People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research

University M'Hamed BOUGARA – Boumerdes

Institute of Electrical and Electronic Engineering

Department of Power and Control

Final Year Project Report Presented in Partial Fulfilment of the Requirements for the Degree of

MASTER

In **Electrical and Electronic Engineering**

Option: **Control**

Title:

Identification of A Distributed Parameter System using The Least Squares Technique

Presented by:

- **DAHLOUK Abdelkader**
- **AZZOUG El-Mehdi**

Supervisor:

Dr. KOUADRI Abdelmalek

Registration Number: 2015/2016

Dedication

It is our genuine gratefulness and warmest regard that we dedicate this work to the Mighty God, to our beloved parents, the reason of what we have become today and to our colleagues and friends without whom it was difficult for us to complete our thesis work.

Acknowledgements

We are highly appreciative of the effort of our supervisor Dr KOUADRI Abdalmalek for his invaluable assistance and guidance, for his contribution and taking time to read through this report, and for his encouragements and positive criticism of the project.

Sincere acknowledgement is given to all the teachers who taught us and were the essential key of this success.

Our deep appreciation to INELEC students and family for their efforts, patience and concern throughout this project.

Abstract

The identification of such systems represents one of the important directions of theoretical and practical research, due to large class of applications; fluid flow systems, heat diffusion systems, etc. In this context, we consider the linear parametric identification of distributed parameters systems using the Least Square. Two numerical examples of heat transfer systems are presented where linear and nonlinear models are obtained from heat difference equation. The third application is about the identification of a heat conduction in a cement rotary kiln using experimental data. We anticipate that this work be intuitive for practical applications in the areas of controls and signal processing.

Table of contents

List of tables

List of figures

Nomenclature

Symbol:	**Meaning:**
c	Specific heat
K	Heat diffusivity
l	length of tube (or rod)
T	Temperature
Greek Symbols:	
α	System parameter
σ	Standard deviation of the noise
ϵ	Prediction Error
Abbreviations:	
DPS	Distributed Parameter System
MSE	Mean Square Error
PDE	Partial Differential Equations
ODE	Ordinary Difference Equation
LS	Least Square
IC	Initial Condition
BC	Boundary Condition

General Introduction

Control systems are ubiquitous in the modern world where the instruments of our scientific and industrial society are applied to an increasingly wide range of processes. Such control intervention is undertaken with many different objectives in mind; e.g. steering the process to a desired state, minimizing the effects of various disturbances tending to move the system in undesirable directions, stabilizing systems which are inherently unstable or improving the stability properties of system with weak stability characteristics; etc. While it is rarely possible, in a mathematical model, to account for all the factors affecting the performances of a real world system mathematical modeling of the system is, nevertheless, ordinarily essential for efficient and effective design and implementation of control procedures.

The first concern with determining the dynamical structure for a practical system is the problem of system parameter identification. Up to the present time, the principal line of research activities is directed to treat lumped parameter systems described by ordinary differential equations. A lumped parameter representation of a physical system implies that its local spatial variations are disregarded. However dynamic behaviors of almost all real physical systems are, in fact, distributed. A distributed parameter system DPS is characterized by a class of partial differential equations, integral equations or integro-differential equations. In this thesis, we restrict our attention to physical systems described by partial differential equations, because the great majority of research including efforts of practical applications is concerned only with this kind of mathematical models. Although a class of partial differential equations may usually be derived by considering basic conservation principles, system parameters remain to be determined in many examples. From practical viewpoints, determination of heat transfer coefficients, specific heats, electro-magnetic properties, gas properties, chemical reaction rate constants, diffusion constants, etc. is a requisite effort throughout state estimates and/or optimal controls of distributed parameter systems.

The objective of this report is to identify heat conduction systems as an application for distributed parameter systems. The parameters of the system model are obtained using Least Square Method, a parametric approach based on the minimization of the sum of squared residuals; the difference between the observed value and fitted value by a model.

To achieve this objective, the report is divided into three chapters. The first chapter is about the techniques used in system identification. These techniques are divided according to the system model; parametric and non-parametric approaches for linear and nonlinear models

in time and frequency domains. The second chapter is about the distributed parameter system, its features and characteristics. The third is about application of the chosen approach (Least Square Method) in two numerical examples and to experimental data obtained from a cement rotary kiln.

Chapter One

System Identification Techniques

1. Introduction

This chapter gives a brief introduction to system identification. System identification is an immense topic, and the contents of this chapter are neither new nor complete. The purpose is merely to introduce concepts, ideas and algorithms that are useful in system modeling.

Modeling is the abstraction of a real process to characterize its behavior. It is concerned with the determination of particular models for systems that are intended for a certain purpose such as control. The idea is to build accurate, simplified models of complex systems from physical insights (input/output) measurements. Whether it is parametric identification, we determine the parameters of some chosen models, or non-parametric identification, it gives the general behavior of the system and sometimes it is essential for parametric identification **[1].** The best model is the simplest model that accurately describes the dynamics of system.

2. The Procedures of System Identification

The System Identification allows us to estimate a model of a system based on observed input-output data. Several ways to describe a system and to estimate such descriptions exist. This section gives a brief description of the most important approaches**.**

The procedure to determine a model of a dynamical system from observed input-output data involves three basic ingredients:

- Input-output data
- A set of candidate models (A model structure)
- A criterion to select a particular model in the set.

The identification process amounts to repeatedly selecting a model structure, computing the best model in the structure, and evaluating this structured model's properties to see if they are satisfactory. The cycle can be as follows:

1- We design an experiment and collect input-output data from the process to be identified.

2- We examine the data. Polish it to remove trends and outliers, select useful portions of the original data, and apply filtering to enhance important frequency ranges.

3- We select and define a model structure

4- We compute the best model in the model structure according to the input-output data and a given criterion of fit.

5- We examine the obtained model's properties

6- To validate the model, we take a close look at the model's output and compare it to the measured one on a data set that was not used for the fit. If it is good enough, we select the model. Otherwise, we go back and select another model set. Possibly, we also try other estimation methods or we work further on the input-output data. **[2].** Fig 1.1 illustrates different procedures of system identification.

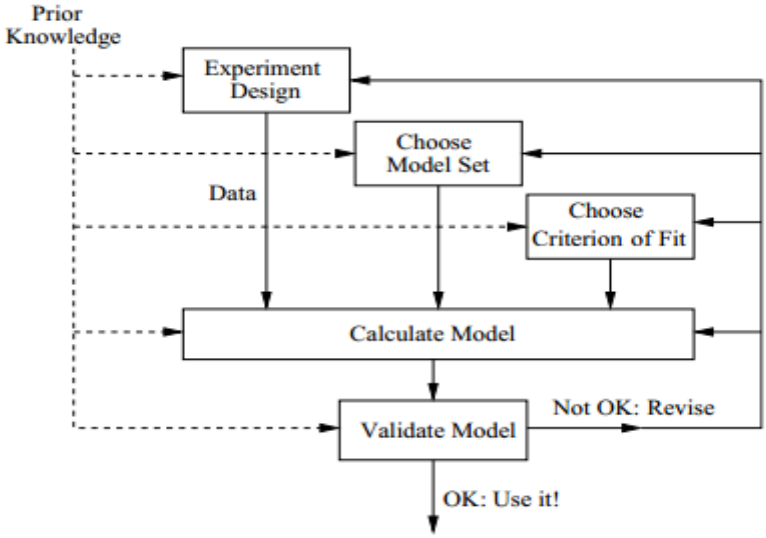

Fig 1.1 System identification loop

MATLAB allows us to build mathematical models of a dynamic system using the system identification toolbox. This toolbox features a flexible graphical user interface that aids in building accurate, simplified models of complex systems from observed data.

This Toolbox contains also all the common techniques to adjust parameters in all kinds of linear models **[2].** It also allows you to examine the models' properties, and to check if they are any good, as well as to preprocess and polish the measured data.

Estimating models for control systems is based on measured data. The models will describe the behavior of the observed data, which makes them related directly to the measurements. The data is treated as a time series that is why there are two types of estimation algorithms; online and offline estimation. Online estimation algorithms estimate the parameters of a model when new data is available during the operation of the model. In offline estimation, we first collect all the input/output data then we estimate the model parameters. Parameter values estimated using online estimation can vary with time, but parameters estimated using offline estimation do not [4].

3. Parametric Identification

It is also called conventional time domain identification; it deals with models containing a vector of parameters, which are to be adjusted so that the model mimics the system behavior as close as possible. Parametric identification relies on a model previously defined by a set of parameters that must be calculated to accomplish a given quality criterion. The system characteristics can have a parametric representation through a polynomial of a finite and known degree **[1]**. The model structure can be obtained by physical modeling (grey box) or it may be a standard one (black box). In the latter case, a set of generic standard structures must be taken into consideration such as Auto Regressive Moving Average eXogeneous(ARMAX) and its variants, Finite Impulse Response(FIR), Box-Jenkins (BJ),and Output Error (OE) models.(see Fig 1.2)

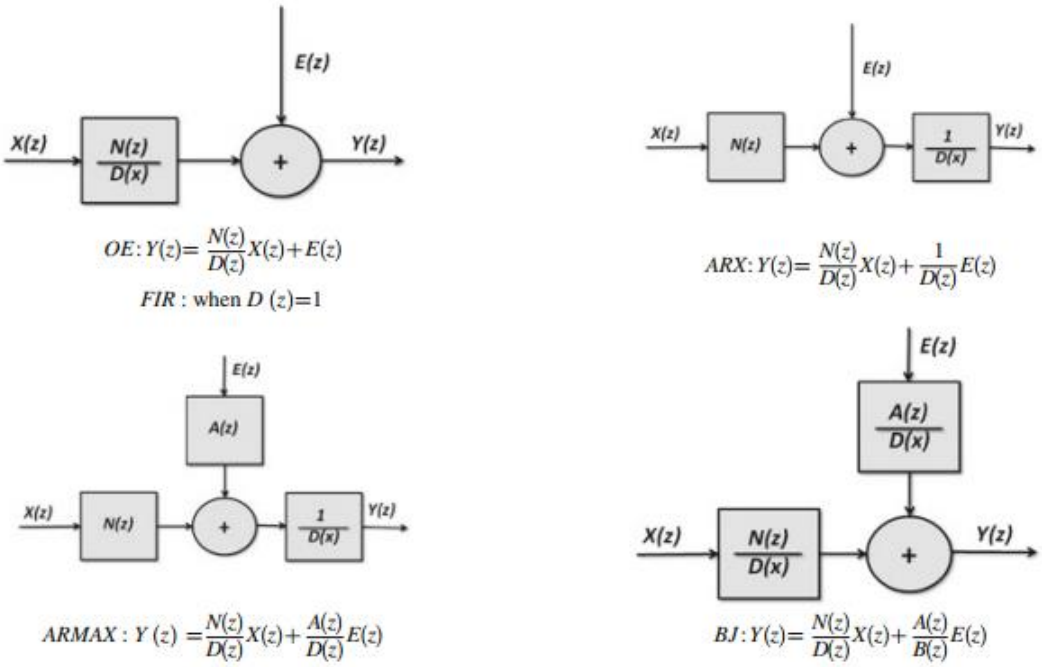

Fig. 1.2 System Model Structures

3.1. Linear Identification

Parametric identification techniques depend mostly on Prediction-Error Methods (PEM). The output of system y(t) can be expressed based on the z-transform as :

$$Y(z) = G(z)X(z) + W(z) \quad (1.1)$$

The expression (1.1) can be rewritten as follows:

$$Y(z) = G(z)X(z) + H(z)E(z) =(N(z))/(D(z)\ ) X(z) +(A(z))/(B(z)\ ) E(z) \quad (1.2)$$

Where E(z) is the transform of a white noise, $\epsilon(t)$. G(z) is the transfer function of the system, H(z) is the stochastic model of noise, and Y(z), X(z) are the z-transform of the output input, respectively.

The relationship between both functions defines several model structures.

Fig.1.2 shows the most common ones: Auto Regressivee Xogeneous(ARX) model, Auto Regressive Moving Average eXogeneous(ARMAX) model, Box-Jenkins (BJ) model and Output Error (OE) models.

The ARX model uses the past inputs and past outputs as regressors. This results in linear least square description where the cost function needs to be minimized.

$$\min_{\theta} E\left\{\sum_{k=1}^{N-\delta}(\hat{y}_{k+\frac{2}{k}}(\theta) - y_{k+2})^2\right\} \quad (1.3)$$

Where $(\hat{y}_{k+\frac{2}{k}}(\theta)$ is the estimate δ-step ahead predicted output of the system, θ is vector of unknown parameters, y_{k+2} represents the measured output, and E is the expected value of the squared z-step ahead prediction error.

The foremost disadvantage is that the disturbance model 1/N(z) comes along with the system's poles. It is, therefore, easy to get an incorrect estimate of the system dynamics because the A(z) polynomial can also include the disturbance properties. So, higher orders in A(z)and B(z) coefficients may be required. If the signal-to-noise ratio is good, this disadvantage is less important **[1].**

The ARMAX model has more flexibility in the handling of disturbance modeling than the ARX model. For this reason, ARMAX is a widespread used model and performs well in many engineering applications.

The FIR is the simplest model structure to be considered. The past inputs are used as regressors. The structure results again in a linear least square problem for minimizing the cost function (1.3). It requires many regressors and the convergence rate is slow **[1]**.

The OE model has the advantage that the system dynamics can be described separately and that no parameters are wasted on a disturbance model. If the system operates without feedback during the data collecting, a correct description of the transfer function G(z) = N(z)/D(z) can be obtained regardless of the nature of the disturbance **[1]**.

In the BJ model, the disturbances properties are modeled separately from the system dynamics.

Model validation is carried out by comparing the model behavior with the system's one and evaluating the difference. All models have a certain domain of validity. This may determine how exactly they are able to describe the system behavior. There are a number of different methods to set a criterion e.g., least squares **[3]**, generalized least squares **[3]**, maximum likelihood **[3]** or instrumental variables **[3]**. Some of them will be discussed in details in the following section.

3.1.1. Parameter estimation

Assuming that a certain model structure $\mathcal{M}$ has been chosen. $\mathcal{M}(\theta)$ denotes a particular model in the model set parameterized using the parameter setθ. The aim is to find the "best" model within the model set $\mathcal{M}^*$

$$\mathcal{M}^* = \{ \mathcal{M}(\theta) | \theta \in \mathcal{D}_{\mathcal{M}} \} \tag{1.3}$$

Where $\mathcal{D}_{\mathcal{M}}$is a closed subset of $\mathcal{R}^d$, and is the dimension of the parameter vector θ. We assume an experiment that is carried out on the process and measurement data are collected with equally spaced time intervals. Let the time index $k \in \{0, 1, 2, \ldots, N_s\}$.

The following notations are introduced:

$$u^k = [u_0, u_1, u_2, \ldots, u_k] \tag{1.4}$$

$$y^k = [y_0, y_1, y_2, \ldots, y_k] \tag{1.5}$$

i.e., u^k represents measured input up to timet_k, and y^k represents measured output up to timet_k. Denote the joint set of the input and output data as:

$$z^k = (u^k, y^k) \tag{1.6}$$

The parameter estimation problem is to use the data set z^N to select a suitable value $\hat{\theta}_N$, and hence a model $\mathcal{M}(\hat{\theta}_N)$ from the model set $\mathcal{M}^*$. Note that this is named off-line estimation **[4]**, as all the measured inputs and outputs at all sampling times are used for parameter estimation. In some cases, it may be of interest to estimate parameters at some certain time t_k, using all measurement data up to this time point (z^k), this is termed on-line estimation **[4]**, or recursive identification **[5].**

3.1.2. Prediction error method

The performance of a model is judged by its ability to predict the outputs of the system. According to Ljung (1999), most of the methods used for parametric estimation can be characterized as general prediction error methods (PEM). The prediction error is defined as the difference between the predicted output and the measured output

$$\epsilon(t_k, \theta) = y^k - \widehat{y^k} \tag{1.7}$$

where y^k and $\widehat{y^k}$ are the measured and estimated outputs at time $t_k$respectively. By assumption, the estimated output depends on the parameter setθ, i.e., $\widehat{y^k} = \hat{y}(t_k|\,\theta)$.
The criterion may be further expanded by filtering the prediction error, through a stable linear filter L(q):

$$\epsilon_f(t_k|\theta) = L(q)\epsilon(t_k|\theta) \tag{1.8}$$

where L acts like a frequency weighting of the criterion. By doing this, the system properties in frequencies of specific interest can be emphasized.

The principle of fitting parameterized models to data is based on choosing some norm of the prediction error, which is often termed the loss function, and then find the parameter vector θ that minimizes this loss function.

$$\hat{\theta}_{N_s} = \arg\min V_{N_s}(\theta, z^{N_s}) \tag{1.9}$$

A general form of loss function can be represented as:

$$V_{N_s}(\theta, z^{N_s}) = \frac{1}{N_s}\sum_{k=1}^{N_s} l(\epsilon_f(t_k|\,\theta)) \tag{1.10}$$

where l denotes a scalar valued, positive function which is used to measure the norm of the filtered prediction error. Different ways of forming the function l and choosing the prefilter L(q) lead to different estimation methods, e.g., the least squares (LS) method and the maximum likelihood (ML) method. To simplify the notation, we omit N_s in the following since we are here mainly concerned with off-line estimation and the whole sequence of measurement data are known.

3.1.3. Least squares (LS) method

In the least squares method, the loss function is defined as a quadratic residual function, dependent on the parameter vectorθ.

$$V(\theta) = \frac{1}{N_s}\sum_{k=1}^{N_s} \epsilon^2(t_k|\theta)) \tag{1.11}$$

Thus, the quadratic form of the residual between the estimated outputs and the measured outputs is minimized

$$\hat{\theta} = \arg\min \frac{1}{N_s}\sum_{k=1}^{N_s} \epsilon^2(t_k|\theta)) \tag{1.12}$$

which is known as a least square estimator.

when each residual is multiplied with a certain weight factor, the estimator is called weighted least squares (WLS) estimator.

$$\min_{\theta} V(\theta) = \frac{1}{N_s}\sum_{k=1}^{N_s} \alpha_k \epsilon^2(t_k|\theta)) \tag{1.13}$$

where $\alpha_k$is the weight factor.

3.1.4. Maximum likelihood (ML) method

The way of forming the loss function in the (LS) method is completely deterministic. In contrast, to a statistical sense, the measurement data are stochastic realizations, since the system in general can be seen to be corrupted by stochastic disturbances and/or measurement noise. A probabilistic approach to constructing the loss function is the well-known maximum likelihood method. The objective of maximum likelihood estimation is to find the parameter set that maximizes the probability that the observed data are explained by the model. More explicitly, a likelihood function is defined as the joint probability density of all the observation data assuming that the parameter set is known

$$L(\theta) = \prod_{k=1}^{N_s} p(y^k|y^{k-1},\theta) \tag{1.14}$$

Where $p(y^k|y^{k-1},\theta)$is the probability density function (PDF) of y^k depending on previous measurement data set $y^{k-1}$and parameter setθ. Assuming that the prediction error sequence $\{\epsilon_k\}$ consist of zero-mean, independent stochastic variables with the probability density function expression $p(\epsilon_k(\theta)|\theta)$.

The eq (1.14) can then be rewritten as:

$$L(\theta) = \prod_{k=1}^{N_s} p(\epsilon_k(\theta)|\theta) \quad (1.15)$$

The maximum likelihood estimator determines the parameter set by maximizing the likelihood function (1.15), which is equivalent to minimizing

$$V(\theta) = -\log L(\theta) = -\sum_{k=1}^{N_s} \log\left(p(\epsilon_k(\theta)|\theta)\right) \quad (1.16)$$

The maximum likelihood method is often considered to be the optimal method for parameter estimation, in the sense that it asymptotically approaches the best achievable results, namely, lowest estimate variances, under the assumption that the true system is within the model set. Furthermore, its asymptotic properties make it very useful for model validation by using different kinds of likelihood based statistical tests.

Note that the ML-estimator is equivalent to the WLS-estimator under some specific condition, i.e., when the measurement errors are normally distributed with known covariance matrix and the elements of the inverse covariance matrix are used as weights.

4. Non-Parametric Identification

Nonparametric identification techniques provide a very effective and simple way of finding model structure in data sets without the imposition of a parametric one **[1].** Its methods aim at determining the system functions without first selecting set of confined possible models. Such methods are often called nonparametric since they do not employ a finite-dimensional parameter vector in the search of the best description **[1].** Commonly, the initial process to carry out is the nonparametric identification, and then, if it were suitable, the parametric identification should be performed. The next sections review the non-parametric identification methods from time domain and frequency domain perspectives.

It is difficult to establish a clear identification methodology of nonlinear systems, since analysis is usually more intricate than in the identification of linear models, because of the variety of nonlinear model structures and nonlinear behaviors. For instance Donoho and Johnstone **[6]** and Donoho **[7]** introduced nonlinear wavelet estimators in nonparametric regression through thresholding, i.e., the term-by-term assessment of coefficients in the

wavelet expansion. Only coefficients that exceed a predetermined threshold are taken into account. This produces the wavelet shrinkage. Bendat describes procedures to identify and analyze the properties of many types of nonlinear systems as Zero-Memory Nonlinear Systems and Parallel Nonlinear System, with analysis of Nonlinear System Input/Output Relationships **[8]**.

Zhang applied wavelet theory for nonlinear system identification with a wavelet basis as a universal function approximator, with a neural network used to determine the resolution, and the translation coefficients of the wavelet **[9].** This nonparametric estimator named wavelet neural network has a neural network like structure that makes use of techniques of regressor selection completed with back propagation procedure **[9].**

4.1. Non-parametric Identification in Time Domain

4.1.1. The Impulse Response

The notion of characterizing a dynamical system by its impulse (or pulse) response dates from the earliest forays of process engineers into system identification.

Let us assume the following system

$$y(t) = G_0(q).u(t) + v(t) \tag{1.20}$$

or equivalently

$$y(t) = \sum_{k=0}^{\infty} g_0(k).u(t-k) + v(t) \tag{1.21}$$

We subject our system to a pulse input

$$u(t) = \begin{cases} \alpha, & t = 0 \\ 0, & t \neq 0 \end{cases} \tag{1.22}$$

Then the output will be

$$y(t) = \alpha g_0(t) + v(t) \tag{1.23}$$

If the noise is low, it is thus possible to determine the impulse-response coefficient $g_0(t)$ from an experiment with a pulse input. The estimate will be

$$\hat{g}(t) = \frac{y(t)}{\alpha} \tag{1.24}$$

and the error is v(t)/ α. Its basic weakness is that many physical processes do not allow pulse inputs of such amplitude that the error is insignificant compared to the impulse-response coefficients. The possible damage inflicted by direct use of an impulsive test signal on control system hardware and the presence of output noise has led to a decline in its use **[4].**

4.1.2. Cross-Correlation Approach

Cross-covariance is a non-parametric identification technique and is related with the impulse response. Correlation means how two variables are related together.

The correlation between the input and output with the assumption that the mean is zero for y and u is the eq. (1.21):

$$y(t) = \sum_{k=0}^{\infty} g_0(k).u(t-k) + v(t)$$

v is the noise in the system.

If the input and output are uncorrelated, the cross covariance between them is:

$$R_{uy}(\tau) = g(\tau) * R_{uu}(\tau) \qquad (1.25)$$

That is, the cross correlation is the convolution between the impulse response and the autocorrelation of the input. Thus, the impulse response can be estimated from the covariance (correlation if both signals have zero mean) if the input is a white noise **[4].**

If the input is chosen as white noise so that:$R_{uu}(\tau) = \alpha\delta_{T_0}(\tau)$

then $$g_0(\tau) = \frac{R_{uy}(\tau)}{\alpha} \qquad (1.26)$$

where g_0 is an estimate for the impulse response is obtained from an estimate of R_{uy}.

4.1.3. Step-Response Analysis

The response of the system can be determined by applying a step input:

$$u(t) = \begin{cases} \alpha, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

as follows

$$y(t) = \alpha \sum_{k=1}^{t} g_0(k) + v(t) \qquad (1.27)$$

For this the estimate of $g_0(k)$ could be obtained as:

$$g_0(\tau) = \frac{y(t)-y(t-1)}{\alpha} \quad (1.28)$$

This method is useful for obtaining qualitative information about the system; more specifically, it shows the transient behavior of the system through Dead Time, Static Gain and time constant **[4].**.

4.2. Non-Parametric Identification in Frequency Domain

The frequency domain characterization of system dynamics has, like its time domain counterpart, a long and varied history **[10].** As far as control engineering is concerned, however, frequency domain identification gained deep relevance with the development of stability and design methods based upon frequency response measurements **[10].**

4.2.1. Transfer Function

The fundamental physical interpretation of the transfer function G(z) is that the complex number $G(e^{jw})$ bears information about what happens to an input sinusoid.

$$u(t) = \alpha \ cos\, wt \quad t=0, 1, 2\ldots \quad (1.29)$$

then

$$y(t) = \alpha \left|G_0(e^{jw})\right| cos(wt + \varphi) + v(t) \quad (1.30)$$

where $$\varphi = arg\, G_0(e^{jw}) \quad (1.31)$$

This is known as frequency analysis and it is a simple method for obtaining detailed information about the system.

4.2.2. Frequency Response using Correlation

With the noise component v(t), it may be cumbersome to determine $\left|G_0(e^{jw})\right|$ and φ accurately by graphic methods. Since the interesting component of y(t) is a sine function of known frequency, it is possible to correlate it out from the noise in the following way. From the sums

$$I_c(N) = \frac{1}{N}\sum_{t=1}^{N} y(t)\, cos\, wt \quad (1.32)$$

$$I_s(N) = \frac{1}{N}\sum_{t=1}^{N} y(t)\, sin\, wt \quad (1.33)$$

We substitute y(t) by its expression (1.32) in (1.33), we obtain

$$I_c(N) = \frac{1}{N}\sum_{t=1}^{N} \alpha\left|G_0\left(e^{jw}\right)\right| cos(wt + \varphi)\, cos\, wt + \frac{1}{N}\sum_{t=1}^{N} v(t)\, cos\, wt$$

$$= \frac{\alpha}{2N}\left|G_0\left(e^{jw}\right)\right| \sum_{t=1}^{N}[cos(2wt+\varphi)+cos\,\varphi] + \frac{1}{N}\sum_{t=1}^{N} v(t)\,cos\,wt$$

$$= \frac{\alpha}{2}\left|G_0\left(e^{jw}\right)\right| cos\,\varphi + \frac{\alpha}{2N}\left|G_0\left(e^{jw}\right)\right| \sum_{t=1}^{N}[cos(2wt+\varphi)] + \frac{1}{N}\sum_{t=1}^{N} v(t)\,cos\,wt \tag{1.33}$$

The second term tend to zero as N tends to infinity.

Similarly, for (1.32)

$$I_s(N) = -\frac{\alpha}{2}\left|G_0\left(e^{jw}\right)\right| \sin\varphi + \alpha\left|G_0\left(e^{jw}\right)\right| \frac{1}{2}\frac{1}{N}\sum_{t=1}^{N}[\sin(2wt+\varphi)] + \frac{1}{N}\sum_{t=1}^{N} v(t)\sin wt \tag{1.34}$$

These two expressions (1.33) and (1.34)suggest the following estimates of $\left|G_0\left(e^{jw}\right)\right|$ and φ, respectively

$$\left|\widehat{G}_N\left(e^{jw}\right)\right| = \frac{2\sqrt{I_c^2(N)+I_s^2(N)}}{\alpha} \tag{1.35}$$

$$\widehat{\varphi} = \arg\widehat{G}_N\left(e^{jw}\right) = -\arctan\frac{I_s(N)}{I_c(N)} + k\pi \qquad k=0.1.2...N \tag{1.36}$$

Repeating this procedure for a number of frequencies, a good picture of $G_0(e^{jw})$ over the frequency domain of interest can be obtained.

This method allows us to obtain easily the Bode plot of the system, but many industrial processes do not admit sinusoidal inputs in normal operations.

5. Conclusion

In this first chapter of the thesis, we have discussed different techniques of system identification, such as Parametric and non-parametric methods, for linear and nonlinear systems, and in Time Domain and Frequency Domain.

We have also discussed the different procedures of system identification, and how to identify using MATLAB toolbox starting from measuring the input and output signals of a given system in time or frequency domain.

When, model structure is selected, an estimator is applied online and offline in order to find estimate values for a parametric system or to determine system functions defining a non-parametric system.

Chapter Two

Distributed Parameter System

1. Introduction

Distributed parameter system (DPS**)** is an established area of research in control that can trace its roots back to the sixties **[11]**. While the general aims are the same as for lumped parameter systems, to adequately describe the distributed nature of the system one needs to use partial differential equation (PDE) models, or the Ordinary Difference Equation (ODE).

In this chapter, we are going to present the Distributed Parameter System; starting with its history and first research about it, then we speak about its characteristics and what makes the Distributed Parameter Systems different than Lumped Parameter Systems. Finally we state some examples of such systems in the industry and control domain.

2. Brief Definitions of Distributed Parameter System

Distributed effects are present in almost all physical systems. In some cases, these can be safely ignored but there are many interesting problems where these effects must be taken into account.

Distributed Parameter Systems are systems whose state space is infinite-dimensional; that is why they are called infinite-dimensional systems. They can be represented or described using Partial Differential Equations or Delay Differential Equations, this gives rise to distinctive features.

3. Mathematical Description of DPS

It appears from the literature that Paraskevopoulos and Bounas, 1978 were the first investigators to study the identification of distributed parameter systems via orthogonal functions **[12]**. They used Walsh functions. The main drawback of the Walsh function approach is the selection of Walsh functions based on 2^k where k is any positive integer. For a moderately large value of k this approach becomes computationally laborious.

In the same year Tzafestas **[13]**, 1978 has also investigated the general distributed parameter system identification by first transforming the model of a distributed parameter system into its equivalent lumped form by using Galerkin expansion .In 1983, Paraskevopoulos and Kekkeris, revisited the same problem of Paraskevopoulos and Bounas. This time they employed Tchebycheff polynomials of the first kind and restricted their investigations to only first-order systems.

3.1. Unified Approach for Identification

Consider the model of a linear time-invariant distributed parameter system described by:

$$a_{tt}\frac{\partial^2 y(x,t)}{\partial t^2} + a_{xx}\frac{\partial^2 y(x,t)}{\partial x^2} + a_{xt}\frac{\partial^2 y(x,t)}{\partial x \partial t} +$$
$$a_t \frac{\partial y(x,t)}{\partial t} + a_x \frac{\partial y(x,t)}{\partial x} + a y(x,t) = u(x,t) \quad (2.1)$$

With initial conditions $y(x,t_0)$ and $\frac{\partial y(x,t)}{\partial t}|_{t=t_0}$ and boundary conditions $y(x_0,t)$ and $\frac{\partial y(x,t)}{\partial t}|_{x=x_0}$

To identify the system given by (2.1), the input $u(x,t)$ and the output $y(x,t)$ of the system are assumed to be available over the region x ϵ (x_0, x_f), (t_0, t_f).

The aim is to estimate the system parameters $a_{tt}, a_{xx}, a_{xt}, a_x, a_t, a$

It may be noted that depending upon the value of $\Delta = a_{xt}{}^2 - 4a_{tt}a_{xx}$, the system described by (2.1) turns out to be

i) An elliptic system for $\Delta < 0$.

ii) A parabolic system for $\Delta = 0$ and

iii) A hyperbolic system for $\Delta > 0$.

3.2. DPS Subjected to Additive Noise

Let y(x,t) be the scalar system state.

$$\frac{\partial y(x,t)}{\partial t} = L_x y(x,t) + f(x,t) + G(x,t)\gamma(x,t), \quad t\epsilon T =]0, t_f[, \quad x\epsilon D, \quad (2.2)$$

with initial and boundary conditions

I.C. $y(x,0) = y_0(x)$ $x\epsilon D$ (2.3)

B.C $B_x\, y(x,t) = 0$ $t\epsilon]0, t_f[$, $x\epsilon D$ (2.4)

Where both L_x and B_x are well-posed linear spatial differential operators, $\gamma(x,t)$ is the zero-mean Gaussian white (with respect to t) distributed noise process, f and G are known functions and $y_0(x)$ is the initial state function which is usually assumed to be the Gaussian stochastic variable for each x independent of $\gamma(x,t)$.There are many cases where one or more parameters of L_x are unknown.

A mathematical model of a general class of nonlinear DPSs is written by using a conventional function F in the form:

$$\frac{\partial y(x,t)}{\partial t} = F(x, t, y, y_x, y_{xx}, \dots) + G(x, t)\gamma(x,t) \quad (2.5)$$

Where, for the mathematical security, the function F is usually assumed to be sufficiently smooth with respect to its arguments, and where u_x is the partial derivative with respect to x. In order to handle practical problems, Eq.(2.9) is written in a more concrete form,

$$\frac{\partial y(x,t)}{\partial t} = N_s(x, t, u) + G(x, t)\gamma(x,t) \quad (2.6)$$

With the boundary condition which is considered to be also nonlinear, e.g.

$$N_b(x, t, y)=0 \qquad t\epsilon]0, t_f[, \qquad x\epsilon \partial D \quad (2.7)$$

and with the same initial condition as given by (2.3). It should be noted that, if there exists an additive noise at the system boundary, then the noise term $\varepsilon_b(x,t)$ appears on the right hand side of both Eqs.(2.4) and (2.7).

3.3. DPS with Stochastic Coefficients

In many practical problems, one or more coefficients in differential operators are random because measurements of physical properties of the system considered inherently exhibit greater uncertainty. It may therefore be more realistic to consider system parameters as stochastic variables with a probability distribution. Many different forms have been proposed for DPSs and each is useful for different tasks **[14].** A general form using the parabolic type, which arises directly from fundamental physical axioms, is as follows:

$$\frac{\partial y(x,t)}{\partial t} = A(x, t, w;\ D_x)y(x, t) + f(x, t) \qquad (x,t)\ \epsilon\ T \times D \quad (2.8)$$

With the boundary condition

$$B(x, t; D_x)y(x, t)=0 \qquad (x,t)\ \epsilon\ T \times D \quad (2.9)$$

and with the same initial condition as given by (2.3), where A is a partial differential operator containing stochastic coefficients, B is a boundary operator with deterministic non vanishing coefficients and *w* is the generic point of the sample space. Since unknown stochastic

coefficients are involved in A, the problem at hand belongs to the context of stochastic inverse problems and is motivated by the oil exploration survey, the experimental physiology, cardiology, etc.

3.4. DPS with Free Boundary

Recent advances in control technique have stirred a great deal of enthusiasm in the development of parameter identification and state estimate of DPSs with free boundaries **[14].** A number of systems of practical interest involve boundaries moving by phase change such as melting or solidification, chemical reaction, heat transfer and so on. Considering a spatial region consisting of water and ice, and choosing the system state as the temperature distribution of the water region, the water-ice interface becomes a moving boundary denoted by y(t). For example, consider a mathematical model of the system:

$$\frac{\partial y(x,t)}{\partial t} = a\,\frac{\partial^2 y(x,t)}{\partial x^2} \qquad (x,t) \in TxD \tag{2.10}$$

with I.C $$y(x,0) = y_o(x) \geq 0 \qquad \text{for } x \in D \tag{2.11}$$

On the boundary i.e. the ice-water interface, by considering the latent heat of fusion and the rate at which ice is converted into water, it follows that, for $0 < t < tf$,

$$ay(0,t) = g(t),\ y(t,x) = 0,\ a\frac{\partial y(t,x)}{\partial x} = -\frac{dx}{dt}, \tag{2.12}$$

where a is a positive constant and the ice region is assumed to be bounded. Fig. (2.1) illustrates the construction of a DPS with free boundary. A fundamental difference between regular boundary problems and free boundary ones is that the domain of solutions to the basic state equation is not known but should be determined by additional information through underlying properties of physical systems under consideration **[14]**.

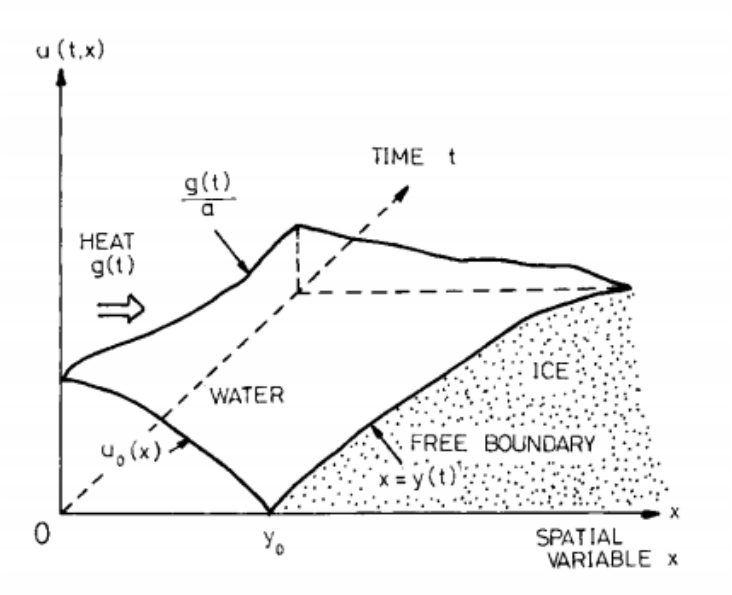

Fig.2.1 DPS with free boundaries.

3.5. DPS with Inequality Boundary Conditions

There has recently been much practical interest shown in problems of the DPS with a permeable wall as its system boundary **[14]**. Identification problems of system parameters in which the osmotic pressure dues to the nature of membrane are of practical importance. More basically, in the field of synovial pint biomechanics, the permeability plays an important role of analyzing locomotion of the musculoskeletal system. From theoretical viewpoints, a remarkable feature of the problem considered here relates to a mathematical modelling of the system boundary.

Suppose that the region D is filled up by the liquid as shown in Fig.(2.2). Then, the state variable y(x,t) is reasonably selected to be the pressure of the fluid. The boundaries of D are respectively denoted by Γ_0 at x = 0 and by Γ_1 at x = 1. The boundary conditions peculiar to the system shown in Fig.(2.2). are considered as follows. At the boundary Γ_0, the well-known Neumann condition is set in a form of:

$$-a\frac{\partial y(x,t)}{\partial x}\Big|_{x=0} = f(t) \quad \text{on } \Gamma_0 \tag{2.13}$$

Where a is a positive constant expressing the diffusion coefficient and f(t) is the known input function applied to the boundary Γ_0.

On the other hand, the boundary Γ_1 consists of a semi-permeable membrane whose thickness is supposed to be negligible. This implies that, as shown in Fig.(2.2), although the fluid flow leaving D freely passes Γ_1, the counter flow is prevented by the existence of the membrane. Let h(t) be the given fluid pressure in the outside of D.

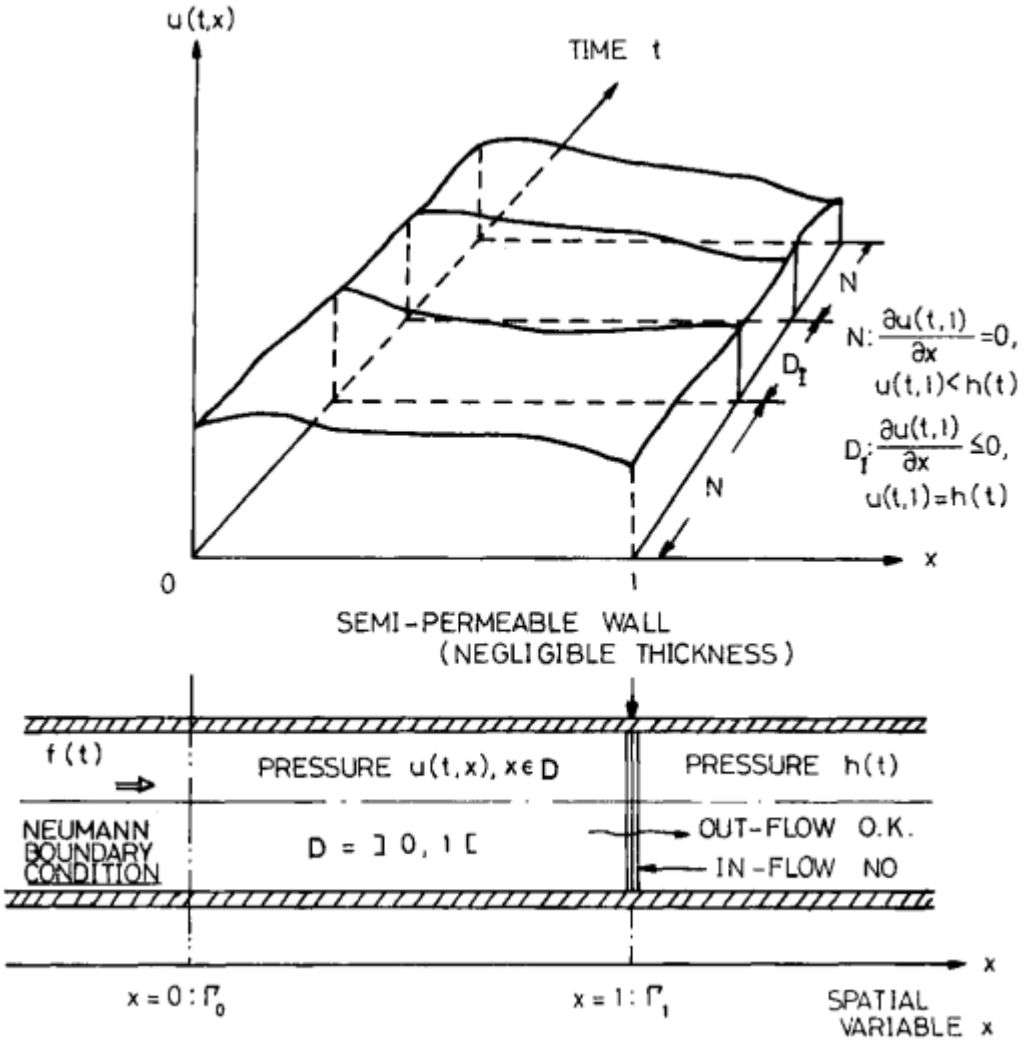

Fig.2.2 System with semi-permeable wall.

First, suppose that

Case-1: y(1,t) < h(t); There exists a fight where the semi-permeable wall prevents the fluid flow against the fluid trying to enter into D. A success of the fight allows us to write

$$\frac{\partial u(x,t)}{\partial x}|_{x=0} = 0 \quad \text{on } \Gamma_0 \qquad (2.14)$$

Case-2: y(1,t)> h(t); This case implies that the fluid tries to leave D.

Hence $$\frac{\partial y(x,t)}{\partial x}|_{x=1} \leq 0 \quad \text{on } \Gamma_1 \qquad (2.15)$$

However, since the thickness of the semi-permeable wall is negligible, y(1,t) is not greater than h(t), i.e., y(l,t) = h(t). The results mentioned above are summarized with the same state equation as given by Eq. (2.10) and the same initial conditions as given by (2.3), the boundary conditions are modelled by (2.13), (2.14) and (2.15).

4. Lumped Vs Distributed Parameter System

Lumped Parameter Systems are modeled by ODEs (Ordinary Difference Equations). Their state space are finite dimensional. It is a system in which the dependent variables of interest are a function of time alone. Whereas Distributed Parameter Systems are modeled by PDEs or DDEs (Delay Differential Equations). Their state spaces are infinite-dimensional. It is a system in which all dependent variables are functions of time and one or more spatial variables.

For example, consider the following two systems illustrated in the following figure:

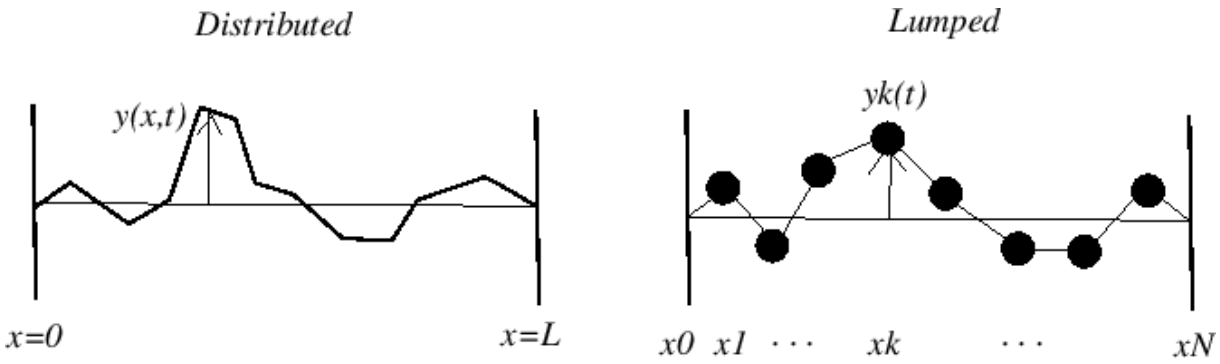

Fig.2.3 Distributed Vs Lumped System

The first system is a distributed system, consisting of an infinitely thin string, supported at both ends; the dependent variable, the vertical position of the string $y(x,t)$is indexed continuously in both space and time.

The second system, a series of ``beads" connected by massless string segments, constrained to move vertically, can be thought of as a lumped system, perhaps an approximation to the continuous string.

For electrical systems, consider the difference between a lumped RLC network and a transmission line (see figure 2.4)

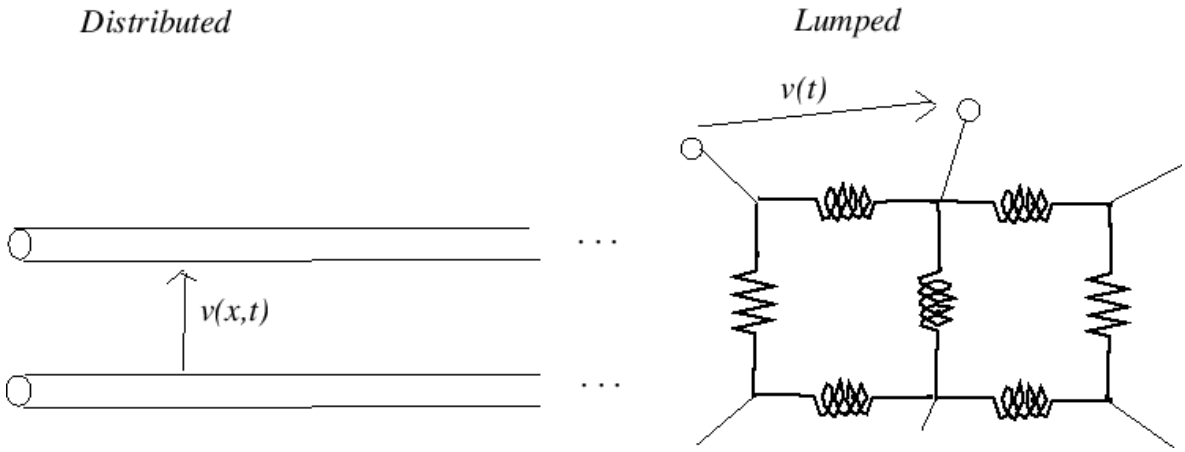

Fig.2.4 Transmission line Vs Lumped RLC system

The importance of lumped approximations to distributed systems will become obvious later, especially for waveguide-based physical modeling, because it enables one to cut computational costs by solving ODEs at a few points, rather than a full PDE (generally much more costly) **[15].**

5. Examples

A wide variety of phenomena can be represented by the PDE of Distributed Parameter Systems, from sound, to heat, electrostatics, electrodynamics, fluid flow, etc. These seemingly distinct physical phenomena can be formalized similarly in terms of PDEs. Just as ordinary differential equations often model one-dimensional dynamical systems, partial differential equations often model multidimensional systems.

5.1. Wave Equation

The wave equation is an equation for an unknown function y(k, x) of the form:

$$y_{kk} = m^2 y_{xx} \tag{2.16}$$

Here *y* might describe the displacement of a stretched string from equilibrium, or the difference in air pressure in a tube, or the magnitude of an electromagnetic field in a tube, and *m* is a number that corresponds to the velocity of the wave**[16]**.

5.2. Euler-Tricomi Equation

In mathematics, the Euler–Tricomi equation is a linear partial differential equation useful in the study of transonic flow **[17]**. It is named for Leonhard Euler and Francesco Giacomo Tricomi and is given by

$$y_{xx} = z y_{zz} \tag{2.17}$$

It is hyperbolic in the half plane x > 0, parabolic at x = 0 and elliptic in the half plane x < 0. Its characteristics are:

$$x dx^2 = dz^2 \tag{2.18}$$

Which have the integral

$$z \pm \frac{2}{3} x^{\frac{3}{2}} = C \tag{2.19}$$

where C is a constant of integration. The characteristics thus comprise two families of semi cubical parabolas, with cusps on the line x = 0, the curves lying on the right hand side of the z-axis.

5.3. Heat Equation

The equation for conduction of heat in one dimension for a homogeneous body has

$$y_t = \alpha y_{xx} \tag{2.20}$$

where y(t,x) is temperature, and α is a positive constant that describes the rate of diffusion.

We are going to discuss and study this application in details later **[16]**.

6. Conclusion

In this part, we have presented the distributed parameter systems with its characteristics and features. We have discussed its history and mentioned the first studies about the topic done by Scientists and mathematicians. Then we have introduced a comparison of lumped and distributed parameter systems; their models and applications. Finally, we have shown some examples and physical phenomena that can be represented by PDE.

Chapter Three

Heat Conduction Application

1. Introduction

In this chapter, we present an application of parametric identification of a distributed parameter system which represent the dynamic behavior of the heat conduction. The parameters of the heat equation are identified using Least Square method.

The application of the least square estimator is divided into three main parts. The first two parts are divided into the estimation of parameters sets for two different numerical simulations. These estimates are obtained for different noise amplitude and using different structures. However, the third part is about the identification of conduction of the heat in a cement rotary kiln.

2. Description of Heat Conduction

Heat transfer is of particular interest to engineers, who attempt to understand and control the flow of heat through the use of thermal insulation, heat exchangers, and other devices. Heat transfer is defined as the flow of thermal energy within solids and non-flowing fluids, driven by thermal non-equilibrium (i.e. the effect of a non-uniform temperature field), commonly measured as a heat flux (vector), i.e. the heat flow per unit time (and usually unit normal area) at a control surface.

In a metal rod with non-uniform temperature, heat (thermal energy) is transferred from regions of higher temperature to regions of lower temperature. Three physical principles are used here:

1. Heat (or thermal) energy of a body with uniform properties

$$\text{Heat energy} = c*m*T \tag{3.1}$$

where m is the body mass, T is the temperature, c is the specific heat which is the amount of heat required to change the body's temperature by one degree.

2. Fourier's law of heat transfer represent rate of heat transfer proportional to negative temperature gradient,

$$\frac{\text{Rateofheattransfer}}{\text{area}} = -K_0 \frac{\partial T}{\partial x} \tag{3.2}$$

where K_o is the thermal conductivity. In other words, heat is transferred from areas of high temperature to low temperature.

3. Conservation of energy:

Consider a uniform rod of length *l* with non-uniform temperature lying on the *x*-axis from x =0 to x = l. By uniform rod, we mean the density ρ, specific heat c, thermal conductivity K_o, cross-sectional area A are all constant. Assume the sides of the rod are insulated and only the ends may be exposed. Also, assume there is no heat source within the rod. Consider an arbitrary thin slice of the rod of width Δx between x and x+Δx. The slice is so thin that the temperature throughout the slice is T(x,t).

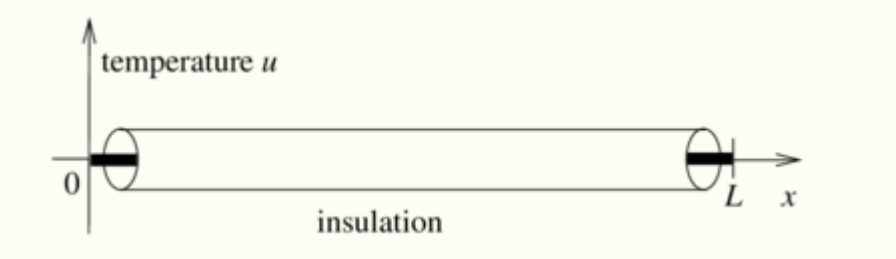

Fig 3.1 uniform insolated rod of length L.

Thus,

$$\text{Heat energy of segment} = c \times \rho A \Delta x \times T = c\rho A \Delta x T(x, t) \qquad (3.3)$$

By conservation of energy,

Change of heat energy of segment in time Δt	=	Heat in form Left boundary	-	Heat out form Right boundary

From Fourier equation (3.2)

$$c\rho A \Delta x T(x, t + \Delta t) - c\rho A \Delta x T(x, t) = \Delta t A \left(\frac{\partial T}{\partial x}\right)\Big|_{x} - \Delta t A \left(\frac{\partial T}{\partial x}\right)\Big|_{x+\Delta x} \qquad (3.4)$$

Rearranging (3.4) yields (recall ρ, c, A, K0 are constant),

$$\frac{y(x,t+\Delta t) - y(x,t)}{\Delta t} = -\frac{K_0}{c\rho}\left(\frac{\left(\frac{\partial y}{\partial x}\right)_x - \left(\frac{\partial y}{\partial x}\right)\Big|_{x+\Delta x}}{\Delta x}\right) \qquad (3.5)$$

Taking the limit Δt,Δx→0 gives the heat equation:

$$\frac{\partial y}{\partial t} = k\frac{\partial^2 y}{\partial x^2} \tag{3.6}$$

where $k = \frac{K_0}{c\rho}$ is called the thermal diffusivity. Since the slice was chosen arbitrarily, the Heat Equation (3.6) applies throughout the rod.

4. Initial condition and boundary conditions

To make use of the Heat Equation (3.6), we need more information:

1. Initial Condition(IC): In this case, the initial temperature distribution in the rod is T(x, 0).
2. Boundary Conditions (BC):

In this case, the temperature of the rod is affected by what happens at the ends, x =0, and x=*l*. What happens to the temperature at the end of the rod must be specified. In reality, the BCs can be complicated.

Temperature prescribed at a boundary

$$T(0,t)= T_1(t)$$

$$T(L,t)=T_l(t)$$

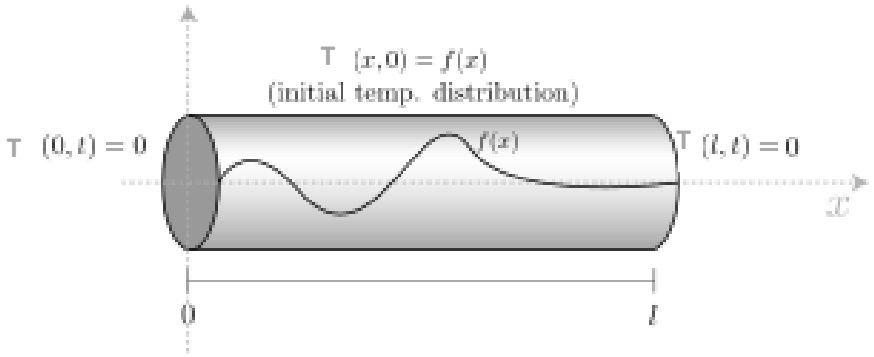

Fig 3.2 Description of Initial and boundary conditions in the rod.

3. Discretization of Heat equation using Euler Method

For the following problem involving the heat equation with a source term u(x,t)

$$T_t = \beta T_{xx} + f(x,t) + u(x,t), \quad a < x < b, t > 0, \tag{3.7}$$

$$BC: T(a,t) = g_1(t), \quad T(b,t) = g_2(t), \quad \text{IC:} \quad T(x,0) = T_0(x) \tag{3.8}$$

Let us seek a numerical solution for T(x, t) at a particular time t> 0 or at certain times in the interval $0 < t < t_f$.

As the first step, we expect to generate a grid

$$x_i = a + ih, \quad i=0, 1,\ldots, m, \quad h = \frac{b-a}{m}$$

$$t^k = k\Delta t, k=0, 1,\ldots, n, \quad \Delta t = \frac{T}{n}$$

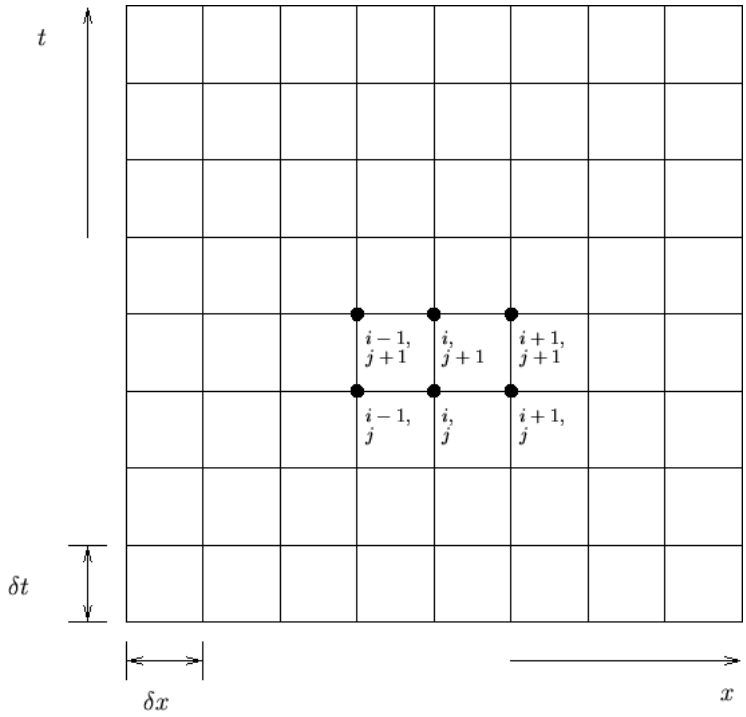

Fig 3.3 generation of the grid.

It turns out that any arbitrary Δt cannot be used for explicit methods because of numerical instability concerns. The second step is to approximate the derivatives with finite difference (FD) approximations. Since we already know how to discretize the spatial derivatives, let us focus on possible FD formulas for the time derivative. At grid point (x_i, t^k), $k \geq 0$ using the forward FD approximation for $T_i$central and FD approximation for T_{xx} we have:

$$\frac{T(x_i,t^k+\Delta t)-T(x_i,t^k)}{\Delta t}=\beta\frac{T(x_{i-1},t^k)-2T(x_i,t^k)+T(x_{i+1},t^k)}{h^2}+f(x_i,t^k)+u(x_i,t^k) \quad (3.9)$$

The discretization is first order in time and second order in space, when the FD equation is:

$$\frac{(T_i^{k+1})-(T_i^k)}{\Delta t}=\beta\frac{(T_{i-1}^k)-2(T_i^k)+(T_{i+1}^k)}{h^2}+f_i^k+u_i^k \quad (3.10)$$

where $f_i^k = f(x_i,t^k)$ with T_i^k again denoting the approximate values for the true solution $T(x_i,t^k)$. When k =0, T_i^0 is the initial condition at the grid point $(x_i,0)$ and from the values T_i^k at the time level k the solution of the FD equation at the next time level k+ 1 is

$$(T_i^{k+1}) = (T_i^k) + \Delta t\,(\beta\frac{(T_{i-1}^k)-2(T_i^k)+(T_{i+1}^k)}{h^2}+f_i^k+u_i^k), \quad i=1,2\ldots,m-1 \quad (3.11)$$

Let $\alpha_1 = \frac{\Delta t\beta}{h^2}$ and $\alpha_2 = 1-2\frac{\Delta t\beta}{h^2}$

$$Tes_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_1 T_{i-1}^j \quad (3.12)$$

4. Numerical Simulation

4.1. Homogenous heat equation

The first application deals with a linear homogenous heat equation. This example is solved using the separation of variables method. The exact solution of the heat equation is used to generate data that is employed in the identification process.

The example is about an insulated unit wire, such that its ends are embedded in ice (temperature 0°). Let k=0.003 and initial distribution is T(x,0)=50x(1-x). The Heat equation is:

$$\frac{\partial T}{\partial t} = 0.003\frac{\partial^2 T}{\partial x^2} \quad (3.13)$$

with Boundary Conditions: $T(0,t) = T(1,t)=0$

The exact solution of th equation (3.13) is:

$$T(x,t) = \begin{cases} \sum_{n=1}^{\infty}\frac{400}{\pi^3 n^3}\sin(n\pi x)e e^{-n^2\pi^2 0.003t} & ,n\ odd \\ 0 & ,n\ even \end{cases} \quad (3.14)$$

The spatial-temporal evolution of temperature is shown in the figure 3.4

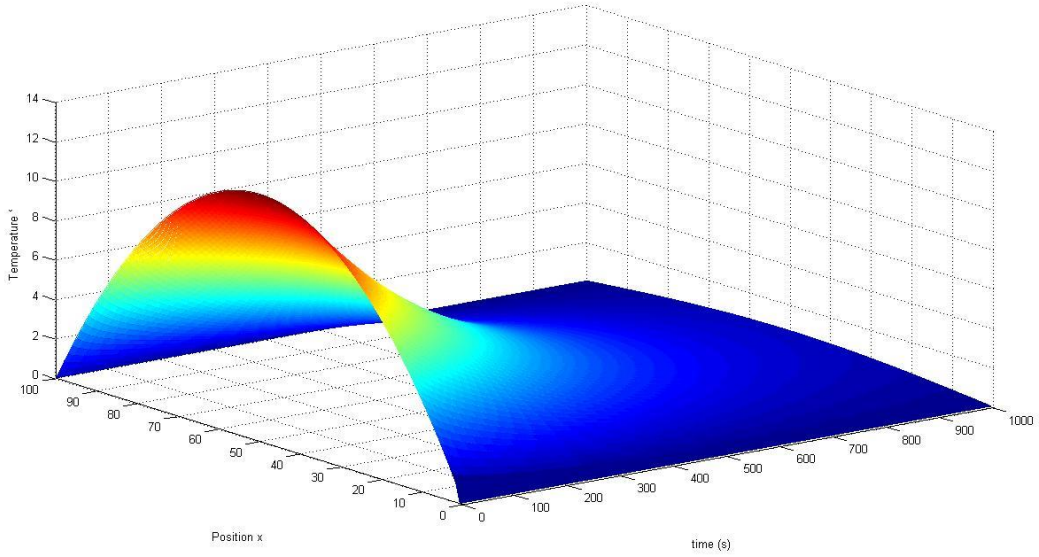

Fig 3.4 Temperature Evolution.

Using the Least Square Method, we estimate the parameters α_1,$\alpha_2$and α_3 of the FD model (3.12) which yields to one step-ahead prediction error at location *i*:

$$\varepsilon_i^k = (\hat{\mathrm{T}}_i^{k+1/k} - \mathrm{T}_i^{k+1})^2 \tag{3.15}$$

where the estimate $\hat{\mathrm{T}}_i^{k+1/k}$ uses different locations.

The cost function to be minimized in order to obtain a better fit is of type:

$$\min_{\theta}(\sum_{i=0}^{m-1}(\sum_{j=0}^{n-1}(\hat{\mathrm{T}}_i^{k+1/k} - \mathrm{T}_i^{k+1})^2)) = 0 \tag{3.16}$$

Where $\theta = [\alpha_1, \alpha_2, \alpha_3]^T$

The vector of parameters θ is the root of:

$$\nabla_{\theta}(\sum_{i=0}^{m-1}(\sum_{j=0}^{n-1}(\hat{\mathrm{T}}_i^{k+1/k} - \mathrm{T}_i^{k+1})^2)) = 0 \tag{3.17}$$

This is equivalent to:

$$\begin{bmatrix} \frac{\partial \sum\sum}{\partial \alpha_1} \\ \frac{\partial \sum\sum}{\partial \alpha_2} \\ \frac{\partial \sum\sum}{\partial \alpha_3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.18}$$

This yieds to the following set of algebraic equations:

$$\begin{bmatrix} \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T_i^{j+1}T_{i+1}^{j} \\ \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T_i^{j+1}T_i^{j} \\ . \\ \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T^{j+1}T_{i-1}^{j} \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} {T_{i+1}^{j}}^2 & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T_i^{j}T_{i+1}^{j} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T_{i-1}^{j}T_{i+1}^{j} \\ \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T_i^{j}T_{i+1}^{j} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} {T_i^{j}}^2 & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T_{i-1}^{j}T_i^{j} \\ \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T_{i-1}^{j}T_{i+1}^{j} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T_{i-1}^{j}T_i^{j} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} {T_{i-1}^{j}}^2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

Considering the data generated using spatial-temporal temperature of the equation (3.14) of 100 times samples and 100 positions in both unit time and unit length , respectively All the estimate parameters using different noisy data are regrouped in the table (3.1). In addition, the mean value of the cost function (3.16) is also given. The noise is normally distributed with zero mean and a given varianceσ^2.

Table 3.1 Identification's results of the linear example.

Noise $\boldsymbol{\sigma^2}$	α1	α2	α3	MSE
0	0.009047	0.979214	0.008742	$2.400530.10^{-5}$
0.0001	0.009060	0.979189	0.008754	$2.402542.10^{-5}$
0.0004	0.009276	0.978762	0.008965	$2.430057.10^{-5}$
0.0025	0.018451	0.960715	0.017847	$3.592499.10^{-5}$
0.01	0.109866	0.781078	0.106161	$1.903710.10^{-4}$

The estimated data is generated using the identified parameters, first by assuming the same Neumann boundary conditions and initial conditions as for the original data, then estimating the boundary conditions using the identified parameters.

The spatial-temporal squared one step-ahead prediction error is shown in figures 3.5 and 3.6 using corrupted and uncorrupted generated data, respectively.

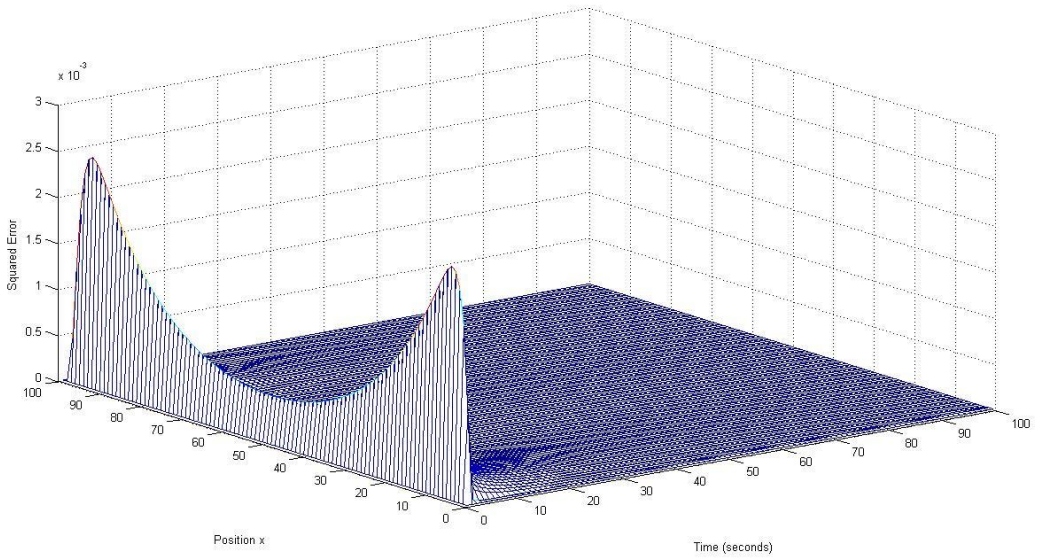

Fig3.5 Uncorrupted One step-ahead prediction error.

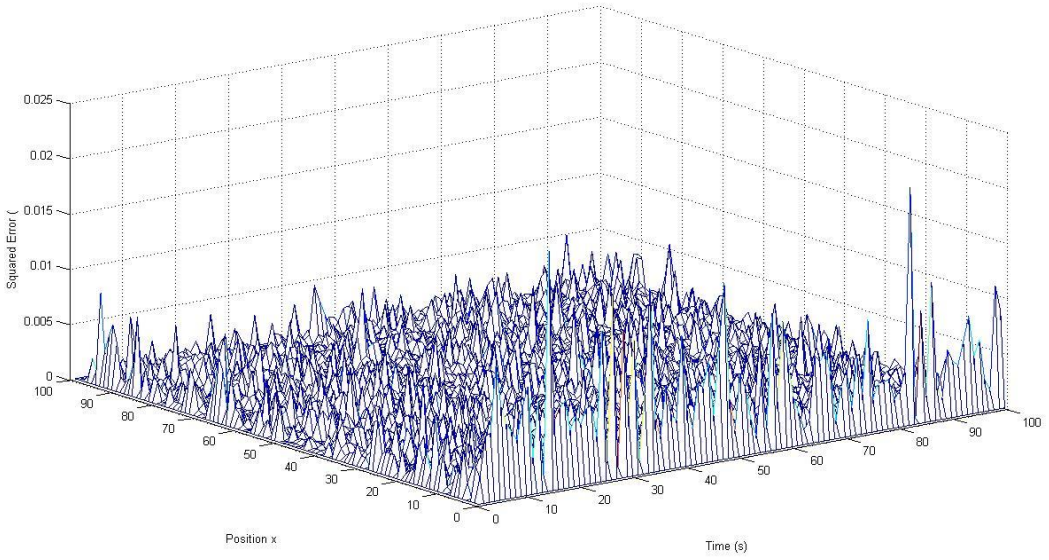

Fig3.6 Uncorrupted One step-ahead prediction error ($\sigma^2 = 0.01$).

Discussion of the results

According to the results obtained, the estimated data is close enough to the exact data; $\boldsymbol{\alpha_1}$ approximately equals to $\boldsymbol{\alpha_3}$ and $\boldsymbol{\alpha_2} \cong 1 - 2\boldsymbol{\alpha_1}$, which satisfies the linear difference equation.

The obtained error (E=$\hat{T}$-T) has a small variance which means that the least square based-estimator performs well even in the case of noise; increasing the standard deviation of noise led to a small increase in the values of parameters and a very small change in the one step-ahead error.

4.2. Non-Linear Heat equation

The second application deals with a nonlinear heat equation. The equation is described as follows:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} - 3T^3 \tag{3.19}$$

with Boundary Conditions: $T(0,t) = \frac{1}{6t+1}$

$$T(1,t) = \frac{1}{2t+1}$$

And initial conditions: $T(x,0) = \frac{1+2x}{x^2+x+1}$

This example is solved using the separation of variable method (used to solve PDE). The analytic solution of the heat equation is used to generate data that is employed in the identification process.

Using the separation of variables, we get the solution:

$$T(x,t) = \frac{1+2x}{x^2+x+6t+1} \tag{3.20}$$

This temperature is illustrated in figure 3.7 on the same way and by using the data generated by the equation 3.14, the estimate parameter values are regrouped in table 3.2

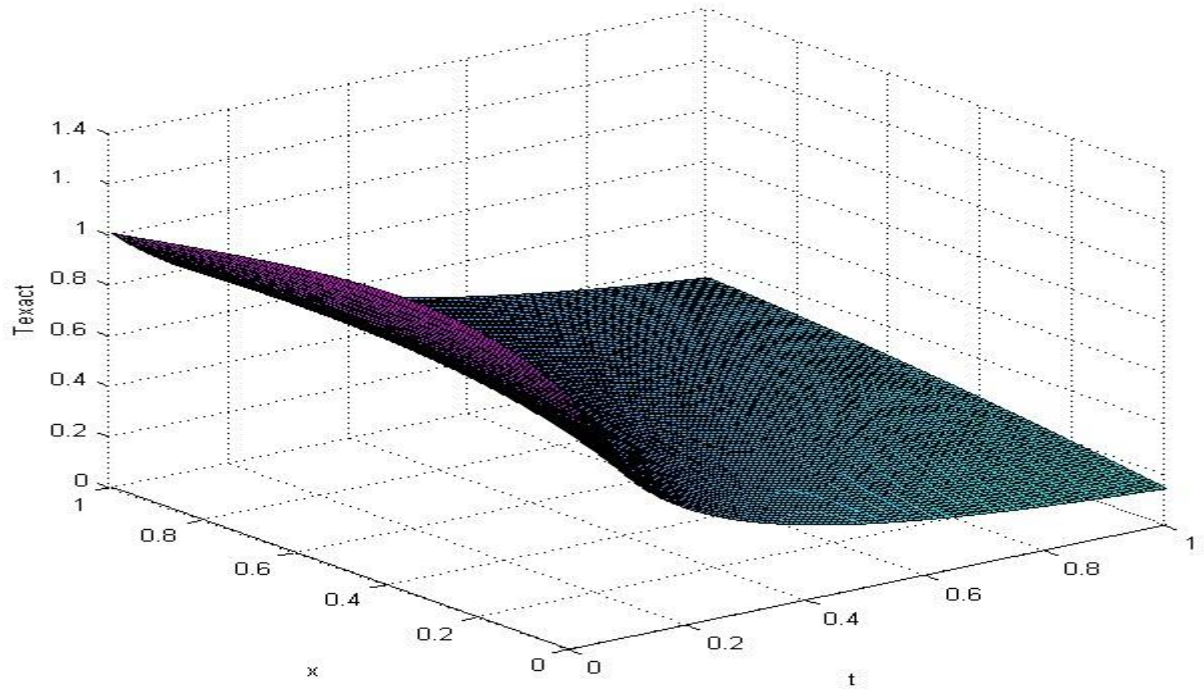

Fig3.7 Temperature Evolution of the nonlinear example.

Using the least square method, we obtained the parameters α_1, α_2, α_3, and the mean square error between the estimated and the original data as it is well shown in the table below:

Table 3.2 Identification's results of the nonlinear example.

Noise σ^2	α1	α2	α3	MSE
0	0.077692	0.323855	0.560467	$1.836192.10^{-4}$
0.0001	0.077594	0.324086	0.560299	$1.836049.10^{-4}$
0.0004	0.078078	0.324334	0.559561	$1.838098.10^{-4}$
0.0025	0.086394	0.335814	0.539715	$1.918635.10^{-4}$
0.01	0.164831	0.365850	0.430015	$3.153481.10^{-4}$

The spatial-temporal squared one step-ahead prediction error is shown in figures 3.8 and 3.9 using corrupted and uncorrupted generated data, respectively.

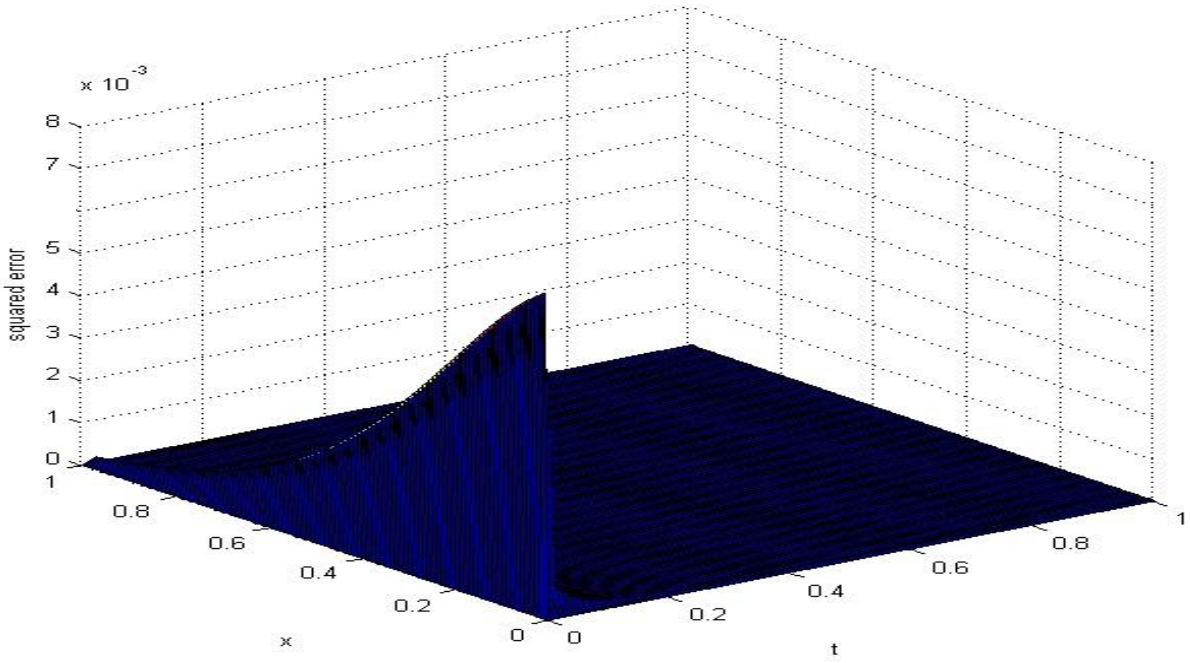

Fig3.8 Uncorrupted One step-ahead prediction error of the nonlinear example.

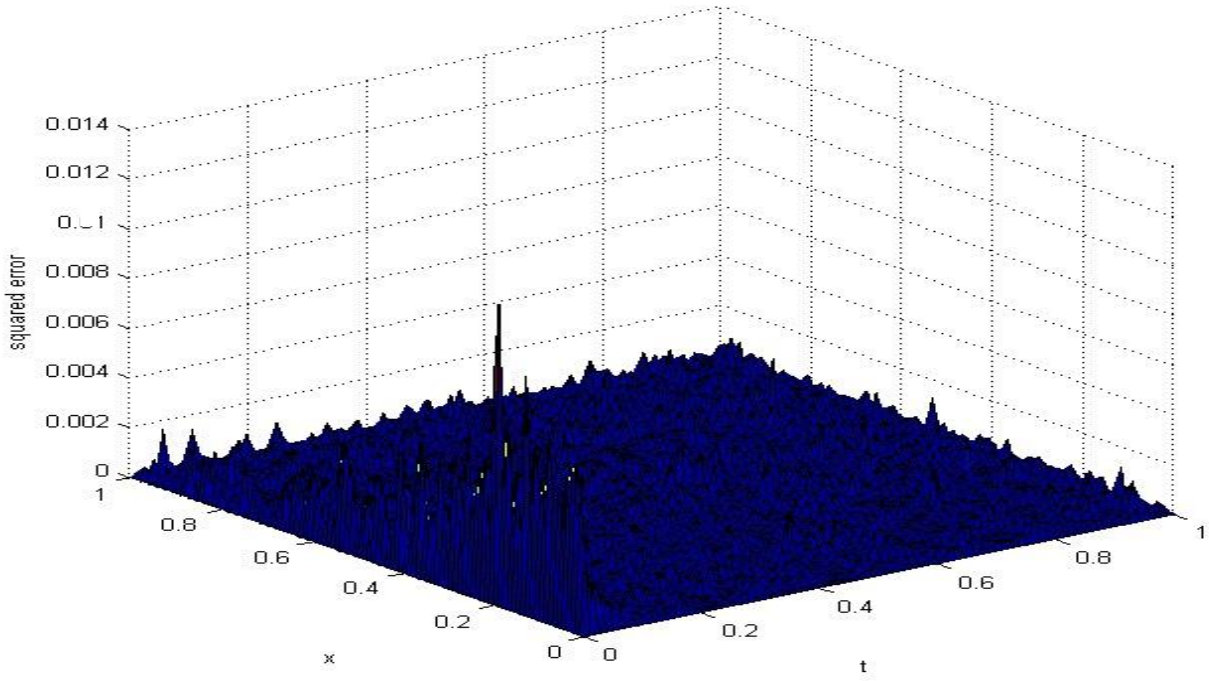

Fig3.9 Corrupted One step-ahead prediction error (σ^2 = 0.01) of the nonlinear example.

Discussion of the results

According to the results it is clear from the error that the estimation using the identified parameters is good since the error is small.

Another case is presented by increasing the estimation size using the squared values of the temperature at three positions so the model is:

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 {T_{i+1}^j}^2 + \alpha_5 {T_i^j}^2 + \alpha_6 {T_{i-1}^j}^2 \quad (3.21)$$

The results are presented in the following table obtained using the least square method:

Table 3.3 Identification's results using nonlinear squared terms.

Noise σ^2	α1	α2	α3	α4	α5	α6	MSE
0	0.002204	0.295866	0.740488	0.028501	-0.141175	-0.162372	3.2535255 .10^{-6}
0.0001	0.246025	0.384823	0.424849	-0.248543	-0.054781	-0.010352	1.532014. 10^{-4}
0.0004	0.376466	0.359778	0.319535	-0.339205	-0.049534	0.066557	5.424368. 10^{-4}
0.0025	0.354820	0.374944	0.315453	-0.178662	-0.155882	0.002804	0.003214
0.01	0.332525	0.339789	0.333005	-0.119730	-0.148761	-0.062380	0.012406

The squared errors for noisy data and non-noisy data are well sketched in the following figures:

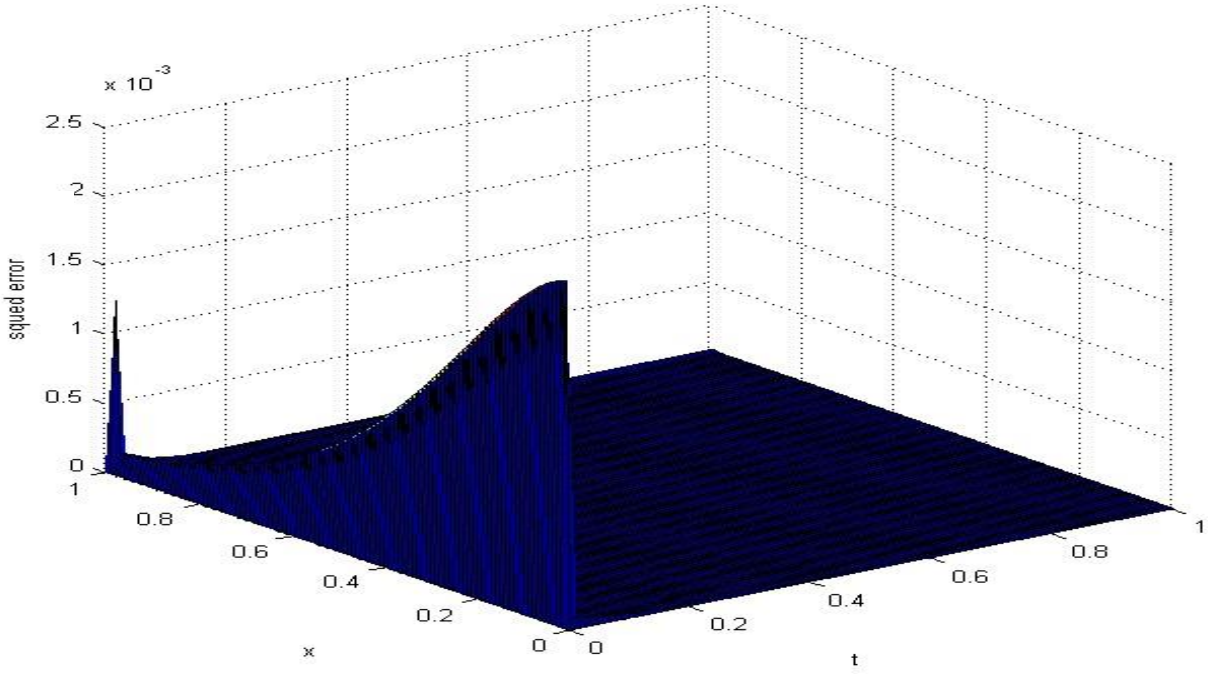

Fig3.10 Uncorrupted One step-ahead prediction error

of the nonlinear example using squared terms.

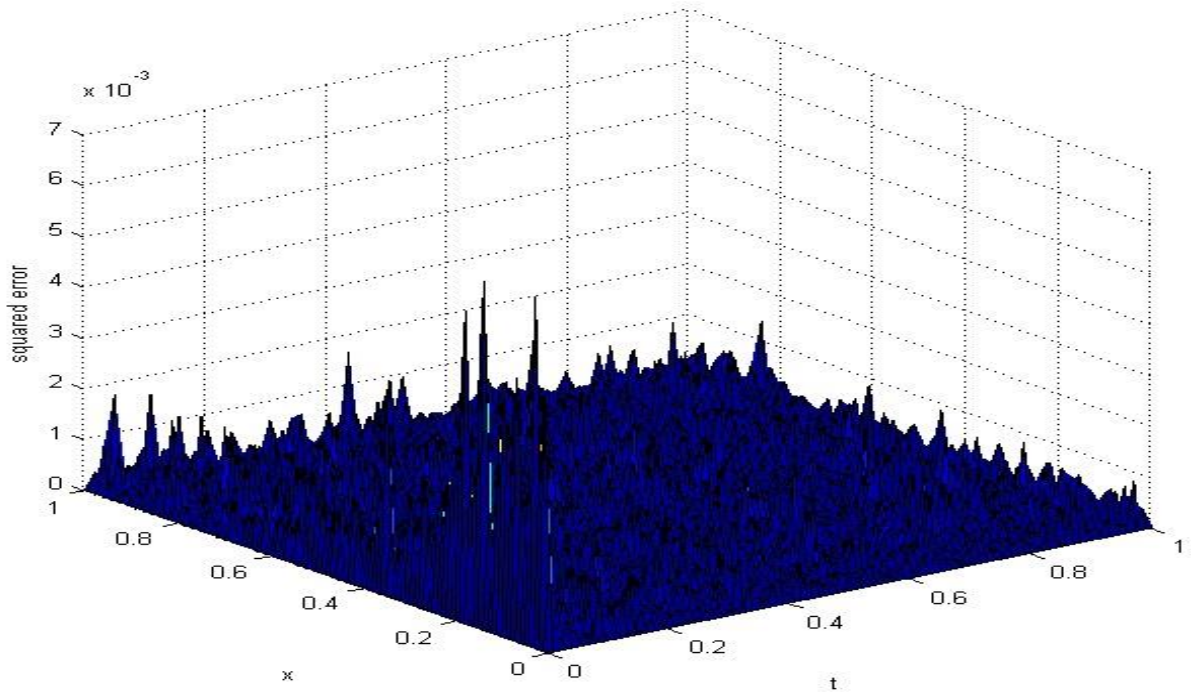

Fig3.11 Corrupted One step-ahead prediction error ($\boldsymbol{\sigma}^2 = 0.01$)

of the nonlinear example using squared terms.

.

For the next case, another set of three cubed temperature elements are added to the model:

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 {T_{i+1}^j}^2 + \alpha_5 {T_i^j}^2 + \alpha_6 {T_{i-1}^j}^2 + \alpha_7 {T_{i+1}^j}^3 + \alpha_8 {T_i^j}^3 + \alpha_9 {T_{i-1}^j}^3 \quad (3.22)$$

The parameters and the mean square error obtained are presented in the following table:

Table 3.4 Identification's results using nonlinear cubed terms.

Noise σ^2	α1	α2	α3	α4	α5
0	-0.011983	1.036778	-0.028711	0.109010	-1.708045
0.0001	-0.060703	1.088736	-0.037042	0.265553	-1.704312
0.0004	-0.061138	1.090273	-0.038166	0.266867	-1.708045
0.0025	0.055541	0.790685	0.145776	0.035042	-1.090097
0.01	0.325670	0.434422	0.232882	-0.325400	-0.276423

Noise σ^2	α6	α7	α8	α9	MSE
0	1.500813	-0.278446	0.924481	-1.004313	$2.538763.10^{-6}$
0.0001	1.498265	-0.277354	0.922114	-1.002890	$2.550344.10^{-6}$
0.0004	1.500813	-0.278446	0.924481	-1.004313	$2.538763.10^{-6}$
0.0025	1.110322	-0.145846	0.580426	-0.789244	$1.156777.10^{-5}$
0.01	0.647844	-0.025461	0.170059	-0.491869	$1.267321.10^{-4}$

The squared error is well sketched in the following figures:

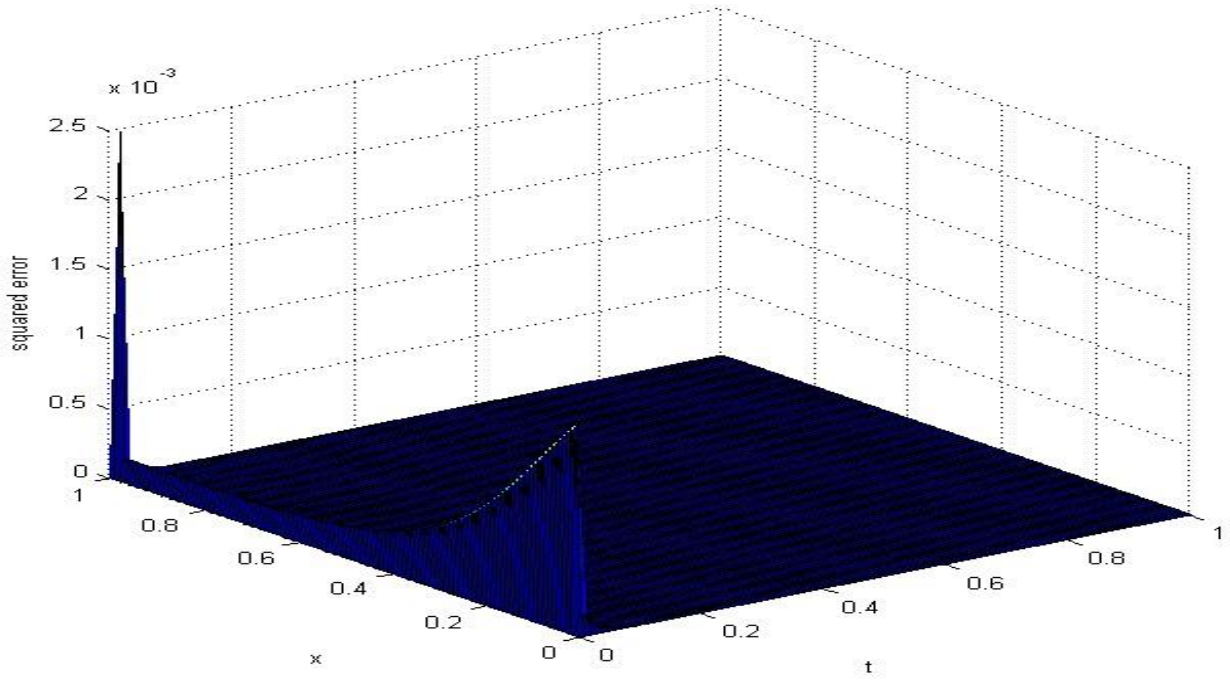

Fig3.12 Uncorrupted One step-ahead prediction error

of the nonlinear example using cubed terms.

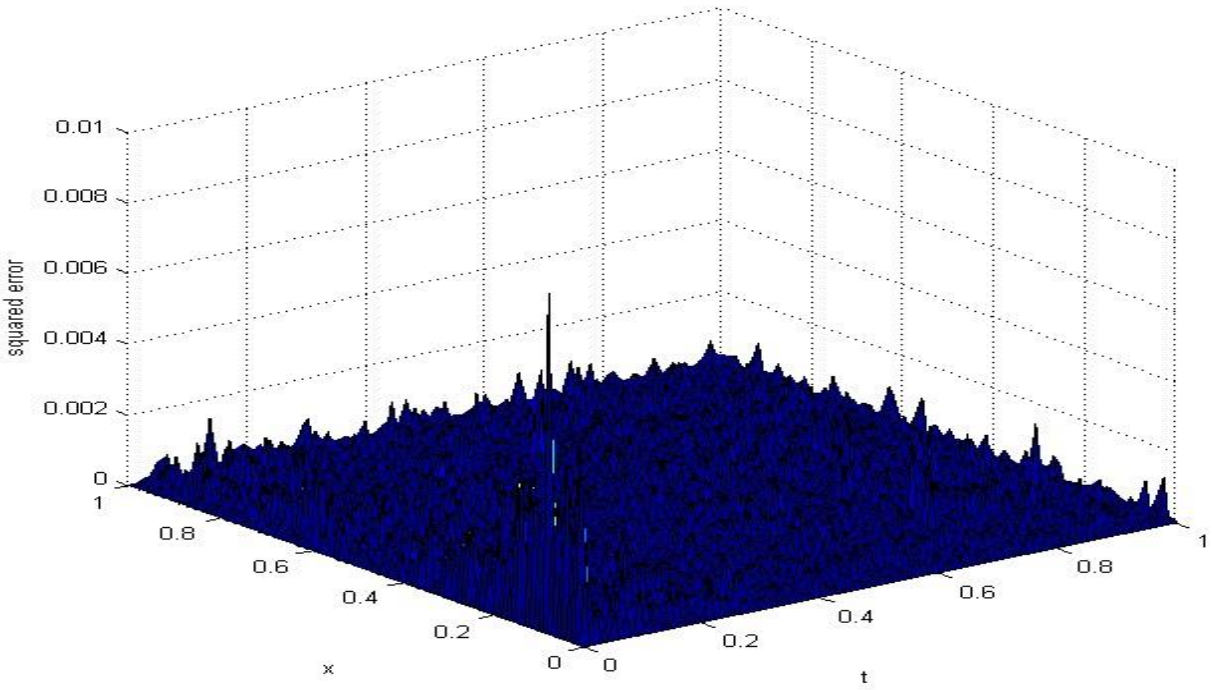

Fig3.13 Corrupted One step-ahead prediction error ($\boldsymbol{\sigma}^2 = 0.01$)

of the nonlinear example using cubed terms.

Discussion of the results

As it is shown in tables and figures, the parameters do not satisfy the linearity of the difference equation since there are significant values for parameters of nonlinear terms; it is expected because the difference equation is nonlinear.

The calculated error ($E=\hat{T}-T$) has a small MSE (a small variance) which means that the estimated parameters matches the system characteristics. It gets smaller when adding the nonlinear terms to the identification algorithm (difference equation). So the one step-ahead prediction error shows that the nonlinear model is better fitted representation than the linear model.

The noise has a significant effect on the values of the parameters. It affects the error as well; by increasing the standard deviation (noise), the error is increased as well.

4.3. The Real Data identification

4.3.1. Description of the Cement Rotary kiln

The rotary kiln is the most important machine in the cement plant; its purpose is to process materials at an extreme heat in order to derive cement. The cement production industry produces over a billion tons of cement, making this type of rotary kiln a very important component within this industry.

The cement plant of Ain Kbira has a rotary kiln that consists of a hollow cylindrical metallic shell, lined using refractory bricks. It is 80 meter long, and 5 meters in height. At one end fuel, in the form of gas, oil, or pulverized solid fuel, is blown in through the "burner pipe", producing a large concentric flame in the lower part of the kiln tube. As material moves under the flame, it reaches its peak temperature, before dropping out of the kiln tube into the cooler. Air is drawn first through the cooler and then through the kiln for combustion of the fuel. In the cooler, the cooling clinker heats the air, so that it may be 400 to 800 °C before it enters the kiln, thus causing intense and rapid combustion of the fuel.

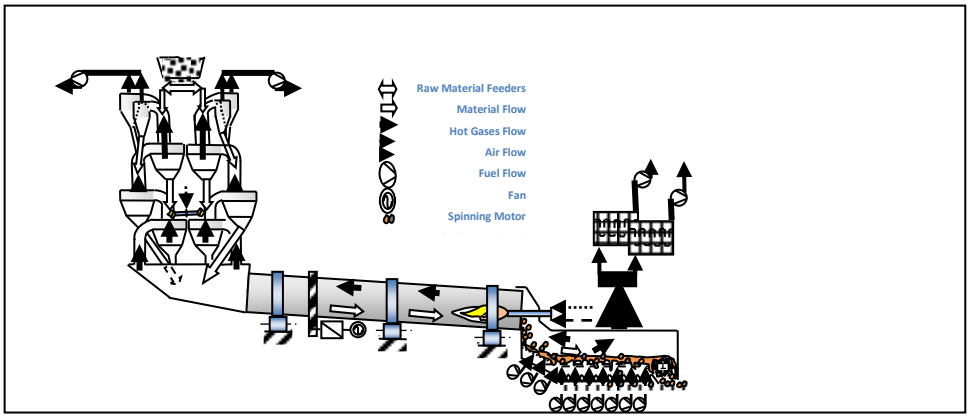

Fig 3.14 Schematic of a Cement Rotary Kiln**.**

The third part of the application is about the identification of the experimental data obtained from the cement rotary kiln. The rotary kiln has a burner placed 1.2 meter from the beginning of the tube. This burner provides energy of 70560 thermie per hour. The rows represent the temperature along the kiln length while the columns are the observation of the temperature:

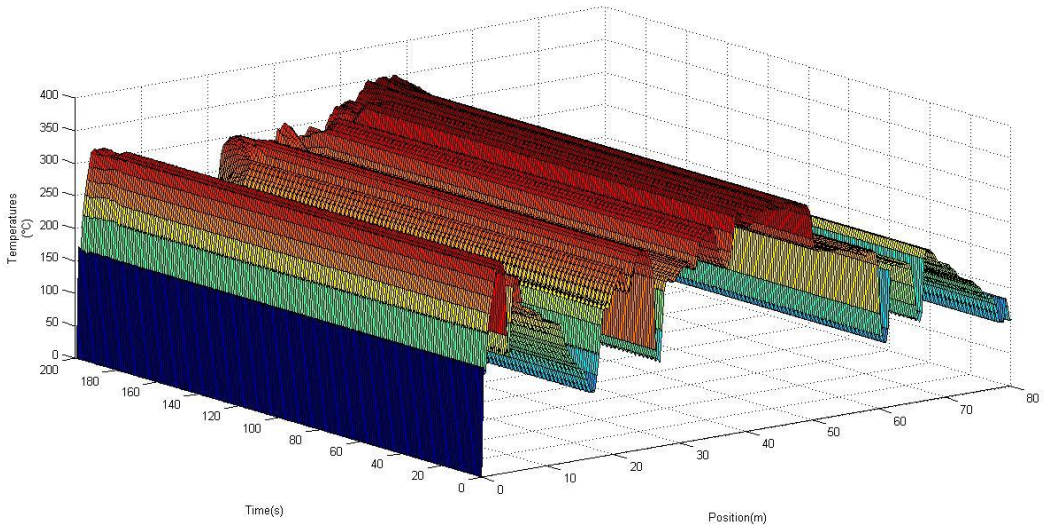

Fig3.15 Real data evolution.

Using the linear difference model with a second term u(x,t):

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 U_i^j \qquad (3.23)$$

The identified parameters and the means square error are presented in the table below:

Table 3.5 Identification's results of the real data.

Parameters	α1	α2	α3	α4	MSE
Values	0.004499	0.991775	0.003690	-0.000004	3.833575

The squared error is well sketched in the following figure:

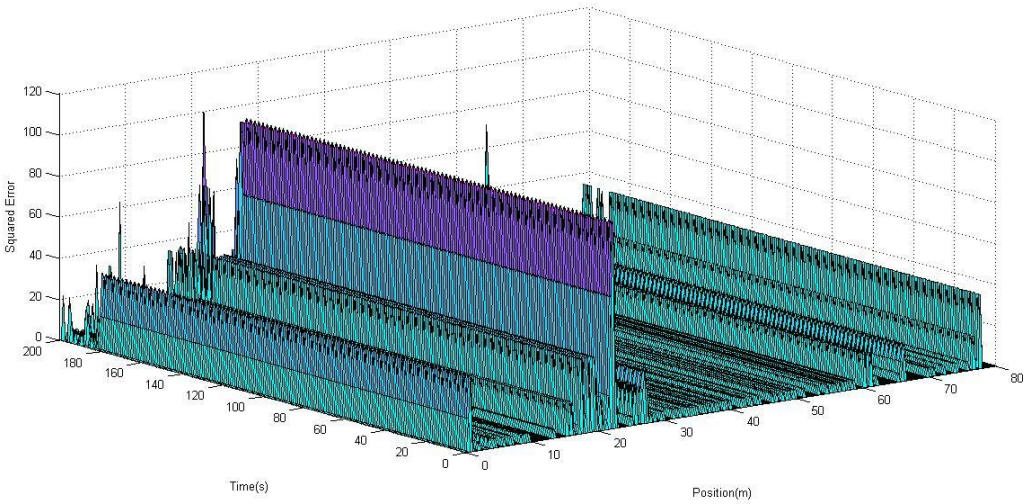

Fig3.16 One step-ahead prediction error.

Using the nonlinear difference model:

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 {T_{i+1}^j}^2 + \alpha_5 {T_i^j}^2 + \alpha_6 {T_{i-1}^j}^2 + \alpha_7 U_{i+1}^j \qquad (3.24)$$

The identified parameters and the means square error are presented in the table below:

Table 3.6 Identification's results of the real data using squared terms.

Parameters	α1	α2	α3	α4	α5	α6	α7	MSE
Values	0.000086	1.003201	-0.003397	0.000013	-0.000034	0.000021	-0.000004	3.830502

The squared error is well sketched in the following figure:

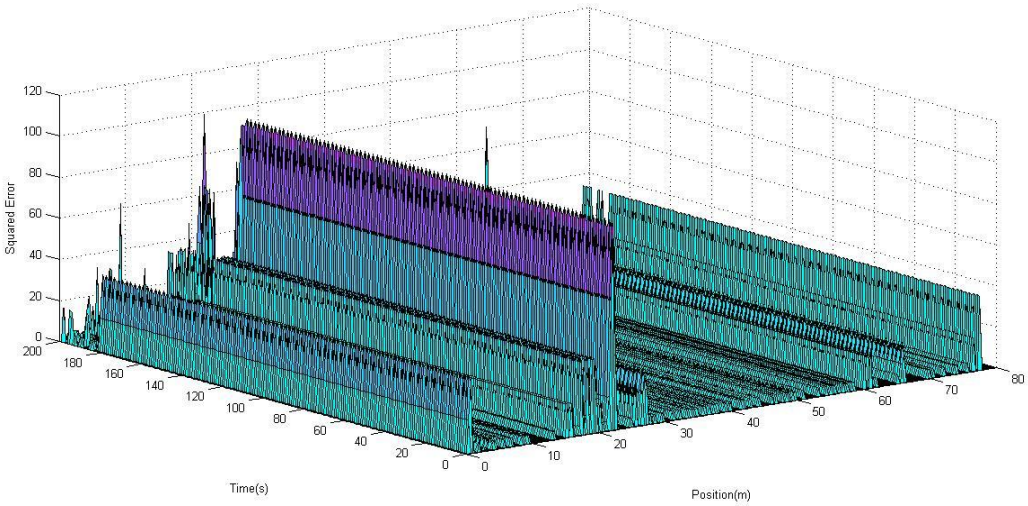

Fig3.17 One step-ahead prediction error using squared terms.

Using the nonlinear difference model

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 {T_{i+1}^j}^2 + \alpha_5 {T_i^j}^2 + \alpha_6 {T_{i-1}^j}^2 + \alpha_4 {T_{i+1}^j}^3 + \alpha_5 {T_i^j}^3 + \alpha_6 {T_{i-1}^j}^3 + \alpha_{10} U_{i+1}^j \quad (3.25)$$

The identified parameters and the means square error are presented in the table below:

Table 3.7 Identification's results of the real data using cubed terms.

Parameters	α1	α2	α3	α4	α5
Values	0.043992	0.953320	0.003237	-0.000212	0.000236

Parameters	α6	α7	α8	α9	α10	MSE
Values	-0.000028	0.0000003	-0.0000004	0.0000001	-0.000002	3.827386

The squared error is well sketched in the following figure:

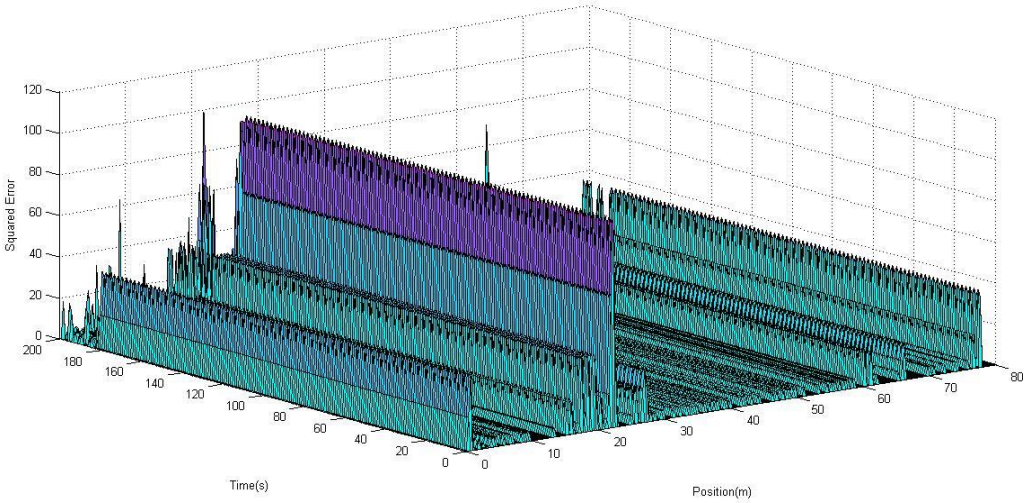

Fig 3.18 One step-ahead prediction error using cubed terms

Discussion of the results

According to the results obtained, the error is significant for linear models, and by adding the nonlinear terms to the models (square and cube), even though we have noticed a very small decrease in the error, and that doesn't prove the linearity of the data.

The identified parameters have small values which show that the linear model is not fitted with the data.

To conclude the experimental data are nonlinear and to model it, a nonlinear identification technique is required.

5. Conclusion

To conclude, this chapter has demonstrated the Least Square approach used for many heat diffusion applications (homogenous and heterogeneous examples) and for the experimental data as well. The least square approach has been useful for three different difference equations. The homogenous application has satisfied the linearity. While the heterogeneous application has been proven to be nonlinear.

The third part of the simulations was dealing with real data obtained from a rotary kiln. Using the least square approach, the results have shown that the data are nonlinear.

The obtained results show the effectiveness, sensitivity and robustness of the proposed approach.

Conclusion

In this thesis, an identification of distributed parameters systems (**DPS)** approach in a cement rotary kiln is presented. The rotary kiln demonstrates the heat transfer phenomena through its tube.

Using the Least Square approach we have built accurate and simplified models of distributed parameter systems, from numerical applications to time-series data obtained from the rotary kiln. The numerical solutions in the examples depend on the associated Neumann Boundary conditions and initial conditions. The measure of performance for the identification was the error (mean square error) between the model and the system. It has shown small values confirming the accuracy of the models.

Further Researches:

We had an immense interest to go further with distributed parameters systems and identify their parameters using other estimators such as variants of least square, Maximum likelihood…etc.

References

[1] D. Chinarro, *System Engineering Applied to Fuenmayor Karst Aquifer (San Julián de Banzo, Huesca) and Collins Glacier (King George Island, Antarctica)*, Springer. 2014

[2] Lennart Ljung, ‘’System Identification Toolbox user’s guide’’, The MathWorks, Inc.1988-1997

[3] Yi Liu ‘’Grey-box Identification of Distributed Parameter Systems’’, Doctoral Thesis. Automatic Control. Department of Signals. Sensors and Systems. Royal Institute of Technology. Stockholm. Sweden 2005

[4] F. van der Heijden, R.P.W. Duin , D Ridder , D.M.J. Tax, ‘’Classification, Parameter Estimation and State Estimation’’ John Wiley & Sons,Ltd 2004

[5] Lennart Ljung, ‘’System Identification Theory for the user ’’ University of Linkoping Sweden.1987

[6] Donoho D, Johnstone I (1994) Ideal spatial adaptation by wavelet shrinkage. Biomefrika 81:425–455

[7] Donoho D (1995) De-Noising by soft-thresholding. IEEE Trans Inf Theory 41(3):613–627

[8] Bendat J (1990) Nonlinear system analysis and identification. Willey, New York

[9] Zhang Q (1997) Using wavelet network in nonparametric estimation. IEEE Trans Neural Network 8(2):227–236

[10] J. LINDFORS, "Advances in control education 2003", University of Lapland, Finland, 2003

[11] Denis Matignon, Joseph Winkin, Denis Dochan "Control of distributed parameter system", University of Namur, Namur, Belguim,2007

[12] P. N. Paraskevopoulos and A. C. Bounas. Distributed parameter system identification via walsh functions. *Int. J. Systems Science,* 9(l):75-83, January 1978.

[13] S. G. Tzafestas. Walsh series approach to lumped and distributed system identification. *J. Franklin Institute,* 305(4):199-220, April 1978.

[14] Y. Sunahara, ‘’Identification of distributed parameter systems’’ Kyoto Institute of Technology, Matsugasaki, Kyoto, Japan.

[15] Stefan Bilbao and Julius O. Smith III ,``Discrete-Time Lumped Models", (From Lecture Overheads) Stanford University.

[16] John, F. (1982), Partial Differential Equations (4th ed.), New York: Springer-Verlag, ISBN 0-387-90609-6.

[17] A. D. Polyanin, Handbook of Linear Partial Differential Equations for Engineers and Scientists, Chapman & Hall/CRC Press, 2002.

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research

University M'Hamed BOUGARA – Boumerdes

Institute of Electrical and Electronic Engineering
Department of Power and Control

Final Year Project Report Presented in Partial Fulfilment of
the Requirements for the Degree of

MASTER

In **Electrical and Electronic Engineering**

Option: **Control**

Title:

Identification of A Distributed Parameter System using Least Square Technique

Presented by:

- **DAHLOUK Abdelkader**
- **AZZOUG El-Mehdi**

Supervisor:

Dr. KOUADRI Abdelmalek

Registration Number: 2015/2016

Dedication

It is our genuine gratefulness and warmest regard that we dedicate this work to the Mighty God, to our beloved parents, the reason of what we have become today and to our colleagues and friends without whom it was difficult for us to complete our thesis work.

Acknowledgements

We are highly appreciative of the effort of our supervisor Dr KOUADRI Abdalmalek for his invaluable assistance and guidance, for his contribution and taking time to read through this report, and for his encouragements and positive criticism of the project.

Sincere acknowledgement is given to all the teachers who taught us and were the essential key of this success.

Our deep appreciation to INELEC students and family for their efforts, patience and concern throughout this project.

Abstract

The identification of such systems represents one of the important directions of theoretical and practical research, due to large class of applications; fluid flow systems, heat diffusion systems, etc. In this context, we consider the linear parametric identification of distributed parameters systems using the Least Square. Two numerical examples of heat transfer systems are presented where linear and nonlinear models are obtained from heat difference equation. The third application is about the identification of a heat conduction in a cement rotary kiln using experimental data. We anticipate that this work be intuitive for practical applications in the areas of controls and signal processing.

Table of Contents

List of Tables

List of Figures

Nomenclature

Symbol:	**Meaning:**
c	Specific heat
K	Heat diffusivity
l	length of tube (or rod)
T	Temperature
Greek Symbols:	
α	System parameter
σ	Standard deviation of the noise
ϵ	Prediction Error
Abbreviations:	
DPS	Distributed Parameter System
MSE	Mean Square Error
PDE	Partial Differential Equations
ODE	Ordinary Difference Equation
LS	Least Square
IC	Initial Condition
BC	Boundary Condition

General Introduction

Control systems are ubiquitous in the modern world where the instruments of our scientific and industrial society are applied to an increasingly wide range of processes. Such control intervention is undertaken with many different objectives in mind; e.g. steering the process to a desired state, minimizing the effects of various disturbances tending to move the system in undesirable directions, stabilizing systems which are inherently unstable or improving the stability properties of system with weak stability characteristics; etc. While it is rarely possible, in a mathematical model, to account for all the factors affecting the performances of a real world system mathematical modeling of the system is, nevertheless, ordinarily essential for efficient and effective design and implementation of control procedures.

The first concern with determining the dynamical structure for a practical system is the problem of system parameter identification. Up to the present time, the principal line of research activities is directed to treat lumped parameter systems described by ordinary differential equations. A lumped parameter representation of a physical system implies that its local spatial variations are disregarded. However dynamic behaviors of almost all real physical systems are, in fact, distributed. A distributed parameter system DPS is characterized by a class of partial differential equations, integral equations or integro-differential equations. In this thesis, we restrict our attention to physical systems described by partial differential equations, because the great majority of research including efforts of practical applications is concerned only with this kind of mathematical models. Although a class of partial differential equations may usually be derived by considering basic conservation principles, system parameters remain to be determined in many examples. From practical viewpoints, determination of heat transfer coefficients, specific heats, electro-magnetic properties, gas properties, chemical reaction rate constants, diffusion constants, etc. is a requisite effort throughout state estimates and/or optimal controls of distributed parameter systems.

The objective of this report is to identify heat conduction systems as an application for distributed parameter systems. The parameters of the system model are obtained using Least

Square Method, a parametric approach based on the minimization of the sum of squared residuals; the difference between the observed value and fitted value by a model.[wiki]

To achieve this objective, the report is divided into three sections (chapters). The first chapter is about the techniques used in system identification. These techniques are divided according to the system model; parametric and non-parametric approaches for linear and nonlinear models in time and frequency domains. The second chapter is about the distributed parameter system, its features and characteristics. The third is about application of the chosen approach (Least Square Method) in two numerical examples and to experimental data obtained from a cement rotary kiln.

Chapter One

System Identification Techniques

1. Introduction:

This chapter gives a brief introduction to system identification. System identification is an immense topic, and the contents of this chapter are neither new nor complete. The purpose is merely to introduce concepts, ideas and algorithms that are useful in system modeling.

Modeling is the abstraction of a real process to characterize its behavior. It is concerned with the determination of particular models for systems that are intended for a certain purpose such as control. The idea is to build accurate, simplified models of complex systems from physical insights (input/output) measurements. Whether it is parametric identification, we determine the parameters of some chosen models, or non-parametric identification, it gives the general behavior of the system and sometimes it is essential for parametric identification **[1].** The best model is the simplest model that accurately describes the dynamics of system.

2. The Procedures of System Identification:

The System Identification allows us to estimate a model of a system based on observed input-output data. Several ways to describe a system and to estimate such descriptions exist. This section gives a brief description of the most important approaches**.**

The procedure to determine a model of a dynamical system from observed input-output data involves three basic ingredients:

- Input-output data
- A set of candidate models (A model structure)
- A criterion to select a particular model in the set.

The identification process amounts to repeatedly selecting a model structure, computing the best model in the structure, and evaluating this structured model's properties to see if they are satisfactory. The cycle can be as follows:

1- We design an experiment and collect input-output data from the process to be identified.
2- We examine the data. Polish it to remove trends and outliers, select useful portions of the original data, and apply filtering to enhance important frequency ranges.
3- We select and define a model structure
4- We compute the best model in the model structure according to the input-output data and a given criterion of fit.
5- We examine the obtained model's properties

6- To validate the model, we take a close look at the model's output and compare it to the measured one on a data set that was not used for the fit. If it is good enough, we select the model. Otherwise, we go back and select another model set. Possibly, we also try other estimation methods or we work further on the input-output data. **[2].** Fig 1.1 illustrates different procedures of system identification.

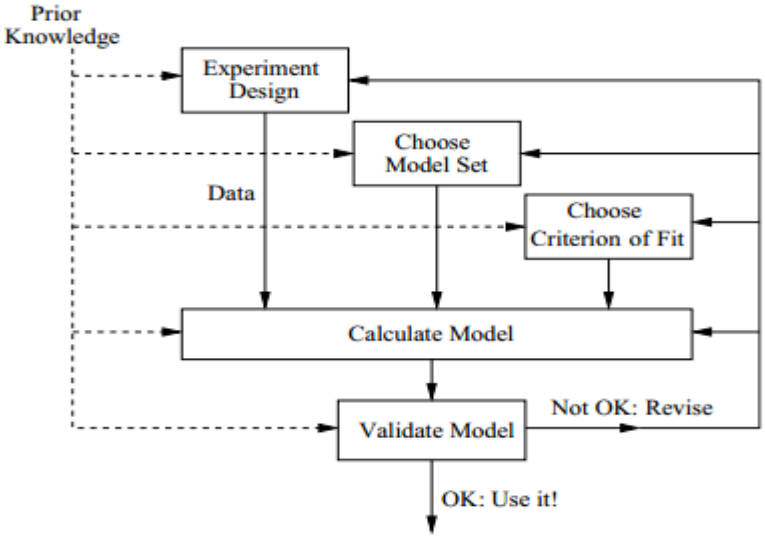

Fig 1.1 System identification loop

MATLAB allows us to build mathematical models of a dynamic system using the system identification toolbox. This toolbox features a flexible graphical user interface that aids in building accurate, simplified models of complex systems from observed data.

This Toolbox contains also all the common techniques to adjust parameters in all kinds of linear models **[2].** It also allows you to examine the models' properties, and to check if they are any good, as well as to preprocess and polish the measured data.

Estimating models for control systems is based on measured data. The models will describe the behavior of the observed data, which makes them related directly to the measurements. The data is treated as a time series that is why there are two types of estimation algorithms; online and offline estimation. Online estimation algorithms estimate the parameters of a model when new data is available during the operation of the model. In offline estimation, we first collect all the input/output data then we estimate the model parameters. Parameter values estimated using online estimation can vary with time, but parameters estimated using offline estimation do not [4].

3. Parametric Identification:

It is also called conventional time domain identification; it deals with models containing a vector of parameters, which are to be adjusted so that the model mimics the system behavior as close as possible. Parametric identification relies on a model previously defined by a set of parameters that must be calculated to accomplish a given quality criterion. The system characteristics can have a parametric representation through a polynomial of a finite and known degree **[1]**. The model structure can be obtained by physical modeling (grey box) or it may be a standard one (black box). In the latter case, a set of generic standard structures must be taken into consideration such as AutoRegressive Moving Average eXogeneous(ARMAX) and its variants, Finite Impulse Response(FIR), Box-Jenkins (BJ),and Output Error (OE) models.(see Fig 1.2)

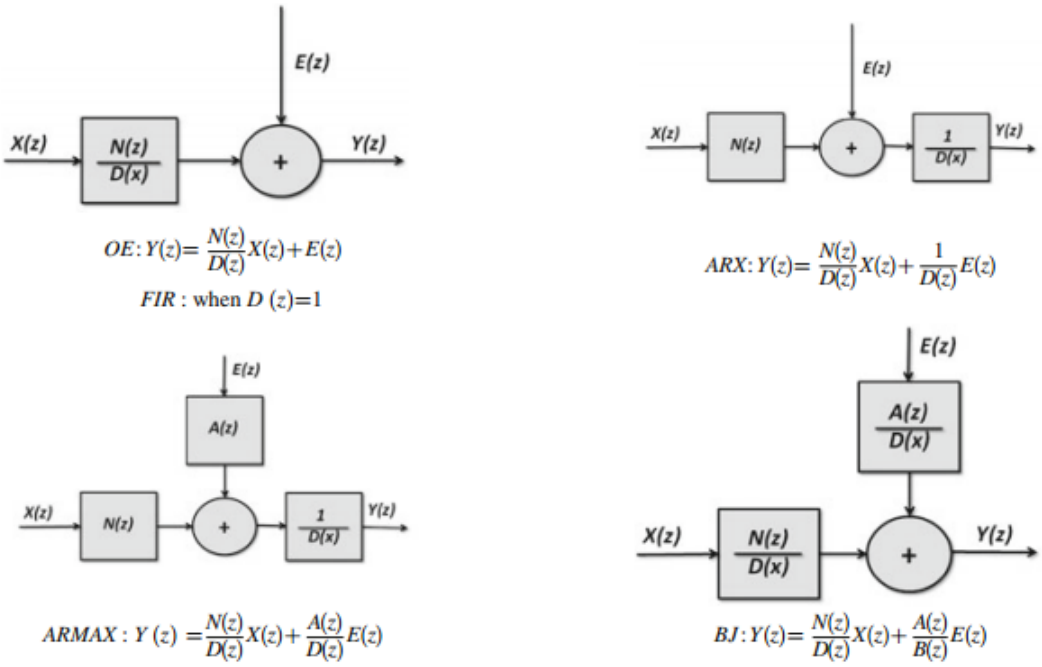

Fig. 1.2 System Model Structures

3.1. Linear Identification:

Parametric identification techniques depend mostly on Prediction-Error Methods (PEM). The output of system y(t) can be expressed based on the z-transform as :

$$Y(z) = G(z)X(z) + W(z) \quad (1.1)$$

The expression (1.1) can be rewritten as follows:

$$Y(z) = G(z)X(z) + H(z)E(z) = (N(z))/(D(z)\ )\ X(z) + (A(z))/(B(z)\ )\ E(z) \quad (1.2)$$

Where E(z) is the transform of a white noise, $\epsilon(t)$. G(z) is the transfer function of the system, H(z) is the stochastic model of noise, and Y(z), X(z) are the z-transform of the output input, respectively.

The relationship between both functions defines several model structures. Fig.1.2 shows the most common ones: AutoRegressiveeXogeneous(ARX) model, AutoRegressive Moving Average eXogeneous(ARMAX) model, Box-Jenkins (BJ) model and Output Error (OE) models.

The ARX model uses the past inputs and past outputs as regressors. This results in linear least square description where the cost function needs to be minimized.

$$\min_{\theta} E\left\{\sum_{k=1}^{N-\delta}(\hat{y}_{k+\frac{2}{k}}(\theta) - y_{k+2})^2\right\} \quad (1.3)$$

Where $(\hat{y}_{k+\frac{2}{k}}(\theta)$ is the estimate δ-step ahead predicted output of the system, θ is vector of unknown parameters, y_{k+2} represents the measured output, and E is the expected value of the squared z-step ahead prediction error.

The foremost disadvantage is that the disturbance model 1/N(z) comes along with the system's poles. It is, therefore, easy to get an incorrect estimate of the system dynamics because the A(z) polynomial can also include the disturbance properties. So, higher orders in A(z)and B(z) coefficients may be required. If the signal-to-noise ratio is good, this disadvantage is less important **[1].**

The ARMAX model has more flexibility in the handling of disturbance modeling than the ARX model. For this reason, ARMAX is a widespread used model and performs well in many engineering applications.

The FIR is the simplest model structure to be considered. The past inputs are used as regressors. The structure results again in a linear least square problem for minimizing the cost function (1.3). It requires many regressors and the convergence rate is slow **[1]**.

The OE model has the advantage that the system dynamics can be described separately and that no parameters are wasted on a disturbance model. If the system operates without feedback during the data collecting, a correct description of the transfer function G(z) = N(z)/D(z) can be obtained regardless of the nature of the disturbance **[1]**.

In the BJ model, the disturbances properties are modeled separately from the system dynamics.

Model validation is carried out by comparing the model behavior with the system's one and evaluating the difference. All models have a certain domain of validity. This may determine how exactly they are able to describe the system behavior. There are a number of different methods to set a criterion e.g., least squares **[3]**, generalized least squares **[3]**, maximum likelihood **[3]** or instrumental variables **[3]**. Some of them will be discussed in details in the following section.

3.1.1. Parameter estimation:

Assuming that a certain model structure $\mathcal{M}$ has been chosen. $\mathcal{M}(\theta)$ denotes a particular model in the model set parameterized using the parameter setθ. The aim is to find the "best" model within the model set $\mathcal{M}^*$

$$\mathcal{M}^* = \{ \mathcal{M}(\theta) | \theta \in \mathcal{D}_{\mathcal{M}} \} \quad (1.3)$$

Where $\mathcal{D}_{\mathcal{M}}$is a closed subset of $\mathcal{R}^d$, andd is the dimension of the parameter vector θ. We assume an experiment that is carried out on the process and measurement data are collected with equally spaced time intervals. Let the time index $k \in \{0, 1, 2,..., N_s\}$.

The following notations are introduced:

$$u^k = [u_0, u_1, u_2, \dots, u_k] \quad (1.4)$$

$$y^k = [y_0, y_1, y_2, \dots, y_k] \quad (1.5)$$

i.e., u^k represents measured input up to timet_k, and y^k represents measured output up to timet_k. Denote the joint set of the input and output data as:

$$z^k = (u^k, y^k) \quad (1.6)$$

The parameter estimation problem is to use the data set z^N to select a suitable value $\hat{\theta}_N$, and hence a model $\mathcal{M}(\hat{\theta}_N)$ from the model set $\mathcal{M}^*$. Note that this is named off-line estimation **[4]**, as all the measured inputs and outputs at all sampling times are used for parameter estimation. In some cases, it may be of interest to estimate parameters at some certain time t_k, using all measurement data up to this time point (z^k), this is termed on-line estimation **[4]**, or recursive identification **[5].**

3.1.2. Prediction error methods:

The performance of a model is judged by its ability to predict the outputs of the system. According to Ljung (1999), most of the methods used for parametric estimation can be characterized as general prediction error methods (PEM). The prediction error is defined as the difference between the predicted output and the measured output

$$\epsilon(t_k, \theta) = y^k - \widehat{y^k} \tag{1.7}$$

Where y^k and $\widehat{y^k}$ are the measured and estimated outputs at time $t_k$respectively. By assumption, the estimated output depends on the parameter setθ, i.e., $\widehat{y^k} = \hat{y}(t_k|\theta)$.

The criterion may be further expanded by filtering the prediction error, through a stable linear filter L(q):

$$\epsilon_f(t_k|\theta) = L(q)\epsilon(t_k|\theta) \tag{1.8}$$

Where L acts like a frequency weighting of the criterion. By doing this, the system properties in frequencies of specific interest can be emphasized.

The principle of fitting parameterized models to data is based on choosing some norm of the prediction error, which is often termed the loss function, and then find the parameter vector θ that minimizes this loss function.

$$\hat{\theta}_{N_s} = \arg\min V_{N_s}(\theta, z^{N_s}) \tag{1.9}$$

A general form of loss function can be represented as:

$$V_{N_s}(\theta, z^{N_s}) = \frac{1}{N_s}\sum_{k=1}^{N_s} l(\epsilon_f(t_k|\theta)) \tag{1.10}$$

Where l denotes a scalar valued, positive function which is used to measure the norm of the filtered prediction error. Different ways of forming the function l and choosing the prefilter L(q) lead to different estimation methods, e.g., the least squares (LS) method and the maximum likelihood (ML) method. To simplify the notation, we omit N_s in the following since we are here mainly concerned with off-line estimation and the whole sequence of measurement data are known.

3.1.3. Least squares (LS) method:

In the least squares method, the loss function is defined as a quadratic residual function, dependent on the parameter vectorθ.

$$V(\theta) = \frac{1}{N_s}\sum_{k=1}^{N_s} \epsilon^2(t_k|\theta)) \quad (1.11)$$

Thus, the quadratic form of the residual between the estimated outputs and the measured outputs is minimized

$$\hat{\theta} = \arg\min \frac{1}{N_s}\sum_{k=1}^{N_s} \epsilon^2(t_k|\theta)) \quad (1.12)$$

Which is known as a least square estimator.

When each residual is multiplied with a certain weight factor, the estimator is called weighted least squares (WLS) estimator.

$$\min_{\theta} V(\theta) = \frac{1}{N_s}\sum_{k=1}^{N_s} \alpha_k \epsilon^2(t_k|\theta)) \quad (1.13)$$

Where $\alpha_k$is the weight factor.

3.1.4. Maximum likelihood (ML) method:

The way of forming the loss function in the (LS) method is completely deterministic. In contrast, to a statistical sense, the measurement data are stochastic realizations, since the system in general can be seen to be corrupted by stochastic disturbances and/or measurement noise. A probabilistic approach to constructing the loss function is the well-known maximum likelihood method. The objective of maximum likelihood estimation is to find the parameter set that maximizes the probability that the observed data are explained by the model. More explicitly, a likelihood function is defined as the joint probability density of all the observation data assuming that the parameter set is known

$$L(\theta) = \prod_{k=1}^{N_s} p(y^k|y^{k-1},\theta) \quad (1.14)$$

Where $p(y^k|y^{k-1},\theta)$is the probability density function (PDF) of y^k depending on previous measurement data set $y^{k-1}$and parameter setθ. Assuming that the prediction error sequence $\{\epsilon_k\}$ consist of zero-mean, independent stochastic variables with the probability density function expression $p(\epsilon_k(\theta)|\theta)$.

The eq (1.14) can then be rewritten as:

$$L(\theta) = \prod_{k=1}^{N_s} p(\epsilon_k(\theta)|\theta) \quad (1.15)$$

The maximum likelihood estimator determines the parameter set by maximizing the likelihood function (1.15), which is equivalent to minimizing

$$V(\theta) = -\log L(\theta) = -\sum_{k=1}^{N_s} \log\left(p(\epsilon_k(\theta)|\theta)\right) \quad (1.16)$$

The maximum likelihood method is often considered to be the optimal method for parameter estimation, in the sense that it asymptotically approaches the best achievable results, namely, lowest estimate variances, under the assumption that the true system is within the model set. Furthermore, its asymptotic properties make it very useful for model validation by using different kinds of likelihood based statistical tests.

Note that the ML-estimator is equivalent to the WLS-estimator under some specific condition, i.e., when the measurement errors are normally distributed with known covariance matrix and the elements of the inverse covariance matrix are used as weights.

4. Non-Parametric Identification:

Nonparametric identification techniques provide a very effective and simple way of finding model structure in data sets without the imposition of a parametric one **[1].** Its methods aim at determining the system functions without first selecting set of confined possible models. Such methods are often called nonparametric since they do not employ a finite-dimensional parameter vector in the search of the best description **[1].** Commonly, the initial process to carry out is the nonparametric identification, and then, if it were suitable, the parametric identification should be performed. The next sections review the non-parametric identification methods from time domain and frequency domain perspectives.

It is difficult to establish a clear identification methodology of nonlinear systems, since analysis is usually more intricate than in the identification of linear models, because of the variety of nonlinear model structures and nonlinear behaviors. For instance Donoho and Johnstone **[6]** and Donoho **[7]** introduced nonlinear wavelet estimators in nonparametric regression through thresholding, i.e., the term-by-term assessment of coefficients in the

wavelet expansion. Only coefficients that exceed a predetermined threshold are taken into account. This produces the wavelet shrinkage. Bendat describes procedures to identify and analyze the properties of many types of nonlinear systems as Zero-Memory Nonlinear Systems and Parallel Nonlinear System, with analysis of Nonlinear System Input/Output Relationships **[8]**.

Zhang applied wavelet theory for nonlinear system identification with a wavelet basis as a universal function approximator, with a neural network used to determine the resolution, and the translation coefficients of the wavelet **[9].** This nonparametric estimator named wavelet neural network has a neural network like structure that makes use of techniques of regressor selection completed with back propagation procedure **[9].**

4.1. Non-parametric Identification in Time Domain:

4.1.1. The Impulse Response:

The notion of characterizing a dynamical system by its impulse (or pulse) response dates from the earliest forays of process engineers into system identification.

Let us assume the following system

$$y(t) = G_0(q).u(t) + v(t) \tag{1.20}$$

or equivalently

$$y(t) = \sum_{k=0}^{\infty} g_0(k).u(t-k) + v(t) \tag{1.21}$$

We subject our system to a pulse input

$$u(t) = \begin{cases} \alpha, & t = 0 \\ 0, & t \neq 0 \end{cases} \tag{1.22}$$

Then the output will be

$$y(t) = \alpha g_0(t) + v(t) \tag{1.23}$$

If the noise is low, it is thus possible to determine the impulse-response coefficient $g_0(t)$ from an experiment with a pulse input. The estimate will be

$$\hat{g}(t) = \frac{y(t)}{\alpha} \tag{1.24}$$

and the error is v(t)/ α. Its basic weakness is that many physical processes do not allow pulse inputs of such amplitude that the error is insignificant compared to the impulse-response coefficients. The possible damage inflicted by direct use of an impulsive test signal on control system hardware and the presence of output noise has led to a decline in its use **[4].**

4.1.2. Cross-Correlation Approach:

Cross-covariance is a non-parametric identification technique and is related with the impulse response. Correlation means how two variables are related together.

The correlation between the input and output with the assumption that the mean is zero for y and u is the eq. (1.21):

$$y(t) = \sum_{k=0}^{\infty} g_0(k).u(t-k) + v(t)$$

v is the noise in the system.

If the input and output are uncorrelated, the cross covariance between them is:

$$R_{uy}(\tau) = g(\tau) * R_{uu}(\tau) \qquad (1.25)$$

That is, the cross correlation is the convolution between the impulse response and the autocorrelation of the input. Thus, the impulse response can be estimated from the covariance (correlation if both signals have zero mean) if the input is a white noise **[4].**.

If the input is chosen as white noise so that:$R_{uu}(\tau) = \alpha\delta_{T_0}(\tau)$

then $$g_0(\tau) = \frac{R_{uy}(\tau)}{\alpha} \qquad (1.26)$$

Where g_0 is an estimate for the impulse response is obtained from an estimate of R_{uy}.

4.1.3. Step-Response Analysis:

The response of the system can be determined by applying a step input:

$$u(t) = \begin{cases} \alpha, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

as follows

$$y(t) = \alpha \sum_{k=1}^{t} g_0(k) + v(t) \qquad (1.27)$$

For this the estimate of $g_0(k)$ could be obtained as:

$$g_0(\tau) = \frac{y(t)-y(t-1)}{\alpha} \quad (1.28)$$

This method is useful for obtaining qualitative information about the system; more specifically, it shows the transient behavior of the system through Dead Time, Static Gain and time constant **[4].**.

4.2. Non-Parametric Identification in Frequency Domain:

The frequency domain characterization of system dynamics has, like its time domain counterpart, a long and varied history **[10].** As far as control engineering is concerned, however, frequency domain identification gained deep relevance with the development of stability and design methods based upon frequency response measurements **[10].**

4.2.1. Transfer Function:

The fundamental physical interpretation of the transfer function G(z) is that the complex number $G(e^{jw})$ bears information about what happens to an input sinusoid.

$$u(t) = \alpha \ cos\, wt \quad t=0, 1, 2\ldots \quad (1.29)$$

then

$$y(t) = \alpha \left|G_0(e^{jw})\right| cos(wt + \varphi) + v(t) \quad (1.30)$$

where $$\varphi = arg\, G_0(e^{jw}) \quad (1.31)$$

This is known as frequency analysis and it is a simple method for obtaining detailed information about the system.

4.2.2. Frequency Response using Correlation:

With the noise component v(t), it may be cumbersome to determine $\left|G_0(e^{jw})\right|$ and φ accurately by graphic methods. Since the interesting component of y(t) is a sine function of known frequency, it is possible to correlate it out from the noise in the following way. From the sums

$$I_c(N) = \frac{1}{N}\sum_{t=1}^{N} y(t)\, cos\, wt \quad (1.32)$$

$$I_s(N) = \frac{1}{N}\sum_{t=1}^{N} y(t)\, sin\, wt \quad (1.33)$$

We substitute y(t) by its expression (1.32) in (1.33), we obtain

$$I_c(N) = \frac{1}{N}\sum_{t=1}^{N} \alpha\left|G_0\left(e^{jw}\right)\right| cos(wt + \varphi)\, cos\, wt + \frac{1}{N}\sum_{t=1}^{N} v(t)\, cos\, wt$$

$$= \frac{\alpha}{2N}\left|G_0\left(e^{jw}\right)\right| \sum_{t=1}^{N}[cos(2wt+\varphi)+cos\,\varphi] + \frac{1}{N}\sum_{t=1}^{N} v(t)\,cos\,wt$$

$$= \frac{\alpha}{2}\left|G_0\left(e^{jw}\right)\right| cos\,\varphi + \frac{\alpha}{2N}\left|G_0\left(e^{jw}\right)\right| \sum_{t=1}^{N}[cos(2wt+\varphi)] + \frac{1}{N}\sum_{t=1}^{N} v(t)\,cos\,wt \quad (1.33)$$

The second term tend to zero as N tends to infinity.

Similarly, for (1.32)

$$I_s(N) = -\frac{\alpha}{2}\left|G_0\left(e^{jw}\right)\right| \sin\varphi + \alpha\left|G_0\left(e^{jw}\right)\right| \frac{1}{2}\frac{1}{N}\sum_{t=1}^{N}[\sin(2wt+\varphi)] + \frac{1}{N}\sum_{t=1}^{N} v(t)\sin wt \quad (1.34)$$

These two expressions (1.33) and (1.34)suggest the following estimates of $\left|G_0\left(e^{jw}\right)\right|$ and φ, respectively

$$\left|\hat{G}_N\left(e^{jw}\right)\right| = \frac{2\sqrt{I_c^2(N)+I_s^2(N)}}{\alpha} \quad (1.35)$$

$$\hat{\varphi} = \arg\hat{G}_N\left(e^{jw}\right) = -\arctan\frac{I_s(N)}{I_c(N)} + k\pi \qquad k=0.1.2...N \quad (1.36)$$

Repeating this procedure for a number of frequencies, a good picture of $G_0(e^{jw})$ over the frequency domain of interest can be obtained.

This method allows us to obtain easily the Bode plot of the system, but many industrial processes do not admit sinusoidal inputs in normal operations.

5. Conclusion:

In this first chapter of the thesis, we have discussed different techniques of system identification, such as Parametric and non-parametric methods, for linear and nonlinear systems, and in Time Domain and Frequency Domain.

We have also discussed the different procedures of system identification, and how to identify using MATLAB toolbox starting from measuring the input and output signals of a given system in time or frequency domain.

When, model structure is selected, an estimator is applied online and offline in order to find estimate values for a parametric system or to determine system functions defining a non-parametric system.

Chapter Two

Distributed Parameter System

1. Introduction:

Distributed parameter system (DPS**)** is an established area of research in control that can trace its roots back to the sixties **[11]**. While the general aims are the same as for lumped parameter systems, to adequately describe the distributed nature of the system one needs to use partial differential equation (PDE) models, or the Ordinary Difference Equation (ODE).

In this chapter, we are going to present the Distributed Parameter System; starting with its history and first research about it, then we speak about its characteristics and what makes the Distributed Parameter Systems different than Lumped Parameter Systems. Finally we state some examples of such systems in the industry and control domain.

2. Brief Definitions of Distributed Parameter System:

Distributed effects are present in almost all physical systems. In some cases, these can be safely ignored but there are many interesting problems where these effects must be taken into account.

Distributed Parameter Systems are systems whose state space is infinite-dimensional; that is why they are called infinite-dimensional systems. They can be represented or described using Partial Differential Equations or Delay Differential Equations, this gives rise to distinctive features.

3. Mathematical Description of DPS:

It appears from the literature that Paraskevopoulos and Bounas, 1978 were the first investigators to study the identification of distributed parameter systems via orthogonal functions **[12]**. They used Walsh functions. The main drawback of the Walsh function approach is the selection of Walsh functions based on 2^k where k is any positive integer. For a moderately large value of k this approach becomes computationally laborious.

In the same year Tzafestas **[13]**, 1978 has also investigated the general distributed parameter system identification by first transforming the model of a distributed parameter system into its equivalent lumped form by using Galerkin expansion .In 1983, Paraskevopoulos and Kekkeris, revisited the same problem of Paraskevopoulos and Bounas. This time they employed Tchebycheff polynomials of the first kind and restricted their investigations to only first-order systems.

3.1. Unified Approach for Identification:

Consider the model of a linear time-invariant distributed parameter system described by:

$$a_{tt}\frac{\partial^2 y(x,t)}{\partial t^2} + a_{xx}\frac{\partial^2 y(x,t)}{\partial x^2} + a_{xt}\frac{\partial^2 y(x,t)}{\partial x \partial t} +$$
$$a_t \frac{\partial y(x,t)}{\partial t} + a_x \frac{\partial y(x,t)}{\partial x} + a y(x,t) = u(x,t) \quad (2.1)$$

With initial conditions $y(x,t_0)$ and $\frac{\partial y(x,t)}{\partial t}|_{t=t_0}$ and boundary conditions $y(x_0,t)$ and $\frac{\partial y(x,t)}{\partial t}|_{x=x_0}$

To identify the system given by (2.1), the input $u(x,t)$ and the output $y(x,t)$ of the system are assumed to be available over the region xϵ $[x_0,x_f]$ x $[t_0,t_f]$.

The aim is to estimate the system parameters $a_{tt}, a_{xx}, a_{xt}, a_x, a_t, a$

It may be noted that depending upon the value of $\Delta = a_{xt}{}^2 - 4a_{tt}a_{xx}$, the system described by (2.1) turns out to be

i) An elliptic system for $\Delta < 0$.

ii) A parabolic system for $\Delta = 0$ and

iii) A hyperbolic system for $\Delta > 0$.

3.2. DPS Subjected to Additive Noise:

Let y(x,t) be the scalar system state.

$$\frac{\partial y(x,t)}{\partial t} = L_x y(x,t) + f(x,t) + G(x,t)\gamma(x,t), \quad t\epsilon T =]0,t_f[, \quad x\epsilon D, \quad (2.2)$$

With initial and boundary conditions

I.C. $y(x,0) = y_0(x)$ $x\epsilon D$ (2.3)

B.C $B_x y(x,t) = 0$ $t\epsilon]0,t_f[$, $x\epsilon D$ (2.4)

Where both L_x and B_x are well-posed linear spatial differential operators, $\gamma(x,t)$ is the zero-mean Gaussian white (with respect to t) distributed noise process, f and G are known functions and $y_o(x)$ is the initial state function which is usually assumed to be the Gaussian stochastic variable for each x independent of $\gamma(x,t)$.There are many cases where one or more parameters of L_x are unknown.

A mathematical model of a general class of nonlinear DPSs is written by using a conventional function F in the form:

$$\frac{\partial y(x,t)}{\partial t} = F(x, t, y, y_x, y_{xx}, \dots) + G(x, t)\gamma(x,t) \quad (2.5)$$

Where, for the mathematical security, the function F is usually assumed to be sufficiently smooth with respect to its arguments, and where u_x is the partial derivative with respect to x. In order to handle practical problems, Eq.(2.9) is written in a more concrete form,

$$\frac{\partial y(x,t)}{\partial t} = N_s(x, t, u) + G(x, t)\gamma(x,t) \quad (2.6)$$

With the boundary condition which is considered to be also nonlinear, e.g.

$$N_b(x, t, y)=0 \qquad t\epsilon]0, t_f[, \qquad x\epsilon \partial D \quad (2.7)$$

and with the same initial condition as given by (2.3). It should be noted that, if there exists an additive noise at the system boundary, then the noise term $\varepsilon_b(x,t)$ appears on the right hand side of both Eqs.(2.4) and (2.7).

3.3. DPS with Stochastic Coefficients:

In many practical problems, one or more coefficients in differential operators are random because measurements of physical properties of the system considered inherently exhibit greater uncertainty. It may therefore be more realistic to consider system parameters as stochastic variables with a probability distribution. Many different forms have been proposed for DPSs and each is useful for different tasks **[14].** A general form using the parabolic type, which arises directly from fundamental physical axioms, is as follows:

$$\frac{\partial y(x,t)}{\partial t} = A\,(x, t, w;\ D_x)y(x, t) + f(x, t) \qquad (x,t)\ \epsilon\ T \times D \quad (2.8)$$

With the boundary condition

$$B(x, t; D_x)y(x, t)=0 \qquad (x,t)\ \epsilon\ T \times D \quad (2.9)$$

and with the same initial condition as given by (2.3), where A is a partial differential operator containing stochastic coefficients, B is a boundary operator with deterministic non vanishing coefficients and *w* is the generic point of the sample space. Since unknown stochastic

coefficients are involved in A, the problem at hand belongs to the context of stochastic inverse problems and is motivated by the oil exploration survey, the experimental physiology, cardiology, etc.

3.4. DPS with Free Boundary:

Recent advances in control technique have stirred a great deal of enthusiasm in the development of parameter identification and state estimate of DPSs with free boundaries **[14].** A number of systems of practical interest involve boundaries moving by phase change such as melting or solidification, chemical reaction, heat transfer and so on. Considering a spatial region consisting of water and ice, and choosing the system state as the temperature distribution of the water region, the water-ice interface becomes a moving boundary denoted by y(t). For example, consider a mathematical model of the system:

$$\frac{\partial\, y(x,t)}{\partial t} = a\,\frac{\partial^2 y(x,t)}{\partial x^2} \qquad (x,t) \in TxD \tag{2.10}$$

with I.C $$y(x,0) = y_o(x) \geq 0 \qquad \text{for } x \in D \tag{2.11}$$

On the boundary i.e. the ice-water interface, by considering the latent heat of fusion and the rate at which ice is converted into water, it follows that, for 0 < t < tf,

$$ay(0,t) = g(t),\ y(t,x) = 0,\ a\frac{\partial y(t,x)}{\partial x} = -\frac{dx}{dt}, \tag{2.12}$$

where a is a positive constant and the ice region is assumed to be bounded. Fig. (2.1) illustrates the construction of a DPS with free boundary. A fundamental difference between regular boundary problems and free boundary ones is that the domain of solutions to the basic state equation is not known but should be determined by additional information through underlying properties of physical systems under consideration **[14]**.

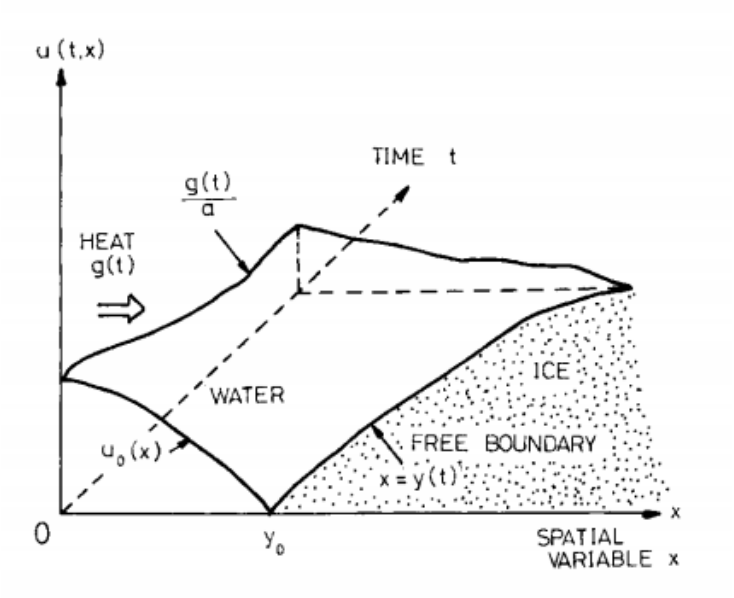

Fig.2.1 DPS with free boundaries.

3.5. DPS with Inequality Boundary Conditions:

There has recently been much practical interest shown in problems of the DPS with a permeable wall as its system boundary **[14]**. Identification problems of system parameters in which the osmotic pressure dues to the nature of membrane are of practical importance. More basically, in the field of synovial pint biomechanics, the permeability plays an important role of analyzing locomotion of the musculoskeletal system. From theoretical viewpoints, a remarkable feature of the problem considered here relates to a mathematical modelling of the system boundary.

Suppose that the region D is filled up by the liquid as shown in Fig.(2.2). Then, the state variable y(x,t) is reasonably selected to be the pressure of the fluid. The boundaries of D are respectively denoted by Γ_0 at x = 0 and by Γ_1 at x = 1. The boundary conditions peculiar to the system shown in Fig.(2.2). are considered as follows. At the boundary Γ_0, the well-known Neumann condition is set in a form of:

$$-a\frac{\partial y(x,t)}{\partial x}\Big|_{x=0} = f(t) \quad \text{on } \Gamma_0 \tag{2.13}$$

Where a is a positive constant expressing the diffusion coefficient and f(t) is the known input function applied to the boundary Γ_0.

On the other hand, the boundary Γ_1 consists of a semi-permeable membrane whose thickness is supposed to be negligible. This implies that, as shown in Fig.(2.2)., although the fluid flow leaving D freely passes Γ_1, the counter flow is prevented by the existence of the membrane. Let h(t) be the given fluid pressure in the outside of D.

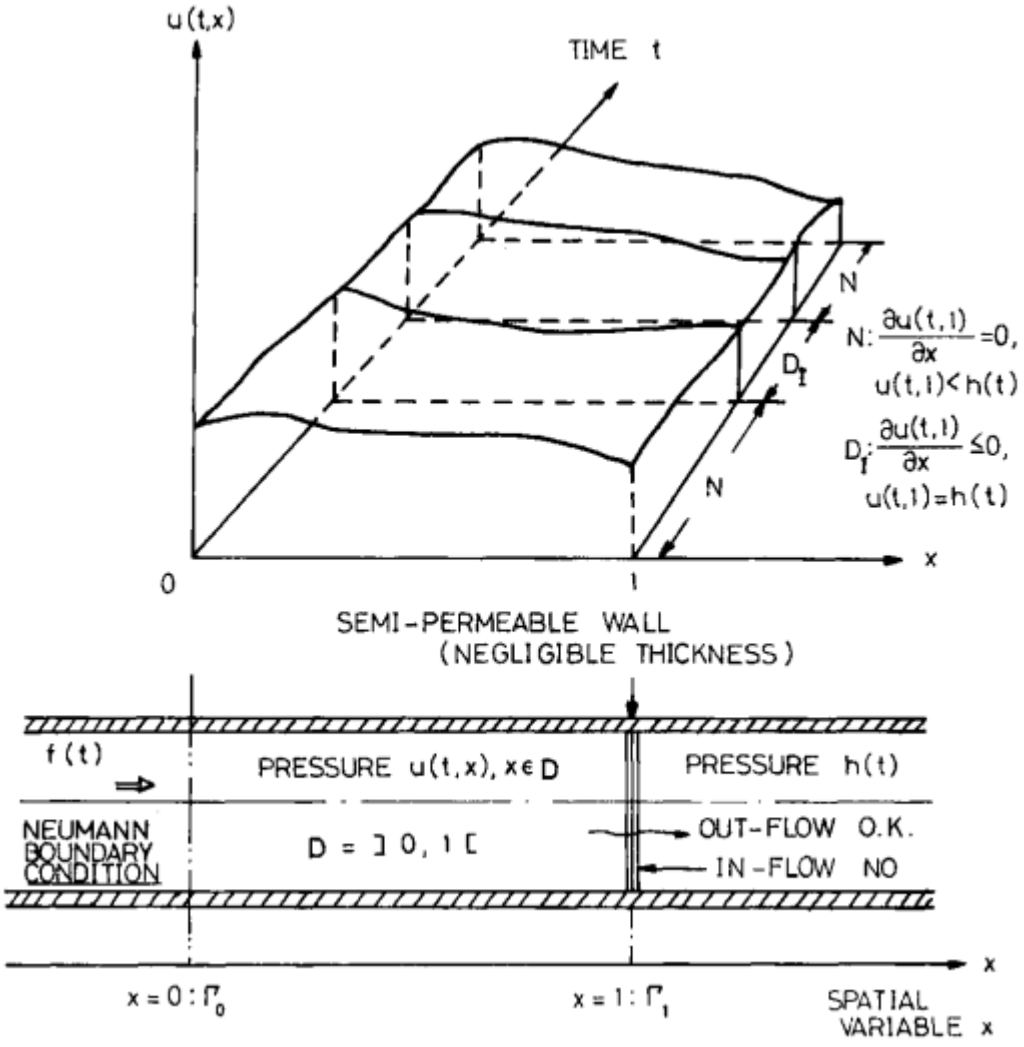

Fig.2.2 System with semi-permeable wall.

First, suppose that

Case-1: y(1,t) < h(t); There exists a fight where the semi-permeable wall prevents the fluid flow against the fluid trying to enter into D. A success of the fight allows us to write

$$\frac{\partial u(x,t)}{\partial x}|_{x=0} = 0 \quad \text{on } \Gamma_0 \tag{2.14}$$

Case-2: y(1,t)> h(t); This case implies that the fluid tries to leave D.

Hence $$\frac{\partial y(x,t)}{\partial x}|_{x=1} \leq 0 \quad \text{on } \Gamma_1 \tag{2.15}$$

However, since the thickness of the semi-permeable wall is negligible, y(1,t) is not greater than h(t), i.e., y(l,t) = h(t). The results mentioned above are summarized with the same state equation as given by Eq. (2.10) and the same initial conditions as given by (2.3), the boundary conditions are modelled by (2.13), (2.14) and (2.15).

4. Lumped Vs Distributed Parameter System:

Lumped Parameter Systems are modeled by ODEs (Ordinary Difference Equations). Their state space are finite dimensional. It is a system in which the dependent variables of interest are a function of time alone. Whereas Distributed Parameter Systems are modeled by PDEs or DDEs (Delay Differential Equations). Their state spaces are infinite-dimensional. It is a system in which all dependent variables are functions of time and one or more spatial variables.

For example, consider the following two systems illustrated in the following figure:

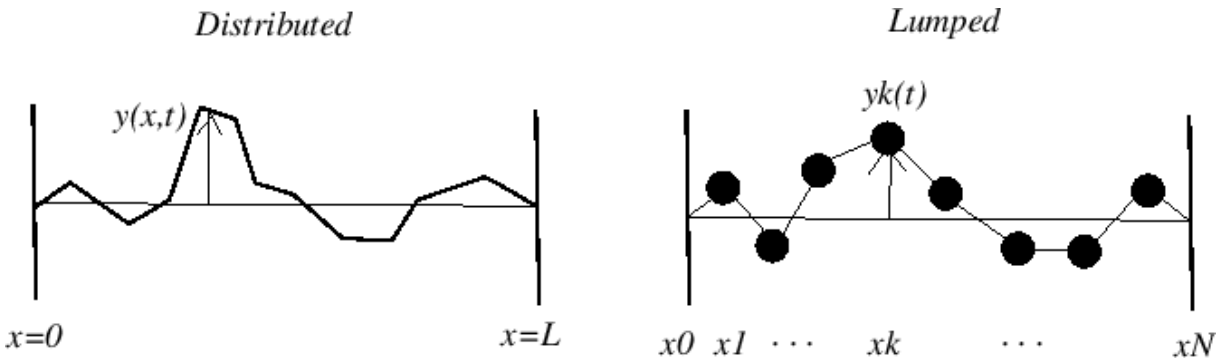

Fig.2.3 Distributed Vs Lumped System

The first system is a distributed system, consisting of an infinitely thin string, supported at both ends; the dependent variable, the vertical position of the string $y(x,t)$is indexed continuously in both space and time.

The second system, a series of ``beads" connected by massless string segments, constrained to move vertically, can be thought of as a lumped system, perhaps an approximation to the continuous string.

For electrical systems, consider the difference between a lumped RLC network and a transmission line (see figure 2.4)

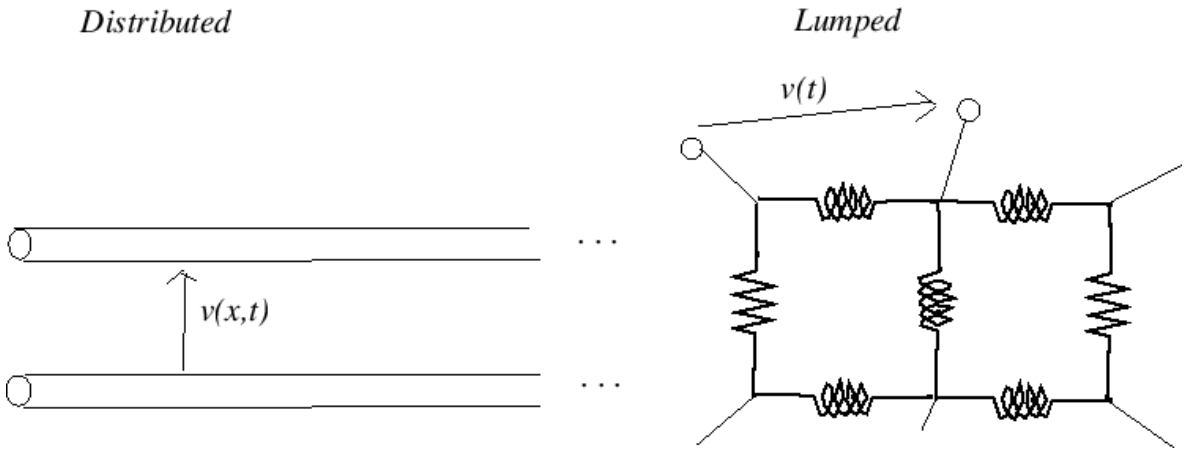

Fig.2.4 Transmission line Vs Lumped RLC system

The importance of lumped approximations to distributed systems will become obvious later, especially for waveguide-based physical modeling, because it enables one to cut computational costs by solving ODEs at a few points, rather than a full PDE (generally much more costly) **[15].**

5. Examples:

A wide variety of phenomena can be represented by the PDE of Distributed Parameter Systems, from sound, to heat, electrostatics, electrodynamics, fluid flow, etc. These seemingly distinct physical phenomena can be formalized similarly in terms of PDEs. Just as ordinary differential equations often model one-dimensional dynamical systems, partial differential equations often model multidimensional systems.

5.1. Wave Equation:

The wave equation is an equation for an unknown function y(k, x) of the form:

$$y_{kk} = m^2 y_{xx} \tag{2.16}$$

Here y might describe the displacement of a stretched string from equilibrium, or the difference in air pressure in a tube, or the magnitude of an electromagnetic field in a tube, and *m* is a number that corresponds to the velocity of the wave**[16]**.

5.2. Euler-Tricomi Equation:

In mathematics, the Euler–Tricomi equation is a linear partial differential equation useful in the study of transonic flow **[17]**. It is named for Leonhard Euler and Francesco Giacomo Tricomi and is given by

$$y_{xx} = z y_{zz} \tag{2.17}$$

It is hyperbolic in the half plane x > 0, parabolic at x = 0 and elliptic in the half plane x < 0. Its characteristics are:

$$x dx^2 = dz^2 \tag{2.18}$$

Which have the integral

$$z \pm \frac{2}{3} x^{\frac{3}{2}} = C \tag{2.19}$$

Where C is a constant of integration. The characteristics thus comprise two families of semi cubical parabolas, with cusps on the line x = 0, the curves lying on the right hand side of the z-axis.

5.3. Heat Equation:

The equation for conduction of heat in one dimension for a homogeneous body has

$$y_t = \alpha y_{xx} \tag{2.20}$$

Where y(t,x) is temperature, and α is a positive constant that describes the rate of diffusion.

We are going to discuss and study this application in details later **[16]**.

6. Conclusion:

In this part, we have presented the distributed parameter systems with its characteristics and features. We have discussed its history and mentioned the first studies about the topic done by Scientists and mathematicians. Then we have introduced a comparison of lumped and distributed parameter systems; their models and applications. Finally, we have shown some examples and physical phenomena that can be represented by PDE.

Chapter Three

Heat Conduction Application

1. Introduction:

In this chapter, we present an application of parametric identification of a distributed parameter system which represent the dynamic behavior of the heat conduction. The parameters of the heat equation are identified using Least Square method.

The application of the least square estimator is divided into three main parts. The first two parts are divided into the estimation of parameters sets for two different numerical simulations. These estimates are obtained for different noise amplitude and using different structures. However, the third part is about the identification of conduction of the heat in a cement rotary kiln.

2. Description of Heat Conduction:

Heat transfer is of particular interest to engineers, who attempt to understand and control the flow of heat through the use of thermal insulation, heat exchangers, and other devices. Heat transfer is defined as the flow of thermal energy within solids and non-flowing fluids, driven by thermal non-equilibrium (i.e. the effect of a non-uniform temperature field), commonly measured as a heat flux (vector), i.e. the heat flow per unit time (and usually unit normal area) at a control surface.

In a metal rod with non-uniform temperature, heat (thermal energy) is transferred from regions of higher temperature to regions of lower temperature. Three physical principles are used here:

1. Heat (or thermal) energy of a body with uniform properties

$$\text{Heat energy} = c*m*T \tag{3.1}$$

Where m is the body mass, T is the temperature, c is the specific heat which is the amount of heat required to change the body's temperature by one degree.

2. Fourier's law of heat transfer represent rate of heat transfer proportional to negative temperature gradient,

$$\frac{\text{Rateofheattransfer}}{\text{area}} = -K_0 \frac{dT}{\partial x} \tag{3.2}$$

Where K_o is the thermal conductivity. In other words, heat is transferred from areas of high temperature to low temperature.

3. Conservation of energy:

Consider a uniform rod of length *l* with non-uniform temperature lying on the *x*-axis from x =0 to x = l. By uniform rod, we mean the density ρ, specific heat c, thermal conductivity K_o, cross-sectional area A are all constant. Assume the sides of the rod are insulated and only the ends may be exposed. Also, assume there is no heat source within the rod. Consider an arbitrary thin slice of the rod of width Δx between x and x+Δx. The slice is so thin that the temperature throughout the slice is T(x,t).

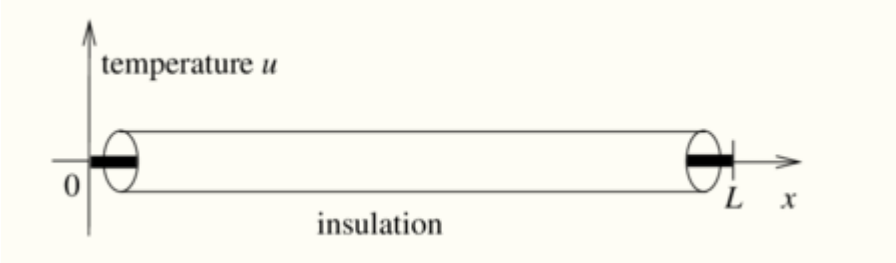

Fig 3.1 uniform insolated rod of length L.

Thus,

$$\text{Heat energy of segment} = c \times \rho A \Delta x \times T = c\rho A \Delta x T(x, t) \qquad (3.3)$$

By conservation of energy,

Change of heat energy of segment in time Δt	=	Heat in form Left boundary	-	Heat out form Right boundary

From Fourier equation (3.2)

$$c\rho A \Delta x T(x, t + \Delta t) - c\rho A \Delta x T(x, t) = \Delta t A \left(\frac{\partial T}{\partial x}\right)\Big|_{x} - \Delta t A \left(\frac{\partial T}{\partial x}\right)\Big|_{x+\Delta x} \qquad (3.4)$$

Rearranging (3.4) yields (recall ρ, c, A, K0 are constant),

$$\frac{y(x,t+\Delta t) - y(x,t)}{\Delta t} = -\frac{K_0}{c\rho}\left(\frac{\left(\frac{\partial y}{\partial x}\right)_x - \left(\frac{\partial y}{\partial x}\right)\Big|_{x+\Delta x}}{\Delta x}\right) \qquad (3.5)$$

Taking the limit $\Delta t, \Delta x \rightarrow 0$ gives the heat equation:

$$\frac{\partial y}{\partial t} = k\frac{\partial^2 y}{\partial x^2} \tag{3.6}$$

where $k = \frac{K_0}{c\rho}$ is called the thermal diffusivity. Since the slice was chosen arbitrarily, the Heat Equation (3.6) applies throughout the rod.

4. Initial condition and boundary conditions

To make use of the Heat Equation (3.6), we need more information:

1. Initial Condition(IC): In this case, the initial temperature distribution in the rod is T(x, 0).
2. Boundary Conditions (BC):

In this case, the temperature of the rod is affected by what happens at the ends, x =0, and x=*l*. What happens to the temperature at the end of the rod must be specified. In reality, the BCs can be complicated.

Temperature prescribed at a boundary

$$T(0,t)= T_1(t)$$

$$T(L,t)=T_l(t)$$

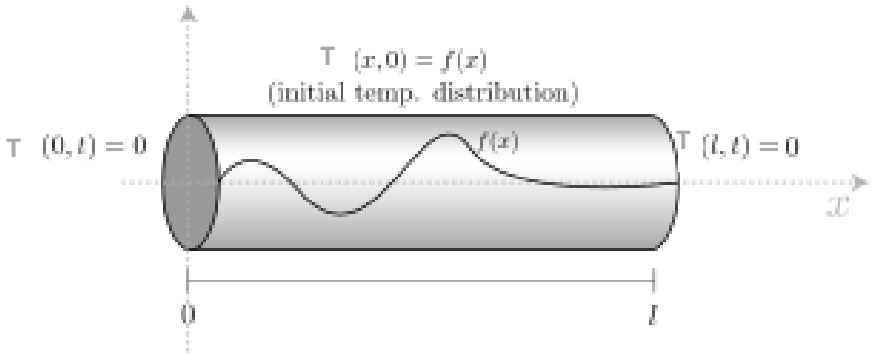

Fig 3.2 Description of Initial and boundary conditions in the rod.

3. Discretization of Heat equation using Euler Method:

For the following problem involving the heat equation with a source term u(x,t),

$$T_t = \beta T_{xx} + f(x,t) + u(x,t), \quad a < x < b, t > 0, \tag{3.7}$$

$$BC: T(a,t) = g_1(t), \quad T(b,t) = g_2(t), \quad \text{IC:} \quad T(x,0) = T_0(x) \tag{3.8}$$

Let us seek a numerical solution for T(x, t) at a particular time t> 0 or at certain times in the interval $0 < t < t_f$.

As the first step, we expect to generate a grid

$$x_i = a + ih, \quad i=0, 1,\ldots, m, \quad h = \frac{b-a}{m}$$

$$t^k = k\Delta t, k=0, 1,\ldots, n, \quad \Delta t = \frac{T}{n}$$

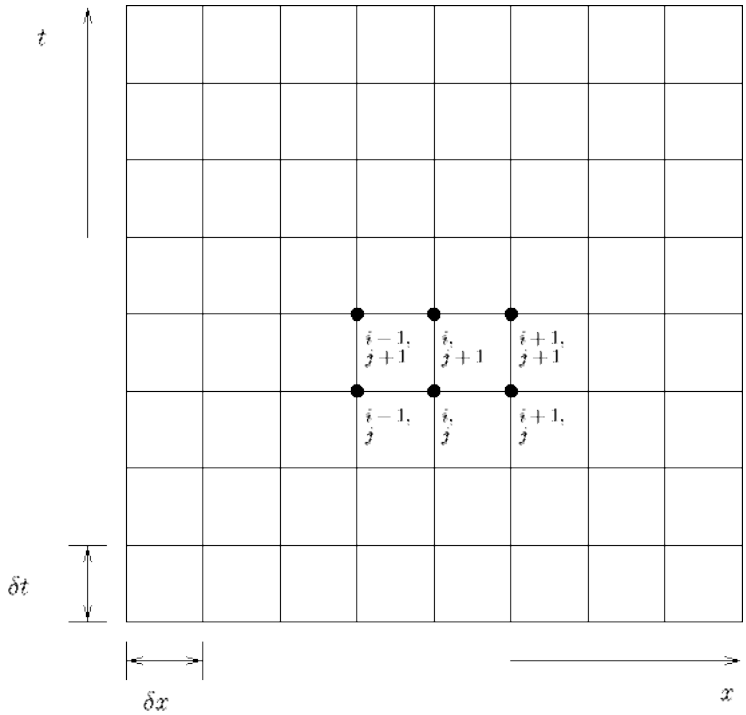

Fig 3.3 generation of the grid.

It turns out that any arbitrary Δt cannot be used for explicit methods because of numerical instability concerns. The second step is to approximate the derivatives with finite difference (FD) approximations. Since we already know how to discretize the spatial derivatives, let us focus on possible FD formulas for the time derivative. At grid point (x_i, t^k), $k \geq 0$ using the forward FD approximation for $T_i$central and FD approximation for T_{xx} we have:

$$\frac{T(x_i,t^k+\Delta t)-T(x_i,t^k)}{\Delta t}=\beta\frac{T(x_{i-1},t^k)-2T(x_i,t^k)+T(x_{i+1},t^k)}{h^2}+f(x_i,t^k)+u(x_i,t^k) \quad (3.9)$$

The discretization is first order in time and second order in space, when the FD equation is:

$$\frac{(T_i^{k+1})-(T_i^k)}{\Delta t}=\beta\frac{(T_{i-1}^k)-2(T_i^k)+(T_{i+1}^k)}{h^2}+f_i^k+u_i^k \quad (3.10)$$

where $f_i^k= f(x_i,t^k)$ with T_i^k again denoting the approximate values for the true solution $T(x_i,t^k)$. When k =0, T_i^0 is the initial condition at the grid point $(x_i,0)$ and from the values T_i^k at the time level k the solution of the FD equation at the next time level k+ 1 is

$$(T_i^{k+1}) = (T_i^k) + \Delta t\,(\beta\frac{(T_{i-1}^k)-2(T_i^k)+(T_{i+1}^k)}{h^2}+f_i^k+u_i^k), \quad i=1,2\ldots,m-1 \quad (3.11)$$

Let $\alpha_1 = \frac{\Delta t\beta}{h^2}$ and $\alpha_2 = 1-2\frac{\Delta t\beta}{h^2}$

$$Tes_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_1 T_{i-1}^j \quad (3.12)$$

4. Numerical Simulation:

4.1. Homogenous heat equation:

The first application deals with a linear homogenous heat equation. This example is solved using the separation of variables method. The exact solution of the heat equation is used to generate data that is employed in the identification process.

The example is about an insulated unit wire, such that its ends are embedded in ice (temperature 0°). Let k=0.003 and initial distribution is T(x,0)=50x(1-x). The Heat equation is:

$$\frac{\partial T}{\partial t} = 0.003\frac{\partial^2 T}{\partial x^2} \quad (3.13)$$

with Boundary Conditions: $T(0,t) = T(1,t)=0$

The exact solution of th equation (3.13) is:

$$T(x,t) = \begin{cases} \sum_{n=1}^{\infty}\frac{400}{\pi^3 n^3}\sin(n\pi x)e e^{-n^2\pi^2 0.003t} & ,n\ odd \\ 0 & ,n\ even \end{cases} \quad (3.14)$$

The spatial-temporal evolution of temperature is shown in the figure 3.4

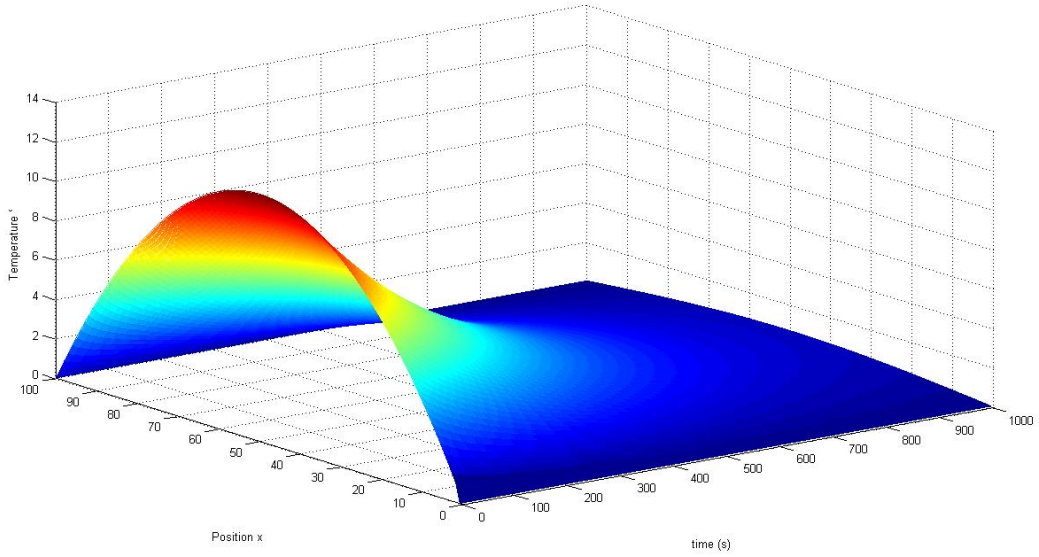

Fig 3.4 Temperature Evolution.

Using the Least Square Method, we estimate the parameters α_1,$\alpha_2$and α_3 of the FD model (3.12) which yields to one step-ahead prediction error at location *i*:

$$\varepsilon_i^k = (\hat{\mathrm{T}}_i^{k+1/k} - \mathrm{T}_i^k)^2 \tag{3.15}$$

where the estimate $\hat{\mathrm{T}}_i^{k+1/k}$ uses different locations.

The cost function to be minimized in order to obtain a better fit is of type:

$$\min_{\theta}\left(\sum_{i=0}^{m-1}\left(\sum_{j=0}^{n-1}(\hat{\mathrm{T}}_i^{k+1/k} - \mathrm{T}_i^k)^2\right)\right) = 0 \tag{3.16}$$

Where $\theta = [\alpha_1, \alpha_2, \alpha_3]^T$

The vector of parameters θ is the root of:

$$\nabla_{\theta}\left(\sum_{i=0}^{m-1}\left(\sum_{j=0}^{n-1}(\hat{\mathrm{T}}_i^{k+1/k} - \mathrm{T}_i^k)^2\right)\right) = 0 \tag{3.17}$$

This is equivalent to:

$$\begin{bmatrix} \frac{d\sum\sum}{d\alpha_1} \\ \frac{d\sum\sum}{d\alpha_2} \\ \frac{d\sum\sum}{d\alpha_3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.18}$$

This yiekds to the following set of algebraic equations:

$$\begin{matrix}\frac{dLS}{d\alpha_1}=0\\ \frac{dLS}{d\alpha_2}=0\\ \frac{dLS}{d\alpha_3}=0\end{matrix}\rightarrow\begin{bmatrix}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_i^{j+1}T_{i+1}^{j}\\ \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_i^{j+1}T_i^{j}\\ \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T^{j+1}T_{i-1}^{j}\end{bmatrix}=\begin{bmatrix}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_{i+1}^{j\,2} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_i^{j}T_{i+1}^{j} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_{i-1}^{j}T_{i+1}^{j}\\ \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_i^{j}T_{i+1}^{j} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_i^{j\,2} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_{i-1}^{j}T_i^{j}\\ \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_{i-1}^{j}T_{i+1}^{j} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_{i-1}^{j}T_i^{j} & \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}T_{i-1}^{j\,2}\end{bmatrix}\begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{bmatrix}$$

Table 3.1 Identification's results of the linear example.

Noise σ^2	α1	α2	α3	MSE
0	0.009047	0.979214	0.008742	$2.400530.10^{-5}$
0.0001	0.009060	0.979189	0.008754	$2.402542.10^{-5}$
0.0004	0.009276	0.978762	0.008965	$2.430057.10^{-5}$
0.0025	0.018451	0.960715	0.017847	$3.592499.10^{-5}$
0.01	0.109866	0.781078	0.106161	$1.903710.10^{-4}$

The estimated data is generated using the identified parameters, first by assuming the same Neumann boundary conditions and initial conditions as for the original data, then estimating the boundary conditions using the identified parameters.

The spatial-temporal squared one step-ahead prediction error is shown in figures 3.5 and 3.6 using corrupted and uncorrupted generated data, respectively.

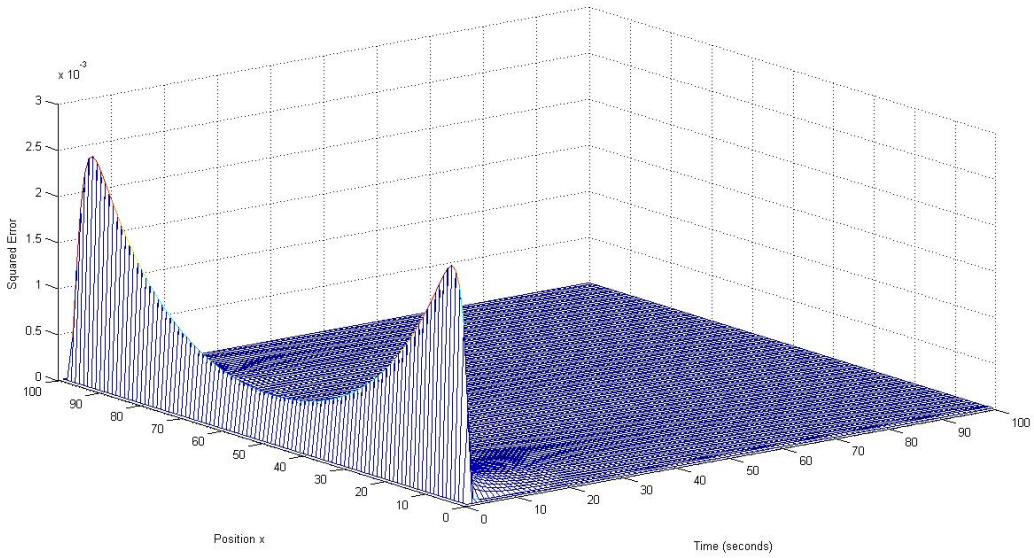

Fig3.5 Uncorrupted One step-ahead prediction error.

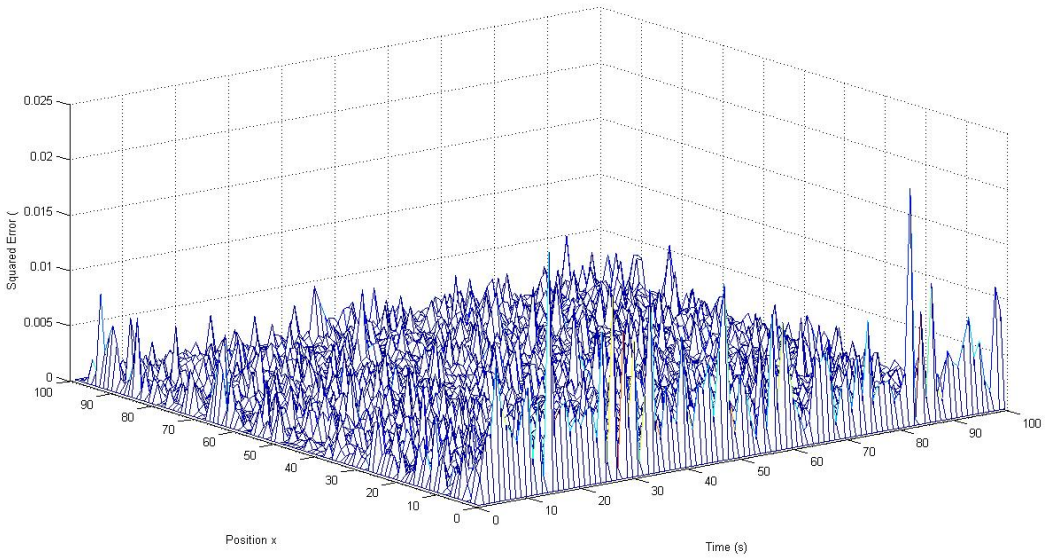

Fig3.6 Uncorrupted One step-ahead prediction error ($\boldsymbol{\sigma}^{\mathbf{2}}$ = 0.01).

Discussion of the results:

According to the results obtained, the estimated data is close enough to the exact data; $\boldsymbol{\alpha_1}$ approximately equals to $\boldsymbol{\alpha_3}$ and $\boldsymbol{\alpha_2} \cong 1 - 2\boldsymbol{\alpha_1}$, which satisfies the linear difference equation.

The obtained error (E=$\hat{T}$-T) has a small variance which means that the least square based-estimator performs well even in the case of noise; increasing the standard deviation of noise led to a small increase in the values of parameters and a very small change in the one step-ahead error.

4.2. Non-Linear Heat equation:

The second application deals with a nonlinear heat equation. The equation is described as follows:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} - 3T^3 \tag{3.19}$$

with Boundary Conditions: $T(0,t)=\frac{1}{6t+1}$

$$T(1,t)=\frac{1}{2t+1}$$

And initial conditions: $T(x,0) = \frac{1+2x}{x^2+x+1}$

This example is solved using the separation of variable method (used to solve PDE). The analytic solution of the heat equation is used to generate data that is employed in the identification process.

Using the separation of variables, we get the solution:

$$T(x,t) = \frac{1+2x}{x^2+x+6t+1} \tag{3.20}$$

This temperature is illustrated in figure 3.7 on the same way and by using the data generated by the equation 3.14, the estimate parameter values are regrouped in table 3.2

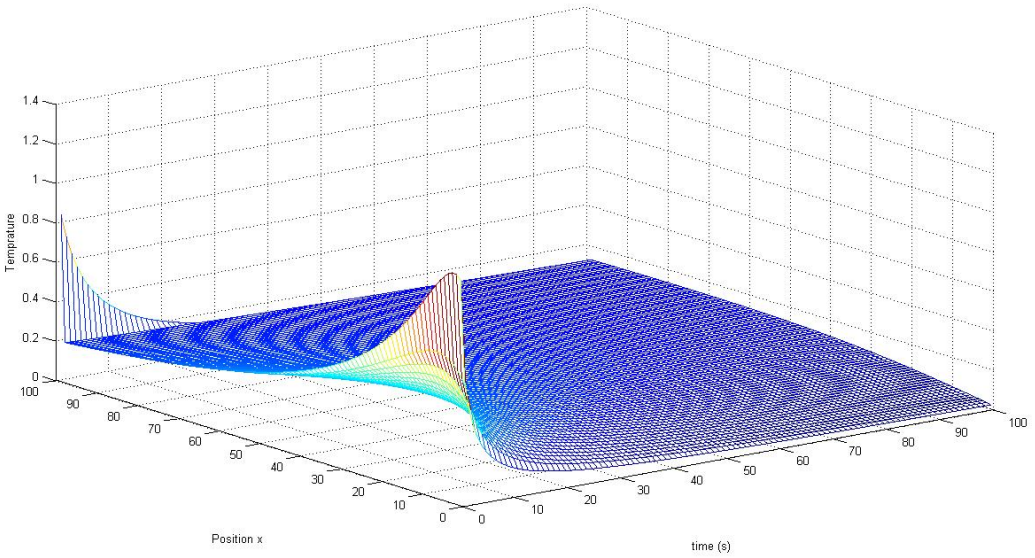

Fig3.7 Temperature Evolution of the nonlinear example.

Table 3.2 Identification's results of the nonlinear example.

Noise σ^2	α1	α2	α3	MSE
0	0.077655	0.323855	0.560467	$1.836192.10.^{-4}$
0.0001	0.077594	0.324086	0.560299	$1.836049.10^{-4}$
0.0004	0.078078	0.324334	0.559561	$1.838098.10^{-4}$
0.0025	0.086394	0.335814	0.539715	$1.918635.10^{-4}$
0.01	0.164831	0.365850	0.430015	$3.153481.10^{-4}$

The spatial-temporal squared one step-ahead prediction error is shown in figures 3.8 and 3.9 using corrupted and uncorrupted generated data, respectively.

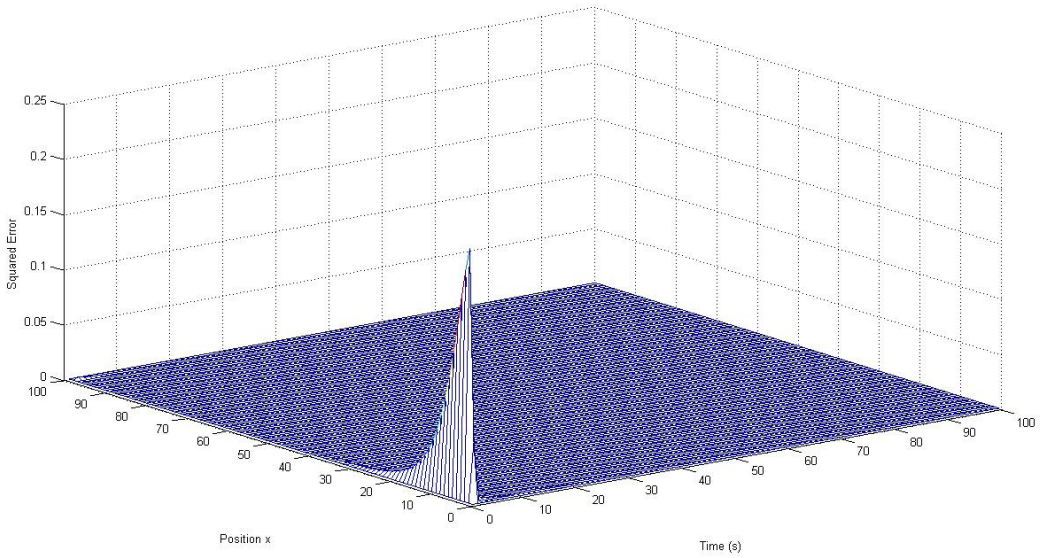

Fig3.8 Uncorrupted One step-ahead prediction error of the nonlinear example.

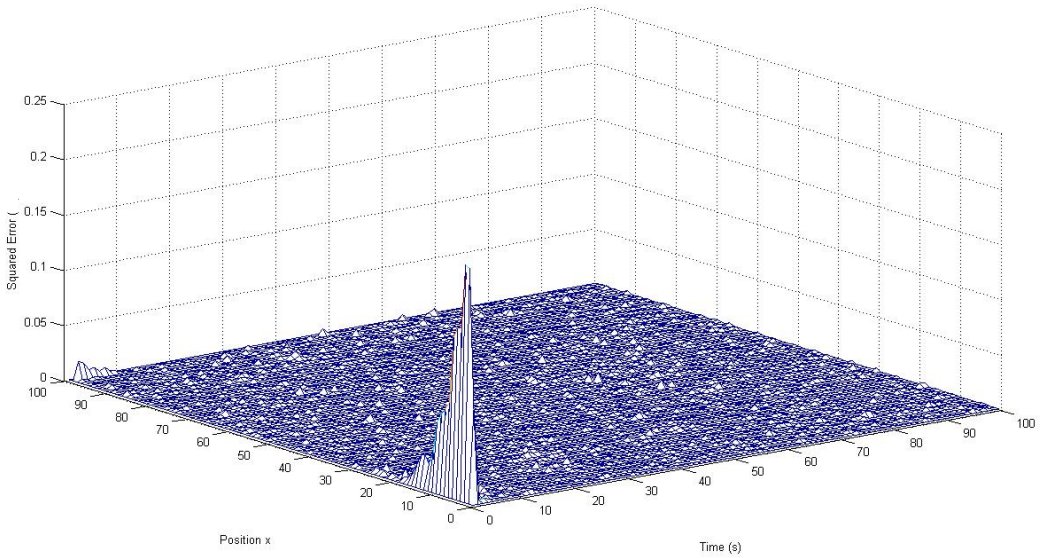

Fig3.9 Corrupted One step-ahead prediction error ($\boldsymbol{\sigma}^2$ = 0.01) of the nonlinear example.

Discussion of the results:

According to the results, the error has slightly increased when the estimation of boundary conditions is included; It is clear from the error at position ***x=0*** and ***x=1*** that the estimation using the identified parameters is good since the error is small.

Another case is presented by increasing the estimation size using the squared values of the temperature at three positions so the model is:

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 {T_{i+1}^j}^2 + \alpha_5 {T_i^j}^2 + \alpha_6 {T_{i-1}^j}^2 \quad (3.21)$$

The results are presented in the following table obtained using the least square method:

Table 3.3 Identification's results using nonlinear squared terms.

Noise σ^2	α1	α2	α3	α4	α5	α6	MSE
0	0.002204	0.295866	0.740488	0.028501	-0.141175	-0.162372	3.2535255 $.10^{-6}$
0.001	0.246025	0.384823	0.424849	-0.248543	-0.054781	-0.010352	1.532014. 10^{-4}
0.002	0.376466	0.359778	0.319535	-0.339205	-0.049534	0.066557	5.424368. 10^{-4}
0.004	0.354820	0.374944	0.315453	-0.178662	-0.155882	0.002804	0.003214
0.01	0.332525	0.339789	0.333005	-0.119730	-0.148761	-0.062380	0.012406

The squared error is well sketched in the following figures:

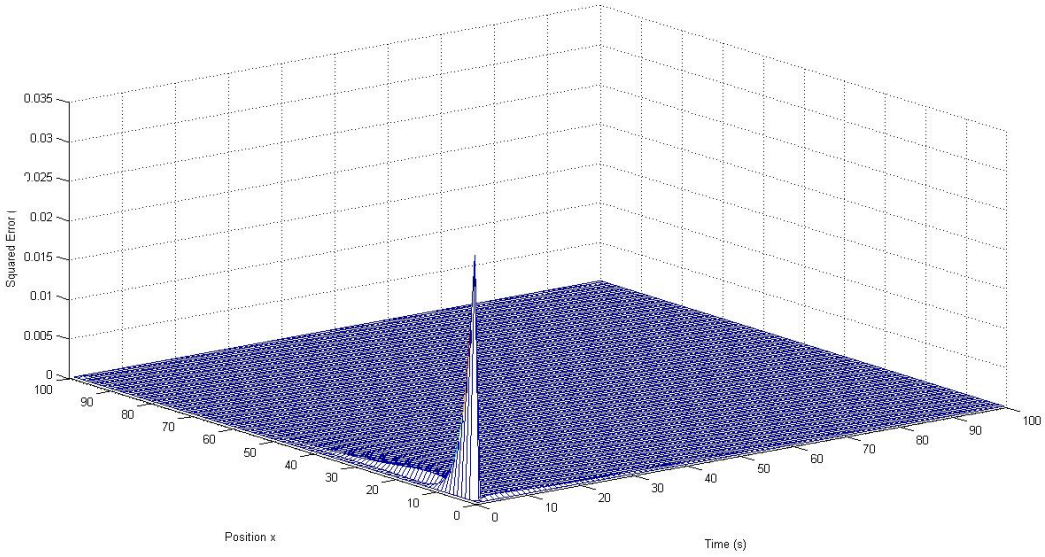

Fig3.10 Uncorrupted One step-ahead prediction error

of the nonlinear example using squared terms.

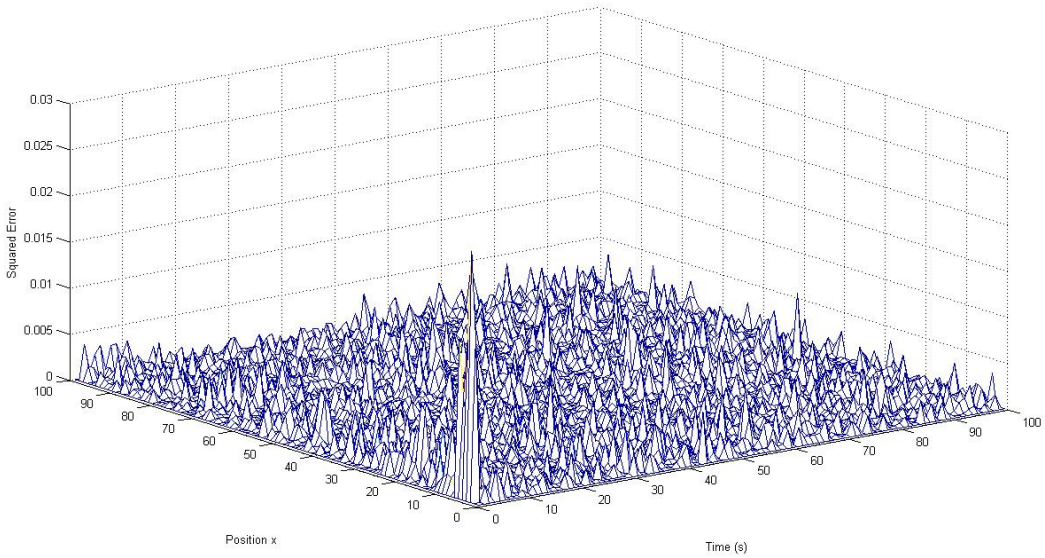

Fig3.11 Corrupted One step-ahead prediction error ($\sigma^2 = 0.01$)

of the nonlinear example using squared terms.

.

For the next case, another set of three cubed temperature elements are added to the model:

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 {T_{i+1}^j}^2 + \alpha_5 {T_i^j}^2 + \alpha_6 {T_{i-1}^j}^2 + \alpha_4 {T_{i+1}^j}^3 + \alpha_5 {T_i^j}^3 + \alpha_6 {T_{i-1}^j}^3 \quad (3.22)$$

The results are presented in the following table:

Table 3.4 Identification's results using nonlinear cubed terms.

Noise σ^2	α1	α2	α3	α4	α5
0	-0.011983	1.036778	-0.028711	0.109010	-1.708045
0.0001	-0.060703	1.088736	-0.037042	0.265553	-1.704312
0.0004	-0.061138	1.090273	-0.038166	0.266867	-1.708045
0.0025	0.055541	0.790685	0.145776	0.035042	-1.090097
0.01	0.325670	0.434422	0.232882	-0.325400	-0.276423

Noise σ^2	α6	α7	α8	α9	MSE
0	1.500813	-0.278446	0.924481	-1.004313	$2.538763.10^{-6}$
0.0001	1.498265	-0.277354	0.922114	-1.002890	$2.550344.10^{-6}$
0.0004	1.500813	-0.278446	0.924481	-1.004313	$2.538763.10^{-6}$
0.0025	1.110322	-0.145846	0.580426	-0.789244	$1.156777.10^{-5}$
0.01	0.647844	-0.025461	0.170059	-0.491869	$1.267321.10^{-4}$

The squared error is well sketched in the following figures:

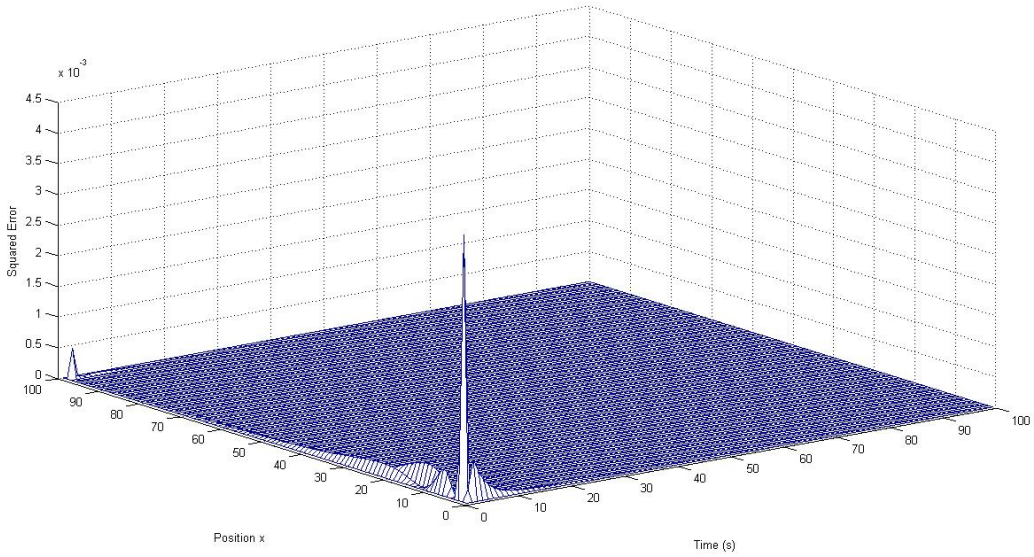

Fig3.12 Uncorrupted One step-ahead prediction error

of the nonlinear example using cubed terms.

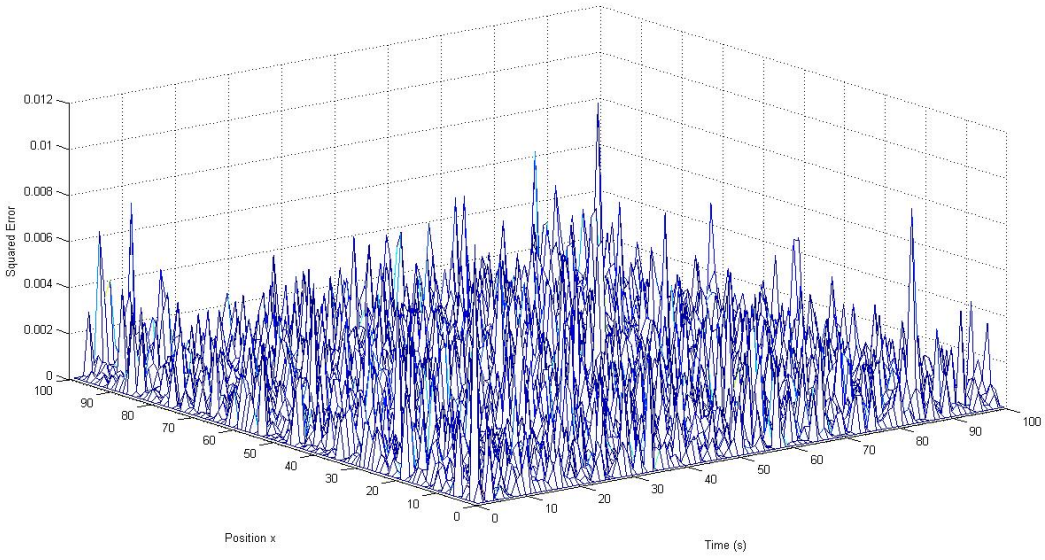

Fig3.13 Corrupted One step-ahead prediction error ($\boldsymbol{\sigma}^2 = 0.01$)

of the nonlinear example using cubed terms.

Discussion of the results:

As it is shown in tables and figures, the parameters do not satisfy the linearity of the difference equation since there are significant values for parameters of nonlinear terms; it is expected because the difference equation is nonlinear.

The calculated error (E=$\hat{T}$-T) has a small MSE (a small variance) which means that the estimated parameters matches the system characteristics. It gets smaller when adding the nonlinear terms to the identification algorithm (difference equation). So the one step-ahead prediction error shows that the nonlinear model is better fitted representation than the linear model.

The noise has a significant effect on the values of the parameters. It affects the error as well; by increasing the standard deviation (noise), the error is increased as well.

4.3. The Real Data identification:

4.3.1. Description of the Cement Rotary kiln:

The rotary kiln is the most important machine in the cement plant; its purpose is to process materials at an extreme heat in order to derive cement. The cement production industry produces over a billion tons of cement, making this type of rotary kiln a very important component within this industry.

The cement plant of Ain Kbira has a rotary kiln that consists of a hollow cylindrical metallic shell, lined using refractory bricks. It is 80 meter long, and 5 meters in height. At one end fuel, in the form of gas, oil, or pulverized solid fuel, is blown in through the "burner pipe", producing a large concentric flame in the lower part of the kiln tube. As material moves under the flame, it reaches its peak temperature, before dropping out of the kiln tube into the cooler. Air is drawn first through the cooler and then through the kiln for combustion of the fuel. In the cooler, the cooling clinker heats the air, so that it may be 400 to 800 °C before it enters the kiln, thus causing intense and rapid combustion of the fuel.

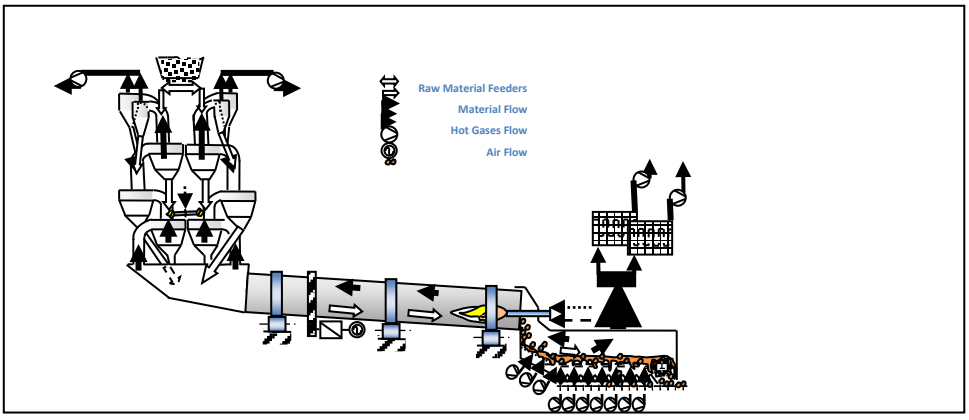

Fig 3.14 Schematic of a Cement Rotary Kiln**.**

The third part of the application is about the identification of the experimental data obtained from the cement rotary kiln. The rotary kiln has a burner placed 1.2 meter from the beginning of the tube. This burner provides energy of 70560 thermie per hour. The rows represent the temperature along the kiln length while the columns are the observation of the temperature:

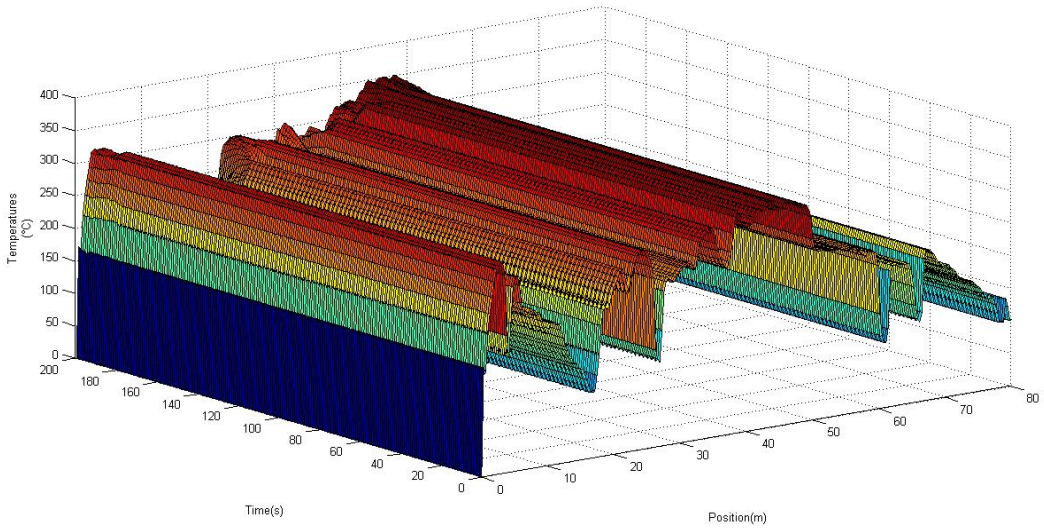

Fig3.15 Real data evolution.

Using the linear difference model with a second term u(x,t):

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 U_i^j \qquad (3.23)$$

The identified parameters and the means square error are presented in the table below:

Table 3.5 Identification's results of the real data.

Parameters	α1	α2	α3	α4	MSE
Values	0.004499	0.991775	0.003690	-0.000004	3.833575

The squared error is well sketched in the following figure:

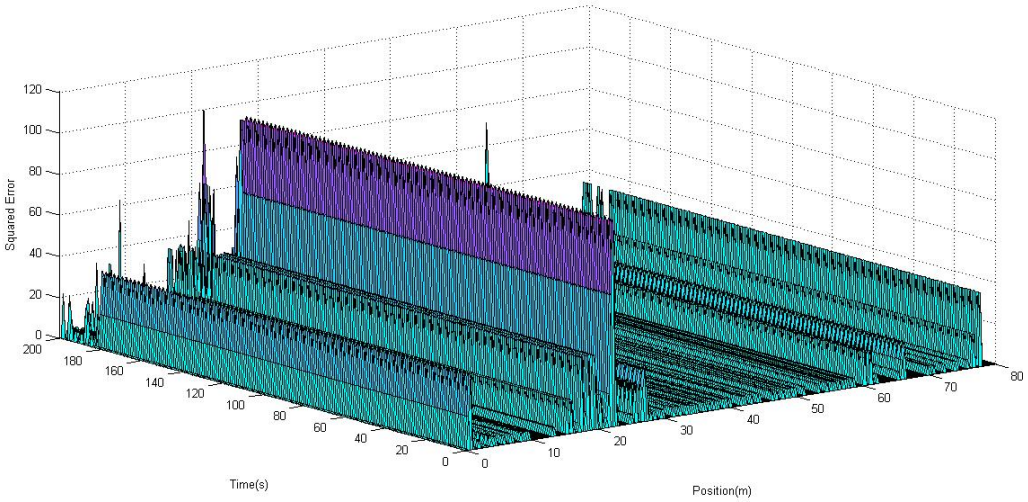

Fig3.16 One step-ahead prediction error.

Using the nonlinear difference model:

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 {T_{i+1}^j}^2 + \alpha_5 {T_i^j}^2 + \alpha_6 {T_{i-1}^j}^2 + \alpha_7 U_{i+1}^j \qquad (3.24)$$

The identified parameters and the means square error are presented in the table below:

Table 3.6 Identification's results of the real data using squared terms.

Parameters	α1	α2	α3	α4	α5	α6	α7	MSE
Values	0.000086	1.003201	-0.003397	0.000013	-0.000034	0.000021	-0.000004	3.830502

The squared error is well sketched in the following figure:

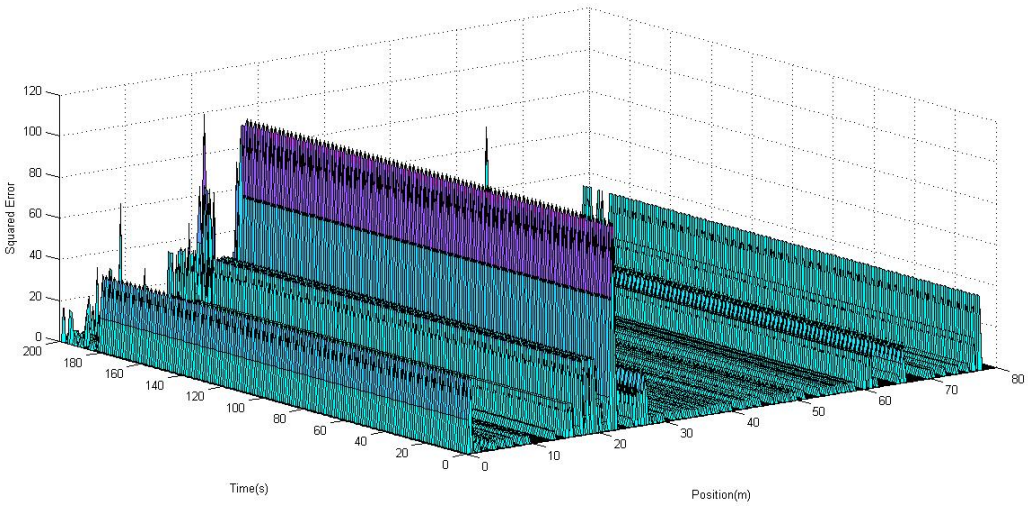

Fig3.17 One step-ahead prediction error using squared terms.

Using the nonlinear difference model

$$T_i^{j+1} = \alpha_1 T_{i+1}^j + \alpha_2 T_i^j + \alpha_3 T_{i-1}^j + \alpha_4 {T_{i+1}^j}^2 + \alpha_5 {T_i^j}^2 + \alpha_6 {T_{i-1}^j}^2 + \alpha_4 {T_{i+1}^j}^3 + \alpha_5 {T_i^j}^3 + \alpha_6 {T_{i-1}^j}^3 + \alpha_{10} U_{i+1}^j \quad (3.25)$$

The identified parameters and the means square error are presented in the table below:

Table 3.7 Identification's results of the real data using cubed terms.

Parameters	α1	α2	α3	α4	α5
Values	0.043992	0.953320	0.003237	-0.000212	0.000236

Parameters	α6	α7	α8	α9	α10	MSE
Values	-0.000028	0.0000003	-0.0000004	0.0000001	-0.000002	3.827386

The squared error is well sketched in the following figure:

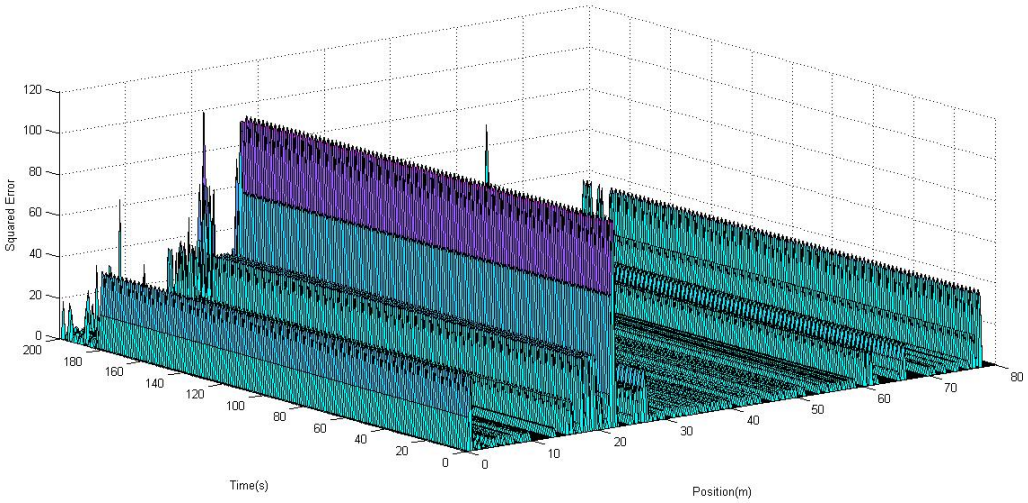

Fig 3.18 One step-ahead prediction error using cubed terms

Discussion of the results:

According to the results obtained, the error is significant for linear models, and by adding the nonlinear terms to the models (square and cube), even though we have noticed a very small decrease in the error, and that doesn't prove the linearity of the data.

The identified parameters have small values which show that the linear model is not fitted with the data.

To conclude the experimental data are nonlinear and to model it, a nonlinear identification technique is required.

5. Conclusion:

To conclude, this chapter has demonstrated the Least Square approach used for many heat diffusion applications (homogenous and heterogeneous examples) and for the experimental data as well. The least square approach has been useful for three different difference equations. The homogenous application has satisfied the linearity. While the heterogeneous application has been proven to be nonlinear.

The third part of the simulations was dealing with real data obtained from a rotary kiln. Using the least square approach, the results have shown that the data are nonlinear.

The obtained results show the effectiveness, sensitivity and robustness of the proposed approach.

Conclusion

In this thesis, an identification of distributed parameters systems (**DPS)** approach in a cement rotary kiln is presented. The rotary kiln demonstrates the heat transfer phenomena through its tube.

Using the Least Square approach we have built accurate and simplified models of distributed parameter systems, from numerical applications to time-series data obtained from the rotary kiln. The numerical solutions in the examples depend on the associated Neumann Boundary conditions and initial conditions. The measure of performance for the identification was the error (mean square error) between the model and the system. It has shown small values confirming the accuracy of the models.

Further Researches:

We had an immense interest to go further with distributed parameters systems and identify their parameters using other estimators such as variants of least square, Maximum likelihood…etc.

References

[1] D. Chinarro, *System Engineering Applied to Fuenmayor Karst Aquifer (San Julián de Banzo, Huesca) and Collins Glacier (King George Island, Antarctica)*, Springer. 2014

[2] Lennart Ljung, ‘’System Identification Toolbox user’s guide’’, The MathWorks, Inc.1988-1997

[3] Yi Liu ‘’Grey-box Identification of Distributed Parameter Systems’’, Doctoral Thesis. Automatic Control. Department of Signals. Sensors and Systems. Royal Institute of Technology. Stockholm. Sweden 2005

[4] F. van der Heijden, R.P.W. Duin , D Ridder , D.M.J. Tax, ‘’Classification, Parameter Estimation and State Estimation’’ John Wiley & Sons,Ltd 2004

[5] Lennart Ljung, ‘’System Identification Theory for the user ’’ University of Linkoping Sweden.1987

[6] Donoho D, Johnstone I (1994) Ideal spatial adaptation by wavelet shrinkage. Biomefrika 81:425–455

[7] Donoho D (1995) De-Noising by soft-thresholding. IEEE Trans Inf Theory 41(3):613–627

[8] Bendat J (1990) Nonlinear system analysis and identification. Willey, New York

[9] Zhang Q (1997) Using wavelet network in nonparametric estimation. IEEE Trans Neural Network 8(2):227–236

[10] J. LINDFORS, "Advances in control education 2003", University of Lapland, Finland, 2003

[11] Denis Matignon, Joseph Winkin, Denis Dochan "Control of distributed parameter system", University of Namur, Namur, Belguim,2007

[12] P. N. Paraskevopoulos and A. C. Bounas. Distributed parameter system identification via walsh functions. *Int. J. Systems Science,* 9(l):75-83, January 1978.

[13] S. G. Tzafestas. Walsh series approach to lumped and distributed system identification. *J. Franklin Institute,* 305(4):199-220, April 1978.

[14] Y. Sunahara, ‘’Identification of distributed parameter systems’’ Kyoto Institute of Technology, Matsugasaki, Kyoto, Japan.

[15] Stefan Bilbao and Julius O. Smith III ,``Discrete-Time Lumped Models", (From Lecture Overheads) Stanford University.

[16] John, F. (1982), Partial Differential Equations (4th ed.), New York: Springer-Verlag, ISBN 0-387-90609-6.

[17] A. D. Polyanin, Handbook of Linear Partial Differential Equations for Engineers and Scientists, Chapman & Hall/CRC Press, 2002.