An Improved Monthly Oil Palm Yield Predictive Model in Malaysia

Check for updates

Jen Feng Khor, Zulkifli Yusop, and Lloyd Ling

Abstract Oil palm crop is sensitive to the heat stress. A new model is developed with 36 years of national monthly yield data to quantify the impact of past El Niño events on the Malaysian palm oil industry, namely Fresh Fruit Bunch Index (FFBI) model. The FFBI model shows significant correlation with the National Oceanic and Atmospheric Administration (NOAA), Oceanic Niño Index (ONI) and higher predictive accuracy (adjusted R-squared $= 0.9312$) than the conventional FFB model (adjusted R-squared $= 0.8274$). The FFBI model suggests that oil palm yields in Malaysia could be affected after 2–16 months of the occurrence of El Niño events. The FFBI model also forecasts an oil palm under yield concern in Malaysia from July 2021 to December 2023 and matches with the actual national oil palm under yield trend to date (July 2021–April 2022). Malaysian oil palm yields failed to recover from the 2015/16 very strong El Niño and showed a production downtrend pattern even before the pandemic market lock down. This strongly suggests that there are other hidden threats that have plagued the Malaysian palm oil industry for years, other than the climatic factor.

Keywords Oil palm · El Niño · Yield modelling and prediction

J. F. Khor
Centre of Disaster Risk Reduction (CDRR), Lee Kong Chian Faculty of Engineering and Science, Department of Civil Engineering, Universiti Tunku Abdul Rahman, Jalan Sungai Long, Bandar Sungai Long, 43000 Cheras, Malaysia

Z. Yusop
Centre for Environmental Sustainability and Water Security (IPASA), Faculty of Civil Engineering Department, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

L. Ling (✉)
Centre of Disaster Risk Reduction (CDRR) and Centre for Mathematical Sciences (CMS), Lee Kong Chian Faculty of Engineering & Science, Department of Civil Engineering, Universiti Tunku Abdul Rahman, Jalan Sungai Long, Bandar Sungai Long, 43000 Cheras, Malaysia
e-mail: linglloyd@utar.edu.my

187

1   Introduction

There are 17 varieties of oils and fats in the edible oil and fat sector, including vegetable oils and animal fats. Among them, palm oil is by far the most manufactured oil, accounting for around 31.3% of total edible oils and fats produced in 2019 [1]. Oil palm has the highest efficiency in terms of edible oil yields per hectare, with yields of nearly 4.27 tonnes per hectare per year (t/ha/yr) (combined palm and palm kernel oil), which is at least six times more productive than any other oilseed [2, 3]. However, oil palm crop yields which are sensitive to heat stress are affected by El Niño, a warm phase of a larger phenomenon called the El Niño-Southern Oscillation (ENSO) which causes drought in Malaysia. During the occurrence of El Niño in Malaysia, oil palm crops which are sensitive to prolonged drought periods face a high level of water stress due to reduced rainfall and increased temperature [4, 5].

Many researchers have looked into the relationship between rainfall, temperature, and palm oil production in Malaysia. This study investigates the direct relationship between Oceanic Niño Index (ONI) (an ENSO index used to measure El Niño) and the monthly oil palm yields in Malaysia. Furthermore, this study uses an improved model, called Fresh Fruit Bunch Index (FFBI) [6] to model the impact of El Niño on the palm oil production in Malaysia and to forecast the Malaysian oil palm yields in the near future.

2   The Fresh Fruit Bunch Index (FFBI) Model

ONI is based on the monthly sea surface temperatures (SST) tracked by the National Oceanic and Atmospheric Administration (NOAA) [7]. Fresh Fruit Bunch Index (FFBI) [6] follows a similar calculation method of ONI and is created using the monthly Malaysian oil palm fresh fruit bunches (FFB) yield data [8–10]. In this study, monthly oil palm FFB yields from January 1986 to July 2021 (total 427 data points) in Malaysia were used. The national average FFB yield for each month was calculated on the month-to-month basis from 1986 to 2020 in Table 1.

This study adopted the way that NOAA calculates ONI and developed Fresh Fruit Bunch Index (FFBI) model with following calculations. $FFB_{anomaly}$, the difference between FFB yield and average FFB yield (Table 1) at each corresponding month was calculated using Eq. (1):

$$FFB_{anomaly} = FFB - FFB_{average} \qquad (1)$$

FFBI was then calculated by taking the 3-month running mean of FFBanomaly, as shown in Eq. (2):

$$FFBI_i = \left(FFB_{anomaly,i-1} + FFB_{anomaly,i} + FFB_{anomaly,i+1}\right)/3, \qquad (2)$$

Table 1 Average FFB yield for each month from 1986 to 2020 in Malaysia [8–10]

Month	Average FFB yield (t/ha)
January	1.3194
February	1.1903
March	1.3234
April	1.4049
May	1.4829
June	1.5109
July	1.6329
August	1.7420
September	1.8543
October	1.8146
November	1.6746
December	1.4743

where $FFB_{anomaly,\ i-1}$ and $FFB_{anomaly,\ i+1}$ were the FFB anomalies one month before and after the month that were being calculated. Hence, from 427 points of FFB yields, a total of 425 points of FFBI was derived, from February 1986 to June 2021.

2.1 Correlation Test Between Malaysian Oil Palm Yields and Oceanic Niño Index (ONI)

The correlations between FFB and FFBI with ONI were tested using nonparametric Spearman's rho correlation test. A nonparametric correlation test was used because not all tested variables were normally distributed. The normality of the parameters was tested using Shapiro–Wilk test (suggested for dataset with less than 2,000 samples) [11].

A Spearman's rho correlation is also referred to as Spearman correlation or Spearman rank correlation. The strength of association between the tested variables is expressed in a bivariate correlation, which is between −1 and +1. The two variables have a positive relationship (as values of one variable increases, values of the other variable also increase) if the correlation coefficient is positive. In a negative relationship (negative correlation coefficient), as the values of one variable increases, the value of the other variable decreases. No relationship exists between the tested variables if the correlation coefficient is zero [12]. All tests were conducted using IBM SPSS Statistics version 26.0 [13].

2.2   Time Series Forecasting

FFB and FFBI time series forecasting models with ONI as its predictor were created using the Expert Modeler in IBM SPSS Statistics version 26.0 [13]. Using monthly Malaysian FFB data points from February 1986 to June 2021 ($N = 425$), the final FFB and FFBI time series forecasting models were created to forecast for another 30 months from July 2021 until December 2023.

Adjusted *R*-squared was used in this study to benchmark the models' predictive accuracy. The adjusted *R*-squared was derived from the *R*-squared produced by the models using Eq. (3):

$$R_{\text{adj}}^2 = 1 - \left[\left(1 - R^2 \right) (n-1)/(n-k-1) \right], \tag{3}$$

where R^2 is the *R*-squared of the models, N is the number of data points, and K is the number of independent regressors [14].

To compare the FFB and FFBI time series models, residual analyses were conducted using IBM SPSS Statistics version 26.0. Residual between the predicted and observed data point was calculated [15, 16]. Descriptive statistics of the residuals from both models were calculated for model comparison, including residuals' skewness, range, median, standard deviation, and variance. Shapiro–Wilk normality tests were also conducted for the residuals to determine whether the median or mean residuals should be referred for model prediction accuracy comparison. As the residuals from both models were determined to be non-normally distributed, the FFB and FFBI model's median residuals were used for comparison assessment. The nonparametric inferential statistics of the bias corrected and accelerated (BCa) bootstrapping method were conducted with 2000 random samples (with replacement) to compute the 99% confidence intervals (CI) of the residuals' median, standard deviation, and variance.

The FFBI time series model was also validated repeatedly using a 30-year moving time frame. Six FFBI time series models were created and validated using different block periods (1986–2015, 1987–2016, 1988–2017, 1989–2018, 1990–2019, and 1991–2020). Each model forecasted FFBI one year ahead of the respective block period to validate the predictive model's accuracy with the observed oil palm yield data.

3   Results and Discussion

Based on the nonparametric Spearman's rho correlation test, monthly FFB data has significant correlation (at least 0.05 alpha level) with ONI data at lag periods from 6 to 13 months, with the highest correlation at -0.217 at 0.01 alpha level. On the other hand, FFBI shows higher correlation with ONI. The correlation between FFBI and

Table 2 FFB and FFBI time series models' comparison [6]

Model	FFB	FFBI
Adjusted R-squared	0.8274	0.9312
Residual sum of squares (RSS)	4.6876	0.5459
Residual median	0.0077	0.0051
Residual range	0.7644	0.2790
Residual standard deviation	0.1067	0.0359
Residual variance	0.0114	0.0013

[1]Modelled period: February 1986–June 2021 ($N = 425$)
[2]ONI is used as predictor for both models

ONI is significant at lag periods from 2 to 16 months, with the highest correlation of -0.399 at 0.01 alpha level [6].

The negative correlations between oil palm yields and ONI show that palm oil production reduced when ONI increases during the occurrence of El Niño events. However, the impact could only be seen after a period of time and is not constrained to a single fixed period. The FFBI model suggests that oil palm yields in Malaysia could be affected after 2–16 months of the occurrence of El Niño events [6].

Based on the statistical assessment of the FFB and FFBI time series model (Table 2), the FFBI model has higher predictive accuracy than the conventional FFB model. The adjusted R-squared of the FFBI model (0.9312) is significantly higher than the FFB model (0.8274). Furthermore, the residual analyses further support that FFBI is a better oil palm yields predictive model in Malaysia. The residual sum of squares (RSS), median, range, standard deviation, and variance of the FFBI model are also lower than the FFB model. The FFBI model provides an improved oil palm yield predictive tool for the Malaysian oil palm industry stakeholders to better understand the impact due to El Niño events [6].

The six FFBI validation models using different time frames have consistent model predictive accuracy as well. The adjusted R-squared of all the validation models are in the range of 0.9160–0.9354, showing that the newly developed FFBI model has stable and consistent predictive accuracy [6].

Looking at the forecasts of the oil palm yields predictive models from July 2021 to December 2023 in Malaysia (Fig. 1), FFB model forecasted that the monthly oil palm yields to fluctuate around the BCa 99% confidence interval range during the aforementioned time period. Unlike FFB model, the FFBI model revealed a monthly under yield concern of Malaysian oil palm yields in the near future. Both models' forecasts were benchmarked and validated recently to the actual national oil palm monthly yields in Malaysia (July 2021–April 2022) and warranted the FFBI model's under yield prediction concern. This suggests that there might be other underlying factors that affect the Malaysian palm oil production [6] as there were no El Niño recently.

Based on the historical data, oil palm yields in Malaysia are recovered from the huge drop in yields caused by the 1997/98 very strong El Niño about a year after the event. However, national oil palm yields failed to recover from the 2015/16 very
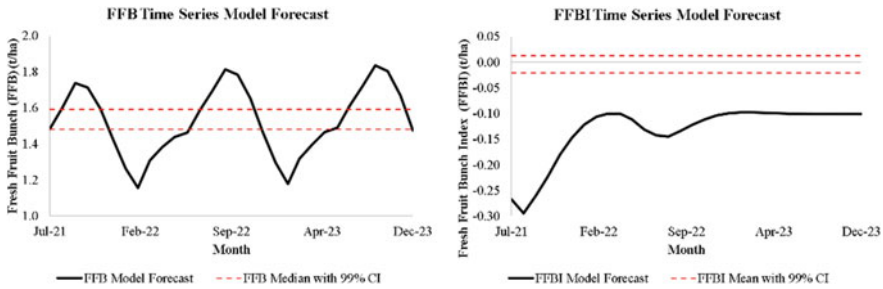
Fig. 1 FFB and FFBI time series models' forecasts [6]. 1. Forecasted period: July 2021–December 2023 ($N = 30$ months). 2. Both models' forecasts are validated using monthly national oil palm yields in Malaysia to date. FFBI model has the closer prediction

strong El Niño and maintained a production downtrend ever since. Since 2010, palm oil industry has been listed as one of the important sectors by Malaysian government under the Economic Transformation Programme (ETP). Target was set to achieve oil palm national average yield of 26.2 t/ha/yr by 2020 [17]. However, the performance of Malaysian oil palm yield is still far from the target. This strongly suggests that there are other hidden threats that have plagued the Malaysian palm oil industry for years, other than the climatic factor.

4 Conclusions

This study developed an improved oil palm yield predictive model, namely Fresh Fruit Bunch Index (FFBI) model that is derived from the monthly oil palm FFB yields in Malaysia. Using the FFBI model, the oil palm yields are better correlated with ONI to model the impact of El Niño on the palm oil production in Malaysia. In terms of predictive accuracy, the FFBI time series model has significantly higher adjusted R-squared (0.9312) compared to the FFB model (0.8274). The forecasts of the FFBI model are also closer to the actual Malaysian oil palm monthly yield to date when compared to the conventional FFB model.

The production downtrend in the Malaysian palm oil industry validates the forecasted under yield concern by the FFBI model. It is strongly suggested that El Niño is not the sole factor in affecting palm oil production in Malaysia. Other underlying factors should be investigated, such as ageing of oil palm crops, so that Malaysia can remain competitive in the global palm oil industry.

References

1. Mielke T (2018) Global supply, demand and price outlook of oils and fats in 2018/19. In: Globoil, Mumbai
2. Malaysian palm oil industry. http://www.palmoilworld.org/about_malaysian-industry.html, last accessed 17 Aug 2021
3. World Energy Council (2010) Biofuels: policies, standards and technologies. World Energy Council, London
4. Kamil NN, Omar SF (2016) Climate variability and its impact on the palm oil industry. Oil Palm Industr Econ J 16(1)
5. Oettli P, Behera SK, Yamagata T (2018) Climate based predictability of oil palm tree yield in Malaysia. Sci Rep 8(2271). https://doi.org/10.1038/s41598-018-20298-0
6. Khor JF, Ling L, Yusop Z, Tan WL, Ling JL, Soo EZX (2021) Impact of El Niño on oil palm yield in Malaysia. Agronomy 11(2189)
7. Cold & warm episodes by season. https://origin.cpc.ncep.noaa.gov/products/analysis_monito ring/ensostuff/ONI_v5.php, last accessed 25 Oct 2021
8. Palm Oil Registration & Licensing Authority (PORLA) (1986–1999) PORLA palm oil statistics. PORLA, Kelana Jaya
9. Malaysian Palm Oil Board (MPOB) (2000–2020) Malaysian oil palm statistics. MPOB, Bangi
10. Economics and industry development division. https://bepi.mpob.gov.my/index.php/en/, last accessed 17 Aug 2021
11. Ling L, Yusop Z, Yap WS, Tan WL, Chow MF, Ling JL (2019) A calibrated, watershed-specific SCS-CN method: application to Wangjiaqiao watershed in the three gorges area, China. Water 12(60)
12. Artusi R, Verderio P, Marubini E (2002) Bravais-Pearson and Spearman correlation coefficients: meaning, test of hypothesis and confidence interval. Int J Biol Mark 17(2)
13. Downloading IBM SPSS Statistics 26. https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-26, last accessed 27 Aug 2021
14. Adjusted R2/Adjusted R-Squared: What is it used for? https://www.statisticshowto.com/pro bability-and-statistics/statistics-definitions/adjusted-r2/, last accessed 27 Aug 2021
15. Ling L, Yusop Z, Chow MF (2020) Urban flood depth estimate with a new calibrated curve number runoff prediction model. IEEE Access 8:10915–10923
16. Ling L, Yusop Z, Ling JL (2021) Statistical and type II error assessment of a runoff predictive model in Peninsula Malaysia. Mathematics 9(812)
17. Performance Management and Delivery Unit (PEMANDU) (2010) Economic transformation programme: a roadmap for Malaysia. Prime Minister's Department of Malaysia, Putrajaya