MJFAS
Malaysian Journal of Fundamental and Applied Sciences

Jamaliyatul *et al.* | Malaysian Journal of Fundamental and Applied Sciences, Vol. 20 (2024) 302-319

RESEARCH ARTICLE

Modelling Wind Speed, Humidity, and Temperature in Butterworth and Melaka during Southwest Monsoon in 2020 with a Simultaneous Linear Functional Relationship

Nur Ain Al-Hameefatul Jamaliyatul[a], Nurkhairany Amyra Mokhtar[a]*, Basri Badyalina[a], Adzhar Rambli[b], Yong Zulina Zubairi[c]

[a]Mathematical Sciences Studies, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA (UiTM) Johor Branch, Segamat Campus, 85000 Segamat, Johor, Malaysia; [b]School of Mathematical Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia; [c]Institute of Advanced Studies, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

Abstract The extension of parameter estimation from a bivariate linear functional relationship model (LFRM) to simultaneous LFRM for linear variables using the maximum likelihood estimation (MLE) method is explored in this paper. The covariance matrix of the parameter estimates is derived through the Fisher information matrix. A simulation study was done to investigate the performance of the parameter estimation. According to the simulation study, the estimated parameters have a small bias. The beauty of simultaneous LFRM lies in developing the model to study the relationship between more than two linear variables while considering error terms for all variables. The applicability of the proposed simultaneous model is demonstrated using wind speed, humidity, and temperature data from Butterworth and Melaka during the southwest monsoon season of 2020.

Keywords: Simultaneous Linear Functional Relationship Model, Fisher Information Matrix, Maximum Likelihood Estimation, Linear Variables.

Introduction

A standard statistical procedure is to discover the relationships between variables. The relationship between these variables can be estimated using regression. If these variables have relationships, they can be linear or non-linear. The variables are commonly divided into an independent variable, X and a dependent variable, Y. Regression analysis is concerned with expressing the dependent variables as functions of the independent variables (Gillard, 2010).

The errors-in-variables model, EIVM is a linear regression model extension in which variables X and Y are continuously linear and measured with errors (Arif *et al*., 2020a). Suppose the variables X and Y are related by $Y = \alpha + \beta X$. There is no statistical problem in obtaining values of α and β if both X and Y are correctly observed. When both X and Y are subject to error, the EIVM is applied. In practice, measurement errors occur when neither variable is precisely recorded (Arif *et al*., 2019). Measurement errors can arise in various fields, including econometrics, environmental sciences, engineering, and manufacturing. For example, instrument issues could occur in the industrial sector due to variations in the measuring process (Arif *et al*., 2021). Adcock (1878) investigated the problem of fitting a linear relationship when both the dependent and independent variables were subject to error in the late 18th century, leading to the EIVM's development (Arif *et al*., 2020a; Fah *et al*., 2010).

***For correspondence:**
nurkhairany@uitm.edu.my

Received: 20 Dec. 2023
Accepted: 8 Feb. 2024

Error-in-variables problems typically develop when modelling is used to get physical insight into an operation (Mokhtar *et al*., 2021a, 2022b). Because it considers the presence of error across all parameters, EIVM is the most statistically relevant tool for predicting reactivity ratios. EIVM is classified into three types: functional, structural, and ultrastructural (Ghapor *et al*., 2014; Jamaliyatul *et al*., 2023). A functional relationship model between X and Y is when X is a mathematical variable or fixed constant. Meanwhile, the structural relationship model between X and Y is when X is a random variable. An ultrastructural relationship model is when a functional and structural relationship is combined. This study will focus on the linear functional relationship model, LFRM, which defines the X variable as a mathematical variable or fixed constant.

In the last five decades, researchers have focused on estimating parameters within the bivariate LFRM. To the best of the author's knowledge, the bivariate LFRM model has been extended to the simultaneous LFRM to study the relationship between multiple circular variables, representing a novel contribution to existing literature (Mokhtar *et al*., 2015). The existing research uses the bivariate LFRM to study the relationship between two linear variables (Arif *et al*., 2020b; Ghapor *et al*., 2014). Therefore, this study extends the work of Ghapor *et al*. (2014) by extending the bivariate LFRM to simultaneous LFRM, allowing for the statistical examination of relationships among more than two linear variables, while considering measurement errors.

Simultaneous Linear Functional Relationship Model

In this study, a simultaneous linear functional relationship model (LFRM) is extended from the bivariate LFRM for linear data. This enhancement is done so that the relationship between more than two linear variables can be studied statistically. Suppose the variable $Y_{ji}(j=1,...,q;i=1,...,n)$ and $X_i(i=1,...,n)$ related by simultaneous LFRM of $Y_j=\alpha_j+\beta_j X$, where the α_j is the *y*-intercept, β_j is the slope of the function, n is the number of observations or data point in dataset and q is the number of response variables. Let the observation be (x_i, y_{ji}), and the observation corresponds to the measurement of the true values of (X_i, Y_{ji}), made with some random error. The random error δ_i and ε_{ji} are assumed to be normally distributed with $\delta_i \sim N(0,\sigma_i^2)$ and $\varepsilon_j \sim N\left(0,\tau_j^2\right)$, respectively. σ_i^2 and τ_j^2 is the error variance. The model of simultaneous LFRM can be written as follows:

$$Y_j=\alpha_j+\beta_j X \quad (1)$$

where $x_i = X+\delta_j$ and $y_{ji}=Y_j+\varepsilon_j$ for $j=1,...,q;i=1,...,n$.

Parameter Estimation using Maximum Likelihood Estimation for Simultaneous LFRM

In this study, we consider the case when the ratio of error variances, λ is known where $\lambda=\frac{\tau_j^2}{\sigma_i^2}$ for all observations on both variables. Thus, there are $(n+q+1)$ parameters to be estimated which $\alpha_1,...,\alpha_q$, $\beta_1,...,\beta_q$, $\sigma_i^2,...,\sigma_n^2$ and $X_i,...,X_n$ by using the maximum likelihood estimation (MLE) method. The log-likelihood function of the model is given by

$$\log L = -n\log(2\pi)-\frac{n}{2}\log\lambda - n\log\sigma_i^2 - \frac{1}{2\sigma_i^2}\left\{\begin{array}{l}\sum_{i=1}^{n}(x_i-X_i)^2+\\ \frac{1}{\lambda}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)^2\end{array}\right\} \quad (2)$$

a) Maximum Likelihood Estimation of α_j

Differentiating equation (2) with respect to α_j:

$$\frac{\delta}{\delta\alpha_j}(\log\ L) = -\frac{1}{2\sigma_i^2}\frac{\delta}{\delta\alpha_j}\left[\sum_{i=1}^{p}(x_i-X_i)^2+\frac{1}{\lambda}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)^2\right]$$

First, let's focus on the second summation term with respect to α_j :

$$\frac{\delta}{\delta\alpha_j}(y_{ji}-\alpha_j-\beta_j X_i)^2=-2(y_{ji}-\alpha_j-\beta_j X_i)$$

Now, substitute this back into the expression:

$$\frac{\delta}{\delta\alpha_j}(\log\ L)=\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)$$

By setting $\frac{\delta}{\delta\alpha_j}(\log\ L)=0$

$$\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)=0$$

Now, let's isolate α_j by multiplying both sides by $\lambda\sigma_i^2$:

$$\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)=0$$

Now, let's rearrange the terms:

$$\sum_{i=1}^{n}y_{ji}-\sum_{i=1}^{n}\alpha_j-\sum_{i=1}^{n}\beta_j X_i=0$$

Combine the summations:

$$\sum_{i=1}^{n}y_{ji}-n\alpha_j-\beta_j\sum_{i=1}^{n}X_i=0$$

Now, solve for α_j

$$\alpha_j=\frac{1}{n}\left(\sum_{i=1}^{n}y_{ji}-\beta_j\sum_{i=1}^{n}X_i\right)$$

Simplify then we get

$$\hat{\alpha}_j=\bar{y}_j-\hat{B}_j\bar{x}\quad(3)$$

where $\bar{y}_j=\frac{1}{n}\sum_{i=1}^{n}y_{ji}$ and $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$

b) Maximum Likelihood Estimation of X_i

Differentiating equation (2) with respect to X_i :

$$\frac{\delta}{\delta X_i}(\log\ L)=-\frac{1}{2\sigma_i^2}\frac{\delta}{\delta X_i}\left[\sum_{i=1}^{n}(x_i-X_i)^2+\frac{1}{\lambda}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)^2\right]$$

First, differentiate the first summation term with respect to X_i

$$\frac{\delta}{\delta X_i}(x_i-X_i)^2=-2(x_i-X_i)$$

Second, differentiate the first summation term with respect to X_i

$$\frac{\delta}{\delta X_i}(y_{ji}-\alpha_j-\beta_j X_i)^2=-2\beta_j(y_{ji}-\alpha_j-\beta_j X_i)$$

Now, substitute this back into the expression:

$$\frac{\delta}{\delta X_i}(\log\ L)=-\frac{1}{\sigma_i^2}\sum_{i=1}^{n}(x_i-X_i)-\frac{1}{\lambda\sigma_i^2}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)$$

By setting $\frac{\delta}{\delta X_i}(\log\ L)=0$

$$-\frac{1}{\sigma_i^2}\sum_{i=1}^{n}(x_i-X_i)-\frac{1}{\lambda\sigma_i^2}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)=0$$

Next, isolate the term involving X_i :

$$-\frac{1}{\sigma_i^2}\sum_{i=1}^{n}(x_i-X_i)=\frac{1}{\lambda\sigma_i^2}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}(y_{ji}-\alpha_j-\beta_j X_i)$$

Now, multiply both sides by σ_i^2 to simplify:

$$-\sum_{i=1}^{n}(x_i - X_i) = \frac{1}{\lambda}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}(y_{ji} - \alpha_j - \beta_j X_i)$$

Expand the summation terms:

$$-\sum_{i=1}^{n}x_i + \sum_{i=1}^{n}X_i = \frac{1}{\lambda}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}y_{ji} - \frac{1}{\lambda}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}\alpha_j - \frac{1}{\lambda}\sum_{j=1}^{q}\beta_j^2\sum_{i=1}^{n}X_i$$

Now, isolate the terms involving X_i:

$$\left(1+\frac{1}{\lambda}\sum_{j=1}^{q}\beta_j^2\right)\sum_{i=1}^{n}X_i = \sum_{i=1}^{n}x_i + \frac{1}{\lambda}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}y_{ji} - \frac{1}{\lambda}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}\alpha_j$$

Finally, solve for X_i:

$$nX_i = \frac{\sum_{i=1}^{n}x_i + \frac{1}{\lambda}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}y_{ji} - \frac{1}{\lambda}\sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}\alpha_j}{1+\frac{1}{\lambda}\sum_{j=1}^{q}\beta_j^2}$$

Multiply both the numerator and denominator by λ to get rid of the fraction in the denominator:

$$X_i = \frac{\lambda\sum_{i=1}^{n}x_i + \sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}y_{ji} - \sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}\alpha_j}{\lambda+\sum_{j=1}^{q}\beta_j^2}$$

$$\hat{X}_i = \frac{\lambda\sum_{i=1}^{n}x_i + \sum_{j=1}^{q}\beta_j\sum_{i=1}^{n}\left(y_{ji} - \alpha_j\right)}{\lambda+\sum_{j=1}^{q}\beta_j^2} \quad (4)$$

c) Maximum Likelihood Estimation of β_j

Differentiating equation (2) with respect to β_j:

$$\frac{\delta}{\delta\beta_j}(\log\ L) = -\frac{1}{2\sigma_i^2}\frac{\delta}{\delta\beta_j}\left[\sum_{i=1}^{n}(x_i - X_i)^2 + \frac{1}{\lambda}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_{ji} - \alpha_j - \beta_j X_i)^2\right]$$

Now, let's differentiate the second summation term with respect to β_j:

$$\frac{\delta}{\delta\beta_j}(y_{ji} - \alpha_j - \beta_j X_i)^2 = -2X_i(y_{ji} - \alpha_j - \beta_j X_i)$$

Now, substitute this back into the expression:

$$\frac{\delta}{\delta\beta_j}(\log\ L) = \frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i(y_{ji} - \hat{\alpha}_j - \hat{\beta}_j\hat{X}_i)$$

and by setting $\frac{\delta}{\delta\beta_j}(\log\ L) = 0$

$$\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i(y_{ji} - \hat{\alpha}_j - \hat{\beta}_j\hat{X}_i) = 0$$

Multiply both sides by $\lambda\sigma_i^2$ to simplify:

$$\sum_{i=1}^{n}\hat{X}_i(y_{ji} - \hat{\alpha}_j - \hat{\beta}_j\hat{X}_i) = 0$$

Now, solve for $\hat{\beta}_j$:

$$\sum_{i=1}^{n}\hat{X}_i y_{ji} - \sum_{i=1}^{n}\hat{X}_i\hat{\alpha}_j - \hat{\beta}_j\sum_{i=1}^{n}\hat{X}_i^2 = 0$$

$$\hat{\beta}_j = \frac{\sum_{i=1}^{n} \hat{X}_i y_{ji} - \hat{\alpha}_j \sum_{i=1}^{n} \hat{X}_i}{\sum_{i=1}^{n} \hat{X}_i^2}$$

Now, substitute this expression for X_i and α_j into the equation for $\hat{\beta}_j$

$$\hat{\beta}_j = \frac{\sum_{i=1}^{n} y_{ji} \left(\frac{\lambda \sum_{i=1}^{n} x_i + \sum_{j=1}^{q} \beta_j \sum_{i=1}^{n} (y_{ji} - \alpha_j)}{\lambda + \sum_{j=1}^{q} \beta_j^2} \right) - \hat{\alpha}_j \sum_{i=1}^{n} \left(\frac{\lambda \sum_{i=1}^{n} x_i + \sum_{j=1}^{q} \beta_j \sum_{i=1}^{n} (y_{ji} - \alpha_j)}{\lambda + \sum_{j=1}^{q} \beta_j^2} \right)}{\sum_{i=1}^{n} \left(\frac{\lambda \sum_{i=1}^{n} x_i + \sum_{j=1}^{q} \beta_j \sum_{i=1}^{n} (y_{ji} - \alpha_j)}{\lambda + \sum_{j=1}^{q} \beta_j^2} \right)^2}$$

Now, we can simplify this expression by cancelling out some terms:

$$\hat{\beta}_j = \frac{\left(\lambda + \sum_{j=1}^{q} \beta_j^2 \right) \left(\lambda S_{xy} + \sum_{j=1}^{q} \beta_j S_{yy} + \lambda n \bar{x}^2 \sum_{j=1}^{q} \beta_j + n \bar{x}^2 \sum_{j=1}^{q} \beta_j^3 \right)}{\lambda^2 \sum_{i=1}^{n} x_i^2 + 2\lambda \sum_{j=1}^{q} \beta_j S_{xy} + \sum_{j=1}^{q} \beta_j^2 S_{yy} + 2\lambda n \bar{x}^2 \sum_{j=1}^{q} \beta_j^2 + n \bar{x}^2 \sum_{j=1}^{q} \beta_j^4}$$

where $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$, $S_{yy} = \sum_{i=1}^{n} (y_{ji} - \bar{y}_j)^2$, and $S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$.

This implies that

$$\sum_{j=1}^{q} \hat{\beta}_j^2 S_{xy} + \sum_{j=1}^{q} \hat{\beta}_j (\lambda S_{xx} - S_{yy}) - \lambda S_{xy} = 0$$

Solving the quadratic equation, where , $a = S_{xy}$, $b = \lambda S_{xx} - S_{yy}$, and $c = -\lambda S_{xy}$ yields

$$\hat{\beta}_j = \frac{-(\lambda S_{xx} - S_{yy}) \pm \sqrt{(\lambda S_{xx} - S_{yy})^2 - 4(S_{xy})(-\lambda S_{xy})}}{2(S_{xy})}$$

$$\hat{\beta}_j = \frac{(S_{yy} - \lambda S_{xx}) + \sqrt{(\lambda S_{xx} - S_{yy})^2 + 4\lambda S_{xy}^2}}{2S_{xy}} \quad (5)$$

The positive sign is used in equation (5) because it gives a maximum to the likelihood function in equation (2) as shown below. From the previous result, we have

$$\frac{\delta}{\delta \beta_j} (\log L) = \frac{1}{\lambda \sigma_i^2} \sum_{i=1}^{n} \hat{X}_i (y_{ji} - \hat{\alpha}_j - \hat{\beta}_j \hat{X}_i)$$

and the second-order derivative yields

$$\frac{\delta^2}{\delta \beta_j^2} (\log L) = -\frac{1}{\lambda \sigma_i^2} \sum_{i=1}^{n} \hat{X}_i^2$$

Since $\sum_{i=1}^{n} \hat{X}_i^2 > 0$, (practically $\hat{X}_i \neq 0$) and $\lambda > 0$, this implies that $\frac{\delta^2}{\delta \beta_j^2} (\log L) < 0$. The $\hat{\beta}_j$ are local maximum points. Now, we let

$$\hat{\beta}_j = \frac{(S_{yy} - \lambda S_{xx}) + \sqrt{(\lambda S_{xx} - S_{yy})^2 + 4\lambda S_{xy}^2}}{2S_{xy}} = \frac{\Delta}{2S_{xy}}$$

It could be shown that $\Delta = 2\hat{\beta}_j S_{xy} \geq 0$ must be non-negative and therefore the positive square root must always be taken.

d) Maximum Likelihood Estimation of σ_i^2

Differentiating equation (2) with respect to σ_i^2 we get

$$\frac{\delta}{\delta\sigma_i^2}(\log\ L) = -\frac{n}{\sigma_i^2} + \frac{1}{2\left(\sigma_i^2\right)^2}\left\{\sum_{i=1}^{n}(x_i - X_i)^2 + \frac{1}{\lambda}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_i - \alpha - \beta X_i)^2\right\}$$

To estimate $\hat{\sigma}_i^2$, let $\frac{\delta}{\delta\sigma_i^2}(\log L) = 0$

$$-\frac{n}{\sigma_i^2} + \frac{1}{2\left(\sigma_i^2\right)^2}\left\{\sum_{i=1}^{n}(x_i - X_i)^2 + \frac{1}{\lambda}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_i - \alpha - \beta X_i)^2\right\} = 0$$

$$\hat{\sigma}_i^2 = \frac{1}{2n}\left\{\sum_{i=1}^{n}(x_i - X_i)^2 + \frac{1}{\lambda}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_{ji} - \alpha_j - \beta_j X_i)^2\right\} \quad (6)$$

Based on Kendall and Stuart (1973), knowledge of the ratio of error variances has enabled us to evaluate the MLE of the parameter estimate, but the trouble is $\hat{\sigma}_i^2$ is not a consistent estimator of σ_i^2 as Lindley (1947) showed. The inconsistency of the MLE is therefore a reflection of the small-sample bias of the MLE in general. This particular inconsistent estimator causes no difficulty, a consistent estimator of σ_i^2 being given by replacing the number of observations, $2n$, by the number of degrees of freedom, $2n-(n+2) = n-2$, in the divisor of $\hat{\sigma}_i^2$. The consistent estimator is therefore

$$\hat{\sigma}_i^2 = \frac{1}{n-2}\left\{\sum_{i=1}^{n}(x_i - X_i)^2 + \frac{1}{\lambda}\sum_{j=1}^{q}\sum_{i=1}^{n}(y_{ji} - \alpha_j - \beta_j X_i)^2\right\} \quad (7)$$

Fisher Information Matrix of the Simultaneous LFRM

The Fisher Information Matrix of parameters $\hat{\alpha}_j$ and $\hat{\beta}_j$ are used to obtain the variance-covariance matrix of $\hat{\alpha}_j$ and $\hat{\beta}_j$. The second derivative of equation (2) with respect to α_j is then:

$$\frac{\delta^2}{\delta\alpha_j^{\ 2}}(\log L) = -\frac{n}{\lambda\sigma_i^2}$$

Therefore, $E\left(-\frac{\delta^2}{\delta\alpha_j^{\ 2}}(\log L)\right) = \frac{n}{\lambda\sigma_i^2}$

The second derivative of equation (2) with respect to β_j is then:

$$\frac{\delta^2}{\delta\beta_j^{\ 2}}(\log L) = -\frac{1}{\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i^{\ 2}$$

Therefore, $E\left(-\frac{\delta^2}{\delta\beta_j^{\ 2}}(\log L)\right) = \frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i^{\ 2}$

$$\frac{\delta^2}{\delta\alpha_j\ \delta\beta_j}(\log L) = -\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i$$

Therefore, $E\left(-\frac{\delta^2}{\delta\alpha_j\ \delta\beta_j}(\log L)\right) = \frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i$

Thus, the estimated Fisher information matrix, F for $\hat{\alpha}$ and $\hat{\beta}$ is as follows,

$$F = \begin{pmatrix} \frac{n}{\lambda\sigma_i^2} & \frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i \\ \frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i & \frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i^{\ 2} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad (8)$$

where A is a 1×1 matrix, B is a $1\times n$ matrix, C is a $n\times1$ matrix, and D is a $n\times n$ matrix. From the theory of partitioned matrices (Nelder, 1977), the inverse of F is

$$F^{-1}=\begin{pmatrix}\left(A-BD^{-1}C\right)^{-1} & -A^{-1}B\left(D-CA^{-1}B\right)^{-1}\\ -D^{-1}C\left(A-BD^{-1}C\right)^{-1} & \left(D-CA^{-1}B\right)^{-1}\end{pmatrix} \quad (9)$$

$$V\hat{a}r(\hat{\alpha}_j)=\left(A-BD^{-1}C\right)^{-1}$$

$$V\hat{a}r(\hat{\alpha}_j)=\left(\frac{n}{\lambda\sigma_i^2}-\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i\right)\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i^{\,2}\right)^{-1}\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i\right)\right)^{-1}$$

$$V\hat{a}r(\hat{\alpha}_j)=\frac{\left(\lambda+\hat{\beta}_j^{\,2}\right)\hat{\sigma}_i^2\hat{\beta}_j}{S_{xy}}\left\{\bar{x}^2\left(1+\hat{T}\right)+\frac{S_{xy}}{n\hat{\beta}_j}\right\} \quad (10)$$

$$V\hat{a}r(\hat{\beta}_j)=\left(D-CA^{-1}B\right)^{-1}$$

$$V\hat{a}r(\hat{\beta}_j)=\left(\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i^{\,2}\right)-\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i\right)\left(\frac{n}{\lambda\sigma_i^2}\right)^{-1}\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i\right)\right)^{-1}$$

$$V\hat{a}r(\hat{\beta}_j)=\frac{\left(\lambda+\hat{\beta}_j^{\,2}\right)\hat{\sigma}_i^2\hat{\beta}_j}{S_{xy}}\left\{1+\hat{T}\right\} \quad (11)$$

Thus, the covariance of $\hat{\alpha}_j$ and $\hat{\beta}_j$ are

$$C\hat{o}v(\hat{\alpha}_j,\hat{\beta}_j)=-D^{-1}C\left(A-BD^{-1}C\right)^{-1}$$

$$C\hat{o}v(\hat{\alpha}_j,\hat{\beta}_j)=-\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i^{\,2}\right)^{-1}\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i\right)\left(\begin{matrix}\frac{n}{\lambda\sigma_i^2}-\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i\right)\left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i^{\,2}\right)^{-1}\\ \left(\frac{1}{\lambda\sigma_i^2}\sum_{i=1}^{n}\hat{X}_i\right)\end{matrix}\right)^{-1}$$

$$C\hat{o}v(\hat{\alpha}_j,\hat{\beta}_j)=\frac{\left(\lambda+\hat{\beta}_j^{\,2}\right)\hat{\sigma}_i^2\hat{\beta}_j\bar{x}}{S_{xy}}\left\{1+\hat{T}\right\} \quad (12)$$

where $\hat{T}=\dfrac{n\lambda\hat{\beta}_j\hat{\sigma}_i^2}{\left(\lambda+\hat{\beta}_j^{\,2}\right)S_{xy}}$.

Materials and Methods

This simultaneous LRFM will be evaluated through a bias measure and a Monte Carlo simulation study is carried out in MATLAB software to assess the performance measure of the parameter estimation. The results from the bias measure would indicate the adequacy of the model's parameter estimates. Mean, estimated bias, and mean absolute percentage error, MAPE are evaluated for the parameter estimates of $\hat{\alpha}_j$, $\hat{\beta}_j$ and $\hat{\sigma}_i^2$. Without loss of generality, the number of simulations is set to be $s=10000$, and the response variables, $j=1,\ldots,q$, are set to be $q=2$, which is two response variables, y_1 and y_2 (Arif *et al*., 2019; Ghapor *et al*., 2015). The values of α_1 and α_2 = 5 and 10 while β_1 and β_2 = 1 and also σ_1^2 and σ_2^2 = 1. The sample size is set to be n = 40, 80, 120 and 160, where $i=1,\ldots,n$. In the simulation, the value of λ considered is 1 (Arif *et al*., 2022, 2021a). For simplicity, θ_r represent the parameter of α_j, β_j and σ_i^2, and $\hat{\theta}_r$ be the estimated value of $\hat{\alpha}_j$, $\hat{\beta}_j$ and $\hat{\sigma}_i^2$, respectively. $\bar{\hat{\theta}}$ represent the mean value of the estimated value. The following are the measures used to assess the estimation quality of $\hat{\alpha}_j$, $\hat{\beta}_j$ and $\hat{\sigma}_i^2$.

- Mean of $\hat{\theta}$, $\bar{\hat{\theta}}=\frac{1}{s}\sum_{r=1}^{s}\hat{\theta}_r$
- Estimated bias of $\hat{\theta}$, $EB\left(\bar{\hat{\theta}}\right)=\bar{\hat{\theta}}-\theta_r$

- Mean absolute percentage error of $\hat{\theta}$, $MAPE\left(\hat{\theta}\right)=\frac{1}{s}\sum_{r=1}^{s}\left|\frac{\theta-\hat{\theta}_r}{\theta}\right|$

This general formula will be used to obtain bias measures of $\hat{\alpha}_j$, $\hat{\beta}_j$ and $\hat{\sigma}_i^2$.
The simulation design can be described as follows:

Step 1: Generate random X_i of size n, with $i=1,2,3,\ldots,n$. Without loss of generality, the slope, β_j and y -intercept of the parameter, α_j of simultaneous LFRM, for y_1 and y_2 are fixed at $\beta_1=1$, $\beta_2=1$, $\alpha_1=5$, $\alpha_2=10$, $\sigma_1^2=1$, $\sigma_2^2=1$, and $\lambda=1$, respectively.

Step 2: Generate three random error terms δ_1 from $N(0,\sigma_i^2)$ while ε_1 and ε_2 from $N(0,\tau_j^2)$, respectively. δ_1 is the error term of x while ε_1 and ε_2 is the error terms of y_1 and y_2, respectively.

Step 3: Calculate the observed value of x, y_1 and y_2 using equation (1), where $x=X+\delta_1$, $y_1=Y_1+\varepsilon_1$, and $y_2=Y_2+\varepsilon_2$ for $q=2;i=1,\ldots,n$.

Step 4: Calculate the mean of x, y_1 and y_2.

Step 5: Calculate the parameter estimates $\hat{\alpha}_1, \hat{\alpha}_2, \hat{X}_1, \hat{X}_2, \hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$.

Step 6: Calculate the mean, estimated bias and mean absolute percentage error (MAPE) of $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$.

Simulation Results and Discussion

The result of a simulation in observing the accuracy of the parameter estimation for the simultaneous LFRM is given below.

a) Simulation result of $\hat{\alpha}_1$

Table 1 shows the simulation result of $\hat{\alpha}_1$ when $\alpha_1=5$, $\beta_1=1$ and $\sigma_1^2=1$ while Table 2 shows the simulation result of $\hat{\alpha}_1$ when $\alpha_1=10$, $\beta_1=1$ and $\sigma_1^2=1$.

Table 1. Performance measurement for $\hat{\alpha}_1$ when $\alpha_1=5$, $\beta_1=1$ and $\sigma_1^2=1$

n	Mean ($\hat{\alpha}_1$)	EB ($\hat{\alpha}_1$)	MAPE ($\hat{\alpha}_1$)
40	4.9744	-0.0256	0.0750
80	4.9870	-0.0130	0.0525
120	4.9959	-0.0041	0.0424
160	4.9976	-0.0024	0.0366

Table 2. Performance measurement for $\hat{\alpha}_1$ when $\alpha_1=10$, $\beta_1=1$ and $\sigma_1^2=1$

n	Mean ($\hat{\alpha}_1$)	EB ($\hat{\alpha}_1$)	MAPE ($\hat{\alpha}_1$)
40	9.9865	-0.0135	0.0373
80	9.9877	-0.0123	0.0261
120	9.9978	-0.0022	0.0182
160	9.9981	-0.0019	0.0212

Both Table 1 and Table 2 have the mean of $\hat{\alpha}_1$ closer to their respective actual value of α_1 as we increase the value of n. This suggests that, on average, the estimates in both tables are centred around the actual values. The estimated bias of $\hat{\alpha}_1$ is closer to to zero in both tables, indicating better estimation accuracy. The MAPE of $\hat{\alpha}_1$ is decreasing from both tables. A decreasing MAPE of $\hat{\alpha}_1$ as n increases and a lower value in Table 2 indicate an improvement in the accuracy of estimates when $\alpha_1 = 10$. Therefore, the estimation seems adequate for most values of n.

b) Simulation result of $\hat{\alpha}_2$

Table 3 shows the simulation result of $\hat{\alpha}_2$ when $\alpha_2 = 5$, $\beta_2 = 1$ and $\sigma_2^2 = 1$ while Table 4 shows the simulation result of $\hat{\alpha}_2$ when $\alpha_2 = 10$, $\beta_2 = 1$ and $\sigma_2^2 = 1$.

Table 3. Performance measurement for $\hat{\alpha}_2$, when $\alpha_2 = 5$, $\beta_2 = 1$ and $\sigma_2^2 = 1$

n	Mean ($\hat{\alpha}_2$)	EB ($\hat{\alpha}_2$)	MAPE ($\hat{\alpha}_2$)
40	4.9858	-0.0142	0.0752
80	4.9868	-0.0132	0.0532
120	4.9938	-0.0062	0.0421
160	4.9968	-0.0032	0.0366

Table 4. Performance measurement $\hat{\alpha}_2$ when $\alpha_2 = 10$, $\beta_2 = 1$ and $\sigma_2^2 = 1$

n	Mean ($\hat{\alpha}_2$)	EB ($\hat{\alpha}_2$)	MAPE ($\hat{\alpha}_2$)
40	9.9880	-0.0120	0.0372
80	9.9919	-0.0081	0.0264
120	9.9964	-0.0036	0.0210
160	9.9970	-0.0030	0.0185

Both Table 3 and Table 4 have the mean of $\hat{\alpha}_2$ closer to their respective actual value of α_2 as we increase the value of n. This suggests that, on average, the estimates in both tables are centred around the actual values. The estimated bias of $\hat{\alpha}_2$ is closer to to zero in both tables, indicating better estimation accuracy. The MAPE of $\hat{\alpha}_2$ is decreasing from both tables. A decreasing MAPE of $\hat{\alpha}_2$ as n increases and a lower value in Table 4 indicate an improvement in the accuracy of estimates when $\alpha_2 = 10$. Therefore, the estimation seems adequate for most values of n.

c) Simulation result of $\hat{\beta}_1$

Table 5 shows the simulation result of $\hat{\beta}_1$ when $\alpha_1 = 5$, $\beta_1 = 1$ and $\sigma_1^2 = 1$ while Table 6 shows the simulation result of $\hat{\beta}_1$ when $\alpha_1 = 10$, $\beta_1 = 1$ and $\sigma_1^2 = 1$.

Table 5. Performance measurement for $\hat{\beta}_1$ when $\alpha_1 = 5$, $\beta_1 = 1$ and $\sigma_1^2 = 1$

n	Mean ($\hat{\beta}_1$)	EB ($\hat{\beta}_1$)	MAPE ($\hat{\beta}_1$)
40	1.0054	0.0054	0.0642
80	1.0021	0.0021	0.0452
120	1.0012	0.0012	0.0371
160	1.0011	0.0011	0.0319

Table 6. Performance measurement for $\hat{\beta}_1$ when $\alpha_1 = 10$, $\beta_1 = 1$ and $\sigma_1^2 = 1$

n	Mean ($\hat{\beta}_1$)	EB ($\hat{\beta}_1$)	MAPE ($\hat{\beta}_1$)
40	1.0029	0.0029	0.0634
80	1.0015	0.0015	0.0448
120	1.0011	0.0011	0.0367
160	1.0010	0.0010	0.0316

Both Table 5 and Table 6 have the mean of $\hat{\beta}_1$ closer to their respective actual value of β_1 as we increase the value of n. This suggests that, on average, the estimates in both tables are centred around the actual values. The estimated bias of $\hat{\beta}_1$ is closer to to zero in both tables, indicating better estimation accuracy. The MAPE of $\hat{\beta}_1$ is decreasing from both tables. A decreasing MAPE of $\hat{\beta}_1$ as n increases and a lower value in Table 6 indicate an improvement in the accuracy of estimates when $\alpha_1 = 10$. Therefore, the estimation seems adequate for most values of n.

d) Simulation result of $\hat{\beta}_2$

Table 7 shows the simulation result of $\hat{\beta}_2$ when $\alpha_2 = 5$, $\beta_2 = 1$ and $\sigma_2^2 = 1$ while Table 8 shows the simulation result of $\hat{\beta}_2$ when $\alpha_2 = 10$, $\beta_2 = 1$ and $\sigma_1^2 = 1$.

Table 7. Performance measurement for $\hat{\beta}_2$ when $\alpha_2 = 5$, $\beta_2 = 1$ and $\sigma_2^2 = 1$

n	Mean ($\hat{\beta}_2$)	EB ($\hat{\beta}_2$)	MAPE ($\hat{\beta}_2$)
40	1.0030	0.0030	0.0642
80	1.0022	0.0022	0.0446
120	1.0017	0.0017	0.0369
160	1.0010	0.0010	0.0321

Table 8. Performance measurement $\hat{\beta}_2$ when $\alpha_2 = 10$, $\beta_2 = 1$ and $\sigma_2^2 = 1$

n	Mean ($\hat{\beta}_2$)	EB ($\hat{\beta}_2$)	MAPE ($\hat{\beta}_2$)
40	1.0023	0.0023	0.0636
80	1.0019	0.0019	0.0445
120	1.0013	0.0013	0.0364
160	1.0005	0.0005	0.0315

Both Table 7 and Table 8 have the mean of $\hat{\beta}_2$ closer to their respective actual value of β_2 as we increase the value of n. This suggests that, on average, the estimates in both tables are centred around the actual values. The estimated bias of $\hat{\beta}_2$ is closer to to zero in both tables, indicating better estimation accuracy. The MAPE of $\hat{\beta}_2$ is decreasing from both tables. A decreasing MAPE of $\hat{\beta}_2$ as n increases and a lower value in Table 8 indicate an improvement in the accuracy of estimates when $\alpha_2 = 10$. Therefore, the estimation seems adequate for most values of n.

e) Simulation result of $\hat{\sigma}_1^2$

Table 9 shows the simulation result of $\hat{\sigma}_1^2$ when $\alpha_1 = 5$, $\beta_1 = 1$ and $\sigma_1^2 = 1$ while Table 10 shows the simulation result of $\hat{\sigma}_1^2$ when $\alpha_1 = 10$, $\beta_1 = 1$ and $\sigma_1^2 = 1$.

Table 9. Performance measurement for $\hat{\sigma}_1^2$ when $\alpha_1 = 5$, $\beta_1 = 1$ and $\sigma_1^2 = 1$

n	Mean ($\hat{\sigma}_1^2$)	EB ($\hat{\sigma}_1^2$)	MAPE ($\hat{\sigma}_1^2$)
40	0.9963	-0.0037	0.1814
80	0.9984	-0.0016	0.1290
120	0.9987	-0.0013	0.1037
160	0.9992	-0.0008	0.0903

Table 10. Performance measurement for $\hat{\sigma}_1^2$ when $\alpha_1 = 10$, $\beta_1 = 1$ and $\sigma_1^2 = 1$

n	Mean ($\hat{\sigma}_1^2$)	EB ($\hat{\sigma}_1^2$)	MAPE ($\hat{\sigma}_1^2$)
40	0.9968	-0.0032	0.1813
80	0.9989	-0.0011	0.1284
120	0.9993	-0.0007	0.1026
160	0.9998	-0.0002	0.0895

Both Table 9 and Table 10 have the mean of $\hat{\sigma}_1^2$ closer to their respective actual value of $\hat{\sigma}_1^2$ as we increase the value of n. This suggests that, on average, the estimates in both tables are centred around the actual values. The estimated bias of $\hat{\sigma}_1^2$ is closer to to zero in both tables, indicating better estimation accuracy. The MAPE of $\hat{\sigma}_1^2$ is decreasing from both tables. A decreasing MAPE of $\hat{\sigma}_1^2$ as n increases and a lower value in Table 10 indicate an improvement in the accuracy of estimates when $\alpha_1 = 10$. Therefore, the estimation seems adequate for most values of n.

f) Simulation result of $\hat{\sigma}_2^2$

Table 11 shows the simulation result of $\hat{\sigma}_2^2$ when $\alpha_2 = 5$, $\beta_2 = 1$ and $\sigma_2^2 = 1$ while Table 12 shows the simulation result of $\hat{\sigma}_2^2$ when $\alpha_2 = 10$, $\beta_2 = 1$ and $\sigma_2^2 = 1$ $\sigma_1^2 = 1$.

Table 11. Performance measurement for $\hat{\sigma}_2^2$ when $\alpha_2 = 5$, $\beta_2 = 1$ and $\sigma_2^2 = 1$

n	Mean ($\hat{\sigma}_2^2$)	EB ($\hat{\sigma}_2^2$)	MAPE ($\hat{\sigma}_2^2$)
40	0.9918	-0.0082	0.1806
80	0.9981	-0.0019	0.1270
120	0.9997	-0.0003	0.1034
160	1.0001	0.0001	0.0898

Table 12. Performance measurement $\hat{\sigma}_2^2$ when $\alpha_2 = 10$, $\beta_2 = 1$ and $\sigma_2^2 = 1$

n	Mean ($\hat{\sigma}_2^2$)	EB ($\hat{\sigma}_2^2$)	MAPE ($\hat{\sigma}_2^2$)
40	0.9990	-0.0010	0.1815
80	0.9997	-0.0003	0.1260
120	0.9998	-0.0002	0.1029
160	0.9999	-0.0001	0.0894

Both Table 11 and Table 12 have the mean of $\hat{\sigma}_2^2$ closer to their respective actual value of $\hat{\sigma}_2^2$ as we increase the value of n. This suggests that, on average, the estimates in both tables are centred around the actual values. The estimated bias of $\hat{\sigma}_2^2$ is closer to to zero in both tables, indicating better estimation accuracy. The MAPE of $\hat{\sigma}_2^2$ is decreasing from both tables. A decreasing MAPE of $\hat{\sigma}_2^2$ as n increases and a lower value in Table 12 indicate an improvement in the accuracy of estimates when $\alpha_2 = 10$. Therefore, the estimation seems adequate for most values of n.

Application of Simultaneous Linear Functional Relationship Model to Real Data

Coastal weather stations commonly collect data on wind speed, temperature, and humidity, which are essential for understanding local weather conditions and atmospheric environments. This data is valuable for gaining insights into meteorological and climatic patterns, and it plays a crucial role in predicting and comprehending monsoon phenomena. Monsoons, characterised by seasonal shifts in wind direction along with notable changes in temperature and humidity, are closely observed using these collected data. In summary, the information gathered from coastal weather stations, precisely wind speed, temperature, and humidity data, is vital for monitoring and comprehending monsoon patterns. This knowledge has diverse applications, ranging from aiding in agricultural planning to facilitating disaster preparedness and response in regions affected by monsoons. In this study, we demonstrate the applicability of the simultaneous linear functional relationship model by utilising wind speed, temperature, and humidity data from the Butterworth and Melaka stations. Figure 1 shows the location of Melaka and Butterworth in Malaysia.

Figure 1. Location Melaka and Butterworth in Malaysia

Melaka town is located on the southwestern peninsula of Malaysia (2.16 °N, 102° 15'E) and encounters high temperatures and humidity without much fluctuation most days of the year (Manteghi *et al*., 2020). Next, Butterworth is located on the northwest coast of Peninsular Malaysia (5° 27' N, 100° 23' E). Wind conditions are generally light and variable, originating from the Andaman Sea and the Straits of Melaka (Hussin *et al*., 2015). The dataset, sourced from the Malaysian Meteorological Department (MDD) and presented in Microsoft Excel, includes daily maximum wind speed, 24-hour mean relative humidity, and 24-hour mean temperature during the southwest monsoon season in 2020 starting from 18th May 2020 until 22nd September 2020 (*Laporan Tahunan 2020*, 2020). With the sample size of n =128, the wind speed of Melaka is addressed as the variable x_1 while the wind speed of Butterworth is as x_2. Next, the mean relative humidity of Melaka is let as y_1 while y_3 is the mean relative humidity for Butterworth. The variable y_2 is the mean temperature data for Melaka and y_4 for Butterworth. The relationship between the variable x_1, x_2, y_1, y_2, y_3 and y_4 is shown in Table 13:

Table 13. Relationship of wind speed, temperature, and humidity data of Melaka and Butterworth

Type of relationship/Station	Melaka	Butterworth
Humidity data with wind speed	$x_1 = X_1 + \delta_1$ and $y_1 = Y_1 + \varepsilon_1$, where $Y_1 = \alpha_1 + \beta_1 X_1$	$x_2 = X_2 + \delta_2$ and $y_3 = Y_3 + \varepsilon_3$, where $Y_3 = \alpha_3 + \beta_3 X_1$
Temperature data with wind speed	$x_1 = X_1 + \delta_1$ and $y_2 = Y_2 + \varepsilon_2$, where $Y_2 = \alpha_2 + \beta_2 X_2$	$x_2 = X_2 + \delta_2$ and $y_4 = Y_4 + \varepsilon_4$, where $Y_4 = \alpha_4 + \beta_4 X_2$

The normal distribution will be used to model the relationship between wind speed, humidity and temperature data throughout the southwest monsoon in 2020 for Butterworth and Melaka. The normality of the wind speed, humidity and temperature data is tested using the Kolmogorov-Smirnov test. Kolmogorov-Smirnov test is a well-known and widely used method to test whether the data is normally distributed (Zakaria, 2022). This test is applicable when the population distribution function is continuous (Hawkins & Kanji, 1995). The following are the null, H_0 and alternative hypotheses, H_A used in a Kolmogorov-Smirnov test:

H_0: The distribution of the data is normal.

H_A: The distribution of the data is not normal.

The Kolmogorov-Smirnov statistic (D) is defined as

$$D = \max_{1 \le i \le n}\left(F(Y_i) - \frac{i-1}{n}, \frac{i}{n} - F(Y_i) \right) \quad (13)$$

where F is the theoretical cumulative distribution. H_0 is rejected if D exceeds the critical value determined from the table obtained by Massey (Lo Brano *et al.*, 2011; Massey, 1951). The critical value is derived from the maximum absolute difference between sample and population cumulative distributions for a sample size n (Massey, 1951). The equation critical value of D when α = 0.05, and the sample size is over 35, is $\frac{1.36}{\sqrt{n}}$ based on Massey (1951) (Hawkins & Kanji, 1995). Insert the value of the sample size, 128, into the equation critical value of $D = \frac{1.36}{\sqrt{128}}$. Hence, the critical value is 0.1202.

The Kolmogorov-Smirnov statistic, D for wind speed, humidity, and temperature throughout the southwest monsoon in 2020 for Melaka and Butterworth is shown in Table 14.

Table 14. Kolmogorov-Smirnov statistic (D) for wind speed, humidity, and temperature throughout the southwest monsoon season in 2020 for Melaka and Butterworth

Data	Melaka	Butterworth
Wind speed	0.1186	0.0811
Humidity	0.0480	0.0793
Temperature	0.0713	0.0700

From Table 14, all the D values are below the critical value for Melaka and Butterworth; hence the H_0 cannot be rejected. This indicates that Melaka and Butterworth's wind speed, humidity and temperature data throughout the southwest monsoon in 2020 can be assumed to be normally distributed. Therefore, the extended model in this study is normally distributed and can be used to describe the relationship between wind speed, humidity and temperature data throughout the southwest monsoon in 2020 for Melaka and Butterworth.

Q-Q plots for wind speed, humidity and temperature data in Melaka and Butterworth throughout the southwest monsoon in 2020 are constructed to show the data's goodness-of-fit to the normal distribution. Q-Q plot illustrates the data distribution. The points will fall on a reference line if the two data sets are from the normal distribution. The Q-Q plot for wind speed, humidity and temperature data in Melaka and

Butterworth throughout the southwest monsoon in 2020 are displayed in Figure 2, Figure 3, Figure 4, Figure 5, Figure 6, and Figure 7, respectively.

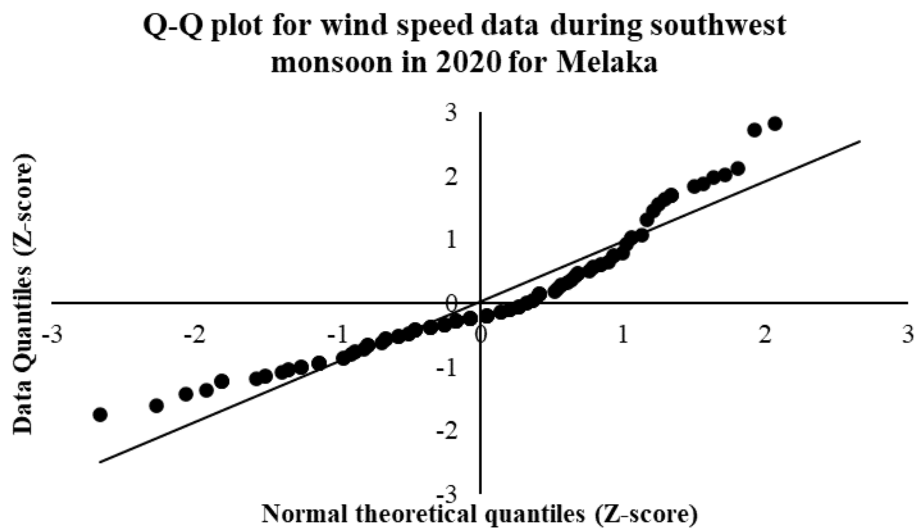

Figure 2. Q-Q plot for wind speed data during the southwest monsoon in 2020 for Melaka

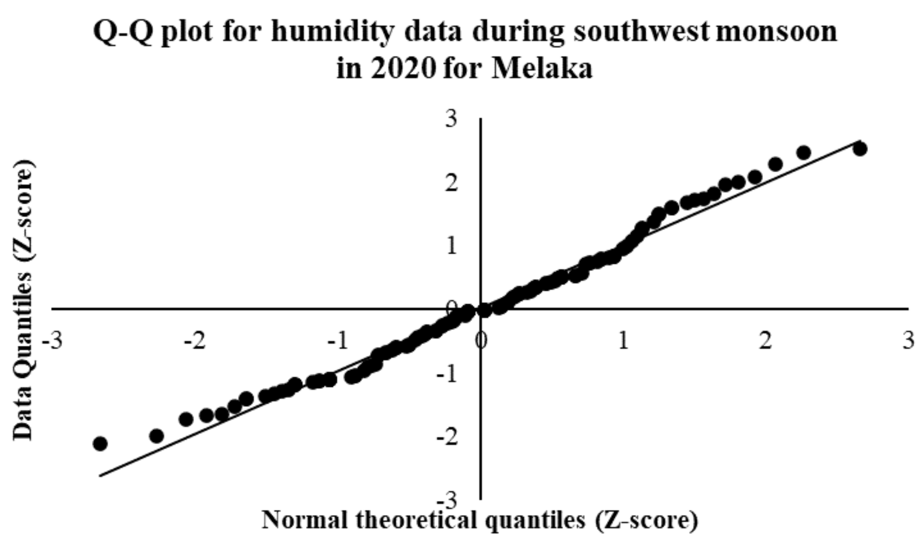

Figure 3. Q-Q plot for humidity data during the southwest monsoon in 2020 for Melaka

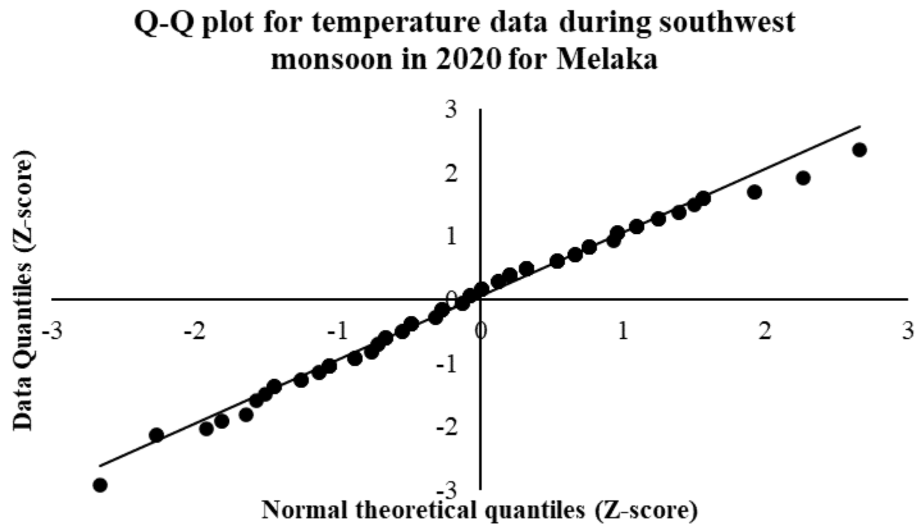

Figure 4. Q-Q plot for temperature data during the southwest monsoon in 2020 for Melaka

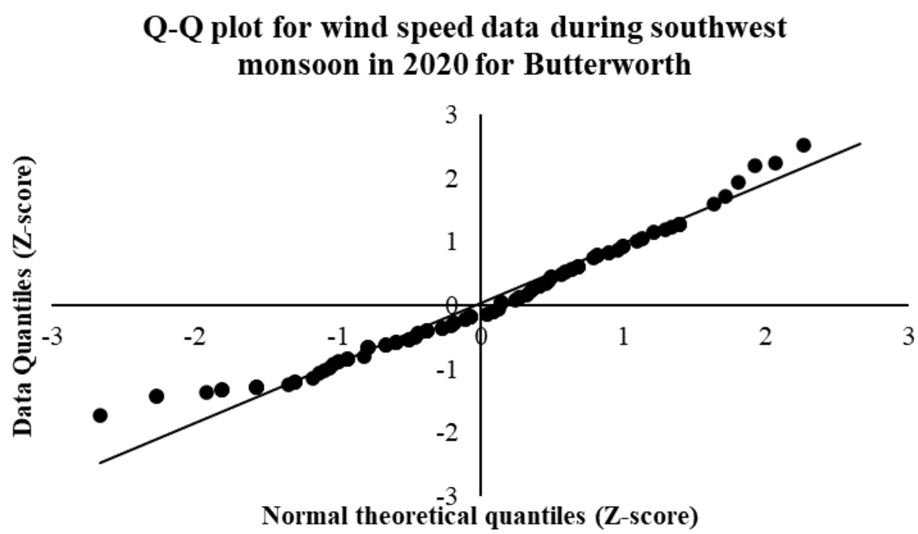

Figure 5. Q-Q plot for wind speed data during the southwest monsoon in 2020 for Butterworth

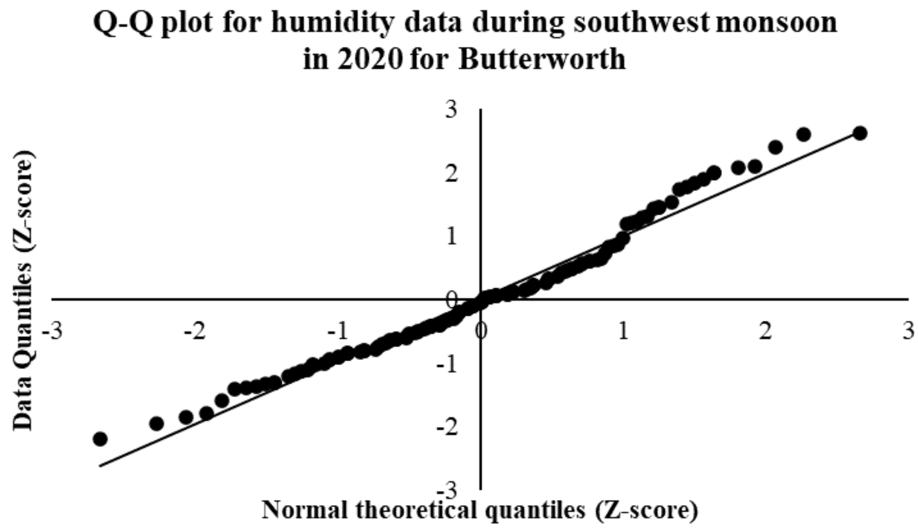

Figure 6. Q-Q plot for humidity data during the southwest monsoon in 2020 for Butterworth

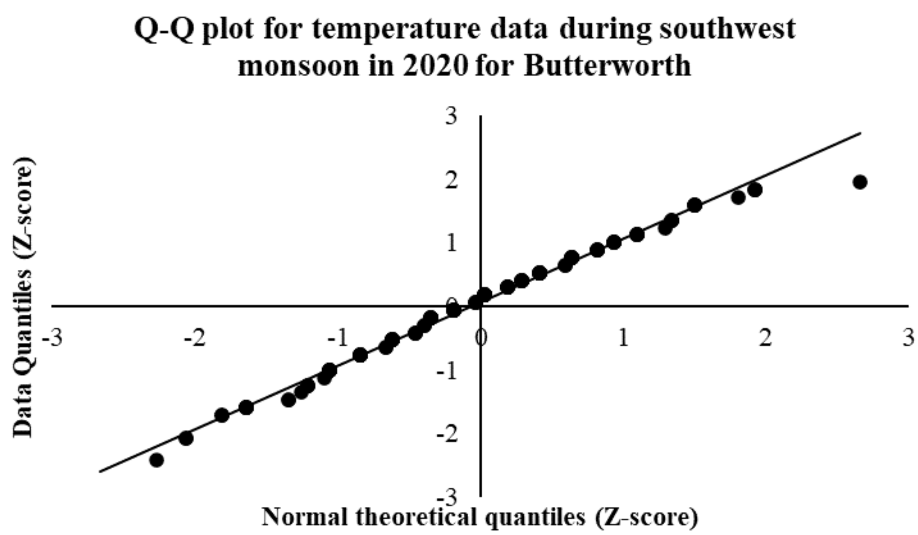

Figure 7. Q-Q plot for temperature data during the southwest monsoon in 2020 for Butterworth

Kolmogorov-Smirnov test and the Q-Q plot support that Melaka and Butterworth's wind speed, humidity, and temperature data during the southwest monsoon of 2020 will be treated with normal distribution. The detail for real data simultaneous LFRM for each station can be described as follows:

Step 1:Insert x_1, x_2, y_1, y_2, y_3 and y_4. Let $\lambda = 1$.

Step 2:Calculate the mean of x_1, x_2, y_1, y_2, y_3 and y_4.

Step 3:Fit the data by using simultaneous LFRM from equation (2).

Step 4:Calculate the parameter estimates $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\alpha}_3$, $\hat{\alpha}_4$, $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, $\hat{\beta}_4$, $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, $\hat{\sigma}_3^2$ and $\hat{\sigma}_4^2$.

Step 5: Calculate the Var ($\hat{\alpha}_1$), Var ($\hat{\alpha}_2$), Var ($\hat{\beta}_1$), Var ($\hat{\beta}_2$), Cov($\hat{\alpha}_1$, $\hat{\beta}_1$), and Cov($\hat{\alpha}_2$, $\hat{\beta}_2$), respectively.

Table 15 presents the parameter estimates for wind speed, humidity and temperature collected from Melaka and Butterworth during the southwest monsoons of 2020 when fitted with a simultaneous functional relationship model for linear variables.

Table 15. Parameter estimates of Melaka and Butterworth wind speed, humidity and temperature during the southwest monsoon 2020

	Melaka		Butterworth	
Variable	Humidity	Temperature	Humidity	Temperature
$\hat{\alpha}_j$	$\hat{\alpha}_1$ = 13.9429	$\hat{\alpha}_2$ = 28.9451	$\hat{\alpha}_3$ = 13.6790	$\hat{\alpha}_4$ = 29.1486
$\hat{\beta}_j$	$\hat{\beta}_1$ = 6.7436	$\hat{\beta}_2$ = -0.1217	$\hat{\beta}_3$ = 7.1225	$\hat{\beta}_4$ = -0.1059
σ_j^2	$\hat{\sigma}_1^2$ = 4.3346	$\hat{\sigma}_2^2$ = 0.7793	σ_3^2 = 4.9464	σ_4^2 = 0.6738
Var ($\hat{\alpha}_j$)	Var ($\hat{\alpha}_1$) = 1113.2165	Var ($\hat{\alpha}_2$) = 0.2005	Var ($\hat{\alpha}_3$) = 1259.2059	Var ($\hat{\alpha}_4$) = 0.1285
Var ($\hat{\beta}_j$)	Var ($\hat{\beta}_1$) = 11.5343	Var ($\hat{\beta}_2$) = 0.0020	Var ($\hat{\beta}_3$) = 14.1740	Var ($\hat{\beta}_4$) = 0.0014
Cov ($\hat{\alpha}_j$, $\hat{\beta}_j$)	Cov ($\hat{\alpha}_1$, $\hat{\beta}_1$) = 0	Cov ($\hat{\alpha}_2$, $\hat{\beta}_2$) =0	Cov ($\hat{\alpha}_3$, $\hat{\beta}_3$) = 0	Cov ($\hat{\alpha}_4$, $\hat{\beta}_4$) =0

From Table 15, the model for wind speed, humidity and temperature collected from Melaka and Butterworth during the southwest monsoons of 2020 are $y_1 = 13.9429 + 6.7436x_1$, $y_2 = 28.9451 - 0.1217x_1$, $y_3 = 13.6790 + 7.1225x_2$ and $y_4 = 29.1486 - 0.1059x_2$. Next, the variance of $\hat{\alpha}_1$, $\hat{\alpha}_3$, $\hat{\beta}_1$, and $\hat{\beta}_3$ are relatively high. A surprisingly high variance may prompt a closer examination of the data to ensure its quality. Outliers or errors in measurement could contribute to unusually high variance values. The variance of $\hat{\alpha}_2$, $\hat{\alpha}_4$, $\hat{\beta}_2$, and $\hat{\beta}_4$ are rather small and indicates good estimation for $\hat{\alpha}_2$, $\hat{\alpha}_4$, $\hat{\beta}_2$, and $\hat{\beta}_4$.

Conclusions

In conclusion, this paper extends parameter estimation from a simple linear functional relationship model (LFRM) to simultaneous LFRM for linear variables using the maximum likelihood estimation (MLE) method and the covariance matrix of the parameters is derived using the Fisher Information matrix. Results from the simulation study showed that the mean of the parameter estimate becomes closer to the real value of the parameter as we increase the value of sample size, n. Therefore, the estimation seems adequate for most values of n. The applicability of the simultaneous linear functional relationship model is demonstrated using Melaka and Butterworth wind speed, humidity, and temperature data throughout the southwest monsoon season in 2020 using this simultaneous linear functional relationship model (LFRM). Error terms are taken into consideration for the multivariate data. The MLE method estimates the wind speed, humidity and temperature parameters with errors in every variable. The simultaneous LFRM for wind speed, humidity and temperature collected from Melaka and Butterworth during the southwest monsoon of 2020 are $y_1 = 13.9429 + 6.7436x_1$, $y_2 = 28.9451 - 0.1217x_1$, $y_3 = 13.6790 + 7.1225x_2$ and $y_4 = 29.1486 - 0.1059x_2$. From the findings, we can study the relationship between more than two linear variables which are wind speed, humidity and temperature for two locations which are Melaka and Butterworth during the southwest monsoon season in 2020 with error considerations for each variable. Understanding these relationships is crucial for drawing meaningful conclusions in statistical analysis. This model can assist in managing outdoor activities while considering weather and safety by calculating the wind speed, mean relative humidity and temperature in Melaka and Butterworth throughout the southwest monsoon. This model may be used in future research to investigate the relationship between wind speed data, mean relative humidity, and temperature and to describe it as a simultaneous functional relationship model at several other locations with consideration of outliers.

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Acknowledgement

The authors gratefully acknowledge the Ministry of Higher Education Malaysia for supporting this study through the Fundamental Research Grant Scheme (FRGS/ 1/ 2021/ STG06/ UITM/ 02/ 5).

References

[1] Arif, A. M., Zubairi, Y. G., & Hussin, A. G. (2019). On Robust estimation for slope in linear functional relationship model. *Sains Malaysiana*, *48*(1), 237-242. https://doi.org/10.17576/jsm-2019-4801-27.

[2] Arif, A. M., Zubairi, Y. Z., & Hussin, A. G. (2020a). Parameter estimation in replicated linear functional relationship model in the presence of outliers. *Malaysian Journal of Fundamental and Applied Sciences*, *16*(2), 158-160. https://doi.org/10.11113/mjfas.v16n2.1633.

[3] Arif, A. M., Zubairi, Y. Z., & Hussin, A. G. (2020b). Parameter estimation in replicated linear functional relationship model in the presence of outliers. *Malaysian Journal of Fundamental and Applied Sciences*, *16*(2), 158-160. https://doi.org/10.11113/mjfas.v16n2.1633.

[4] Arif, A. M., Zubairi, Y. Z., & Hussin, A. G. (2021). COVRATIO Statistic for replicated linear functional relationship model. *Journal of Physics: Conference Series*, *1988*(1). https://doi.org/10.1088/1742-6596/1988/1/012100.

[5] Arif, A. M., Zubairi, Y. Z., & Hussin, A. G. (2022). Outlier detection in balanced replicated linear functional relationship model. *Sains Malaysiana*, *51*(2), 599-607. https://doi.org/10.17576/jsm-2022-5102-23.

[6] Arif, A. M., Zubairi, Yong Zulina, & Hussin, Abdul Ghapor. (2021a). Maximum likelihood estimation of replicated linear functional relationship model. *Applied Mathematics and Computational Intelligence*, *10*(1), 301-308.

[7] Fah, C. Y., Rijal, O. M., & Abu Bakar, S. A. R. (2010). Multidimensional unreplicated linear functional relationship model with single slope and its coefficient of determination. *WSEAS Transactions on Mathematics, 9*(5), 295-313.

[8] Ghapor, A. A., Zubairi, Y. Z., Mamun, A. S. M. A., & Imon, A. H. M. R. (2014). On detecting outlier in simple linear functional relationship model using covratio statistic. *Pakistan Journal of Statistics*, *30*(1), 129-142.

[9] Ghapor, A. A., Zubairi, Y. Z., Mamun, A. S. M. A., & Imon, A. H. M. R. (2015). A robust nonparametric slope estimation in linear functional relationship model. *Pakistan Journal of Statistics*, *48*(1), 339-350. https://doi.org/10.17576/jsm-2019-4801-27.

[10] Gillard, J. (2010). An overview of linear structural models in errors in variables regression. *Revstat Statistical Journal*, *8*(1), 57-80.

[11] Hawkins, D. I., & Kanji, G. K. (1995). 100 Statistical Tests. *Journal of Marketing Research*, *32*(1), 112. https://doi.org/10.2307/3152117.

[12] Hussin, A., Salleh, E., Chan, H. Y., & Mat, S. (2015). The reliability of predicted mean vote model predictions in an air-conditioned mosque during daily prayer times in Malaysia. *Architectural Science Review*, *58*(1), 67-76. https://doi.org/10.1080/00038628.2014.976538.

[13] Jamaliyatul, N. A. A. H., Badyalina, B., Mokhtar, N. A., Rambli, A., Zubairi, Y. Z., & Abdul Ghapor, A. (2023). Modelling wind speed data in Pulau Langkawi with functional relationship. *Sains Malaysiana*, *52*(8), 2419-2430.

[14] Kendall, M. G., & Stuart, A. (1973). *The Advanced Theory of Statistics.* London: Griffin.

[15] *Laporan Tahunan 2020* (p. 73). (2020). Jabatan Meteorologi Malaysia.

[16] Lindley, D. V. (1947). Regression Lines and the linear functional relationship. *Journal of the Royal Statistical Society*, *9*(2), 218-244.

[17] Lo Brano, V., Orioli, A., Ciulla, G., & Culotta, S. (2011). Quality of Wind speed fitting distributions for the urban area of Palermo, Italy. *Renewable Energy*, *36*(3), 1026-1039. https://doi.org/10.1016/j.renene.2010.09.009.

[18] Manteghi, G., Mostofa, T., & Md. Noor, M. P. B. (2020). A field investigation on the impact of the wider water body on-air, surface temperature and physiological equivalent temperature at Malacca Town. *International Journal of Environmental Science and Development*, *11*(6), 286-289. https://doi.org/10.18178/ijesd.2020.11.6.1264.

[19] Massey, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, *46*(253), 68. https://doi.org/10.2307/2280095.

[20] Mokhtar, N. A., Badyalina, B., & Zubairi, Y. Z. (2022b). Functional model of wind direction data in Kuching, Sarawak, Malaysia. *Applied Mathematical Sciences*,*16*(7), 349-357.

[21] Mokhtar, N. A., Zubairi, Y. Z., & Hussin, A. G. (2015). Parameter estimation of simultaneous linear functional relationship model for circular variables assuming equal error variances. *Pakistan Journal of Statistics*, *31*(2).

[22] Mokhtar, N. A., Zubairi, Y. Z., Hussin, A. G., Badyalina, B., Ghazali, A. F., Ya'Acob, F. F., Shamala, P., & Kerk, L. C. (2021a). Modelling wind direction data of Langkawi Island during Southwest monsoon in 2019 to 2020 using bivariate linear functional relationship model with von Mises distribution. *Journal of Physics: Conference Series*, *1988*(1). https://doi.org/10.1088/1742-6596/1988/1/012097.

[23] Nelder, J. (1977). Theory and application of linear-model-Graybill, FA. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, *140*, 384-385.

[24] Zakaria, M. N. (2022). The limitation of widely used data normality tests in clinical research. *Auditory and Vestibular Research*, *31*(1), 1-3. https://doi.org/10.18502/avr.v31i1.8127.