# CHAPTER III
# METHODOLOGY

## 3.1 Data Preparation

It is essential to select the effective inputs to develop an ANN model. However, availability of data is another major factor for choosing the input parameters since ANN demands large volume of data to be used for training and cross-validation in order to solve complex, nonlinear problems accurately. In this study, PVT data are collected from available publications. Data collected are checked based on following criteria. Redundant data points are removed. Some published correlations are used to filter some data sets for more quality of the data. The data point resulted with errors over 15% by the majority of the filtering correlations is removed (Mohammadpoor et al., 2010). Then the selected data are randomly divided into data sets for developing and testing ANNs and correlations as shown in Figure 3.1.
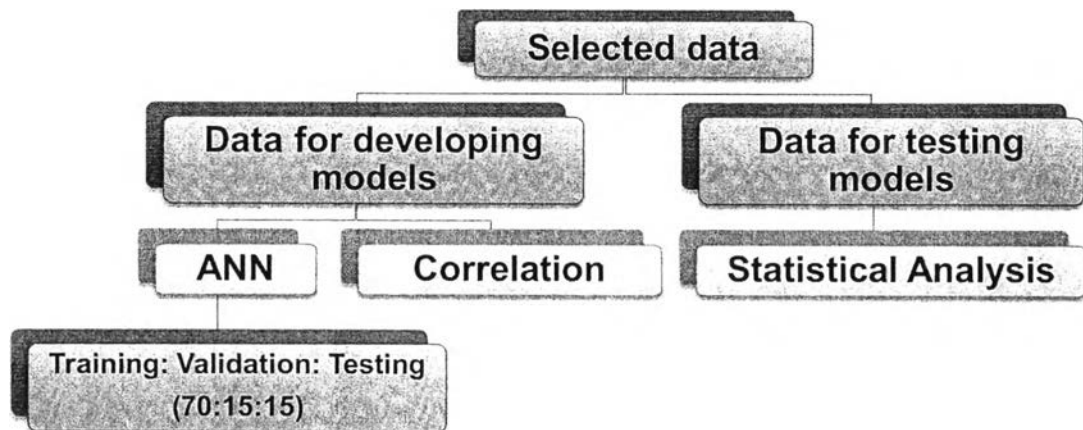


**Figure 3.1** Classification of the selected data.

## 3.2 Developing ANNs and Correlations

The prepared data sets for developing model are used for developing ANNs using neural network toolbox (nntool), which is graphical user interface (GUI) embedded in Matlab software as depicted in Figure 3.2. Moreover, the data sets for developing models are used for developing correlations using nonlinear regression technique from Minitab software. Figure 3.3 shows nonlinear regression tool in Minitab. Finally, the best correlations and ANNs resulted from numerous trials are selected.
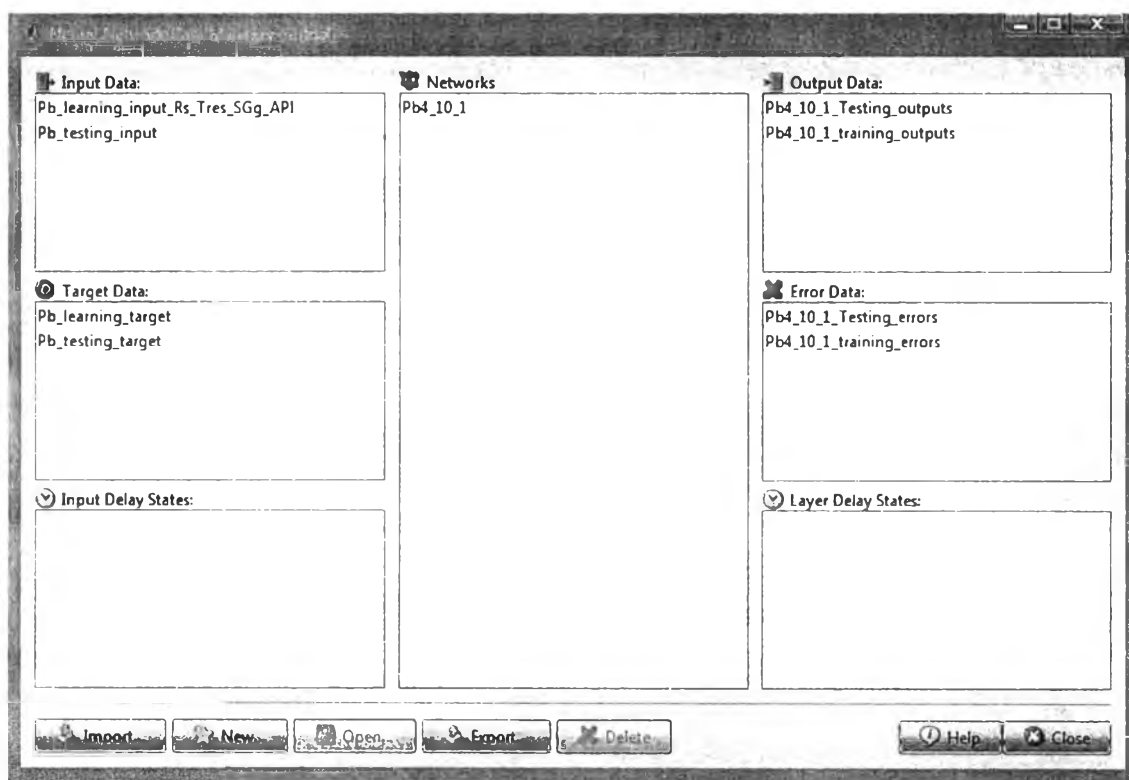


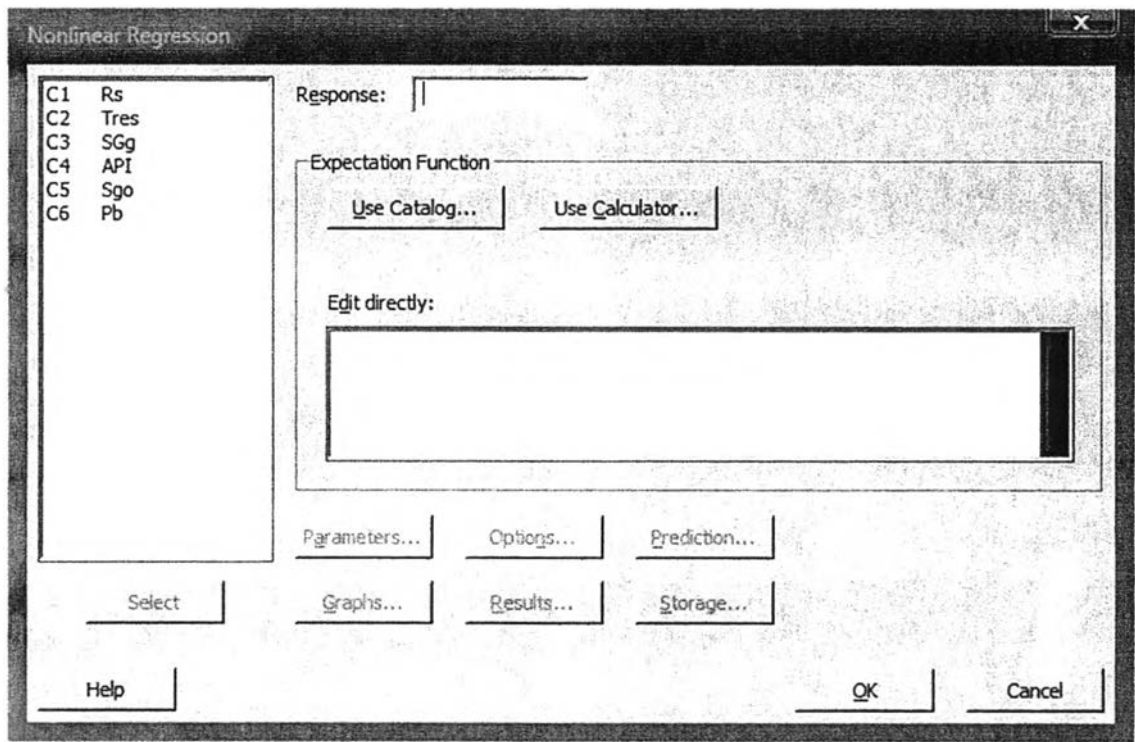**Figure 3.2** Neural network toolbox (nntool) embedded in Matlab.

**Figure 3.3** Nonlinear regression tool in Minitab.

## 3.3 Statistical Analysis

The accuracy of the developed models depends on the data used in developing process. The developed models have to be rechecked with data sets for testing, which have not been used in developing the models. As a consequence of testing the developed ANNs and correlations, the statistical parameters, such as minimum error ($Er_{min}$), maximum error ($Er_{max}$), average absolute error ($AEr_{avg}$), and coefficient of determination ($R^2$) are determined and compared with results from some published correlations (Appendix A).