

# PREDICTING PURIFIED TEREPHTHALIC ACID (PTA) PRICE USING REGRESSION ANALYSIS AND SAME SLOPE MODEL

Ajdin Vatres, Hadis Bajric, Edin Kadric, Mirza Pasic



**This Publication has to be referred as:** Vatres, A[jdin]; Bajric, H[adis]; Kadric, E[din] & Pasic, M[irza] (2018). Predicting Purified Terephthalic Acid (PTA) Price Using Regression Analysis and Same Slope Model, Proceedings of the 29th DAAAM International Symposium, pp.0846-0852, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-20-4, ISSN 1726-9679, Vienna, Austria  
DOI: 10.2507/29th.daaam.proceedings.122

## Abstract

In this paper, the relationship between PTA (purified terephthalic acid) price and oil price was tested and a linear regression model was established. The following prediction models were tested on a historical oil price time series: the Same Slope, the Weighted Moving Averages, the Simple Exponential Smoothing, and the Exponential Smoothing With Additive Trend or the Holt's model. In order to determine the unknown parameters of the used models, nonlinear mathematical programming was used. Performance analysis on the given time series showed that the Same Slope model yielded better results than other tested models. The oil forecast given by the Same Slope model was then combined with the initial regression equation to establish the final PTA predicting model.

**Keywords:** predicting PTA price; regression analysis; same slope model; forecasting

## 1. Introduction

Nearly all purified terephthalic acid (PTA) is consumed in polyester production including polyester fibre [1]. This means that PTA price is an essential information for polyester fibre manufacturers, those making products from polyester yarn and competing textile manufacturers. In order to be able to adjust the cost of their products and services in a timely manner and to be able to make proper long-term strategic decisions, manufacturers need a prediction system. The aim of this paper was to create a model that would predict the price of PTA.

There is a well-documented widespread amount of research that establishes the link between oil prices and the price of different products. Thus, for example, [2] creates a regression model that explains variations in crude palm oil (CPO) and oil prices. Furthermore, in [3] the effects of oil price shocks on cotton prices have been examined and it was found that the monthly changes in cotton prices were significantly affected by unexpected increases in the global oil demand driven by increased global real economic activity.

In the study [4] impact of oil price variation on 14 industries in the following six markets: Canada, China, France, India, the U.K., and the U.S., from June 1998 to December 2011 is investigated and authors conclude that, in general, when the price of oil rises or the oil market is unstable, the production costs increase as well.

Taking into account the fact that oil is a resource for the production of PTA, the relationship between their prices is tested in this paper and a model that predicts PTA price on the market has been developed. Theoretical basics of conducted regression analysis are given in [5]. The following models were used and compared: SS defined in [6], [7], [8], WMAM with model information available in [9], and SES and Holt model defined in [10].

Nomenclature	
Symbol	Definition
SS	Same Slope model
WMAM	Weighted Moving Averages model
SES	Simple Exponential Smoothing model
Holt's model	Exponential Smoothing With Additive Trend model
MSE	Mean Square Error
MAPE	Mean Absolute Percent Error
$y$	Predicted PTA price as output
$x$	Predicted oil price as input
$X_t$	Observed value of the time series in period $t$
$\hat{X}_t(m)$	Forecast for $m$ periods ahead from origin $t$
$m$	Number of periods in the forecast lead-time
$\beta$	Coefficient of the SS model
$w$	Weight coefficient of the WMAM model
$S_t$	Smoothed level of the series, computed after $X_t$ is observed
$\alpha$	Smoothing parameter for the level of the series
$T_t$	Smoothed additive trend at the end of period $t$
$\gamma$	Smoothing parameter for the trend

**2. The forecasting model**

Since crude petroleum is used as one of the principle raw materials for the production of substances later used for the production of yarn, a model was designed which would take information about the raw material price on the market as the input data. To establish a mathematically explicit link between crude petroleum price and granulate prices, a regression model was developed using linear regression. As the independent variable, the price of crude petroleum was used, and the dependent variable was the price of granules.

Before the regression analysis was performed a graph showing the relationship between PTA price and oil price was plotted and it is shown in Fig. 1. The results obtained by regression analysis for analyzed granules of PTA are shown in Table 1.

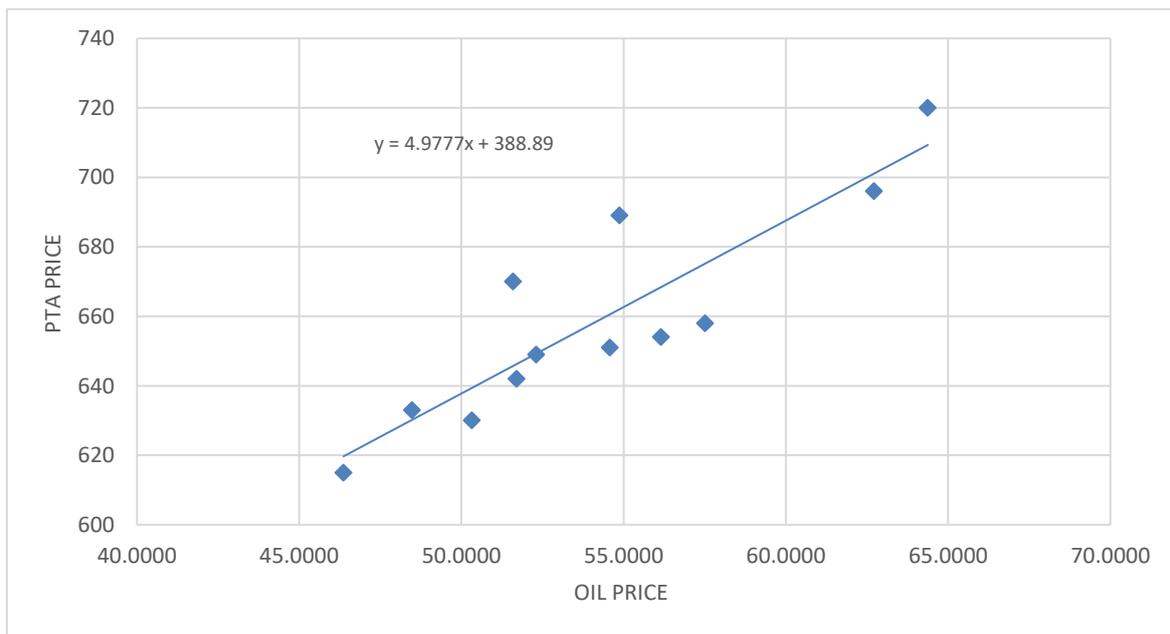


Fig. 1. Price of PTA vs. price of oil

<i>Regression Statistics</i>	
Multiple R	0.884368233
R Square	0.782107171
Adjusted R Square	0.760317888
Standard Error	14.79283449
Observations	12

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	7854.637143	7854.637143	35.89412165	0.000133674
Residual	10	2188.279524	218.8279524		
Total	11	10042.91667			

Table 1. Results of regression analysis

From the results presented in Table 1 it can be seen that the coefficient R Square is 0.78 which indicates a strong correlation between the dependent and independent variables and that the regression model is significant.

To test how fast the PTA market responds to the change in oil prices and whether a delay in response to change in oil prices does exist, three additional regression analysis were conducted in which the price data for oil was moved one step backward, that is one, two and three months back. The results of those analysis are shown in Tables 2, 3 and 4.

<i>Regression Statistics</i>	
Multiple R	0.856575547
R Square	0.733721668
Adjusted R Square	0.707093835
Standard Error	16.35301532
Observations	12

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	7368.705566	7368.705566	27.55468918	0.000373787
Residual	10	2674.2111	267.42111		
Total	11	10042.91667			

Table 2. Results of regression analysis with data shifted one month back

<i>Regression Statistics</i>	
Multiple R	0.578237443
R Square	0.334358541
Adjusted R Square	0.267794395
Standard Error	25.85533157
Observations	12

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	3357.934962	3357.934962	5.023102694	0.048902255
Residual	10	6684.981705	668.4981705		
Total	11	10042.91667			

Table 3. Results of regression analysis with data shifted two months back

<i>Regression Statistics</i>	
Multiple R	0.015978571
R Square	0.000255315
Adjusted R Square	-0.099719154
Standard Error	31.68651537
Observations	12

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	2.564104637	2.564104637	0.002553799	0.960691118
Residual	10	10040.35256	1004.035256		
Total	11	10042.91667			

Table 4. Results of regression analysis with data shifted three months back

Based on test results shown in Tables 2, 3 and 4, it can be seen that the value of the coefficient R Square is dropping, thus the first analysis will be used, as no indication that any latency in the price adjustment on this time scale exists. The resulting model given by the first analysis is given by (1):

$$y = 4,9777 \cdot x + 388,89 \tag{1}$$

For the model defined by (1) to give PTA price forecasts in the desired time period, it is necessary to know the price of oil for the same time period. As the analysed time frame is in the future, and the values of oil on the market are currently unknown, the required data will be obtained by predicting the value of oil prices for the required period.

### 3. Forecasting oil prices in the required time period

To create a prediction as precise as possible, four simple models, based on time series analysis, were selected: SS, WMAM, SES and Holt. The selection of the most suitable model is performed in such a way that the created models obtained by applying each of the above methods are mutually matched, by comparing their MSE, MAPE and bias values. The time series with historical data on oil prices, which was used to train the models was obtained from [11], was windowed between 2000 and 2017. The graphical interpretation of the models is shown in Figures 2, 3, 4 and 5, and the comparison of the MSE, MAPE and bias values of the models is shown in Table 5.

SS model used in this paper assumes that the gradient (slope) present in the previous time interval will continue to have the same effect and value in the following time period. This model is defined by (2). The idea and theory of this model as well as its settings and the comparison of its effectiveness are presented in [6], [7], [8].

$$\hat{X}_t(1) = X_t - \beta(X_{t-1} - X_t) \tag{2}$$

In order to determine the optimal value of the coefficient  $\beta$ , nonlinear mathematical programming was used, where standard deviation or MSE of the model was minimized.



Fig. 2. Graphical presentation of SS model

WMAM model used in this paper assumes that the date used to calculate moving average are not of equal importance, that is why their values have been modified using weight factors. WMAM model is defined by (3), while the theory of moving average models is given in [9].

$$\hat{X}_t(1) = \frac{\sum_{i=1}^n w_{n+1-i} X_{t-i}}{\sum_{i=1}^n w_{n+1-i}}, \quad (3)$$

where  $n$  is the number of previous data points used to create the prediction. In the case of this paper  $n=3$  was chosen in order to create a balance between using the newest information and smoothing out the prediction. In order to determine the optimal values of the weight coefficients  $w$ , nonlinear mathematical programming was used, where standard deviation or MSE of the model was minimized.



Fig. 3. Graphical presentation of WMAM model

SES is the standard type of exponential smoothing model defined by (4) and is listed among models talked about in [9], [10].

$$\left. \begin{aligned} S_t &= \alpha X_t + (1 - \alpha)S_{t-1} \\ \hat{X}_t(m) &= S_t \end{aligned} \right\} \quad (4)$$

In order to determine the optimal value of the coefficient  $\alpha$ , nonlinear mathematical programming was used, where standard deviation or MSE of the model was minimized.



Fig. 4. Graphical presentation of SES model

Holt's model is defined by (5) and is listed among models talked about in [10].

$$\left. \begin{aligned} S_t &= \alpha X_t + (1 - \alpha)(S_{t-1} + T_{t-1}) \\ T_t &= \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1} \\ \hat{X}_t(m) &= S_t + mT_t \end{aligned} \right\} \quad (5)$$

In order to determine the optimal value of the coefficient  $\alpha$  and  $\gamma$ , nonlinear mathematical programming was used, where standard deviation or MSE of the model was minimized.



Fig. 5. Graphical presentation of Holt's model

Nonlinear mathematical programming problems were solved using Solver in Microsoft Excel application. In order to better understand the model's behavior, Fig. 2, 3, 4 and 5 show the real oil price as well as the price obtained by using the appropriate prediction model for 2017 (from January to December).

Model	MSE	MAPE	Bias
SS	27.62	7.17%	22.74
WMAM	35.59	7.59%	64.37
SES	28.19	7.1%	29.45
Holt	28.36	7.16%	24.2

Table 5. Comparison of the MSE, MAPE and bias values of the models

From Table 5 it can be seen that the SS model has the smallest MSE and bias value, SES model has the smallest MAPE value. It is important to notice that all models have a positive bias meaning they overestimate the oil price. One of the most commonly used measures of the quality of the model is MSE and it was chosen as the deciding factor. An important advantage of the SS model is that it does not require non-linear mathematical programming to determine the optimal parameter values, and in [6], [7], [8] it is shown that in other cases (using other time series) the SS model gives reliable results, often better than models with similar rank of complexity. The SS model was used to generate the oil price prediction.

**4. Conclusion**

Since oil is one of the basic raw materials for the production of PTA, the paper has confirmed the existence of a strong linear correlation between oil prices and PTA prices, on the basis of which a model for predicting PTA price has been developed. It was found that there was no delay on the monthly basis of the price change of the PTA with the change in the price of oil. It is possible that such an occurrence exists on a daily basis, which is not explored in this paper and requires further consideration. Various models for predicting oil prices based on time series analysis have been tested. The used models were mutually compared by their MSE, MAPE and bias values, and it is shown that all models have a positive bias value, indicating that the models tend to overestimate the predicted value at a given time.

The best results were obtained by the SS, with the lowest value of the MSE, and the SES model, with the lowest MAPE value. As the differences in the values for the two models mentioned above are minimal, one can conclude that they are of the same level of reliability, but the advantage in this case is given to the SS model due to lower values of MSE and bias. In addition to the oil price, as the main resource for PTA production, other potentially significant factors such as: energy prices, logistical costs, worker fees associated with living standards, etc. could be affecting its cost. It should be tested which other factors will have a significant impact. The developed SS model can be incorporated into a company ERP system.

**5. References**

[1] <https://www.icis.com/resources/news/2007/11/06/9076461/purified-terephthalic-acid-pta-uses-and-market-data/>, ICIS, Accessed on: 2018-05-09

[2] Khaeri, H. & Andriyana, Y. (2018). Crude Oil Price Effect on Crude Palm Oil (CPO) Pricing Model by Nonparametric B-Splines Regression Approach, Available from: <https://www.researchgate.net/publication/326762786>, Accessed on: 2018-08-10

[3] Mutuc, M.; Pan, S. & Hudson, D. (2010). What Drives Commodity Prices More: Oil Demand or Supply Shocks?, Available from: <https://www.researchgate.net/publication/254383758>, Accessed on: 2018-07-12

- [4] Randjbaran, E., Tahmoorespour, R., Rezvani, M., & Safari, M. (2018). The Impact of Oil Price Fluctuations on Industry Stock Returns: Evidence from International Markets, *Journal of Management Info*, Vol. 5, No. 1, (March 2018) page numbers (1 – 12), <https://doi.org/10.31580/jmi.v5i1.26>
- [5] Dekking, F.M.; Kraaikamp, C.; Lopuhaä, H. P. & Meester, L. E. (2007). *A Modern Introduction to Probability and Statistics: Understanding Why and How*, Springer, ISBN 1852338962, London
- [6] Bajric, H.; Kadric, E. & Pasic, M. (2016). A Comparison of Same Slope Seasonality and Holt-Winters Smoothing Forecasting Models, *Proceedings of 26th DAAAM International Symposium*, ISBN 978-3-902734-07-5, ISSN 1726-9679, Katalinic, B. (Ed.), pp. 0062-0068, Published by DAAAM International, Vienna, Austria, DOI: 10.2507/26th.daaam.proceedings.009
- [7] Bajric, H.; Bijelonja, I. & Pasic, M. (2009). A Comparison of Same Slope Seasonality and Exponential Smoothing Forecasting Models, *Proceedings of 20th DAAAM International Symposium*, Katalinic, B. (Ed.), pp. 1343-1344, Published by DAAAM International, Vienna, Austria
- [8] Pasic, M.; Bijelonja, I.; Sunje, A. & Bajric, H.; (2007). Same Slope Forecasting Method, *Proceedings of 18th DAAAM International Symposium*, Katalinic, B. (Ed.), pp. 547-548, Published by DAAAM International, Vienna, Austria
- [9] Nau, R. (2014). Forecasting with moving averages, Available from: [http://people.duke.edu/~rnau/Notes\\_on\\_forecasting\\_with\\_moving\\_averages--Robert\\_Nau.pdf](http://people.duke.edu/~rnau/Notes_on_forecasting_with_moving_averages--Robert_Nau.pdf), Accessed on: 2018-08-04
- [10] Gardner E. S. (2006). Exponential smoothing: The state of the art – Part II, *International Journal of Forecasting*, Vol. 22, No. 4, (October – December 2006) page numbers (637 – 666), ISSN 0169-2070
- [11] <http://tonto.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RB RTE&f=M>, The U.S. Energy Information Administration (EIA), Accessed on: 2018-05-09