

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

This chapter presents the introduction to this thesis. It begins by describing the overall research background followed by a brief history of the oil palm industry in Malaysia. Research objectives, the scope of this study, research framework and discussion on the research contribution are also given. Finally, the brief of each chapter is outlined.

1.2 RESEARCH BACKGROUND

In the oil palm industry, modelling plays an important role in understanding various issues. It is used in decision making and the advance in computer technology has created new opportunity for the study of modelling. Modelling can be categorised into statistical and heuristic modelling. Statistical modelling is defined as the analysis of the relationship between multiple measurements made on groups of subjects or objects, and the model usually contains systematic elements and random effects. As a mathematical aspect, statistical modelling can be defined as a set of probability distributions on the sample space. Modelling involves the appropriate application of statistical analysis techniques with certain assumptions on hypothesis testing, data interpretation, and applicable conclusion.

Statistical analysis requires careful selection of analytical techniques, verification of assumptions and verification of the data. In conducting statistical

analysis, it is normal to begin with the descriptive statistics, graphs, and relationship plots of the data to evaluate the legitimacy of the data, identify possible outliers and assumption violations, and form preliminary ideas on variable relationships for modelling.

The heuristic approach is defined as pertaining to the use of general knowledge based on experimentation, evaluating possible answers or solutions, or trial-and-error methods relating to solving problems by experience rather than theory. Heuristic is also the problem-solving procedure that involves conceiving a hypothetical answer to a problem at the outset of an inquiry for purposes of giving guidance or direction to the inquiry. One of the heuristic approaches is the neural network model, which is based on the rules of thumb and widely used in various fields. A very important feature of neural networks is their adaptive nature where 'learning by example' replaces 'programming' in solving problems. This feature renders these computational models very appealing in application domains, where one has little or incomplete understanding of the problem to be solved, but where training data or examples are available.

Neural networks are viable and very important computational models for a wide variety of problems. These include pattern classification, function approximation, image processing, clustering, forecasting and prediction. It is common practice to use the trial and error method to find a suitable neural networks architecture for a given problem. A number of neural networks are successfully used and reported in literature (Zuhaimy and Azme, 2001; Zuhaimy and Azme, 2002). Neural network also has been applied in various fields such as in environmental (Corne *et al.*, 1998; Hsieh and Tang, 1998; Navone and Ceccatto, 1994), in economy and management (Boussabaine and Kaka, 1998; Franses and Homelen, 1998; Garcia and Gency, 2000; Indro *et al.*, 1999; Klein and Rossin, 1999b; Tkacz and Hu, 1999; Yao *et al.*, 2000) and in agronomy (Shearer *et al.*, 1994; Drummond *et al.*, 1995; Liu *et al.*, 2001; Kominakis *et al.*, 2002; Shrestha and Steward, 2002).

There are different types of the network are perceptron network, multiple layer perceptron, radial basis function network, Kohonen network, Hopfield network etc. However, the multiple layer perceptron is widely reported and used neural

networks in application. The most popular architecture, in the class of multiple layer perceptron, is the feedforward neural network.

The developments of models for agriculture are normally divided into three steps. The first step is to develop a preliminary model, which is inadequate. This preliminary model does not have to be a good model but it acts as a basis. This leads to further research, to develop a comprehensive model incorporating all the processes that appear to be important. Such a model is valuable for research, but far too complex for everyday use. To overcome this, a set of summary models is produced, each containing enough detail to answer limited questions. For example, there might be a summary model to predict the response to fertilisers on different soil types. Another model might be used to predict cyclic variation in yield. Modelling helps to make predictions more accurate. There is no doubt that modelling will maintain its importance in oil palm research as the problems set more complex and difficult. This study proposes the development of statistical model and neural network in modelling oil palm yield.

1.3 BRIEF HISTORY OF OIL PALM INDUSTRY IN MALAYSIA

Oil palm (*Eleais guineensis. Jacq.*), is a plant of African origin and is grown commercially in Africa. In the early 19th century the oil palm was brought into this country by the British. The oil palm was first planted in 1848 in Bogor-Indonesia and in Malaysia in 1870, at the same time rubber seeds were brought in (Hartley, 1977). Due to lower profitability of oil palm in comparison to rubber, the development of oil palm industry was rather slower. The first commercial planting of oil palm in Malaysia took place in 1917, six years after its systematic cultivation in Sumatera. The early planting was undertaken by European plantations, including Tannamaran Estate in Selangor and Oil Palm Malaya Limited. The 1960s and 1970s were marked by extensive development of oil palm undertaken largely by private

plantations and the Federal Land Development Authority (FELDA). In addition, a number of State Agriculture Development Corporations (SADC) became involved in oil palm cultivation after learning about its good prospects. The Rubber Industry Smallholders' Development Authority (RISDA) and the Federal Land Consolidation and Rehabilitation Authority (FELCRA) were also involved in cultivating abandoned and idle rubber and paddy areas with oil palm (Teoh, 2000).

From year 1975 to year 2000, the worldwide area planted with oil palm (*Elaeis guineensis* Jacq.) has increased by more than 150 percent. Most of this increase has taken place in Southeast Asia, with a spectacular production increase in Malaysia and Indonesia. The production of crude palm oil (CPO) in 2003 increased markedly, by 12.1 percent or 1.4 million tonnes to 13.35 million tonnes from 11.91 million tonnes in 2002 (Figure 1.1) (Teoh, 2000).

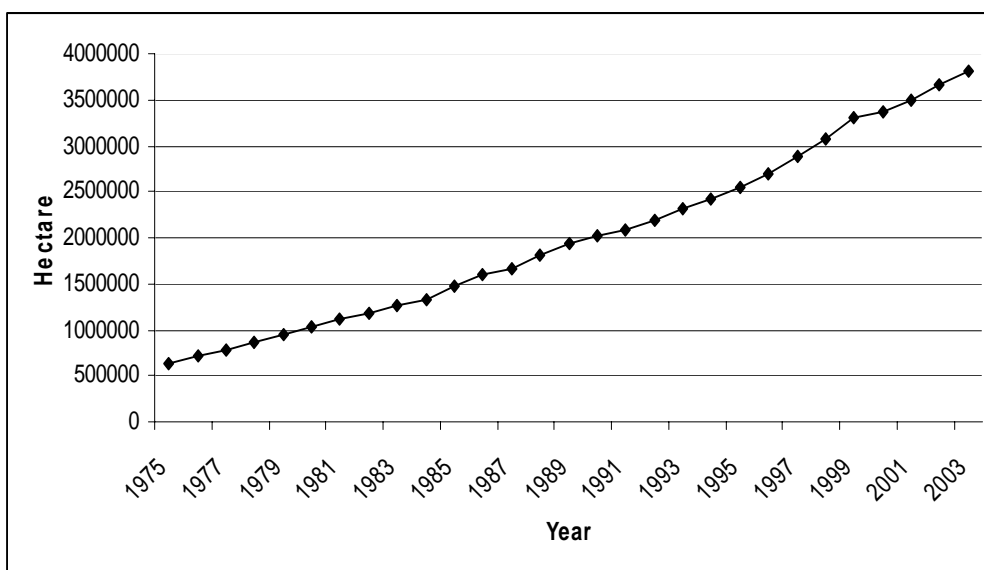


Figure 1.1: Oil palm planted area: 1975 – 2003 (hectares) including Peninsular Malaysia, Sabah and Sarawak

(Source: Department of Statistics, Malaysia: 1975-1989; MPOB: 1990-2003)

The production of crude palm kernel also rose substantially by 11.6 percent in to 1.6 million tonnes year 2003 from 1.47 million tonnes in year 2002. The increase was mainly attributed to the expansion in the matured area (Figure 1.2), favourable weather conditions and rainfall distribution as well as constant sunshine throughout the year. Exports of palm oil increased by 12.5 percent or 1.36 million tonnes to 12.25 million tonnes from 10.89 million tonnes in 2002 (Figure 1.3) (MPOB, 2003).

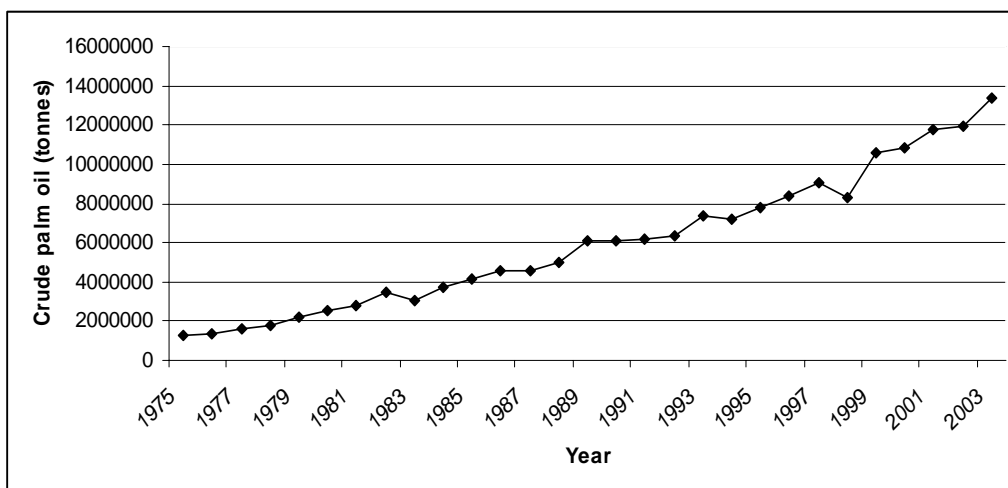


Figure 1.2: Annual production of crude palm oil (1975-2003) including Peninsular Malaysia, Sabah and Sarawak. (Source: Department of Statistics, Malaysia: 1975-1989; MPOB: 1990-2003)

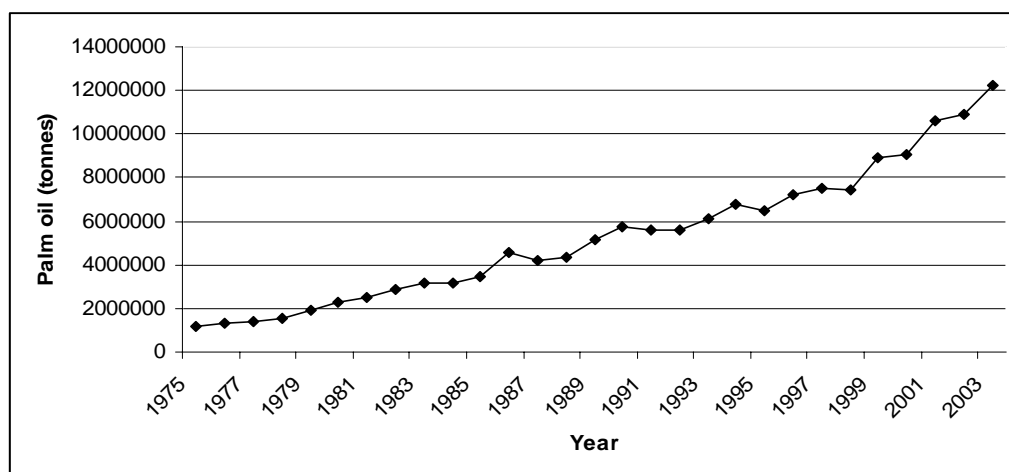


Figure 1.3: Annual export of palm oil: 1975 – 2003 in tonnes. (Source: MPOB)

Malaysia is the major producer and exporter of palm oil in the world (Teoh, 2000). Figure 1.4 shows Malaysian production of palm oil compared to Indonesia and other countries from 1999 to 2003. It shows that Malaysia and Indonesia recorded an increase in production every year. While Figure 1.5 presents the world's major palm oil exporters of palm oil from year 1999 to 2003, it also indicates that Malaysia and Indonesia also recorded the higher volume. In 2003, the Malaysian palm oil exporting industry has increased by around 12.5 percent to 12,248 million tonnes, from 10,886 million tones the previous year. Indonesia only recorded a 7.07 percent increase over the same period. The development of the oil palm industry is growing at a fast rate and requires a lot of research. This study took the challenge to contribute our knowledge to the development of the oil palm industry.

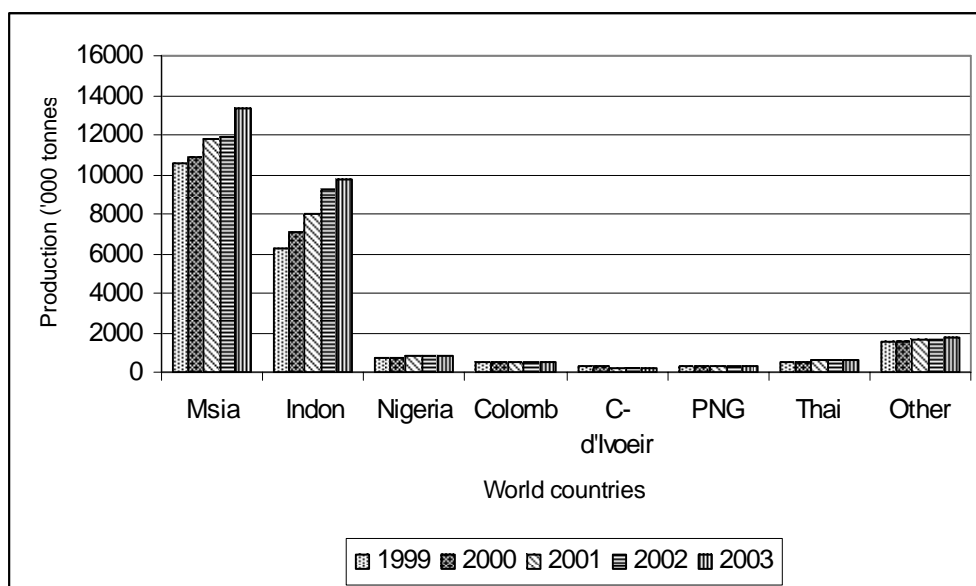


Figure 1.4: World major producers of palm oil ('000 tonnes)

Source: Oil World (December 12, 2003), Oil World Annual (1999-2003)

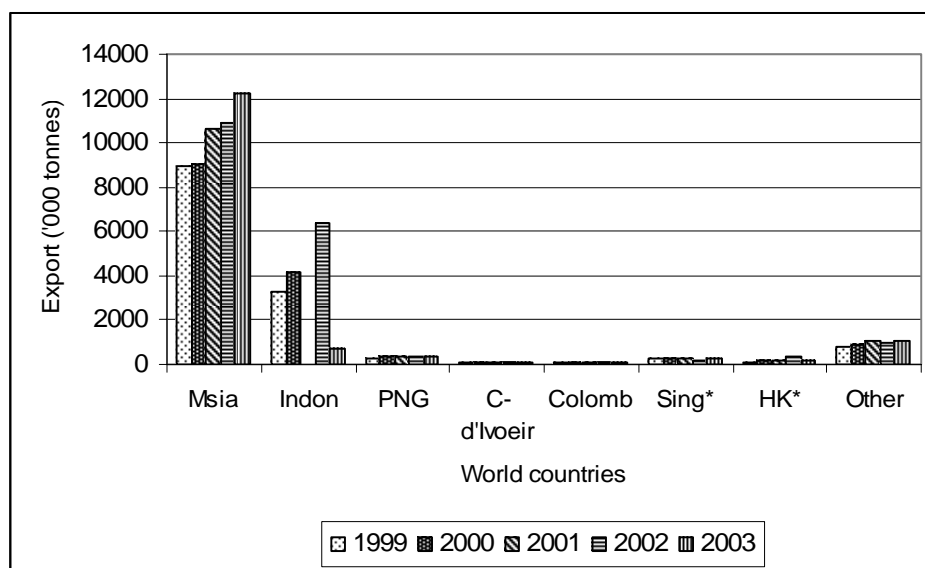


Figure 1.5: World major exporter of palm oil, including re-exporting country (*)

Source: Oil World (December 12, 2003), Oil World Annual (1999-2003)

1.4 PROBLEM DESCRIPTIONS

The problem in modelling oil palm yield growth is that it does not follow a linear model. It normally follows a nonlinear growth curve. In modelling a nonlinear curve, the complexity of the problem increases with the increase in the number of independent variables. The function of a growth curve has a sigmoid form, ideally its origin is at (0,0), a point of inflection occurring early in the adolescent stage and either approaching a maximum value, an asymptote or peaking and falling in the senescent stage (Philip, 1994). Normally, oil palm can be harvested after three years of planting. The oil palm yield will increase vigorously until the tenth year of planting. The yield will then increase at a low increment until the twenty-fifth year. From our exploratory study on modelling practices, little work has been reported on modelling the oil palm yield growth (Corley and Gray, 1976).

In most cases, researchers focused their study on the effect of environmental factors, such as evapotranspiration, moisture and rainfall to the oil palm growth. Chan *et al.* (2003) conducted a study on the effect of climate change to fresh fruit bunches (FFB) yield, and found that climate change has significantly affected oil palm yield. The most popular method used in the oil palm industry is multiple linear

regression. This model is used to investigate the causal effect of the independent variables to the dependent variable. The literature shows that the foliar nutrient composition can be used as an indicator to estimate the oil palm yield. Nevertheless the foliar nutrient composition is also dependent on several factors, such as climate, soil nutrients, fertilisers, pest and diseases, but little had been done on modelling these factors. This study explores the possibility of improving the model but in particular, in improving the level of accuracy it can produce. The proposed model should give smaller error values than previous model (Multiple Linear Regression, MLR).

The response surface analysis is the technique used to model the relationship between the response variable (Fresh Fruit Bunch yield, FFB) and treatment factors (fertilisers). The factor variables are sometimes called independent variables and are subject to the control by the experimenter. In particular, response surface analysis also emphasises on finding a particular treatment combination, which causes the maximum or minimum response. For example, in the oil palm industry there is a relationship between the response variable (oil palm yield) and the four fertiliser treatments, namely nitrogen (N), phosphorus (P), potassium (K) and magnesium (Mg). The expected yield can be described as a continuous function of the application level of fertiliser used. A continuous second-degree-function (N^2 , P^2 , K^2 or Mg^2) is often a sufficient description of the expected yield over the range of factor levels applied (Verdooren, 2003). If the fertiliser application rates are greater or smaller than the optimum application rate it may result in reduced yields. Fertilisers are wasted if the amount applied is more than the optimum rate. The advantage of this technique is that the effects of treatment combinations that have not been carried out in the experiment may still be estimated.

The use of response surface analysis is necessary to obtain the optimum level of fertiliser requirements. In response surface analysis, the eigenvalues will determine whether the solution gives a maximum, minimum or saddle point of the response curve. From our exploratory study on the use of response surface analysis, there is no solution if the stationary point is a saddle. This study will propose to use ridge analysis as an alternative solution to overcome the saddle point problem.

1.5 RESEARCH OBJECTIVES

Even though Malaysia is the largest producer of palm oil in the world, studies on modelling yields have been very limited. The modelling of Malaysian oil palm yield has been a recent phenomenon for decades. Literature reviews on research conducted in this field are confined to simple models. The oil palm industry is currently under going a structural change and is becoming more complex due to technological advances, agricultural management, product demand and planting areas (Teoh, 2000).

This research is an attempt to present a proper methodology for modelling oil palm yield. The model may then be used for estimating and managing the oil palm industry.

We further refine the objectives as follows:-

- To study current modelling and estimating practices in the oil palm industry.
- To explore and propose the best model for oil palm yield growth.
- To explore the use of neural network to model oil palm yield.
- To optimise fertiliser level which will generate optimum yield.

These objectives will be achieved by following the research framework as presented in Figure 1.6.

1.6 SCOPE OF THE STUDY

This section is divided into three subsections. The first section will discuss the scope of the data, followed by a discussion on the model scope, and finally the discussion on statistical testing deployed in this study.

1.6.1 Data Scope

For modelling oil palm yield growth data used in this study is secondary data taken from research done by Foong (1991; 1999). The research was conducted at Serting Hilir in Negeri Sembilan with relatively wet weather. The annual rainfall in this area is between 1600 mm to 1800 mm with two distinct droughts in January to March and June to August. The data used here is the average fresh fruit bunches (tonnes/hectare) from 1979 to year 1997.

The Malaysian Palm Oil Board (MPOB) provided us with a data set taken from several estates in Malaysia. The factors included in the data set were foliar composition, fertiliser treatments and FFB yield. The variables in foliar composition include percentage of nitrogen concentration N, percentage of phosphorus concentration P, percentage of potassium concentration K, percentage of calcium concentration Ca, and percentage of magnesium concentration Mg. The fertiliser treatments included N, P, K and Mg fertilisers, and they were measured in kg per palm per year, example 3.7 kg N fertilisers were needed for one palm per year. The foliar composition data was presented in the form of measured values while the fertiliser data in ordinal levels, from zero to three.

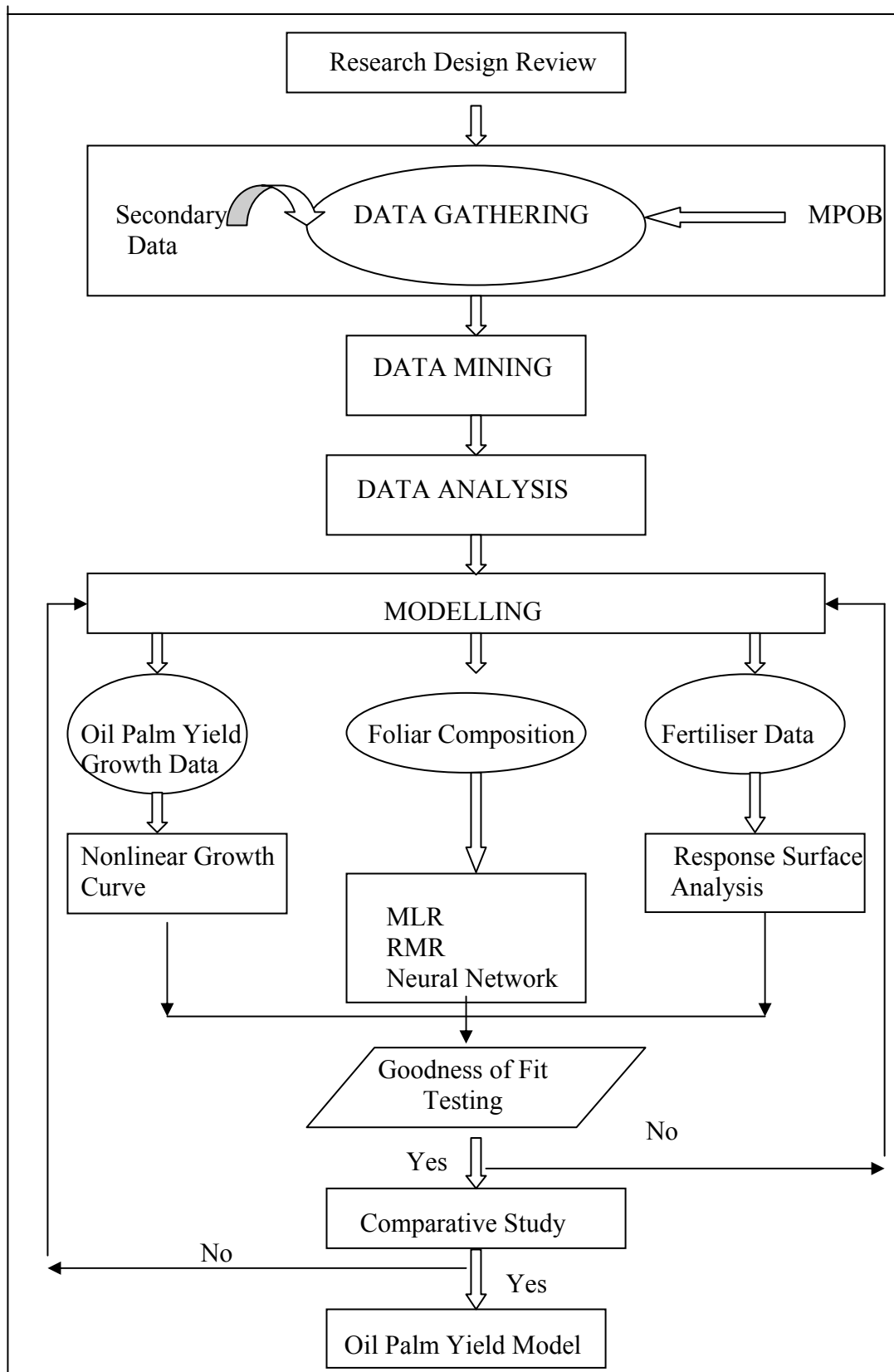


Figure 1.6: Summary of research framework and research methodology used in this study

1.6.2 Model Scope

This study will confine the scope of models, namely the nonlinear growth model (NLGM), multiple linear regression (MLR), robust M-regression (RMR), response surface analysis (RSA) and neural network (NN) models. The nonlinear growth model will be used to model the data of oil palm yield growth. Using foliar analysis data we employ the multiple linear regression and robust M-regression to estimate the oil palm yield. In the MLR model the independent variables are N, P, K, Ca and Mg concentration (or as we call it, major nutrient component, MNC) and the dependent variable is fresh fruit bunches (FFB) yield. Aside from MNC concentration, we also introduce the use of nutrient balance ratio (NBR), critical leaf phosphorus concentration (CLP), total leaf basis (TLB), deficiency of K (defK) and deficiency of Mg as independent variables in the second part in MLR. In MM regression we only consider N, P, K, Ca and Mg concentration as independent variables and FFB yield as the dependent variable.

We propose the use of the neural network to model oil palm yield. The discussion on the selection of neural network architecture and some statistical analysis will be given in Chapter 6. Chapter 7 will describe the use of response surface analysis to obtain the optimum fertiliser rate to produce an optimum FFB yield. Following this is a simple economic analysis to select the best combination of fertilisers input that generates the maximum profit.

1.6.3 Statistical Testing Scope

In this study we considered several statistical tests. They are the error model, sum of squares error (SSE), root mean squares error (RMSE), determination coefficient (R^2), coefficient of correlation (r), t-test, F test and chi-square test. The discrepancy between the predicted value from the model fitted, \hat{y}_i and actual value y_i is used to measure the model goodness of fit. The difference between the actual and the estimated value as known as the model error, and can be written as follows;

$$e_i = y_i - \hat{y}_i \quad i = 1, 2, \dots, n$$

where e_i is the model error in observation i . y_i is the actual observation i , and \hat{y}_i is the estimated value at i observation. If the model performance is ‘good’, the model error will be relatively small.

For the purposes of measuring the accuracy of model fitting, we consider the four measurements commonly used in any research on model fitting. Namely sum squares error, root mean squares error, determination coefficient R^2 and correlation coefficient. All formulas are given below;

$$(i) \quad \text{Sum Squares Error, SSE} = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad i = 1, 2, \dots, n$$

$$(ii) \quad \text{Mean Squares Error, MSE} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}, \quad i = 1, 2, \dots, n$$

$$(iii) \quad \text{Root Mean Squares Error, RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad i = 1, 2, \dots, n$$

$$(iv) \quad \text{Determination of coefficient, } R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad i = 1, 2, \dots, n$$

and

$$(v) \quad \text{Coefficient of correlation, } r = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\text{Var}(x)\text{Var}(y)}}}, \quad i = 1, 2, \dots, n$$

where y_i observed value, \hat{y} predicted value, n number of observation, \bar{x} and \bar{y} are the mean of x_i observation and y_i observation, respectively, $\text{var}(x)$ is the variance of X and $\text{var}(y)$ is the variance of Y . SSE, MSE and RMSE are used to measure the model accuracy. The R^2 value is a measure of how well the explanatory variables explain the response variable. Correlation coefficient is used to identify the strength of the relationship between any two variables.

In the case of more than two samples, one-way analysis of variance (anova) can be used to test the difference between the groups using F-test. The anova F-test is

calculated by dividing an estimate of the variability between the groups by the variability within the groups;

$$F = \frac{\text{Variance between groups}}{\text{Variance within groups}}$$

A high value of F , therefore, is evidence against the null hypothesis of equality of all population means. If the test shows the mean difference between groups to be statistically significant, the Multiple Duncan test can be used to examine which groups are different to each other (Montgomery, 1991). Another alternative to one-way analysis of variance is the Chi-square test, which is a nonparametric test which can be used when assumption of normality is not needed.

The model performance will be measured using sum squares error, mean squares error, mean absolute, root mean squares error, mean absolute percentage error, coefficient of determination and coefficient of correlation.

1.7 DATA GATHERING

The Malaysian Palm Oil Board (MPOB) provided data from the MPOB database of oil palm fertiliser treatments, which have been carried out from fourteen oil palm estates. All the data from each estate has been collected, recorded and compiled by MPOB researchers in the Research Database Center. All treatments were based on a factorial design with at least three levels of N, P and K fertiliser rates. Although different types of fertiliser were used in the treatments, the rates quoted in the final analysis will be equalized to the amounts of ammonium sulphate (AS), muriate of potash (KCI), Christmas Island Rock Phosphate (CIRP) and kieserite (Kies). Cumulative yields obtained over a period of two to five years in each trial were analyzed. The data of this study is experimental basic and was collected for a certain period of time and differs for each experiment. We study fourteen experimental stations (including Peninsular Malaysia and East Malaysia), seven stations in inland areas and seven stations in coastal areas. Appendix A presents the background of the experimental stations including age of oil palm, soil type and the location of the station.

Fresh fruit bunches (FFB) yield data used in this study was measured in tonnes per hectare per year or the average of FFB yield in one year. Foliar analysis was only done once a year and the samples are taken either on March or July every year. For example, if this year foliar analysis conducted in July, the next sample also conducted in July next year, and so on. The type of FFB yield data and foliar analysis data is continuous, and a fertiliser input is in coded form (0, 1, 2, and 3). If recode data is needed, the coded value will be recoded to the exact value (Appendix B). The detail of the leaf analysis procedure is presented in section 1.8.

1.8 LEAF ANALYSIS

The best method of determining the kind and amount of fertiliser to apply to fruit trees is by leaf analyses. It effectively measures macro and micronutrients and indicates the need for changes in fertiliser programs (Cline, 1997). Leaf analyses integrate all the factors that might influence nutrient availability and uptake. The essentials of macronutrients to oil palm tree were listed in Appendix C. However, leaf analysis indicates the nutritional status of the crop at the time of sampling (Pushparajah, 1994). It also shows the balance between nutrients for example, magnesium (Mg) deficiency may be the result of a lack of Mg in the soil or due to antagonistic effect with excessive K levels or both of these conditions. It also shows hidden or incipient deficiencies. Adding N, for example, when K is low may result in a K deficiency because the increased growth requires more K (Fairhurst and Mutert, 1999).

The leaf analysis was conducted to determine the nutritional status of leaflets from frond 9 on immature palms and frond 17 on mature palms (Corley, 1976). This is conducted to assist the preparation of annual fertiliser programmes. In each nominated leaf sampling the appropriate frond is correctly sampled for each leaf sampling unit (LSU). Frond 17 is sampled from the labeled reference LSU palm in some or all fields in a LSU and prepared for analysis. Cleanliness is essential at all

stages to prevent sample contamination and sampling time between 6.30 am and 12.00 noon.

A frond 17 is identified by counting from the first fully open frond in the center of the crown (frond 1) (and moved three steps downward (frond 1, 9, 17) with the same stack) and removed with a sickle. The frond is cut into approximately three equal sections (to get the average of the nutrient concentration). The top and base sections are discarded and placed in the frond stack. Twelve leaflets are selected and removed from each frond. Six leaflets are cut from each site at the mid-point of the frond section (Corley, 1976). Ensure that the 12 leaflets comprise of three from the upper rank and three from the lower rank from each side of the rachis. The leaflets samples from each field (or smaller area if required) are put together in a large labeled plastic bag. About 500 leaflets are collected from each field of 30 hectare.

The samples are then sent to the estate laboratory or sample preparation room for further preparation. The leaflets are bundled and trimmed to retain the 20-30 cm mid-section; it is not necessary to wash the leaves. The mid-rib of each leaflet's section is removed and discarded. The remaining parts of the leaflet's (lamina) are then cut into small pieces 2 cm long and placed on aluminium trays to be dried. The leaflets are dried in a fan-assisted oven for 48 hours (65⁰C) or 24 hours (105⁰C). The leaf N concentration will be reduced if the temperature exceeds 105⁰C.

After drying, the leaflets are placed in a labeled plastic bag. Half of the sample retained as a backup for future reference (stored in a cool, dry place) while the other is submitted for analysis. The LSU sample results from the laboratory are then formatted as a spreadsheet and the variability is calculated. Leaf samples are analyzed for N, P, K, Ca and Mg. Other nutrients may be included for palms planted on particular soil types.

Leaf sampling is carried out once each year. Sampling is frequently conducted to examine specific areas or to investigate particular nutritional problems. Leaf sampling should be done at the same time each year and not during wet or very dry periods. Complete the sampling procedure in the shortest possible time.

Because of the synergism between nitrogen (N) and phosphorus (P) uptake, leaf concentration must be assessed in ratio to leaf N concentration (Ollagnier and Ochs, 1981). This is due to the constant ratio between N and P in protein compounds found in plant tissue (Fairhurst and Mutert, 1999). A critical curve has been developed where CLC_p is defined as;

$$\text{Critical Leaf P concentration, } CLC_p = 0.0487 \times \text{Leaf N concentration} + 0.039$$

A different approach to determine whether potassium (K) and magnesium (Mg) are deficient taking into account the relative concentrations of the leaf cations K, Mg and calcium (Ca). First, the total amount of bases in leaf (TLB) is calculated and K and Mg are assessed as a percentage of TLB (Foster 1999). TLB can be derived from equation below;

$$TLB \text{ (cmol/kg)} = (\% \text{ leaf K}/39.1 + \% \text{ leaf Mg}/12.14 + \% \text{ leaf Ca}/20.04) \times 1000$$

roughly, K and Mg deficiency can then be assessed individually, based on their percentage of TLB. The deficiency of K and Mg can then be obtained

as $\left(\frac{X}{TLB}\right) \times 100$, where X is partial to TLB of K and Mg. The K and Mg deficiency

can be rated into three categories; If the value is below than 25 the rating is deficient, a low rating is between 25 to 30 and a rating more than 30 is considered sufficient. Nutrient Balance Ratio, NBR is defined as the ratio between the foliar nutrient composition and another foliar nutrient composition. For example, the NBR between N and K in foliar, is defined as the ratio between N and K concentration. The range of the NBR values for oil palm presented in Table 1.1.

Table 1.1: The optimum value of nutrient balance ratio (NBR) for foliar analysis

Nutrient ratio	NBR
N/K	2.50 – 3.00
N/Mg	14.00– 18.00
N/P	11.00 – 17.00
N/C	4.00 – 9.00
K/Mg	4.00 – 10.00
K/Ca	2.00 – 5.00
Mg/Ca	0.25 – 0.55

1.9 RESEARCH IMPORTANCE

The nonlinear growth models are used in modelling the nonlinear phenomenon. Since the nonlinear growth model has not yet been explored in oil palm industry (Foong, 1999 and Ahmad Tarmizi *et al.*, 2004), we proposed the use of the nonlinear growth model in the oil palm yield growth study. Here we will provide some mathematical basis in parameter estimation for modelling oil palm yield growth. Then from the results and analysis we can study the biological process of oil palm yield growth.

Multiple linear regression can be used to find the relationship between the dependent variable and the independent variable. There can be more than one independent variable, which allows for the additional relevance of the independent variable to the model. In these sense, multiple linear regression is rather flexible. Our study emphasizes the proposed new independent variables into the model, an area yet to be explored by researchers. In real life, nothing seems to work linearly all the time. Data are sometime inclusive of outlier or unusual observation. We proposed the use of multiple robust regression to overcome the negative impact of outlier to the model's development.

To improve the models, there are various new heuristic methods suggested in this literature. We explore the flexibility of the neural network to improve the estimated performance and the model's accuracy. Previous studies in oil palm stopped when the stationary point was saddle (Ahmad Tarmizi, 1986). This caused did not make allowances for the possibility an incomplete inference from the model than produce inefficient decision. It also caused difficulties in implementing improvements in practice outcomes. This study proposes the use of ridge analysis when the stationary point is saddle to improve data analysis.

1.10 RESEARCH CONTRIBUTION

There are many contributions in this study. Since it is an area of high importance for the sustainability of the oil palm industry, the contributions can be categorized as follows;

- Identifying several nonlinear growth models for oil palm yield growth.
- The investigation on the relationship between foliar nutrient composition and yield was conducted using MLR and RMR. A practical model and procedure were developed for this purpose.
- Development of neural networks model to predict the oil palm yield and NN results more reliable compared with the MLR and RMR models.
- This study proposes statistical testing to evaluate the factors that influence NN performance. The findings indicated that the combination activation and number of hidden nodes have a significant effect on the NN performance. However, the learning rate, momentum term and number of runs do not give any effect on the NN performance.
- This study investigates the effects of outliers on NN performance. The findings show that percentage-outliers and magnitude-outliers significantly affect the NN performance.
- The response surface analysis when combined with the ridge analysis was used to obtain the optimum level of foliar nutrient composition and fertiliser input to produce optimum oil palm yield.

Several of the contribution demonstrated above has been published in various form as described in Appendix D

1.11 THESIS ORGANISATION

This thesis contains eight chapters. Chapter 1 is the introduction. This chapter gives an introduction to the problem's description, research objectives, research scopes, research importance, research data and a brief description on the usage of the data in this research.

Chapter 2 is the Literature Review. This chapter contains a discussion on the current and past research on oil palm yield. Here we present the application of neural network modelling in several fields, such as economic, management and agronomy. A summary is included at the end of the chapter.

Four main models used in the thesis are explained in Chapter 3. It discusses the statistical methods such as nonlinear growth models, multiple linear regression, response surface analysis and the neural networks model. This chapter also proposes the research framework

In Chapter 4 the use of the nonlinear growth curve to model the oil palm yield growth is considered. Twelve nonlinear growth models are presented and the partial derivative for each models are provided. Comparisons among the model is done and given at the end of the chapter.

Chapter 5 discusses the development of multiple linear regression and robust M-regression to investigate the relationship between fresh fruit bunch and the nutrient foliar composition. The use of nutrient balance ratio, deficiency of magnesium, deficiency of potassium and critical leaf phosphorus as independent variables are proposed in this chapter. The numerical results from both methods are presented and compared in terms of modelling performance.

Chapter 6 presents the development of neural network to oil palm yield modelling. The experimental design is conducted to investigate the effect of the number of hidden nodes, the number of runs, momentum terms learning rate and outliers data to the NN performance. The results and conclusion of model selection have been carried out. The results from multiple regression analysis and neural network model are compared in terms of goodness of fit and model accuracy.

Numerical results of the foliar nutrient composition and fertiliser treatments performed by response surface analysis are reported in Chapter 7. The use of ridge analysis is discussed to overcome the 'saddle point' problem at the stationary point. This chapter ends with a simple economics analysis to generate the optimum fertilisers level in order to maximise the profit.

Chapter 8 concludes the relevant and important findings from this research. Recommendations on areas related to the findings and possible directions for future research are presented.