ISSN 1913-1844

MODERN APPLIED SCIENCE

Vol. 3, No. 5
May 2009

Canadian Center of Science and Education

Editorial Board

Ahmad Mujahid Ahmad Zaidi	Universiti Tun Hussein Onn Malaysia, Malaysia
Hamimah Adnan	Universiti Teknologi MARA, Malaysia
J S Prakash	Sri Bhagawan Mahaveer Jain College of Engineering, India
Lim Hwee San	Universiti Sains Malaysia, Malaysia
Moussaoui Abdelkrim	University of Guelma, Algeria
Musa Mailah	Universiti Teknologi Malaysia, Malaysia
Stefanos Dailianis	University of Patras, Greece
Sujatha. C.H	Cochin University of Science and Technology, India
Sundus H Ahmed	Ministry of Science and Technology, Iraq
Susan Sun	Canadian Center of Science and Education, Canada
Sutopo Hadi	University of Lampung, Indonesia

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Contents

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Contents

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Estimating a Nonlinear Mixed Volume-age Model with and without Taking into Account Serially-correlated Errors: Differences and Implications

Shongming Huang, Shawn X. Meng & Yuqing Yang

Biometrics Unit, Forest Management Branch

Alberta Department of Sustainable Resource Development

8th Floor, 9920-108 Street, Edmonton, Alberta T5K 2M4, Canada

Tel: 1-780-422-5281 E-mail: shongming.huang@gov.ab.ca

Abstract

In this study we estimated a lodgepole pine (*Pinus contorta* var. *latifolia* Engelm.) volume-age model with and without taking into account serially-correlated errors arisen from permanent sample plots. The estimations were based on the first-order (FO) and first-order conditional expectation (FOCE) methods of the nonlinear mixed model technique. Among the correlated error structures considered, the spatial power structure was found to be the most appropriate. Model predictions were obtained and evaluated on model fitting data, as well as on independent validation data collected from a different ecoregion. Results showed that the model estimated with the independent and identically distributed (iid) error structure performed much better than the model estimated with the correlated error structure. This is true on both model fitting and validation data sets, and for both FO and FOCE methods. It implies that, if the main purpose of a study is to develop models for predictions, there is no real benefit to consider more elaborate and complex error structures to account for the correlated errors. The iid error structure is a sound choice for dealing with correlated errors under the nonlinear mixed model framework.

Keywords: Longitudinal data, Error structure, Correlation, Model prediction, First-order (FO) method, First-order conditional expectation (FOCE) method, Volume-age model

1. Introduction

For repeatedly measured, unequally spaced and unbalanced longitudinal data collected from permanent sample plots (or stem analysis trees), various within- and between-plot correlations and heterogeneous variances may occur. In addition, the correlations can be both temporal and spatial.

Violation of the independent and identically distributed (iid) error assumption could have important statistical consequences for analyses based on the least squares principle. It may invalidate hypothesis testing and interval estimation. The problem may be solved using the generalized least squares (GLS) method (Judge et al. 1985) or mixed model technique (Davidian and Giltinan 1995). Examples of GLS in forestry have been shown by many (e.g., Monserud 1984, Gregoire 1987). One issue related to GLS is that it generally applies to equally spaced and balanced data. It may create other problems when used for unequally spaced and unbalanced data. Lappi and Bailey (1988) stated that ordinary least squares (OLS) are "deemed to be appropriate" for large samples, even though GLS could be used. Borders et al. (1988), West (1995), and Huang (1997) cited other studies which concluded that OLS is adequate, and it is unnecessary to apply more elaborate procedures when faced with correlated errors. Knowledge of the correlation parameter in GLS is typically ignored in predictions (Monserud 1984).

In more recent years, more researchers have chosen the mixed model approach to deal with correlated errors and heterogeneous variances (e.g., Lappi 1986, Ojansuu 1987, Gregoire et al. 1995, Huang 1997, Fang and Bailey 2001, Schabenberger and Pierce 2002, Wang et al. 2007). This approach allows for a more flexible error covariance structure. More importantly, knowledge of the correlation parameter and the entire error covariance structure from the mixed model estimation can be used in predictions.

In spite of the large amount of effort invested in addressing correlated errors, many questions remain. The primary objectives of this study are to: (1) compare a volume-age model for lodgepole pine (*Pinus contorta* var. *latifolia*

Engelm.) estimated using nonlinear mixed model (NLMM) technique with and without taking into account correlated errors; (2) evaluate the impacts of different error structures on model predictions based on model fitting data as well as on independent model validation data; (3) clarify some methodological confusions and miscues about the NLMM technique in forest modeling; and (4) recommend a sound and practical approach to use when faced with correlated errors.

2. Data and base model

The lodgepole pine data used in this study were collected by the Alberta Forest Service over the last 45 years as a part of the provincial Permanent Sample Plots (PSP) system. They were obtained from pine dominated forests in the upper foothills and lower foothills ecoregions of Alberta. The most common PSP size is 1000 m^2, but it can range from 200 to 2000 m^2 depending on the density of a stand. Within each PSP, diameters (always at 1.3 m above ground) were measured for all trees taller than 1.3 m, and heights were measured for all trees with diameters or portions of representative trees (Alberta Forest Service 2000). The missing tree heights were predicted using the "bias-free" height-diameter models (Huang et al. 2008a). Individual tree volumes were calculated following the standard taper functions (Huang 1994), and were summarized to give plot volume. Plot age (at breast height) was obtained from cored trees within the plot or sectioned trees in the buffer areas around the plot.

The volume-age data from the upper foothills ecoregion were used as model fitting data, and the data from the lower foothills ecoregion were used as model validation data. The two data sets (Figure 1) can be considered independent of each other because of the geographic separation. The model fitting data consist of 1,249 measurements from 288 plots (subjects), with a mean (minimum-maximum) volume (m^3/ha) of 302.694 (13.909-642.471) and a standard deviation (SD) of 108.902 (m^3/ha). The validation data consist of 1,011 measurements from 265 plots, with a mean volume of 287.611 (1.748-564.329) and an SD of 108.581. Both data sets are unequally spaced and unbalanced longitudinal data.

Based on a detailed examination of the overall and subject-specific (SS) volume-age trends exhibited by the data in Figure 1, two NLMMs were selected for evaluation:

(1) $\mathbf{Vol}_i = (b_1 + u_{1i})\mathbf{Age}_i^{(b_2+u_{2i})} \exp[-(b_1 + u_{1i})\mathbf{Age}_i] + \boldsymbol{\varepsilon}_i$

(2) $\mathbf{Vol}_i = (b_1 + u_{1i})\mathbf{Age}_i^{(b_2+u_{2i})} \exp[-(b_3 + u_{3i})\mathbf{Age}_i] + \boldsymbol{\varepsilon}_i$

where $\mathbf{Vol}_i$ and $\mathbf{Age}_i$ are observed volumes and ages for the ith plot, b_1, b_2 and b_3 are fixed parameters common to all plots, u_{1i}, u_{2i} and u_{3i} are random parameters unique to the ith plot, and $\boldsymbol{\varepsilon}_i$ is a normally distributed within-plot error term. Both (1) and (2) can be sigmoidal. They can also describe declining volumes at old ages, which were observed in some plots likely due to mortality exceeding growth at old ages (Figure 1). Preliminary analyses indicated that (2) failed to achieve convergence in some cases because of over-parameterization. Therefore, (1) was chosen in this study for further analyses.

3. Error structure and parameter estimation

Model (1) can be written as a generalized NLMM of the form:

(3) $\mathbf{y}_i = f(\mathbf{x}_i, \mathbf{b}, \mathbf{u}_i) + \boldsymbol{\varepsilon}_i$

where $\mathbf{y}_i = [y_{i1}, y_{i2}, ..., y_{in_i}]'$ is a vector of plot volumes for plot i, $\mathbf{x}_i$ is a design matrix of the covariate age, $\mathbf{b}$ is a vector of fixed parameters, $\mathbf{u}_i = [u_{1i}, u_{2i}]'$ is a vector of subject-specific random parameters, and $\boldsymbol{\varepsilon}_i$ is a vector of within-plot errors. The $\mathbf{u}_i$ and $\boldsymbol{\varepsilon}_i$ are assumed to be uncorrelated and normally distributed with mean zero and variance-covariance matrices $\mathbf{D}$ and $\mathbf{R}_i$, respectively. The $\mathbf{D}$ is typically assumed to be unstructured and identical for all subjects within the population:

(4) $$\mathbf{D} = \begin{bmatrix} \sigma^2_{u_1} & \sigma_{u_1u_2} \\ \sigma_{u_1u_2} & \sigma^2_{u_2} \end{bmatrix}$$

where $\sigma^2_{u_1}$, $\sigma^2_{u_2}$, and $\sigma_{u_1u_2}$ are variances and covariance for random parameters u_1 and u_2.

As noted earlier, for the unequally spaced and unbalanced longitudinal data collected from the PSPs, the errors may not be iid. Instead, they could be correlated and unequally varied. To account for these, a total of more than 20 error covariance structures described in SAS Institute Inc. (2004) and Littell et al. (2006) were examined. The spatial power error structure produced the most reasonable fit to the data. It was chosen for $\mathbf{R}_i$. As an example, for a plot repeatedly measured five times, the $\mathbf{R}_i$ is defined as:

(5) $$\mathbf{R}_i = \sigma^2\boldsymbol{\Psi} = \sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} & \rho^{d_{15}} \\ \rho^{d_{21}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} & \rho^{d_{25}} \\ \rho^{d_{31}} & \rho^{d_{32}} & 1 & \rho^{d_{34}} & \rho^{d_{35}} \\ \rho^{d_{41}} & \rho^{d_{42}} & \rho^{d_{43}} & 1 & \rho^{d_{45}} \\ \rho^{d_{51}} & \rho^{d_{52}} & \rho^{d_{53}} & \rho^{d_{54}} & 1 \end{bmatrix} = \text{SP(POW)} = \text{power}$$

where σ^2 is the overall error variance, $\mathbf{\Psi}$ is a generalized error covariance structure, $d_{\iota\eta}$ is the distance between two measurements at times ι and η ($d_{\iota\eta} = | t_\iota - t_\eta |$), and ρ is the correlation parameter. The spatial power error structure captured the unequally spaced and unbalanced nature of the volume-age data. Other error structures examined but not presented here, e.g., compound symmetry, heterogeneous Toeplitz, Toeplitz with different bands, unstructured correlation, first-order ante-dependent, and several spatial correlation structures evaluated in Yang and Huang (2008), were not converged or performed poorer than the SP(POW). The time-series structures such as AR(1), MA(2) and ARMA(1,1), which assume that the measurements are taken at equally spaced time intervals, may not be adequate without further adjustments, even though some of them produced seemingly reasonable results.

For models assumed to follow an iid error structure, the $\mathbf{R}_i$ is simplified to: $\mathbf{R}_i = \sigma^2\mathbf{\Psi} = \sigma^2\mathbf{I}_{n_i}$, where $\mathbf{I}_{n_i}$ is a $n_i \times n_i$ identity matrix of 1s.

Different methods have been developed to estimate the generalized NLMM (3). They are described in detail in Davidian and Giltinan (1995), and Vonesh and Chinchilli (1997). Since the estimation of the fixed parameters is readily available in most statistical software packages, we focused on the estimation of subject-specific random parameters based on the first-order (FO) method of Beal and Sheiner (1982) and the first-order conditional expectation (FOCE) method of Lindstrom and Bates (1990).

For the FO method, a first-order Taylor series expansion of (3) around zero, the expected value of the random parameters, is used to approximate (3). The random parameters are predicted by:

(6) $\hat{\mathbf{u}}_i = \hat{\mathbf{D}}\mathbf{Z}_i'(\mathbf{Z}_i\hat{\mathbf{D}}\mathbf{Z}_i' + \hat{\mathbf{R}}_i)^{-1}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \mathbf{0})]$

where $\hat{\mathbf{D}}$ and $\hat{\mathbf{R}}_i$ are estimates of $\mathbf{D}$ and $\mathbf{R}_i$, respectively, $\hat{\mathbf{b}}$ is an estimate of the fixed parameters $\mathbf{b}$, and $\mathbf{Z}_i$ is the derivative matrix with respect to the random parameters, defined by

(7) $$\mathbf{Z}_i = \left.\frac{\partial f(\mathbf{x}_i, \mathbf{b}, \mathbf{0})}{\partial \mathbf{u}_i'}\right|_{\hat{\mathbf{b}}, \mathbf{0}}$$

For the FOCE method, the first-order Taylor series expansion of model (3) is iterated around a current predictor of the random parameters (and a current $\mathbf{Z}_i$) until the convergence criterion is met. The final random parameters are obtained numerically through the following equation:

(8) $\hat{\mathbf{u}}_i = \hat{\mathbf{D}}\mathbf{Z}_i'(\mathbf{Z}_i\hat{\mathbf{D}}\mathbf{Z}_i' + \hat{\mathbf{R}}_i)^{-1}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \hat{\mathbf{u}}_i) + \mathbf{Z}_i\hat{\mathbf{u}}_i]$

where $\hat{\mathbf{D}}$, $\hat{\mathbf{R}}_i$ and $\hat{\mathbf{b}}$ are as defined above, but $\mathbf{Z}_i$ is given by

(9) $$\mathbf{Z}_i = \left.\frac{\partial f(\mathbf{x}_i, \mathbf{b}, \mathbf{u}_i)}{\partial \mathbf{u}_i'}\right|_{\hat{\mathbf{b}}, \hat{\mathbf{u}}_i}$$

Since $\hat{\mathbf{u}}_i$ appears on both sides of (8), for the FOCE method, $\hat{\mathbf{u}}_i$ must be solved iteratively. This is the fundamental concept innate to the FOCE method (Schabenberger 1994, Vonesh and Chinchilli 1997, Huang 2008, Temesgen et al. 2008, Meng et al. 2008, Meng and Huang 2009). Unfortunately, many forest modelers used (6) from the FO method to predict $\hat{\mathbf{u}}_i$ while implementing the FOCE method, without realizing that this could seriously bias the outcome. The methodological mixing should be avoided in future studies as it may lead to erroneous conclusions.

All model fittings were carried out using the SAS macro %NLINMIX, with expand = ZERO for the FO method and expand = EBLUP for the FOCE method (Littell et al. 2006). The default fittings assume iid error structure (i.e., $\mathbf{R}_i = \sigma^2\mathbf{I}_{n_i}$). For the spatial power error structure illustrated in (5), the fittings must be implemented using the keyword TYPE = SP(POW)(age) within the REPEATED statement.

4. Predictions

Once the $\hat{\mathbf{b}}$ and $\hat{\mathbf{u}}_i$ are estimated, subject-specific volume predictions for plot *i* can be made:

(10) $\hat{\mathbf{y}}_i = f(\mathbf{x}_i, \hat{\mathbf{b}}, \mathbf{0}) + \mathbf{Z}_i\hat{\mathbf{u}}_i$ (FO method)

(11) $\hat{\mathbf{y}}_i = f(\mathbf{x}_i, \hat{\mathbf{b}}, \hat{\mathbf{u}}_i)$ (FOCE method)

where $\mathbf{Z}_i$ for FO is defined in (7), and $\mathbf{x}_i$ is the $\mathbf{x}$ matrix for new or old (modeling) observations. Note the $\hat{\mathbf{y}}_i$s for FO and FOCE are different (Vonesh and Chinchilli 1997). Mixing them will likely result in biased predictions (Huang 2008, Meng and Huang 2009). It is possible that, following Judge et al. (1985)'s descriptions on prediction, we can also derive the following alternatives to obtain the adjusted predictions from the use of SP(POW) or any other generalized error covariance structure:

(10a) $\hat{\mathbf{y}}_{\mathbf{0}i} = f(\mathbf{x}_{\mathbf{0}i}, \hat{\mathbf{b}}, \mathbf{0}) + \mathbf{Z}_i\hat{\mathbf{u}}_i + \hat{V}'\hat{\mathbf{\Psi}}^{-1}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \mathbf{0}) - \mathbf{Z}_i\hat{\mathbf{u}}_i]$ (FO method)

(11a) $\hat{\mathbf{y}}_{\mathbf{0}i} = f(\mathbf{x}_{\mathbf{0}i}, \hat{\mathbf{b}}, \hat{\mathbf{u}}_i) + \hat{V}'\hat{\mathbf{\Psi}}^{-1}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \hat{\mathbf{u}}_i)]$ (FOCE method)

where $\mathbf{x}_{\mathbf{0}i}$ is a known $\mathbf{x}$ matrix of new observations for plot *i*, $\hat{\mathbf{y}}_{\mathbf{0}i}$ corresponds to $\mathbf{x}_{\mathbf{0}i}$, $\mathbf{x}_i$ is an $\mathbf{x}$ matrix from the old (existing) data, $\hat{V}$ is the estimated covariances between elements of old and new errors, and $\hat{\mathbf{\Psi}}$ is the estimated

generalized within-plot error covariance structure. For the SP(POW) structure, for instance, more explicit prediction equations can be derived following Judge et al. (1985) (detailed derivations are omitted here. Interested readers may contact the senior author to obtain a copy):

(10b) $\hat{\mathbf{y}}_{i,\kappa} = f(\mathbf{x}_{i,\kappa}, \hat{\mathbf{b}}, \mathbf{0}) + \mathbf{Z}_{i,\kappa}\hat{\mathbf{u}}_i + \rho^{d_{1(\kappa+1)}}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \mathbf{0}) - \mathbf{Z}_i\hat{\mathbf{u}}_i]$ (FO method)

(11b) $\hat{\mathbf{y}}_{i,\kappa} = f(\mathbf{x}_{i,\kappa}, \hat{\mathbf{b}}, \hat{\mathbf{u}}_i) + \rho^{d_{1(\kappa+1)}}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \hat{\mathbf{u}}_i)]$ (FOCE method)

where κ is the number of projection intervals (or periods), (κ+1) is the number of measurements, $\mathbf{x}_{i,\kappa}$ and $\mathbf{Z}_{i,\kappa}$ are the $\mathbf{x}_i$ and $\mathbf{Z}_i$ values κ intervals ahead, and $d_{1(\kappa+1)}$ is the distance (or time) between the first and (κ+1)th measurements (e.g., $d_{1(\kappa+1)} = d_{15}$ for the 4th projection period of the SP(POW) structure given in (5)). For a simpler structure such as AR(1), the above equations reduce to:

(10c) $\hat{\mathbf{y}}_{i,\kappa} = f(\mathbf{x}_{i,\kappa}, \hat{\mathbf{b}}, \mathbf{0}) + \mathbf{Z}_{i,\kappa}\hat{\mathbf{u}}_i + \rho^{\kappa}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \mathbf{0}) - \mathbf{Z}_i\hat{\mathbf{u}}_i]$ (FO method)

(11c) $\hat{\mathbf{y}}_{i,\kappa} = f(\mathbf{x}_{i,\kappa}, \hat{\mathbf{b}}, \hat{\mathbf{u}}_i) + \rho^{\kappa}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \hat{\mathbf{u}}_i)]$ (FOCE method)

For the reasons discussed at the end of Section 5, however, we opted not to implement the alternative equations in the present study before further assessment and simulations are conducted.

There are different ways to compare the differences between the observed and predicted volumes, and various statistics exist for reporting the goodness-of-fit of NLMMs (Vonesh and Chinchilli 1997). Here, we are mostly interested in the criteria to evaluate the prediction errors (residuals), their average and variations. Thus, the following statistics were computed:

(12) $\bar{\mathrm{e}} = \frac{1}{\mathrm{N}}\sum_{i=1}^{m}\sum_{j=1}^{n_i} e_{ij} = \frac{1}{\mathrm{N}}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})$

(13) $\mathrm{SD} = \sqrt{\frac{1}{\mathrm{N}}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(e_{ij} - \bar{\mathrm{e}})^2}$

(14) $\delta = \bar{\mathrm{e}}^2 + \mathrm{SD}^2$

(15) $\bar{\mathrm{e}}\% = \frac{\bar{\mathrm{e}}}{\bar{y}} \times 100\%$

where $\bar{\mathrm{e}}$ is the overall mean bias, y_{ij} and $\hat{y}_{ij}$ are the *j*th observed and predicted values for subject *i*, m is the number of subjects, N is the total number of observations from m subjects ($\mathrm{N} = \sum_{i=1}^{m}\mathrm{n}_i$), SD is standard deviation, δ is an overall accuracy measure combining the bias ($\bar{\mathrm{e}}$) and precision (SD) of the errors (Cochran 1977), $\bar{\mathrm{e}}\%$ is percent bias, and $\bar{y}$ is the arithmetic mean of the observed values.

Because the predominant interest of a NLMM is in subject-specific predictions, it is more important that we also compute the above statistics by subject:

(16) $\bar{\mathrm{e}}_i = \frac{1}{\mathrm{n}_i}\sum_{j=1}^{n_i} e_{ij} = \frac{1}{\mathrm{n}_i}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})$

(17) $\mathrm{SD}_i = \sqrt{\frac{1}{\mathrm{n}_i}\sum_{j=1}^{n_i}(e_{ij} - \bar{\mathrm{e}}_i)^2}$

(18) $\delta_i = \bar{\mathrm{e}}_i^2 + \mathrm{SD}_i^2$

(19) $\bar{\mathrm{e}}_i\% = \frac{\bar{\mathrm{e}}_i}{\bar{y}_i} \times 100\%$

where all variables are as defined before, except that they apply to subject *i*, not the entire population.

For a population of m subjects, the frequency distribution for any one of the SS measures expressed in (16)-(19) can be constructed, and the percentage of subjects that fall below or above certain thresholds such as the 5th or the 95th percentile, can be determined. As will be shown later, examination of such frequency distributions is exceedingly useful in helping modelers obtaining a realistic understanding of the goodness-of-fit a NLMM when it is used for subject-specific predictions. More importantly, it also helps revealing the differences between the models estimated with different error structures.

It is worthwhile to note here that many forest modelers have used the fixed parameters estimated as a part of a NLMM to predict the population-averaged (PA) responses, i.e., by setting $\hat{\mathbf{u}}_i = \mathbf{0}$ and using $\hat{\mathbf{y}}_i = f(\mathbf{x}_i, \hat{\mathbf{b}}, \mathbf{0})$. This is likely a result of some misunderstanding. The fixed parameters estimated as a part of a NLMM do not fully characterize the mean responses of the population, especially when the inter-subject variability is large (Davidian and Giltinan 2003, Fitzmaurice et al. 2004). Consequently, they generally provide a biased partial representation of the true population mean responses. It is very important to recognize that the predictions obtained from the fixed parameters estimated as a

part of a NLMM generally do not represent the "unbiased" PA responses (Huang 2008).

5. Results and discussion

Results of the estimated parameters for model (1) with iid and spatial power error structures are shown in Table 1. The Akaike's information criterion (AIC) is computed by (Littell et al. 2006):

(20) $\mathrm{AIC} = -2\mathrm{LL} + 2\mathrm{p}$

where -2LL is -2 times the log-likelihood function, and p is the total number of estimated parameters.

Studentized residual plots were obtained from the FO and FOCE methods using the iid and correlated (spatial power) error structures. Only those from the FOCE method are shown here (Figure 2), as those from the FO method display similar patterns. The differences in the studentized residual plots between the iid errors and the correlated errors are easy to detect. The studentized residuals from the iid errors are scattered evenly around the zero line, whereas the studentized residuals from the correlated errors are not. This suggests that judged by visual means, the iid error structure is more satisfactory.

To facilitate the discussion of the upcoming results, and to help readers grasp the essences of the issues, we first demonstrate the computations involved in predicting the random parameters and in deriving the predicted volumes for a "new" plot measured five times (j = 5) from the validation data (Table 2a), based on the FOCE method with the SP(POW) error structure. Actual computation algorithms for this demonstration are provided in Appendix 1, and were explained in detail in Huang (2008), Meng et al. (2008), and Meng and Huang (2009). The computations for the iid errors and for the FO method are relatively simpler and can be derived directly from the algorithms given in Appendix 1 (Huang 2008).

For the FOCE method, the derivatives of model (1) with respect to the two random parameters are:

(21) $der\_u_1 = \partial f(\mathbf{x}_i, \mathbf{b}, \mathbf{u}_i)/\partial u_1 = \mathbf{Age}_i^{(\hat{b}_2+\hat{u}_2)}(1-(\hat{b}_1+\hat{u}_1)\mathbf{Age}_i)\exp(-(\hat{b}_1+\hat{u}_1)\mathbf{Age}_i)$

(22) $der\_u_2 = \partial f(\mathbf{x}_i, \mathbf{b}, \mathbf{u}_i)/\partial u_2 = (\hat{b}_1+\hat{u}_1)\exp(-(\hat{b}_1+\hat{u}_1)\mathbf{Age}_i)\mathbf{Age}_i^{(\hat{b}_2+\hat{u}_2)}\ln(\mathbf{Age}_i)$

Therefore, for the ith subject with j measurements, the estimated $\mathbf{Z}_i$ matrix is constructed as follows:

(23) $$\mathbf{Z}_i = \begin{bmatrix} \dfrac{\partial f(x_{i1}, \mathbf{b}, \mathbf{u}_i)}{\partial u_1} & \dfrac{\partial f(x_{ij}, \mathbf{b}, \mathbf{u}_i)}{\partial u_2} \\ \vdots & \vdots \\ \dfrac{\partial f(x_{ij}, \mathbf{b}, \mathbf{u}_i)}{\partial u_1} & \dfrac{\partial f(x_{ij}, \mathbf{b}, \mathbf{u}_i)}{\partial u_2} \end{bmatrix} = \begin{bmatrix} der\_u_{11} & der\_u_{21} \\ \vdots & \vdots \\ der\_u_{1j} & der\_u_{2j} \end{bmatrix}$$

where der_u_{11}, …, der_u_{1j} are given by (21) and der_u_{21}, …, der_u_{2j} are given by (22), and j = 5. The estimated **D** matrix for any subject i is (from eq. (4) and Table 1):

(24) $$\hat{\mathbf{D}} = \begin{bmatrix} \sigma_{u_1}^2 & \sigma_{u_1u_2}^2 \\ \sigma_{u_1u_2}^2 & \sigma_{u_1}^2 \end{bmatrix} = \begin{bmatrix} 0.00030 & 0.000385 \\ 0.000385 & 0.007794 \end{bmatrix}$$

The estimated $\mathbf{R}_i$ matrix for the example subject is (from eq. (5) and Tables 1 and 2a):

(25) $$\hat{\mathbf{R}}_i = 6361.73\begin{bmatrix} 1 & 0.9977^7 & 0.9977^{21} & 0.9977^{31} & 0.9977^{41} \\ 0.9977^7 & 1 & 0.9977^{14} & 0.9977^{24} & 0.9977^{34} \\ 0.9977^{21} & 0.9977^{14} & 1 & 0.9977^{10} & 0.9977^{20} \\ 0.9977^{31} & 0.9977^{24} & 0.9977^{10} & 1 & 0.9977^{10} \\ 0.9977^{41} & 0.9977^{34} & 0.9977^{20} & 0.9977^{10} & 1 \end{bmatrix}$$

For the FOCE method, there is no simple algebraic solution for $\hat{\mathbf{u}}_i$ from (8). An iterative procedure involving the following computation steps is used to numerically solve for $\hat{\mathbf{u}}_i$:

Step 1: Obtain an initial estimate, termed $\hat{\mathbf{u}}_{i,0}$, of the random parameters. A reasonable first guess of this estimate is $\hat{\mathbf{u}}_{i,0}$ = E($\mathbf{u}_i$) = **0**, the expected value of the random parameters for the FO method. Thus, $\hat{\mathbf{u}}_{i,0}$ can be computed directly using eq. (6) for the FO method, as follows:

(26) $\hat{\mathbf{u}}_{i,0} = \hat{\mathbf{D}}\mathbf{Z}_{i,0}'(\mathbf{Z}_{i,0}\hat{\mathbf{D}}\mathbf{Z}_{i,0}' + \hat{\mathbf{R}}_i)^{-1}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\mathbf{b}}, \mathbf{0})]$

where $\mathbf{Z}_{i,0}$ is a first estimate of the $\mathbf{Z}_i$ matrix (equivalent to the $\mathbf{Z}_i$ in (7) for the FO method):

$$\mathbf{Z}_{i,0}=\begin{bmatrix}\dfrac{\partial f(x_{i1},\mathbf{b},\mathbf{0})}{\partial u_1} & \dfrac{\partial f(x_{ij},\mathbf{b},\mathbf{0})}{\partial u_2}\\ \vdots & \vdots\\ \dfrac{\partial f(x_{ij},\mathbf{b},\mathbf{0})}{\partial u_1} & \dfrac{\partial f(x_{ij},\mathbf{b},\mathbf{0})}{\partial u_2}\end{bmatrix}=\begin{bmatrix}Age_{i1}^{\hat{b}_2}(1-\hat{b}_1Age_{i1})\exp(-\hat{b}_1Age_{i1}) & \hat{b}_1\exp(-\hat{b}_1Age_{i1})Age_{i1}^{\hat{b}_2}\ln(Age_{i1})\\ \vdots & \vdots\\ Age_{ij}^{\hat{b}_2}(1-\hat{b}_1Age_{ij})\exp(-\hat{b}_1Age_{ij}) & \hat{b}_1\exp(-\hat{b}_1Age_{ij})Age_{ij}^{\hat{b}_2}\ln(Age_{ij})\end{bmatrix}$$

Step 2: Once the $\hat{\mathbf{u}}_{i,0}$ is calculated, the next estimation of $\hat{\mathbf{u}}_i$, termed $\hat{\mathbf{u}}_{i,1}$, is obtained based on (8), with $\hat{\mathbf{u}}_i$ on the right-hand side replaced by $\hat{\mathbf{u}}_{i,0}$ from Step 1:

(27) $\hat{\mathbf{u}}_{i,1}=\hat{\mathbf{D}}\mathbf{Z}_{i,1}^{'}(\mathbf{Z}_{i,1}\hat{\mathbf{D}}\mathbf{Z}_{i,1}^{'}+\hat{\mathbf{R}}_i)^{-1}[\mathbf{y}_i-f(\mathbf{x}_i,\hat{\mathbf{b}},\hat{\mathbf{u}}_{i,0})+\mathbf{Z}_{i,1}\hat{\mathbf{u}}_{i,0}]$

where $\mathbf{Z}_{i,1}$ is the $\mathbf{Z}_i$ expressed in (23) evaluated at $\hat{\mathbf{u}}_{i,0}$, with its elements given by:

(28) $der\_u_1=\partial f(\mathbf{x}_i,\mathbf{b},\hat{\mathbf{u}}_{i,0})/\partial u_1=\mathbf{Age}_i^{(\hat{b}_2+\hat{u}_{2,0})}(1-(\hat{b}_1+\hat{u}_{1,0})\mathbf{Age}_i)\exp(-(\hat{b}_1+\hat{u}_{1,0})\mathbf{Age}_i)$

(29) $der\_u_2=\partial f(\mathbf{x}_i,\mathbf{b},\hat{\mathbf{u}}_{i,0})/\partial u_2=(\hat{b}_1+\hat{u}_{1,0})\exp(-(\hat{b}_1+\hat{u}_{1,0})\mathbf{Age}_i)\mathbf{Age}_i^{(\hat{b}_2+\hat{u}_{2,0})}\ln(\mathbf{Age}_i)$

Step 3: Having the calculated $\hat{\mathbf{u}}_{i,1}$ from (27), the new estimation for $\hat{\mathbf{u}}_i$, termed $\hat{\mathbf{u}}_{i,2}$, is computed using (8) again, with the $\hat{\mathbf{u}}_i$ on the right-hand side replaced by the updated $\hat{\mathbf{u}}_{i,1}$ from Step 2:

(30) $\hat{\mathbf{u}}_{i,2}=\hat{\mathbf{D}}\mathbf{Z}_{i,2}^{'}(\mathbf{Z}_{i,2}\hat{\mathbf{D}}\mathbf{Z}_{i,2}^{'}+\hat{\mathbf{R}}_i)^{-1}[\mathbf{y}_i-f(\mathbf{x}_i,\hat{\mathbf{b}},\hat{\mathbf{u}}_{i,1})+\mathbf{Z}_{i,2}\hat{\mathbf{u}}_{i,1}]$

where $\mathbf{Z}_{i,2}$ is evaluated at $\hat{\mathbf{u}}_{i,1}$, with its elements defined in eqs. (28)-(29) updated using $\hat{\mathbf{u}}_{i,1}$.

This process is iterated *k* times until a desired precision, e.g., $\left|\hat{\mathbf{u}}_{i,k}-\hat{\mathbf{u}}_{i,(k-1)}\right|<0.0000001$, is achieved. The predicted final random parameters for subject *i* is: $\hat{\mathbf{u}}_i=\hat{\mathbf{u}}_{i,k}$.

For the example plot shown in Table 2a, the predicted final random parameters for the FOCE–pow are: u_1 = -0.00338906 and u_2 = -0.059555. The final elements (*der*_u_1 and *der*_u_2) of the $\mathbf{Z}_i$ matrix are listed in Table 2a, along with the $\mathbf{Z}_i$ matrices for the other error structure and the FO method.

Once the $\hat{\mathbf{u}}_i$ are known, the predicted volumes for the FOCE method are obtained directly using (11). Results are shown in Table 2a. For the FO method, the predicted volumes are obtained differently using (10). Results are also shown in Table 2a. Once the prediction errors and their SD are known, various intervals can be constructed (Hahn and Meeker 1991). Some common ones are:

(31) $\bar{\mathrm{e}}_i\pm t_{(1-\alpha/2),(n_i-1)}SD\sqrt{1+\dfrac{1}{n_i}}$ (prediction interval for a future observation)

(32) $\bar{\mathrm{e}}_i\pm z_{\frac{1+\gamma}{2}}\left(1+\dfrac{1}{2n_i}\right)\sqrt{\dfrac{n_i-1}{\chi^2_{\alpha,(n_i-1)}}}$ (tolerance interval to contain at least proportion γ of the subject)

(33) $\bar{\mathrm{e}}_i\pm t_{(1-\alpha/2),(n_i-1)}\dfrac{SD}{\sqrt{n_i}}$ (confidence interval for the subject mean)

(34) $SD\sqrt{\dfrac{n_i-1}{\chi^2_{(1-\alpha/2),(n_i-1)}}}\rightarrow SD\sqrt{\dfrac{n_i-1}{\chi^2_{(\alpha/2),(n_i-1)}}}$ (confidence interval for the subject standard deviation)

where z is the standard normal distribution (and α = 0.05 throughout this study).

Table 2b lists the normality test results and the calculated intervals for the example plot. All four tests conducted following Yang et al. (2004) showed that normality was met. However, it must be noted that since the sample size for this example plot is small, the interval calculations serve only as illustrations. The "preferred" calculations in fact require that the sample size should not be smaller than 8 (Hahn and Meeker 1991). In the cases where the normality assumption is not met, distribution-free intervals could be constructed (Hahn and Meeker 1991). One could also use a 10% trimmed mean as an estimate of $\bar{\mathrm{e}}_i$ and a jackknifed standard deviation expressed in (35) as an estimate of SD to construct approximated intervals (Efron and Tibshirani 1986, Hahn and Meeker 1991):

(35) $SD_{jack}=\sqrt{\dfrac{n_i-1}{n_i}\sum_{i=1}^{n_i}(\bar{e}_{i(j)}-\bar{e}_i)^2}$

where $\bar{e}_{i(j)}$ is the sample average of the data set without the jth point. It is straightforward to infer that (31)-(35) can also be applied to the entire population of N observations (N = $\sum_{i=1}^{m}\mathrm{n}_i$).

A fundamental difference between the FOCE and FO methods is that, the random parameters for the FOCE method must be solved numerically, whereas they can be obtained directly from (6) for the FO method. This fundamental difference appeared to have largely been neglected in previous applications of NLMMs in forestry. It is essential to recognize that when implementing the FOCE method, the $\hat{\mathbf{u}}_i$ calculated by (6) is just a first initial "guesstimate" (and sometimes an inappropriate guesstimate) in the search of a true empirical best linear unbiased predictor of $\mathbf{u}_i$.

Table 3 lists the overall prediction errors (e_{ij}) on both model fitting and validation data sets. For both FO and FOCE methods, the prediction errors (absolute values) and their variations from the correlated error structure are much larger than their counterparts from the iid error structure. For example, for the FO method on the model fitting data, the δ and $\bar{e}\%$ values for the iid error structure are 107.965 and 0.226%, respectively, whereas for the spatial power error structure, they are 4696.876 and 10.906%, respectively. The differences between the values from the iid and correlated error structures are striking. They are also consistent across the methods and data sets. All these suggest that the iid error structure produced much more accurate and precise predictions than the correlated error structure.

For the more important SS predictions, the statistics calculated according to (16)-(19) by different error structures and data sets are summarized in Table 4. The number of plots where the percent bias ($\bar{e}_i\%$) exceeded ±2.5% of the observed mean were identified and listed separately in Table 4. The frequency distributions of the mean biases, percent biases and standard deviations from all 288 plots of the model fitting data and 265 plots of the model validation data were obtained. For the sake of brevity, only the frequency distributions for $\bar{e}_i\%$ from different error structures and data sets are shown (Figure 3).

Judging from the results shown in Table 4, the contrasts between the iid and power error structures are remarkable. The power error structure produced much larger biases, SDs and δs. More specifically, for instance, for the FO method, among the 288 plots of the modeling data, only two plots produced biases exceeding ±2.5% of the observed means when the iid error structure was used. But when the power error structure was used, 272 plots produced biases exceeding ±2.5% of the observed means. Similar results are also apparent on the validation data, and for the FOCE method. The frequencies of poorer predictions from the power error structure are much greater than those from the iid error structure.

The poorer performance of the power error structure is more readily seen in Figure 3, where the percent biases from different plots are centered around zero but those from the power error structure are spread out much more than those from the iid error structure, indicating lower accuracies and larger variations from the power error structure. This is true on both model fitting and validation data sets, and for both FO and FOCE methods.

Figure 4 shows the observed volumes against the volumes predicted from different methods with iid and power error structures. We only show the plots from the validation data because the plots from the modeling data are similar. More importantly, it is more telling on validation data than on modeling data because the validation data are "new" independent data that a fitted model is most likely to be applied to make predictions in real-world applications. The poorer performance of the power error structure is rather apparent in Figure 4. It is consistent for both FO and FOCE methods.

An examination of the "spaghetti" or "chow-mein" plots (Figure 5) of the prediction trajectories of the validation data indicates that the prediction trajectories from the iid error structure closely mimic the observed trajectories shown in Figure 1(b), whereas the prediction trajectories from the power error structure "compress" the observed trajectories. This is again consistent for both FO and FOCE, even though the degree of the compression varies slightly.

Since much of the data collected for forest modeling is unequally spaced and unbalanced longitudinal data, where repeated measurements are taken on the same experimental units, the lack of independence of such data is inborn. The correlated nature of the data could violate an important assumption required by certain statistical techniques, mostly notably the least squares principle, for "optimal" performance and inference. Forest modelers have studied for many years, and have used the old and new techniques brought to light by statisticians, to address the dependence problem, believing in many cases, that the successful removal of the correlation would result in a "better" model.

However, contrary to the common belief, the results obtained in this study from the NLMM methods showed that although accounting for the serial correlation appeared to have produced better fits when judged by the AIC values in Table 1 (a smaller AIC value is said to mean a better fit), it produced much worse predictions on both the model fitting and validation data sets. The predictions obtained with the iid error structure had consistently much smaller and better distributed biases than those obtained with the correlated (spatial power) error structure. This created a dilemma, as the successful accounting of the correlated errors produced much poorer predictions, while ignoring the correlated errors produced much better predictions but could potentially invalid hypothesis testing and interval estimation. It also asked for the question of whether the common practice of using AIC (and its modified forms) to select a preferred model and/or a preferred error covariance structure is correct.

While there is no "best" answer to the apparent contradiction, and opinions may understandably vary depending on the focus of a study or a particular researcher, we believe in general that, because the vast majority of forestry models are developed to be used as predictive tools once the parameters have been estimated, and because much of the practical, real-world emphasis of forest modeling has been on the predictive capabilities and biases from the applications of the models, the prediction results, preferably obtained on independent validation data, should be given the predominant consideration. They should be used as the deciding factor in determining the appropriateness of a fitted model, and in

comparing alternative models and/or error structures.

We also believe that, there may be some debatable issues and misconceptions about hypothesis testing and interval estimation in forest modeling. We understand that, in the presence of correlated (and/or heteroscedastic) errors, the ordinary estimators of coefficients ignoring the correlation are still unbiased. However, the standard errors of the coefficients are biased and inconsistent, and the estimators are not efficient (Judge et al. 1985). When a correct error covariance structure has been identified and a consistent estimator of this error covariance structure is obtained, it can be used to obtain a consistent estimator of the standard errors of the estimated coefficients that are also efficient. The question though is that, for a forest model developed as predictive tools, model users are (arguably) not really concerned much about the hypothesis testing and interval estimation on the coefficients that have been estimated in an unbiased manner. Instead, they are generally more concerned about the hypothesis testing and interval estimation on the predictions illustrated in eqs. (31)-(35) and Table 2b, which have little to do with the standard errors of the coefficients obtained on the modeling data. To ward off some potential statistical traps, forest modelers should be aware of and be mindful of some common miscues when conducting hypothesis testing and interval estimation under normal or non-normal error assumption, on different data sets (i.e., modeling vs validation/application), at different levels (i.e., population average vs SS), and for different purposes (e.g., model coefficients vs model predictions).

For the vast majority of forest models built on repeated measurement data and to be used as predictive tools on new data or data ranges beyond those used in modeling, one needs to understand the pertinent prediction scenario first before making any prediction and judging the goodness of the prediction. As illustrated in Figure 6, various prediction scenarios exist, and different evaluation measures can be used to judge the goodness of predictions. We examined the scenarios and conducted simulations based on the fitted volume-age model, and found that accurate local predictions could be achieved from 2, 3 or more prior observations. We also found that, under the NLMM framework, with or without adjusting the predictions from the use of SP(POW) produced varied outcome, dependent on the 'best' covariance structure chosen, the direction of the predictions, the length and number of the prediction intervals, the level on which the predictions are made, the estimation technique (FO vs FOCE) used, the evaluation measures selected, and more importantly, the number of prior observations available for predicting the SS random parameters (we used all available prior observations in our analyses). In general, however, when SS random parameters are obtained from 2 or more prior observations, adjusting the predictions degenerated the predictions for our data. While the results from our study are obvious, we recognize that this is just one such study to demonstrate that the iid error structure is a sound choice for dealing with correlated data under the NLMM framework, if prediction is the primary focus of the study. We realize the potential options and variations in making predictions (as illustrated in Figure 6), and plan to conduct additional studies to further test this phenomenon on larger data sets, and for other tree species and other types of models.

6. Conclusions

Based on the FO and FOCE methods, we evaluated the volume-age model for lodgepole pine estimated with and without taking into account the serially-correlated errors. We found that the model estimated with the iid error structure outperformed the model estimated with the correlated (spatial power) error structure by a large margin. This observation was consistent for both FO and FOCE methods, and on both model fitting and validation data sets. It means that a NLMM estimated with the iid error structure is better in predictions than the model estimated with the correlated error structure. The result of this study can have some important practical implications, as a better model can be estimated using a simpler analysis without the need to account for the correlated error structure. In fact, accounting for the correlated error structure within the NLMM context not only did not improve model predictions, it actually degenerated the predictions substantially. Unless the main objective of modeling is not to develop models for prediction purposes, we believe that there is little or no concrete benefit to consider more elaborate error structures that account for the correlated errors. The iid error structure is simply a better choice for the lodgepole pine model, and for a similar model for black spruce (*Picea mariana*) (Huang et al. 2008b). Future studies should be conducted to see if the conclusions reached in this study still hold for other types of models and species, particularly on independent validation data sets.

Acknowledgements

This research was supported by the Alberta Government and Forest Resource Improvement Association of Alberta (the FRIAA-GYPSY project). Special thanks to the 10 project partners led by West Fraser Mills, Weyerhaeuser Canada and Canadian Forest Products, and to Mr. R. Briand, Dr. W.R. Dempster and Mr. D.J. Morgan for their help and contributions to this project.

References

Alberta Forest Service. (2000). Permanent sample plots field procedures manual. Forest Management Branch, Alberta Sustainable Resource Development, Edmonton, Alberta, Pub. FMOPC 83-03, 106 p.

Beal, S. L., & Sheiner, L.B. (1982). Estimating population kinetics. *Critical Reviews in Biomedical Engineering* 8: 195-222.

Borders, B.E., Bailey, R.L., & Clutter, M.L. (1988). Forest growth models: parameter estimation using real growth series. USDA For. Serv. Gen. Tech. Rep. NC-120. pp. 660-667.

Cochran, W.G. (1977). *Sampling techniques* (3rd ed.). Wiley, New York, 428 p.

Davidian, M., & Giltinan, D.M. (1995). *Nonlinear models for repeated measurement data*. Chapman & Hall, New York, 360 p.

Davidian, M., & Giltinan, D.M. (2003). Nonlinear models for repeated measurement data: an overview and update. *J. Agricultural, Biological, and Environmental Statistics* 8: 387–419.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat. Sci.* 1: 54-77.

Fang, Z., & Bailey, R.L. (2001). Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *For. Sci.* 47: 287–300.

Fitzmaurice, G.M., Laird, N.M., & Ware, J.H. (2004). *Applied longitudinal analysis*. Wiley, New York, 506 p.

Gregoire, T.G. (1987). Generalized error structure for forestry yield models. *For. Sci.* 33: 423-444.

Gregoire, T.G., Schabenberger, O., & Barrett, J. (1995). Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. *Can. J. For. Res.* 25: 137-156.

Hahn, G.J., & Meeker, W.Q. (1991), Statistical intervals: a guide for practitioners, John Wiley & Sons, New York. 392 p.

Huang, S. (1994). Individual tree volume estimation procedures for Alberta: methods of formulation and statistical foundations. Alberta Environmental Protection Tech. Rep. Pub. No. T/288. 80 p.

Huang, S. (1997). Development of compatible height and site index models for young and mature stands within an ecosystem-based management framework. *In* Empirical and process-based models for forest tree and stand growth simulation. Lisbon, Portugal, pp. 61-98.

Huang, S. (2008). A generalized procedure for bias-free predictions at population and local levels. A discussion paper, Pub-Feb-22-2008, Forest Management Branch, ASRD, Edmonton, Alberta, 61 p.

Huang, S., Meng, S.X., & Yang, Y. (2008a). Selected GYPSY components released for evaluation. Pub-Oct-23-2008, Forest Management Branch, ASRD, Edmonton, Alberta, 43 p.

Huang, S., Meng, S.X., & Yang, Y. (2008b). Selected GYPSY components released for evaluation: total volume models. Pub-Dec-19-2008, Forest Management Branch, ASRD, Edmonton, Alberta, 18 p.

Judge, G.G., Griffiths, W.E., Hill, R.C., Lütkepohl, H., & Lee, T.C. (1985). *The theory and practice of econometrics*. 2nd ed., John Wiley & Sons, New York. 1056 p.

Lappi, J. (1986). Mixed linear models for analyzing and predicting stem form variation of Scots pine. Communicationes Inst. Forestalis Fennicae 134. 69 p.

Lappi, J., & Bailey, R.L. (1988). A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. *For. Sci*. 34: 907-927.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., & Schabenberger, O. (2006). *SAS system for mixed models* (2nd ed.), SAS Institute Inc., Cary, NC, 814 p.

Lindstrom, M.J., & Bates, D.M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics* 46: 673–687.

Meng, X.S., Huang, S. Lieffers, V.J., Nunifu, T., & Yang, Y. (2008). Wind speed and crown class influence the height–diameter relationship of lodgepole pine: nonlinear mixed effects modeling. *For. Ecol. Manage.* 256: 570-577.

Meng, X.S., & Huang, S. (2009). Improved calibration of nonlinear mixed-effects models demonstrated on a height growth function. *For. Sci.* (in revision).

Monserud, R.A. (1984). Height growth and site index curves for inland Douglas-fir based on stem analysis and forest habitat type. *For. Sci*. 30: 943-965.

Ojansuu, R. (1987). Mixed linear models for stem size and form development. USDA For. Serv. Gen. Tech. Rep. NC-120. pp. 724-730.

SAS Institute Inc. (2004). *SAS/STAT 9.1 User's Guide*. SAS Institute Inc., Cary, NC. 5136 p.

Schabenberger, O. (1994). Nonlinear mixed effects growth models for repeated measures in ecology. In *Proc. of ASA, Section on Stats. & Environment*, Toronto, Canada, pp.156-161.

Schabenberger, O., & Pierce, F.J. (2002). *Contemporary statistical models for the plant and soil sciences*. CRC Press, Boca Raton, FL, 760 p.

Temesgen, H., Monleon, V.J., & Hann, D.W. (2008). Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Can. J. For. Res*. 38: 553–565.

Vonesh, E.F., & Chinchilli, V.M. (1997). *Linear and nonlinear models for the analysis of repeated measurements*. Marcel Dekker, New York, 560 p.

Wang, Y., Lemay, V.M, & Baker, T.G. (2007). Modeling and prediction of dominant height and site index of eucalyptus *globulus* plantations using a nonlinear mixed-effects models approach. *Can. J. For. Res.* 37: 1390-1403.

West, P.W. (1995). Application of regression analysis to inventory data with measurements on successive occasions. *For. Ecol. & Manage*. 71 (1995): 227-234.

Yang, Y., & Huang, S. (2008). Modeling percent stocking changes for lodgepole pine stands in Alberta. *Can. J. For. Res. 38*: 1042-1052.

Yang, Y., Monserud, R.A., & Huang, S. (2004). An evaluation of diagnostic tests and their roles in validating forest biometric models. *Can. J. For. Res.* 34: 619-629.

Appendix 1. A generalized SAS program for predicting random parameters based on the FOCE method with the serially-correlated (spatial power) error structure

-This program applies to unequally spaced and unbalanced longitudinal data for any number of plots.

-The spatial power error structure is illustrated in (5) for a plot (subject) measured five times.

```
*Input the data, where variable 'measure' refers to the measurement number;
data tt1;
input plotid measure age vol;
cards;
38001   1   53.234   155.057
38001   2   60.234   163.343
38001   3   74.234   226.229
38001   4   84.234   268.278
38001   5   94.234   290.151
more datalines…;
run;
proc sort data=tt1; by plotid;
run;

data wed3;
set tt1;
by plotid;
j+1;
if first.plotid then do; i+1; j=1; end;
run;
proc means data=wed3; var i j;
run;

*Input the estimated parameters (from Table 1), and execute the proc iml;
proc iml;
```

```
use wed3;
read all var {j} into tobs; read all var {i age} into age; read all var {vol} into vol;
covar={0.000030, 0.000385, 0.007794, 0.9977, 6361.73};
fixp={0.02396 2.5972};
d=j(2,2,.);
d[1,1]=covar[1]; d[1,2]=covar[2]; d[2,1]=covar[2]; d[2,2]=covar[3];
s=covar[5]; cor=covar[4]; bb=fixp[1,]; b=bb`;
tn=max(age[,1]);/*number of plots*/
q=2; /*number of parameters*/
bx={1 1}; /*number of random effect parameters*/
nn=nrow(vol); u=j(tn,q,0); mc=max(tobs);
start sm1 (tn,age,s,nn,cor,cov,mc);
cov=j(nn,mc,0); start=j(nn,1,.);
do i=1 to tn;
c=0; start[,1]=.;
do j=1 to nn; if age[j,1]=i then;
do;
start[j]=j; c=c+1;
end; end;
st=min(start);
do h=1 to c;
do k=1 to c;
a=st+h-1; b=st+k-1;
cv=s*cor**(abs(age[a,2]-age[b,2]));
cov[a,k]=cv;
end; end;
end;
finish sm1;
run sm1 (tn,age,s,nn,cor,cov,mc);
start sm (tn,bx,q,u,z,b,s,d,age,vol,nn,res,uv,cov);
z=j(nn,q,0); res=j(nn,1,0); uv=j(tn,q,0); xx=1;
do k=1 to tn;
z[1:nn,]=.; res[1:nn,]=.;
do j=1 to nn;   if age[j,1]=k then;
 do;
u1=u[k,1];   u2=u[k,2];
agem=age[j,2]; volm=vol[j,1];
 zb1 = agem**(b[2]+u2)*(1-(b[1]+u1)*agem)*exp(-(b[1]+u1)*agem);
 z[j,1]=zb1;
 zb2 = (b[1]+u1)*exp(-(b[1]+u1)*agem)*log(agem)*agem**(b[2]+u2);
 z[j,2]=zb2;
 re=volm-(b[1]+u1)*agem**(b[2]+u2)*exp(-(b[1]+u1)*agem)+zb1*u1+zb2*u2;
 res[j]=re;
end; end;
```

```
r1=z;
r2=r1[loc(r1[,1]^=.),]#bx;
w1=res; w2=w1[loc(w1[,1]^=.),];
mm=nrow(w2); yy=xx+mm-1;
rr=cov[xx:yy,1:mm]; xx=mm+xx;
uu=d*r2`*INV(r2*d*r2`+rr)*w2; uk=uu`;
uv[k,]=uk;
end;
finish sm;

diff=j(tn,q,0); diff1=1; diff2=1;
eps=1E-7;
do iter=1 to 200 until (diff1<eps & diff2<eps);
run sm (tn,bx,q,u,z,b,s,d,age,vol,nn,res,uv,cov);
diff=abs(uv-u);
diff1=max(diff[,1]);
diff2=max(diff[,2]);
u=uv;
end;

*u is the predicted random parameters;
ubu=u;
bf=j(tn,q,0);
do i=1 to tn;
bf[i,1]=b[1];
bf[i,2]=b[2];
end;
ub=ubu||bf; cnm={ub1i,ub2i,b1,b2};
create rpm from ub[colname=cnm];
append from ub;
quit;

*predicting the volumes for all plots;
data rpm1;
set rpm;
i=_n_;
run;
proc sort data=wed3;by i;run;
proc sort data=rpm1;by i;run;

*y_pred and y_res are predicted volume and residual from the FOCE method with spatial power errors;
data allx ;
 merge wed3 rpm1;
 by i ;
```

```
y_pred = (b1+ub1i)*age**(b2+ub2i)*exp(-(b1+ub1i)*age);
y_res= vol - y_pred;
run;
```

Table 1. Parameter estimates and fit statistics for model (1) based on the first-order (FO) and first-order conditional expectation (FOCE) methods assuming iid or correlated (spatial power) error structures.

Method	b_1	b_2	$\sigma^2_{u_1}$	$\sigma_{u_1u_2}$	$\sigma^2_{u_2}$	ρ	σ^2	AIC
FO-iid	0.02758	2.6258	0.000149	0.001766	0.02693		188.62	12448.0
FO-pow	0.02329	2.5799	0.000028	0.000364	0.008311	0.9980	6952.33	11890.9
FOCE-iid	0.02691	2.6261	0.000098	0.001176	0.01947		171.97	12367.2
FOCE-pow	0.02396	2.5972	0.000030	0.000385	0.007794	0.9977	6361.73	11913.6

Note: iid refers to independent and identically distributed error structure, power refers to spatial power error structure, b_1 and b_2 are fixed parameters, $\sigma^2_{u_1}$ and $\sigma^2_{u_2}$ are variances for random parameters u_1 and u_2, respectively, and $\sigma_{u_1u_2}$ is the covariance between u_1 and u_2, ρ is correlation parameter, σ^2 is residual variance, and AIC is Akaike's information criterion (defined in eq. (20)).

Table 2a. Example computations of prediction errors for a plot (subject) measured five times.

Method	Age	Vol_o	der_u_1	der_u_2	Vol_p	Res	$\bar{e}_i$	SD
FO-iid	53.234	155.057	-3676.41	860.79	144.299	10.759	0.021	8.434
	60.234	163.343	-5921.11	1012.11	174.680	-11.337		
	74.234	226.229	-11034.88	1251.58	230.671	-4.442		
	84.234	268.278	-14744.17	1362.55	263.790	4.488		
	94.234	290.151	-18154.98	1423.50	289.511	0.640		
FO-pow	53.234	155.057	-1973.51	761.18	155.821	-0.764	-7.351	9.475
	60.234	163.343	-3872.25	917.03	185.356	-22.013		
	74.234	226.229	-8668.21	1192.64	238.085	-11.856		
	84.234	268.278	-12553.09	1347.42	268.531	-0.253		
	94.234	290.151	-16498.20	1461.78	292.020	-1.869		
FOCE-iid	53.234	155.057	302.23	579.59	145.821	9.236	-0.144	8.116
	60.234	163.343	-862.75	717.57	175.093	-11.750		
	74.234	226.229	-4343.18	988.36	229.466	-3.237		
	84.234	268.278	-7603.33	1165.34	262.842	5.436		
	94.234	290.151	-11310.55	1320.80	290.555	-0.404		
FOCE-pow	53.234	155.057	-763.71	656.79	165.244	-10.187	-18.020	9.564
	60.234	163.343	-2275.16	802.30	195.767	-32.424		
	74.234	226.229	-6391.37	1074.44	249.451	-23.222		
	84.234	268.278	-9968.73	1240.74	279.850	-11.572		
	94.234	290.151	-13816.33	1376.67	302.845	-12.694		

Note: the methods are defined in Table 1, Vol_o and Vol_p are observed and predicted volumes (m^3/ha), der_u_1 and der_u_2 are derivatives with respect to random parameters u_1 and u_2, respectively, Res is residual, $\bar{e}_i$ is the arithmetic mean of the residuals, and SD is the standard deviation of the residuals. The algorithms given in Appendix 1 show the computations for the FOCE-pow method.

Table 2b. Lower and upper limits of confidence intervals for the example plot shown in Table 2a.

Md.	Normality test				Prediction		Tolerance		CL for $\bar{e}_i$		CL for SD	
	W	KS	W^2	A^2	lower	upper	lower	upper	lower	upper	lower	upper
1	0.997	0.129	0.018	0.135	-25.63	25.67	-43.12	43.16	-10.45	10.49	5.05	24.24
2	0.818	0.319	0.086	0.492	-36.17	21.47	-55.81	41.11	-19.12	4.41	5.68	27.23
3	0.977	0.154	0.024	0.171	-24.83	24.54	-41.65	41.37	-10.22	9.93	4.86	23.32
4	0.846	0.311	0.076	0.437	-47.11	11.07	-66.94	30.90	-29.90	-6.15	5.73	27.48

Note: Md. (method) is defined in Table 1, W, KS, W^2 and A^2 are Shapiro-Wilk, Kolmogorov-Smirnov, Cramér-von Mises and Anderson-Darling statistics, respectively, for testing normality. Prediction intervals, tolerance intervals, and confidence limits (CL) containing the mean ($\bar{e}_i$) and the standard deviation (SD) are computed according to eqs. (31)-(34), respectively.

Table 3. A summary of overall prediction errors from different methods and error structures.

Data	Method	N	Mean	Min.	Max.	SD	δ	$\bar{e}$ %
Modelling	FO-iid	1249	0.685	-45.859	50.767	10.368	107.965	0.226
	FO-pow	1249	33.012	-119.206	195.721	60.059	4696.876	10.906
	FOCE-iid	1249	0.288	-50.877	42.649	9.922	98.529	0.095
	FOCE-pow	1249	16.243	-168.589	166.271	61.505	4046.700	5.366
Validation	FO-iid	1011	0.333	-38.308	38.004	8.157	66.648	0.116
	FO-pow	1011	12.438	-90.274	134.873	42.618	1970.998	4.325
	FOCE-iid	1011	-0.003	-38.336	35.250	7.769	60.357	-0.001
	FOCE-pow	1011	-2.211	-124.540	120.759	43.656	1910.735	-0.769

Note: the methods are defined in Table 1, N is the total number of observations, min., max. and SD are minimum, maximum and standard deviation, δ is an overall accuracy measure calculated by eq. (14), and $\bar{e}$ % is the percent bias calculated by eq. (15).

Table 4. A summary of subject-specific prediction errors from different methods and error structures.

Data	Method	m	Mean	Min.	Max.	SD	δ	Freq. of $\lvert\bar{e}_i\%\rvert$>2.5
Modelling	FO-iid	288	0.250	-1.410	25.603	1.574	2.54	2
	FO-pow	288	6.181	-106.839	53.433	23.357	583.754	272
	FOCE-iid	288	0.085	-2.659	25.864	1.615	2.615	3
	FOCE-pow	288	-0.540	-145.735	50.470	28.404	807.079	260
Validation	FO-iid	265	-0.090	-16.637	5.187	1.549	2.408	12
	FO-pow	265	-0.364	-86.454	66.863	20.288	411.735	228
	FOCE-iid	265	-0.289	-20.144	3.256	1.820	3.396	13
	FOCE-pow	265	-6.627	-119.641	60.429	24.014	620.589	226

Note: the methods are defined in Table 1, m is the number of subjects (plots), min., max. and SD are minimum, maximum and standard deviation, δ_i is the subject-specific accuracy measure calculated by eq. (18), and $\bar{e}_i$ % is the subject-specific percent bias calculated by eq. (19). The frequency (freq.) of $\lvert\bar{e}_i\%\rvert$>2.5 refers to the number of plots whose $\bar{e}_i$% exceeded ±2.5 (the 2.5 threshold was chosen by dividing the one-sided 5% significance level commonly used in statistical inference into two-sides).

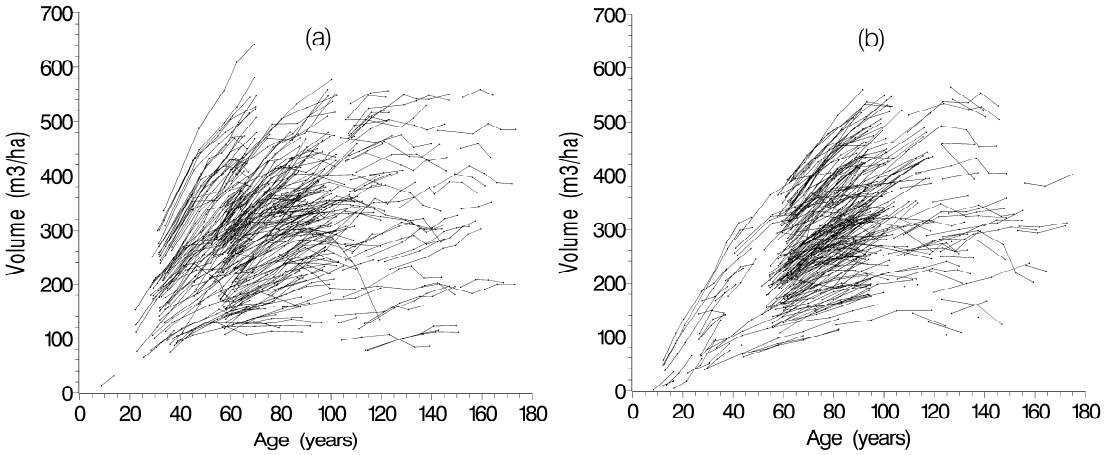


Figure 1. Model fitting (a) and validation (b) data sets. The model fitting data are from the upper foothills ecoregion. The model validation data are from the lower foothills ecoregion.

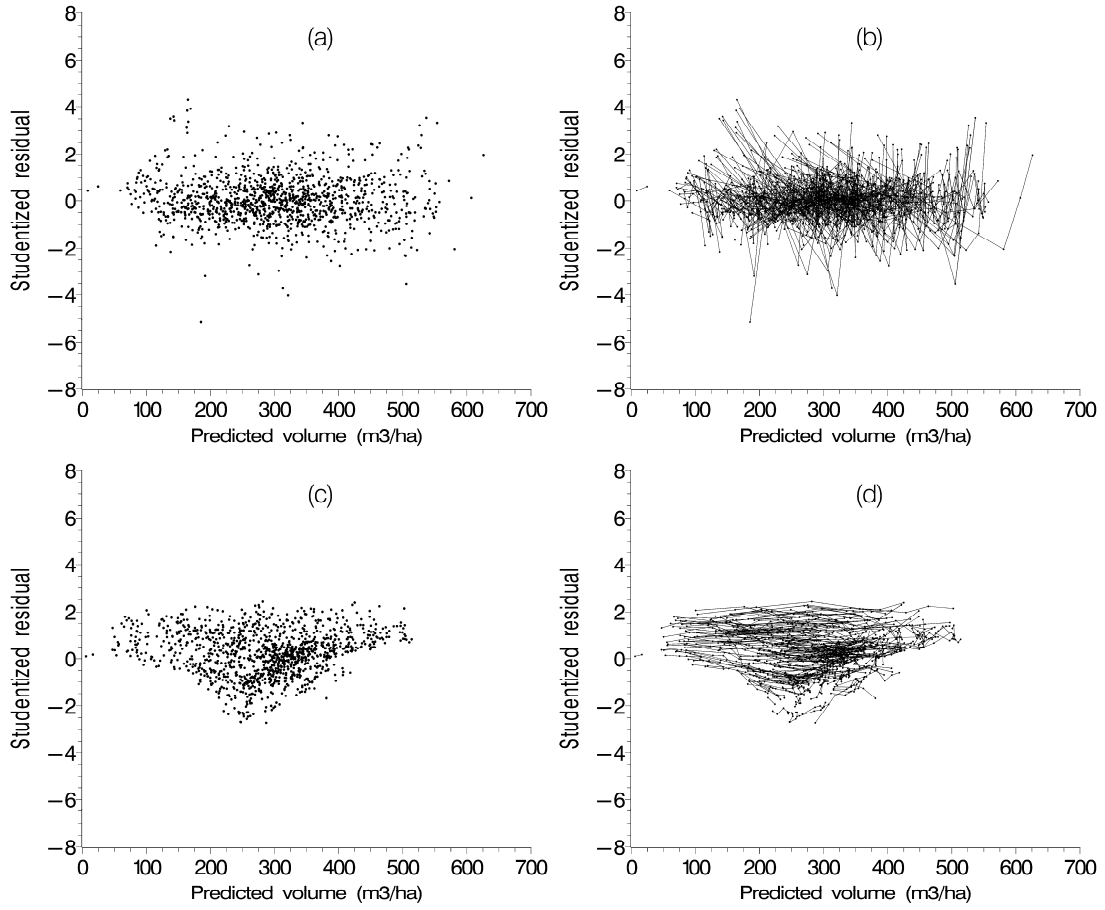


Figure 2. Studentized residual plots from the first-order conditional expectation (FOCE) method with the iid error structure (a, b), or correlated (spatial power) error structure (c, d). The right-hand side graphs connect the scatter points of the left-hand side graphs by subject.

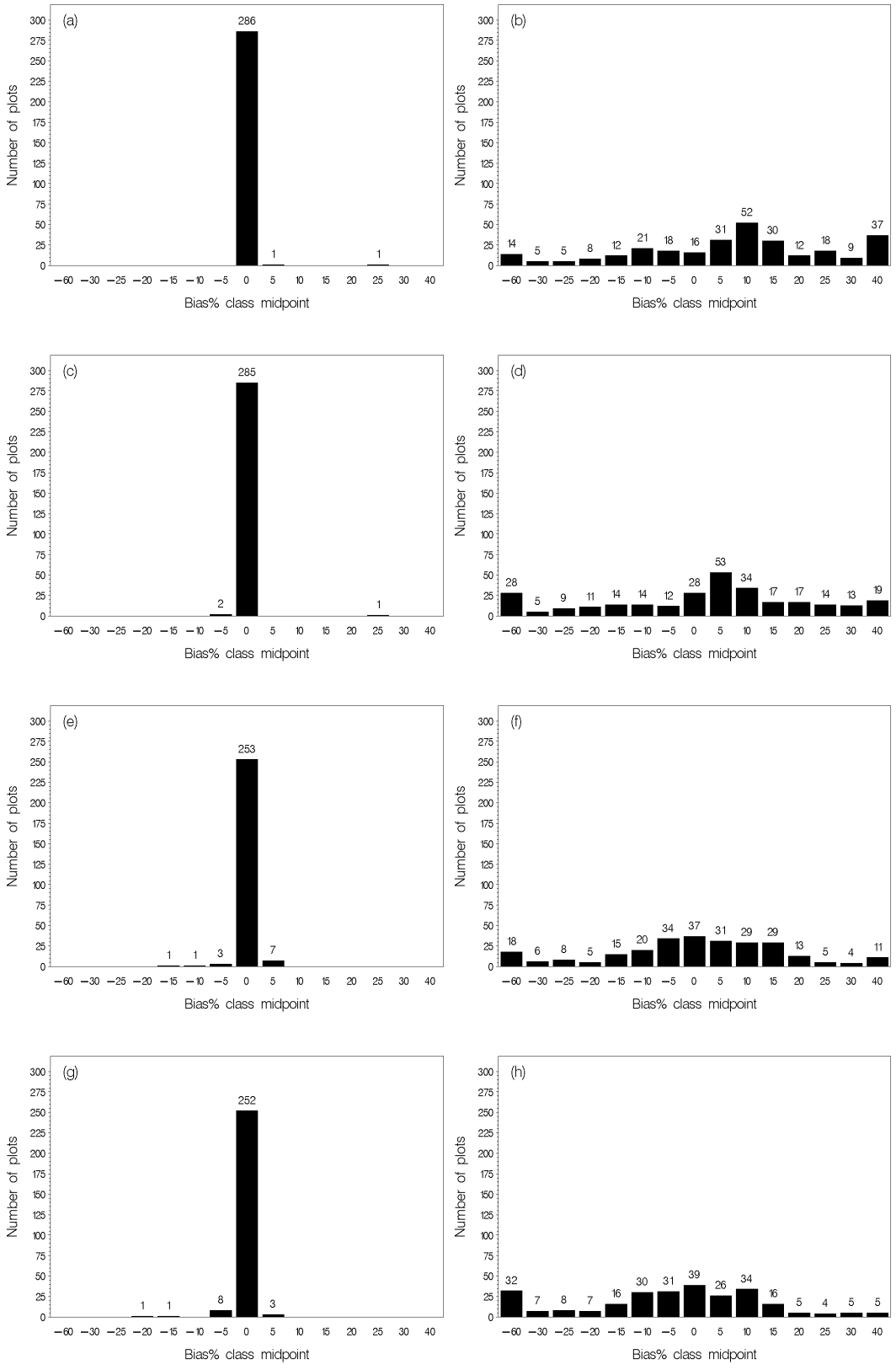


Figure 3. Frequency distributions of subject-specific percent biases ($\bar{e}_i$%s) for the model fitting data (a, b, c, d) and model validation data (e, f, g, h), based on the first-order (FO) and first-order conditional expectation (FOCE) methods: (a) and (e) – FO with iid errors; (b) and (f) - FO with correlated (spatial power) errors; (c) and (g) – FOCE with iid errors; and (d) and (h) - FOCE with spatial power errors.

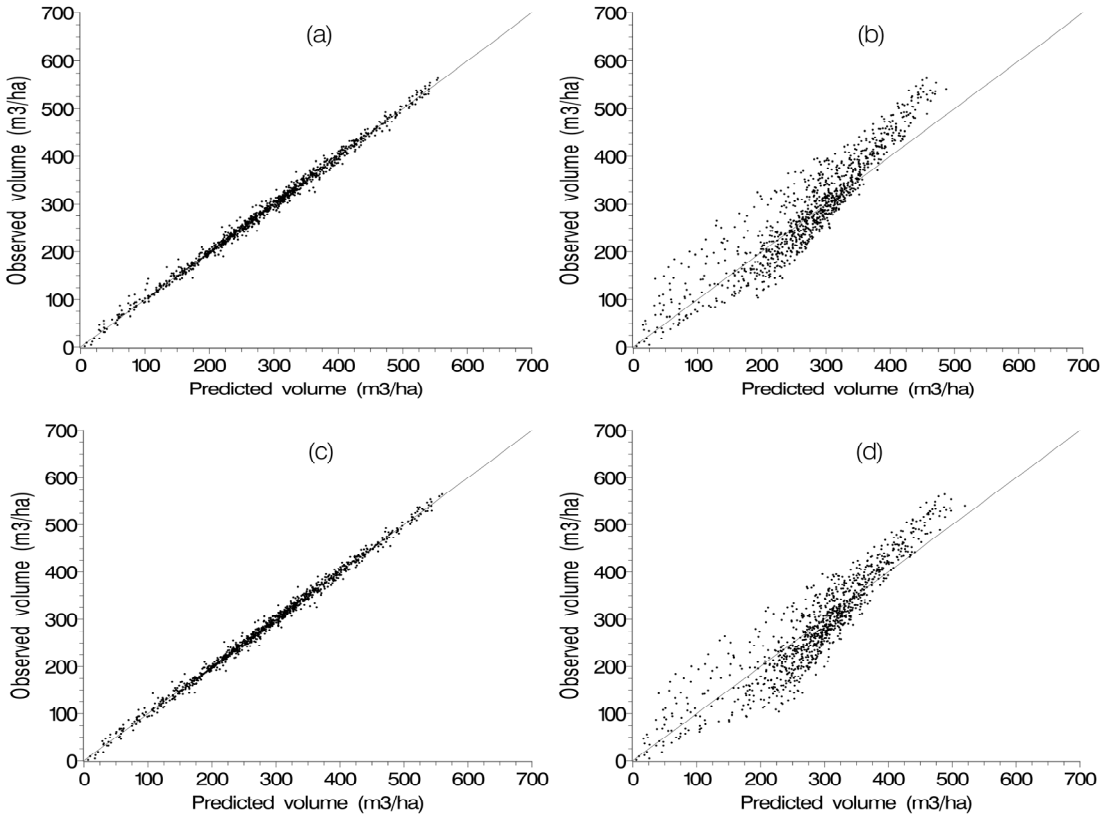


Figure 4. Observed versus predicted volumes for the validation data: (a) – first-order (FO) method with iid errors; (b) - FO method with spatial power errors; (c) – first-order conditional expectation (FOCE) method with iid errors; and (d) - FOCE method with spatial power errors. The line represents equality.

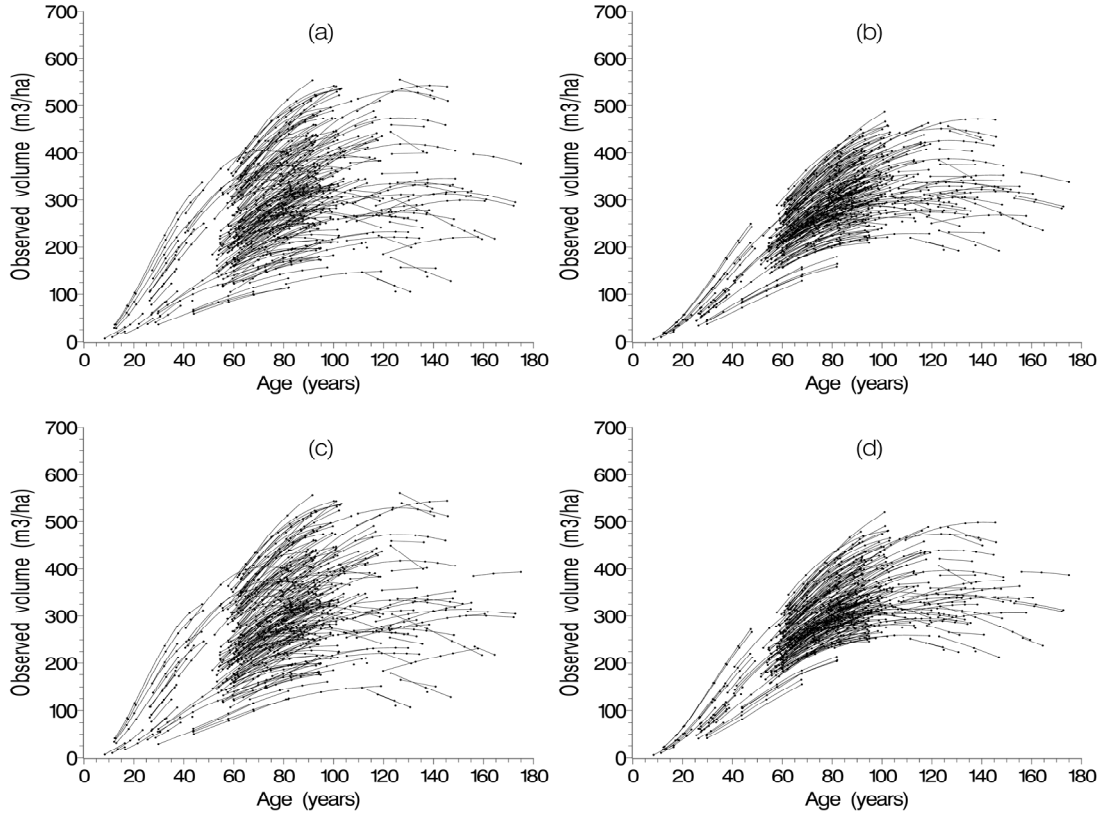


Figure 5. Predicted volume-age trajectories for the validation data: (a) – FO with iid errors; (b) – FO with spatial power errors; (c) – FOCE with iid errors; and (d) - FOCE with spatial power errors.

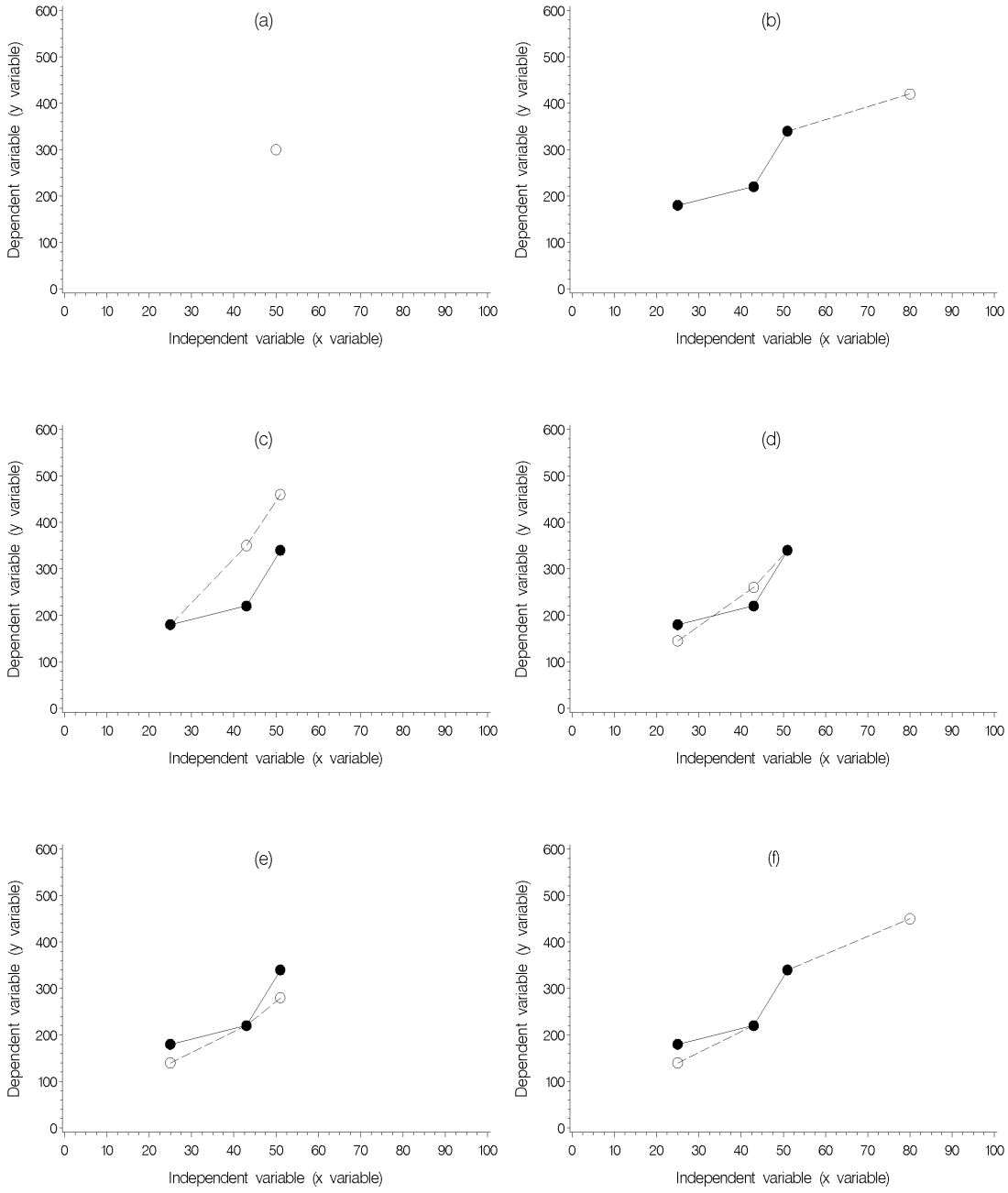


Figure 6. An illustration of selected prediction scenarios, where '●' represents an observed value and '○' represents a predicted value: (a) – predicting a *y* based on an observed *x* (equivalent to an overall population mean prediction); (b) – predicting a new (future) *y* based on the last *y*-*x* pair of an observed trajectory; (c) – predicting *y*s based on the first observed *y*-*x* pair; (d) – predicting previous *y*s based on the last observed *y*-*x* pair; (e) – predicting *y*s based on an observed *y*-*x* pair in the middle of a trajectory; and (f) – predicting a *y* or *y*s based on two observed *y*-*x* pairs. Note that (b), (c), (d) and (e) correspond to SS predictions based on one prior observation, and (f) corresponds to the predictions based on two prior observations. Depending on the number of observations available from an observed trajectory, forward and backward predictions can also be made using 3, 4,…, *k* observations randomly selected from the trajectory. In our analyses, using (c) to illustrate, we used all 3 observations to get the random parameters. We then made predictions using (10) or (11) (we also examined the adjusted predictions, starting with the 1st observation and taking into account the spatial power structure when predicting the 2nd and 3rd observations through (10b) or (11b)).

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Investigation of Radial Swirler Effect on Flow Pattern inside a Gas Turbine Combustor

Yehia A. Eldrainy

Department of Aeronautical Engineering, Faculty of Mechanical Engineering

Universiti Teknologi Malaysia, 81310 Skudai, Johor Malaysia

Tel: 6-013-751-3221 E-mail: yeldrainy@yahoo.com

Mohd Fairuz bin Ahmad @ Ibrahim

Department of Aeronautical Engineering, Faculty of Mechanical Engineering

Universiti Teknologi Malaysia, 81310 Skudai, Johor Malaysia

Tel: 6-013-719-0498 E-mail: mohdfairuz.ahmadibrahim@malakoff.com.my

Mohammad Nazri Mohd Jaafar (Corresponding author)

Department of Aeronautical Engineering, Faculty of Mechanical Engineering

Universiti Teknologi Malaysia, 81310 Skudai, Johor Malaysia

Tel: 6-07-553-4851 E-mail: nazri@fkm.utm.my

Abstract

A study was conducted to investigate the flow pattern in a gas turbine combustor using numerical and experimental approaches. The main function of a combustor is to burn the air-fuel mixture then to admit the high energy burnt gases, with uniform and limited temperature, to drive the turbine blades. The gas temperature must not exceed a certain allowable temperature to prevent any damage of the blades. Flow pattern within the combustor has a great effect on the self sustaining flame, mixing of fuel and air, combustion intensity and combustor exit temperature uniformity. For this reason, a radial vaned swirler was used in this study to demonstrate its effect on the flow pattern within the combustor. The radial swirler vanes had an aerodynamically curved profile to allow the incoming axial flow to turn gradually and hence to inhibit the flow separation on the suction side of the vane. Therefore, smooth flow turning, higher tangential and radial-velocity components can be generated at the swirler exit with lower pressure loss. Three swirler vane angles 40°, 50° and 60° corresponding to swirl numbers of 0.35, 0.54 and 1.13 were used to evaluate their effect on the combustor aerodynamics. The results show that the swirl number has a direct proportion relationship with the size and shape of the central recirculation zone as well as the corner recirculation. It is concluded that with the use of 40°, 50° and 60° vane angles swirler, the vortex breakdown phenomena occurs at different intensities and sizes.

Keywords: Vane swirlers, Recirculation zone, Swirl number, Combustor flow pattern, CFD simulation

1. Introduction

The main focus of this research is to investigate the dynamics of the swirling flow in a can type gas turbine combustor. Aerodynamic curved vane radial swirlers with various vane angles were used. Flow pattern characteristics include velocity components and turbulent stresses, which are the main characteristics of the swirling flows, have been studied to understand the physical process both by experimental investigations and by numerical modeling CFD software Fluent 6.3 (Fluent user manual 2006).

In gas turbine combustor, it is necessary to maintain a stationary flame within a high velocity air stream. The flame propagates through the unburned mixture at a certain speed, accordingly the mixture have to be maintained within the same certain limits for flame stabilization. I.e. For flame travel in speed higher than mixture speed, the flame blowout occurs. In order to overcome this difficulty, a region of recirculation downstream the main burner or the bluff body should be established. This region creates areas of local low stream velocity equal to the flame speed which is used to hold the flame, hence to sustain the combustion.

Air swirlers are used as a flame holder by adding swirl to the incoming air. The presence of swirl results in setting up of radial and axial pressure gradients, which in turn influence the flow fields. In case of strong swirl the adverse axial

pressure gradient is sufficiently large to result in reverse flow along the axis and generating an internal circulation zone (Lefebvre, 1983; Mellor, 1990; Mattingly, 1996; Ganesan, 2003).

In addition swirling flows are used to improve and control the mixing process between fuel and air streams and enhance heat release rate (Beer and Chigier, 1972; Gupta *et al*., 1984).

1.1 Geometry

It is shown in the literature (Lefebvre, 1983) that curved-vane swirlers produce a larger recirculation region and a higher recirculated mass with lower pressure loss. Furthermore, strength of shear layer zones and turbulence intensities are augmented by changing flat vane to curved vanes (Lefebvre, 1983). Therefore curved vane swirlers are selected in this research to show its effect on the flow pattern within the combustor.

The geometrical details of the model swirler, which was carefully machined from mild steal, are given in Figure 1. The swirler has eight curved vanes, equally distributed around the circumference of its plate. The axial depth of inlet passage is 20 mm. Three different vane angles of 40°, 50° and 60 ° were used for this study.

1.2 Swirler number

The degree of swirl for a swirling flow is usually characterized by the swirl number, *S*, as was originally proposed by (Chigier and Beer, 1964):

$$S = \frac{2G_\phi}{R\,G_x} \tag{1}$$

where G_ϕ stands for the axial flux of the tangential momentum, G_x is the axial flux of the axial momentum, and *R* is the outer radius of the annulus. These two parameters, G_ϕ and G_x can be expressed as:

$$G_\phi = \int_0^R \rho(Wr)U\ 2\pi r\,dr \tag{2}$$

$$G_x = \int_0^R \rho U^2\ 2\pi r\,dr \tag{3}$$

where *U, W,* ρ are the axial velocity, tangential velocity and density respectively.

Shee *et al.* (1996) experimentally derived a modified correlation for swirl number that was found to be strongly dependent on Reynolds number (Re) and the vane angle for a radial-type swirl generator. The modified swirl number *S* correlated with both the Reynolds number and the vane angle. The derived correlation is:

$$S = C_1(\mathrm{Re})\,\sigma(\theta) \tag{5}$$

where C_1 is a function of Re and can be obtained from the following correlation:

$$C_1 = 0.28 \text{ x } [1 - \mathrm{sech}(0.026\mathrm{Re}^{0.67})] \tag{6}$$

The values of C_1 against Re is shown in Figure 2. It can be seen that for high Reynolds number, C_1 becomes constant and can be taken as 0.28.

σ is the ratio of the mean tangential and radial velocity components at the swirler exit, which has a radius R_1 as shown in Figure 3. Since σ is dependent only on the geometrical dimensions of the guide vane and the vane angle θ, the angle of the guide vane cord line with respect to the radial direction, the value of σ can be obtained from the following relations:

$$\sigma(\theta) = \frac{1}{1-\psi}\left(\frac{\tan\theta}{1+\tan\theta\tan(\pi/z)}\right) \tag{7}$$

where $\psi = z\,s/2\pi R_1 \cos\theta$ is a blockage factor that comes from the finite thickness of the guide vanes, z is the number of guide vanes, and s is the thickness of the guide vanes. The technical data of the three swirlers used in this study are listed in Table 1.

2. Test facility

The water model test facility was used in the present study for isothermal investigation of flow patterns. The test facility was an open circuit arrangement. The main components consisted of a storage tank, pump, flow meters, a control valve, a combustor model and an outlet pipe. The tubular combustor model was the main test section and was placed inside the main water channel test rig as shown in Figure 4. The whole water channel was filled with water in order to reduce distortion and improve visibility. The combustor model used in the water model test facility was geometrically equal to the one used numerical simulation. It was made from 3mm thick Perspex tube with 140 mm internal diameter and 400 mm length.

For flow visualization of the flow pattern, coconut milk liquid was introduced as tracer into the test section using syringe pump. Particle Image Velocimetry (PIV) system was used to measure the velocity vector in the combustor model. The system includes: 20 mJ, 10 Hz pulse Nd-YAG laser and 1024x1024 pixels Kodak CCD camera. The CCD ship inside CCD camera is exposed to first frame scattered light of seeding material particles when they exposed to Laser light sheet and full 1k x 1k pixel image is acquired. The CCD chip is then cleared and exposed to scattered light from the second pulse of the light sheet, and second full 1k x 1k pixel image is acquired. Both images are then transferred to the FlowMap Processor via digital connector. The minimum time between these two pulses is 1 μs.

The inlet volume flow rate into the water tunnel was 200 lit/min which is equivalent to 3.3 kg/sec. The water mass flow was kept constant for all the swirlers.

3. Numerical simulation

The isothermal flow inside the combustor was simulated using FLUENT 6.3 CFD codes. For the prediction of the isothermal flow fields within the combustion chambers, FLUENT solves the equations for conservation of mass and momentum in their time averaged form.

3.1 Geometry specification and grid set-up

The flow through eight vanes radial swirlers was divided into 8 separated volumes; each volume was simulated using hexahedral cooper approach Figure 5.a. The physical domain was decomposed to several volumes to facilitate meshing with cooper hexahedral structured grid. The geometry meshing was done to have a variable density distribution by mean of small mesh size was incorporated in high gradient zone and bigger size in low gradient zone. The grid set-up of the present work is shown in Figure 5.

3.2 Turbulence modeling

As mentioned above, FLUENT solves the equations for conservation of mass and momentum in their time averaged form for the prediction of isothermal flow fields. For the process of Reynolds decomposition and time averaging results in unknown correlation of the fluctuating velocity components, a turbulence model is required for equations closure purpose. In the present simulation, k-ε turbulence model with the default setting were used. Turbulence is represented by the standard k-ε model, which provides an optimal choice and economy for many turbulent flow (Kim, 1989).

Menzies (2005) had studied the behavior of five k-ε variants in modeling the isothermal flow inside a gas turbine combustor and compared the results with the experimental data of Da Palma (1988) for the velocity and turbulence fields. The studied models were the standard, the RNG, the realizable, the Durbin modified, and the nonlinear k-ε models. The results showed that the standard and the Durbin k-ε models gave the best agreement with the experimental data. This supported the finding of Jiang and Shen (1994) and Durst and Wennerberg (1991) where good agreement between k-ε model predictions and experimental result were reported.

3.3 Boundary conditions

The appropriate choice of boundary conditions is essential and is a critical part in modeling a flow accurately. Typical boundary conditions for FLUENT simulation were the inlet, the wall and the outlet boundaries. At the inlet of the computational region, the inlet boundary condition is defined as mass flow inlet while the exit boundary is defined as pressure outlet (gauge pressure at model outlet is 0.0). Some assumptions about boundary conditions that were not directly measured had to be made as follows:

- Velocity components and turbulence quantities at the inlet were constant
- Turbulence at inlet is calculated from the following equations (Versteeg and Malalasekera, 1995):

$$k_{inlet} = 0.002\,(u^2)_{inlet} \tag{8}$$

$$\varepsilon = \frac{k_{inlet}^{1.5}}{0.3\,D} \tag{9}$$

4. Results and discussion

In order to investigate the effect of vane angle on the flow pattern within the combustor model, the prediction results using Fluent code for 40°, 50° and 60° vane swirler are presented and discussed in the following section.

4.1 Numerical results

The swirl number was numerically computed for the three radial swirlers using the integration listed in Equations 2 and 3 on the swirler exit plane. The different values of swirl number are listed in Table 1. It can be observed that the calculated swirl number has a good agreement with that obtained from experimental correlation 5.

Figures 6, 7 and 8 show velocity contours within the recirculation zone for the various vane swirlers. These contours are obtained by plotting the axial velocities which ranges from maximum negative value to zero within the recirculation zone. The flow downstream of 40° swirler shows maximum reverse velocity of 0.63 m/s as seen from Figure 6 which shows the formation of recirculation zone. Maximum reverse velocity magnitude of 1.02 and 1.04 m/s are found for 50° and 60◦ swirlers respectively as seen from Figure 7 and Figure 8. The region with reverse velocity is termed as central recirculation zone (CRZ), which is mainly responsible for flame stabilization. As it can be seen from Figure 6 the recirculation zone has smaller size and shorter length with a width of 1.53 D_1 and length 4.85 D_1. This recirculation zone is critical as far the residence time for the reactant species and the heat transfer rate are concerned. The above parameters, in turn, depend on the volume and surface area of the recirculation zone. Hence, it will be useful to investigate the effect of the vane angle on the geometry namely the length, width, surface area and the volume of the recirculation zone. For the 50° vane swirler, it is seen from Figure 7 that the length and width of the recirculation zone has increased to 1.86 D_1 and 5.6 D_1 respectively compared to 40° swirler, and the longest length of 9.6 D_1 and widest width of 1.93 D_1 are obtained for the 60° vane angle swirler. Figures 6, 7 and 8 also show that the size of the corner zone circulation decreases with the increase of the swirl number.

Total Pressure loss is a dimensionless number and is defined as:

$$\text{Pressure Loss } \% = \frac{P_{o2} - P_{o1}}{P_{o1}} \tag{10}$$

where P_{o1} is the inlet stagnation pressure, measured at the inlet plenum upstream of the swirler and P_{o2} is the exit stagnation pressure, obtained in the combustor exit section. The variation of pressure loss with the swirl number is shown in Figure 9. It is seen that with increase in vane angle, pressure loss also increases. Furthermore, Figure 9 shows that there is a steep increase in pressure when the vane angle is increased from 50° to 60°. For efficient design of combustion chamber the pressure loss should be as minimum as possible, however, it should have more recirculated gases in the recirculation zone. So a compromise is required between the recirculation zone formed and the total pressure loss across the swirler. Taking these factors into account it is found that the total pressure loss across 60° swirler is nearly 1.7 times higher than that for 50° swirler as well the recirculation zone size of 45° is quite sufficient.

Briefly, the volume of recirculation zone formed for 50° and 60° swirlers are close with pressure loss for 60° swirler is approximately twice that of the 50° swirler. Hence it is concluded that the 50° vane swirler produces good recirculation zone with reasonable pressure drop, which may aid better combustion.

4.2 PIV results

Figures 10, 11 and 12 illustrate the streamlines profiles for 40°, 50° and 60° vane angle swirlers. The scale for PIV image is 150 pixel per 100 mm according with the datum sets at the value of zero. These figures show that, there are three recirculation zone created inside the combustor model. One recirculation was formed in the corner and two were formed in combustor central zone. The poor agreement between the experimental and numerical results owes to the following reasons:

- High swirl flow which makes residency time of the seeding particle in measuring stream-wise plane for two successive frames unfeasible.
- The curved surfaces such used in the circular combustor model reflect a fraction of the laser light sheet which affects the quality of the image hence the accuracy of measurement.

5. Conclusion

The study of the flow pattern inside a gas turbine combustor model using numerical simulation and experimental approach has been achieved. Numerical simulation was done using Fluent 6.3 and based on standard k-epsilon turbulence model while experimental work was accomplished using PIV system.

The three radial swirlers employed in this research show a significant effect on the flow pattern within the combustor model. The 40° vane angle swirler produced a small volume of recirculation zone while 50° and 60° vane angle swirler produced larger recirculation zone size. From the parametric study carried out with varying angle of vanes, it is found that 50° swirler is the best for producing appropriate recirculation zone with reasonable pressure drop.

It can be concluded also that with the increase of the swirler vane angle, the corner circulation or dead zone size decreased as the tangential to axial momentum ratio increased. This dead zone has a negative impact on the combustion process and the homogeneity of the wall temperature as well.

Finally, it is recommended for high swirling flow to use 3D PIV system instead of 2D PIV used in this paper.

References

Beer, J.M. & Chigier, N.A. (1972). *Combustion Aerodynamics*. Applied Science Publisher, London

Chigier, N. A., & Beer, J. M. (1964). Velocity and Static-Pressure Distributions in Swirling Air Jets Issuing from Annular and Divergent Nozzles. *Journal Basic Eng*., 788-796.

Da Palma, J. M. L. M. (1988). Mixing in Non-Reacting Gas Turbine Combustor Flows. *PhD Thesis,* University of London, UK.

Durst, F. & Wennerberg D. (1991). Numerical Aspects of Calculation of Confined Swirling Flows with Internal Recirculation. *Int. J. for Numerical methods in Fluids*, Vol. 12, 203-224.

FLUENT 6.3 User's Guide, Fluent Inc. 2006.

Ganesan, V. (2003). *Gas Turbines*. Tata McGraw-Hill, New Delhi.

Gupta, A.K, & Lilley, D.G., & Syred, N. (1984). *Swirl Flows.* Energy and Engineering Science Series, Abacus Press, England.

Jack D. Mattingly, (1996). *Elements of Gas Turbine Propulsion*. McGraw-Hill International Edition, Singapore.

Jiang, T. L. & Shen, C. H. (1994). Numerical Predictions of the Bifurcation of Confined Swirling Flows. *Int. J. for Numerical methods in Fluids*, Volume 19, 961-979.

Kim, Y.M. and Chung, T.J. (1989). Finite-Element Analysis of Turbulent Diffusion Flames. *AIAA Journal*, Vol. 27, No. 3, pp. 330-339.

Lefebvre, A. H. (1983). *Gas Turbine Combustion*. Hemisphere Publishing Corporation, first edition.

Mellor, M. (1990). *Design of Modern gas Turbine Combustors.* Academic Press.

Menzies, K. R. (2005). An Evaluation of Turbulence Models for The Isothermal Flow in A Gas Turbine Combustion System. *6th International Symposium on Engineering Turbulence Modeling and Experiments*, Sardinia, Italy.

Shee, H. J. & Chen W. J.& Jeng, S.Y. & Huang, T. L. (1996). Correlation of swirl number for a radial-type swirl generator. *Experimental Thermal and Fluid Science*, 12, 444-451.

Versteeg, H. K., & Malalasekera, W. (1995). *An Introduction to Computational Fluid Dynamics, the Finite Volume Method.* Longman Group Ltd.

Table 1. Technical data of the swirlers

Swirler	A	B	C
Vane angle, deg	40	50	60
Vane depth, mm	20		
No. of vanes	8		
Vane thickness, mm	6		
Outlet radius R_1, mm	20		
C_1 based Re_{in} = 21900	0.28		
S based on Equation	0.35	0.54	1.13
Calculated from CFD	0.45	0.587	0.97

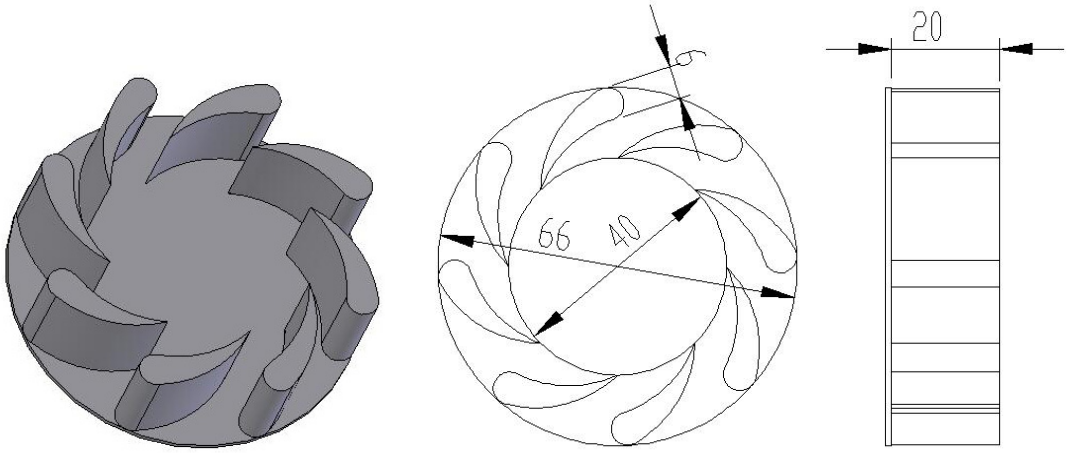


Figure 1. Geometrical details of the model swirler (all dimensions in mm)

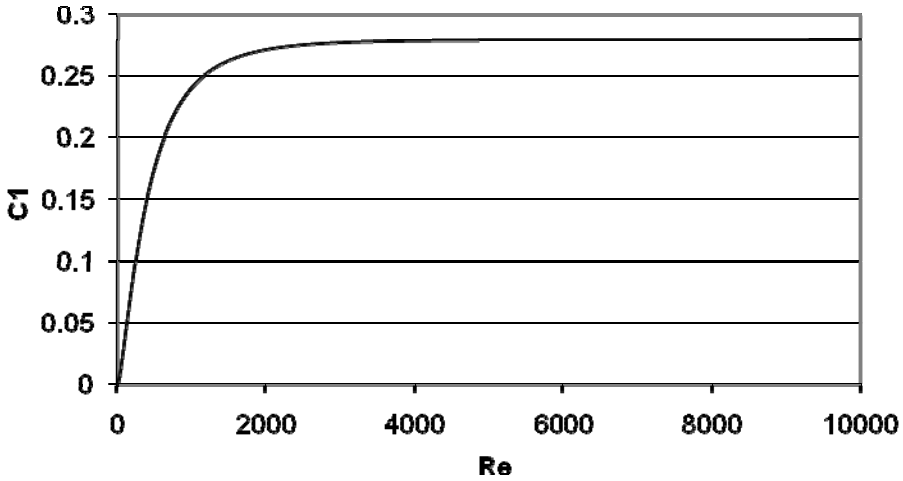


Figure 2. Correlation for the constant C_1 Reynolds number.

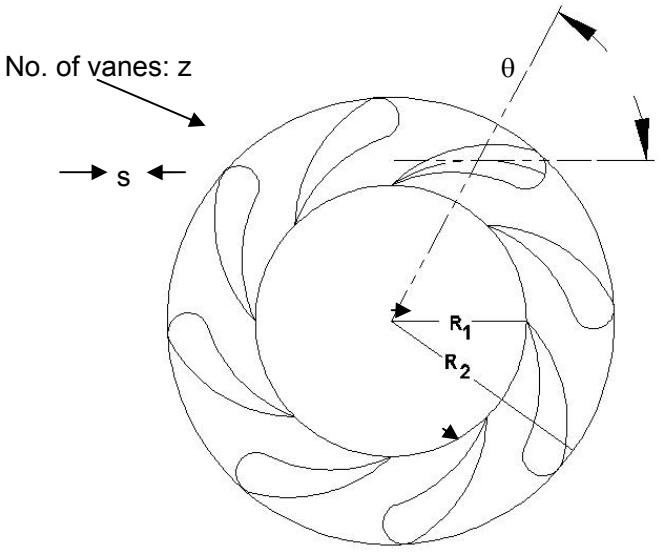


Figure 3. Guide vane cascade in the radial swirler

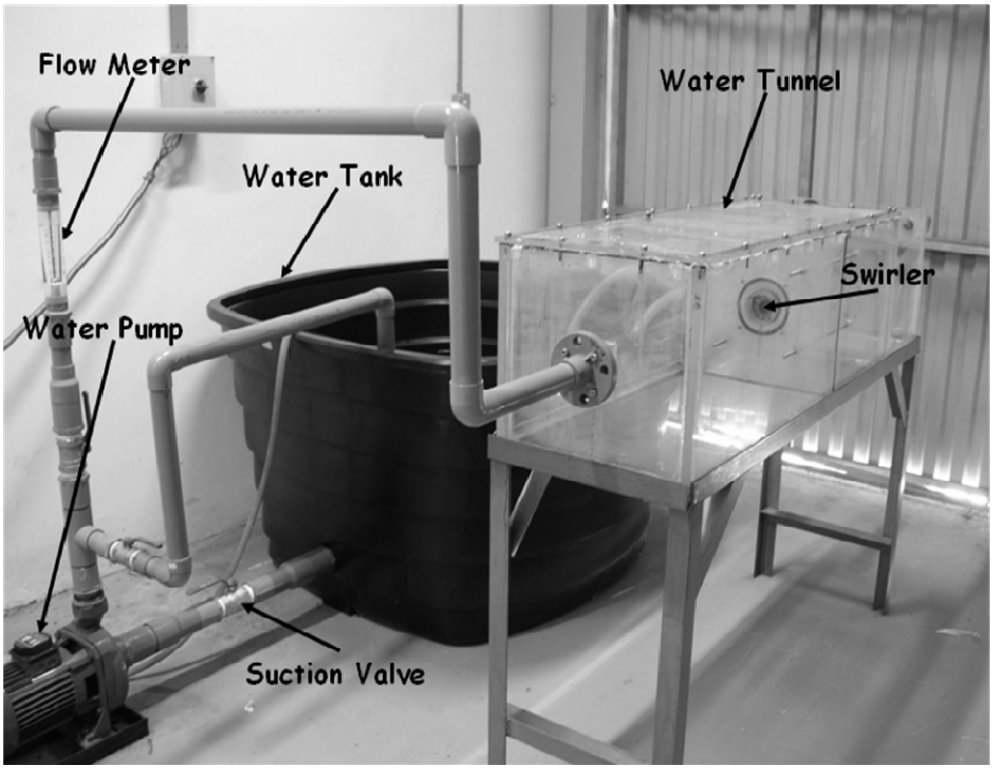
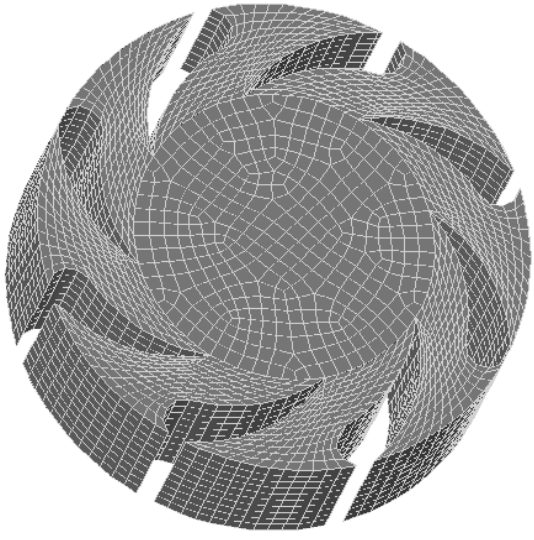
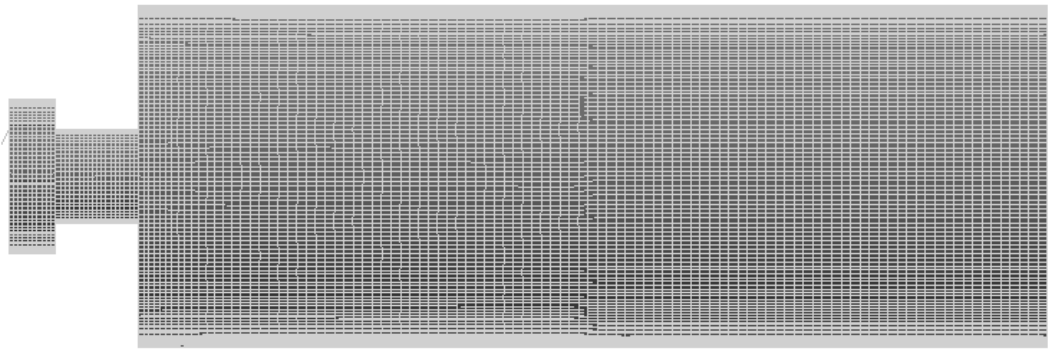


Figure 4. Water channel test rig



(a)



(b)

Figure 5. Grid set-up (a) Swirl passage volume mesh (b) Combustor model mesh

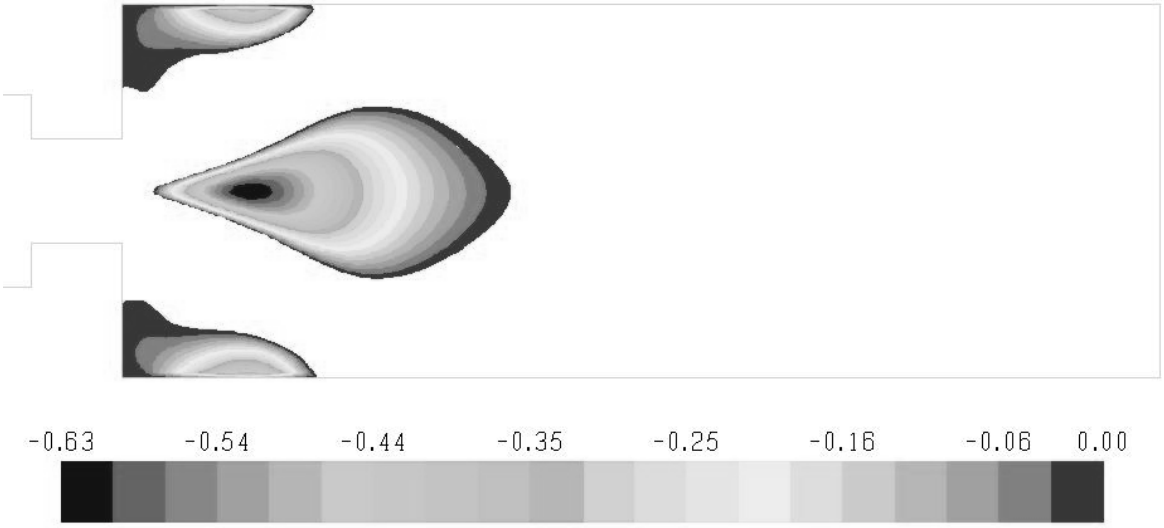


Figure 6. Velocity contour within recirculation zone for 40° swirler

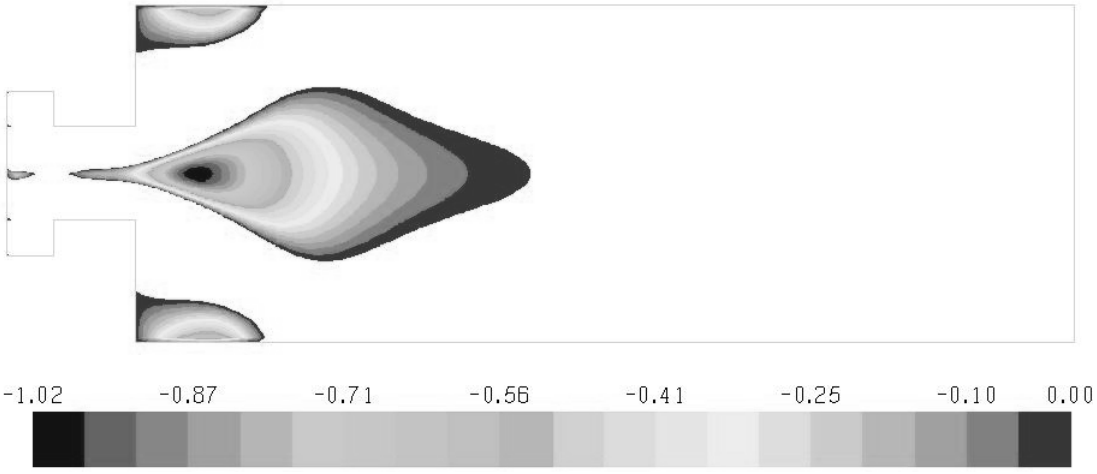


Figure 7. Velocity contour within recirculation zone for 50° swirler

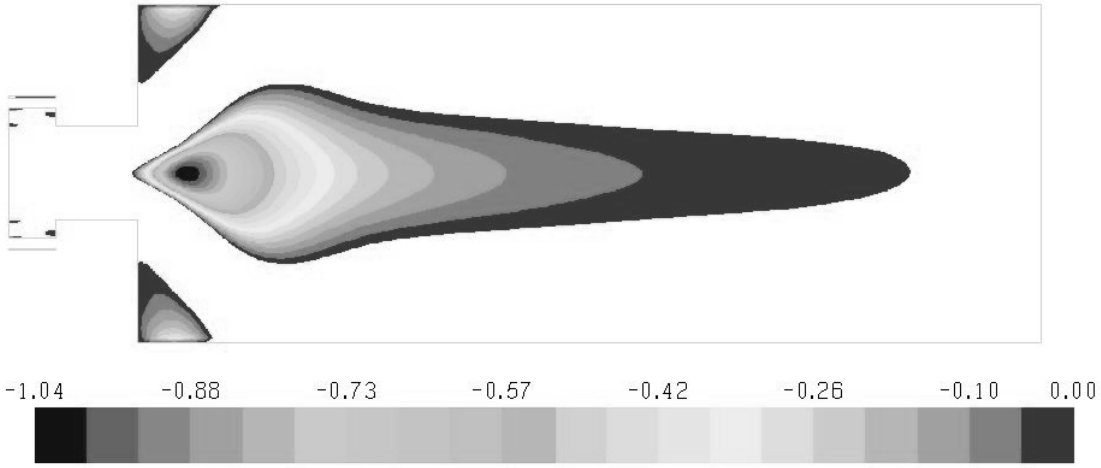


Figure 8. Velocity contour within recirculation zone for 60° swirler

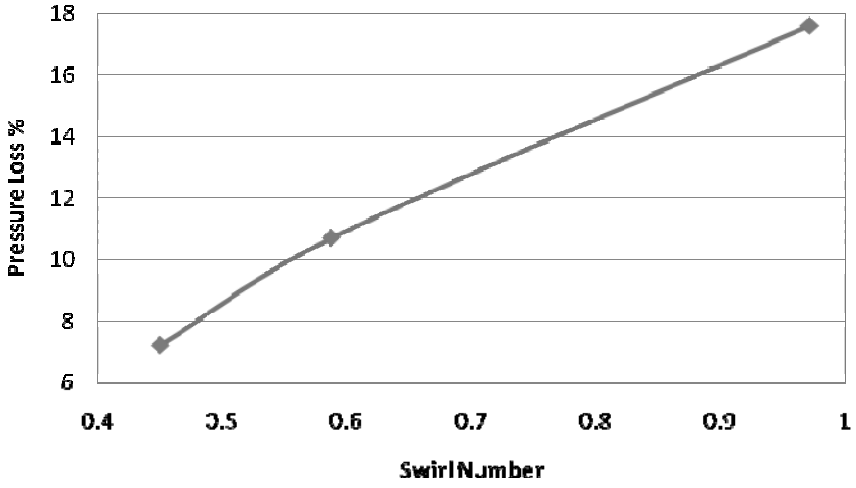


Figure 9. Total pressure loss % of various swirlers.

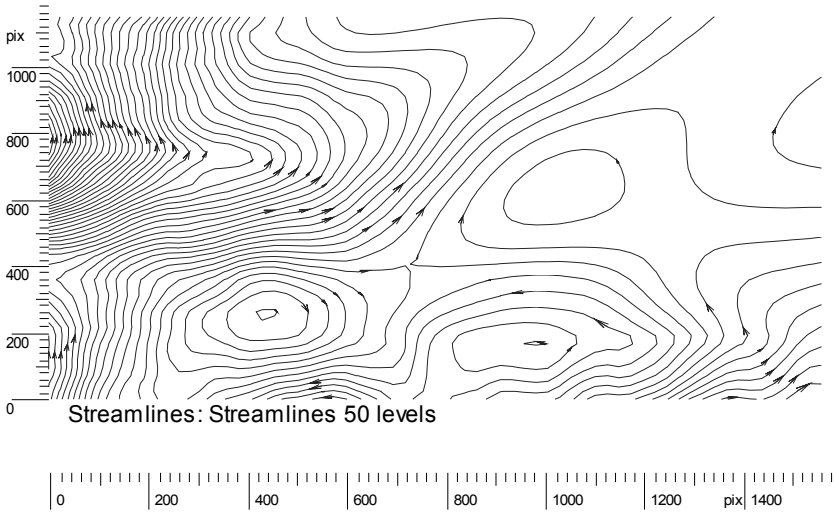


Figure 10. Streamlines of 40° vane angle swirler in the mid plane

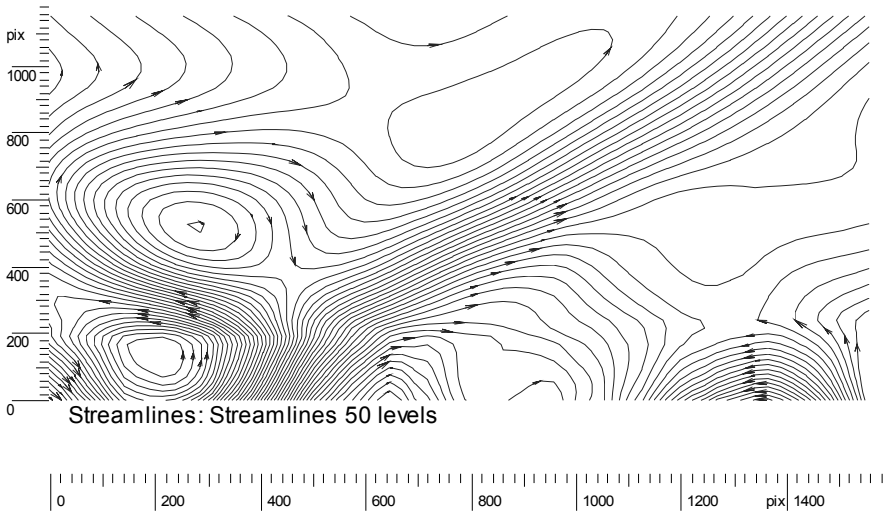


Figure 11. Streamlines of 50° vane angle swirler in the mid plane

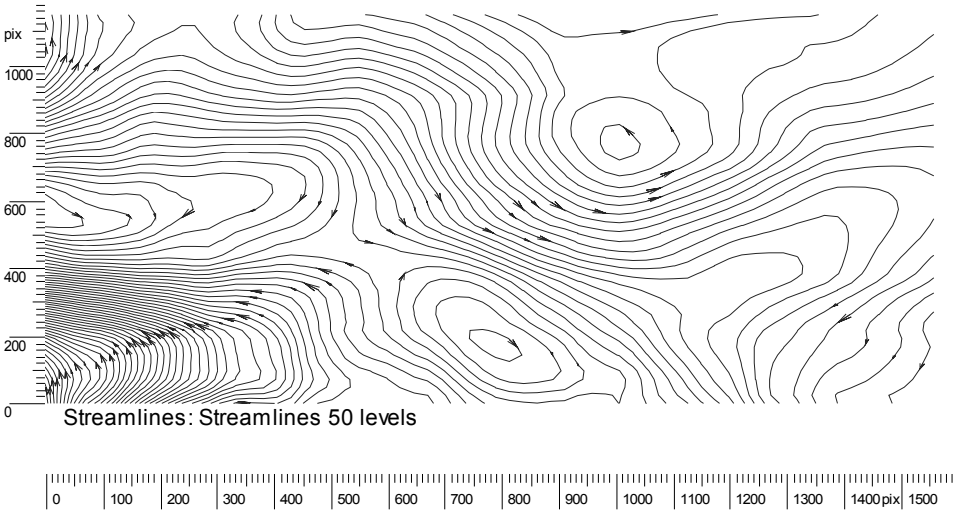


Figure 12. Streamlines of 60° vane angle swirler in the mid plane

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Research on Dynamic Facial Expressions Recognition

Xiaoning Peng & Beiji Zou

School of Information Science and Engineering

Central South University

Changsha 410083, China

E-mail: hhpxn@mail.csu.edu.cn

Department of Computer Science and Technology of Huaihua College

Huaihua 418008, China

E-mail: hhpxn@163.com

Lijun Tang

School of Information Science and Engineering

Central South University

Changsha 410083, China

E-mail: hhpxn@mail.csu.edu.cn

Ping Luo

Department of Computer Science and Technology

Tsinghua University

Beijing 100084, China

E-mail: luop@mail.tsinghua.edu.cn

This work is supported in parts by the National Natural Science Foundation of China under Grant No.60673093 and supported by Hunan Provincial Natural Science Foundation of China under Grant No.06JJ2065 and supported by Program for Changjiang Scholars and Innovative Research Team in University of China.

Abstract

Human-computer intelligent interaction (HCII) is usually based on facial expression recognition. A dynamic facial expression recognition method based on video sequence is proposed in this paper, which uses Gaussian of Mixture Hidden Markov Model. Firstly, we get some special facial expression regions, in which the motion features are extracted and described as phase form and then constituted to eigen-sequences. Secondly we use Gaussian of Mixture Hidden Markov Model to learn and test these eigen-sequences, and recognize six universal facial expressions: angry, disgust, fear, happy, sad and surprise. And we developed an experimental system based on our algorithm. The experimental results show that the computing time and the error of vector quantization is reduced, while the classification efficiency is improved.

Keywords: Facial expressions recognition, Optical flow, Gaussian of Mixture Hidden Markov Model, Eigen-sequences

1. Introduction

Human-computer intelligent interaction (HCII) is an emerging field of science aimed at providing natural ways for humans to use computers as aids. To interact with the human, the computer must be equipped with human communication skills, such as understanding, distinguishing and identifying the capacity of human emotional state. Facial expression includes rich information about human emotions, it is the main carriers of emotions, and it is an important way of understanding emotions. Facial expression plays a very important role in human life, and it is a main way of human nonverbal communication. In recent years, the reasons for renewed interest in facial expression recognition are multiple, but mainly due to people have more interests about Human-computer intelligent interaction.

Facial expression recognition deals with the classification of facial motion and facial feature deformation into abstract classes that are purely based on visual information(B. Fasel, J. Luettin, 2003). The expression recognition method has very much, according to expression data these approaches can be divided into two main categories: one kind is based on

the static image method, another kind is based on the dynamic image sequence method. In static approaches recognition of facial expression is performed using a single face image. These approaches are, mainly, based on the Facial Action Coding System (FACS)(Ekman P, Friesen WV, 1978), a presentation developed in order to allow human psychologists to code expressions from still images. Such works fail to incorporate the timing of expressions, which is a critical parameter in emotion recognition. Mase(K.Mase, 1991) used optical flow to estimate facial muscle action and obtained 80% recognition rate in user-trained mode, which also belongs to this approach. But the temporal characteristics of facial expressions were not fully utilized in Mase's method. In order to manifest motion information between frames in facial expression changing, another approach is produced, which is based on dynamic image sequence. J.Lien(J. Lien, 1998) used this method, he analyzed facial expression and estimated expression intensity based on dense flow and Hidden Markov Model (HMM). But because he used discrete HMM, much important information was lost in quantifying vector process, complex computation was needed in determining quantification code process.

In view of insufficiency in the expression recognition based on the static image method, this article uses the dynamic facial expression recognition method, which will fully use the temporal and spatial information when expression changes. We get some special facial expression regions, in which the motion features are extracted and described as phase form and then constituted to eigen-sequences. In view of insufficiency in vector quantization based on discrete HMM, this paper describes facial expression time-based eigen-sequences with 1-st order left-right Gaussian of Mixture Hidden Markov Model (MHMM), and recognize six universal facial expressions: angry, disgust, fear, happy, sad and surprise. Some examples is as follow(Neutral Anger Disgust Fear Happiness Sadness Surprise

Figure 1):

2. The Basic Structure of Facial Expression Recognition Experimental System

Facial expression recognition deals with the classification of facial motion and facial feature deformation into abstract classes that are purely based on visual information. Our basic structure of facial expression recognition experimental system consists of three steps (Figure 2): face acquisition, facial data extraction and representation, and facial expression recognition.

Face acquisition is a processing stage to automatically find the face region for the input images or sequences. Adaboost algorithm is used in this paper, which just detect face in the first frame and then track the face in the remainder of the video sequence. After the face is located, the next step is to extract and represent the facial changes caused by facial expressions. In this paper, the motion features are extracted and described as phase form and then constituted to eigen-sequences, which is useful to the dimension of motion feature. Facial expression recognition is the last stage of facial expression recognition experimental system. The facial changes can be identified as facial action units or prototypic emotional expressions. In this paper, we assume the facial expression sequences satisfy the Gaussian of Mixture model, and use Hidden Markov Model to train and test them.

3. Feature Extraction based on dynamic image sequences

Feature extraction methods can be categorized according to whether they focus on motion and deformation of faces and facial features. Motion extraction approaches directly focus on facial changes occurring due to facial expressions, whereas deformation-based methods do have to rely on neutral face images in order to extract facial features. This paper researches facial expression based on dynamic image sequences, which uses Motion extraction approaches to extract facial motion features between facial expression changing.

The use of optical flow to track motion is advantageous because facial features and skin naturally have a great deal of texture. Using feature vector construction, a low-dimensional weight vector in eigenspace can be obtained to represent the high-dimensional dense flows of each frame. Based on the displacement and weight vectors, the motion information is converted to symbol sequences from which we can recognize facial expressions.

3.1 Motion Feature Extraction

Image processing is performed in two steps. In the first step, a velocity vector is obtained from every two successive frames by using a gradient based optical flow algorithm(B. Horn and B. Schunck, 1981)(Freund Y, Schapire R E., 1997)(Gao Wen, Chen Xilin, 1999). To improve performance, the region for processing is confined to two small regions, one is eye-brow region, and another is mouth region. The angry expression's velocity vector of eye-brow region is as Figure 3. The eye-brow and mouth regions are defined as the characteristic parts in the extraction of features from different facial expressions. These regions were selected based on the results from three-dimensional measurements of expressions, and as such they proved to be the regions in which the changes were most pronounced.

3.2 Feature Vector Construction

In the second step, a feature vector construction processing is applied to a vertical and a horizontal component of the velocity vector field at the regions around an eye and the mouth.

(1)Compute the phase form of a vertical v_k and a horizontal u_k component of the velocity vector field, such as

$V_k = a_k e^{j\theta_k}$, $0 \le \theta_k < 2\pi$.

(2)Let R denote the facial area, R_1 denote the eye-brow sub-area, and R_2 denote the mouth sub-area. Let $A_i(R_j)$ denote the i-th group of motion vectors in the sub-area R_j, ($1 \le j \le 2$). Grouping of the motion vector is performed according to their orientation:

$$A_i(R_j) = \{V_k \mid V_k \in R_j, 2\pi i/8 \le \theta_k < 2\pi(i+1)/8\} \tag{1}$$

$0 \le i < 8, 1 \le j \le 2$

(3)The energy $E_i(R_j)$, within each direction in sub-area Rj is computed according to the equation:

$$E_i(R_j) = \sum_{V_k \in A_i(R_j)} \| V_k \| \tag{2}$$

The feature vector describing the motion of two consecutive frames is formulated from the set of E_i over the whole area R. The above formulation leads to a 8-dimensional feature vector of averaged powers, within each sub-area and all directions: $E(R_j)=[E_1(R_j),E_2(R_j),\ldots,E_8(R_j)], 1 \le j \le 2$.

4. Facial Expression Recognition based on dynamic image sequences

Facial expression Recognition can be regarded as a pattern recognition problem. It is necessary to model dynamic facial feature vector sequence in order to analyze facial expression sequence. Modeling facial expression needs to take into account the stochastic nature of human facial expression involving both the human mental state, which is hidden or immeasurable, and the human action, which is observable or measurable. For example, different people with the same emotion may exhibit very different facial actions, expression intensities and durations. Individual variations notwithstanding, a human observer can still recognize what emotion is being expressed, indicating that some common element underlies each motion. Therefore, the purpose of facial expression modeling is to uncover the hidden patterns associated with specific expressions from the measured (observable) data. Facial expression modeling requires a criterion for measuring a specific expression. It is desirable to analyze a sequence of images to capture the dynamics. Expressions are recognized in the context of an entire image sequence of arbitrary length. We will develop a recognition system based on the stochastic modeling of the encoded time series describing facial expressions, which should perform well in the spatio-temporal domain, analogous to the human performance.

Some other advantages of using HMMs are: HMM computations converge quickly making it practical for real time processing, it may evaluate an input sequence of uncertain category to present a low output probability, and a multi-dimensional HMM may be developed to integrate individual HMMs to give a robust and reliable recognition.

We usually divide HMM into two types, one is discrete HMM, and another is continuous HMM. As facial expression eigen-sequences are continuous signals. If we use discrete HMM to model eigen-sequences, it might be serious degradation association with vector quantization. Hence it would be advantageous to be able to use HMMs with continous observation densities.

Continuous HMM model that was based on a probability density distribution of the vector, such models are good or bad depends on the assumption that the probability distribution is consistent with the actual situation. Generally speaking, the probability distribution of some commonly used as Gauss distribution does not accurately describe its distribution, and therefore the Gauss mixed Hmm model (Gaussian of Mixture Hmm, in short MHMM), which uses several different centres, separated for different combinations of Gauss distribution vector is approaching the actual characteristics of the distribution. Generally speaking, the probability distribution of some commonly used models such as the Gauss distribution MHMM assumptions short periods expression vector characteristics mixed with Gauss distribution. Assuming that the Observer series $O=(O_1,O_2,\ldots,O_T)$ expression vector characteristics, whose probability distribution B meet Gauss mixed distribution, the form is:

$$b_j(O_t) = \sum_{m=1}^{M} c_{jm} \Gamma(O_t, \mu_{jm}, U_{jm}), \quad 1 \le j \le N \tag{3}$$

$$\Gamma(O,\mu,U) = \frac{1}{(2\pi)^{\frac{|O|}{2}} \sqrt{|U|}} \exp\{\frac{1}{2}(O-\mu)^T U^{-1}(O-\mu)\} \tag{4}$$

where O is the vector being modeled, c_{jm} is the mixture coefficient for the mth mixture in state j and Γ is any log-concave or elliptically symmetric density, with mean vector μ_{jm} and covariance matrix U_{jm} for the mth mixture component in the state j. Usually a Gaussian density is used forΓ. The mixture gains c_{jm} satisfy the stochastic constraint

$$\sum_{m=1}^{M} c_{jm} = 1 \quad \text{and} \quad c_{jm} \ge 0; \tag{5}$$

so that the pdf is properly normalized. i.e.,

$$\int_{-\infty}^{+\infty} b_j(x)dx = 1, \quad 1 \leq j \leq N \tag{6}$$

The facial expression image sequence is the same as human's pronunciation sequence characteristic, which has the time order also the irreversible process. So this article uses the left-right MHMM model (1st-order left-right MHMM) to describe the facial sequences. The model structure is as Figure 4. Along with the time increase, each condition only can shift from this condition to the right side neighboring next condition or the maintenance is same, in corresponding transition matrix A only then host diagonal element a_{ii} and the right vice-diagonal element a_{ii+1} permission non-vanishing, other elements all are the zero.

5. Experiment and Result

We adopt the six universal expressions defined by psychologists Ekman to do experiments, such as angry, disgust, fear, happiness, sadness and surprise. We collected separately for each group of 15 video series, a total of 90 video series. psychology research indicates that typical changes of muscular activities are brief, lasting for a few seconds, but rarely more than five seconds or less than 250ms0. So the duration of experiment video series is approximately 3 seconds. The change of expression for the experimental model is neural->apex->neural.

Firstly, we get some special facial expression regions according to the human face structure characteristic. And then carries on the normalization and the standardization. Secondly the motion features are extracted and described as phase form and then constituted to eigen-sequences with the method of part 3.2 described in this paper. Thirdly the eigen-sequences is divided into the training data and the test data, which is used to train the MHMM model parameter and test experiment result separately. We have gathered 90 groups of images sequences. After feature vector construction processing, we get 90 corresponding eigen-sequences. And 60 subassemblies were the training data sets, the other 30 subassemblies for test data set.

To estimate the parameters of a HMM the Baum_ Welch algorithm is used; here, re-estimation is repeatedly performed so as to maximize the generation probability of the training data. For the discrete case a reasonable probability can be obtained from initial values that are set at random. However, for the continuous case the constraints of the parameterized density would limit the range of parameters reachable from a initial value. Therefore, the random setting of initial values will converge into a local maximum in the generation probability, which is far away from the real maximum to be aimed at. To solve this problem, we apply a clustering operation on the training data and the statistical parameters of each cluster are used as an initial value for the parameter of the output probability density. The processing of training HMM parameter is as Figure 5.

In the system, the mix number M of MHMM is defined by the experiment. In order to select optimum mix number M, we have carried on many experiments. The experimental results are as Figure 6. The horizontal axis is mixture number M and the vertical axis is error recognition rate. When mix number M=4, 6, 8, 12, the error recognition rate all might achieve the minimum value. In order to reduce computation complexness, we selected M=4. Therefore, the mixture number M is 4 in this system.

To carry on the expression classification experiment with MHMM, we have designed a MHMM for each kind of expression, altogether six MHMMs, respectively is: anger(1), disgust(2), fear(3), happiness(4), sadness(5) and surprise(6). Each MHMM is a left-right Gaussian of mixture Hidden Markov Model, six such MHMMs constitutes a person face expression Maximum Likelihood classifier, see Figure 2. Given an observation sequence O= (O_1, O_2,…, O_T), the probability of the observation given each of the six models $P(O|\lambda_j)$, where $j \in [1,6]$ is computed using the Forward-backward algorithm0. The sequence is classified as the emotion corresponding to the model that yielded the highest probability, i.e., $c^*=\text{argmax}[P(O|\lambda_j)]$. We have carried on the test to 30 groups of characteristics sequences, the experimental result is as Table 1.

The experiment proved that, the system in the recognition dynamic expression sequence process, the recognition speed and the recognition result both is ideal. The happiness and disgust expression recognition rate is good, the anger expression recognition rate is relative bad, the overall recognition rate achieves 86.7%.

Use the [6], calculates this sequence separately in six kind of MHMM under model probability P (O| lamda j), j ∈ [1,6]. Takes probability value P (O| lamda j) the biggest MHMM correspondence expression for this observation sequence expression, namely takes c*=argmax [P (O| lamda j)] the human face expression serial number which distinguishes for the system.

6. Conclusion

An improved dynamic facial expression recognition experiment system based on the MHMMs model is realized in this paper. And six universal facial expressions: angry, disgust, fear, happiness, sadness and surprise is recognized. We describe facial motion feature based on phase form and construct feature vector, which compress the motion feature and simplified the computing process. In view of the dynamic facial expression analysis, we use Gaussion of mixture model to describe facial expression succession characteristic sequences, which avoided the vector quantification error in the

discrete HMM model. In order to obtain optimum initial value of Gaussion of mixture model, we use k-means algorithm. Which is more precisely describe expression eigen-sequences' probability distribution compares with the traditional model.

But this experimental system is only based on the small scope data acquisition work, uses in training and the test facial expression video frequency sequence is limited. In order to strengthen the experimental system vigorous and healthy, the next step of work may consider the expanded expression database, for example uses the Cohn-Kanade database(Takeo Kanade, Yingli Tian, Jeffrey F. Cohn. 2000). In this experimental system, in view of the different face characteristic region, used the same characteristic number to describe its motion feature, the next step of work might consider using different characteristic number in different face characteristic region separately.

References

B. Fasel, J. Luettin. (2003). Automatic Facial Expression Analysis: A Survey. *Pattern Recognition*, 2003, 36(1): 259~275.

B. Horn and B. Schunck. (1981). Determining Optical Flow. *Artificial Intelligence*, 1981, 17: 185~203.

Ekman P, Friesen WV. (1978). Facial Action Coding System. Palo Alto: Consulting Psychologists Press Inc., 1978.

Freund Y, Schapire R E. (1997). A Decision-Theoretic Generalization of on-line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 1997, 55(1): 119-139.

Gao Wen, Chen Xilin. (1999). Computer Vision: *Algorithm and System Principle*. Beijing: Tsinghua University Press, 1999. 83~91

J. Lien. (1998). Automatic Recognition of Facial Expression Using Hidden Markov Models and Estimation of Expression Intensity, The Robotics Institute, CMU.

K.Mase. (1991). *Recognition of Facial Expressions From Optical Flow*. IEICE Transactions, 1991, 74(10): 3474~3483.

Rabiner L R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE, 1989, 77(2): 257- 286.

Takeo Kanade, Yingli Tian, Jeffrey F. Cohn. (2000). *Comprehensive Database for Facial Expression Analysis*. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000: 46~53.

Table 1. The recognition result with MHMMs

	An	Di	Fe	Ha	Sa	Su
An	3	0	0	0	0	0
Di	0	5	0	0	0	0
Fe	0	0	4	0	0	0
Ha	1	0	1	5	0	0
Sa	0	0	0	0	5	1
Su	1	0	0	0	0	4

Neutral Anger Disgust Fear Happiness Sadness Surprise

Figure 1. Universal Facial Expressions from JAFFE

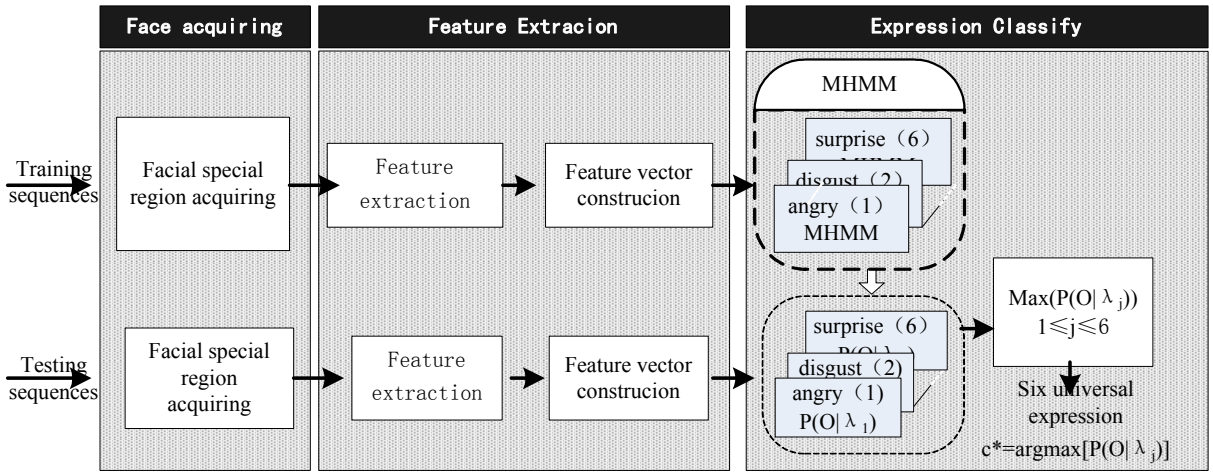


Figure 2. The Basic Structure of Facial Expression Recognition Experimental System

(a) Eye-brow Region (b) u_k Optical Flow (c) v_k Optical Flow

Figure 3. Example of velocity vector

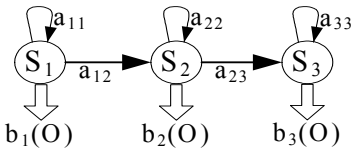


Figure 4. The structure of 1st-order left-right MHMM

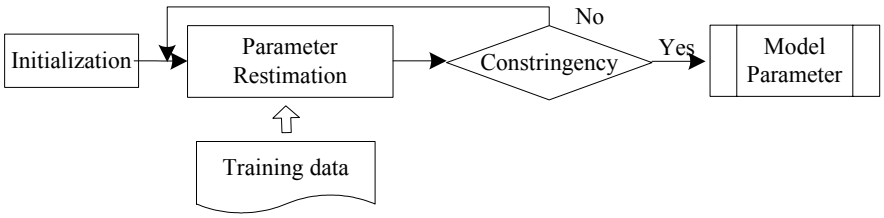


Figure 5. Training MHMM parameter

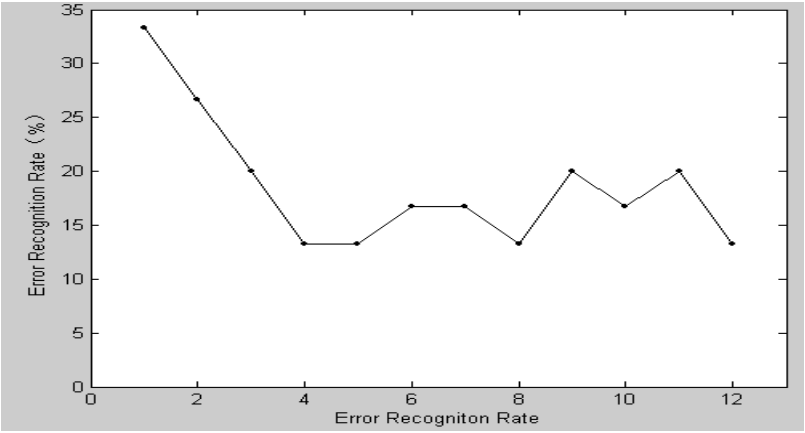


Figure 6. The error rate versus the mixture number M in the MHMMs

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Real Time Implementation of A New CDM-PI Control Scheme in A Conical Tank Liquid Level Maintaining System

Bhaba. P.K

Department of Chemical Engineering, Annamalai University

Annamalai Nagar-608002, TamilNadu, India

E-mail: drpkbhaba@gmail.com

Somasundaram. S

Department of Instrumentation Engineering, Annamalai University

Annamalai Nagar-608002, TamilNadu, India

E-mail: ssm98@rediffmail.com

Abstract

The work focuses on the development and real time implementation of Coefficient Diagram Method (CDM) based PI (CDM–PI) controller for a Conical Tank Liquid Level Maintaining System (CTLLMS). The process exhibits severe static non- -linear characteristics. Based on the Polynomial approach (CDM), the elements of PI controller are designed and implemented in the conical tank liquid level maintaining control loop. Performance of the CDM-PI controller and its robustness are analyzed and reported.

Keywords: CDM-PI, CTLLMS, Real time implementation

1. Introduction

In most of the chemical plants, level control is extremely important because desired production rates and inventories are achieved through proper control of flow and level. The performance of some processes such as chemical reactors depends critically on the residence time in the vessel which in turn depends on the level. At this point it is clear that level control is an important control objective. Due to the pronounced non linear nature of several chemical processes, interest in non linear feed back control has been steadily increasing over the last several years [Bhaba,P.K et al,2007]. Linear controllers can yield a satisfactory performance if the process is operated close to a nominal steady state or is fairly linear. But the performance of the controller degrades with change in operating point and process parameters. Advance controllers such as adaptive or predictive controllers works well even with model mismatch, but the design and implementation require on-line identification of the model. A nonlinear gain scheduling controller works satisfactorily only when the gain is changing for different operating points and time delay is not significant. Further more the non linear gain scheduling controller has to be tuned at every sampling time, which is relatively complex [Chidambaram, M, 1997]. Thus, there is an incentive to develop and implement feed back control schemes that takes the process non linearity in control calculations.

In the present work, a new attempt is made to design a PI Controller for CTLLMS using CDM. The explicit tuning formula of CDM-PI controller is proposed. The important features of CDM are adaptation of the polynomial representation for the plant and the controller, nonexistence or existence of very small overshoot in the closed loop response, obtaining the characteristic polynomial of the closed loop system efficiently by taking a good balance of stability. This technique leads to good robustness of the control system with uncertainty in the plant parameters. The strength of CDM is simple and can be designed for any plant [Hamamci, S.E, 2002].

The paper is organized as follows. Section 2 gives the basics of CDM and the CDM controller design procedure. In section 3, the proposed CDM-PI tuning formula is presented. Experimental set up and experimental works are dealt in section 4. Results and discussions are described in section 5. Finally concluding remarks are given in section 6.

2. Coefficient Diagram Method

2.1 Basics of CDM

The polynomial algebraic method namely CDM was developed and introduced in control engineering in the year 1998 [Manabe, S, 1998]. The merits of the classical and modern control techniques are integrated with the basic principles of

CDM. CDM uses polynomial expressions for both the plant and the controller. In this representation, all equations are dealt with numerator and denominator polynomials independent from each other, better results can be achieved against pole-zero cancellations. In this approach, the type and degree of the controller polynomials and characteristic polynomial of the closed loop system are defined at the beginning. Considering the design specifications, coefficients of the controller polynomials are found later. CDM is an efficient and fertile control tool with which very good control systems can be designed. It is easy to realize a controller under the conditions of stability, time domain performance and robustness. The close relations between these conditions and coefficients of the characteristic polynomial can be easily found. That means CDM is not only effective for control system design but also for controller tuning.

2.2 CDM controller design procedure

The standard block diagram of the CDM control system is shown in Figure 1, where y is the output, r is the reference input, u is the controller signal and d is the external disturbance signal. $N(s)$ and $D(s)$ are numerator and denominator polynomials of the transfer function of the plant. $A(s)$ is the forward denominator polynomial while $F(s)$ and $B(s)$ are the reference numerator and the feedback numerator polynomials of the controller transfer function. Since the transfer function of the controller has two numerators, it resembles to a 2DOF system structure. $A(s)$ and $B(s)$ are designed as to satisfy the desired transient behavior, while pre-filter $F(s)$ is determined as zero order polynomial and used to provide the steady-state gain [Manabe, S. and Kim,Y.C, 2000].

From Figure 1, the output of the CDM control system is given by

$$y = \frac{N(s)F(s)}{P(s)} r + \frac{A(s)N(s)}{P(s)} d \tag{1}$$

where $P(s)$ is the characteristic polynomial of the closed-loop system. This polynomial is a Hurwitz polynomial with real positive coefficients and defined by

$$P(s) = A(s)D(s) + B(s)N(s) = \sum_{i=0}^{n} a_i s^i , \qquad a_i>0 \tag{2}$$

The polynomials, $A(s)$ and $B(s)$ appearing in the CDM control structure are given as

$$A(s) = \sum_{i=0}^{p} l_i s^i \text{ and } B(s) = \sum_{i=0}^{q} k_i s^i \tag{3}$$

where the condition p ≥ q must be satisfied for practical realization.

The CDM design parameters, namely equivalent time constant (τ) and stability indices (γ_i) are chosen as

$$\tau = t_s / (2.5 \approx 3), \tag{4a}$$

Where t_s is the user specified settling time.

$$\gamma_I = [2.5 \;\; 2 \;\; 2 \;\; \ldots] \tag{4b}$$

The above γ_i values are from the standard form [Manabe, S, 1998] and these values can be changed in order to satisfy the desired performance.

The controller polynomials defined in (3) are replaced in

$$P(s) = A(s)D(s) + B(s)N(s) = \sum_{i=0}^{n} a_i s^i , a_i>0 \tag{5}$$

Hence the coefficients of this characteristic polynomial P(s) are expressed in terms of K_i and l_i (i.e.) P(s) is expressed in terms of the coefficients of the controller polynomials. Using the design parameters ($\tau$and γ_i), a target characteristic polynomial ($P_{target}(s)$) is determined as

$$P_{target}(s) = a_0 \left[\{ \sum_{i=2}^{n} (\prod_{j=1}^{i-1} \frac{1}{\gamma_{i-j}^{j}}) (\tau s)^i \} + \tau s + 1 \right] \tag{6}$$

Equating the two polynomials represented in (5) and (6), a Diophantine equation [Kucera, V, 1993] of

$$A(s)D(s) + B(s)N(s) = P_{target}(s) \tag{7}$$

is obtained. The controller parameters (K_i and l_i) are computed by solving this equation easily.

3. Proposed New CDM-PI Controller

The CDM –PI control structure is shown in Figure 2. Here $C(s)$ is the main controller and $C_f(s)$ is feed forward controller. It can be shown that the steady state error to unit step change and unit step disturbance become zero [Hamamci,S.E,et al.2007] if the conditions

$$\lim_{s\to 0} C(s) = 0 \qquad \text{and} \qquad \lim_{s\to 0} \frac{C_f(s)}{C(s)} = 0 \tag{8}$$

imposes on the controller. The most important result satisfied this condition is that C(s) must include an integrator. In this case, *C(s)* can be chosen as

$$C(s) = K_c \left(1 + \frac{1}{T_i s}\right) \tag{9}$$

in the type of the conventional PI element and $C_f(s)$ is an appropriate element satisfying in (8).

When Figure 2 is connected with Figure 1, the polynomials of $C(s)$ and $C_f(s)$ in Figure 2 are expressed by $B(s)/A(s)$ and $F(s)/B(s)$ respectively. Here the CDM controller polynomials are chosen as follows:

$$A(s) = s, \tag{10a}$$

$$B(s) = k_1 s + k_0 \tag{10b}$$

The numerator polynomial *F(s)* which is defined as the set-point filter element is chosen to be

$$F(s) = P(s)/N(s)\big|_{s=0} = P(0)/N(0) = 1/K = k_0. \tag{11}$$

This way, the value of the error that may occur in the steady-state response of the closed-loop system is reduced to zero.

Now substituting the expression for C(s) in (9) we have

$$\frac{B(s)}{A(s)} = K_c \left(1 + \frac{1}{T_i s}\right) \tag{12}$$

Finally, equating the coefficients of the terms of equal power, the parameters of PI controller in terms of CDM controller polynomials are obtained as follows:

$$K_c = k_1 \quad \text{and} \quad \mathrm{T_i} = k_1/K_0 \tag{13}$$

Due to the structure of PI controller and approximation of the process dead time to be of first order by pade approximation, the parameters of PI controller designed by CDM can be obtained only by specifying the stability index γ_i because the equivalent time constant τ has been defined implicitly.

The parameters of feed forward controller is found to be

$$C_f(s) = \frac{F(s)}{B(s)} = \frac{k_0}{k_1 s + k_0} = \frac{1}{T_i s + 1} \tag{14}$$

Note that the parameters of $C_f(s)$ depend on PI parameters directly. Therefore, the designer does not need the extra calculation for the feed forward controller.

4. Experiments and Analysis

4.1 Experimental Set up

The schematic diagram of experimental set up is shown in Figure 3. The setup consists of a mild steel column of 34 cm diameter, height of 60 cm, opened to the atmosphere at the top. Flow rate is metered with Rota meter at the inlet. An RF capacitance level transmitter is used to measure the level in the tank (0-25cm). The output current signal (4-20 mA) from the sensor is processed using a VAD104, a multifunction, high-speed Analog and Digital Converter (ADC) interface board, to digital value. This digital value is read back as level and compared to the set point .The real time PI control algorithm written in "C" provides an appropriate control signal, which is again a digital value. This digital signal is converted to an analog

(4- 20 mA) signal in a digital to analog converter using VAD104.This current signal is converted to a pneumatic signal in an ABB make I/P converter. This pneumatic signal is employed for actuation of the final control element, an RK make control valve that is normally open with C_v of 2.0. Control algorithm is implemented using a P4 - PC, which is interfaced to the CTLLMS. In open loop scheme, after the level reaches a steady state, a step magnitude of +10% DAC output to control valve is given. The level in the tank varies and this variation in level (through RF capacitance Level transmitter) is recorded against time until a new steady state is reached. This recorded data are converted into fractional response and plotted against time to get process reaction curve. Using the results of S-K identification method [Sunderasan,K.R. and Krishnaswamy,P.R , 1978] the parameters for the model of CTLLMS are estimated from the reaction curve. Similarly at different steady state values and at different step magnitudes in DAC output the model parameters are identified and tabulated in Table 1.

4.2 Experimental Works

From Table 1, Worst case of the model parameters (larger process gain (K_p = 1.88), larger time delay (L = 4.66s) and smaller time constant of the process (τ_p = 56.95s)) are chosen, and the CDM-PI controller settings (Kc, Ti and $C_f(s)$) for the stability indices $\gamma_1 = 2.5$, $\gamma_2 = 2$are determined and tabulated in Table 2.

Experimental runs for set point tracking of ±5% and ±10% at the operating point of 40% in CTLLMS is carried out. The tracking responses are recorded in Figure 4.

5. Results and Discussion

5.1 Performance of CDM - PI controller in CTLLMS

The performance of the CDM - PI Controller is evaluated for various set points tracking at the operating point of 40% in CTLLMS. The performance measures (ISE and IAE) derived from Figure 4 is reported in Table 3.Total variation (TV) of the output (y) is also considered for evaluating the performance. By adopting the expression of TV = $\sum_{k=1}^{\infty} |y(k+1) - y(k)|$ the TV indices for CDM - PI controller is calculated and presented in Table 3. The minimum TV index clearly indicates the smoothness and more consistent output signal [Chen, D and Seborg, D.E, 2002]. In addition, Table 4 gives the time domain performance of the proposed controller. Together with the mentioned Figure, the tables indicate that CDM–PI controller gives minimum ISE, IAE, minimum total variation, fast settling time and minimum overshoot in all set point tracking cases.

To analyze the robustness of the proposed controller, an experimental run at other operating points of 60% and 72% level in CTLLMS are carried out. The results of set point tracking of ±5% and ±10% at these operating levels are recorded in Figures 5 and 6. The performance measures tabulated in Tables 3and 4 clearly indicate the supremacy of the CDM-PI controller.

6. Conclusions

In this work, a CDM - PI control schemes for a CTLLMS is designed. Real time implementation of this control scheme is carried out in a Conical Tank Liquid Level Maintaining System. The performance of the control schemes in set point tracking cases is analyzed. The results clearly favor CDM-PI control scheme. In addition, the Robustness of the controller is also investigated. It concludes that CDM - PI control techniques works well for non linear systems.

References

Bhaba,P.K., SathishBabu, S., Asokan, A., & Karunanith, T. (2007). Real time implementation of WMPI controller for conical tank liquid level system. *J.Appled Sci*, 7 (15), 2194 – 2197.

Chen, D., & Seborg, D.E. (2002). PI/PID controller design based on Direct synthesis and disturbance rejection. *Ind.Eng.Chem.Res*, 41, 4807 – 4822.

Chidambaram, M., Reddy, B.C., & Al-Gobaisi, D.M.K. (1997). Design of robust SISO PI controllers for a MSF desalination plant. *J.Desalination*, 109, 109 -119.

Hamamci, S.E., Koksal, M., & Manabe,S. (2002). On the control of some nonlinear systems with the Coefficient Diagram Method. *4th Asian Control Conference*, Singapore.

Hamamci,S.E., Bhaba,P.K., Somasundaram,S., & Karunanith,T. (2007). Two Degree of Freedom Transient response control for unstable Processes. *TIMA*, 145 – 150.

Kucera, V. (1993). Diophantine Equations in control-A survey. *Automatica*, 29,1361- 1375.

Manabe, S. (1998). The Coefficient Diagram Method. *14th IFAC symposium on Automatic control in Aerospace, 24 – 28.*

Manabe, S. & Kim, Y.C. (2000). Recent development of Coefficient Diagram Method. *ASSC'2000 3rd Asian Control Conference*, Shanghai.

Sunderasan,K.R. & Krishnaswamy, P.R. (1978). Estimation of time delay, time constant parameters in time, frequency and lap lace domains. *Can.J.Chem.Eng*, 56,257 – 262.

Table 1. Identified model parameters at different steady state conditions

Step size	K_p	τ	L
36-46	1.88	85.76	16.41
38-48	1.51	76.38	12.39
42-52	1.66	111.89	33.89
46-56	1.00	80.94	34.66
50-60	1.42	56.95	23.32

Table 2. CDM - PI Controller Parameters

Tuning Rules	K_c	T_i	C_f (s)
CDM - PI	2.2	59.07	1/(59.07s + 1)

Table 3. Performance measures in terms of ISE, IAE and Total Variation at operating point of 40%, 60% and 72% for CDM – PI controller

Set Point Tracking cases	%ISE			%IAE			Total Variation (Output)		
	CDM - PI Controller- CTLLMS								
	Operating Points (%)								
	40	60	72	40	60	72	40	60	72
+05%	581.47	700.64	616.88	192.7	262.4	298	21.7	14.3	23.7
-05%	386.16	499.81	781.27	148.4	215.7	307.3	18.1	15.8	18.3
+10%	1979.56	2455.28	2499.7	311.8	441	531.8	24.2	24.7	31.3
-10%	1461.5	1828.78	2461.3	230.6	335.8	448.4	27.1	22	23

Table 4. Time domain Performance measures at operating point of 40%, 60% and 72% for CDM – PI controller

Set Point Tracking cases	Settling Time (Sec)			%Maximum overshoot			Rise Time (Sec)		
	CDM - PI Controller- CTLLMS								
	Operating Points (%)								
	40	60	72	40	60	72	40	60	72
+05%	480	321	744	1.3	1.5	1.8	123	174	147
-05%	459	273	690	1.7	2	2	84	99	135
+10%	411	492	768	1.6	2.5	3.1	138	153	159
-10%	474	639	678	1	2.5	3.1	102	108	132

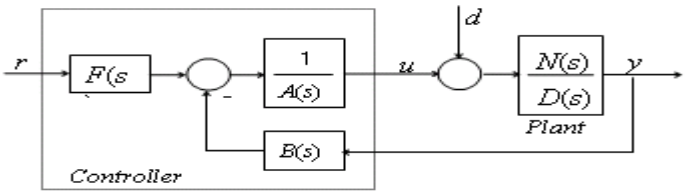


Figure 1. Block diagram of CDM control system

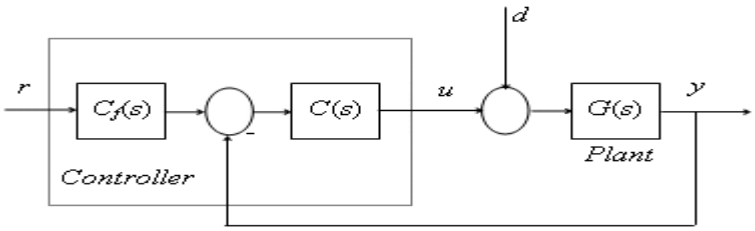


Figure 2. CDM based PI control system

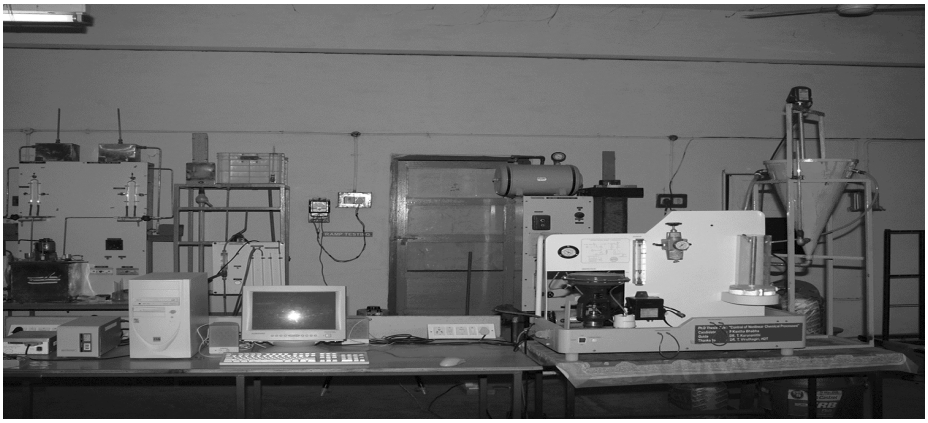


Figure 3. Experimental set up of CTLLM

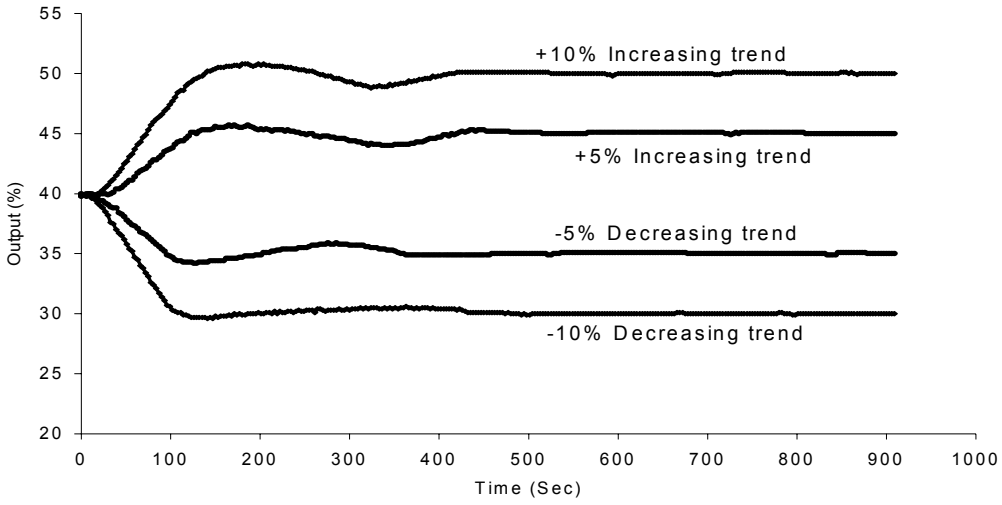


Figure 4. CDM-PI servo responses for step sizes of ±5% and ±10% at the operating point of 40%

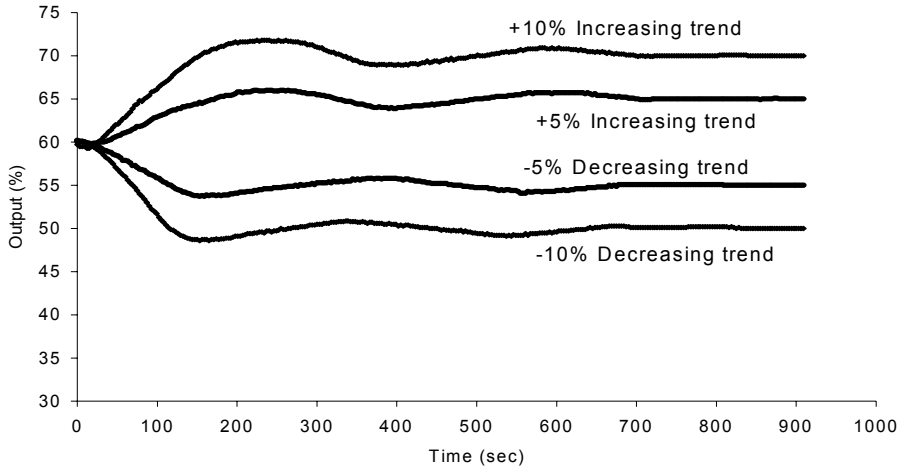


Figure 5. CDM-PI servo responses for step sizes of ±5% and ±10% at the operating point of 60%

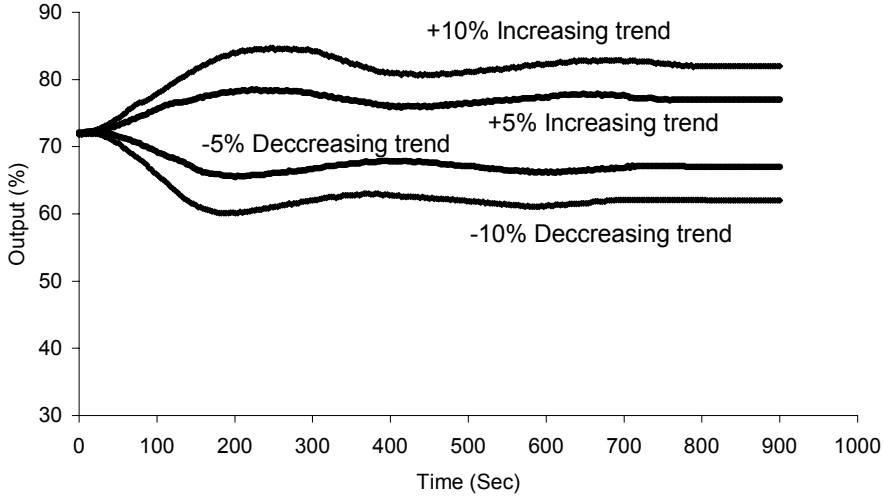


Figure 6. CDM-PI servo responses for step sizes of ±5% and ±10% at the operating point of 72%

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

The Testing Method of the Co-Movement of A+H Stock Prices

Jing Xiang, Susheng Wang & Jiaji Hao

Department of Economics and Management

Harbin Institute of Technology Shenzhen Graduate School, Shenzhen 518055, China

E-mail: carol.hsiang@163.com

Abstract

In this paper, the technical methods utilized for testing the co-movement of A+H stock prices of dual listed companies were analyzed through co-integration test, the Granger causality test, variance decomposition and Geweke indicators. Currently, empirical research on A+H stock markets are mainly based on two perspectives: empirical analysis of international asset price equalization theory and empirical analysis of measurement technology. Commonly there are four base unit test methods: Dickey-Fuller test, Augmented Dickey-Fuller test, Phillips and Perron test and KPSS test. These indicators and methods are then analyzed and evaluated. Research has not yet found an ideal approach to link A+H stock markets, although the above mentioned methods complement each other. In short, these test theories for development are only used in practice. With expanding research, in the future, there will be more powerful theoretical tools for empirical analysis.

Keywords: The dual listed companies stock, Co-movement, Testing method

1. INTRODUCTION

Co-movement refers to a strong correlation among yields of different markets, or that different stock prices share a common long-run equilibrium relationship or possess a long-term trend of synchronized movement. With communication technology and data processing technology improvement, finance theory and technology development, and development of derivatives market, the world capital market is deregulating, as well as the rapid expansion of institutional investors for spreading the risks. These forces will inevitably promote the linkage of A+H stock markets.

Research on the linkage of the trends of dual-listed stocks, on the micro level, will help vast numbers of investors in securities to perform portfolio analysis, and forecast prices to spread risks and raise revenue. At the same time, will also help listed companies specify financing strategy, and realize capital internationalization. From a macro perspective, a linkage of the stock reflects the overall linkage of the stock markets. It is valuable for regulatory authorities to enhance policy and better safeguard against financial risks. Many of the financial forecasts are based on the long-run equilibrium and short-term dynamic relationship. Inspection of the direction and the strength of the trend will help establish future profit strategy. If no co-integration exists, it means that diversified investment in these markets will be accessible for long-term hedging. If there is co-integration relationship, we can create short-term profit models.

2. MAIN INDICATORS FOR TESTING THE CO-MOVEMENT

Currently, empirical research on A+H stock markets are mainly based on two perspectives; empirical analysis of international asset price equalization theory and empirical analysis on measurement technology. The former analyzes whether the same risk securities have the same prices or not. If the test is passed, this shows that the linkage effects of markets are so high that the realization of asset price movements is in absolute synchronization. According to different research methods, the model is subdivided into the international capital asset pricing model and international arbitrage pricing model. In fact, the research literature can be found in Stehle (1977) fixed proceeds CAPM, Errunza and Losq (1985) moderate market segmentation theory, Cho, Eun and Senbet (1986) of fixed-income IAPT, Bekaert and Harvey (1995) time-varying expected return IntCAPM and so on. The conclusions of these models are not the same because their conditions and forms are different. International asset price equalization theory is impeccable from a purely theoretical point of view, but in empirical testing, inevitably encounters problems on the asset pricing model and linkage testing among A+H stock markets. Therefore, the usefulness of the theory for testing the market linkage is a greatly reduced. However, the rapid development of measurement technology makes up for the shortcomings of securities market linkage measure tools.

Some scholars try to bypass the asset pricing model, from a practical point of view of technical indicators, to inspect the linkage of the securities markets, and bring out the relevant parameters, co-integration, Granger causality, such as a series of theoretical methods and their expansion.

3. CO-INTEGRATION THEORY

Before introducing the co-integration theory, we first introduce several important basic concepts. a) Stationary time series, set Y is time series, and if $E(Y)=\lambda$ and $Var(Y)=\delta_y^2$, Y is a stationary time series, that is $I(0)$. For stationary series, its mean and variance are unchanged, even by the temporary impact of external factors. But in the long term, it will return to its average level. b) Non-stationary time series: variance sequence does not return to the average level of convergence and the variance unlimited increases. c) Single co-integration: If a non-stationary time series experience k times difference, it constitutes a stationary time series, then called it k single-band, that is $I(k)$.

At present, co-integration test has become a major method to test linkage of A+H stock markets.

A. Unit Root Test

Before the statistical analysis of time sequence, we must first determine series integration level using the unit root test, to determine whether the time series is smooth, and avoid the emergence of pseudo-regression. Normally, there are four unit root test methods: Dickey-Fuller (DF) test, Augmented Dickey-Fuller (ADF) test, Phillips and Perron (PP) test and KPSS test.

1)Dickey-Fuller(DF)Test:

Standard DF test is hypothesis testing the following relationship:

$$\Delta y_t = \alpha y_{t-1} + x_t' \delta + \varepsilon_t \tag{1}$$

The replacement of the original assumption:

$$H_0 : \alpha = 0 \quad H_1 : \alpha < 0$$

Use traditional statistics t to test α: $t_\alpha = \hat{\alpha} / (se(\hat{\alpha}))$

The above is a simple unit root test and this method is only effective when the sequence is $AR(1)$. If the data sets have a correlation in a higher time lag, then hypothesis that residual value is a white noise has not been established.

2)Augmented Dickey-Fuller (ADF)Test

Dickey and Fuller (1979) proposed the Dickey-Fuller test, which is based on the assumption that regression residuals are white noise, but in reality the regression residuals may have a serial correlation. In order to resolve this problem, Said and Dickey supposed an Augmented Dickey-Fuller Test. There are three test models:

Model 1: Sequence with no intercept and no time trend items:

$$\Delta z_t = \gamma z_{t-1} + \sum_{i=1}^{R} a_u \Delta z_{t-i} + \varepsilon_t \tag{2}$$

Model 2: Sequence with intercept but no time trend of items:

$$\Delta z_t = a_0 + \gamma z_{t-1} + \sum_{i=1}^{R} a_u \Delta z_{t-i} + \varepsilon_t \tag{3}$$

Model 3: sequence with intercept and time trend items:

$$\Delta z_t = a_0 + a_1 t + \gamma z_{t-1} + \sum_{i=1}^{R} a_u \Delta z_{t-i} + \varepsilon_t \tag{4}$$

Time series data z_t, and ε_t is the error term. In these three styles it is white noise. R is lag number, and the choice of multi-lag can be selected through different principle (such as: SBC or AIC) to determine ε_t is white noise, t is the trend of the time.

Null hypothesis $H_0 : \gamma = 0$, if refuse null hypothesis, it means that the time series is a smooth sequence, which is $I(0)$; If cannot refuse null hypothesis, it means it is a non-stationary sequence. When a non-stationary time series experience became a smooth sequence after many times difference, this sequence is $I(d)$.

3)Phillips and Perren(PP)Test

In ADF test, has put time-series variable self-related issues into account. The implied residual seizure model must be non-self relevant and have the same qualitative change. Philips and Perron (1988) proposed non-parameter method based on the function of the central limit theory to amend the DF and ADF, relaxing the basic assumptions that the same qualitative change in the DF test, and developed the PP test, allowing residual items may have relevance and the existence of heterogeneous variability. Its model is:

$$Y_t = \mu + \beta(t - T/2) + \alpha Y_{t-1} + \mu_t \quad (5)$$

1 test $\alpha = 1$ statistic:

$$Z(t_\alpha) = (s/\sigma_n)t_\alpha - \lambda'\sigma_n / M^{\frac{1}{2}}$$

2 test $\mu = 0$ statistic:

$$Z(t_u) = (s/\sigma_n)t_\alpha - \lambda'\sigma_n m_y / M^{\frac{1}{2}}\left(M + m_y^2\right)^{\frac{1}{2}}$$

3 test $\beta = 0$ statistic:

$$Z(t_\beta) = (s/\sigma_n) - \lambda'\sigma_n\left(\frac{1}{2}m_y - m_{ty}\right)/(M/12)^{\frac{1}{2}} m_{yy}^{\frac{1}{2}}$$

4 At the same time test $\alpha = 1, \beta = 0$, namely test $t_\mu = 0$. Statistics:

$$Z(t_\mu) = (s/\sigma_n)t_\mu + \lambda'\sigma_n / \overline{m_{yy}}^{\frac{1}{2}} m_{yy}^{\frac{1}{2}}$$

5 Does not consider the trend, test $\alpha = 1$ statistic:

$$Z(t_\alpha) = (s/\sigma_n)t_\alpha + \lambda'\sigma_n / m_{yy}^{\frac{1}{2}}$$

Among them, the sample size T and the standard deviation s, $\lambda = \frac{1}{2}\left(\sigma_n^2 - s_n^2\right), \lambda' = \lambda/\sigma_n^2$

$$m_y = T^{-\frac{3}{2}}\sum y_t$$

$$m_{yy} = T^{-2}\sum y_t^2$$

$$m_{ty} = T^{-\frac{5}{2}}\sum t y_t$$

$$M = \left(1 - T^{-2}\right)m_{yy} - 12m_{yy}^2 + 12\left(1 + T^{-1}\right)m_{ty}m_y$$

$$-\left(4 + 6T^{-1} + 2T^{-2}\right)m_y^2$$

4)KPSS Test

Since in the ADF test, the null hypothesis is based on the assumption the existence of a single root in the sequence, there must be significant evidence to suggest that there is no single root, then null hypothesis can be refused. The test results tend to think that sequence has a single root; Sims (1988) pointed out that when a smooth sequence is very close to a single process, ADF test power will be lower; Kwiatkowski, Phillips, Schmidt and Shin (1992) supposed a method, which the opposite of zero hypothesis testing methods (referred to KPSS single-root test), the null hypothesis is based on the assumption do not have a single root in the sequence (of the stationary series).

KPSS test formulas have two models:

$$z_t = r_t + \varepsilon_t;\ r_t = r_{t-1} + \mu_t$$

$$z_t = \theta t + \gamma_t + \varepsilon_t;\ r_t = r_{t-1} + \mu_t$$

r_t is a random item, $u_t \sim iid\left(0, \sigma_u^2\right)$, from above get the errors $e_t, t = 1, 2, 3\cdots T$, set $s_t = \sum_{t=1}^{r} e_t, \hat{\sigma}^2_\varepsilon$ the estimation value of the error variable, the LM statistic is as follows: $LM = \sum_{t=1}^{r} s_t^2 / \hat{\sigma}^2_\varepsilon$.

The null hypothesis of KPSS test: H_0 sequence is a smooth sequence. When refuse null hypothesis, it shows that the detection sequence is a smooth sequence.

B. EG Two-Stage Method of Co-integration

Engle and Granger (1987) supposed two-stage method of co-integration to test whether the two variables are co-integration or not. This method of co-integration test is performing unit root testing the residual of regression equation. On the ideological point of co-integration theory, independent variables and dependent variable have co-integration relationship.

The first step, using OLS estimation to estimate co-integration regression equation

. $Y_t = b_0 + b_1 X_t + \mu_t$ (6)

Get residual sequence e_t, as estimated value of balance error μ_t.

The second step, test the stability of e_t. If e_t is a smooth, then X_t and Y_t are co-integration, and on the other hand, X_t and Y_t are not co-integration. This is because if X_t and Y_t are not co-integration, then any linear combination is non-stationary, so residuals e_t will be non-smooth. In other words, test whether residual sequence e_t is stable, that is, test the existence of co-integration test.

There are some shortcomings of the method in practice, as we only can confirm the existence of co-integration relationship, but we do not know the exact number of co-integration vectors.

4. GRANGER CAUSALITY TEST

After doing the unit root test on the non-stationary time series data and examining the co-integration relationships and co-integration vector, which can be clear and accurate understanding of several non-stationary time series data and the long-run equilibrium relationship, and then perform the Granger causality test on the time-series data, to determine whether there are leading and lag relationship in the time series, which is great significance for actual economic and financial analysis and operation. The Granger causality concept was first supposed in 1969 by Granger. Sims (1972) also made a similar definition. Granger causality is based on whether prediction error can be reduced to determine standards, that when the variables X past data will help forecast variables Y, namely, that the statistics Y is for the result of X, the definition of causal relationship have a more definite conclusion. Following that Granger causality test pattern:

$$X_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-1} + \sum_{i=1}^{p} b_i Y_{t-1} + \mu_{1t} \tag{7}$$

$$Y_t = \beta_0 + \sum_{i=1}^{p} c_i X_{t-1} + \sum_{i=1}^{p} d_i Y_{t-1} + \mu_{2t} \tag{8}$$

The use of statistical F test, respectively test the above two type of null hypothesis:

$H_0 : b_1 = b_2 = \cdots = b_p = 0$

$H_0^{'} : c_1 = c_2 = \cdots = c_p = 0$

Test results, if can not be refused the zero assumption H_0 and $H_0^{'}$ at the same time, said that causal relationship does not exist in X_t and Y_t. If refuse zero $H_0^{'}$ but do not refuse assume H_0, means that X_t leads Y_t; If refused null hypothesis H_0 but not rejected $H_0^{'}$, said Y_t leads X_t; And if refusing two a zero assumption at the same time , means that a feedback causal relationship exists in X_t and $\tilde{Y_t}$

5. VARIANCE DECOMPOSITON

If said that the impulse response function described the effects from the information shock of a new endogenous variable on other variables of VAR model, then variance decomposition will decompose the impact of other variables shocks. We can characterize relative importance (contribution) of each variables impactation. Composed by k variables, VAR model (1) s -step prediction error was:

$$\begin{aligned} &Var\left[Y_{t+s} - E\left(Y_{t+s} / Y_t, Y_{t-1}, Y_{t-2} \cdots \right) \right] \\ &= \xi_{t+s} + \Psi_1 \xi_{t+s-1} + \Psi_2 \xi_{t+s-2} + \cdots + \Psi_{s-1} \xi_{t+1} \end{aligned} \tag{9}$$

Its mean-square error was:

$$\begin{aligned} &Var\left[Y_{t+s} - E\left(Y_{t+s} / Y_t, Y_{t-1}, Y_{t-2} \cdots \right) \right] \\ &= \xi_{t+s} + \Psi_1 \xi_{t+s-1} + \Psi_2 \xi_{t+s-2} + \cdots + \Psi_{s-1} \xi_{t+1} \\ &= PP^{'} + \Psi_1 PP^{'} \Psi_1^{'} + \cdots + \Psi_{S-1} PP^{'} \Psi_{S-1}^{'} \\ &= \sum_{j=1}^{k} \left(P_j P_j^{'} + \Psi_1 P_j P_j^{'} \Psi_1^{'} + \cdots + \Psi_{S-1} P_j P_j^{'} \Psi_{S-1}^{'} \right) \end{aligned} \tag{10}$$

Which P_j is the j column matrix vector, the expression in brackets shows that the impact of j orthogonal (or new) contribution to s step prediction. Any endogenous variables forecast MSE could composited into the impact of the random variables contributions, and then calculated the relative importance of impact of each of variables.

6. GEWEKE INDICATORS

Geweke (1982) supposed a direct measure of information feedback and interdependence among the securities markets. The method first assumes that in two countries stock markets return rate depends on another market rate of return over past, self market rate of return over past and noise. In this model, the two equations would be experienced seemingly unrelated Regression. Secondly, assumption same period and one-way simultaneity information feedback do not exist in two markets, that is parameters restrictions for first established equation and the use least squares to estimate the constraint equations. Finally, use estimated the variance-covariance residual of constraint equations and non-binding equation to structure a likelihood ratio. Statistics test the assumption, and determine information feedback and dependence relationship between the two markets.

7. CONCLUSION

There are four indicators when examining the linkage between the A+H securities markets: the relevant parameters, co-integration relationships, Grange causality, and Geweke indicators. Related parameters are intuitive and simple. Co-integration can test long-run equilibrium relationships existing in a number of non-stationary time series. In its basis, we can establish the vector error correction model, in which we can also observe the short-term market movements, but cannot observe the details of long-run equilibrium relationship and the changes in characteristics and time-varying: The Granger causality test can directly determine the leading or lagging relationship in two markets, but the test prerequisite is a smooth time series, and non-stationary time series must be differenced before Grange causality test, but the information contained in variable differential suffered losses, and Granger causality test can only accept or reject the null hypothesis. LR statistic called Geweke-based indicators measure the A+H stock markets, which provides a measure of degree of direct linkage.

Research has also not yet found a linkage between A+H stock markets perfect approaches, although less than the above method, but can complement each other. In short, the test theory of development is only used in practice. With expanding research, in the future there will also be more powerful theoretical tools for empirical analysis.

REFERENCES

B Alessandra and A. F. Carlo, (2005). “Explaining co-movements between stock markets: The case of US and Germany,” *Journal of International Money and Finance,* vol. 24, Nov. 2005 pp. 1299-1316.

C. Kao and M. H. Chiang. (2000). “On the Estimation and Inference of a Co-integrated Regression in Panel Data,” *Advances in Econometrics*, vol. 15, Feb. 2000, pp. 179-222.

C. Kao. (2001). “Spurious Regressions and Residual Based Tests for Co-integration in Panel Data,” *Journal of Econometrics*, vol. 90, Jan. 2001, pp. 1-44.

G. Huang and F M. Song. (2005). “The Financial and Operating Performance of China’s Newly Listed H-firms,” *Pacific Basin Financial Journal*, vol. 13, Dec. 2005, pp. 53-83.

G. Meric and P. L. Ricardo, (2001). “Co-movements of US and Latin American equity markets before and after the 1987 crash,” *International Review of Financial Analysis*, vol. 78, Oct. 2001 pp. 219-235.

H. Q Zhang and M Yang, (2007). “Investment Opportunities in A - H Stock Market,” *Economic Information*, vol. 68, Jul. 2007, pp. 53-78.

P. B. Phillips and B. E. Hansen, (1990). “Statistical Inference in Individual Variables Regression with I(1) Process,” *Review of Economic Studies*, vol. 57, Jun. 1990, pp. 99-125.

P. Pedroni (1999). “Critical Values for Co-integration Tests in Heterogeneous Panels with Multiple Regressors,” *Oxford Bulletin of Economics and Statistics*, vol. 61, Mar. 1999, pp. 653-678.

S. Johansen, (1991). “Estimation and Hypothesis Testing of Co-integration Vectors in Gaussian Vector Autoregressive Models,” *Econometrics*, vol. 59, Jun. 1991, pp. 1551-1580.

S. Johansen, (1988). “Statistical Analysis of Co-integrating Vectors,” *Journal of Economic Dynamics and Control*, vol. 94, Dec. 1988, pp. 231-254.

S. M. Strauss, (2000). “Exchange traded funds-the wave of the future?” *Investment Lawyer*, vol. 7, Apr. 2000, pp. 15-21.

S. McCoskey and C. Kao, (1998). “A Residual Based Test of the Null of Co-integration in Panel Data,”*Econometric Reviews*, vol. 17, May. 1998, pp. 57-84.

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Performance of CHROM Agar and Oxacillin Resistant Screening Agar Base Media for Detection of Methicillin Resistant *Staphylococcus aureus* (MRSA) from Chronic Wound

E.S. Karthy (Corresponding author)

Research Department of Microbiology

Sengunthar Arts and Science College

Tiruchengode 637 205, Tamilnadu, INDIA

E-mail: karthy.es@gmail.com

P. Ranjitha & A. Mohankumar

Research Department of Microbiology

Sengunthar Arts and Science College

Tiruchengode 637 205, Tamilnadu, INDIA

Abstract

CHROM agar *Staphylococcus aureus* and Oxacillin Resistant Screening Agar Base (ORSAB) media with oxacillin were evaluated for the screening of Methicillin Resistant *Saphylococcus aureus* (MRSA). Among 190 samples, totally 126 confirmed *Staphylococcus aureus* strains were used for screening of MRSA used CHROM agar and ORSAB media were compared with the other MRSA screening media like Baird Park agar (BPA) with ciprofloxacin, Mannitol salt agar (MSA) with oxacillin, Blood agar (BA) with oxacillin and Muller Hinton agar (MHA) with oxacillin. Totally 54 MRSA strains were confirmed using PCR, among that 83% and 92% of the MRSA stains were isolated as mauve colonies on CHROM agar and blue colonies on ORSAB medium in 24 hrs incubation, compare with 64%, 61%, 50%, and 42% of the strains that were isolated on BPA, MSA, BA and MHA respectively. After 48 hrs of incubation 100%, 98%, 77%, 77%, 72% and 68% of the MRSA strains were isolated on CHROM agar, ORSAB, BPA, MSA, BA and MHA respectively. CHROM agar *Staphylococcus aureus* and ORSAB agar proved to be more sensitivity and specificity than other MRSA selective media. These provide an alternative for the detection of MRSA in clinical laboratories, especially when PCR is unavailable.

Keywords: MRSA, Specificity, Sensitivity, PCR, *mec*A gene

1. Introduction

In 1961 the first case of Methicillin Resistant *Staphylococcus aureus* (MRSA) was documented (Barber., 1961). Over the last three decades Methicillin Resistant *Staphylococcus aureus* (MRSA) has caused major problems in hospitals throughout the world (Waldvogel., 1995). Detection of MRSA in clinical samples continues to be important, since infection due to MRSA have high morbidity and mortality rates (Wertheim *et al.,* 2001). Approximately 75% of hospital strains are resistant to methicillin (York *et al*., 1996). Mechanisms of methicillin resistant in *S. aureus* is based on the production of an additional low affinity penicillin binding protein (PBP; PBP2a), which is encoded by the *mec*A gene (Chambers., 1997). Accurate and quick identification of Methicillin Resistant *Staphylococcus aureus* (MRSA) in clinical specimen is essential for timely decision on isolation procedures and effective antimicrobial chemotherapy (Jonas *et al*., 1999).

The isolation, recovery and identification methods used in any routine clinical laboratory are crucial for the detection of MRSA. PCR based methods have recently been developed for the direct detection of MRSA in clinical specimen. However the use of these assays is largely restricted to reference centers and they are not currently utilized by most routine diagnostic laboratories. In this study, we evaluated the performance of CHROM agar and ORSAB as compared to that other conventional media with proper supplement for recovery of MRSA strains in clinical specimens, also studied specificity and sensitivity of the selected media compared with PCR analysis detection of *mec*A gene.

2. Experiments

2.1 Isolation of Clinical Samples

Totally 190 samples were obtained for the infected wound included accident, bite, septic and burn wounds. No duplicate isolates from a single patient were included in this study. Isolates were plated onto Mannitol Salt Agar (MSA) and incubate at 37ºC, after 48 hrs of incubation results were observed.

2.2 Identification and sensitivity test

Method of identification included Gram staining, morphology, catalase, mannitol salt fermentation, slide and tube coagulase. The antibiotic sensitivity profile of the 126 *S. aureus* isolates were determined according to the method of Bauer-Kirby (Bauer *et al*., 1966) using discs of antibiotics places on the surface of MHA medium seeded with the test organism. Inhibition zones were measured after 48 hrs of incubation at 37ºC. Interpretation of resistance was based on the National Committee for Clinical Laboratory Standards (NCCLS) criteria. The antibiotics used were ciprofloxacin, methicillin, oxacillin and vancomycin.

2.3 Screening of MRSA

All confirmed *Staphylococcus aureus* isolates were inoculated onto a CHROM agar *Staphylococcus arueus* media with Oxacillin 4mg/l, Oxacillin Resistant Screening Agar Base (ORSAB) supplement with 50 000 IU of polymyxin B and 2.0 mg/l of oxacillin and 5.5% sodium chloride, Baird Park Agar (BPA) with 8mg/l ciprofloxacin and 7% sodium chloride, Mannitol Salt Agar (MSA) with 2mg/l oxacillin, Blood agar with 2mg/l oxacillin, Muller Hinton Agar (MHA) with 6mg/l oxacillin. Confirmed *Staphylococcus aureus* strains were streaked on the plates and incubated at 37ºC. All the culture plates were interpreted after 24 hrs and 48 hrs of incubation. If any growth with expected color was detected, the isolates were considered as MRSA.

2.4 Detection of mecA gene

The PCR procedure was based on a modification by Unal *et al*., (1992) and this were used as the gold standard for all isolates. Oligonuclotide used were *mec*A F primer 1282 (5'–AAA–ATC–GAT–GGT–AAA–GGT–TGG–C–3') and *mec*A R primer 1793 (5'– AGT-TCT-GCA-GTA-CCG-GAT-TTG-C-3'), which gives a PCR products of 533bp. PCR was performed on cooled thermocycler 5333, Eppendorf version 2.30.33-09, using a reaction mixture of 20 µl consisting of *Taq* polymerase buffer 2µl, 1 µl of each primer, DNA sample 1 µl, *Taq* polymerase enzyme 0.2 µl and distilled water 12.8 µl. 20µl of PCR product was then analyzed by 1.2% agarose gel electrophoresis.

3. Results and Discussion

3.1 Isolation and Sensitivity test

Totally 190 clinical specimens were directly tested on MSA. Out of these specimens 126 *S. aureus* were isolated after 48 hrs of incubation *S. aureus* colonies appeared yellow on MSA plate and later confirmed by tube and slide coagulase test. Four antibiotics discs were used for antimicrobial susceptibility test against *S. aureus*. It showed 49%, 44%, 39% and 50% of isolates resistant to ciprofloxacin, methicillin, oxacillin and vancomycin respectively. All *S. aureus* strains were resistant to any one of the above antibiotics.

3.2 Screening of MRSA

All *S. aureus* isolates were cultured on CHROM agar *S. aureus* with Ox, ORSAB with Ox, Baird Parker agar with ciprofloxacin, Mannitol Salt agar with 4-Ox, Blood agar with Ox, and Muller Hinton agar media with Ox and incubated for 48 hrs at 37°C. The plates were examined at 24 and 48 hrs, which were mauve color on CHROM agar, blue on ORSAB, yellow on MSA-4 Ox, black on Baird Park agar, white on Blood agar and Ox-MH agar. 83.30% of the MRSA strains were recovered by the use of CHROM agar after 24 hrs of incubation and this rate increased to 100% after the plate had been incubated for 48 hrs. Of 126 *S. aureus* isolates tested on 6 different MRSA selective media, 45 and 54 isolates were grown on CHROM agar after 24 and 48 hrs respectively. Similarly ORSAB, BPA, MSA Ox, BA, Ox-MHA were successfully in identifying MRSA in 24 and 48 hrs (Table 1).

3.3 Sensitivity and Specificity of MRSA

Multiple antibiotic resistant *S. aureus* strains constitute a major health care problem, therefore the availability of sensitive and specific method for the accurate detection of antibiotic resistance in these bacteria has become an important tool in clinical diagnosis. CHROM agar and ORSAB media showed a substantially better performance than any of the other media tested, and its sensitivity after 24 hrs of incubation were superior to that of other BPA, MSA-4Ox, BA and Ox-MHA. ORSAB was slightly less sensitive than CHROM agar after both 24 and 48 hrs of incubation. BPA, MSA-Ox, BA, Ox-MHA media were less sensitive than ORSAB and CHROM agar in 24 hrs 64.80%, 61.10%, 50% and 42.60%, in 48 hrs 77.80%, 77.80%, 72.20% and 68.50% respectively (Table I). Six media were showed high specificity (>90%) at 24 hrs incubation, but it was reduced after 48 hrs incubation except CHROM agar. Simor *et al*., (2001) reported that the incubation time of 24 hrs should be extended to 48 hrs for the detection of MRSA.

CHROM agar MRSA achieved 100% in both specificity and sensitivity. MSA and ORSAB both achieved 91.50% and 91.50% respectively (Taguchi *et al*., 2004). Apfalter *et al*., (2002) also examined the performance of ORSAB at 24 and 48 hrs and found that the sensitivity of ORSAB increased from 50.80% to 68.20%. Blanc *et al*., (2003) found that 38% of blue MRSA colonies on ORSAB were visible only after 48 hrs of incubation. Out of 114 *S. aureus* isolates tested on CHROM agar *S. aureus*, all the isolates were grew and were identified chromogenically as *S. aureus* by a pink to mauve color change after 18 to 24 hrs of incubation and also DNase and MSA were successful in identifying *S. aureus* in 112 of 114 isolates (98%) (Merlino *et al*., 2002).

The similar results were observed by Lally *et al*., (1985) that the use of MSA-4Ox resulted in a small number of false positive and false negative results. Brown and Walpole (2001) reported that a positive results were markedly lower for test on colonies from Blood agar and Baird Parker agar than from those on MSA and also reported blood agar plates with and without oxacillin were incubated at 30°C and 37°C respectively and the lower temperature of incubation of the plates with oxacillin may have contributed to better expression at PBP2a. Baird Parker agar contains ciprofloxacin was a useful medium for isolation of MRSA in situations where ciprofloxacin resistant strains were endemic (Davies and Zadik, 1997).

MDR-MRSA isolates were grew on CHROM agar *S. aureus* and ORSAB when supplemented with oxacillin 4mg/l, 2mg/l respectively in 48 hrs as well as on the isolation of NMDR-MRSA made difficult phenotypic interpretation. These isolates were usually resistant to β-lactam antibiotic but often susceptible to agents such as gentamicin, tetracycline, trimethoprim and variably erythromycin and ciprofloxacin (Maguire *et al.,* 1998).

3.4 Correlation between antibiotics groups and PCR

The correlation between major three antibiotics with Methicillin resistant (M, Ox), (M, Cf), (M, V) and the presence and absence of the *mec*A gene. In MRSA, there were 34 strains of Ox^r /*mec*A^+ , 40 of Cf^r /*mec* A^+ and 39 of V^r /*mec*A^+ and Ox, Cf, V sensitive and the presence of *mec*A gene were 14, 8 and 9 isolates respectively. In MSSA, these were none of the strain of Ox^r /*mec*A^- , 2 of Cf^r /*mec*A^- and 5 of V^r /*mec*A^- and Ox, Cf, V sensitive and the absence of *mec*A gene were 8, 6 and 3 isolates respectively. In MRSA, the correlation between three antibiotics with Methicillin sensitive and the presence of *mec*A gene isolates were observed very low percentage (<3.2%). In MSSA class methicillin sensitive and absence of *mec*A gene were 12 strains of Ox^r, 16 of Cf^r and 17 of V^r and sensitivity of all three antibiotic 52, 48 and 47 isolates respectively (Table II).

3.5 Detection of mecA gene by PCR

Currently, multiple antibiotic resistant *S. aureus* strains constitute a major healthcare problem, since they are the etiologic agent of several nosocomical and skin infection. For that reason, accurate detection of resistant isolates constitutes a critical goal of clinical microbiology. The phenotypic expression of pre-confirmed 126 clinical wounds *S. aureus* isolates were examined in this study. These isolates were tested for the phenotypic confirmation for MRSA and hence were all genetically confirmed to be MRSA using PCR. Among 126 *S. aureus*, 54 isolates were positive for the *mec*A gene by PCR, remaining 74 *S. aureus* were negative (Table I).

Schmitz *et al.,* (1997) and Tokue *et al.,* (1992) previously reported that the utility of PCR for the accurate detection of the *mec*A gene and the possibility of simultaneous identification of *S. aureus* and detection of *mec*A gene. During the last decade, many studies have demonstrated the extremely high capacity of PCR for specifically detecting bacteria and genes of interest (Salisbury *et al.,* 1996). That ability has revealed PCR as a powerful tool in clinical microbiology studies (Cockerill, 1999). Several authors have already shown the feasibility of the PCR methodology for the identification of *S. aureus* strains and for the detection of antibiotic resistance genes (Cockerill, 1999; Jonas *et al.,* 1999). We concluded that the procedure still require a minimum of 2 days before final results can be obtained, but these media proved to the more sensitive and specific than other media. It may be an alternative for the detection of MRSA in clinical laboratories, especially when the PCR is unavailable.

References

Apfalter, P., Assadian, O., Kalczyk, A., Lindenmann, V., Makristathis, A., Mustafa, S., Rotter, M., & Hirschl, A. M. (2002). Performance of a new chromogenic oxacillin resistance screen medium (oxoid) in the detection and presumptive identification of methicillin-resistant *Staphylococcus aureus*. Diagn. Microbiol. Infect. Dis. 44: 209-211.

Barber, M. (1961). Methicillin resistant staphylococci. J. Clin. Pathol. 14: 385-393.

Bauer, A. W., Kirby, W. M. M., Sherris, J. C., & Turck, M. (1966). Antibiotic susceptibility testing by standardized single disk method. Am. J. Clin. Pathol. 45: 493-496.

Blanc, D. S., Wenger, A., & Bille, J. (2003). Evaluation of a novel medium for screening specimens from hospitalized patients to detect methicillin resistant *Staphylococcus aureus.* J. Clin. Microbiol. 41 (8): 3499-3502.

Brown, D. F. J., & Walpole, E. (2001). Evaluation of the Mastalex latex agglutination test for methicillin resistance in *Staphylococcus aureus* grown on different screening media. J. Antimicrob. Chemother. 47: 187-189.

Chambers, H. F. (1997). Methicillin resistance in Staphylococci: molecular and biochemical basis and clinical implications. Clin. Microbiol. Rev. 10: 718-791.

Cockerill, F. R. **(**1999). Genetic methods for assessing antimicrobial resistance. Antimicrob. Agent Chemother. 43: 199-212.

Davies, S., & Zadik, P. M. (1997). Comparison of methods for the isolation of methicillin-resistant *Staphylococcus aureus*. J. Clin. Pathol. 50: 257-8.

Jonas, D., Grundmann, H., Hartung, D., Daschner, F. D., & Towner, J. K. (1999). Evaluation of the *mec*A, *fem*B duplex polymerase chain reaction for detection of methicillin-resistant *Staphylococcus aureus*. Eur. J. Clin. Microbiol. Infect. Dis. 18: 643-647.

Lally, R. T., Ederer, M. T., & Woolfrey, B. F. (1985). Evaluation of Mannitol Salt Agar with oxacillin as a screening medium for methicillin - resistant *Staphylococcus aureus*. J. Clin. Microbiol. 22(4): 501-504.

Maguire, G. P., Arthur, A. D., Boustead, P. J., Dwyer, B., & Currie, B. J. (1998). Clinical experience and outcomes of community-acquired and nosocomial methicillin resistant *Staphylococcus aureus* in a Northern Australian hospital. J. hosp. Infect. 38: 273-281.

Merlino, J., Watson, J., Rose, B., Beard, M. P., Gottlieb, T., & Harbour, C. (2002). Detection and expression of Methicillin/ Oxacillin resistance in multidrug resistant and non-multidrug resistant *Staphylococcus aureus* in central Sydney, Australia. J. Anti. Chem. 49: 793-801.

Salisbury, S. M., Sabatini, L. M., & Spiegel, C. A. (1996). Identification of methicillin resistant staphylococci by multiplex polymerase chain reaction assay. Microbiol. Infect. Dis. 107: 368-373.

Schmitz, F. J., Mackenzie, C. R., Hofmann, B., Verhoef, J., Finken-Eigen, M., Heinz, H. P., & Kohrer, K. (1997). Specific information concerning taxonomy, pathogenecity and methicillin resistance of Staphylococci obtained by multiplex PCR. J. Med. Microbiol. 46: 773-778.

Simor, A. E., Goodfellow, J., Louie, L., & Louie, M.(2001). Evaluation of a new medium. Oxacillin resistance screening agar base, for the detection of methicillin-resistant *Staphylococcus aureus* from clinical specimens. J. Clin. Microbiol. 39: 3422 (Letter).

Taguchi, H., Kaneko, T., Onozaki, M., Kubo, R., & Kamiya S. (2004). Evaluation of a new chromogenic medium for isolation of MRSA. J. Japanese Association for Infectious Diseases. 54-58.

Tokue, Y., Shoji, S., Satoh, K., Watanabe, A., & Motomiya, M. (1992). Comparison of polymerase chain reaction assay and conventional microbiological method for detection of methicillin-resistant *Staphylococcus aureus*. Antimicrob. Agents Chemother. 36: 6-9.

Unal, S., Hoskins, J., Flokowitsch, J, E., Wu, C. Y. E., Perston, D. A., & Skatrud P. L. (1992). Detection of methicillin-resistant Staphylococci by using the Polymerase Chain Reaction. J. Clin. Microbiol. 30: 1685-1691.

Waldvogel, F.A. (1995). *Staphylococcus aureus* (including toxic shock syndrome), p. 1754-1777. In G.L. Mandell, J.E. Bennett, and R. Dolin (ed.), Mandell, Douglas and Bennett's principles and practice of infectious diseases, 4th ed. Churchill Livingstone, New York, N.Y.

Wertheim, H., Verbrugh, H, A., Pelt, C., Man, P., Belkum, A., & Margreet C. V. (2001). Improved detection of methicillin-resistant *Staphylococcus aureus* using phenyl mannitol broth containing Aztreonam and Ceftizoxime. J. Clin. Microbiol. 39(7): 2660-2662.

York, M.K., Gibbs, L., Chehab, F & Brooks, G. F. (1996). Comparison of PCR detection of *mec*A with standard susceptibility testing methods to determine methicillin resistance in coagulase negative staphylococci. J. Clin. Microbial. 34. 249-253.

Table 1. COMPARISON OF THE LEVELS AND ACCURACY OF REACTIONS OF MRSA AND MSSA ISOLATES ON DIFFERENT MEDIA AFTER 24 AND 48 HRS

S. No	Media	24 hrs						48hrs					
		Positive Test Results				Accuracy of media		Positive Test Results				Accuracy of media	
		MRSA (mec+) n=54 Isolates	%	MSSA (mec-) n=72 Isolates	%	Sensitivity (%)	Specificity (%)	MRSA (mec+) n=54 Isolates	%	MSSA (mec-) n=72 Isolates	%	Sensitivity (%)	Specificity (%)
1.	CHROM Agar[a]	45	83.3	0	0	83.30	100.00	54	100.0	0	0	100.00	100.00
2.	ORSAB[b]	42	77.8	1	1.4	77.80	98.60	53	98.1	5	6.9	98.10	93.10
3.	BPAgar[c]	35	64.8	0	0	64.80	100.00	42	77.8	3	4.2	77.80	95.80
4.	MSA-4 Ox[d]	33	61.1	2	2.7	61.10	97.30	42	77.8	8	11.1	77.80	88.90
5.	BAgar[e]	27	50.0	4	5.5	50.00	94.50	39	72.2	13	18.1	72.20	81.90
6.	Ox- MH Agar[f]	23	42.6	5	7	42.60	93.00	37	68.5	16	22.2	68.50	77.80

n - Number of strains

[a] - CHROM agar *Staphylococcus aureus* with oxacillin 4mg/l.

[b] - Oxacillin Resistant Screening Agar Base (ORSAB) with Oxacillin 2mg/l.

[c] - Baird Parker Agar (BPA) with Ciprofloxacin 8mg/l.

[d] - Mannitol Salt Agar (MSA-4 Ox) with Oxacillin 4mg/l.

[e] - Blood Agar (BA) with Oxacillin 2mg/l.

[f] - Mueller Hinton Agar (Ox- MH Agar) with 4% NaCl, 6mg/l Oxacillin.

Table 2. CORRELATION BETWEEN ANTIBIOTIC GROUPS AND PCR RESULTS

S. No	Bacterial Class	Antibiotics M	Ox	No. of Isolates	*mecA* gene	%	Antibiotics M	Cf	No. of Isolates	*mecA* gene	%	Antibiotics M	Va	No. of Isolates	*mecA* gene	%
1	MRSA	R	R	34	+	27.00	R	R	40	+	31.70	R	R	39	+	31.00
2	MRSA	R	S	14	+	11.00	R	S	8	+	6.35	R	S	9	+	7.10
3	MRSA	S	R	4	+	3.20	S	R	4	+	3.20	S	R	2	+	1.60
4	MRSA	S	S	2	+	1.60	S	S	2	+	1.60	S	S	4	+	3.20
5	MSSA	R	R	0	–	0.00	R	R	2	–	1.60	R	R	5	–	4.00
6	MSSA	R	S	8	–	6.30	R	S	6	–	4.80	R	S	3	–	2.40
7	MSSA	S	R	12	–	9.50	S	R	16	–	12.70	S	R	17	–	13.50
8	MSSA	S	S	52	–	41.30	S	S	48	–	38.10	S	S	47	–	37.30

MRSA : Methicillin Resistant *Staphylococcus aureus*

MSSA : Methicillin Sensitive *Staphylococcus aureus*

+: Positive, - : Negative,

R: Resistant, S: Sensitive, %: Percentage

M: Methicillin, Ox: oxacillin, Va: Vancomycin, Cf: Ciprofloxacin

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Study on the Application of CIPS in the Pulp and Papermaking Engineering

Zhongjun Xiao

College of Electric and Information Engineering, Shaanxi University of Science & Technology, Xi'an 710021, China

Tel: 86-029-8616-8631 E-mail: xiaodongjun@sust.edu.cn

Xiuqin Wang

China Liaoyang Paper Machinery Co. Ltd, Liaoyang 111004, China

Tel: 86-0419-252-0750 E-mail: lyzj@lyzj.com

Abstract

For the modern pulp and papermaking industry with high technical requirement, huge energy consumption and large pollution output in, the computer integrated process system (CIPS) of the flow industry which could effectively integrate corporate management concept, information technology and automatic control technology is necessary. Based on the analysis of the characteristics of various workshop sections of the pulp and papermaking industry and its demands of CIPS, we put forward the CIPS three-layer system structure (ERP, MES and PCS) which could be applied in the pulp and papermaking industry, and a new CIPS optimized application project based on SIMATIC PCS7 for the papermaking corporation.

Keywords: Pulp and papermaking, Computer integrated process system (CIPS), Process control system, SIMATIC PCS7

1. Introduction

In 1973, American Dr. Joseph Harrington put forward a sort of enterprise production organization theory, CIM (Computer Integrated Manufacturing). CIMS (Computer Integrated Manufacturing System) is the system based on CIM theory, i.e. it is the large-scale complex system which comprehensively utilizes modern management technology, manufacturing technology, information technology, automatization technology and system engineering technology to integrate and optimize correlative personnel, technology and management in the whole production process of the enterprise and the information flow and material flow organically through computer hardware and software. In 1986, European Community also developed the concept of CIMS to CIPS in time, and formed the concept of CIPS in the flow industry, and advocated implementing CIPS plan in the successive production processes such as petrifaction and energy sources.

In this article, we mainly analyzed the characteristics of the pulp and papermaking industry and its demands for CIPS, put forward the CIPS three-layer system structure which could be applied in the pulp and papermaking industry, and a new CIPS optimized application project based on SIMATIC PCS7 for the papermaking corporation.

2. Characteristics of the pulp and papermaking industry and the demands of CIPS

The pulp and papermaking industry is an important industry which is closely linked to the development of national economy, and it related to many industries such as the forest industry, agriculture, mechanical manufacturing, chemical industry, electric automatization, traffic transportation and environmental protection. The whole pulp and papermaking process flow can be described as that the wheat straw, cotton stalks, bulrush and other papermaking materials are prepared through the dry and wet process, and are digested in the digestion section mixing with digestion liquors (black liquor and white liquor) under high temperature, and produce coarse pulps through the digestion reaction, and through washing and filtering, the black liquors, residues and coarse fibers could be removed from the coarse pulps which could be turned into accepted pulps, and the accepted pulps are sent to the whitening section and turned into paper pulps through CEH three-section continual whitening, and the pulps from the washing, filtering and whitening workshop mix with other pulps according to certain proportion, and mix with fillings, gluing materials, coloring matters and other accretions, the mixed pulps could be formed, and the mixed pulps are put into the pulp pool, and diluted as the paper pulp with regulated concentration by clear water and white water, and mixed with fillings such as alums, and filtered through the intermediate stuff box, desander, steady flow box and stock discharge box, and the finished rolls could be

formed through procedures such as pressing, drying, gluing, press polishing, taking up and distorting (Wang, 2003).

Thus it can be see that the pulp and papermaking industry is the industry with intensive technologies, large resource consumption and many pollution outputs. To adapt the drastic market competition and realize the modernization of the pulp and papermaking industry, when we enhance its economic benefits, we should try to reduce pollution and energy consumption and protect the environment in the production process, so we need effectively integrating modern management technology, manufacturing technology, automatization technology and information technology, i.e. under the guidance of flow industry CIPS, realizing CIPS in the pulp and papermaking industry aiming at the production characteristics of the pulp and papermaking corporation (Li, 2003, P.35-37).

3. CIPS three-layer system structure of the papermaking corporation

The CIPS of the papermaking corporation is divided by three layers, i.e. ERP (Enterprise Resource Planning), MES (Manufacturing Execution System) and PCS (Process Control System) (Zhao, 1999, P.137-139 & Hu, 2005, P.35-38) (seen in Figure 1).

The layer of ERP is the technology of the total resource optimization which emphasizes the planning character of the corporation, pays attention to the production plan, dispatching, cost and logistics, and takes the financial analysis and decision-making as the core. ERP could effectively realize the material stock, repertory management, code management, order arrangement, production management, quality management, sales management, transportation management, client complaint management, cost accounting and financial analysis of the papermaking industry, and support the special management of the distribution in the papermaking industry. The keystone of ERP of the papermaking corporation is the logistic management, and the raw materials, finished product quality and quality traceability should fulfill very high management requirements, and the system should follow from the finished goods to semi finished articles and to material suppliers, so the ERP system must effectively help the corporation to follow the quality responsibility, and offer strong quality management and following function, and allow the corporation to follow the quality from the finished goods to raw materials. In a word, for the papermaking corporation, to judge whether one ERP system is effective, we should judge whether it produce significant effects for reducing cost, enhancing quality and increasing benefits through the information-based reconstruction.

The layer of MES is the technology about the production process optimization control, production execution optimization operation and optimization management which emphasizes the execution of plan, pays attention to the equipment management and takes the production comprehensive index as the objective. Concretely speaking, the layer of MES mainly takes the realization of the three-optimization of the production comprehensive index, i.e. optimized operation, optimized control and optimized management, as the objective, and it possesses the backbone function which could link the preceding and the following, and it could realize many functions such as production dispatching, material following, resource distribution and management in the production, quality management, flow simulation, system data acquirement in the production process, model computation and process optimization, and it is also the basic data processing platform. Many papermaking corporations have their bottom production control system, and many of them implement the ERP system, but the communication is still lacked between the management layer and the control layer, and MES could just fill the blank to realize the integration of management and control. MES is a set of system which could transmit and produce correlative information in time and optimize a series of production activities from receiving order to completing finished-product. MES takes the software as the base, takes the real-time database and historical database in the network environment as the supports, adopts the client/server system structure and the objective-oriented design method to realize the information management of the whole lifecycle of product and achieve the function of information hub between the upper planning layer and the lower control layer.

The layer of PCS is the technology which emphasizes the equipment monitoring (if the field-bus technology is adopted, the information of equipment fault diagnosis will be more abundant), takes the quality and quantity of the product (such as the finished rolls) and the technical requirement fulfillment as the objective, and takes the comprehensive management control of equipment as the core. The layer of PCS could automatically control the technical process of the production with high efficiency, high precision and high level to ensure the quality, cost and delivery term of the products and make the products possess strong competitive power in the market. It is mainly used to control the pulp and papermaking production process including process test, soft measurement, meter control, electric control and execution institution.

By the comprehensive automatic system structure composed by three associative systems such as PCS, MES and ERP, the papermaking corporation could realize the comprehensive integration of three systems under the drives of control technology, computer technology and optimization technology, really complete the comprehensive application of information, and realize the integration of management and control.

4. A new CIPS application in the papermaking corporation based on PCS7

To fully integrate ERP, MES and PCS, Siemens pushed its new standard automatic control system, i.e. SIMATC PCS7,

and SIMATC PCS7 is the comprehensive automatic solution based on process control. Traditional process automatization includes the electric control system PLC and the instrument control system DCS, and through several years' development and technical advancement, Siemens developed the concept of whole integrated TIA which integrated PLC and DOS, combined advanced electric manufacturing technology, figure and image processing technology, field-bus technology, computer technology and advanced automatization control theory, and it was the advanced process control system oriented to all process control applications. The application of SIMATIC PSC7 in the papermaking corporation could realize the whole integrated automatization, and the core of the system integration includes integrated allocation and program configuration, integrated data management and integrated communication. It could offer complete automatic solution for the papermaking industry, and reduce stop time, spare parts and interfaces of the paper machine to the largest extent, and offer intelligent service for the factory.

Except for controlling the production line, the SIMATIC PSC7 in the papermaking corporation could monitor and control the running of the assistant factories. In the system, operators could control and monitor all processes and obtain the important information of the whole corporation in time by the consistent user interface. The standard software modules about MES, process control layer and quality management could ensure the coherence of the data flow between the process and the production layout. From planning and running production to reducing costs, this system could offer powerful supports and guarantees for the production corporations, strengthening the system performance largely and ensure achieving the requirements of the environmental protection. Figure 2 is the CIPS application system structure of the papermaking corporation.

In ERP, the system use special enterprise resource programming ERP software to comprehensively manage the supply, sale, repertory, finance, personnel and equipments, and make comprehensive plan and management decisions, and the major control objective of ERP is the capital flow of the corporation.

According to the middle and long planning requirements from the market and the corporation, MES comprehensively harmonizes and balances the production systems of the whole corporation, allocates the resources, deals with the accidents to obtain largest comprehensive benefits. The exertion layer mainly considers the logistic balance, abilities (reaction ability, repertory ability and transportation ability) balance and environmental protection requirement, and the major control objective of this layer is the logistics of the corporation.

PCS mainly includes the optimized control of technology, fault diagnosis and system monitoring. We adopt SIMATIC PCS7 as the concrete implement tool to realize the whole integrated automatization control. The integrated allocations includes S7-400 series, distributed I/O, ET200M and Profibus, and the integrated program configuration includes that the PCS7 of STEP7+WinCC in the host computers of OS and ES.

The hardware and software systems are the base to compose the automatic control system. Thousands of I/O signals are processed by the several severs in the dispersedly designed automatic system, and the concrete implementation approaches can be described as follows.

(1) ES. In the engineer station (ES), according to the pulp and papermaking technical flow and the process control requirements of various loops, engineer uses PCS7 to compile CFC/SFC programs and make the WinCC human-computer interface. CFC is used in the configuration controlled by the continual process, and PCS7 offers a great lot of standard process industrial function modules such as PID control, valve control and motor control. SFC is used in sequence control, especially for the configurations of the control system which are exerted in turn.

(2) OS. Operator station (OS) connects with the server by the high speed industrial Ethernet, and operator monitors and control the operation by the WinCC of the host computer. WinCC is the multi-functional operation system based on Windows NT, and it can quickly and highly effectively respond the process events and alarm, and ensure the reliability of data effectively. The human-computer interface software of WinCC is the base to control and monitor the operation for various operation stations. It could make PC turn into the monitoring system with high efficiencies. This integrated data management system could optimize the control loop and solve problems in the debugging process.

(3) AS. Process control station (AS) controls the locale equipments such as transducer and control valve by the Profibus interface, and implements the net connection by the optical fiber or cable according to their corresponding distance. Data are transmitted by the 100Mbps Ethernet which is very fit for the papermaking corporation needing huge data exchanges, and these high speed data could largely enhance the performance of the network.

The corporation process analysis system could acquire data by the industrial standard interface OPC. When WinCC server is failed, OPC communication mode will be switched automatically to avoid data loss. The ISDN connection with the automatization system could implement remote diagnosis and remote configuration to ensure the reliable production of the corporation.

5. Conclusions

At present, many problems such as numerous corporate quantity, small scale, deficient raw materials, large resource and

energy consumption, lower production efficiency and serious pollution still exist in Chinese pulp and papermaking industry, and the key reason is that the technologies (including the industrial technology and the equipment automatization technology) are lagged. To enhance the competitive power of Chinese pulp and papermaking industry in the world, when we introduce, digest and absorb foreign advanced papermaking technology, we should adopt the high and new technologies with measured strokes, and adjust and optimize the structure to implement clean production, save natural resources and enhance the production efficiency and economic benefit. Therefore, based on the theories about foreign and domestic CIPS, it is very meaningful to exert Chinese advantages in the domain of CIMS and implement CIPS in major corporations in Chinese pulp and papermaking industry.

References

Hu, Nanjiang & Wang, Mengxiao. (2005). Research on CIPS of Papermaking Industry. *China Pulp & Paper Industry.* No.26(12). P.35-38.

Li, Jigeng, Liu, Huanbin et al. (2003). Contemporary Integrated Processing System of Modern Paper Industry. *China Pulp & Paper.* No.22(4). P.35-37.

Wang, Mengxiao, Sun, Yu & Tang, Wei. (2003). *Pulp and Paper Process Monitor and Control System and Engineering.* Beijing: Chemical Industry Press.

Zhao, Xinyu, Chai, Tianyou & Houbin. (1999). Research of the Architecture of CIPS. *Journal of Northeastern University.* No.20(2). P.137-139.

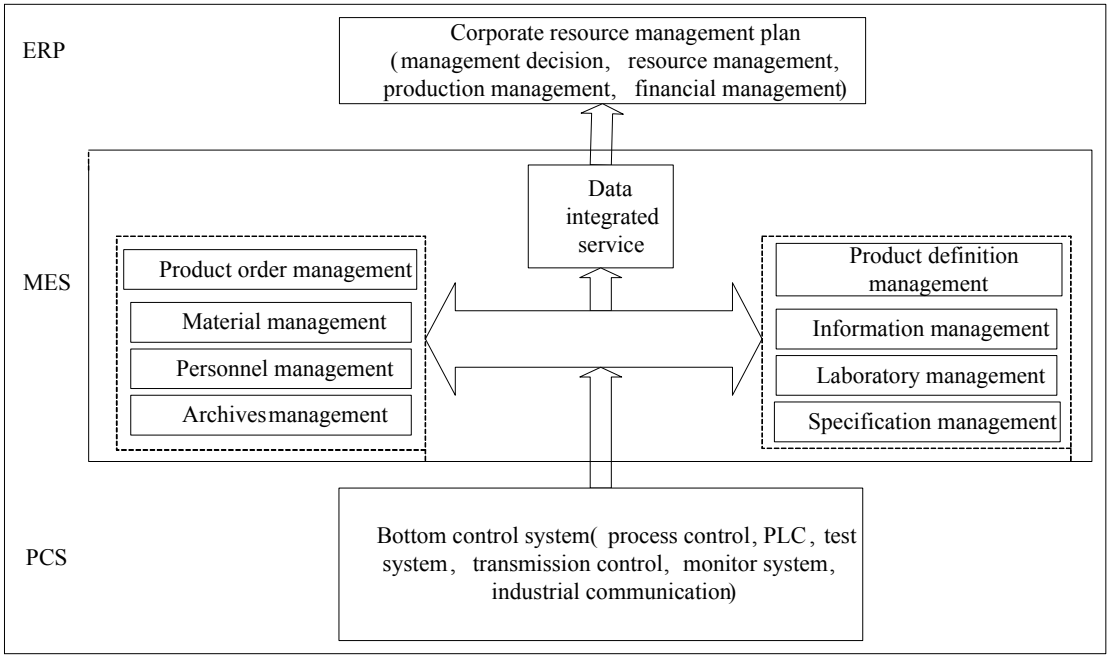


Figure 1. CIPS Three-layer Structure System of Papermaking Corporation

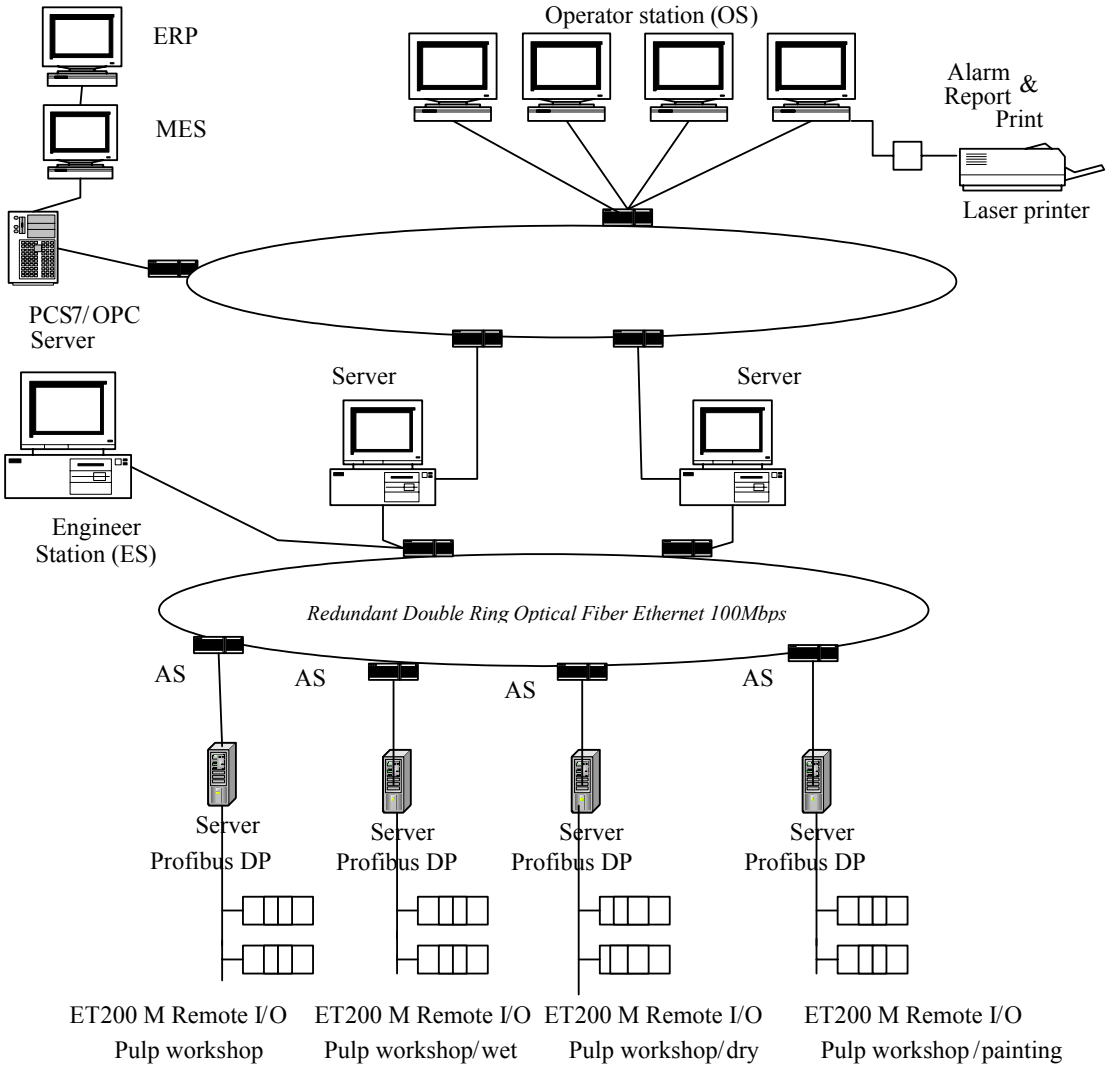


Figure 2. CIPS Application Project of Papermaking Corporation

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

A Baseline Study on Groundwater Quality of the Tourist Island, Pulau Tiga, Sabah, Malaysia

Chin Yik Lin (Corresponding Author)
School of Science and Technology, Universiti Malaysia Sabah
Beg Berkunci 2073, 88999, Kota Kinabalu, Sabah, Malaysia
Tel: 60-88-325-808 E-mail: cy_lin_ars@hotmail.com

Mohd Harun Abdullah
School of Science and Technology, Universiti Malaysia Sabah
Beg Berkunci 2073, 88999, Kota Kinabalu, Sabah, Malaysia
Tel: 60-88-325-808 E-mail: harunabd@ums.edu.my

Ahmad Zaharin Aris
Faculty of Environmental Studies, Universiti Putra Malaysia
43400 UPM, Serdang, Selangor, Malaysia
Tel: 60-88-325-808 E-mail: zaharin@env.upm.edu.my

Sarva Mangala Praveena
School of Science and Technology, Universiti Malaysia Sabah
Beg Berkunci 2073, 88999, Kota Kinabalu, Sabah, Malaysia
Tel: 60-88-325-808 E-mail: smpraveena@gmail.com

This research was financially supported by the Ministry of Higher Eduction, Malaysia through the Fundamental Research Grant No. FRG0050-ST-1/2006. Permission from the Sabah Park Trustees for the study site is highly acknowledged. Besides, the authors would also opt to acknowledge Chua Li Ying, Lim Kuan Leang and Ooi Chun Wei for their assistance during the fieldworks.

Abstract

Pulau Tiga is a group of small islands located at Kimanis Bay off the western coast of Sabah. This preliminary study on groundwater of Pulau Tiga is undertaken to provide guidance and baseline data for future references. An understanding of the chemistry and behavior of water in the island aquifer is crucial for the determination of the availability of freshwater, and also to create a platform for considering proper management and successful remediation of groundwater resources on island to prevent over exploitation of such limited resources. Groundwater samples were collected from five wells on Pulau Tiga to understand the groundwater chemistry based on various ion composition (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, HCO_3^{-}, SO_4^{2-}, Cl^{-}, NO_3^{-}), *in situ* parameters such as DO (Dissolved oxygen), EC (Electrical conductivity), TDS (Total dissolved solids), pH, salinity and temperature. Even though groundwater in Pulau Tiga was constantly subject to abstraction activities for its tourism purposes, preliminary results showed that the groundwater still remains freshwater type. The Piper diagram suggested that the groundwater facies generally range from Ca-HCO_3 to Ca-Cl waters type. The processes influencing the groundwater chemistry are mainly dissolution of minerals.

Keywords: Groundwater, Pulau Tiga, Hydrochemical, Hydrochemical Facies, Groundwater Chemistry

1. Introduction

Pulau Tiga is a group of small uninhabited Malaysian islands in Kimanis Bay off the western coast of Sabah. Pulau Tiga is located between 5º 42′ and 5º 43′ and 115º 37′ and 115º 39’, South West from Kota Kinabalu with a distance of 55 km and 12 km North East from Klias peninsular (Sanudin *et al*., 1999). Which covers approximately 725 hectares with length of approximately 5.2 km and width of 1.8 km. Pulau Tiga is the biggest island among those three islands. Land use is typically residential or undeveloped. Pulau Kalampunian Besar and Pulau Kalampunian Kecil are located at 1 km North from Pulau Tiga, respectively. At present, approximately thousands of people visit Pulau Tiga every year. Groundwater is the only source for regular water supply in the island. With the growing number of tourists and human activities on the island, the demand for groundwater in this island has increased apparently. Since its existence has become the most important source for the small island’s community, pumping from the upper phreatic zone is widely practiced on the island which may lead to deterioration of groundwater quality. Furthermore, there is lack of studies done in this island to understand the change of groundwater chemistry in such condition. Thus, this preliminary research was set up to study the groundwater quality on the island.

1.1 Site Description

Pulau Tiga has unique hydrometeorological characteristics with wide variation in the rainfall pattern (average 2816.4 mm). The mean temperature is approximately 23 °C – 32 °C. Most of its water comes from precipitation. The island received heavy precipitation during the southwest monsoon (April – October) and the southeast monsoon (October – February) prevail. This high volume of rainfall is mainly depends with transition of monsoon seasons. However, some rains also occur during the month of March. The annual average rainfall distribution from 1995 to 2007 is shown in Figure 2. While, the monthly average rainfall for year 2007 recorded at Kuala Penyu is shown in Figure 3, respectively. The recharge of aquifer comes from direct infiltration of rainwater and lateral inflow from the hills. On the contrary, the discharge of the aquifer is a result of pumping activities, outflow to the sea (subterranean or direct), and drainage to the swampy land. During the wet season there is one tidal channel that allows the mixing of seawater into the brackish mangrove swamp water on this island. The geology of Pulau Tiga is unique in view of the presence of mud volcano activity. The most active mud volcano activity is located at the highest point of Puncak Mat Salleh. Geologically, the study area is part of Quaternary sedimentary terrain. Based on the coastal structure analysis, Pulau Tiga is located at the junction of two widely spread formation in this region, namely the Crocker and Belait formations. The island consists of two distinct lithological units. According to Sanudin *et al*., (1999) the uppermost formation is Belait underlain by Quaternary thick mud volcanic sequences. The presence of Setap Shale formation on the island is well-known for its actively petroleum production. Quaternary sedimentary formation of Pulau Tiga unconformably overlay Neogen. Pulau Tiga is constituted by material of sedimentary origin that forms a chaotic accumulation of fluvial and volcanic materials located at the vent of mud volcano. Therefore, the island is mostly covered with mud volcanic soil. This mud volcanic soil is covered with mud and shale. Although there are a great variety of rock types within the study area, the vast majority is made up of Crocker, Setap Shale and Belait rock fragments. These rock fragments being ejected out together with mud volcano during the mud volcano eruption events. In general, these kinds of bedrocks are known as clastic sedimentary aquifer. Clastic sedimentary rocks consist of mudstone (e.g. shale), sandstone and conglomerate. Mudstone and sandstone are by far the most common of these, and sandstone is the most common aquifer type because it is normally more permeable than mudstone. There are trace amount of carbonates rock found in Pulau Tiga. Few outcrops of carbonate rocks can be found within the stratigraphy of the study area. Presence of carbonate terraces are sighted within the cliff of Pulau Tiga. However, the presence of these rocks is discontinuous. Therefore, it is believe that the carbonate rocks are just single block fragments. Microscopy study indicates there are various kinds of foram’s exoskeleton, algae and coral cemented by carbonate or calcite as matrix in carbonate rocks. Besides, there is presence of aragonite as trace minerals. Few kinds of micro fossils such as foraminifer plankton and corals had been found embedded in the carbonate terraces. The foraminifer species from the carbonate terraces found in the coastal plain of Pulau Tiga aged from mid Miocene (Sanudin *et al*., 1999). In this paper, the authors have opted to examine the groundwater quality, major dissolved cation and anion constituents in the groundwater of Pulau Tiga and as well as to classify the types of groundwater chemistry in the island aquifer.

2. Material and Methods

A total of 30 samples of groundwater in the study area were collected from 5 wells located on Pulau Tiga, respectively. The first sampling was conducted in August 2007 (dry period), and the second in November 2007 (wet period). The sampling wells are located at the low elevation area of the island. Currently, there are only 3 production wells (PT 1, PT 2 and PT 4) that are still constantly provide freshwater supply to the inhabitants on that island. The other two wells (PT 3 and PT 5) were temporary abandoned due to high salinity. Polyethylene bottles that soaked with 10 % HNO_3 acid wash and pre-rinsed with distilled water were used to store groundwater samples based on the methods described in APHA (1995). Groundwater samples were collected by using Wildco Water Sampler. Triplicate samples were used to verify the sampling protocol (Allen *et al*., 2001). The precise locations of the sampling points were determined in the

field through the development of the GARMIN 12 Channel Instrument, based on the principles of Global Positioning System (GPS). The exact longitude and latitude of the sampling points, hydrogeological data and well's features were summarised in Table 1. Through this study, graphical analyses were used to display the areal distribution of dissolved constituents throughout the Pulau Tiga aquifer. Statistical techniques were applied to analyse the groundwater quality data. Besides, the statistical analyses also provide us a clear view of the general chemical character of the groundwater of Pulau Tiga. Statistical analysis was undertaken by using correlation matrix. This analysis was undertaken by using the computer software SPSS 15 in order to verify and describe the relationships between the various determinants.

2.1 Physico-chemical parameters

In situ parameters such as temperature, electrical conductivity (EC), total dissolved solids (TDS), pH, dissolved oxygen (DO), salinity and redox potential (mV) were measured using probe Eutech EcoScan, Oakton Cond., EC-TDS Scan, YSI 650 and Eutech Salttester. Groundwater were analysed immediately following sample collection in order to acquire representatives values of ambient aquifer conditions. The samples were stored at 1-4 ºC temperature prior to analysis in the laboratory. Water samples collected were filtered through a 0.45 µm membrane filter paper (Millipore) using glass filtration unit and acidified with concentrated nitric acid (HNO_3) to lower the pH of samples < 2 in order to prevent the adsorption occurred upon the container surface and precipitation of cations. Cation and anion such as chloride (Cl^-), bicarbonate (HCO_3^-), calcium (Ca^{2+}), magnesium (Mg^{2+}), sodium (Na^+) and potassium (K^+) were analysed in the Water Quality Laboratory of the School of Science and Technology, Universiti Malaysia Sabah, Malaysia. The major cations (Na^+, K^+, Mg^{2+} and Ca^{2+}) were determined using atomic absorption spectrophotometer (AAS Perkin Elmer – 4100). Sulfate (SO_4^{2-}) and nitrate (NO_3^-) were detected using HACH (DR/2040) meter. Cl^- and HCO_3^- were analysed using argentometric ($AgNO_3$) and titration methods (HCl 0.1 N) as suggested by APHA (1995), respectively.

3. Results and Discussion

3.1 In-situ parameters

In-situ results for both sampling periods (August 2007 and November 2007) from triplicate samples were averaged into one number for inclusion in Table 2. The major cation (Ca, Mg, Na, K), major anion (Cl, HCO_3^-, SO_4^{2-}, NO_3^-) of the groundwater and WHO (2004) guidelines for drinking water are shown in Table 4. The pH values of groundwater were between 6.84 to 7.33 and a median value of 7.02 for both sampling periods (August 2007 and November 2007). A look on the values (Table 2) reveals that all the samples had pH value within the permissible range of 6.5 – 8.5 (WHO, 2004), no distinct groupings of these values were observed. The groundwater temperature is generally between 26.7 °C to 29.9 °C and with an average value at 28.3 °C, which correspond with mean annual temperature.

Groundwater in Pulau Tiga has low mineralisation, as shown by the electrical conductivity (EC) measurements. EC of these groundwater sources varies from 861 µScm^{-1} to 1005 µScm^{-1} (August 2007) and 329 µScm^{-1} to 828 µScm^{-1} (November 2007). The EC decrease from 1005 µScm-1 to 828 µScm-1 is probably due to the dilution of groundwater by heavy precipitation (Mondal *et al*, 2007). Table 2 indicates that all the groundwater samples are below the WHO (2004) guideline of 1400 µS/cm.

Total dissolve solids (TDS) varies from 164 mgl^{-1} to 502 mgl^{-1} in August 2007 and November 2007, respectively. The lowest value was recorded at station PT 5 (165 mg/L) in November 2007. The result indicates that the groundwater belongs to fresh water type and presently not under the influence of seawater intrusion (Mondal *et al*., 2007). Since there is absence of river and stream in Pulau Tiga, therefore, the groundwater recharge in Pulau Tiga is depends mainly on precipitation. As a general observation, lower TDS values were observed in November 2007 while higher TDS values were observed in August 2007, indicating the possibility of dilution effect of dry and wet period.

Low dissolved oxygen (DO) concentrations were recorded ranging between 1.08 mgl^{-1} to 2.06 mgl^{-1} on August 2007. There was a slight increased in concentration of DO during dry period (November 2007) which ranging from 4.20 mgl^{-1} to 6.12 mgl^{-1}. The salinity of groundwater varies between 0.1 ppt to 0.5 ppt. Due to proximity with a tidal swamp channel, it was noted that groundwater samples in station PT 3 showed a comparatively higher salinity with other sampling locations during the sampling campaign in August. This can be explained by during dry season, which rainfall is relatively low, the saline water finds its way through tidal channels and it admixes with shallow coast aquifers. Hence, the salinity of groundwater in PT 3 becomes much higher in August.

3.2 Laboratory Chemical Analysis

3.2.1 Anion constituents

Groundwater sample for both sampling periods suggested HCO_3^- ion concentration varies between 120 mgl^{-1} to 330 mgl^{-1}. The results apparently show that HCO_3^- is the dominant anion in the groundwater. Generally, dissolution of carbonate minerals by fresh water is the main source of calcium, magnesium and bicarbonates (Allen & Matsuo, 2002). Hence, high concentration of HCO_3^- may be explain by the dissolution of carbonate minerals which presence in the aquifer. Limestone and sandstone may also contribute to the presence of dissolved HCO_3^- and Ca^{2+} (Chadha, 1999). In

freshly recharged groundwater, HCO_3^- is typically the dominant anion (Aris *et al*., 2008). On the contrary, the comparatively low HCO_3^- concentration in station PT 3 and PT 5 may be explained by the occurrence of biological breakdown of organic carbon in the soil and the aquifer (Appelo & Postma, 1993; Drever, 1997).

Sulfate contents in all sampling stations are said to be at normal level, with its concentration range between 25 mgl^{-1} and 82 mgl^{-1} for both sampling periods. Sulfate concentration in the area is generally low and therefore poses no problem to the groundwater quality. These low values are most probably due to the removal of SO_4^{2-} by the action of bacteria (Amadi *et al*., 1989). From the correlation coefficient matrix for August 2007, a high positive correlation was found between electrical conductivity, sulfate and potassium contents, suggesting the large contribution of these elements to the groundwater chemical load. Nevertheless, occulted by sulfate and potassium the groundwater in Pulau Tiga is highly bicarbonate as a result of the presence of carbonate compounds in the reservoirs matrix (Fernandes *et al*., 2005). Figure 4 shows the dissolved SO_4^{2-} are well correlated with Cl^- concentration for both sampling period in August and November 2007 (r = 0.776, and r = 0.717, respectively).

In this study, chloride (Cl^-) concentrations obtained were in the range of 98.97 mgl^{-1} to 287.84 mgl^{-1} with an average value of 160.10 mgl^{-1}. Chemical analysis results indicate that 20 % of the samples (n = 30) had concentrations of chloride above 250 mgl^{-1} that what is recommended by WHO (2004) as the maximum permissible limit for drinking water. PT 3 recorded a Cl^- concentration of 266.95 mgl^{-1} in August and 241.42 mgl^{-1} in November 2007, respectively, which shows the highest chloride concentration among all sampling stations. High chloride concentration may be due to the proximity of the well to a tidal channel and the poor muddy sediments present in the aquifer system which further infers saline intrusion. According to Cidu *et al*., (2006), Cl^- may also derive from the interaction of water with marine-derived sediment. All these factors might be responsible for the relatively high Cl^- concentration (> 50 mgl^{-1}) in the groundwater of the area. Analyses of all groundwater samples in Pulau Tiga shows a decrease in chloride concentration with an average from 162.06 mgl^{-1} (dry period) to 99.24 mgl^{-1} (wet period). This suggests the possibility of hydrochemical processes such as seawater intrusion and that dilution with fresh groundwater has been taken place simultaneously in various places on this island in August to November 2007. Due to the nature of the island's aquifer which surrounded by marine environment, the possible source of chloride in groundwater in Pulau Tiga may be due to precipitation of marine aerosols and intrusion of saline water into the aquifer (Aris *et al*., 2006).

Nitrate concentration reveals a relatively low value with range from 3.70 mgl^{-1} to 0.80 mgl^{-1} for both sampling periods. Nitrate concentration in the area is far below the WHO permissible limit of 50 mgl^{-1} (WHO, 2004). As a general observation, lower nitrate concentrations were observed during the sampling event in November 2007. A possible explanation for the diminution of nitrate in November 2007 may be due to the flushing of nitrate down the soil column and through the aquifer, and dilution of nitrate within the aquifer by heavy precipitation (Richerson & Cole, 1999). Natural processes that may add nitrate to the groundwater include precipitation, mineral weathering, and decay of organic matter (Wynne *et al*., 1989). Hence, the range of nitrate contents in the samples reflects on the hydrogeochemical conditions in the aquifer, which result in "natural" nitrate concentrations between 0.80 and 3.70 mgl^{-1} as suggested by Kim *et al*., (2003).

3.2.2 Cation constituents

Potassium (K^+) concentration varies with range from 1.36 mgl^{-1} to 7.53 mgl^{-1} for August sampling and between 0.90 mgl^{-1} to 3.790 mgl^{-1} for November sampling as shown in Table 3. The K^+ concentration in the sample is less than the permissible limit (200 mgl^{-1}, WHO, 2004). Potassium may derived from the interaction of groundwater with the aquifer's matrix (high-Ca-calcite) which is very common in coastal aquifer (Appelo & Postma, 2005), part of which may derived from seawater intrusion in the area (Al-Agha, 2004). The positive and significance correlation of K^+ with both Cl^- and SO_4^{2-} on August 2007 (Table 4) may be interpreted in the remnant of saline water intrusion as well (Aris *et al*., 2007).

Analyses of all groundwater samples show a low concentration of sodium (Na^+) with concentrations below 10 mgl^{-1}. The sodium concentrations in all sampling stations were found to be relatively constant. Sodium concentrations recorded during August sampling were between 7.63 mgl^{-1} to 8.12 mgl^{-1}. While second sampling on November 2007 (dry period) found that the sodium concentrations for all groundwater samples were range between 5.81 mgl^{-1} to 7.59 mgl^{-1}. Weak positive correlation between Na^+ and Cl^- on November 2007 as shown in Table 5 suggests the groundwater might be influenced by marine sources such as result of rain water, sea-salt spray and marine deposits in a lesser extent. Dakin *et al*., (1983) calculated that Na^+ and Cl^- could originate in the mudstone units on Pulau Tiga at concentrations that could support long-term release to the circulating groundwater, and be released by diffusion through sandstone and into the fractures over many thousands of years.

In the study area, dissolved magnesium (Mg^{2+}) is generally lower than calcium and sodium. Magnesium concentration in groundwater shows a narrow range from 3.31 mgl^{-1} to 3.43 mgl^{-1} in August and ranges from 1.70 mgl^{-1} to 3.08 mgl^{-1} for the re-sampling in November 2007. The groundwater's calcium concentration in Pulau Tiga is less than the WHO permissible limit (200 mgl^{-1}, WHO, 2004). From the chemical analyses, calcium (Ca^{2+}) concentrations on August were

between 42.85 mgl^{-1} to 51.87 mgl^{-1}. During second sampling, the concentration values obtained were between 21.30 mgl^{-1} to 42.20 mgl^{-1}. Dissolution of carbonate minerals by fresh water is the main source of calcium, magnesium and bicarbonates, and water has a low content of dissolved solids. The calcium bicarbonate and magnesium calcium bicarbonate compositions usually coupled with TDS not more than 600 mgl^{-1} and pH not exceeded 7.5 (Mokrik & Baublyte, 2005). Besides, magnesium can be released upon weathering of olivines, pyroxenes and hornblende (van Sambeek *et al*., 2000).

3.3 Hydrogeochemical facies

The principal characters of the hydrochemical facies can be illustrated by methods similar to those used in lithofacies and geochemical studies – trilinear diagrams (Piper, 1944). From the Piper plot shown in Figure 6, it can be concluded that 80 % of the five sampling sites on the island had Ca-HCO_3 water types. While the remaining 20 % had Ca-Cl water types as shown in Table 5. Based on Figure 5, apparently the groundwater is mainly fresh, as indicated by its Ca-HCO_3 water type. Large proportions of the groundwater showed Ca-HCO_3, which generally indicates a strong freshwater content in the aquifer (Aris *et al*., 2008). Conversely, the Ca-Cl facies indicates the occurrence of active seawater intrusion (Gimenez & Morell, 1997). The formation of $CaCl_2$ waters is a result of progressive salinisation of the waters. The process of reverse cation exchange may create $CaCl_2$ waters at higher salinities due to the removal of Na out of solution for bound Ca. Alternatively, $CaCl_2$ type waters could also be a result of the process of mixing between "younger" fresh water and saline older water (Adams, 2001). However, the Piper diagram elucidate that there is no significant changes in the hydrochemical facies in different seasons as shown in Figure 5. Such steady condition in the hydrochemical facies clearly reflects the homogeneity of the geological condition for the area (Leung *et al*., 2005). Table 6 summarized the general groundwater facies of each sampling periods and locations.

The groundwater components of Pulau Tiga were dominated by $Ca^{2+} > Na^{+} > Mg^{2+} > K^{+}$; and $HCO_3^{-} > Cl^{-} > SO_4^{2-} > NO_3^{-}$, except station PT 3 where the dominant ions order is followed by $Ca^{2+} > Na^{+} > K^{+} > Mg^{2+}$; and $Cl^{-} > HCO_3^{-} > SO_4^{2-} > NO_3^{-}$. Ca^{2+} and HCO_3^{-} account 50% - 70% of the ions. HCO_3^{-} accounts around 53 % of the total major ions in any given analysis depending on the well. While Cl^{-} accounts around 29 % of the total major ions. The percentage of ion distributions calculated using the following equation:

$$X\% = \frac{X\ mg/L}{\sum Ions\ mg/L} \text{ X } 100$$

where X = specific ion (Ca^{2+}, Na^{+}, Mg^{2+}, K^{+}, Cl^{-}, HCO_3^{-}, NO_3^{-} or SO_4^{2-}), and ions = total for Ca^{2+}, Na^{+}, Mg^{2+}, K^{+}, Cl^{-}, HCO_3^{-}, NO_3^{-} or SO_4^{2-} (Aris *et al*., 2007).

3.4 Statistical Analysis

ANOVA analysis shows significant change in certain chemical composition of groundwater when the groundwater is sampled at different climatic conditions, it can be observed that the concentrations of main ions generally decrease in November 2007 at rainy season. This decrease may be due to the high contribution of rainwater in November which lead to flushing effect (dilution) and input of more fresh water into the aquifer. From the ANOVA analysis, pH, temperature, EC, DO, salinity, TDS, HCO_3^{-}, NO_3^{-}, SO_4^{2-}, Ca^{2+}, Mg^{2+}, K^{+} and Na^{+} shows significant variations among dry and wet period. Conversely, both Eh values and Cl^{-} (r = 0.838, and r = 0.213, respectively) do not show significant variations between seasons.

4. Conclusions

This study shows that the groundwater quality of Pulau Tiga is generally good in quality for domestic and other purposes with respect to the parameters considered. The water quality as well as the main chemical composition of groundwater, do change significantly when water sampled at dry and wet condition, this observation apparently indicates that climate has a significant influence on the water chemistry. Nevertheless, the amount of the rainwater component can be an important factor in some geochemical processes affecting the aqueous system, and therefore, climatic conditions should be considered when planning water sampling campaigns. A high rainwater contribution to the aquifer leads to a decrease in the concentration of many chemical species due to dilution processes. In this study, there are two types of groundwater identified on the basis of by water chemistry via trilinear diagrams. That is the Ca-HCO_3 and Ca-Cl water types. The groundwater of Pulau Tiga is rather fresh, with chloride concentration varying between 50.48 mgl^{-1} and 489.85 mgl^{-1}. The main possible source of chloride ions presents in the groundwater of Pulau Tiga may be remnant of seawater, the influence of sea spray and other marine sources. In PT 3, due to the encroachment of brackish tidal channel water into the aquifer, the salinity, TDS and chloride concentration increased. As a result, the water quality deteriorates. Generally, groundwater presents in shallow aquifers of all stations were good in quality and below potable limit as set by WHO (WHO, 2004). Since this is a baseline study, it is recommended that groundwater analysis be carried out frequently to monitor the rate and kind of contaminations.

References

Adams, S., Titus, R., Pietersen, K., Tredoux, G. & Harris, C., (2001). Hydrochemical characteristic s of aquifers near Sutherland in the Western Karoo, South Africa. *Journal of Hydrology* 241, 93-103.

Al-Agha, M.R., (2004). Hydrochemical Faciens of Groundwater in the Gaza Strip. *Hydrol Sci-Journal-des-sciences Hydrologiques.* 49, pg359-372.

Allen, D.M. & Matsuo, G.P., (2002). *Results of the groundwater geochemistry Study on Hornby Island, British Columbia*. Final Report, Simon Fraser University & Island Trust, Victoria, British Columbia.

Amadi, P.A., Ofoegbu C.O. & Morrison, T., (1989). Hydrogeochemical assessment of groundwater quality in parts of the Niger Delta, Nigeria. *Environ Geol Water Sci*, 14(3): 195-202.

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

Appelo, C.A.J. & Postma, D., (1993). *Geochemistry Groundwater and Pollution.* Balkema, Rotterdam, The Netherlands, pg 536.

Appelo, C.A.J., & Postma, D., (2005). *Geochemistry, Groundwater and Pollution.* 2nd. Balkema, Rotterdam.

Aris, A.Z., Abdullah, M.H., & Wong, K.K., (2006). Hydrochemical Analysis on Seawater Intrusion of Small Carbonate Islands; Manukan and Sipadan, Sabah. *Proceedings of the 2nd SANREM Conference,* pg40-44.

Aris, A.Z., Abdullah, M.H., & Kim, K.W., (2007). Hydrogeochemistry of Groundwater in Manukan Island, Sabah. *The Malaysian Journal of Analytical Sciences,* 11(2), pg407-413.

Aris, A.Z., Abdullah, M.H., Kim, K.W. & Praveena, S.M., (2008). Hydrochemical Changes in a Small Tropical Island's Aquifer: Manukan Isalnd, Sabah, Malaysia. *Environ Geol,* DOI 10.1007/s00254-008-1275-3.

Chadha, D.K., (1999). A Proposed New Diagram for Geochemical Classification of Natural Waters and Interpretation of Chemical Data. *Hydrogeology Journal,* 7, pg431-439.

Cidu, R., Biddau, R., Manca, F. & Piras, M., (2006). Hydrogeochemical Features of the Sardinian Rivers. *Per Mineral* 76(1), pg41-57.

Dakin, R.A., Farvolden, J.A., Cherry, J.A. & Fritz, P., (1983). "Origin of dissolved solids in groundwater of Mayne Island, British Columbia, Canada, *Journal of Hydrology*, 63, 233-270, 1983.

Drever, J.I., (1997). *The geochemistry of Natural Waters*: *Surface and Groundwater Environments*. 3rd Ed. Prentice-Hall, New Jersey.

Fernandes, P.G., Bahir, M., Mendonca, Carreira, P., Fakir & Silva, M.O., (2005). Anthropogenic features in the Sines (Portugal) and Essaouira (Morocco) Coastal aquifers: A comparative study of their hydrochemical evolution by a principal component analysis. *Estudios Geol*., 61, pg 207-219.

Gimenez, E. & Morell, I., (1997). Hydrogeochemical analysis of salinization processes in the coastal aquifer of Oropesa (Castellon, Spain). *Environmental Geology*, 29, pg 118 – pg 131.

Gonzalez, L.N., Gonzalez, A., Tume, P. & Silva, A., (2001). Hydrogeochemistry and Behavior of Groundwater in the Rio Laja Basin, Bio-Bio Region, Chile. *7th International Conference on Environmenta Science and Technology, Ermoupolis, Syros Island, Greece,* pg268-274.

Hem, J.D., (1970). Study and interpretation of the chemical characteristics of natural water. *US Geol.Surv.Water Supply Pap*. 1473, pp. 363.

Kim, J.H., Yum, B.W., Kim, R.H., Koh, D.C., Cheong, T.J., Lee, J.H. & Chang, H.W., (2003). Application of Cluster Analysis for the Hydrogeochemical Factors of Saline Groundwater in Kimje, Korea. *Geosciences Journal* 7(4), pg313-322.

Leung, C.M., Jiao, J.J., Malpus, J., Chan, W.T., Wang, Y.X., (2005). Factors Affecting the Groundwater Chemistry in a Highly Urbanized Costal Area in Hong Kong: an Example from the Mid-Levels Area. *Environ Geol* 48, pg480-495.

Mokrik, R., & Baublyte, A., (2005). Water Geochemistry in the Sventoji-Arukula Aquifer System, Lithuania. *Geologija.* 52, pg 55-64.

Mondal, N.C., Singh, V.S., Saxena V.K. & Prasad, R.K., (2007). Improvement of Groundwater Quality due to Fresh Water Ingress in Potharlanka Island, Krishna Delta, India. *Environ Geol,* DOI 10.1007/s 00254-007-1010-5.

Park, S.C., Yun S.T., Chae G.T., Yoo I.S., Shin K.S., Heo C.H. & Lee S.K., (2005). Regional Hydrochemical Study on Salinization of Coastal Aquifers, Western Coastal Area of South Korea. *Journal of Hydrology* 313, pg182-194.

Piper, A.M., (1944). *A Graphical procedure in the geochemical interpretation of water analysis*: *American Geophysical Union Trans*., 25, p.914-923.

Richerson, P., & Cole, D., (2000). *April 1999 Milton-Freewater Groundwater Quality Study*. Department of Environmental Quality, State of Oregon

Sanudin, H.T., Sahibin, A.R., & Baba, M., (1999). *Geology Pulau Tiga*, Universiti Malaysia Sabah.

van Sambeek, M.H.G., Eggenkamp, H.G.M., & Vissers, M.J.M., (2000). The Groundwater Quality of Aruba, Bonaire and Cutacao: a Hydrogeochemical Study. *Geologie en Mijnbouw/ Netherlands Journal of Geosciences* 79(4), pp 459-466.

WHO (World Health Organization), (2004). Guideliness for Drinking-water quality. Volume 1 Recommendations. 3rd Ed. World Health Organization, Geneva.

Wynne, J.B., Hopkins, P., Houchins, J.O., Beinke, A., (1989). Groundwater quality of Texas – An overview of natural and man-affected conditions. *Report 89-01*. Texas Water Commision.

Table 1. The hydrogeological data, station features and GPS location of each station in Pulau Tiga

Station Code	GPS Location	Well's Features
PT 1	05º43'12.2" N 115 º 39'01.1" E	Year of Construction = 1981 Well's Diameter = 1 x 1 m Water Table = 1.07 m Screen = Cement and pebbles Distance from Shoreline = 50 m
PT 2	05 º 43'12.6" N 115 º 39'01.6" E	Year of Construction = 1985 Well's Diameter = 1.5 x 1.5 m Water Table = 1.24 m Screen = Cement and pebbles Distance from Shoreline = 57 m
PT 3	05 º 43'13.4" N 115 º 39'02.9" E	Year of Construction = 2004 Well's Diameter = 1.53 x 1.53 m Water Table = 0.93 m Screen = Cement and red brick Distance from Shoreline ≈ 95.5 m
PT 4	05 º 43'19.5" N 115 º 38'55.33" E	Year of Construction = 1999 Well's Diameter = Electrical Pump Water Table = 1.53 m Screen = Plastic ring Distance from Shoreline = 60 m
PT 5	05 º 43'23.8" N 115 º 38'49.3" E	Year of Construction = 2006 Well's Diameter = 1.3 x 1.3 m Water Table = 0.54 m Screen = Cement ring Distance from Shoreline = 45 m

Table 2. Pulau Tiga first sampling (August 2007) *in-situ* data

	Station Code	pH	Temp. (°C)	Cond. (μScm^{-1})	DO (mgl^{-1})	Sal. (ppt)	Eh (mV)	TDS (ppm)
August 2007	PT 1	6.84	28.80	956.00	2.00	0.50	-24.00	476.00
	PT 2	6.84	28.80	861.00	2.06	0.40	-26.00	430.00
	PT 3	7.27	28.60	999.00	1.84	0.50	-47.00	499.00
	PT 4	6.98	29.90	987.00	1.08	0.50	-30.00	490.00
	PT 5	7.08	29.80	1005.00	1.53	0.50	-34.00	502.00
	Mean	7.01	29.18	961.60	1.70	0.48	-32.20	479.33
	St. Dev.	0.16	0.57	54.93	0.37	0.04	8.44	27.36
	Median	6.98	28.80	987.00	1.84	0.50	-30.00	490.00
	Maximum	7.27	29.90	1005.00	2.06	0.50	-24.00	502.00
	Minimum	6.84	28.60	861.00	1.08	0.40	-47.00	429.00
November 2007	PT 1	6.93	27.70	817.00	4.20	0.40	-28.00	412.00
	PT 2	7.13	27.70	828.00	4.22	0.40	-37.00	414.00
	PT 3	6.86	27.60	505.00	4.69	0.20	-23.00	253.00
	PT 4	7.33	27.40	786.00	6.01	0.40	-46.00	393.00
	PT 5	6.94	26.90	330.00	5.69	0.10	-23.00	165.00
	Mean	7.04	27.45	653.13	4.96	0.30	-31.53	327.40
	St. Dev.	0.19	0.31	207.88	0.78	0.13	9.26	104.58
	Median	6.93	27.60	786.00	4.69	0.40	-28.00	393.00
	Maximum	7.35	27.70	828.00	6.12	0.40	-21.00	416.00
	Minimum	6.86	26.70	329.00	4.20	0.10	-46.00	164.00
WHO limits (2004)		6.5 – 8.5		1400.00				

Table 3. Average value of major anion-cation (mg l^{-1}) of groundwater at different sampling locations and period along with WHO (2004) maximum permissible level

	Sampling Station	HCO_3^- (mgl^{-1})	NO_3^- (mgl^{-1})	SO_4^{2-} (mgl^{-1})	Cl^- (mgl^{-1})	Ca^{2+} (mgl^{-1})	Mg^{2+} (mgl^{-1})	K^+ (mgl^{-1})	Na^+ (mgl^{-1})
August 2007	PT 1	284.00	3.30	31.00	147.08	49.26	3.43	1.40	7.75
	PT 2	289.00	3.00	25.00	101.17	48.48	3.36	1.75	7.92
	PT 3	205.00	1.60	79.00	266.95	43.22	3.41	7.39	7.89
	PT 4	327.00	1.50	34.00	145.43	51.48	3.33	2.55	7.92
	PT 5	322.00	2.10	32.00	149.65	51.41	3.32	2.36	8.29
	Average	285.13	2.30	40.07	160.06	48.77	3.37	7.95	3.09
	St. Dev.	45.61	0.78	20.30	50.83	3.14	0.05	0.32	2.27
	Median	285.00	2.10	32.00	142.68	49.23	3.35	7.90	2.37
	Maximum	330.00	3.70	82.00	287.84	51.87	3.43	8.89	7.53
	Minimum	203.00	1.40	25.00	98.97	42.85	3.31	7.59	1.35
November 2007	PT 1	217.00	1.60	20.30	73.15	40.93	3.06	1.00	7.47
	PT 2	272.00	1.80	19.70	68.98	41.39	3.01	0.98	7.34
	PT 3	120.00	1.80	41.00	241.42	22.26	2.44	3.51	7.29
	PT 4	287.00	1.00	20.30	57.65	41.21	2.88	1.75	7.44
	PT 5	125.00	1.20	3.00	54.98	22.37	1.84	1.47	6.08
	Average	217.67	1.48	20.87	115.80	33.63	2.65	7.12	1.74
	St. Dev	80.65	0.37	12.75	120.99	9.61	0.48	0.56	0.98
	Medium	273.00	1.50	21.00	68.98	39.86	2.90	7.36	1.49
	Maximum	293.00	1.90	42.00	489.85	42.20	3.08	7.59	3.79
	Minimum	119.00	0.80	3.00	50.48	21.30	1.70	5.81	0.90
WHO limits (2004)			50.00	400.00	250.00	200.00	100.00	200.00	

Table 4. Correlation coefficient matrix of analysed parameters for August 2007

Dry	pH	Temp	EC	DO	Salinity	mV	ORP	TDS	HCO_3^-	NO_3^-	SO_4^{2-}	Cl^-	Ca^{2+}	Mg^{2+}	K^+	Na^+
pH	1.000	0.584	0.000**	0.058	0.196	0.000**	0.723	0.000**	0.641	0.002**	0.002**	0.141	0.642	0.187	0.000**	0.268
Temp	-0.154	1.000	0.585	0.007**	0.518	0.584	0.264	0.585	0.000**	0.408	0.604	0.224	0.000**	0.002**	0.858	0.312
EC	0.797	0.153	1.000	0.004**	0.003**	0.000**	0.724	0.000**	0.483	0.032*	0.005**	0.206	0.282	0.120	0.020*	0.301
DO	-0.500	-0.667	-0.697	1.000	0.003**	0.058	0.723	0.004**	0.013*	0.001**	0.024*	0.217	0.013*	0.013*	0.021*	0.389
Salinity	0.354	0.181	0.705	-0.707	1.000	0.196	0.003**	0.003**	0.490	0.113	0.004**	0.004**	0.263	1.000	0.205	0.785
mV	-1.000	0.154	-0.797	0.500	-0.354	1.000	0.723	0.000**	0.641	0.002**	0.002**	0.141	0.642	0.187	0.000**	0.268
ORP	-0.100	-0.308	0.100	-0.100	0.707	0.100	1.000	0.724	0.285	0.985	0.094	0.001**	0.483	0.008**	1.000	0.078
TDS	0.798	0.153	0.994	-0.698	0.705	-0.798	0.100	1.000	0.517	0.030*	0.005**	0.195	0.276	0.109	0.020*	0.198
HCO_3^-	-0.131	0.937	0.196	-0.624	0.193	0.131	-0.295	0.182	1.000	0.428	0.565	0.255	0.000**	0.005**	0.703	0.642
NO_3^-	-0.735	-0.231	-0.556	0.746	-0.427	0.735	-0.005	-0.560	-0.221	1.000	0.004**	0.089	0.651	0.080	0.000**	0.628
SO_4^{2-}	0.732	-0.146	0.679	-0.579	0.696	-0.732	0.448	0.679	-0.161	-0.697	1.000	0.001**	0.950	0.894	0.002**	1.000
Cl^-	0.399	-0.334	0.347	-0.339	0.696	-0.399	0.782	0.354	-0.314	-0.454	0.776	1.000	0.270	0.195	0.048*	0.344
Ca^{2+}	-0.131	0.901	0.297	-0.622	0.309	0.131	-0.196	0.301	0.834	-0.127	-0.180	-0.304	1.000	0.012*	0.594	0.221
Mg^{2+}	-0.360	-0.739	-0.419	0.622	0.000	0.360	0.655	-0.430	-0.687	0.466	0.038	0.355	-0.631	1.000	0.216	0.019*
K^+	0.884	-0.050	0.591	-0.589	0.347	-0.884	0.000	0.593	-0.107	-0.846	0.739	0.517	-0.150	-0.340	1.000	0.524
Na^+	0.306	0.280	0.286	-0.240	-0.077	-0.306	-0.469	0.352	0.131	-0.136	0.000	-0.263	0.336	-0.595	0.179	1.000

** Represents correlation is significant at the 0.01 level (2-tailed);

* Represents correlation is significant at the 0.05 level (2-tailed)

Table 5. Correlation coefficient matrix of analysed parameters for November 2007

Wet	pH	Temp	EC	DO	Salinity	mV	ORP	TDS	HCO_3^-	NO_3^-	SO_4^{2-}	Cl^-	Ca^{2+}	Mg^{2+}	K^+	Na^+
pH	1.000	0.872	0.050*	0.305	0.003**	0.000**	0.752	0.079	0.002**	0.305	0.696	0.060	0.006**	0.163	0.413	0.099
Temp	0.045	1.000	0.000**	0.000**	0.010**	0.668	0.523	0.000**	0.431	0.004**	0.119	0.010**	0.081	0.000**	0.065	0.025*
EC	0.514	0.789	1.000	0.018*	0.000**	0.039*	0.718	0.000**	0.047*	0.054	0.506	0.370	0.001**	0.000**	0.015*	0.023*
DO	0.284	-0.855	-0.601	1.000	0.224	0.351	0.486	0.012*	0.954	0.002**	0.480	0.015*	0.449	0.004**	0.010**	0.328
Salinity	0.714	0.643	0.892	-0.334	1.000	0.001**	0.426	0.000**	0.001**	0.479	0.428	0.649	0.000**	0.000**	0.101	0.002**
mV	-0.983	-0.121	-0.536	-0.259	-0.743	1.000	0.436	0.063	0.003**	0.455	0.985	0.175	0.012*	0.143	0.548	0.085
ORP	-0.089	-0.179	-0.102	-0.195	-0.222	0.218	1.000	0.763	0.540	0.508	0.000**	0.051	0.980	0.969	0.007**	0.207
TDS	0.468	0.855	0.964	-0.632	0.889	-0.490	-0.085	1.000	0.032*	0.092	0.509	0.339	0.000**	0.000**	0.016*	0.005**
HCO_3^-	0.730	0.220	0.520	-0.016	0.773	-0.706	0.172	0.554	1.000	0.395	0.435	0.299	0.016*	0.013**	0.102	0.067
NO_3^-	-0.284	0.693	0.507	-0.733	0.198	0.209	-0.185	0.450	-0.237	1.000	0.119[a]	0.005**	0.995	0.206	0.443	0.934
SO_4^{2-}	-0.110	0.420	0.186	-0.198	0.221	-0.005	-0.792	0.185	-0.218	0.420	1.000	0.003**	0.980	0.467	0.170	0.108
Cl^-	-0.496	0.641	0.250	-0.614	0.128	0.369	-0.511	0.165	-0.288	0.681	0.717	1.000	0.554	0.229	0.751	0.423
Ca^{2+}	0.675	0.465	0.778	-0.212	0.805	-0.629	-0.007	0.817	0.609	0.002	0.007	-0.166	1.000	0.002**	0.062	0.002**
Mg^{2+}	0.380	0.841	0.887	-0.694	0.878	-0.397	-0.011	0.933	0.621	0.346	0.203	0.331	0.721	1.000	0.010*	0.001**
K^+	-0.229	-0.487	-0.615	0.643	-0.440	0.169	-0.662	-0.608	-0.438	-0.214	0.374	0.089	-0.493	-0.642	1.000	0.516
Na^+	0.442	0.575	0.582	-0.271	0.732	-0.459	-0.345	0.685	0.485	-0.024	0.432	0.223	0.721	0.746	-0.182	1.000

** Represents correlation is significant at the 0.01 level (2-tailed);

* Represents correlation is significant at the 0.05 level (2-tailed)

Table 6. Water types of Pulau Tiga

	PT 1	PT 2	PT 3	PT 4	PT 5
August 2007	Ca-HCO_3	Ca-HCO_3	Ca-Cl	Ca-HCO_3	Ca-HCO_3
November 2007	Ca-HCO_3	Ca-HCO_3	Ca-Cl	Ca-HCO_3	Ca-HCO_3

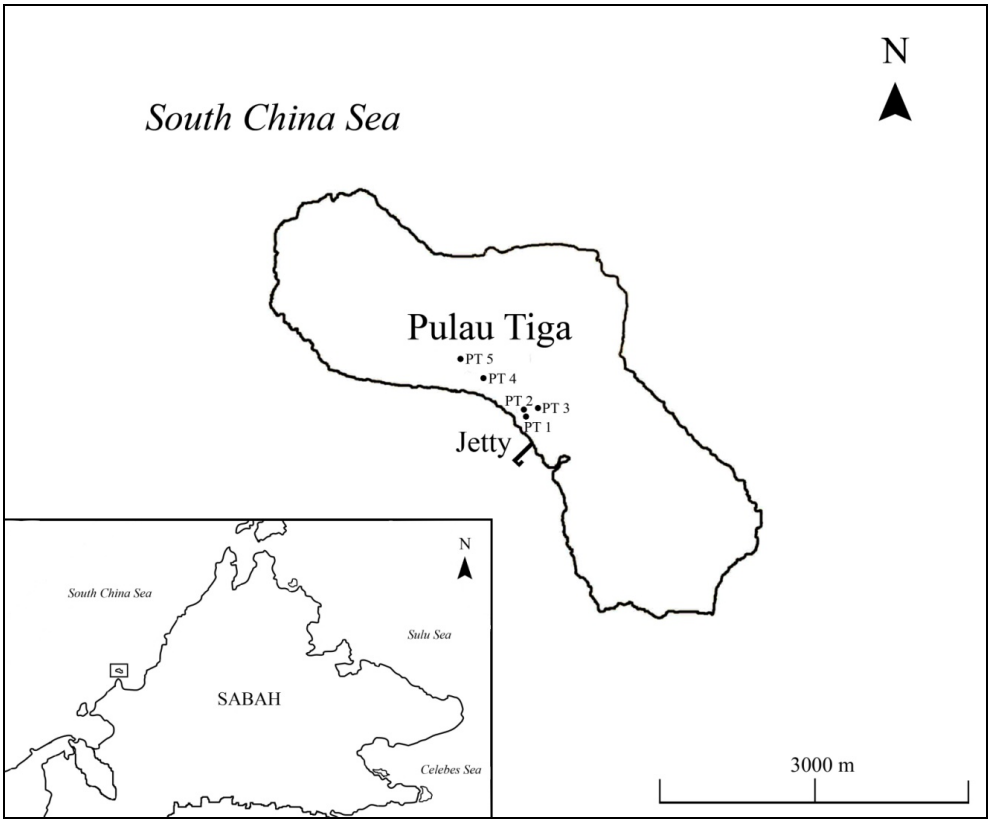


Figure 1. Schematic map showing the geographical locality of Pulau Tiga and its sampling locations

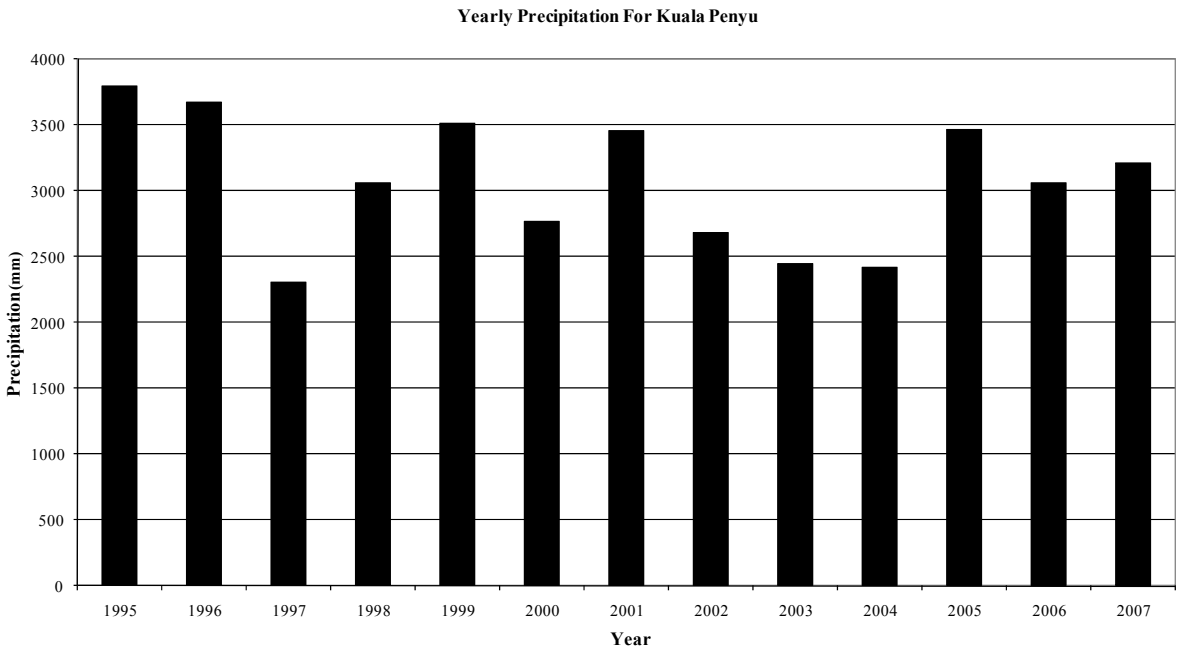


Figure 2. The average annual rainfall of Pulau Tiga for 1995 – 2007 (Sabah Meteorology Department, 2008)

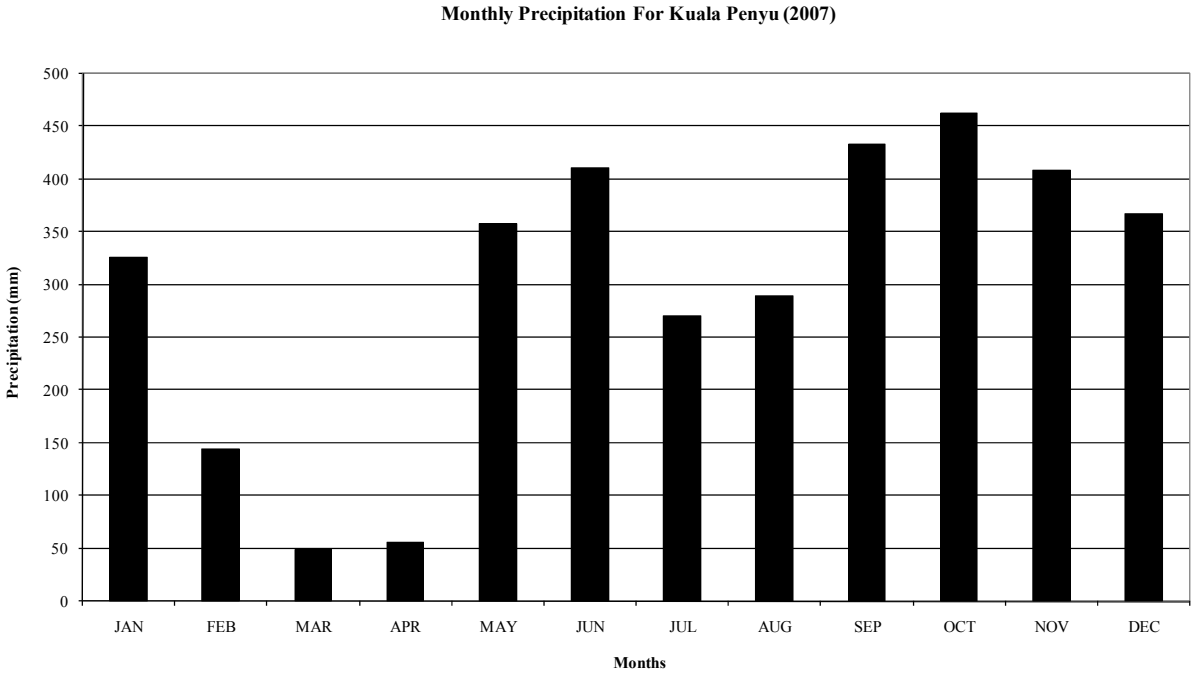


Figure 3. The average monthly rainfall on Pulau Tiga for 2007 (Sabah Meteorology Department, 2008)

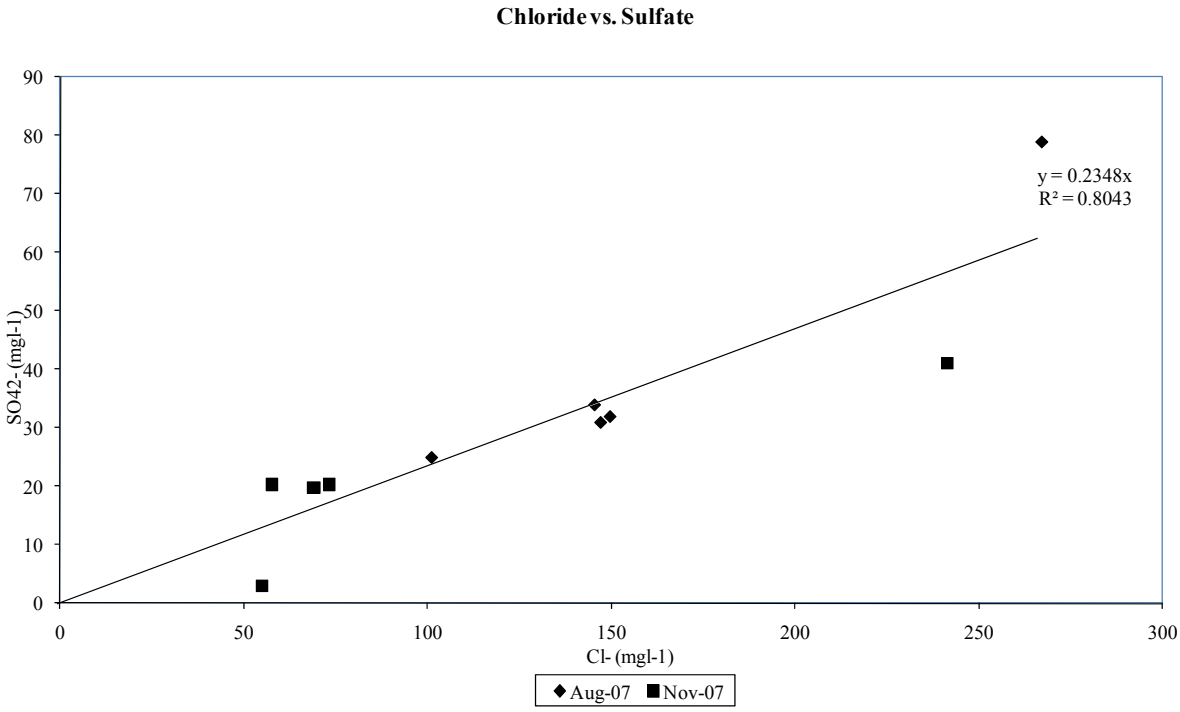


Figure 4. The dissolved SO_4^{2-} are well correlated with Cl^- concentration for both sampling period in August and November 2007 (r = 0.776, and r = 0.717, respectively)

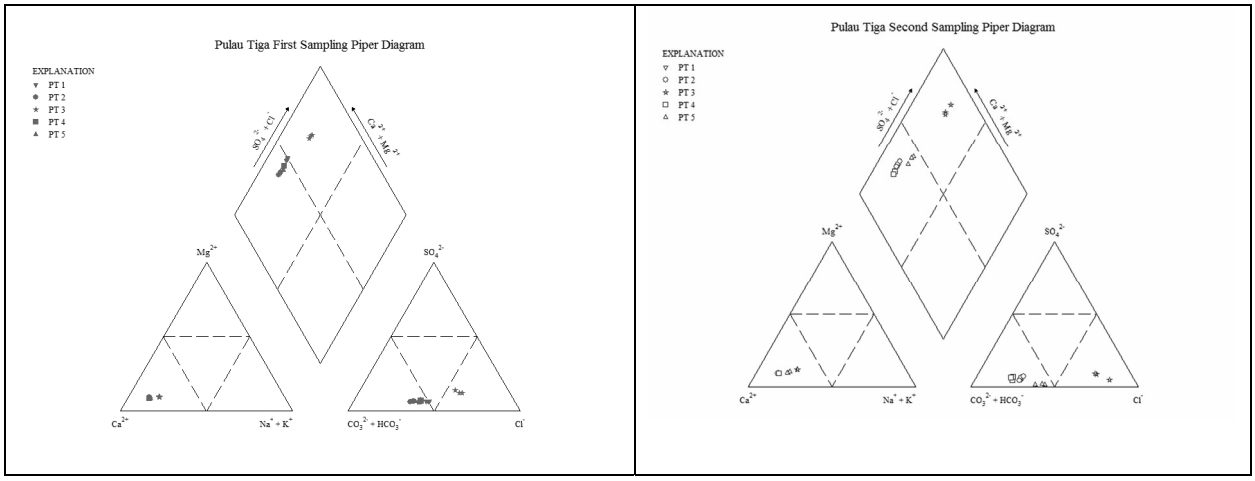


Figure 5. Piper Plot of groundwater samples collected during August 2007 (left) and November 2007 (right).

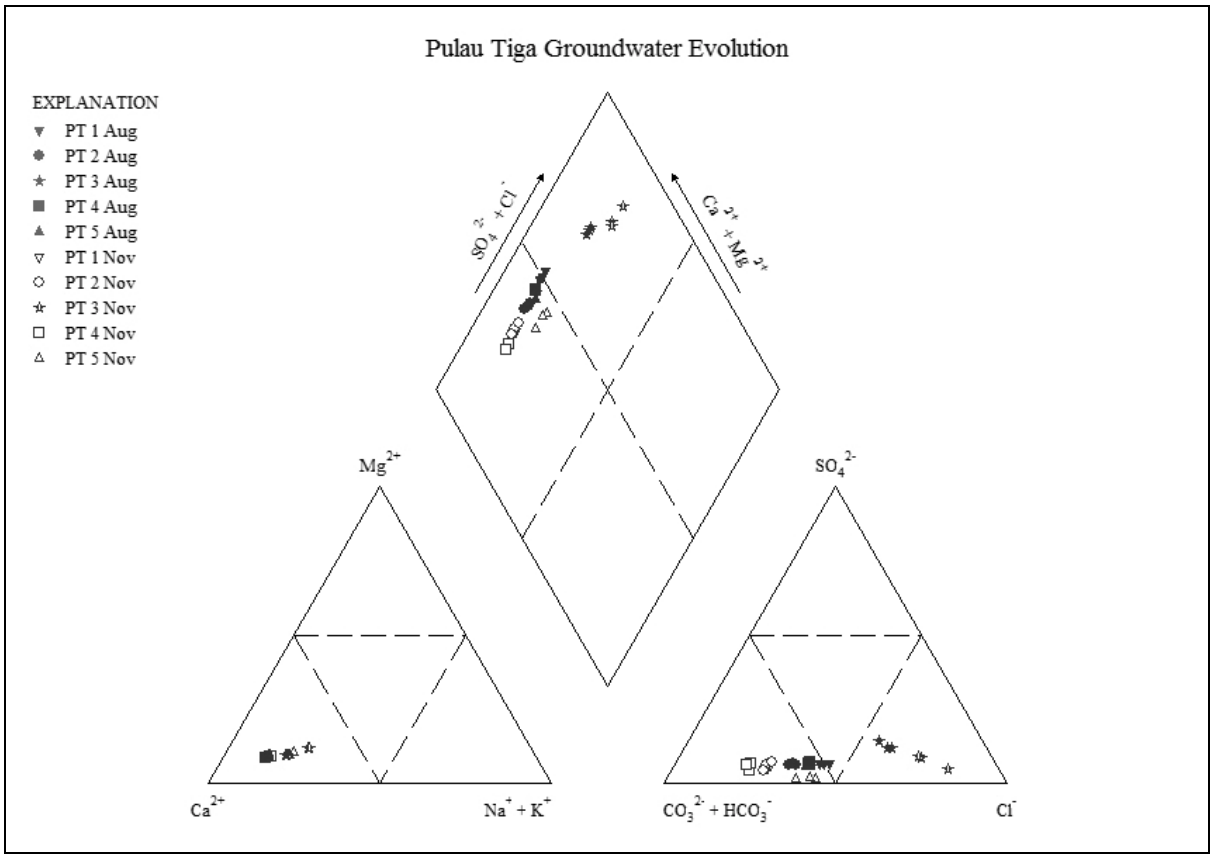


Figure 6. Piper diagram showing overall composition of groundwater samples collected during August and November 2007

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

The Construction of pDH25-pcpC-Vgb as a Recombinant DNA System for the Intracellular Expression of Vitreoscilla Hemoglobin in *Cephalosporium Acremonium*

Yubin Liu

College of Life Science and Technology

Beijing University of Chemical Technology

Beijing 100029, China

E-mail: Lau.yubin@gmail.com

Qiang Li (Corresponding Author)

Laboratory of Antibody and Biocatalysts Engineering

Department of Chemical Engineering

Tsinghua University

Beijing 100084, China

E-mail: liqiang@mail.tsinghua.edu.cn

Qipeng Yuan

College of Life Science and Technology

Beijing University of Chemical Technology

Beijing 100029, China

E-mail: yuanqp@mail.buct.edu.cn

Abstract

A recombinant DNA system for the intracellular expression of a bacterial heme-binding protein (*Vitreoscilla hemoglobin*, Vgb) was constructed and named as pDH25-pcpC-Vgb. It could be introduced into a cephalosporin C-producing strain of *Cephalosporium acremonium*. The Vgb-expressing transformants will provide higher internal oxygen concentrations, which will cause higher yields of cephalosporin C.

Keywords: *Cephalosporium acremonium*, *Vitreoscilla hemoglobin*(Vgb), pDH25, Recombinant DNA

1. Introduction

Cephalosporium acremonium is used for industrial production of the β-lactam antibiotic cephalosporin C. For decades, cephalosporin C yields have increased through development of better production strains by classical mutagenesis procedures and optimization of the growth conditions. Recently recombinant DNA technology has been used to overexpress the potentially rate-limiting enzyme deacetoxycephalosporin C synthetase which resulted in increased cephalosporin C production in pilot scale fermentation (Skaturd, 1989, p.477-485). Thus, engineering of filamentous fungi through genetic manipulation represents a promising technology for rational development of improved strains.

Cephalosporin C biosynthesis is regulated by environmental factors such as the phosphate (Zhang, J., 1998, p.242-247), carbon (Martin, J. F., 1982, p.258-268) and oxygen (Hilgendorf, P., 1987, p.247-251) contents of the medium. The overall rate of cephalosporin C synthesis is severely reduced under conditions of low oxygen. Reduction of oxygen supply leads to accumulations of penicillin N, a precursor of cephalosporin C. The mechanism of oxygen control over cephalosporin C biosynthesis is not well understood. Possibly, low oxygen levels directly affect the biosynthesis pathway of cephalosporin C, which includes three oxidation reactions. It is also possible that a more efficient overall metabolism provided by higher oxygen levels indirectly results in higher cephalosporin C yields. Regardless of the mechanism, technology that improve aerobic metabolism in these organisms should have a positive effect on cephalosporin C production (Kallio, P. T., 1994, p.201-218).

Intracellular expression of a bacterial heme-binding protein (*Vitreoscilla hemoglobin*, Vgb) has resulted in higher productively of industrial cell types. For example, the expression of Vgb in the filamentous bacterium *streptromyces coelicolor* resulted in tenfold higher yield of the polyketide actinorhodin in bench scale batch fermentation run under reduced aeration (Magnnolo, S. K., 1991, p.473-476). Also, the efficiency of cloned protein synthesis by oxygen-limited *Escherichia coli* was increased in Vgb-expressing strains (Cullen, D., 1987, p.21-26). Possible mechanisms for Vgb action include increasing the flux of oxygen to the respiratory apparatus, providing higher internal oxygen concentrations, altering the internal redox state, or functioning as an efficient terminal oxidase.

The present work describes the construction of a recombinant DNA system for the intracellular expression of Vgb in filamentous fungi. The recombinant DNA system was named as pDH25-pcpC-Vgb.

2. Experiment

Fungal strain *Cephalosporium acremonium* AS 3.3502 was obtained from The Committee on Type Culture Collection of Chinese Academy of Sciences. Plasmid pDH25 was donated by Pro. Cullen, D. from University of Wisconsin-Madison (Figure 1.). pDH25 provide the structure and hygromycin B resistance gene of recombinant plasmid. Plasmid pVgb and pMW I was obtained from China Agricultural University. The pVgb provides Vitreoscilla hemoglobin gene (Vgb) and the pMW I provides the promoter of isopenicillin N synthetase gene (pcpC) from *Penicillium chrysogenum*. *Escherchia coli* strains Top10 and DH5α were used as recipient hosts for all DNA work. Enzyme and Chemical medicines used in the experiment was purchased from commercial sources, such as Takara Co., Ltd. LB plates with resistance were used for selection. All DNA manipulations were performed by standard protocol (Sambrook, J., 1989).

3. Methods and Results

3.1 Ligation of promoter pcpC and Vgb gene by over-lap PCR

The over-lap PCR premiers was designed as follow: pcpC-OL-Xba I primer1: 5'-GCTCTAGAGGGGCGTCGAGTTGC-3', pcpC-OL primer2: 5'-TTGCTGGTCTAACATGGTGACGGTTTGTCC-3', Vgb-OL primer1: 5'-GGACAAACCGTCACCATGTTAGACCAGCAA-3', Vgb-OL-Xba I primer2: 5'-GCTCTAGATTATTCAACCGCTTGAGCG-3', The sequences of pcpC and Vgb were joined together by the overlapped 15 bp in primer and introduced the Xba I site at each end, thus pcpC-Vgb was constructed. Over-lap PCR was performed by two steps. Firstly, pcpC-OL-Xba I primer1and pcpC-OL primer2 was used for amplifying pcpC, and Vgb-OL primer1 and Vgb-OL-Xba I primer2 for Vgb gene. The amplifying program was showed in Table 1., Table 2. and Figure 2. Secondly, the above product was amplified as template. Table 3. and Figure 2. showed the amplifying program. After amplification and purification, pcpC-Vgb was acquired with Xba I site at each end. The right fragment pcpC-Vgb was confirmed by restriction enzyme digestion and sequencing (Figure 3.).

3.2 Construction of plasmid pDH25-pcpC-Vgb for intracellular expression of Vgb

Plasmid PDH25-pcpC-Vgb was constructed to intracellular express of Vgb in *Cephalosporium acremonium*. Figure 4. graphically illustrates the construction of this plasmid. Plasmid pDH25 and fragment pcpC-Vgb were digested by restriction enzymes Xba I, then were ligated together and named as pDH25-pcpC-Vgb. Plasmid pDH25 was modified to form plasmid pDH25 by addition a 890bp Xba I restriction fragment pcpC-Vgb, which was constructed above. It contains a promoter of isopenicillin N synthetase (pcpC), which can be identified by *Cephalosporium acremonium* ribosome, and can express Vgb successfully if the plasmid integration successful. Plasmid pDH25 contains hygromycin B resistance gene encoding hygromycin B phosphotransferase, which provide a selectable marker for transformation of *Cephalosporium acremonium*. The ampicillinase gene of pBR322 provides the selectable maker of the construct work in *Escherichia coli*. Hygromycin B resistance gene and Vgb have the same transcriptional termination region (Ttrp C) from the tryptophan C gene of *Aspergillus nidulans* in pDH25-pcpC-Vgb. The right plasmid was identified by digestion with restriction enzyme Mlu I (Figure 5.).

4. Conclusion

(1) Plasmid pDH25-pcpC-Vgb can be transformed to *Cephalosporium acremonium*. The transformation will involve chromosomal integration of the plasmid. If it transformed successful, the plasmid will express Vgb and Hygromycin B resistance gene separately, which use its own promoter. The plasmid has the structure of polycistron. The level of Vgb protein expression will be determined by the strain and the location of integration. Further, the EcoR I-Cla I trpC fragment of pDH25 can be conveniently replaced with other promoters for expression Vgb in other filamentous fungi.

(2) The important role played by bacterial hemoglobin in metabolic activities makes this protein a suitable candidate for providing higher internal oxygen concentrations. According the reference (Cullen, D., 1987, p.21-26), Vgb improves antibiotics productivity under low aeration. For both aeration conditions studied, there was no difference in overall growth between Vgb-expressing cells and control cells.

(3) Oxygen plays a major role in cephalosporin C biosynthesis. The tripeptide cyclase, ring expansion, and hydroxylation of deacetoxycephalosprion C steps are all oxygen-requiring processes (DeModena, J. A.,1993, p.926-929). In addition, oxygen regulation may be more significant in the later steps of the pathway due to the accumulation of penicillin N in oxygen-deprived cultures (Khosla, C., 1990, p.849-853). It is possible that Vgb may directly benefit these reactions through its ability to provide higher levels of intracellular oxygen. In a related study, synthesis of the polyketide antibiotics actinorhodin by *streptomyces coelicolor*, another oxygen-intensive pathway, was improved ten-fold in Vgb-expressing cells grown under oxygen-limited conditions (Herold, T., 1988, p.168-173). Possible roles suggested for other unicellular hemoglobins, include oxygen buffering and electron transfer. To help elucidate the role of Vgb in filamentous fungi, it would be of interest to determine its cellular location. Future studies in *Cephalosporium acremonium* will also examine the effect of Vgb on specific enzyme and intermediates of the cephalosporin C pathway, since, to be maximally useful, Vgb must affect cell metabolism in a manner complementary to mutations that may already exist in high-producing industrial strains.

5. Acknowledgement

We would like to thank professor D. Cullen at University of Wisconsin-Madison for donation us the plasmid pDH25.

References

Cullen, D., Leong, S. A., Wilson, L. J. and Henner, D. J. (1987). Transformation of *Aspergillus nidulans* with the hygromycin- resistance gene, hph. *Gene*, 57, 21-26.

DeModena, J. A., Gutierrez, S., Vetasco, J., Femandez, F. J., Fachini, R. A., Galazzo, J. L., Hyghes, D. E. and Martin, J. F. (1993). The production of cephalosporin C by *Acremonium Chrysogenum* is improved by intracellular expression of a bacterial hemoglobin. *Bio/Technology*, 11, 926-929.

Herold, T., Bayer, T. and Schugert, K. (1988). Cephalosporin production in a stirred tank reacter. *Appl. Microbiol. Biotechnol*, 29, 168-173.

Hilgendorf, P., Heiser, V., Diekmann, H. and Thoma, M. (1987). Constant dissolved oxygen concentration in cephalosporin C fermentation: Applicability of different controllers and effect on fermentation parameters. *Appl. Microbiol. Biotechnol*, 27, 247-251.

Kallio, P. T., Kim, D. J., Tsai, P. S. and Bailley, J. E. (1994). Intracellular expression of Vitreoscilla hemoglobin alters *Escherichia coli* energy metabolism under oxygen limited condition. *Eur. J. Biochem*, 219,201-218.

Khosla, C., Curtis, J. E., DeModena, J. A., Rinas, U. and bailey, J. E. (1990). Expression of intracellular hemoglobin improve protein synthesis in oxygen-limited *Escherichia coli*. *Bio/Technology*, 8, 849-853.

Magnnolo, S. K., Leenutaphong, D. L., DeModena, J. A., Curtis, J. E., Bailey, J. E., Galazzo, J. L. and Hughes, D. E. (1991). Actinorhodin production by *Streptomyces coelicolor* and growth of *Streptomyces lividans* are improved by the expression of a bacterial hemoglobin. *Bio/technology*, 9, 473-476.

Martin, J. F., Revilla, G., Zanca, D. M. and Lopez-Nieto, M. J. (1982). *Carbon catabolite regulation of penicillin and cephalosporin biosynthesis*. Tokyo: Antibiotics Research Association. pp. 258-268.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A laboratory manual.* (2nd ed). Cold Spring Harbor: Cold Spring Harbor laboratory Press.

Skaturd, P. L., Teitz, A. J., Ingolia, T. D., Cantwell, C. A., Fisher, D. L., Chapman, J. L. and Queener, S. W. (1989). Use of recombinant DNA to improve production of cephalosporin C by *Cephalosporin acremonium*. *Bio/technology*, 7, 477-485.

Zhang, J., Wolfe, S. and Demain, A. L. (1998). Phosphate repressible and inhabitable β-lactam synhetases in *cephalosporin acremonium* strain C-10. *Appl. Microbiol. Biotechnol*, 29, 242-247.

Table 1. Reaction System of pcpC PCR in Over-lap PCR Step 1

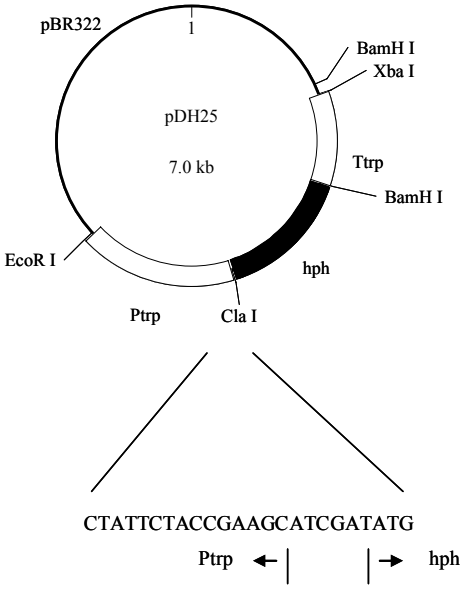
Reactant	Dosage (uL)
Sterile water	13.8
10 *LA Taq Buffer	2
dNTPs (each dNTP is 10 mM)	1
PcpC-OL-Xba I primer1	1
PcpC-OL primer2	1
pMW I	1
LA Taq	0.2
Total	20

Table 2. Reaction System of Vgb PCR in Over-lap PCR Step 1

Reactant	Dosage (μL)
Sterile water	13.8
10 *LA Taq Buffer	2
dNTPs (each dNTP is10 mM)	1
Vgb-OL primer1	1
Vgb-OL-Xba I primer2	1
pVgb	1
LA Taq	0.2
Total	20

Table 3. Reaction System of Over-lap PCR Step 2

Reactant	Dosage (μL)
Sterile water	76.8
10 *LA Taq Buffer	14
dNTPs (each dNTP is10 mM)	7
pcpC-OL-Xba I primer1	7
Vgb-OL-Xba I primer2	7
Vgb(production of PCR above)	7
pcpC(production of PCR above)	7
LA Taq	1.4
Total	140



Ptrp(1250bp): the promoter of *Aspergillus nidulans trpC gene*

hph(1040bp): hygromycin B resistance gene (encoding hygromycin B phosphotransferase)

Ttrp(710bp): the terminator of *Aspergillus nidulans trpC* gene

pBR322(3990bp): BamH I-EcoR I fragment of pBR322 followed by the Xba I-BamH I linker deviced from the PUC18 polylinker

Figure 1. Structure of Vgb Express Vector pDH25

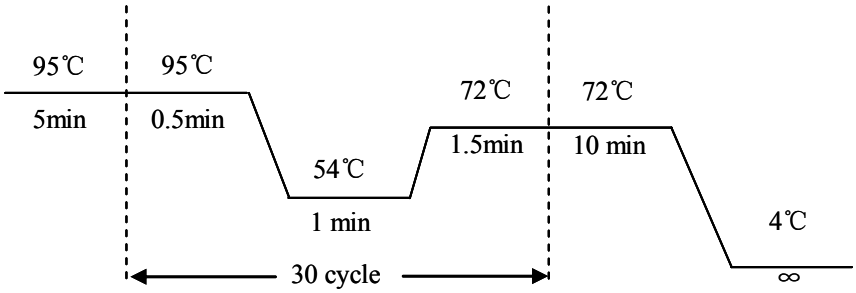


Figure 2. Set of Temperature, Time and Cycle in Over-lap PCR

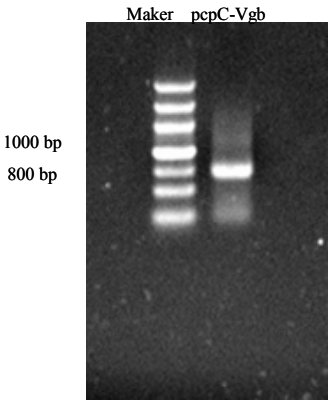


Figure 3. Polymerase Chains Reaction of pcpC and Vgb Over-lap

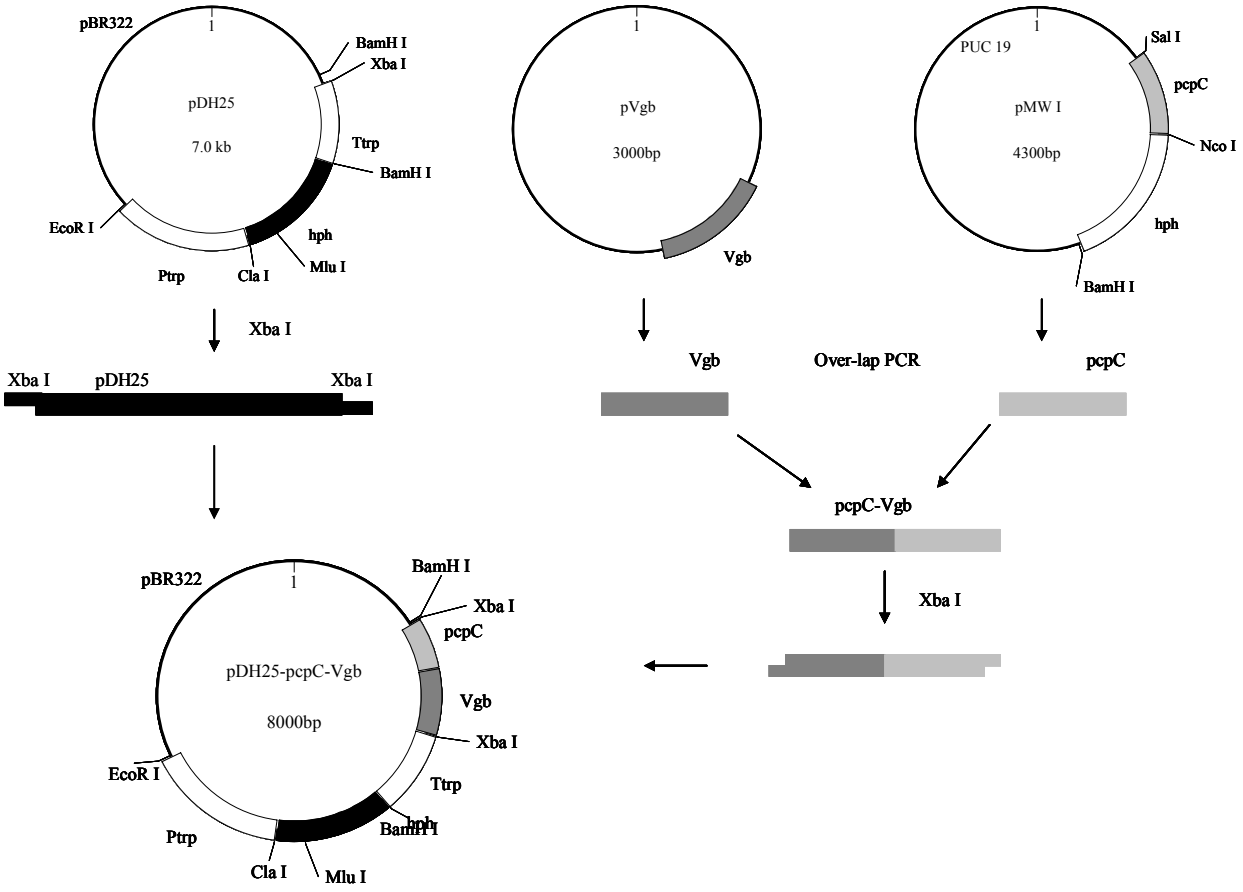


Figure 4. Construction of Plasmid pDH25-pcpC-Vgb.

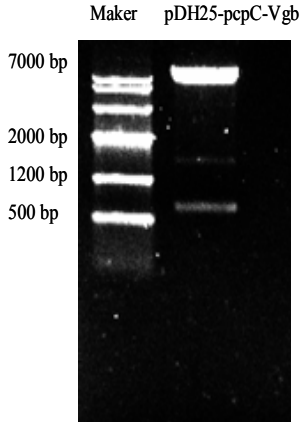


Figure 5. Clave of pDH25-pcpC-Vgb by Mlu I

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

High Spatial Resolution Land Cover Mapping Using Remotely Sensed Image

H. S. Lim, M. Z. MatJafri & K. Abdullah

School of Physics

Universiti Sains Malaysia

11800 Penang, Malaysia

Tel: 60-4-653-3888, Fax: 60-4-657-9150

E-mail: hslim@usm.my, mjafri@usm.my, khirudd@usm.my

The research is financed by Ministry of Science, Technology and Innovation (MOSTI) Science Funds MOSTI Science Funds and USM short term grants.

Abstract

Land cover classification from remotely sensed data is an important topic in remote sensing applications. This paper present an economical analysis of land covers in Mebok Estuary, Penang, Malaysia. This study investigated the potential of using digital camera for land cover mapping Mebok Estuary, Penang, Malaysia. Airborne digital imagery has proved to be an effective tool for land cover studies. A Kodak camera, model DC290, was used to capture images from an elevation of 2438.4 meter on board Cessna 172Q. The use of digital camera as a sensor to capture digital images is cheaper and economical compared to the use of other airborne sensor. This technique overcomes the problem of the difficulty in obtaining cloud-free scenes in the Equatorial region from a satellite platform. Supervised classification techniques were used in the classification analysis. Supervised classification technique (Parallelepiped with Maximum Likelihood as tie breaker classifier, PML, Maximum Likelihood, ML, Minimum Distance-to-Mean, MDM, and Parallelepiped, P) was applied to the digital camera spectral bands (red, green and blue) to extract the thematic information from the acquired scenes. The best supervised classifier was chosen based on the highest overall accuracy and Kappa statistic. The accuracy of the classified images was validated using a reference data set. The study revealed that the Parallelepiped with Maximum Likelihood as tie breaker classifier produced superior result and achieved a high degree of accuracy. The classified land cover map was geometrically corrected to provide a geocode map. The results produced by this study indicated that land cover features could be clearly identified and classified into a land cover map. This study suggested that the land cover types of Mebok Estuary, Penang, Malaysia can be accurately mapped.

Keywords: Classification, Land Cover, Airborne, Supervised classification, Kodak

1. Introduction

Land cover is a fundamental variable that impacts on and links many parts of the human and physical environments (Foody, 2002). This is important for it provides useful information for planning the area. Remote sensing technology is able to prepare the latest data with multiple scales and can be used in many purposes. Land cover mapping at coarse spatial resolution provides key environmental information needed for scientific analyses, resource management and policy development at regional, continental and global levels (Latifovic, et al., 2004). The increasing availability of remote-sensing images, acquired periodically by satellite sensors on the same geographical area, makes it extremely interesting to develop the monitoring systems capable of automatically producing and regularly updating land-cover maps of the considered site (Bruzzone, et al., 2002).

A basic problem for any country or region in economic planning, environmental studies, or resource management is accurate, current information. The need for basic surface characteristic information, such as land use and land cover, is critical to both scientific analysis and decision-making activities. Without accurate information, scientists cannot

complete valid studies and decision-makers often fail to make decisions or make incorrect decisions (Haack and Bechdol, 2000). The availability of remote sensing data applicable for global, regional and local environment monitoring has greatly increased over recent years (Ehlers, et al., 2003). Land cover is a fundamental parameter describing the Earth's surface. With sufficient calibration, a land cover map can be used to identify spatial patterns of physical quantities such as carbon storage or vegetation cover as well as more abstract phenomena such as land use. Accurate information on land cover is required for both scientific research (eg, climate change modeling, flood prediction) and management (eg, city planning, disaster mitigation) [Tatem, et al., 2001].

The sensor used in this study was a normal digital camera model Kodak DC290, and a light aircraft was used as a platform to capture the digital images. The cost of the present sensor is very much cheaper than available airborne sensors. Besides, this study is also good for collecting data in planning and management. Digital cameras have been used in many researchers in remote sensing application (Ellis and Dodd 2001, Nakada and Chikatsu 2003, Levesque and King 1999, Heier and Hinz, 1999, White, et al. 2000, Mohammed et al. 2007, Yamamoto and Takagi 2007, Fugate and Aaron 2008). A Canon Camera EOS 1D Mark II 8 MPixel digital camera was used by Reinartz, et al., (2006) for traffic monitoring. Dean, et al, (2000) and Mason, et al., (1997) also used a Kodak DCS460c in their study of vegetation analysis and small area mapping respectively. A Nikon Coolpix 885, CP885 was used by Goddijn and White, (2006) for water quality measurements in Galway Bay.

The objective of this study is to investigate the potentiality of using digital camera imagery for land cover mapping. In this study, digital images were acquired from a digital camera for land cover mapping. Supervised classification of remote-sensing images has been widely used as a powerful means to extract various kinds of information concerning earth environment. Supervised classification methods were applied to the digital images. Maximum Likelihood classifier was found to produce the best accuracy in this study. Many researchers choose the Maximum Likelihood method in their studies [Saura and Miguel-Ayanz, (2002), Pal and Mohanty, (2002), Donoghue and Mironnet, (2002), Thiemann and Kaufmann, (2000) and Guerschman, et al., (2003)]. The monitoring task can be accomplished by supervised classification techniques, which have been proven to be effective categorization tools (Bruzzone, et al., 2002). The accuracy assessment of the classified images also has been done in this study. The traditional method of collecting data for planning is surveying samples at field. Remote sensing technique is a useful method for classifying the image. With process data available, a quick decision about the area can made.

2. Study area and Data Acquisition

The study area was the Merbok river estuary, located within latitudes and 5^o 39' N to 5^o 41' N and longitudes 100^o 20' E to 100^o 24' E, respectively (Figure 1). The study area was located at the north of Peninsular Malaysia. The whole length of Merbok River, being tidally influenced, is mangroved-fringed where saltwater intrusion can be experienced until the upstream end of the river. The estuary width ranges from about 2000 m at the estuary mouth to about 15 meter towards the upper reaches with depths varying from 15 meter to 3 meter (except for a few 20 meter deep holes where tributaries join the Merbok). The catchment area of Merbok Estuary measures 500,000,000 m^2. The dry season is in December, January and May with rainfall maxima in October-November and April. The mean annual discharge is estimated to be 20 m^3s^{-1}. Freshwater discharge into the mangrove estuary is not from a single river but rather from numerous small streams that feed into a number of estuarine tributaries. So even if the main river is gauged, this discharge does not represent the total freshwater input to the estuary. A digital camera, Kodak DC 290 (Figure 2) was used to capture digital images from a light aircraft, Cessna 172Q, at 2438.4 meter of altitude. The digital images were captured during the flight between 9 a.m. to 11 p.m. on 9 March 2002 from a light aircraft flying at an altitude of 2438.4 meter on 9 March 2002.

3. Remote Sensor

Kodak is without question one of the dominant companies in the field of photography. While the bulk of their enormous business is built upon conventional film-based photography, they clearly recognize that the future lies with digital (The Imaging Resource, 1999). In this study, red–green–blue (RGB) digital outputs from a digital camera, a Kodak DC 290, were used to estimate water quality parameters. The front of the Kodak DC290 holds the 3x optical zoom lens, autofocus sensor, flash, self-timer light, optical viewfinder, light and flash sensors and a nice, thick hand grip (Figure 2). The Kodak DC 290 is a 2.3-mega pixels digital camera and storage digital images in JPEG and uncompressed TIFF file formats. The technical specification for the Kodak DC 290 can be obtained from KODAK DC290 Zoom Digital Camera User's Guide Camera Specifications - http://wwwde.kodak.com/global/en/service/digCam/dc290/ownerManual/ch14.shtml#76207. Its sensor is capable of acquiring data in multispectral mode (3-bands: red, green and blue). According to Ahmad, et al., (2002), Kodak DC 290 is categorized as a high-resolution digital camera.

The technical specification for the Kodak DC 290 is shown in Table 1. Its sensor is capable of acquiring data in multispectral mode (3-bands: red, green and blue). There are many types of digital cameras available in the market. Any person could possess the camera either for personal use, for producing reports or for research and other purposes. The

use of digital camera is preferable since digital images could be stored on-board and easy to operate. Digital cameras are manufactured with different sensor sizes and resolutions. Sensor size is defined by width and height of sensor array in millimeter or inches. On the other hand, sensor resolution is defined by number of horizontal pixel multiplied by number of vertical pixel. A digital camera could be categorized as (1) low resolution if the number of pixels is less than 500, 000; (2) medium resolution between 500, 000 to 1.5 million and (3) high resolution if number of pixel is beyond 1.5 million. The cost of the digital camera depends on the resolution. Normally, digital camera of low resolution is much cheaper compared to digital camera of medium and high resolution (Ahmad, et al., 2002). According to Ahmad, et al., (2002), Kodak DC 290 is categorized as a high-resolution digital camera. Recently, the number of pixels of amateur digital cameras is amazingly increased by modern semiconductor and digital technology. Nevertheless, the highest pixel of amateur digital camera was 0.8 Mega pixels in 1996 and the highest pixel was 5.24 Mega pixels in 2001. It is expected that high-resolution amateur digital camera will become a useful tool in various photogrammetric fields (Nakada and Chikatsu, 2003).

4. Methodology

The aim of the classification analysis is to categorize all of the pixels in the digital camera imagery into land cover classes. Figure 2 illustrates the classification analysis flow chart. Basically, the process can be divided into three simple steps, the pre-processing, data classification and output. In the pre-processing, the classes were established by using polygons for training sites. They are delineated by spectrally homogeneous sub areas, which have, class name given. Accuracy assessment was done in this study to compute the probability of error for the classified map. Many methods of accuracy assessment have been discussed in remote sensing literatures. Three measures of accuracy were tested in this study, namely overall accuracy, error matrix and Kappa coefficient. The confusion matrix is currently at the core of the accuracy assessment literature (Foody, 2002). The most widely promoted and used accuracy measure, however, may be derived from a confusion or error matrix. The Kappa statistic is a statistical method of assessing accuracy that takes into account the chance of random agreement. This statistic has been used by many researchers in their studies [Selamat, et al., (2002), Dymond and Johnson, (2002)].

5. Data Analysis and Results

The size of the airborne colour digital image of the Merbok River estuary is 1200 pixels by 1792 lines (Figure 3). All image-processing analysis was carried out using PCI Geomatica version 10.1.3 software at the School of Physics, Universiti Sains Malaysia (USM), Malaysia. Three supervised classification methods were performed to the digital images (Maximum Likelihood, Minimum Distance-to-Mean, and Parallelepiped). The digital image was classified into four classes using all the three bands. Training sites were needed for supervised classification. Selection of training areas in this study was based on the colour image. A total of 12 sample-training areas were studied in this analysis. A total of 100 samples were chosen randomly for the accuracy assessment.

The digital image was classified into 6 classes, namely forest, water, land, urban, cloud and cloud Shadow. Kappa coefficient and overall accuracy results of the three measures of accuracy are shown in Table 2. The overall accuracy is expressed as a percentage of the test-pixels successfully assigned to the correct classes. Producer's accuracy indicates how well training set pixels were classified, user accuracy indicates the probability that a classified pixel actually represents that category in reality and overall accuracy is the total number of correctly classified pixels divided by the total number of reference pixels (Rogan, et al., 2002).

Parallelepiped with Maximum Likelihood as tie breaker classifier produced the highest degree of accuracy with overall accuracy of 89.5%, Maximum Likelihood gave overall accuracy of 88.0%, Minimum Distance-to-Mean gave overall classification accuracy of 75.5%, and Parallelepiped resulted in the overall classification accuracy of 26.5%. Table 2 shows the results of overall classification accuracy and Kappa coefficient using supervised classification techniques of PML, ML, MDM and P classifier. The classified images using supervised classification techniques of PML, ML, MDM and P classifier are shown in Figure 4.

6. Conclusion

From the four classified maps, Parallelepiped with Maximum Likelihood as tie breaker gives a good result for land cover mapping. The digital camera imagery can be used to provide useful data for planning and management. The digital images used in this study were captured by a Kodak DC290 digital camera from a low-altitude flying aircraft that is cheap and economical compared with using a scanner/sensor placed on either airborne or satellite. Besides, this technique can also overcome the problem of difficulty in obtaining cloud-free images from satellites in equatorial region. This study provide an alternative to obtain the remotely data for land cover mapping.

Acknowledgements

This project was carried out using the USM short term grants and Ministry of Science, Technology and Innovation (MOSTI) Science Funds. We would like to thank the technical staff who participated in this project. Thanks are also extended to USM for support and encouragement.

References

Ahmad, A., Busu, I. and Desa, G., (2002). Assessment of The Impact of Digital Camera Resolution on Accuracy and Precision of Targeted Points. Proceeding of the International Symposium and Exhibition on Geoinformation.

Bruzzone, L., Cossu, R. and Vernazza, G. (2002). Combining parametric and non-parametric algorithms for a partially unsupervised classification of multitemporal remote-sensing images. Information Fusion 3, 289 –297.

Dean, C., Warner, T. A. and Mcgraw, J. B., (2000). Suitability of the DCS460c colour digital camera for quantitative remote sensing analysis of vegetation. ISPRS Journal of Photogrammetry & Remote Sensing, 55, 105–118.

Donoghue, D. N. M. and Mironnet, N. (2002). Development of an integrated geographical information system prototype for coastal habitat monitoring. Computers and Geosciences, 28, 129-141.

Dymond, C. C. and Johnson, E. A. (2002). Mapping vegetation spatial patterns from modeled water, temperature and solar radiation gradients. Journal of Photogrammetry and Remote Sensing, 57, 69-85.

Ehlers, M., Gahler, M. and Janowsky, R., (2003). Automated analysis of ultra high resolution remote sensing data for biotope type mapping: new possibilities and challenges, ISPRS Journal of Photogrammetry & Remote Sensing, 57, 315– 326.

Ellis, J. M. and Dodd, H., Using Airborne Digital Cameras for Environmental Applications. [Available Online]. http://www.hjw.com/pdf-articles/airborne-digital-apps2.pdf.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. Remote Sensing and Environment, 80, 185-201.

Fugate, B. and Aaron, L, (2008). Automated Feature Extraction Using Digital Camera Imagery, Proceeding of the 2008 West Virginia GIS Conference Asian. [Available Online] http://mapwv.gov/wvagp/conference/presentations/Fugate.ppt.

Goddijn, L. M. and White, M., (2006). Using a digital camera for water quality measurements in Galway Bay, Estuarine, Coastal and Shelf Science**,** 66(3-4), 429-436.

Guerschman, J. P., Paruelo, J. M., Di Bella, C., Giallorenzi, M. C. and Pacin, F., (2003). Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data, International Journal of Remote Sensing, 24(17), 3381–3402.

Guershman, J. P., Paruelo, J. M., Di Bella, C., Giallorenzi, M. C. and Pacin, F., (2003). Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data, *International Journal of Remote Sensing*, 24(17), 3381–3402.

Haack, B. and Bechdol, M., (2000). Integrating multisensor data and RADAR texture measures for land cover mapping, Computers & Geosciences, 26, 411 – 421.

Heier, H. and Hinz, A., A Digital Airborne Camera System for Photogrammetry and Thematic Applications. [Online] http://www.ipi.uni-hannover.de/html/publikationen/1999/isprs-workshop/cd/pdf-papers/heier.pdf.

Kodak DC290 Zoom Digital Camera User's Guide Camera Specification, [Available Online] http://wwwde.kodak.com/global/en/service/digCam/dc290/ownerManual/ch14.shtml#76207.

Kodak DC290 Zoom Digital Camera User's Guide Camera Specification, [Available Online] http://wwwde.kodak.com/global/en/service/digCam/dc290/ownerManual/ch14.shtml#76207.

Latifovic, R., Zhu, Z. L., Cihlar, J., Giri, C. and Olthof, I., (2004). Land cover mapping of North and Central American-Global Land Cover 2000, Remote sensing of environment, 89, 116 – 127.

Levesque, J. and King, D. J., (1999). Airborne Digital Camera Image Semivariance for Evaluation of Forest Structural Damage at an Acid Mine Site. Remote Sensing of Environment, 68, 112−124.

Mason, S., Ruther, H. and Smit, J., (1997). Investigation of the Kodak DCS460 digital camera for small-area mapping. ISPRS Journal of Photogrammetry & Remote Sensing, 52, 202−214.

Mohamme, A., White, L and Aiken, C., (2007). Capturing, Visualizing, and Analyzing Geological Outcrops with 3d Laser Rangefinders and Terrestrial Laser Scanners and Digital Camera, Proceeding of the Asian Conference on Remote Sensing 2007 (ACRS). [Available Online] http://www.aars-acrs.org/acrs/proceeding/ACRS2007/Papers/TS31.6.pdf.

Nakada, R. and Chikatsu, H., (2003). Generating 3D Model of "Meguro Residence" using Digital Armature Camera. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV−5/W10.

Pal, S. R. and Mohanty, P. K., (2002). Use of IRS-1B data for change detection in water quality and vegetation of Chika lagoon, east coast of India. International Journal of Remote Sensing, 23(6), 1027–1042.

Reinartz, P., Lachaise, M., Schmeer, E., Krauss, T. and Runge, H., (2006). Traffic monitoring with serial images from airborne cameras**,** ISPRS Journal of Photogrammetry and Remote Sensing**,** 61 (3-4), 149-158.

Rogan, J., Franklin, J. and Roberts, D. A. (2002). A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. Remote Sensing and Environment, 80, 143-156.

Saura, S. and Miguel-Ayanz, J. S., (2002). Forest cover mapping in Central Spain with IRS-WIFS images and multi-extent textual-contextual measures. International Journal of Remote Sensing, 23(3), 603–608.

Selamat, I., Nordin, L., Hamdan, N., Mohti, A. and Halid, M. (2002). Proceeding of the Seminar Kumpulan Pengguna TiungSAT-1, Jabatan Remote Sensing dan Sains Geoinformasi, Fakulti Kejuruteraan Dan Sains Geoinformasi Universiti Teknologi Malaysia and Astronautic Technology (M) Sdn Bhd.

Stargate Syetems, INC. Digital camera - Kodak, (1999). Kodak - Digital cameras. [Online]. http://www.stargate−usa.com/digitalcameras/stargate.dc.kodak.html.

Tatem, A., J., Lewis, H. G., Atkinson, P. M. and Nixon, M. S., (2001). Multiple-class land-cover mapping at the sub-pixel scale using a Hopfield neural network, JAG, V 3(2), 184 – 190.

The Imaging Resource, (1999). [Available Online]. http://www.imaging-resource.com/PRODS/DC290/DC290P.HTM.

Thiemann, S. and Kaufmann, H., (2000). Determination of Chlorophyll Content and Trophic State of Lakes Using Field Spectrometer and IRS-1C Satellite Data in the Mecklenburg Lake District, Germany. Remote Sensing of Environment, 73, 227–235.

White, M. A., Asner, G. P., Nemani, R. R., Privette, J. L. and Running, S. W., (2000). Measuring Fractional Cover and Leaf Area Index in Arid Ecosystems: Digital Camera, Radiation Transmittance, and Laser Altimetry Methods. Remote Sensing of Environment, 74, 45–57.

Yamamoto, T and Takagi, M., (2007). Registration for Digital Camera Image of Paper Map Distorted By Transparent Filem, Proceeding of the Asian Conference on Remote Sensing 2007 (ACRS). [Available online]. http://www.aars-acrs.org/acrs/proceeding/ACRS2007/Papers/TS37.4.pdf.

Table 1. The technical specification for the digital camera model Kodak DC 290

Pixel Resolution	CCD	1901 x 1212=2.3 millions of pixels (total number of pixels)
	Ultra	2240 x 1500=3,360,000
	High	1792 x 1200=2,150,400
	Medium	1440 x 960=1,382,400
	Standard	720 x 480=345,600
Color		24-bit, millions of colors
Picture File Format		Exif version 2.1 (JPEG base) or TIFF
Picture Storage		External memory only: ATA compatible Compact Flash card
Viewfinder		Real image
ASA/ISO Sensitivity		100
Flash Range	Wide	1.6 ft to 13.1 ft (0.5 to 4.0 m)
	Telephoto	1.6 ft to 8.2 ft (0.5 to 2.5 m)
Lens	Type	Optical quality glass
	Maximum Aperture	Wide: F/3 Telephoto: F/4.7
	Zoom	6X: 3X Optical zoom 2X Digital zoom
	Focal Length	38.4 to 115.2 mm (equivalent to 35 mm camera) 8 to 24 mm (actual)
	Auto Focus	Wide/Telephoto: 1.0 ft (0.3 m) to infinity
	Manual Focus	Wide/Telephoto: 1.6 ft (0.5 m) to infinity
Power	Batteries	AA size 1.5-volt alkaline, or AA size 1.2-volt Ni-MH rechargeable
	DC Input	AC Adapter for KODAK DC200 Series Digital Cameras
Tripod Socket		.25 in. (.006 m) threaded
Video Out		NTSC or PAL
Dimensions	Width	4.6 in. (118 mm)
	Length	2.5 in. (63 mm)
	Height	4.2 in. (106 mm)
Weight		1.2 lbs (525 g) without batteries
Operating Temperature		32 to 104° F (0 to 40° C)

(Source: KODAK DC290 Zoom Digital Camera User's Guide Camera Specifications)

Table 2. The overall classification accuracy and Kappa coefficient.

Classification method	Overall classification accuracy (%)	Kappa coefficient
Maximum Likelihood	88.00	0.799
Minimum Distance-to-Mean	75.50	0.609
Parallelepiped with Maximum Likelihood as tie breaker	89.50	0.824
Parallelepiped	26.50	0.198

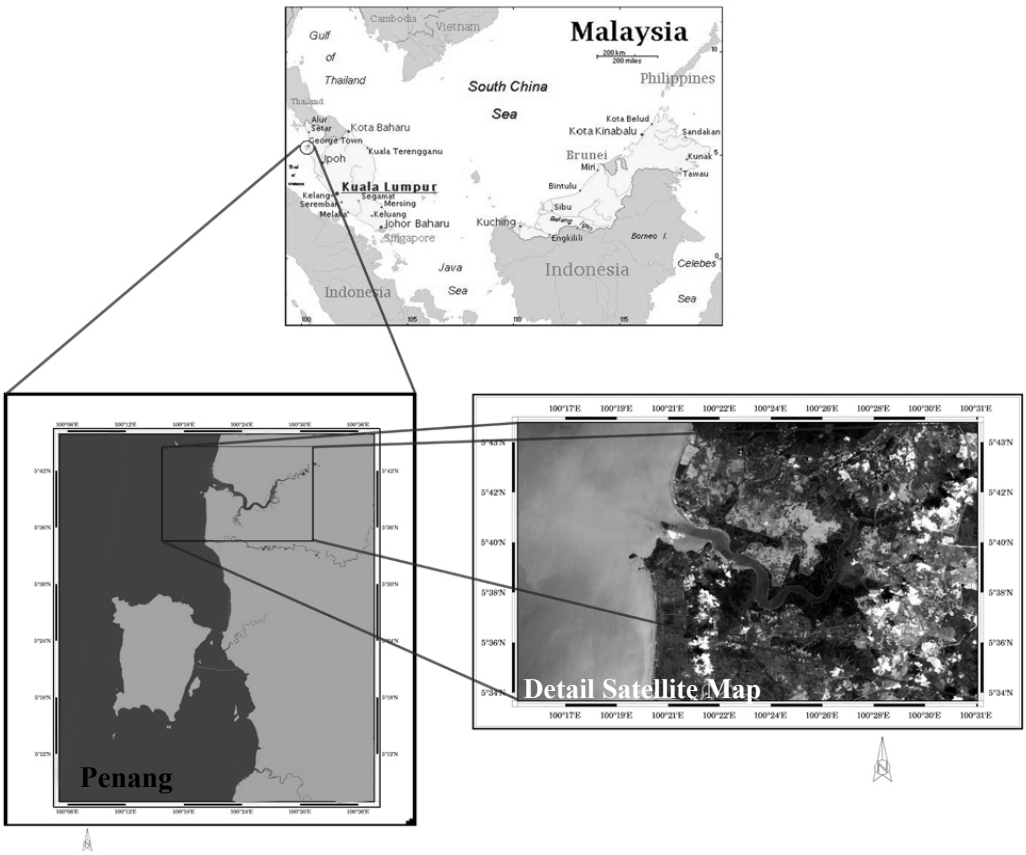


Figure 1. The study area

Figure 2. Digital Camera- Kodak DC290
(Source: Stargate Systems, INC: Digital Cameras - Kodak, 1999)

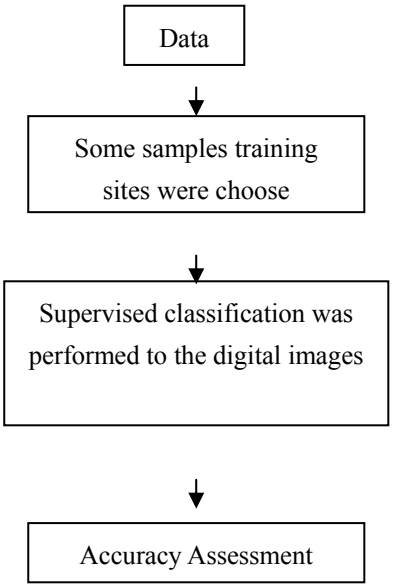
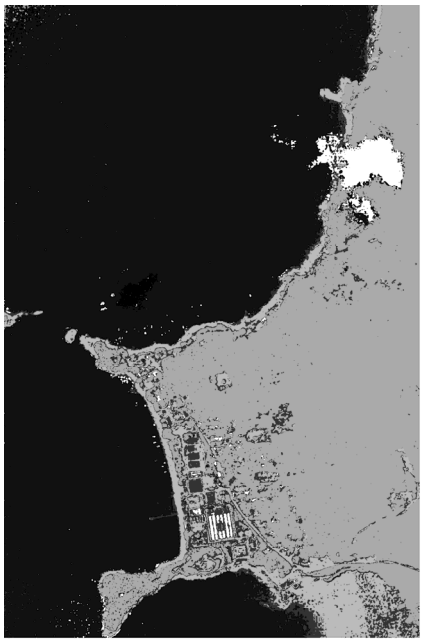
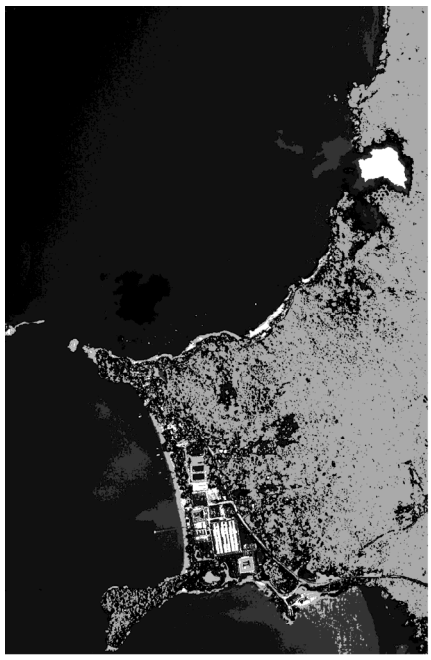


Figure 3. Flow chart for data processing of the images

Figure 4. The image used in the classification.

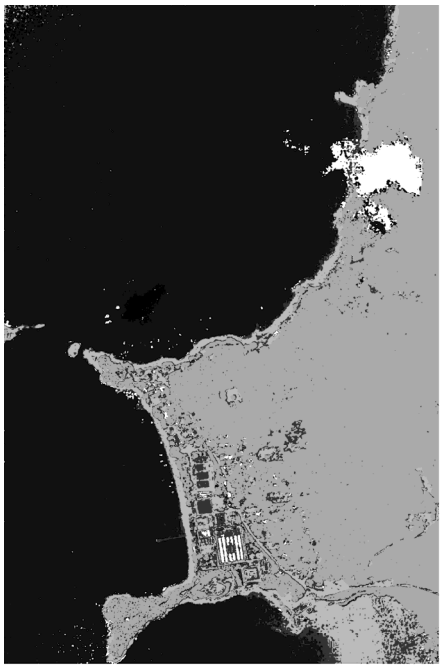


(a) Maximum Likelihood



(b) Minimum Distance-to-Mean

(c) Parallelepiped



(d) Parallelepiped with Maximum Likelihood as tie breaker

Figure 5. The classified images using supervised technique (a) Minimum Likelihood, (b) Minimum Distance-to-Mean, (c) Parallelepiped and (d) Parallelepiped with Maximum Likelihood as tie breaker (Forest = Green, Water = Blue, Land = Orange, Urban = Red, Cloud = White and Cloud Shadow = Black).

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Cognition and Investigation on the Manufacturing of Sesame-Flavor Liquor

Yuanmin Qi

Shandong Jingangshan Liquor Industry Co. Ltd., China

Abstract

Sesame-flavor liquor with the flavor hallmarks of thickness, purity and aromatic type and technique characteristic, is exquisite, unique in its own style, and belongs to the top grade of the liquors.

Technique outline of sesame-flavor liquor is: "steam and remain dregs, muddy bottom and brick furnace, with large bran, ferment with several microorganism, four higher, one more and one longer, blend elaborately", which highlights effects of Maillard reaction.

In the manufacturing process, we pay attention to the variety, quantity and fermentation environment of proteins and raw material ratio; Strengthen researches of effects of proteins in several amylopectin on components formation of sesame-flavor liquor; control the excessive consumption of nutrients and assure double bumpers of protein and starch fermentations; attempt to find new insights into the healthy factors of sesame-flavor liquor earlier.

Keywords: Sesame-flavor liquor, Maillard reaction, Process links, Cognition and investigation, Flavor and yield

1. Development and flavor traits of sesame-flavor liquor

Nowadays, liquors in China are classified into 11 species according to the flavor type, where there are 8 products of national standard, 1 of industry standard, 1 of the origin place standard and 1 of enterprise standard. Sesame-flavor liquor is one of the two intra-industry self-developed new liquor flavor types.

Sesame-flavor was first unconsciously found from Jingzhi white trunk by Yu Shuming, liquor-making expert in Shandong province in 1957. Sesame-flavor was first claimed by Xiong Zhishu in 1965. Sesame-flavor liquor have been investigated and developed in due form by many institutes and enterprises of provinces or industries ever since 1980s. Industry standard of "Sesame-flavor liquor" was issued in 1995, and subsequently national standard of that was published in 2006. Zhou Lixiang of Jinzhi Distillery, my collegemate, was one drafter of these two standards. Therefore, sesame-flavor liquor, as a banner of Shandong liquor, is popular over China, and deeply favored by many consumers. Meanwhile, remarkable progress was made by long-term research and remorseless efforts of Shandong liquor industry known for their representatives such as Jinzhi and Bandaojing.

Flavor hallmarks of thickness, purity and aromatic type was absorbed in the process of Sesame-flavor liquor, as well as a great deal of modern technological factors. With basic mature process, sesame-flavor liquor possesses property of purity and elegance from fen-flavor liquor, softness and fullness of highly-flavor liquor and refinement and nuance of aromatic type liquor, and also owns the compound fragrance of roasting sesame after comprehensive sensory evaluation. People feel relaxed and happy after drinking it. Sesame-flavor liquor was also advanced blending liquor of highly-flavor type.

Shandong sesame-flavor liquor was divided into three style characteristics by Huang Yeli; one is elegant type, Jinzhi Shenliang and Bandaojing sesame-flavor as representative, with higher content of ethyl acetate(160~200mg/100ml) and lower content of ethyl caproate(50~80mg/100ml), possesses the properties of serenity and elegance, purity and harmony, sesame-flavor grace; two is fragrant type, Baotu Spring sesame-flavor liquor as representative, with equivalent content of ethyl acetate and ethyl caproate(110~130mg/100ml), similar to Jiugui liquor, possesses the properties of mellow-bold and fullness, refinement and nuance and sesame-flavor; three is pit-flavor type, Shuihu 108 liquor as representative, based on traditional brewing of highly-flavor liquor, modified in the domain of material, fermentation pit and starter etc, with sesame-flavor due to long-term storage, with higher content of ethyl caproate than ethyl acetate, possessed the properties of pit-flavor and elegance, mellow-bold and fullness, typical sesame-flavor, and is popular in the southwest of Shandong province.

For owing characteristics of various traditional flavor types, fragrant but not too brilliant, sesame-flavor liquor has its own style and soon becomes fashionable all over China and top grade of liquors. Due to the elusive studies of its

fermentation mechanism and aroma components, gleaming typical sesame-flavor which was tough to control in the formation process, sesame-flavor liquor has intensively and extensively attracted many peers and experts' attention to further and elaborate investigation.

2. Process characteristics and links of sesame-flavor liquor

During the long period of practice, technology points of sesame-flavor liquor has been generally formed, which could be outlined as follows: "steam and remain dregs, muddy bottom and brick furnace, with large bran, ferment with several microorganism, four higher, one more and one longer(ingredient of high nitrogen, accumulation at high temperature, fermentation at high temperature, liquor flow at high temperature, more percentages of koji and long storage period), blend elaborately".

2.1 Basic process and fermentation vessel qualified for thickness, purity and aromatic type

Steaming is used to shiitake foodstuff without unhappy odor, remaining dregs is to inherit traditional process and beneficial to fermentation of glutinous rice and accumulation of aroma components. Referring to steaming and burning process of fen-flavor liquor, fresh sense is strengthened and quietly elegant style is highlighted.

Brick furnace is favorable to the formation of sesame-flavor with refinement and nuance, and muddy bottom can increase components of high-flavor liquor. Proper percentage of ethyl caproate plays a pivotal role in aroma releasing of sesame-flavor, and also in finesse and palatability of liquor body.

The application of muddy bottom and brick furnace is both different from the fermentations in muddy pit of highly-flavor liquor and in land cylinder of fen-flavor liquor, and similar to the fermentation in detritus or stone band pit of fen-flavor liquor. Artificial pit with muddy bottom where numerous pit mud microorganisms such as caproic acid bacteria and methane bacteria etc inhabit is beneficial to produce certain aroma constituents of ethyl caproate and caproic acid.

2.2 High nitrogen ingredients

Reducing sugar and amino acids as precursor matters of Maillard reaction originated from starch conversion and plant and animal protein, respectively. Certain proportion of bran and wheat is added to increase plant protein, and autolysates resulting from the death of numerous yeasts due to the accumulation process at high temperature are the source of animal protein.

Application of protein in production is closely associated with the variety, quantity, fermentation environment, pH value and utilization ratio of raw proteins. In practice, variety or ration of raw material should be paid attention to for its effects on microbial strains. Bacteria and yeast need more nitrogen nutrition while mould needs more carbonaceous nutrients. Bran contains high content of ammoniacal nitrogen, and high content of ferulic acid is transferred into phenolic compounds such as vanillin, vanillic acid, vanillic acid ester, 4-ethyl guaiacol and so on.Such variety and ratio of raw material play a vital role in regulating the propagation and metabolism of microbial strains and forming and enlarging sesame-flavor compositions.

To adjust the ratio of carbon to nitrogen and introduce microorganism qualified for glutinous rice circumstance is attempted to provide material base for Maillard reaction and produce more nitrogen compounds, especially pyrazine heterocyclic compounds.

Monomers of pyrazine compounds are mainly baking flavor, and some are popcorn flavor, similar to the so-called "scorched-scent ". "Scorched-scent " is not identical to sesame-flavor unless they are in combination with furans(sweet taste), phenols(tarry), thiazole(nut scent), sulfur compounds(onion scent) in proper proportion. Accordingly, sulphur amino acids should be paid more attention excluding increasing nitrogen content of material to highlight the typical sesame-flavor, for their metabolites, sulfur compounds release aroma intensively and everlastingly, and important components of the formation of sesame-flavor. Millet contains a large quantity of sulphur amino acids such as methionine, cysteine and so on.

2.3 Common fermentation by various microorganisms

In combination with high temperature daqu, medium temperature starter, hanoi white qu, aroma-producing yeast and bacteria, and synergistic fermentation have gradually been the unique process characteristics of sesame-flavor liquor brewing, which is the product of combination of traditional technology and modern science.

Hanoi white qu is acid and temperature resistance and suitable at the pH values ranged from 2.5~6.5 which is quite similar to the pH scope of brewing. At such pH values, it can decompose protein into L- amino acids for direct utilization by microorganism. Excessive fermentation ascribed to the reducing sugars produced by exuberant in the process of accumulation at high temperature white qu enzyme family to a great extent, and addition amount should be adjusted.

Compound aroma-producing yeast is the second important microbial strains for sesame-flavor. Candida and Hansen

yeast are temperature resistance, and have good ability to produce esters with strong enzyme system at approximate 45℃. Liquors fermented by Hansen yeast have obvious sauce flavor, and good aftertaste; As for Torulopsis glabrata, strong scorched-scent; candida, strong ester scent; alcohol yeast, sweet taste and little super alcohol taste; lichen yeast and Saccharomyces italicus, produce various kinds of organic acids, esters, alcohols and sulfur methylpropanol etc. During the period of accumulation, large quantity of yeasts propagate and die, and provide excellent protein substrates for Maillard reaction after autolysis.

Thermophilic bacilluses are the cardinal strains for sauce-flavor in the brewing of fragrant type liquor; as for sesame-flavor liquor, they do not play a cardinal role but essential. Monomer culture of these strains could usually produce sesame-flavor which is related to raw material and its protein components. We speculated that effects of proteins in several amylopectin crops on the formation of sesame-flavor should be taken further investigation.

2.4 Accumulation and fermentation at high temperature

Accumulation and fermentation at high temperature as production process of fragrant type liquor, was induced to the produce of sesame-flavor, which played a critical role in forming aroma constituents of sesame-flavor liquor, providing precursor matter and reaction environment for Maillard reaction and affording basic condition for fermentation at high temperature etc. Actually, accumulation at high temperature is the process of koji-making for the second time. As same accumulation, brewing of sesame-flavor and sauce-flavor liquors are either common or rather different in some way. As for the similar required temperature of 45~50℃, sesame-flavor liquor needs only about 10 hours while sauce-flavor 2~3 days. There are reasons that the former relies on aroma-producing yeast and hanoi white qu to grow and propagate while the latter on microorganism enriched from natural environment with the assistance of adding high temperature daqu(bacterial qu) to natural amplification culture. Consequently, resulting in different process of heating, these two manners both incubate lots of yeasts and other microorganisms, and domesticate and eliminate them.

Main flavor components of sesame-flavor liquor are nitrogen-containing pyrazine heterocyclic compounds while those of sauce-favor liquor are oxygen-containing furans heterocyclic compounds. Due to different effects of accumulation, process conditions should be different. Generally, initial temperature is imposed as follows: about 28℃ in Summer and Autumn, 30~32℃ in Spring and Winter. Accumulation was less than 24 hours at 45~50℃ with the height of about 50cm and the appropriate appearance of long strip. When little white strains emerged numerously in the surface layer of accumulation grains, you feel hot once you put your hands in it and bring out strong fruit aroma and accumulation could be transferred into pool. Fermentation grains of accumulation are usually covered with straw matting and Bingsi cloth in order to avoid the contamination of bacteria and keep temperature and humidity, for accumulation doesn't need to enrich exogenous bacteria from the air.

Appropriate accumulation time could avoid the excessive consumption of nutrients and assure double bumpers of protein and starch fermentations. Heating should be regulated in the scope of 10~12℃ according to the rules of "first delay, then elevate and finally decline " during the period of fermentation. The top temperature and time of fermentation is approximate 40℃ and 40 days, respectively.

As far as steaming is concerned, the temperature of liquor flow is very crucial and usually controlled at about 35℃ in favor of the enrichment of sesame-flavor components.

2.5 Layered steaming, subsection brewing and long storage according to the grades and sorts

Due to the unique fermentation pit of sesame-flavor liquor, trace components vary in the different layers of fermentation grains. Ethyl caproate level is high in the bottom due to the effects of pit muds and thus the resultant liquor is rather dense; ethyl acetate level is hight in the middle layer and thus the liquor is quite clear; ethyl lactate level is high and the liquor has strong scorched-scent. Flavor of sections in "early, middle and later stage" in each layer varies greatly, and therefore, according to flavor hallmarks of fermentation grains, steaming is undertaken according to three layers respectively. In each layer, liquor is obtained according to three periods of "early, middle and later stage". Due to the existence of high 3-methylthio-propanol level in sections of later stage, alcohol degree should be paid attention when obtaining liquor, and degree of the last section should be about 54. After over one-month storage, liquor with identical flavor was packaged with jars for long preservation. The storage vessel should be pottery jar which is beneficial to the formation of stale taste of sesame-flavor due to liquor aging. Virginal liquor should be preserved for over 3 years.

According to the production process of sesame-flavor, virginal liquor could be classified into following 5 types: 1, quite clear; 2, typical and complete; 3, rather dense; 4, with sauce-flavor; 5, special blending liquor.

In the blending process, elaborate combination and arrangement and repeated degustation are required to make perfect liquor.

3. Suggestions

Sesame-flavor liquor developed late, and many issues are at the beginning stage. Therefore, we hope that many industries and enterprises could communicate and interact with one another, and we also could discuss with peers and

experts for further investigation.

3.1 Due to the large consumption of qu, complicated process, and lots of uncertain factors, sesame-flavor liquors are produced along with low yield, high cost, large labor quantity and long work hours. Considering excessive consumption of nutrients, in the premise of keeping flavor, increasing yield of liquor is a cardinal point for further investigation.

3.2 Single 3-methylthio-propanol as the standard characteristic brings controversial among many peers and experts. We think that the point is practicability, facility for detecting and perfection of detection method at present. Certainly, we should pay more attention to the investigation of other heterocyclic compounds and make sure of numerous trace characteristic components.

3.3 Since sesame-flavor liquor has countless ties with sauce-flavor liquor and is so popular with consumers, we look forward to the new insight into the healthy factors of sesame-flavor liquor imposed by peers and experts.

References

Hu, G.D. (1992). Studies on the characteristic aroma components of Jinzhi Whit Trunk. *Liang Making*, (1):83~88.

Lai, G.X. (2005). *Handbook of liquor tasting judge*. Beijing: China Alcoholic Drinks Industry Association.

Zhuang, M.Y. (1997). Maillard reaction and fragrant type liquor. *Liquor-Making Science & Technology*, (1):73~77.

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Defense Spending and Income Inequality: Evidence from Selected Asian Countries

M.T. Hirnissa

Department of Economic, Universiti Putra Malaysia

43400 Serdang, Selangor Darul Ehsan, Malaysia

Tel: 60-3-8946-7635 E-mail: nissa_tahir16@yahoo.com

Muzafar Shah Habibullah (Corresponding author)

Department of Economic, Universiti Putra Malaysia

43400 Serdang, Selangor Darul Ehsan, Malaysia

Tel: 60-3-8946-7635 E-mail: muzafar@econ.upm.edu.my

A.H. Baharom

Department of Economic, Universiti Putra Malaysia

43400 Serdang, Selangor Darul Ehsan, Malaysia

Tel: 60-3-8946-7751 E-mail: baharom@econ.upm.edu.my

Abstract

This paper examines the causality between defense spending and income inequality in selected Asian countries namely Malaysia, Indonesia, Singapore, Philippines, India and South Korea for the period 1970-2005. Autoregressive Distributed Lag (ARDL) bounds testing procedure is employed to (1) analyze the impact of defense spending on income inequality and (2) the impact of income inequality on defense spending as well. Interestingly our results indicate one way causality running from defense spending to income inequality only for the case of Malaysia and bidirectional causality for the case of Singapore. As for the remaining countries, no meaningful relationship could be detected and it can be seen as sign of good governance in these countries.

Keywords: Defense spending, Income inequality, Asian, Bounds testing

1. Introduction

Causality relationship between defense spending and income inequality has been subject of interest for many parties; however the lack of availability of information on its statistics and data has been a stumbling block to more researches being conducted. Out of the few studies that have been done, results are often mixed. Ali (2007) made one of the early attempts on a global scale, to identify the relationship between defense spending and income inequality. They treat economic growth as a control variable rather than a dependent variable and emphasize on the impact of defense spending on income inequality only. In this study we went a step ahead by treating income inequality, as both regressor (control variable) and regresand (dependant).

Theoretically it is believed that there are number of ways by which defense spending may be cointegrated with income inequality: (1) Any increase in defense spending could be at the expense of public spending on social programs such as health and education which in turn will have an equalizing effect, (2) The taxes required to support military spending may fall disproportionately on the middle classes; if so, post-tax income inequality might be at a risk of increasing. (3) High levels of military spending may reflect the use of violence as a means of social control, notably against trade unions and other egalitarian social forces thus, it is not surprising to witness that higher military spending means more societal control and a sacrifice of egalitarian values.

On the other hand, looking at it from another perspective (4) military experience may cut in the other direction. The military absorbs low-skilled labor, which may raise wages for the young and unskilled. Mobilization for war may require equalizing concession to labor's interests. In general, the more equipment-intensive defense spending, the more we expect the income inequality-increasing effects to dominate; the more labor-intensive the military and home grown the military production, the more we might expect to find inequality-reduction effects in the data. It can even be (5) no-cointegration at all, when there are good governance, respective governments carefully planning their policies and budget, so that defense spending would not stand in the way of spending on other important aspects, such as education, health, public amenities etc. Caputo (1975) was one of the earlier studies on public policy implications of military and welfare expenditures. The subject became more popular and much more researches were conducted, however most of these researches were centered around defense spending and economic growth, such as to name a few, Hassan et al. (2003), Al-Yousif (2002), Shieh et al. (2002), and Kollias et al (2004a and 2004b). As for the defense spending and income inequality, as mentioned above, Ali (2007) was one of the few papers other than Boswell and Dixon (1990), Auvinen and Nafziger (1999), and Jorgensen (2005)

2. Trend of Defense Spending and Income Inequality in Asian Countries

Defense spending and income inequality has been an important component in economy. Figure 1 displays the trend of defense spending in six selected Asian countries; Indonesia, Malaysia, Philippines, Singapore, India and South Korea. It can clearly be seen that, the volatility is quite high for almost all the selected countries for the period 1970 to 1988, however, it stabilizes after 1988.

As for Figure 2, Singapore and South Korea show declining pattern in income inequality (better income distribution) for the period 1972 to 1997, while Malaysia, quite the contrary, shows an increasing (worsening income distribution) for the period 1982 to 1990. While for the case of Indonesia, there are fluctuations in income inequality pattern from 1974 to 1990 and declining after that and finally, the Philippines show an increasing trend.

Figure 3 show the defense spending as a percentage of gross domestic products in these six countries for three different times, albeit, 1970, 1990 and 2006. As can bee seen for all three different point of time, Singapore is the highest spender in terms of ratio to GDP. Malaysia was second highest in 1970, dropped to fourth among these six countries in 1990 and remained fourth in 2006 as well. Indonesia ranked fifth in all three points of time, similar to the Philippines who ranked sixth in all. South Korea ranked third in 1970, climbed to second in 1990 and dropped back to third in 2006.And finally India ranked fourth in 170, climbed to third in 1990 and remained there for 2006.

3. Review of Related Literature

Ali (2007) examines the effect of military spending on income inequality for the period 1987-1997, controlling for the size of armed forces, GDP growth, per capita income and other possible determinants. Their hypothesis is that as per capita defense spending increases, income inequality increase, controlling for the size of armed forces, and for regional and economic variables. They found consistent estimates that there is positive effect of defense spending on income inequality and it is robust across variable definitions and model specifications. Given the close relationship, this result suggests that an increase in the defense spending's of a country will worsen the income distribution (increase the income inequality). The same results were shared by Jorgensen (2005), Auvinen and Nafziger (1999), Auvinen and Nafziger (2002), Jayadev and Bowles (2006) but was contrary to Henderson et al. (2008)

Auvinen and Nafziger (1999) explained that there is a high correlation between high ratio of defense spendings to income and high income inequality in 124 less developed countries (LDCs) for the period 1980-1995, using various causality regressions, and ultimately this can turn into source of humanitarian emergency, a view that was supported by their following paper, Auvinen and Nafziger (2002) in their study on developing countries. Jayadev and Bowles (2006), in their study on participation in Guard Labor in the United States based on empirical data from even 1890s, using classical model on power and growth, claimed that these people could have been employed in other productive sectors, and by serving in the less productive sector (Guard Labor), it contributed to a higher income inequality (worsening income distribution). However the finding of Henderson et al. (2008) was on the contrary, in their study on the transition countries of Eastern Europe and Central Asia, they found that these countries during their transition, with a cut budget on their defense spending still turned out worse off, with a higher income inequality. They then suggested that there could be elements of hidden income inequality in these countries in their past history.

4. Methodology

4.1 ARDL Approach to Causality Test

In order to test for causality between defense spending and economic growth we utilized the autoregressive distributed lag model (ARDL) popularize by Pesaran et al. (2001). The ARDL has numerous advantages. Firstly, the ARDL approach is able to examine the presence of short run as well as long run relationship between the independent variables and the dependent variable. Secondly, the ARDL model takes a sufficient numbers of lags to capture the data generating process in a general to specific modeling framework (Laurenceson and Chai, 2003). Apart from that, unrestricted

error-correction model (UECM) is likely to have better statistical properties than the two-step Engle-Granger method because, unlike the Engle-Granger method, the UECM does not push the short –run dynamics into the residual term (Banerjee et al., 1998). Finally, the ARDL approach provides robust result in a small sample size. Since the sample size of our study is small, this provides more motivation for this study to adopt this model.

The ARDL unrestricted error correction model (UECM) is shown below:

$$\Delta LDS_t = \alpha_0 + \alpha_1 LDS_{t-1} + \alpha_2 LI_{t-1} + \sum_{i=1}^{m} \alpha_{3,i} \Delta LDS_{t-i} + \sum_{i=1}^{m} \alpha_{4,i} \Delta LI_{t-i} + \varepsilon_t \quad (1)$$

$$\Delta LI_t = \beta_0 + \beta_1 LI_{t-1} + \beta_2 LDS_{t-1} + \sum_{i=1}^{m} \beta_{3,i} \Delta LDS_{t-i} + \sum_{i=1}^{m} \beta_{4,i} \Delta LI_{t-i} + \mu_t \quad (2)$$

whereby DS is the ratio of defense spending to GDP, I is income inequality, Δ is the first difference operator, L denote variables in logarithm and ε_t and μ_t are serially independent random errors.

To examine the long- run relationship, the bound cointegration test based on *F*-statistic taken from Narayan and Narayan, (2005) will be used. The null hypothesis for no cointegration among the variables in Eq. (1) is ($Ho: \alpha_1 = \alpha_2 = 0$) denoted by F_{MILEX} against the alternative ($H_1: \alpha_1 \neq \alpha_2 \neq 0$). Similarly, for Eq. (2) the null hypothesis for no long-run meaningful relationship among the variables is ($Ho: \beta_1 = \beta_2 = 0$) as denoted by F_I against the alternative ($H_1: \beta_1 \neq \beta_2 \neq 0$).

The two asymptotic critical values bound provide a test for cointegration when the independent variables are *I(d)* (where 0≤ d ≤ 1): a lower value assuming the regressors are *I(0)*, and an upper value assuming purely *I(1)* regressors. If the test statistic exceed the upper critical value, we can conclude that a long – run relationship exist regardless of whether the underlying order of integration of variable are zero or one. If the test statistics fall below the lower critical values we cannot reject the null hypothesis of no cointegration. However, if the statistic fall between these two bound, inference would be inconclusive.

4.2 Description and sources of data

The data used in this study are annual data on defense spending and income inequality for the selected Asian countries. The countries are Malaysia, Indonesia, Philippine, Singapore, India and Korea. DS is measure by the defense spending as a percentage of GDP. This data was obtained from various issues of SIPRI Yearbook and SIPRI online database. Meanwhile the data for the income inequality, for the corresponding period was obtained from University of Texas, which is estimates of gross household income inequality, computed from a regression relationship between the Deininger and Squire Inequality measures and the UTIP-UNIDO pay inequality measures. All the data used in the study were transformed into logarithm.

5. Empirical results

We tested for the order of integration for defense spending and income inequality before proceeding to testing for cointegration by using the ARDL bounds testing procedure. Table 1(A and B) show the results of the unit root test for the test of the order of integration of the economic time series under investigation. Clearly the augmented Dickey-Fuller test (Dickey and Fuller, 1981) statistics indicate that both the defense spending and income inequality economic series in selected Asian countries are stationary after first differencing (*I*(1)) thus our relevant critical values are the upper bound of purely *I*(1) regressors. These results are tabulated in Table 2 (Panel A and Panel B). Whereby in Panel A, the dependent variable is income inequality and in Panel B, the dependent variable is defense spending. It can be summarized that there seems to be unidirectional causality from defense spending to income inequality in Malaysia while for the case of Singapore there seems to be bidirectional causality. As for the other countries, the null hypothesis of no cointegration cannot be rejected in all the cases (Panel A and Panel B); these results suggest that there are no long-run relationships between defense spending and income inequality in these countries namely, India, South Korea, Thailand and Philippines.

Table 3 Panel A and Panel B) display the long run coefficients results. For both Malaysia and Singapore case, it is positively significant; any increase in defense spending will increase income inequality (worsening income distribution) as for panel B (defense spending as a dependant variable) Singapore's income inequality is also positively related with defense spending.

Figure 4 display the results of the impulse response of counties, based on VECM for Malaysia and Singapore, while for the remaining countries based on VAR, and again the results are robust. It clearly shows that any shock in the defense spending does not constitute any shocks to income inequality *vice versa* for India, South Korea, Thailand and Philippines. On the other hand, any shock to defense spending does causes shock to income inequality for Malaysia and for Singapore it is both way.

As for variance decomposition, the results shown in Table 4 to Table 9 are similar to prior finding whereby showing the same pattern of results, there are no meaningful relationship between these variables (defense spending and income inequality) for India, South Korea, Thailand and Philippines (in fact percentage changes that contributed to the other

variable is too small and it stabilizes after a few periods). While for Malaysia and Singapore the results are similar to ARDL and IRF. These results are very consistent in nature.

6. Conclusion

In this study the autoregressive distributed lag (ARDL) bounds testing procedure was employed to investigate the long-run relationship between defense spending and income inequality in six selected Asian countries, namely Malaysia, Singapore, Thailand, Philippine, South Korea and India. A bivariate analysis on the impact of income inequality on defense spending, *vice versa* the impact of defense spending on income inequality was conducted. The sample period was 1970 – 2005 and the data was annual. All the data went through log-log transformation so that the estimates will be less sensitive to outliers or influential observations and also in order to reduce the data range.

The results suggest that all the variables chosen are *I*(1) or in other words they are non-stationary variables and achieved stationarity only after first differencing. The cointegration analysis using the ARDL bounds testing approach clearly indicates that only in the case of Malaysia and Singapore, the military spending are cointegrated with income inequality. Though the results are interesting, not much comparison could be made because not many researches done on this issue, even the few researches made, they normally treat income inequality as the dependant variable only as in the case of Ali (2007).

However our results for the case of Malaysia and Singapore are concurrent with his finding, whereby any increase in Defense spending will worsen of income distribution (higher income inequality. as also supported by Caputo (1975) who explained that there is a trade off between defense and welfare expenditure. Another paper with similar result is of Jayadev and Bowles (2006), however their argument is different, they claimed that being in the lower productivity sector (Guard Labor) deprives the nation of their contribution in other higher productivity sectors, thus worsening income distribution resulting higher income inequality. And as for the remaining countries, no trace of cointegration among these variables can be concluded as a sign of good governance and good policy making, whereby the decisions of defense spending is independent and does not have any whatsoever impact on income distribution.

References

Al-Yousif, Y.K. (2002). Defense spending and economic growth: Some empirical evidence from the Arab Gulf region, *Defence and Peace Economics,* 13(3), 187-197.

Ali, H.E. (2007) Military expenditure and inequality: Empirical evidence from global data, D*efense and Peace,* 18(6), 519-535.

Auvinen, J. & Nafziger, E.W. (1999). The sources of humanitarian emergencies. *Journal of Conflict Resolution,* 43(3), 267-290.

Auvinen, J. & Nafziger, E.W. (2002). Economic development, inequality, war, and state violence. *World Development,* 30(2), 153-163.

Banerjee, A., Dolado, J. & Mestre, R. (1998) Error-correction mechanism tests for cointegration in a single equation framework. *Journal of Time Series Analysis,* 19, 267-283.

Boswell, T. & Dixon, W.J. (1990). Dependency and rebellion: A cross-national analysis. *American Sociological Review,* 55(4), 540-559.

Caputo, D.A. (1975). New perspectives on the public policy implications of defense and welfare expenditures in four modern democracies: 1950-1970. P*olicy Sciences,* 6, 423-446.

Dickey, D. & Fuller, W.A. (1981). Likehood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49, 1057-1072.

Hassan, M.K, Waheeduzzaman, M. & Rahman, A. (2003). Defense expenditure and economic growth in the SAARC countries. *The Journal of Social, Political and Economic Studies,* 28(3), 275-282.

Henderson, D.R., McNab, R.M. & Rozsas, T. (2008). Did inequality increase in transition: An analysis of the transition countries of Eastern Europe and Central Asia, *Eastern European Economics,* 46 (2), 28-49.

Jayadev, A. & Bowles, S (2006). guard Labor. *Journal of Development Economics*, 79, 328– 348.

Jorgensen, A.K. (2005). Unpacking international power and the ecological footprints of nations: A quantitative cross-national study. *Sociological Perspectives,* 48(3), 383-402.

Kollias, C., Naxakis, C. & Zarangas, L. (2004a). Defence spending and growth in Cyprus: a causal analysis, *Defence and Peace Economics*, 15(3), 299-307.

Kollias, C., Manolas, G. & Paleologou, S. Z. (2004b). Defence expenditure and economic growth in the European Union: a causality analysis, *Journal of Policy Modeling*, 26, 553-569.

Laurenceson, J. & Chai, J. C. H. (2003). *Financial reform and economic development in China*. Chelthenham, Edward Elgar.

Narayan, P.K. & Narayan, S. (2005). Estimating income and price elasticities of imports for Fiji in a cointegration framework. *Economic Modelling,* 22, 423-438.

Pesaran, M.H., Shin, Y. & Smith, R.J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics,* 16, 289-326.

Shieh, J.Y, Lai, C.C & Chang, W.Y (2002). Endogenous growth and defense expenditures: A new explanation of the Benoit hypothesis, *Defence and Peace Economics,* 13(3), 179-186.

SIPRI. SIPRI Yearbook 1975, 1977, 1985, 1990, 1999 and 2006. Oxford: Stockholm International Peace research Institute, Oxford University Press. Stockholm International Peace Research Institute (SIPRI) database.

Table 1A. Results of Unit Root Test for Series in Level

Asian	LI		LDS	
	ADF *t*-statistic	Lag	ADF *t*-statistic	Lag
Indonesia	-2.485	0	-2.593	2
	[0.33]		[0.28]	
Malaysia	-2.174	1	-2.360	0
	[0.48]		[0.39]	
Philippine	-2.971	0	-1.887	1
	[0.15]		[0.63]	
Singapore	-1.835	1	-3.309	1
	[0.66]		[0.08]	
India	-1.651	0	-1.972	0
	[0.75]		[0.59]	
Korea	-1.754	0	-0.981	0
	[0.70]		[0.93]	

Notes: Asterisk (*) denotes statistically significant at 5% level.

Table 1B. Results of Unit Root Test for Series in First Difference

Asian	LI		LDS	
	ADF *t*-statistic	Lag	ADF *t*-statistic	Lag
Indonesia	-5.874 [0.00]*	0	-5.021 [0.00]*	0
Malaysia	-3.808 [0.00]*	0	-5.097 [0.00]*	0
Philippine	-7.474 [0.00]*	0	-4.140 [0.00]*	1
Singapore	-3.912 [0.00]*	1	-4.466 [0.00]*	1
India	-5.211 [0.00]*	0	-4.833 [0.00]*	0
Korea	-7.399 [0.00]*	0	-5.941 [0.00]*	0

Notes: Asterisk (*) denotes statistically significant at 5% level

Table 2. Bounds Test for Cointegration Analysis Based on the Equation 1 and Equation 2

Panel A

Dependent variable LI, Independent variable LDS

n	Critical value	Lower Bound Value	Upper Bound Value
30	5%	4.090	4.663
35	5%	3.957	4.530

Computed *F*- statistic

Countries	*F*-Statistic
Indonesia	3.2073
Malaysia	8.1759*
Philippines	1.2587
Singapore	4.5901*
India	3.2941
Korea	0.6370

Panel B

Dependent variable LDS, Independent variable LI

n	Critical value	Lower Bound Value	Upper Bound Value
30	5%	4.090	4.663
35	5%	3.957	4.530

Computed *F*- statistic

Countries	*F*-Statistic
Indonesia	1.6459
Malaysia	0.4302
Philippines	1.6126
Singapore	5.4879*
India	3.0022
Korea	3.7224

Notes: Asterisk (*) denotes statistically significant at 5% level.

Table 3. Long – run coefficient

Panel A		
Dependent : LI Independent: LDS	Coefficient	*t*-statistic
Malaysia	0.1516*	2.8874
Singapore	0.3299	2.0727

Notes: Asterisk (*) denotes statistically significant at 5% level.

Panel B		
Dependent : LDS Independent: LI	Coefficient	*t*-statistic
Singapore	1.2251*	3.1538

Notes: Asterisk (*) denotes statistically significant at 5% level.

Table 4. Variance Decomposition for Indonesia

Variance Decomposition of LI:				Variance Decomposition of LDS:		
Period	S.E.	LI	LDS	S.E.	LI	LDS
1	0.022925	100	0	0.165022	2.653384	97.34662
		0	0		-6.82756	-6.82756
2	0.027867	99.10883	0.891174	0.227362	7.580501	92.4195
		-5.26673	-5.26673		-11.842	-11.842
3	0.031397	93.88101	6.118991	0.271018	12.07511	87.92489
		-11.0515	-11.0515		-13.9801	-13.9801
4	0.034703	86.87199	13.12801	0.307294	15.82857	84.17143
		-16.2466	-16.2466		-15.456	-15.456
5	0.037922	80.13026	19.86974	0.339722	18.84457	81.15543
		-19.2965	-19.2965		-16.904	-16.904
6	0.041052	74.34026	25.65974	0.369778	21.24321	78.75679
		-20.9493	-20.9493		-18.3616	-18.3616
7	0.044088	69.55419	30.44581	0.398207	23.15797	76.84203
		-21.8963	-21.8963		-19.6217	-19.6217
8	0.047034	65.62939	34.37061	0.425443	24.70219	75.29781
		-22.5405	-22.5405		-20.6567	-20.6567
9	0.049901	62.39588	37.60412	0.451764	25.96374	74.03626
		-23.0639	-23.0639		-21.5038	-21.5038
10	0.052698	59.70614	40.29386	0.477362	27.00834	72.99166
		-23.5083	-23.5083		-22.1952	-22.1952

Notes: Cholesky Ordering: LI LDS, Standard Errors: Monte Carlo (100 repetitions)

Table 5. Variance Decomposition for Malaysia

	Variance Decomposition of LDS:			Variance Decomposition of LI:		
Period	S.E.	LDS	LI	S.E.	LDS	LI
1	0.266557	95.64534	4.354663	0.014783	0	100
2	0.377034	96.01499	3.985006	0.017077	0.273237	99.72676
3	0.44892	89.52627	10.47373	0.018914	11.43213	88.56787
4	0.491782	89.77204	10.22796	0.022343	35.85366	64.14634
5	0.520347	90.43182	9.568181	0.030903	64.65095	35.34905
6	0.541778	90.87705	9.122947	0.040849	77.10891	22.89109
7	0.552736	91.22093	8.779066	0.049864	81.27815	18.72185
8	0.561965	91.49932	8.500684	0.057151	83.53265	16.46735
9	0.573034	91.82453	8.175469	0.062367	85.01956	14.98044
10	0.585868	92.14817	7.85183	0.066048	86.07387	13.92613

Notes: Cholesky Ordering: LI LDS, Standard Errors: Monte Carlo (100 repetitions)

Table 6. Variance Decomposition for Philippines

	Variance Decomposition of LI:			Variance Decomposition of LDS:		
Period	S.E.	LI	LDS	S.E.	LI	LDS
1	0.023289	100	0	0.145828	1.057613	98.94239
		0	0		-4.47486	-4.47486
2	0.026771	99.56198	0.438021	0.217376	0.532303	99.4677
		-3.38495	-3.38495		-4.91326	-4.91326
3	0.028245	97.12752	2.872483	0.254707	3.255674	96.74433
		-5.44155	-5.44155		-8.26713	-8.26713
4	0.0293	93.48713	6.512867	0.278368	9.86498	90.13502
		-8.44007	-8.44007		-12.6839	-12.6839
5	0.030191	90.31968	9.680322	0.295458	16.35462	83.64538
		-11.2921	-11.2921		-15.5529	-15.5529
6	0.030915	88.22972	11.77028	0.307406	20.77633	79.22367
		-12.7343	-12.7343		-17.1077	-17.1077
7	0.031485	87.00087	12.99913	0.315548	23.35761	76.64239
		-13.6331	-13.6331		-18.0559	-18.0559
8	0.031923	86.26725	13.73275	0.321304	24.82582	75.17418
		-14.4606	-14.4606		-18.7374	-18.7374
9	0.032254	85.78099	14.21901	0.3256	25.72254	74.27746
		-15.2145	-15.2145		-19.2824	-19.2824
10	0.032503	85.42064	14.57936	0.328911	26.34074	73.65926
		-15.8594	-15.8594		-19.7411	-19.7411

Notes: Cholesky Ordering: LI LDS, Standard Errors: Monte Carlo (100 repetitions)

Table 7. Variance Decomposition for Singapore

Variance Decomposition of LI:				Variance Decomposition of LDS:		
Period	S.E.	LI	LDS	S.E.	LI	LDS
1	0.01596	100	0	0.094554	5.824267	94.17573
2	0.031067	98.26474	1.735263	0.128454	31.46932	68.53068
3	0.042983	96.65819	3.341806	0.14217	42.28278	57.71722
4	0.050975	96.7781	3.221896	0.147478	41.43657	58.56343
5	0.056207	97.28511	2.714888	0.148118	41.92341	58.07659
6	0.060201	97.61928	2.380721	0.153195	43.96428	56.03572
7	0.064082	97.77378	2.226215	0.161674	48.69535	51.30465
8	0.068315	97.75849	2.241512	0.16817	52.55435	47.44565
9	0.072621	97.77204	2.227961	0.171954	54.601	45.399
10	0.076588	97.86095	2.139051	0.174805	56.06379	43.93621

Cholesky Ordering: LI LDS

Variance Decomposition of LI:				Variance Decomposition of LDS:		
Period	S.E.	LI	LDS	S.E.	LI	LDS
1	0.01596	94.17573	5.824267	0.094554	0	100
2	0.031067	87.41166	12.58834	0.128454	17.25636	82.74364
3	0.042983	83.56941	16.43059	0.14217	32.35628	67.64372
4	0.050975	83.60554	16.39446	0.147478	33.89552	66.10448
5	0.056207	84.9227	15.0773	0.148118	34.45745	65.54255
6	0.060201	85.89814	14.10186	0.153195	35.45686	64.54314
7	0.064082	86.19503	13.80497	0.161674	39.16031	60.83969
8	0.068315	86.05097	13.94903	0.16817	43.08837	56.91163
9	0.072621	85.99287	14.00713	0.171954	45.42669	54.57331
10	0.076588	86.15368	13.84632	0.174805	46.93114	53.06886

Notes: Cholesky Ordering: LI LDS, Standard Errors: Monte Carlo (100 repetitions)

Table 8. Variance Decomposition for India

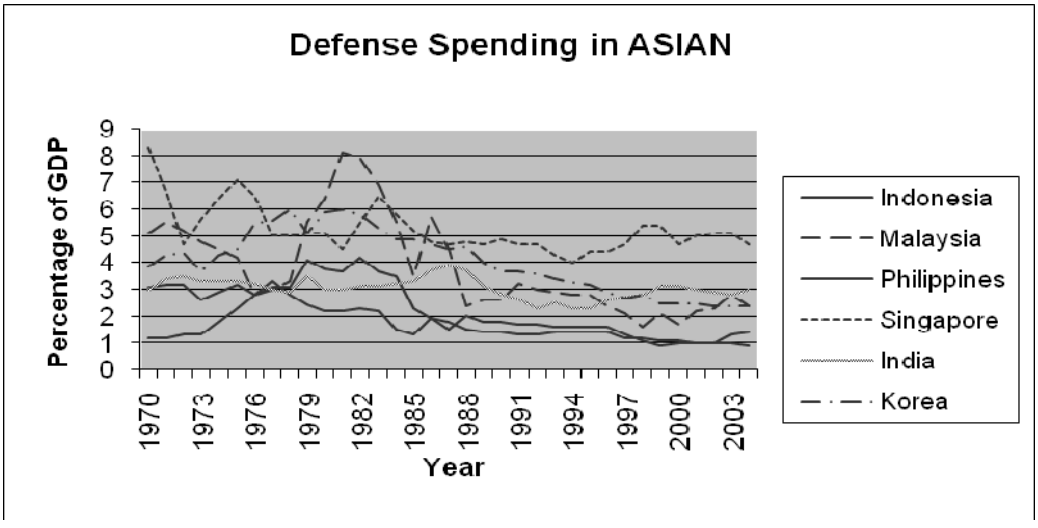
Variance Decomposition of LI:				Variance Decomposition of LDS:		
Period	S.E.	LI	LDS	S.E.	LI	LDS
1	0.01017	100	0	0.075708	15.45394	84.54606
		0	0		-11.3684	-11.3684
2	0.014074	96.44746	3.552535	0.106976	10.31343	89.68657
		-5.45812	-5.45812		-10.7977	-10.7977
3	0.016629	93.68182	6.318178	0.123263	8.319606	91.68039
		-8.69645	-8.69645		-11.0174	-11.0174
4	0.018296	92.53814	7.461861	0.130617	7.556469	92.44353
		-11.1295	-11.1295		-11.3122	-11.3122
5	0.019361	92.37748	7.622519	0.133372	7.304724	92.69528
		-12.6125	-12.6125		-11.6441	-11.6441
6	0.020041	92.59423	7.40577	0.134181	7.253678	92.74632
		-13.5325	-13.5325		-11.9032	-11.9032
7	0.020479	92.85893	7.141069	0.134351	7.268909	92.73109
		-14.264	-14.264		-12.077	-12.077
8	0.020766	93.05509	6.944907	0.134383	7.299587	92.70041
		-14.9908	-14.9908		-12.1927	-12.1927
9	0.020958	93.17355	6.826446	0.134408	7.330241	92.66976
		-15.6841	-15.6841		-12.2812	-12.2812
10	0.021087	93.2387	6.761301	0.134438	7.357096	92.6429
		-16.267	-16.267		-12.3509	-12.3509

Notes: Cholesky Ordering: LI LDS, Standard Errors: Monte Carlo (100 repetitions)

Table 9. Variance Decomposition for Korea

Variance Decomposition of LI:				Variance Decomposition of LDS:		
Period	S.E.	LI	LDS	S.E.	LI	LDS
1	0.019765	100	0	0.081054	3.767607	96.23239
		0	0		-7.55895	-7.55895
2	0.023102	99.76627	0.23373	0.102724	2.377794	97.62221
		-4.28299	-4.28299		-7.08835	-7.08835
3	0.026931	98.76076	1.239241	0.1224	4.198122	95.80188
		-3.99355	-3.99355		-8.19928	-8.19928
4	0.029722	97.23416	2.765838	0.138483	5.705997	94.294
		-5.29946	-5.29946		-9.23975	-9.23975
5	0.032179	95.0728	4.927203	0.152848	8.241795	91.7582
		-6.53097	-6.53097		-10.9795	-10.9795
6	0.034287	92.38633	7.613674	0.165664	11.11894	88.88106
		-8.78309	-8.78309		-12.7173	-12.7173
7	0.036155	89.23916	10.76084	0.177299	14.36836	85.63164
		-11.0019	-11.0019		-14.5966	-14.5966
8	0.037831	85.73493	14.26507	0.187899	17.82596	82.17404
		-13.4092	-13.4092		-16.2771	-16.2771
9	0.039358	81.98362	18.01638	0.197593	21.40274	78.59726
		-15.5849	-15.5849		-17.7794	-17.7794
10	0.040769	78.10419	21.89581	0.206465	25.00394	74.99606
		-17.582	-17.582		-18.9809	-18.9809

Notes: Cholesky Ordering: LI LDS, Standard Errors: Monte Carlo (100 repetitions

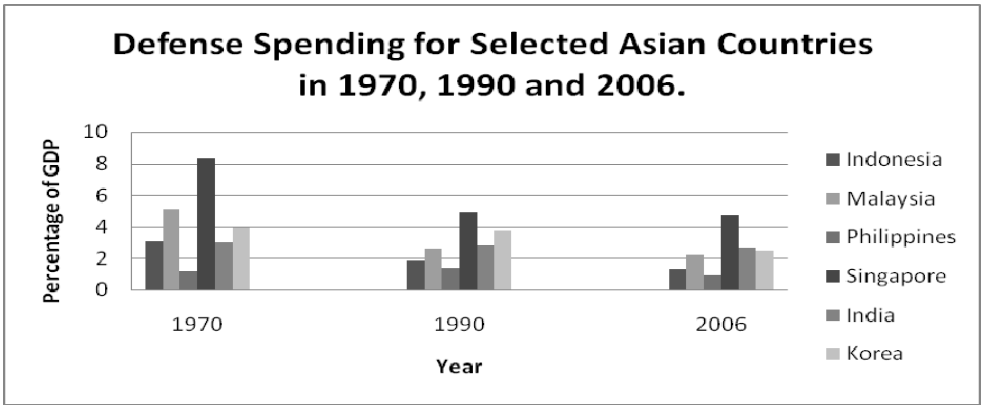


Sources: SIPRI yearbook, various issues

Figure 1. Defense spending in Asian countries

Sources: UTIP-UNIDO

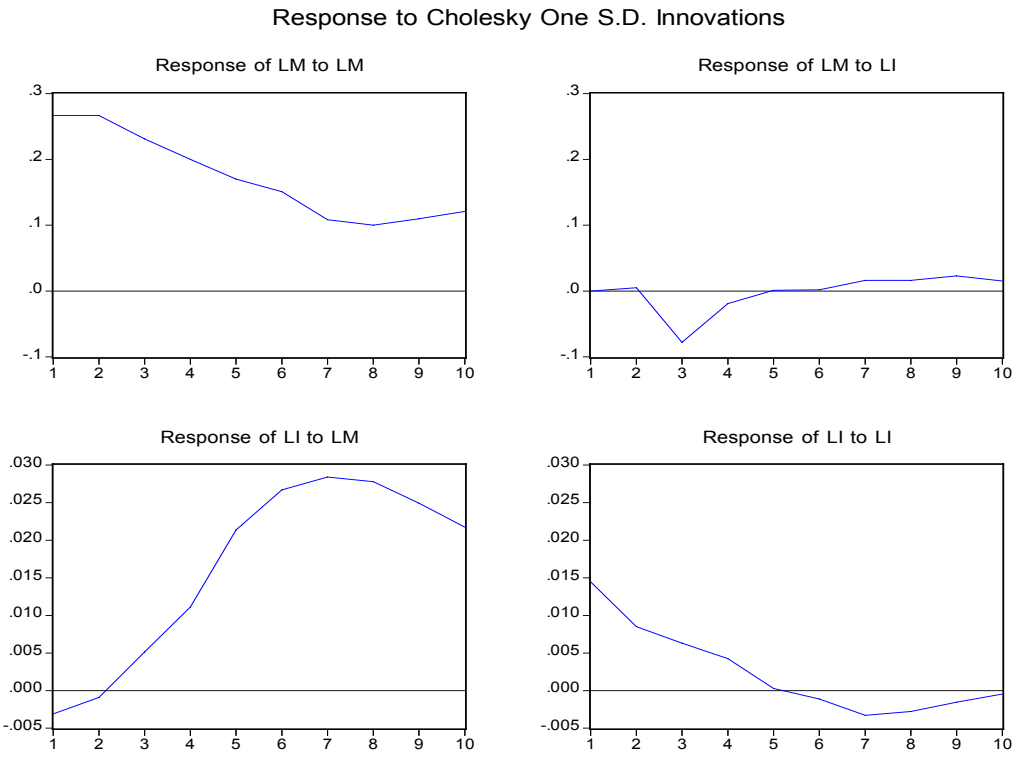
Figure 2. Income inequality in Asian countries



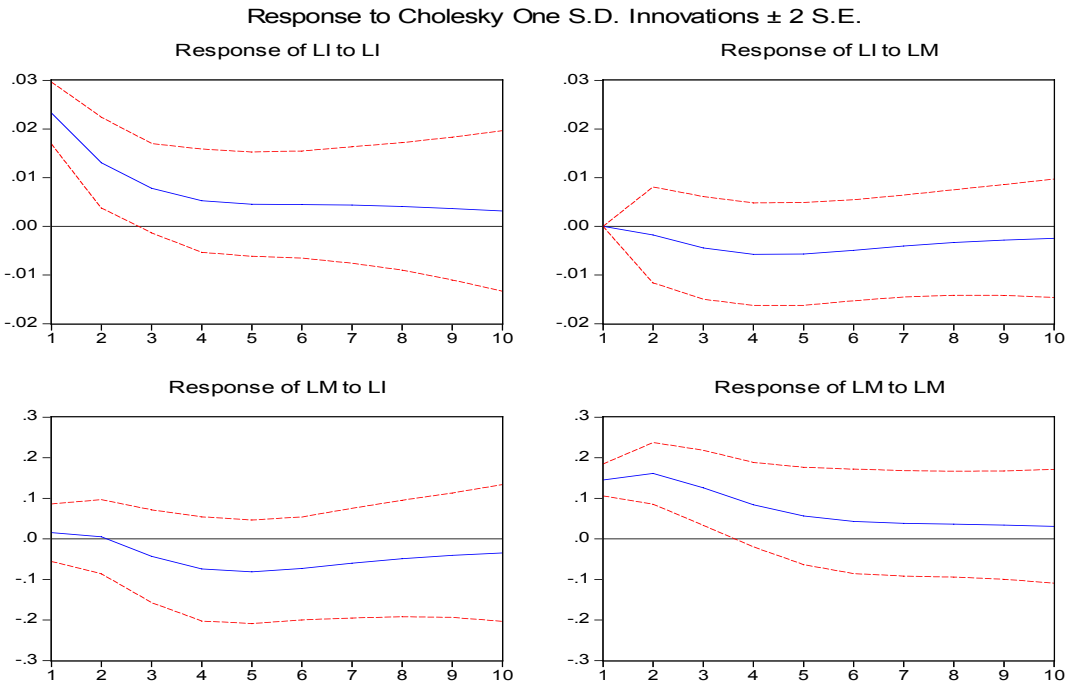
Sources: SIPRI yearbook, various issues

Figure 3. Defense spending for Selected Asian Countries in 1970, 1990 and 2006

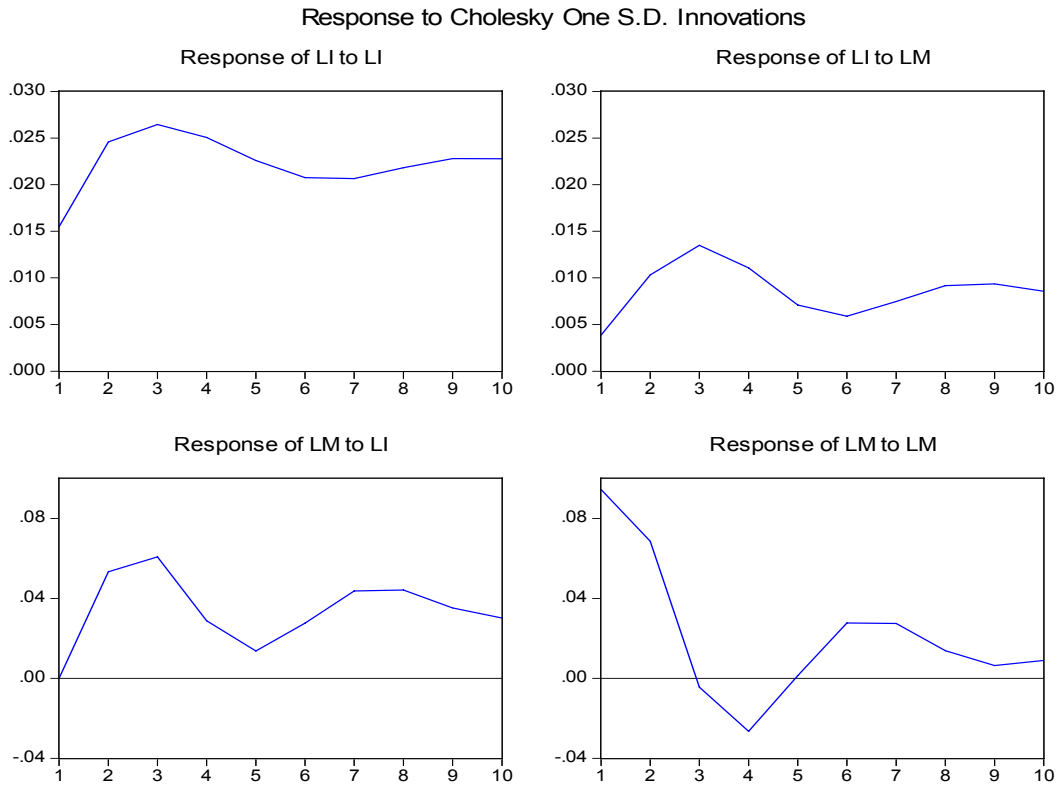
Malaysia



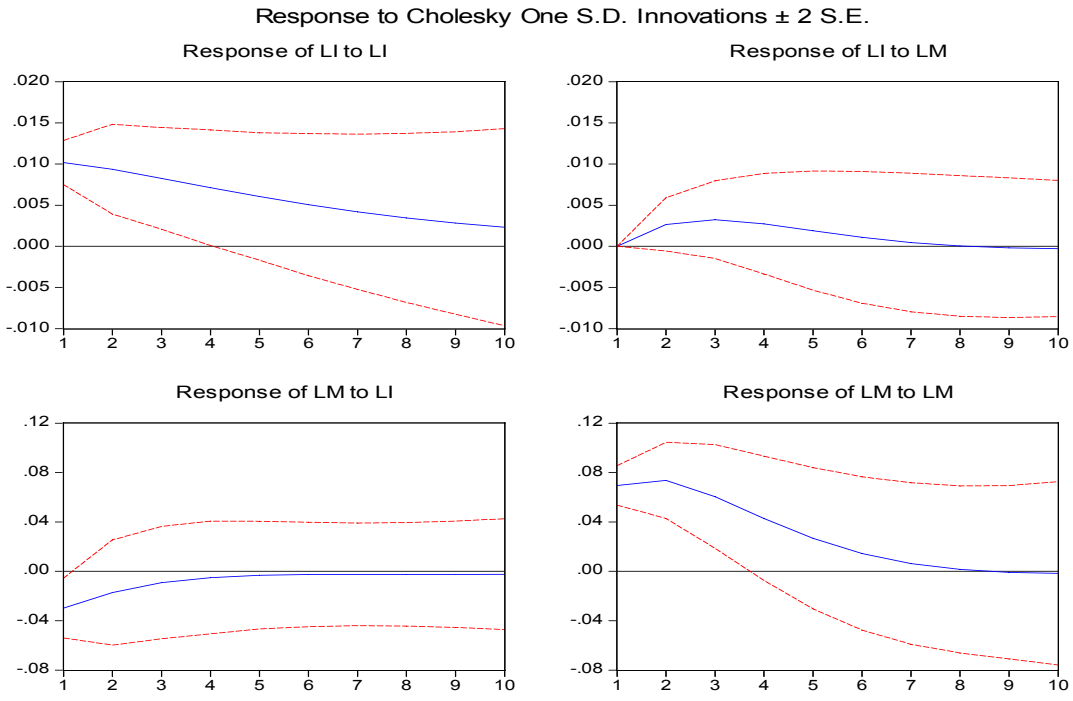
Philippines



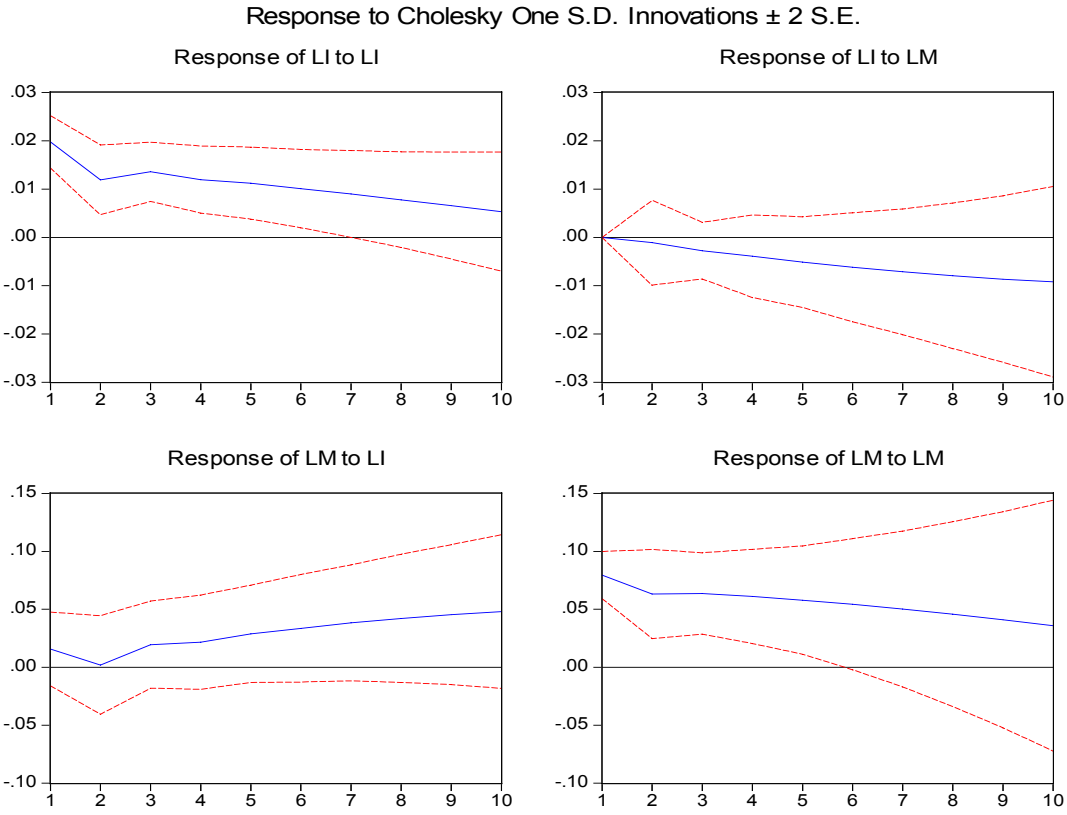
Singapore



India



Korea



Notes: LM denotes defense spending. LI denotes income inequality.

Figure 4. The Results of Impulse Response for Asian Countries

Modern Applied Science

www.ccsenet.org/journal.html

Vol. 3, No. 5

May 2009

Study on the Simulation of the Soft Start of the Asynchronous Motor with Pump Control Function Based on Thyristor

Tingjian Zhong, Li Zhu & Minghua Zhou

Jiangxi Vocational & Technical College of Electricity, Nanchang 330032, China

E-mail: jxdlztj@163.com

Yongtao Dai

School of Information Engineering, Nanchang University, Nanchang 330029, China

E-mail: dytnc2007@126.com

Abstract

To reduce the impact of the high current on the power grid from the start process of the asynchronous motor and eliminate the negative influences of traditional reduced voltage start on the electronic and mechanic equipments and enhance the start character of the motor, we studied and improved the soft start of the motor, and put forward the soft starter with the function of pump control which could reduce the water hammers induced in the start and stop of the pump, reduce the oscillation of the pipeline system and perfect the functions of the soft starter. In the article, we used MATLAB to establish simulation module of the three phase current voltage regulation circuit system, and tested the soft starter designed in the article, and the simulation result indicated that the soft starter with pump control function could not only effectively reduce the starting impact of the motor to the power grid, but reduce the impact from the water hammer and optimize the start performance of the motor.

Keywords: Soft starter, Pump control, Torque, Thyristor

1. Introduction

The traditional start mode of motor can be divided into direct start and reduced voltage start which can be divided into Y-Δ start, auto-coupling reduced voltage start, stator string reactor start and so on, and because traditional start modes are hard to adjust the motor parameters exactly to fulfill the smooth start according to certain requirement in the start process, and induce certain impact to the power grid (Li, 2006, 185-187 & Zhang, 1999, P. 3-5), we begin to study the soft start of the asynchronous motor. The general soft starters all have the voltage ramp start mode and limited current start mode (Xu, 2001, P. 41-44), and the soft starter designed in the article takes the thyristor as the execution component, and emits PWM wave to control the thyristor to trigger the pulse by the control unit, and realize the control of the motor start and control the connection of the thyristor. The voltage regulation circuit of thyristor has very important meaning for the implementation of the functions of the soft starter, and we mainly utilize the voltage regulation principle and voltage regulation circuit of the thyristor to study the characteristics of the soft start.

2. Mathematical model of asynchronous motor

To study the relationships among various variables such as the voltage, current and torque in the start and stop of the asynchronous motor, we give the mathematical model of the asynchronous motor here. When we study the frequency conversion speed regulation, we mainly adopt the mathematical model based on the state equation, and for the soft start, the mathematical model of the asynchronous motor based on centralized parameters equivalent circuit is generally adopted (Wang, 1990 & Gao, 1993) (seen in Figure 1), where, U_1 is the phase voltage virtual value of the power grid, r_m and x_m respectively are the excitation resistance and the excitation reactance, r_1 and $x_{1\sigma}$ respectively are the resistance and the leakage reactance of the stator, $r_2^{'}$ and $x_{2\sigma}^{'}$ respectively are the resistance reduced value and the leakage reactance reduced value of the rotor, and s is the slip ratio.

From the principle of the motor theory, the mechanical character equation expression of three phase asynchronous motor is

$$T_{em}=\frac{P_{em}}{\Omega_1}=\frac{3pU_1^2\frac{r_2^{'}}{s}}{2\pi f_1[(r_1+\frac{r_2^{'}}{s})^2+(x_{1\sigma}+x_{2\sigma}^{'})^2]} \tag{1}$$

When the motor just starts, the rev of rotor is $n_2=0$, and the slip ratio is $s=1$, and the mechanical character equation of the motor here is

$$T_{st}=\frac{3pU_1^2r_2^{'}}{2\pi f_1[(r_1+r_2^{'})^2+(x_{1\sigma}+x_{2\sigma}^{'})^2]} \tag{2}$$

Because $\left|(r_1+r_2^{'})+j(x_{1\sigma}+x_{2\sigma}^{'})\right|$ is much less than $\left|(r_1+r_m)+j(x_{1\sigma}+x_m)\right|$, so when the motor is started,

$$Z_k\approx\sqrt{(r_1+r_2^{'})^2+(x_{1\sigma}+x_{2\sigma}^{'})^2} \tag{3}$$

$$I_1\approx\frac{U_1}{Z_k} \tag{4}$$

Because the currents of the rotor and the stator are much bigger than the rated currents when starting the motor, the leakage reactance of the rotor and the stator makes the iron magnetic part in the leakage magnetic circuit saturated, and the leakage magnetic resistance becomes bigger, which makes $x_{1\sigma}$ and $x_{2\sigma}^{'}$ diminished, Z_k diminished further, and the current higher. From formula (2) and formula (4), the start current is proportional to the terminal voltage, and the start torque is proportional to the square of the stator terminal voltage. When the start voltage is low, the start torque and the current are low, and if the voltage is high, the start torque and the impact current are high. If the asynchronous motor adopts the direct start mode, the voltage break from zero to the voltage of the power grid will produce large impact torque and impact current which will bring many negative influences to the electronic equipments and mechanical equipments.

3. Principle of pump control soft start

Through the analysis of the traditional start, the voltage ramp start (Li, 2000, P. 16-19) and the limited current start (Zenginobuz, 2000, P.1593-1604 & Sen P, 1900, P.1102-1107), there is a common character in the general reduced voltage start, i.e. they all can not exactly adjust the motor parameters to adapt special start requirements. Though the initial torque and the initial current are reduced, but if the start is required to be implemented in the constant accelerating torque and below the special value of the current, the reduced voltage start can not fulfill these requirements. And for this start mode, the maintenance is complex, and the installation charge is high, and it can not completely avoid the current impact and torque impact, and the apparatus in the power grid may be damaged. So we put forward the pump control soft start which takes the thyristor as the execution component, and emits PWM wave to control the thyristor to trigger the pulse by the control unit, and realizes the control of the motor start and control the connection of the thyristor (Lin, 2002 & Wang, 2001). The closed loop control frame chart is seen in Figure 2.

The pump control is actually a sort of torque control, and it adopts the PID control to adjust the electromagnetic torque (Wang, 1998 & Zhang, 2005, P. 19-21), and it requires that the electromagnetic torque should ascend according to the pump character curve when the motor starts, i.e. the accelerating torque should be tried to keep in a stable range, and the value should not be too high, and the electromagnetic torque just exceeds the load torque. The torque/rev figure of the soft starter with pump control function is seen in Figure 3, and to comparing with the general voltage ramp start and the limited current start, we also give the torque character curves of the direct start and the voltage ramp start. We can compare the torque/rev curves of the direct start, the electric reduced voltage star and the special pump control soft start in the figure, and from the figure, we can see that after the pump control function is used, a small difference value will exist between the motor torque and the torque of pump load, which will largely reduce the accelerating torque in the start and perfectly control the output torque.

4. Simulation research

We used the Simulink base and SimPowerSystems in MATLAB6.5 to establish the soft starter simulation model of the asynchronous motor pump control start mode. As seen in Figure 4, the simulation module of the pump control soft starter is mainly composed by the three phase voltage power module, the triggering pulse module, the thyristor module, the measurement module and the control module. The three-AC voltage power module is connected in start by three single phase AC voltage supplies, and the relationship between the phase and the amplitude is set up correctly. The peak amplitudes of phase A, phase B and phase C are all 312V, and the voltage frequencies are all 50Hz, and the phases respectively are 0^0, 120^0 and 240^0, and the trigger circuit is composed by the synchronization stage, sawtooth wave

forming stage and the phase-shift control stage, and the input port of the circuit is the input port of the synchronous voltage, and the synchronous voltage forms the square wave with the width of the half circle of the synchronization voltage after passing the relay stage, and the square wave produces the sawtooth wave by the rate limiter, and the sawtooth wave folds with the phase shift control voltage (input port In2) and adjusts the zero-pass point of the sawtooth wave, and after passing Relay 1, the thyristor emitting pulses which front part can be adjusted and which back part is fixed are produced, and the parameter setting of various modules is seen in Table 1.

The waveforms including the synchronization signal, half-cycle and equal width square wave, sawtooth wave, folded phase shift control and the trigger signal in turn of the various branch modules in the trigger circuit is seen in Figure 5. The thyristor module is composed by three trigger angle control modules, and it produces six ways trigger pulses, and each trigger pulse differs in 60^0, and the simulation waveform is seen in Figure 6. And the thyristor module is the core of the main loop of the soft starter, and it is the AC voltage regulation circuit composed by three groups including 6 thyristors.

Because the system is mainly controlled by the pump load, so the objective of the simulation is not generally constant load, but the pump load which load torque changes with the rev. The load torque of the pump load is proportional with the square of the rev, so a pump character curve of the pump load can be approximately fitted (seen in Figure 7). The current waveform and the torque waveform of the pump control start mode are seen in Figure 8 and Figure 9, and from the figures, we can clearly see that the pump control mode can not only further limit the start current (about 30A), but better control the output torque of the motor (the maximum torque is less than $55\,N\cdot m$), and the electromagnetic torque and the load torque keep simultaneously ascending, which can largely reduce the acceleration torque and the hammer impact.

5. Conclusions

From the simulation result, the soft start with pump control could reduce the accelerating torque and the start current to the largest extent, and its performance is much better than general starter, and it could better control the start process of the asynchronous motor.

References

Gao, Jingde, Wang, Xiangxian & Li, Fahai. (1993). *AC Motor and Its System Analysis.* Beijing: Tsinghua University Press.

Li, Guohou & Gao, Shuping. (2006). Soft Starting Technology and Its Application. *Coal Mine Machinery.* No.7. P. 185-187.

Lin, Weixun. (2002). *Modern Power Electronics Circuit.* Hangzhou: Zhejiang University Press.

Li, Wujiu. (2000). Intelligent Soft Starter. *Low Voltage Apparatus.* No.1. P. 16-19.

Sen P. Landa. (1900). Derating of Induction Motors Due to waveform Distoration. IEEE Trans. On Industry Application. No.11. Vol.26(3). P.1102-1107.

Wang, Yudong. (1990). *Motor Theory.* Hangzhou: Zhejiang University Press.

Wang, Zhao'an & Huang, Jun. (2001). *Power Electronics Technology.* China Machine Press.

Wang, Zhao'an, Yangjun & Liu, Jinjun. (1998). *Harmonic Suppression and Reactive Power Compensation.* China Machine Press.

Xu, Liaosong, Qian, Cunyuan & Xie, Weida. (2001). A Soft-stop Control Method of Asynchronous Motor Soft-starter. *Power Electronics.* No.35(3). P. 41-44.

Zenginobuz, G. Cedircil, Ermis, M & Barlark. C. (2000). Soft starting of large induction motors at constant current with minimized starting torque pulsations. *Conference Re Cordofthe2000 IEEE.* No.3(11). P.1593-1604.

Zhang, Xiaoquan. (1999). New Concept of Motor Control Protection: the Characteristics and Application Technology of the Super Soft Start Equipment. *Electric Power Construction.* No.8. P. 3-5.

Zhang, Xueping. (2005). Application of Single-chip Microcomputer in Soft Start of AC Motor. *Industry and Mine Automation.* No.4. P. 19-21.

Table 1. Parameters setting of the triggering angle control module

Module	Power Uin	Relay1, Relay3		Rate Limiter1, Rate Limiter2		Relay2, Relay4	
Parameter setting	220V	Switch On point	eps	Rising slew rate	1000	Switch On point	eps
	50Hz	Switch Off point	eps	Falling slew rate	-1e8	Switch Off point	eps
		Output when on	10			Output when on	1
		Output when off	0			Outputwhen off	0

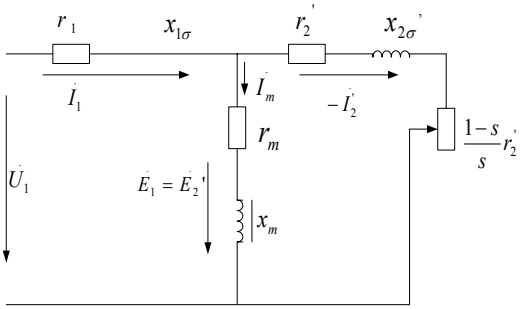


Figure 1. Single Phase Equivalent Circuit of Asynchronous Motor

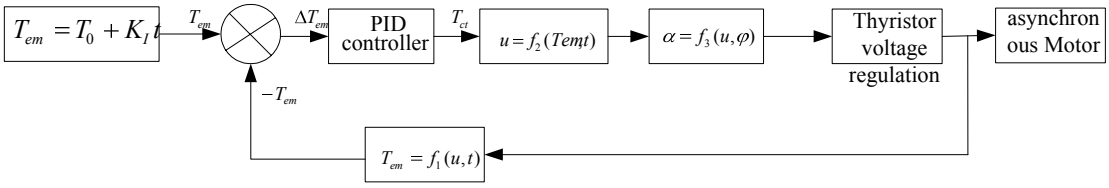


Figure 2. Torque Closed Loop Control Strategy

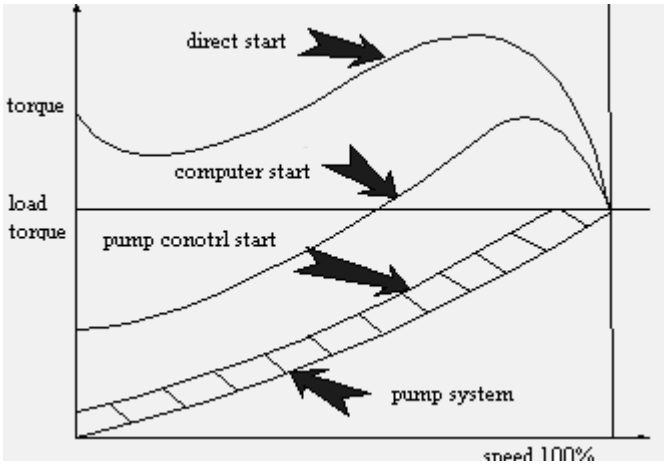


Figure 3. Torque/Rev Curve of Pump Control Soft Starter

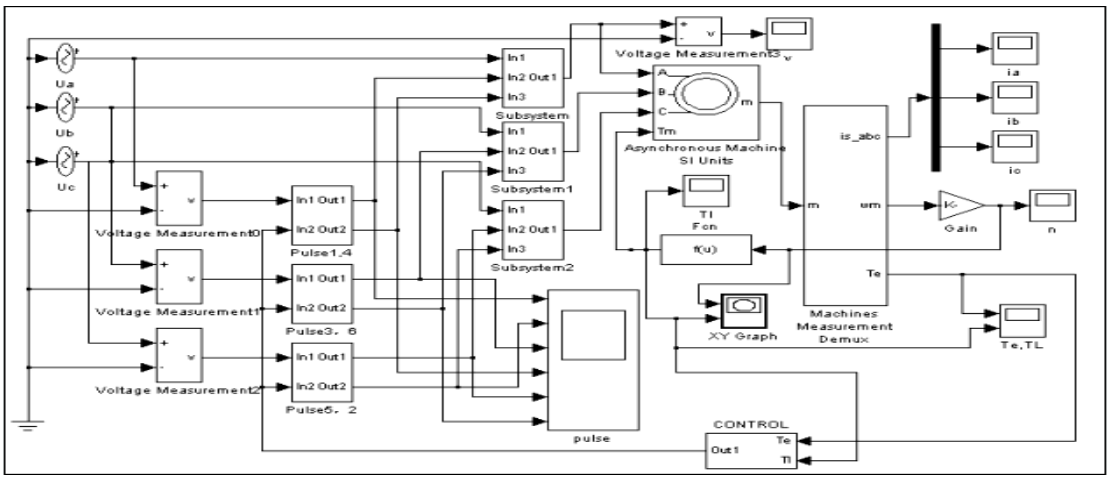


Figure 4. Simulation Principle

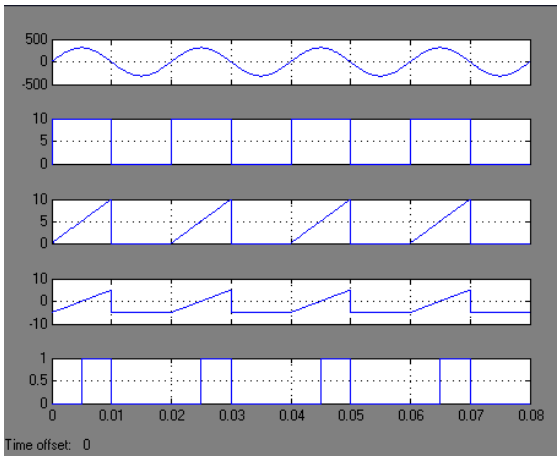


Figure 5. Various Branch Module Waveforms of Trigger Circuit

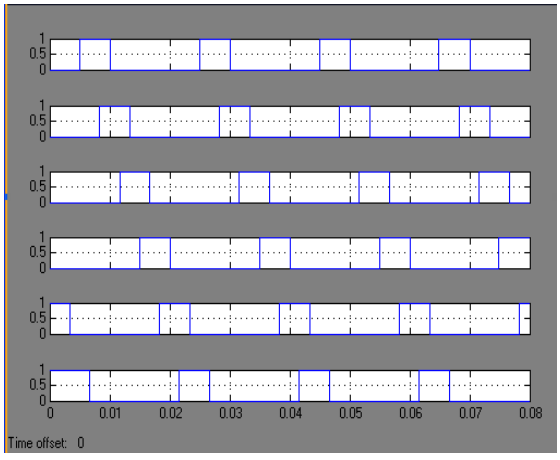


Figure 6. Six-way Pulse Outputs

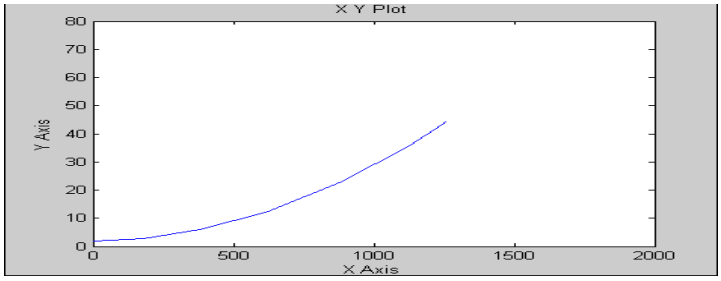


Figure 7. Speed/Torque Curve of Pump Load

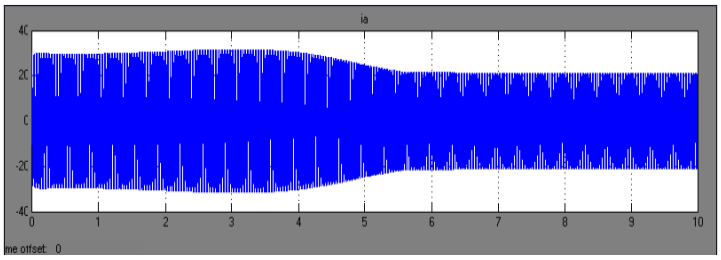


Figure 8. Pump Control Start Circuit i_a

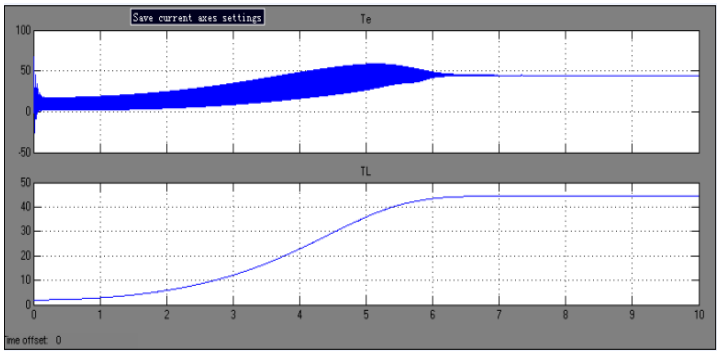


Figure 9. Pump Control Start T_e, T_L