Chapter 1

INTRODUCTION

Chapter 1 introduces the basic concepts of nonlinear system identification/modeling, the current status of the issue, motivation for the current work, objectives and methodologies adopted organization and outline of the thesis etc.
1.1 System Identification

Identification and Control of Non-linear dynamical systems are challenging problems to the control engineers. The problem of system identification and modeling consists of computing a suitably parameterized model, representing a process [1, 2, 3]. The parameters of the model are adjusted to optimize a performance function, based on error between the given process output and identified process/model output. Most of the real world systems are nonlinear in nature and wide applications are there for nonlinear system identification/modeling. The linear system identification field is well established with many classical approaches whereas most of those methods cannot be applied for nonlinear system identification [4, 5]. The problem becomes more complex if the system is completely unknown but only the output time series is available. The thesis concentrates on such problems. Capability of Artificial Neural Networks to approximate all linear and nonlinear input-output maps makes it predominantly suitable for the identification of nonlinear systems, where only the time series is available [7-13]. Different algorithms are available to train the Neural Network model. A comprehensive study of the models using different algorithms and the comparison among them to choose the best technique is not yet available in
any of the published books or technical papers. This thesis is an attempt to develop and implement few of the well known and newly proposed algorithms, in the context of stochastic (where only time series is known) modeling of nonlinear systems, and to make a comparison to establish the relative merits and demerits. When the output time series alone is available, the process is also termed blind identification/modeling [33-36].

Two basic types of modeling problems arise. In the first type, one can associate with each physical phenomenon, a small number of measurable causes (inputs) and a small number of measurable effects (outputs). The outputs and the inputs can generally be related through a set of mathematical equations, in most cases nonlinear partial differential equations. The determination of these equations is the problem of modeling in such cases. These can be obtained either by writing a set of equilibrium equations based on mass and energy balance and other physical laws, or one may use the black box approach which may consists of determining the equations from the past records of the inputs and outputs. Modeling problems of this type appear quite often in engineering practice. Some typical problems are modeling of (i) a stirred – tank chemical reactor, (ii) a multi machine electrical power system, (iii) a synchronous orbit communications satellite and (iv) the control mechanism of a nuclear power reactor [62-64]. In each of these examples one can easily identify certain
input and output quantities, and then obtain mathematical model relating them.

Another type of modeling problem arises in those situations where although it is possible to identify a certain quantity as a definite measurable output or effect, the causes are not so well defined. Some typical examples are (i) the annual population of a country, (ii) the annual rainfall in a certain country, (iii) the average annual flow of a river, and (iv) the daily value of a certain stock in the stock market. In all these cases, one have a sequence of outputs, which will be called a time series, but the inputs or causes are numerous and not quite known in addition to often being unobservable. The models in such cases are called stochastic models, due to a certain amount of uncertainty which is unavoidable [32, 33].

1.1.1 System description

A system can be described by one of the following.

- A transfer function
- A linear differential equation with constant coefficient that relates the input and output of the system.
- An impulse response.
- A set of state equations.

By knowing the input of the system, one can determine the response of the system. But in many cases one may not be having the system description.
system transfer function, impulse response, differential equation; state equation etc has to be derived from a sample of input and output [13-14].

Another type of modeling problem arise in those situation where one can identify a certain quantity as a definite measurable output or effect, the causes are not well defined. This is called time series modeling, where inputs or causes are numerous and not quite known in addition to often being unobservable. This type of modeling is also called stochastic modeling. System identification is concerned with the determination of the system models from records of system operation. The problem can be represented diagrammatically as below.

where \( x(t) \) is the known input vector of dimension ‘m’
\( z(t) \) is the output vector of dimension ‘p’
\( \omega(t) \) is the input disturbance vector

---

Fig. 1.1: A general system configuration
v(t) is the measured output vector of dimension ‘p’

Thus the problem of system identification is the determination of the system model from records of x(t) and y(t).

1.1.2 System identification using neural networks

For linear systems system identification and control are well developed. For non-linear systems the theory is not well defined significantly. Properties such as controllability, observability and stability are well defined for linear system model, but it is not straightforward in the case of non-linear systems.

Artificial Neural networks are a powerful tool for many complex applications such as function approximation, optimization, nonlinear system identification and pattern recognition. This is because of its attributes like massive parallelism, adaptability, robustness and the inherent capability to
handle nonlinear system. It can extract information from heavy noisy corrupted signals. Fig. 1.2 shows the model of a nonlinear neuron. System identification can be either state space model or input-output model.

1.1.3 The Input-Output modeling

An I/O model can be expressed as \( y(t) = g(\phi(t, \theta)) + e(t) \), where, \( \theta \) is the vector containing adjustable parameters which in the case of neural network are known as weights, \( g \) is the function realized by neural network and \( \phi \) is the regression vector. Depends on the choice of regression vector different model structures emerge.

Using the same regressors as for the linear models, a corresponding family of nonlinear models was obtained which are named NARX, NARMAX as in equations 1.1 and 1.2 below. Different model structures in each model family can be obtained by making a different assumption about noise.

\[
\text{NARX, } \phi(t, \theta) = \left[ y(t-1), y(t-2), \ldots, y(t-n), u(t-1), \ldots, u(t-m) \right]^{T} \quad (1.1)
\]

\[
\text{NARMAX } \phi(t, \theta) = \left[ y(t-1), \ldots, y(t-n), u(t-1), \ldots, u(t-m), e(t-1), \ldots, e(t-k) \right]^{T} \quad (1.2)
\]

Where \( y(t) \) is the output, \( u(t) \), the input and \( e(t) \) is the error. For the implementation of the above system, Feed forward neural networks can be used [19-21].
1.1.4 State Space modeling

Suppose that the given plant is described by state space model.

\[ x(n + 1) = f(x(n), u(n)) \]  
(1.3)

\[ y(n) = h(x(n)) \]  
(1.4)

where \( f(.) \) and \( h(.) \) are vector valued nonlinear functions both of which are unknown. \( x(n) \) and \( y(n) \) are the models estimate of the plant state and output at time step \( n \). For the implementation of the above state space equations, recurrent neural networks are used. i.e. a single RNN is used to model both process nonlinearity \( f \) and measurement function \( g \). Also the model incorporates the past residual in the regression [12, 79-82]. This structure is called Neural network State Space Innovation Function (NNSSIF). State space analysis characterizes dynamics of a system in terms of attractors, geometric description of recurrent trajectories and Lyapunov exponents [130].

1.2. Current status

Many researchers have addressed the problem for dynamic nonlinear black box modeling. Different approaches can be used for solving the problem. Among them Artificial Neural Networks is a powerful tool. The system identification then goes down to estimation of the model parameters. Neural network is best suited where unknown dynamics can be constructively approximated. During the past few years, several authors have suggested
neural network implementation for nonlinear dynamical black box modeling [19, 20, 78]. When the mathematical model of the process cannot be derived with an analytical method, the only way for modeling is by deriving the model function using the relationship between input and output of the process. In modeling, a neural network that emulates the behavior of the plant is trained based on the known nonlinear models [9, 11, 14]. Thus dynamical system information is stored in the neural network function. During modeling simulations, the input-output behavior of the neural network is compared to that of the nonlinear plant under study.

Neural network Black Box modeling can be performed using nonlinear Feed Forward (FF) and Recurrent structures. Recurrent Neural Networks (RNN) is fundamentally different from the feed forward architecture, in the sense that they not only operate in the input space but also in the internal state space. Because of the dynamical structure exhibited by them, these networks have been successfully applied to system characterization problems [19, 80, 82].

The classical approach of training neural network is by using the Back Propagation algorithm. Back propagation was created by generalizing the Windrow-Hoff learning rule to multiplayer networks [61] and has been widely used to train neural networks in many applications. Standard back propagation is a gradient descent algorithm. However the convergence could
be slow and appropriate learning parameters need to be chosen; their tuning is not trivial.

Since the development of well-known Kalman filter (KF) [92, 93, 94], the method of linear stochastic state estimation has been widely studied in the literature and applied to many problems in tracking. The Kalman Filter has been extended to the nonlinear systems, which linearises the nonlinear function around the point of interest. The resultant filter is called Extended Kalman Filtering (EKF), which can be implemented in estimating the network parameters in both FF and RNN. The estimation algorithm converges faster than the back propagation algorithms [95, 96]. Also the predictor – corrector approach helps to reduce the computational requirements. Many alternative approaches have been proposed for realizing the Kalman estimation like Decoupled EKF and Unscented Kalman Filter [101]. Computational complexity is quite low when the Decoupled EKF [112] is used.

Expectation Maximization Algorithm (EM) is a method to calculate the initial states and covariance avoiding the difficulty in setting proper values for these by trial and error [113]. Maximum Likelihood Estimation (MLE) is a well established procedure for statistical estimation. In this procedure first formulate a log likelihood function and then optimize it with respect to the parameter vector of the probabilistic model under consideration [114-117].
In classical approaches the search for the optimal approximation model is carried out within a parameterized identification family such as Moving average(MA), Auto Regressive(AR) and their combination (ARMA) or ARMAX (X for exogenous) [21, 68] and it is chosen to optimize a given cost function(e.g. Mean square error). Because of its simplicity linear models does not always approximate a nonlinear system throughout its working environment. Therefore to improve approximation accuracy various solutions have been envisaged which generally encompass system linearization around the working environment. Obviously, difficulties increases when the system is completely unknown, is considered to be the black box models.

In fact, the nonlinear parametric family obtainable with neural structures extends the linear ones by nonlinear models, among them are NAR, NARX, NARMAX subfamilies. Neural networks of the multi layer feed forward and recurrent types are employed for system identification. There are different structures and several algorithms for training neural networks for achieving global minima and the selection of these depends upon the problem one have to analyze. There is a wide gap between applications of these methods in real time and simulation. Issues such as stability, processor speed, learning time, type of algorithm etc arise when it comes to real time implementations. Adaptive designs of neural network are capable of optimization over time.
under conditions of noises and uncertainty.

A large number of literatures and published papers are available for the different techniques of system identification discussed so far. But a cumulative study of all the techniques together and comparative analysis is yet to come. Here in this Thesis, few important techniques are implemented and compared for system identification especially for stochastic modeling of nonlinear systems.

Recently several new approaches to recursive nonlinear filtering have appeared in literature. Particle filters (PF) are suboptimal filters belonging to this category of methods. They perform Sequential Monte Carlo (SMC) estimation based on point mass (or “particle”) representation of probability densities [131-137]. The SMC ideas in the form of sequential importance sampling had been introduced in statistic back in the 1950s. Although these ideas continued to be explored sporadically during the 1960s and 1970s, they were largely overlooked and ignored. Most likely the reason for this was the modest computational power available at that time. In addition, all these early implementations were based on plain sequential importance sampling, which as we shall describe later, degenerates over time. The major contribution to the development of the SMC method was the inclusion of the re-sampling step, which, coupled with the faster computers, made the particle filters useful in practice for the first time. Since then research
activity in the field has dramatically increased, resulting in many improvements of particle filters and their numerous applications especially for nonlinear system modeling [77].

1.3 Motivation

The problem of system modeling and identification has attracted considerable attention during the past few years mostly because of a large number of applications in diverse fields like chemical processes, biomedical systems, transportation, ecology, electric power systems, hydrology, aeronautics and astronautics. An accurate on-line estimate of critical system states and parameters are needed in a variety of engineering applications like in automatic control, signal processing, echo cancellation, SONAR, fault detection, tracking etc. They are used in many commercial products such as modems, image processing, speech recognition, front end signal processors and biomedical instrumentation [62-65].

The amazing challenges in statistical estimation along with an opportunity to learn different techniques in solving the well known problem motivated to take up the study of system identification technique. The rich literature available on the subject offered an opportunity to dig out solutions in situations that are difficult. Since a comprehensive study of the well known techniques and the comparison of their performance is necessary to choose
an efficient technique for particular applications. It is attempted to develop some new approaches and their evaluations based on various criterions for blind identification of nonlinear systems. It is expected that, such comprehensive study and the comparison process can be of great relevance in many fields including control, chemical, electrical, biological, financial and weather data analysis. More specifically the aim of the thesis is to:

- Implement various identification/modeling techniques for nonlinear systems.
- Develop and suggest certain new approaches for the blind identification of nonlinear system and improve some of the currently available techniques.
- Provide a comprehensive evaluation report of these methods based on a number of evaluation criterion/performance measures.

### 1.4 Objectives and the methodologies

The system identification process using neural network can be represented by the block diagram shown in Fig 1.3. The objective is to implement the following algorithms for nonlinear system identification and compare the performance of the models in order to evaluate the relative merits and demerits of the algorithms.

- Back Propagation (gradient – descent)
- Radial Basis Function networks (gradient – descent)
• Extended Kalman Filter
• Extended Kalman Filter with Expectation Maximization.
• Decoupled Extended Kalman Filter
• Maximum Likelihood Estimation
  Gauss Newton
  Conjugate Gradient

• Identification with particle filter approach
• State space modeling

Given below in Fig. 1.3 is an illustration of system identification.

![Block diagram of system identification using neural network](image)

The state space modeling is done to extract the dynamics of the system which is very helpful in the error detection and control of the plant or process. The model behavior and performance are evaluated in terms of Mean Square Error and also in terms of two well known methods (i) Lyapunov exponents
Development and Evaluation of Blind Identification Techniques for Nonlinear Systems

(for stability check) and (ii) Cramer Rao Lower Bound (CRLB) (for efficiency check). The statistical parameter estimation insists that the estimate should be well within the CRLB [121-124].

![Diagram of NARX model](image)

**MLFFN**

Fig 1.4 NARX modeling for system identification

NARX model is well suited for Input-Output modeling of stochastic nonlinear systems [39]. So in this work, NARX model is chosen as the system model in which the model structure is a Multi Layer Feed Forward Neural Network (MLFFN) as shown in Fig. 1.4 for all the nonlinear systems (using different algorithms).

Many nonlinear systems are modeled using each of the algorithms. Four entirely different systems are selected in order to check the consistency in performance of the algorithms. If the model performs equally well for all the
four systems it is assumed to perform well for any other nonlinear systems. The selected nonlinear systems are.

\[ y = \sin(t^2 + t) \]  

Real world systems: Ambient noise in the sea  
Acoustic source ‘A’  
Acoustic source ‘B’

1.5. Organization of the thesis

An introductory review of the available literature is given in chapter 2. Chapter 3 introduces the Neural Network approach using Back Propagation algorithm to estimate the parameters. Due to the local minima problem of BPA, an alternate approach based on Kalman Estimation is explored in chapter 4. Though Kalman Estimation is found good for estimation, the optimality depends on the apriori statistics of states and covariance. To eliminate this problem, the method based on Expectation Maximization is used which is also discussed in chapter 4. The stochastic method based on Maximum Likelihood Estimation is often described as a very standard approach in parameter estimation. Chapter 5 discusses about MLE. In chapter 6 a novel approach for the identification problem with nonlinear filtering method, namely particle filter, has been presented. In order to make the study of system identification problem comprehensive, the state space modeling approach has also been taken up to assess the dynamic behavior of
the systems as discussed in chapter 7. The efficacy of the model is demonstrated by plotting the phase plane plots for the systems identified. The Lyapunov exponents are calculated for the models in order to evaluate the convergence nature of the systems which is also included in chapter 7. Since the recommended procedure in the statistical parameter estimation insists that the estimate should be well within the CRLB, it is evaluated in chapter 8 for all the systems modeled in previous chapters. Chapter 9 includes the comparison of performance of different approaches along with their relative merits of implementation and it also summarizes the thesis with discussions, conclusion and the scope for future work.