# PREDICTION OF PHYSICAL PROPERTIES OF OIL PALM BIOMASS REINFORCED POLYETHYLENE: LINEAR REGRESSION APPROACH

**Syamsiah Abu Bakar [1], Rosma Mohd Dom [1], Ajab Bai Akbarally[1] and Wan Hasamudin Wan Hassan[2]**

[1]*Faculty of Computer & Mathematical Sciences, University Technology MARA (UiTM), Malaysia,*
*rosma@tmsk.uitm.edu.my*
[2]*Biomass Technology Centre, Engineering & Processing Division, MPOB (Malaysia),*
*wanhaswh@mpob.gov.my*

**ABSTRACT**. In recent years, there has been an increasing interest on renewable resources for consumer products and biodegradable materials. Traditional polymeric materials derived from petro-chemical sources do not degrade and disposal of such materials is a major concern in minimizing the environmental problems. Currently, experiments are carried out in laboratories to determine the physical properties of degradable plastics which include melt flow index (MFI), melting point (MP) and Density. Oil palm biomass (OPB) is used as bio-active components in the formulation with Polyethylene (PE). Alternatively, a different approach is required as to minimize the time consume, the cost of production and the cost of labor. In this study, Linear Regression model has been developed and used to predict the physical properties of degradable plastics. The ability of Linear Regression model is assessed by comparing the theoretical results with the actual lab results using correlation coefficient ($r$) and coefficient of determination ($R^2$). The result showed that the percentage prediction accuracy for MFI is 93%, 71% for the prediction of MP and 24% for the prediction of Density respectively using linear regression. The study proves that the use of Linear Regression model for predicting the physical properties of degradable plastics is highly feasible.

**Keywords:** Oil palm biomass (OPB), polyethylene (PE), Linear Regression

## INTRODUCTION

Currently experiments are carried out in laboratories to determine the physical properties of degradable plastics which include Melt flow index, Melting point, and Density by using Oil palm biomass as bio-active components in the formulation. The whole process is very time consuming and costly. Thus there is a need for an alternative method of modeling the properties of fibre-reinforced polymer composites. Traditional mathematical equations derived provide a way for solving real world problems. A study by Sin *et al*. (2010) showed that Linear Regression method is capable of predicting the properties of fibre-reinforced polymer composites.

### Fundamental of Linear Regression Model

Regression model is a statistical method commonly used for modeling relationships between variables. Introduced by Sir Francis Golton in 1985's, the idea of simple linear regression was later extended into other methods of regression such as multiple linear regression, logistic regression and nonlinear regression. Linear Regression is a model with a response variable $y$ that has a relation with single regressor $x$. A typical equation of Linear Regression is as follows:

$$y_i = B_o + B_1 x_i + \in_\mathrm{i}\,, \qquad i = 1, 2, \dots, n \qquad (1)$$

where, $y_i$ is the dependent variable, $B_o$ is the intercept, $B_1$ is the slope and $\in_\mathrm{i}$ is the error. The present study uses Linear Regression model in predicting the physical properties of degradable plastics. Linear Regression model is being used in this study since reports have shown that Linear Regression produces good prediction results through simple computation (Sin *et al*., (2010).

**Background of Natural Fibres**

Studies of natural fibre as filler in producing degradable plastics have attracted numerous researchers and academician in pursuing industrial research. Examples include the investigation of the mechanical and physical properties of degradable plastics (Mungara *et al*., 2002); studies on degradable plastics' thermal behaviour (Mangal *et al*., 2003) and the studies on the effect of treated and untreated fibers on degradable composites (Sreekala and Thomas, 2003).

Recently, there have been an increasing number of studies on oil palm biomass (OPB). Different types of oil palm biomass that can be used as fillers are oil palm empty fruit bunch and oil palm trunk wood flour (Badri *et al*., 2005; Zaini *et al*., 1994). The rationale behind the interest in using OPB as fillers is due to several benefits which include the renewable nature of OPB, OPB's lower density and its amenability to chemical modification as well as lower cost of accumulation (Badri *et al*., 2005).

Numerous mathematical models have been applied in the study of degradable plastics. such as regression model (Sin *et al*., 2010); ANFIS model (Lee *et al*., 2008); Fuzzy model (Muc and Kedziora, 2001); modified rule of mixtures (ROM) strength and simple ROM strength (Facca *et al*., 2007); micromechanical model (Facca *et al*., 2006) and, Tobias and Y. Agari model (Mangal *et al*., 2003). The present study focuses on the application of Linear Regression in predicting the physical properties of oil palm biomass (OPB) reinforced polyethylene (PE).

The objective of this study is to develop Linear Regression model for the prediction of MFI, MP and Density of degradable plastics. The ability of Linear Regression model in predicting MFI, MP and Density of degradable plastics will be assessed by comparing theoretical results with actual lab results using correlation coefficient ($r$) and coefficient of determination ($R^2$).

**METHODOLOGY**

Linear Regression model is developed and used to identify the suitable level of factors affecting the MFI, MP and Density of degradable plastics using MATLAB programming. The method of this study is divided into 3 steps as follows:

Step 1: Identify input and output values.

The results obtained from the Malaysian Palm Oil Board (MPOB) laboratories are used to develop the Linear Regression model. The two inputs are Polyethylene (PE) (%) and Oil palm biomass (OPB) (%).The outputs are the physical properties of degradable plastic composites namely the Melt flow index (g/10min), Melting point ($^o$C) and Density (g/cm$^3$). A resampling technique known as bootsrapping was applied to the data collected by MPOB to produce a sample of 220 data. The data were divided into two sets, the training and testing data sets with the partition of 50%-50%. The training data were used to generate Linear Regression formula while the testing data were used to assess the ability of Linear Regression model based on the Corelation Coefficient ($r$) and the Coefficient of Determination ($R^2$) values.

Step 2: Generate Linear Regression formula.

To create Linear Regression equation, the parameter $\beta_o$ and $\beta_1$ has to be estimated. Here, the method of least square was used such that the sum of the squares of the difference between the observation $y_i$ and the straight line is minimized (Montgomery, 1991).

Let say $\hat{\beta}_0$ and $\hat{\beta}_1$ are the estimated values of $\beta_o$ and $\beta_1$ respectively. Hence the predicted values, $\hat{y}_i$ is given as follow:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (2)$$

The difference between the observations, $y_i$ and the predicted value, $\hat{y}_i$ is called residual, $e_i$. The smaller the residuals, the better fitted value is achieved. The residual is given as:

$$e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \quad (3)$$

The least square criterion is

$$s(\beta_0, \beta_1) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 \quad (4)$$

The least square estimator for $\beta_o$ and $\beta_1$ , say $\hat{\beta}_0$ and $\hat{\beta}_1$ has to satisfy the condition below:

$$\frac{\partial s}{\partial \beta_0}|_{\hat{\beta}_0,\hat{\beta}_1} = -2\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \quad (5)$$

and

$$\frac{\partial s}{\partial \beta_0}|_{\hat{\beta}_0,\hat{\beta}_1} = -2\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)x_i = 0 \quad (6)$$

Simplifying equations (5) and (6) gives

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \quad (7)$$

and

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i\, x_i \quad (8)$$

Equation (9) is called the least squares normal equation. The solution for $\hat{\beta}_0$ and $\hat{\beta}_1$ are as follow:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (9)$$

and

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \frac{(\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} x_i)}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}} \quad (10)$$

where

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \text{and} \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (11)$$

Now, $\bar{y}$ and $\bar{x}$ are the averages of $y_i$ and $x_i$ respectively. Thus from equations (9 ) and (10), $\hat{\beta}_0$ and $\hat{\beta}_1$ are the least squares estimators of the intercept and slope. Below is the fitted simple linear regression model:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \quad (12)$$

From equation (10), the numerator is the corrected sum of the cross products of $x_i$ and $y_i$. The denominator is the sum of the cross products of $x_i$ and $y_i$. These two quantities are repesented as:

$$s_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} = \sum_{i=1}^{n}(x_i - \bar{x})^2, \quad i = 1,2,\ldots,n \quad (13)$$

and

$$s_{xy} = \sum_{i=1}^{n} y_i\, x_i - \frac{(\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} x_i)}{n} = \sum_{i=1}^{n} y_i(x_i - \bar{x}) \quad (14)$$

Thus, equation (10) can be rewriten as

$$\hat{\beta}_i = \frac{s_{xy}}{s_{xx}} \quad (15)$$

Step 3: Assessing Ability of Linear Regression.

This study assesses the ability of Linear Regression model in the prediction of the physical properties of degradable plastics based on the correlation coefficient (*r*) and the coefficient of determination ($R^2$) values calculated in MATLAB R2009a environment. The formulae of $r$ and $R^2$ used are as follows (Wang and Elhag, 2008):

$$r = \frac{\sum_{i=1}^{N}(A_t - \bar{A})\,(F_t - \bar{F})}{\sqrt{\sum_{t=1}^{N}(A_t - \bar{A})^2 \cdot \sum_{t=1}^{N}(F_t - \bar{F})^2}}$$

Where $A_t$ and $F_t$ are the actual and predicted values, *N* is number of training and testing data set, $\bar{A} = \frac{1}{N}\sum_{t=1}^{N} A_t$ and $\bar{F} = \frac{1}{N}\sum_{t=1}^{N} F_t$ are the average values of $A_t$ and $F_t$ over the training or testing data.

And $\sqrt{R^2}$ = Absolute value of $r$

In general, the formula of $r$ is used to investigate the strength of the linear relationship between *X* and *Y* (Pardoe, 2006). The larger value of $r$ indicates the better performance of the model understudied (Wang and Elhag, 2008). On the other hand, $R^2$ is used to measure the goodness of fit of the model.

## RESULTS AND DISCUSSION

In this study, Linear Regression models were developed using the training data set and used to predict the Melt flow index (MFI), Melting point (MP) and Density of degradable plastics based on two input parameters PE and OPB using the testing data set. The Linear Regression equations for the models developed using the training data set are:

MFI = -0.0066 PE – 0.0080 OPB + 0.7592

MP = 0.0986 PE + 0.1016 OPB + 119.6384

Density = 0.0027 PE + 0.0029 OPB +0.6469

The testing data set was then used to predict the MFI, MP and Density of degradable plastics using the three derived equations. The predicted results are then compared to the actual lab results. The discrepancies between the predicted results and the actual lab results are used to calculate the correlation coefficient (*r*) and the coefficient of determination ($R^2$). The correlation coefficient (*r*) and the coefficient of determination ($R^2$) values are tabulated in Table 1. The result reveals that, the prediction accuracy of Melt flow index is 93%, 71% for the prediction accuracy of Melting point and 24% for the prediction accuracy of Density respectively. The results indicate that the use of Linear Regression model for predicting the physical properties of degradable plastics is highly feasible.

**Table 2. Correlation Coefficient (*r*) and Coefficient of Determination ($R^2$).**

| Physical Property of Degradable Plastic Predicted | *r* | $R^2$ |
|---|---|---|
| Melt Flow Index | 0.9656 | 0.9324 |
| Melting Point | 0.8422 | 0.7093 |
| Density | 0.4946 | 0.2446 |

## CONCLUSION AND RECOMMENDATION

Linear Regression model has been developed to study the physical properties of degradable plastics (Melt flow index, Melting point and Density) with different percentages of PE and OPB. The findings from this study indicate the feasibility of Linear Regression model as alternative tool to be used in predicting the physical properties of degradable plastics. It is also recommended that further research be carried out on other aspects of degradable plastics using others mathematical applications. The findings of this research will certainly benefit the industry related to biopolymer products.

**ACKNOWLEDGEMENT**

This research is supported by the Malaysia Ministry of Science, Technology and Innovation for the Grant No. 600-RMI/ST/FRGS 5/3/Fst (79/2010).

**REFERENCES**

Badri. K., Othman. Z., & Mohd Razali. I. (2005). Mechanical properties of polyurethane composites from oil palm resources. *Iranian polymer*. 441-448.

Facca. A. G., Kortschot. M. T., & Yan. N. (2006). Predicting the elastic modulus of natural fibre reinforced thermoplastics. *Composites: Part A* 37. 1660-1671.

Facca. A. G., Kortschot. M. T., & Yan. N. (2007). Predicting the tensile strength of natural fibre reinforced thermoplastics. *Composite Science and Technology*, 67, 2454- 2466.

Lee. S. Y., Hanna. M. A., Jones. D. D. (2008). An adaptive neuro-fuzzy inference system for modeling mechanical properties of tapioca starch-poly (lactic acid) nanocomposite foams. *Starch/ Starke*. 159-164.

Mangal. R., Saxena. N. S., Sreekala. M. S., Thomas. S., & Singh. K. (2003). Thermal properties of pineapple leaf fiber reinforced composites. *Material Sciences and Engineering*. 281-285.

Montgomery. D. C., Peck. E. A. (1991). Introduction to linear regression analysis. Canada, John Wiley & Sons, Inc.

Muc. A., & Kedziora. P. (2001). A fuzzy set analysis for a fracture and fatigue damage response of composite materials. *Composite Structures.*283-287.

Mungara. P., Chang. T., Zhu. J., & Jane. J. (2002). Processing and physical properties of plastics made from soy protein polyester blends. *Journal of Polymers and Environment* , 31-37.

Pardoe. I. (2006). Applied regression modeling. New Jersey, John Wiley & Sons, Inc.

Sin. L. T., Rahman. W.A. W. A., Rahmat. A., Morad. N. A., & Salleh. M. S. N. (2010). A study of specific heat capacity functions of polyvinyl alcohol- cassava starch blends. *Int J Thermophys.* 525-534.

Sreekala. M. S., & Thomas. S. (2003). Effect of fiber surface modification on water- sorption characteristics of oil palm fibres. *Composites Sciences and Technology.* 861-869.

Wang. Y. M., & Elhag. T. M. S. (2008). An adaptive neuro- fuzzy inference system for bridge risk assessment. *Expert Systems with Applications*.3099- 3106.

Zaini. M. J., Ismail. Z., Fuad. M. Y. A., & Mustafah. J. (1993). Application of oil palm wood flour as fillers in polypropylene. *Polymer Journal.* 637-642.