Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Bayesian Network Design for Crude Palm Oil (CPO) Price Prediction Driven by Fluctuation Patterns and Trends

Masnida Hussin[1,2,*], Zurita Ismail[2], Iszuanie Syafidza Che Ilias[2]

[1] Department of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[2] Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 4 February 2023
Received in revised form 16 June 2023
Accepted 24 June 2023
Available online 11 July 2023

Keywords:
Bayesian probability; crude palm oil; dynamic approach; price prediction

ABSTRACT

Crude Palm Oil (CPO) becomes the alternative to petroleum because it can be refined into different products, such as biodiesel. The price of crude palm oil can be considered the future direction of the world economy. CPO prices are based on market demand and the oil supply from the CPO producer, where they fluctuate and have a huge influence on economy. The price pattern of CPO requires a variety of factors to predict accurately. It needs a lot of data analytics to predict and respond in quick time to the highly variable CPO market. In this work, a dynamic probability model is proposed to predict CPO prices using the Bayesian rule. The rule is formed to be effective in responding to changeable market demands and supplies. It formulates the approximation factors for indicating the expected price of the CPO for 12 months. The dynamic training data has been modelled to fit the Bayesian rule. The process of Prediction and hypothesizing price fluctuations is identified according to several level of Bayesian rule. It is then used to generate a measure of deviation between the actual and probable prices. Our Bayesian-based prediction pricing model is able to predict the price pattern that are comparable to existing benchmark data and shows a lower standard error of regression. The prediction approach can help the traders have a better analysis of price fluctuations in CPO demand and supply. Using Bayesian probability not only improves prediction rules, but it can also forecast CPO trades in high fluctuation situations.

1. Introduction

Crude Palm Oil (CPO) prices are based on market demand and the oil supply from the CPO producers, where they fluctuate and have a huge influence on Malaysia's economy. Because of inconsistency in supply and competition from other vegetable-based oils such as soya bean oil, rapeseed oil, canola oil, and sunflower oil, CPO prices can fluctuate in the global market explained by K. Chuangchid *et al*., and N. M. N. Ab Rahman *et al*., [1-3]. The reason for using the probability strategy for determining the expected CPO price is that it is efficient at measuring mess and dynamic information. The adaptive price prediction models heavily influence the dynamic market pricing

* *Corresponding author.*
E-mail address: masnida@upm.edu.my

https://doi.org/10.37934/araset.31.2.117129

determining process mentioned by several researchers [4-6]. Several authors [5-9] utilized Bayesian probability in their pricing applications. The Bayesian probability rules specify how to create inferences about hypotheses from the dataset. The identifying and learning processes are part of the inference in the prediction model. Since the Bayes' theorem is an effective algorithm [10,11], it can be used for the real-time Prediction that is commonly used to predict fast-evolving data such as CPO prices. It can be used to provide a forecasting strategy by utilizing prior knowledge about the possibility of selling CPO at any type of event, whether through dealers, exchanges, or brokers.

In this work, we mainly aim to design a pricing prediction model for CPO using Bayesian probability. We also compared the value of CPO prices generated through our prediction model against the benchmark prices that are given in https://bepi.mpob.gov.my/ [12]. Particularly, it is proof that the Bayes-based pricing estimation is equivalent to the pricing prediction scheme that is applied in the real market. There are two stages in performing the prediction process. The first stage describes in detail the process of designing a price prediction model using Bayesian probability. It is basically for harmonizing the Bayes' theorem with the information given from the dataset. Next, we describe in detail how to use the prediction model to make inferences about the price hypotheses that follow a price fluctuation pattern. We then compared the Bayes-based price value against the real pricing data from the dataset.

2. Related Works

It is important to monitor and accurately forecast the price of crude palm oil (CPO) for the benefit of the Malaysian palm oil industry, as palm oil growth plays a significant role in the Asian economy. The fluctuation of the CPO price posed a big risk to farmers, producers, traders, consumers, and others. The increase in its global CPO prices, as reviewed by the researchers [13-15] in increased the marketing margins of both entrepreneurs and oil palm farmers while encouraging the producer countries to increase their production by expanding plantations. Due to the rapidly increasing number of productions and the exploration of more space for plantations, it effected environmental destruction in the short- and long-term. Due to this matter, the authors in Syahril *et al.*, [16] studied the effect of global CPO prices, marketing margins, and palm oil plantations on environmental destruction within the period of time. They concluded that the global CPO price has a positive influence on the environmental quality index. Earlier, as studied by Karia and Bujang [17] proposed the artificial neural network (ANN) by utilizing daily, hourly, minutely, and per-second basis to perform better for the high frequency data. ANN is also able to generate more forecasted error in linear time series data. They are comparing their work with that of Box-Jenkins, which only gives better predictions when the Prediction deals with the low frequency of the time series data. Supported by the work in Kanchymalay, K. *et al.,* [18] that they also forecast the CPO price using multivariate time series. From their experiment, the support vector regression (SVR) with sequential minimal optimization (SMO) algorithm showed better results. Due to better palm management, i.e., fertilizing and cropping processes, it is challenging to predict the CPO price with various features to consider. Hence, a dynamic mechanism in price prediction becomes one solution to accurately measuring CPO price patterns.

The studies by the authors in Go and Lau [19] examined and compared the influence of the price-volume relation in the CPO futures market during the precrisis, crisis, and post-crisis periods. They used the price-volume interaction as a basic framework for determining demand and supply of a commodity. They claimed that it is more effective for providing Prediction from transmission information than underlying markets. The authors in Rahim *et al.*, [20] forecasted CPO prices by using a data driven fuzzy rule-based system. It is based on the time series forecasting approach, which increases the accuracy of the predictions. Meanwhile, the authors in [21,22] proposed the Neural Network Backpropagation (NNBP) scheme to optimize the weight of Crude Palm Oil commodity price. They proposed a prediction model based on the NNBP for precisely predicting and minimizing differences between the current and forecast prices. Most of the existing studies on CPO prediction price that used a time-series-based method involved a variety of parameters in their prediction approaches. It leads to high calculation complexity. In our work, the price prediction is proposed using a probability approach that merely considers two factors: the current value and the period of forecast. The use of a Bayesian network provided satisfactory results without the need for a large amount of historical data and at a low computational cost. It means reducing calculation complexity while reducing processing delay. Note that today's CPO price is very important for preparing tomorrow's supplies.

3. Design of Bayesian as Prediction Model

Bayesian probability is one of the machine learning methods that is simple and effective at minimize computational complexity. It is scalable and particularly useful for uncertain and dynamic datasets by J. L. Ticknor., [4]. The Bayesian probability is derived using Bayes' theorem, which assumes all features are conditionally dependent. It assumes there is a relationship between the presence of a feature in the training class and the presence of any other feature. In Eq. (1) by Scutari and Denis [23], the fundamental notion of Bayes' theorem is given.

Bayes' Theorem

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \tag{1}$$

The constant A means hypothesis made from the actual data, while B represents the actual data. In the Bayes theorem, P(A|B) means to find the combination between the hypothesis and the actual data. Meanwhile, P(B|A) is a conditional probability that will occur when certain data is analyzed. P(A) is the probability of the data that is required to be found, where it refers to the CPO prices in this work. The P(B) refers to all the actual data combinations. The generation between hypotheses combinations and probability expressions will then produce a value.

By using Bayesian probability, it is not necessary to have a large dataset for training purposes. This method gives benefits in our study when the available real datasets are slightly broad and have dynamic size values. Furthermore, by the Bayes theorem, there are only two main parameters needed to measure the probability output. It helps the CPO commodity players reduce the burden of collecting many data parameters for forecasting the next accurate CPO price. Since, CPO prices are forecasted on a daily basis for the next 12 months, reducing computing complexity will help in fast and effective price assessment. We used Naïve Bayes (NB) probability because it is used for discrete counts. The probability model will be regularly traced the current price and checked the fluctuation patterns of current prices. By understanding the current price patterns, the Bayesian model will analyze each data and make predictions accordingly. Each output from the prediction process is then used to identify whether it falls into an increase or decrease category. Each class of price is then processed again using the earlier process (that with NB) to balance pattern fluctuation; hence the predicted price is proposed.

In response to Bayes' theorem given in Eq. (1), we modified the theorem to be aligned with our prediction scenario and data set. Our CPO pricing scenario required us to embed the parameters of price increases or decreases in the context of supply and demand. The expected price increases will lead to low future demand and sales of CPO, while price decreases are expected to grow future demand. These two main parameters i.e., price increase and decrease must interact with one another in order to produce a new parameter, the fluctuation pattern. Hence, we invent the Bayes theorem links for measuring the degree of belief in the price proposition (either increase or decrease) to forecast market demand and supply over the CPO. The fluctuation pattern is readied prior to the prediction process starting, which required adaptive and adjusting processes for allowing the existence of the Bayes theorem links in the data set. The prices will then be predicted for each row of data and analyzed to forecast each price in the month for the next 12 months.

The modified Bayes theorem to account for the CPO price in market scenario is given as follows.

$$(PD|PM) = \frac{P(PM|PD) \times P(PD)}{P(PM)} \tag{2}$$

It is where PD and PM are the prices of a day and a month, respectively. Both variables are significant in predicting the price changes over time. It is advantageous to predict CPO supply and demand for the following day. Meanwhile, the probabilities *PD* and *PM*, i.e., P(*PD*) and P(*PM*), are the CPO pricing probabilities that can be observed in a day and a month where they are conditionally independent of each other. The P(*PD*|*PM*) measured the price pattern probability when another pattern has occurred; the price *PD* occurred when the *PM* value is true. The P(*PM*|*PD*) is used for measuring the possibility of a price pattern In *PM* when given that *PD* is true. The condition of being true or false is modelled by comparing the price range that was produced from data training. Note that P(*PM*) cannot be 0. It means that the CPO price must be predicted for every month without fail. However, it is not necessary to estimate the price every single day of the month. This consideration is applied in response to the available dataset that was read prior to producing the predicted price.

4. Evaluate the Bayes-based Price Value

We performed data training, filtering, and range processes by using the available dataset from https://bepi.mpob.gov.my/ [12]. In determining the pricing set through the Bayesian probability model, several steps of the mathematical process are involved, as follows

4.1 Identification

Knowledge of the parameters (coefficients and factors) is required to study the information relation and pattern in order to design the prediction model. This step collects the existing dataset related to the price fluctuation (event) from the available dataset. According to the CPO real price dataset, the accuracy of the CPO price for the next 12 months has been entered in the current day to meet current and future CPO demands or sales. We identified two important factors in the dataset that are used for forecasting future CPO prices are

i. the date or day of the week, and
ii. the month of the year.

Such factors are critical for comprehending the price fluctuation pattern (event). We then designed data modelling to understand how both factors are related to each other. The data modelling (Table 1) demonstrated that the price is influenced by the price in the next day, *X,* and in a month, *Y,* as follows

P (*X*) → the CPO price in the next day
P (*Y*) → the CPO price in a month

It is also identified that *X* is not a compulsory value to fulfil *Y*. This means that it can have an empty probability value for the CPO price in the next day, and it will not affect the pattern of the CPO price for the month. However, the *Y* must occur and be predicted for every month. In other words, the probability of *Y* is constant in a normal distribution, while the probability of *X* depends on the pricing and market activity in the trading centre. In response, we denote the indication for the data modelling scheme as follows

Table 1
Indication for data modelling

No.	Factor	Denote	Remark
1	Event	E_C where C = {*add* or *cut*}	Price fluctuation; increase or decrease
2	Current day	*inp*	The actual day when the data been entered
3	Day in a month	D_X where X = {1, 2…, *d*}	Sequence of days in a month, *X*
4	Month in a year	M_Y where Y = {1, 2…, *12*}	Sequence of months in a year, *Y*

4.2 Modelling

After the related factors are identified, the Bayes mathematical formula comes into the picture for processing the hypothesis from the available factors. In this step, to achieve accuracy, many probability events are needed to execute and manipulate. It is because using diverse types of events helps in finding accurate probability sampling.

Based on the given dataset, the price fluctuation patterns (events) are diverse and dynamic. It means that the price increases and decreases are not in an exponential, linear, or identical pattern. In the real trading market, it is hard to assess the CPO prices' patterns in prior. Hence, we have classified a set of events in response to the factors (i.e., D_X and M_Y). Note that, each event consists of both factors to indicate the price fluctuation; an increase or decrease. We assumed that the price fluctuation patterns (events) might give a uniform pattern if most of the prices were in the factor M_Y show small differences in their prices increasing and decreasing. In order to execute the event, the current CPO prices need to be collected as an input for factor M_Y and assigned as *inp*. Such values become the main hypothesis in the modelling process.

4.3 Quantifying

Each event is measured by comparing all prices in the dataset according to the respective data modelling above. In our work, the value of *Y* is always 12; it represents 12 months of duration. As a result, the average CPO price for the M_Y is given as pe / 12; where pe is the current price. However, the occurrence of the event on the next day, *X* for the D_X factor might be different from other events. It is due to the fact that the quantity of *X* is not always the same, e.g., it can be 28, 30, 31 days, or less than that (e.g., if there is no trading operation). Therefore, we employ pricing analysis by using the following rule.

i. Variable → the input day volume, *n* and the input month volume, *m*
ii. Product variables → the total input occurrence, *n x m*
iii. Input → the value of the existing price, p_e.
iv. Output P(*Y*) → the average price of a month, $AM = \sum p_{e \rightarrow} Y / 12$

v. Output P(*X*) → the average price of a day, $AD = \sum p_{e \rightarrow} X / n$

From the given dataset, there is a diverse range of prices for every day of a month. Hence, to predict the incoming price for the next 12 months in a day, we identified the probabilities of observing CPO prices in a day, P(*PD*), as follows.

$$\boldsymbol{P(PD)} = \frac{\boldsymbol{P(PM).\sum_m(p_m)}}{\boldsymbol{P(PM).\sum_m(p_m)+P(PD).\sum_n(p_n)}} \tag{3}$$

It is where *PD* refers to the number of events that have been collected for 12 months in a single day. Meanwhile, for P(*PM*), the probabilities of observing CPO prices in a month are given as follows; where *PM* denotes the number of events collected every day for a month.

$$\boldsymbol{(PM)} = \frac{\boldsymbol{P(PD).\sum_n(p_n)}}{\boldsymbol{P(PM).\sum_m(p_m)+P(PD).\sum_n(p_n)}} \tag{4}$$

The probabilities P(PD) and P(PM) are reflecting each other using the Bayes rule; however, the values of PD and PM are independent. As a result, the price fluctuation pattern is inconsistent, with the price increment and decrease swinging heavily for both X and Y. However, in this work, we assumed that the reduction in the next CPO price would lead to higher demand and sales. Meanwhile, the increase in price for the next day or month will reduce sales or demand. It means that there is a chance of a price in a day occurring given that a price in a month is required.

By using that assumption, the probability P(PM|PD) can be identified as follows.

$$P(PM|PD) = \frac{P(PM).\sum_m(p_m)+P(PD).\sum_n(p_n)}{P(PD)} \tag{5}$$

Then, we extend the Bayes theorem Eq. (1) to formulate the probability of the CPO price that has been forecast for each day for the next 12 months as follows.

$$P(PD|PM) = \frac{P(PM).\sum_m(p_m)}{P(A)} \cdot \frac{P(PM).\sum_m(p_m)+P(PD).\sum_n(p_n)}{P(PD).\sum_n(p_n)} \tag{6}$$

4.4 Analyzing

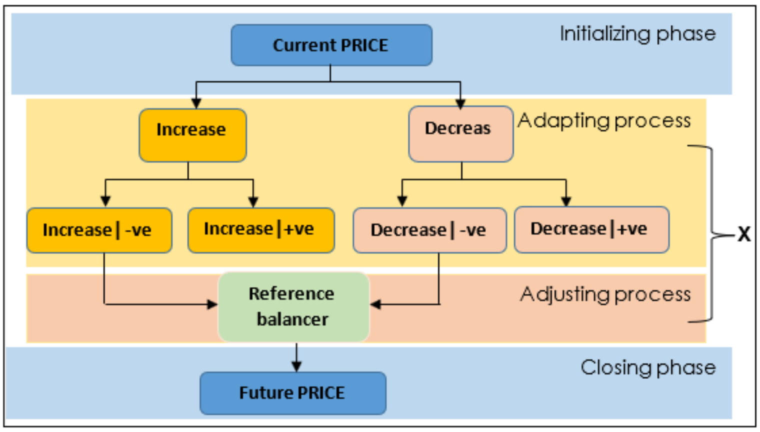
Predicting CPO prices is difficult when dealing with rapidly changing events. With the dynamic demand and supply in the following day, the price must be forecasted quickly. The increase and decrease of the CPO's next (near future) prices are influenced by the current CPO price. Hence, in order to make the price in the market steady and stable within some duration of time, there must be a control mechanism in the event that it changes. One way to achieve that is by ensuring the variation between increases and decreases in CPO pricing does not become a huge breach. The price proposition, whose increases or decreases are further applied into the Bayes theorem to forecast the CPO prices.

In this work, we designed a classifier method using the Multinomial Naïve Bayes model, where it will be regularly counting how often the event of either an increase (price add) or decrease (price cut) occurs in a week (Figure 1). This model will group the price variances according to the number of events produced within one week. The total number from the increase and decrease events then becomes one of the variables in the training and learning process (Figure 2). It is used as a reference (threshold) for predicting the final/testing dataset. It further aims to establish market stability for CPO demand and supply. Note that, in the CPO marketplace, the stability in price pattern is preferable for long-term evaluation [24].

Algorithm 1 Threshold Identifier

```
Import test file and assign to test procedure
Clean raw data
Make a learning dictionary from the data given
for each entry in test do
    Calculate probability prediction using MultinomialNb model creates.
    Calculate probability score.
    Identify average score (threshold)
    if score > threshold then
        label a word of add, for addition process;
    else
        label a word of cut, for subtracting process;
    end if
    Insert prediction and new score into test file.
end for
Display and save test file.
```

Fig. 1. Threshold identifier algorithm



X = repetitive process between the input and balancer

Fig. 2. Data training design

5. Result and Discussion

The result comparison is given between our prediction model and the actual value from the dataset (as the benchmark): https://bepi.mpob.gov.my [8]. Table 2 shows our price prediction approach against CPO price in 2021. It is the CPO local price for 2021 in Ringgit Malaysia (RM) per tan. It is worth noting that the CPO local price has more stable fluctuation patterns in 2021 (Figure 3), with only two times of change; in May to June and November to December. It shows a considerable difference between our result and the benchmark's result, with the average difference merely about 3.55%. The difference indicates that the proposed price prediction model using Bayesian probability can provide (nearly) comparable results.

We also executed the prediction model for the CPO local price for 2017 (in Table 3) to investigate how our model works when the fluctuation patterns are high. Note that the CPO local price in 2017 has changed six times in its pattern (Figure 4). Even though the price fluctuation pattern is high, our prediction model can still predict as the differences are less than 20%. It shows a 12.15% difference between our result and the benchmark's result when there is a high fluctuation in pricing. The price probability model can work with different price fluctuation patterns while producing stable predictions.

Table 2
Comparison results for year 2021

Month	Benchmark Data	Fluctuation Pattern	Bayes Prediction	Contradictory (%)
January	3,748.50	-	3,700.00	0.03
February	3,895.50	↑	3,800.50	0.06
March	4,041.50	↑	4,030.00	0.28
April	4,220.00	↑	4,200.00	0.05
May	4,572.00	↑	4,580.00	0.17
June	3,830.50	↓	3,850.50	0.52
July	4,128.50	↑	4,120.00	0.21
August	4,555.00	↑	4,560.00	0.33
September	4,556.00	↑	4,570.00	0.31
October	5,051.00	↑	5,000.00	1.00
November	5,341.00	↑	5,330.00	0.21
December	5,070.00	↓	5,050.50	0.38
Total				3.55

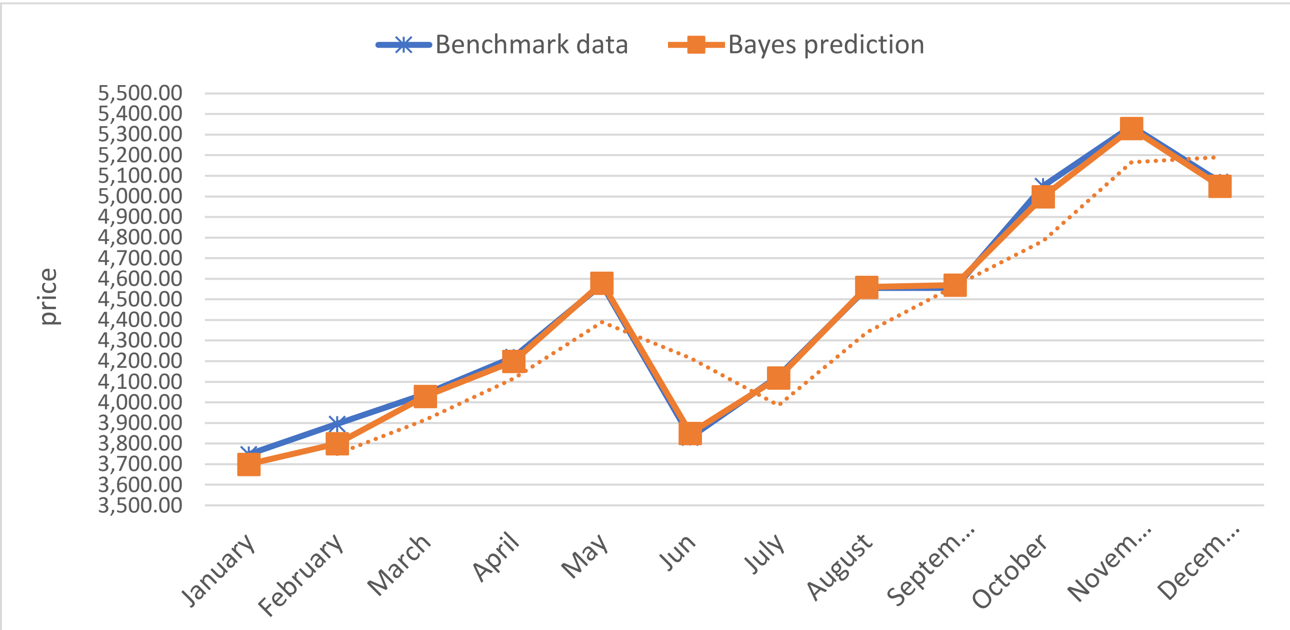


Fig. 3. Price fluctuation pattern for year 2021

Table 3
Comparison results for year 2017

Month	Benchmark Data	Fluctuation Pattern	Bayes Prediction	Contradictory (%)
January	3268.00	-	3262.00	0.18
February	3233.00	↓	3219.00	0.43
March	2955.50	↑	2915.00	1.37
April	2752.50	↓	2710.00	1.54
May	2803.50	↑	2800.00	0.12
June	2686.00	↓	2682.00	0.15
July	2629.50	↓	2692.00	2.38
August	2633.00	↑	2591.00	1.60
September	2780.50	↑	2790.00	0.38
October	2736.00	↓	2709.00	0.98
November	2689.00	↓	2609.00	2.98
December	2407.00	↓	2408.00	0.04
Total				12.15

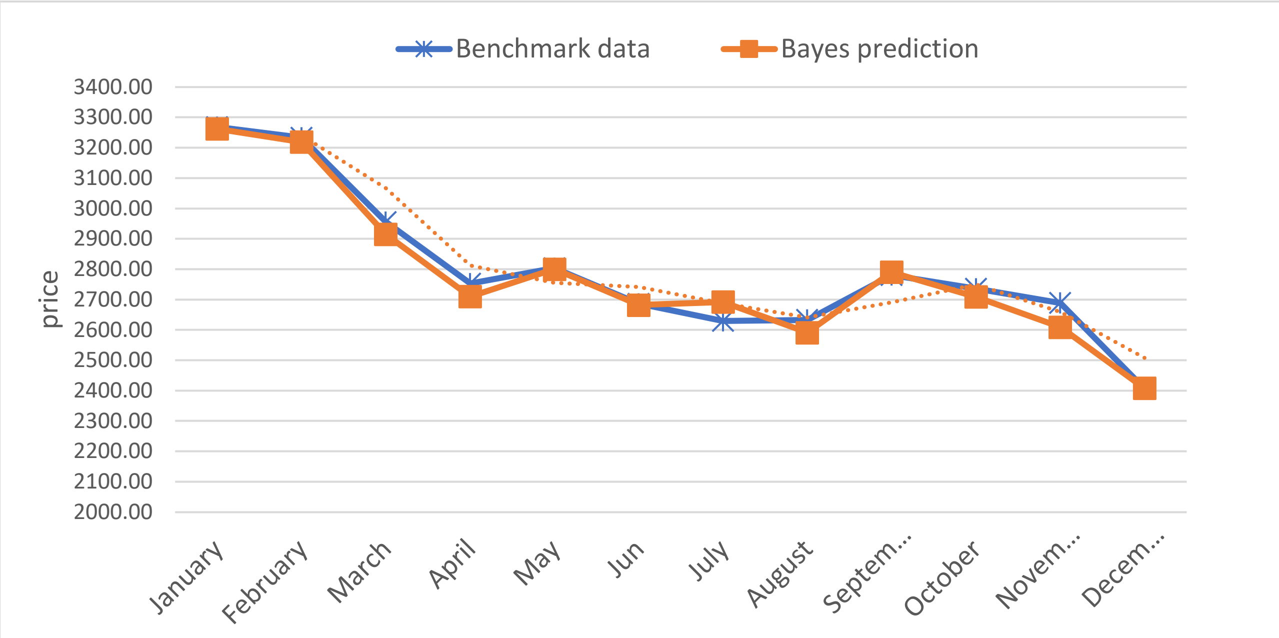


Fig. 4. Price fluctuation pattern for the year 2017

The linear regression analysis for the data for 2021 in Table 4 shows the *significance F* result is 1.6352E-13. Since the p-value is less than the significance level of 0.05, the regression model fits the data better than the intercept model. In the context of this specific data, it means that using the predictor variables Benchmark data and Bayesian Prediction in the model allows fitting the data better than if it is left out and used the intercept model.

R^2 is the proportion of the variance in the response variable that the predictor variable can explain. In this case, 99.63% of the variance in the data can be explained by the monthly CPO price. The standard error of the regression is the average distance the observed values fall from the regression line. The observed values fall an average of 33.55 units from the regression line.

Table 4
Simple linear regression analysis for Bayesian Prediction and Benchmark Data of the year 2021

Regression Statistics	
Multiple R	0.998164601
R Square	0.99633257
Adjusted R Square	0.995965827
Standard Error	33.55667126
Observations	12

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3059146.227	3059146	2716.705051	1.63521E-13
Residual	10	11260.50186	1126.05		
Total	11	3070406.729			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	76.2783	83.8502	0.9097	0.3844	-110.5516	263.1083	-110.5516	263.1083
X Variable 1	0.9868	0.01893	52.1220	1.6352E-13	0.9446	1.0289	0.9446	1.0290

For 2017 data, the *significance F* in Table 5 shows the result is 8.3774E-10. Since the p-value is less than the significance level of 0.05, it concludes that the regression model in 2017 also fits the data better than the intercept model. It also means that using the predictor variables, Benchmark data and Bayesian Prediction in the model allows fitting the data better than if it is left out and used the intercept model. The R^2 is 97.97% in this case, and the monthly CPO price can explain the data variance. The observed standard error of the regression values falls an average of 36.92 units from the regression line.

Table 5
Simple linear regression analysis for Bayesian Prediction and Benchmark Data in the year 2017

Regression Statistics	
Multiple R	0.989841137
R Square	0.979785477
Adjusted R Square	0.977764025
Standard Error	36.92092545
Observations	12

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	660712.7026	660712.7026	484.6938393	8.37739E-10
Residual	10	13631.54736	1363.154736		
Total	11	674344.25			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	12.38875667	126.26328	0.098118445	0.9237771	-268.94336	293.720875	-268.943362	293.7208753
X Variable 1	0.990016975	0.04496855	22.01576343	8.377E-10	0.8898208	1.09021315	0.8898208	1.09021315

Even though the quality of CPO can be one of the parameters used to assess its price, with the Bayes theorem model, the differences between the current and predicted prices can minimize. It helps the CPO commodity players reduce the burden of collecting many data parameters for forecasting the next accurate CPO price. Even though the actual CPO prices complement the Bayes-prediction model, there is still some price fluctuation. The fluctuation price can be explained by the limited factors investigated in this work, where demand and supply (volume) factors were not considered. Such factors are not part of the Bayesian probability because no clear identical fluctuation pattern can be determined through the volume of data in the real dataset. The investigation of the itemization in the dataset towards CPO price changes will be considered in our future work. Explicitly, external factors such as the supply of harvesters, fertilizer, and weather are significant aspects that must be thoroughly studied to produce more stable CPO prices.

6. Conclusion and Future Work

Crude palm oil (CPO) is one of the major contributors to Malaysia's economy by bringing investment, technology, and management knowledge. Therefore, the use of palm oil in the industry has become increasingly significant and substantial. However, the price of CPO keeps fluctuating over time. Because of the volatility of demand and supply, the agricultural sector heavily relies on product and innovation. In situations of considerable uncertainty, the CPO price forecasts are necessary to facilitate decision-making as there is a time lag that intervenes between making 'selling and buying 'decisions and the actual output in the CPO market. In this paper, we used Bayesian probability to predict CPO prices. We forecasted CPO prices daily and monthly, with a predictive period of up to 12 months. Specifically, we compared our Bayesian-based price prediction output with the benchmark data. A low standard error of the regression reveals that our model is comparable with the CPO prices produced by the existing CPO pricing system. In future, study the other external factors, such as CPO quality, that influenced the prices to determine pricing patterns automatically. The invention of a price prediction method can be observed throughout the different research areas. It will bring up agricultural and process improvements while ensuring the industry's sustainable development.

Acknowledgement
This work is supported by the Geran Putra (GP-IPM/2021/9703700), funded by Universiti Putra Malaysia. Masnida Hussin, Zurita Ismail and Iszuanie Syafidza Che Ilias contributed to the design and implementation of the research, to the analysis of the results and the writing of the manuscript. All authors provided feedback throughout the work and contributed to the final manuscript.

References

[1] Chuangchid, Kantaporn, Aree Wiboonpongse, Songsak Sriboonchitta, and Chukiat Chaiboonsri. "Factors affecting palm oil price based on extremes value approach." *International Journal of Marketing Studies* 4, no. 6 (2012): 54. https://doi.org/10.5539/ijms.v4n6p54

[2] Ab Rahman, Nik Muhammad Naziman, Abdol Samad Nawi, and Yusrina Hayati Nik Muhd Naziman. "The price discovery of the Malaysian crude palm oil futures markets." *Journal of Applied Finance and Banking* 2, no. 4 (2012): 25. https://doi.org/10.5539/ijef.v4n7p95

[3] Ab Rahman, Nik Muhammad Naziman. "The cointegration analysis on the spot prices of the Malaysian Crude Palm Oil futures market." *International Journal of Economics and Finance* 4, no. 7 (2012). https://doi.org/10.5539/ijef.v4n7p95

[4] Ticknor, Jonathan L. "A Bayesian regularized artificial neural network for stock market forecasting." *Expert systems with applications* 40, no. 14 (2013): 5501-5506. https://doi.org/10.1016/j.eswa.2013.04.013

[5] Khodakarami, Vahid, and Abdollah Abdi. "Project cost risk analysis: A Bayesian networks approach for modeling dependencies between cost items." *International Journal of Project Management* 32, no. 7 (2014): 1233-1245. https://doi.org/10.1016/j.ijproman.2014.01.001

[6] Huang, Yusheng, and Yong Deng. "A new crude oil price forecasting model based on variational mode decomposition." *Knowledge-Based Systems* 213 (2021): 106669. https://doi.org/10.1016/j.knosys.2020.106669

[7] Wang, Lili, Zitian Wang, Shuai Zhao, and Shaohua Tan. "Stock market trend prediction using dynamical Bayesian factor graph." *Expert Systems with Applications* 42, no. 15-16 (2015): 6267-6275. https://doi.org/10.1016/j.eswa.2015.01.035

[8] Herrera-Granda, Israel D., Leandro L. Lorente-Leyva, Diego H. Peluffo-Ordóñez, and M. M. E. Alemany. "A forecasting model to predict the demand of roses in an Ecuadorian small business under uncertain scenarios." In *Machine Learning, Optimization, and Data Science: 6th International Conference, LOD 2020, Siena, Italy, July 19–23, 2020, Revised Selected Papers, Part II 6*, pp. 245-258. Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-64580-9_21

[9] Kanchymalay, Kasturi, Naomie Salim, Anupong Sukprasert, Ramesh Krishnan, and Ummi Raba'ah Hashim. "Multivariate time series forecasting of crude palm oil price using machine learning techniques." In *IOP Conference Series: Materials Science and Engineering*, vol. 226, no. 1, p. 012117. IOP Publishing, 2017. https://doi.org/10.1088/1757-899X/226/1/012117

[10] Kaur, Harleen, Ritu Chauhan, and Siri Krishan Wasan. "A Bayesian network model for probability estimation." In *Encyclopedia of Information Science and Technology, Third Edition*, pp. 1551-1558. IGI Global, 2015. https://doi.org/10.4018/978-1-4666-5888-2.ch148

[11] Xiao, Qinkun, Chu Chaoqin, and Zhao Li. "Time series prediction using dynamic Bayesian network." *Optik* 135 (2017): 98-103. https://doi.org/10.1016/j.ijleo.2017.01.073

[12] Economics and Industry Development Division, Malaysian Palm Oil Board. "Economics and Industry Development Division," n.d. https://bepi.mpob.gov.my/.

[13] Kushairi, A., Ong-Abdullah, M., Nambiappan, B., Hishamuddin, E., Bidin, M. N. I. Z., Ghazali, R., et al. "Oil palm economic performance in Malaysia and R&D progress in 2018". *J. Oil Palm Res*, 31(2)(2019): 165-194.

[14] Nordin, Norhafiza, Sabariah Nordin, and Rusmawati Ismail. "The impact of commodity prices, interest rate and exchange rate on stock market performance: An empirical analysis from Malaysia." *Malaysian Management Journal* 18 (2014): 39-52. https://doi.org/10.32890/mmj.18.2014.9015

[15] Khowarizmi, Al, Rahmad Syah, Mahyuddin KM Nasution, and Marischa Elveny. "Sensitivity of MAPE using detection rate for big data forecasting crude palm oil on k-nearest neighbor." *International Journal of Electrical and Computer Engineering* 11, no. 3 (2021): 2696. https://doi.org/10.11591/ijece.v11i3.pp2696-2703

[16] Syahnur, Sofyan, Shabri A. Majid, T. Zulham, Jumadil Saputra, and Saiful Badli. "The effect of global prices of crude palm oil, marketing margins and palm oil plantations on the environmental destruction: An application of Johansen cointegration approach." (2019). https://doi.org/10.32479/ijeep.8010

[17] Karia, Abdul Aziz, and Imbarine Bujang. "Progress accuracy of CPO price prediction: Evidence from ARMA family and artificial neural network approach." *International Research Journal of Finance and Economics* 64, no. 64 (2011): 66-79.

[18] Kanchymalay, Kasturi, Naomie Salim, Anupong Sukprasert, Ramesh Krishnan, and Ummi Raba'ah Hashim. "Multivariate time series forecasting of crude palm oil price using machine learning techniques." In *IOP Conference Series: Materials Science and Engineering*, vol. 226, no. 1, p. 012117. IOP Publishing, 2017. https://doi.org/10.1088/1757-899X/226/1/012117

[19] Go, You-How, and Wee-Yeap Lau. "The impact of global financial crisis on informational efficiency: Evidence from price-volume relation in crude palm oil futures market." *Journal of Commodity Markets* 17 (2020): 100081. https://doi.org/10.1016/j.jcomm.2018.10.003

[20] Rahim, Nur Fazliana, Mahmod Othman, Rajalingam Sokkalingam, and Evizal Abdul Kadir. "Forecasting crude palm oil prices using fuzzy rule-based time series method." *IEEE Access* 6 (2018): 32216-32224. https://doi.org/10.1109/ACCESS.2018.2846809

[21] Salman, Nur, Armin Lawi, and Syafruddin Syarif. "Artificial neural network backpropagation with particle swarm optimization for crude palm oil price prediction." In *Journal of Physics: Conference Series*, vol. 1114, no. 1, p. 012088. IOP Publishing, 2018. https://doi.org/10.1088/1742-6596/1114/1/012088

[22] Aini, Hijratul, and Haviluddin Haviluddin. "Crude Palm Oil Prediction Based on Backpropagation Neural Network Approach." *Knowledge Engineering and Data Science* 2, no. 1 (2019): 1-9. https://doi.org/10.17977/um018v2i12019p1-9

[23] Scutari, Marco, and Jean-Baptiste Denis. *Bayesian networks: with examples in R*. CRC press, 2021. https://doi.org/10.1201/9780429347436

[24] Thambiah, Seethaletchumy. "Price forecasting methodology of the Malaysian palm oil market." *The International Journal of Applied Economics and Finance* 7, no. 1 (2013): 23-36. https://doi.org/10.3923/ijaef.2013.23.36