اونيۏرسيتي مليسيا ڤهڠ
UNIVERSITI MALAYSIA PAHANG

ORIGINAL ARTICLE

Modeling the impacts of climate change and air pollutants on the agricultural production yields in Malaysia using Random-Effects Error Components Regression model

Z.L. Chuan[1,*], S-.F. Fam[2], Q.H. Lee[3], J.S. Kok[3], M.N.B.M. Azam[1]

[1]Centre for Mathematical Sciences, College of Computing and Applied Sciences, Universiti Malaysia Pahang, Lebuh Persiaran Tun Khalil Yaakob, 26300, Kuantan, Pahang, Malaysia

[2]Technopreneurship Department, Faculty of Technology Management and Technopreneurship, Universiti Teknikal Melaka Malaysia, Hang Tuah Jaya, 76100 Melaka, Malaysia

[3]Faculty of Industrial Sciences and Technology, College of Computing and Applied Sciences, Universiti Malaysia Pahang, Lebuh Persiaran Tun Khalil Yaakob, 26300, Kuantan, Pahang, Malaysia

ABSTRACT – The occurrence of climate change is attributable to anthropogenic emissions of greenhouse gases (GHG) which have affected the C3 plants' agricultural production yields in past decades. Therefore, this article aims to model the linear association among these C3 plants' agricultural production yields with several climatic and non-climatic explanatory variables using one-way random-effects error components regression model. To be congruent with the main objective of this study, the balanced longitudinal dataset period 1980 to 2018 under big data was acquired. The analysis results revealed that merely maximum temperature (T_{max}) statistically significant detrimental affected the agricultural production yields, while vice versa for other explanatory variables. From a scientific perspective, temperature extreme can be massively detrimental affecting the pollination and developmental stages of C3 plants. Consequently, the high temperature decreases agricultural production yields. On the other hand, the proposed model in this study can be beneficial to national policymakers and smallholder plantations.

INTRODUCTION

Malaysia is a developing country in which the top three economic activities are including services, manufacturing, and agriculture with the main contribution to the national Gross Domestic Product (GDP). In principle, agricultural commodities such as oil palm (*Elaeis guineensis*), natural rubber (*Hevea brasiliensis*), cocoa (*Theobroma cacao*), and rice (*Oryza sativa*) have contributed respectively RM48.31 billion, RM14.53 billion, RM1.64 billion, and RM2.44 billion to the national GDP in the agricultural economic sector during 2020 [1-2]. Oil palm and natural rubber respectively have made Malaysia the second and fourth-largest worldwide exporter in the year 2020. According to the International Trade Statistics Database of United Nations COMTRADE, Malaysia's export of cocoa and cocoa preparations was US$1.48 billion in 2020 [3]. This report conveys that Malaysia plays a substantial role as a global cocoa exporter. On the other hand, rice is the staple food for Malaysians. However, the self-sufficiency rate (SSR) has dropped drastically from 69% to 63% in the year 2019 to 2020 due to the increase in the population [4-5]. This has alarmed the threat to Malaysian food security.

For Malaysia to be continuing prosperous in the agricultural sector and to achieve food security, extensive research on climate change and air pollution are excessively needed attention. The yields of agricultural commodities can be massively affected by climatic factors and air pollution. In specific, climate change is mostly attributable to anthropogenic emissions of greenhouse gases (GHG), which have impacted society and the economy such as food insecurity, human health, and damage to infrastructure [6]. The temperature rises, changes in rainfall patterns, high carbon dioxide (CO_2) concentration emissions, and extreme climatic events, including droughts, floods, heatwaves, and landslides have all favorable and detrimentally impacted agricultural commodities, creating vulnerability in the food supply. Unfortunately, there are limited previous studies involving Malaysia that have been carried out in modeling the impact of several climates' numerical variables such as temperature [7-8], rainfall amount [7-8], and the occurrence of extreme climatic events such as drought, flood, and sea-level rise [9] towards the agricultural production yields.

For instance, Murad *et al.* [10] analyzed the time series dataset to investigate the linear association between climate change and CO_2 air pollutant emission toward agricultural production yields using the ordinary least square (OLS) regression model. Their empirical analysis results revealed that climate change detrimentally affected the agricultural growth rate. Contrarily, CO_2 air pollutant emission has a beneficial effect on agricultural production yields. However, the agricultural production yields taken into account in this study do not state in their publication. Moreover, Chizari *et al.* [7] also analyzed the time series dataset to model the linear association between the climatic and non-climatic

variables towards the cocoa production yields using Auto-regressive Distributed Lag (ARDL) model. The climatic variables taken into account in their study are including the temperature and average annual rainfall, while the non-climatic variables are including the cocoa farm price, the fertilizer price, and the technology trend. Their empirical results showed that the rainfall amount has a statistically significant benefits effect on the cocoa production yields and vice versa for the fertilizer price. In contrast, there is statistically insignificant for the temperature, cocoa farm price, and technology trend.

Hazir et al. [9] carried out a preliminary study on projecting the effect of drought, flood, and sea-level rise on rubber production yields focusing on Peninsular Malaysia. Their study indicated that the occurrence of these hydro-meteorological events may not severely affect rubber production yields in Peninsular Malaysia. On the other hand, Tan et al. [8] carried out an empirical study to investigate the linear association between rice production yields and three climatic variables such as minimum and maximum temperatures and precipitation. However, their study found that precipitation is not statistically significant associated with the rice production yields in the wet and dry seasons, and vice versa for both minimum and maximum temperatures. In particular, both minimum and maximum temperatures respectively have a favorable and detrimental impact on the rice production yields during the wet and dry seasons. Houma et al. [11] also conducted an empirical study about the association between the climatic variables and farming practices regarding the rice production yields for the lowland located at the coastal plain of Kuala Selangor district. Their analysis results using the AquaCrop model revealed that the rising temperature, water scarcity, and worst weed control detrimentally impacted the rice production yields. In addition, their findings also indicated that the rice production yields may be increased with the appropriate adoption of weed control and water management in the condition of the rise of temperature.

In a similar year, Roslan et al. [12] proposed the predictive models for forecasting the paddy production yields using elliptical (normal and t) and Archimedean (Joe, Clayton, and Gumbel) copula families for five selected the Association of Southeast Asian Nations (ASEAN) countries such as Indonesia, Malaysia, Myanmar, Thailand, and Vietnam. The climatic and non-climatic variables taken into account in their proposed models are including an average annual maximum temperature, annual rainfall amount, planted areas, and fertilizer usage. Their findings conveyed that the best-fitted copula model in modeling and forecasting the rice production yields for Indonesia, Malaysia, Myanmar, Thailand, and Vietnam respectively are normal, Gumbel, Clayton, Gumbel, and Gumbel copulas. Moreover, this study indicated that these best-fitted predictive models require to be included all the aforementioned climatic and non-climatic variables. Meanwhile, Abubakar et al. [13] carried out systematically reviewed studies in the context of climate change towards oil palm production yields. Their analytical results showed that those climatic and non-climatic variables, including temperature, rainfall, extreme climatic events such as El Niño, La Niña, drought and flooding, soil fertility, water management, fertilizer use, the emergence of disease can detrimentally impact the oil palm production yields.

In brief, previous Malaysian studies merely focused on a single type of agricultural production yields such as oil palm, natural rubber, cocoa, and rice. In principle, a large number of Malaysia's studies, merely respectively focused on oil palm and rice yields can be found in reputable databases [13] rather than natural rubber and cocoa. Therefore, the main purpose of this present article is to model the linear association between the several annual agricultural production yields (tonnes per hectare) and several climatic, air pollutant, and planted areas (PA) (in thousand hectares) variables using an error components regression model, which the dataset employed in this study is under big data [14-15]. In categorizedspecific, the agricultural production yields taken into account in this study are including oil palm, natural rubber, cocoa, and rice, which are four agricultural categorised as C3 plants, and the main contributors to the agricultural economic sector in Malaysia's national GDP. Moreover, Malaysia also plays a substantial role as an exporter of these three agricultural commodities across the globe with exception of rice. Meanwhile, the climatic, and air pollutant variables taken into account in this study are the average annual minimum (T_{min}), and maximum (T_{max}) temperatures (in degrees Celsius), annual rainfall amount (RA) (in millimetres), and annual CO_2 emissions (in knots). The regression model yielded in this study is competent to provide insight into the national impacts of recent climate trends on the agricultures' production yields and would aid in anticipating the impacts of climate changes on food security in the future. Furthermore, the proposed model also can be beneficial to the smallholder agriculture plantations in adopting the future plantation strategy based on the predicted insights of the effects of climate change on the C3 plants' agricultural production yields.

RESEARCH METHODOLOGY AND THEORETICAL BACKGROUND

In achieving the main purpose of this study, this section provided an overview of the research methodology and the theoretical background of the predictive model employed in this article. In particular, Figure 1 depicted the schematic of the research methodology involved in this study.
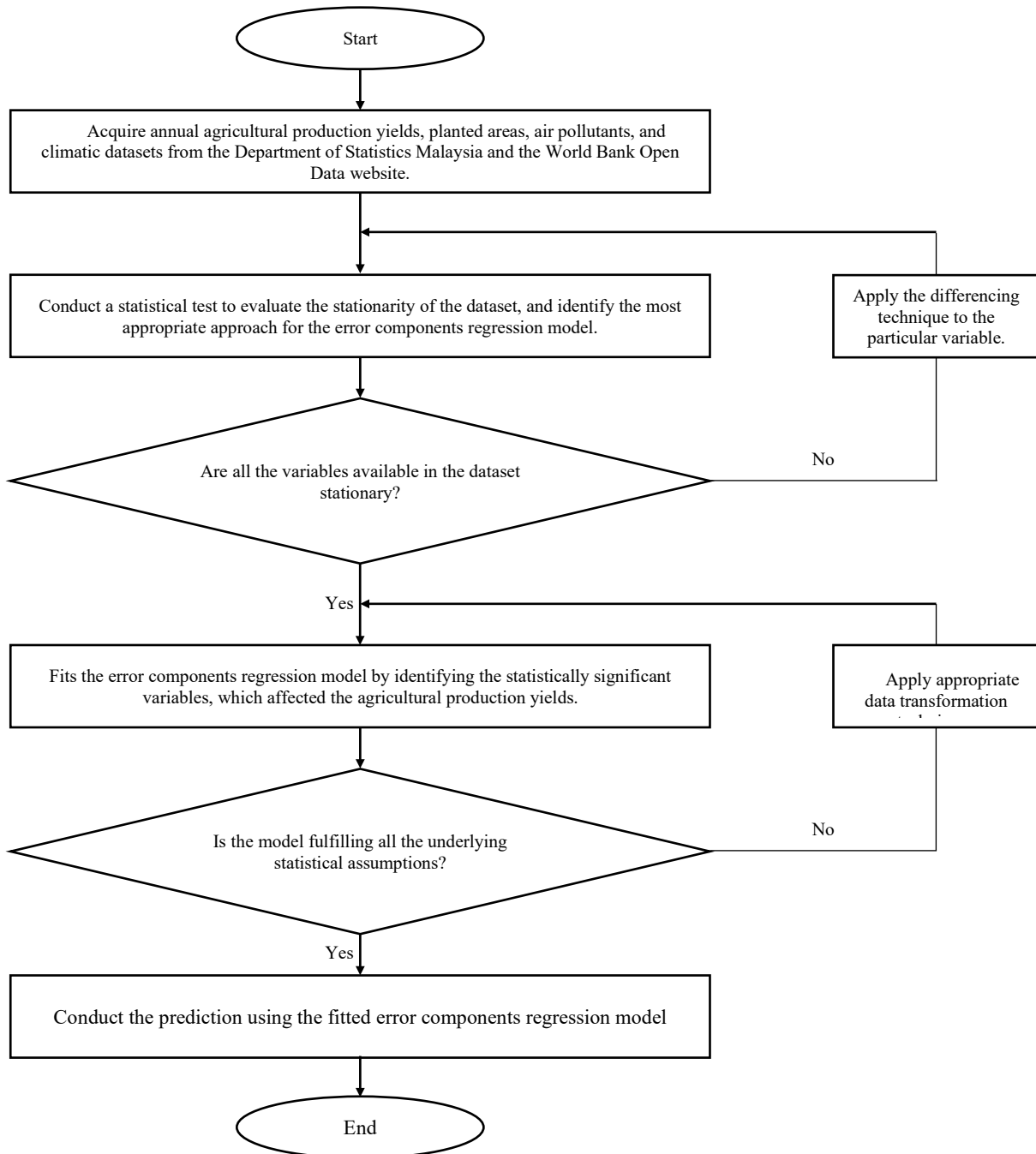
```
                            ┌─────────────┐
                            │    Start    │
                            └─────────────┘
                                   │
        ┌──────────────────────────────────────────────────────┐
        │ Acquire annual agricultural production yields, planted│
        │ areas, air pollutants, and climatic datasets from the │
        │ Department of Statistics Malaysia and the World Bank   │
        │ Open Data website.                                    │
        └──────────────────────────────────────────────────────┘
                                   │
        ┌──────────────────────────────────────────────────────┐    ┌──────────────────┐
        │ Conduct a statistical test to evaluate the stationarity│   │ Apply the        │
        │ of the dataset, and identify the most appropriate     │    │ differencing     │
        │ approach for the error components regression model.   │    │ technique to the │
        └──────────────────────────────────────────────────────┘    │ particular       │
                                   │                                  │ variable.        │
                                   ▼                                  └──────────────────┘
                      Are all the variables available        No
                      in the dataset stationary?  ──────────────────────►
                                   │
                                  Yes
                                   ▼
        ┌──────────────────────────────────────────────────────┐    ┌──────────────────┐
        │ Fits the error components regression model by         │    │ Apply appropriate│
        │ identifying the statistically significant variables,  │    │ data transformation│
        │ which affected the agricultural production yields.    │    │                  │
        └──────────────────────────────────────────────────────┘    └──────────────────┘
                                   │
                                   ▼
                      Is the model fulfilling all the         No
                      underlying statistical assumptions? ──────────────────►
                                   │
                                  Yes
                                   ▼
        ┌──────────────────────────────────────────────────────┐
        │ Conduct the prediction using the fitted error         │
        │ components regression model                           │
        └──────────────────────────────────────────────────────┘
                                   │
                            ┌─────────────┐
                            │     End     │
                            └─────────────┘
```

Figure 1. The schematic of the research methodology in modelling the impacts of climatic change and air pollutant emissions

DATA SOURCES AND STUDY AREAS DESCRIPTION

As the preliminary step of this study, the balanced longitudinal dataset under big data, which comprises climatic and non-climatic explanatory variables period 1980 to 2018 has been acquired from the Department of Statistics Malaysia (DOSM) and the World Bank Open Data website. In particular, all the C3 plants' agricultural production yields, and the corresponding PA variables have been acquired from the Department of Statistics Malaysia website. Meanwhile, the variables such as T_{min}, T_{max}, RA, and CO_2 are acquired from the World Bank Open Data website. The acquired longitudinal dataset comprises all four C3 plants, while the corresponding C3 plants' agricultural production yields are the response variable for the error components regression model.

Planetary, Malaysia is the 66th largest nation covering approximately a total of land and water areas respectively 328,657 and 1,190 square kilometres located practically to the equator of the earth planet. Therefore, Malaysia experienced an equatorial climate, which is hot and humid with a uniform annual ground temperature [8], [16].

Malaysia also experienced dry and wet seasons during the Southwest (May-August) and Northeast monsoons (November-March), respectively. In particular, Malaysia received a relatively small annual rainfall amount during the Southwest monsoon. Contrarily, Malaysia encountered the risk of occurrences of extreme rainfall events, which are frequently associated with short-duration thunderstorms, especially in the East Coast region [6], [16].

From the geographical perspective, Malaysia is composed of two main non-contiguous regions such as Peninsular Malaysia and Malaysian Borneo, which these two regions are isolated by the South China Sea as depicted in Figure 2. In particular, Peninsular Malaysia can be sub-divided into four main regions, including Central (Kuala Lumpur, Putrajaya, Selangor), East Coast (Kelantan, Pahang, Terengganu), Northern (Kedah, Perak, Perlis, Pulau Pinang), and Southern (Johor, Melaka, Negeri Sembilan). Meanwhile, Malaysian Borneo can be sub-divided into two main regions, including Sabah (Labuan, Sabah), and Sarawak. In principle, all the agricultural production yields taken into account in this article are mainly contributed from the selected states of Malaysia due to the appropriateness of biophysical plantations [17-18]. For instance, the main plantation areas for oil palm, and cocoa respectively are the Malaysian Borneo region as summarised in Figure 3 [1]. In contrast, the main plantation areas for natural rubber, and paddy are Peninsular Malaysia, where 85.5% of the paddy is from Peninsular Malaysia [18-19].
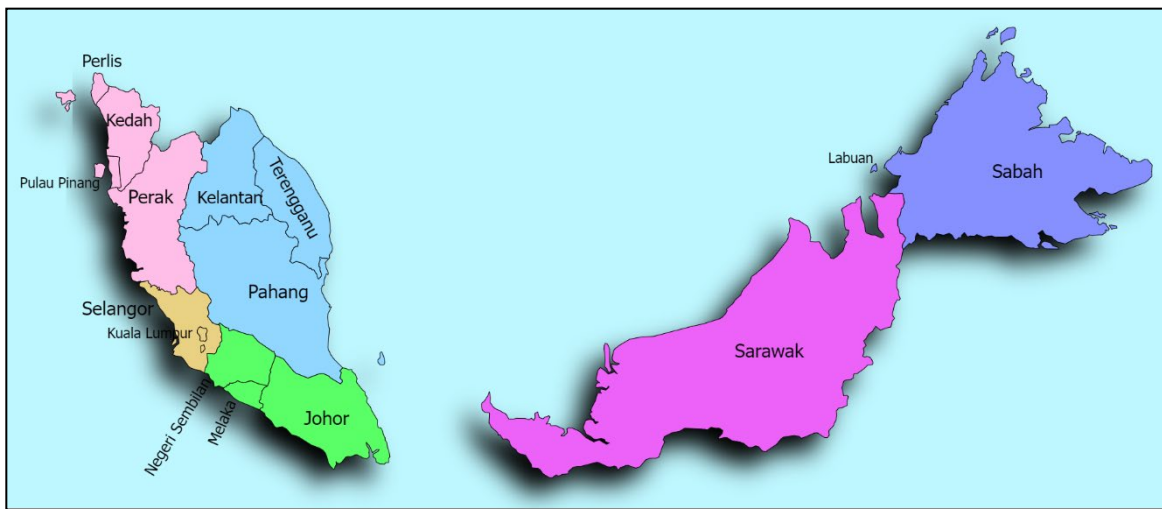
Figure 2. The geographical insight of 6 main sub-divided regions in Malaysia

(a)   (b)

(c)
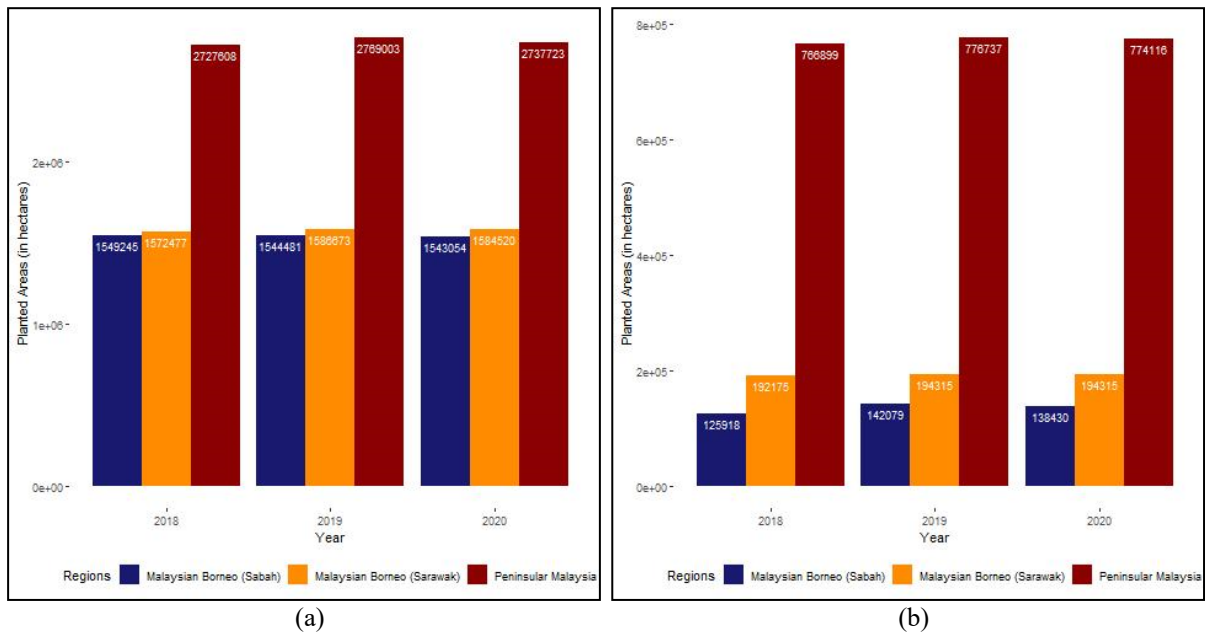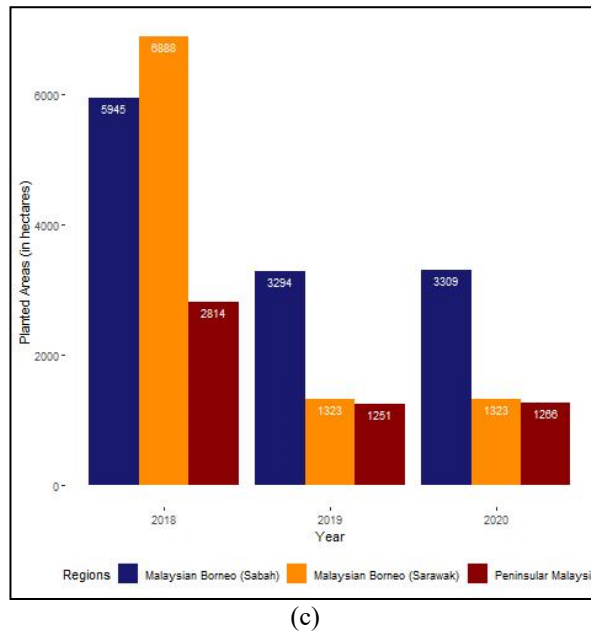
Figure 3. The planted areas (in hectares) of (a) oil palm, (b) natural rubber, and (c) cocoa in Peninsular Malaysia and Malaysian Borneo regions period 2018 to 2020
Source: Ministry of Plantation Industries and Commodities, 2020

ONE-WAY RANDOM-EFFECTS ERROR COMPONENTS REGRESSION MODEL

In the worldwide previous studies [8], [20-23], an error components regression model has been widely employed in analysing the longitudinal datasets with one of the main focused research areas on climate change towards agricultural yields. This is due to the aforementioned statistical model which can be taken into account both cross-sectional and temporal dimensions simultaneously. In principle, there are three main approaches for error components regression models. These include pooled least squares, fixed-effects, and random-effects models. For fitting a precise error components regression model, this study has employed Hausman [24] and Lagrange Multiplier [25] statistical tests. The analysis results for these two statistical tests (Hausman test: p-value = 0.9956; Lagrange Multiplier test: p-value = 0.0372) consistently led to the random-effects model being fitted to the longitudinal dataset in this study. In addition, this study also employed Fisher Augmented Dickey-Fuller (Fisher-ADF) statistical test [26] in investigating the panel unit root with and without trends corresponding to individual series for the acquired dataset. From a statistical perspective, the failure of the rejection for the null hypothesis of the Fisher-ADF statistical test implied that the individual series required differencing transformation, such that the individual series is stationary.

Moreover, suppose that $Y = [y_{st}]_{ST \times 1}; s, (t) = 1, 2, \cdots, S, (T)$ represents a vector of length $S \times T$, which comprises S types of C3 agricultural plants' production yields with each type of plant having a T length of time series. In addition, denotes that $X = [X_k]_{K \times 1}; k = 1, 2, \cdots, K$ represents a $K \times 1$ vector with K explanatory variables, which $X_k = [x_{st}]_{ST \times 1}$. In mathematics, the one-way random-effects error components regression model can be expressed as equation (1).

$$Y = X\theta + \epsilon \tag{1}$$

where $\theta = [\theta_k]_{K \times 1}$ represents the corresponding regression coefficients of X_k, and $\epsilon = [\epsilon_{st}]_{ST \times 1}; \epsilon_{st} = \varepsilon_s + \varepsilon_{st}$ represents the vector of length $S \times T$ for the disturbance. In particular, ε_s is the unobservable specific effect of the C3 plants which is constant through time, and ε_{st} is the remainder disturbance which varies among the C3 plants regardless at a given time point and through time.

In this study, θ of equation (1) is estimated using the best linear unbiased estimator, namely the generalised least square (GLS) estimator in relating Swamy-Arora estimator of error components [27-28]. The main reason the Swamy-Arora estimator is employed in this study due to this estimator has been widely employed in previous socio-economic studies [29-31], which involved the random-effect error components regression model. Furthermore, the error components regression model is applied in this study rather than the time series model due to the several advantages of this model. These are including the capability in controlling for the cross-sectional and temporal-invariant variables, being more efficient and reliable in parameters estimation in the models, and the capability to identify and detect the unobserved heterogeneity when the heterogeneity is constant over time and uncorrelated with explanatory variables [29], [31-32].

DIAGNOSTICS CHECKING

In equation (1), the disturbance has been assumed to be $\epsilon_{st} = \varepsilon_s + \varepsilon_{st}$, such that $E(\varepsilon_s) = 0$, $E(\varepsilon_{st}) = 0$, $E\left(\varepsilon_{s_i}\varepsilon_{s_j}\right) = \begin{cases} \sigma_s^2, & s_i = s_j \\ 0, & s_i \neq s_j \end{cases}$, and $E\left(\varepsilon_{s_i t_i}\varepsilon_{s_j t_j}\right) = \begin{cases} \sigma_{st}^2, & s_i = s_j; t_i = t_j \\ 0, & \text{otherwise} \end{cases}$, where $i, (j) = 1,2,\cdots,S,(T)$, and ε_s and ε_{st} assumed to be unknown and independent to each other. In particular, $\varepsilon_s \sim IID(0, \sigma_s^2)$ and $\varepsilon_{st} \sim IID(0, \sigma_{st}^2)$ with respectively independent and identically (IID) distributed with zero mean and unit variance. To validate the effectiveness of the best-fitted random-effects error components regression model, this study has employed Breusch-Godfrey and Pesaran's CD statistical test in diagnosing the serial correlation for the idiosyncratic component of the disturbance and the cross-sectional dependence, respectively. Moreover, this study also employed the Breusch-Pagan statistical test to validate the assumption of homogeneity of variance is fulfilled. The main objective of this diagnostic checking is to ensure that all the aforementioned assumptions of the best-fitted one-way random-effects error components regression model in this study are fulfilled. In contrast, the failure of the fitted model in fulfilling these statistical assumptions can be detrimental to the prediction performance. The analysis results of this diagnosis checking are presented in the next section.

RESULTS AND DISCUSSION

To be congruent with the main purpose and methodology presented in this present article, this section provided the interpretation and discussion of this research findings. In specific, this section is subdivided into two main sections, such as exploratory data analysis, and statistical modelling and diagnostic checking. All the statistical analysis presented in this section has been fully analysed using R statistical software for computing and graphics.

Exploratory data analysis

Figure 4 and Table 1 depicted the time series graphs and the descriptive summaries for all individual series periods 1980 to 2018 taken into account in this study, respectively. These include the C3 plants' production yields (oil palm, natural rubber, cocoa, and rice), climatic (T_{min}, T_{max}, RA, and CO_2), and non-climatic (PA corresponding to each C3 plant) variables. This study highlighted that climatic variable such as solar radiation is excluded as the explanatory variable in the regression analysis although this variable can affect the agricultural production yields. This is due to the previous study [33] conveyed that the inclusion of this explanatory variable into the regression model is affected by the presence of multicollinearity issues. In particular, solar radiation is highly positively associated with temperature.

Moreover, Figure 4 characterised that the oil palm yields (Figure 4(a)) and rice (Figure 4(d)) productions averagely have increased over the year, and vice versa for the natural rubber (Figure 4(b)) and cocoa (Figure 4(c)). As a result of the collapse in natural rubber and cocoa prices, as well as the growing demand for oil palm globally at the same time, the smallholder farmers have shifted their plantation crops primarily focus on the oil palm [17], [34]. In addition, the high demand for oil palm worldwide also resulted in the deforestation for the expansion of the plantation of oil palm [17], [34]. Simultaneously, deforestation in Malaysia brought an increase in temperature and CO_2 emission, which is evidenced in Figures 4(e), 4(f), and 4(h), respectively. Inlines with the increase of plantation for oil palm, and decreasing of plantation for natural rubber and cocoa, the PA for the oil palm (Figure 4(i)) has been increased, while vice versa for the PA of natural rubber (Figure 4(j)) and cocoa (Figure 4(k)). On the other hand, Figure 4(l) depicted that the PA for rice has increased due to the increase of the population, as well as rice, being the staple of Malaysian cuisine. Moreover, Figure 4(g) also illustrates the total annual rainfall amounts have been averagely increased during the periods 1980 to 2018.

Meanwhile, Table 1 depicted the descriptive summaries of the individual series respectively characterised using the appropriate statistical descriptive measurements such as median (MD), interquartile range (IQR), skewness (SKEW), and kurtosis (KURT). This study has characterised the average and dispersion of the individual series using MD and IQR, respectively. Since the Shapiro-Wilks statistical test (SWST) indicated that some of the individual series, including oil palm and cocoa production yields, CO_2, and PA for natural rubber and cocoa are non-normal distributed, therefore, MD and IQR are robust statistical descriptive measurements rather than the arithmetic mean and variance, respectively. For instance, the skewness value for the oil palm production yields revealed that the shape of the distribution for this series is negatively skewed. However, the box-and-whisker plot indicated that the distribution of this series is positively skewed with the presence of potential outliers. This contradiction showed the not robustness of the arithmetic mean to the potential outliers, which the involvement of the arithmetic mean as the reference point in computing the moments of skewness. In mathematical, $\text{SKEW} = \left(\frac{\sqrt{T}\sum_{t=1}^{T}(y_t-\bar{y})^3}{\sum_{t=1}^{T}(y_t-\bar{y})^2}\right)$.

In practice, Malaysia plays a substantial role as the exporter of both oil palm and natural rubber production yields in the world. Therefore, the average oil palm and natural rubber production yields are higher compared to cocoa and rice. In contrast, Table 1 presented that the rice production yields are lower than the natural rubber, however, the descriptive analysis conveyed that the average of the planted areas for rice is higher rather than the natural rubber, as well as rice, is the staple food for most of Malaysia's population. In specific, the daily consumption of Malaysia's adults on white rice is 2.5 plates on average [35]. On the other hand, the temperature (T_{min} and T_{max}) showed the lowest variation among the individual series depicted in Table 1. This is due to Malaysia experiencing a uniform annual ground temperature as its geographical location at the equator of the earth's planet.
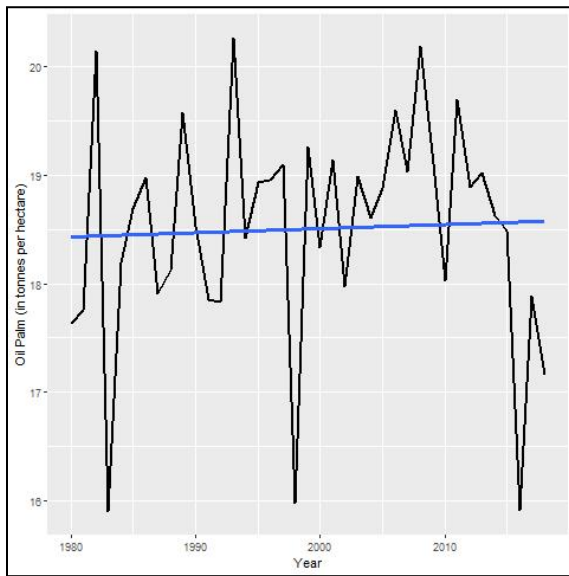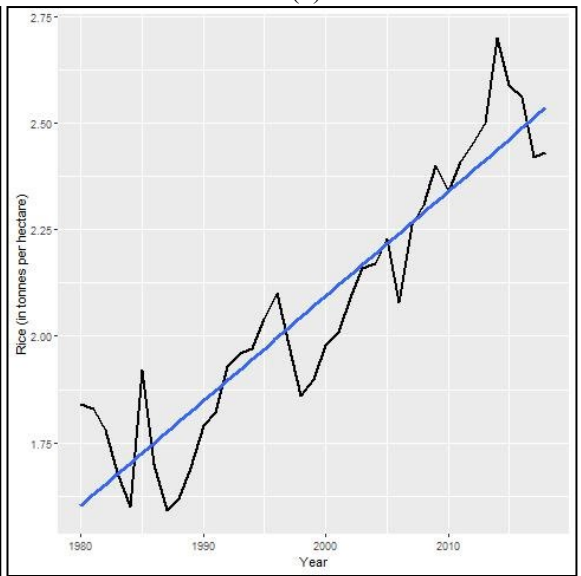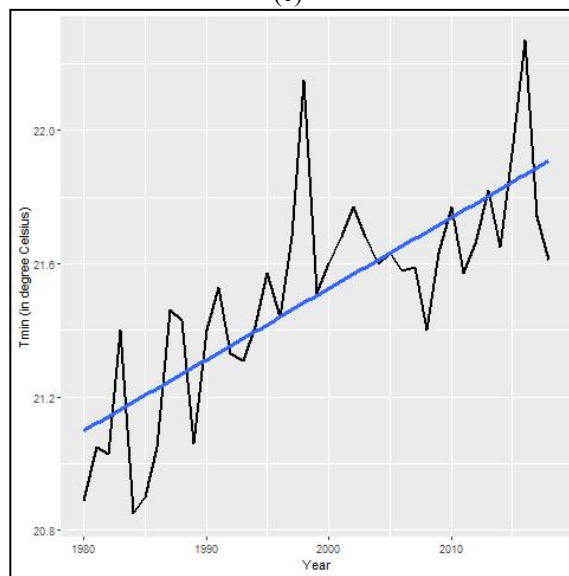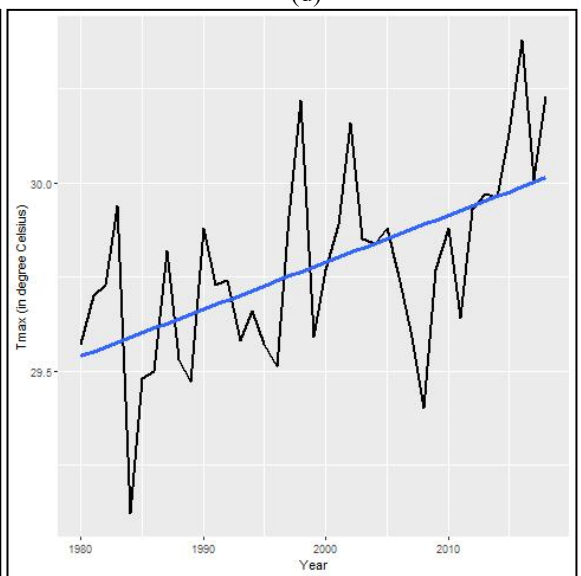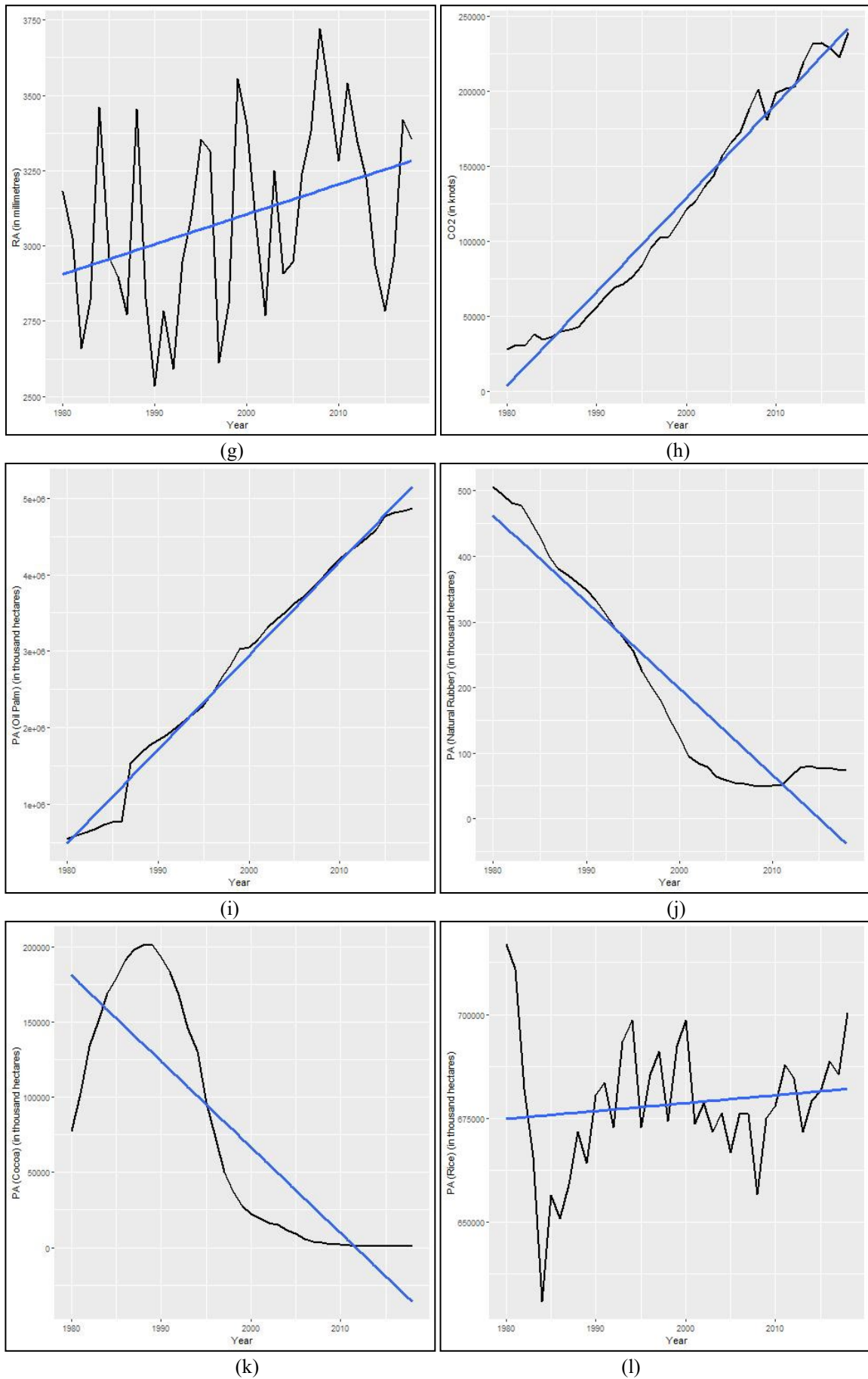
(a)
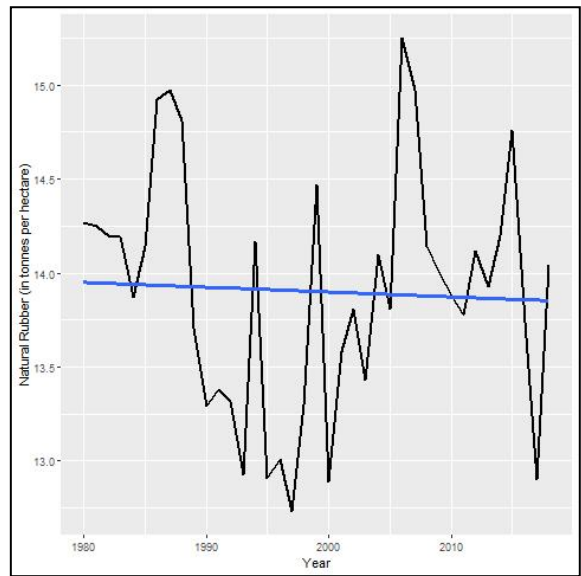
(b)
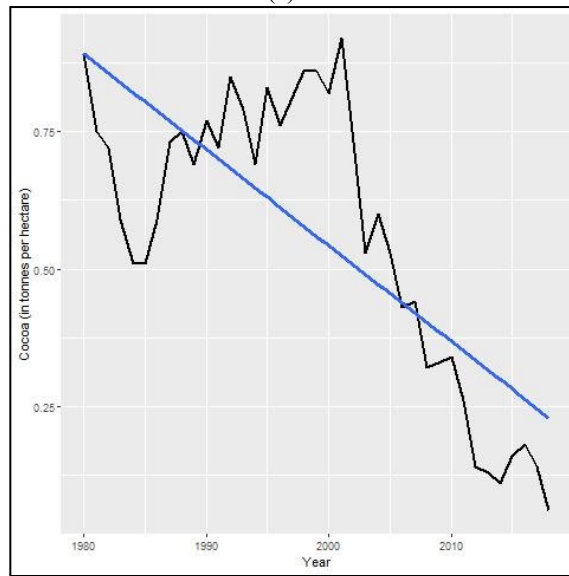
(c)

(d)

(e)

(f)

Figure 4. The trend of the individual series periods 1980 and 2018 corresponding to (a) oil palm production yields; (b) natural rubber production yields; (c) cocoa production yields; (d) rice production yields; (e) T_{min}; (f) T_{max}; (g) RA; (h) CO_2; (i) $PA_{oil\ palm}$; (j) $PA_{natural\ rubber}$; (k) PA_{cocoa}; and (l) PA_{rice}

Table 1. The descriptive summaries use the statistical measurements corresponding to the individual series

Individual Series	MD	IQR	SKEW	KURT	p-value$_{SWST}$
Oil palm production yield	18.6300	1.1250	-0.7982	0.6324	0.0141*
Natural rubber production yield	13.9300	0.7950	0.0522	-0.7068	0.2866
Cocoa production yield	0.6000	0.4300	-0.4908	-1.1885	0.0034*
Rice production yield	2.0100	0.4900	0.2579	-1.0562	0.2183
T_{min}	21.5700	0.2700	-0.1699	0.0651	0.0669
T_{max}	29.7700	0.3350	0.0831	0.0090	0.8508
RA	3086.9100	531.8350	0.0264	-1.1792	0.2156
CO_2	111820.0000	140688.8990	0.1859	-1.5155	0.0029*
PA (Oil palm)	3026880.0000	2206567.0000	-0.1696	-1.2978	0.0228
PA (Natural rubber)	151.2000	280.5500	0.4922	-1.3635	$\approx$ 0.0000*
PA (Cocoa)	27937.0000	145220.0000	0.5255	-1.4672	$\approx$ 0.0000*
PA (Rice)	677884.0000	14956.0000	-0.2291	0.9128	0.5093

Note: "" represents the presence of a statistically significant in the SWST.

Furthermore, verifying the individual series exhibited the statistical properties that the individual series is constant over time, which is also known as stationary is indeed much needed. This is because stationary is one of the preliminary statistical assumptions required to fulfil in fitting the one-way random-effects error components regression model. Hence, this study employed Fisher-ADF statistical test with and without trends in verifying this preliminary assumption of the individual series as the analysis results are depicted in Table 2. Based on Table 2, all the individual series have fulfilled the stationary assumption, as well as all the *p*-values, which are less than the significance level of 0.05 except for CO_2. In ensuring the individual series of CO_2 achieved stationary, this study employed the differencing transformation. The Fisher-ADF statistical test authenticated this series has fulfilled the stationary assumptions after applying the first-order of differencing transformation.

Table 2. The analysis results of the panel unit root statistical test with and without trend

Individual series	Panel unit root (*p*-value)	
	Fisher-ADF (With trend)	**Fisher-ADF (Without trend)**
Yield	$\chi^2 =$ 52.6620 ($\approx$ 0.0000)	$\chi^2 =$ 51.1160 ($\approx$ 0.0000)
T_{min}	$\chi^2 =$ 31.3540 (0.0001)	$\chi^2 =$ 66.0750 ($\approx$ 0.0000)
T_{max}	$\chi^2 =$ 45.5520 ($\approx$ 0.0000)	$\chi^2 =$ 73.6830 ($\approx$ 0.0000)
RA	$\chi^2 =$ 56.3780 ($\approx$ 0.0000)	$\chi^2 =$ 60.2050 ($\approx$ 0.0000)
ΔCO_2	$\chi^2 =$ 188.0500 ($\approx$ 0.0000)	$\chi^2 =$ 182.9200 (0.0000)
PA	$\chi^2 =$ 32.9630 ($\approx$ 0.0000)	$\chi^2 =$ 23.9110 (0.0024)

*Note: Δ represents the first-order differencing transformation

Statistical modelling and diagnostics checking

The fulfilment of the preliminary stationary assumption in the previous sub-section allowed for a further analysis, which is fitting the one-way random-effects error components regression model. The analysis result for the full model is presented in Table 3.

Table 3. The analysis result of the full model for the one-way random-effects error components regression model

Exploratory Variable	Estimated regression coefficient	Statistical test value (p-value)
Intercept	28.7220	$z =$ 3.1248 (0.0018)*
T_{min}	0.0944	$z =$ 0.3038 (0.7612)
T_{max}	−0.7433	$z =$ −1.8028 (0.0714)
RA	≈ 0.0000**	$z =$ 0.0647 (0.9484)
ΔCO_2	≈ 0.0000**	$z =$ 0.5179 (0.6045)
PA	≈ 0.0000**	$z =$ 1.2715 (0.2035)

*Note: "**" indicated the estimated regression coefficient is approximate zero after rounding to the four nearest decimal places; "*" represents the presence of a statistically significant in the statistical test.

In particular, the analysis result for this study showed that merely T_{max} has a statistically significant detrimental impact on agricultural production yields after eliminating started from the most statistically insignificant explanatory variable from the model. In mathematics, the result of the best-reduced model for one-way random-effects error components regression can be expressed as equation (2).

$$\hat{y} = 25.0917 - 0.5485T_{max} \tag{2}$$

This analysis results lead that the average agricultural production yields are expected to decrease by 0.5485 tonnes per hectare for a 1 degree Celcius rise in T_{max}, while T_{min}, RA, ΔCO_2, and PA do not have statistically significant effects on agricultural production yields. However, the diagnostic checking analysis presented that there are violated the assumptions of the one-way random-effects error components regression model as depicted in Table 4. These include the presence of the serial correlation for the idiosyncratic component of the disturbance and no cross-sectional dependence. To overcome these issues, this study employed the Box-Cox transformation technique.

Table 4. The diagnostic checking for the resulting one-way random-effects error components regression model

Statistical Test	Model	
	Before transformation *Statistical test value (p-value)*	**After transformation** *Statistical test value (p-value)*
Breusch-Godfrey	$\chi^2 =$ 58.6110 (0.0226)*	$\chi^2 =$ 51.6420 (0.0847)
Pesaran's CD	$z =$ −2.2985 (0.0215)*	$z =$ −2.2985 (0.0215)
Breusch-Pagan	$\chi^2 =$ 1.6666 (0.1967)	$\chi^2 =$ 2.3346 (0.1265)

Note: "" represents the assumption that is violated.

Equation (3) is the resulting one-way random-effects error components regression model after applying the Box-Cox transformation technique.

$$\hat{y}^2 = 753.9490 - 20.7711T_{max} \tag{3}$$

with the resulting model has fulfilled all the assumptions of the one-way random-effects error components regression model (Table 4). In particular, T_{max} has detrimentally affected agricultural yields, which this finding is inlined with previous studies [8], [11], [13]. This finding also can be supported based on a practical perspective. In a natural science study, the pollination of plants is one of the most susceptible phenological stages to temperature extremes, which affect pollen viability, fertilisation and grain or fruit formation [36]. In addition, extreme temperature also can massively detrimentally affect C3 plants' agricultural production yields during the developmental stage [36]. As a result, the adoption of more effective strategies such as utilising a more appropriate irrigation system is indeed much needed to mitigate the impacts of greater temperature extremes events associated with climate change.

CONCLUSION

In summary, this study proposed a prediction of agricultural production yields using the one-way random-effects error components regression model. The analysis results revealed that T_{max} is statistically significant affected towards the agricultural production yields such as oil palm, natural rubber, cocoa, and rice. The proposed model in this study can be beneficial to national policymakers and smallholder farmers. In particular, the national policy can use the proposed model to predict the future climate change impact on the agricultural production yields to anticipate the impacts of future climate changes on food security, while smallholder farmers can use the proposed model in adapting the future plantation strategy. However, the goodness-of-fits measurement followed by the prediction accuracy of the proposed model can be improved by taking into account more climatic and non-climatic variables as well as the intercept is indicated as statistically significant to the resulting predictive model. These climatic variables are including relative humidity, wind speed, and sunshine duration, while the non-climatic variables are including the emergence of disease, labour force skills, soil fertility, water management, weed control, and fertiliser use. This is because the previous empirical Malaysia studies [7], [11-13], [22] conveyed that these variables also can be affected towards agricultural production yields. In future, this article suggested conducting a comparative study in fitting the model respectively taking into account the temporal and both cross-sectional and temporal with focusing on a single agricultural production yield.

ACKNOWLEDGEMENT

REFERENCES

[1]   Ministry of Plantation Industries and Commodities, *Q4/2020 pocketstats agricommodity*. Putrajaya, Malaysia: Ministry of Plantation Industries and Commodities, 2020.

[2]   Statista, "Gross domestic product (GDP) from paddy in Malaysia from 2016 to 2020," Available: https://www.statista.com/statistics/952735/malaysia-gdp-from-rice-industry/, 2021. [Accessed Jan. 25, 2022].

[3]   Trading Economics, "Malaysia exports of cocoa and cocoa preparations," Available: https://tradingeconomics.com/malaysia/exports/cocoa-cocoa-preparations, 2022. [Accessed Jan. 28, 2022].

[4]   Department of Statistics Malaysia, "Supply and utilization accounts selected agricultural commodities, Malaysia 2015-2019," Available: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=164&bulid=OTM1TDMzS1IvYm5mU1JiU1Fwekt3UT09&menu_id=Z0VTZGU1UHBUT1VJMFlpaXRRR0xpdz09, 2020. [Accessed Jan. 28, 2022].

[5]   Department of Statistics Malaysia, "Supply and utilization accounts selected agricultural commodities, Malaysia 2016-2020," Available: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=164&bul_id=cHgwanhNdU4vWXRvc3pnZU9xSjZTUT09&menu_id=Z0VTZGU1UHBUT1VJMFl__paXRRR0xpdz09, 2021. [Accessed Jan. 28, 2022]

[6]   Z.L. Chuan, S.M. Deni, S-.F. Fam and N. Ismail, "The effectiveness of a probabilistic principal component analysis model and expectation maximisation algorithm in treating missing daily rainfall data," Asia-Pacific Journal of Atmospheric Sciences., vol. 56, pp. 119-129, 2020.

[7]   A. Chizari, Z. Mohamed, M.N. Shamsudin and K.W.K. Seng, "The effects of climate change phenomena on cocoa production in Malaysia," International Journal of Environment, Agriculture and Biotechnology., vol. 2, no. 5, pp. 2599-2604, 2017.

[8]   B.T. Tan, P.S. Fam, R.B.R. Firdaus, M.L. Tan and M.S. Gunaratne, "Impact of climate change on rice yield in Malaysia: a panel data analysis," Agriculture., vol. 11, no. 6, pp. 569, 2021.

[9]   M.H.M. Hazir, R.A. Kadir and Y.A. Karim, "Projections on future impact and vulnerability of climate change towards rubber areas in Peninsular Malaysia," in IOP Conference Series: Earth and Environmental Science., vol. 169, no. 012053, 2018.

[10]   M.W. Murad, R.I. Molla, M.B. Mokhtar and M.A. Raquib, "Climate change and agricultural growth: an examination of the link in Malaysia," International Journal of Climate Change Strategies and Management., vol. 2, no. 4, pp. 403-417, 2010.

[11]   A.A. Houma, M.R. Kamal, M.A. Mojid, A.F.B. Abdullah and A. Wayayok, "Climate change impacts on rice yield of a large-scale irrigation scheme in Malaysia," Agricultural Water Management., vol. 252, no. 106908, 2021.

[12]   N.M. Roslan, W.L. Shinyie and S.S. Ling, "Modelling high dimensional paddy production data using copulas," Pertanika Journal of Science & Technology., vol. 29, no. 1, pp. 263-284, 2021.

[13]   A. Abubakar, M.Y. Ishak and A.A. Makmo, "Impacts of and adaption to climate change on the oil palm in Malaysia: a systematic review," Environmental Science and Pollution Research., vol. 28, pp. 54339-54361, 2021.

[14]   M.S. Fahmy, "The value of big data to the world economy," Arab Journal of Administration, vol. 40, no. 4, pp. 307-322, 2020.

[15]   Y. Ge and H. Wu, "Prediction of corn price fluctuation based on multiple linear regression analysis model under big data," Neural Computing and Applications, vol. 32, no. 1, pp.16843-16855, 2020.

[16] M.L. Tan, L. Juneng, F.T. Tangang, J.X. Chung and R.B.R. Firdaus, "Changes in temperature extremes and their relationship with ENSO in Malaysia from 1985 to 2018," International Journal of Climatology, vol. 41, no. S1, pp. E2564-E2580, 2021.

[17] V.S. Shevade and T.V. Loboda, "Oil palm plantations in Peninsular Malaysia: determinants and constraints on expansion," PLoS ONE, vol. 14, no. 2, pp. e0210628, 2019.

[18] A. Simon, V.K. Subbiah, C.F. Tyng and N.H.M. Yusuf, "Genetic diversity of Sabah rice cultivars using random amplified polymorphic DNA (RAPD) markers," Borneo International Journal of Biotechnology (BIJB), vol. 1, pp. 35-43, 2020.

[19] N. Vaghefi, M.N. Shamsudin, A. Radam and K.A. Rahim, "Impact of climate change on food security in Malaysia: economic and policy adjustments for rice industry," Journal of Integrative Environmental Sciences, vol. 13, no. 1, pp. 19-35, 2015.

[20] J. Lu, G.J. Carbone, X. Huang, K. Lackstrom and P. Gao, "Mapping the sensitivity of agriculture to drought and estimating the effect of irrigation in the United States, 1950-2016," Agricultural and Forest Meteorology, vol. 292-293, no. 1, pp. 108124, 2020.

[21] W. Mahrous, "Climate change and food security in EAC region: a panel data analysis," Review of Economics and Political Science, vol. 4, no. 4, pp. 270-284, 2019.

[22] S.H. Mosavi, S. Soltani and S. Khalilian, "Coping with climate change in agriculture: evidence from Hamadan-Bahar plain in Iran," Agricultural Water Management, vol. 241, no. 1, pp. 106332, 2020.

[23] M.A.R. Sarker, K. Alam and J. Gow, "Assessing the effects of climate change on rice yields: an econometric investigation using Bangladeshi panel data," Economic Analysis and Policy, vol. 44, no. 4, pp. 405-416, 2014.

[24] J.A. Hausman, "Specification tests in econometrics," Econometrica, vol. 46, no. 6, pp. 1251-1271, 1978.

[25] B.H. Baltagi and Q. Li, "Testing AR(1) against MA(1) disturbances in an error component model," Journal of Econometrics, vol. 68, no. 1, pp. 133-151, 1995.

[26] G.S. Maddala and S. Wu (1999), "A comparative study of unit root tests with panel data and a new simple test," Oxford Bulletin of Economics and Statistics, vol. 61, no. S1, pp. 631-652, 1999.

[27] S.S. Arora, "Error components regression models and their applications," Annals of Economic and Social Measurement, vol. 2, no. 4, pp. 451-461, 1973.

[28] P.A.V.B. Swamy and S.S. Arora, "The exact finite sample properties of the estimators of coefficients in the error components regression models," Econometrica, vol. 40, no. 2, pp. 261-275, 1972.

[29] P.M. Aliha, A.P.D.T. Sarmidi and D.F.F. Said, "Applying panel data model for comparing static and dynamic forecasts of an autoregressive money demand incorporating financial innovation in Asean countries," International Journal of Accounting, Finance and Business, vol. 5, no. 28, pp. 22-36, 2020.

[30] R. Cellmer, A.B. Cichulska and M. Belej, "The regional spatial diversity of housing prices and market activity-evidence from Poland," Acta Scientiarum Polonorum Administratio Locorum, vol. 20, no. 1, pp. 5-18, 2021.

[31] R. Stefko, B. Gavurova, M. Kelemen, M. Rigelsky and V. Ivankova, "Relationships between renewable energy and the prevalence of morbidity in the countries of the European Union: a panel regression approach," International Journal of Environmental Research and Public Health, vol. 18, no. 12, pp. 6548, 2021.

[32] B.H. Baltagi, Econometric Analysis of Panel Data, 6th Edition, Cham, Switzerland AG: Springer Nature, 2021.

[33] N.P. Joshi, K.L. Maharjan and L. Piya, "Effect of climate variables on yield of major food-crops in Nepal: a time-series analysis," Journal of Contemporary India Studies: Space, Society, vol. 1, pp. 19-26, 2011.

[34] Y. Xu, L. Yu, W. Li, P. Ciais, Y. Cheng and P. Gong, "Annual oil palm plantation maps in Malaysia and Indonesia from 2001 to 2016," Earth System Science Data, vol. 12, no. 2, pp. 847-867, 2020.

[35] R.B.R. Firdaus, M.L. Tan, S.R. Rahmat and M.S. Gunaratne, "Paddy, rice and food security in Malaysia: a review of climate change impacts," Cogent Social Sciences, vol. 6, no. 1, pp. 1818373.

[36] J.L. Hatfield and J.H. Prueger, "Temperature extremes: effect on plant growth and development," Weather and Climate Extremes, vol. 10, no. Part A, pp. 4-10, 2015.