Kwame Nkrumah University Of Science And Technology

State and Parameter Estimation Using Unscented Kalman Filter

by

Andrew Yeboah Nortey (BSc. Hons)

A Thesis Submitted To Department Of Mathematics

In Partial Fulfilment Of The Requirements For The Degree

Of

Master Of Philosophy

College of Science

December 2012

Declaration

I hereby declare that this submission is my own work towards the Master of Philosophy (MPhil.) and that, to the best of my knowledge, it contains no material previously published by another person nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in the text.

Andrew Yeboah Nortey	....	....
Student	Signature	Date

Certified by:

E. Owusu-Ansah	....	....
Supervisor	Signature	Date

Certified by:

K. F. Darkwah (Mr.)	....	....
Head of Department	Signature	Date

Abstract

The Extended Kalman Filter is a bayesian state estimation used for nonlinear models or systems. This filter may however fail to produce accurate results depending on the degree of nonlinearity of the system. The Unscented Kalman Filter on the other hand can be applied to highly nonlinear systems or models. Comparisons between the two filters are made using two systems. The first system is a four degrees of freedom shear building with time-varying system parameters and the second is a nonlinear hysteric damping system with unknown system parameters. The results indicated that the latter provides consistent as well as more accurate state and parameter estimates than the extended kalman filter for nonlinear systems.

Contents

Declaration ii

Abstract iii

List of Tables vii

List of Figures viii

Acknowledgements x

1 Introduction 1

1.1 Background 2

1.2 Problem Statement 3

1.3 Objectives 3

1.4 Methodology 3

1.5 Justification of Problem 4

1.6 Structure of the Thesis 4

2 Literature Review 6

2.1 Introduction 6

2.2 Data Assimilation 7

2.3 State Estimation in Civil Engineering 10

3 Methodology **16**

3.1 Introduction . . . 16

3.2 Numerical Methods for Ordinary Differential Equations (ODE) . . . 16

3.2.1 Introduction . . . 16

3.2.2 The Explicit Euler Method . . . 19

3.2.3 The Implicit Euler Method . . . 20

3.2.4 Consistency, Order, Stability and Convergence . . . 21

3.2.5 Consistency and Order . . . 22

3.2.6 Consistency and Order Analysis for the Explicit Euler FDE 22

3.2.7 Stability Analysis . . . 23

3.2.8 Second-Order Single-Point Methods . . . 24

3.2.9 Runge-Kutta Methods . . . 25

3.3 Data Assimilation . . . 27

3.3.1 Introduction . . . 27

3.3.2 Types of Data Assimilation . . . 28

3.3.3 Sequential Data Assimilation . . . 29

3.3.4 Variational Data Assimilation . . . 29

3.3.5 Kalman Filter . . . 30

3.3.6 The Process to be Estimated . . . 30

3.3.7 The Computational Essence of the Filter . . . 31

3.3.8 The Probabilistic Essence of the Filter . . . 33

3.3.9 The Discrete Kalman Filter Algorithm . . . 33

3.3.10 The Extended Kalman Filter . . . 37

3.3.11 The Unscented Kalman Filter . . . 40

3.3.12 Unscented Transformation . . . 40

3.4 Numerical Example of Data Assimilation Methods in State Estimation 41

4 Results and Discussion **49**

4.1 Introduction . . . 49

4.2 Planer Four-Story Shear Building . . . 49

4.2.1 Data . . . 49

4.2.2 Results . . . 54

4.3 Nonlinear SDOF System . . . 57

4.3.1 Data . . . 57

4.3.2 Results . . . 61

5 Conclusion and Recommendation **65**

5.1 Conclusion . . . 65

5.2 Recommendation . . . 66

References **67**

List of Tables

3.1 Mean and Variance of RMSE for state estimate of the two filters . . 45

3.2 Compare effect of different process noise variance Q on the performance of the various filters holding other parameters fixed 45

3.3 Compare effect of different observation noise R on the performance of the various filters holding other parameters fixed 46

3.4 Compare effect of different initial error covariance P_0 on the performance of the various filters holding other parameters fixed 47

3.5 Compare effect of different initial error covariance P_0 further on the performance of the various filters holding other parameters fixed . . 47

4.1 Mean of the RMSE of the states and parameters 64

List of Figures

3.1 The ongoing discrete Kalman filter cycle. The *time update* projects the current state estimate ahead in time. The *measurement update* adjusts the projected estimate by an actual measurement in time (source: Welch and Bishop (2006)) . . . 34

3.2 A complete picture of the operation of the Kalman filter (source: Welch and Bishop (2006)) . . . 36

3.3 Time series of the noisy observations and the true state for the first experiment . . . 44

3.4 EKF and UKF estimates for the first experiment . . . 44

3.5 RMSE of EKF and UKF for the first experiment . . . 45

4.1 Time series of the inter-story stiffness and viscous damping coefficients 50

4.2 Time series of the simulated observation and excitation data . . . 52

4.3 Estimation of displacements x_t . . . 54

4.4 Estimation of velocities $\dot{x}_t$ . . . 55

4.5 Estimation of inter-story stiffness k_t . . . 56

4.6 Estimation of damping coefficients c_t . . . 57

4.7 RMSE of EKF and UKF for x_t and $\dot{x}_t$ . . . 58

4.8 RMSE of EKF and UKF for k_t and c_t . . . 59

4.9 Time series of simulated acceleration and input force . . . 59

4.10 Estimation of the states of the SDOF system 62
4.11 Estimation of the parameters of the SDOF system 62
4.12 Estimation of the uncertain parameter h_t 63
4.13 Time propagation of the RMSE of EKF and UKF 63

Acknowledgements

I cannot thank anybody without first expressing my greatest appreciation to the Almighty God for my life and how far He has brought me as far as my education is concerned. I also want to thank my supervisor, E. Owusu-Ansah for his time, help and assistance from the very first moment I met him up until now. I also want to use this opportunity to thank Jianye Ching for providing me with matlab codes that aided in my simulations. To Dr. P. Amoako-Yirenkyi, I say God bless you for your continued assistance and guidance.

I am also grateful to my parents for their love and support shown to me throughout these years. God richly bless you. I also want to thank the lecturers at the Department of Mathematics, KNUST for the knowledge imparted in me and also to my course mates and friends who in one way or the other helped in the completion of this thesis.

Chapter 1

Introduction

State and parameter estimation has been a subject of research for over four decades. One of the ways of simultaneously estimating the state and parameters of dynamical systems is Data Assimilation (DA), where observations from measurements are combined with the dynamic principles (model) governing the system under discussion. It also adds value to the observations by filling in the observational gaps and to the model by constraining it with observations.

There are two main categories of data assimilation, the Sequential Data Assimilation where observations are incorporated into the model at each time they are available and a best estimate is produced and used to predict future states and the Variational Data Assimilation which seeks an optimal fit of the model solution to observations over a period by adjusting the estimation states in this period simultaneously and also the estimated states are somehow influenced by all the observations distributed in time. The Kalman filter (KF) which is a sequential data assimilation method is believed to be an optimal technique for estimating linear dynamical systems. This filter is basically a set of mathematical equations used to estimate the state of a dynamical system and at the same time minimizing the mean of the squared error.

Most real-life problems are in fact nonlinear and so the Kalman filter is unable to predict or estimate the state of these nonlinear systems or models accurately. An improvement of the Kalman filter known as the Extended Kalman filter (EKF) is the most popular method used to estimate the state as well as parameters of nonlinear systems. The Extended Kalman Filter can sometimes be very difficult to implement because of the calculation of Jacobians used for the linearizaiton and also the filter sometimes diverge if the dynamical model is highly nonlinear. The Unscented Kalman filer (UKF) and its variations is an alternative method to the EKF for the estimation of nonlinear dynamical systems.

1.1 Background

Knowledge of the dynamic loads is crucial to design purposes as far as civil engineering is concerned. Very often these dynamic loads are not well known or cannot be measured directly. In these cases, inverse identification techniques may be used for identifying the unknown excitation forces from the measured responses.

State as well as parameter estimations have been applied in civil engineering for a few decades. For instance, it can be used in structural health monitoring to detect changes of dynamical properties of structural systems during earthquakes and, more generally, it can be used for system identification to better understand the nonlinear behavior of structures subject to seismic loading (earthquake). The ability to estimate system states in real time may help to accomplish an efficient control strategy as considered in structural control. Again in performance-based earthquake engineering, state estimation can provide some crucial information to assess the seismic performance of an instrumented structure in terms of repair costs, casualties and repair duration shortly after the cessation of strong motion (e.g. after an earthquake). Various studies have been carried out because of the wide applicability of state and parameter estimation in civil engineering.

1.2 Problem Statement

The Unscented Kalman Filter is an alternative to the Extended Kalman Filter for estimating nonlinear system states and parameters which involves the use of sigma points and the nonlinear function to obtain the required statistics (mean and covariance).

The performance and accuracy of this filter is investigated using two structural systems. The first is a planar four degrees of freedom shear building with time varying system parameters and the second is a single degree-of-freedom (SDOF) Bouc-Wen hysteretic damping system.

1.3 Objectives

The objectives of this study are to

- Test the feasibility of the Unscented Kalman filter in simultaneously estimating state and parameters.
- Compare the performance of the Unscented and Extended Kalman filters on two structural systems.

1.4 Methodology

This research involves the state and parameter estimation of two structural dynamical systems (a linear system and a nonlinear system) using the Unscented Kalman filter. The first system is a planar four-story shear building with time varying system parameters and the second system is a single degree-of-freedom

(SDOF) damping system. The models used for this two systems are initially differential equations and as a result numerical methods for solving ordinary differential equations are also used in this research.

The simulations for this research were performed using MATLAB (version 7.8.0.347, R2009a). It is also appropriate to mention that, the two structural systems considered in this thesis were obtained from Ching et al. (2006) and the matlab codes used with some modifications were obtained from Jianye Ching. Information as well as references needed for this research were obtained from the the KNUST library and the internet.

1.5 Justification of Problem

Most researchers in the Civil and Structural Engineering fields have studied the state and parameter estimation using the the popular Extended Kalman filter due to the fact that most of these estimation problems are nonlinear. On the other hand, a lot of work has been carried out in other scientific areas using the Unscented Kalman filter and its variations. Much work has not not be done in civil engineering using the Unscented Kalman filter (most people used the well-known Extended Kalman filter). This research focuses on the application of the Unscented Kalman filter in civil engineering which has not seen much work there.

1.6 Structure of the Thesis

This thesis is written in five chapters. The first chapter introduces the thesis. That is, the general overview of state and parameter estimation and its application in structural engineering. It also talks about the problem statement, objectives of the thesis, methods used and then the justification of the problem. Chapter two deals with the review of the literature related to the thesis followed by the third chapter which focuses on the methods used in the thesis as well as a numerical example. Chapter four talks about the results and discussion and the final chapter deals with the conclusion and recommendations drawn from the thesis.

Chapter 2

Literature Review

2.1 Introduction

The main challenges to society, for example, climate change, impact of extreme weather, environmental degradation and ozone loss, require information for an intelligent response, including making choices on future action. Regardless of its source, we wish to be able to use this information to make predictions for the future, test hypotheses, and attribute cause and effect. In this way, we are able to take action according to information provided on the future behaviour of the system of interest, and in particular future events (prediction); test our understanding of the system, and adjust this understanding according to new information (hypothesis testing); and understand the cause of events, and obtain information on possible ways of changing, mitigating or adjusting to the course of events (attribute cause and effect)(Lahoz et al., 2010).

However, we still need two ingredients: a means of gathering information, and methods to build on this information gathered. Roughly speaking, *observations* (measurements) provide the first ingredient, and *models* (conceptual, numerical or otherwise) provide the second ingredient. Lahoz et al. (2010) went on to ask a very important question. “What combination of the model and observation information is said to be optimal?”. Mathematics provides an answer to this question as well as an estimate of the errors of the “optimum” or “best” estimate. This method is known as **data assimilation**. The data assimilation adds value to the observations by filling in the observational gaps and to the model by constraining it with observations. Accurate inferences are then made from the observations. Bouttier and Courtier (1999) also described data assimilation as an analysis technique in which observed information is accumulated into the model state by taking advantage of consistency constraints with laws of time evolution and physical properties.

2.2 Data Assimilation

There are two main categories of data assimilation, namely Sequential data assimilation which considers observation made in the past until the time of analysis, which is the case of real-time assimilation systems and Variational data assimilation where observation from the future can be used, for instance in a reanalysis exercise (Bouttier and Courtier, 1999).

This research only deals with the sequential data assimilation. The most well known Sequential data assimilation or bayesian state estimation algorithm is the Kalman filter (KF) named after Rudolph E. Kalman who published is famous paper in 1960 describing a recursive solution to the discrete-data linear filtering problem (Kalman, 1960; Kalman and Bucy, 1961).

Welch and Bishop (2006) describes Kalman filter as a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process, in a way that minimizes the mean of the squared error. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown.

The Kalman filter however has limitations, one of such is the fact that it is unable to predict or estimate the state vector of *nonlinear* model. That is, the Kalman filter breaks down when applied to nonlinear systems. Most of real world problems are nonlinear systems and so the Kalman filter is unable to predict the state or parameters or even both for such systems (Julier et al., 1995; Julier and Uhlmann, 1996, 1997; van der Merwe et al., 2000; van der Merwe and Wan, 2003; Ching et al., 2006).

Fortunately, there are improved forms of the Kalman filter used for nonlinear state equations. The famous of such algorithms is the *Extended Kalman Filter* (Jazwinski, 1970) which extends the kalman filtering through a procedure of linearization by making use of the taylor series expansion of the the nonlinear model. Once a linear model is obtained the kalman filter equations are then applied to obtain the estimate. This filter has become a standard technique used in state and parameter estimation.

Again the Extended Kalman Filter has a couple of limitations. One of such is the fact that it is not always possible to calculate the *Jacobian* and also it does not produce accurate results if the nonlinear function is not well approximated by a linear model or function. That is, this model works well when the nonlinear model is close to a linear model.

Julier et al. (1995) pointed out that although the EKF is conceptually simple it has, in practice, three well-known drawbacks:

1. Linearization can produce highly unstable filter performance if the time step intervals are not sufficiently small.

2. The derivation of the Jacobian matrices are nontrivial in most applications and often lead to significant implementation difficulties and finally,

3. Sufficiently small time step intervals usually imply high computational overhead as the number of calculations demanded for the generation of the Jacobian and the predictions of state estimate and covariance are large.

Although the EKF (in its many forms) is a widely used filtering strategy, over thirty years of experience with it has led to a general consensus within the tracking and control community that it is difficult to implement, difficult to tune, and only reliable for systems which are almost linear on the time scale of the update intervals and many of these difficulties arise from its use of linearization (Julier and Uhlmann, 1997; Julier et al., 2000; Julier and Uhlmann, 2004). A central and vital operation performed in the Kalman Filter is the propagation of a Gaussian random variable (GRV) through the system dynamics. In the EKF, the state distribution is approximated by a GRV, which is then propagated analytically through the first-order linearization of the nonlinear system. This can introduce large errors in the true posterior mean and covariance of the transformed GRV, which may lead to sub-optimal performance and sometimes divergence of the filter (Wan and van der Merwe, 2000).

Another improved form of the Kalman filter used for the estimation of nonlinear systems is the *Unscented Kalman Filter*. This method was first introduced by Julier et al. (1995); Julier and Uhlmann (1996, 1997) and further improved by Wan et al. (2000); Wan and van der Merwe (2000) which is obtained from the basic idea of the *Unscented Transform*.

This method is used for calculating the statistics (mean and covariance) of a random variable which undergoes a *nonlinear* transformation. It is based on the fact that, it is easier to approximate a Gaussian distribution than it is to approximate a nonlinear function. The idea of this transform is to obtain a set of points known as Sigma points so that the sample mean and sample covariance are $\bar{x}$ and P_x respectively. The nonlinear function is applied to these set of points to produce a set of transformed sigma points whose statistics are $\bar{y}$ and P_y (Julier and Uhlmann, 1997).

2.3 State Estimation in Civil Engineering

For civil engineering structures, knowledge of the dynamical loads is crucial to design purposes. Very often these dynamical loads are not well known or cannot be measured directly. In these cases, inverse identification techniques may be used for identifying the unknown excitation forces from the measured responses.

State as well as parameter estimations have been applied to civil engineering for some time now. For instance, it can be used in structural health monitoring to detect changes of dynamical properties of structural systems during earthquakes and, more generally, it can be used for system identification to better understand the nonlinear behavior of structures subject to seismic loading (earthquake). The ability to estimate system states in real time may help to accomplish an efficient control strategy is considered in structural control. Again in performance-based earthquake engineering, state estimation can provide some crucial information to assess the seismic performance of an instrumented structure in terms of repair costs, casualties and repair duration shortly after the cessation of strong motion(e.g. after an earthquake). Various studies have been carried out because of the wide applicability of state and parameter estimation in civil engineering.

Distefano and Rath (1975) reported the sequential identification methods of control and optimization theory and the way in which the parameters can be identified other than the natural circular frequency of a bilinear system. Since their approach is generally a sequential processing, the same problem can be accessed by the use of Kalman filter.

Carmichael (1979) incorporated the model parameter estimation by the extended Kalman filter within the state estimation problem by suitably augmenting the state vector of dynamic behavior of the model. However, his problem is based on a single degree-of-freedom system, and the adaptability of the reference state vector is not markedly stable due to the finite difference approximation.

Yun and Shinozuka (1980) used an extended Kalman filter to study nonlinear fluid-structure interaction and to identify the hydrodynamic coefficient matrices associated with nonlinear drag and linear inertia forces appearing in the equations of motion of offshore structures subjected to wave forces. They reported a multiple degree-of-freedom linear model of an offshore tower in which the response of each mass was observed.

Loh and Tsaur (1988) also applied the extended Kalman to a system identification problem of seismic structural systems which presented an identification method for an equivalent linear system, a bilinear hysteric restoring system and a bilinear hysteretic restoring system with stiffness degradation effect. For the accuracy of their proposal, the justification of the method was investigated on numerically simulated data on response of a known system as well as a known degrading system. It was then applied to identify the hysteresis behaviour of two buildings which were subjected to earthquake loads. They showed that good estimates of the test-response time history and reliable values of the structural parameters can be obtained.

In order to obtain the stable and convergent solutions (i.e. for stable estimation), Hoshiya and Saito (1984) proposed a weighted global iteration procedure with an objective function which was incorporated into the extended Kalman filter algorithm. For the effectiveness of the proposal, the identification problems were investigated for multiple degree-of-freedom linear systems, bilinear hysteretic systems, and equivalent linearization of bilinear hysteretic systems. The simulated results showed that the weighted global iteration procedure may be useful to identification problems for structural system identification.

It is generally recognized that damage will result in degradation of system characteristics, such as stiffness or damping. Iemura and Jennings (1973) and Udwadia and Jerath (1980) have observed that a degradation of the stiffness from 50% to 70% can occur, based on data obtained from the Millikan Library Building, located on the campus of the California Institute of Technology, during the 1971 San Fernando earthquake. It was found that the ratio of the time-dependent natural frequency to the original value was a smooth time-varying function, which dropped to about half its value during the strong motion part of the excitation. Several authors (Chen et al., 1977; Foutch and Housner, 1977; Meyer and Roufaiel, 1984; Mihai et al., 1980; Vasilescu and Diaconu, 1980) have also indicated that stiffness degrades in both full-scale structures and small-scale models as a result of seismic damage.

Ogawa and Abe (1980) and Carydis and Mouzakis (1986) attempted to correlate stiffness degradation to the severity of damage. Based on all these observations, structures under strong environmental loads are thus expected to undergo nonlinear and time-dependent degrading behavior. Hence, time-varying behavior of system parameters can occur, and on-line identification becomes a real issue under these conditions to permit real-time corrective action, repair, and control to minimize the possibility of further damage.

Lin et al. (1990) developed a general, real time-domain technique to identify the time-varying system parameters for better understanding of the degrading behavior of structures subject to dynamic loads. They tested their identification methodology on two numerical examples to demonstrate the use and efficiency of the method. The formulation of the identification procedure and the simulated results from the two numerical examples indicated that the proposed technique which identifies time-varying physical system parameters (e.g., stiffness, damping) rather than the modal parameters (e.g., modal stiffness, natural frequency), can be effective in detecting the damage of the structure as to its occurrence and location. Since current system state is updated at each identification time instant, no identification error is accumulated and for that matter, the knowledge of the initial system parameters is not so critical to the accuracy of their identification. The numerical results in the second example also demonstrated that if the response is sensitive to a particular system parameter, that parameter will be estimated with higher accuracy.

Hoshiya and Sutoh (1993) investigated the extended Kalman filter - weighted local iteration procedure with an objective function (EK-WLI procedure) for system identification in geotechnical engineering problems. They incorporated the EK-WLI procedure with the finite element method in order to identify unknown parameters. For the effectiveness of this procedure, parameter identification problems were numerically analyzed for an elastic plane strain field represented by the finite element models under several conditions. The results from numerical examples proved that a weighted local iteration procedure of Kalman filter with an objective function is found to be effective for stable estimation of state vector and also showed the usefulness of this method in parameter identification.

Koh and See (1994) developed an improved version of the commonly used extended Kalman filter (EKF) by incorporating an adaptive filter procedure. The system noise covariance is updated in time segments in order to ensure statistical consistency between the predicted error covariance and the mean square of actual residuals. Comprising two stages in a cycle, the adaptive EKF method not only identifies the parameter values but also gives a useful estimate of uncertainties. They tested the method on two numerical examples of simulation with noise. The first example illustrated the superior statistical performance of the proposed method over the conventional EKF method. The second example demonstrated the numerical accuracy and efficiency of this method, with and without modeling error, in comparison with the recursive least-squares approach.

Ghanem and Shinozuka (1995a) presented and reviewed several structural-identification algorithms which included the extended kalman filter, maximum likelihood technique, recursive least squares method and recursive instrumental variable method. They applied these algorithms to experimental data obtained in controlled laboratory conditions. The data pertained to the acceleration records from two building models subjected to various loading conditions. The performance of the various identification algorithms was critically assessed, and guidelines were obtained regarding their suitability to various engineering applications (Ghanem and Shinozuka, 1995b).

Glaser (1996) used the Kalman filter to identify the time-varying natural frequency and damping of a liquefied soil to get insight into the liquefaction phenomenon. They investigated the Wildlife Site in California, subject to two large earthquakes (Elmore Ranch and Superstition Hills) on November 24, 1987 as the associated data were the only publicly available record of buried and surface motions.

Sato and Qi (1998) derived an adaptive H_∞^+ filter and applied it to time-varying linear and nonlinear structural systems in which displacements and velocities of the floors are measured.

Smyth et al. (1999) formulated an adaptive least-squares algorithm for identifying multi-degree-of-freedom nonlinear hysteretic systems for control and monitoring.

Ching et al. (2004) presented a real-time Bayesian state-estimation algorithm that employs a stochastic simulation approach called the particle filter (PF) as well as introduced and discussed some techniques that improve the convergence of the particle filter simulations. They made comparisons between the particle filter and the extended Kalman filter using several numerical examples of nonlinear systems. The results indicated that the particle filter provides consistent state and parameter estimates for highly nonlinear systems, while the extended Kalman filter does not. They also applied the particle filter to strong motion data recorded in the 1994 Northridge earthquake in a seven-story hotel (The Van Nuys hotel) whose structural system consisted of non ductile reinforced-concrete moment frames, two of which were severely damaged during the earthquake. The particle filter once again provided consistent state and parameter estimates, in contrast to the extended Kalman filter, which provided inconsistent estimates. They concluded that for a state estimation procedure to be successful, at least two factors are essential: an appropriate estimation algorithm and a suitable identification model. Finally, recorded motions from the 1994 Northridge earthquake were used to illustrate how to do real-time performance evaluation by computing estimates of the repair costs and probability of component damage for the hotel.

During the review of literature, it was observed that the Unscented Kalman filter has not been applied much to civil or structural engineering. Most researchers rather used the Extended Kalman filter. It is of this view that the Unscented Kalman filter was applied to dynamical structural systems and the performance of this filter was compared with that of the Extended Kalman filter.

Chapter 3

Methodology

3.1 Introduction

This chapter discusses the methods used in this thesis. They are the Numerical Methods for Ordinary Differential Equations and Data Assimilation using Unscented Kalman Filter. The models used to estimate the state as well as the parameter of the system under discussion are ordinary differential equations. These equations are solved using numerical methods and then the unscented kalman filter is used to estimate the state and parameters of the system.

3.2 Numerical Methods for Ordinary Differential Equations (ODE)

3.2.1 Introduction

A lot of mathematical models that arise in many branches of science, engineering, economics, etc are differential equations. The equations unfortunately rarely have solutions which can be expressed in closed form or have analytical solutions, so it is common to seek approximate solutions by means of numerical methods. Nowadays this can usually be achieved very inexpensively to high accuracy and with a reliable bound on the error between the analytical solution and its numerical approximation.

An *Ordinary differential equation (ODE)* is an equation that specifies a relationship between a function of a single independent variable and the total derivatives of this function with respect to the independent variable or a differential equation in which the unknown function (also known as the dependent variable) is a function of a single independent variable. The variable y is actually used as a generic dependent variable is this thesis. The independent variable is either time t or space x in most problems in engineering and science. Usually, if more than one independent variable exists, then partial derivatives occur, and partial differential equations (PDE) are obtained which are not covered in this thesis.

The *order* of an ODE is the highest derivative of the dependent variable with respect to the independent variable appearing in the equation. The general first-order ODE is

$$\frac{dy}{dt} = f(t, y) \tag{3.1}$$

where $f(t, y)$ is called the *derivative function.* Also the general nth-order ODE for $y(t)$ has the form

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_2 y^{''} + a_1 y^{'} + a_0 y = F(t) \tag{3.2}$$

A *linear* ODE is one in which all of the derivatives appear in linear form and none of the coefficients depends on the dependent variable or an ODE in which there are no products of the unknown function $y(t)$ and its derivatives and neither the function or its derivatives occur to any power other than the first power. For example

$$y' + 2y = F(t) \tag{3.3}$$

is a linear, constant-coefficient first-order ODE but

$$y' + \alpha t y = F(t) \tag{3.4}$$

is a linear, variable-coefficient first-order ODE. If the coefficients depend on the dependent variable, or the derivatives have powers more than one, then the ODE is *nonlinear*. For example, the following equations;

$$yy' + \alpha y = 0 \tag{3.5}$$

$$(y')^2 + \alpha y = 0 \tag{3.6}$$

are *nonlinear* first-order ODEs.

A *homogeneous* differential equation is differential equation (or DE) in which each term involves only the dependent variable and one of its derivatives but a *nonhomogeneous* differential equation contains additional terms known as source, force or nonhomogeneous terms which do not involve the dependent variable. For example

$$y' + \alpha t y = 0 \tag{3.7}$$

is a linear, first-order homogeneous ODE and

$$y' + \alpha t y = F(t) \tag{3.8}$$

is a linear, first-order nonhomogeneous ODE with the nonhomogeneous term being $F(t)$.

Unfortunately, many practical problems are not as simple as the equations described above but rather involve several dependent variables, each of which is a function of the same single independent variable and one or more of the dependent variables, and each of which is governed by an ordinary differential equation. These sets of ODE are called *systems of differential equations*. The general solution of a differential equation is usually a family of solutions due to the fact that it contains one or more constants of integration. A particular member of that family which is of interest is obtained by auxiliary conditions. The number of auxiliary conditions must equal the number of constants of integration, which is the same as the order of the differential equation.

There are two distinct classes of ordinary differential equations depending on the type of auxiliary conditions specified. If all the auxiliary conditions are specified at the same value of the independent variable and the solution is to be marched forward from that initial point, the differential equation is an *initial-value* ODE but if the auxiliary conditions are specified at two different values of the independent variable, usually the end points or boundaries of the domain of interest, the differential equation is a *boundary-value* ODE.

The differential equations used in the thesis are initial-value ODEs so the subsequent subsections talks about some of the methods used in solving initial value ODEs.

3.2.2 The Explicit Euler Method

Consider the general nonlinear first-order ODE:

$$y' = f(t, y) \qquad y(t_0) = y_0 \tag{3.9}$$

Choosing point n as the base point and developing a finite difference approximation for eqn. (3.9), the first-order forward-difference finite difference approximation of y' is written as

$$y'|_n = \frac{y_{n+1} - y_n}{\Delta t} - \frac{1}{2}y''(\tau_n)\Delta t \tag{3.10}$$

Substituting eqn. (3.10) into eqn. (3.9), evaluating $f(t, y)$ at point n and finally solving for y_{n+1} which also includes truncating the remainder term, which is $O(\Delta t^2)$ (Hoffman, 2001) yields the explicit Euler finite difference equation (FDE):

$$y_{n+1} = y_n + \Delta t f_n \qquad O(\Delta t^2) \tag{3.11}$$

where $O(\Delta t^2)$ is the order of the local truncation error.

A few points to note about this equation is as follows:

- The FDE is clearly explicit, since the function f_n, does not depend on y_{n+1}.
- The FDE is a single point method since it requires only one know point as well as one derivative function evaluation per step.
- The error in calculating y_{n+1} for a single step, the local truncation error, is $O(\Delta t^2)$ whiles the global (i.e., total) error accumulated after N steps is $O(\Delta t)$.

The last point indicates that this FDE is of first order since the global error is of order one. Finally, the algorithm based on the repetitive application of the explicit Euler FDE to solve initial-value ODEs is called the *explicit Euler method*.

3.2.3 The Implicit Euler Method

Consider the general nonlinear first-order ODE:

$$y' = f(t, y) \qquad y(t_0) = y_0 \tag{3.12}$$

Choosing point n as the base point and developing a finite difference approximation for eqn. (3.12), the first-order backward-difference finite difference approximation of y' is given by

$$y'|_{n+1} = \frac{y_{n+1} - y_n}{\Delta t} - \frac{1}{2}y''(\tau_{n+1})\Delta t \tag{3.13}$$

Substituting eqn. (3.13) into eqn. (3.12), evaluating $f(t, y)$ at point $n+1$ and finally solving for y_{n+1} which also includes trancating the remainder term, which is $O(\Delta t^2)$ (Hoffman, 2001) yields the *implicit Euler FDE*:

$$y_{n+1} = y_n + \Delta t f_{n+1} \qquad O(\Delta t^2) \tag{3.14}$$

where $O(\Delta t^2)$ is the order of the local truncation error.

It can be observed that:

- Since f_{n+1} depends on y_{n+1}, it implies that the FDE above is implicit. If $f(t, y)$ is linear in y, then f_{n+1} is linear in y_{n+1}, and eqn. (3.14) is a linear FDE which can be solved directly for y_{n+1}. If $f(t, y)$ is nonlinear in y, then eqn. (3.14) is a nonlinear FDE, and additional effort is required to solve for y_{n+1}.

- Again the FDE is a single point method since it requires only one know point as well as one derivative function evaluation per step.

- Only one derivative function evaluation is needed per step if $f(t, y)$ is linear in y but several evaluations of the derivative function may be required to solve the nonlinear FDE if it is nonlinear and eqn. (3.14) is nonlinear for that matter.

- Similar to the explicit euler FDE, its implicit form also has a single truncation error of $O(\Delta t^2)$ and its global error is $O(\Delta t)$.

The algorithm based on the repetitive application of the implicit Euler FDE to solve initial-value ODEs is called the *implicit Euler method*.

3.2.4 Consistency, Order, Stability and Convergence

An FDE is *consistent* with an ODE if the difference between them (i.e., the truncation error) vanishes as $\Delta t \to 0$ or simply put, the FDE approaches the ODE. The rate at which the global error decreases as the grid size approaches zero is the *order* of the FDE. An FDE is *stable* if it produces a bounded solution for a stable ODE and is *unstable* if it produces an unbounded solution for a stable

ODE. A finite difference method is *convergent* if the numerical solution of the FDE approaches the exact solution of the ODE as $O(\Delta t)$ approaches zero.

3.2.5 Consistency and Order

Consistency analysis is done by changing all terms in the FDE to a Taylor series having the same base point as the FDE which results in an infinite-order differential equation known as the modified differential equation (MDE). Making $\Delta t \to 0$ in the MDE results in a finite-order differential equation and if this equation is identical to the exact differential equation whose solution is desired, then the FDE is *consistent*. The *order* of a FDE is the order of the lowest-order terms in the MDE.

3.2.6 Consistency and Order Analysis for the Explicit Euler FDE

Using this first-order ODE

$$y' + \alpha y = F(t) \tag{3.15}$$

The MDE is obtained by first substituting eqn. (3.15) into eqn. (3.11). The taylor series for y_{n+1} is also substituted into the resulting equation. Finally cancelling the y_n terms, dividing by Δt and rearranging terms yields the modified differential equation (MDE):

$$y'|_n + \alpha y_n = F_n - \frac{1}{2}\Delta t y''|_n - \frac{1}{6}(\Delta t)^2 y'''|_n - \cdots \tag{3.16}$$

Let $\Delta t \to 0$ in eqn. (3.16) it can easily be proved that the explicit euler FDE is consistent. The order of the FDE is the order of the lowest-order term in eqn.(3.16). This results in

$$y'|_n + \alpha y_n = F_n + O(h) + \cdots \tag{3.17}$$

The conclusion is that the explicit euler FDE is of order $O(h)$.

3.2.7 Stability Analysis

The exact solution of the FDE can be expressed as

$$y_{n+1} = G y_n \tag{3.18}$$

where G is known as the amplification factor of the FDE and its generally a complex number. The general solution for the FDE at $T = N\Delta t$ is

$$y_N = G^N y_0 \tag{3.19}$$

For y_n to remain bounded as $N \to \infty$

$$|G| \leq 1 \tag{3.20}$$

Therefore, stability analysis reduces to first of all determining the amplification factor G of the FDE and then determining the conditions to ensure that $|G| \leq 1$.

Consider the *explicit Euler method* in eqn. (3.11) the amplification factor G is given by

$$G = (1 - \alpha\Delta t)$$

For stability, $|G| \leq 1$ and so $-1 \leq (1 - \alpha\Delta t) \leq 1$. The left-hand inequality is only satisfied if $\Delta t \leq 2/\alpha$ and that of the right-hand side if $\Delta t \geq 0$. Therefore the explicit Euler method is *conditionally stable*.

Consider the *implicit Euler method* in eqn. (3.14) the amplification factor G is given by

$$G = \frac{1}{1 - \alpha\Delta t}$$

For stability, $|G| \leq 1$ is true for all values of $\alpha\Delta t$. Therefore the implicit Euler method is *unconditionally stable*.

3.2.8 Second-Order Single-Point Methods

Again consider eqn. (3.9), expressing y_{n+1} and y_n in Taylor series about the point $n+1/2$, subtracting the equation involving y_n from that of y_{n+1} and solving for $y|_{n+1/2}$ gives

$$y|_{n+1/2} = \frac{y_{n+1} - y_n}{\Delta t} - \frac{1}{24} y'''(\tau)\Delta t^2 \tag{3.21}$$

where $t_n \leq \tau \leq t_{n-1}$. Substituting eqn. (3.21) into eqn. (3.9) and solving for y_{n+1} results in the *implicit midpoint FDE*:

$$y_{n+1} = y_n + \Delta t f_{n+1/2} + O(\Delta t^3) \tag{3.22}$$

However y_{n+1} is obtained by first predicting $y_{n+1/2}$ using the first-order explicit Euler method and $f_{n+1/2}$ is evaluated using the value of $y_{n+1/2}$ obtained earlier. Therefore two equations are needed which are called the *modified midpoint FDEs* and given as:

$$y_{n+1/2} = y_n + \frac{\Delta t}{2} f_n \tag{3.23}$$

$$y_{n+1} = y_n + \Delta t f_{n+1/2} \tag{3.24}$$

The single-step FDE corresponding to the midpoint needed for the amplification factor G is given as

$$y_{n+1} = \left[1 - \alpha\Delta t + \frac{(\alpha\Delta t)^2}{2}\right] y_n$$

and so amplification factor is

$$G = 1 - \alpha\Delta t + \frac{(\alpha\Delta t)^2}{2}$$

. $|G| \leq 1$ if only $\alpha\Delta t \leq 2$

The algorithm based on the repetitive application of the modified midpoint FDEs is called the *modified midpoint method*. The modified midpoint method has the following features

- This method is an explicit predictor-corrector set of FDEs which requires two derivative function evaluations per step.

- This method is consistent, $O(\Delta t^3)$ locally and $O(\Delta t^2)$ globally.

- This method is conditionally state (i.e., $\alpha\Delta t \leq 2$)

- Finally, this method is convergent since its consistent and conditionally stable.

3.2.9 Runge-Kutta Methods

Runge-Kutta methods are a group of single-point methods which evaluate the difference between two solutions at points n and $n+1$ denoted by $\Delta y = (y_{n+1} - y_n)$ as the weighted sum of several Δy_i (for $i = 1, 2, \ldots$), where each Δy_i is evaluated as Δt multiplied by the derivative function $f(t, y)$, evaluated at some point in the range $t_n \leq t \leq t_{n+1}$, and the C_i (for $i = 1, 2, \ldots$) are the weighting factors. That is,

$$y_{n+1} = y_n + \sum_{i=1}^{v} C_i \Delta y_i \tag{3.25}$$

When Δy is a weighted of two Δy's, The second-order Runge-Kutta method is obtained

$$y_{n+1} = y_n + C_1 \Delta y_1 + C_2 \Delta y_2$$

where Δy_1 is given by the explicit Euler FDE

$$\Delta y_1 = \Delta t f(t_n, y_n) = \Delta t f_n \tag{3.26}$$

and Δy_2 is based on $f(t, y)$ evaluated somewhere in the interval $t_n \leq t \leq t_{n+1}$:

$$\Delta y_2 = \Delta t f[t_n + (\alpha \Delta t), y_n + (\beta \Delta y_1)] \tag{3.27}$$

Substituting Δy_1 and Δy_1 into y_{n+1} above and letting $\Delta t = h$ results in

$$y_{n+1} = y_n + C_1(hf_n) + C_2 hf[t_n + (\alpha h), y_n + (\beta \Delta y_1)] \tag{3.28}$$

There are infinite number of possibilities depending on the values of C_1, C_2, α and β. The modified Euler method is obtained if $C_1 = \frac{1}{2}$, $C_2 = \frac{1}{2}$, $\alpha = 1$ and $\beta = 1$. Assuming $C_1 = 0$, $C_2 = 1$, $\alpha = \frac{1}{2}$ and $\beta = \frac{1}{2}$ yields the *modified midpoint method*.

If $k_i = \Delta y_i$, then the *Second-Order Runge-Kutta* method is given by

$$y_{n+1} = y_n + \frac{1}{2}(k_1 + k_2) \tag{3.29}$$

where

$$k_1 = hf(t_n, y_n) = hf_n$$

and

$$k_2 = hf(t_n + h, y_n + k_1) = hf(t_n + h, y_n + hf(t_n, y_n))$$

The *Classical Fourth-Order Runge-Kutta* popularly known as the *RK4* is given by

$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{3.30}$$

where

$$k_1 = hf(t_n, y_n) = hf_n$$
$$k_2 = hf(t_n + \frac{h}{2}, y_n + \frac{1}{2}k_1)$$
$$k_3 = hf(t_n + \frac{h}{2}, y_n + \frac{1}{2}k_2)$$

and

$$k_4 = hf(t_n + h, y_n + k_3)$$

Rk4 is consistent with an order of $O(\Delta t^4)$. The amplification factor G needed for the stability analysis is given by

$$G = 1 - (\alpha h) + \frac{1}{2}(\alpha h)^2 + \frac{1}{6}(\alpha h)^3 + \frac{1}{24}(\alpha h)^4$$

Finally $|G| \leq 1$ if $(\alpha h) \leq 2.785$ which makes this method *conditionally stable.* Therefore the conclusion is that the Classical Fourth-Order Runge-Kutta method is convergent.

3.3 Data Assimilation

3.3.1 Introduction

Information is always needed as far as the challenges to the environment are concerned. Most at times, this information is used to make predictions for the future, test hypotheses, and attribute cause and effect. This enables us to take action according to information provided on the future behaviour of the system of interest, and in particular future events (*prediction*), then test our understanding of the system, and adjust this understanding according to new information (*hypothesis testing*) and also understand the cause of events, and obtain information on possible ways of changing, mitigating or adjusting to the course of events (*attribute cause and effect*) irrespective of the origin of the information used.

However, we need to address some important issues. These are, a means of gathering information, and methods to build on this information gathered. The former is catered for by *observations* (measurements) whiles *models* (conceptual, numerical or otherwise) provide help us with the latter. The means of obtaining this information actually distinguishes observations from models, otherwise, the two are not distinct. In other words, observations have a roughly *direct link* with the system of interest via the measurement process whereas models have a roughly *indirect link* with the system of interest, being an embodiment of information received from measurements, experience and theory.

One very important and logical question to ask is; "What combination of the model and observation information is said to be optimal?". Mathematics provides an answer to this question as well as an estimate of the errors of the “optimum” or “best” estimate. This method is known as **data assimilation**. Data assimilation adds value to the observations by filling in the observational gaps and to the model by constraining it with observations. Accurate inferences are then made from the observations. In other words, data assimilation is the technique whereby observational data are combined with output from a numerical model to produce an optimal estimate of the evolving state of the system as a result of the need to improve the output of our models.

Data assimilation has strong links to several mathematical disciplines, including control theory and Bayesian statistics. Scientists are aware of the impossibility to create models that would reproduce to the perfection the behaviour of nature.

Even though the computers become everyday more sofisticated, they still cannot cope with the complexity of the world, and especially our inability to capture all the details of the system we want to model. Model users have to deal with those imperfections and try to correct them as efficiently as possible. In the simplest cases, a review of the system to be modeled, combined with a thorough calibration of the model lead to results that are acceptable.

In the most complex cases, real life data is incorporated to correct the model behaviour. This is what data assimilation is about: combining model predictions and real world data to make a better estimate of the state of the system we want to model.

3.3.2 Types of Data Assimilation

There are basically two categories of data assimilation namely Variational Data Assimilation and Sequential Data Assimilation.

3.3.3 Sequential Data Assimilation

With sequential assimilation, a *priori* estimates for the initial states x_0 are chosen and the model is evolved forward to the time t_k where the first observations are available. The predicted states of the system at this time are known as the background states and are denoted by x_k^b.

The difference between the predicted observation vector given by the background states and the measured observations vector at this time $(Hx_{k+1}^b - y_{k+1})$, known as the innovation vector, is then used to correct the background state vector in order to obtain improved state estimates x_k^a, known as the analysis states. The model is then evolved forward again from the analysis states to the next time where an observation is available and the process is repeated several times.

Some of the Sequential Data Assimilation methods include Kalman Filter, Extended Kalman Filter, Ensemble Kalman Filter and the Unscented Kalman Filter.

3.3.4 Variational Data Assimilation

The expression variational data assimilation designates a class of assimilation algorithms in which the fields to be estimated are explicitly determined as minimizers of a scalar function, called the objective function, that measures the misfit to the available data.

The variational assimilation usually seeks an optimal fit of the model solution to observations over a period by adjusting the estimation states in this period simultaneously. The estimated states over this period are somehow influenced by all the observations distributed in time. The information is propagated both from the past into the future and also from the future into the past.

The variational approach has, however, been extensively used in data assimilation for meteorological models and shows promising results for Numerical Weather

Prediction (NWP). This approach includes Three-dimensional variational data assimilation (3D-Var) and the four-dimensional variational data assimilation (4D-Var). The 4D-Var searches for an optimal set of model parameters (e.g., optimal initial state of the model) which minimizes the discrepancies between the model forecast and time distributed observational data over the assimilation window. Contrary to sequential data assimilation, which evolves the model one step at a time and updates the estimated states each time an observation is available, the four-dimensional assimilation schemes use all the observations available over a given time window to provide improved estimates for all the states in that window.

3.3.5 Kalman Filter

The *Kalman Filter* is basically a systematic procedure or a set of mathematical equations that uses the idea of predictor-corrector estimator where the estimator is optimal. The estimator that is obtained is said to be optimal because the filter minimizes the estimated error covariance. The Kalman filter is named after Rudolph E. Kalman who published is famous paper in 1960 describing a recursive solution to the discrete-data linear filtering problem (Kalman, 1960).

The filter is very powerful in several aspects. It supports estimations of past, present, and even future states (recursive), and it can do so even when the precise nature of the modeled system is unknown (Welch and Bishop, 2006).

3.3.6 The Process to be Estimated

The Kalman filter addresses the general problem of trying to estimate the state $x_k \in \mathbb{R}^n$ of a discrete-time controlled process that is governed by the linear stochastic difference equation at time k

$$x_k = Ax_{k-1} + Bu_{k-1} + \omega_{k-1}, \tag{3.31}$$

with a measurement $z \in \mathbb{R}^m$ that is

$$z_k = Hx_k + \nu_k. \tag{3.32}$$

The random variables ω_k and ν_k cater for the process and measurement noise (respectively) which are assumed to be independent, white, and with normal probability distributions

$$p(\omega_k) \sim N(0, Q_k), \tag{3.33}$$

$$p(\nu_k) \sim N(0, R_k). \tag{3.34}$$

The matrix in the difference equation (3.31) relates the state at the previous time step $k-1$ to the state at the current step, in the absence of either a driving function or process noise. Note that in practice A might change with each time step, but here we assume it is constant. The $n \times l$ matrix B relates the optional control input $u \in \mathbb{R}^l$ to the state x. The $m \times n$ matrix H in the measurement equation (3.32) relates the state to the measurement z_k. In practice H might change with each time step or measurement, but here we assume it is constant.

3.3.7 The Computational Essence of the Filter

We define $\hat{x}_k^- \in \mathbb{R}^n$ to be our *a priori* state estimate at step k given knowledge of the process prior to step k, and $\hat{x}_k \in \mathbb{R}^n$ to be our *a posteriori* state estimate at step k given measurement z_k. Defining *a priori* and *a posteriori* estimate errors respectively as

$$e_k^- \equiv x_k - \hat{x}_k^-$$

and

$$e_k \equiv x_k - \hat{x}_k$$

The *a priori* estimate error covariance is then given by

$$P_k^- = E[e_k^- e_k^{-T}], \tag{3.35}$$

and the *a posteriori* estimate error covariance is

$$P_k = E[e_k e_k^T]. \tag{3.36}$$

In deriving the equations for the Kalman filter, we begin with the goal of finding an equation that computes an *a posteriori* state estimate $\hat{x}_k$ as a linear combination of an *a priori* estimate $\hat{x}_k^-$ and a weighted difference between an actual measurement z_k and a measurement prediction $H\hat{x}_k^-$ as shown below in (3.37).

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-) \tag{3.37}$$

The difference $(z_k - H\hat{x}_k^-)$ in (3.37) is called the measurement innovation, or the residual. The residual reflects the discrepancy between the predicted measurement $H\hat{x}_k^-$ and the actual measurement z_k.

The $n \times m$ matrix K in (3.37) is chosen to be the *gain* or *correcting factor* that minimizes the *a posteriori* error covariance (3.36). This minimization can be accomplished by first substituting (3.37) into the above definition for e_k, substituting that into (3.36), performing the indicated expectations, taking the derivative of the trace of the result with respect to K, setting that result equal to zero, and then solving for K. (Maybeck, 1979; Brown and Hwang, 1992; Jacobs, 1993). One form of the resulting K that minimizes (3.36) is given by

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R_k)^{-1} \\ &= \frac{P_k^- H^T}{H P_k^- H^T + R_k} \end{aligned} \tag{3.38}$$

From (3.38), as the measurement error covariance approaches zero, the gain K weights the residual more heavily. Specifically,

$$\lim_{R_k \to 0} K_k = H^{-1}$$

On the other hand, as the *a priori* estimate error covariance P_k^- approaches zero, the gain K weights the residual less heavily. That is,

$$\lim_{P_k^- \to 0} K_k = 0$$

Alternatively, as the measurement error covariance R_k approaches zero, the actual measurement z_k is "preferred" more and more, while the predicted measurement $H\hat{x}_k^-$ is trusted less and less. On the other hand, as the a priori estimate error covariance P_k^- approaches zero the actual measurement z_k is trusted less and less, while the predicted measurement $H\hat{x}_k^-$ is trusted more and more.

3.3.8 The Probabilistic Essence of the Filter

The justification for (3.37) is rooted in the probability of the *a priori* estimate x_k^- conditioned on all prior measurements z_k (Bayes' rule). For now let it suffice to point out that the Kalman filter maintains the first two moments of the state distribution,

$$E[x_k] = \hat{x}_k$$

$$E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] = P_k.$$

The *a posteriori* state estimate (3.37) reflects the mean (the first moment) of the state distribution. It is normally distributed if the conditions of (3.33) and (3.34) are met. The *a posteriori* estimate error covariance (3.36) reflects the variance of the state distribution. In other words,

$$\begin{aligned} p(x_k|z_k) &\sim N(E[x_k], E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]) \\ &= N(\hat{x}_k, P_k) \end{aligned}$$

Maybeck (1979), Brown and Hwang (1992) and Jacobs (1993) provide more information.

3.3.9 The Discrete Kalman Filter Algorithm

This part of the thesis begins with a broad overview, covering the operation of one form of the discrete Kalman filter after which the focus is narrowed to the specific equations and their use in this version of the filter.

The filter estimates the process state at some time and then obtains information in the form of (noisy) measurements. As such, the equations for the Kalman filter fall behaves like a predictor-corrector process which are the *time update* equations and *measurement update* equations. The former are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the *a priori* estimates for the next time step whereas the latter are responsible for the correction - i.e. for incorporating a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate as shown below in figure 3.1.

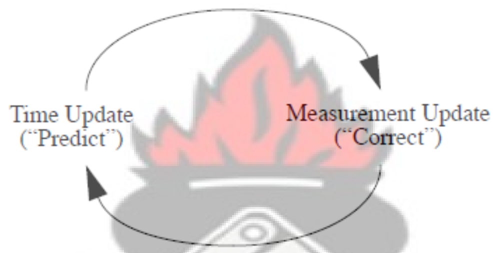

Figure 3.1: The ongoing discrete Kalman filter cycle. The *time update* projects the current state estimate ahead in time. The *measurement update* adjusts the projected estimate by an actual measurement in time (source: Welch and Bishop (2006))

The first task during the measurement update is to compute the *Kalman gain*, K_k. The next step is to actually measure the process to obtain z_k, and then to generate an *a posteriori* state estimate by incorporating the measurement. The last step is to obtain an a posteriori error covariance estimate.

After each time and measurement update pair, the process is repeated with the previous *a posteriori* estimates used to project or predict the new *a priori* estimates. This recursive nature is one of the very appealing features of the Kalman

Algorithm 1 *Kalman Filter*

1. Set initial estimates for $\hat{x}_0$ and P_0
2. For $k = 1$ to maximum number of iterations
   3. For $i = 1$ to n

      Time Update

      Project the state forward

      $\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}$

      Project the error covariance forward

      $P_k^- = AP_{k-1}A^T + Q_k$

   4. End of loop for i

      Measurement Update

      Compute the Kalman gain

      $K_k = P_k^- H^T(HP_k^- H^T + R_k)^{-1}$

      Update state estimate with measurement y_k

      $\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$

      Update the error covariance

      $P_k = (I - K_kH)P_k^-$

5. End of loop for k

filter. It makes practical implementations much more feasible than (for example) an implementation of a Wiener filter (Brown and Hwang, 1992) which is designed to operate on all of the data *directly* for each estimate. The Kalman filter instead recursively conditions the current estimate on all of the past measurements. Figure 3.2 offers a complete picture of the operation of the filter.

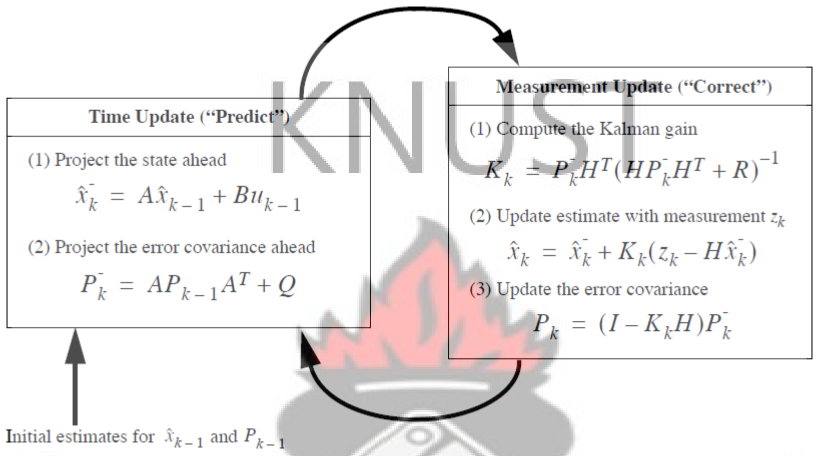

Figure 3.2: A complete picture of the operation of the Kalman filter (source: Welch and Bishop (2006))

Under conditions where Q_k and R_k are in fact constant, both the estimation error covariance P_k and K_k the Kalman gain will stabilize quickly and then remain constant. If this is the case, these parameters can be pre-computed by either running the filter off-line, or for example by determining the steady-state value of P_k as described in Grewal and Andrews (1993).

The Kalman filter is the most well known sequential data assimilation scheme. It has been developed in the sixties by R. E. Kalman to try to solve the Wiener problem in a generally easier way. The filter has the advantage to be sequential. It needs only the system variables of the previous time step and the forcing terms and observations of the current time step.

The Kalman filter however has limitations, one of such is the fact that it is unable to predict or estimate the state vector of *nonlinear* model accurately. That is the Kalman filter breaks down. Fortunately, there are improved forms of the Kalman filter used for nonlinear state equations.

The most famous of such methods is the *Extended Kalman Filter*. This method extends the kalman filtering through a procedure of linearization by making use of the taylor series expansion of the the nonlinear model. Once a linear model is obtained the kalman filter equations are then applied to obtain the estimate.

3.3.10 The Extended Kalman Filter

This method extends the kalman filtering through a procedure of linearization by making use of the taylor series expansion of the the nonlinear model. Once a linear model is obtained the kalman filter equations are then applied to obtain the estimate.

Assuming the process has a state vector $x \in \mathbb{R}^n$, but the process is now governed by the nonlinear stochastic difference equation

$$x_k = f(x_{k-1}, u_{k-1}, \omega_{k-1}) \tag{3.39}$$

with a measurement $z \in \mathbb{R}^m$ whose equation is given by

$$z_k = h(x_k, \nu_k) \tag{3.40}$$

The random variable ω_k and ν_k again represent the process and measurement noises respectively which are still white Gaussian. Their covariances are respectively Q_k and R_k.

The function f is used to compute the *a priori* state from the previous estimate and similarly the function h is also used to compute the predicted measurement from the predicted state. The two functions cannot be applied to the covariance directly. Instead a matrix of partial derivatives (*the Jacobian*) is computed due to the linearization using taylor series.

Defining the state transition and observation matrices by the following jacobians:

$$F_{k-1} = \frac{\partial f}{\partial x}\bigg|_{x^+_{k-1}}$$

$$H_k = \frac{\partial h}{\partial x}\bigg|_{x^-_k}$$

. The extended kalman filter has its time update equations given by

$$\hat{x}^-_k = f(\hat{x}_{k-1}, u_{k-1}) \tag{3.41}$$

$$P^-_k = F_{k-1}P_{k-1}F^T_{k-1} + Q_{k-1} \tag{3.42}$$

and that of the measurement update are

$$K_k = P^-_k H_k(H_k P^-_k H^T_k + R_k)^{-1} \tag{3.43}$$

$$\hat{x}_k = \hat{x}^-_k + K_k(z_k - h(x^-_k)) \tag{3.44}$$

$$P_k = (I - K_k H_k)P^-_k \tag{3.45}$$

The entire EKF process is described below in algorithm 2;

Again the Extended Kalman filter has a couple of short-falls. One of such is the fact that it is not always possible to calculate the *jacobian* and also it does not produce accurate results if the nonlinear function is not well approximated by a linear model or functions. That is, this model works well when the nonlinear model is close to a linear model. Another improved form of the Kalman filter used for the estimation of nonlinear model is the *Unscented Kalman Filter*.

Algorithm 2 *Extended Kalman Filter*

1. Set initial estimates for $\hat{x}_0$ and P_0
2. For $k = 1$ to maximum number of iterations
   3. For $i = 1$ to n

      Time Update

      Project the state forward

      $\hat{x}_k^- = f(\hat{x}_{k-1}, u_{k-1})$

      Linearize the process equation by computing the jacobian

      $F_{k-1} = \frac{\partial f}{\partial x}\big|_{x_{k-1}^+}$

      Project the error covariance forward

      $P_k^- = F_{k-1} P_{k-1} F_{k-1}^T + Q_{k-1}$

   4. End of loop for i

      Measurement Update

      Linearizing the measurement equation

      $H_k = \frac{\partial h}{\partial x}\big|_{x_k^-}$

      Compute the Kalman gain

      $K_k = P_k^- H_k (H_k P_k^- H_k^T + R_k)^{-1}$

      Update state estimate with measurement y_k

      $\hat{x}_k = \hat{x}_k^- + K_k(z_k - h(x_k^-))$

      Update the error covariance

      $P_k = (I - K_k H_k) P_k^-$

5. End of loop for k

3.3.11 The Unscented Kalman Filter

The *Unscented Kalman Filter* was first introduced by Julier et al. (1995); Julier and Uhlmann (1996, 1997) and further improved by Wan et al. (2000); Wan and van der Merwe (2000). This filter is obtained from the basic idea of the *Unscented Transform*

3.3.12 Unscented Transformation

The *Unscented transformation* is a method used for calculating the statistics (mean and covariance) of a random variable which undergoes a nonlinear transformation. It is based on the fact that it is easier to approximate a Gaussian distribution than it is to approximate a nonlinear function. The idea of this transform is to obtain a set of points known as Sigma points so that the sample mean and sample covariance are $\bar{x}$ and P_x respectively. The nonlinear function is applied to these set of points to produce a set of transformed sigma points whose statistics are $\bar{y}$ and P_y.

The n-dimensional random variable with statistics, $\bar{x}$ and P_x representing the mean and covariance respectively is approximated by $2L+1$ points given by

$$\begin{aligned}
\chi_0 &= \bar{x} \\
\chi_i &= \bar{x} + (\sqrt{(L+\lambda)+P_x})_i \qquad i = 1,\dots,L \\
\chi_i &= \bar{x} - (\sqrt{(L+\lambda)+P_x})_{i-L} \quad i = L+1,\dots,2L
\end{aligned}$$

where $\lambda \in \mathbb{R}$, $\lambda = \alpha^2(L+\kappa) - L$ is a scaling parameter. The constant α determines the spread of the sigma points around $\bar{x}$ and is usually set to a small positive value (e.g., 1e-3). The constant κ is a secondary scaling parameter which is usually set to 0 or $3-L$ and β is used to incorporate prior knowledge of the distribution of x (for Gaussian distributions, $\beta = 2$ is optional). $(\sqrt{(L+\lambda)+P_x})_i$ is the *ith* row or column of the matrix square root of $(L+\lambda)+P_x$. A numerically efficient and stable method such as the Cholesky decomposition method should be used to calculate the matrix square root. The nonlinear function is applied to these set of points obtained above to yield the transformed sigma points given by

$$Y_i = f[\chi_i] \quad i = 0, \ldots, 2L$$

whose mean and covariance are respectively

$$\bar{y} = \sum_{i=0}^{2L} W_i^{(m)} Y_i \text{ and } P_y = \sum_{i=0}^{2L} W_i^{(c)} \{Y_i - \bar{y}\}\{Y_i - \bar{y}\}^T$$

where the weights W_i are given as

$$\begin{aligned} W_0^{(m)} &= \lambda/(L+\lambda) \\ W_0^{(c)} &= \lambda/(L+\lambda) + (1-\alpha^2+\beta) \\ W_i^{(m)} &= W_i^{(c)} = 1/\{2(L+\lambda)\} \quad i = 0, \ldots, 2L \end{aligned}$$

See algorithm 3 below for the entire UKF process.

3.4 Numerical Example of Data Assimilation Methods in State Estimation

The purpose of this section is to reconstruct the trajectories and dynamics of the scaler state-space problem described below using UKF and compare its performance with that of the EKF.

State Equation

$$x_k = \frac{1}{2}x_{k-1} + \frac{25x_{k-1}}{1+x_{k-1}^2} + 8\cos(1.2(k-1)) + \omega_k \tag{3.46}$$

Measurement Equation

$$y_k = \frac{1}{20}x_k^2 + \nu_k \tag{3.47}$$

Algorithm 3 *Unscented Kalman Filter*

1. Set initial estimates for $\hat{x}_0$ and P_0
2. For $k = 1$ to maximum number of iterations
   3. For $j = 1$ to n

      Time Update

      Compute sigma points

      $\chi_{k-1} = \left[\hat{x}_{k-1} \quad \hat{x}_{k-1} + \sqrt{(L+\lambda) + P_{k-1}} \quad \hat{x}_{k-1} - \sqrt{(L+\lambda) + P_{k-1}}\right]$

      Compute transformed sigma points

      $\chi_k^* = F[\chi_{k-1}, u_{k-1}]$

      Project the state forward and error covariance forward

      $\hat{x}_k^- = \sum_{i=0}^{2L} W_i^{(m)} \chi_{i,k}^*$ and $P_k^- = \sum_{i=0}^{2L} W_i^{(c)} \{\chi_{i,k}^* - \hat{x}_{i,k}^-\}\{\chi_{i,k}^* - \hat{x}_{i,k}^-\}^T$
   4. End of loop for j

      Measurement Update

      Compute sigma points

      $\chi_{k-1} = \left[\hat{x}_k^- \quad \hat{x}_k^- + \sqrt{(L+\lambda) + P_k^-} \quad \hat{x}_k^- - \sqrt{(L+\lambda) + P_k^-}\right]$

      Compute transformed sigma points

      $Z_{k-1} = H[\chi_{k-1}]$

      Compute the mean and covariance of Z_{k-1}

      $\hat{z}_k^- = \sum_{i=0}^{2L} W_i^{(m)} Z_{i,k-1}$ and $P_{z_k z_k} = \sum_{i=0}^{2L} W_i^{(c)} \{Z_{i,k-1} - \hat{z}_{i,k}^-\}\{Z_{i,k-1} - \hat{z}_{i,k}^-\}^T$

      Compute the cross covariance of χ_k^* and Z_{k-1}

      $P_{x_k z_k} = \sum_{i=0}^{2L} W_i^{(c)} \{\chi_{i,k-1} - \hat{x}_{i,k}^-\}\{Z_{i,k-1} - \hat{z}_{i,k}^-\}^T$

      Compute the Kalman gain

      $K = P_{x_k z_k} P_{z_k z_k}^{-1}$

      Update state estimate with measurement y_k

      $\hat{x}_k = \hat{x}_k^- + K_k(z_k - \hat{z}_k^-)$

      Update the error covariance

      $P_k = P_k^- - K_k P_{z_k z_k} K_k^T$
5. End of loop for k

where ω_k and ν_k are both Gaussian white noise sequences. A synthetic truth data is generated by adding white Gaussian noise to the initial mean and covariance using the process model. This truth data is also perturbed with white Gaussian noise and the help of the measurement model to obtain the observations (also synthetic). The simulation length used is one hundred (100) time period. The Kalman filter cannot accurately estimate the above system due to the high degree of nonlinearity in the process as well as the measurement equations.

A couple of experiments are performed to estimate the state of the system above using EKF and UKF and most importantly compare their performances. The first experiment performed is to estimate the state using EKF and UKF with the initial values described below. The process and measurement noises and the initial error covariance are also varied to investigate their effects on the performance of the two filters in subsequent experiments.

The first experiment was carried out with an initial state estimate and covariance of $x_0 = 0.1$ and $P_0^+ = 1$ respectively and also the variances of both the state and the measurement noises were one. A plot of the noisy observations and the true state is given in figure 3.3. Figure 3.4 shows the plot of the EKF and UKF for this experiment. This plot also contains the noisy observations and the true state. It can be seen clearly that the UKF estimates the state better than the EKF. The performance can also be compared by using the mean and variance of the root mean square error (RMSE) of the two filters in table 3.1. This table indicates that the RMSE of UKF is smaller than that of EKF. A better performance of the UKF can also be seen in the plot of the RMSE of the two filters in figure 3.5.

The second experiment is to establish the effect of the process noise on the two filters. The variance of the process noise is varied for the estimation of the state using the two filters whiles keeping other parameters constant. The result of this experiment is shown in table 3.2. The experiment showed that the RMSE

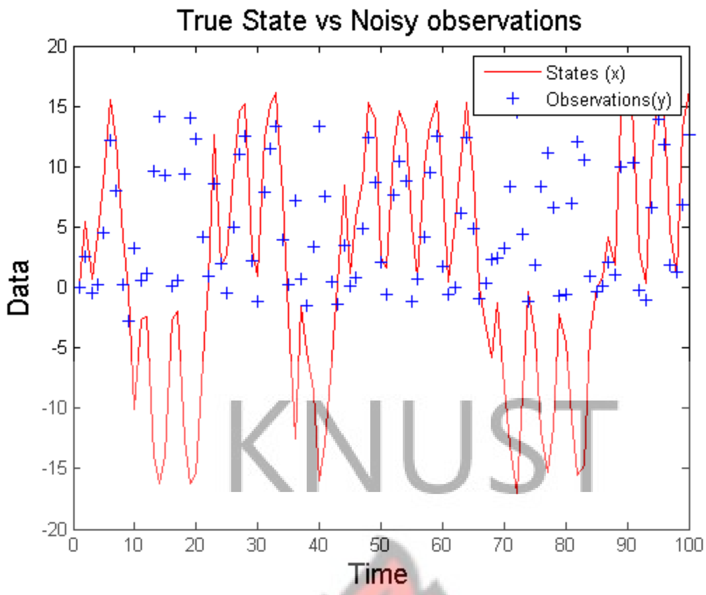

Figure 3.3: Time series of the noisy observations and the true state for the first experiment

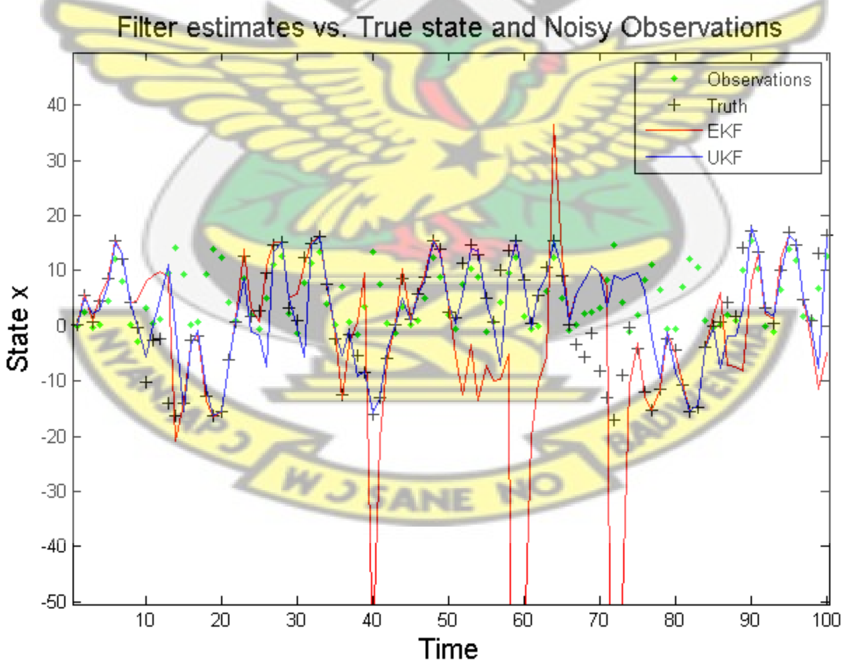

Figure 3.4: EKF and UKF estimates for the first experiment

Table 3.1: Mean and Variance of RMSE for state estimate of the two filters

	RMSE	
Algorithm	**Mean**	**Variance**
Extended Kalman Filter (EKF)	2.9269	19.6201
Unscented Kalman Filter (UKF)	0.9021	0.6872

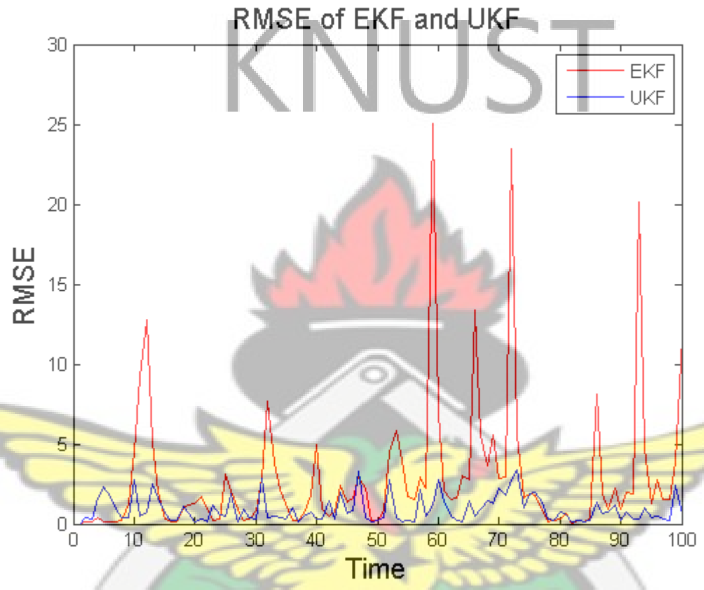

Figure 3.5: RMSE of EKF and UKF for the first experiment

Table 3.2: Compare effect of different process noise variance Q on the performance of the various filters holding other parameters fixed

	$Q = 10^{-2}$		$Q = 10^{-1}$		$Q = 10^{1}$		$Q = 10^{2}$	
Filters	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**
EKF	0.6987	3.9352	0.9949	2.3690	6.1988	51.5068	9.5933	142.2488
UKF	0.2039	0.2127	0.5332	0.4119	1.5985	0.9806	3.0093	2.3602

increased with increase in the variance of the process noise but of course, the UKF performed better than the EKF.

The next experiment is to establish the effect of the measurement noise on the two filters. That is, the variance of the measurement noise R is varied whiles keeping the other parameters constant. The result of this experiment is summarized in table 3.3. It was observed that the performance of the filters was not affected much by the variance of the measurement covariance since the mean of the RMSE of EKF was between 2 and 4 and that of UKF was between 1 and 2 even though it increased steadily with increase in the measurement variance. This result also showed that the UKF performed better than the EKF. The final experiment is to find the effect of the initial estimation covariance P_0 on the two filters. The result of the experiment shown in table 3.4 revealed that the filters slightly deteriorated with increase in the initial estimation covariance. A further test summarized in table 3.5 also showed that the performance of the two filters was not much affected by the value of the initial estimation covariance. The performance of the two filters increased slowly for higher values of P_0. The UKF breaks down for values of $P_0 \geq 10^{160}$.

Table 3.3: Compare effect of different observation noise R on the performance of the various filters holding other parameters fixed

	$R = 10^{-2}$		$R = 10^{-1}$		$R = 10^{1}$		$R = 10^{2}$	
Filters	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**
EKF	3.2160	14.3793	3.5132	32.4918	2.3519	8.4098	2.5222	6.1755
UKF	1.0385	1.7343	1.2049	1.0691	1.2367	0.5034	1.6283	0.8508

Table 3.4: Compare effect of different initial error covariance P_0 on the performance of the various filters holding other parameters fixed

	$P_o = 10^{-2}$		$P_0 = 10^{-1}$		$P_0 = 10^{1}$		$P_0 = 10^{2}$	
Filters	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**
EKF	2.5802	14.6112	2.7513	14.9827	2.9719	44.974	3.2867	23.7786
UKF	0.97259	0.65879	1.0393	1.10832	0.87909	0.5737	1.27415	0.59991

Table 3.5: Compare effect of different initial error covariance P_0 further on the performance of the various filters holding other parameters fixed

	$P_o = 10^{-160}$		$P_0 = 10^{-50}$		$P_0 = 10^{50}$		$P_0 = 10^{160}$	
Filters	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**
EKF	2.2052	9.2556	2.6689	9.7787	3.2928	25.3028	3.3867	23.7786
UKF	0.76574	0.38944	0.79011	0.52138	1.6292	1.1591	NaN	NaN

The trajectories and dynamics of the state-space problem above have been reconstructed using the UKF. Further experiments performed revealed that the UKF estimated the state better than the EKF for the above nonlinear problem.

Chapter 4

Results and Discussion

4.1 Introduction

Two systems are discussed in this chapter, a linear system and a non-linear system. The first is a planar four-story shear building with time varying system parameters and the second is a single degree-of-freedom (SDOF) system. Data contaminated with noise is generated for these dynamical systems. With these simulated data, identification models are derived for these dynamical system which are then used for the state and parameter estimation using the Unscented Kalman filter. Finally, the performance of the UKF is compared with that of the EKF and inferences are made on the results obtained from the two filtering techniques.

4.2 Planer Four-Story Shear Building

4.2.1 Data

An idealized planer four degrees of freedom (four-story) shear building is considered. The four floors of this building have equal masses of 250,000kg. That is $m_1 = m_2 = m_3 = m_4$ (where the subscript denote the floor number). The time-variant inter-story viscous damping coefficients are c_1, c_2, c_3 and c_4 while the inter-story stiffness are k_1, k_2, k_3 and k_4 and they also change as far as time is concerned. It is very important to mention here that the time evolutions of the inter-story viscous damping coefficients and the inter-story stiffness with the exception of k_2 are brownian motions with coefficient of variation equal to 2% which is clearly seen in figure 4.1.

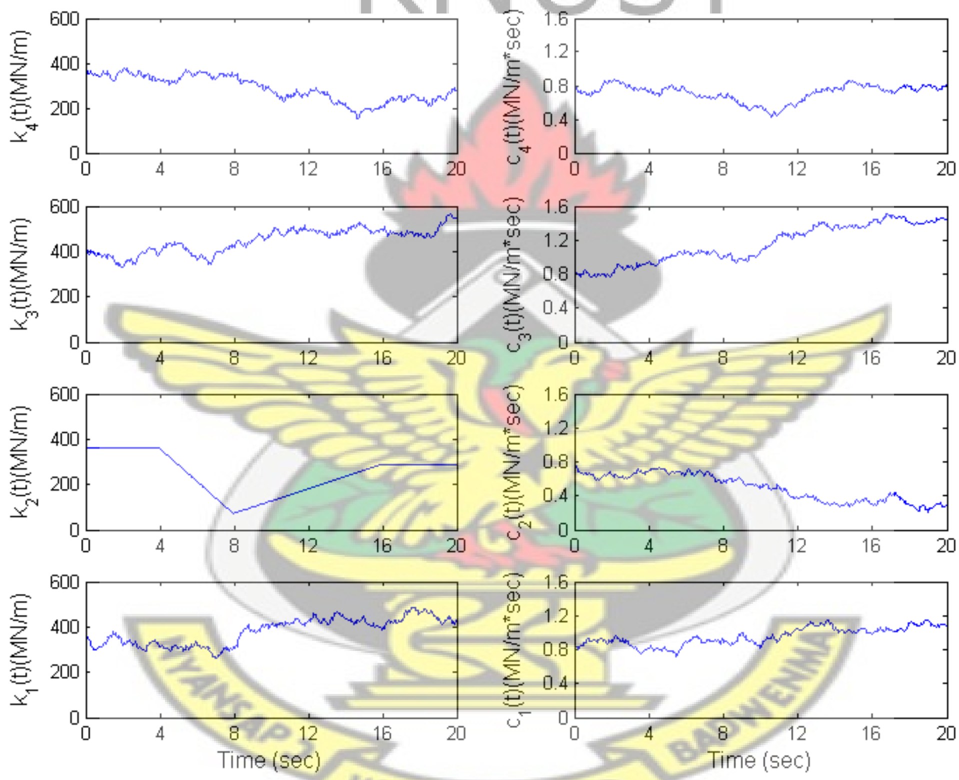

Figure 4.1: Time series of the inter-story stiffness and viscous damping coefficients

The state or process for this system subject to base excitation is given by

$$M\ddot{x}_t + C_t\dot{x}_t + K_t x_t = F u_t \tag{4.1}$$

where

$$x = \begin{bmatrix} x_{4,t} \\ x_{3,t} \\ x_{2,t} \\ x_{1,t} \end{bmatrix} \qquad F = \begin{bmatrix} -m_4 \\ -m_3 \\ -m_2 \\ -m_1 \end{bmatrix} \qquad M = \begin{bmatrix} m_4 & 0 & 0 & 0 \\ 0 & m_3 & 0 & 0 \\ 0 & 0 & m_2 & 0 \\ 0 & 0 & 0 & m_1 \end{bmatrix}$$

$$C_t = \begin{bmatrix} c_{4,t} & -c_{4,t} & 0 & 0 \\ -c_{4,t} & c_{4,t} + c_{3,t} & -c_{3,t} & 0 \\ 0 & -c_{3,t} & c_{3,t} + c_{2,t} & -c_{2,t} \\ 0 & 0 & -c_{2,t} & c_{2,t} + c_{1,t} \end{bmatrix} \tag{4.2}$$

$$K_t = \begin{bmatrix} k_{4,t} & -k_{4,t} & 0 & 0 \\ -k_{4,t} & k_{4,t} + k_{3,t} & -k_{3,t} & 0 \\ 0 & -k_{3,t} & k_{3,t} + k_{2,t} & -k_{2,t} \\ 0 & 0 & -k_{2,t} & k_{2,t} + k_{1,t} \end{bmatrix}$$

$x_{i,t}$ is the (i+1)-th floor relative to the ground where the fifth floor is the taken as the roof. u_t is the acceleration at the first floor of the building. with i representing the i-th floor, $c_{i,t}$ and $k_{i,t}$ are respectively the inter-story damping coefficient and the inter-story stiffness.

Data is generated using white gaussian noise for the excitation u_t. The observed output or observation y_t is the absolute accelaration time at the four stories given by:

$$\begin{aligned} y_t &= \begin{bmatrix} \ddot{x}_{1,t} + u_t \\ \ddot{x}_{2,t} + u_t \\ \ddot{x}_{3,t} + u_t \\ \ddot{x}_{4,t} + u_t \end{bmatrix} + \Gamma.\nu_t \\ &= -M^{-1}[C_t \dot{x}_t + K_t x_t] + \Gamma.\nu_t \end{aligned} \tag{4.3}$$

where $\nu_t \in \mathbb{R}^4 \sim N(0, I)$ are the measurement uncertainties for y_t which are stationary. $\Gamma = \text{diag}(\gamma_1, \ldots \gamma_4)$ is such that the overall signal/noise root mean square amplitude ratios for each channel is roughly equal to ten (10). With the observation $y_t : t = 1, \ldots, t$ and excitation $u_t : t = 1, \ldots, 4$ of the system which are both sampled at an interval of 0.02s (see figure 4.2), the ultimate is to estimate the system states (i.e. displacements x_t and velocities $\dot{x}_4$) and also the system parameters (i.e inter-story damping coefficients $c_{i,t}$ and inter-story stiffness $k_{i,t}$) in real time with the help of an identification model.

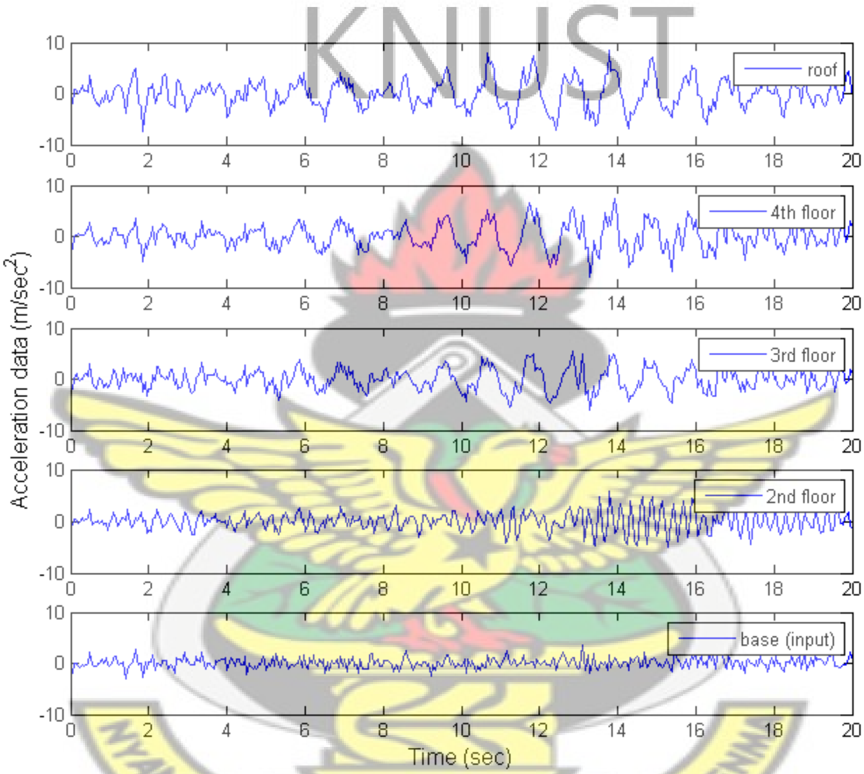

Figure 4.2: Time series of the simulated observation and excitation data

Equation (4.1) is initially transformed into a system of first-order differential equations. Let

$$w_1 = x_t$$

$$w_2 = \dot{w}_1 = \dot{x}_t$$

and finally

$$\dot{w}_2 = \ddot{w}_1 = \ddot{x}_t$$

Equation (4.1) is then written as:

$$\begin{aligned} \dot{w}_1 &= w_2 \\ \dot{w}_2 &= -M^{-1}C_t w_2 - M^{-1}K_t w_1 + M^{-1}F u_t \end{aligned} \tag{4.4}$$

Changing the w's back to x_t results in

$$\frac{d}{dt}\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} = \begin{bmatrix} \dot{x}_t \\ -M^{-1}C_t\dot{x}_t - M^{-1}K_t x_t \end{bmatrix} + \begin{bmatrix} 0 \\ M^{-1}F \end{bmatrix} u_t \tag{4.5}$$

or better still:

$$\frac{d}{dt}\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -M^{-1}K_t & -M^{-1}C_t \end{bmatrix}\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} + \begin{bmatrix} 0 \\ M^{-1}F \end{bmatrix} u_t \tag{4.6}$$

An identification model helps us to include the system parameters to our existing model described earlier. Using a brownian motion prior PDF for the parameter evolution, that is governed by the following equation;

$$\dot{\theta}_t = G.\omega_t \tag{4.7}$$

the identification model that is used for this system is given by:

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} x_t \\ \dot{x}_t \\ \theta_t \end{bmatrix} &= \begin{bmatrix} 0 & I & 0 \\ -M^{-1}K_t & -M^{-1}C_t & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_t \\ \dot{x}_t \\ \theta_t \end{bmatrix} + \begin{bmatrix} 0 \\ M^{-1}F \\ 0 \end{bmatrix} u_t + \begin{bmatrix} 0 \\ 0 \\ G \end{bmatrix} \omega_t \\ y_t &= -M^{-1}K_t x_t - M^{-1}C_t\dot{x}_t + \nu_t \end{aligned} \tag{4.8}$$

where $\omega_t \sim N(0, I)$, $G \in \mathbb{R}^{8\times 8}$ is a diagonal matrix whose diagonals must be specified and $\theta_t \in \mathbb{R}^8$ is made up of the system's parameters (i.e. four inter-story damping coefficients and four inter-story stiffness).

The dimension of the augmented state of the model in (4.8) is sixteen (16) since it includes four displacements x_t, four velocities $\dot{x}_t$, four damping coefficients c_t and four stiffness k_t and the dimension for the observed output y_t is four (4).

4.2.2 Results

The linear system described above was estimated using EKF and UKF and the results shown in figures 4.3 - 4.8. The system's states (i.e. displacements x_t and velocities $\dot{x}_t$) were very well estimated using the two filters in figures 4.3 and 4.4 respectively. The two filters again successfully tracked the system parameters. The inter-story stiffness $k_{(2,t)}$ was well estimated along with the other inter-story stiffness by both EKF and UKF even though it was not brownian (see figure 4.5). The estimates of the damping coefficients c_t in figure 4.6 is worse compared to the accuracy of the inter-story stiffness.

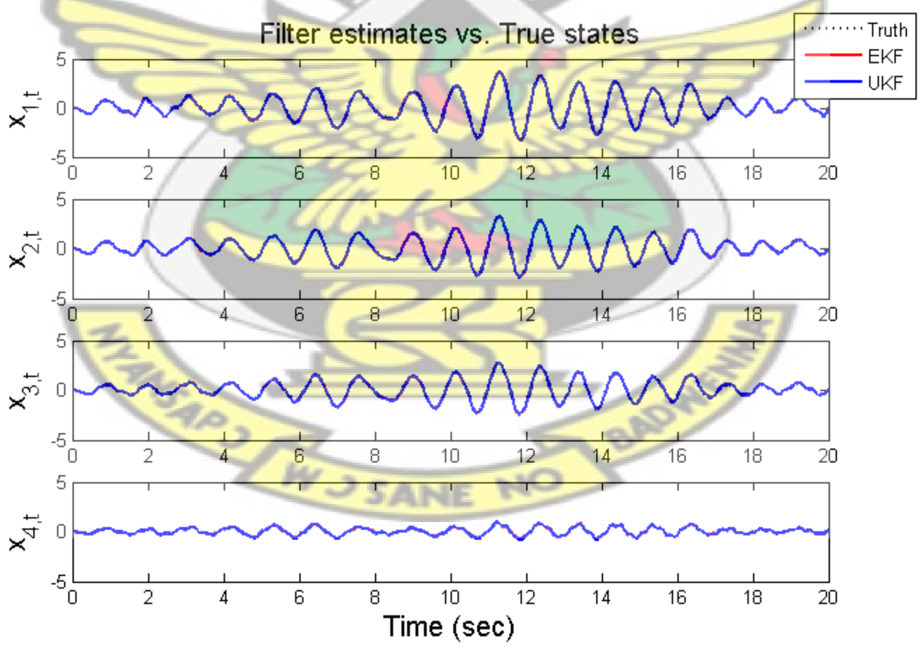

Figure 4.3: Estimation of displacements x_t

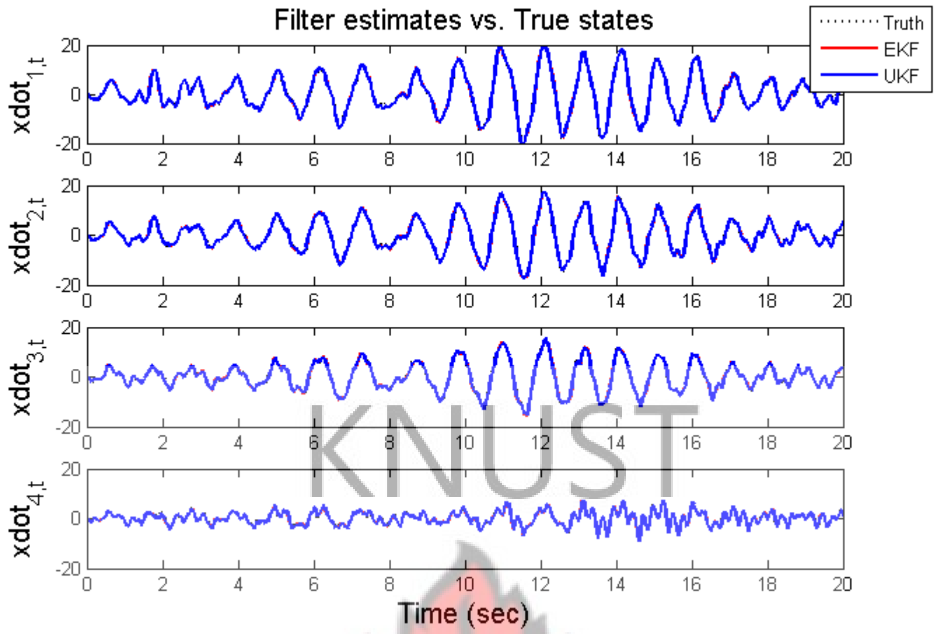

Figure 4.4: Estimation of velocities $\dot{x}_t$

The performance of the two filters was also compared using the RMSE for the states and parameters of the system. Figure 4.7 showed that the RMSE for the system states are close, indicating the same level of accuracy for the two filters. This observation is however not the same with the system parameters where UKF recorded lower rmse values than EKF for some parameters and vice versa (see figure 4.8).

The state as well as parameters of the planer four-story shear building were successfully estimated using the Unscented Kalman filter and its performance was compared with the Extended Kalman filter. The results obtained indicated an approximately equal performance for the two filters.

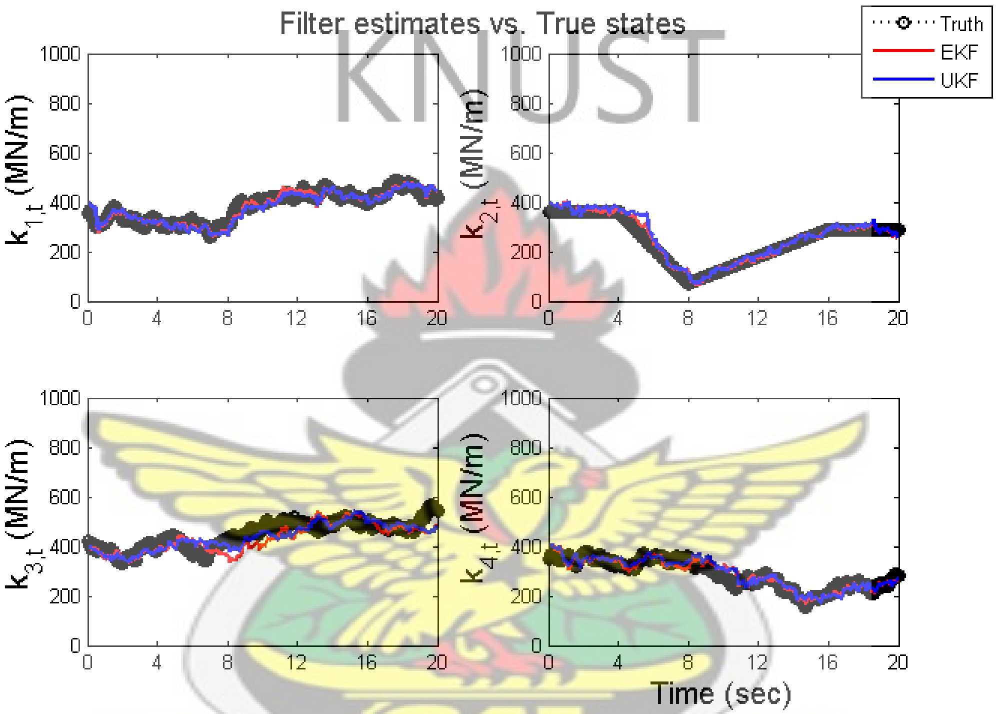

Figure 4.5: Estimation of inter-story stiffness k_t

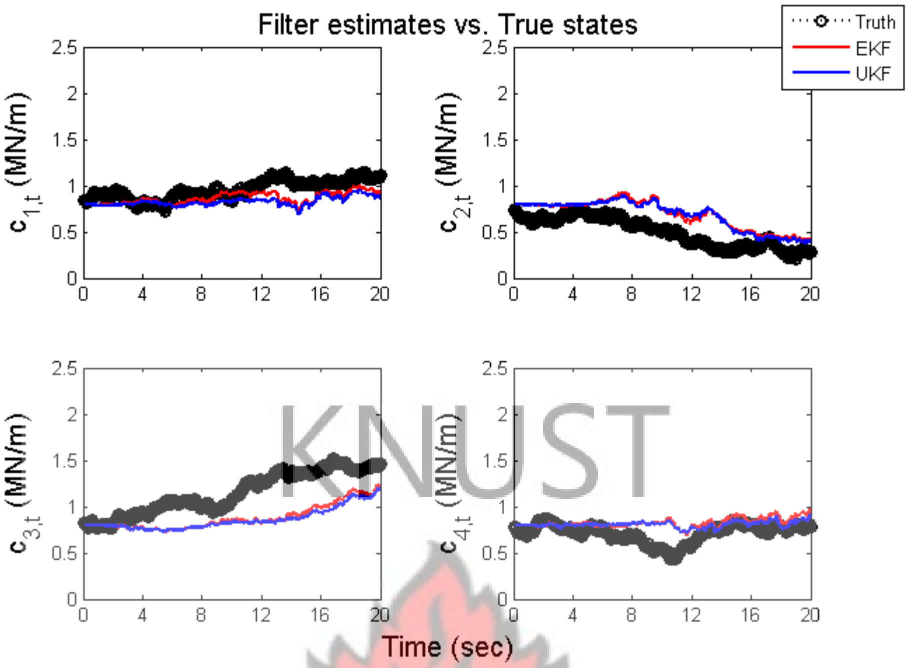

Figure 4.6: Estimation of damping coefficients c_t

4.3 Nonlinear SDOF System

4.3.1 Data

The system used to generate the synthetic data for the SDOF is given by;

$$\frac{d}{dt}\begin{bmatrix} x_t \\ \dot{x}_t \\ r_t \end{bmatrix} = \begin{bmatrix} \dot{x}_t \\ -1/m \cdot r_t + 1/m \cdot u_t \\ \theta_{1,t} \cdot \dot{x}_t - \theta_{2,t} \cdot |\dot{x}_t||r_t|^{\theta_{4,t}-1} r_t + \theta_{3,t} \cdot \dot{x}_t |r_t|^{\theta_{4,t}} \end{bmatrix} \tag{4.9}$$
$$y_t = -1/m \cdot r_t + 1/m \cdot u_t + \nu_t$$

where m is the mass of the SDOF system, r_t is the restoring force, u_t is a white-noise excitation force on the mass. y_t is the acceleration measured on the mass, ν_t is stationary such that the overall signal/noise amplitude ratio is ten (10). The time-varying system parameters are $\theta_{1,t}$, $\theta_{2,t}$, $\theta_{3,t}$ and $\theta_{4,t}$. These parameters

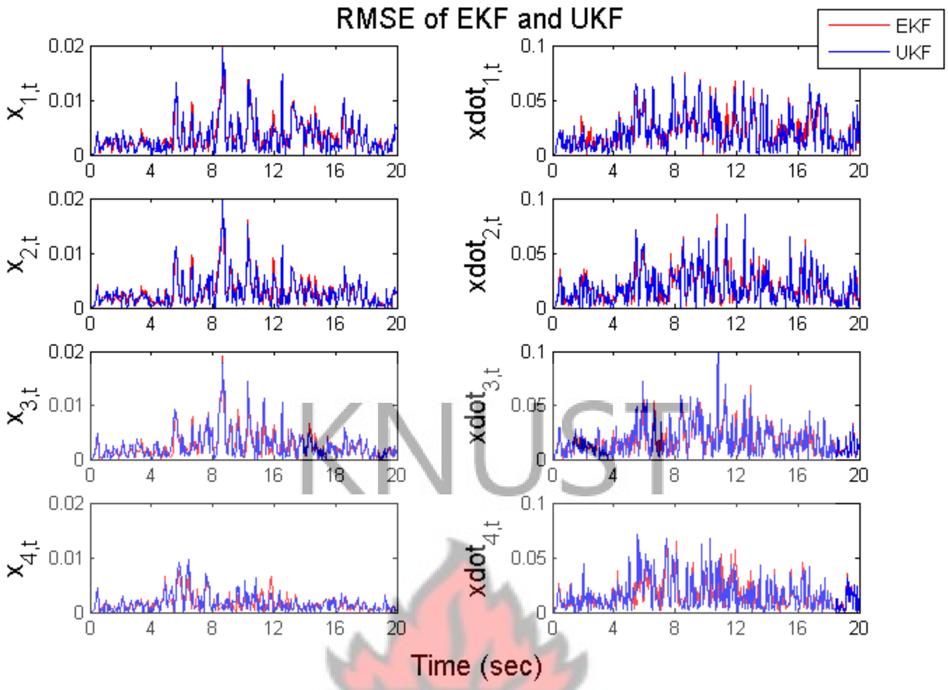

Figure 4.7: RMSE of EKF and UKF for x_t and $\dot{x}_t$

are brownian motions with drift coefficient of variance equal to 2% during each sampling interval and they actually fine tune the shape of the hysteric loop. The observation (or acceleration) y_t and excitation u_t are sampled at an interval of $0.5s$ as shown in figure 4.9.

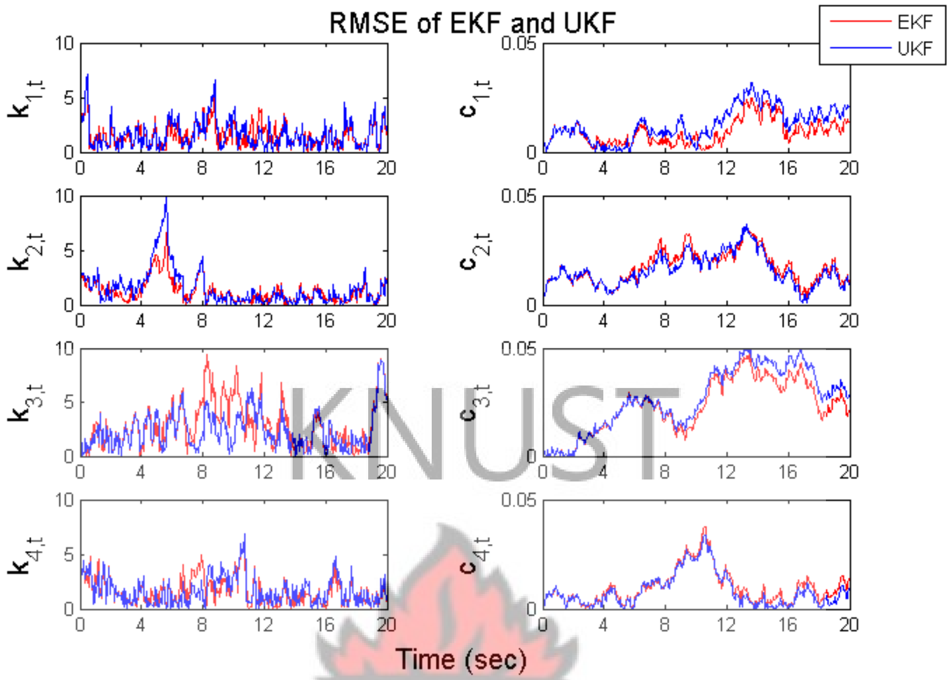

Figure 4.8: RMSE of EKF and UKF for k_t and c_t

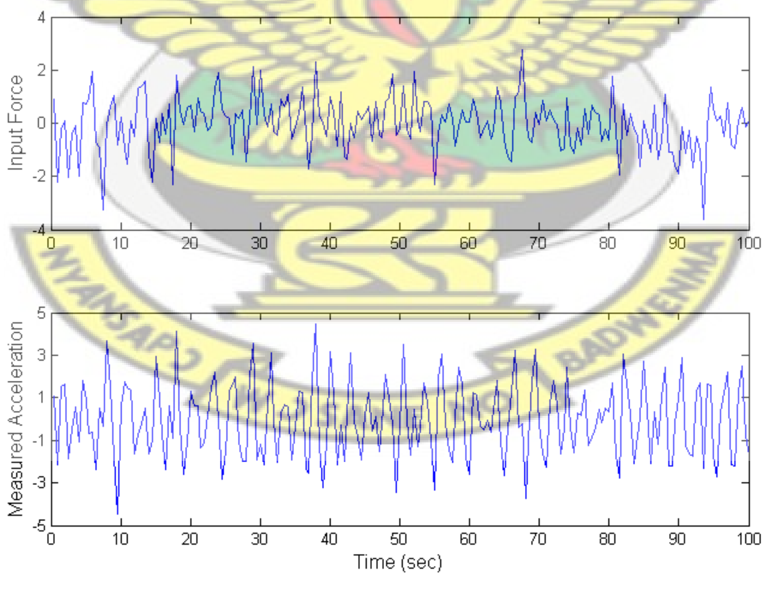

Figure 4.9: Time series of simulated acceleration and input force

The identification model used for this system is given by

$$\frac{d}{dt}\begin{bmatrix} x_t \\ \dot{x}_t \\ r_t \\ \theta_{1,t} \\ \theta_{2,t} \\ \theta_{3,t} \\ \theta_{4,t} \\ h_t \end{bmatrix} = \begin{bmatrix} \dot{x}_t \\ -1/m \cdot r_t + 1/m \cdot u_t \\ \theta_{1,t} \cdot \dot{x}_t - \theta_{2,t} \cdot |\dot{x}_t||r_t|^{\theta_{4,t}-1} r_t + \theta_{3,t} \cdot \dot{x}_t |r_t|^{\theta_{4,t}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.10}$$

$$+ \begin{bmatrix} 0 \\ 0 \\ 0 \\ G \end{bmatrix} \cdot \omega_t$$

$$y_t = -1/m \cdot r_t + 1/m \cdot u_t + h_t \cdot \nu_t$$

where $\omega_t \in \mathbb{R}^5 \sim N(0, I)$, $G \in \mathbb{R}^{5\times 5}$ and $\nu_t \in \mathbb{R} \sim N(0, 1)$. G is a 5×5 diagonal matrix given by

$$G = \begin{bmatrix} G_{1,t} & 0 & 0 & 0 & 0 \\ 0 & G_{2,t} & 0 & 0 & 0 \\ 0 & 0 & G_{3,t} & 0 & 0 \\ 0 & 0 & 0 & G_{4,t} & 0 \\ 0 & 0 & 0 & 0 & G_{5,t} \end{bmatrix}$$

and ω_t is given by

$$\omega_t = [\omega_{1,t}, \omega_{2,t}, \omega_{3,t}, \omega_{4,t}, \omega_{5,t}]^T$$

.

The initial PDF of the states x_0 and $\dot{x}_0$ is taken to be zero-mean Gaussian with large variances and that of the parameters $\theta_{1,0}$, $\theta_{2,0}$, $\theta_{3,0}$ and $\theta_{4,0}$ are also taken to be Gaussian. As mentioned earlier, each parameter is allowed to drift with a coefficient of variation equal to 2% except for the uncertain parameter h_t, whose actual propagation is a constant instead of a brownian motion.

4.3.2 Results

Figures 4.10 - 4.12 show the results of the estimation of the state and the parameters of the above system using EKF and UKF. The estimations of the system states (i.e. displacement x_t, velocity $\dot{x}_t$ and restoring force r_t) are shown in figure 4.10. The performances of EKF and UKF with regards to the estimation of the system's states are at par since their estimations are almost the same. This however is not the same with the system parameters. The plot of the estimation of the system parameters $\theta_{1,t}$, $\theta_{2,t}$, $\theta_{3,t}$ and $\theta_{4,t}$ in figure 4.11 revealed that the UKF performed better than the EKF. UKF again performed better than the EKF in estimating the uncertain parameter h_t which is shown in figure 4.12.

The performances of the two filters are again compared using the root mean squared error (RMSE). The plots of the RMSE of the two filters are shown in figure 4.13. The plots of the RMSE of EKF and UKF are similar for the system's states and $\theta_{1,t}$ (i.e. the first four plots). UKF however had lower RMSE for the rest of the parameters. Finally, the mean of the RMSE of the states and parameters for two filters were calculated and shown in table 4.1. This table indicates a better performance of UKF for the states and parameters except for the displacement (x_t) whose value was higher than that of the EKF.

The nonlinear SDOF system described above was successfully estimated using the Unscented Kalman filter and its performance was compared with the Extended Kalman filter. The results revealed that the UKF performed better than the EKF.

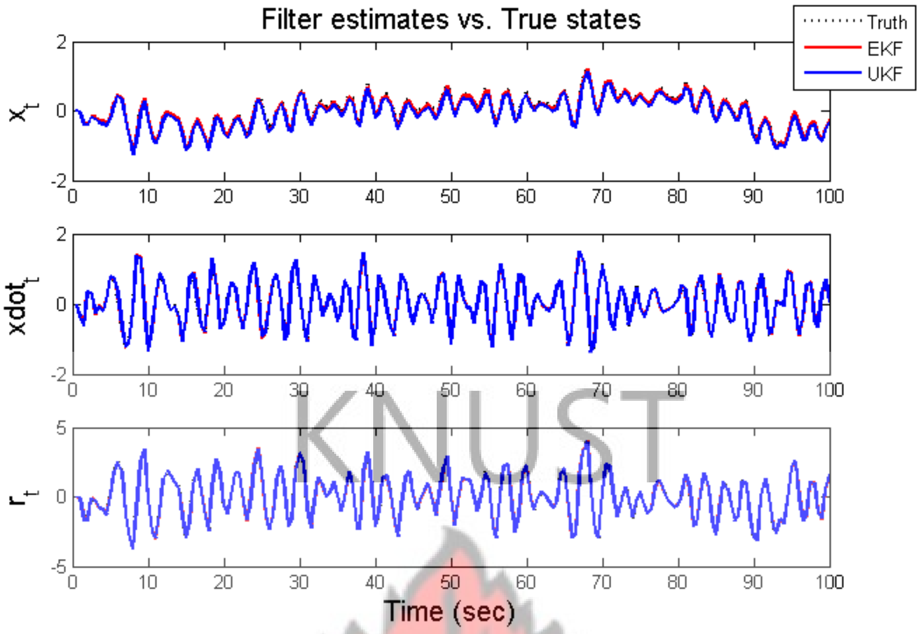

Figure 4.10: Estimation of the states of the SDOF system

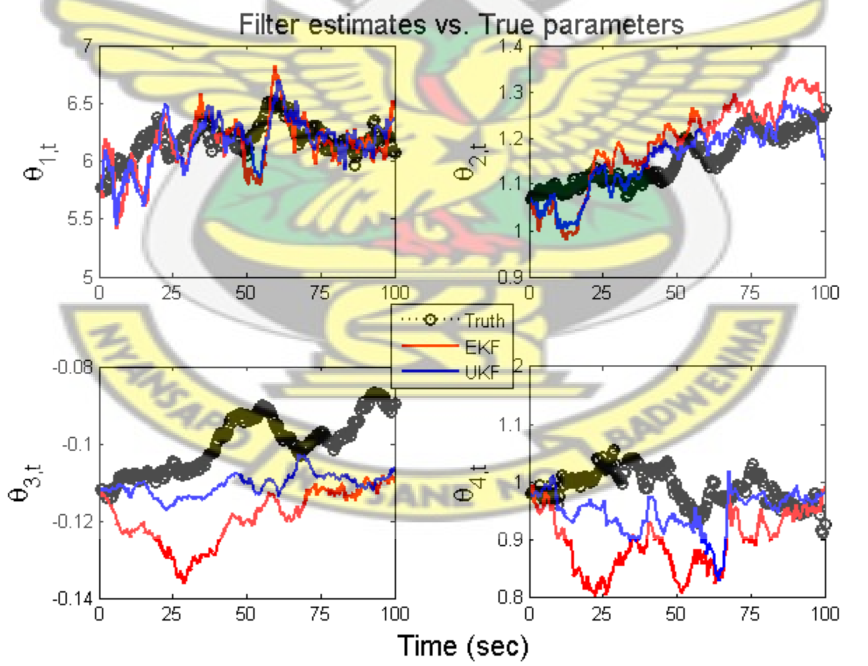

Figure 4.11: Estimation of the parameters of the SDOF system

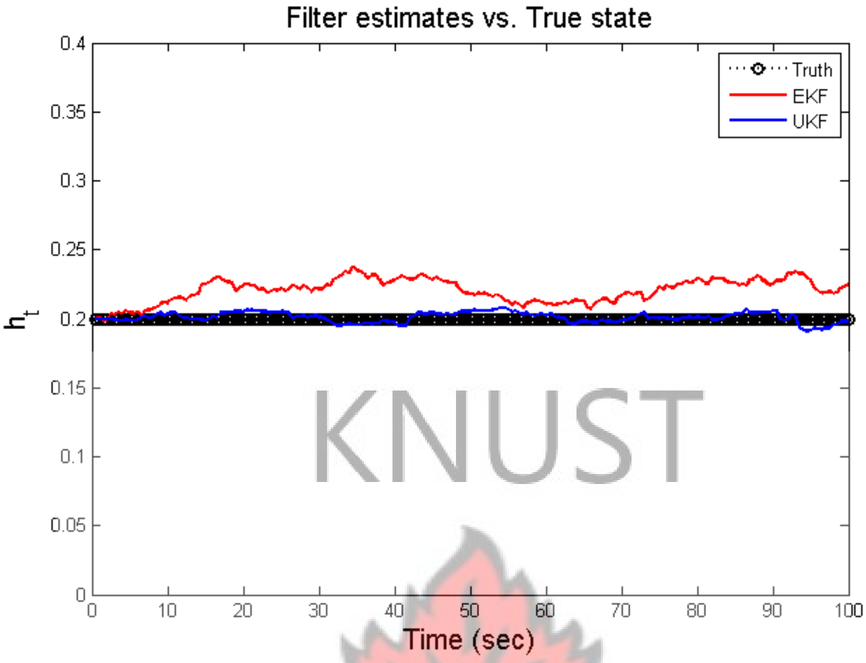

Figure 4.12: Estimation of the uncertain parameter h_t

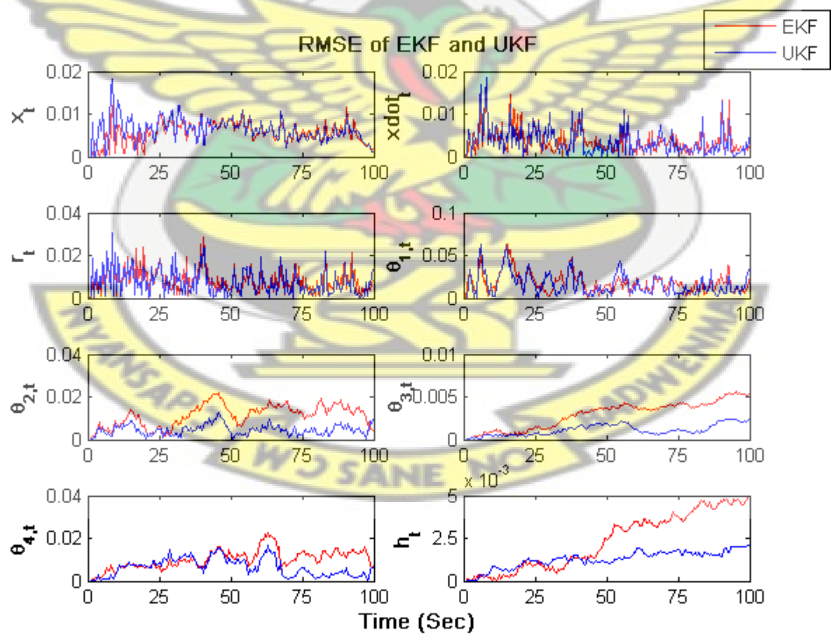

Figure 4.13: Time propagation of the RMSE of EKF and UKF

Table 4.1: Mean of the RMSE of the states and parameters

State / Parameter	Algorithm	Mean RMSE
x_t	EKF	0.0057104
	UKF	0.0123760
$\dot{x}_t$	EKF	0.0039881
	UKF	0.0035723
r_t	EKF	0.0080489
	UKF	0.0075815
$\theta_{1,t}$	EKF	0.0194950
	UKF	0.015327
$\theta_{2,t}$	EKF	0.0077102
	UKF	0.0048823
$\theta_{3,t}$	EKF	0.0016484
	UKF	0.0011671
$\theta_{4,t}$	EKF	0.0103990
	UKF	0.0068425
h_t	EKF	0.0014238
	UKF	0.0011282

Chapter 5

Conclusion and Recommendation

5.1 Conclusion

The trajectories and dynamics of a nonlinear state-space problem were successfully reconstructed using the UKF in section 3.4 of chapter 3. Varying the initial error covariance, process noise as well as measurement noise were further experiments performed to compare the performance of the UKF and EKF. The outcomes of these experiments revealed that the UKF estimated the state better than the EKF.

The next test performed was the feasibility of the Unscented Kalman filter in simultaneously estimating state and parameters. The Unscented Kalman filter successfully estimated the linear and the nonlinear systems considered in chapter 4. Results from the estimation of the states and parameters of the linear planer four-story shear building indicated that the performance of the UKF was approximately equal to that of the EKF whereas the UKF performed better than the EKF in the state and parameter estimation of the nonlinear SDOF system.

Results obtained from the estimation of the two nonlinear systems in this research indicated that the Unscented Kalman filter is able to estimate the state and parameters of nonlinear systems and actually produces better results than the famous Extended Kalman filter.

5.2 Recommendation

The Unscented Kalman filter is recommended as a more accurate and better method as compared to the Extended Kalman filter for nonlinear systems. It is also recommended that the process noise, measurement noise and the initial error covariance should be minimal when using these filters as they sometimes affect their performances. Finally, it is recommended that other factors that may affect the performance of the UKF and not considered in this thesis be investigated in further research. An example is to investigate the effect of varying observation frequency on the performance of the UKF.

References

Bouttier, F. and Courtier, P. (1999). *Data assimilation concepts and methods.*

Brown, R. G. and Hwang, P. Y. C. (1992). *Introduction to Random Signals and Applied Kalman Filtering.* John Wiley & Sons, Inc, second edition.

Carmichael, D. G. (1979). The state estimation problem in experimental structural mechanics. In *Application of Statistics & Probability in Soil and Structural Engineering*, number 3rd Ed., pages 802–15. International Conference.

Carydis, P. and Mouzakis, H. P. (1986). Small amplitude vibration measurement of buildings undamaged, damaged and repaired after earthquakes. *Earthquake Spectra*, 3:123–142.

Chen, C. K., Czarnecki, R. M., and Scholl, R. E. (1977). Vibration test of a 4-story concrete structure. In *Proc. 6th World Conference on Earthquake Engineering*, volume 3, pages 2753–2758, New Delhi, India. International Association of Earthquake Engineering.

Ching, J., Beck, J. L., and Porter, K. A. (2006). Bayesian state and parameter estimation of uncertain dynamical systems. *Probabilistic Engineering Mechanics*, 21:81–96.

Ching, J., Beck, J. L., Porter, K. A., and Shaikhutdinov, R. (2004). Real-time bayesian state estimation of uncertain dynamical systems. Technical Report

EERL 2004-01, Earthquake Engineering Research Laboratory, California Institute of Technology, Pasadena, California.

Distefano, N. and Rath, A. (1975). Sequential identification of hysteretic and viscous models in structural seismic dynamics. *Computer Methods in Applied Mechanics and Engineering*, 6(2):219 – 232.

Foutch, D. A. and Housner, G. W. (1977). Observed change in the natural period of vibration of a nine-story steel frame building. In *Proc. 6th World Conference on Earthquake Engineering*, volume 3, pages 2698–2704, New Delhi, India. International Association of Earthquake Engineering.

Ghanem, R. and Shinozuka, M. (1995a). Structural-system identification, i. theory. *Journal of Engineering Mechanics*, 121(2):255–264.

Ghanem, R. and Shinozuka, M. (1995b). Structural-system identification, ii: experimental verification. *Journal of Engineering Mechanics*, 121(2):265–273.

Glaser, S. D. (1996). Insight into liquefaction by system identification. *Geotechnique*, 46(4):641–655.

Grewal, M. S. and Andrews, A. P. (1993). *Kalman Filtering Theory and Practice.* Upper Saddle River, NJ USA, Prentice Hall.

Hoffman, J. D. (2001). *Numerical Methods for Engineers and Scientists.* Marcel Dekker, Inc., second edition.

Hoshiya, M. and Saito, E. (1984). Structural identification by extended kalman filter. *Journal of Engineering Mechanics*, 110(12):1757–1770.

Hoshiya, M. and Sutoh, A. (1993). Kalman filter - finite element method in identification. *Journal of Engineering Mechanics*, 119(2):197–210.

Iemura, H. and Jennings, P. C. (1973). Hysteretic response of a nine-story reinforced concrete building. Technical Report EERL 73-07, Earthquake Engineering Research Laboratory, California Institute of Technology, Pasadena, California.

Jacobs, O. L. R. (1993). *Introduction to Control Theory*. Oxford University Press, 2nd edition.

Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. Academic Press, New York.

Julier, S. J. and Uhlmann, J. K. (1996). A general method for approximating nonlinear transformations of probability distributions. Technical report, Department of Engineering Science, University of Oxford.

Julier, S. J. and Uhlmann, J. K. (1997). A new extension of the kalman filter to nonlinear systems. In *Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defence Sensing, Simulation and Controls*.

Julier, S. J. and Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. In *Proceedings of the IEEE*, volume 92.

Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. (1995). A new approach for filtering nonlinear systems. In *Proceedings of the American Control Conference*, pages 1628–1632.

Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F. (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–82.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(Series D):35–45.

Kalman, R. E. and Bucy, R. S. (1961). New results in linear filtering and prediction problems. *Journal of Basic Engineering*, 83(Series D).

Koh, C. G. and See, L. M. (1994). Identification and uncertainty estimation of structural parameters. *Journal of Engineering Mechanics*, 120:1219–1236.

Lahoz, W., Khattotov, B., and Ménard, R., editors (2010). *Data Assimilation: Making Sense of Observations*. Springer-Verlag Berlin Heidelberg.

Lin, C. C., Soong, T. T., and Natke, H. G. (1990). Real-time system identification of degrading structures. *Journal of Engineering Mechanics*, 116(10):2258–74.

Loh, C. H. and Tsaur, Y. H. (1988). Time domain estimation of structural parameters. *Engineering Structures*, 10:95–105.

Maybeck, P. S. (1979). *Stochastic Models, Estimation, and Control*, volume 1. Academic Press, Inc.

Meyer, C. and Roufaiel, M. S. L. (1984). Reliability of damaged concrete frames. In *Proc. 8th World Conference on Earthquake Engineering*, volume 4, pages 535–543, San Francisco, California. International Association of Earthquake Engineering.

Mihai, C., Manolovici, M., Carlan, S., Marinescu, S., and Iticovici, M. (1980). Static and seismic model behavior of a new type of mixed structure for reinforced concrete industrial buildings. In *Proc. 7th World Conference on Earthquake Engineering*, volume 7, pages 107–110, Istanbul, Turkey. International Association of Earthquake Engineering.

Ogawa, J. and Abe, Y. (1980). Structural damage and stiffness degradation of buildings caused by severe earthquakes. In *Proc. 7th World Conference on*

Earthquake Engineering, volume 7, pages 527–534, Istanbul, Turkey. International Association of Earthquake Engineering.

Sato, T. and Qi, K. (1998). Adaptive h_∞^+ filter: Its application to structural identification. *Journal of Engineering Mechanics*, 124(11):1233–1240.

Smyth, A. W., Masri, S. F., Chassiakos, A. G., and Caughey, T. K. (1999). On-line parametric identification of mdof nonlinear hysteretic systems. *Journal of Engineering Mechanics*, 125(2):133–142.

Udwadia, F. E. and Jerath, N. (1980). Time variations of structural properties during strong ground shaking. *Journal of the Engineering Mechanics Division*, 106(1):111–121.

van der Merwe, R., de Freitas, N., Doucet, A., and Wan, E. A. (2000). The unscented particle filter. Technical Report CUED / F-INFENG / TR 380, Cambridge University Engineering Department.

van der Merwe, R. and Wan, E. A. (2003). Gaussian mixture sigma-point particle filters for sequential probabilistic inference in dynamic state-space models. In *Proceedings of IEEE international conference on acoustics, speech and signal processing (ICASSP)*, Hong Kong.

Vasilescu, D. and Diaconu, D. (1980). Stiffness degradation and energy dissipation in reinforced concrete lamellar structures under earthquake excitation. In *Proc. 7th World Conference on Earthquake Engineering*, volume 6, pages 247–254, Istanbul, Turkey. International Association of Earthquake Engineering.

Wan, E., van der Merwe, R., and Nelson, A. (2000). Dual estimation and the unscented transformation. In *S.A. Solla, T.K. Leen, and K.-R. Müller, Eds. Advances in Neural Information Processing Systems 12*, pages 666–672. MIT Press, Cambridge, MA.

Wan, E. A. and van der Merwe, R. (2000). The unscented kalman filter for nonlinear estimation. In *Proceedings of Symposium 2000 on Adaptive Systems for Signal Processing, Communication and Control (AS-SPCC), IEEE*, Lake Louise, Alberta, Canada.

Welch, G. and Bishop, G. (2006). An introduction to the kalman filter. Technical Report TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3175.

Yun, C. B. and Shinozuka, M. (1980). Identification of nonlinear dynamic systems. *Journal of Structural Engineering*, 8(2):187–203.