CHAPTER 4

MODELING TECHNIQUES

4.1 GENERAL

Process model relates the input variables to the response variable of the process and thereby it is possible to predict the overall response of the process. The relationship between input and response variables could be found out by developing an empirical model, derived from experimental data that expresses the relationship between input response variables. Once the model is formed, the values of the actual process variable are substituted in the equation to predict the response of the process and measured on a continuous scale. A large variety of modeling techniques is available and the selection of a particular technique depends on constraints of the experiments. Furthermore, each technique has its own demands regarding the time and effort needed to apply the technique and interpret the results, consequently, each technique has its own strengths and limitations regarding the type of insight it can provide. The accuracy of prediction depends not only on the process conditions, but also on the choice of the modeling technique. A brief review is presented for the modeling techniques, which is then followed by details of proposed modeling capabilities and issues.

The experimental approach includes conducting a number of bending experiments and creating bending tables for general usage. Eventhough, the accuracy depends on the precision of the measurements before and after the bending process, this approach can be very accurate.
However, testing can be very time consuming and costly at the beginning. The classical experimental investigation adapts one factor at a time strategy and it fails to consider any possible interaction between the parameters.

Several analytical approaches are available in the earlier studies. Most of those methods also make simplifying assumptions such as plain strain deformation, rigid-plastic material models, isotropic materials, rigid tooling etc. The major difficulty with the analytical method is due to lack of understanding of stress distribution through the sheet, which limits the analytical approach to simple geometries and simple deformation. But the analytical model leads to a simple expression that predicts the response.

Finite element method (FEM) is usually computationally expensive. Furthermore, with the involvement of predicting the response, iterative solver usually has to use explicit-implicit transition, which can be challenging and losing its practicality. Furthermore, shortcomings in the numerical modeling arise because current finite element based simulation methods lack the resolution and smoothness to effectively capture the mechanics of the bending process.

Much of the research in engineering, science and industry is empirical and makes extensive use of experimentation. Statistical methods can greatly increase the efficiency of these experiments and often strengthen the conclusion so obtained. A number of statistical techniques is used to predict responses and the multiple regression analysis (MRA) and response surface methodology (RSM) are its offshoots.

MRA is widely used for various types of statistical analysis. Stepwise regression removes and adds variables to the regression model for the purpose of identifying a useful subset of the predictors. Best subsets regression identifies the best fitting regression models that can be constructed
with the predictor variables. Response surface methodology is a collection of mathematical and statistical techniques for empirical model building. Artificial neural networks have gained prominence recently among researchers of manufacturing systems. The ANN prediction is able to effectively capture the non-linear relationships and interaction of process parameters in the process. ANN is fault tolerant and robust, amenable to parallel implementation and faster than conventional computing.

The problem of finding the equation of curve which represents the exact relationship between two variables is called curve fitting. The evaluation of empirical constants occurring in the equation is by three standard methods, namely, group averages, moments and least squares. The method of group averages will not give a unique answer for a given problem, as division of data points into groups can be done in different ways. The method of moments will give a unique answer, but it cannot be applied where the values of the independent variable are not equally spaced. The method of least squares is the best, as the other methods assume that the data points are almost equally distributed on both sides of the best fitting curve, whereas the method of least squares assumes that the data points lie close to the best fitting curve (Veerarajan and Ramachandran 2004). Correlation is a statistical technique which measures and analyse the degree or extent to which two or more variables fluctuate with reference to one another. Scatter diagram is a method of diagrammatic representation of a bi-variable distribution for ascertaining the nature of correlation between the variables. It is very useful in identifying the relation between two variables.

This thesis is concerned with curve fitting techniques and three modeling techniques namely multiple regression analysis (MRA) based on Taguchi design of experiment, response surface methodology (RSM) based on central composite method, artificial neural network (ANN) based on back
propagation method and their theoretical aspects and they are depicted in the following sections.

4.2 FITTING CURVE EQUATIONS

In order to obtain equation of the best fitting of curve, five kinds of models are considered in this investigation (Veerarajan and Ramachandran 2004). The straight-line behavior suggests a power law relationship between response and bend angle factor. The first model is expressed as

\[ S_{T1} = aB^c \]  

(4.1)

where, \( S_{T1} \) is the response, \( B \) is the new bend angle factor and ‘\( a \)’ and ‘\( c \)’ are empirically determined constants. The equation of the best fitting curve (equation 4.1) is not in linear form. The curve can be reduced to the linear form by the following transformations.

Taking logarithms to the base 10 on both sides of equation (4.1).

\[ \log_{10} S_{T1} = \log_{10} a + c \log_{10} B \]  

(4.2)

The transformations are \( X = \log_{10} B \), \( Y = \log_{10} S_{T1} \)

\[ Y = A + cX \]  

(4.3)

where \( A = \log_{10} a \)

As per the least square method, the normal equations are

\[ NA + c \sum x = \sum y \]  

(4.4)
\[ A \sum x + c \sum x^2 = \sum xy \quad (4.5) \]

\[ a = 10^A \quad (4.6) \]

The equations (4.2) and (4.3) can be solved by Cramer’s rule (using determinants). According to the power law equation, simulations can be made to evaluate relation between response and bend angle.

The second model is expressed as

\[ S_{T2} = ac^B \quad (4.7) \]

where \( S_{T2} \) is the response, \( B \) is the new bend angle factor and ‘a’ and ‘c’ are empirically determined constants. Equation (4.7) is converted into a linear relation by taking logarithms to the base 10 on both sides and putting

\[ Y = \log_{10} S_{T2} \]

\[ \log_{10} S_{T2} = \log_{10} a + B \log_{10} c \quad (4.8) \]

The third model is defined as

\[ S_{T3} = ae^{cB} \quad (4.9) \]

Since the ‘a’ is taken as constant, equation (4.9) shows the relationship between the response and the exponential power. Equation (4.9) is not linear and converts it into a linear relation by taking logarithms to the base ‘e’ on both sides \( Y = \log_e S_{T3} \)

\[ \log_e S_{T3} = \log_e a + cB \quad (4.10) \]
The forth model is expressed as

\[ S_{T4} = aB^2 + bB + c \]  

(4.11)

where \( S_{T4} \) is the response, \( B \) is the new bend angle factor and ‘a’, ‘b’ and ‘c’ are empirically determined constants. As per the method of least squares, the normal equations that give the values of a, b and c in the required equation are

\[ Nc + b \sum B + a \sum B^2 = \sum S_{T4} \]  

(4.12)

\[ c \sum B + b \sum B^2 + a \sum B^3 = \sum BS_{T4} \]  

(4.13)

\[ c \sum B^2 + b \sum B^3 + a \sum B^4 = \sum B^2 S_{T4} \]  

(4.14)

The fifth model is expressed as

\[ S_{T5} = a \ln(B) - c \]  

(4.15)

where \( S_{T5} \) is the response, \( B \) is the new bend angle factor and ‘a’ and ‘c’ are empirically determined constants.

### 4.2.1 Standard Deviation and Dispersion Effect

The standard deviation indicates the extent of dispersion for a group as a whole. The standard deviation of responses is vital because it impacts the overall model efficiency. To find appropriateness of the model, a statistical analysis was proposed by Li (2002). In this method, the standard deviations were determined on the basis of calculated and measured responses by

\[ \langle \sigma \rangle = \sqrt{\frac{\sum_{i=1}^{N} \left( \Delta \theta_{\text{measured}}^{i} - \Delta \theta_{\text{calculated}}^{i} \right)^2}{N}} \]  

(4.16)
where $\langle \sigma \rangle$ - Standard deviation of response

$\Delta \theta^i_{\text{measured}}$ - Measured response

$\Delta \theta^i_{\text{calculated}}$ - Calculated response

$N$ - Number of trials

Inequality of variance at different levels needs transformations to stabilize the variance to improve inference on the location effects and check the levels affect variability. The standard deviation is used as response variable. In this method, the dispersion effects (Montgomery 2007) were determined by

$$y = -\ln\langle \sigma \rangle$$

*(4.17)*

### 4.3 MULTIPLE REGRESSION MODELS

Regression analysis is used to model and investigate the relationship between a response variable and predictors. It is a common tool for experimental analysis of data, when the response variable is continuous. It is also very helpful to present the results of many experiments in terms of an empirical model, that is, an equation derived from the data that expresses the relationship between the response and important design factors.

#### 4.3.1 Design of Experiments

An important aspect of MRA is the design of experiments. These strategies were originally developed for model fitting of physical experiments. The choice of the design of experiments can have a large influence on the accuracy of prediction. The design of experiment (DOE) is not a simple one step process but is actually a series of steps which must follow a certain sequence for the experiment to yield an improved understanding of process
performance. The major activities in the DOE process are the determination of the combination of factors and the levels which will provide the experimenter with the desired information. It requires powerful and efficient base for design of experiments. Generally, for a large number of variables, the number of experiments grows exponentially and becomes impractical. If the number of design variables becomes large, Taguchi’s design of experiments (orthogonal array) can be used. Taguchi DOE presents an alternative to factorial designs in the construction of models and reduces the number of experiments as compared to a full factorial design. The following assumptions are considered for multiple regression analysis

- The errors are normally distributed
- The mean of the errors is zero
- Errors have constant variance
- The model errors are independent

4.3.2 Design Structures

With multiple linear regression analysis, the basic relationship (Montgomery 2007) is written as

\[ Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \epsilon \]

(4.18)

where

- \( Y \) - Dependent variable (Predicted values)
- \( \beta_0 \) - Estimator of intercept, \( \beta_1, \beta_2, \ldots, \beta_k \) are linear terms,
- \( k \) - Number of factors
- \( x_i \) - \( i^{th} \) independent variable
- \( \epsilon \) - Random error
The various linear models of equation (4.18) in terms of the estimated response can be written as (4.19)

Linear terms can be expressed as

$$\hat{Y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i$$ \hspace{1cm} (4.20)

Linear-square terms are presented as

$$\hat{Y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \sum_{i=1}^{k} \hat{\beta}_{ii} x_i^2$$ \hspace{1cm} (4.21)

Linear-interaction terms are represented as

$$\hat{Y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \sum_{i=1, j=1, j \neq i}^{k} \hat{\beta}_{ij} x_i x_j$$ \hspace{1cm} (4.22)

Quadratic terms are given by

$$\hat{Y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \sum_{i=1}^{k} \hat{\beta}_{ii} x_i^2 + \sum_{i=1, j=1, j \neq i, j}^{k} \hat{\beta}_{ij} x_i x_j$$ \hspace{1cm} (4.23)

where $\hat{\beta}_{11}, \hat{\beta}_{12}, \ldots, \hat{\beta}_{kk}$ are the square terms, and $\hat{\beta}_{12}, \hat{\beta}_{13}, \ldots, \hat{\beta}_{k-1,k}$ are the second order interaction terms.

This equation allows identifying the given value of parameters and the expected value of the response in the process. Various models are compared to find the best fit and an appropriate model can be chosen since it has higher adjusted R-squared values. The selected mathematical model can be tested for its accuracy with analysis of variance (ANOVA) through F-test and R-squared values. The ANOVA is the statistical treatment most
commonly applied to the results of the experiments to determine the adequacy of the model. For modeling, MRA is simple, economical and relatively easy for use. One of the major drawbacks with the regression method is that these studies did not apply the factorial experimentation approach to design the experiments. Therefore, the data can be biased and factor interactions can not be clearly examined. It offered some more efficient and powerful alternative strategies.

4.4 RESPONSE SURFACE METHODOLOGY

Response surface methods are used to examine the relationship between one or more response variables and a set of quantitative experimental variables or factors. These methods are often employed to find the factor settings that produce the best response using controllable factors. Many response surface applications are sequential in nature that they require more than one stage of experimentation and analysis. Exploration of response surfaces is a very important aspect of experimental design. A response surface can be linear or non-linear and is typically classified as first order and second order methods. For non-linear response surfaces, interaction terms between input are considered. Fitting and analyzing response surfaces are greatly facilitated by the proper choice of an experimental design (Montgomery 2007). There are two main types of response designs namely, central composite designs and Box-Behnken designs. The central composite design is the most widely used experimental design for modeling a second order response surface.

4.4.1 Central Composite Design

CCD introduced by Box and Wilson is one of most popular response surface designs. It is designed to fit a second order model. It exists for spherical or cubical regions for \( k \geq 2 \), where \( k \) is the number of factors.
Generally, this design has a small to moderate sample size with excellent estimation properties. The CCD has three components: the $2^k$ factorial points, the $2k$ axial points and the $n_c$ centre runs. As the number of variables increase rapidly in an experiment, the number of runs needed to perform the full factorial design may become unrealistically large. Moreover, it is expensive and time consuming. Hence the fractional factorial designs were introduced. The optimal fractions are typically selected according to the resolution criteria. The axial points also called the star points are at distance of $\alpha$ from the design centre. They mainly contribute to the estimation of quadratic terms. The centre runs are located in the centre of the design space. They help estimating the quadratic terms as well as pure error. The choice of the number of the centre runs provides flexibility to get a better estimate of the pure error and better power of tests. Moreover, the choice of the number of center runs affects the distribution of the Scaled Prediction Variance (SPV). The choice of $\alpha$ is based on the region of interest. Depending on the value of $\alpha$, the design may achieve rotatability. Box and Hunter suggested that a second order response surface design should be rotatable. Rotatability is attained when a design has the same value of SPV for any two points on the same distance from the centre. This can be achieved for $\alpha = \sqrt[n_F]{n_c}$, where $n_F$ is the number of points used in the factorial portion of the design, for spherical regions, $\alpha = \sqrt{k}$ and $n_c = 3$ to 5 helps keep the distribution of the SPV. For cuboidal regions, $\alpha = 1$ is used and no further improvement is achieved for $n_c > 2$. The CCD of second order has been found to be the most efficient tool in RSM to establish the mathematical relation of the response surface using the smallest possible number of experiments without losing its accuracy (Montgomery 2007).
4.4.2 Selection of Levels

To reduce the total number of experiments and to obtain data uniformly from all the regions of the selected working area, a factorial design procedure has been adapted. The analysis was carried out using actual values. Three levels should always be selected if the vast experimental region has to be covered and to get accurate results. The selection of the limits of the process variables is followed by choosing of the number of levels within the limits. For the convenience of recording and processing experimental data, the upper level is coded as 2 and the lower level as –2 and intermediate levels can be calculated by using the following expression.

\[
Z_i = \frac{2Z - [Z_{\text{Max}} + Z_{\text{Min}}]}{Z_{\text{Max}} - Z_{\text{Min}}} \tag{4.24}
\]

where

- \( Z_i \) - Required coded value of a variable \( Z \)
- \( Z \) - Any value of the variable from \( Z_{\text{Max}} \) to \( Z_{\text{Min}} \)
- \( Z_{\text{Min}} \) - Lower limit of the variable, and
- \( Z_{\text{Max}} \) - Upper limit of the variable

4.4.3 Design Structures

For five independent variables \( x_1, x_2, x_3, x_4 \) and \( x_5 \), the response \( Y \) (Montgomery 2007) can be represented as a function of \( x_1, x_2, x_3, x_4 \) and \( x_5 \) as follows

\[
Y = f(x_1, x_2, x_3, x_4, x_5) + \varepsilon \tag{4.25}
\]

where \( \varepsilon \) represents an error component.
The second order RSM model is adequate and it can be represented by the following equation

\[ Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_i^2 x_i^2 + \sum_{i=1, j=1, j>i}^{k} \beta_i x_i x_j + \varepsilon \quad (4.26) \]

where \( \beta_i \) (i = 0, 1, \ldots, k) are coefficients that have to be estimated and \( \varepsilon \) represents a normally distributed random error that accounts for all source of variability.

The expected response is denoted by

\[ (Y - \varepsilon) = \hat{Y} \quad (4.27) \]

Then the surface represented by \( \hat{Y} = f(x_1, x_2, x_3, x_4, x_5) \) is termed as response surface.

The second order polynomial (regression) equation used to represent the response surface for \( k \) factors is given by

\[ \hat{Y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \sum_{i=1}^{k} \hat{\beta}_i^2 x_i^2 + \sum_{i=1, j=1, j>i}^{k} \hat{\beta}_i x_i x_j \quad (4.28) \]

where \( \hat{\beta}_0 \) is the estimator of intercept, \( \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k \) are linear terms, \( \hat{\beta}_{11}, \hat{\beta}_{22}, \ldots, \hat{\beta}_{kk} \) are the square terms, and \( \hat{\beta}_{12}, \hat{\beta}_{13}, \ldots, \hat{\beta}_{k-1,k} \), are the second order interaction terms.

The magnitude of the regression coefficient is a good indication of the significance of the parameters. The unknown coefficients \( \beta_i \) (i=0,1,2\ldots,k) are estimated by a linear multiple regression. The linear multiple regression model is rewritten in matrix form as
\[ Z = X\beta + \epsilon \] (4.29)

where

\[ Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_k \end{bmatrix} \]

\[ X = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{31} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & x_{32} & \cdots & x_{k2} \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1k} & x_{2k} & x_{3k} & \cdots & x_{kk} \end{bmatrix} \]

where \( \epsilon \) is the error vector. The unbiased ‘b’ of the coefficient vector \( \beta \) is obtained using the least square error method as

\[ b = \left( X^T X \right)^{-1} X^T Z \] (4.30)

By obtaining coefficients vector \( b \), the response surface is prepared. The tests of significance are performed using student’s ‘t’ test. According to this test, the calculated value of ‘t’ corresponding to a coefficient exceeds the standard tabulated value for the desired levels of probability, say 95%. Thus the coefficient becomes significant. After determining the significant coefficients, the final models are constructed by using only these coefficients. The estimated ‘t’ values for particular process parameter can be obtained from the following equation

\[ t_{\text{estimated}} = \frac{\text{Coefficient of a process parameter}}{\sigma_\beta} = \frac{\hat{\beta}}{\sigma_\beta} \] (4.31)
where \[ \sigma_{\beta}^2 = \frac{\text{Estimate of Error}}{N} \]

N - Number of trials

The significance of the coefficients is evaluated and insignificant coefficients are eliminated in the model. The tests of significance have been performed at different levels of significance using student’s ‘t’ - test. The second order model is widely used in RSM for several reasons. They are as follows

- The second order model is flexible. It can take on a wide variety of functional forms, so it will often work well as an approximation to the true response surface.
- It is easy to estimate the parameters in the second order model. The method of least squares can be used for this purpose.
- There is considerable practical experience indicating that the second order models work well in solving real response surface problems.
- The advantage of RSM is reasonably accurate and the interaction factors provide additional contribution to the responses. The limitation of the statistical model is that it is based on experimental results. For prediction, the response surface methodology requires slightly more number of experimentations.

4.5 ARTIFICIAL NEURAL NETWORK

Neural networks are composed of highly interconnected, simple processing units which are inspired by neural process observed in human
brain. Successful applications can be found in areas such as process engineering, process control and estimation, pattern recognition, fault detection and image analysis. A common theme of these applications is the ability of the artificial neural networks (ANN) to learn complex input/output relationships. They require no clearly defined algorithm or theory. Rather they have the property of acquiring knowledge through the presentation of examples. The ANN prediction is able to effectively capture the non-linear relationships and interactions of process parameters in a process.

Artificial neural networks are biologically inspired, that is, they are composed of elements that perform in a manner, analogous to the most elementary functions of the biological neurons. ANN has a parallel distributed architecture with a large number of neurons and connections. Each connection points from one node to another and is associated with a weight. The artificial neural networks are characterized by their topology, weight vectors and activation function that are used in the hidden layers and output layer. The ANN has emerged as a problem solving technique for many metal forming problems.

4.5.1 Architecture

Neural networks are computational models that share some of the properties of the brain. These networks consist of many simple “units” working in parallel with no central control, and learning takes place by modifying the weights between connections. The basic components of an ANN are “neurons”, weights, and learning rules. In general, neural networks are utilized to establish a relationship between a set of inputs and a set of outputs. Artificial neural networks are made up of three different types of “neurons”: (a) input neurons, (b) hidden neurons, and (c) output neurons. Inputs are provided to the input neurons, such as process parameters, and outputs are provided to the output neurons. These outputs may be a
measurement of the performance of the process, such as part measurements. The network is trained by establishing the weighted connections between the input neurons and output neurons via the hidden neurons. Weights are continuously modified until the neural network is able to predict the outputs from the given set of inputs within an acceptable user-defined error level.

The activation function used in both hidden layer and output layer is a non-linear function, whereas for the input layer, no activation function is used since no computation is involved in the input layer. All neurons in a layer are fully connected to neurons in the adjacent layers. Information flows from one layer to other layer in a feed forward manner. The feed forward back propagation network is a popular architecture among different types of neural networks and finds applications in several areas of engineering. The activation function $f(X)$ is a non-linear function and is given by

$$f(X) = \frac{1}{1 + \exp(-X)}$$  \hspace{1cm} (4.32)

where $f(X)$ is differentiable.

$$X = \sum_{i=1}^{W_{ij}} U_i + \text{threshold}$$

where $W_{ij}$ is the weight vector connecting the $i^{th}$ neuron of the input layer to the $j^{th}$ neuron of the hidden layer and $U_i$ is the input vector of the $j^{th}$ neuron of the hidden layer.

Therefore, the output of a neuron in the successive layer is given by

$$\text{Output}, \quad Y_1 = \frac{1}{1 + \exp\left[-\left(\sum_{i=1} W_{ij} U_i + \text{threshold}\right)\right]}$$  \hspace{1cm} (4.33)
Artificial neural network consists of a number of layers (generally three) each with a number of nodes. These nodes process the data and pass it to the next layer. The input is given to the input layer and the network, after processing, it gives the result via the output layer. It has been found that they have the capacity of learning a complex polynomial equation of any degree. Since the working of the network is based on simple mathematical equations, they have a low response time which is extremely useful in controlling dynamic processes. The ANN methodology represents a useful alternative to classical modeling technique, when applied to variable data sets presenting non-linear relationships. There is no need to assume an underlying data distribution such as usually done in statistical modeling. The accuracy of predictions depends more on the number of training patterns used than on the ANN architecture. It requires larger number of training patterns for achieving better results.