A TOOL FOR FUNCTIONAL DATA ANALYSIS AND EXPERIMENTATION

By

Dennis Rosario Román

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCES
in
INDUSTRIAL ENGINEERING

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS
2005

Approved by:

______________________________ _____________
Noel Artiles León, Ph.D. Date
Member, Graduate Committee

______________________________ _____________
José R. Deliz Álvarez, Ph.D. Date
Member, Graduate Committee

______________________________ _____________
David González Barreto, Ph.D. Date
President, Graduate Committee

______________________________ _____________
Rafael Rodríguez Solís, Ph.D. Date
Representative of Graduate Studies

______________________________ _____________
Agustín Rullán Toro, Ph.D. Date
Chairperson of the Department

Abstract

The objective of this research is to demonstrate the necessity of the industry of the use of functional data analysis techniques in order to analyze experiments. The type of experiment analyzed has the peculiarity that the response is measured repeatedly through time or through a specific signal factor. Two case studies are used in order to test the three methods proposed. The first method is a Point-Wise approach in which a classical ANOVA is performed in each level of the signal factor. The second uses a basis to represent the collection of all the response functions in order to relate the coefficients of the basis representation with the factors of the experiment. The third approach is a modification of the second method. The only difference is that regions are predetermined and the basis is applied and analyzed in each region separately. The three methods are proved in order to determine their effectiveness.

Resumen

El objetivo de esta investigación es el demostrar la necesidad de utilizar el análisis de datos funcionales en experimentos industriales. El tipo de experimentos analizados tiene la peculiaridad de que la respuesta se mide repetidamente a lo largo de un factor señal. Dos casos de estudio fueron utilizados para probar los tres métodos propuestos. El primer método se basa en conducir un análisis de varianza en cada nivel del factor señal. El segundo se utiliza una base para representar todas las funciones. Los coeficientes de la base están asociados a los factores del experimento. El tercer método es una modificación del segundo; la única diferencia es que se crean regiones con respecto al factor señal. En cada región se aplica la base y se analizan los coeficientes de la misma. Los tres métodos fueron probados con el propósito de determinar su efectividad.

Dedicatory

This thesis is dedicated in first place to God, thanks for all your blessings, and chances provided during my entire life. To my parents, for been with me, every time. For teaching me the real meaning of honesty, passion, determination, and work. For been such an example of what great persons are. To my brother Domingo, for been an example of good academic achievement and hard work. To my sister Maricelys, for teaching me that I can do everything that I want, that the real key to succeed is to have a balance between hard work and fun. To my girlfriend Myriam, for been so lovely, for been such a marvelous woman, for been the woman that I love.

Acknowledgments

I am greatly indebted to my advisor, David Gonzalez, for his guidance, patience and support during the investigation. Thanks for been a great example of a good, practical industrial engineer. I also want to express my gratitude to Dr. Noel Artiles, for been such a great professor, for push me all the time up to my limits, for challenge me all the time. I want to thank Dr. José Deliz, for his wisdom, and endless happiness. For been such a great professor and friend. I also want to thank the Industrial Engineering Department of the University of Puerto Rico at Mayagüez, for all the support during these years.

A special mention to my friend Zuriel. Thanks for helping me with the programming of the knots search. For been such a good friend in the last eight years. Thanks to my friend Jannette for always been such a nice person. I will never forget all the things that you have done for me. To my classmates and friends: Damaris, Geovannie, Juan Guillermo, Jessica, Cathe, Martha, Giannina, Carlos Andrés, Heidi Romero, Heidi Mejía, Karina, Eduardo, María, Chucho, Andrew, Paola, Wilfredo Huertas, Wilfredo Ortiz, Ana Maria, Lina, Wilfredo Yushimito, Andres Uribe, Efrain, Alexis, Anthony, Professor Mercedes Ferrer, Mabel, Litza, Vanessa, Beatriz, Marlene, Nancy, Yaleidy, Omar, Alejandro, Julián, Juan, Rose, and Roberto. Thanks for all the great moments, the jokes, the hard work, the passion, the intensity that we shared during all these years. In addition, I want to thank my best friend Melissa Serrano, just for been the person that you are.

I want to express gratitude to my new family at Hewlett-Packard Aguadilla. To the members of PRSSC (Puerto Rico Software Services Center), for been so nice to me. I am delighted to express my gratitude to Ida Calero, for giving me the chance to prove what I can do as a quality engineer.

To all my friends at Jayuya, but especially to Quevedo, Jose Omar and Gil for all the great moments, the support, the loyalty during all these years. To my friends Steven and Saul for been the best partners on earth. To all my friends, for been part of my life, for help me at every time, for make me a better person.

Table of Contents

List of Tables

List of Figures

List of Appendixes

List of Tables for the Appendixes

List of Figure for the Appendixes

1 Introduction

1.1 Justification

The technological developments in information processing have made real time process monitoring possible. The result of these developments is the collection of huge amounts of data. For that reason, new techniques are required in order to analyze and take advantage of the data available. Today it is possible to analyze a process or a system considering the input or output variables as functions instead of as discrete points. For this type of analysis some techniques have been developed, for example longitudinal data analysis and the functional data analysis (FDA). These analyses are appropriate when each individual is measured repeatedly through time or through a specific signal factor for example (frequencies, rotating speeds or compression loads [3]). The techniques previously mentioned have been applied successfully in the biological sciences, psychology, and social sciences. Only now are these techniques starting to be applied in engineering problems that affect industry.

In every manufacturing process, it is necessary to establish standards, and monitor the performance of the process in order to ensure the highest quality to the customer. In most of the manufacturing process Statistical Process Control (SPC), the set of tools used to control the process accuracy and precision. However, the implementation of an SPC program is not enough to ensure quality. Sometimes problems occur and it is necessary to have the proper mechanisms to detect the root causes in order to take the necessary corrective actions. One of the most widely used tools to find the root causes of the problems is Design of Experiments (DOE). DOE provides the mechanisms to find the factors that affect directly the process; it can be used as an optimization tool.

For all the reasons mentioned above, it is necessary to develop the proper set of tools to introduce FDA concepts to problems faced by industries. These tools must integrate the analysis of functional data in order to perform successful experiments.

1.2 Objectives

Functional data analysis is very complex; it involves a lot of computational effort and it requires the understanding of topics like, basis functions and Fourier series among others. The main objective of this research is to integrate the capabilities of the functional data analysis to industrial experiments in an efficient manner. Other specific objectives are

- Develop some methods to simplify the experimentation with functional data.
- Analyze and compare the methods proposed
- Develop a series of applications to ease the implementation of the proposed methods in industry

1.3 Organization

Chapter 2 presents the literature review and the background of the most relevant concepts related to the work completed in this research. These concepts include design of experiments (DOE), an introduction to functional data analysis, which is the most important concept, presented in this chapter and some other important topics such as linear regression, and functional analysis of variance. The methodology used to complete the objectives of this thesis is explained in Chapter 3. This chapter includes the discussion of the three methods proposed for this research in detail.

Chapters 4 and 5 present the results for the two case studies selected for this research. The first case is a simulated one; the second is a real world application. Chapter 6 explains the computer applications developed during this research. Chapter 7 presents the conclusions and future work to expand this research.

2 Basic Concepts

Several concepts required for the FDA experimental integration are discussed in this chapter. The concepts are design of experiments, functional data analysis, linear regression, functional analysis of variance, and high dimensional analysis of variance.

2.1 Design of Experiments

2.1.1 General Definition and Objectives

In general terms it is possible to define the concept of design of experiments as the systematic manipulation of certain input variables (factors) to observe their respective impact in an output variable (response variable). The main objectives of the design of experiments are the following

1. Obtain the maximum amount of information using the minimum of resources.
2. Detect the factors that shift the mean of the response variable.
3. Find the factors that affect the dispersion or variability of the response variable.
4. Detect the factors that do not have effect any effect in the response variable.
5. Construct an empirical model that relates the factors with the response variable.
6. Find the proper levels of the main factors to optimize the process.

2.1.2 Factorial Experiments

Factorial experiments are one of the most widely used experimental designs when two or more factors are involved. The basic characteristics of a factorial design are

1. All the possible level combinations of every factor are studied.
2. It is necessary to investigate the interaction effects among factors.

Among the factorial experiments, the most widely used is the 2^k in which each factor has only two levels. This type of experiment has certain properties desired in experimentation like orthogonality and projection among others.

2.2 Functional Data Analysis

2.2.1 Goals of the Functional Data Analysis

Functional data analysis (FDA) was developed for analyzing functional (or curve) data [2]. In FDA, the data consists of functions not vectors. Samples $y_1, y_2 \ldots y_n$ taken at time points $t_1, t_2, \ldots$ are converted into functions $\{x(t_j)\}, j = 1,2,\ldots$ as shown on Figure 2-1. The goals of the functional data analysis are

1. Represent the data in ways that facilitate further analysis.
2. Display the data to highlight various characteristics.
3. Explain variation in an outcome or dependent variable by using input or independent variable information.
4. Compare two or more sets of data with respect to certain types of variation, where two sets of data can contain different sets of replicates of the same functions, or different functions for a common set of replicates

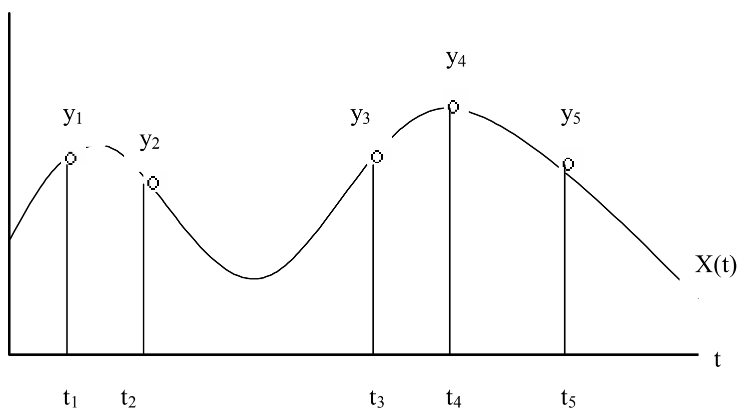


Figure 2-1 Graphical Representation of the Functional Data Problem

2.2.2 Main Steps in an Functional Data Analysis

Assuming that a functional datum for replication i arrives as a set of discrete measured values, $y_{i1}, y_{i2} \ldots y_{in}$, the tasks required to perform the functional data analysis are [2]

1. The raw data is collected, cleaned and organized.

2. Data are converted to functional form. That is, the raw data for observation *i* are used to define a function x_i that can be evaluated at all values of t over some interval.
3. Summary statistics and plots can be generated in order to ease the analysis.
4. The functions may be registered or aligned in some way; so important features found in each curve occur at roughly the same argument values.
5. Exploratory analyses can be carried out on the registered data, for example principal components analysis.
6. Models can be constructed to establish the relationship between a dependent variable with respect one or more independent variable.

2.2.3 Representing the Functional Data as a Smooth Function

The simplest way to convert the raw data into a functional object is using interpolation [2], [4]. This technique can be applied in cases when the measures do not have too much observational noise. When the raw data have, considerable noise is necessary to apply a smoother to reduce the effect of the noise in calculations and analysis. There are several types of smoothers that can be applied to functional data for example linear smoothing and smoothing based in basis-function methods. Those two types are explained with more details in the following sections.

2.2.3.1 Linear Smoothing

A linear smoother estimates the function value x(t) by a linear combination of discrete observations

$$\hat{x}(t) = \sum_{j=1}^{n} S_j(t) y_j \,. \qquad t \in T \tag{2.1}$$

The behavior of the smoother at t is determined by the weights $S_j(t)$. Linear smoothers can be represented in a matrix form. Suppose that the sequence $s_1 < s_2 < \ldots < s_m$ of evaluation points in T at which the function x is to be estimated, is on hand. Notice that the evaluation points do not need to be the same as the observation values t_j. Let $\hat{x}$ be the m-vector of values $x(s_i)$ and y for the vector of observed data y_j. Is possible to write

$$\hat{x} = Sy \tag{2.2}$$

where $\mathbf{S_{ij}}=S_j(s_i)$.

Many widely used smoothers are linear. The linearity of a smoother is a desirable feature for various reasons: The linearity property

$$S(ay + bz) = aSy + bSy$$

is important for obtaining various properties of the smooth representation. Simplicity of the smoother implies relatively fast computation. The concept of linear smoothing it is shown in Figure 2-2

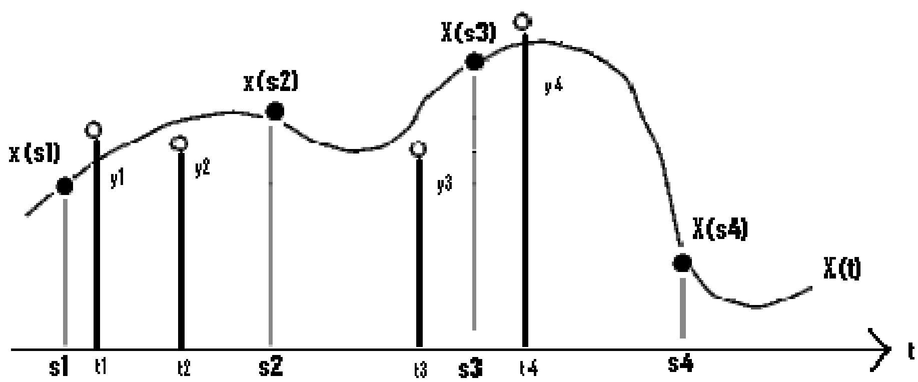


Figure 2-2 Graphical Explanation of the Linear Smoothing approach

From Figure 2-2 it is possible to observe the observed points y_i, i = 1,2,3,…, at times t_i, i = 1,2,3,…, and the predetermined evaluation points are S_i , i = 1,2,3,…

2.2.3.2 Smoothing Based in Basis Function Methods

The function x_i can be represented by a basis function expansion, which is defined by a set of basis functions, $\phi_k, k = 1,\ldots,K$ [2]. In this approach, a functional observation x_i is expressed as

$$x_i(t) = \sum_k^K c_{ik}\phi_k(t) \tag{2.3}$$

When these basis functions ϕ_k are specified, then the conversion of the data into a functional data object involves computing and storing the coefficients of the expansion, c_{ik}, into a coefficient matrix.

There are many bases possible, and many considerations to take into account. The following list provides a number of the more common bases:

1. Fourier Basis, typically used for periodic data.
2. B-Spline Basis, typically used for nonperiodic data.
3. Polygonal Basis, defining a function made up of straight-line segments.
4. Monomial Basis, consisting of the power of t: 1, t, t^2, t^3…
5. Exponential basis, a set of exponential functions, $e^{\alpha_k t}$ each with a different rate parameter α_k.

Of these basis functions, the first two are by far the most important [2]. The Fourier and the Polygonal basis are used when the data does not present many local features in extremely stable functions.

2.2.4 Summary Statistics for Functional Data

The classical summary statistics apply equally for functional data [2]. The mean, variance, covariance and correlation are shown in this section.

1. Mean

$$\bar{x}(t) = N^{-1}\sum_{i=1}^{N} x_i(t) \tag{2.4}$$

2. Variance

$$\mathrm{var}_x(t) = (N-1)^{-1}\sum_{i=1}^{N}\left[x_i(t) - \bar{x}(t)\right]^2 \tag{2.5}$$

3. Covariance

$$\mathrm{cov}_x(t_1,t_2) = (N-1)^{-1}\sum_{i=1}^{N}\{x_i(t_1) - \bar{x}(t_1)\}\{x_i(t_2) - \bar{x}(t_2)\} \tag{2.6}$$

4. Correlation

$$corr_x(t_1,t_2) = \frac{\mathrm{cov}_x(t_1,t_2)}{\sqrt{Var_x(t_1)Var_x(t_2)}} \tag{2.7}$$

2.2.5 Functional Linear Regression

Sometimes it is necessary to establish the relationship between one response variable and two or more independent variables. There are a few possible situations in which the functional linear regression applies [2], and these are

1. Functional response with non-functional independent variables.
2. Non-functional response with functional independent variables.
3. Functional response and functional independent variables.

In this thesis, the situation under study is the first one. This is because in the design of experiments, the set of independent variables (factors) is fixed. The design matrix (in a 2^k factorial design) contains ones and minus ones representing the two levels (low and high) of each factor. To complete the linear regression, which is the fundament of the analysis conducted in every design of experiment, is necessary to understand the linear regression when the data is not functional.

2.2.5.1 Linear Regression

Linear regression analysis is probably the most widely used technique to establish the relationship between a response variable and one or more decision (independent) variables in the form of

$$y = \mathrm{X}\beta + \varepsilon \tag{2.8}$$

where X is a matrix n x p, β is a vector p x 1 of the regression coefficients that has to be estimated by $\hat{\beta}$, and ε is the random error that are assumed to be independent are normally distributed with constant variance.

The essence of the linear regression is obtain a model which minimizes the sum of squared errors that are defined as

$$\text{sum of square errors (SSE)} = \sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 . \tag{2.9}$$

It can be proved that the coefficients that minimize the sum of the square errors can be obtained using the formula

$$\hat{\beta} = (X^T X)^{-1}(X^T y) . \tag{2.10}$$

A serious problem that may dramatically affect the usefulness of a regression model is multicollinearity, or near linear dependence among the regression variables [5]. Regression models fit to data by the method of least of squares when strong muticollinearity is present are notoriously poor prediction equations, and the values of the coefficients are often very sensitive to the data in the particular sample collected [5]. Another effect of the multicollinearity problem is the physical interpretation of the coefficients of the model obtained. The model can fit the data, but the coefficients that are used to determine which factors are more relevant in the experiment can be seriously affected not only in magnitude, the sign of the coefficient can affected. In order to detect the multicollinearity problem the variance inflation factors (VIF) are used. Variance inflation factors greater than 10 imply a serious multicollinearity problem.

In linear regression, not all the regression variables are relevant to the model all the times. Sometimes it is necessary to eliminate some variables from the model. One of the most widely used techniques for variable selection is the backward elimination stepwise procedure. The procedure begins with all **K** candidate regressors. Then a partial F-statistic is computed for each regressor as if it were the last variable to enter the model. The smallest of these partial F-statistics is compared with a pre-selected value F_{out} (or F-to-remove), for example, and if the smallest partial F value is less than F_{out}, the regressor is removed from the model. Now a regression with **K**-1 regressors is fit, the partial F-statistics for this new model calculated, and the procedure repeated. The backward elimination algorithm terminates when the smallest partial F value is no less than the preselected cutoff value F_{out} [5]. The stepwise procedure can be used to improve models that have the multicollinearity problem; due the elimination of the variables that are correlated, the model can be improved dramatically.

2.2.5.2 Functional Analysis of Variance (FANOVA)

The concepts of linear regression can be applied in a Point-Wise manner [2], [3] to functional data following these steps

1. Convert the response variable into a functional form; this implies the execution of most of the steps mentioned previously for example smoothing, registration, etc.

2. For each selected level of the signal factor, the coefficients of the model have to be calculated.
3. Each coefficient is converted into a function of the signal factor.

In other words the result of this procedure is a group of coefficients that are functions of the signal factor in the same manner that the response variable. In addition, of the functional linear regression a functional analysis of variance can be used to analyze the effects of some variables over the response. FANOVA considers the problem as a univariate ANOVA problem for each specific level of the signal factor. A crucial drawback to this approach is that an enormous number of hypothesis (the number of data points per curve can be hundreds or thousands) has to be tested simultaneously that causes a serious multiplicity problem [9].

2.2.5.3 High-Dimensional Analysis of Variance (HANOVA)

It is a powerful overall test for functional hypothesis testing based on the decomposition of the original functional data into Fourier or wavelet series, and applied the adaptive Neyman and wavelet thresholding procedures to the resulting empirical Fourier and wavelet coefficients respectively. The underlying idea based on the sparcity of the underlying signal's representation in the Fourier or wavelet domains that allows a significant reduction of dimensionality [6], [9]. This procedure is not used in the methods proposed, but it is a good example of how complex the methods developed for curves comparison are.

2.2.5.3.1 Testing Differences among Multiple Groups of Curves

Consider the observed curves from *I* different groups:

$$\{X_{ij}(t), \mathrm{i}=1,2\ldots,\mathrm{I}, \mathrm{j}=1,\ldots,\mathrm{n_i}, \mathrm{t}=1,\ldots,\mathrm{T}\}$$

It is possible to assume that

$$X_{ij}(t) = f_i(t) + \varepsilon_{ij}(t) \tag{2.11}$$

where $\{e_{ij}(t), \mathrm{t}=1,...,\mathrm{T}\}$ are stationary time series with mean 0. One is interested in testing hypothesis:

$$\begin{aligned} Ho &: \mathrm{f_i}(t) = f(t) \\ H_1 &: \mathrm{f_i}(t) \neq f(t) \end{aligned} \quad \text{for i} = 1,\dots,\text{I} \quad \text{and t} = 1,\dots,\text{T}$$

Let $\{X_{ij}^*(t)\}$ be the direct Fourier transform of the vector $\{X_{ij}(t)\}$. Then $\{X_{ij}^*(t)\}$ satisfied the ideal model

$$X_{ij}^*(t) = f_i^*(t) + \varepsilon_{ij}^*(t) \tag{2.12}$$

Then the previous hypothesis is equivalent to the following problem $Ho: f_i^*(k) = f^*(k)$ for $\text{i} = 1,\dots,\text{I}$ and $\text{k} = 1,\dots,\text{T}$ then it is possible to apply the Adaptive Analysis of Variance [6].

2.2.5.3.2 Adaptive Analysis of Variance

For simplicity of notation, it is possible to state the HANOVA as follows: let $X_{ij} \sim N(\mu_{ij}, \sigma_{ij}^2)$ be independent and random variables [6]. One wants to test

$$Ho: \mu_{ij} = \mu_j \quad \text{for i} = 1,\dots,\text{I} \text{ and j} = 1,\dots,\text{n}$$

where n is large. It is assumed that $\{\sigma_{ij}^2\}$ are known. Suppose that prior knowledge indicates that useful information is concentrated on the first *m* cells. Then the following sub-problem is considered

$$Ho: \mu_{ij} = \mu_j \quad \text{for j} = 1,\dots,\text{m}$$

The maximum likelihood ratio statistic for the sub-problem is

$$X^2 = \sum_{j=1}^{m} \sum_{i=1}^{I} \sigma_{ij}^{-2} \left(X_{ij} - \overline{X}_{.j} \right)^2 \tag{2.13}$$

with

$$\overline{X}_{.j} = \frac{\sum_{i=1}^{I} \sigma_{ij}^{-2} X_{ij}}{\sum_{i=1}^{I} \sigma_{ij}^{-2}} \tag{2.14}$$

Thus a level-α test is to reject Ho when

$$F_m = \frac{1}{\sqrt{2(I-1)m}} \left\{ \sum_{j=1}^{m} \sum_{i=1}^{I} \sigma_{ij}^{-2} \left(X_{ij} - \overline{X}_{.j} \right)^2 - (I-1)m \right\}$$
$$\geq \frac{1}{\sqrt{2(I-1)m}} \left\{ X^2_{(I-1)m}(1-\alpha) - (I-1)m \right\} \tag{2.15}$$

Note that when the degrees of freedom (I-1)m are large, F_m is normally distributed with mean $\delta_m^{*2} = \frac{\delta_m^2}{\sqrt{2(I-1)m}}$ and variance 1. Where $\delta_m^2 = \sum_{j=1}^{m} \sum_{i=1}^{I} \sigma_{ij}^{-2} \left(\mu_{ij} - \overline{\mu}_{.j} \right)^2$ with $\overline{\mu}_{.j} = \frac{\sum_{i=1}^{I} \sigma_{ij}^{-2} \mu_{ij}}{\sum_{i=1}^{I} \sigma_{ij}^{-2}}$. In practice, m must be determined as $\hat{m} = \arg\max_{1 \leq m \leq n} F_m$ leading to the adaptive testing statistic, which defines the HANOVA

$$F_{\hat{m}} = \max_{1 \leq m \leq n} \frac{1}{\sqrt{2(I-1)m}} \left\{ \sum_{j=1}^{m} \sum_{i=1}^{I} \sigma_{ij}^{-2} \left(X_{ij} - \overline{X}_{.j} \right)^2 - (I-1)m \right\} \tag{2.16}$$

Specifically when I=2, the test statistic reduces to

$$F_{\hat{m}} = \max_{1 \leq m \leq n} \frac{1}{\sqrt{2m}} \left\{ \sum_{j=1}^{m} \frac{\left(X_{1j} - X_{2j} \right)^2}{\sigma_{1j}^2 + \sigma_{2j}^2} - m \right\} \tag{2.17}$$

From the previous sections, it is possible to observe the amount of concepts that will be integrated and simplified on this work. As mentioned on the first chapter, the intention of this research is to make feasible the use of these tools in industry. The next chapter presents the details of the methodology that will be used during this work. Three methods were proposed based on the concepts presented on this chapter.

3 Methodology

3.1 Point-Wise method

The Point-Wise method is based on the ideas of the functional data analysis (FDA) proposed by Ramsay and Silverman [2] in their book and papers in specific in the functional analysis of variance (FANOVA) mentioned in the previous chapter. A linear regression model is generated at each level of the signal factor. Then curves of coefficients are obtained. In other words, dynamic models are generated. The main difference of this method with respect the FDA is that no smoothing techniques have been used. In the type of industrial experiments considered there is no interest with respect the derivatives of the response variable; this is the justification for not using the smoothers; also simplification is highly desired in a technique that will be applied in the industry. In addition to the coefficients, the respective statistical tests to validate the regression are performed at each level of the signal factor. To be more specific, an F-test to verify the significance of the regression and a T-test to verify the contribution of each factor are performed at each level of the signal factor resulting in curves for the F and T tests. In addition, the residuals and the coefficients of determination are calculated at each level of the signal factor.

3.1.1 Model Validation and Inference

As mentioned previously, an F test is applied to validate the significance of the regression at each level of the signal factor. The F test is used to complete the following hypothesis test

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0$$

$$H_1 : \beta_j \neq 0 \text{ for at least one j.}$$

The statistical test is given by

$$F_0 = \frac{MSR}{MSE} \tag{3.1}$$

where the MSE is the mean squared error and the MSR is the mean squared regression. The hypothesis is rejected if $F_0 > F_{\alpha,p-1,n-p}$.

The coefficients are verified in order to know which ones are making a real contribution to the model. This is very important because the coefficients of the regression are associated with a factor or combinations of factors in the experiment. The hypothesis for the coefficients is given by

$$H_0 : \beta_j = 0$$
$$H_1 : \beta_j \neq 0$$

The statistical test t used is given by

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \tag{3.2}$$

where C_{jj} is the diagonal element of $(\mathbf{X'X})^{-1}$ corresponding to $\hat{\beta}_j$. The null hypothesis is rejected if $|t_o| > t_{\alpha/2,n-k-1}$ [1].

An important measure of performance for the linear regression models is the determination coefficient R^2 that is a measure of the total variability of the data explained by the model. The formula for this coefficient is

$$R^2 = \frac{SSR}{SSE} \tag{3.3}$$

Because R^2 always increases as more terms enter to the model [1]. It is preferred to use the adjusted R^2 defined as

$$R^2_{adj} = 1 - \left(\frac{n-1}{n-p}\right)\left(1 - R^2\right) \tag{3.4}$$

In addition to the previous measures of performance, the Matlab® applications created generate the plots for the residuals in order to validate the stochastic assumptions for the linear regression. The model validation procedure is also part of the other two methods that are going to be presented in this chapter.

3.2 Basis Representation Model

The focus of this method is to represent the response functions as a sum of basis functions. However, in this representation the coefficients of the basis functions are going to be dependent on the factors of the experiment. The purpose of this representation is to provide a direct way to capture the factors that have more relevance in the experiment. In the FDA, the basis representation is used to estimate the response curves. In this method, the idea is to relate directly the response to the factors. Several bases can be applied. The basis selection is going to depend on the behavior of the responses. In the Matlab, applications developed in this work three type of basis are considered: Monomial basis, Fourier basis and a Cubic Spline basis. These bases have been widely used. To illustrate this method, an example is presented in the next section.

3.2.1 An Example for the Basis Representation Method

Suppose that an experiment with two factors (x_1 and x_2) and a functional response is being analyzed. The analyst considers important the interaction between the factors and considers that a Fourier basis will be appropriate for the data due the periodical behavior of the response. The general form for a Fourier basis expansion is the following:

$$y(t) = c_0 + c_1 \sin(\omega t) + c_2 \cos(\omega t) + c_3 \sin(2\omega t) + c_4 \cos(2\omega t) + \dots \tag{3.5}$$

where $\omega = \frac{2\pi}{\mathrm{T}}$ and T is the highest level of the signal factor [2]. Since the interaction is going to be considered, the first term of the expansion will be $c_0 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$ and the general basis representation model for this example will be

$$\begin{aligned} y(t) = {} & (\beta_{00} + \beta_{10} x_{10} + \beta_{20} x_2 + \beta_{30} x_1 x_2) + \\ & (\beta_{01} + \beta_{11} x_1 + \beta_{21} x_2 + \beta_{31} x_1 x_2) \sin(\omega t) + \\ & (\beta_{02} + \beta_{12} x_1 + \beta_{22} x_2 + \beta_{32} x_1 x_2) \cos(\omega t) + \\ & (\beta_{03} + \beta_{13} x_1 + \beta_{23} x_2 + \beta_{33} x_1 x_2) \sin(2\omega t) + \\ & (\beta_{04} + \beta_{41} x_1 + \beta_{24} x_2 + \beta_{34} x_1 x_2) \cos(2\omega t) + \dots \end{aligned} \tag{3.6}$$

The dependence of the coefficients of the basis representation on the factors of the experiment can be observed directly. The measures of performance discussed previously

(Section 3.1.1) can be applied in order to determine the effectiveness of the model. If the design matrix **X** have coded variables then the magnitudes of the resulting coefficients can be used to determine which factors are the most important in the experiment. The same procedure can be applied to the other bases; the key element is to pick the right basis and the best number of terms for the basis expansion.

3.3 Piece-wise Method

In all the previously designed methods, the dimensionality is an issue. As the number of levels of the signal factor increases so does the complexity of the analysis. It is necessary to develop an approach able to deal with the dimensionality problem and detect which factors are more relevant in the experiment. A Piece-Wise approach has been proposed the idea of this method is the following:

- Divide the range of the signal factor in a series of regions
- Use a common basis representation in each region.
- Perform a stepwise procedure in order to simplify the model and eliminate the non-relevant terms in the regressions.
- Verify the coefficients of the regression in order to determine which factors are more relevant in the experiment per region.

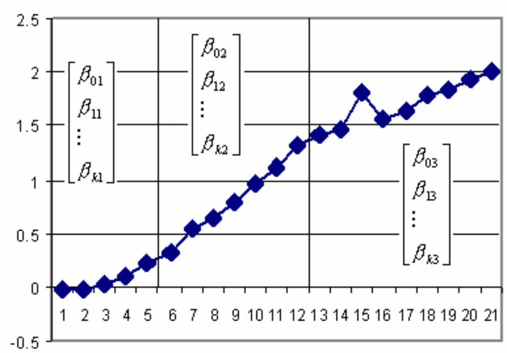


Figure 3-1 Illustration of the Piece-Wise method. In each region, there is a set of coefficients for one experimental condition

This procedure is a modification to the basis representation method. The only difference between both procedures is the division of the levels of the signal factor. This procedure is used when one desires to know which factors are affecting the response in some specific regions of the signal factor.

A challenge behind this method is to find an optimal method to obtain the knots that are used to divide the levels of the signal factor in regions. For a predetermined number of knots k that divides the levels of the signal factor in k+1 region, it is necessary to determine the optimal position of those knots. Let *knots* be the vector of positions for the different levels of the signal factor, that are equally spaced. Let SSE_i be the Sum Squared of Error for the region i. The optimization problem is stated as follows

$$\min z = \sum_i SSE_i \tag{3.7}$$

Subject to

$$knots(j+1) - knots(j) \geq \mathrm{B} \qquad \forall\, \mathrm{j}, \mathrm{j} = 1,2,\ldots \mathrm{k} \tag{3.8}$$

$$length(knots) \leq floor\left(\frac{length(t)}{B}\right) - 1 \tag{3.9}$$

$$knots(1) \geq B \tag{3.10}$$

$$knots(k) \leq length(t) - B \tag{3.11}$$

$$B \geq 0 \tag{3.12}$$

where *length(t)* is the number of levels of the signal factor *t,* and B is a parameter that sets the minimum distance between knots in order to ensure the feasibility of the regressions.

The first constraint forces the knots to keep a distance of B levels in order to make the regression estimation possible. The second restriction delimits the number of knots to be used. This number cannot be more than the total levels of the signal factor divided by the constant B minus one; this constraint guarantees that there are enough points for the last region. The third constraint forces the first knot to be at the B position or higher in the levels of the signal factor, as the previous restriction this forces the regions to have enough points to estimate the regressions required at each region. The value for the constant B was selected to be equal to three. That value was selected in order to limit the maximum number of knots to be less than a third part of the signal factor levels. Equation 3.11 ensures enough points for the regressions in the last region. The last restriction is for the values of B that must be positive. All the values of the vector *knots* have to be integer due the definition of the variable. This is a limitation in terms of the software that is been used because the optimizations tools of Matlab do not deal with integer variables.

There is another constraint that has to be considered, the continuity of the functions at the knots. In order to force the models to obey this restriction the following final procedure was used.

1. Divide the response matrix into regions delimited by the knots
2. Obtain the models for each region using the ordinary least of squares
3. Find the average of the estimated responses per experimental condition at the last signal factor level of the first (previous) region
4. Use the averages previously calculated as constraints for the indicator variables of the next region. The number of indicator variables to use will be the number of experimental conditions minus one.
5. Calculate the coefficients of the regression using the restricted least of squares procedure based on the following formula

$$b_r = b + (X^T X)^{-1} R^T [R(X^T X)^{-1} R^T]^{-1} (r - Rb) \tag{3.13}$$

   where b_r is the vector of the restricted coefficients obtained from Equation 3.13, b is the vector of coefficients obtained by the ordinary least of squares regression, R and r are the restrictions expressed in the following form $\mathbf{R}\beta = \mathbf{r}$. The number of rows of the matrix R will be equal to the number of the experimental conditions minus one. The number of columns of the R matrix it is going to be equal to the number of terms of the basis expansion plus the number of indicator variables.
6. Repeat steps 3 to 5 on the other regions

A procedure was developed in order to find the optimal set of knots given a desired number of knots. This method is not efficient but it is effective. The main idea is to evaluate all the possible combinations of knots for a pre-selected quantity; and select the combination that minimizes the total sum of squared errors. The computational effort increases as the number of knots increases. The optimal knots are used to delimit the regions and perform the rest of the procedure.

In each region a basis, it is going to be used in order to model the behavior of the responses. The bases that are going to be used are the monomial basis and a Fourier basis, the cubic spline is not going to be considered. Since this basis in the previous method was used to incorporate some delimitation of regions in the estimates. In this method is not necessary because the signal factor has been divided before completing the basis expansions. The following section presents a detailed example for this method.

3.3.1 An Example for the Piece-wise Method

Suppose that an experiment was conducted considering two factors (x_1 and x_2), with functional response. The analyst wants to study two regions; this means that only one knot is required. Assume tha the optimal knot location is knwon and the initial models were obtained, considering a monomial basis. The Figure 3-2 illustrates the separation of the response curves of the two regions at the knot for two experimental conditions with two replicates.

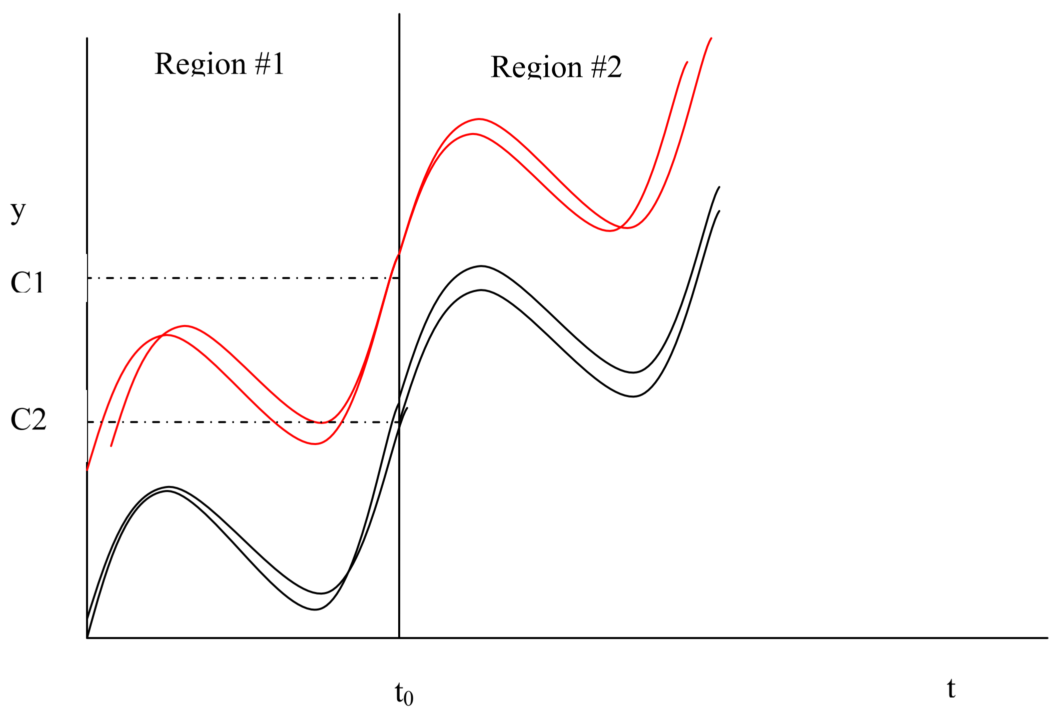


Figure 3-2 Illustration of the separation of two adjacent curves at the knot t_0

The point C1 and C2 represents the averages of the estimated responses at the last level of the first region for two of the experimental conditions. The functions on represents the functions obtained in each region without the continuity constraint. The next step of this procedure is to set C1 or C2 (one of the experimental conditions must be used as a base for the indicator variables) as the constraint for the indicator variable for the next region model; let us take C2 for this example. Suppose that a monomial basis is used, then the expansion for the second region is in the form of the Equation 3.14

$$\begin{aligned} y(t) &= \left(b_{00} + b_{10}x_1 + b_{20}x_2 + b_{30}x_1x_2\right) + \left(b_{01} + b_{11}x_1 + b_{21}x_2 + b_{31}x_1x_2\right)(t - t_0) + \\ &\left(b_{02} + b_{12}x_1 + b_{22}x_2 + b_{32}x_1x_2\right)(t - t_0)^2 \end{aligned} \tag{3.14}$$

From Equation 3.14 it is possible to observe that the terms related to the signal factor the knot is subtracted. The reason for this is in order to let the indicators variables to take all the effect of the imposed restrictions. The restrictions matrix (R) and right hand side (r) have the following form

$$R = (1,0,\ldots 0,1) \quad \text{and} \quad r = (C_2) \tag{3.15}$$

After imposing the restriction, the restricted least of squares procedure is completed for the second region. The coefficients of all the regions except for the first one are obtained using the restricted least of squares.

3.4 Metric of comparison for the all methods

In order to compare all the methods it is necessary to establish a metric or a measure of performance. The metric considers some important quantities such as the number of parameters estimated, the sum of all the squared sums of error and the total number of data points used. The name given to this metric is “pseudo-MSE” and the following equation defines it.

$$\text{Pseudo MSE} = \frac{\text{Total SSE}}{\text{N} - \text{total number of parameters}} \tag{3.16}$$

where *Total SSE* is the total sum of all the SSE. In the basis representation method, this is a single number but in the Point-Wise approach, there is an SSE at every level of the signal factor. The variable N represents the total number of data points used, and the variable *total number of parameters* is the total number of parameters (coefficients) estimated in the procedure. For the basis representation method this is the total number of coefficients of the basis expansion, for the Point-Wise and the Piece-Wise methods is the total number of coefficients calculated for the whole procedure. It is important to mention that Equation 3.16 must include the indicator variables used for the continuity constraint, which implies an additional lost of degrees of freedom. In general, the method with the lowest pseudo-MSE will be preferred.

The next chapter presents the results obtained using the methods presented on a theoretical case study that was developed in order to challenge the proposed methods and verify their efficacy. Also was intended to compare the procedures as mentioned earlier in order to determine which procedure is better than the others.

4 Theoretical Case Study

4.1 Introduction

Two case studies are used to compare and validate the proposed methods. The first case is presented in detail in this chapter. This chapter begins with a description of the experiment. Then the results obtained using each method is going to be presented and discussed.

4.2 Experiment Description

A macro was created in MS Excel® to generate the functions that correspond to each treatment in the experiment. It is important to mention that the function used to create the macro and the nature of the errors was unknown to the author until the end of the study. Appendix #1 shows some details with respect to the function used. The experiment has two factors each one with two levels (a classical 2^2 experiment) and 21 measures of the response variable were generated in each treatment. In addition, a central treatment was performed in each run. A sample of a single run of this experiment is shown in Figure 4-1.

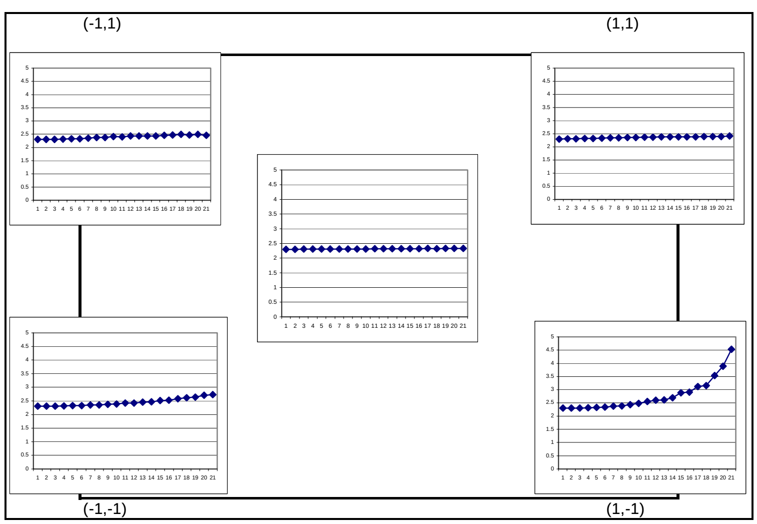


Figure 4-1 A graphical representation of a full run of the theoretical case study

From Figure 4.1 shows that the response function changes at each combination of factors. Five runs of the experiment were used to test the methods. The figures and the tables for all the runs are presented in the Appendix #1. The general model used as a base for all the methods is given by

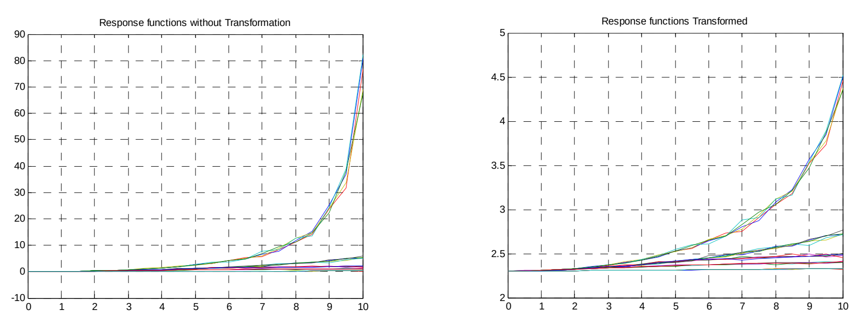
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_1 x_2 \quad (4.1)$$

where x_1 and x_2 are the factors of interest in the experiment. For the Point-Wise method, the coefficients of Equation 4.1 are calculated at each level of the signal factor. For the other two methods, the Equation 4.1 is inside each term in the expansions. In the same way as presented in the example on section 3.2.1. The results for the application of each method are presented next.

4.3 Results for Theoretical Case

4.3.1 Point-Wise Method

A variable transformation y = ln(y+10) was necessary to scale all the functions. The transformation was selected because adding the ten eliminated the possibility of having negatives inside the logarithm and this mathematical function was used due the shape of the original functions. This transformation is used for all the methods. From Figure 4-2 one can observe the need of the variable transformation in order to put all the runs of the experiment in a more suitable scale. A Matlab® program was created in order to ease the analysis of the data. The results of the Point-Wise analysis of variance and the measure of performance for the method are shown in the next pages.



(a) (b)

Figure 4-2 Plot of all the functions for the five runs of the theoretical experiment. (a) Responses without Transformation, (b) Responses Transformed.

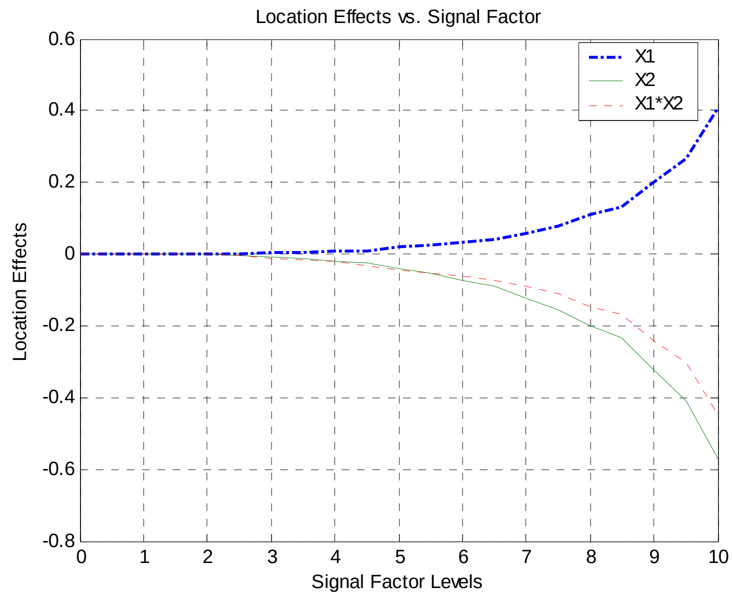


Figure 4-3 Location effects for theoretical case study using the Point-Wise method

The model generated in each level of the signal factor corresponds to Equation 4.1 as mentioned previously and the coefficients at each level are included in Table 4-1. It is possible to observe from Figure 4-3 that both factors are relevant in this experiment. In the first levels of the signal factor, the values for of the effects are very close to zero. The effects that correspond to x_1 have larger magnitudes with respect x_2 and their interaction on the first levels of the signal fact (0 to 2); then x_2 becomes more relevant than x_1 but the interaction effects have the larger magnitudes on the higher levels of the signal factor (2.5 to 10).

Table 4-1 Coefficients Table

t	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
0	2.3016	0.001875	-5.50E-05	3.50E-05
0.5	2.3035	0.00183	-8.00E-05	0.00137
1	2.3078	0.001705	-0.00018	0.001605
1.5	2.3139	0.00178	-0.00079	0.00018
2	2.322	0.00172	-0.00172	-0.0011
2.5	2.3315	0.00207	-0.00402	-0.00549
3	2.3432	0.00302	-0.00616	-0.01015
3.5	2.3545	0.00539	-0.01182	-0.01548
4	2.3681	0.008675	-0.01905	-0.02124
4.5	2.3871	0.010345	-0.02607	-0.0322
5	2.4057	0.01938	-0.0421	-0.04274
5.5	2.4241	0.024105	-0.05247	-0.05179
6	2.4448	0.03193	-0.07309	-0.06225
6.5	2.4643	0.04327	-0.09145	-0.07266
7	2.493	0.05878	-0.12118	-0.09134
7.5	2.5238	0.078595	-0.15351	-0.11118
8	2.5666	0.11049	-0.20129	-0.14681
8.5	2.5963	0.12966	-0.23407	-0.16937
9	2.6718	0.20146	-0.32303	-0.24241
9.5	2.7434	0.2648	-0.40945	-0.30374
10	2.877	0.40942	-0.57222	-0.44798

The residuals plot is next. The importance of this plot is to validate the stochastic assumptions with respect to the behavior of the regression errors. Residuals are shown in Figure 4-4.

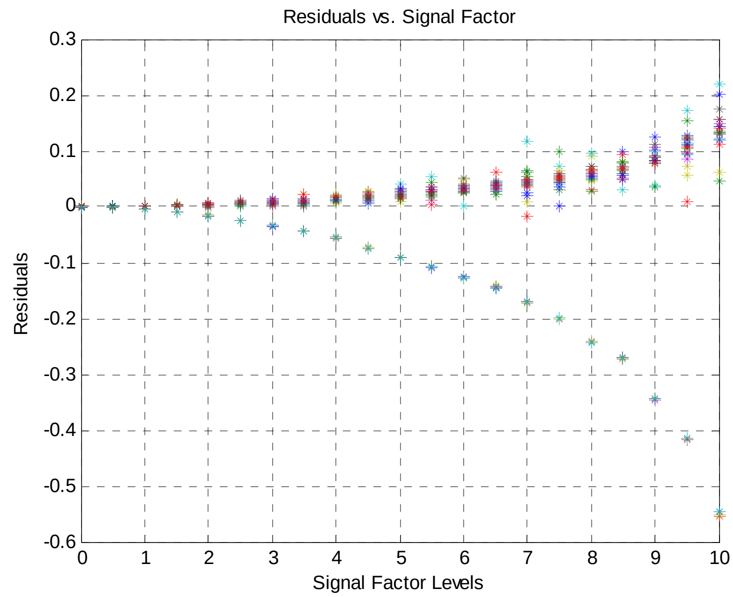


Figure 4-4 Residuals plot for the theoretical case using the Point-Wise method

This graph can be used to verify if the variance of the models is constant. The obtained models variances are not constant. The problem corresponds to a specific experimental condition ($x_1 = 1$, $x_2 = -1$). A dramatic change in the response occurs in that experimental condition. This change can be seen in Figure 4-1 and in Figure 4-2. The models underestimation of the response at the mentioned experimental condition increases with the signal factor. For the rest of the experimental conditions the residuals look approximately constant. The next three figures provide different measures of performance in order to verify the model adequacy. These measures of performance include the F ratio tests, the determination coefficient and the T test discussed previously.

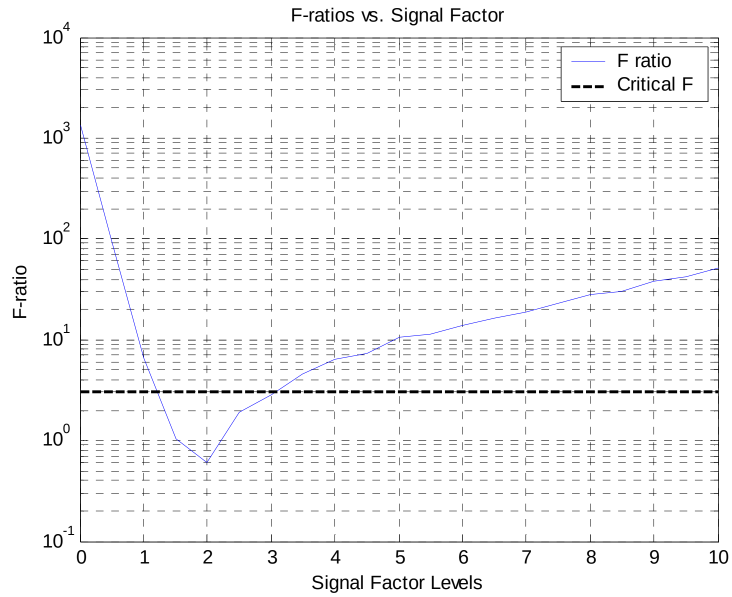


Figure 4-5 F ratio test for the theoretical case using the Point-Wise method

Figure 4-5 show that all the regressions conducted by the Point-Wise procedure for this experiment were significant. The critical F distribution value for all cases is 3.0725.

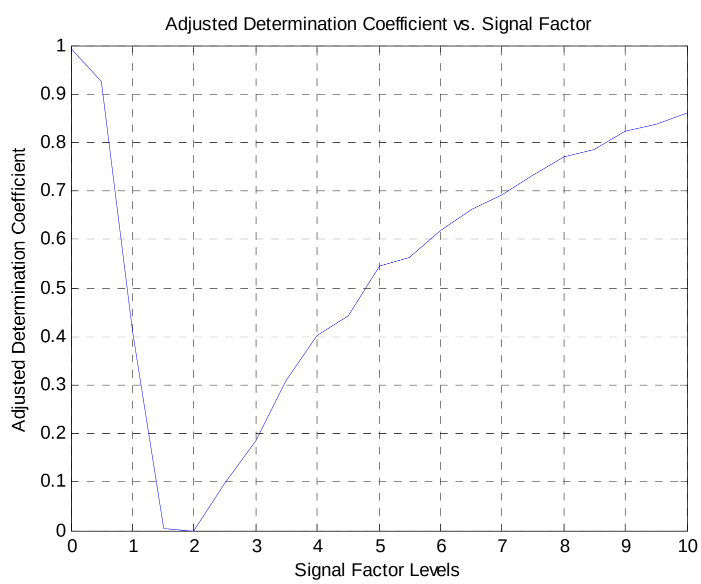


Figure 4-6 Adjusted R^2 for the theoretical case study using the Point-Wise method

Figure 4-6 shows the adjusted determination coefficient for all the regressions executed. It is possible to deduce that most of the regressions are reasonable because the adjusted R^2 is over 0.60, which might be acceptable in practice. For the signal factor levels from 0.5 to 5, the regressions are not very effective and the adjusted R^2's are very low.

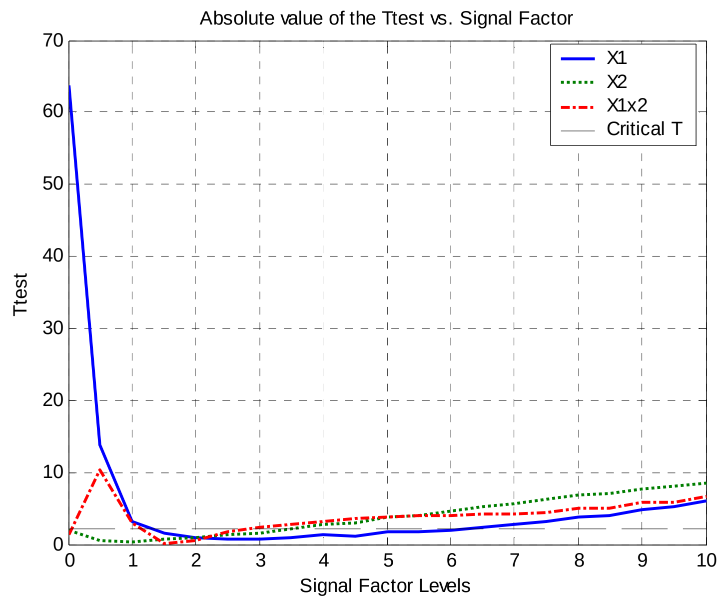


Figure 4-7 Absolute value of the T test for the theoretical case using the Point-Wise method

The T test plot in Figure 4-7 shows that the coefficients were significant for most of the regressions. In the signal factor, levels from 0.5 to 5 most of the coefficients are considered non-relevant.

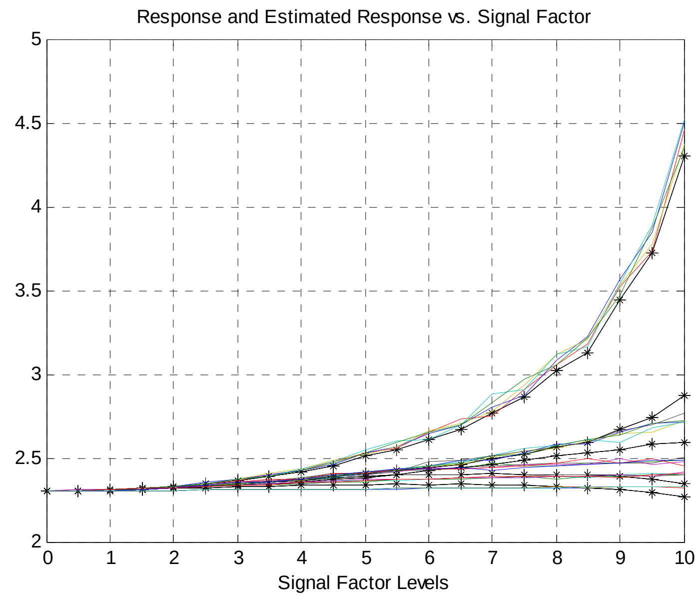


Figure 4-8 Response and Estimated Response Plot for the theoretical case using the Point-Wise method. The asterisks represent the estimated response.

Figure 4-8 shows the response and the estimated response on the same set of axes. The asterisks on the plot represent the estimated response. Form this figure can be observed that the obtained models estimate the response function in a reasonable manner. The next Figure 4-9 shows the lack of fit F ratio test for the Point-Wise analysis of variance realized. This graph potentially indicates a curvature component that has not been considered by the individual analyses of variance. This can be the reason of the low values of the adjusted determination coefficients in this procedure. In addition, it could be the explanation for the behavior of the residuals.

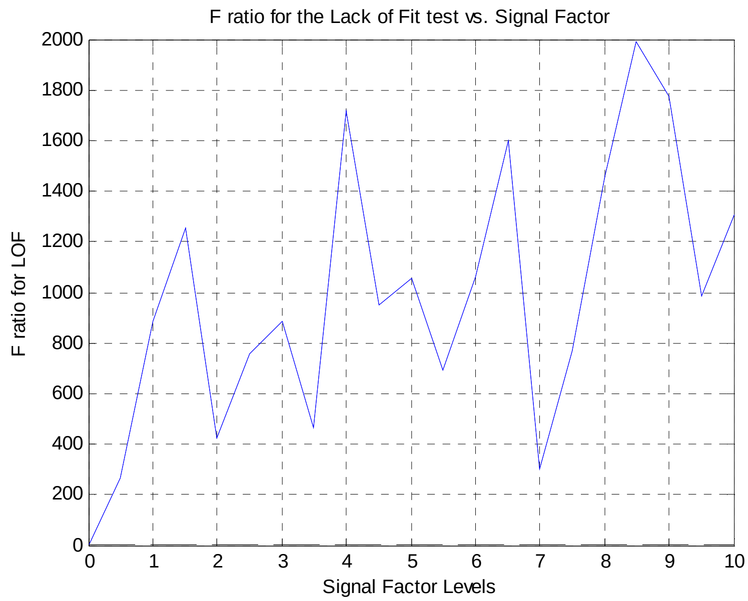


Figure 4-9 F ratio test for the lack of fit for the theoretical case study

The pseudo-MSE for this procedure is summarized in the next table.

Table 4-2 Pseudo-MSE for the theoretical case using the Point-Wise method

Total SSE	Total Number of Parameters	Total Number of Data Points	Pseudo MSE
5.48682	84	525	0.0124

In general, this method worked is adequate for this experiment, with the transformed data. An important fact is that the results obtained using the Point-Wise method are exactly the same results obtained using the functions prepared by Ramsay to perform the Functional Analysis of Variance (FANOVA) with exception of the smoothing of the curves, the shape and the inferences obtained with both methods are the same (see Appendix #2). In the next section, the results for the basis representation method are presented.

4.3.2 Basis Representation Method

4.3.2.1 Monomial Basis

The selected monomial basis has the following general form

$$y(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + \cdots c_k t^k \tag{4.2}$$

Inserting the general model used in the experiment provides the following a basis representation with k =2:

$$\begin{aligned} y(t) &= \left(b_{00} + b_{10}x_1 + b_{20}x_2 + b_{30}x_1x_2\right) + \left(b_{01} + b_{11}x_1 + b_{21}x_2 + b_{31}x_1x_2\right)t + \\ &\left(b_{02} + b_{12}x_1 + b_{22}x_2 + b_{32}x_1x_2\right)t^2 \end{aligned} \tag{4.3}$$

The k =2 was selected because a higher number in this case will produce a serious multicollinearity problem. The results obtained for this model are shown in the next pages. The calculations were performed using an application created in Matlab. The first column of the Matlab output labeled as “Terms” indicates with a number which term of the basis corresponds to each coefficient. The Table 4-3 summarizes the relationship between the column “Term” and the coefficient of the basis that is being represented.

Table 4-3 Relation between the Matlab’s output column “Term” and the coefficients of Equation 4.3

Terms	Coefficient Represented	Related Factor	Basis Term
0	b_{00}	1	1
1	b_{10}	X1	1
2	b_{20}	X2	1
3	b_{30}	X1*X2	1
4	b_{01}	1	t
5	b_{11}	X1	t
6	b_{21}	X2	t
7	b_{31}	X1*X2	t
8	b_{02}	1	t^2
9	b_{12}	X1	t^2
10	b_{22}	X2	t^2
11	b_{32}	X1*X2	t^2

The following is the initial output generated by the application.

Table 4-4 Initial Matlab Output for the Theoretical Case Study Using the Basis Representation Model with a Monomial Basis (K =2)

```
The model is not a good one due a multicollinearity problem
%-------------------------------------------------------------------------%
 Results for the Analysis of Variance for the Basis Representation Model
%-------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-------------------------------------------------------------------------%
                     Coefficients and Variance Inflation Factors
%-------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
   0     2.3266      0.014077   165.2737   0

  Term    Coefficients  SE Coef   T-test  P-value   VIF
   1.0000    0.0402    0.0157    2.5534    0.0110    7.4822
   2.0000   -0.0435    0.0157   -2.7635    0.0059    7.4822
   3.0000   -0.0312    0.0157   -1.9827    0.0479    7.4822
   4.0000   -0.0210    0.0065   -3.2125    0.0014   14.7300
   5.0000   -0.0411    0.0073   -5.6316    0.0000   54.9026
   6.0000    0.0489    0.0073    6.7046    0.0000   54.9026
   7.0000    0.0341    0.0073    4.6806    0.0000   54.9026
   8.0000    0.0068    0.0006   10.8124         0   14.7300
   9.0000    0.0068    0.0007    9.6571         0   32.2149
  10.0000   -0.0091    0.0007  -12.9937         0   32.2149
  11.0000   -0.0066    0.0007   -9.4092         0   32.2149

%-------------------------------------------------------------------------%
                     R^2 and Adjusted R^2
%-------------------------------------------------------------------------%
  R^2,    R^2(adj)
0.85255  0.84939
%-------------------------------------------------------------------------%
                     Analysis of Variance
%-------------------------------------------------------------------------%
Source            DF          SS          MS        F         P
Regression        11     41.2436     3.7494     269.6507     0
Residual Error   513     7.1331      0.013905
Total            524     48.3767
%-------------------------------------------------------------------------%
```

By inspecting the VIF in Table 4-4 one can conclude that the model has a multicollinearity problem. In addition to that problem, some of the terms must be eliminated in order to improve the adequacy of the basis representation model. To complete that task the stepwise command of Matlab is used. This command generates a graphical use interface (GUI) in which the user can select the terms to eliminate. For each term selected, the root mean squared, the determination coefficient R^2, the F ratio and the corresponding p-value are calculated. The Figure 4-10 shows the GUI that corresponds to the stepwise command with all the terms in the model. When a term is eliminated, the text and graphs that correspond to that term change the font color.

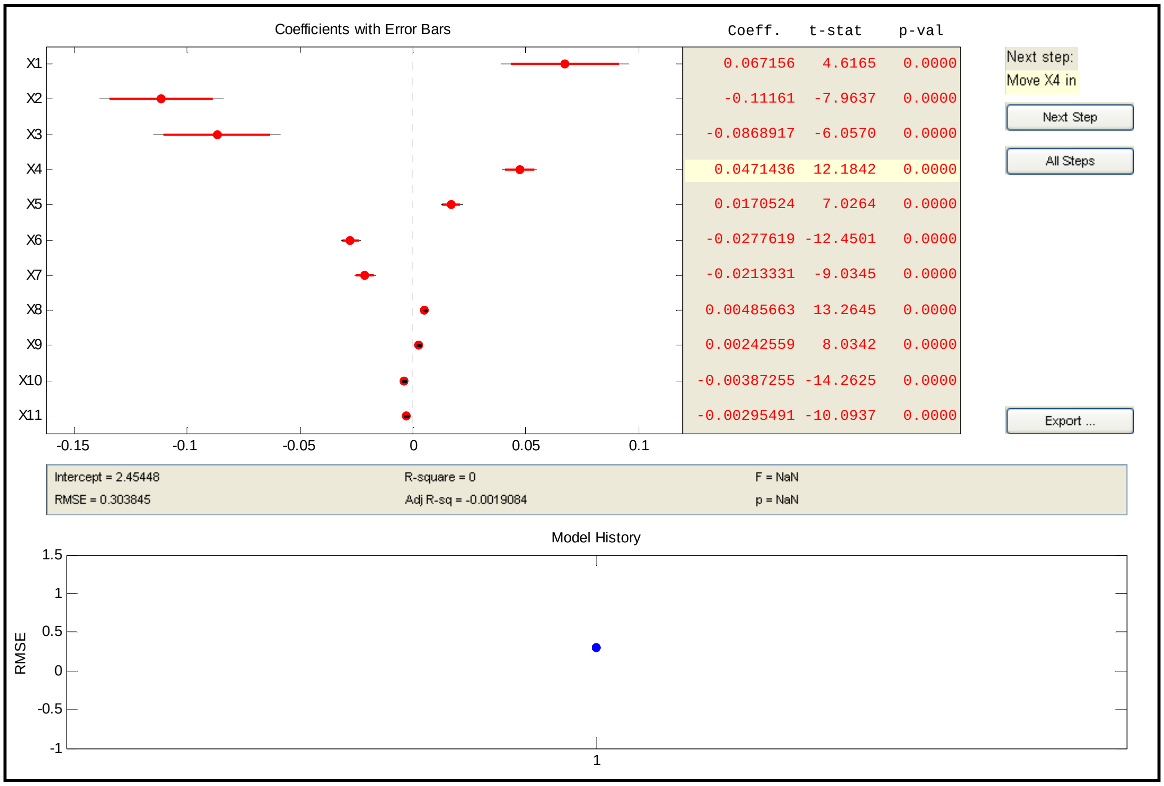


Figure 4-10 Outputs of the Matlab stepwise command.

After the elimination of some terms, the final results using the monomial basis representation model are obtained. The final results are shown in Table 4-5 and in Figure 4-11.

Table 4-5 Final Matlab Output for the Theoretical Case Study Using the Basis Representation Model with a Monomial Basis (K =2)

```
%--------------------------------------------------------------------------%
  Results for the Analysis of Variance for the Basis Representation Model
        Model with the IN variables after the Stepwise Procedure
%--------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%--------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.3266      0.015659   148.5765   0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    2.0000    0.0453    0.0095    4.7834    0.0000    2.1870
    4.0000   -0.0210    0.0073   -2.8879    0.0040   14.7300
    8.0000    0.0068    0.0007    9.7201         0   14.7300
    9.0000    0.0024    0.0001   17.5755         0    1.0000
   10.0000   -0.0046    0.0002  -22.4980         0    2.1870
   11.0000   -0.0030    0.0001  -21.4108         0    1.0000

%--------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%--------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.81577  0.81364
%--------------------------------------------------------------------------%
                      Analysis of Variance
%--------------------------------------------------------------------------%
Source            DF      SS         MS           F        P
Regression        6     39.4642      6.5774     382.2813     0
Residual Error   518     8.9125      0.017206
Total            524    48.3767
%--------------------------------------------------------------------------%
```

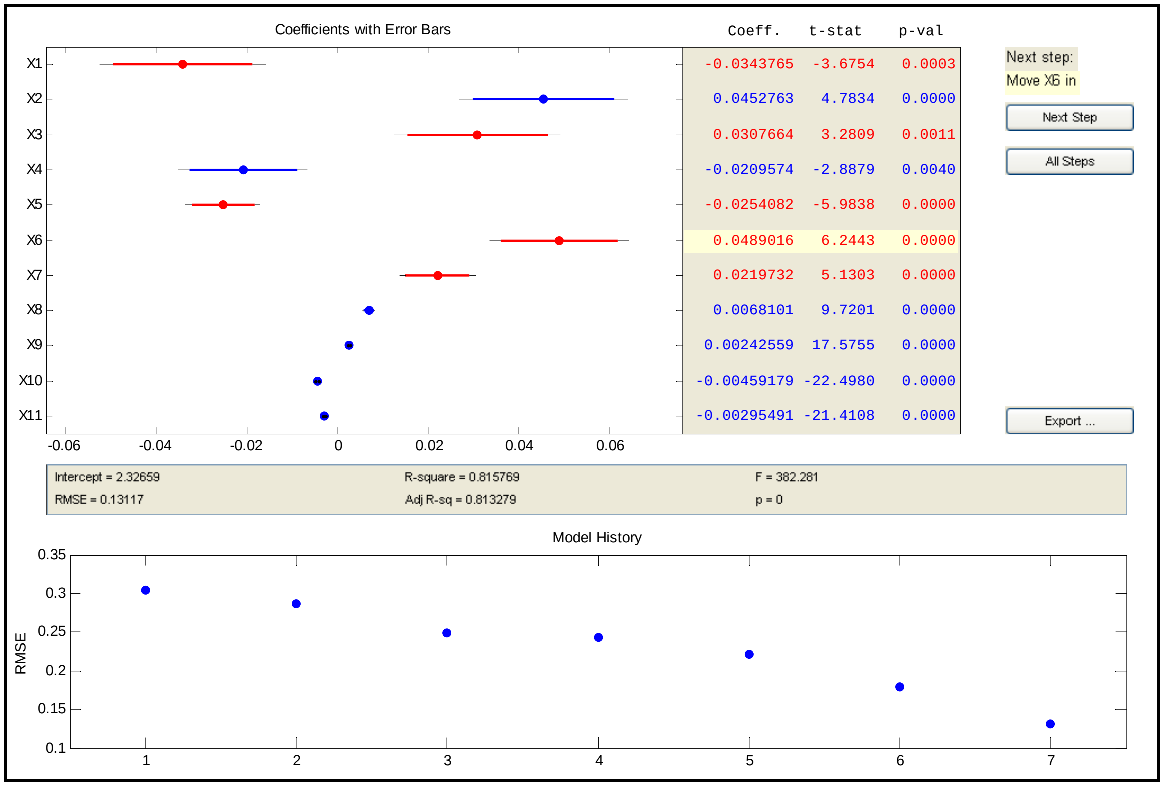


Figure 4-11 Final results for the Monomial basis representation model.

After the elimination of some terms, the multicollinearity problem was almost eliminated. Considering the problem presented by the model, some observations can be realized. The terms 2, 4, 8, 9, 10, and 11 are the coefficients relevant in the model. The Table 4-6 summarizes the results previously mentioned and shows the factors that correspond to each of the coefficients that stayed in the model.

Table 4-6 Results summary for Monomial basis representation for the theoretical case

Terms	Coefficient Represented	Value	Related Factor	Basis Term
2	b_{20}	0.0453	X2	1
4	b_{01}	-0.0210	1	t
8	b_{02}	0.0068	1	t^2
9	b_{12}	0.0024	X1	t^2
10	b_{22}	-0.0046	X2	t^2
11	b_{32}	-0.0030	X1*X2	t^2

Using the results presented in the Table 4-6 some inferences can be done. The more relevant terms corresponds to both factors and their interaction. These results are acceptable because most of the multicollinearity problem was solved using the stepwise procedure. The first model had a maximum Variance Inflation Factor (VIF) over 50 and the final model has a maximum VIF of 15. In addition, the adjusted determination coefficient is over 80%, which is an acceptable number for that adequacy measure. The graphs for the estimated responses and residuals are shown in the next two figures.

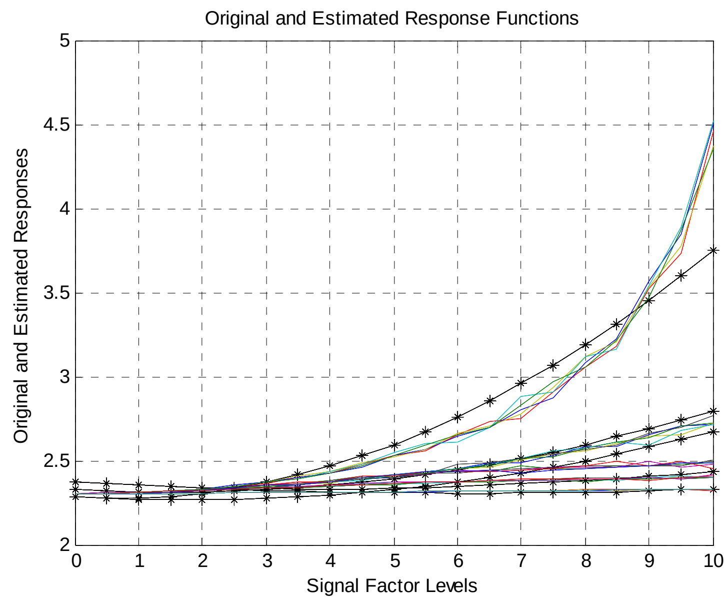


Figure 4-12 Responses and estimated responses for the Basis Representation Final Model with a Monomial Basis for the Theoretical Case Study. The asterisks correspond to the estimated functions.

From Figure 4-12 one can observe that estimated responses are adequate for most of the experimental conditions. The residuals plot presented in Figure 4-13, shows that the model does not obey the stochastic assumption for the regression model of constant variance.

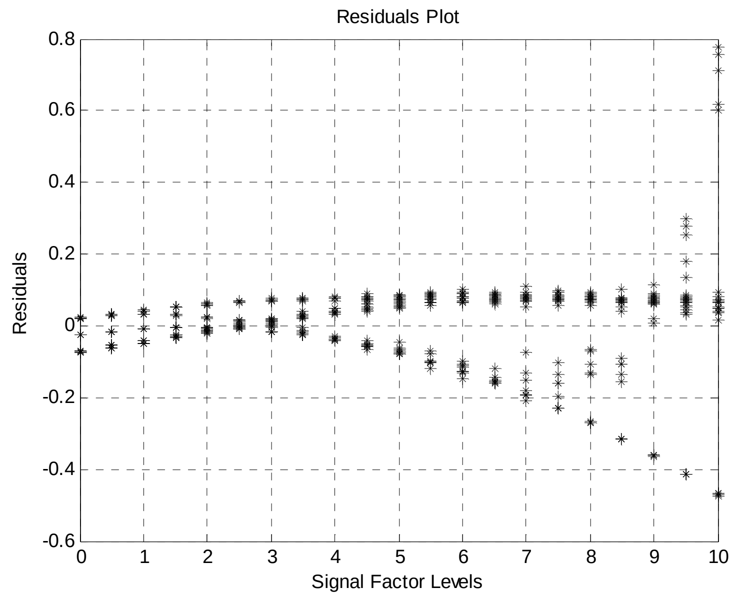


Figure 4-13 Residuals for the Basis Representation Final Model with a Monomial Basis for the Theoretical Case Study

The Table 4-7 presents the pseudo-MSE that is going to be used to compare this procedure with the other methods.

Table 4-7 Pseudo-MSE for the theoretical case using the basis representation method with a Monomial basis

Total SSE	Total Number of Parameters	Total Number of Data Points	Pseudo MSE
8.9124	7	525	0.0172

The next section have the results obtained using a Fourier Basis Representation

4.3.2.2 Fourier Basis

The general Fourier series representation has the following form

$$y(t) = c_0 + c_1 \sin(\omega t) + c_2 \cos(\omega t) + c_3 \sin(2\omega t) + c_4 \cos(2\omega t) + \ldots \qquad (4.4)$$

where $\omega = \dfrac{2\pi}{\mathrm{T}}$ and T is the highest level of the signal factor. The model used for the experiment was the following

$$y(t) = (b_{00} + b_{10}x_1 + b_{20}x_2 + b_{30}x_1x_2) + (b_{01} + b_{11}x_1 + b_{21}x_2 + b_{31}x_1x_2)\sin(wt) + \\ (b_{02} + b_{12}x_1 + b_{22}x_2 + b_{32}x_1x_2)\cos(wt) + (b_{03} + b_{13}x_1 + b_{23}x_2 + b_{33}x_1x_2)\sin(2wt) + \\ (b_{04} + b_{14}x_1 + b_{24}x_2 + b_{34}x_1x_2)\cos(2wt) \quad (4.5)$$

This type of basis is widely used when the data shows a periodical behavior. In this case, there is no periodical behavior but the basis was used just to illustrate the procedure. As done in the previous section the following table summarizes the terms of the model with the Matlab output and factors related with each term.

Table 4-8 Relation between the Matlab's output column "Term" and the coefficients of Equation 4.5

Terms	Coefficient Represented	Related Factor	Basis Term
0	b_{00}	1	1
1	b_{10}	X1	1
2	b_{20}	X2	1
3	b_{30}	X1*X2	1
4	b_{01}	1	sin(wt)
5	b_{11}	X1	sin(wt)
6	b_{21}	X2	sin(wt)
7	b_{31}	X1*X2	sin(wt)
8	b_{02}	1	cos(wt)
9	b_{12}	X1	cos(wt)
10	b_{22}	X2	cos(wt)
11	b_{32}	X1*X2	cos(wt)
12	b_{03}	1	sin(2wt)
13	b_{13}	X1	sin(2wt)
14	b_{23}	X2	sin(2wt)
15	b_{33}	X1*X2	sin(2wt)
16	b_{04}	1	cos(2wt)
17	b_{14}	X1	cos(2wt)
18	b_{24}	X2	cos(2wt)
19	B_{34}	X1*X2	cos(2wt)

The Matlab application created for the Basis Representation Model was used setting the basis to be a Fourier one. The initial output produced by the application is shown next.

Table 4-9 Initial Matlab Output for the Theoretical Case Study Using the Basis Representation Model with a Fourier Basis (K =2)

```
%------------------------------------------------------------------------%
   Results for the Analysis of Variance for the Basis Representation Model
%------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.4499      0.008836   277.2606   0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    1.0000    0.0626    0.0099    6.3336    0.0000    1.0080
    2.0000   -0.1055    0.0099  -10.6832         0    1.0080
    3.0000   -0.0824    0.0099   -8.3365    0.0000    1.0080
    4.0000   -0.1291    0.0128  -10.1228         0    1.0000
    5.0000   -0.0584    0.0143   -4.0934    0.0000    1.0000
    6.0000    0.1058    0.0143    7.4202    0.0000    1.0000
    7.0000    0.0773    0.0143    5.4196    0.0000    1.0000
    8.0000    0.0655    0.0122    5.3576    0.0000    1.0076
    9.0000    0.0650    0.0137    4.7495    0.0000    1.0120
   10.0000   -0.0900    0.0137   -6.5777    0.0000    1.0120
   11.0000   -0.0625    0.0137   -4.5666    0.0000    1.0120
   12.0000   -0.0730    0.0128   -5.7238    0.0000    1.0000
   13.0000   -0.0431    0.0143   -3.0232    0.0026    1.0000
   14.0000    0.0677    0.0143    4.7450    0.0000    1.0000
   15.0000    0.0506    0.0143    3.5464    0.0004    1.0000
   16.0000    0.0310    0.0122    2.5359    0.0115    1.0076
   17.0000    0.0314    0.0137    2.2929    0.0223    1.0120
   18.0000   -0.0375    0.0137   -2.7437    0.0063    1.0120
   19.0000   -0.0328    0.0137   -2.3975    0.0169    1.0120

%------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.57551  0.55954
%------------------------------------------------------------------------%
                      Analysis of Variance
%------------------------------------------------------------------------%
Source            DF          SS          MS         F         P
Regression        19     27.8412       1.4653     36.0346     0
Residual Error   505     20.5355     0.040664
Total            524     48.3767
%------------------------------------------------------------------------%
```

From Table 4-9 one can observe that this model does not have the multicollinearity problem. In addition, the regression model is significant and has an adjusted determination coefficient close to a 56%, which is not acceptable in many applications. In this case, the stepwise procedure was not necessary because all the terms were relevant in the model and the multicollinearity problem was not present.

Table 4-10 Results Summary for Fourier basis representation for the theoretical case

Terms	Coefficient Represented	Value	Related Factor	Basis Term
0	b_{00}	2.4498	1	1
1	b_{10}	0.0626	X1	1
2	b_{20}	-0.1055	X2	1
3	b_{30}	-0.0824	X1*X2	1
4	b_{01}	-0.1291	1	sin(wt)
5	b_{11}	-0.0584	X1	sin(wt)
6	b_{21}	0.1058	X2	sin(wt)
7	b_{31}	0.0773	X1*X2	sin(wt)
8	b_{02}	0.0655	1	cos(wt)
9	b_{12}	0.065	X1	cos(wt)
10	b_{22}	-0.09	X2	cos(wt)
11	b_{32}	-0.0625	X1*X2	cos(wt)
12	b_{03}	-0.073	1	sin(2wt)
13	b_{13}	-0.0431	X1	sin(2wt)
14	b_{23}	0.0677	X2	sin(2wt)
15	b_{33}	0.0506	X1*X2	sin(2wt)
16	b_{04}	0.031	1	cos(2wt)
17	b_{14}	0.0314	X1	cos(2wt)
18	b_{24}	-0.0375	X2	cos(2wt)
19	b_{34}	-0.0328	X1*X2	cos(2wt)

The model with the Fourier expansion indicates that both factors and the interaction are relevant in the model. However, this model is not very reliable because can explain only the 50% of the variability of the data. The next plots correspond to the estimated responses and the residuals.

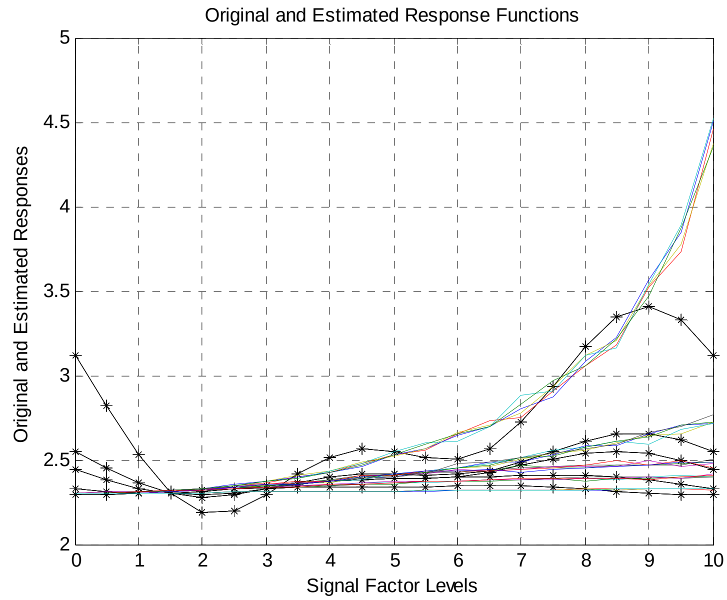


Figure 4-14 Responses and estimated responses for the Basis Representation Final Model with a Fourier Basis for the Theoretical Case Study. The asterisks correspond to the estimated functions.

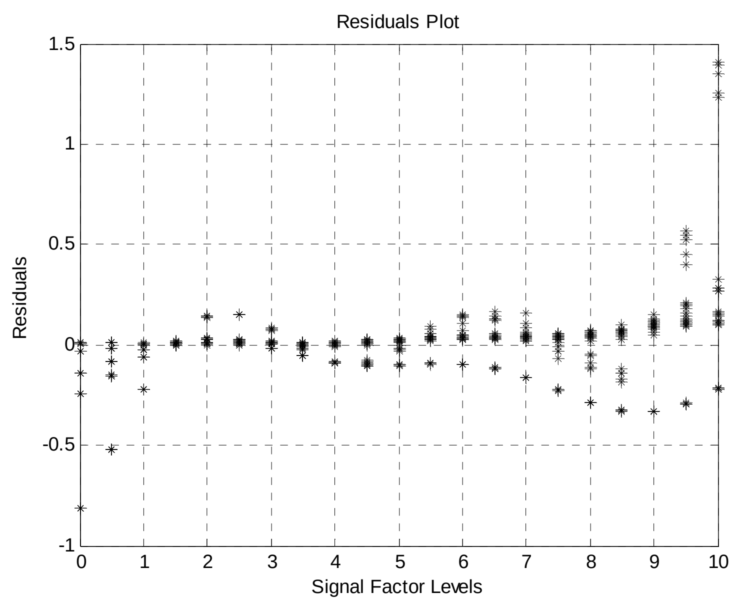


Figure 4-15 Residuals for the Basis Representation Final Model with a Fourier Basis for the Theoretical Case Study

From Figure 4-14, one can observe the lack of fit of the obtained model. The residuals plot shows the non-constant variance presented by the model. Table 4-11 presents the estimated pseudo-MSE.

Table 4-11 Pseudo-MSE for the theoretical case basis using the representation Method with a Fourier basis

Total SSE	Total Number of Parameters	Total Number of Data Points	Pseudo MSE
20.5355	20	525	0.0406

The next section shows the last type of basis selected which is a Cubic Spline.

4.3.2.3 Cubic Spline Basis

The general cubic spline basis follows the form

$$y(t) = c_{00} + c_{01}t + c_{02}t^2 + c_{03}t^3 + c_1(t-\tau_1)_+^3 + c_2(t-\tau_2)_+^3 \ldots c_k(t-\tau_k)_+^3 \qquad (4.6)$$

where k is the number of basis functions and $(t-\tau_i)_+^3$ is defined if $t-\tau_i$>0 otherwise is zero; τ represents the knots selected for the spline.

The model selected for the experiment was the following

$$\begin{aligned} y(t) &= (b_{00} + b_{10}x_1 + b_{20}x_2 + b_{30}x_1x_2) + (b_{01} + b_{11}x_1 + b_{21}x_2 + b_{31}x_1x_2)t + \\ &(b_{02} + b_{12}x_1 + b_{22}x_2 + b_{32}x_1x_2)t^2 + (b_{03} + b_{13}x_1 + b_{23}x_2 + b_{33}x_1x_2)t^3 + \\ &(b_{04} + b_{14}x_1 + b_{24}x_2 + b_{34}x_1x_2)(t-\tau_1)_+^3 + (b_{05} + b_{15}x_1 + b_{25}x_2 + b_{35}x_1x_2)(t-\tau_2)_+^3 + \\ &(b_{06} + b_{16}x_1 + b_{26}x_2 + b_{36}x_1x_2)(t-\tau_3)_+^3 \end{aligned} \qquad (4.7)$$

The selected knots for this model are shown in Table 4-12. The values presented were selected looking for levels of the signal factor where the response started to change dramatically. This selection could have a great impact in the performance of the models.

Table 4-12 Knots Location for Cubic Spline Basis

τ_1	2
τ_2	6
τ_3	8

The relation between the terms and the Equation 4.7 are shown in Table 4-13.

Table 4-13 Relation between the Matlab's output column "Term" and the coefficients of Equation 4.7

Terms	Coefficient Represented	Related Factor	Basis Term
0	b_{00}	1	1
1	b_{10}	X1	1
2	b_{20}	X2	1
3	b_{30}	X1*X2	1
4	b_{01}	1	t
5	b_{11}	X1	t
6	b_{21}	X2	t
7	b_{31}	X1*X2	t
8	b_{02}	1	t^2
9	b_{12}	X1	t^2
10	b_{22}	X2	t^2
11	b_{32}	X1*X2	t^2
12	b_{03}	1	t^3
13	b_{13}	X1	t^3
14	b_{23}	X2	t^3
15	b_{33}	X1*X2	t^3
16	b_{04}	1	$(t-2)^3$
17	b_{14}	X1	$(t-2)^3$
18	b_{24}	X2	$(t-2)^3$
19	b_{34}	X1*X2	$(t-2)^3$
20	b_{05}	1	$(t-6)^3$
21	b_{15}	X1	$(t-6)^3$
22	b_{25}	X2	$(t-6)^3$
23	b_{35}	X1*X2	$(t-6)^3$
24	b_{06}	1	$(t-8)^3$
25	b_{16}	X1	$(t-8)^3$
26	b_{26}	X2	$(t-8)^3$
27	b_{36}	X1*X2	$(t-8)^3$

The initial output for the cubic spline it is shown in Table 4-14.

Table 4-14 Initial Matlab Output for the Theoretical Case Study Using the Basis Representation Model with a Cubic Spline Basis

```
The model is not a good one due a multicollinearity problem
%---------------------------------------------------------------------------%
 Results for the Analysis of Variance for the Basis Representation Model
%---------------------------------------------------------------------------%
Type of Basis:   Cubic_spline
Number of basis functions or knots if the basis is a Cubic Spline:   2 6 8
%---------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%---------------------------------------------------------------------------%
   Term  Intercept    SE Coef        T-test      P-value
    0     2.3017      0.020158      114.1847     0

   Term  Coefficients  SE Coef        T-test      P-value         VIF
    1    0.0021043     0.022537      0.09337      0.92565       19.223
    2  -0.00020305     0.022537   -0.0090096      0.99282       19.223
    3  -0.00016942     0.022537   -0.0075177      0.994         19.223
    4   0.00072865     0.055133     0.013216      0.98946       1318.2
    5   -0.0022255      0.06164    -0.036105      0.97121       4913.3
    6    0.0013893      0.06164     0.022539      0.98203       4913.3
    7    0.0058583      0.06164      0.09504      0.92432       4913.3
    8    0.0056588      0.03855      0.14679      0.88336        69141
    9     0.001777       0.0431      0.04123      0.96713  1.5121e+005
   10   -0.0011361       0.0431    -0.026359      0.97898  1.5121e+005
   11    -0.004196       0.0431    -0.097355      0.92248  1.5121e+005
   12  -0.00046446    0.0075006    -0.061924      0.95065  2.4526e+005
   13  -0.00031308    0.0083859    -0.037334      0.97023  4.2866e+005
   14 -1.3964e-005    0.0083859   -0.0016652      0.99867  4.2866e+005
   15   0.00039213    0.0083859      0.04676      0.96272  4.2866e+005
   16   0.00054622    0.0084327     0.064774      0.94838        77521
   17   0.00074768    0.0094281     0.079303      0.93682   1.183e+005
   18  -0.00056515    0.0094281    -0.059943      0.95223   1.183e+005
   19  -0.00048651    0.0094281    -0.051603      0.95887   1.183e+005
   20    0.0012965    0.0042662       0.3039      0.76133       235.11
   21   0.00067352    0.0047698      0.14121      0.88776       286.36
   22  -0.00041589    0.0047698    -0.087193      0.93055       286.36
   23   -0.0017558    0.0047698     -0.36812      0.71294       286.36
   24     0.012767     0.012917      0.98838      0.32345       25.925
   25     0.017029     0.014442       1.1791       0.2389       28.722
   26    -0.017546     0.014442      -1.2149      0.22499       28.722
   27    -0.014239     0.014442     -0.98593      0.32465       28.722

%---------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%---------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.88599  0.8798
%---------------------------------------------------------------------------%
                      Analysis of Variance
%---------------------------------------------------------------------------%
Source            DF          SS          MS      F         P
Regression        27     42.8614     1.5875     143.051     0
Residual Error   497     5.5153      0.011097
Total            524     48.3767
%--------------------------------------------------------------------------%
```

Table 4-14 has a warning indicating that there is multicollinearity problems present in the model, as the Variance Inflation Factors (VIF) are extremely high. Performing the same, procedure for variables elimination, a new and more reduce model is obtained. The Figure 4-16 shows the results. The following is the Table that corresponds to the final model.

Table 4-15 Final Matlab Output for the Theoretical Case Study Using the Basis Representation Model with a Cubic Spline Basis

```
%--------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
          Model with the IN variables after the Stepwise Procedure
%--------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%--------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.3212      0.0060996   380.5503   0

   Term    Coefficients    SE Coef      T-test    P-value     VIF
     12       0.000508     1.52E-05     33.407        0          1
     13       0.000136     2.76E-05      4.929    1.11E-06     4.594
     14       -0.00031     2.76E-05    -11.421        0        4.594
     15       -0.00023     2.76E-05    -8.5119    2.22E-16     4.594
     21       0.003977     0.000606     6.5598    1.31E-10     4.594
     22       -0.00373     0.000606    -6.1512    1.54E-09     4.594
     23       -0.00296     0.000606     -4.879    1.42E-06     4.594

%--------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%--------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.88057  0.87895
%--------------------------------------------------------------------------%
                      Analysis of Variance
%--------------------------------------------------------------------------%
Source            DF       SS          MS            F       P
Regression         7     42.5989     6.0856     544.5379     0
Residual Error   517      5.7778     0.011176
Total            524    48.3767
%--------------------------------------------------------------------------%
```

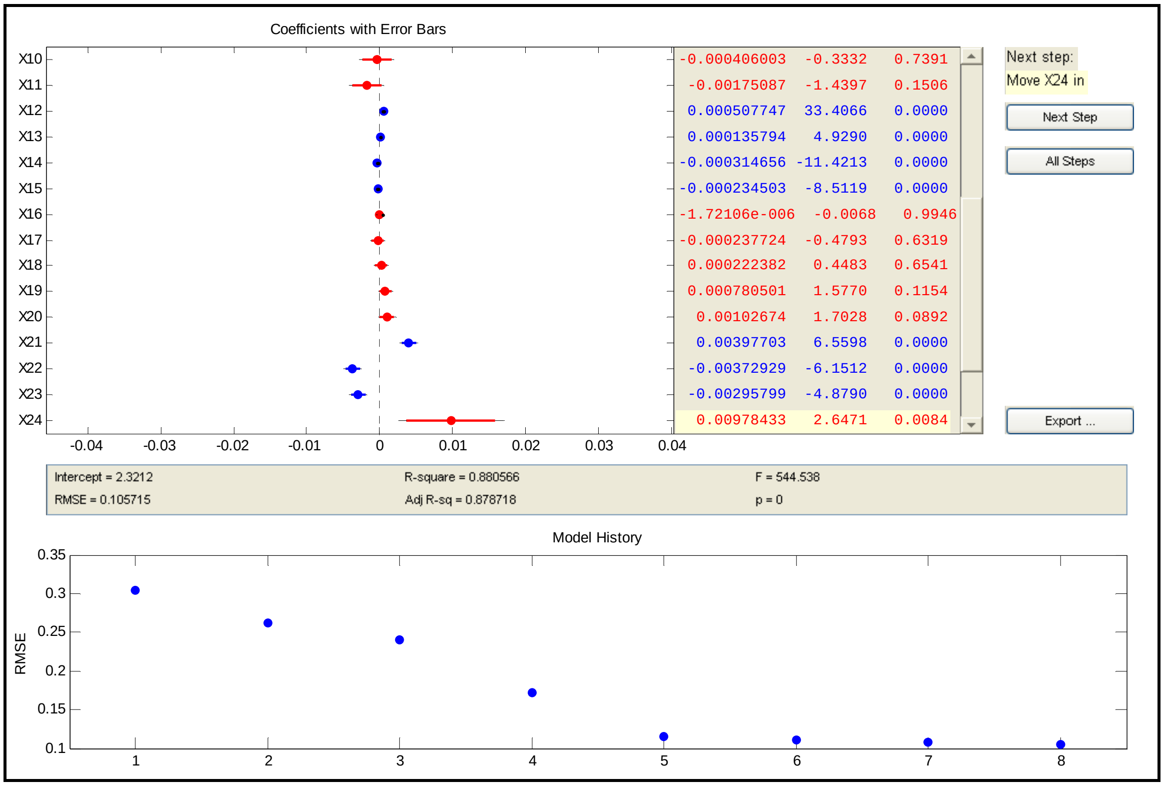


Figure 4-16 Final results for the Cubic Spline basis representation model.

The next figures correspond to the estimated responses and the residuals for the final model obtained.

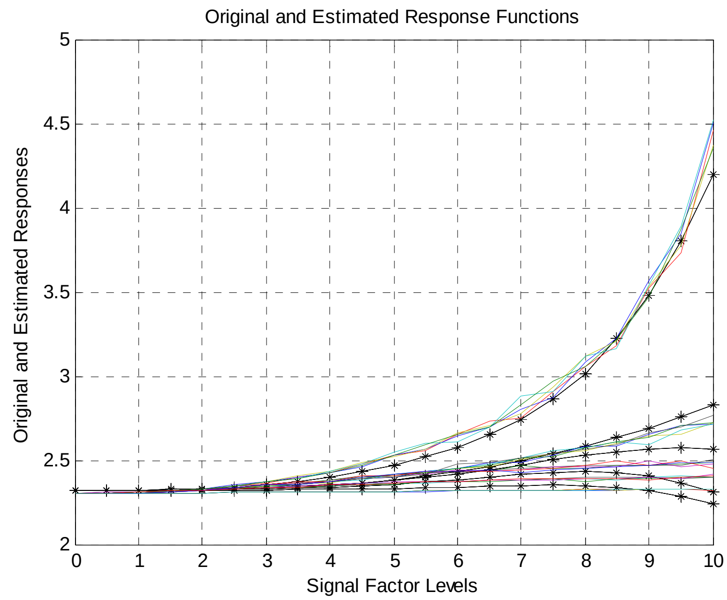


Figure 4-17 Responses and estimated responses for the Basis Representation Final Model with a Cubic Spline Basis for the Theoretical Case Study. The asterisks correspond to the estimated functions

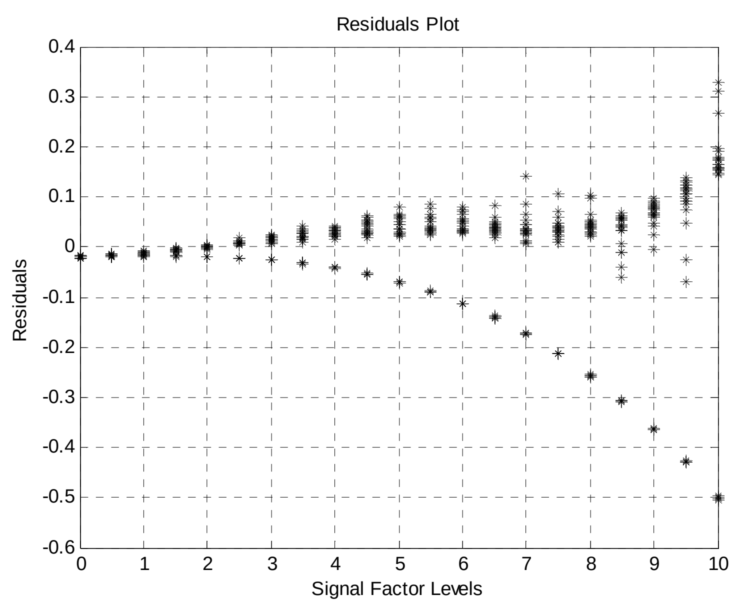


Figure 4-18 Residuals for the Basis Representation Final Model with a Cubic Spline Basis for the Theoretical Case Study

From Figure 4-17 it is possible to observe the fit of the estimated functions. The estimation using the Cubic Spline looks better in comparison with the two previous bases. Table 4-16 shows the terms that were considered relevant and the factors that are related to those terms.

Table 4-16 Results Summary for Cubic Spline basis representation for the theoretical case

Terms	Coefficient Represented	Value	Related Factor	Basis Term
12	b_{03}	0.000508	1	t^3
13	b_{13}	0.000136	X1	t^3
14	b_{23}	-0.00031	X2	t^3
15	b_{33}	-0.00023	X1*X2	t^3
21	b_{15}	0.003977	X1	$(t-6)^3$
22	b_{25}	-0.00373	X2	$(t-6)^3$
23	b_{35}	-0.00296	X1*X2	$(t-6)^3$

These terms confirm once again that factors x_1 and x_2, and their interaction are relevant in the experiment. This model only has 8 terms and a good adjusted determination coefficient, over 87%. Table 4-17 presents the pseudo MSE as in the previous sections.

Table 4-17 Pseudo-MSE for the theoretical case basis using the representation method with a Cubic Spline basis

Total SSE	Total Number of Parameters	Total Number of Data Points	Pseudo MSE
5.7778	8	525	0.0111

The total SSE of this model is almost equal to the total SSE of the Point-Wise method, but the difference in the number of terms is dramatic. In the next section, the Piece-Wise method is applied to this Theoretical case study and their results and discussion are shown.

4.3.3 Piece-Wise Method without the Continuity Constraint

4.3.3.1 Monomial Basis

The models that are going to be used in each region are identical to the Equation 4.3. In order to complete, this procedure it is necessary to find the combination of knots that provides the minimal sum of squares of error for all the regions. The objective function and the restrictions considered for optimization were presented on Chapter 3. The equation that corresponds to the objective function is the Equation 3.7, and the restrictions are represented by equations 3.8 to 3.12. The Table 4-18 shows the results obtained for this case.

Table 4-18 Summary of the Knots Search for the Theoretical Case Study using a monomial basis

Monomial Basis K=2

Number of knots	Objective function	optimal knots				
5	5.4869	5	9	12	15	18
4	5.4872	9	12	15	18	*
3	**5.4879**	**10**	**15**	**18**	*	*
2	5.4929	10	18	*	*	*
1	5.5294	16	*	*	*	*

Due the small difference that exists among the objective functions for the models with 3, 4, and 5 knots the one with the smaller number of knots was selected. In other words, the knots selected for the procedure were 10, 15, and 18. The plots that correspond to each region are shown in Figure 4-19. This figure shows the different behaviors of the response functions in each region.

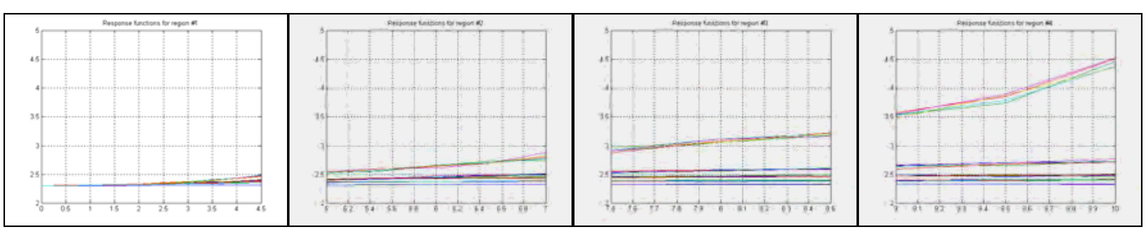


Figure 4-19 Plots for all the regions delimited before using the Piece-Wise procedure for the theoretical case study with a monomial basis

The results for this procedure are shown in the next pages. The results are presented region by region. At the beginning of each section, the tables that correspond to the initial outputs of Matlab are shown. Then the stepwise plots and final model details are shown. To conclude, the results are discussed and the relevant factors for each region are presented.

4.3.3.1.1 Results for Region #1

Table 4-19 is the initial Matlab output for the procedure in this region

Table 4-19 Initial Output for the Piece-Wise Method using a Monomial basis for region #1

```
%------------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for the
region #1
%------------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%------------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%------------------------------------------------------------------------------%
   Terms   Coefficients   SE Coef     T-test       P-value         VIF
     0       2.3013    0.0028939      795.25            0             0
     1    0.0024691    0.0032354     0.76314      0.44614       6.1818
     2   -0.0012226    0.0032354    -0.37788      0.70586       6.1818
     3   -0.0001135    0.0032354   -0.035081      0.97204       6.1818
     4      0.00327     0.002995      1.0918      0.27601       13.656
     5   -0.0021578    0.0033485    -0.64442      0.51993       47.176
     6    0.0041098    0.0033485      1.2274       0.2209       47.176
     7    0.0041849    0.0033485      1.2498      0.21261       47.176
     8    0.0034566   0.00064071      5.3948  1.6537e-007       13.656
     9   0.00087436   0.00071634      1.2206      0.22345        29.04
    10   -0.0021076   0.00071634     -2.9421    0.0035816        29.04
    11   -0.0024629   0.00071634     -3.4381   0.00069138        29.04

%------------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%------------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.76278  0.75182
%------------------------------------------------------------------------------%
                      Analysis of Variance
%------------------------------------------------------------------------------%
Source           DF          SS          MS          F       P
Regression        11     0.25919     0.023563     69.5728     0
Residual Error   238     0.080605    0.00033867
Total            249     0.33979
%------------------------------------------------------------------------------%
```

The model has a serious multicollinearity problem. In order to solve this problem and reduce the model a stepwise procedure was performed. The following figure shows the stepwise procedure conducted.

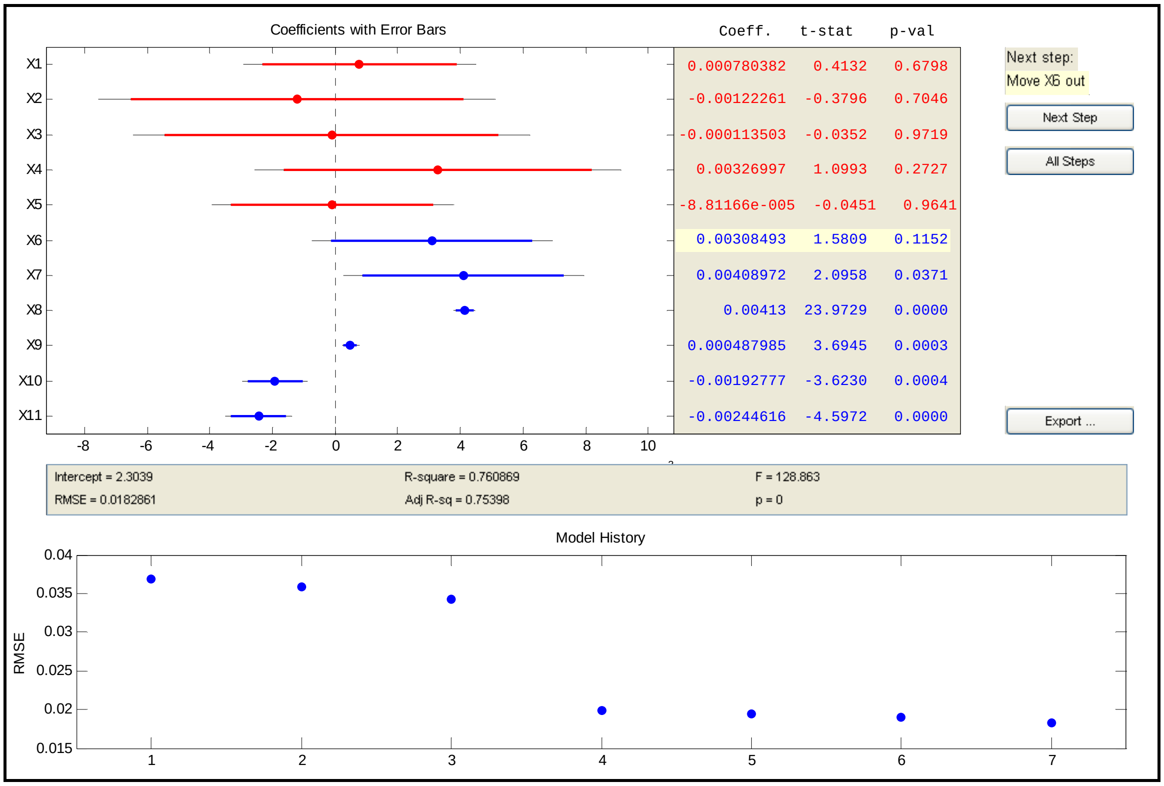


Figure 4-20 Final results for the Piece-Wise method using a monomial basis for region #1.

Table 4-20 correspond to the final model after the completion of the stepwise procedure

Table 4-20 Final Output for the Piece-Wise Method using a Monomial basis for region #1

```
%------------------------------------------------------------------------------%
   Results for the Analysis of Variance for the Basis Representation Model
%------------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%------------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
   0     2.3039      0.0016865   1366.0952   0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    6.0000    0.0031    0.0020    1.5809    0.1152   16.2281
    7.0000    0.0041    0.0020    2.0958    0.0371   16.2281
    8.0000    0.0041    0.0002   23.9729         0    1.0000
    9.0000    0.0005    0.0001    3.6945    0.0003    1.0000
   10.0000   -0.0019    0.0005   -3.6230    0.0004   16.2281
   11.0000   -0.0024    0.0005   -4.5972    0.0000   16.2281

%------------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%------------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.76087  0.75496
%------------------------------------------------------------------------------%
                      Analysis of Variance
%------------------------------------------------------------------------------%
Source            DF       SS          MS            F       P
Regression         6     0.25854     0.04309     128.8631     0
Residual Error   243    0.081255    0.00033438
Total            249    0.33979
%------------------------------------------------------------------------------%
```

The model still having the multicollinearity problem but is not as serious as the initial one. Table 4-21 shows the most relevant terms and the factors that are associated with the coefficients.

Table 4-21 Results summary for Region #1

Terms	**Coefficient Represented**	**Value**	**Related Factor**	**Basis Term**
6	b_{21}	0.0031	X2	t
7	b_{31}	0.0041	X1*X2	t
8	b_{02}	0.0041	1	t^2
9	b_{12}	0.0005	X1	t^2
10	b_{22}	-0.00019	X2	t^2
11	b_{32}	-0.0024	X1*X2	t^2

The main factors and their interaction are relevant in the experiment. The following pages show the results for the second region. The next figures correspond to the estimated responses and the residuals. Figure 4-21 shows the estimated responses that look adequate for this region and Figure 4-22 , shows the non-constant variance problem presented by this model.

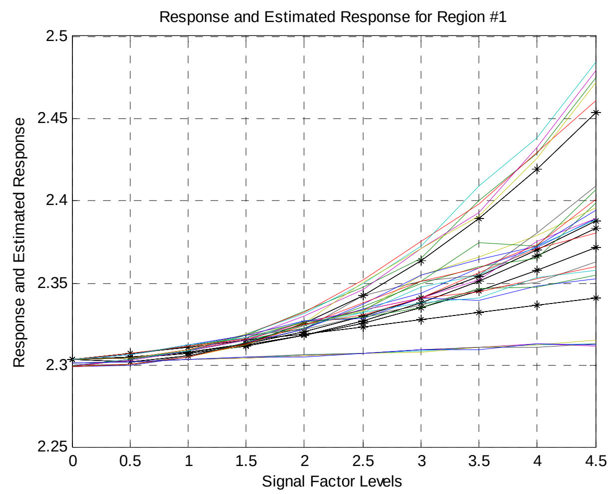


Figure 4-21 Response and estimated response for the region #1. The asterisks correspond to the estimated functions.

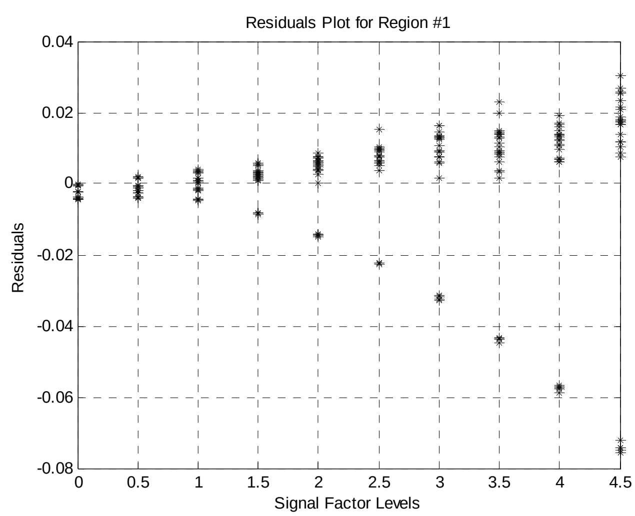


Figure 4-22 Residuals plot for region #1

4.3.3.1.2 Results for Region #2

In the same way as done with the previous region, region #2 was also analyzed in detail. The initial Matlab output it is shown in Table 4-22.

Table 4-22 Initial Output for the Piece-Wise Method using a Monomial basis for region #2

```
%-----------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for
the region #2
%-----------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
  Terms   Coefficients   SE Coef      T-test       P-value        VIF
    0       2.3847      0.53493        4.458  1.9605e-005           0
    1       0.17257     0.59807      0.28855      0.77345       7274.4
    2      -0.20932     0.59807     -0.34999        0.727       7274.4
    3      -0.11727     0.59807     -0.19608       0.8449       7274.4
    4     -0.023315     0.18013     -0.12944      0.89724       412.43
    5     -0.066486     0.20139     -0.33013      0.74191        30107
    6      0.085584     0.20139      0.42497      0.67167        30107
    7      0.042236     0.20139      0.20972      0.83426        30107
    8      0.005523    0.014993      0.36838      0.71328       412.43
    9     0.0071735    0.016762      0.42796       0.6695       8025.3
   10     -0.010418    0.016762     -0.62149      0.53553       8025.3
   11    -0.0054877    0.016762     -0.32738      0.74398       8025.3

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.70997  0.68173
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF          SS        MS         F          P
Regression        11     1.3601     0.12365     25.1465     0
Residual Error   113     0.55562    0.004917
Total            124     1.9157
%-----------------------------------------------------------------------------%
```

The problem of multicollinearity is present in this model. In this region the problem is even higher than in the previous one. The stepwise procedure was completed in order to improve the obtained model. In Figure 4-23, the plots for this procedure are presented.

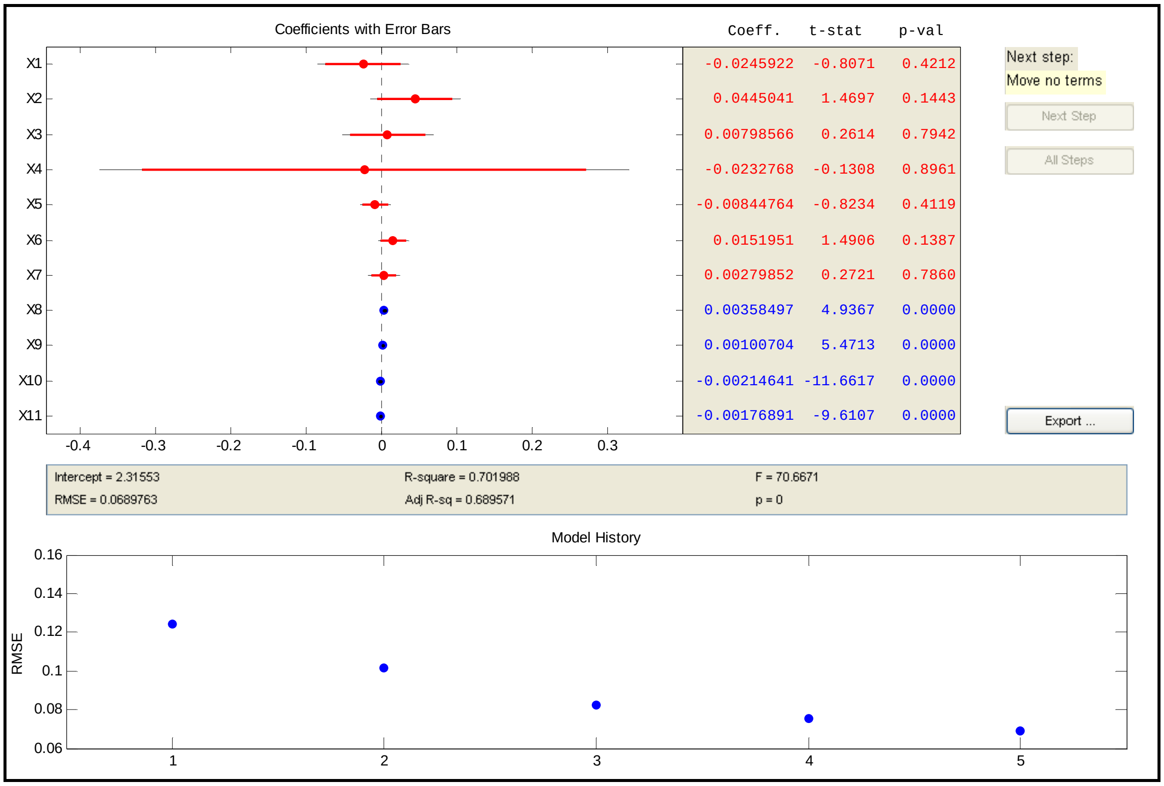


Figure 4-23 Final results for the Piece-Wise method using a monomial basis for region #2

The output for the final model it is shown in Table 4-23.

Table 4-23 Final Output for the Piece-Wise Method using a Monomial basis for region #2

```
%------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.3155      0.027215   85.0822   0

   Term  Coefficients  SE Coef    T-test   P-value       VIF
    8    0.0035848   0.00072621   4.9363   2.5973e-006     1
    9     0.001007   0.00018406   5.471    2.4845e-007     1
   10   -0.0021463   0.00018406   -11.661            0     1
   11   -0.0017689   0.00018406   -9.6105  2.2204e-016     1

%------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.70196  0.69203
%------------------------------------------------------------------------%
                      Analysis of Variance
%------------------------------------------------------------------------%
Source            DF          SS          MS          F      P
Regression        4     1.3448         0.33619     70.6577     0
Residual Error   120     0.57096      0.004758
Total            124     1.9157
%------------------------------------------------------------------------%
```

The results summary it is presented in Table 4-24

Table 4-24 Results summary for Region #2

Terms	**Coefficient Represented**	**Value**	**Related Factor**	**Basis Term**
8	b_{02}	0.0041	1	t^2
9	b_{12}	0.0005	X1	t^2
10	b_{22}	-0.00019	X2	t^2
11	b_{32}	-0.0024	X1*X2	t^2

In this case, the design matrix for this model it is orthogonal and this eases dramatically the analysis. Both factors are relevant, but the cross-term is the most relevant in this case. The following plots correspond to the estimated response and residuals. From Figure 4-24 one can observe that this model is not as good as the model of obtained for the first region. The residuals plot shows a better behavior in terms of the model variance if is compared to the first region model.

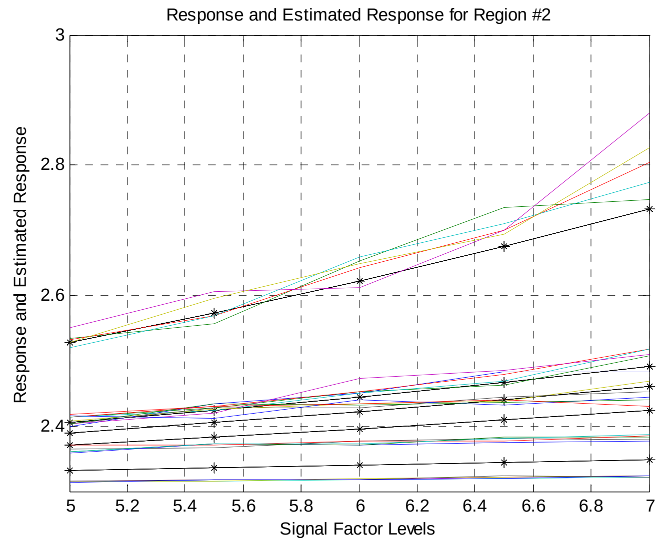


Figure 4-24 Response and estimated response for the region #2. The asterisks correspond to the estimated functions.

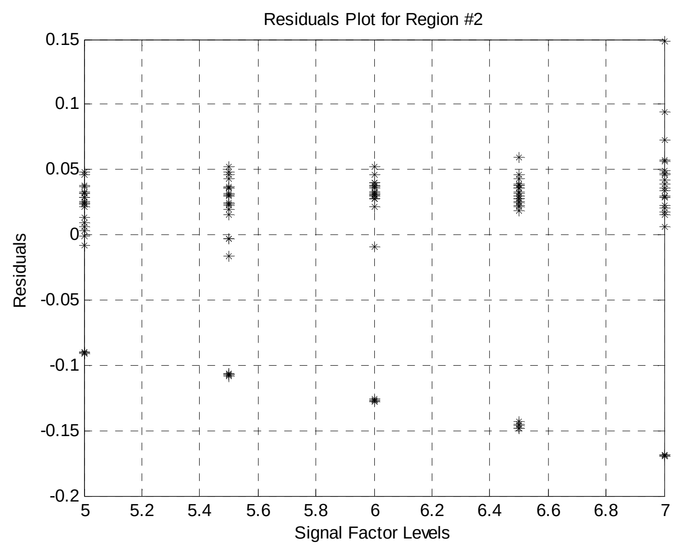


Figure 4-25 Residuals plot for region #2

The next section presents the results for the region #3.

4.3.3.1.3 Results for Region #3

The Matlab initial output for this region is presented in Table 4-25

Table 4-25 Initial Output for the Piece-Wise Method using a Monomial basis for region #3

```
%----------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for
the region #3
%----------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%----------------------------------------------------------------------------%
   Terms  Coefficients   SE Coef     T-test       P-value         VIF
    0       0.3001        8.246     0.036394      0.97108            0
    1      -1.9237       9.2193     -0.20866      0.83539  2.9376e+005
    2       2.3625       9.2193      0.25626      0.79859  2.9376e+005
    3       1.9939       9.2193      0.21627      0.82947  2.9376e+005
    4      0.49411       2.0659      0.23918      0.81174         3073
    5      0.45747       2.3097      0.19806      0.84363  1.1831e+006
    6     -0.56041       2.3097     -0.24263      0.80908  1.1831e+006
    7     -0.47699       2.3097     -0.20652      0.83705  1.1831e+006
    8     -0.02635       0.1291     -0.20411      0.83892         3073
    9      -0.0254      0.14433     -0.17598      0.86087  2.9952e+005
   10     0.029992      0.14433      0.20779      0.83606  2.9952e+005
   11     0.026175      0.14433      0.18135      0.85667  2.9952e+005

%----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.80091  0.76615
%----------------------------------------------------------------------------%
                      Analysis of Variance
%----------------------------------------------------------------------------%
Source            DF           SS          MS         F       P
Regression        11     4.3998     0.39998     23.0405    0
Residual Error   63     1.0937      0.01736
Total            74     5.4935
%----------------------------------------------------------------------------%
```

Once again, the initial model is not a good one due a serious multicollinearity problem. Figure 4-26 shows the plots that correspond to the stepwise procedure for this region.

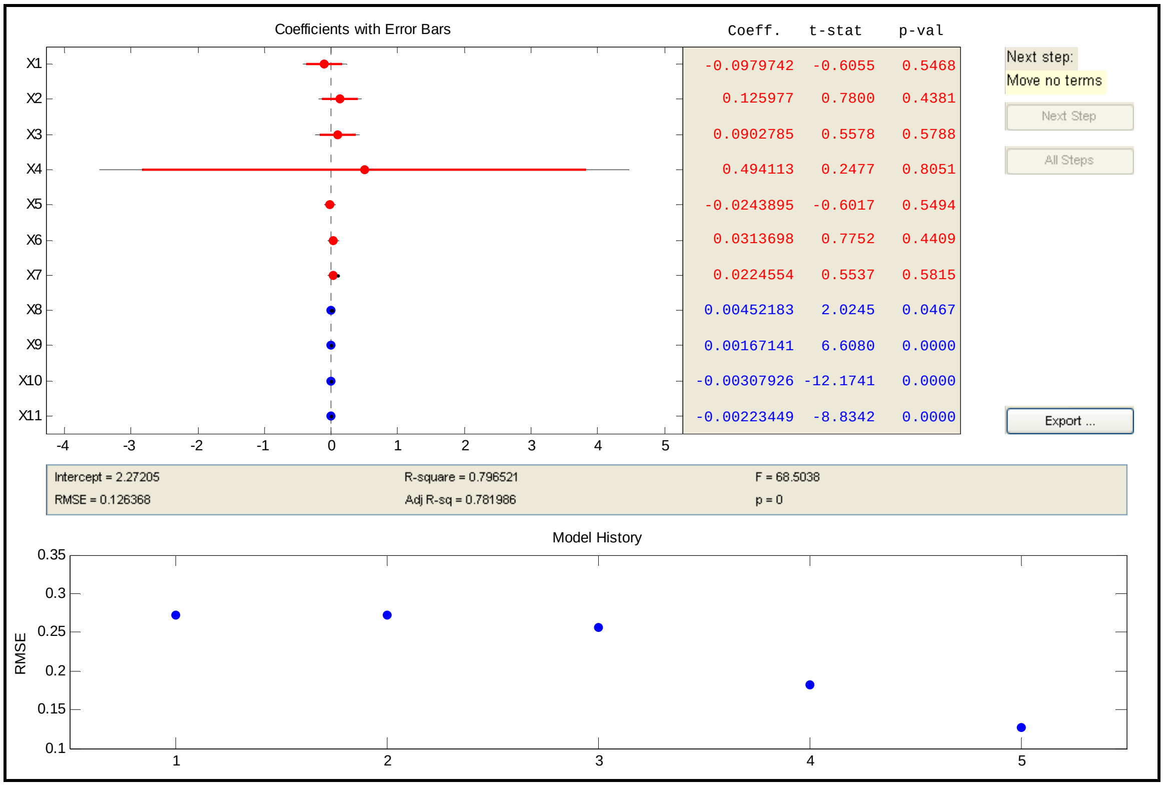


Figure 4-26 Final results for the Piece-Wise method using a monomial basis for region #3.

Table 4-26 corresponds to the final model obtained.

Table 4-26 Final Output for the Piece-Wise Method using a Monomial basis for region #3

```
%----------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%----------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.272      0.14406   15.7717   0

   Term   Coefficients   SE Coef   T-test    P-value      VIF
    8    0.0045218     0.0022335    2.0245   0.046732      1
    9    0.0016714     0.00025294   6.608    6.4236e-009   1
   10   -0.0030793     0.00025294   -12.174            0   1
   11   -0.0022345     0.00025294   -8.8342  5.3269e-013   1

%----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.79652  0.78489
%----------------------------------------------------------------------------%
                      Analysis of Variance
%----------------------------------------------------------------------------%
Source            DF          SS    MS           F         P
Regression         4     4.3757     1.0939     68.5038     0
Residual Error    70     1.1178     0.015969
Total             74     5.4935
%----------------------------------------------------------------------------%
```

In this model, the terms considered in the previous region were also considered. Once again, the final design matrix is orthogonal and the coefficients can be compared easily. Table 4-27 is a summary for these results.

Table 4-27 Results summary for Region #3

Terms	Coefficient Represented	Value	Related Factor	Basis Term
8	b_{02}	0.0045218	1	t^2
9	b_{12}	0.0016714	X1	t^2
10	b_{22}	-0.0030793	X2	t^2
11	b_{32}	-0.0022345	X1*X2	t^2

The factors and their interaction are relevant in this region. The term that has the most influence in the model corresponds to the factor x_2. The estimated response plot (Figure 4-27) shows a better fit for the responses compared to the previous two regions.

The residuals plot presents a better behavior in terms of the models variance with respect the models of the previous regions.

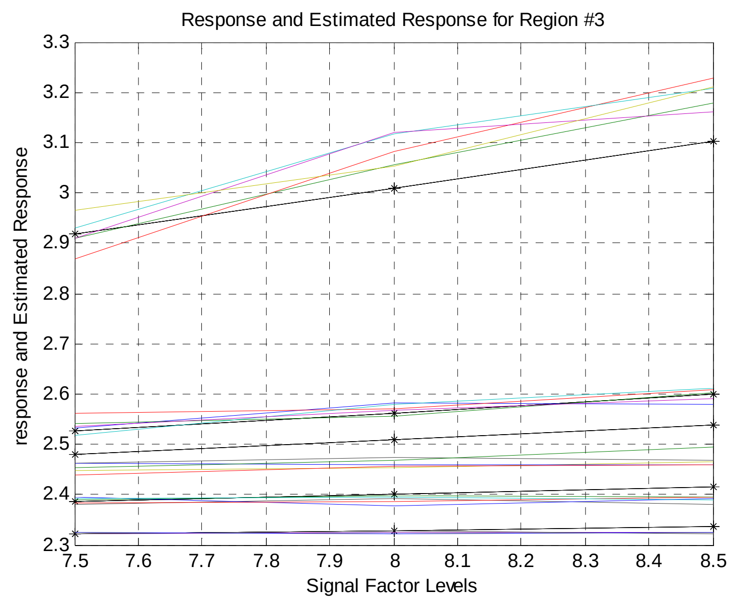


Figure 4-27 Response and estimated response for the region #3. The asterisks correspond to the estimated functions.

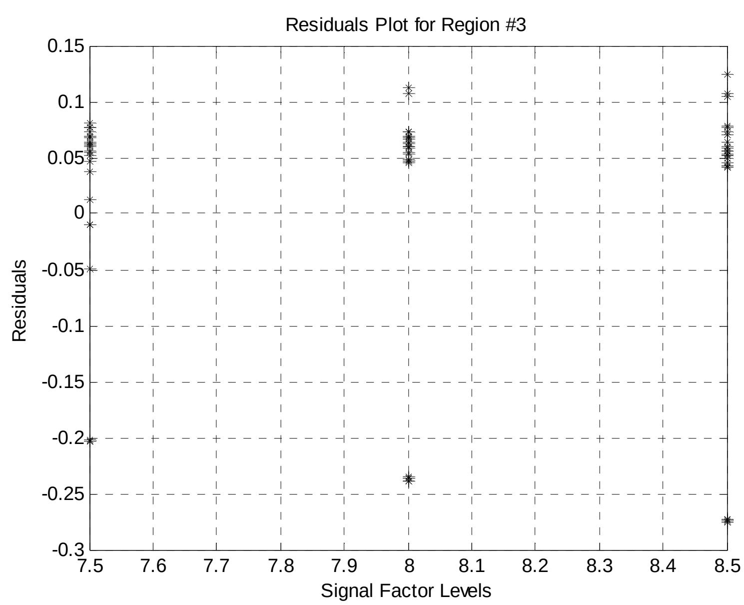


Figure 4-28 Residuals plot for region #3

The results for the fourth and last region are shown in the next section.

4.3.3.1.4 Results for Region #4

Table 4-28 corresponds to the initial model obtained for this region.

Table 4-28 Initial Output for the Piece-Wise Method using a Monomial basis for region #4

```
%-----------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for
the region #4
%-----------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
   Terms Coefficients   SE Coef      T-test       P-value       VIF
    0       12.001       21.567      0.55646      0.57987            0
    1       12.954       24.113      0.53725      0.59299  5.8482e+005
    2      -11.816       24.113     -0.49005       0.6258  5.8482e+005
    3      -13.317       24.113     -0.55228      0.58271  5.8482e+005
    4      -2.1542       4.5472     -0.47374      0.63732         4333
    5      -2.8794       5.0839     -0.56638      0.57315  2.3507e+006
    6       2.6506       5.0839      0.52137      0.60393  2.3507e+006
    7       2.9452       5.0839      0.57931      0.56445  2.3507e+006
    8      0.12418       0.2393      0.51893      0.60563         4333
    9       0.1625      0.26754      0.60736       0.5458  5.9295e+005
   10     -0.15262      0.26754     -0.57045       0.5704  5.9295e+005
   11     -0.16583      0.26754     -0.61981      0.53762  5.9295e+005

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.86992  0.84721
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF          SS          MS            F         P
Regression        11     25.1315     2.2847     38.3014     0
Residual Error   63     3.758      0.05965
Total            74     28.8895
%-----------------------------------------------------------------------------%
```

As in the previous regions, the model has a serious multicollinearity problem. The next page shows the plots for the stepwise procedure.

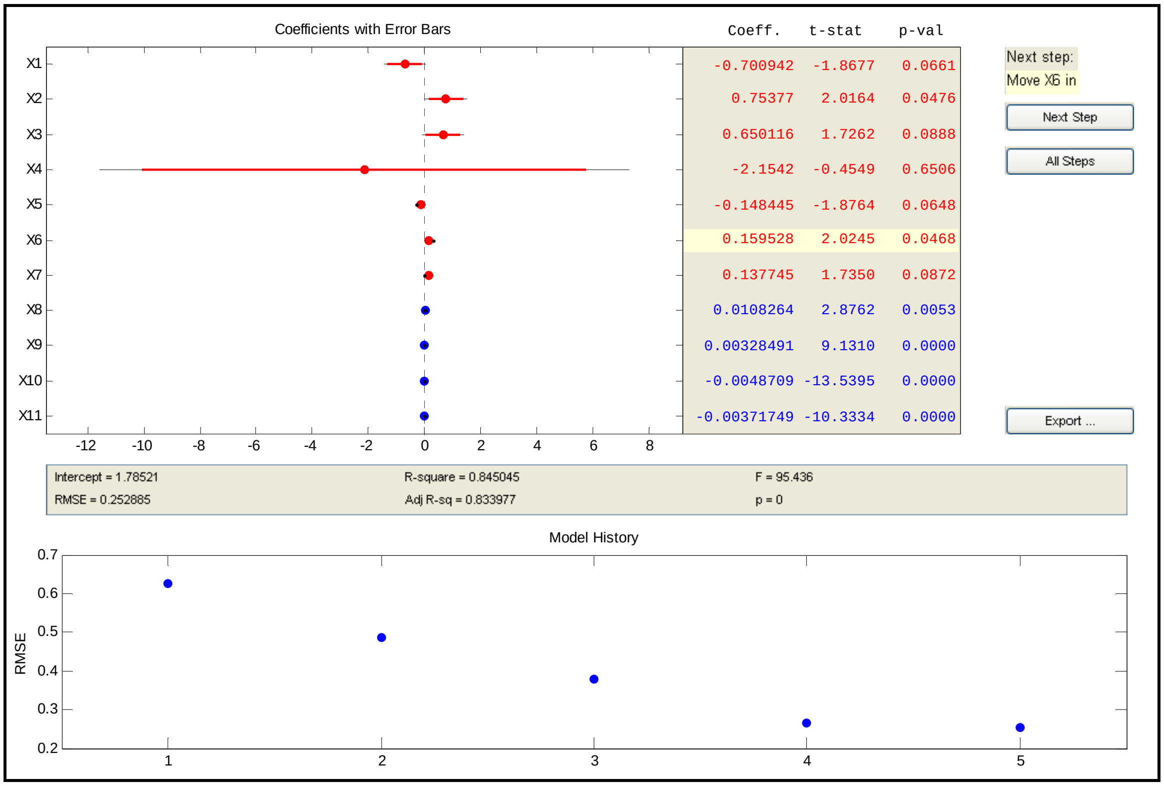


Figure 4-29 Final results for the Piece-Wise method using a monomial basis for region #4.

Table 4-29 output shows the final results for this region.

Table 4-29 Final Output for the Piece-Wise Method using a Monomial basis for region #4

```
%-----------------------------------------------------------------------------%
     Results for the Analysis of Variance for the Basis Representation Model
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
   0     1.7852      0.34159   5.2262   1.7057e-006

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    8.0000    0.0108    0.0038    2.8762    0.0053    1.0000
    9.0000    0.0033    0.0004    9.1310    0.0000    1.0000
   10.0000   -0.0049    0.0004  -13.5395         0    1.0000
   11.0000   -0.0037    0.0004  -10.3334    0.0000    1.0000

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.84504  0.83619
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF          SS     MS           F      P
Regression         4     24.4129     6.1032     95.436     0
Residual Error    70     4.4766      0.063951
Total             74     28.8895
%-----------------------------------------------------------------------------%
```

The same terms obtained for the previous two regions were selected again by the stepwise procedure. Table 4-30 summarizes the results.

Table 4-30 Results summary for Region #4

Terms	Coefficient Represented	Value	Related Factor	Basis Term
8	b_{02}	0.0108	1	t^2
9	b_{12}	0.0033	X1	t^2
10	b_{22}	-0.0049	X2	t^2
11	b_{32}	-0.0037	X1*X2	t^2

This model has the peculiarity that has an adjusted determination coefficient over 80%. The most relevant terms in this model correspond to the factor x_2 and the cross-term. The following plots correspond to the estimated responses and residuals for this region. Figure 4-30 shows the estimated responses for this region. This looks to be the region with the best fit. The residuals plot shows a behavior very close to a model with constant variance, this is presented in Figure 4-31

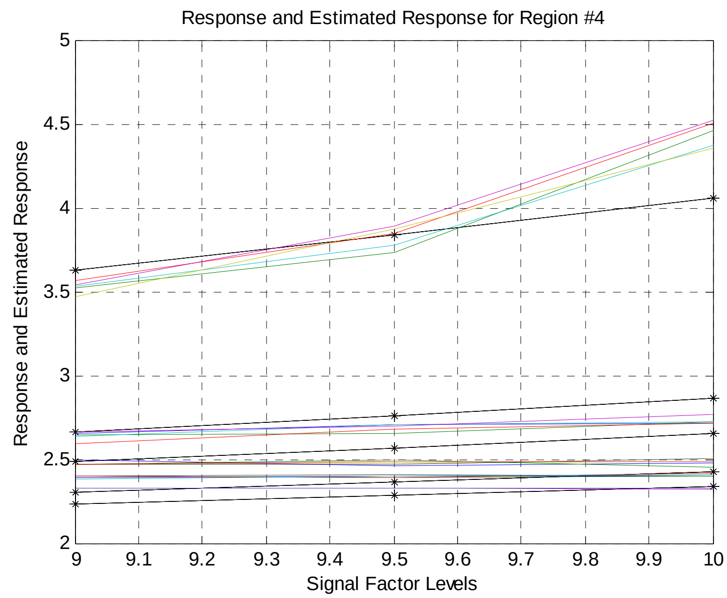


Figure 4-30 Response and estimated response for the region #4. The asterisks correspond to the estimated functions.

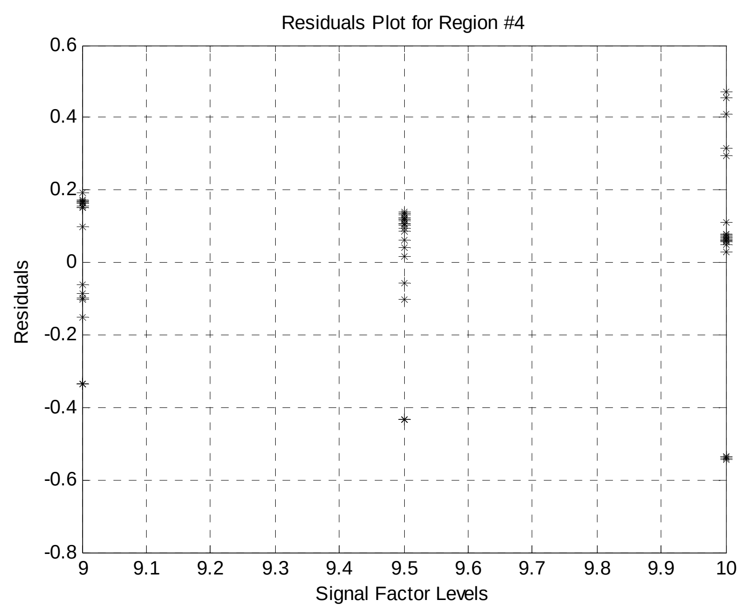


Figure 4-31 Residuals plot for region #4

The next plots correspond to the estimated responses and the residuals for all the regions combined.

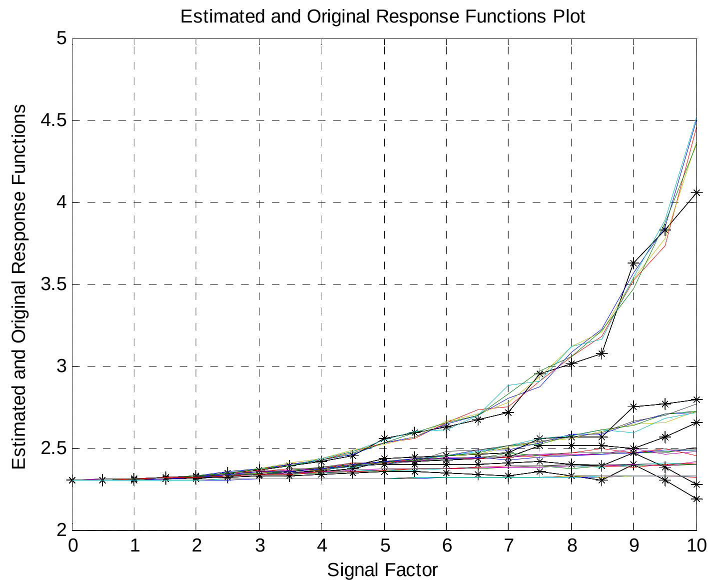


Figure 4-32 Estimated and Original Functions for the Piece-Wise Method for all the regions Combined

The models fit relatively well the response functions. The problems occur on the last region due to the dramatic change on the curves behavior. The next plot is the residuals plot for the combined model.

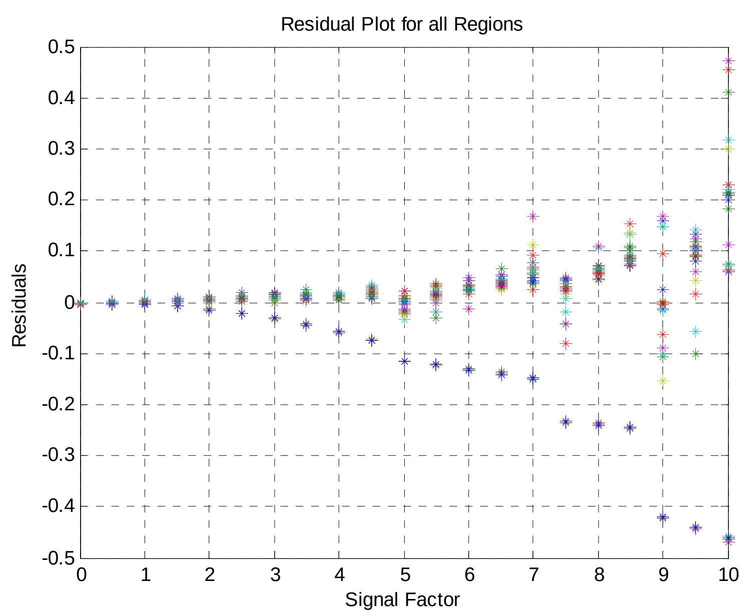


Figure 4-33 Residuals Plot for all the Regions

Figure 4-33 shows that the residuals present the same behavior than the obtained with the basis representation method. In the next pages, the results for the same procedure using a Fourier Basis are shown.

4.3.3.2 Fourier Basis

In order to begin with the procedure the knots location was completed. Table 4-31 presents the results of the knots search.

Table 4-31 Summary of the Knots Search for the Theoretical Case Study using a Fourier basis

Fourier Basis	**K=2**					
Number of knots	**Objective function**	**optimal knots**				
5	7.5641	3	6	9	12	16
4	5.5241	3	7	11	16	*
3	**5.4986**	**6**	**11**	**16**	*	*
2	5.6749	11	16	*	*	*
1	6.7496	16	*	*	*	*

The selected quantity of knots is 3 and the selected knots are 6, 11, and 16. Due the use of a Fourier Basis, the models in each region are going to equal to the Equation 4.5. The Figure 4-34 presents the plots for each of the four regions delimited by the knots.

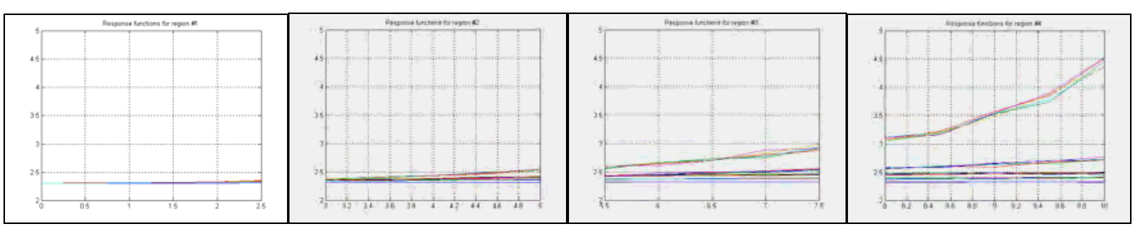


Figure 4-34 Plots for all the regions delimited before using the Piece-Wise procedure for the theoretical case study with a Fourier basis

The next sections show the results obtained in each region in detail.

4.3.3.2.1 Results for Region #1

As done with the previous basis, the initial Matlab output is presented in order to show the results obtained for this region.

Table 4-32 Initial Output for the Piece-Wise Method using a Fourier basis for region #1

```
%-------------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for
the region  #1
%-------------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-------------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-------------------------------------------------------------------------------%
   Terms     Coefficients  SE Coef   T-test  P-value   VIF
     0       2.3127   0.00099084      2334.1            0             0
     1    0.0018039    0.0011078      1.6284      0.10587          1.08
     2  -0.00096109    0.0011078    -0.86757      0.38723          1.08
     3  -0.00013464    0.0011078    -0.12153      0.90346          1.08
     4   -0.0084468    0.0014771     -5.7187  7.0121e-008             1
     5   2.415e-005    0.0016514    0.014624      0.98835             1
     6    0.0007692    0.0016514     0.46579      0.64215             1
     7    0.0012705    0.0016514     0.76932       0.4431             1
     8    0.0027616    0.0013211      2.0903     0.038536        1.0667
     9   9.657e-005    0.0014771     0.06538      0.94797          1.12
    10   -0.0007232    0.0014771    -0.48962      0.62523          1.12
    11   -0.0016348    0.0014771     -1.1068      0.27044          1.12
    12   -0.0020372    0.0014771     -1.3792       0.1702             1
    13  5.1963e-005    0.0016514    0.031466      0.97495             1
    14   0.00015014    0.0016514     0.09092       0.9277             1
    15  3.9023e-005    0.0016514     0.02363      0.98118             1
    16     0.001068    0.0013211     0.80841      0.42033        1.0667
    17  7.6214e-005    0.0014771    0.051598      0.95893          1.12
    18  -0.00035591    0.0014771    -0.24096      0.80996          1.12
    19   -0.0009557    0.0014771    -0.64703      0.51875          1.12

%-------------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-------------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.27208  0.1657
%-------------------------------------------------------------------------------%
                      Analysis of Variance
%-------------------------------------------------------------------------------%
Source            DF          SS           MS            F         P
Regression        19     0.0066259     0.00034873     2.5575     0.00098176
Residual Error   130     0.017726      0.00013636
Total            149     0.024352
%-------------------------------------------------------------------------------%
```

This model does not have the multicollinearity problem presented by the monomial basis, but it has a series of non-relevant terms that can be eliminated. This in order to improve and simplify the model the stepwise procedure was completed. Figure 4-35 correspond to this procedure.

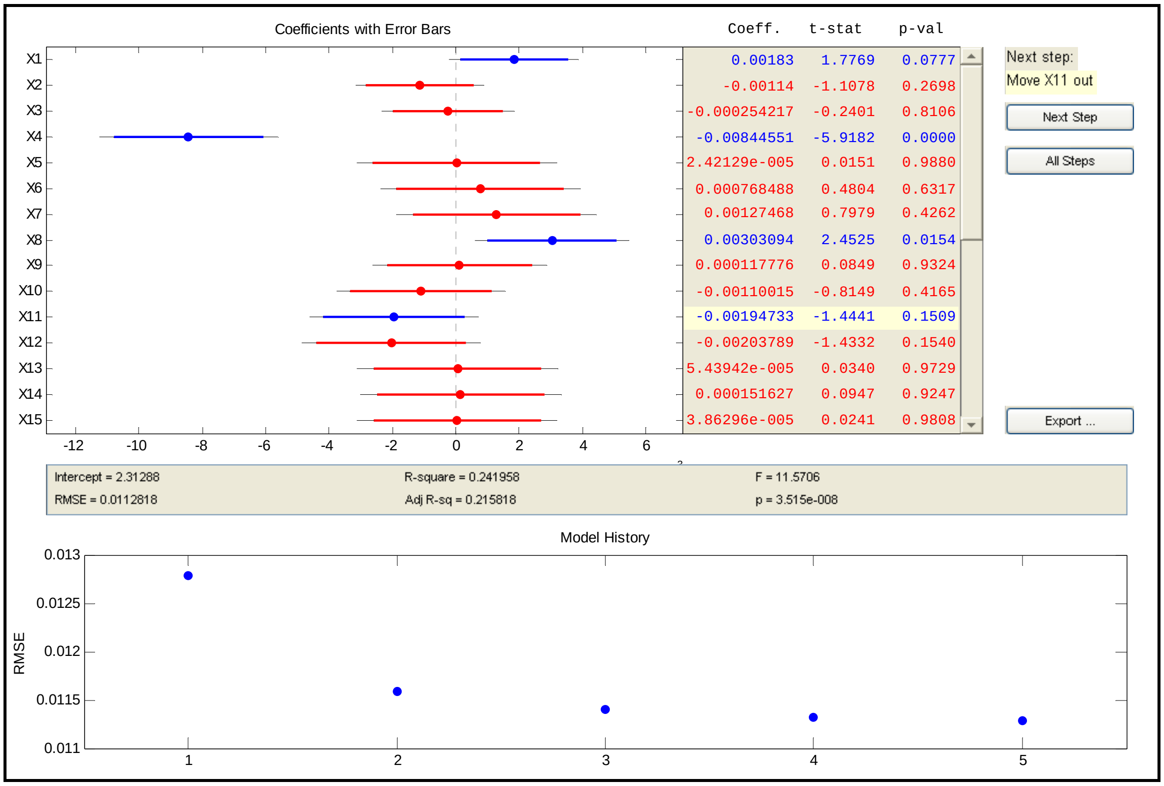


Figure 4-35 Final results for the Piece-Wise method using a Fourier basis for region #1.

Table 4-33 corresponds to the final model it is shown next.

Table 4-33 Final Output for the Piece-Wise Method using a Fourier basis for region #1

```
%----------------------------------------------------------------------------%
     Results for the Analysis of Variance for the Basis Representation Model
%----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%----------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
   0     2.3129      0.00094403   2450.0134   0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    1.0000    0.0018    0.0010    1.7793    0.0773    1.0000
    4.0000   -0.0084    0.0014   -5.9183    0.0000    1.0000
    8.0000    0.0030    0.0012    2.4503    0.0155    1.0000
   11.0000   -0.0019    0.0013   -1.4432    0.1511    1.0000

%----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.24195  0.22103
%----------------------------------------------------------------------------%
                      Analysis of Variance
%----------------------------------------------------------------------------%
Source            DF       SS            MS            F           P
Regression         4     0.0058919     0.001473     11.5698     3.5191e-008
Residual Error   145    0.01846       0.00012731
Total            149    0.024352
%----------------------------------------------------------------------------%
```

The model is not a good one; it only describes 24% of the variability of the data. Inferences using this model are not appropriate only two factor related terms were considered relevant by the model and two additional terms corresponding to the intercept of the model. The next plots (Figure 4-36 and Figure 4-37) show the lack of fit and variance problems that this model has. This is expected to happen in all the regions, due the basis selection.

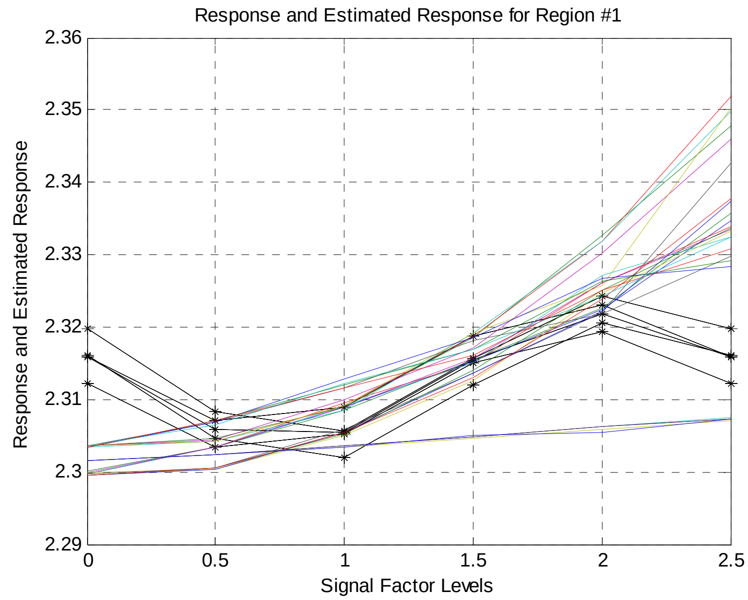


Figure 4-36 Response and estimated response for the region #1. The asterisks correspond to the estimated functions.

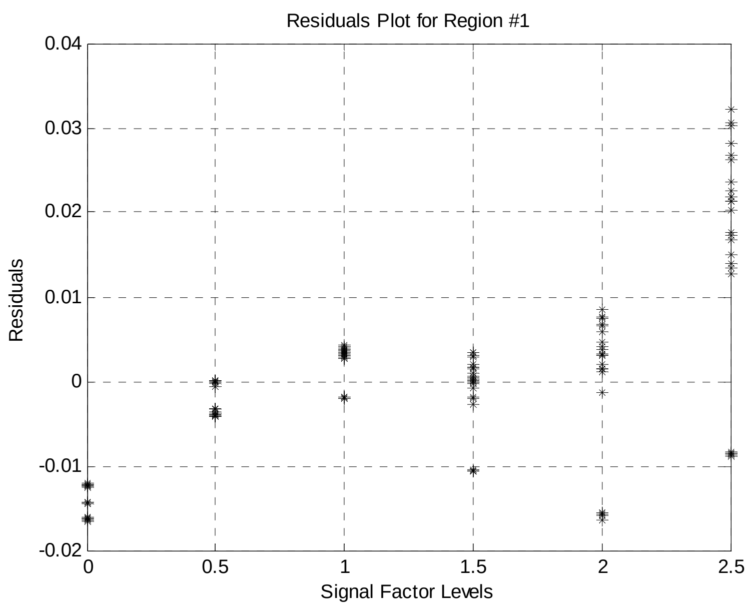


Figure 4-37 Residuals plot for region #1

4.3.3.2.2 Results for Region #2

Table 4-34 is the initial Matlab output; the stepwise plots and the final model are presented in the next pages.

Table 4-34 Initial Output for the Piece-Wise Method using a Fourier basis for region #2

```
%--------------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for the
region  #2
%--------------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%--------------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%--------------------------------------------------------------------------------%
  Terms   Coefficients   SE Coef        T-test        P-value       VIF
     0       2.3711     0.0032493       729.72            0           0
     1    0.0089025     0.0036328       2.4506      0.015914      1.0937
     2    -0.020265     0.0036328      -5.5783    1.905e-007      1.0937
     3    -0.023839     0.0036328      -6.5621   2.0578e-009      1.0937
     4     0.031921     0.0054923        5.812   6.6951e-008        1.25
     5    0.0065675     0.0061406       1.0695       0.28728        1.25
     6    -0.016107     0.0061406      -2.6231      0.010009        1.25
     7    -0.016505     0.0061406      -2.6879     0.0083636        1.25
     8    -0.003161     0.0042543     -0.74301       0.45913        1.05
     9   -0.0012625     0.0047565     -0.26543        0.7912       1.125
    10    0.0025425    0.0047565      0.53454        0.5941       1.125
    11     0.002605    0.0047565      0.54768       0.58508       1.125
    12  -1.2757e+014   2.0057e+013    -6.3606   5.3344e-009       39.65
    13  -3.3397e+013   2.2424e+013    -1.4894       0.13939       41.25
    14   7.3368e+013   2.2424e+013     3.2719     0.0014463       41.25
    15   6.6529e+013   2.2424e+013     2.9669     0.0037255       41.25
    16     -0.12477    0.019917       -6.2645   8.3644e-009       39.45
    17    -0.031685    0.022267       -1.4229       0.15772      41.094
    18     0.070557    0.022267        3.1686     0.0020065      41.094
    19     0.065179    0.022267        2.9271     0.0041952      41.094

%--------------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%--------------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.62116  0.5526
%--------------------------------------------------------------------------------%
                      Analysis of Variance
%--------------------------------------------------------------------------------%
Source           DF          SS          MS            F          P
Regression       19     0.20773     0.010933     9.061     1.0325e-014
Residual Error   105    0.12669      0.0012066
Total            124    0.33442
%--------------------------------------------------------------------------------%
```

It is possible to observe that the model has a multicollinearity problem. In addition, the determination coefficient for the model is relatively low. The stepwise plots are shown in Figure 4-38.

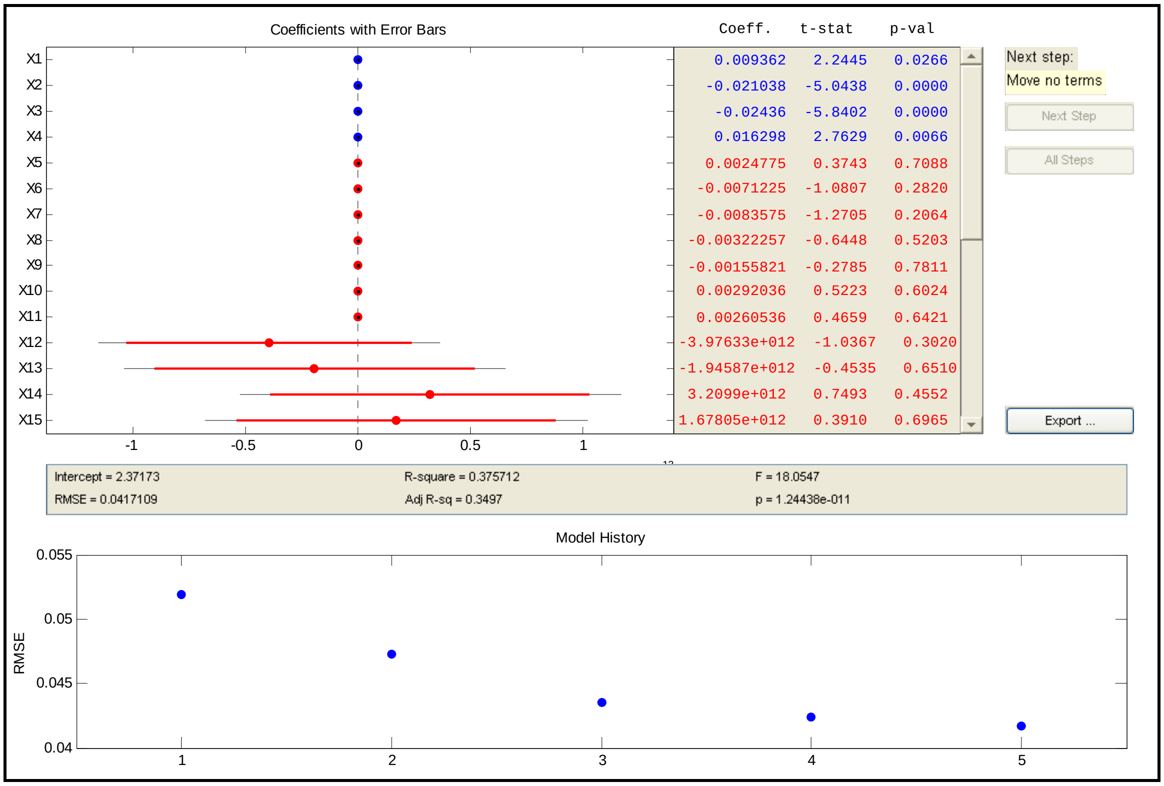


Figure 4-38 Final results for the Piece-Wise method using a Fourier basis for region #2.

Table 4-35 provides the details with respect the final model obtained after the execution of the stepwise procedure.

Table 4-35 Final Output for the Piece-Wise Method using a Fourier basis for region #2

```
%-----------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.3717      0.0037311   635.6675   0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    1.0000    0.0094    0.0042    2.2447    0.0266    1.0000
    2.0000   -0.0210    0.0042   -5.0424    0.0000    1.0000
    3.0000   -0.0244    0.0042   -5.8393    0.0000    1.0000
    4.0000    0.0163    0.0059    2.7625    0.0066    1.0000

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.37563  0.35482
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF          SS          MS         F         P
Regression        4     0.12563     0.031406     18.0484     1.254e-011
Residual Error   120     0.20881    0.0017401
Total            124     0.33444
%-----------------------------------------------------------------------------%
```

The final model is not a good one. Only around 35% of the total variability of the data is explained. Only the first four terms are relevant and the one that correspond to the cross-term in is the most relevant. In general, this model cannot be used to conclude with respect this region. Plots presented in figures 4-39 and 4-40 show the lack of fit presented by this model. The next section shows the results for region #3.

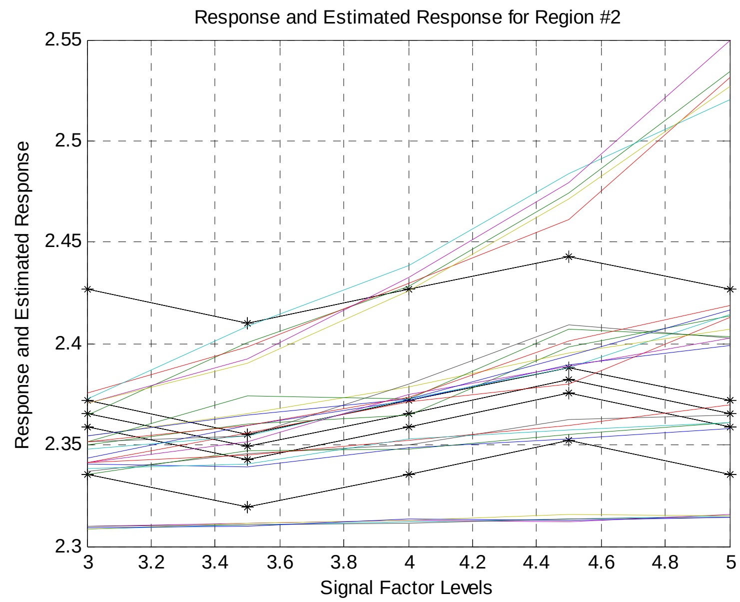


Figure 4-39 Response and estimated response for the region #2. The asterisks correspond to the estimated functions.

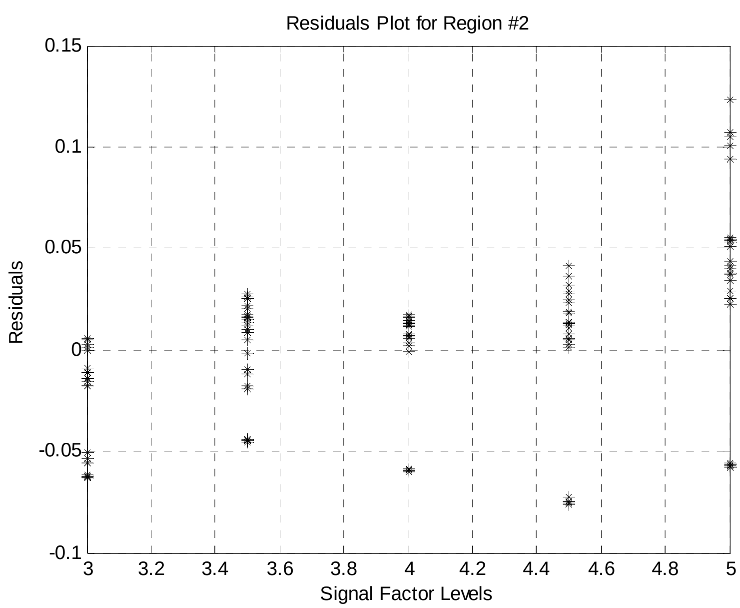


Figure 4-40 Residuals Plot for region #2

4.3.3.2.3 Results for Region #3

Table 4-36 corresponds to the initial model for this region

Table 4-36 Initial Output for the Piece-Wise Method using a Fourier basis for region #3

```
%-----------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for the
region  #3
%-----------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
Terms   Coefficients  SE Coef        T-test      P-value          VIF
   0        2.469    0.0079599      310.18            0            0
   1     0.046331    0.0088994      5.2061  9.6428e-007       1.0937
   2    -0.097173    0.0088994     -10.919            0       1.0937
   3    -0.076934    0.0088994     -8.6449  6.5947e-014       1.0937
   4      0.71764      0.17487      4.1039  8.0472e-005       295.61
   5      0.39114      0.19551      2.0006     0.048012       316.72
   6     -0.72718      0.19551     -3.7194   0.00032255       316.72
   7     -0.42635      0.19551     -2.1807     0.031436       316.72
   8    -0.048996     0.013455     -3.6416   0.00042265         1.25
   9    -0.027054     0.015043     -1.7984     0.074979         1.25
  10     0.049308     0.015043      3.2779    0.0014187         1.25
  11     0.029391     0.015043      1.9538      0.05338         1.25
  12  2.0336e+014  4.9133e+013      4.1389  7.0606e-005       472.63
  13  1.1123e+014  5.4933e+013      2.0249     0.045415       516.75
  14 -2.0631e+014  5.4933e+013     -3.7557   0.00028408       516.75
  15 -1.2125e+014  5.4933e+013     -2.2072     0.029475       516.75
  16      0.32367     0.078626      4.1166  7.6768e-005       102.45
  17      0.17611     0.087907      2.0034     0.047713       106.72
  18     -0.32841     0.087907     -3.7359    0.0003045       106.72
  19      -0.1929     0.087907     -2.1944     0.030412       106.72

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.73981  0.69272
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source           DF          SS          MS            F         P
Regression       19     2.1618       0.11378        15.7129     0
Residual Error   105    0.76031      0.0072411
Total            124    2.9221
%-----------------------------------------------------------------------------%
```

This model has a better initial determination coefficient than the initial models for the first two regions. The model has a serious multicollinearity problem; in addition there are some terms that are not relevant in the model. The plots that correspond to the stepwise procedure are presented by Figure 4-41.

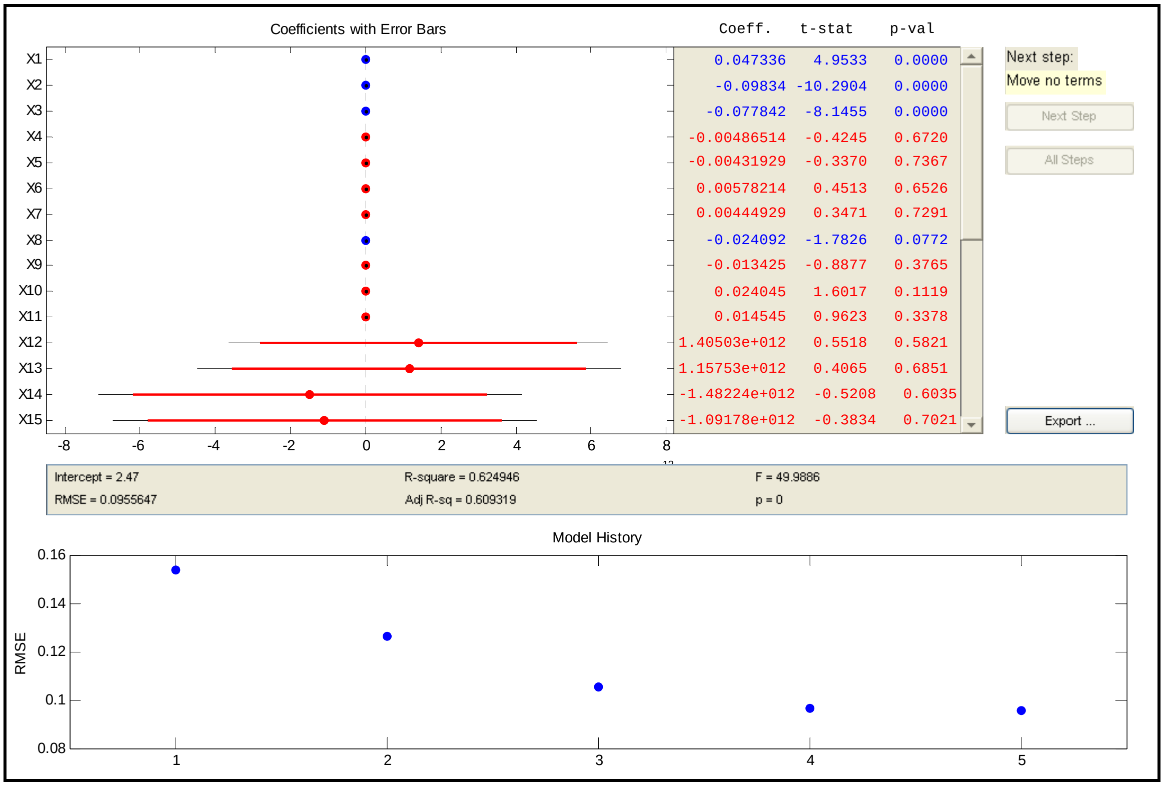


Figure 4-41 Final results for the Piece-Wise method using a Fourier basis for region #3.

Table 4-37 corresponds to the stepwise procedure it is shown in the next page.

Table 4-37 Final Output for the Piece-Wise Method using a Fourier basis for region #3

```
%-----------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.47      0.008548   288.9553   0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    1.0000    0.0473    0.0096    4.9528    0.0000    1.0000
    2.0000   -0.0983    0.0096  -10.2896         0    1.0000
    3.0000   -0.0778    0.0096   -8.1452    0.0000    1.0000
    8.0000   -0.0241    0.0135   -1.7825    0.0772    1.0000

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.62491  0.61241
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF          SS      MS          F       P
Regression         4     1.8261     0.45651      49.9819     0
Residual Error   120     1.096      0.0091336
Total            124     2.9221
%-----------------------------------------------------------------------------%
```

This is model it is better than the previous ones in terms of the adjusted determination coefficient. The model describes 63% of the total variability of the data. This quantity is not excellent but is acceptable in some applications. Table 4-38 summarizes the obtained results.

Table 4-38 Results summary for region #3 of the Piece-Wise method using a Fourier basis

Terms	Coefficient Represented	Value	Related Factor	Basis Term
1	b_{10}	0.0473	X1	1
2	b_{20}	-0.0983	X2	1
3	b_{30}	-0.0778	X1*X2	1
8	b_{02}	-0.0241	1	cos(wt)

Terms 1 to 3 correspond to the factors without a sine or cosine term. By the magnitudes of the coefficient the factor, x_2 is the most relevant in this region.

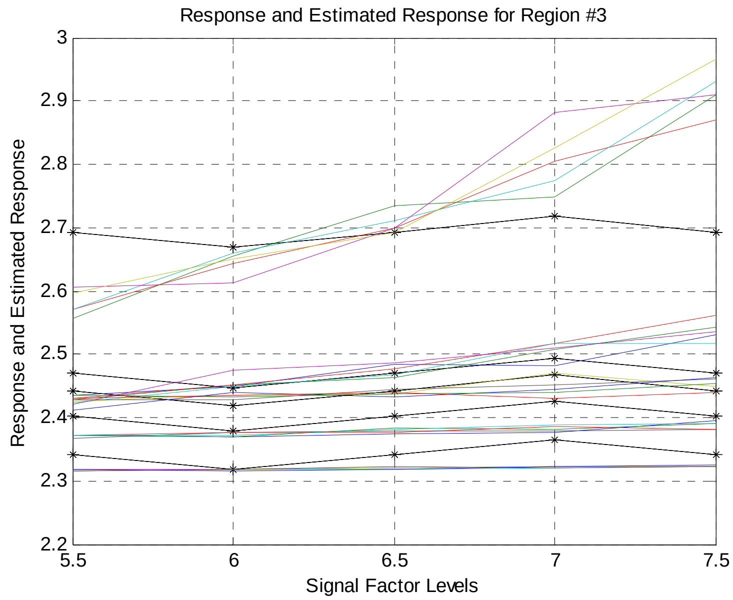


Figure 4-42 Response and estimated response for the region #3. The asterisks correspond to the estimated functions.

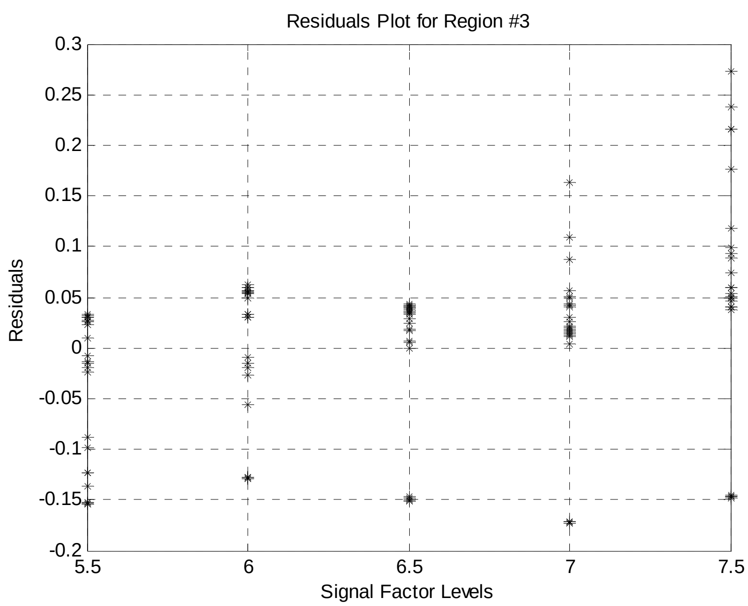


Figure 4-43 Residuals for region #3

The estimated responses plot presents a better fit compared to the previous two. The residuals plot shows the problem of non-constant variance for this model. The next section shows the results for the region #4, which is the last one due the number of knots used.

4.3.3.2.4 Results for Region #4

As done with all the regions in this procedure Table 4-39 corresponds to the initial model for this region.

Table 4-39 Initial Output for the Piece-Wise Method using a Fourier basis for region #4

```
%-----------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for the
region  #4
%-----------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
 Terms  Coefficients  SE Coef        T-test      P-value          VIF
   0       2.6833     0.019566       137.14            0            0
   1      0.21396     0.021875       9.7812  2.2204e-016       1.0937
   2     -0.33833     0.021875      -15.466            0       1.0937
   3     -0.25323     0.021875      -11.576            0       1.0937
   4      -2.4028      0.44484      -5.4015  4.1446e-007       226.15
   5      -2.3104      0.49735      -4.6453  9.8854e-006       226.15
   6       2.8705      0.49735       5.7716  8.0325e-008       226.15
   7       2.3266      0.49735       4.6781  8.6648e-006       226.15
   8     0.024987     0.025618      0.97538      0.33161         1.05
   9     0.029249     0.028641       1.0212      0.30949        1.125
  10    -0.031858     0.028641      -1.1123      0.26855        1.125
  11    -0.027494     0.028641     -0.95996      0.33928        1.125
  12 -6.3378e+014  1.2077e+014      -5.2478   8.065e-007       420.55
  13 -6.1025e+014  1.3503e+014      -4.5195  1.6321e-005       428.65
  14  7.5719e+014  1.3503e+014       5.6077  1.6724e-007       428.65
  15  6.1479e+014  1.3503e+014       4.5531  1.4287e-005       428.65
  16      -1.3836      0.26694      -5.1831   1.064e-006       195.45
  17      -1.3285      0.29845      -4.4513  2.1344e-005       203.59
  18       1.6526      0.29845       5.5371  2.2862e-007       203.59
  19       1.3385      0.29845        4.485    1.87e-005       203.59

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.86579  0.84151
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF       SS           MS          F        P
Regression        19     29.6349     1.5597     35.6505     0
Residual Error   105     4.5938      0.043751
Total            124     34.2287
%-----------------------------------------------------------------------------%
```

It is not surprising to obtain a model with the presence of the multicollinearity problem. In addition, some terms are not relevant in the model. For those reasons, once again the stepwise procedure was used. The plots and final model are shown in the next pages.

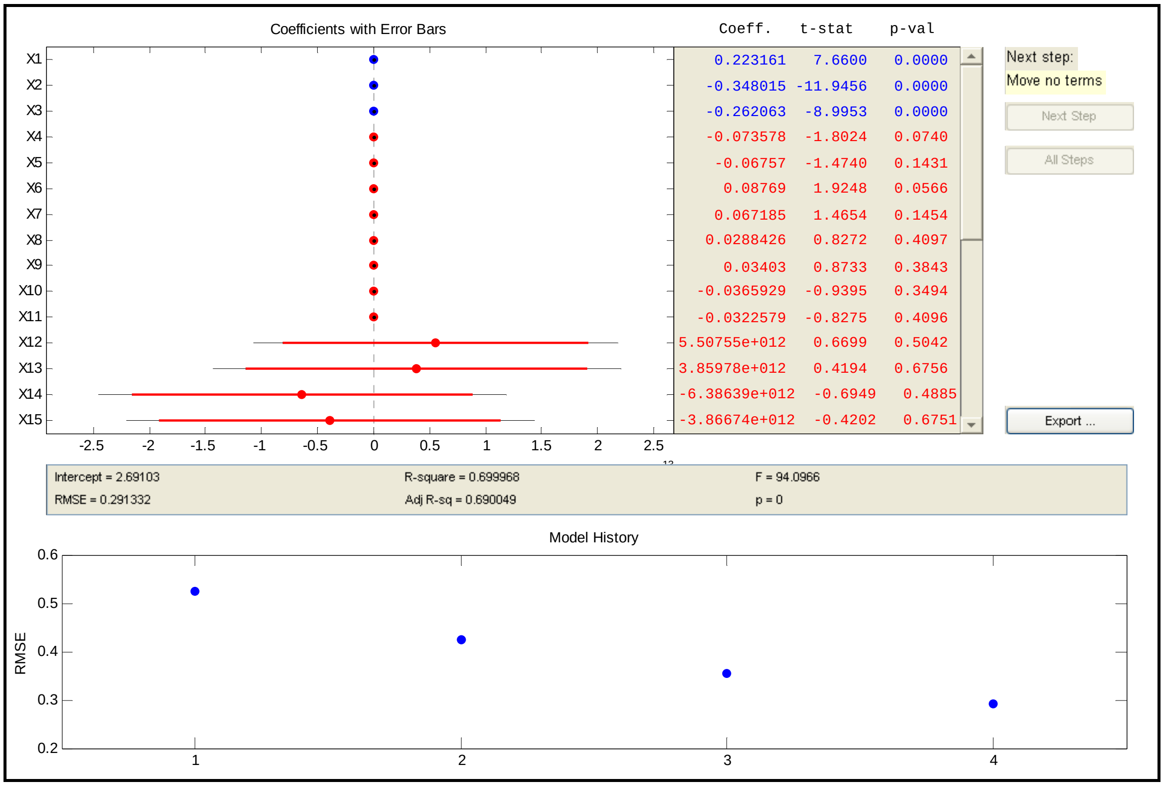


Figure 4-44 Final results for the Piece-Wise method using a Fourier basis for region #4.

The details related with the final model for this region are shown in Table 4-40.

Table 4-40 Final Output for the Piece-Wise Method using a Fourier basis for region #4

```
%----------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%----------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%----------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.691      0.026058   103.2727   0

   Term    Coefficients  SE Coef   T-test  P-value       VIF
     1      0.22316     0.029133    7.66   5.0999e-012     1
     2     -0.34801     0.029133  -11.946            0     1
     3     -0.26206     0.029133  -8.9952  3.9968e-015     1
%----------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%----------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.69996  0.69253
%----------------------------------------------------------------------%
                      Analysis of Variance
%----------------------------------------------------------------------%
Source            DF          SS      MS        F         P
Regression         3     23.9589     7.9863     94.0954     0
Residual Error   121      10.2698    0.084874
Total            124      34.2287
%----------------------------------------------------------------------%
```

The model obtained is the best one for the Fourier basis. The adjusted determination coefficient is close to a 70% and the model has an orthogonal design matrix. The next table shows the results summarized.

Table 4-41 Results summary for region #4 of the Piece-Wise method using a Fourier basis

Terms	Coefficient Represented	Value	Related Factor	Basis Term
1	b_{10}	0.22316	X1	1
2	b_{20}	-0.34801	X2	1
3	b_{30}	-0.26206	X1*X2	1

The relevant terms correspond to the terms that not are multiplied by a sine or cosine term. This is exactly what happened with the previous region. The term that corresponds to the factor x_2 is the most relevant in this model. Figure 4-45 shows the estimated responses as constant lines; this behavior is caused by the elimination of all the terms related to the signal factor. The residuals plot shows the same problems presented by the models of the previous regions.

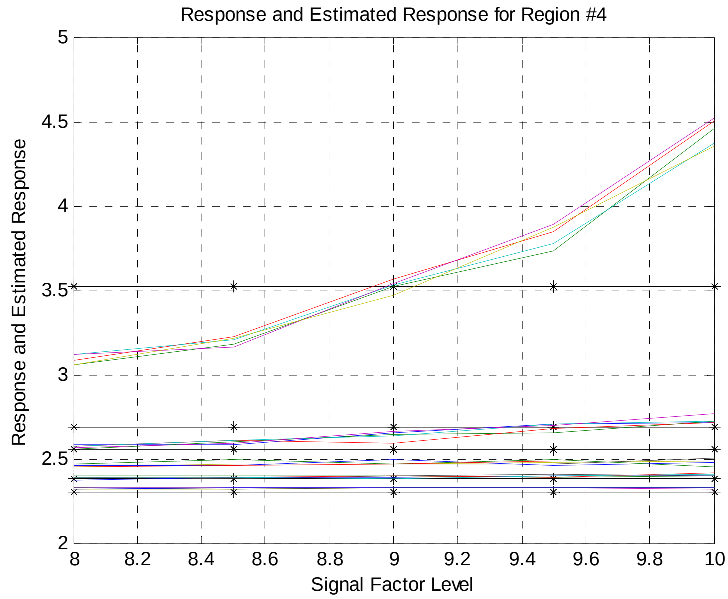


Figure 4-45 Response and estimated response for the region #4. The asterisks correspond to the estimated functions.

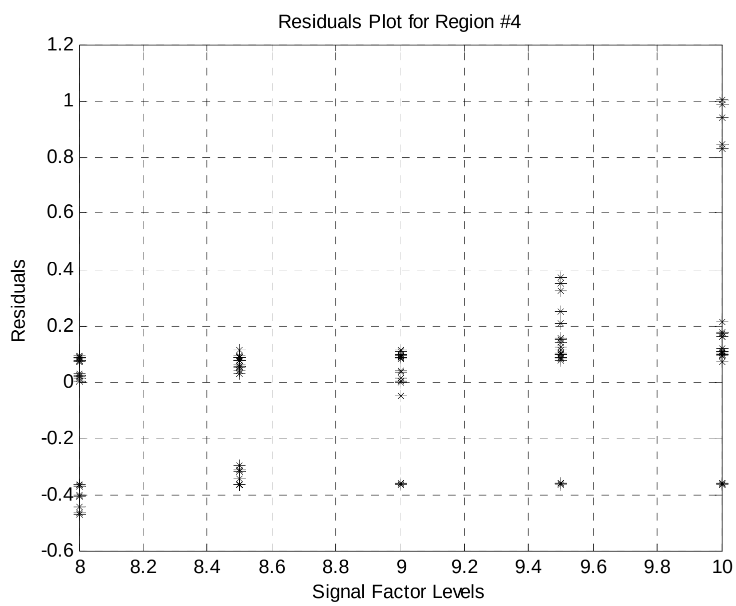


Figure 4-46 Residuals plot for region #4

The next section summarizes the results for the Piece-Wise method without the continuity constraint.

4.3.3.3 Results Summary for the Piece-Wise Method without the Continuity Constraint

After performing, the Piece-Wise method for the theoretical case study is important to summarize and make some remarks with respect the results obtained. The next tables show a summary for the two bases used by region. These tables also present the results for the pseudo-MSE for both bases.

Table 4-42 Summary for the Piece-Wise procedure using a Monomial Basis

Monomial Basis	Number of Terms	MSE	Number of Data Points	Adjusted R^2
Region 1	7	0.00033438	250	0.75496
Region 2	5	0.004758	125	0.69203
Region 3	5	0.015969	75	0.78489
Region 4	5	0.063951	75	0.83619
Pseudo-MSE	0.012418718			

Table 4-43 Summary for the Piece-Wise procedure using a Fourier Basis

Fourier Basis	Number of Terms	MSE	Number of Data Points	Adjusted R^2
Region 1	7	0.00012672	150	0.22469
Region 2	7	0.0017288	125	0.35902
Region 3	8	0.0090161	125	0.61098
Region 4	4	0.084874	125	0.69253
Pseudo-MSE	0.023173788			

From tables 4-42 and 4-43 one can notice the difference in the adjusted determination coefficients for both bases. The monomial basis has a better performance compared to the Fourier basis. This result was expected due the behavior of the response functions that do not have a periodical behavior. The pseudo-MSE for the monomial basis it is approximately the half of the pseudo-MSE of the Fourier basis. Both bases worked better in their respective regions 3 and 4. Also for both bases, the factor x_2 was the most relevant for the last region. In general, the method worked out decently well for this case in specific when the monomial basis was used. The next section of this chapter presents the results for the Piece-Wise method with the incorporation of the continuity constraints.

4.3.4 Piece Wise Method with the Continuity Constraint

In order to ensure the continuity among the regions for all the curves, a set of constraints were imposed to the models. As mentioned on Chapter 3, these constraints are incorporated to the models using a series of indicator variables, after the stepwise procedure. The number of those variables is equal to the number of experimental conditions minus one. The reasoning of subtracting one experimental condition is to create a reference or a base for the rest of the indicator variables. Then the restrictions to these indicator variables are the averages of the responses at the last level of the signal factor of the previous region. The restricted least of squares algorithm it is used to find the regression coefficients for each region. The following sections are going to present the implementation of this procedure, it is important to understand that the terms considered for the final regressions are the same terms selected by the stepwise procedure for the models without the continuity constraint. The results obtained are next.

4.3.4.1 Monomial Basis

4.3.4.1.1 Results for Region #1

The results obtained for this region are exactly the same results obtained for the model without the continuity constraint, because this region does not involve the restrictions. The next tables show the results summary for this region.

Table 4-44 Final Matlab Output for Region #1

```
%-----------------------------------------------------------------------------%
   Results for the Analysis of Variance for the Basis Representation Model
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     2.3039      0.0016865   1366.0952   0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    6.0000    0.0031    0.0020    1.5809    0.1152   16.2281
    7.0000    0.0041    0.0020    2.0958    0.0371   16.2281
    8.0000    0.0041    0.0002   23.9729         0    1.0000
    9.0000    0.0005    0.0001    3.6945    0.0003    1.0000
   10.0000   -0.0019    0.0005   -3.6230    0.0004   16.2281
   11.0000   -0.0024    0.0005   -4.5972    0.0000   16.2281

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.76087  0.75496
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF       SS          MS           F       P
Regression        6     0.25854     0.04309     128.8631     0
Residual Error   243    0.081255    0.00033438
Total            249    0.33979
%-----------------------------------------------------------------------------%
```

Table 4-45 Results Summary for Region #1

Terms	Coefficient Represented	Value	Related Factor	Basis Term
6	b_{21}	0.0031	X2	t
7	b_{31}	0.0041	X1*X2	t
8	b_{02}	0.0041	1	t^2
9	b_{12}	0.0005	X1	t^2
10	b_{22}	-0.00019	X2	t^2
11	b_{32}	-0.0024	X1*X2	t^2

As mentioned previously the model has the multicollinearity problem but both factors were considered as relevant.

4.3.4.1.2 Results for Region #2

The results of the analysis of variance for this region are shown on the next tables

Table 4-46 Matlab Output for Region #2

```
%-----------------------------------------------------------------------------%
Analysis of Variance for the Final Model with the restricted coefficients for
the region #2
%-----------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
    Term       Coefficient      SE Coef       T-test      P-value        VIF
     0            2.4561         0.0196      125.0281        0
     8            0.0242         0.0042       5.7108         0             1
     9            0.0205         0.0047       4.3445         0          2.0345
    10           -0.0194         0.0047      -4.1053      0.0001       2.0345
    11           -0.0239         0.0047      -5.0538         0          2.0345
   ind 1         -0.0725         0.0281      -2.5766      0.0112       3.2552
   ind 2         -0.0024         0.0281      -0.0856      0.9319       3.2552
   ind 3         -0.115          0.0281      -4.0868      0.0001       3.2552
   ind 4         -0.0686         0.0232      -2.9517      0.0038       2.2207

%-----------------------------------------------------------------------------%
                           R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
                                  R^2    adj R^2
                                   0        0
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
      Source          DF          SS          MS          F            P
    Regression         8       -0.4375     -0.0547        0            1
  Residual Error      116       0.5644      0.0049
      Total           124       0.1269
%-----------------------------------------------------------------------------%
```

This model is completely inadequate if is compared to the model without the continuity constraints. Inferences with this model are not reliable. The next section presents the results for the third region.

4.3.4.1.3 Results for region #3

As done on the previous sections the results for the third region are in the following tables.

Table 4-47 Matlab Output for the region #3

```
%--------------------------------------------------------------------------%
Analysis of Variance for the Final Model with the restricted coefficients for
the region #3
%--------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%--------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%--------------------------------------------------------------------------%
    Term       Coefficient      SE Coef      T-test      P-value       VIF
     0           2.5404           Inf          0            1
     8           0.2427           Inf          0            1            0
     9           0.2433           Inf          0            1            0
    10          -0.0349           Inf          0            1         9.7965
    11           0.0349           Inf          0            1         9.7965
   ind 1        -0.1244           Inf          0            1            0
   ind 2         0.2654           Inf          0            1            0
   ind 3        -0.1938           Inf          0            1            0
   ind 4        -0.0563           Inf          0            1            0
                                                                 1.0e+015 *
%--------------------------------------------------------------------------%
                           R^2 and Adjusted R^2
%--------------------------------------------------------------------------%
                                  R^2    adj R^2
                                0.7069   0.6714
%--------------------------------------------------------------------------%
                      Analysis of Variance
%--------------------------------------------------------------------------%
      Source           DF          SS          MS           F            P
    Regression          8        2.4026      0.3003      19.8983     6.77E-15
  Residual Error       66        0.9961      0.0151
      Total            74        3.3988
%--------------------------------------------------------------------------%
```

The multicollinearity problem is extremely severe. The effect of the indicator variables in the model was very adverse on it inference capabilities.

4.3.4.1.4 Results for region #4

The results for the last region analyzed are presented on the next couple of tables.

Table 4-48 Matlab Output for Region #4

```
%---------------------------------------------------------------------------%
Analysis of Variance for the Final Model with the restricted coefficients for
the region #4
%---------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%---------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%---------------------------------------------------------------------------%
        Term      Coefficient    SE Coef    T-test    P-value      VIF
          0             2.631     Inf            0          1
          8            0.6381     Inf            0          1       0
          9            0.5668     Inf            0          1       0
         10           -0.0288     Inf            0          1    9.7965
         11            0.0288     Inf            0          1    9.7965
       ind 1          -0.2854     Inf            0          1       0
       ind 2           0.4175     Inf            0          1       0
       ind 3          -0.0417     Inf            0          1       0
       ind 4          -0.2173     Inf            0          1       0
                                                                1.0e+015 *

%---------------------------------------------------------------------------%
                          R^2 and Adjusted R^2
%---------------------------------------------------------------------------%
                               R^2        adj R^2
                                0.4034       0.3311

%---------------------------------------------------------------------------%
                      Analysis of Variance
%---------------------------------------------------------------------------%
      Source          DF           SS           MS            F             P
    Regression         8           7.1038        0.888       5.5794      2.11E-05
  Residual Error      66         10.5039       0.1592
      Total           74          17.6077
%---------------------------------------------------------------------------%
```

Similar to the previous two regions the cost of the imposition of the continuity constraint has been a dramatic loss on inference capabilities. It is important to mention that the constraint was not incorporated to the models obtained using the Fourier bases because from the shape of the response functions it is clear that this type of basis is not appropriate. The Table 4-49 summarizes the results obtained

Table 4-49 Results Summary for the Piece-Wise method with Continuity Constraints

Monomial Basis	Number of Terms	SSE	MSE	Number of Data Points	Adjusted R^2
Region 1	7	0.081255	0.00033438	250	0.75496
Region 2	9	0.5644	0.0049	125	0
Region 3	9	0.9961	0.0151	75	0.6714
Region 4	9	10.5039	0.1592	75	0.3311
Pseudo-MSE	0.024737				

The next and last section of this chapter summarizes and compares the obtained results for all the methods applied on this case. The comparison it is based on the Pseudo MSE presented on Chapter 3.

4.4 Comparison of Results for the Three Methods for the Theoretical Case Study

After completing the discussion of the three methods individually, it is important to make some comparisons in order to verify the adequacy of the methods to determine which procedures are better. In general, the conclusions driven by the three procedures were the same. Both factors and their interaction were considered as relevant by every procedure. An important inference obtained by the Point-Wise procedure is the significance of a curve component that was not captured or considered in the experiment. The main purpose of the pseudo-MSE as explained in Chapter 3 is to ease the comparison of the three methods. The following table presents a summary of the pseudo-MSE for all the methods used to analyze the experiment.

Table 4-50 Pseudo-MSE for all the methods applied to the theoretical case study

Method	Basis Used	Total SSE	Total Number of Parameters	Total Number of Data Points	Pseudo MSE
Point-Wise	N/A	5.48682	84	525	0.01244177
Basis Representation	Monomial	8.9124	7	525	0.01720541
Basis Representation	Fourier	20.5355	20	525	0.04066436
Basis Representation	Cubic Spline	5.7778	8	525	0.01117563
Piece-Wise without continuity constraints	Monomial	6.246615	22	525	0.01241872
Piece-Wise without continuity constraints	Fourier	11.56372	26	525	0.02317379
Piece-Wise with continuity constraint	Monomial	12.14566	34	525	0.024737

The procedures that involved a Fourier bases are less effective in comparison with the other procedures. They have the highest SSE totals. As previously mentioned this behavior was expected due the nature of the response functions. The Point-Wise method, the Piece-Wise using a monomial basis and the basis representation using the cubic spline basis provided the best results in terms of the pseudo-MSE, considering the number of terms that each model has. It is highly desired that the final models have the lowest quantity of terms as possible, this in order to have a simple model able to describe the data under study. The basis representation model with the cubic spline basis provided the second lowest number of terms and the second lowest pseudo-MSE can be considered the selected model for this case.

This model has a good determination coefficient (approximately 83%), a low number of terms and a low SSE. The Piece-Wise method can be an alternative if is desired to investigate some specific regions of the signal factor. With the monomial basis in some regions, the factor x_2 was the most relevant and in other regions the cross-term had the greatest impact. This is very interesting because the factors can affect the response functions differently depending on the levels of the signal factor. This can be a real contribution for some engineering applications. The Piece Wise method provided a good fit for all the regions but the inferential capabilities of this method are questionable.

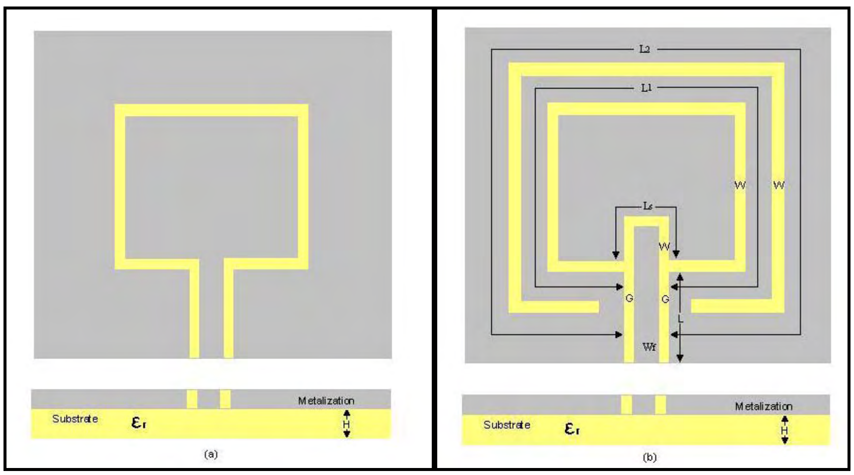
The work with the theoretical case study was completed successfully. The three methods were executed, analyzed and compared. The results were consistent among the methods and conclusions could be obtained for each procedure. The next chapter presents the application of all the techniques used in the theoretical case study in a real world application, the analysis of the reflection coefficient of rectangular slot ring antennas.

5

An Applied Case Study: Reflection Coefficient Analysis for Rectangular Slot Ring Antennas

5.1 Introduction

Recently the Design of Experiments (DOE) has been used in the analysis and characterization of Antennas [13]. In this chapter, the methods applied to the Theoretical case study are going to be applied to an experiment in order to analyze the behavior of the Reflection Coefficient for this structure in a range of frequencies. A set of runs where simulated in order to collect data for the analysis.



(a) (b)

Figure 5-1 (a) A single RSRA. (b) Two concentric RSRA. Two examples of the antennas used for the experiments

5.2 Experiment Description

Initially, some simulations were performed for the single Rectangular Slot Ring Antenna (RSRA) shown in Figure 5-1(a). Two different substrates were used. One with a relative permittivity of 3 and a thickness of 0.76 mm and the other with a relative permittivity of 6.15 and a thickness of 0.635 mm. The perimeter and the width of the slot of the antenna were varied in order to have an initial data to compare with. In the frequencies where the observations were made the reflection coefficient was very high for both substrates.

In order to improve the matching of the antenna to 50 Ω an open circuit stub is implemented in a new set of antennas. Then, a third set of antennas with an open circuit stub and an exterior ring are simulated. For some, cases a low reflection coefficient was achieved as it is shown in Figure 5-2.

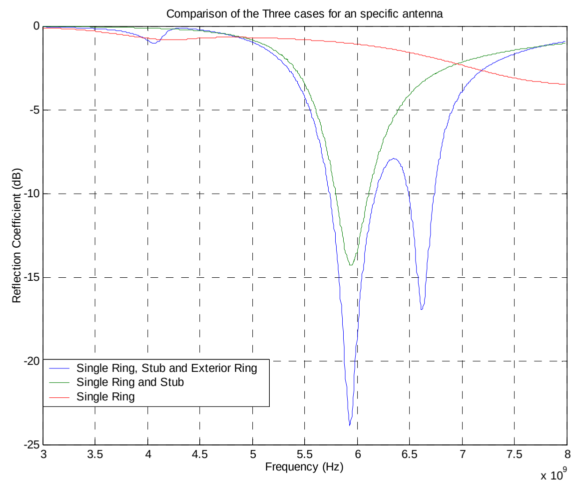


Figure 5-2 A case where a low reflection coefficient is observed with the structure proposed.

According to Figure 5-1(b), the parameters that can be varied are the substrate permittivity (ε_r), the substrate thickness (H), the slot ring perimeters (L_1 y L_2), the slot width (W), the open circuit stub (L_s), the fed slot width (G) and the fed conductor width (W_f). The factors to be considered in this design with its respective levels are shown in Table 5-1.

Table 5-1 Factors and levels considered for the experiments

Factor	Low level	High level
ε_r	3	6.15
W	0.25mm	1mm
L_1	$1\lambda_g$	$2\lambda_g$
L_2	$2.2\lambda_g$	$2.6\lambda_g$
L_s	$0.25\lambda_g$	$0.3\lambda_g$

The factors L_1, L_2, and L_s depend on λ_g. For these DOE designs the design frequency (f_d) is 5.77 GHz. The other variables such as W_f and G depend on the substrate permittivity and are obtained as follows. For ε_r=3 and *H*= 0.76 mm a 50 Ω CPW line is used with a slot width *G*, of 0.25 mm, a central conductor width W_f, of 6.435 mm and slot length *L*, of 10.32 mm. For ε_r=6.15 and *H*= 0.635 mm a 50 Ω CPW line is used with a slot width *G*, of 0.25 mm, a central conductor width W_f, of 1.725 mm and slot length *L*, of 7.634 mm. Considering all the experimental conditions of the five factors, it is necessary to run 32 simulations. In order to make regression coefficients comparable, factor physical dimensions need to be transformed into coded factors: low level as -1 and high level as 1. Additionally, the combination of ε_r and *W* will be coded as shown in Table 5-2.

Table 5-2 Coding for the combination of ε_r and W

ε_r	*W*	Indicator Variable
-1	-1	T_1
1	-1	T_2
-1	1	T_3
1	1	T_4

Only one of these combinations is present in the model:

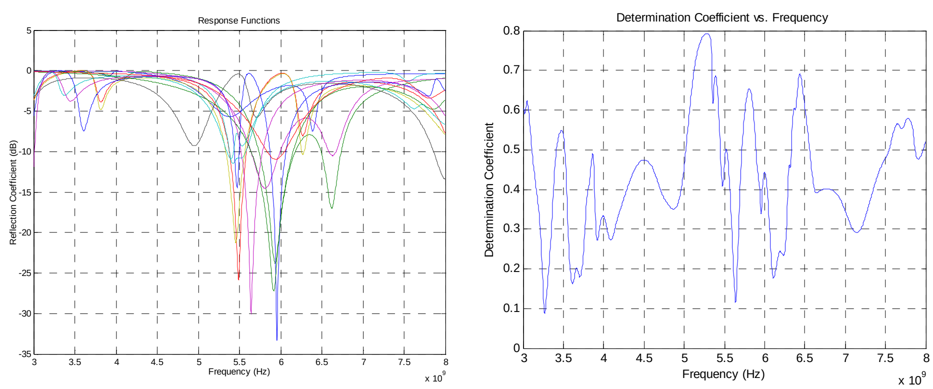
$$y = B_0 + B_1 T_1 + B_2 T_2 + B_3 T_3 + B_4 L_1 + B_5 L_2 + B_6 L_s \tag{5.1}$$

Where *y* is the response to be analyzed, in this case it is the reflection coefficient; $T_i, i = 0,...4$ is one of the experimental condition shown in table 5.2, and only three of these variables are shown in Equation 5.1 because are indicator variables that used T_4 as the base and $B_j, j = 0,...3$ are the regression coefficients in the model. The range of frequency where the experiments were performed is from 5.7 GHz to 6GHz.

5.3 Results for the Applied Case Study

5.3.1 Point-Wise Method

Figure 5-3(a) shows some of the 32 responses that correspond to each experimental condition. The frequency range in the graphs is form 3 GHz to 8 GHz. The Point-Wise method was applied. Only in some frequency intervals the model was reliable enough to derive reasonable conclusions. The selection of the frequency range was based in the adjusted determination coefficient for the model and the frequency of design. This coefficient measures the variability of the data explained by the model and its value is between zero and one. Figure 5-3(b) shows the determination coefficients for the frequency considered.



(a) (b)

Figure 5-3 (a) Observed responses, (b) Determination Coefficient.

The selected region is from 5.7 to 6 GHz. This range corresponds to the frequencies used for the design of the antenna. The Figure 5-4 shows the response functions for the selected range.

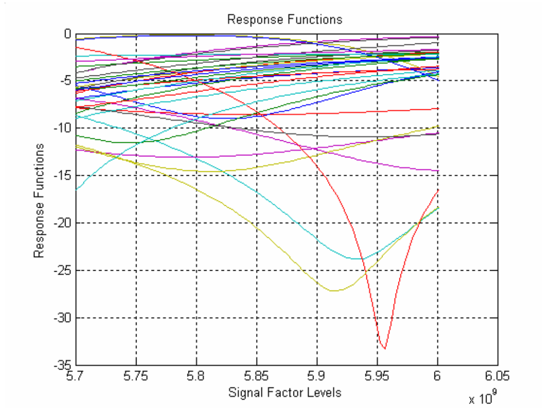
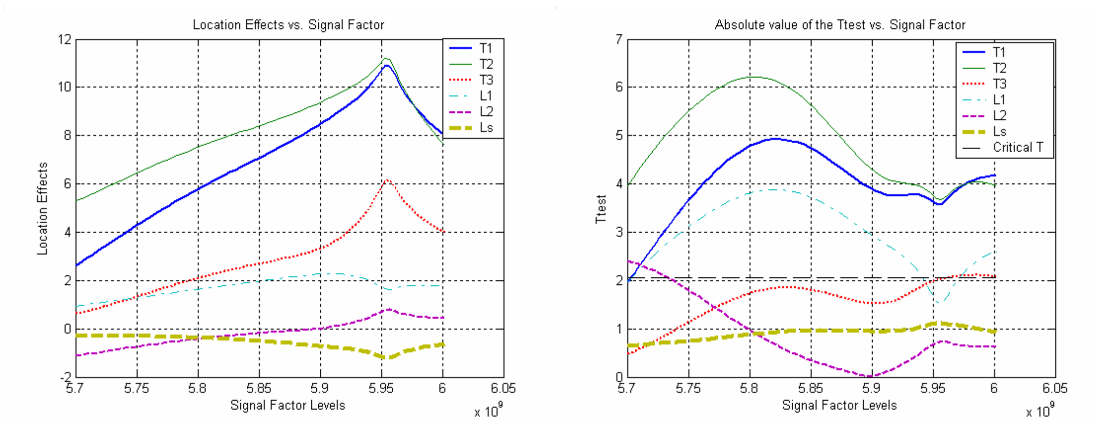


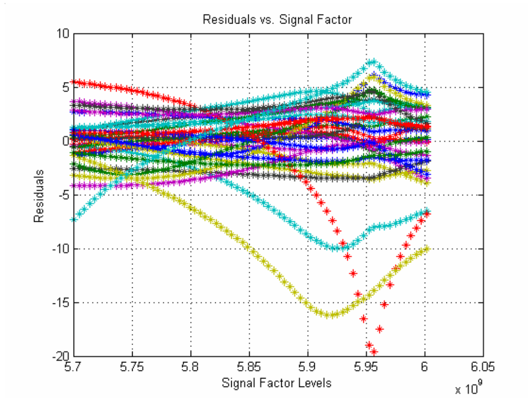
Figure 5-4 Response functions for the selected range of frequencies



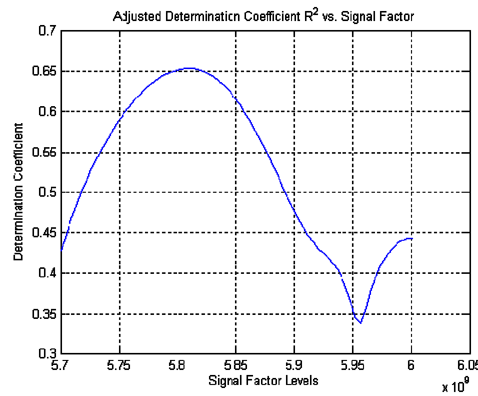
(a) (b)

Figure 5-5 (a) Factor effects, (b) Absolute value of the T test.

From Figure 5-5(b) it is possible to infer that the factors T_2, T_1 and L_1 are the most significant effects in the selected region. Because T_1 and T_2 are indicator variables that represent levels of the same factor T, which is a combination of ε_r and W they cannot coexist. T_2 has the strongest influence on the response and that is the level that has to be selected over this region. The others factors such as L_s, L_2, and T_3 have low impact on the reflection coefficient. The intercept is not shown in the figures in order to ease the visualization of the effects; this is because the magnitude of the intercept is greater than the rest of the effects. In Figure 5-5(a), the four highest effects are positive for this reason it is necessary to set these factors to their low level in order to minimize the reflection coefficient as it is desired. Some additional plots related with the implementation of this method are presented in the next figures.



(a)



(b)

Figure 5-6 Residuals and Adjusted Determination Coefficients Plots

From Figure 5-6(a) it is possible to observe that the residuals look constant for almost all the responses. A few ones make some changes for frequencies that are higher than 5.9 GHz. The Figure 5-6(b) shows the determination coefficients for the selected range of frequencies. This coefficients drop dramatically for frequencies higher than 5.85 GHz. Table 5-3 shows the pseudo-MSE generated by the applications, necessary to ease the comparison of the methods.

Table 5-3 Pseudo-MSE for the Applied Case Using the Point-Wise Method

Total SSE	Total Number of Parameters	Total Number of Data Points	Pseudo MSE
22985.9	476	2176	13.521

5.3.2 Basis Representation Method

5.3.2.1 Monomial Basis

This type of basis was the first attempt to fit the selected range of frequencies the following series of figures tables and outputs shows the results obtained. The first monomial base used was expanded two times (k = 2) terms in the following form

$$\begin{aligned} y(t) &= B_{00} + B_{10}T_1 + B_{20}T_2 + B_{30}T_3 + B_{40}L_1 + B_{50}L_2 + B_{60}L_s + \\ &\left(B_{01} + B_{11}T_1 + B_{21}T_2 + B_{31}T_3 + B_{41}L_1 + B_{51}L_2 + B_{61}L_s\right)t + \\ &\left(B_{02} + B_{12}T_1 + B_{22}T_2 + B_{32}T_3 + B_{42}L_1 + B_{52}L_2 + B_{62}L_s\right)t^2 \end{aligned} \tag{5.2}$$

The Table 5-4 shows all the terms and the factors associated with them.

Table 5-4 Relation between the Matlab's output column "Term" and the coefficients of Equation 5.2

Terms	Coefficient Represented	Related Factor	Basis Term
0	B_{00}	1	1
1	B_{10}	T1	1
2	B_{20}	T2	1
3	B_{30}	T3	1
4	B_{40}	L1	1
5	B_{50}	L2	1
6	B_{60}	Ls	1
7	B_{01}	1	t
8	B_{11}	T1	t
9	B_{21}	T2	t
10	B_{31}	T3	t
11	B_{41}	L1	t
12	B_{51}	L2	t
13	B_{61}	Ls	t
14	B_{02}	1	t^2
15	B_{12}	T1	t^2
16	B_{22}	T2	t^2
17	B_{32}	T3	t^2
18	B_{42}	L1	t^2
19	B_{52}	L2	t^2
20	B_{62}	Ls	t^2

The initial output for this expansion it is presented in Table 5-5.

Table 5-5 Initial results for the applied case using the basis representation method with a Monomial Basis (K = 2)

```
%-------------------------------------------------------------------------%
  Results for the Analysis of Variance for the Basis Representation Model
%-------------------------------------------------------------------------%
Type of Basis:   monomial
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-------------------------------------------------------------------------%
  Term    Intercept   SE Coef   T-test   P-value
    0     1249.2824   691.0816   1.8077  0.070789

  Term   Coefficients SE Coef   T-test   P-value    VIF
       1     -2356.3     977.34    -2.411  0.015993   3.61E+07
       2     -2585.8     977.34   -2.6458   0.00821   3.61E+07
       3     -31.053     977.34  -0.03177   0.97466   3.61E+07
       4     -1140.1     345.54   -3.2995  0.000985   2.40E+07
       5     -138.89     345.54  -0.40194   0.68777   2.40E+07
       6      -18.24     345.54  -0.05279   0.95791   2.40E+07
       7   -4.18E-07   2.36E-07    -1.771  0.076694      87744
       8    7.83E-07   3.34E-07    2.3446  0.019135   1.44E+08
       9    8.72E-07   3.34E-07    2.6089  0.009146   1.44E+08
      10   -4.23E-09   3.34E-07  -0.01266    0.9899   1.44E+08
      11    3.87E-07   1.18E-07    3.2783  0.001061   9.62E+07
      12    4.16E-08   1.18E-07   0.35248   0.72451   9.62E+07
      13    8.62E-09   1.18E-07  0.072996   0.94182   9.62E+07
      14    3.47E-17   2.02E-17     1.719  0.085767      87744
      15   -6.48E-17   2.86E-17   -2.2712  0.023236   3.61E+07
      16   -7.32E-17   2.86E-17   -2.5636  0.010427   3.61E+07
      17    1.71E-18   2.86E-17  0.060034   0.95213   3.61E+07
      18   -3.28E-17   1.01E-17    -3.252  0.001164   2.41E+07
      19   -3.06E-18   1.01E-17  -0.30352   0.76152   2.41E+07
      20   -9.58E-19   1.01E-17  -0.09486   0.92443   2.41E+07

%-------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.58379  0.57992
%-------------------------------------------------------------------------%
                      Analysis of Variance
%-------------------------------------------------------------------------%
Source            DF         SS          MS            F       P
Regression        20    32673.959    1633.698      151.131     0
Residual Error   2155    23295.1436    10.8098
Total            2175    55969.1026
%-------------------------------------------------------------------------%
```

The outputs shows that a serious multicollinearity problem. As an attempt to solve this situation, some terms are eliminated. The final output for the simplified model is shown next.

Table 5-6 Stepwise procedure results for the applied case using the basis representation method with a Monomial Basis (K = 2)

```
The model is not a good one due a multicollinearity problem
%-------------------------------------------------------------------------%
  Results for the Analysis of Variance for the Basis Representation Model
%-------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-------------------------------------------------------------------------%
 Term    Intercept    SE Coef   T-test   P-value
    0     16.5949      6.6728   2.4869   0.01296

 Term    Coefficients SE Coef  T-test    P-value     VIF
       1     -1123.7     698.67   -1.6083    0.10792   1.80E+07
       2     -1353.1     698.67   -1.9367   0.052914   1.80E+07
       4     -1140.1     349.32   -3.2638  0.0011166   2.40E+07
       5     -34.019     4.7184   -7.2099   7.71E-13     4385.1
       6      14.536     4.7184    3.0806  0.0020918     4385.1
       7   -4.65E-09   1.14E-09   -4.0794   4.68E-05     2.0005
       8    3.70E-07   2.39E-07    1.5475    0.12188   7.22E+07
       9    4.58E-07   2.39E-07    1.9172   0.055343   7.22E+07
      10    4.99E-10   3.44E-11    14.494          0     1.5001
      11    3.87E-07   1.19E-07    3.2428  0.0012015   9.62E+07
      12    5.78E-09   8.06E-10    7.1742   9.95E-13     4385.1
      13   -2.58E-09   8.06E-10    -3.203  0.0013796     4385.1
      15   -3.01E-17   2.04E-17   -1.4768    0.13988   1.81E+07
      16   -3.85E-17   2.04E-17   -1.8859   0.059443   1.81E+07
      18   -3.28E-17   1.02E-17   -3.2168  0.0013154   2.41E+07

%-------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.57365  0.57068
%-------------------------------------------------------------------------%
                      Analysis of Variance
%-------------------------------------------------------------------------%
Source            DF         SS          MS           F         P
Regression        15     32106.4339     2140.4289   193.7473     0
Residual Error   2160    23862.6688      11.0475
Total            2175    55969.1026
%-------------------------------------------------------------------------%
```

Table 5-6 is the result after using the stepwise procedure of the application. The graphs that correspond to the stepwise procedure are shown in Figure 5-7.

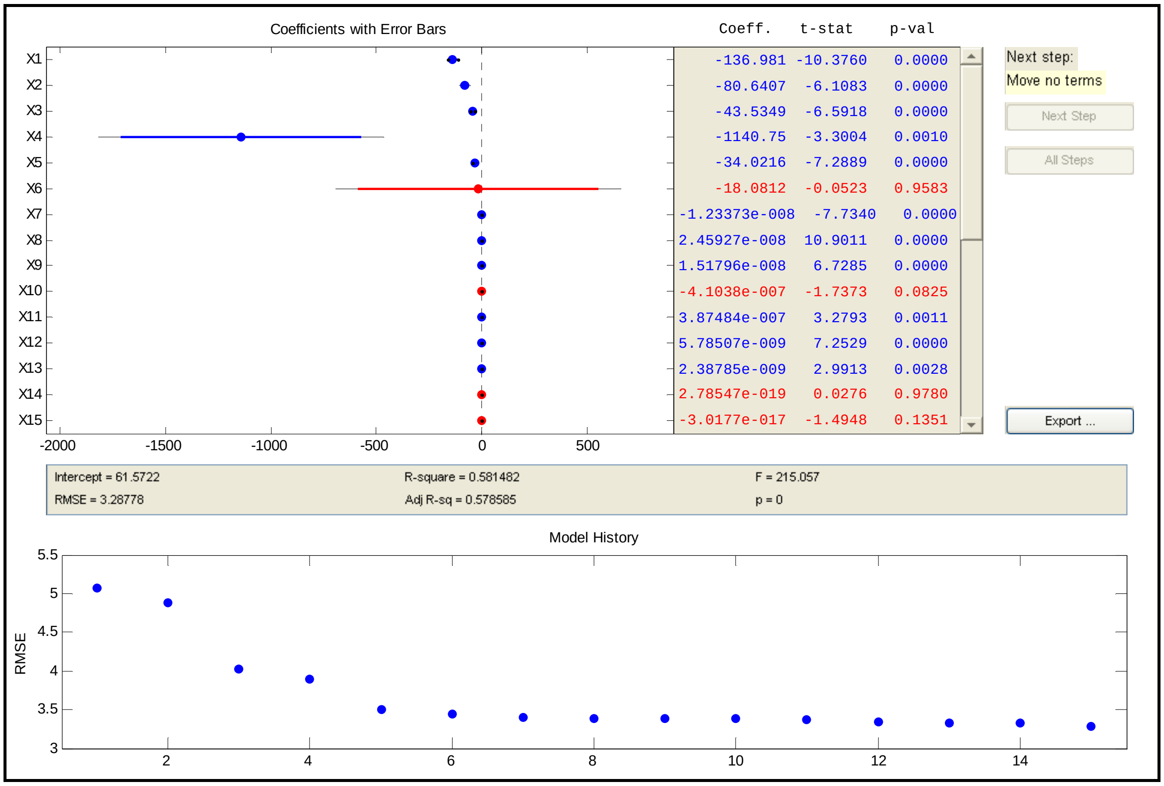


Figure 5-7 Final Results for the Monomial Basis Representation Model with k =2.

The results for the basis representation method are not reliable with the use of a monomial basis with k = 2. A severe multicollinearity problem is present and for that reason, the inferences that can be drawn for the model are not correct. It is not appropriate to use this model to describe the region under study. In general, the monomial basis is not a good one to analyze the range of frequencies selected in this experiment. In the next section, a Fourier Basis is used to analyze the experiment under study.

5.3.2.2 Fourier Basis

As a second option a Fourier basis it is going to be used in order to analyze the experiment under study. The expansion of terms selected was k = 2 and the general model has the form shown in Equation 5.3

$$\begin{aligned}
y(t) = {} & B_{00} + B_{10}T_1 + B_{20}T_2 + B_{30}T_3 + B_{40}L_1 + B_{50}L_2 + B_{60}L_s + \\
& (B_{01} + B_{11}T_1 + B_{21}T_2 + B_{31}T_3 + B_{41}L_1 + B_{51}L_2 + B_{61}L_s)\sin(\omega t) + \\
& (B_{02} + B_{12}T_1 + B_{22}T_2 + B_{32}T_3 + B_{42}L_1 + B_{52}L_2 + B_{62}L_s)\cos(\omega t) + \\
& (B_{03} + B_{13}T_1 + B_{23}T_2 + B_{33}T_3 + B_{43}L_1 + B_{53}L_2 + B_{63}L_s)\sin(2\omega t) + \\
& (B_{04} + B_{14}T_1 + B_{24}T_2 + B_{34}T_3 + B_{44}L_1 + B_{54}L_2 + B_{64}L_s)\cos(2\omega t)
\end{aligned} \tag{5.3}$$

Table 5-7 relates the terms in Equation 5.3 with the Matlab output

Table 5-7 Relation between the Matlab's output column "Term" and the coefficients of Equation 5.3

Terms	Coefficient Represented	Related Factor	Basis Term
0	B_{00}	1	1
1	B_{10}	T1	1
2	B_{20}	T2	1
3	B_{30}	T3	1
4	B_{40}	L1	1
5	B_{50}	L2	1
6	B_{60}	Ls	1
7	B_{01}	1	sin(wt)
8	B_{11}	T1	sin(wt)
9	B_{21}	T2	sin(wt)
10	B_{31}	T3	sin(wt)
11	B_{41}	L1	sin(wt)
12	B_{51}	L2	sin(wt)
13	B_{61}	Ls	sin(wt)
14	B_{02}	1	cos(wt)
15	B_{12}	T1	cos(wt)
16	B_{22}	T2	cos(wt)
17	B_{32}	T3	cos(wt)
18	B_{42}	L1	cos(wt)
19	B_{52}	L2	cos(wt)
20	B_{62}	Ls	cos(wt)
21	B_{03}	1	sin(2wt)
22	B_{13}	T1	sin(2wt)
23	B_{23}	T2	sin(2wt)
24	B_{33}	T3	sin(2wt)
25	B_{43}	L1	sin(2wt)
26	B_{53}	L2	sin(2wt)
27	B_{63}	Ls	sin(2wt)
28	B_{04}	1	cos(2wt)
29	B_{14}	T1	cos(2wt)
30	B_{24}	T2	cos(2wt)
31	B_{34}	T3	cos(2wt)
32	B_{44}	L1	cos(2wt)
33	B_{54}	L2	cos(2wt)
34	B_{64}	Ls	cos(2wt)

The initial output for this model as follows.

Table 5-8 Initial results for the applied case using the basis representation method with a Fourier Basis (K = 2)

```
%-------------------------------------------------------------------------%
  Results for the Analysis of Variance for the Basis Representation Model
%-------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-------------------------------------------------------------------------%
   Term    Intercept     SE Coef   T-test  P-value
    0       -10.625     0.14205   -74.799        0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    1.0000    6.9295    0.2009   34.4951         0    1.5012
    2.0000    8.1937    0.2009   40.7881         0    1.5012
    3.0000    2.9021    0.2009   14.4464         0    1.5012
    4.0000    1.7321    0.0710   24.3883         0    1.0008
    5.0000   -0.1711    0.0710   -2.4093    0.0161    1.0008
    6.0000   -0.5764    0.0710   -8.1153    0.0000    1.0008
    7.0000    1.0962    0.2013    5.4461    0.0000    4.0019
    8.0000   -1.8493    0.2846   -6.4969    0.0000    2.0012
    9.0000   -1.1199    0.2846   -3.9343    0.0001    2.0012
   10.0000   -1.3876    0.2846   -4.8749    0.0000    2.0012
   11.0000   -0.1000    0.1006   -0.9933    0.3207    1.0006
   12.0000   -0.4431    0.1006   -4.4025    0.0000    1.0006
   13.0000    0.2775    0.1006    2.7572    0.0059    1.0006
   14.0000    1.0479    0.2005    5.2266    0.0000    4.0028
   15.0000   -1.9657    0.2835   -6.9326    0.0000    2.0018
   16.0000   -1.5098    0.2835   -5.3249    0.0000    2.0018
   17.0000   -0.8698    0.2835   -3.0678    0.0022    2.0018
   18.0000   -0.4797    0.1002   -4.7855    0.0000    1.0010
   19.0000   -0.3354    0.1002   -3.3454    0.0008    1.0010
   20.0000    0.1634    0.1002    1.6303    0.1032    1.0010
   21.0000   -0.2485    0.1997   -1.2443    0.2135    4.0032
   22.0000    0.1580    0.2824    0.5596    0.5758    2.0022
   23.0000    0.3145    0.2824    1.1139    0.2655    2.0022
   24.0000   -0.0257    0.2824   -0.0910    0.9275    2.0022
   25.0000    0.0502    0.0998    0.5025    0.6154    1.0012
   26.0000   -0.0362    0.0998   -0.3623    0.7172    1.0012
   27.0000   -0.0258    0.0998   -0.2582    0.7963    1.0012
   28.0000    0.7567    0.2021    3.7446    0.0002    4.0005
   29.0000   -1.3524    0.2858   -4.7323    0.0000    2.0003
   30.0000   -0.9047    0.2858   -3.1657    0.0016    2.0003
   31.0000   -1.0443    0.2858   -3.6540    0.0003    2.0003
   32.0000   -0.0491    0.1010   -0.4864    0.6267    1.0002
   33.0000   -0.3559    0.1010   -3.5227    0.0004    1.0002
   34.0000    0.1408    0.1010    1.3934    0.1636    1.0002

%-------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.58046  0.5738
%-------------------------------------------------------------------------%
                      Analysis of Variance
%-------------------------------------------------------------------------%
Source            DF         SS          MS           F         P
Regression        34     32487.8485     955.525     87.1239     0
Residual Error   2141     23481.2541     10.9674
Total            2175     55969.1026
%-------------------------------------------------------------------------%
```

The stepwise procedure was implemented in order to improve the model. The final output and the stepwise procedure graphs are shown in the next.

Table 5-9 Stepwise procedure results for the applied case using the basis representation method with a Fourier Basis (K = 2)

```
%--------------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%--------------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%--------------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     -10.6173      0.14224   -74.644   0

   Term    Coefficients  SE Coef   T-test  P-value   VIF
    1.0000    6.9236    0.2012   34.4180        0    1.5006
    2.0000    8.1853    0.2012   40.6898        0    1.5006
    3.0000    2.8931    0.2011   14.3834        0    1.5003
    4.0000    1.7323    0.0711   24.3583        0    1.0003
    5.0000   -0.1703    0.0711   -2.3951    0.0167    1.0005
    6.0000   -0.5736    0.0711   -8.0665    0.0000    1.0001
    7.0000    1.1084    0.2016    5.4980    0.0000    4.0005
    8.0000   -1.8821    0.2851   -6.6015    0.0000    2.0006
    9.0000   -1.1408    0.2851   -4.0012    0.0001    2.0006
   10.0000   -1.3945    0.2851   -4.8916    0.0000    2.0003
   12.0000   -0.4507    0.1008   -4.4708    0.0000    1.0004
   13.0000    0.2775    0.1008    2.7534    0.0059    1.0001
   14.0000    0.6100    0.1419    4.2974    0.0000    2.0005
   15.0000   -1.5093    0.2459   -6.1389    0.0000    1.5007
   16.0000   -1.0726    0.2459   -4.3629    0.0000    1.5007
   18.0000   -0.4775    0.1004   -4.7585    0.0000    1.0003
   19.0000   -0.3249    0.1004   -3.2367    0.0012    1.0006
   28.0000    0.7715    0.2023    3.8130    0.0001    4.0002
   29.0000   -1.3618    0.2861   -4.7593    0.0000    2.0002
   30.0000   -0.9227    0.2861   -3.2246    0.0013    2.0002
   31.0000   -1.0480    0.2861   -3.6626    0.0003    2.0001
   33.0000   -0.3536    0.1012   -3.4951    0.0005    1.0001

%--------------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%--------------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.57676  0.57243
%--------------------------------------------------------------------------------%
                      Analysis of Variance
%--------------------------------------------------------------------------------%
Source            DF          SS          MS           F         P
Regression        22     32280.6522    1467.3024     133.36     0
Residual Error   2153     23688.5236      11.0026
Total            2175     55969.1758
%--------------------------------------------------------------------------------%
```

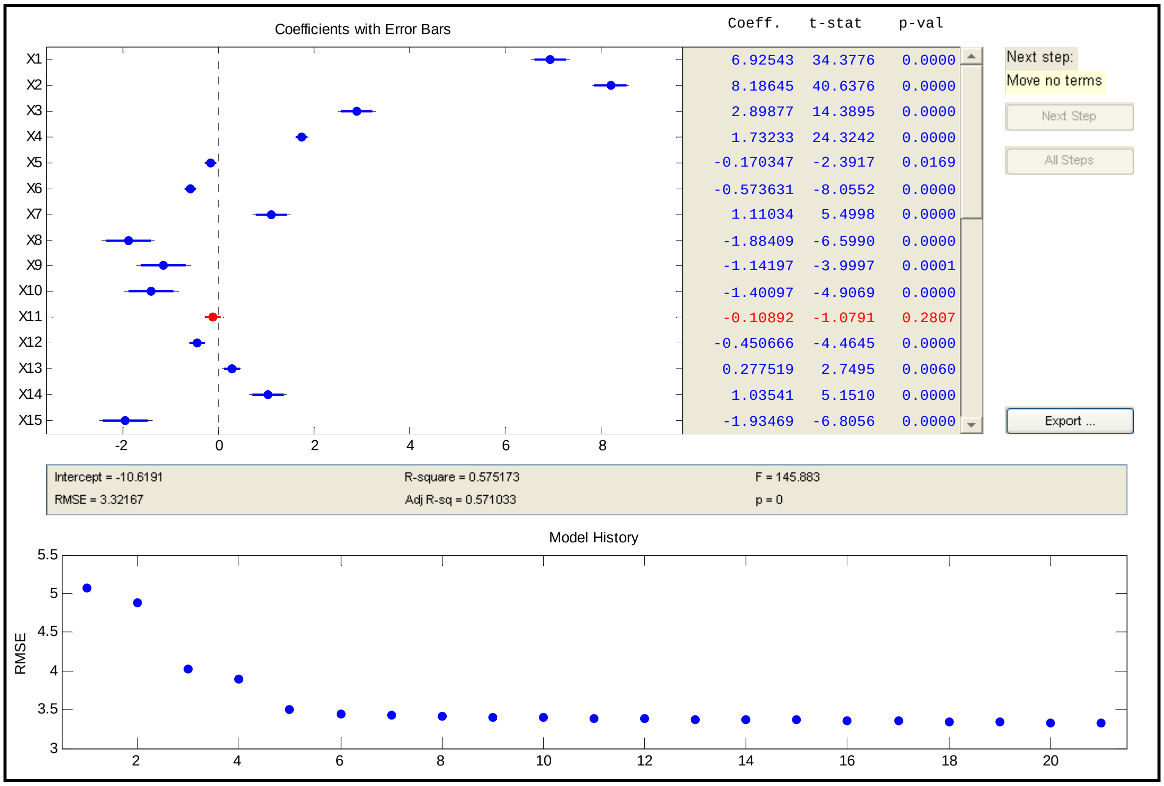


Figure 5-8 Final Results for the Fourier Basis Representation Model with k =2.

The final model obtained is not good at all, but at least it is able to describe approximately the 60% of the variability of the data. The most relevant terms of this model are summarized in Table 5-10.

Table 5-10 Results Summary for Fourier Basis Representation for the Applied Case

Terms	Coefficient Represented	Value	Related Factor	Basis Term
1	B_{10}	6.9223	T1	1
2	B_{20}	8.1844	T2	1
3	B_{30}	2.8924	T3	1
4	B_{40}	1.7321	L1	1
5	B_{50}	-0.1706	L2	1
6	B_{60}	-0.576	Ls	1
7	B_{01}	1.0883	1	sin(wt)
8	B_{11}	-1.8406	T1	sin(wt)
9	B_{21}	-1.1086	T2	sin(wt)
10	B_{31}	-1.3759	T3	sin(wt)
12	B_{51}	-0.4436	L2	sin(wt)
13	B_{61}	0.2771	Ls	sin(wt)
14	B_{02}	0.6157	1	cos(wt)
15	B_{12}	-1.5345	T1	cos(wt)
16	B_{22}	-1.0821	T2	cos(wt)
18	B_{42}	-0.4798	L1	cos(wt)
19	B_{52}	-0.3346	L2	cos(wt)
28	B_{04}	0.7603	1	cos(2wt)
29	B_{14}	-1.3564	T1	cos(2wt)
30	B_{24}	-0.9099	T2	cos(2wt)
31	B_{34}	-1.0496	T3	cos(2wt)
33	B_{54}	-0.3557	L2	cos(2wt)

The terms presented in the table shows that the most relevant factors in the experiment are T1, T2, T3, and L1. These factors are the same factors considered as relevant by the Piece-Wise approach. As mentioned before the factor Ls is completely non-relevant to the experiment. Most of the terms that include Ls were eliminated by the stepwise procedure. The plots that correspond to the residuals, estimated response are presented in the next figures.

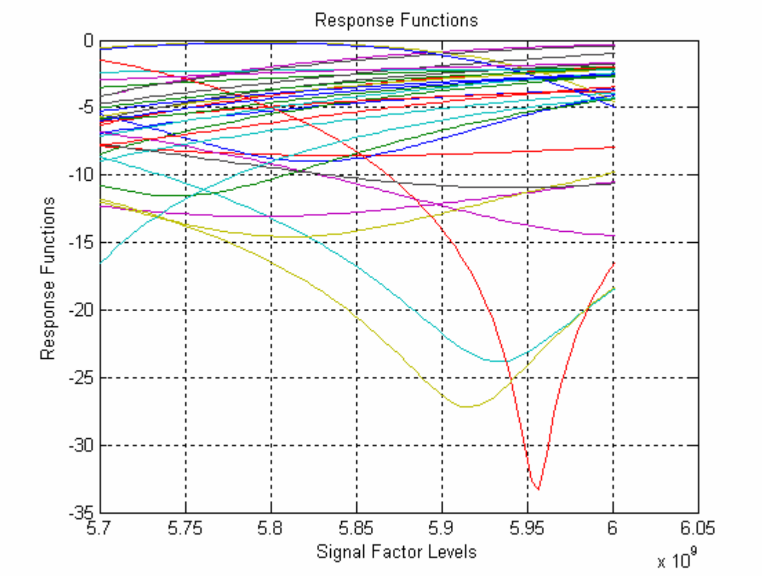


Figure 5-9 Responses for the Basis Representation Final Model with a Fourier Basis for the Applied Case Study.

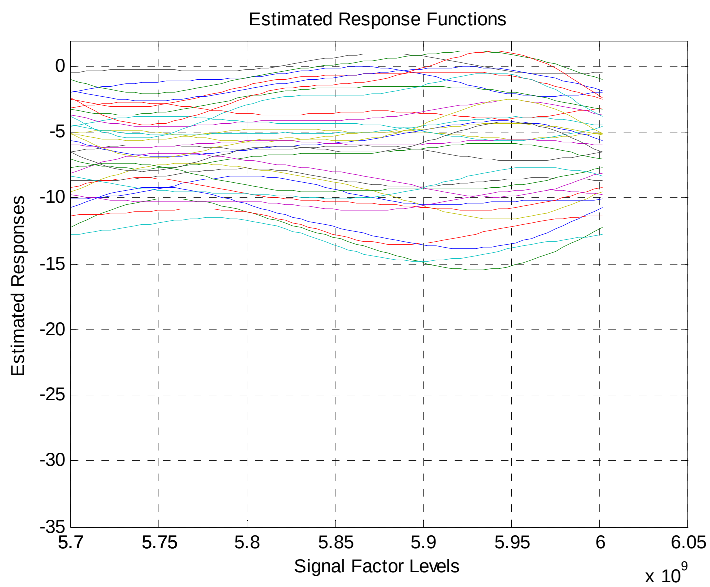


Figure 5-10 Estimated responses for the Basis Representation Final Model with a Fourier Basis for the Applied Case Study.

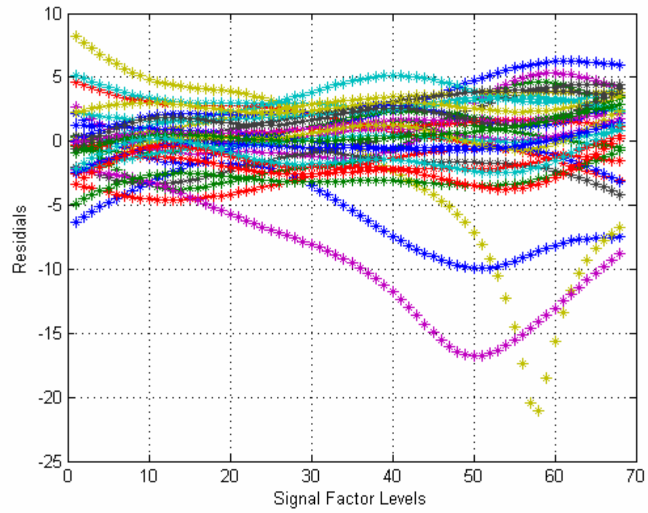


Figure 5-11 Residuals for the Basis Representation Final Model with a Fourier Basis for the Applied Case Study.

From Figure 5-10 it is possible to observe the lack of fit of the models. The selected basis is unable to fit the local features of some of the responses. The metrics for this procedure are summarized in the following table

Table 5-11 Metrics for the Applied Case Using the Basis Representation Method using a Fourier Basis with k=2

Total SSE	Total Number of Parameters	Total Number of Data Points	Pseudo MSE
23617.8	26	2176	10.985

5.3.2.3 Piece-Wise Method without the Continuity Constraint

In order to complete the comparison among the methods, the Piece-Wise method is applied for this applied case study. Only the Fourier basis it is going to be used. The main reason for this is the poor performance of the monomial basis and the cubic spline in the previous method. Due the high computational effort of the knots search procedure the quantities of knots tested were 1, 2, and 3. Table 5-12 shows the results of the search procedure.

Table 5-12 Knots Search for the Piece-Wise method for the applied case study

Fourier Basis K=2

Number of knots	Objective function	optimal knots		
3	**23066**	**10**	**22**	**35**
2	23085	14	31	*
1	23142	24	*	*
Minimal	**23066**			

The selected knots were 10, 22, and 35. The general model in each region is exactly the same model presented as Equation 5.3. Figure 5-12 presents the four regions delimited by the knots search.

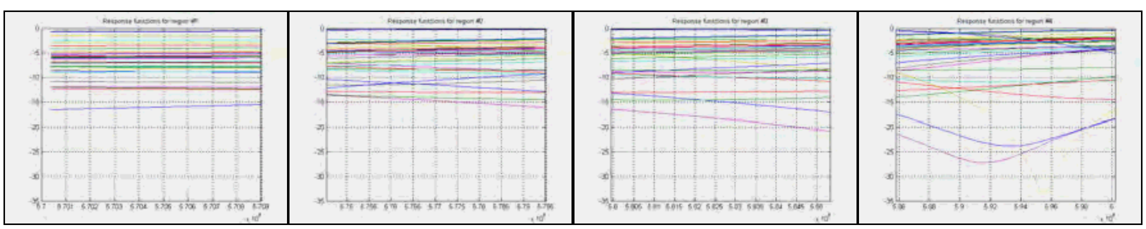


Figure 5-12 Plots for all the regions delimited before using the Piece-Wise procedure for the applied case study with a Fourier basis

The next sections show the results for this procedure in detail. As done in the previous chapter the initial model, stepwise procedure and the final model are presented per each region.

5.3.2.3.1 Results for Region #1

Table 5-13 corresponds to the initial model is shown next.

Table 5-13 Initial Output for the Piece-Wise Method using a Fourier basis for region #1

```
%-----------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model
for the region  #1
%-----------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
   Terms   Coefficients  SE Coef   T-test    P-value         VIF
    0      -8.8963      0.26472     -33.606           0           0
    1       3.2922      0.37437      8.7939  2.2204e-016      1.5476
    2        5.746      0.37437      15.348           0      1.5476
    3      0.89029      0.37437      2.3781    0.018062      1.5476
    4         1.08      0.13236      8.1595   1.088e-014      1.0317
    5      -0.9667      0.13236     -7.3035   2.818e-012      1.0317
    6     -0.29931      0.13236     -2.2613    0.024493      1.0317
    7     0.012943      0.35969    0.035984     0.97132      4.1143
    8   -0.0038653      0.50867  -0.0075988     0.99394      2.0857
    9   -0.0068824      0.50867    -0.01353     0.98921      2.0857
   10    -0.015631      0.50867   -0.030728     0.97551      2.0857
   11    0.0015116      0.17984    0.008405      0.9933      1.0476
   12    0.0031899      0.17984    0.017737     0.98586      1.0476
   13    0.0024999      0.17984      0.0139     0.98892      1.0476
   14      0.14078       0.3885     0.36237     0.71735           4
   15     -0.41488      0.54943    -0.75511      0.4508           2
   16     -0.28839      0.54943    -0.52489     0.60007           2
   17      -0.1692      0.54943    -0.30796     0.75834           2
   18    -0.089061      0.19425    -0.45848     0.64696           1
   19    -0.097349      0.19425    -0.50114     0.61666           1
   20  -0.00076938      0.19425  -0.0039607     0.99684           1
   21    -0.060006       0.3885    -0.15445     0.87736           4
   22      0.17901      0.54943     0.32581      0.7448           2
   23      0.12407      0.54943     0.22582      0.8215           2
   24     0.072177      0.54943     0.13137     0.89558           2
   25     0.038809      0.19425     0.19979     0.84179           1
   26     0.042464      0.19425      0.2186     0.82712           1
   27   0.00034518      0.19425    0.001777     0.99858           1
   28    0.0044164      0.35969    0.012278     0.99021      4.1143
   29   -0.0021857      0.50868  -0.0042968     0.99657      2.0857
   30   -0.0028494      0.50868  -0.0056017     0.99553      2.0857
   31   -0.0053194      0.50868   -0.010457     0.99166      2.0857
   32   0.00032012      0.17984     0.00178     0.99858      1.0476
   33   0.00083329      0.17984   0.0046334     0.99631      1.0476
   34   0.00078824      0.17984   0.0043829     0.99651      1.0476
%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.5986  0.55072
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF          SS          MS           F      P
Regression        34     2309.4382     67.9247     12.5006     0
Residual Error   285     1548.6063      5.4337
Total            319     3858.0445
%-----------------------------------------------------------------------------%
```

The obtained model has many non-relevant terms. The stepwise procedure was executed in order to simplify the model and eliminate the terms that are not contributing to the model.

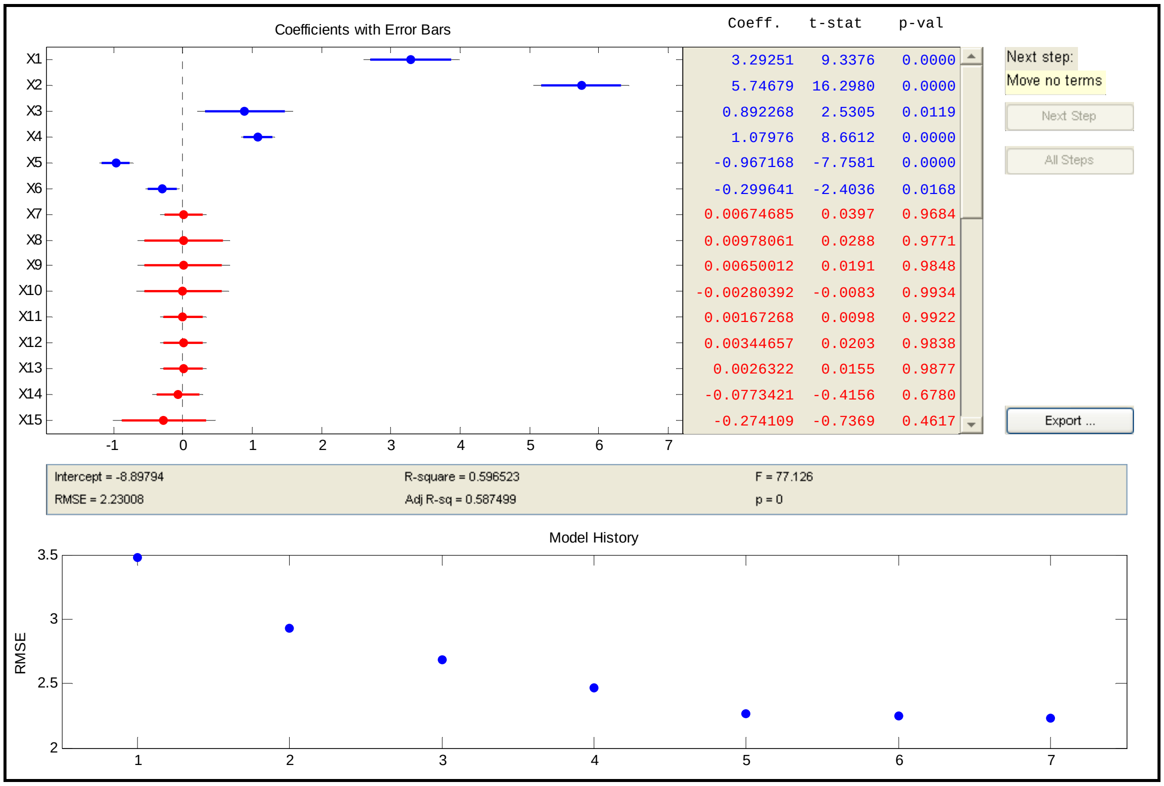


Figure 5-13 Stepwise Plots for the Region #1 of the Applied Case Study Using A Fourier basis (K=2)

The final model obtained after the stepwise procedure it is shown in Table 5-14.

Table 5-14 Final Output for the Piece-Wise Method using a Fourier basis for region #1

```
%-----------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     -8.8979      0.24933   -35.6873   0

   Term    Coefficients  SE Coef      T-test     P-value          VIF
    1       3.2925      0.35261       9.3376           0          1.5
    2       5.7468      0.35261       16.298           0          1.5
    3      0.89227      0.35261       2.5305     0.01188          1.5
    4       1.0798      0.12467       8.6612  2.2204e-016           1
    5     -0.96717      0.12467      -7.7581  1.2279e-013           1
    6     -0.29964      0.12467      -2.4036    0.016818            1

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.59652  0.58879
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source           DF          SS           MS          F       P
Regression        6     2301.4111     383.5685     77.126     0
Residual Error   313     1556.6334      4.9733
Total            319     3858.0445
%-----------------------------------------------------------------------------%
```

After the stepwise procedure all the terms that stayed in the model, are terms that do not are multiplied by a sine or cosine term. Of those terms, the less relevant is the term that correspond to the factor Ls. The factors L1, L2 and the combinations of W and ε_r have a considerable relevance in this model as shown in Table 5-15. The model is not very good. Only explains around a 60% of the total variability of the data, but most of their terms are relevant. The next section shows the results for the next region.

Table 5-15 Results Summary for Region #1

Terms	Coefficient Represented	Value	Related Factor	Basis Term
0	B_{00}	-8.8973	1	1
1	B_{10}	3.2925	T1	1
2	B_{20}	5.7468	T2	1
3	B_{30}	0.8923	T3	1
4	B_{40}	1.0798	L1	1
5	B_{50}	-0.9672	L2	1
6	B_{60}	-0.2996	Ls	1

5.3.2.3.2 Results for Region #2

Table 5-16 presents the initial model obtained for this region.

Table 5-16 Initial Output for the Piece-Wise Method using a Fourier basis for region #2

Results for the Analysis of Variance for the Basis Representation Model
for the region #2

Type of Basis: fourier
Number of basis functions or knots if the basis is a Cubic Spline: 2

Coefficients and Variance Inflation Factors

Terms	Coefficients	SE Coef	T-test	P-value	VIF
0	-9.5316	0.22526	-42.313	0	0
1	4.897	0.31857	15.372	0	1.5341
2	6.8848	0.31857	21.612	0	1.5341
3	1.6482	0.31857	5.1738	3.8761e-007	1.5341
4	1.4254	0.11263	12.655	0	1.0227
5	-0.60844	0.11263	-5.4021	1.2216e-007	1.0227
6	-0.33045	0.11263	-2.9339	0.0035689	1.0227
7	0.18672	0.3254	0.57381	0.56647	4.0266
8	-0.43232	0.46018	-0.93945	0.34815	2.0172
9	-0.30946	0.46018	-0.67246	0.50173	2.0172
10	-0.22049	0.46018	-0.47915	0.63213	2.0172
11	-0.095537	0.1627	-0.5872	0.55745	1.0092
12	-0.094815	0.1627	-0.58276	0.56043	1.0092
13	0.016254	0.1627	0.099904	0.92048	1.0092
14	-0.084179	0.31159	-0.27016	0.7872	4.0808
15	0.18479	0.44065	0.41936	0.67521	2.0572
16	0.13143	0.44065	0.29826	0.76569	2.0572
17	0.096099	0.44065	0.21808	0.82749	2.0572
18	0.043491	0.15579	0.27916	0.78029	1.0314
19	0.040125	0.15579	0.25755	0.7969	1.0314
20	-0.011005	0.15579	-0.070637	0.94373	1.0314
21	0.060941	0.31692	0.19229	0.84762	4.0685
22	-0.14238	0.44819	-0.31768	0.75092	2.0464
23	-0.10187	0.44819	-0.2273	0.82032	2.0464
24	-0.072093	0.44819	-0.16085	0.8723	2.0464
25	-0.031265	0.15846	-0.19731	0.8437	1.0252
26	-0.031201	0.15846	-0.1969	0.84402	1.0252
27	0.0049886	0.15846	0.031482	0.9749	1.0252
28	-0.070607	0.32021	-0.2205	0.82561	4.0555
29	0.16057	0.45284	0.35457	0.72312	2.0369
30	0.11457	0.45284	0.253	0.80042	2.0369
31	0.082108	0.45284	0.18132	0.85622	2.0369
32	0.036335	0.1601	0.22695	0.8206	1.02
33	0.034946	0.1601	0.21827	0.82735	1.02
34	-0.0073094	0.1601	-0.045654	0.96361	1.02

R^2 and Adjusted R^2

R^2,	R^2(adj)
0.69342	0.66355

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	34	3759.7825	110.5818	23.2164	0
Residual Error	349	1662.3155	4.7631		
Total	383	5422.0981			

There is a lot of term that are not contributing to the model. The stepwise procedure is executed in order to eliminate those terms and simplify the model. The plots that correspond to the procedure are shown next.

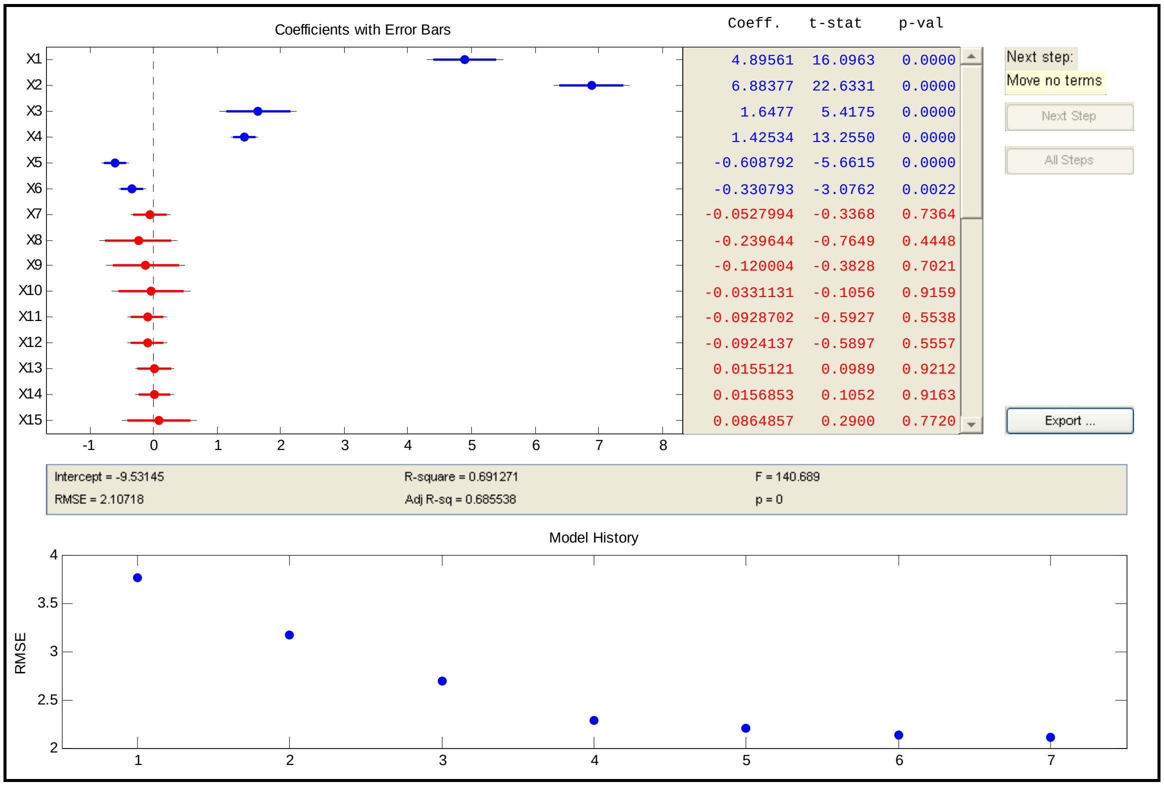


Figure 5-14 Stepwise Plots for the Region #2 of the Applied Case Study Using A Fourier basis (K=2)

The details for the final model are shown in Table 5-17.

Table 5-17 Final Output for the Piece-Wise Method using a Fourier basis for region #2

```
%-----------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%-----------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-----------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     -9.5315      0.21506   -44.3193   0

   Term    Coefficients  SE Coef     T-test       P-value     VIF
    1       4.8956      0.30415      16.096             0     1.5
    2       6.8838      0.30415      22.633             0     1.5
    3       1.6477      0.30415      5.4175  1.0798e-007     1.5
    4       1.4253      0.10753      13.255             0     1
    5     -0.60879      0.10753     -5.6615  2.9772e-008     1
    6     -0.33079      0.10753     -3.0762    0.0022493     1

%-----------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-----------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.69127  0.68636
%-----------------------------------------------------------------------------%
                      Analysis of Variance
%-----------------------------------------------------------------------------%
Source            DF        SS            MS            F       P
Regression         6     3748.1381     624.6897     140.6892     0
Residual Error   377    1673.96          4.4402
Total            383    5422.0981
%-----------------------------------------------------------------------------%
```

Table 5-18 Results Summary for Region #2

Terms	Coefficient Represented	Value	Related Factor	Basis Term
0	B_{00}	-9.5315	1	1
1	B_{10}	4.8956	T1	1
2	B_{20}	6.8838	T2	1
3	B_{30}	1.6477	T3	1
4	B_{40}	1.4253	L1	1
5	B_{50}	-0.6088	L2	1
6	B_{60}	-0.3308	Ls	1

The same terms selected by this procedure in the previous region were selected for this one. Those terms are the terms that relate W and ε_r, and the terms that correspond to L1 and L2. There is a term that corresponds to Ls but it has the lowest magnitude among the other coefficients. This model is better than the previous one in terms of the adjusted determination coefficient, which is close to a 70%. The execution of the Piece-Wise method on region #3 is shown in the next section.

5.3.2.3.3 Results for Region #3

Table 5-19 corresponds to the initial model obtained for this region.

Table 5-19 Initial Output for the Piece-Wise Method using a Fourier basis for region #3

```
%---------------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for
the region  #3
%--------------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%---------------------------------------------------------------------------------%
                    Coefficients and Variance Inflation Factors
%---------------------------------------------------------------------------------%
          Terms   Coefficients  SE Coef   T-test         P-value         VIF
           0      -10.256      0.24538     -41.796             0           0
           1        6.472      0.34702       18.65             0      1.5294
           2       7.9918      0.34702       23.03             0      1.5294
           3       2.4279      0.34702      6.9965  1.1869e-011      1.5294
           4       1.8058      0.12269      14.719             0      1.0196
           5     -0.28722      0.12269      -2.341     0.019746      1.0196
           6     -0.43949      0.12269     -3.5821   0.00038498      1.0196
           7     -0.16884      0.34988    -0.48256      0.62969      4.0435
           8      0.33825       0.4948     0.68361      0.49463      2.0285
           9      0.23096       0.4948     0.46677      0.64093      2.0285
          10      0.15069       0.4948     0.30455      0.76088      2.0285
          11     0.089698      0.17494     0.51274      0.60843      1.0154
          12     0.057168      0.17494     0.32679      0.74401      1.0154
          13    -0.037958      0.17494    -0.21698      0.82834      1.0154
          14     -0.12594      0.34414    -0.36594      0.71461      4.0626
          15      0.26777      0.48669     0.55018      0.58252      2.0426
          16       0.1856      0.48669     0.38136      0.70315      2.0426
          17      0.12941      0.48669     0.26591      0.79045      2.0426
          18      0.06832      0.17207     0.39704      0.69156      1.0232
          19     0.051957      0.17207     0.30195      0.76285      1.0232
          20    -0.023156      0.17207    -0.13457      0.89302      1.0232
          21      0.02703      0.33756    0.080074      0.93622       4.072
          22    -0.051285      0.47739    -0.10743       0.9145       2.052
          23    -0.034725      0.47739    -0.07274      0.94205       2.052
          24    -0.021411      0.47739   -0.044851      0.96425       2.052
          25    -0.013858      0.16878   -0.082105      0.93461      1.0286
          26   -0.0076899      0.16878   -0.045561      0.96368      1.0286
          27    0.0067976      0.16878    0.040274       0.9679      1.0286
          28     0.095049      0.35623     0.26682      0.78975      4.0098
          29     -0.19475      0.50378    -0.38658      0.69928      2.0062
          30     -0.13408      0.50378    -0.26615      0.79027      2.0062
          31    -0.090036      0.50378    -0.17872      0.85825      2.0062
          32    -0.050499      0.17811    -0.28352      0.77693      1.0033
          33    -0.035083      0.17811    -0.19697      0.84396      1.0033
          34     0.019556      0.17811     0.10979      0.91263      1.0033

%---------------------------------------------------------------------------------%
                             R^2 and Adjusted R^2
%---------------------------------------------------------------------------------%
          R^2,    R^2(adj)
        0.70805  0.68199
%---------------------------------------------------------------------------------%
                             Analysis of Variance
%---------------------------------------------------------------------------------%
       Source           DF          SS          MS           F       P
       Regression       34     5674.8544    166.9075     27.1766     0
       Residual Error   381    2339.9442     6.1416
       Total            415    8014.7985
%---------------------------------------------------------------------------------%
```

Once again, the model has a good quantity of terms that are not contributing to the model. The stepwise procedure once again provides a new simplified model. The plots that correspond to this procedure are shown in Figure 5-15.

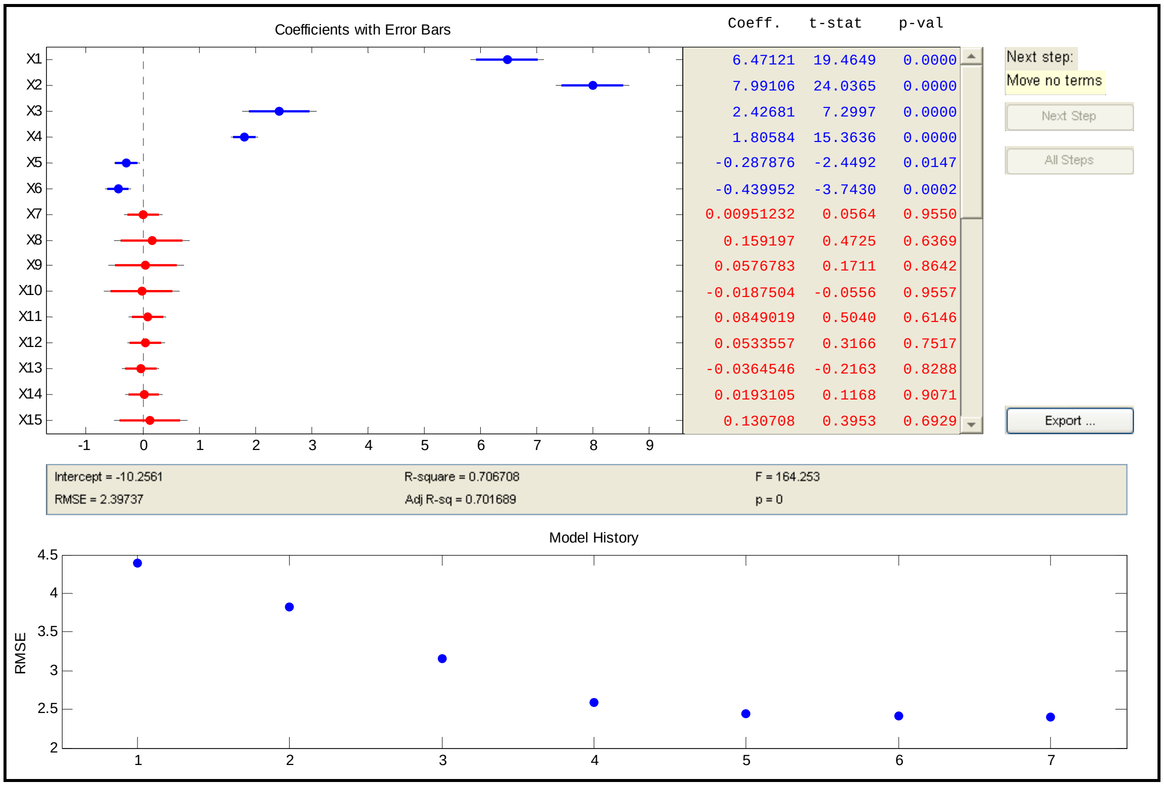


Figure 5-15 Stepwise Plots for the Region #3 of the Applied Case Study Using A Fourier basis (K=2)

The details for the final model are presented in Table 5-20.

Table 5-20 Final Output for the Piece-Wise Method using a Fourier basis for region #3

```
%-------------------------------------------------------------------------------%
      Results for the Analysis of Variance for the Basis Representation Model
%-------------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%-------------------------------------------------------------------------------%
Term    Intercept    SE Coef   T-test    P-value
    0     -10.2561      0.23508   -43.628   0

   Term  Coefficients   SE Coef     T-test        P-value    VIF
    1       6.4712      0.33245       19.465            0    1.5
    2       7.9911      0.33245       24.037            0    1.5
    3       2.4268      0.33245       7.2997  1.5155e-012    1.5
    4       1.8058      0.11754       15.364            0    1
    5     -0.28788      0.11754      -2.4492     0.014738    1
    6     -0.43995      0.11754       -3.743   0.00020787    1

%-------------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%-------------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.70671  0.70241
%-------------------------------------------------------------------------------%
                      Analysis of Variance
%-------------------------------------------------------------------------------%
Source            DF          SS          MS           F       P
Regression         6     5664.1254     944.0209     164.2528     0
Residual Error   409      2350.6731      5.7474
Total            415      8014.7985
%-------------------------------------------------------------------------------%
```

Once again, the constant terms were selected by the procedure. The model explains a 70% of the variability of the data. This is almost equal to the model for region #2. The next table presents a summary of these results. The next table summarizes these results.

Table 5-21 Results Summary for Region #3

Terms	Coefficient Represented	Value	Related Factor	Basis Term
0	B_{00}	-10.2561	1	1
1	B_{10}	6.4712	T1	1
2	B_{20}	7.9911	T2	1
3	B_{30}	2.4268	T3	1
4	B_{40}	1.8058	L1	1
5	B_{50}	-0.2879	L2	1
6	B_{60}	-0.4400	Ls	1

The next section presents the results for the region #4, which is the last one.

5.3.2.3.4 Results for Region #4

In the same way as in the previous section the Matlab output for the initial model it is presented first.

Table 5-22 Initial Output for the Piece-Wise Method using a Fourier basis for region #4

```
%----------------------------------------------------------------------------------%
Results for the Analysis of Variance for the Basis Representation Model for
the region  #4
%----------------------------------------------------------------------------------%
Type of Basis:   fourier
Number of basis functions or knots if the basis is a Cubic Spline:   2
%----------------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%----------------------------------------------------------------------------------%
   Terms   Coefficients   SE Coef       T-test        P-value       VIF
       0      -11.709      0.25535      -45.852             0           0
       1       8.9599      0.36112       24.811             0      1.5052
       2       9.4994      0.36112       26.305             0      1.5052
       3       4.1743      0.36112       11.559             0      1.5052
       4       2.0027      0.12768       15.686             0      1.0035
       5      0.27792      0.12768       2.1767       0.02973      1.0035
       6     -0.80924      0.12768      -6.3382  3.4821e-010      1.0035
       7      -1.0964      0.35937      -3.0509    0.0023407      4.0118
       8        1.342      0.50822       2.6406    0.0084026      2.0077
       9       1.1039      0.50822       2.1721     0.030077      2.0077
      10       1.1733      0.50822       2.3086     0.021167      2.0077
      11    -0.088834      0.17968      -0.4944      0.62113      1.0042
      12      0.30719      0.17968       1.7096     0.087636      1.0042
      13     -0.23103      0.17968      -1.2858      0.19881      1.0042
      14    -0.035729      0.36285    -0.098469      0.92158      4.0071
      15     -0.22989      0.51315     -0.44801      0.65424      2.0046
      16      0.23196      0.51315      0.45204      0.65133      2.0046
      17     -0.70884      0.51315      -1.3814      0.16747      2.0046
      18      0.28085      0.18142        1.548      0.12192      1.0025
      19     -0.25225      0.18142      -1.3904      0.16472      1.0025
      20     0.048552      0.18142      0.26762      0.78905      1.0025
      21      0.37463      0.35712       1.0491       0.2944      4.0127
      22     -0.32811      0.50504     -0.64968      0.51605      2.0088
      23     -0.43287      0.50504      -0.8571      0.39159      2.0088
      24     -0.26835      0.50504     -0.53135      0.59529      2.0088
      25      0.03991      0.17856      0.22351      0.82318      1.0048
      26    -0.045391      0.17856     -0.25421      0.79939      1.0048
      27      0.07497      0.17856      0.41986      0.67467      1.0048
      28      0.21428      0.36505        0.587      0.55734      4.0025
      29    -0.041564      0.51626    -0.080509      0.93585      2.0016
      30     -0.28299      0.51626     -0.54815      0.58371      2.0016
      31    -0.010145      0.51626    -0.019651      0.98433      2.0016
      32     0.020609      0.18253      0.11291      0.91013      1.0009
      33     0.033339      0.18253      0.18266       0.8551      1.0009
      34     0.017307      0.18253     0.094817      0.92448      1.0009

%----------------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%----------------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.54666  0.53157
%----------------------------------------------------------------------------------%
                      Analysis of Variance
%----------------------------------------------------------------------------------%
Source            DF          SS          MS          F         P
Regression        34     21120.4829    621.1907     36.2113     0
Residual Error   1021     17514.8358     17.1546
Total            1055     38635.3188
%----------------------------------------------------------------------------------%
```

The following graphs correspond to the stepwise procedure for this region.

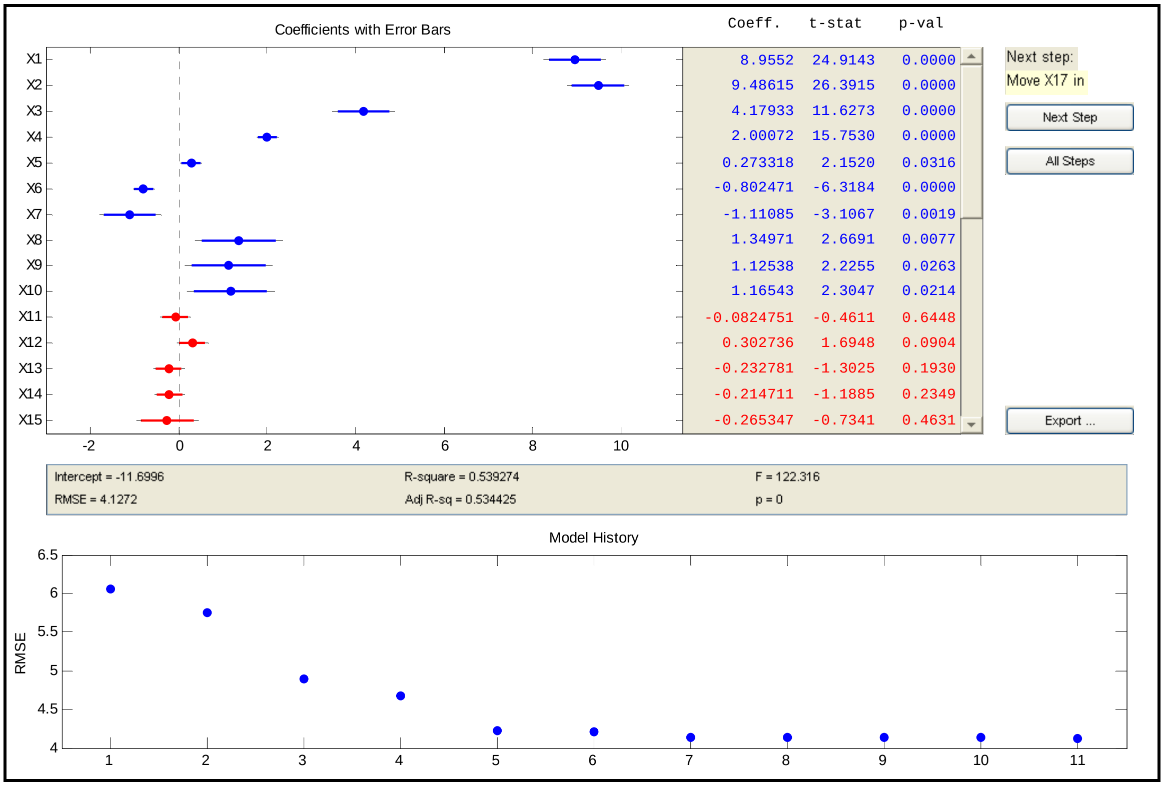


Figure 5-16 Stepwise Plots for the Region #4 of the Applied Case Study Using A Fourier basis (K=2)

The final model details are presented in the Table 5-23

Table 5-23 Final Output for the Piece-Wise Method using a Fourier basis for region #4

```
%------------------------------------------------------------------------------%
    Results for the Analysis of Variance for the Basis Representation Model
%------------------------------------------------------------------------------%
                      Coefficients and Variance Inflation Factors
%------------------------------------------------------------------------------%
Term    Intercept    SE Coef       T-test    P-value
    0     -11.6996     0.25416     -46.032   0

 Term   Coefficients  SE Coef       T-test      P-value          VIF
   1       8.9552      0.35944       24.914            0       1.5018
   2       9.4861      0.35944       26.391            0       1.5018
   3       4.1793      0.35944       11.627            0       1.5018
   4       2.0007      0.12701       15.753            0            1
   5      0.27332      0.12701        2.152     0.031625            1
   6     -0.80247      0.12701      -6.3184  3.9053e-010            1
   7      -1.1108      0.35757      -3.1067    0.0019431            4
   8       1.3497      0.50568       2.6691    0.0077231       2.0018
   9       1.1254      0.50568       2.2255     0.026262       2.0018
  10       1.1654      0.50568       2.3047      0.02138       2.0018

%------------------------------------------------------------------------------%
                      R^2 and Adjusted R^2
%------------------------------------------------------------------------------%
   R^2,    R^2(adj)
 0.53927  0.53487
%------------------------------------------------------------------------------%
                      Analysis of Variance
%------------------------------------------------------------------------------%
Source             DF          SS           MS            F          P
Regression         10     20835.036     2083.5036     122.3161     0
Residual Error   1045     17800.2828      17.0338
Total            1055     38635.3188
%------------------------------------------------------------------------------%
```

This model has more terms considered relevant due the behavior of some of the response functions in this region. For the previous regions, all the curves behave as constant functions. However, in this region some of the responses differ from that pattern. This is shown in the next figure.

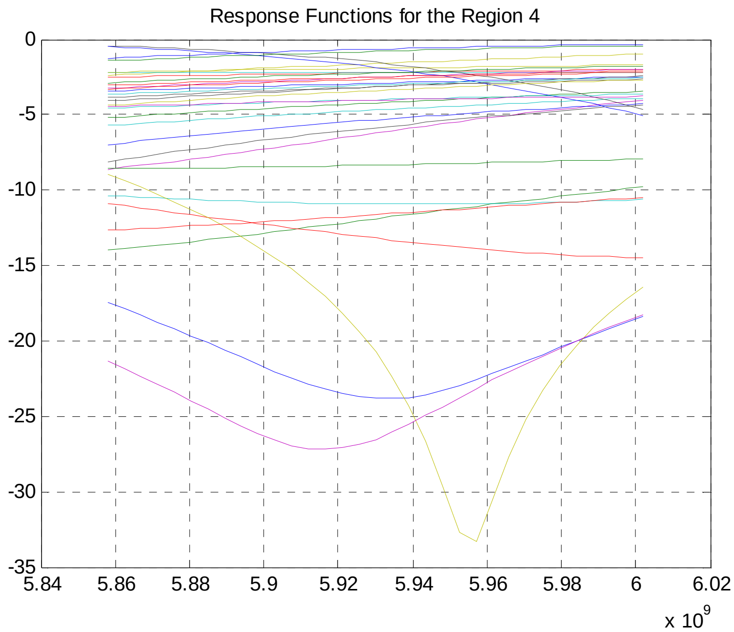


Figure 5-17 Response Functions for Region #4

For the behavior of the response functions in this region, the adjusted determination coefficient is close to a 50%. Table 5-24 summarizes the terms that were considered as relevant by this model.

Table 5-24 Results Summary for Region #4

Terms	Coefficient Represented	Value	Related Factor	Basis Term
0	B_{00}	-11.6996	1	1
1	B_{10}	8.9552	T1	1
2	B_{20}	9.4861	T2	1
3	B_{30}	4.1793	T3	1
4	B_{40}	1.9987	L1	1
5	B_{50}	0.2807	L2	1
6	B_{60}	-0.8025	Ls	1
7	B_{01}	-1.1108	1	sin(wt)
8	B_{11}	1.3497	T1	sin(wt)
9	B_{21}	1.1254	T2	sin(wt)
10	B_{31}	1.1654	T3	sin(wt)

This model considered the same factors that were pointed out as relevant in the previous regions. However, due the low determination coefficient this results are not reliable. In general, this procedure worked out in the same way than the basis representation model. The next section presents a summary of the results obtained using the Piece-Wise method.

5.3.2.4 Results Summary for the Piece-Wise Method without the Continuity Constraints

The Piece-Wise method was performed in order to analyze this applied case study. The next table shows a summary of the results obtained.

Table 5-25 Summary of the Piece-Wise Method for the Applied Case Study

Fourier Basis	Number of Terms	SSE	Number of Data Points	Adjusted R^2
Region 1	7	4.9733	320	0.58879
Region 2	7	4.4402	384	0.68636
Region 3	7	5.7474	416	0.70241
Region 4	11	17.0338	1056	0.53487
Pseudo-MSE	10.90628925			

All the models contained almost the same quantity of terms for all the regions. The only exception was the model for the region #4. The adjusted determination coefficients obtained were relatively low. In general, in all the regions the terms T1, T2, T3, L1 and L2 were considered as the most relevant ones. The next section is a comparison of the three methods under for the case that corresponds to this chapter.

5.3.3 Piece-Wise Method with the Continuity Constraints

As done with the theoretical case study the procedure used to force the curves from two adjacent regions to be joined at the knot was applied to this case. In this case, 31 indicator variables were incorporated to the model. The consequences of the high number of variables added to the model were dramatic. The effects on the design matrix were fatal. The variance-covariance matrices are completely ill-conditioned. For that reasons the obtained models are completely useless to analyze or infer with respect the factors of this experiment. For the reasons previously mentioned the results for this part of the piece-wise method are not presented.

5.4 Comparison of Results for the Three Methods for the Applied Case Study

The three methods were utilized to analyze the experiment. The pseudo-MSE it is used as a metric of comparison for the methods. The next table shows the pseudo-MSE for the three methods.

Table 5-26 Pseudo-MSE Table for all the methods applied for the applied case study

Method	Basis Used	Total SSE	Total Number of Parameters	Total Number of Data Points	Pseudo MSE
Point-Wise	N/A	22985.9	476	2176	13.521
Basis Representation	Fourier	23617.8	26	2176	10.985
Piece-Wise without Continuity Constraints	Fourier	23328.55	37	2176	10.9063

The Point-Wise method provides the lowest total SSE but total of number of parameters was so high that the pseudo-MSE became the highest one among the methods. The basis representation method and the Piece-Wise procedure worked out almost in the same way. The three methods considered ε_r, W, L_1, and L_2 as the most relevant parameters for the design of a rectangular slot ring antenna considering as a measure of interest the reflection coefficient. The next chapter presents the conclusions and future work for the whole research.

6 Computer Applications

6.1 Introduction

One of the objectives of this research is the development of a series of tools to ease the implementation of the concepts developed and presented on the previous chapters in the industry. The following sections describe the applications developed for each of the methods developed in this study.

6.2 Point-Wise Method GUI

In order to ease the use of these tools, graphical user interfaces were created. For each method, a Matlab GUI was created. Figure 6-1 shows the GUI for the Point-Wise method. This GUI provides to the user with all the plots in order to complete the desired analysis. In addition, the “pseudo MSE” is included in the final output of this form. The next page shows the MS Front Page document created to be used as the help for this application. The user only needs to press the “help” button of the application and the html document appears as part of the Matlab help.

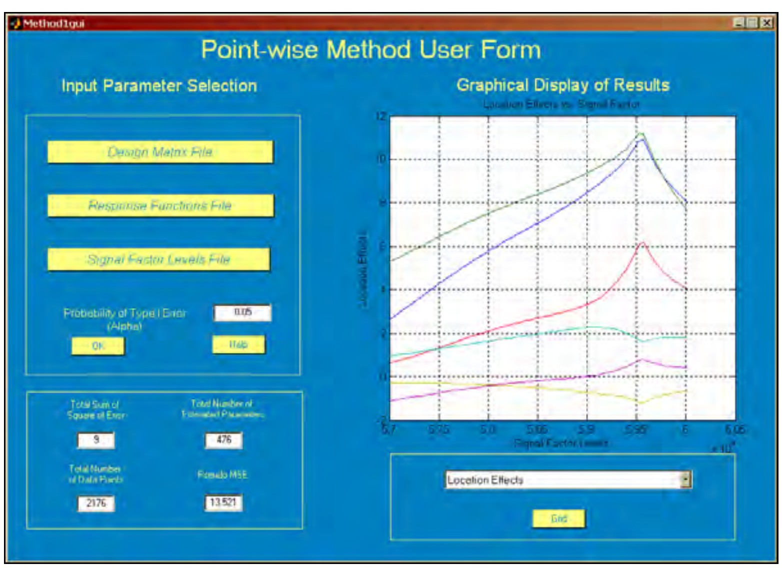


Figure 6-1 Snapshot of the Point-Wise Method user form

6.2.1 Point-Wise Method GUI Help

In order to use the Piece-Wise procedure you need to complete the following series of steps

- Press the "Design Matrix File" button- Select the **text** file with the design matrix
- Press the "Response Functions File" button- Select the **text** file with the responses
- Press the "Signal Factor Levels File" button- Select the **text** file with the responses
- In addition to the files that must be uploaded, a probability must be entered in the edit box provided.
- The "OK" button must be pressed in order to execute the procedure

Note: Each row of the response functions file must correspond to the row of the design matrix, otherwise errors can occur in the analysis.

The following figure shows a snapshot of the GUI for the Piece-Wise procedure

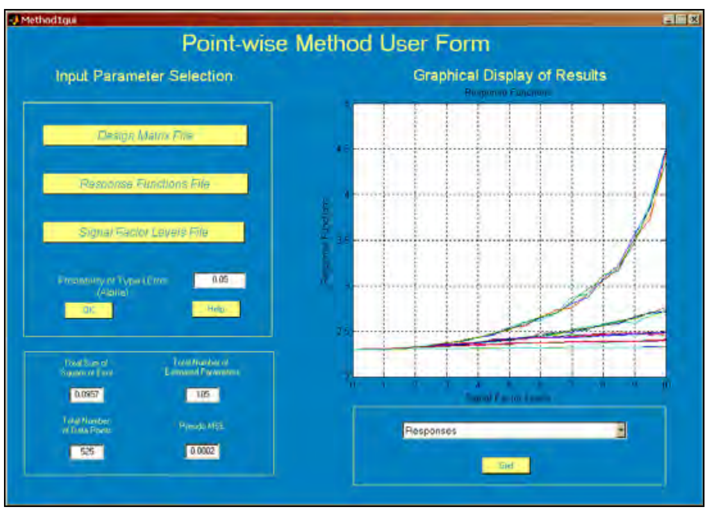


Figure 6-2 Snapshot of the user form for the Point-Wise method presented in the help file of the application

The popup menu has the options required. The user has to select the graphs in order to see them in the axes. In addition, the grid can be added to the graphs if is desired.

6.3 Basis Representation Method GUI

For the Basis Representation Method also a graphical user interface was created. This is quite different with respect the previous one. This GUI has three main sections. The first section, which is located at the left side of the GUI, is the input parameter section. In this section, the user loads the files with the data. After loading the files, the user has to select the type of basis that is going to be used and the number of basis that are desired. If the basis is a Cubic Spline the user has to enter the desired knots for the spline. The second section of this application is the results section, which is right in the center of the application. All the results for the method are presented in this section, including the results after the stepwise procedure. The last section of the GUI includes the controls to initialize the stepwise procedure, the outputs for the “pseudo MSE” and the help button. In the next page the html document for the help it is presented

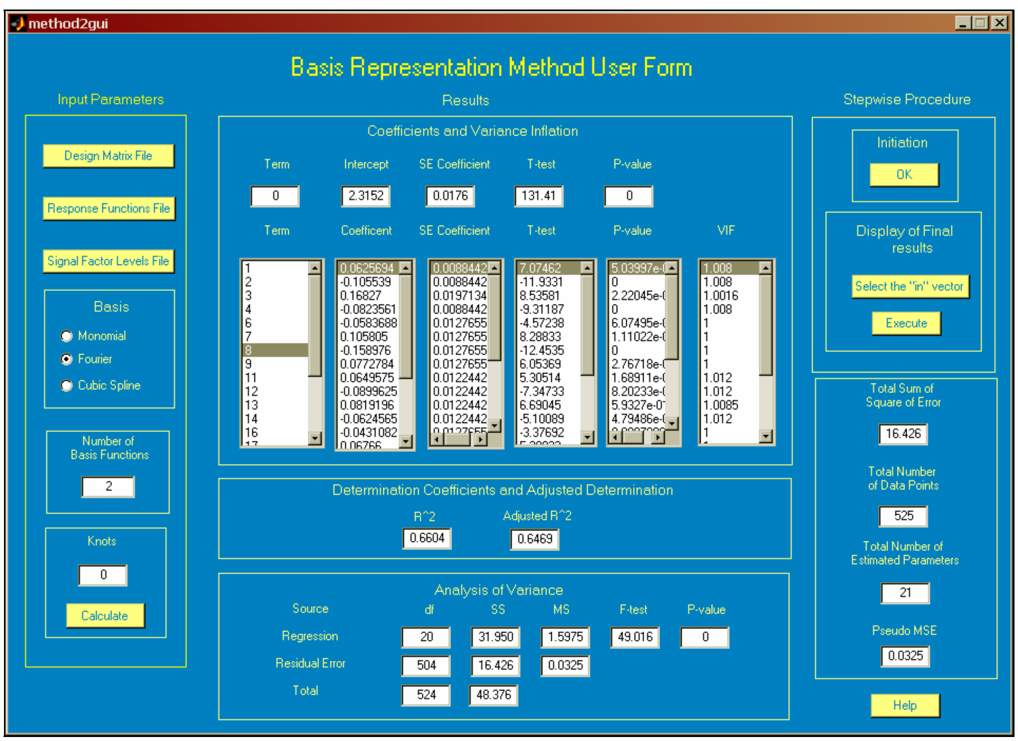


Figure 6-3 Snapshot of the Basis Representation Method user form

6.3.1 Basis Representation Model GUI Help

The design matrix must include the column for the intercept estimation. This column of ones must be the first one in the design matrix. The number of rows of the design matrix and the responses matrix must be the same. In addition, the number of columns of the responses matrix and the length of the signal factor vector must be equal. Otherwise, error messages will appear indicating the errors.

In order to use the GUI for the Basis Representation Model it is required to follow the next sequence of steps.

- Load the file for the design matrix (x)
- Load the file for the responses matrix (y)
- Load the file for the signal factor levels (t)
- Select the type of basis to be used

  (i) If the basis is a monomial, or Fourier basis fill the "knots" edit text with a zero.

  (ii) If the basis is a Cubic Spline fill the "Number of Basis Functions" edit text with a zero
- The knots have to be enclosed using brackets ([ ]) and each element must be separated using spaces or commas. The knots selected must be equal to some levels of the signal factor vector.
- Press the "Calculate" Button to execute.

Note: If you select decide to select other type of basis unselect the previous one.

Figure 6-4 the GUI for the basis Representation Method in detail.

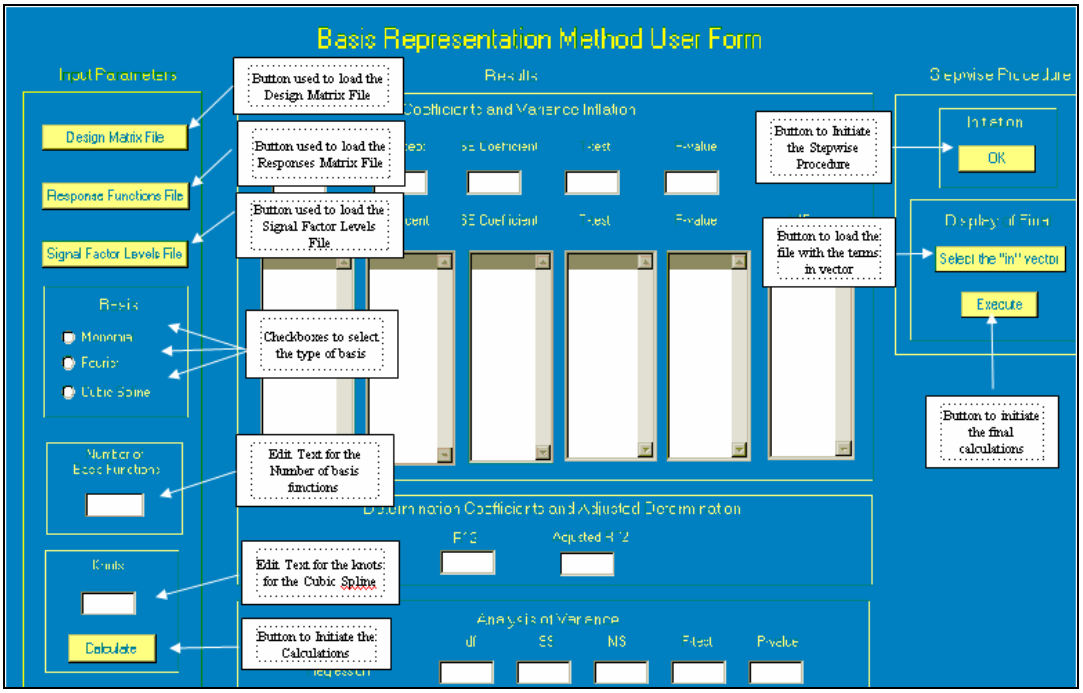


Figure 6-4 Snapshot of the Basis Representation user form used in the help file of the application

To use the stepwise procedure of this application it is necessary to complete the following steps

- Press the "Initiate" button of the application

This Button activates the Matlab stepwise procedure that opens three GUIs that are shown in Figure 6-5

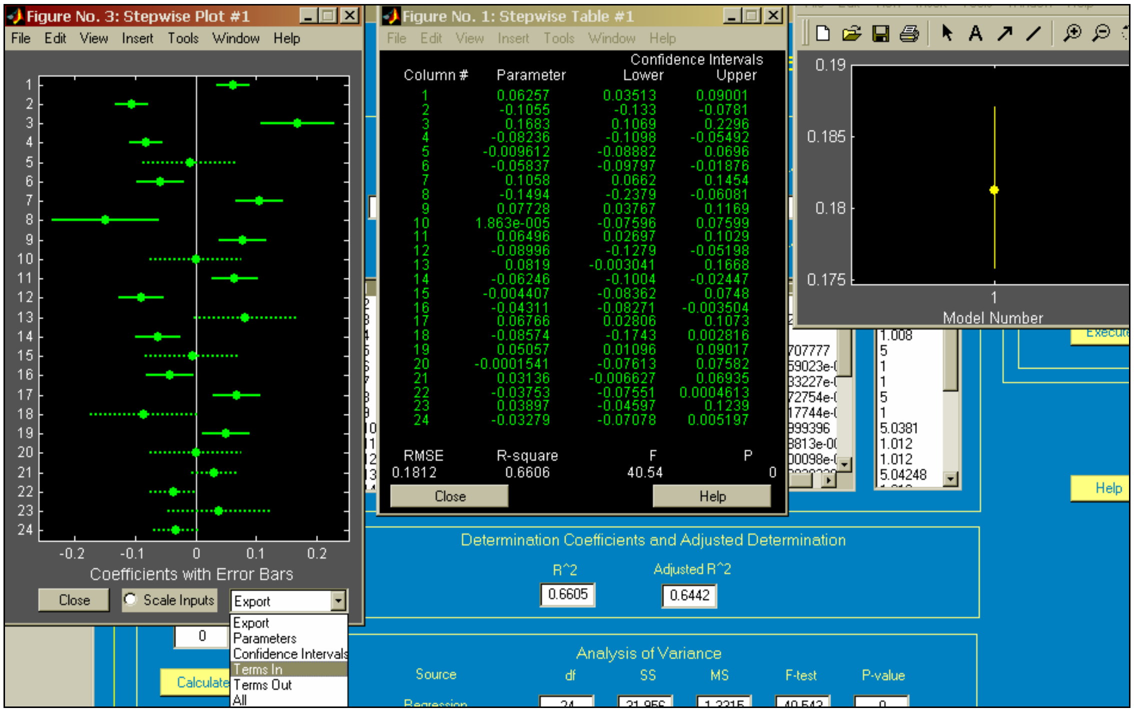


Figure 6-5 Snapshot of the stepwise procedure of the Basis Representation user form, presented in the help file

After eliminating al the terms that are not going to be included in the final model; use the export option of the stepwise plot and select the option "Terms in". This option is going to save that vector of indices in the workspace. Select the vector and create a text file with those indices.

- Press the "Select the "in" vector" button of the application
- Select your file with the Terms in indices
- Press the "Execute" button to present the final results in the GUI

6.4 Piece-wise Method GUI

For this method a GUI was develop in order to ease their implementation. This GUI is similar to the previous ones, but all the results of all the procedures conducted are presented on the Matlab's workspace instead on the GUI. The Piece Wise method is a more complicated procedure compared to the previous two. It can be summarized in three main steps, which are the following

- Knots Search
- Unconstrained Piece Wise Method and variable selection via stepwise procedure
- Constrained Least of Squares

The performance metrics for the model such as the Pseudo-F ratio will appear on the user form. The plot for the knot search it is part of the GUI but the regression plots will appear as separate figures. The next section presents the help for this application.

6.4.1 Piece Wise Method GUI Help

The design matrix must include the column for the intercept estimation. This column of ones must be the first one in the design matrix. The number of rows of the design matrix and the responses matrix must be the same. In addition, the number of columns of the responses matrix and the length of the signal factor vector must be equal. Otherwise, error messages will appear indicating the errors.

In order to use the GUI for the Basis Representation Model it is required to follow the next sequence of steps.

- Load the file for the design matrix (x)
- Load the file for the responses matrix (y)
- Load the file for the signal factor levels (t)
- Select the type of basis to be used
- Define the number of basis functions
- If the knot search it is going to be executed the “Number of Knots” field must be defined
    - Press the “Start Knot Search Button

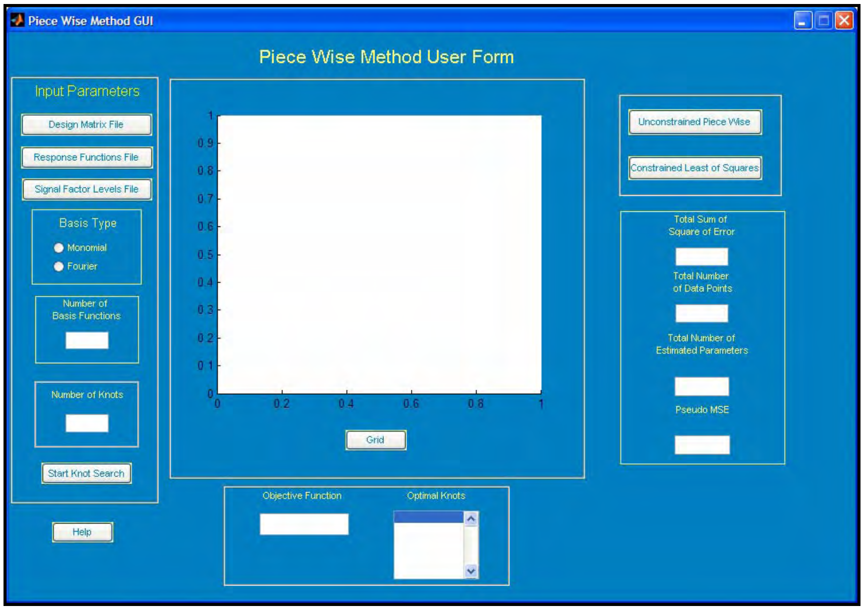


Figure 6-6 Piece Wise Method user form used in the help file of the application

- To execute the unconstrained piece wise method press the button "Unconstrained Piece Wise"
    - A question dialog it is going to appear asking you if you completed the knot search
        - If the answer to this question is "not" then the program, it is going to ask you to enter the knots using the Matlab's vector notation on the command window.
    - The stepwise command GUI it is going to appear for each region of the problem (the number of knots plus one), to move to the next stepwise you must press any key
    - You must export the "in terms" of each stepwise procedure to the workspace
- To execute the Constrained Least of Squares press the button "Constrained Least of Squares"
    - A question dialog it is going to appear asking you if you completed the knot search

        - If the answer to this question is “not” then the program, it is going to ask you to enter the knots using the Matlab’s vector notation on the command window.
    - The program will require you to enter the in terms of the stepwise procedure for each region, using the Matlab’s workspace
    - All the required graphs are going to appear as separate figures
- The GUI should look like the next figure if you completed all the steps mentioned

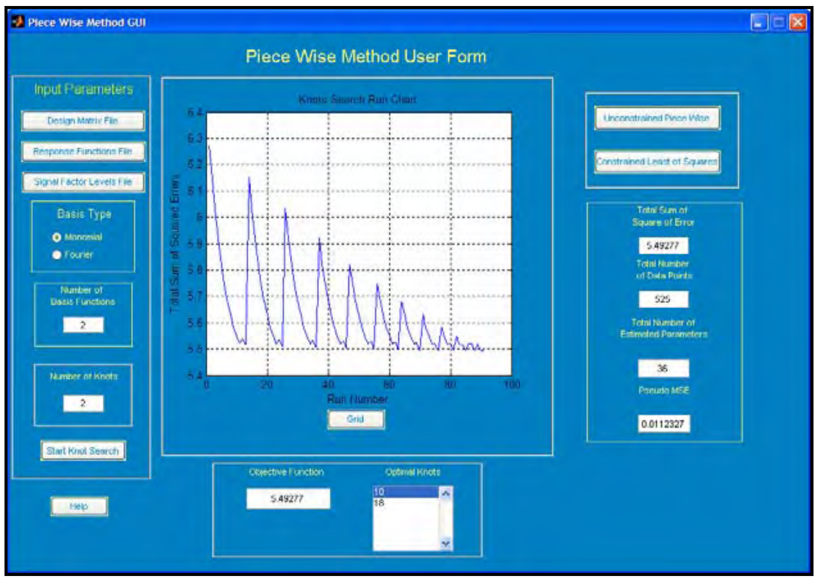


Figure 6-7 Piece Wise Method GUI after Execution

This chapter presented the main aspects of the applications developed in order to ease the implementation of the concepts of this research in industry. The technical details of the applications are on the appendixes. The next chapter presents the conclusions and future work for this research.

7 Conclusions and Future Work

7.1 Conclusions

The main contribution of this research is the demonstration of how powerful can be the functional data analysis in order to analyze industrial experiments. The development of new methods of analysis such as the basis representation method and the Piece-Wise methods are also considerable contributions. The three methods were consistent, with respect the factors considered as relevant in both of the experiments. The point wise procedure can lead to good results but sometimes the interpretation of the graphs can be complex. The basis representation method is more complex than the Point-Wise in terms of computational effort, but the results obtained can be easily interpreted because is a linear regression model. The Piece-Wise can be an option for experiments in which some regions need to be analyzed in detail. The greatest disadvantage that this method has with respect the previous two is the necessity of the knot search. This knot search can take a lot of computational effort in order to find the best locations for the knots. In addition to the knot search, the procedures to ensure the continuity of the curves among the regions can have a serious impact on the inferential capabilities of the models. A possible advantage of the Point-Wise method over the basis representation and the Piece-Wise method is the non-dependence in a basis selection. In addition, the Point-Wise method is very efficient in terms of computational effort. In general, the three methods complete their task. They are able to find the most relevant factors that can affect the functional response. The following table presents a summary of the respective advantages and disadvantages for each of the methods studied in this research.

Table 7-1 Summary of the Advantages and Disadvantages for the Three Methods Used in this Research

Method	Advantages	Disadvantages
Point-Wise	• Computationally Efficient • Independent of a basis selection • Useful for model selection and determination	• Depends on graphical interpretation • High number of terms to be calculated and analyzed • No functional form is obtained for the estimated response • A multiplicity problem that increases with the number of levels of the signal factors
Basis Representation	• Simple interpretation • Low number of terms	• Basis Dependent • Could be computationally intensive
Piece-Wise	• Simple interpretation • Deeper analysis in each region	• Basis Dependent • Computationally extensive due the knot search • The approach used for to establish the continuity constraint induces a severe multicollinearity problem

It is possible to observe that all the methods have advantages and disadvantages. The final decision with respect the method to be applied is going to depend in the nature of the problem. If the problem has a high number of levels for the signal, the Piece-Wise method can be a problem in terms of the knot search part of the procedure. It is recommended to use the Piece-Wise method only for situations in which is required to analyze response functions by regions or when the signal factor does not has a high number of levels. The other two procedures can be applied practically in every desired situation.

7.2 Future Work

There is a lot of work to be done; in specific for the basis representation, and the Piece-Wise methods. More types of bases can be integrated to those methods. Bases such as wavelet, exponential, polygonal and constant could be used for those methods. New procedures in order to complete the knots search are extremely necessary in order to make more efficient this search in order to ease the implementation of the Piece-Wise method. In addition, procedures to obtain the curves continuity between regions need to be developed in order to ensure the continuity without affecting the quality of the regressions. In the cases studied some problems were founded with respect the regressions. To be specific the assumption of constant variance was not followed by most of the models. Different approaches for the parameter estimation could be tested in order to overpass the difficulties presented by the use of the least of squares regression.

In addition to all the work completed in this research, there is a lot of work been doing by many researchers. Statisticians and engineers must develop some robust and practical techniques in order to demonstrate the strengths that the functional data analysis can provide to industrial experiments.

8 References

1. Montgomery, D.C. (2001), *Design and Analysis of Experiments* (5^{th} edition), New York: Wiley

2. Ramsay, J.O., Silverman, B.W. (1997), *Functional Data Analysis,* New York: Springer-Verlag

3. Nair, V.N., Taam, W., Ye, K.G. (2002), "*Analysis of Functional Response Form Robust Design Studies*", Journal of Quality Technology, Vol. 34, No. 4

4. Faraway, J.J., "*Regression Analysis for a Functional Response*", Technometrics, Vol. 39, No. 3

5. Montgomery, D. C. (1994), *Applied Statistics and Probability for Engineers*, New York: Wiley

6. Fan, J., Lin, S. (1998), *"Test of Significance When Data Are Curves"*, Journal of the American Statistical Association, Vol. 93, No. 443, Theory and Methods

7. Cuevas, A., Febrero, M., Fraiman R., (2002), *"Linear Functional Regression: the case of fixed design and functional response"*, The Canadian Journal of Statistics, Vol. 30, No. 2

8. Ramsay, J.O., Silverman, B.W., (2002), *Applied Functional Data Analysis Methods and Case Studies,* New York: Springer-Verlag

9. Abramovich, F., Antoniadis, A., Sapatinas, T., Vidakovic, B. (2002), "*Optimal Testing in Functional Analysis of Variance Models*" www.isye.gatech.edu/~brani/Wavelets/02-02.pdf

10. Montgomery, D.C., Peck Elizabeth A., Vining G. Geoffrey (2001), *Introduction to Linear Regression Analysis* (3^{rd} edition), New York: Wiley

11. Ramsay, J.O., and Dalzzel, C.J. (1991), "*Some Tools for Functional Data Analysis*" (with discussion), Journal of the Royal Statistical Society, Ser. B, 539-572

12. Cuevas, A.; Febrero, M. y Fraiman, R. (2002) "*Linear functional regression: The case of fixed design and functional response***.**" The Canadian Journal of Statistics, Vol. 30, nº 2, pp. 285-300.

13. Salazar, J. (2002),"*Desarrollo de un sensor para el sistema Crosswell Radar Tomography para detección de residuos bajo tierra."* M.S Thesis. ,University of Puerto Rico, Mayagüez, P.R.

Appendix 1 Details form Theoretical Case

In this first appendix, the details for the theoretical case study are going to be presented. As mentioned in chapter 4 a VBA Macro was used to generate the response curves for the experiment. An snapshot of the excel spreadsheet used to generate the data used for the experiment is presented in figure A.1.1

To recalculate press CTRL+ALT+F9

x1	-1	-1	1	1	0
x2	-1	1	-1	1	0
t	Resp(-1,-1)	Resp(-1,1)	Resp(1,-1)	Resp(1,1)	Resp(0,0)
0.00	-0.0261	-0.0317	0.0091	0.0096	-0.0112
0.50	0.0083	-0.0187	0.0162	0.0410	-0.0012
1.00	0.0588	0.0259	0.0643	0.1031	0.0091
1.50	0.1501	0.1103	0.1601	0.1567	0.0201
2.00	0.1971	0.2334	0.2546	0.2106	0.0390
2.50	0.3150	0.3488	0.4008	0.3042	0.0511
3.00	0.4577	0.4776	0.6850	0.3574	0.0668
3.50	0.5032	0.7030	0.9834	0.4220	0.0809
4.00	0.8000	0.7273	1.3696	0.5324	0.0925
4.50	0.9602	1.0304	1.9202	0.5827	0.1156
5.00	1.1380	1.0008	2.5596	0.6782	0.1456
5.50	1.2901	1.0618	2.7642	0.7191	0.1226
6.00	1.4991	1.3723	4.7049	0.8467	0.1637
6.50	1.8385	1.6184	4.8250	0.7877	0.1767
7.00	2.2173	1.2114	7.5715	0.8247	0.1964
7.50	2.0659	1.6399	7.7693	0.8033	0.1857
8.00	2.7472	1.5161	12.0100	0.8910	0.2281
8.50	3.1385	2.0180	16.5359	1.0229	0.2215
9.00	3.3522	1.7976	21.6583	0.9768	0.2457
9.50	4.8005	2.0264	34.1406	1.1356	0.2248
10.00	4.7947	2.0241	78.4544	1.0729	0.2330

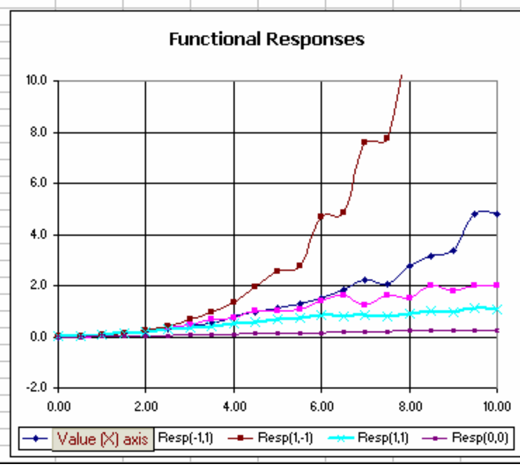


Figure A.1.1 Snapshot of the MS Excel spreadsheet used to generate the experiment for the theoretical case study.

The details of the VBA macro are shown in the next figure.

(General) | Response

```
Private Function Response(t, x1, x2) As Single

    t = Abs(t)
    If t > 10 Then t = 10

    f0 = x1 - 0.5
    f1 = 2 * x1 * x2 + 0.75
    f2 = 2 * x1 ^ 2 + x2 ^ 2 - x1 * x2 + 0.25
    ft = f0 + f1 * t + f2 * t ^ 2

    g0 = 2 * x1 ^ 2 + x2 ^ 2 + 50
    g1 = 2 * x1 * x2 + 2.25
    g2 = x2 + 0.5
    gt = g0 + g1 * t + g2 * t ^ 2

    Eps = (Sqr(-2 * Log(Rnd(1)))) * Sin(2 * 3.1415926 * Rnd(2))

    Response = ft / gt * (0.075 * Eps + 1)

End Function
```

Figure A.1.2 Snapshot of the VBA macro used to generate the functions for the Theoretical Case Study

The numerical results of the five runs used to test the methods of analysis are shown in the next tables.

Table A.1.1 Data for Run #1 of the Theoretical Case Study

x1	-1	-1	1	1	0
x2	-1	1	-1	1	0
	-0.02955	-0.02701	0.009643	0.009401	-0.00974
	0.00794	-0.0224	0.019626	0.043568	-0.00134
	0.06313	0.025469	0.06363	0.090566	0.009705
	0.127897	0.100342	0.161652	0.156327	0.022091
	0.200127	0.224025	0.305872	0.194455	0.037763
	0.326992	0.311946	0.463723	0.275638	0.049903
	0.420874	0.53448	0.644597	0.346922	0.060038
	0.584393	0.649618	1.028072	0.44136	0.083804
	0.716669	0.792269	1.342418	0.488681	0.097133
	0.905562	0.973432	1.876639	0.620892	0.10853
	1.011142	1.104563	2.604896	0.637629	0.125836
	1.414877	1.307627	2.891856	0.660031	0.137948
	1.584432	1.408364	4.208405	0.762732	0.147079
	1.987885	1.452143	5.395805	0.805077	0.165387
	1.972278	1.817569	5.612346	0.789524	0.185538
	2.574979	1.568784	8.346522	0.824069	0.200709
	3.219038	1.631887	11.24417	0.936698	0.226127
	3.20401	1.78047	14.04148	0.822115	0.239182
	4.195609	1.822697	23.79627	0.99046	0.238341
	4.922475	1.931292	31.7241	1.120306	0.258074
	5.06598	2.007657	76.81027	1.042286	0.301072

Table A.1.2 Data for Run #2 of the Theoretical Case Study

x1	-1	-1	1	1	0
x2	-1	1	-1	1	0
	-0.02506	-0.02814	0.009319	0.009185	-0.01058
	0.008322	-0.01996	0.016525	0.042236	-0.00128
	0.059312	0.031915	0.066759	0.101967	0.010312
	0.129266	0.110564	0.162704	0.160596	0.022341
	0.201625	0.196582	0.297258	0.245443	0.035951
	0.337096	0.410005	0.505396	0.262178	0.048042
	0.482517	0.498941	0.758635	0.389001	0.070933
	0.591822	0.534093	1.003673	0.375722	0.087029
	0.643676	0.806376	1.354527	0.469875	0.104813
	1.00613	1.125653	1.717689	0.516822	0.095494
	1.170984	1.049902	2.573347	0.574457	0.131791
	1.342684	1.340854	3.061642	0.715934	0.153145
	1.618439	1.325836	4.053267	0.698535	0.155132
	1.730971	1.521352	4.867299	0.742636	0.180463
	2.270578	1.611667	6.525233	0.774761	0.201395
	2.707817	1.724198	7.626052	0.961136	0.20375
	2.880525	1.861544	11.79702	0.771159	0.224973
	3.502698	1.802845	15.2124	0.936421	0.234723
	4.021999	1.798873	25.33165	0.9183	0.260367
	4.248888	1.776384	36.97179	1.017769	0.244292
	5.298789	2.191868	80.79288	0.982667	0.204118

Table A.1.3 Data for Run #3 of the Theoretical Case Study

x1	-1	-1	1	1	0
x2	-1	1	-1	1	0
	-0.02906	-0.0308	0.009129	0.008319	-0.01112
	0.007969	-0.02145	0.015479	0.045803	-0.00132
	0.069092	0.028228	0.060468	0.095082	0.009102
	0.126612	0.111507	0.166558	0.145239	0.02142
	0.237457	0.197497	0.296702	0.239752	0.033853
	0.318097	0.354262	0.484636	0.270072	0.044545
	0.393697	0.534973	0.726386	0.331751	0.057815
	0.549785	0.640372	1.121206	0.454334	0.087166
	0.710146	0.726084	1.453431	0.460759	0.101167
	0.808599	0.957632	1.990109	0.537481	0.130697
	1.163068	1.204151	2.432114	0.600821	0.123196
	1.359438	1.159892	3.061314	0.71798	0.131949
	1.61253	1.480099	4.293924	0.700131	0.163548
	1.919435	1.377092	5.037033	0.840747	0.191638
	2.395726	1.520627	6.030873	0.827303	0.20146
	2.943201	1.732558	8.75532	0.91076	0.213701
	3.085407	1.704156	12.55411	0.992311	0.200034
	3.58548	1.698999	14.7318	0.968709	0.220278
	3.319357	2.134924	24.04517	1.009499	0.255059
	4.549213	1.680975	33.67556	0.938481	0.271797
	5.10281	1.868435	68.98808	1.00252	0.290429

Table A.1.4 Data for Run #4 of the Theoretical Case Study

x1	-1	-1	1	1	0
x2	-1	1	-1	1	0
	-0.02799	-0.03028	0.009709	0.007986	-0.00927
	0.008035	-0.01951	0.017758	0.045331	-0.00115
	0.065106	0.025831	0.073214	0.090301	0.009807
	0.125944	0.114701	0.14619	0.13574	0.022432
	0.248526	0.228625	0.280081	0.227296	0.037364
	0.303013	0.315163	0.44326	0.285834	0.04778
	0.463151	0.499487	0.706184	0.392888	0.069485
	0.535819	0.74476	0.94271	0.430409	0.081568
	0.725761	0.725605	1.387714	0.511043	0.087297
	0.894978	1.100409	1.932457	0.588839	0.109323
	1.183528	1.064585	2.804713	0.697142	0.118582
	1.294208	1.413398	3.53286	0.714377	0.14214
	1.582925	1.391951	3.64095	0.772333	0.15273
	1.798324	1.431159	4.88521	0.777191	0.210812
	2.390928	1.477567	7.834334	0.859378	0.19127
	2.40153	1.627087	8.339213	0.828463	0.206762
	3.173717	1.803808	12.6681	0.876307	0.247802
	3.622976	2.131194	13.56933	0.954295	0.205997
	4.017898	1.859688	24.40208	0.999199	0.248427
	5.028654	2.13463	39.10254	0.926936	0.258994
	5.236986	1.646408	82.36643	1.182261	0.240994

Table A.1.5 Data for Run #5 of the Theoretical Case Study

x1	-1	-1	1	1	0
x2	-1	1	-1	1	0
	-0.02562	-0.03041	0.009371	0.010595	-0.01093
	0.007637	-0.02082	0.016673	0.03829	-0.00113
	0.06271	0.029197	0.06654	0.0964	0.008378
	0.13141	0.105038	0.166515	0.143538	0.023743
	0.237821	0.215558	0.234498	0.218809	0.029061
	0.313876	0.359247	0.487976	0.303875	0.047052
	0.390985	0.499875	0.706666	0.364817	0.067519
	0.502899	0.586057	0.920314	0.38907	0.073615
	0.751922	0.733302	1.312079	0.519346	0.108103
	0.902728	1.033647	1.842729	0.566822	0.101534
	1.052538	1.229487	2.519927	0.601609	0.113821
	1.245388	1.35897	3.404696	0.723718	0.147072
	1.867583	1.409289	4.152516	0.72506	0.140835
	2.014749	1.475526	4.782989	0.823508	0.164042
	2.303423	1.353883	6.889404	0.883207	0.204301
	2.627156	1.462087	9.424388	0.922087	0.218063
	3.046413	1.651445	11.1958	0.964646	0.207535
	3.347896	1.686946	14.79526	0.90296	0.234474
	4.331627	1.846729	22.23893	0.866161	0.259719
	4.78677	1.991228	38.17537	1.049412	0.260177
	5.896906	2.030268	67.59177	1.095003	0.281483

In the next pages the graphical representation of the response functions at each experimental condition for all the runs are presented.

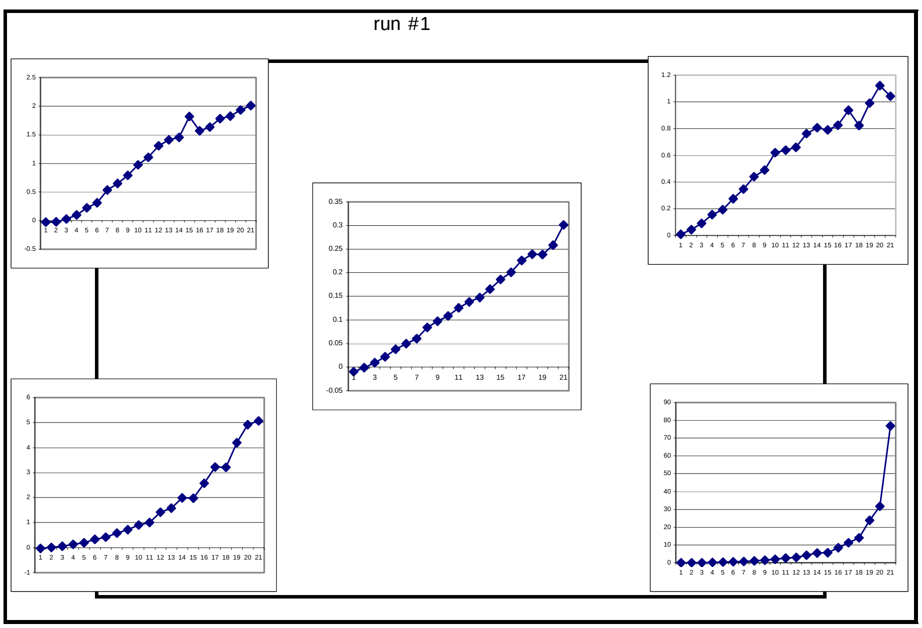


Figure A.1.3 Graphical Representation for run #1

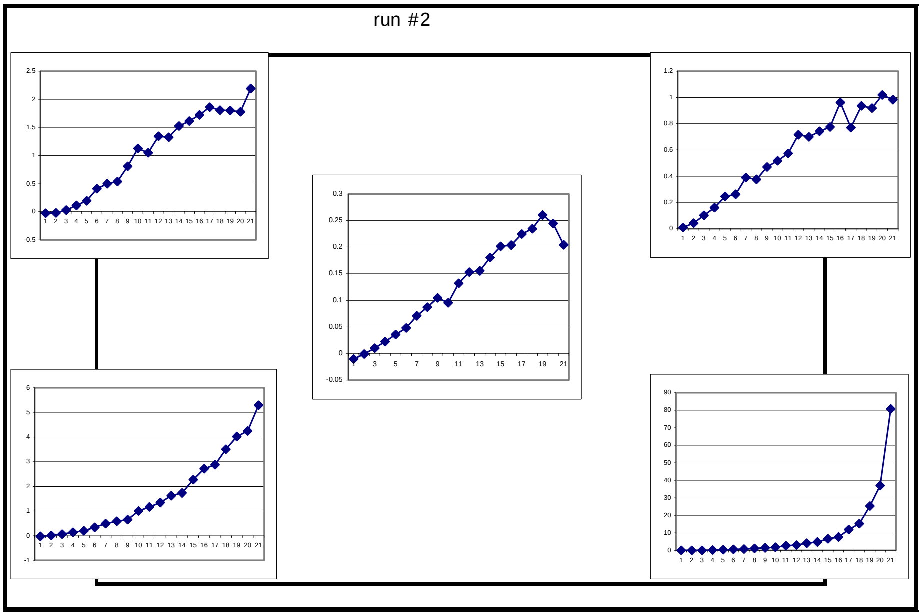


Figure A.1.4 Graphical Representation for run #2

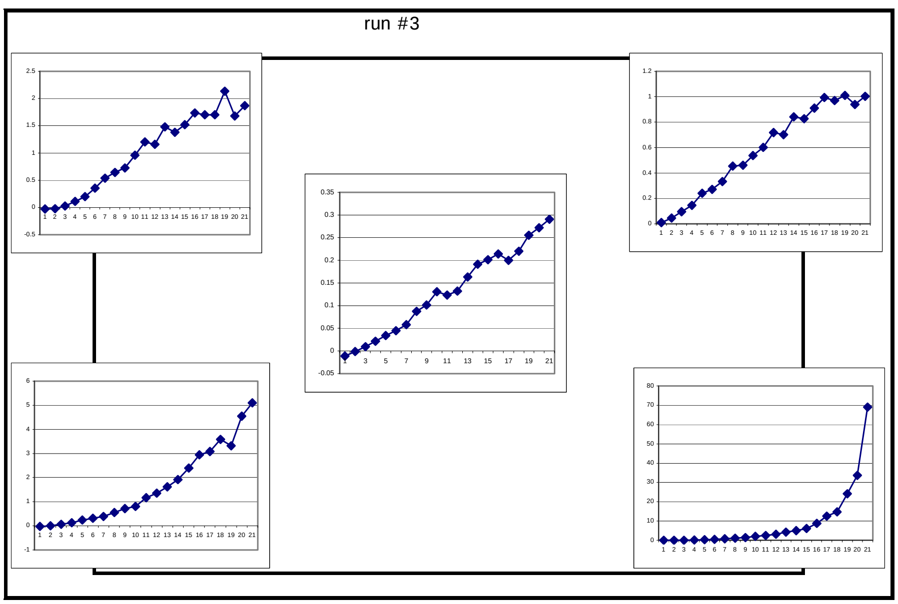


Figure A.1.4 Graphical Representation for run #3

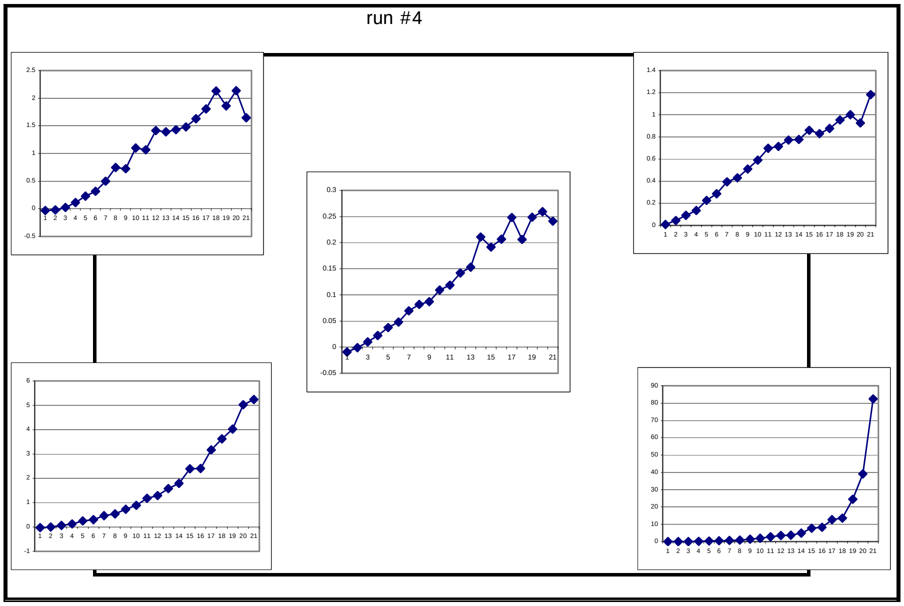


Figure A.1.5 Graphical Representation for run #4

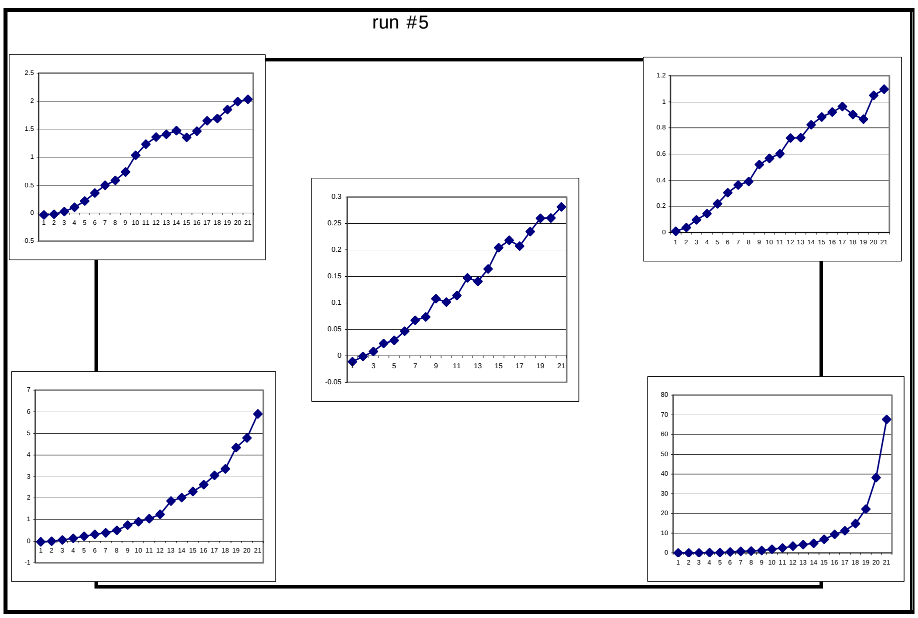


Figure A.1.6 Graphical Representation for run #5

It is important to mention once again that the origin of the function used to generate this theoretical experiment was completely unknown by the analyst of the experiment until the end of the investigation. Also is relevant to mention that Dr. Noel Artiles León, member of the thesis committee, created the function used.

Appendix 2 Results of the Classical Functional Data Analysis for Both Cases

Results of the FDA procedure to the Theoretical Case Study

The classical functional data analysis proposed by Ramsay was also used to analyze the cases studied. In order to complete this part of the analysis the Matlab functions that complement the book Functional Data Analysis were used. The graphical results for the Theoretical Case are shown in the next pages.

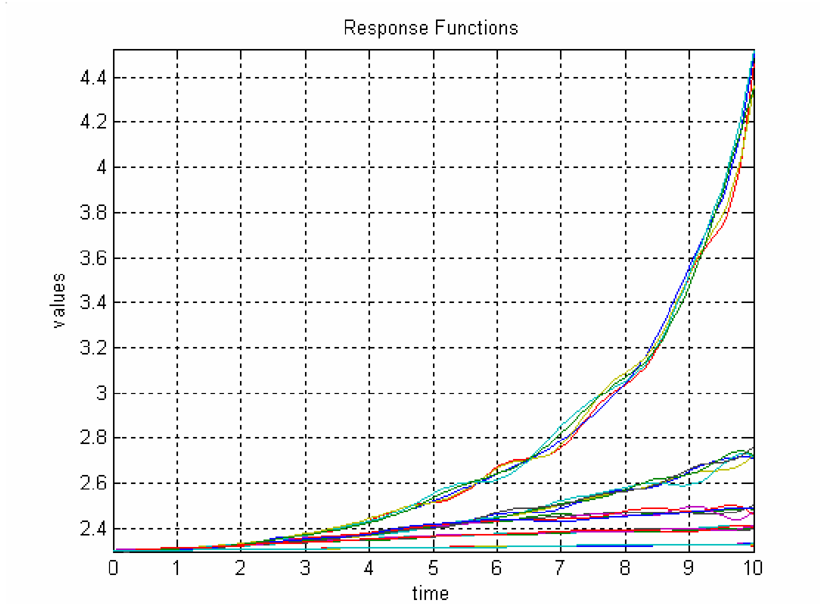


Figure A.2.1 Observed Functions

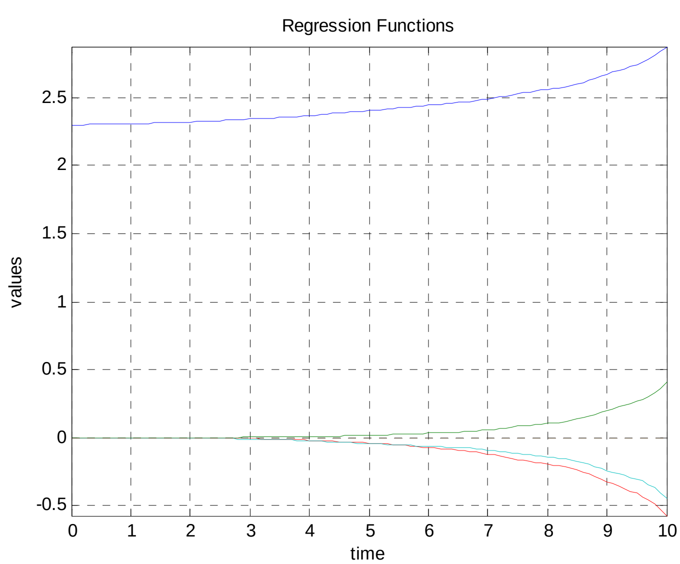


Figure A.2.2 Effects plot using the FDA Methodology

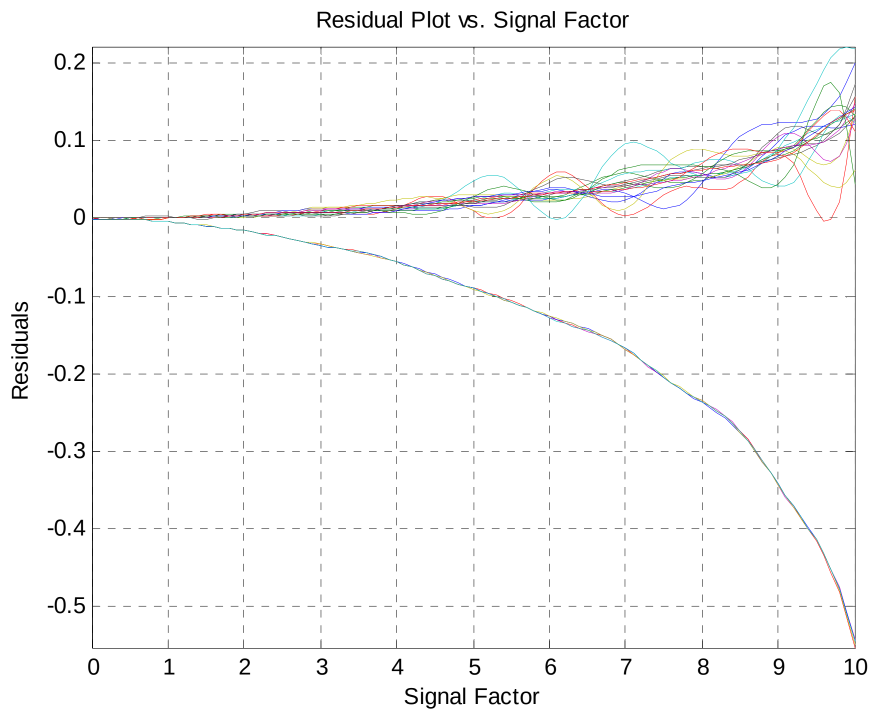


Figure A.2.3 Residuals Plot using the FDA Methodology

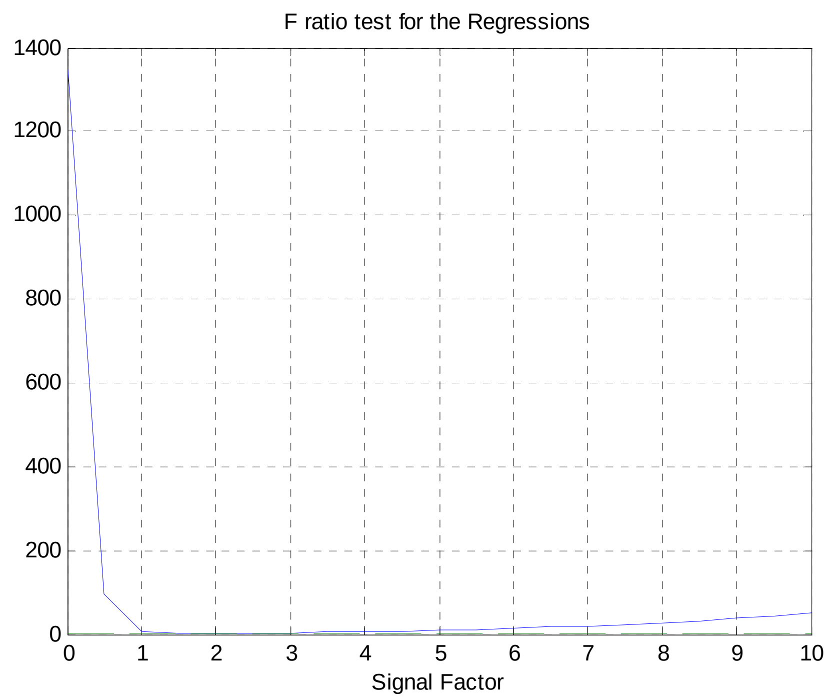


Figure A.2.4 F- Ratio Plot using the FDA Methodology

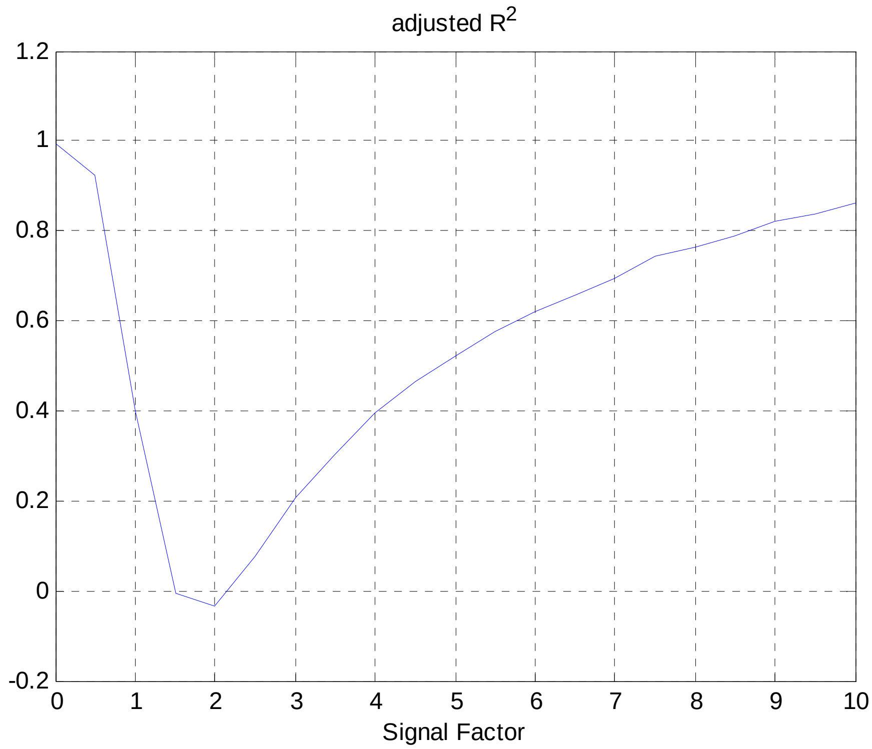


Figure A.2.5 Adjusted Determination Coefficient Plot using the FDA Methodology

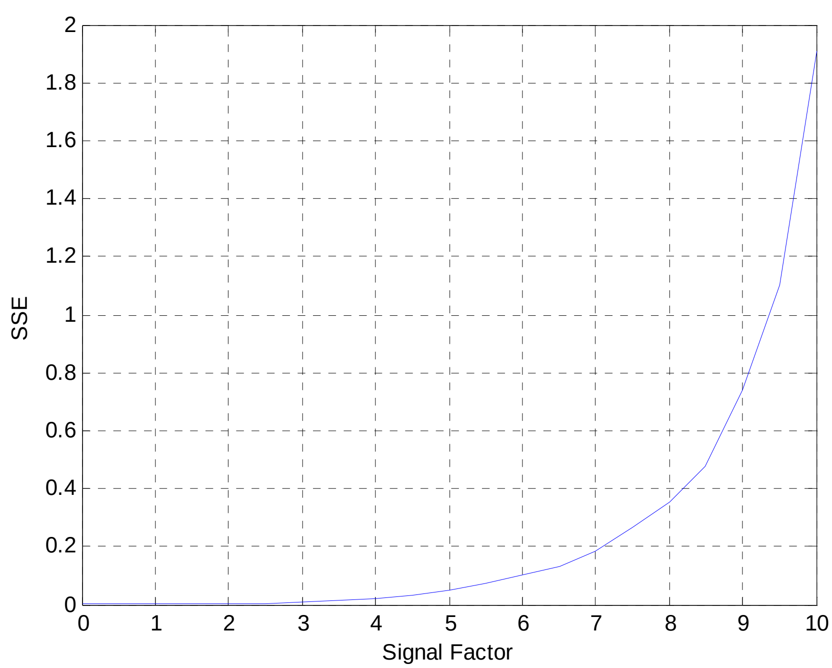


Figure A.2.6 SSE Plot using the FDA Methodology

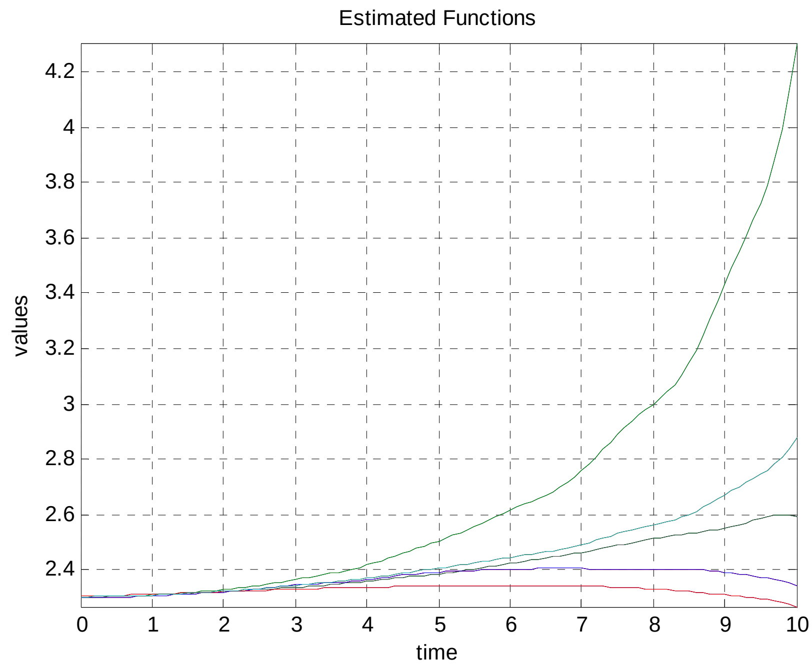


Figure A.2.7 Estimated Responses Plot using the FDA Methodology

It is possible to observe that the plots presented in this section are almost identical to the plots presented with use of the Point-Wise method. The main difference is that the plots look smoother than the previous ones. This is the effect of the use of techniques such as smoothers, registration and roughness of penalty approach among others. However, the basic interpretation and characterization of the response are basically the same.

Results of the FDA procedure to the Applied Case Study

The classical FDA was also applied to the applied study case in order to verify the performance of the Point-Wise procedure. The following graphs are the results obtained.

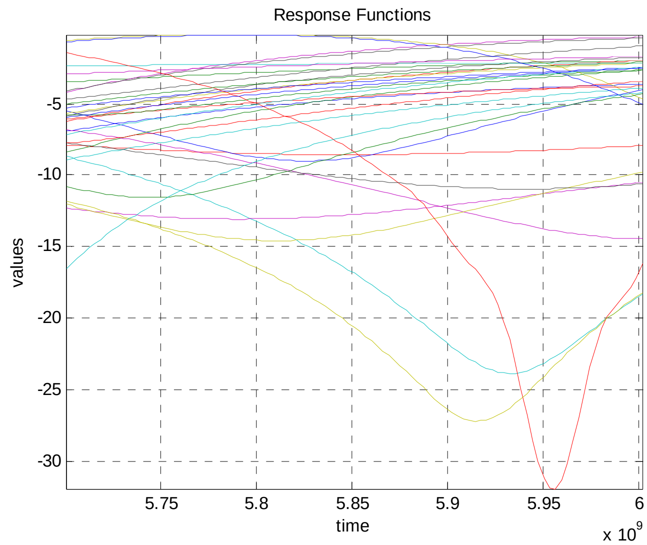


Figure A.2.8 Response Function for the Applied Case Study

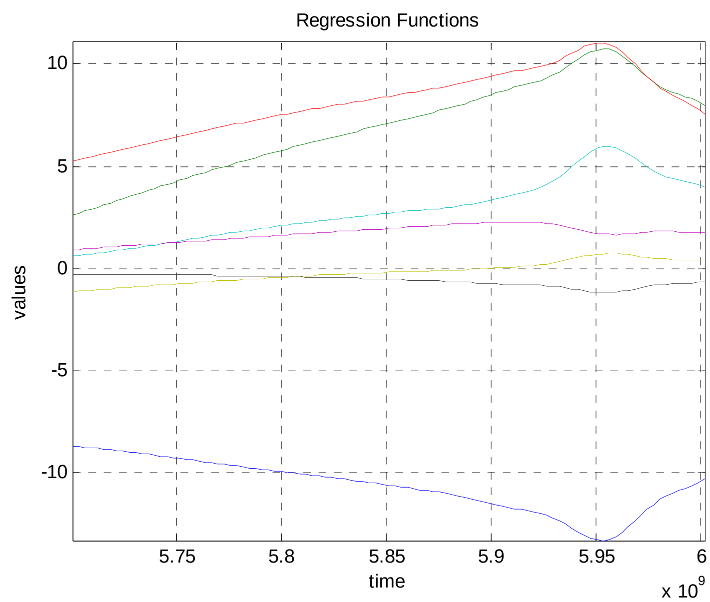


Figure A.2.9 Location Effects for the Applied Case Study using FDA

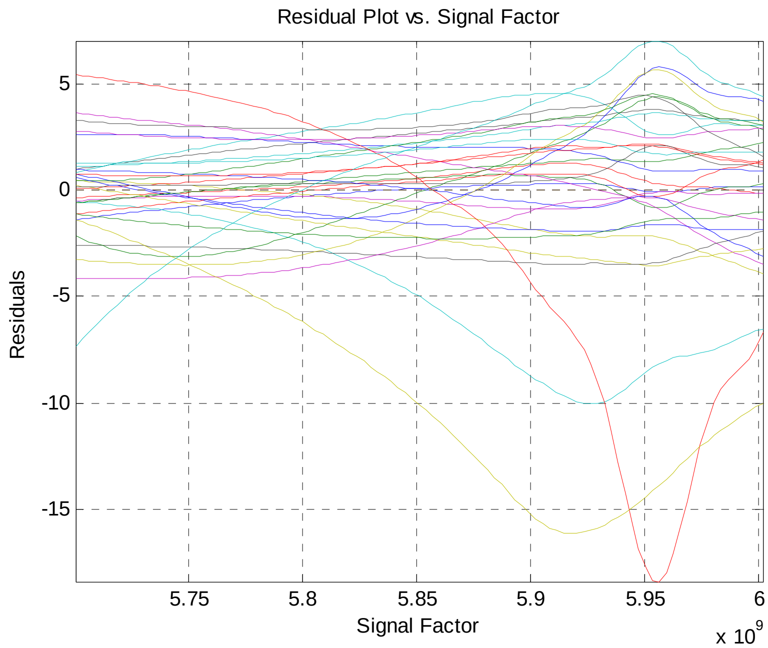


Figure A.2.10 Residuals for the Applied Case Study Using FDA

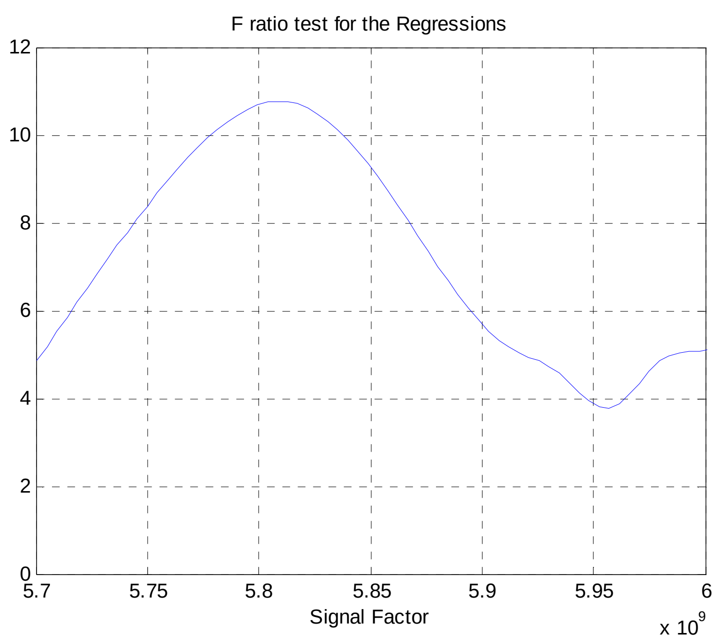


Figure A.2.11 F ratio test for the Regressions for the Applied Case Study

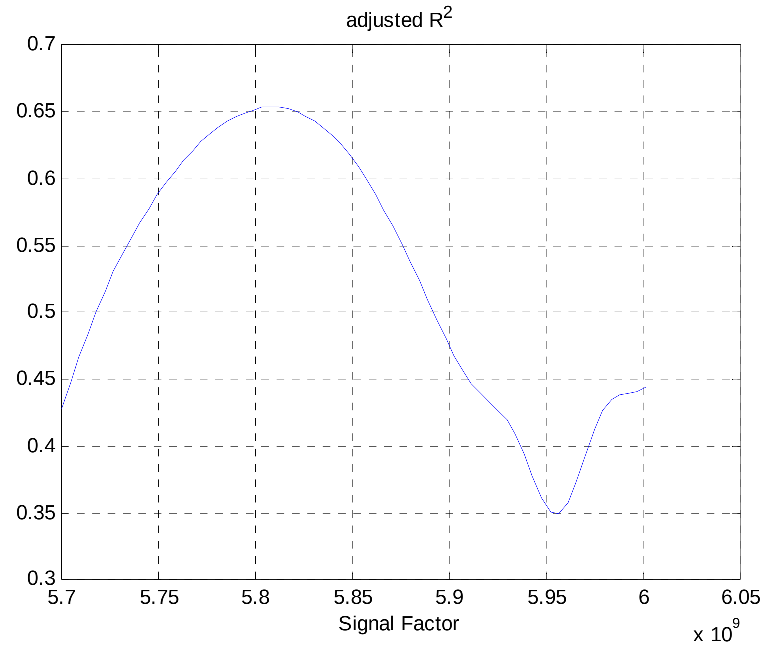


Figure A.2.12 Adjusted Determination Coefficient Plot for the Applied Case Study Using FDA

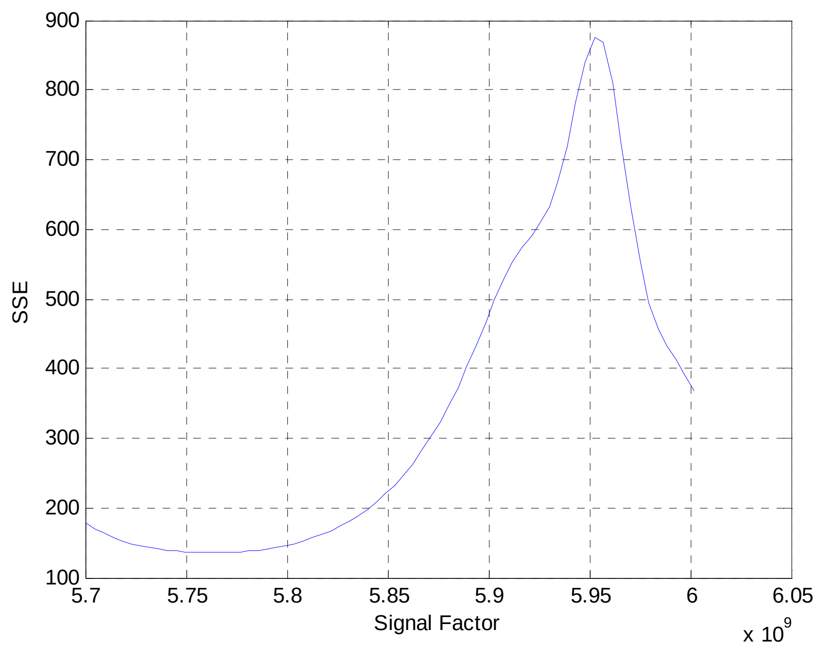


Figure A.2.13 SSE Plot for the Applied Case Study Using the FDA Procedure

Appendix 3 Response functions, Residuals and Estimated functions plots for the Basis Representation Model Using The Theoretical Case Study

Plots for the Monomial Basis Representation

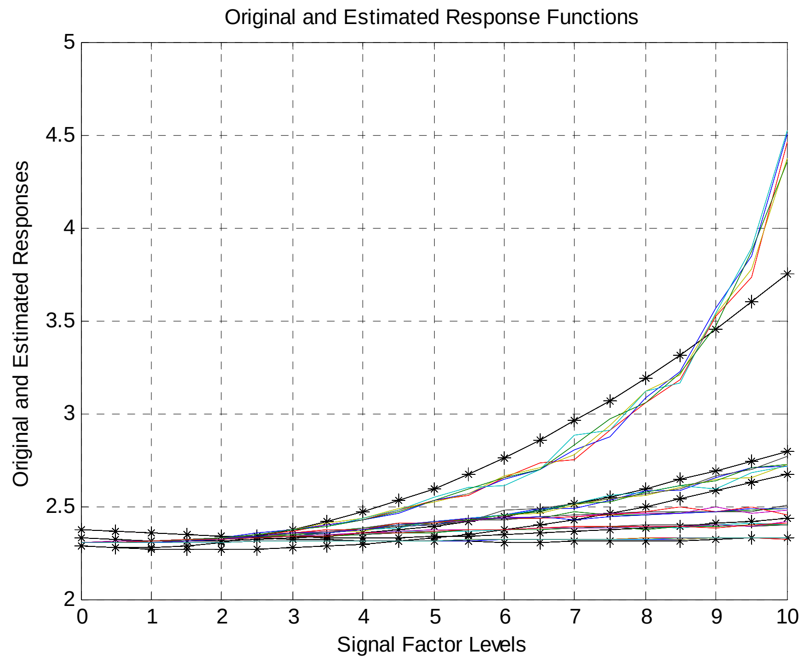


Figure A.3.1 Responses and estimated responses for the Basis Representation Final Model with a Monomial Basis for the Theoretical Case Study. The asterisks correspond to the estimated functions.

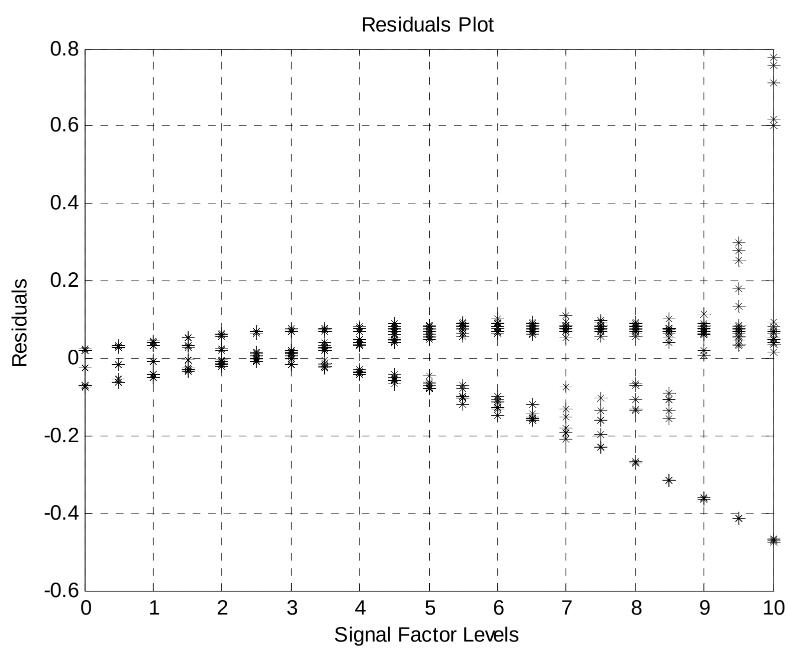


Figure A.3.2 Residuals for the Basis Representation Final Model with a Monomial Basis for the Theoretical Case Study

The following plots correspond to the Fourier Basis Representation model for the theoretical case study.

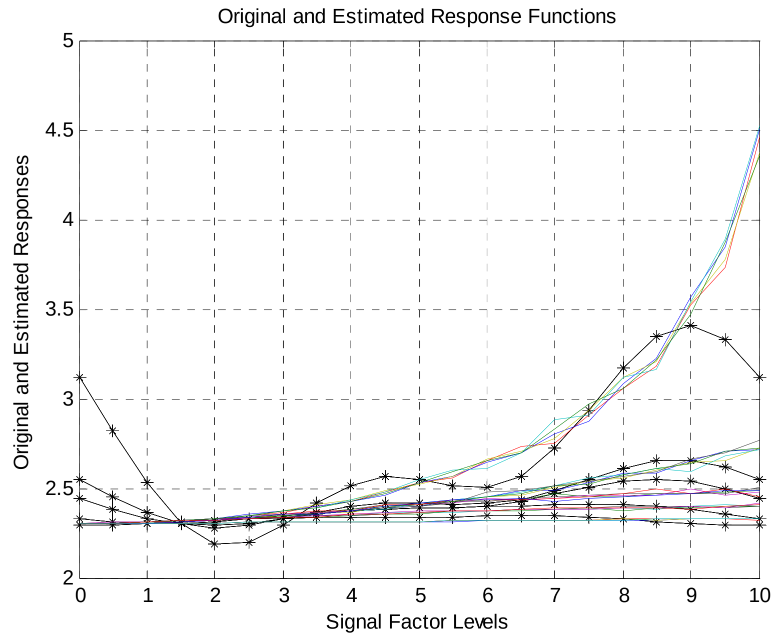


Figure A.3.3 Responses and estimated responses for the Basis Representation Final Model with a Fourier Basis for the Theoretical Case Study. The asterisks correspond to the estimated functions.

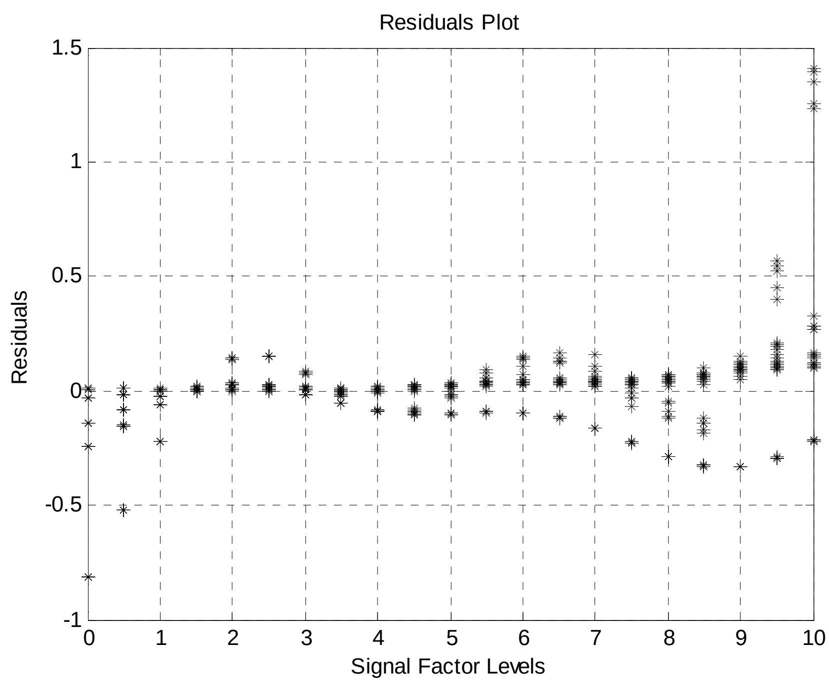


Figure A.3.5 Residuals for the Basis Representation Final Model with a Fourier Basis for the Theoretical Case Study

To conclude this appendix the graphs for the Cubic Spline Basis Representation Model are presented.

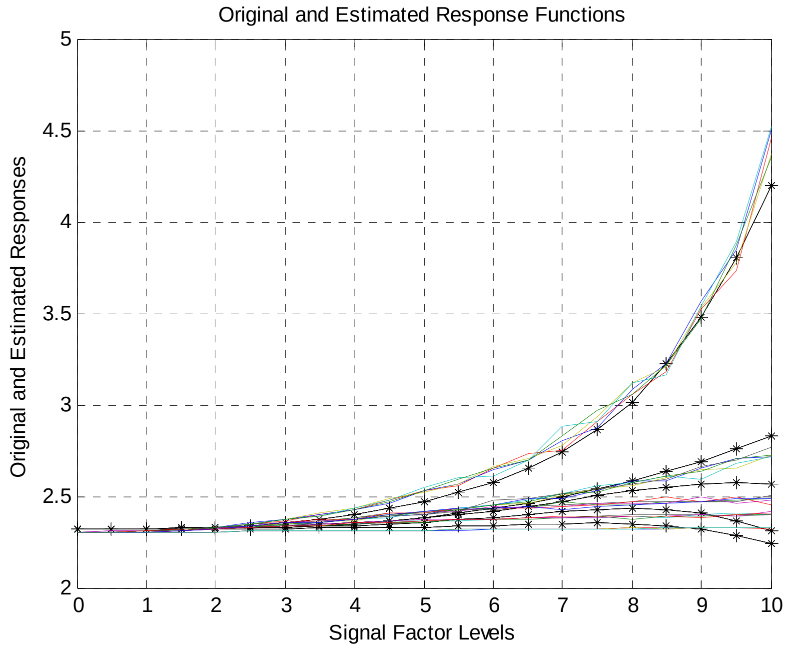


Figure A.3.6 Responses and estimated responses for the Basis Representation Final Model with a Cubic Sppline Basis for the Theoretical Case Study. The asterisks correspond to the estimated functions

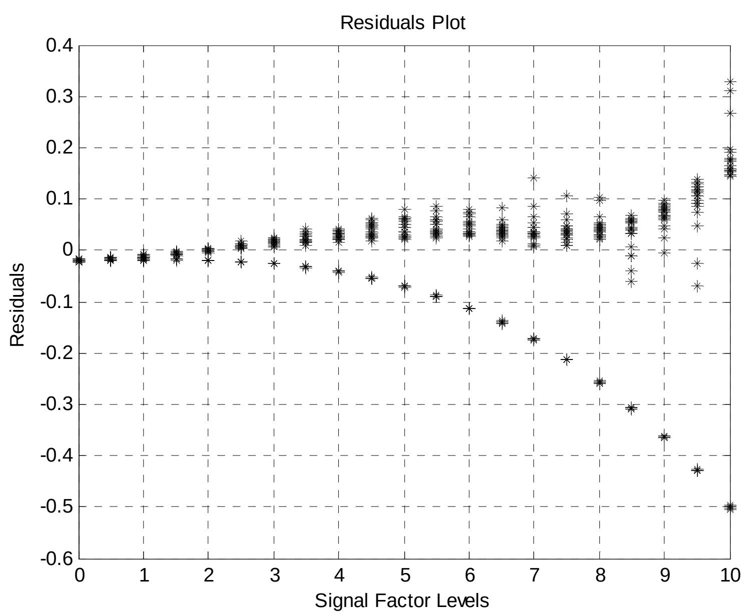


Figure A.3.7 Residuals for the Basis Representation Final Model with a Cubic Spline Basis for the Theoretical Case Study

Appendix 4 Response functions, Residuals and Estimated functions plots for the Basis Representation Model Using The Applied Case Study

The following plots correspond to the point wise analysis to the selected region in the applied case study. The selected region as mentioned in Chapter 5 is from 5.7 to 6 GHz.

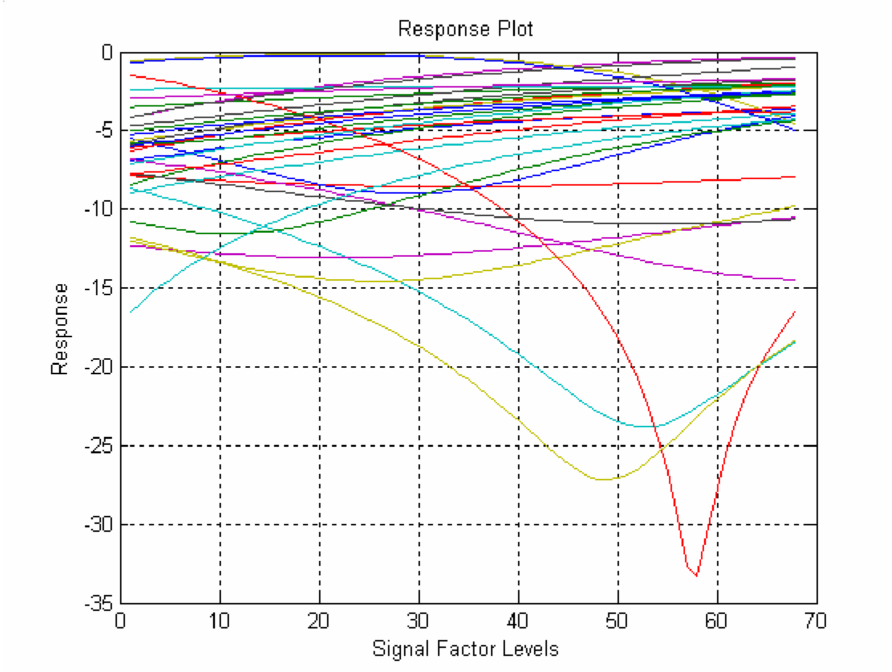


Figure A.4.1 Plot of all the response functions for the applied case study

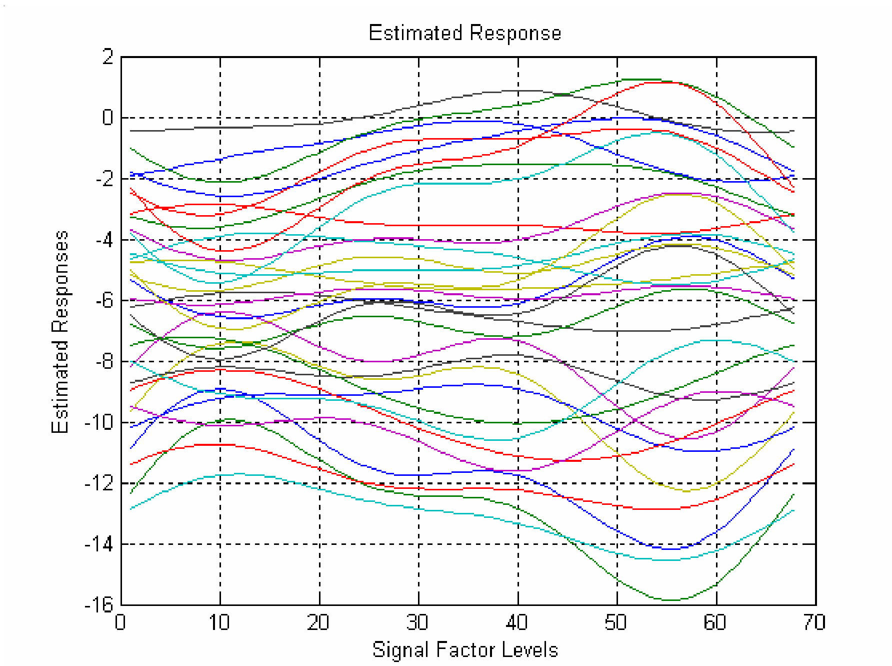


Figure A.4.2 Estimated Responses Plot for the Applied Case Study Using a Fourier Basis (K=2)

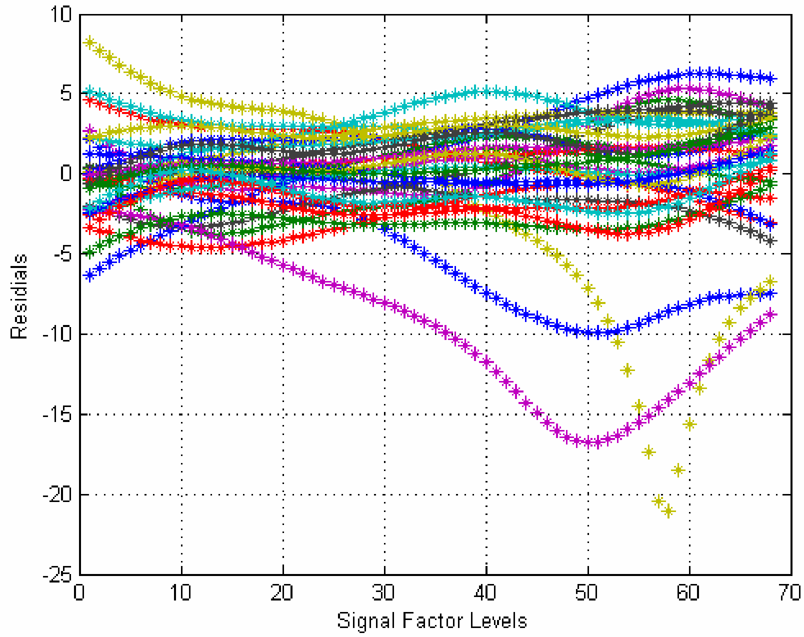


Figure A.4.3 Residuals Plot for the Applied Case Study Using a Fourier Basis (K=2)