

## Forecasting Factors Influencing the Crude Palm Oil Market – A Composite Method

By Pavan Kumar Amarvadi and Juan Carlos Dahbura  
Thesis Advisor: Dr. Asad Ata

**Summary:** First, this research performs regression analysis on various palm oil market indicator variables for the country of Indonesia and discovers the relevant variables that explain the interactions within the market. Second, it uses five different methods to forecast the time series of the identified variables and discovers that Monte Carlo simulation gives the best result for forecasting trend of these variables. Moreover, third, it develops a tool for the practitioner to forecast the demand for their special chemical in the studied palm oil industry.



Pavan Kumar Amarvadi has received his Bachelor's degree in Maritime Technology from Birla Institute of Technology, Mesra, India. Before the SCM program, he was as a Maritime Officer onboard container vessels for a Danish company, Maersk Line sailing worldwide from India.

Juan Carlos  
Energy and  
gua and his  
El Salva-  
companies



Dahbura has received his received his Master's degree in Renewable Environment from the National Autonomous University of Nicaragua. Before the SCM program, he worked with Multinational FMCG in various supply chain roles.

### Key Insight

- Fifty independent variables explaining the ten dependent factors of the Indonesian Palm Oil market are identified.
- Ten market defining factors are forecasted with high accuracy using Monte Carlo Simulation for estimating trend for the independent variables.
- Simulation Tool developed for practitioner to forecast demand of special chemical based on the forecast of ten market factors.

### Research Background:

Chemicals International, aka Chemi, is a multinational specialty chemical company with annual revenue of six billion US dollars last year. They are a key supplier of 'Blearth,' a special chemical, in Indonesia which is used in refining crude palm oil into

refined palm oil. Indonesia is the largest producer of crude palm oil in the world with 54% of total world capacity. Competition for the supply of Blearth is intensifying in Indonesia, and Chemi aims to gain an edge by making tactical and strategic investments in its supply chain. However, to do so, the key market factors for Indonesian palm oil

industry are needed to be identified, quantified and modeled to enable forecasting. If such a forecast for the market factors is achieved, then a tool can be developed which can translate the palm oil market parameters into the demand for the special chemical Blearth.

### **Palm Oil Extraction:**

Oil palm trees bear fruits approximately every ten days and are harvested as fresh fruit bunches. These bunches are milled close to the plantations to produce crude palm oil. This crude palm oil is sent to refineries within a day to be converted in refined palm oil. This conversion process required Blearth which has the special property of removing impurities in the crude palm oil and also eliminate odour and colour. Blearth is very an effective but expensive component of the refining process as it consumes nearly 35% of the feed crude palm oil. To reduce costs, Blearth is manufactured in different grades at different price points. For example, If the refined palm oil is being produced for food industry then the most expensive grade is used. Therefore, it becomes necessary to identify factors of the palm oil industry that influence the demand of different grades of Blearth.

### **Research Method:**



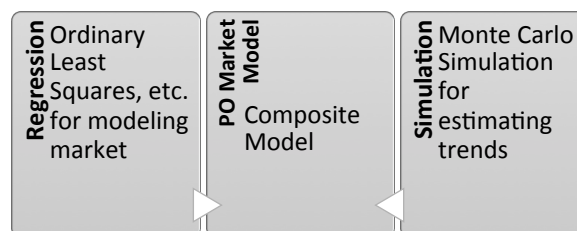
First a qualitative method was used to understand the dynamics of palm oil markets from Chemi executives. Their hypotheses about the causal interactions of various mar-

ket factors were noted upon. Using their expectations as a guideline a literature review was conducted to understand the behaviour of palm oil markets of other countries. It was discovered that time series econometrics and regression analysis types of quantitative research methods were being utilised in other countries. Indonesia on the other hand did not show evidence of any methods being used upon.

Fifty independent variables were identified as part of previous researches or models developed for Palm oil for other countries. Then, all available public databases were examined to capture secondary data of statistical nature for these fifty market indicator variables.

These variables were grouped into ten market factors that describe the primary and secondary areas of the market, (areas are production, domestic demand, exports and imports, to name a few) and regression analysis were performed. The data range available for the variables were not exhaustive, therefore to account for missing data and uncertainty in trends, several time series analyses were done. These results gave quantitative basis to analyse the qualitative understanding of the market and develop accurate forecasts.

### **Literature Review:**



The palm oil market models of Malaysia, Nigeria, Middle East and North Africa show the use of econometric models to capture the

interaction of various market factors. Indonesia on the other hand did not show to have such econometric models. It was also seen that there were several ordinary liner regression models that were also used in Malaysian markets to capture market interactions. Curiously, the Indonesian market did not present extensive statistical data about the palm oil industry market indicators.

Although fifty independent variables were identified, only about ten years' worth of statistical data was available for the region in the complete set of variables previously identified. Therefore, methods of accounting for this limitation were looked into.

It was discovered that combining regression or econometric models with certain time series analysis-simulation techniques could improve the accuracy of the model parameters and provide composite models that take feedback from the system and improve with time.

### Market Modelling:

The data for the identified fifty independent variables was available from 1964 to 2016 for variables like palm oil production but was only available from 2007 to 2016 for variables such as biodiesel production. The data range was made uniform from 2007 to 2016 for all fifty variables.

The palm oil market was broken down into six primary market components of: production, demand, import, export, world price and biodiesel. Biodiesel was selected as a primary market component although it is a part of the demand factor because of the unique emphasis is given by the Indonesian government to develop the biodiesel industry in the country.

Of these six, factors of demand were broken down further into demand for food and demand for industry, a factor of export was broken down into export of crude palm oil and export of refined palm oil. This resulted in ten market factors that needed to be explained, estimated and forecasted to form a model of the market.

1. *Indonesian palm oil market* =  
*production + demand + import + export + world price.*
2. *Demand* = *demand food + demand industrial*
3. *Demand industrial* =  
*other industrial uses + biodiesel.*
4. *Export* =  
*export crude + export refined*

Estimating the above equations would provide a forecast of the entire palm oil market, which will be later translated into the market forecast of special chemical Blearth by combining it with the average dosage of the 'Blearth' in the palm oil industry.

5. *Blearth market* =  
*(crude palm oil production – export crude) \* Blearth dosage*

Based on the literature review, for each of the ten market factors, the relevant independent variables were grouped according to other country market models and inputs from Chemi executives. On these groups three types of multivariate regression analysis were done:

1. Ordinary least squares without time trend
2. Ordinary least squares with time trend
3. Two-stage least squares

The literature showed the use of OLS, 2SLS and ARDL models in econometric analysis to develop complex models that allow for

causal inferences. In this research, only OLS and 2SLS were used to make simplistic models combined with time series analysis methods to only allow for forecasting and not allow for causal inferences.

Upon performing regression analysis, the most relevant variables for each of the factors were selected in two steps:

1. Allow software to identify all the combinations of variables for each factor that provide highest  $R^2$  value.
2. Make manual selection of each of the combination identified above based on following criteria:
  - a.  $R^2 \geq 0.8$
  - b.  $\text{Adj } R^2 \geq 0.8$
  - c.  $p \text{ value} \leq 0.05$
  - d.  $f \text{ test} \leq 0.05$

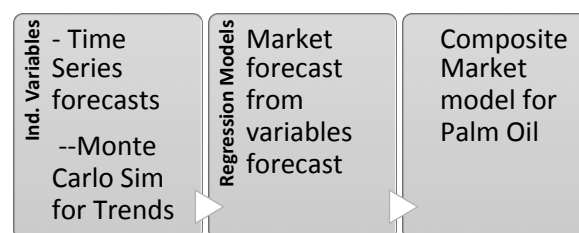
The resulting regression models were validated and found to comply with

1. set  $R^2$  criteria
2. t-statistic shows slope  $> 0$
3. Normality of errors\*
4. domains of predictors are within range of sample data
5. homoscedasticity of error terms

\*Normality was observed for regression results for world price factor but did not show for some factors like CPO demand. It was assumed that this is occurring not because of lack of normality in the errors but because of lack of extensive data for the relevant variables. Since the variables for the world price equation factor had data from 1964 they showed normality, and if the other factor equations had such extensive data they would show normality too.

Methods of incorporating the missing data or increasing granularity from yearly to monthly was considered.

### Forecasting:



The previously mentioned three regression models explained the palm oil market of Indonesia but to use the regression models for forecasting, the future values of the relevant independent variables needed to be known.

Two types of time series analysis were done with two different settings to arrive at a forecast for the fifty independent variables, each resulting in four time series analyses:

1. Moving average,  $M=3$  years
2. Moving average,  $M=5$  years
3. Holt's,  $\alpha=0.1$ ,  $\beta=0.1$
4. Holt's,  $\alpha=0.1$ ,  $\beta=0.2$

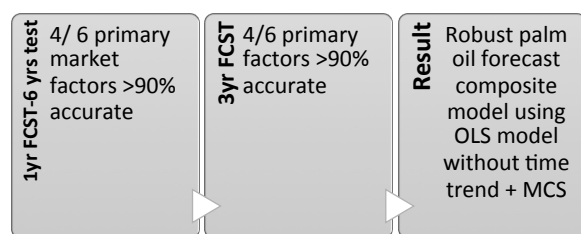
The forecasts arrived using these methods were put into the three market regression models for the ten factors to forecast the market size of palm oil for one year. The MAPE values of these forecasts were compared. It was found that neither of these four models provided consistency in accuracy across all ten market factors in all three-regression analysis.

It was considered that the inconsistency could be due to the high level of uncertainty in variables such as commodity prices, interest rates and exchange rates. To account for this uncertainty Monte Carlo Simulation (MCS) was considered as it could incorporate this uncertainty. Also, as the data set established was limited in range concerning all the variables mentioned in the literature, simulation method was best suited. As mentioned before the world price factor equation showed normality, so it was assumed that the rest of the factors could show similar

characteristic if more data was included in the analysis.

The past average trend values and standard deviation of the trend was used as inputs for the MCS. The forecasted trend values were used to forecast the values of the independent variables. The MAPE value of the forecasts were compared with the forecasts of the previous four time series forecasts and it was found that the MCS average provided a much more consistent accuracy across the ten market factors in all three regression models.

### Validation:



The next step was to identify the best combination of a regression model with a variable forecasting method. This was done by making a one-year forecast for each of the ten market factors using all five forecasting techniques and feeding the values in all three regression models. This one-year forecast was compared for six years, and it was found that the combination of using Monte Carlo Simulation for estimating trends and Ordinary Least Squared regression without time trend provided accuracy greater than 90% consistently for six years.

Specific market factors such as import, and biodiesel did not show this level of accuracy, and it was associated with more uncertainty that is present in the market concerning the recent developments in the industry and dependency on government subsidies.

An attempt was made to use this technique to forecast for a five-year mid to longer term horizon by using the one-year forecasts as input for the subsequent year in a loop until a five-year forecast was complete. It was found that for three primary factors of production, demand and export crude with which the market size can be estimated, accuracy >90% was derived for a range of three years from the baseline year.

Further, these forecasts were compared with the outlooks published by expert industry watchers, and it was found that these results perform as well and income cases better than the industry experts. This validation process provides sufficient confidence in the robustness of the developed model.

### Simulation Tool:

An MS Excel VBA simulation tool was developed for Chemi which allow them to perform a five-year market forecast using the composite model. From a database file, the statistical data for each of the fifty variables was fed into the simulator, which generated trends for the variables and forecasted their values on year by year basis for five years. These variable forecasts were translated into palm oil factor values which were further translated into Blearth market size values using the equations previously mentioned while also including inputs from Chemi about industry average dosage and target market shares. The Blearth market size was further divided into the market size of each grade of Blearth by taking the input of portfolio grade split. Based on the findings of this research it has been advised that a three-year horizon forecast is recommended and for a higher time horizon caution is advised.

### Limitation:

The primary assumption this research stands on is that the missing normality in the regression models is due to restrictions made in the dataset range to accommodate all the variables as suggested in the previous researches and it was intended that MCS would be used along with a normal probability distribution. It is to be seen if over time with the collection of more data by the Indonesian governmental agencies improves these equations and the identified independent variables remain the same.

Further, the identified palm oil models of other countries used time series econometric models of regression of the form ARDL. With the availability of more data in the future other models of higher complexity can be developed using such advanced models. Also, it has been advised that system dynamics type of simulation could be more flexible to needs of some of the market factors like biodiesel, which the current model does not provide for sufficient results.

### Conclusions:

Finally, this research identified the factors governing the Palm Oil Market of Indonesia, developed a composite model for commodities, which uses multilinear regression approach plus a simulation for time series analysis and developed a simulation tool that determines the three-year forecast of palm oil market and Chemi's special chemical Blearth in Indonesia.

Additionally, the accuracy of the forecasting model developed was able to predict this stock reliably for three years ahead. This value was the key result that Chemi has been looking for. By incorporating some company-specific data, this crude oil stock level was converted into a demand for special chemical Blearth.

It is to be seen how this model would perform in the neighboring countries of Vietnam, Thailand etc. In fact, this model could be used for modeling the markets of other edible oils too. Such an analysis will allow to improve the robustness of this model or help identify a weakness that could be improved upon in further iterations.

The authors would like to thank the reader for their time in perusing through this report and would like to direct them to the full version of the thesis for an exhaustive list of citations that have been used to perform this research.

### References:

- Applanaidu Shri Dewi a/p [et al.]** The Relationship among Petroleum Prices, Biodiesel Demand, and Malaysian Palm Oil Prices: Evidence from Simultaneous Equation Approach [Journal] // BANWA VOL. 8, NOS. 1&2. - 2011. - pp. 64-83.
- Abdel-Hameed Amna Awad** An Econometric Study Of Palm Oil Import Demand In The Middle East And North African Countries [Book]. - Kuala Lumpur: Universiti Putra Malaysia, 2005.
- Allen P Geoffrey** Economic forecasting in Agriculture [Journal] // International Journal of Forecasting. - 1994. - pp. 81-135.
- Applanaidu Shri Dewi [et al.]** An Econometrics Model of the Malaysian Palm Oil Market [Conference] // Workshop on Agricultural Sector in Malaysia: Quantitative Models For Policy Analysis. - Johor Bahru: Institute of Agricultural And Food Policy Studies, University Putra Malaysia, 2009.
- Labys W. C.** Commodity Models for Forecasting and Policy Analysis [Book]. - [s.l.]: Nichols Publishing Co., 1984.
- Egwuma Henry [et al.]** A Model for the Palm Oil Market in Nigeria: An Econometrics Approach [Journal] // International Journal of Food and Agricultural Economics. - 2016. - pp. Vol. 4 No. 2, pp. 69-85.

**Food and Agriculture Organization of the United Nations** All FAOSTAT Data [Online] // FAOSTAT . - 26 October 2017. - <http://www.fao.org/faostat/en/#home>.

**Hassan Asari Fadli Fizari Abu [et al.]** A Time Series Analysis of the Relationship Between Total Area Plantes, PalmOil Price and Produciton of Malaysia Palm Oil.

**Forecasting with Regression Analysis** – 9-894-007, Harvard Business School.