resolved especially when there is a high number of fuzzy sets in the same attribute. The Mamdani fuzzy system is therefore the best choice for data analysis and knowledge discovery from the data set. Thus, the choice of the system type is based on the needs of the user as to whether a high accuracy prediction or a high comprehensibility system is needed. For generating a completely transparent credit scoring model, Mamdani fuzzy system should be chosen. This makes the credit scoring model easier to

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

(1) (IF A4 is p(=1)) and (IF A9 is t(=0))(Then customer is bad)

(2) (IF A5 is j(=5)) and (IF A9 is f(=1)) and (IF A12 is s(=1))(Then customer is bad)

(3) (IF A5 is j(=5)) and (IF A8 is f(=1))(Then customer is good)

(4) (IF A5 is aa(=6)) and (IF A9 is t(=0)) and (IF A10 is very low) and (IF A11 is f(=1))(Then customer is bad)

(5) (IF A5 is c(=8)) and (IF A10 is very low) and (IF A14 is average)(Then customer is bad)

(6) (IF A4 is p(=1)) and (IF A8 is f(=1)) and (IF A12 is p(=3)) and (IF A13 is very low)(Then customer is good)

(7) (IF A13 is very high)(Then customer is bad)

Figure 7. Descriptive Mamdani fuzzy rules generated from Australian data-S1

(1) *(IF Credit amount is very low) and (IF unknown/no savings account) and (IF Other debtors/guarantors: guarantor) and (IF Age is very low) and (IF Other installment plans: bank) and (IF Housing :for free)(Then customer is bad)*

(2) *(IF existing credits paid back) and (IF Purpose is radio/television) and (IF savings < 100 DM) and (IF employment since :unemployed) and (IF male: single) and (IF Property is life insurance) and (IF Housing :rent) and (IF Number of credits=1)(Then customer is good)*

(3) *(IF Purpose is car (used)) and (IF unknown/no savings account) and (IF Other debtors/guarantors: none) and (IF Property car or other) and (IF Age is very low)(Then customer is good)*

(4) *(IF Check Account ≥ 200 DM) and (IF Purpose is vacation) and (IF Credit amount is low) and (IF male: married/widowed) and (IF Other debtors/guarantors :none) and (IF Residence=4 years) and (IF Property is life insurance) and (IF Age is very low) and (IF Housing :rent)(Then customer is bad)*

(5) *(IF Check Account < 0 DM) and (IF critical account)and (IF Purpose is education) and (IF savings≥ 1000 DM) and (IF 4 ≤ employment since <7 years) and (IF Other debtors/guarantors:none) and (IF Other installment plans:bank) and (IF Housing :own) and (IF Number of credits=4)(Then customer is bad)*

(6) *(IF Duration is too short) and (IF critical account) and (IF 4≤employment since<7 years) and (IF female:divorced/separated/married) and (IF Other debtors/guarantors:none) and (IF Residence=4 years) and (IF Housing :rent) and (IF Number of credits=1) and (IF Job is unskilled - resident)(Then customer is good)*

(7) *(IF Check Account ≥ 200 DM) and (IF Duration is very long) and (IF 500 ≤ savings <1000 DM) and (IF Other debtors/guarantors:none) and (IF Age is very low)(Then customer is bad)*

(8) *(IF delay in paying off in the past) and (IF male:married/widowed) and (IF Property car or other) and (IF Age is high) and (IF Number of people=1)(Then customer is good)*

Figure 8. Descriptive Mamdani fuzzy rules generated from German data-S1

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

understand, for example the reason behind some decisions such as rejecting a credit application. In such a case, it is reasonable to trade some accuracy for extra transparency and better readability of the credit scoring model. The multi-objective genetic algorithm applied in this study can achieve a maximum trade-off between the accuracy and transparency. Hence, an adequate credit scoring system can be chosen based on the needs of the user, for example in the case where the end user wants only to conduct a data analysis to find out about the main customer attributes that influence the discrimination between the good and bad customers. In this case, the transparency which is measured by the number of selected inputs is more important than the classification accuracy and the recommended choice for him/her is to select an accuracy-transparency level where the accuracy value is acceptable while the transparency value is as high as possible (i.e. select the minimum number of inputs). The acceptable level of accuracy is the minimum level required to have a reliable data analysis while a very high level of transparency allows for better understanding of the key factors that influence the classification process. In another case in which the end user is only interested in the outcome without paying attention to the interpretation of the results, the fuzzy credit system with the highest classification accuracy value is chosen, irrespective of the number of selected inputs. The first approach (TS fuzzy system) is suitable for getting the above-mentioned choices. To further the investigation and the analysis on the relation between the attributes and the outcome, the end user needs to see the variation in the outcome when the values of certain attributes change. In such a case, the values of the fuzzy sets have to be well-defined and distinguishable so that each of the fuzzy sets can be defined using a linguistic value. The linguistic values such as low, average and high are natural and humanly understandable values and can be used as labels for the fuzzy sets to construct the fuzzy system. The second approach (Mamdani fuzzy system) is suitable for this kind of data analysis where the end user is interested in knowing not only the important attributes that contribute to the outcome but also the details on the relation between the attributes and the outcome.

Conclusions

In this paper, the transparency and accuracy of credit scoring model have been investigated using two different fuzzy model types, namely Takagi-Sugeno (TS) and Mamdani. The following conclusions have been drawn from this study.

TS fuzzy system is highly accurate and computationally efficient although lacking in transparency while Mamdani fuzzy system is highly transparent, intuitive, well suited to human input and relatively accurate. Therefore, TS fuzzy system is apparently better in predicting the customer's creditworthiness while the latter is better in data analysis and knowledge discovery. The transparency of the fuzzy systems resulting from clustering techniques is often lost during the learning of parameters and can be evaluated in two levels. The first and most important level of transparency is the capacity to represent the knowledge characterising the relations between the customers' features and creditworthiness in a natural manner, e.g. as a series of linguistic fuzzy rules. The second level is the degree of complexity of the fuzzy system which can be measured by the number of fuzzy rules in the fuzzy system, the number of input variables for each rule, and the

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

number of fuzzy sets per variable. This study illustrates a classical trade-off between accuracy and transparency, and the power of multi-objective learning to find an adequate trade-off between them so the user can choose between different levels of accuracy-transparency based on the end user's needs. This technique can also remove unnecessary complexity that may result from extracting the rules from the data set without affecting the performance of the fuzzy system. Therefore, it can be used in the pre-processing stage of a high dimensional pattern modelling as a feature selection method to reduce the number of inputs in the model. In this case the user can, based on his need, choose between different levels of number of inputs/accuracy of the model.

Acknowledgement

This research was supported by a fundamental research grant scheme from Ministry of Higher Education, Malaysia.

References

1. D. West, "Neural network credit scoring models", *Comput. Operat. Res.*, **2000**, *27*, 1131-1152.
2. L. C. Thomas, "A survey of credit and behavioral scoring: Forecasting financial risks of lending to customers", *Int. J. Forecast.*, **2000**, *16*, 149-172.
3. A. Lahsasna, R. N. Ainon and T. Y. Wah, "Credit scoring models using soft computing methods: A survey", *Int. Arab J. Inform. Technol.*, **2010***, 7,* 115-123.
4. H. L. Jensen, "Using neural networks for credit scoring", *Managerial Finance*, **1992**, *18*, 15-26.
5. V. S. Desai, J. N. Crook and G. A. Overstreet, "A comparison of neural networks and linear scoring models in the credit union environment", *Eur. J. Operat. Res.*, **1996**, *95*, 24-37.
6. B. Baesens, R. Setiono, C. Mues and J. Vanthienen, "Using neural network rule extraction and decision tables for credit-risk evaluation", *Manage. Sci.*, **2003**, *49*, 312-329.
7. S. Piramuthu, "Financial credit-risk evaluation with neural and neuro-fuzzy systems", *Eur. J. Operat. Res.*, **1999**, *112*, 310-321.
8. R. Malhotra and D. K Malhotra, "Differentiating between good credits and bad credits using neuro-fuzzy system", *Eur. J. Operat. Res.*, **2002**, *136*, 190-211.
9. F. Hoffmann, B. Baesens, J. Martens, F. Fput and J. Vanthienen, "Comparing a genetic fuzzy and a neuro-fuzzy classifier for credit scoring", *Int. J. Intell. Syst.*, **2002**, *17*, 1067-1083.
10. D. Nauck, U. Nauck and R. Kruse, "Generating classification rules with neuro-fuzzy system NEFCLASS", Proceedings of the Biennial Conference of the North American Fuzzy Information Processing Society, **1996**, Berkeley, CA, USA, pp.466-470.
11. F. Herrera, "Genetic fuzzy systems: taxonomy, current research trends and prospects", *Evol. Intell.*, **2008**, *1*, 27-46.

12. H. Ishibuchi, "Multiobjective genetic fuzzy systems: review and future research directions", Proceedings of the 2007 IEEE International Conference on Fuzzy Systems, **2007**, London, UK, pp. 913–918.
13. A. Asuncion and D. J. Newman, "UCI Machine Learning Repository", School of Information and Computer Science, University of California, Irvine, CA, http://www.ics.uci.edu/~mlearn/ MLRepository.html. (Retrieved: May 15, **2007**)
14. R. Quinlan, "C4.5: Programs for Machine Learning", Morgan Kaufmann, San Fransisco, **1992**.
15. A. Konak, D. W. Coit and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial", *Reliab. Eng. Syst. Saf.*, **2006**, *91*, 992-1007.
16. C. A. C. Coello, "A comprehensive survey of evolutionary-based multi-objective optimization techniques", *Knowl. Inform. Syst.*, **1999**, *1*, 269-308.
17. D. A. Van Veldhuizen and G. B. Lamont, "Multi-objective evolutionary algorithms: analyzing the state-of-the-art", *Evol. Comput.*, **2000**, *8*, 125-147.
18. K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multi-objective genetic algorithm : NSGA-II," *IEEE Trans. Evol. Comput.*, **2002**, *6*, 182-197.
19. N. Srinivas and K. Deb, 'Multi-objective optimization using non-dominated sorting in genetic algorithms", *Evol. Comput.*, **1994**, *2*, 221-248.
20. K. Deb and T. Goel, "Controlled elitist non-dominated sorting genetic algorithms for better convergence", Proceedings of the 1st International Conference on Evolutionary Multi-Criterion Optimization, **2001**, Berlin, Germany, pp. 67-81.
21. S. L. Chiu, "Fuzzy model identification based on cluster estimation", *J. Intell. Fuzzy Syst.*, **1994**, *2*, 267-278.
22. T. Takagi and M. Sugeno, "Fuzzy identification of systems and its application to modeling and control", *IEEE Trans. Syst. Man Cybern.*, **1985**, *15*, 116-132.
23. The MathWorks, "Matlab Global Optimization Toolbox 3, User's Guide", The MathWorks Inc., Natick, MA, **2010**.
24. E. H. Mamdani, "Applications of fuzzy logic to approximate reasoning using linguistic synthesis", *IEEE Trans. Comput.*, **1977**, *26*, 1182-1191.
25. J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algoritms", Plenum Press, New York, **1981**.
26. H. Ishibuchi, T. Nakashima and T. Murata, "Three-objective genetics-based machine learning for linguistic rule extraction", *Inform. Sci.*, **2001**, *136,* 109-133.
27. R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection", Proceedings of International Joint Conference on Artificial Intelligence, **1995**, Montreal, Canada, pp.1137-1143.
28. M. Setnes, R. Babuska, U. Kaymak and H. R. van Nauta Lemke, "Similarity measures in fuzzy rule based simplification", *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, **1998**, *28*, 376-386.
29. C.-L. Huang, M.-C. Chen and C.-J. Wang, "Credit scoring with a data mining approach based on support vector machines", *Expert Syst. Appl.*, **2007**, *33*, 847-856.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

30. C. S. Ong, J. J. Huang and G. H. Tzeng, "Building credit scoring models using genetic programming", *Expert Syst. Appl.*, **2005**, *29*, 41-47.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

Appendix

Table 9. Attributes for German credit data set

No.	Attribute	Type	Value
1	Status of existing checking account	Categorical	0 : ... < 0 DM ; 1 : 0 <= ... < 200 DM 2 : ... >= 200 DM /salary assignments for at least 1 year 3 : no checking account
2	Duration in month	Continuous	[4 72]
3	Credit history	Categorical	0 : no credits taken/all credits paid back duly; 1 : all credits at this bank paid back duly; 2 : existing credits paid back duly till now; 3 : delay in paying off in the past; 4 : critical account/ other credits existing (not at this bank)
4	Purpose	Categorical	0: car (new); 1 : car (used); 2 : furniture/equipment; 3 : radio/television; 4 : domestic appliances; 5 : repairs; 6 : education; 7 : (vacation - does not exist?); 8 : retraining; 9 : business; 10 : others
5	Credit amount	Continuous	[250 18424]
6	Savings account/bonds	Categorical	0: ... < 100 DM; 1 : 100 <= ... < 500 DM; 2: 500 <= ... < 1000 DM; 3: .. >= 1000 DM; 4:unknown/ no savings account
7	Present employment since	Categorical	0 : unemployed; 1 : ... < 1 years; 2 : 1 <= ... < 4 years ; 3 : 4 <= ... < 7 years; 4 : .. >= 7 years
8	Installment rate	Continuous	[1 4]
9	Personal status and sex	Categorical	0 : male : divorced/separated; 1 : female divorced/ separated /married; 2 : male: single; 3 : male : married/widowed; 4 : female : single
10	Other debtors / guarantors	Categorical	0 : none; 1 : co-applicant; 3 : guarantor
11	Present residence since	Continuous	[1 4]
12	Property	Categorical	0 : real estate; 1:if not 0 : building society savings agreement/ life insurance; 2: if not 0/1 : car or other, not in attribute 6; 3 : unknown / no property
13	Age in years	Continuous	[19 75]
14	Other installment plans	Categorical	0 : bank; 1 : stores; 2 : none
15	Housing	Categorical	0 : rent; 1 : own; 2 : for free
16	Number of existing credits at this bank	Continuous	[1 4]
17	Job	Categorical	0 : unemployed/ unskilled - non-resident; 1 : unskilled – resident; 2 : skilled employee / official; 3 : management/ self-employed/ highly qualified employee/ officer
18	Number of depends	Continuous	[1 2]
19	Telephone	Categorical	0 : none; 1 : yes, registered under the customers name
20	Foreign worker	Categorical	0 : yes; 1 : no

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 136-158

Table 10. Attributes for Australian credit data set

No.	Attribute	Type	Value
1	A1	Categorical	0,1
2	A2	Continuous	[13.75 80.25]
3	A3	Continuous	[0 25.125]
4	A4	Categorical	1,2,3
5	A5	Categorical	1, 2,3,4,5,6 ,7,8,9,10,11, 12, 13,14
6	A6	Categorical	1, 2,3, 4,5,6,7,8,9
7	A7	Continuous	[0 20]
8	A8	Categorical	1,0
9	A9	Categorical	1,0
10	A10	Continuous	[0 23]
11	A11	Categorical	1,0
12	A12	Categorical	1,2,3
13	A13	Continuous	[0 2000]
14	A14	Continuous	[1 100001]

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

Maejo International Journal of Science and Technology

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

Full Paper

Semiclassical three-valley Monte Carlo simulation analysis of steady-state and transient electron transport within bulk $InAs_xP_{1-x}$, InAs and InP

Hadi Arabshahi* , Mahmoud R. Rokn-Abadi and Fatemeh Badieyan

Physics Department, Ferdowsi University of Mashhad, Mashhad, Iran

* Corresponding author, e-mail: arabshahi@um.ac.ir

Received: 2 December 2009 / Accepted: 20 April 2010 / Published: 20 April 2010

Abstract: We have studied how electrons, initially in thermal equilibrium, drift under the action of an applied electric field within bulk zincblende $InAs_xP_{1-x}$, InAs and InP. Calculations are made using a non-parabolic effective-mass energy band model. Monte Carlo simulation includes all of the major scattering mechanisms. The band parameters used in the simulation are extracted from optimised pseudo-potential band calculations to ensure excellent agreement with experimental information and ab-initio band models. The effects of alloy scattering on the electron transport physics are examined. For all materials, it is found that electron velocity overshoot only occurs when the electric field is increased to a value above a certain critical field, unique to each material. This critical field is strongly dependent on the material parameters. Transient velocity overshoot has also been simulated, with the sudden application of fields up to 1600 kVm^{-1}, appropriate to the gate-drain fields expected within an operational field-effect transistor. The electron drift velocity relaxes to the saturation value of about 1.5×10^5 ms^{-1} within 4 pico-seconds for all crystal structures. The steady-state and transient velocity overshoot characteristics are in fair agreement with other recent calculations.

Keywords: Monte Carlo simulation, steady-state electron transport, transient electron transport, pseudo-potential, alloy scattering, velocity overshoot, critical field

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

Introduction

InP and InAs offer the prospect of mobility comparable to GaAs and are increasingly being developed for the construction of optical switches and optoelectronic devices. While GaAs has been extensively studied [1-3], InAs and InP and alloys constructed from them such as $InAs_xP_{1-x}$ have yet to be examined to the same extent. Alloys of InAs and InP have unfortunately proved to be difficult material to work with in practice and very little experimental work on $InAs_xP_{1-x}$ material and devices has been done because of technical problems in forming Schottky contacts with sufficiently high barrier potentials. Nevertheless, some experimental work has been done on other types of InAs and InP field-effect transistor, most notably MISFETs [4-5], and there is every reason to be optimistic that some form of heterojunction under the gate may well overcome the problem of low barrier.

Improved electron transport properties are one of the main targets in the ongoing study of binary and ternary InP, InAs and $InAs_xP_{1-x}$ materials. The Monte Carlo technique has proved valuable for studying non-equilibrated carrier transport in a range of semiconductor materials and devices [6-7]. However, carrier transport modelling of InP and InAs materials has only recently begun to receive sustained attention, now that the progress in compounds and alloys has resulted in the production of valuable materials for the electronics industry. In this communication we present Monte Carlo calculation of steady-state and transient electron transport conditions in InP, InAs and $InAs_xP_{1-x}$. We demonstrate the effect of injection energy and electric field on the transient electron transport. The differences in transport properties are analysed in terms of important material parameters.

Our current approach employs a one-dimensional ensemble Monte Carlo technique to investigate steady-state and transient electron transport in InP, InAs and $InAs_xP_{1-x}$. However, the momentum space treatment is three-dimensional and the scattering events consider all three dimensions. Specifically, our model includes the three lowest valleys of the conduction band with non-parabolicity. Details of the conduction band parameters and the Monte Carlo simulation are presented in the next section and the results of steady-state and transient transport simulations are discussed afterwards.

Model Details

Our ensemble Monte Carlo simulations of electron transport in zincblende InP, InAs and InAsxP1-x are similar to those of Arabshahi et al [8-9]. As indicated earlier, a three-valley model for the conduction band is employed.

In order to calculate the electron drift velocity for large electric fields, consideration of conduction band satellite valleys is necessary. The first-principle band structure of zincblende InAs, InP and $InAs_xP_{1-x}$ predicts a direct band gap located at the Γ point and lowest-energy conduction band satellite valleys at the *X* point and *L* point. In our Monte Carlo simulation, the Γ valley, the three equivalent *X* valleys and the four equivalent *L* valleys are represented by ellipsoidal, non-parabolic dispersion relationships of the following form [10-12]:

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

$$E(k)[1+\alpha_i E(k)] = \frac{\hbar^2 k^2}{2m^*} \quad (1)$$

where m^* is effective mass at the band edge and α_i is the non-parabolicity coefficient of the *i*-th valley given by Kane model [13] as

$$\alpha_i = \frac{1}{E_g}\left[1-\frac{2m^*}{m_0}\right]\left[1-\frac{E_g\Delta}{3(E_g+\Delta)(E_g+2\Delta/3)}\right] \quad (2)$$

where E_g is the band-gap energy and Δ is the spin-orbit splitting. We assume that all donors are ionised and that the free-electron concentration is equal to the dopant concentration. For each simulation, the motion of ten thousand electron particles are examined, the temperature being set to 300 K and the doping concentration being set to 10^{17} cm^{-3}. In the case of the ellipsoidal, non-parabolic conduction valley model, the usual Herring-Vogt transformation matrices are used to map carrier momenta into spherical valleys when particles are drifted or scattered. Electrons in bulk material suffer intravalley scattering by polar optical, non-polar optical and acoustic phonon scattering, intervalley phonons, and ionised impurity scattering.

Acoustic scattering is assumed elastic and the absorption and emission rates are combined under the equipartition approximation, which is valid for lattice temperature above 77 K. Elastic ionised impurity scattering is described using the screened Coulomb potential of the Brooks-Herring model. Band edge energies, effective masses and non-parabolicity are derived from empirical pseudo-potential calculations. Important parameters used throughout the simulations are listed in Tables 1-2.

Table 1. Valley parameter selections for InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ [3-5]

	InAs	InP	$InAs_{0.2}P_{0.8}$	$InAs_{0.8}P_{0.2}$
m_Γ	5667	4810	4981	5495
m_L	4280	5300	5096	4484
m_x	14.6	12.4	12.84	14.16
α_Γ	12.25	9.55	10.09	11.71
α_L	4.9	8.3	7.62	5.58
Γ-X	0.015	0.06	0.05	0.024
Γ-L	1	1	1	1

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

Table 2. Material parameter selections for InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ [3-5]

	InAs	InP	$InAs_{0.2}P_{0.8}$	$InAs_{0.8}P_{0.2}$
Density ρ (kgm^{-3})	5667	4810	4981	5495
Longitudinal sound velocity v_s (ms^{-1})	4280	5300	5096	4484
Low-frequency dielectric constant ε_s	14.6	12.4	12.84	14.16
High-frequency dielectric constant ε_∞	12.25	9.55	10.09	11.71
Acoustic deformation potential D(eV)	4.9	8.3	7.62	5.58
Polar optical phonon energy (eV)	0.015	0.06	0.05	0.024
Intervalley deformation potential (eVm^{-1})	1	1.1	1.05	1.065
Intervalley phonon energies (meV)	11.2	29	25.44	14.76

Results

Figure 1 shows the simulated velocity-field characteristics of zincblende InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ semiconductors at 300 K, with a background doping concentration of 10^{17} cm^{-3} and with the electric field applied along one of the cubic axes. The simulations suggest that the peak drift velocity for zincblende InAs is 3.4 $\times 10^5$ ms^{-1} while that for InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ are about ~ 2.3 $\times 10^5$ ms^{-1}, 2.5×10^5 ms^{-1} and 3.2×10^5 ms^{-1} respectively. At higher electric fields, intervalley optical phonon emission dominates, causing the drift velocity to saturate at around 1.5 $\times 10^5$ ms^{-1} for all materials.

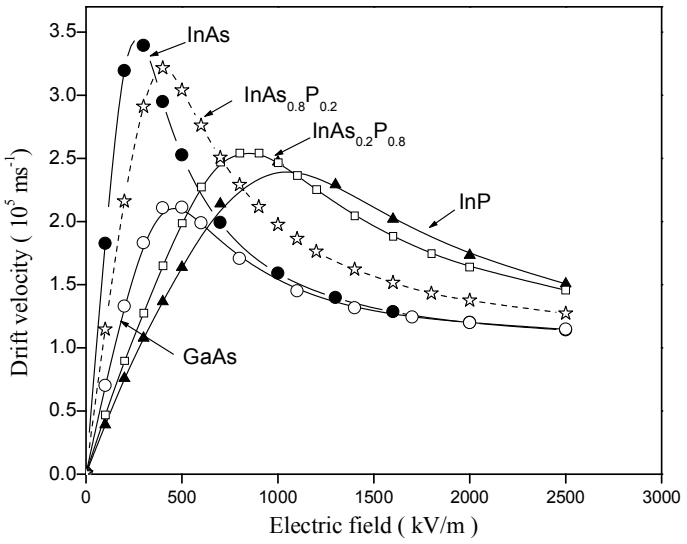


Figure 1. Calculated steady-state electron drift velocity in bulk zincblende InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ at room temperature as per the non-parabolic band model

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

The calculated drift velocities apparent from Figure 1 are fractionally lower than those that have been calculated by Adachi [14-16], who assumed an effective mass in the upper valleys was equal to the free electron mass. The threshold field for the onset of significant scattering into satellite conduction band valleys is a function of the intervalley separation and the density of electronic states in the satellite valleys.

The valley occupancy for the Γ, X and L valleys is illustrated in Figure 2 and shows that the inclusion of the satellite valleys in the simulation is important. Significant intervalley scattering into the satellite valleys occurs for fields above the threshold field for each material. This is important because electrons which are near a valley minimum have small kinetic energy and are therefore strongly scattered. It is apparent that intervalley transfer is substantially larger in InAs over the range of applied electric fields shown, due to the combined effect of a lower Γ effective mass, lower satellite valley separation energy, and a slightly lower phonon scattering rate within the Γ valley.

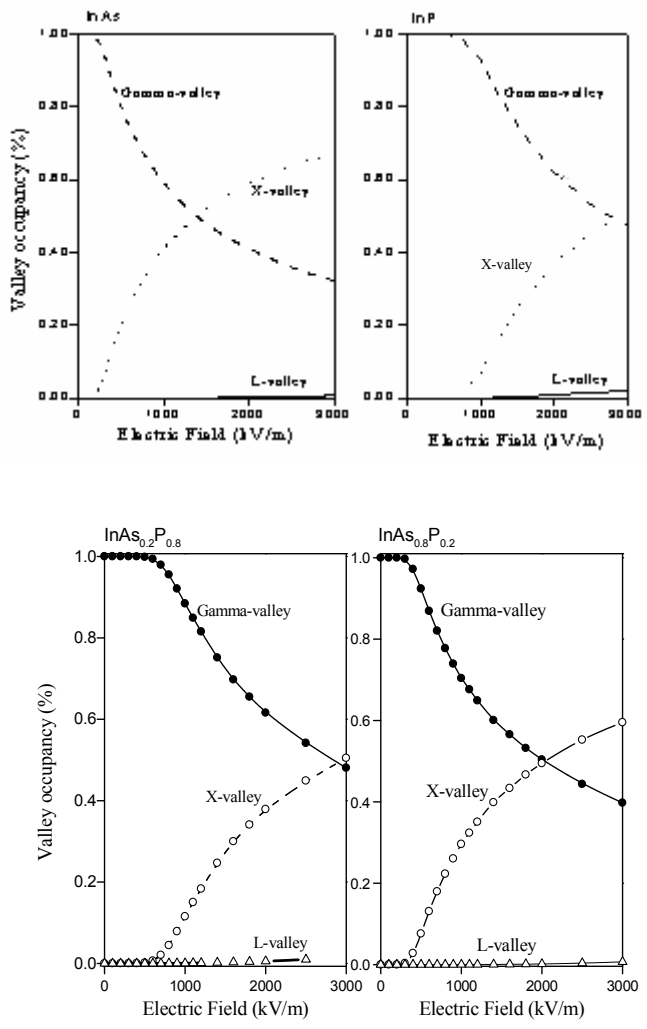


Figure 2. Fractional occupation of the central Γ and satellite valleys of zincblende InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ as a function of applied electric field at room temperature as per the non-parabolic band model

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

We have also examined transient electron transport in bulk InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ semiconductors. The transient response of electrons in these materials are compared in Figure 3 for fields up to 1600 kVm^{-1} strength. In InAs, very little or no overshoot occurs below the threshold field of 400 kVm^{-1}. As the electric field strength is increased to a value above the threshold field, overshoot begins to occur. As the field strength is increased further, both the peak overshoot velocity increases and the time for overshoot relaxation decreases.

In InAs, the velocity overshoot initially increases more rapidly with increasing electric field due to the lower Γ valley effective mass. For example, at 1600 kVm^{-1}, the maximum overshoot velocity for InAs is about 8×10^5 ms^{-1}, whereas for InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ it is about 4×10^5 ms^{-1}, 5×10^5 ms^{-1} and 7×10^5 ms^{-1} respectively. It is found also that for the same value of electric field above the threshold value, the electron drift velocity is always smaller in InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ than in InAs.

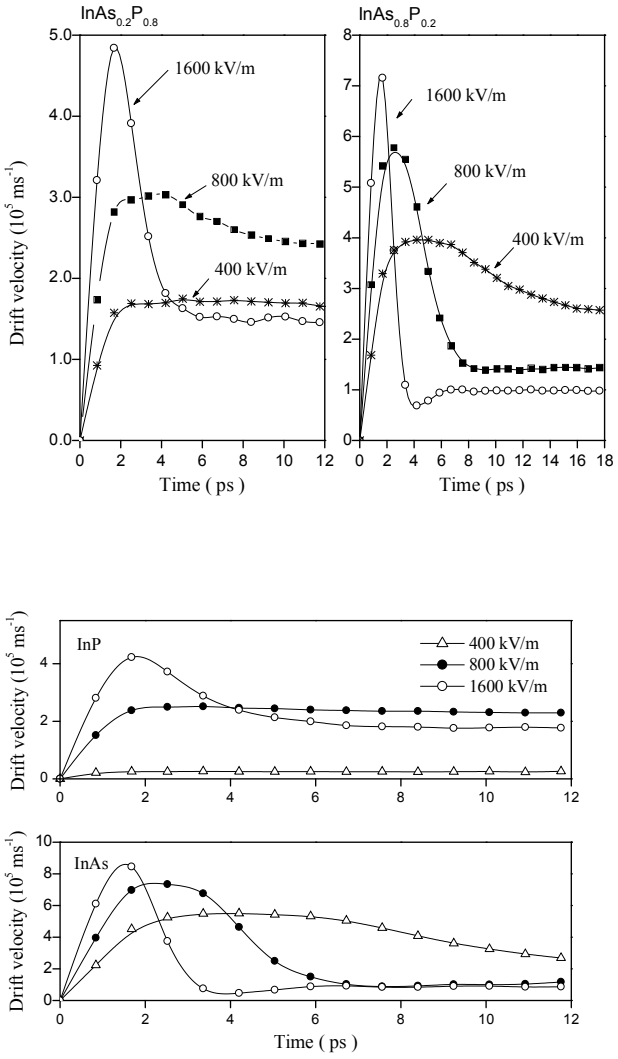


Figure 3. A comparison of the velocity overshoot effect exhibited by InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ semiconductors as calculated by Monte Carlo simulation. The donor concentration is 10^{17} cm^{-3} and the temperature is 300 K.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

Figure 4 shows the average velocity of electrons in InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ as a function of distance. We note that for the applied field of 400 to 1600 kVm^{-1} the average electron velocity reaches steady-state very quickly with little or no velocity overshoot. It is suggested that in $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ 400 kVm^{-1} is the critical field for the onset of velocity overshoot. As mentioned above, 400 kVm^{-1} also corresponds to the peak in the velocity-field characteristic associated with $InAs_{0.2}P_{0.8}$. Steady-state Monte Carlo simulations suggest that this is the point at which significant upper valley occupation begins to occur, as shown in Figure 2. This signifies that velocity overshoot is related to the transfer of electrons to the upper valleys. To optimise device performance, we have to minimise the transit time over a given distance.

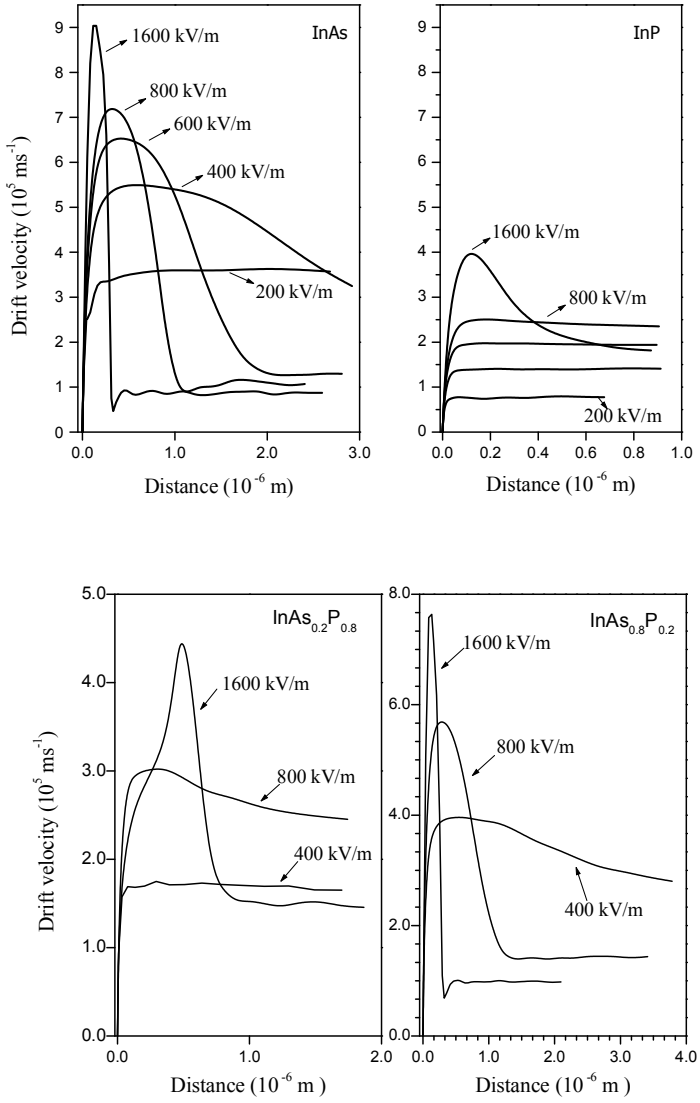


Figure 4. A comparison of the average electron velocity as a function of the displacement for various applied fields in InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ semiconductors. In all cases, an initial zero field distribution, a crystal temperature of 300 K and a doping concentration of 10^{17} cm^{-3} are assumed.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

From Figure 5, it can be seen that there is a trade-off between the peak overshoot velocity and the distance taken to achieve steady state. In particular, when the applied electric field is set to 1600 kVm^{-1}, the peak overshoot velocity of $InAs_{0.8}P_{0.2}$ is 7.8×10^5 ms^{-1} while the corresponding steady-state drift velocity, 1.5×10^5 ms^{-1}, is achieved after just 0.2 μm.

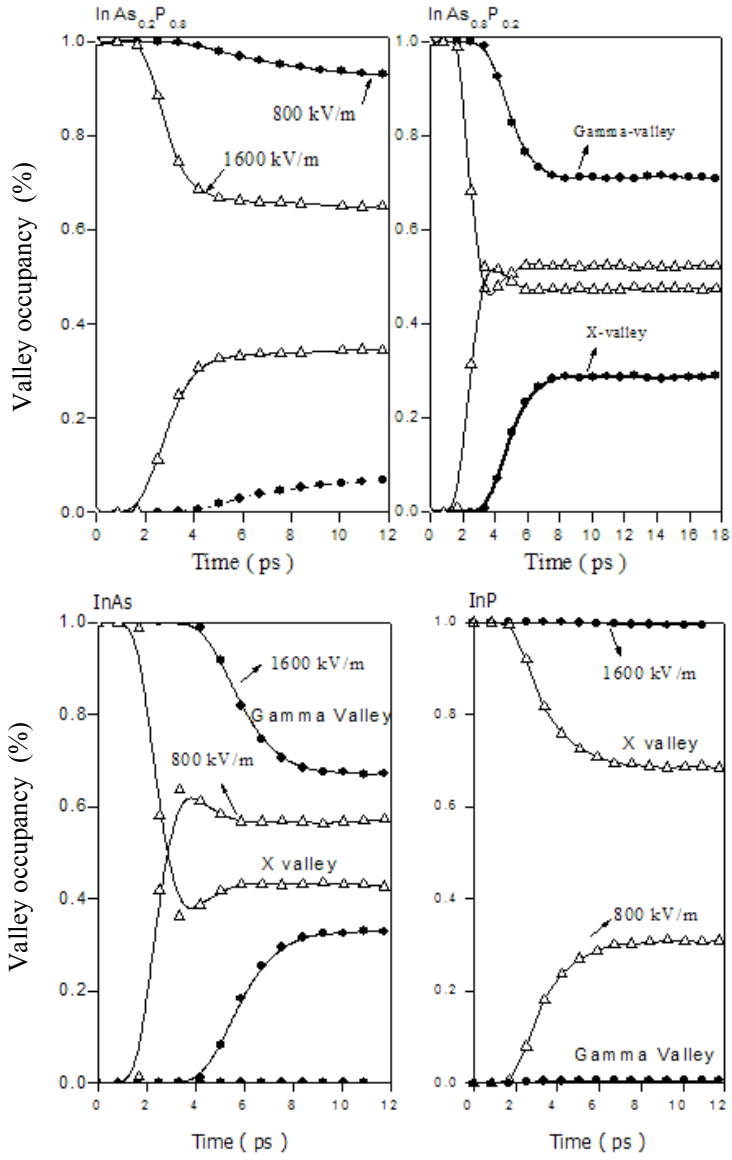


Figure 5. Comparison of the valley occupancy as a function of applied electric field in InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ for Γ, *X* and *L* valleys at room temperature

However, for an applied field of 400 kVm^{-1}, just above the critical field, the peak overshoot velocity is only 4×10^5 ms^{-1} and it takes longer (about 0.8 μm) to achieve the corresponding steady-state drift velocity of 3×10^5 ms^{-1}. Similar results are noted for $InAs_{0.2}P_{0.8}$, as is seen in Figure 5. In particular, the critical field denoting the onset of velocity overshoot coincides almost exactly with the field at which the peak drift velocity in the steady-state velocity field characteristic is found, i.e. 200 kVm^{-1} for InAs and 800 kVm^{-1} for InP. The correspondence

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

between the critical field at which the onset of velocity overshoot effect occurs and the peak in the steady-state velocity field characteristic appears to be valid for the case of other III-V semiconductors as well.

Conclusions

Electron transport at 300 K in bulk zincblende InAs, InP, $InAs_{0.2}P_{0.8}$ and $InAs_{0.8}P_{0.2}$ has been simulated using an ensemble Monte Carlo simulation. Using valley models to describe the electronic bandstructure, calculated velocity-field characteristics are in fair agreement with other calculations. Saturation drift velocities ($\sim 1.5\times10^5$ ms^{-1}) match recent measurements on low-doped bulk samples. The velocity-field characteristics of the materials show similar trends, reflecting the fact that all the semiconductors have satellite-valley effective densities of states several times greater than the central Γ valley. However, the peak velocity in $InAs_{0.2}P_{0.8}$ occurs at a field ~ 700 kVm^{-1}, twice larger than that for $InAs_{0.8}P_{0.2}$. This is a consequence of the large Γ valley effective mass in $InAs_{0.2}P_{0.8}$ structure. This reduced valley effective mass in $InAs_{0.8}P_{0.2}$ permits substantial population of the upper valleys and velocity saturation at far lower electron temperature than that for InP.

Acknowledgement

We would like to thank M. G. Paezi for her useful comments.

References

1. K. Brennan, K. Hess, J. Y. Tang and G. J. Iafrate, "High field electron transport properties", *IEEE Trans. Electron Devices*, **1983**, *30*, 1750-1755.
2. N. Newman, T. Kendelewicz, L. Bowman and W. E. Spicer, "Electron transport properties in nitrides semiconductor devices", *Appl. Phys. Lett.,* **1985**, *46*, 1176-1181.
3. N. Newman, V. Schilfgaarde, T. Kendelewicz and W. E. Spicer, "Monte Carlo studies in two dimensional GaN/AlGaN devices", *Mater. Res. Soc. Symp. Proc.,* **1986**, *54*, 443-449.
4. D. C. Cameron, L. D. Irving and C. R. Whitehouse, "Hot carrier relaxation in InP and GaAs", *Thin Solid Films.,* **1983**, *103*, 61-66.
5. D. C. Cameron, L. D. Irving and C. R. Whitehouse, "Ensemble Monte Carlo simulation in InP and InN", *Electron. Lett.*, **1982**, *18*, 534-538.
6. C. Moglestue, "Monte Carlo Simulation of Semiconductor Devices", Chapman and Hall, New York, **1993**.
7. C. Jacoboni and P. Lugli, "The Monte Carlo Method for Semiconductor and Device Simulation", Springer-Verlag, New York, **1989**.

Maejo Int. J. Sci. Technol. **2010**, *4*(01), 159-168

8. H. Arabshahi, M. R. Benam and B. Salahi, "Comparison of high field and steady-state transient electron transport properties in GaN, InN and AlN", *Modern Phys. Lett. B.*, **2007**, *21*, 1715-1721.
9. H. Arabshahi, "Comparison of high field steady state transport properties in zincblende and wurtzite GaN", *Modern Phys. Lett. B.*, **2007**, *21*, 199-206.
10. B. E. Foutz, L. F. Eastman, U. V. Bhapkar and M. Shur, "Full band Monte Carlo simulation of zincblende GaN MESFET's including realistic impact ionization rates", *Appl. Phys. Lett.,* **1997**, *70*, 2849-2854.
11. U. V. Bhapkar and M. S. Shur, "Ensemble Monte Carlo study of electron transport in wurtzite InN", *J. Appl. Phys.,* **1997**, *82*, 1649-1654.
12. J. D. Albrecht, R. P. Wang, P. P. Ruden and K. F. Brennan, "Monte Carlo simulation of GaN in zincblende and wurtzite structures", *J. Appl. Phys.,* **1998**, *83*, 2185-2190.
13. E. O. Kane, "Band structure calculation in group III and IV materials", *J. Phys. Chem. Solids,* **1957**, *1*, 249-253.
14. S. Adachi, "GaAs and Related Materials, Bulk semiconducting and Superlattice Properties", World Scientific, Singapore, **1994**.
15. S. Adachi, "Physical Properties of III-V Semiconductor Compounds, InP, InAs, GaAs, GaP,AlAs and AlGaAs", Wiley, New York, **1992**.
16. S. Adachi, "Properties of AlGaAs", INSPEC, Stevenage (UK), **1993**.