Pricing Financial Options Using

Ensemble Kalman Filter

By

GASTON EDEM AWASHIE, B.Sc.

A Thesis submitted to the Department of Mathematics,

Kwame Nkrumah University of Science and Technology

in Partial Fulfillment of the Requirements for the Degree of

MASTER OF PHILOSOPHY

College of Science

April, 2012

Declaration

I hereby declare that this submission is my own work towards the award of Master of Philosophy (M.Phil.) 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.

Gaston Edem Awashie (PG5069710)	....	....
Student	Signature	Date

Certified by:

E. Owusu-Ansah	....	....
(Department of Mathematics)	Signature	Date

Supervisor

Mr. K. F. Darkwah ....

Head, Department of Mathematics Signature Date

Abstract

Pricing financial options is amongst the most important problems in the financial industry. In this study we investigated the use of the ensemble Kalman filter for pricing financial options in the Black-Scholes model. The performance and accuracy of the Ensemble Kalman Filter (EnKF) method based on a Monte Carlo simulation approach for propagation of errors is evaluated on two estimation problems. The first is a synthetic estimation problem using the Van der Pol equation and then a real-world estimation problem concerned with pricing financial instruments. The scenarios considered were to compare effect of different process noise, effect of different measurement noise, and the effect of different ensemble sizes on the performance and accuracy of the EnKF. It was found that as the ensemble size grows the performances of the ensemble Kalman filter improves judging from the values of the root mean square errors. With regard to the process noise, the measurement noise and the initial error covariance, decrease in the value of these parameters actually improves the performance of the EnKF. It was also found that the ensemble Kalman filter approaches same or better accuracy than the extended Kalman filter.

Dedication

With love to my

family, who

encourage me to

strive for success

in everything I

do and be

humbled the

higher I go.

Acknowledgements

My greatest thanks and appreciation goes to the Almight God for his grace and mercy that is keep me through out my life. To my supervisor, E. Owusu-Ansah, God richly bless you for the much help, the pressure you mounted on me and the positive criticisms, which have given me much strength and aided me in acquiring knowledge. My appreciation goes to my parents for the much sacrifice, advice and support given to me through out my studies. May the good Lord bless and keep you to see your grandchildren. To all my lectures, in the Department of Mathematics, KNUST, I am highly grateful for the knowledge imparted during my undergraduate and postgraduate studies. Last but not the least, I acknoledge all my postgraduate colleagues for their contribution.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The pricing of options is a very important problem encountered in financial domain. Options are financial instruments that give the right to buy or sell an underlying asset at a set price on or up to a given maturity date. They are widely traded in the financial market and so the need for methods to determine accurately the price of these financial instruments. Different formula have been developed for determining the value of an option. Adopted in this study is the Black-Scholes formula developed in 1973 popularly used for pricing call and put options in the market.

Different methods have been developed for estimation problems over the years. The Ensemble Kalman filter (EnKF) is one of such methods introduced by Evensen (1994). Clearly this method have been widely used and some comparisons on the Extended Kalman filter (EKF) and Ensemble Kalman filter have also been carried out. On the basis of the comparisons and findings, the EnKF is being proposed as an alternative to the EKF for the pricing of financial instruments.

1.1 Background Studies

Data assimilation is a powerful methodology which involves the combination of observational data with the underlying dynamical principles governing the system under observation. It therefore requires the estimation of the states of a dynamical system as a sequence of noisy observations becomes available. Data assimilation schemes are developed to use measured observations in combination with a dynamical system model in order to derive accurate estimates of the current and future states of the system, together with estimates of the uncertainty in the estimated states. The two approaches to data assimilation are the sequential and variational assimilation.

Sequential data assimilation methods have been proven to be very useful in many applications in finance, example is the pricing of financial options. For dynamical systems that are linear, the Kalman filter is an optimal sequential technique, it computes estimates with available observations together with underlying models by minimizing the variance. But most real-world estimation problems are nonlinear and the Kalman filter fails to accurately predict the estimates when the state-space equations are nonlinear.

The extended Kalman filter is a sequential assimilation scheme in which an approximate linearized equation is used for predicting the error statistics and its considered to be one of the most effective methods for both nonlinear state estimation and parameter estimation. In recent times, a number of effective methods have being develop as alternatives to the extended Kalman filter for handling estimation problems, some of which are the Unscented Kalman filter, and the Ensemble Kalman filter.

The basic idea of data assimilation is to quantify errors in both the model predictions and observations. The errors in the model predictions and observations are due to several reasons. Model errors may be due to violations in the assumptions of the model, use of uncertain or parameter values that are not optimal. Due to unavailability of exact solutions to underlying models, continuous dynamical systems are solved using numerical solution thereby transforming the model into a discrete dynamical systems. These computations also introduce errors in model predictions. Observations on the other hand may contain human errors (error from taking readings or collecting the data). Errors in observations could also be due to inaccuracy in the instruments used.

In solving the underlying models employed in this work, we used finite difference schemes for finding numerical solutions to the underlying models governing the systems. In pricing of financial instruments, the introduction of model and observation errors could cause the value of the underlying asset to be under-priced or over-priced. Data assimilation therefore tries to minimize these error and thereby giving more accurate predictions of the states and parameters. Schemes developed for data assimilation therefore address the issues of overpricing or underpricing financial instruments. It also helps in estimating the unknown parameters. This research therefore employs the ensemble Kalman filter for the pricing of financial options in the financial market. We therefore examine the performance of the ensemble Kalman filter on two estimation problems comparing its performance to that of the extended Kalman filter and the ensemble open loop (where there is no update of the model predictions).

1.2 Problem Statement

The Ensemble Kalman Filter is a Monte Carlo approximation of the Kalman filter, representing the distribution of the system state by using random samples (ensembles) and computes the covariance from the ensemble.

In this study, we investigate the performance of the ensemble Kalman filter which was introduced by Evensen (1994) when the process noise is varied keeping the other parameter constant, varying the measurement noise, varying the initial error covariance matrix and last but not the least the effect of different ensemble size. The performance of the EnKF is evaluated on two estimation problems. The first estimation problem being a synthetic estimation problem using the Van der Pol equation and the second is a real world estimation problem concerned with the pricing of financial instruments.

The Ensemble Kalman filter method is proposed as an alternative to the Extended Kalman filter method for the pricing of financial options in the Black-Scholes model on the British FTSE-100 index.

1.3 Objectives

The objective of the study is to examine the performance of the Ensemble Kalman filter and to propose the EnKF as an alternative to the EKF for pricing of financial instruments in the Black-Scholes model. The performance of the EnKF is verified for the following scenarios:

- compare the effect of different process noise (Q) on the filter
- compare the effect of different measurement noise (R) on the filter
- compare the effect of different initial error covariance (P_0) on the filter
- compare the effect of different ensemble size on EnKF

1.4 Methodology

The impact of the process noise, measurement noise, initial error covariance and the ensemble size on the performance and accuracy of the Ensemble Kalman filter is examined using two estimation problems. The EnKF is implemented on state-space representation of the estimation problem. The studies therefore employs the finite difference scheme for the discretization of the underlying differential equation defining the dynamics of the systems being considered, thus the Van der Pol model and the Black-Scholes model.

The algorithms employed in the study were implemented in MatLab and used in performing the various experiments. In the pricing of financial options in the Black-Scholes model, we adopted the British FTSE-100 index which was used by de Freitas et al. (2000), and Wan and van der Merwe (2001).

1.5 Justification

Undoubtedly, much studies have been carried out on the performance of the ensemble Kalman filter as a sequential data assimilation scheme. Data assimilation schemes and Monte Carlo approaches have been applied in the pricing of financial options. The extended Kalman filter is one of the widely used data assimilation schemes for estimating the state of nonlinear dynamical systems. Wan and van der Merwe (2001) proposed the unscented Kalman filter as an alternative to the Kalman filter for pricing financial options. So far what has not been done is implementing the ensemble Kalman filter (a Monte Carlo approach) for the pricing of financial options. The research was investigate how accurately the ensemble Kalman filter can predict the value of derivatives in the financial market.

1.6 Outline of The Thesis

This thesis is presented in five chapters. Chapter 1 introduces the research work by giving the background of the study, problem statement, objectives of the study, methodology employed in the research and the justification of the study.

Review of literature and the framework of the study is presented in Chapter 2 where related studies are discussed. Chapter 3 talks about the methodology employed in the study. This chapter also contains results from the synthetic estimation problem.

Chapter 4 presents and discusses the results and finding from the real-world estimation problem concerned with pricing of financial instruments. Finally the summary of our findings, conclusion and recommendation are presented in the last chapter.

Chapter 2

Literature Review

2.1 Introduction

Derivatives are financial instruments whose value depends on some basic underlying cash products, such as interest rate, equity indices, commodities, foreign exchange, or bonds, (Hull, 2006). An option is an example of a derivative. Black and Scholes (1973) derived the standard equations for pricing European call and put options and there on there have been many models developed for pricing financial options. Many have also work extensively to obtain solution to the Black-Scholes model. The chapter present literature overview of related work in the field and the overview of the thesis is provided.

2.2 Options

An option is a contract that gives the right but not the obligation, to buy or to sell a set quantity of a particular asset at a set price on or up to a given maturity date. A call option gives the holder the right to buy the underlying asset by a certain date (expiration or maturity date) for a certain price (exercise or strike price). A put option gives the holder the right to sell the underlying asset by a certain date for a certain price. If the option can be exercised at any time up to its expiration, it is called American option; if it can be exercised only at its expiration, it is a European option, (Hull, 2006).

According to Hull (2006), there are six major factors affecting stocks option: the current stock price (S_o), the strike price (K), the time of expiration (T), the volatility of the stock price (σ), the risk free interest rate (r), and the dividends expected during the life of the option. Table 2.1 describe how the value of these factors affect the value of options.

Table 2.1: Determinants of option value

Factors	**Call Value**	**Put Value**
Increase in stock price	Increases	Decreases
Increase in strike price	Decreases	Increases
Increase in time of expiration	Increases	Increases
Increase in volatility	Increases	Increases
Increase in interest rate	Increases	Decreases
Increase in dividends	Decreases	Increases

An option is a financial instrument which is an example of a derivative. Derivative financial secrurities such as options are securities whose value is based upon the value of more basic underlying securities, (Bolia and Huneja, 2005). The strike price is specified when the holder and seller enter into the option. If the holder chooses to exercise the option, the writer is obligated to sell or buy the asset at the strike price or the option is allowed to expire.

2.3 Pricing Financial Options

The pricing of options is a very important problem encountered in financial domain. Options are widely traded on financial markets, and so some method of determining the value (or price) of a given option is required. The determination of the price of an option requires the construction of a model for the way in which the asset price changed over time. One of the most popular models is the Black-Scholes model (Black and Scholes, 1973) for pricing a European put and call option.

Black and Scholes (1973) derived a theoretical valuation formula for option pricing. The formula was based on the principle: if options are correctly priced in the market, it should not be possible to make sure profits by creating portfolios of long and short positions in options and their underlying stocks. The model is applicable to corporate liabilities such as common stock, corporate bonds, and warrants since almost all corporate liabilities can be viewed as combination of options. In particular, the Black-Scholes formula can be used to derive the discount that should be applied to a corporate bond because of the possibility of default.

Rubinstein (1983) in his work developed an option pricing formula that pushes the underlying source of risk back to the risk of individual assets of the firm. Relative to the Black-Scholes formula, the displaced diffusion formula has several desirable features. The formula encompasses differential riskiness of the assets of the firm, their relative weights in price determination of the firm, the effects of firm debt and the effects of a dividend policy with constant and random components.

Heston (1993) used a new technique, based on the Black-Scholes formula, to derive a closed-form solution for the price of a European call option on an asset with stochastic volatility. The model allows arbitrary correlation between volatility and spot asset returns. He introduced stochastic interest rate and showed how the model is applicable to bond options and foreign currency options. The result from his work showed that correlation between volatility and the spot asset price is important for explaining return skewness and strike price biases in the Black-Scholes model (Black and Scholes, 1973).

Kumar et al. (2012) obtained an analytic solution of the fractional Black-Scholes European option pricing equation. The Laplace homotopy perturbation method, a combined form of the Laplace transform and the homotopy perturbation method, was used with boundary condition for a European option pricing problem to obtain a quick and accurate solution to the fractional Black-Scholes equation. The analytic solution of the fractional Black-Scholes equation was calculated in the form of a convergent power series with easily computable components.

Gallant et al. (1992), in their work investigated the joint dynamics of price changes and volume on the stock market making use of daily data on the S&P composite index and total NYSE trading volume from 1928 to 1987. Nonparametric methods were used to achieve the set objectives. Gallant et al. (1992) found out that the daily trading volume is positively and nonlinearly related to the magnitude of the daily price change and that price changes lead to volume movements.

Pastorello et al. (2000) delt with the estimation of continuous-time stochastic volatility models of option pricing. They achieved this in a Monte Carlo experiment which compared two very simple strategies based on different information sets. An Ornstein-Uhlenbeck process for log of the volatility, a zero-volatility risk premium, and no leverage effect was assumed. Sticking to the framework with no overidentifying restrictions, it was shown that, given the option pricing model, estimation based on option prices is much more precise in samples of typical size.

2.4 Review of Data Assimilation

Systems can be described by mathematical models of the system dynamics which can be used to predict the future behaviour of the system. Both the models and the available data defining all the state of the system contain inaccuracies and random noise that can lead to significant differences between the predicted states and the actual states of the system. In view of this, observations of the system over time can be incorporated into the model equations to derive improved estimates of the states and also to provide information about the uncertainty in the estimates.

The concept of combining observation with model predictions to obtain improved estimates is referred to as *data assimilation*. Data assimilation schemes are therefore developed to use measured observations in combination to a dynanical system model in order to derive accurate estimates of the current and future states of the system, together with estimates of the uncertainty in the estimated states, (e.g., Kalman (1960); Evensen (1994); Ott et al. (2004)).

By Bouttier and Courtier (1999), data assimilation is 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.

Model error also arise from inaccurate parameters in the model equations. Data assimilation can be used in this case to estimate the parameters (e.g., Evensen et al. (1998); Annan et al. (2003); Etienne and Dombrowsky (2002); Annan and Hargreaves (2004)), and it may also be a dual estimation problem where both the states and parameters are estimated simultaneously, (e.g., Evensen (2009)).

There are two approaches to the problem of data assimilation; sequential and variational assimilation. *Sequential* assimilation considers observation made in the past until the time of analysis, which is the case of real-time assimilation systems, and *variational* assimilation where observation from the future can be used, for instance in a reanalysis exercise. Variational assimilation observations can be processed in small batches, (Bouttier and Courtier, 1999).

This research considers the sequential data assimilation. The Kalman filter was developed by Rudolf Kalman as a recursive solution to the discrete-data linear filtering problem, (Kalman, 1960). One major limitation of this method being that it is only applicable to linear systems. The extended Kalman filter was then developed to take care of this limitation. Thus for nonlinear dynamics, the extended Kalman filter may be used in estimating the state. The extended Kalman filter used a linearized equation for the error covariance evolution, and this linearization can result in unbounded linear instabilities for the error evolution.

The ensemble Kalman filter (a sequential data assimilation method) introduced by Evensen (1994) is a Monte Carlo approximation to the Kalman filter. The EnKF used an an ensemble of forecasts to estimate background-error covariances.

2.5 Application of Data Assimilation in Pricing Financial Options

Several study have been conducted on how data assimilation can be used to price financial options in the stock market. Indragoby and Pironneau (2004) in their study presented a way to calibrate the volatility surface of Black-Scholes and Dupitre's financial model from market observed values of a set of European call options. They used a standard data assimilation technique for this inverse problem representing the volatility surface with splines.

Gupta and Reisinger (2012) considered a general calibration problem for derivative pricing models, which they reformulated into a Bayesian framework to attain posterior distributions for model parameters. Thus introducing the Bayesian framework for calibrating the parameters of financial models to market prices. It was then shown that the posterior distribution can be used to estimate prices for exotic options. They also highlighted the robustness of the pricing method to inaccuracies in the model and prior and mispricings in observed market data. Gupta and Reisinger (2012) then applied the procedure to a discrete local volatility model and worked in great detail through numerical examples to clarify the construction of Bayesian estimators and their robustness to the model specification, number of calibration products, noisy data and misspecification of the prior.

The application of data assimilation to a term structure of commodity prices was carried out by Javaheri et al. (2002). Different assimillation schemes were considered: the Kalman filter, the extended Kalman filter, the particle filter, as well as the unscented Kalman filter. They tackled the subject of Non-Gaussian filters and described the Particle filtering algoeithm. They found out that the approximation introduced in the Extended filter influences the model performance and the estimation results are sensitive to the system containing the errors of the measurement eqaution. The application of the filters to stochastic volatility models show that the Particle filter performs better than the Gaussian ones, however they are also more computationally expensive.

To creat an adptive model, Nguyen and Nabney (2010) used the extended Kalman filter and the particle filter to update the parameters continuouly on the test set. Some forecasting techniques for energy demand and price prediction, one day ahead was presented by Nguyen and Nabney (2010) which combines wavelet transforms with fixed and adaptive machine learning/time series models. The techniques are applied to large sets of real data from UK energy market.

Nikolaev and Smirnov (2009) developed an unscented grid-based filter and a smoother for accurate nonlinear modelling and analysis of time series. They conducted an empirical investigation which show that the Unscented Grid Filter with a Smoother compares favourably ti similar filters on modeling nonstationary series and option pricing modelling. They considered modelling prices by treating the implied volatility and interest rate as unobservables assuming the volatility, interest rate and prices are to be Gaussian with their noises unknown. The filters were applied with two inputs: the stock price divided the strike price and the time to maturity. The ouputs being the call and put prices normalised with respect to the strike price.

Many problems such as time-series analysis, are characterized by data that arrive sequentially. Performing model estimations, model validation, and inference sequentilly are necessary on arrival of each item of data. de Freitas et al. (2000) presents sequential

Monte Carlo algoeithms and proposed a new hybrid gradient descent/sampling importance resampling algorithm (HySIR). They discussed two experiments, the first using a real synthetic data and secondly a real life application involving the pricing of financial options on the FTSE-100 index. The latter experiment used the HySIR algorithm in estimating timevarying latent variables and illustrated the effect of the process noise covariance on the simulations. de Freitas et al. (2000) in their experiment with financial data found that the sequential sampling approach in addition to the one-step-ahead square errors of the best available pricing methods, allowed getting complete estimates of the probability density functions of the one-step-ahead predictions.

Wan and van der Merwe (2001) pointed out the underlying assumptions and flaws in the Extended Kalman filter and proposed the unscented Kalman filter as an alternative with superior performance to that of that of the EKF. The algorithm was applied to several area including the pricing of financial options. The Black-Scoles model for pricing European call and put options (Black and Scholes, 1973) was used in the experiment. The prediction were made making use of five pairs of call and put option contracts on the British FTSE-100 index. Their results show that the UKF consistently achieves an equal or better level of performance at a comparable level of complexity.

We attempt to propose as an alternative to the extended Kalman filter developed to for state as well as parameter estimation with equations being nonlinear. The proposed method is the ensemble Kalman filter which was introduced by Evensen (1994). Two experiments were setup, the first uses synthetic data to predict the dynamics in the Van der Pol model and the second is a real-world application involving the pricing of financial options on the British FTSE-100 index making use of the Black-Scholes model, (Black and Scholes, 1973). The experimental results are compared to the results from the extended Kalman filter. The effect of the process noise, measurement noise and initial error covariance on these filters is investigated. The effect of the number of ensembles on the ensemble Kalman filter is also investigated.

Chapter 3

Methodology

3.1 Introduction

The chapter discusses the methods used in accomplishing the objectives of the research work. In this chapter, the numerical solution to partial differential equations is discussed. Secondly, we looked at Data assimilation and its types. The chapter considered the Kalman filter, extended Kalman filter and ensemble Kalman filter methods, as sequential data assimilation schemes. Last but not the least, we investigate the performance of the extended Kalman filter and ensemble Kalman filter in state estimation of the Van der Pol model.

3.2 The Black-Scholes Model for Pricing Financial Options

The famous Black-Scholes model has been used as the foundation for option pricing. The concept of dynamic helding was introduced by Black and Scholes (Black and Scholes, 1973) and Merton (Merton, 1973), so that the option payoff will be replicated by a trading strategy in the underlying asset.

Before Black and Scholes set their famous Black-Scholes model in 1973, most of the previous work on the valuation of options has been presented in terms of warrants. The Black-Scholes model is used for valuing European or American call and put options on a non-dividend paying stock. Despite subsequent development of option theory, the original Black-Scholes formula for European option remains the most successful and widely used application. This formula is particularly useful because it relates the distribution of spot returns to the cross-sectional properties of option prices.

Black and Scholes (1973) derived the famous theoretical valuation formula for options. The main conceptual idea of Black and Scholes lie in the construction of a riskless portfolio taking positions in bonds (cash), option and the underlying stock.

In formulating the Black-Scholes model the following assumptions were made in the market for the stock and the option: It is assumed in the Black-Scholes model that

1. the stock price follows a random walk in continuous time with a variance rate proportional to the square of the stock price. Thus the distribution of possible stock prices at the end of any finite interval is lognormal.

2. there are no transaction costs in buying or selling the options or stock.

3. the short-term interest rate is known and is constant through time.

4. the stock pays no dividends or other distributions.

5. it is possible to borrow any fraction of the price of security to buy it or to hold it, at the short-term interest rate

6. there are no restriction on short sale and there are no taxes.

The value of the option will depend only on the price of the stock, time and on variables that are taken to be known constants (Black and Scholes, 1973). The Black-Scholes model for the value of an option is described by the partial differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{3.1}$$

where $V(S,t)$ is a European call or put option at asset price S and at time t, $r(t)$ is the risk free interest rate, and σ representing the volatility of underlying asset.

Given the Black-Scholes model of a financial market, the Black-Scholes formula for European call options (C_t), and put options (P_t) on a non-dividend paying stock are

$$C_t = S_t N(d_1) - Xe^{rT} N(d_2) \tag{3.2}$$

$$P_t = -S_t N(-d_1) - Xe^{rT} N(-d_2) \tag{3.3}$$

where X the desired strike price, T is the time until expiration. $N(.)$ is the cumulative normal distribution function and d_1 and d_2 are given by

$$d_1 = \frac{\ln(S_t/X) + (r + \sigma^2/2)t_m}{\sigma\sqrt{t_m}}$$
$$d_2 = \frac{\ln(S_t/X) + (r - \sigma^2/2)t_m}{\sigma\sqrt{t_m}}$$

3.3 Numerical Solution to Parabolic Partial Differential Equations

The section considers finite difference method for solving partial differential equations (PDE). The finite difference is a powerful method for solving PDEs. The objective of a finite difference method for solving PDE is is to transform a calculus problem into an algebraic problem by

1. Discretizing the continuous physical domain into a discrete difference grid
2. Approximating the individual exact partial derivatives in the partial differential equation by algebraic finite difference approximations
3. Substituting the finite difference approximation into the PDE to obtain an algebraic finite difference equation.
4. Solving the resulting algebraic finite difference equations

There are several choices to be made when developing a finite difference solution to a partial differential equation: the choice of the discrete finite difference grid to discretized the continuous physical domain and the choice finite difference approximations used to represent the the individual exact partial derivatives in the PDE.

Depending on the approximation used, the most popular methods are the explicit method, implicit method and the Crank-Nicolson method which is a special case of the θ method.

The finite difference approximations to the derivatives are mostly derived from the Taylor series expansion.

To demonstrate the finite difference method, we use the weighted finite difference approximation or the theta-scheme in this section. Consider the one-dimensional parabolic equation given by

$$\frac{\partial V(x,t)}{\partial t} = a(x,t)\frac{\partial^2 V(x,t)}{\partial x^2} + b(x,t)\frac{\partial V(x,t)}{\partial x} + c(x,t)V(x,t) + f(x,t) \tag{3.4}$$

subject to the Dirichlet boundary conditions given as

$$V_0^k = \gamma_1(t_k) \qquad V_N{}^k = \gamma_2(t_k) \qquad t_0 \le t \le t_f \qquad x_0 \le x \le x_f$$

The finite difference approximations to the derivatives are given as

$$\frac{\partial V}{\partial t} \approx \frac{V_i^{k+1} - V_i^k}{r} \qquad \frac{\partial V}{\partial x} \approx \frac{V_{i+1}^k - V_{i-1}^k}{2h} \qquad \frac{\partial^2 V}{\partial x^2} \approx \frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{h^2} \tag{3.5}$$

Substituting the approximations in equation (3.5) into equation (3.4) we obtain the resulting equation the expression on the right hand side as

$$G_h = a_j^i\left(\frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{h^2}\right) + b_j^i\left(\frac{V_{i+1}^k - V_{i-1}^k}{2h}\right) + c_i^k V_i^k + f_i^k \tag{3.6}$$

Weighting the expression in the above equation, we obtain an expression for finite difference approximation to the PDE in equation (3.4).

$$
\begin{aligned}
\frac{V_i^{k+1}-V_i^k}{r} &= \theta a_i^{k+1}\left(\frac{V_{i+1}^{k+1}-2V_i^{k+1}+V_{i-1}^{k+1}}{h^2}\right)+(1-\theta)a_i^k\left(\frac{V_{i+1}^{k}-2V_i^{k}+V_{i-1}^{k}}{h^2}\right)\\
&+ \theta b_i^{k+1}\left(\frac{V_{i+1}^{k+1}-V_{i-1}^{k+1}}{2h}\right)+(1-\theta)b_i^k\left(\frac{V_{i+1}^{k}-V_{i-1}^{k}}{2h}\right)+\theta c_i^{k+1}V_i^{k+1}\\
&+ (1-\theta)c_i^kV_i^k+\theta f_i^{k+1}+(1-\theta)f_i^k
\end{aligned} \tag{3.7}
$$

Simplifying, we obtain a obtain the theta scheme for the discritization of the partial differential equation.

$$
\begin{aligned}
&\left(-\frac{r\theta a_i^{k+1}}{h^2}+\frac{r\theta b_i^{k+1}}{2h}\right)V_{i-1}^{k+1}+\left(1+\frac{2r\theta a_i^{k+1}}{h^2}-r\theta c_i^{k+1}\right)V_i^{k+1}+\left(-\frac{r\theta a_i^{k+1}}{h^2}-\frac{r\theta b_i^{k+1}}{2h}\right)V_{i+1}^{k+1}\\
=&\left(\frac{r(1-\theta)a_i^k}{h^2}-\frac{r(1-\theta)b_i^k}{2h}\right)V_{i-1}^k+\left(1-\frac{2r(1-\theta)a_i^k}{h^2}+r(1-\theta)c_i^k\right)V_i^k\\
+&\left(\frac{r(1-\theta)a_i^k}{h^2}+\frac{r(1-\theta)b_i^k}{2h}\right)V_{i+1}^k+r\theta f_i^{k+1}+r(1-\theta)f_i^k
\end{aligned} \tag{3.8}
$$

The following variables are introduced

$$
\begin{aligned}
A_i^k &= \frac{ra_i^k}{h^2}-\frac{rb_i^k}{2h} \qquad B_i^k=-\frac{2ra_i^k}{h^2}+rc_i^k\\
C_i^k &= \frac{ra_i^k}{h^2}+\frac{rb_i^k}{2h} \qquad R_i^k=r\theta f_i^{k+1}+r(1-\theta)f_i^k
\end{aligned} \tag{3.9}
$$

Substituting these into equation 3.8, we get

$$
\begin{aligned}
&- \theta A_i^{k+1}V_{i-1}^{k+1}+\left(1-\theta B_i^{k+1}\right)V_i^{k+1}-\theta C_i^{k+1}V_{i+1}^{k+1}\\
&= (1-\theta)A_i^kV_{i-1}^k+\left(1+(1-\theta)B_i^k\right)V_i^k+(1-\theta)C_i^kV_{i+1}^k+R_j^k
\end{aligned} \tag{3.10}
$$

Making use of the conditions imposed on the problem, we now have $N - 1$ equations with $N - 1$ unknowns. The boundary condition plays an important role in the finite difference discretization.

At this instance we note that if $\theta = 0$ we get the explicit finite

difference scheme, $\theta = 0.5$ we get the Crank-Nicolson finite

difference scheme, $\theta = 1$ we get the implicit finite

difference scheme

The explicit method has disadvantage of certain stability conditions. Thus the method is conditionally stable. The implicit methods on the other hand are absolutely stable.

The resulting algebraic system of equations can be solved by way of the iterative schemes: Gauss Jacobi method, Gauss-Seidel method or Relaxation method. One can also make use of the direct methods of finding numerical solution solution to these problems.

3.4 Data Assimilation

The basic idea of data assimilation is to quantify errors in both the model predictions and observations, and update model estimates in a way that optimally combines model simulations with observations.

Data assimilation involves the combination of observational data with the underlying dynamical principles governing the system under observation. Data assimilation scheme are tools used to estimate the quantities of interest of a dynamical system for a sequence of available noisy observations. This helps in obtaining more accurate estimates of these quantities both at current and future times of the dynamical system together with estimates of the uncertainties in the estimated quantities.

Data assimilation is a powerful methodology, which makes possible efficient, accurate and realistic estimations which may improve forecasting or modeling and increase the physical understanding of the system under consideration. Data assimilation is an analysis technique in which the observed information is accumulated into the model state by taking advantage of consistency constraints with laws of time evolution and physical properties.

A data assimilation system is composed of a set of observations, a dynamical model and a data assimilation scheme. The process may improve the forecasting and increase the physical understanding of the system under consideration, thus providing better estimates than can be obtained by only the data or the model.

A variety of models is used to describe the dynamical systems that arise in various fields. These range from simple linear, deterministic differential equations to sophisticated nonlinear stochastic partial differential continuous or discrete models. In this case the equations modeling the dynamical system uniquely determines the state of the system at all times for any given initial state and inputs, and this is referred to as the perfect model assumption.

There are two basic approaches to data assimilation: sequential assimilation, that only considers observation made in the past until the time of analysis, which is the case of realtime assimilation systems, and variational assimilation, where observation from the future can be used, for instance in a reanalysis exercise.

In variational data assimilation, the past observations until the present time, are used simultaneously to obtain the best estimates of the state.

3.5 Sequential Data Assimilation

Sequential data assimilation is the type of data assimilation in which observations are used soon as they become available to correct the present state of the model. Sequential methods leads to discontinuities in the time series of the corrected state.

There are several schemes for sequential data assimilation, some of which are Kalman filter, Extended Kalman filter and Ensemble Kalman filter. For linear dynamics the optimal sequential technique is the Kalman filter. The Extended Kalman filter, in which an approximate linearized equation is used for the prediction of error statistics, as well as the Ensemble Kalman filter are employed when the dynamical system in nonlinear.

3.6 The Kalman Filter Method

The Kalman filter method is a sequential data assimilation scheme which provides a recursive solution for state estimation of linear dynamical systems from a series of noisy measurements. It is a variance-minimizing algorithm that updates the state estimate whenever measurements become available. The update equation in the Kalman Filter are normally derived by minimizing the trace of the error covariance matrix.

The Kalman filter method is an efficient data assimilation method that explicitly accounts for the dynamical propagation of errors in the model. For linear models with known statistics of the system and measurement errors, the Kalman filter provides an optimal estimate of the state of the system, in terms of minimum estimation error covariance.

The method is named after Rudolph E. Kalman who published his famous paper that describes a recursive solution to the linear filtering problem (Kalman, 1960). The scheme can be applicable for state estimation if the state-space equations are linear. That is, if

$$x_k = F_k x_{k-1} + \omega_k \tag{3.11}$$

where F_k is a linear model operator (transition matrix) relating the state at previous time step x_{k-1} to the state at current time step x_k and ω_k being the process noise. The observation equation is given as

$$y_k = H_k x_k + \nu_k \tag{3.12}$$

where H_k is the measurement operator relating the state x_k to the observation y_k, ν_k being the measurement noise.

The Kalman filter method addresses the problem of trying to estimate the state $x \in R^n$ of a dynamical process that is governed by linear state-space dynamics.

3.6.1 Assumptions of the filter

The following assumption were made in the Kalman filter

1. The state dynamics and the observation process are linear.
2. The state noise ω_k and observation noise ν_k are sequences of white, zero mean, Gaussian noises.
3. x_0, ω_k and ν_k are uncorrelated.
4. $x_0 \sim (x_0^+, P_0^+)$

3.6.2 Derivation of Kalman Filter Equations

The Kalman equations are classified into two namely the forecast (time update) equations and measurement update equations. The forecast equations are used to integrated forward thereby providing estimates of the state and error covariance (a priori estimates), until observation becomes available and the measurement update equations used to update the state and error covariance estimates from the forecast step, (Welch and Bishop, 2001).

We denote $x^f_k \in \mathbb{R}^n$ to be the a priori state estimate at step k given knowledge of the process prior to step k, and x^a_k to be the a posteriori state estimate at step k given measurement y_k. The state error vectors for both the a priori and a posteriori estimates are given as

$$e_{fk} = x_k - x_{fk} \qquad e_{ak} = x_k - x_{ak} \tag{3.13}$$

The a priori estimate error covariance and the a posteriori estimate error covariance is then

$$P_k^f = E[e^f_k(e^f_k)^T] \qquad P_k^a = E[e_k^a(e_k^a)^T] \tag{3.14}$$

In deriving the Kalman filter equations, we find an equation that computes the a posteriori estimate, x^a_k as a linear combination of an a priori estimate x^f_k and a weighted difference between an actual measurement and a predicted measurement and this is demonstrated by equation (3.15)

$$x^a_k = x^f_k + K_k(y_k - y_k^f) \tag{3.15}$$

where K_k is referred to as the Kalman gain. The difference $(y_k - y_k^f)$ is referred to as the measurement innovation, or the residual which reflects the discrepancy between the predicted measurement y_k^f and the actual measurement y_k. The a priori state estimate equation can be obtained as

$$x^f_k = E[x_k|z_{k-1}] = E[(F_k x_{k-1} + \omega_k)|z_{k-1}]$$

$$= F_k E[x_{k-1}|z_{k-1}]$$

$$= F_k x^a_{k-1} \quad (3.16)$$

The error, e^f_k, associated with the forecast is then estimated and is given by

$$\begin{aligned} e^f_k &= x_k - x^f_k \\ &= F_k x_{k-1} + \omega_k - F_k x^a_{k-1} \\ &= F_k(x_{k-1} - x^a_{k-1}) + \omega_k \\ &= F_k e^a_{k-1} + \omega_k \end{aligned}$$

where e^a_{k-1} is the error associated with a posteriori estimate at the previous time step.

The degree of uncertainty of the a priori estimate is verified by computing the error covariance matrix, P_k^f, of the a priori estimate. This is given as

$$\begin{aligned} P^f_k &= E[e^f_k (e^f_k)^T] \\ &= E[(F_k e^a_{k-1} + \omega_k)(F_k e^a_{k-1} + \omega_k)^T] \\ &= F_k P^a_{k-1} F^T_k + Q_k \end{aligned} \quad (3.17)$$

Thus given a linear dynamical model, the Kalman filter prediction for the state and its corresponding error covariance matrix for the model errors are obtained respectively from equations (3.16) and (3.17).

Whenever measurements are made available at the kth time step, the a priori estimate is then updated and this is known as the a posteriori estimate. The expected value of the observation, given available observations to the $(k-1)$th time step can be obtained as

$$
\begin{aligned}
E[y_k|z_{k-1}] &= E[H_k x_k + \nu_k|z_{k-1}] \\
&= H_k E[x_k|z_{k-1}] \\
&= H_k x_{fk}
\end{aligned}
$$

Thus, given z_{k-1} to be all available observations upto the $(k-1)$th step,

$$E[y_k|z_{k-1}] = H_k x^f_k \tag{3.18}$$

We then use equations to estimate the innovation which is given by $y_k - y_k^f$. The equation for the innovation is given as

$$y_k - y_{kf} = H_k e_{fk} + \nu_k \tag{3.19}$$

and the corresponding covariance matrix computed as

$$\Omega_k = H_k P_k^f H_k^T + R_k \tag{3.20}$$

The corresponding error in the a posteriori estimate is computed as

$$
\begin{aligned}
e_k^a = x_k - x_k^a &= x_k - x_k^f - K_k[H_k(x_k - x_k^f) + \nu_k] \\
&= (I - K_k H_k)(x_k - x^f_k) - K_k \nu_k
\end{aligned}
$$

giving rise to an uncertainty in the a posteriori estimate

$$E[e^a{}_k(e^a{}_k)^T] = E[\{(I - K_kH_k)(x_k - x_k{}^f) - K_k v_k\}\{(I - K_kH_k)(x_k - x^f{}_k) - K_k v_k\}^T]$$

$$= (I - K_kH_k)P_{kf}(I - K_kH_k)_T + K_kR_kK_{kT}$$

Therefore

$$P_k^a = (I - K_kH_k)P_k^f(I - K_kH_k)^T + K_kR_kK_k^T \tag{3.21}$$

The Kalman is a variance minimizing algorithm which is used to update the state whenever measurements become available. Therefore the criterion is to minimize the sum of the variances which is just the same as minimizing the trace of the covariance matrix, (Sorenson,

1970).

This can be achieved by using the method of least squares, which helps in minimizing the trace of the covariance matrix P_k^a by taking the derivative of the trace with respect to K_k and setting the result to zero.

$$\begin{aligned} min \quad tr(P_k^a) &= tr[(I - K_kH_k)P_k^f(I - K_kH_k)^T + K_kR_kK_k^T] \\ \frac{\partial tr(P_k^a)}{\partial K_k} &= 2(I - K_kH_k)P_k^f(-H_k^T) + 2K_kR_k \end{aligned}$$

$$= -P_{kf}H_{kT} + K_kH_kP_{kf}H_{kT} + K_kR_k = 0$$

The resulting equation is solved to obtain an estimate for the Kalman Gain K_k. The process gives the expression for the Kalman gain to be

$$K_k = P_k^f H_k^T \left(H_k P_k^f H_k^T + R_k \right)^{-1} \tag{3.22}$$

Welch and Bishop (2001), looking at equation (3.22), noticed that as the measurement error covariance R_k approaches zero, the gain K_k weights the residual more heavily. On the other hand, as the a priori estimate error covariance P_k^f approaches zeros, the gian weights the residual less heavily. These are illustrated in the following equations:

$$\lim_{R_k \to 0} K_k = H^{-1} \quad \text{and} \quad \lim_{P_k^f \to 0} K_k = 0$$

Simplifying further, equation (3.21), with results from the method of least squares, gives rise to the a posteriori covariance matrix given as follows

$$\begin{aligned} P_k^a &= (I - K_k H_k) P_k^f (I - K_k H_k)^T + K_k R_k K_k^T \\ &= P_k^f - P_k^f H_k^T K_k^T - K_k H_k P_k^f + P_k^f H_k^T K_k^T \\ &= (I - K_k H_k) P_k^f \end{aligned}$$

The Kalman filter algorithm determines the analyzed estimate by a linear combination of the measurement vector, y_k and the forecast model state vector, x^f_k. The linear combination is chosen to minimize the variance in the analyzed estimate, x^a_k.

Algorithm 1 *Kalman Filter*

1. Set initial estimates for x_0^a and P_0^a
2. For *k* = 1 to maximum number of iterations
3. For *i* = 1 to *n* Time Update

Project the state forward

$$x_k^f = F x_{k-1}^a$$

Project the error covariance forward

$$P_k^f = F_k P_{k-1}^a F_k^T + Q_k$$

4. End of loop for *i* Measurement Update

Compute the Kalman gain

$$K_k = P_k^f H_k^T \left(H_k P_k^f H_k^T + R_k \right)^{-1}$$

Update state estimate with measurement y_k

$$x_k^a = x_k^f + K_k \left(y_k - y_k^f \right)$$

Update the error covariance

$$P_k^a = \left(I - K_k H_k \right) P_k^f$$

5. End of loop for *k*

3.6.3 Some Drawbacks of the Kalman Filter

The Kalman filter works only for linear dynamical systems and is the best optimal filter. But most dynamical systems are nonlinear and in such cases the Kalman filter is impracticable in obtaining estimates of the state and the propagation of the error covariance matrix. Also, the Kalman filter is impracticable in high-dimensional systems due to the huge computational load and storage requirements associated with the propagation of error covariance matrix.

In the case of non-linear model dynamical systems, an approximate Kalman filter algorithm, extended Kalman filter, in which the error propagation is based on linearization of the model equation can be adopted

3.7 Extended Kalman Filter

The Kalman filter method addresses the general problem of trying to estimate the state of processes governed by linear equations. But what happens if the process to be estimated and (or) the measurement relationship to the process is nonlinear? To address this problem a Kalman filter that linearizes about the current mean and covariance was developed. This method is referred to as the extended Kalman filter method.

The idea behind extending the Kalman filter to nonlinear is to perform a Taylor series expansion of the model and measurement functions at each time step and propagate the error covariance matrix along the truncated series.

The Extended Kalman filter is applied for nonlinear dynamical system in which case an approximate linearized equation is used for the prediction of the error statistics. It is a set of mathematical equations which uses underlying process models to make estimate of the current state and then corrects the estimate using any available measurement.

The estimation can be linearized around the current estimate by the use Taylor series. Thus, using the partial derivative of the process and measurement functions to compute estimates. The process is assumed to have a state vector $x \in R^n$, governed by a nonlinear equation represented by

$$x_k = f(x_{k-1}) + \omega_k \tag{3.23}$$

and the measurement equation given as

$$y_k = h(x_k) + v_k \tag{3.24}$$

The degree of nonlinearity does not only depend on the physics of the system, but also on the data sampling frequency.

3.7.1 Derivation of Extended Kalman Filter

Just as was mentioned in the Kalman filter, the extended Kalman filter also comprises of the forest and measurement update equations. In this section, we outline the extended Kalman filter equations.

The extended Kalman filter is derived from the basic Kalman filter to nonlinear dynamics by linearization at each time step. In linearizing the governing dynamics, we employ Taylor series expansion truncated at the second order. Expanding $f(x_{k-1})$ in the process equation by the Taylor series about the estimate $\bar{x}_k$

$$
\begin{aligned}
f\left(\bar{x}_k + (x_{k-1} - \bar{x}_k)\right) &= f(\bar{x}_k) + \frac{\partial f}{\partial x}\Big|_{\bar{x}_k}(x_{k-1} - \bar{x}_k) + \frac{\partial^2 f}{\partial x^2}\Big|_{\bar{x}_k}(x_{k-1} - \bar{x}_k) + \ldots \\
&= \frac{\partial f}{\partial x}\Big|_{\bar{x}_k}(x_{k-1}) + f(\bar{x}_k) + \frac{\partial f}{\partial x}\Big|_{\bar{x}_k}(\bar{x}_k) + \frac{\partial^2 f}{\partial x^2}\Big|_{\bar{x}_k}(x_{k-1} - \bar{x}_k) + \ldots \\
&= F_k x_{k-1} + \xi_k
\end{aligned}
$$

Likewise, expanding $h(x_k)$ in measurement about the estimate $\hat{x}_k$

$$
\begin{aligned}
h\left(\hat{x}_k + (x_k - \hat{x}_k)\right) &= h(\hat{x}_k) + \frac{\partial h}{\partial x}\Big|_{\hat{x}_k}(x_k - \hat{x}_k) + \frac{\partial^2 h}{\partial x^2}\Big|_{\hat{x}_k}(x_k - \hat{x}_k) + \ldots \\
&= \frac{\partial h}{\partial x}\Big|_{\hat{x}_k}(x_k) + h(\hat{x}_k) + \frac{\partial h}{\partial x}\Big|_{\hat{x}_k}(\hat{x}_k) + \frac{\partial^2 h}{\partial x^2}\Big|_{\hat{x}_k}(x_k - \hat{x}_k) + \ldots \\
&= H_k x_k + \zeta_k
\end{aligned}
$$

The linearized equations for both the process and the measurement equation are given as

$$
F_k = \frac{\partial f}{\partial x}\Big|_{\bar{x}_k} \qquad\qquad H_k = \frac{\partial h}{\partial x}\Big|_{\hat{x}_k} \tag{3.25}
$$

An important feature of the extended Kalman filter is that the Jacobian, H_k, serves to correctly propagate the relevant component of the measurement information.

During estimation the *a priori* estimate can be obtained as follows

$$
\begin{aligned}
x_{fk} &= E[F_k x_{k-1} + f(x_{ak-1}) - F_k x_{ak-1} + \omega_k | z_{k-1}] \\
&= F_k x^a_{k-1} + f(x^a_{k-1}) - F_k x^a_{k-1} \\
&= f(x^a_{k-1})
\end{aligned}
$$

Now we derive an expression for the prediction error, and that is given by

$$\begin{aligned} e_k^f = x_k - x_k^f &= F_k\left(x_{k-1} - x_{k-1}^a\right) + f(x_{k-1}^a) + \omega_k - f(x_{k-1}^a) \\ &= F_k\left(x_{k-1} - x_{k-1}^a\right) + \omega_k \\ &= F_k e_{k-1}^a + \omega_k \end{aligned}$$

The *a priori* error covariance matrix is propagated as

$$\begin{aligned} P_k^f = E[e_k^f(e_k^f)^T] &= E[\left(F_k e_{k-1}^a + \omega_k\right)\left(F_k e_{k-1}^a + \omega_k\right)^T] \\ &= E[\left(F_k e_{k-1}^a\right)\left(F_k e_{k-1}^a\right)^T + \omega_k \omega_k^T] \\ &= F_k P_{k-1}^a F_k^T + Q_k \end{aligned}$$

The measurement estimates are obtained as

$$\begin{aligned} \hat{y}_k = E[y_k|zk-1] &= E[H_k x_k + h(x_{k-1}^f) - H_k x_k^f + \nu_k | z_{k-1}] \\ &= H_k E[x_k|z_{k-1}] + h(x_k^f) - H_k x_k^f \\ &= H_k x_k^f + h(x_k^f) - H_k x_k^f \\ &= h(x_k^f) \end{aligned}$$

Accompanying the estimate of the measurement is its predicted error and is given by

$$e_{y_k} = H_k\left(x_k - x_k^f\right) + h(x_k^f) + \nu_k - h(x_k^f) = H_k\left(x_k - x_k^f\right) + \nu_k$$

The error covariance matrix is propagated as

$$
\begin{aligned}
\Omega_k = E[e_{y_k}(e_{y_k})^T] &= E[(H_k(x_k - x^f_k) + \nu_k)(H_k(x_k - x^f_k) + \nu_k)^T] \\
&= E[(H_k e^f_k + \nu_k)(H_k e^f_k + \nu_k)^T] \\
&= H_k E[e^f_k (e^f_k)^T] H_k^T + E[\nu_k \nu_k^T] \\
&= H_k P^f_k H_k + R_k
\end{aligned}
$$

The Kalman gain, K_k is determined from

$$K_k = P^f_k H_k^T \left(H_k P^f_k H_k^T + R_k \right)^{-1} \tag{3.26}$$

The state is estimated by recursively updating the estimate with new measurements and then propagating the state estimate between the update times. The measurements are combined with the previous estimate using the standard form:

$$x^a_k = x^f_k + K_k \left(y_k - h(x^f_k) \right) \tag{3.27}$$

$$P^a_k = (I - K_k H_k) P^f_k \tag{3.28}$$

Algorithm 2 *Extended Kalman Filter*

---

1. Set initial estimates for x^a_0 and P^a_0
2. For k = 1 to maximum number of iterations
3. For i = 1 to n Time Update

Project the state forward

$x^f_k = f(x^a_{k-1})$

Linearize the process equation by computing the jacobian

$$F_k = \left.\frac{\partial f}{\partial x}\right|_{x_k^f}$$

Project the error covariance forward

$$P_k^f = F_k P_{k-1}^a F_k^T + Q_k$$

4. End of loop for *i* Measurement Update

Linearizing the measurement equation

$$H_k = \left.\frac{\partial h}{\partial x}\right|_{x_k^f}$$

Compute the Kalman gain

$$K_k = P_k^f H_k^T \left(H_k P_k^f H_k^T + R_k\right)^{-1}$$

Update state estimate with measurement y_k

$$x_k^a = x_k^f + K_k \left(y_k - y_k^f\right)$$

Update the error covariance

$$P_k^a = \left(I - K_k H_k\right) P_k^f$$

5. End of loop for *k*

3.7.2 Some Limitations of EKF

A major drawback of the EKF for data assimilation is, it requires a linearization when deriving the error covariance equation. For most systems the derivation of the jacobian matrix, the linear approximation to the nonlinear functions, can be complex causing implementation difficulties. Moreover, the linearization leads to a poor covariance evolution and for some models unstable error covariance growth.

The extended Kalman filter may provide poor results in the case of strongly nonlinear dynamics. Evensen (1992) found that the extended Kalman filter for a nonlinear quasigeostrophic ocean model resulted in an unbounded error covariance growth. In practice, for large systems the extended Kalman filter approach is again not feasible.

3.8 The Ensemble Kalman Filter (EnKF) Method

To resolve the nonlinearities in the error propagation, and also address the limitations in the extended Kalman filter, Evensen (1994) introduced an ensemble Kalman filter approach based on Monte Carlo simulations. In this methods, the error covariance matrix is represented by an ensemble of possible states that are propagated according to the full nonlinear dynamics of the system.

The ensemble Kalman Filter is also a sequential data assimilation algorithm which was first introduced by Evensen (1994) as a method of applying the ideas of KF to nonlinear systems. The EnKF originated as a version of the Kalman filter for large problems and method was formulated with nonlinear dynamics in mind, and the emphasis was focused on deriving a method, which properly could handle the error covariance evolution in nonlinear models (maintains the nonlinear features of the error statistics). Thus, instead of updating a state estimate and error covariance matrix as in KF, the EnKF uses an ensemble, or
statistical sample of state estimates.

In the ensemble Kalman filter, model error estimates are produced by assuming that the ensemble mean is truth and computing the variance of the differences between each ensemble member and the ensemble mean. Each individual observation is then updated based on the relative error in both the model and observations. The EnKF integrates an ensemble of model states forward in time using the model equations. This algorithm uses the Monte Carlo approximation for solving the Fokker Planker equation.

3.8.1 Overview of Ensemble Kalman Filter

According to Evensen (1994), the statistical properties of the state vector are represented by an ensemble of possible state vectors. Each of which is propagated according to the dynamic system subjected to model error, and the resulting ensemble then provides estimates of the forecast state and vector and the error covariance matrix.

The analysis scheme in the EnKF uses traditional update equations of the KF, except that the Kalman gain is computed using the error covariances provided by the ensemble of model states. The nonlinearities introduced by the equation are captured well by EnKF because of the sample based computation of covariance matrices. Thus, it gives a systematic way to calculate the time evolution of the forecast error statistics according to the dynamics of the forecast model.

Each member of the ensemble state vectors is propagated forward in time according to the dynamics of the system and specified model error. We denote the ensemble by $X_k^f \in R^{n\times q}$ where

$$X_k^f = \left(x_k^{f_1}, x_k^{f_2}, \ldots, x_k^{f_q}\right) \tag{3.29}$$

and the superscript f_i refers to the i−th forecast ensemble member. The ensemble mean $\bar{x}^f_k \in R^n$ is defined by

$$\bar{x}_k^f = \frac{1}{q}\sum_{i=1}^{q} x_k^{f_i} \tag{3.30}$$

We defined an ensemble error matrix $E_k^f \in R^{n\times q}$ around the ensemble mean by

$$E_k^f = \left[x_k^{f_1} - \bar{x}_k^f, x_k^{f_2} - \bar{x}_k^f, \ldots, x_k^{f_q} - \bar{x}_k^f\right] \tag{3.31}$$

and an ensemble set with observation error $E_{y_k}^f \in R^{m\times q}$ defined by

$$E_{y_k}^f = \left[y_k^{f_1} - \bar{y}_k^f, y_k^{f_2} - \bar{y}_k^f, \ldots, y_k^{f_q} - \bar{y}_k^f\right] \tag{3.32}$$

In the computation of the Kalman gain filter, the covariance matrices P_k^f, $P_{xy}^f{}_k$, and $P_{yy}^f{}_k$ should be introduced. In the Ensemble Kalman filter, those values can be estimated as

$$\hat{P}_k^f = \frac{1}{q-1}E_k^f\left(E_k^f\right)^T, \quad \hat{P}_{xy_k}^f = \frac{1}{q-1}E_k^f\left(E_{y_k}^f\right)^T, \quad \hat{P}_{yy_k}^f = \frac{1}{q-1}E_{y_k}^f\left(E_{y_k}^f\right)^T \tag{3.33}$$

Thus, we interpret the forecast ensemble mean as the best forecast estimate of the state, and the spread of the ensemble members around the mean as the error between the best estimate and the actual state.

The second step is the analysis step: Each ensemble member is updated according to the updating scheme, equation (3.34), and based on the updated ensemble, the updated state vector and error covariance matrix are estimated. To obtain an analysis estimate of the state, the EnKF performs an ensemble of parallel data assimilation cycles, where for $i = 1,\ldots,q$

$$x_k^{a_i} = x_k^{f_i} + \hat{K}_k \left(y_k^i - h\left(x_k^{f_i} \right) \right) \quad (3.34)$$

The perturbed observations y_k^i is given by

$$y_k^i = y_k + \nu_k^i \quad (3.35)$$

where ν_k^i is a zero-mean random variable with a normal distribution and covariance R_k. The sample error covariance matrix is computed from ν_k^i, which converges to R_k as $q \longrightarrow \infty$.

We approximate the analysis error covariance P_k^a by $\hat{P}_k^a$, where

$$\hat{P}_k^a = \frac{1}{q-1} E_k^a \left(E_k^a \right)^T \quad (3.36)$$

and E_k^a is defined by equation (3.31) with $x_k^{f_i}$ replaced by $x_k^{a_i}$ and $\bar{x}_k^f$ also replaced by the mean of the analysis estimate ensemble members, $\bar{x}_k^a$. We use the classical Kalman filter gain expression and the approximations of the error covariances to determine the filter gain $\hat{K}_k$ by

$$\hat{K}_k = \hat{P}_{xy_k}^f \left(\hat{P}_{yy_k}^f \right)^{-1} \quad (3.37)$$

The evaluation of the filter gain $\hat{K}_k$ in the EnKF does not involve an approximation of the nonlinearity of the state-space equations.

The last step is the prediction of the error statistics in the forecast step which is given by

$$x_{k+1}^{f_i} = f\left(x_k^{a_i}, u_k \right) + \omega_k^i \quad (3.38)$$

where the values ω_k^i are sampled from a normal distribution with average zero and covariance Q_k. The sample error covariance matrix computation from the ω_k^i converges to Q_k as

$q \to \infty$.

There are many different EnKF methods, and these differ in the analysis step. Burgers et al. (1998) , described the perturbed observation EnKF generates an ensemble of observa-

---

Algorithm 3 *Ensemble Kalman Filter*

---

1. Set initial estimates for x_0^a and P_0^a
2. Generate the set of ensemble members, $X_k^f = \left(x_k^{f_1}, x_k^{f_2}, ..., x_k^{f_q}\right)$
3. For k = 1 to maximum number of iterations
4. In the Time Update each ensemble member is allowed to act independently
5. For i = 1 to n Time Update

   Project the state forward and obtain estimates of the observation

   $x_k^{f_i} = f\left(x_k^{a_i}, u_k\right) + \omega_k^i$

   $y_k^{f_i} = h\left(x_k^{f_i}, u_k\right)$
6. End of loop for i

   Estimates of the mean and observations as computed as the mean

   $\bar{x}_k^f = \frac{1}{q}\sum_{i=1}^{q} x_k^{f_i}$ $\qquad$ $\bar{y}_k^f = \frac{1}{q}\sum_{i=1}^{q} y_k^{f_i}$

   Measurement Update

   $E_k^f = \langle x_k^{f_1} - \bar{x}_k^f, x_k^{f_2} - \bar{x}_k^f, ..., x_k^{f_q} - \bar{x}_k^f \rangle$ $E_{y_k}^f = \langle y_k^{f_1} - \bar{y}_k^f, y_k^{f_2} - \bar{y}_k^f, ..., y_k^{f_q} - \bar{y}_k^f \rangle$

   Compute the sample covariance matrices

   $\hat{P}_{xy_k}^f = \frac{1}{q-1} E_k^f \left(E_{y_k}^f\right)^T$ $\qquad$ $\hat{P}_{yy_k}^f = \frac{1}{q-1} E_{y_k}^f \left(E_{y_k}^f\right)^T$

   Compute the Kalman gain

   $\hat{K}_k = \hat{P}_{xy_k}^f \left(\hat{P}_{yy_k}^f\right)^{-1}$

   Update state estimate with measurement y_k

$$x_k^{a_i} = x_k^{f_i} + \hat{K}_k \left(y_k^i - h \left(x_k^{f_i} \right) \right)$$

Update the error covariance

$$\hat{P}_k^a = \frac{1}{q-1} E_k^a \left(E_k^a \right)^T$$

7. End of loop for k

---

tions consistent with the error statistics of the observation and assimilates these into each ensemble member.

3.9 Implementing the Data Assimilation Methods in State Estimation of the Van der Pol Equation

In this section, we implement the data assimilation schemes discussed in this chapter on the van der pol oscillator problem. This problem is nonlinear, therefore the extended Kalman filter and ensemble Kalman filter are implemented, since the Kalman filter method fails for nonlinear systems. The ensemble open loop is also considered, in which the estimates are not updated (combined with available observation).

There are different experimental set ups, where we seek to estimate the states and most importantly compare the performance of the extended Kalman filter, ensemble Kalman filter and the ensemble open loop. The effect of the process noise, measurement noise and the initial error covariance on the performance of the filters is also investigated. Lastly we investigate the performance of the ensemble size on performance of the EnKF.

The van der pol equation is described by a second order differential equation as

$$y^{00} - \alpha(1 - y^2)y^0 + y = 0 \quad (3.39)$$

which is equivalent to the system of first-order differential equations give as

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= \alpha(1 - y_1^2)y_2 - y_1 \end{aligned} \quad (3.40)$$

The ensemble Kalman filter is implemented on state-space representation of the dynamical system. A numerical scheme (Euler's method) is used in the descritization of the governing differential equation to obtain the state (process) equation as follows

$$\begin{aligned} x_{1,k+1} &= x_{1,k} - hx_{2,k} + \omega_{1,k} \\ x_{2,k+1} &= x_{2,k} + h[x_{1,k} - \alpha(1 - x_{1,k}^2)x_{2,k}] + \omega_{2,k} \end{aligned} \quad (3.41)$$

where $x_k = [x_{1,k}\, x_{2,k}]$ which represents the state of the dynamical system. $\omega_k = [\omega_{1,k}\, \omega_{1,k}]$ is a Gaussian random variable modeling the process noise, α is scalar parameters and h is the step size. All states were observed with an observation noise n_k drawn from a Gaussian distribution. The process equation of the above model can then be seen as

$$x_{k+1} = f(x_k) + \omega_k$$

whereas the measurement equation is given as

$$y_k = Hx_k + v_k$$

where H is just an identity matrix.

Figure 3.1 contains plot of observation and truth of the time series generated by the process model. The figure demonstrates the sinusoidal pattern displayed by the van der pol equation.

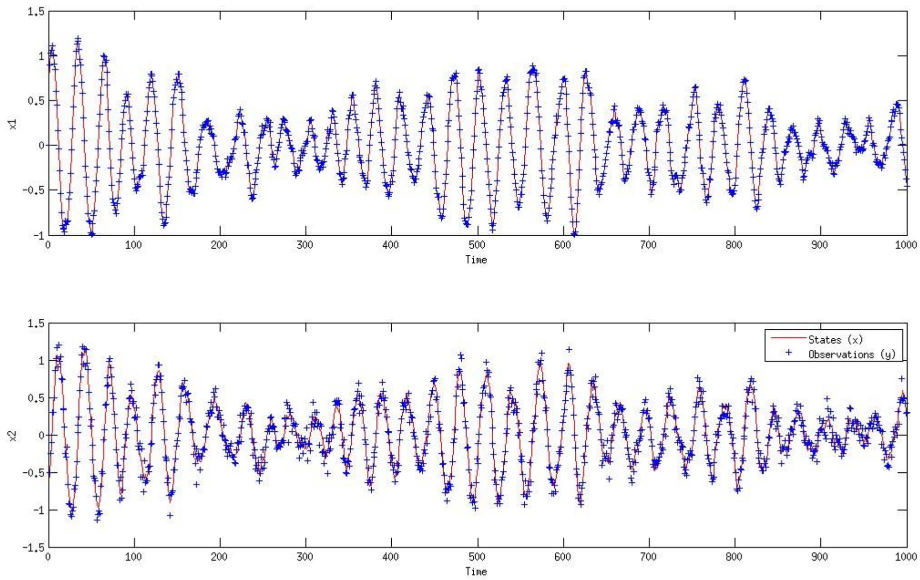

Figure 3.1: Plot of observation and truth generated in the van der pol problem for the first experiment

We carry out some experiments to ascertain the performance of the filters being considered. The filters were used to estimate the state sequence of the van der pol model. The first experiment was set to run with some default values to verify the performance of the filters. From there, we begin to vary parameters like the process noise, measurement noise, initial error covariance matrix and the ensemble size in various experiments to see how far these parameters affect the performance of the filters.

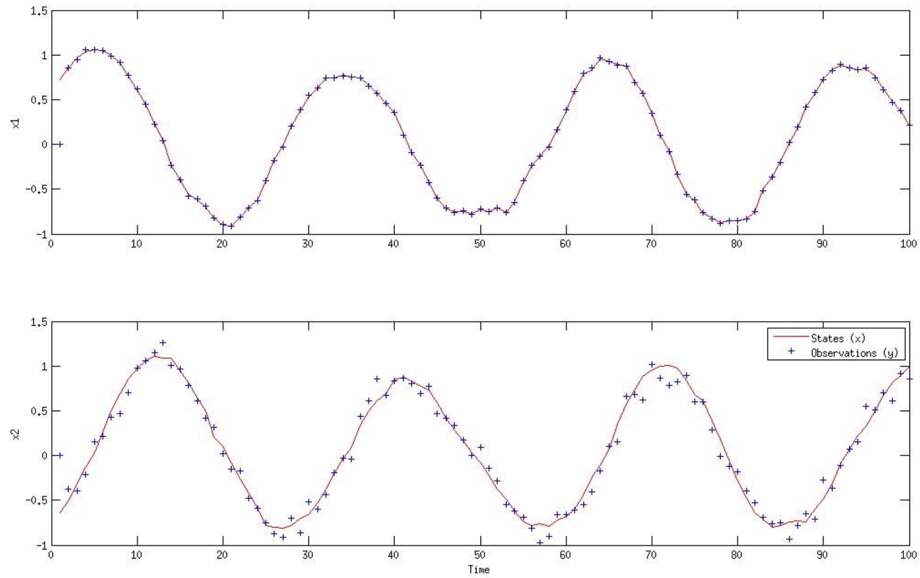

Figure 3.2: Plot of observation and truth generated in the van der pol problem for the first experiment

The experiments use the different filters for the estimation of the state. The first experiment was carried out to estimate the state, $x_k = [x_1\, x_2]$ with $k = 1,2,...,100$ and the data set plotted in figure 3.2. Figure 3.3 compares estimates generated from the various filters to the the true state. It is observed that the ensemble Kalman filter is estimating the true state best, followed by the extended Kalman filter and the the ensemble open loop.

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

	RMSE	
Algorithm	**Mean**	**Variance**

Ensemble Open Loop (EnOL)	0.0281	1.207e-04
Extended Kalman Filter (EKF)	0.0076	6.972e-06
Ensemble Kalman Filter (EnKF)	0.0049	2.679e-06

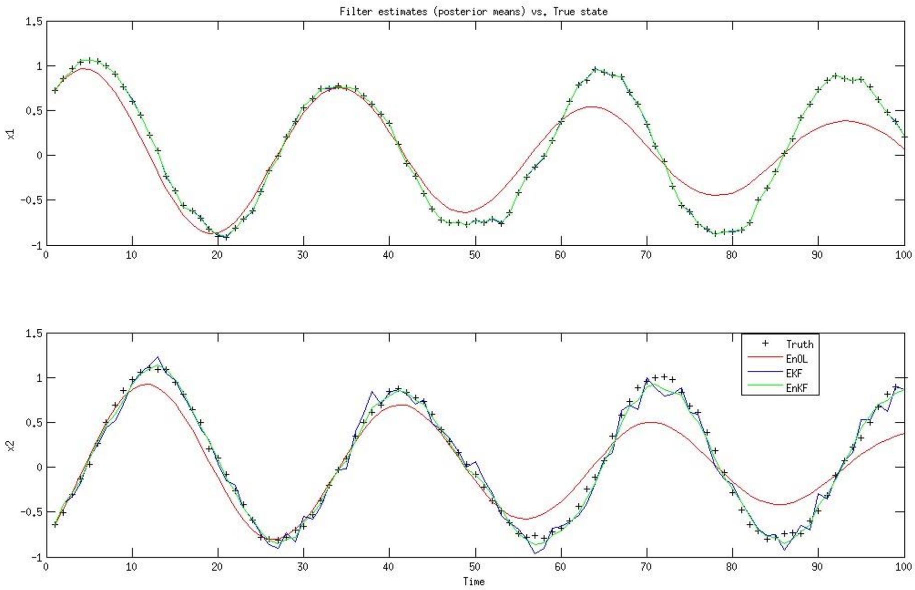

Figure 3.3: Plot of estimates generated by the different filters in the van der pol problem for the first experiment

Table 3.1 contains the summary of the performance of the different filters. The table shows the mean and variance of the root mean square error (RMSE) of the state estimates from the various filters. The RMSE of the EnKF has the least values followed by the EKF and then EnOL. The experiment shows the EnKF is the best performing of the filters considered. The ultimate performance of the ensemble Kalman filter is also proved by the plot of the propagation of the RMSE by the different filters in estimating the state in the van der pol model.

The second experiment tends to vary the process noise while the other parameters are kept fixed. The result summarized in Table 3.2 shows that the performance of the filters in estimating the state improves with decrease in the process noise. Thus the smaller the

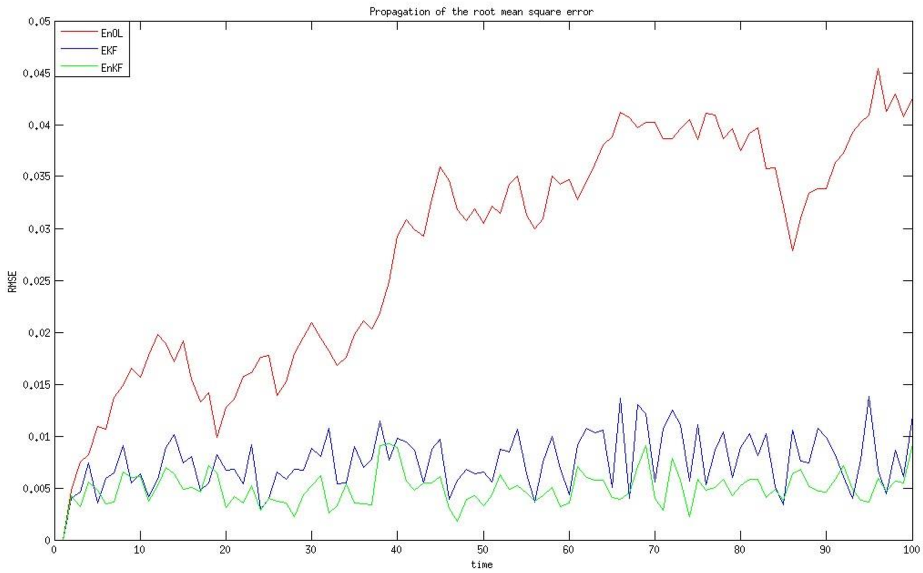

Figure 3.4: Plot of the root mean square error propagation by the different filters in the van der pol problem for the first experiment

process noise, the better the state estimate. The experiment showed that for high values of the process noise, the filters break down as in the case of $Q = I_2$. The ensemble Kalman filter is the best performing followed by the extended and then the the ensemble open loop except for $Q = 10^{-11}I_2$ where the ensemble open loop performed better than the extended Kalman filter. In the next experiment, we vary the measurement noise keeping all the other parameters fixed. The summary of the perform of the filters is contained in Table 3.3. It can be noted from this table that the EnKF is the best performing followed closely by the EKF and then the EnOL. The peformance of the EKF and EnKF improves with decrease in observation noise. It is observed that the mean and variance of the RMSE for

EKF and EnKF at $R = 10^{-4}I_2$ are very close indicating that at this observation noise, the Table 3.2: Compare effect of different process noise Q on the performance of the various filters holding other parameters fixed

	$Q = I_2$		$Q = 10^{-4}I_2$		$Q = 10^{-6}I_2$		$Q = 10^{-11}I_2$	
Filters	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**
EnOL	NaN	NaN	0.03087	1.312e-04	0.00278	1.103e-06	0.00096	4.572e-08
EKF	NaN	NaN	0.00123	9.432e-08	0.00126	1.186e-07	0.00122	9.943e-08
EnKF	NaN	NaN	0.00120	8.836e-08	0.00078	4.520e-08	0.00017	4.869e-08

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

	$R = I_2$		$R = 10^{-1}I_2$		$R = 10^{-2}I_2$		$R = 10^{-4}I_2$	
Filters	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**
EnOL	0.0307	9.152e-05	0.0342	14.951e-05	0.0224	3.717e-05	0.0411	19.422e-05
EKF	0.0252	5.250e-05	0.0214	5.067e-05	0.0109	8.288e-06	0.0012	8.091e-08
EnKF	0.0229	4.615e-05	0.0141	1.081e-05	0.0078	3.935e-06	0.0012	8.791e-08

performance from the both the EKF and EnKF are almost the same.

Whether or not the EKF could perform better than EnKF was unknown from Table 3.3. Further results in Table 3.4 still confirms the improvement in performance of filters with decrease in measurement noise. However the performance of the extended Kalman filter improves beyond the ensemble Kalman filter. In all these cases the ensemble open loop remains as the filter with least performance. Figure 3.5 is a plot of the state estimate from the different filters for $R = I_2$. It can be noted that the estimate from the filters are quite far from the true state. This indicates the bad performance of the filters in estimating the state when the measurement noise is high.

Table 3.4: Further comparison of the effect of different observation noise R on the performance

	$R = 10^{-5}I_2$		$R = 10^{-7}I_2$		$R = 10^{-11}I_2$	
Filters	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**
EnOL	0.0346	6.924e-05	0.0296	7.145e-05	0.0222	4.918e-05
EKF	3.846e-04	1.002e-08	3.947e-05	8.936e-11	3.904e-07	9.634e-15
EnKF	3.891e-04	1.033e-08	3.984e-05	8.626e-11	3.946e-07	9.695e-15

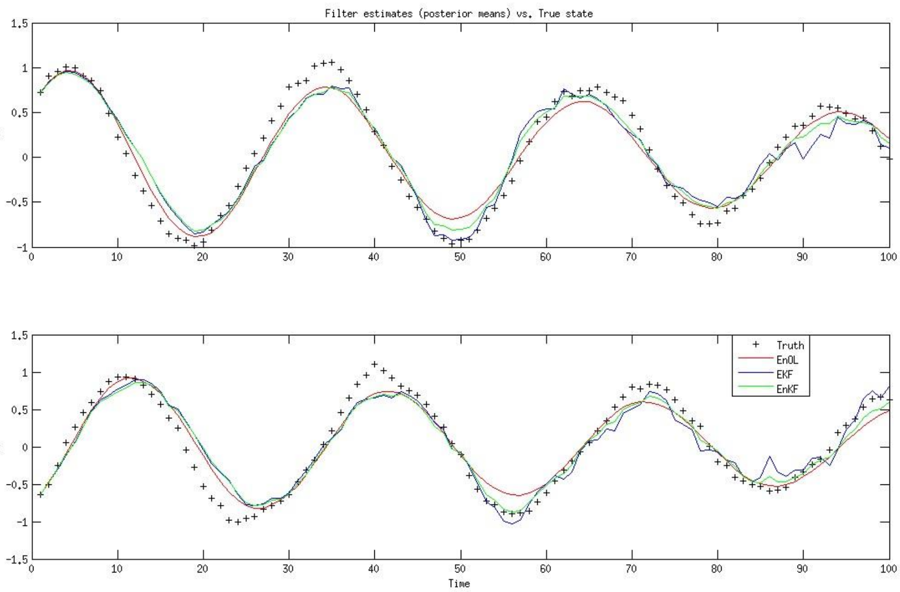

Figure 3.5: Plot of the state estimate from the different filters compared to the true state at $R = I_2$

We have so far seen the effect of both the process noise and the measurement noise on the performance of the filters. The next experimental setup is to investigate the effect of the initial error covariance on the performance of these methods.

The initial error covariance seems to be deteriorating the performance of the filters until

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

	$P_o = I_2$		$P_o = 10^{-2}I_2$		$P_o = 10^{-5}I_2$		$P_o = 10^{-7}I_2$	
Filters	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**	**Mean**	**Var**
EnOL	NaN	NaN	0.02795	1.116e-04	0.03448	9.222e-05	0.02546	8.228e-05
EKF	0.00123	8.833e-08	0.00126	1.288e-07	0.00129	1.004e-07	0.00120	1.014e-07
EnKF	0.00123	9.497e-08	0.00122	1.225e-07	0.00127	9.295e-08	0.00117	8.734e-08

$P_o = 10^{-7}$ where the performance improved drastically. The ensemble open loop breaks down at high values of P_o. Table 3.5 gives an overview of the effect of initial covariance on the performance of the filters. In all these odds, the performance of the EnKF remains the best performing filter, followed by EKF and then EnOL.

The last but not the least of the experiments seeks to investigate the effect of the ensemble size on the performance of the filters. The extended Kalman filter is not considered in this experiment since it does not make use of the ensemble size.

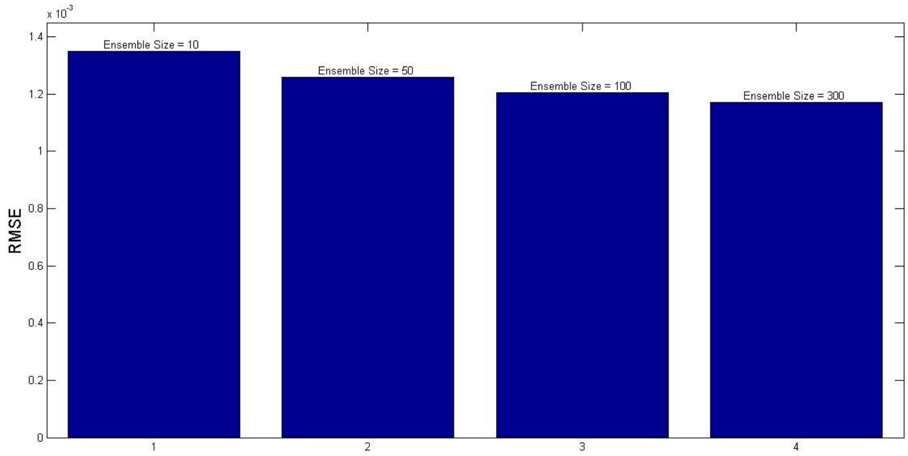

Figure 3.6: Plot of root mean square errors showing the effect of different ensemble size on the performance of ensemble Kalman filter

The experimental results in Figure 3.6 shows a decrease in the RMSE bar as the ensemble sizes increase. That is to say, as the ensemble size grows the performance of the ensemble Kalman filter improves. This confirms the findings in existing literature.

Chapter 4

Results and Discussions

4.1 Introduction

The chapter contains the results of the predictions of financial options in the Black-Scholes model using the ensemble Kalman filter and the extended Kalman filter, comparing the results from both estimates for different strike prices. The estimates were obtained for an ensemble size of 200 and the experiments repeated 100 times reporting the mean value, with corresponding variance identically zero. Typically, options on a particular equity and with the same exercise date are traded with several strike prices. Five strike prices were considered in this research and they are 2925, 3025, 3125, 3225, and 3325 in the evaluation of the pricing algorithm. The strike prices are related to the collected data used in this

research.

The famous Black-Scholes model has been used as the foundation for option pricing. The Black-Scholes model for the value of an option is described by the partial differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{4.1}$$

where $V(S,t)$ is a European call or put option at asset price S and at time t, $r(t)$ is the risk free interest rate, and σ representing the volatility of underlying asset.

Given the Black-Scholes model of a financial market, the Black-Scholes formula for European call options (C_t), and put options (P_t) on a non-dividend paying stock are

$$C_t = S_t N(d_1) - Xe^{rT} N(d_2) \quad (4.2)$$

$$P_t = -S_t N(-d_1) - Xe^{rT} N(-d_2) \quad (4.3)$$

where X the desired strike price, T is the time until expiration. $N(.)$ is the cumulative normal distribution function and d_1 and d_2 are given by

$$d_1 = \frac{\ln(S_t/X) + (r + \sigma^2/2)t_m}{\sigma\sqrt{t_m}}$$
$$d_2 = \frac{\ln(S_t/X) + (r - \sigma^2/2)t_m}{\sigma\sqrt{t_m}}$$

The risk free interest rate and the volatility are in this case treated as hidden states, whereas the call and put options are being considered as the output observations. The current value of the underlying cash product and the time to maturity are treated as input observations. The model setup represents a parameter estimation problem with the observation equation given by equations 4.2 and 4.3 allowing the computation of daily probability distributions for the risk free interest rate and the volatility which helps in deciding whether the current value of an option in the market is over-priced or under-priced.

4.2 Results and Discussion

The results of the prediction for the individual strike prices from the various filters are presented as well as the mean of the normalized root mean square errors. Figure 4.1 shows the plot of call option, put option and stock prices normalized with respect to the strike price. The data indicates that the normalized call prices falls between 0 and 0.25, the

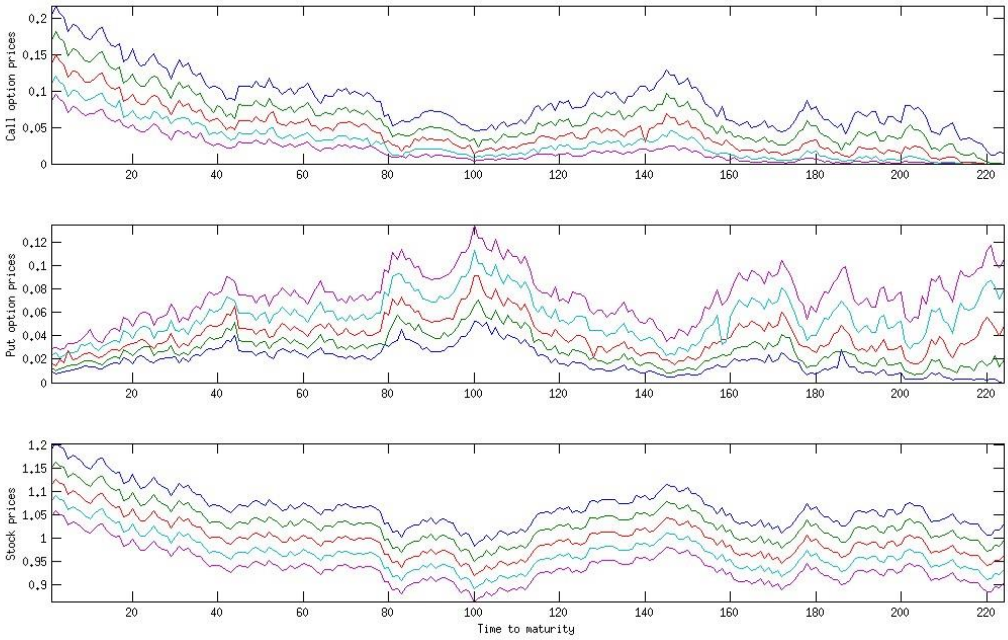

Figure 4.1: Plot of the normalized call prices, put prices and stock prices verse time to maturity.

normalized put is between 0 and 0.15, and that of the stock prices between 0.5 and 1.2. As observed from Figure 4.1, as the call option prices are decreasing with increase in time to maturity, the put option prices increasing with increase in time to maturity and vice versa.

Increasing the stock price cause the call option to worth more and the put option to worth less.

4.3 Ensemble Kalman Filter Estimation of Contract

In this section we look at the ensemble Kalman filter prediction of the call and put options as well as the estimation of the interest rate and volatility for a contract for the different strike prices. The trivial prediction is obtained under the assumption that the price on the following day corresponds to the current price.

4.3.1 Estimates for Contract with Strike Price of 2925

The estimates interest rate and volatility for contract with a strike price of 2925 is presented in Figure 4.2. The interest rate and volatility rises to a sharp peak of about 0.045 and 0.218 respectively and then decreases sharply and begins to rise and drop gradually with changing time to maturity. Table 4.1 contains the mean of the root mean squared errors showing the performance of the EnKF against the Trivial. Figure 4.3 shows the estimated call and put option prices from the ensemble Kalman filter.

Table 4.1: Mean of the normalized RMSE with strike price of 2925

Option type	Algorithm	Mean NSE
Call	Trivial	0.0783
	Ensemble Kalman filter (EnKF)	0.0418
Put	Trivial	0.0354
	Ensemble Kalman filter (EnKF)	0.0269

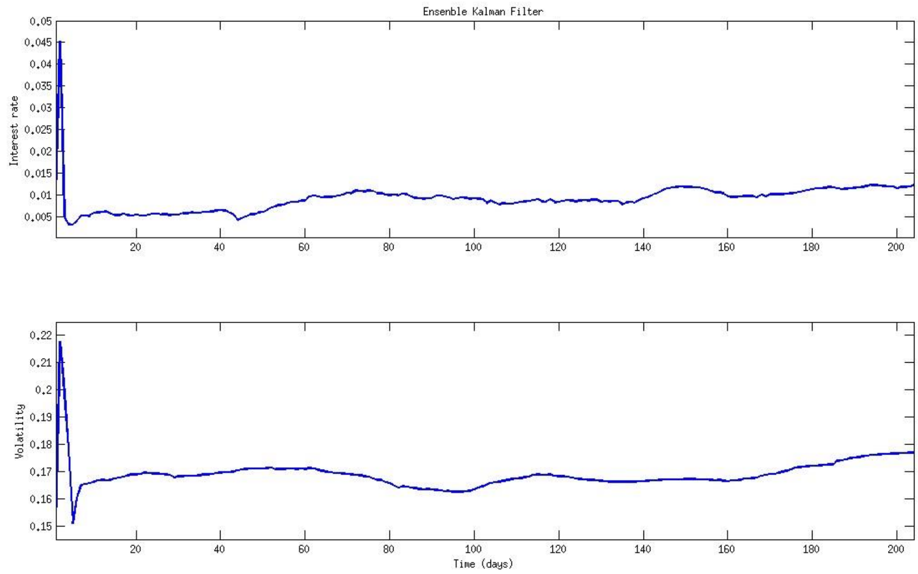

Figure 4.2: EnKF Estimate of interest rate and volatility for a contract with strike price 2925

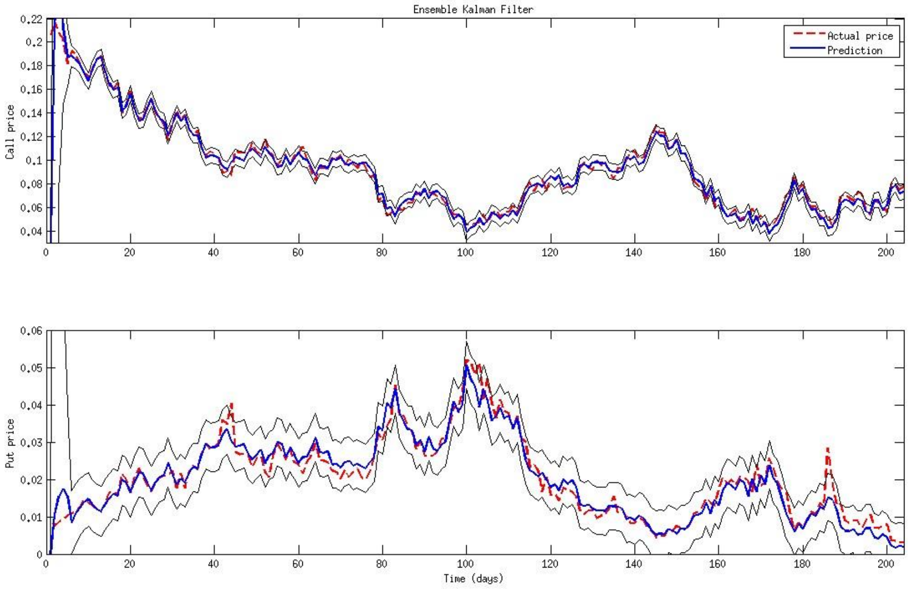

Figure 4.3: Plot of estimated call and put option prices from the ensemble Kalman filter

The estimates of the call and put options closely predict the observed especially in the case of the call option. Due to the sharp rise in volatility, there is a sharp rise in the estimation of the call option indicating that a rise in the volatility increases the value of the call option and this is shown in Figure 4.3. The put on the other hand is worth less with increasing volatility. The options worth less as they approach the time to maturity. The call option attained it maximum value on day 52 while the put option attained on day 100 its maximum value.

4.3.2 Estimates for Contract with Strike Price of 3025

The section presents results base on strike price of 3025 with Figure 4.4 indicating the ensemble Kalman filter estimate of the interest rate and the volatility at this strike price.

Indicated in Figure 4.4 is a sharp rise and fall of interest rate and volatility for the first 3 days in the case of interest rate and 6 days in the case of volatility. Figure 4.5 shows the estimates of the call and put option prices.

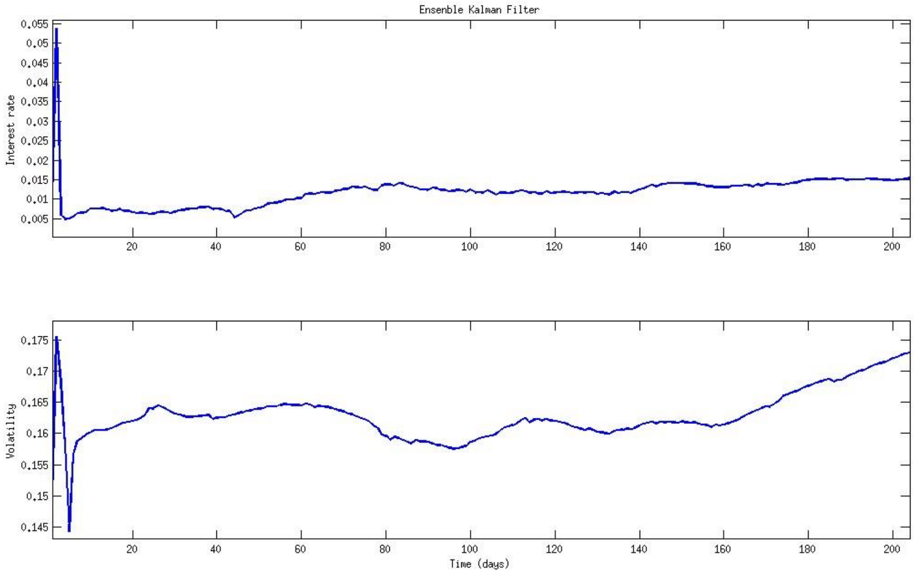

Figure 4.4: EnKF Estimate of interest rate and volatility for a contract with strike price 3025

Table 4.2: Mean of the normalized RMSE with strike price of 3025

Option type	**Algorithm**	**Mean MSE**
Call	Trivial	0.061132
	Ensemble Kalman filter (EnKF)	0.036137
Put	Trivial	0.039403
	Ensemble Kalman filter (EnKF)	0.027626

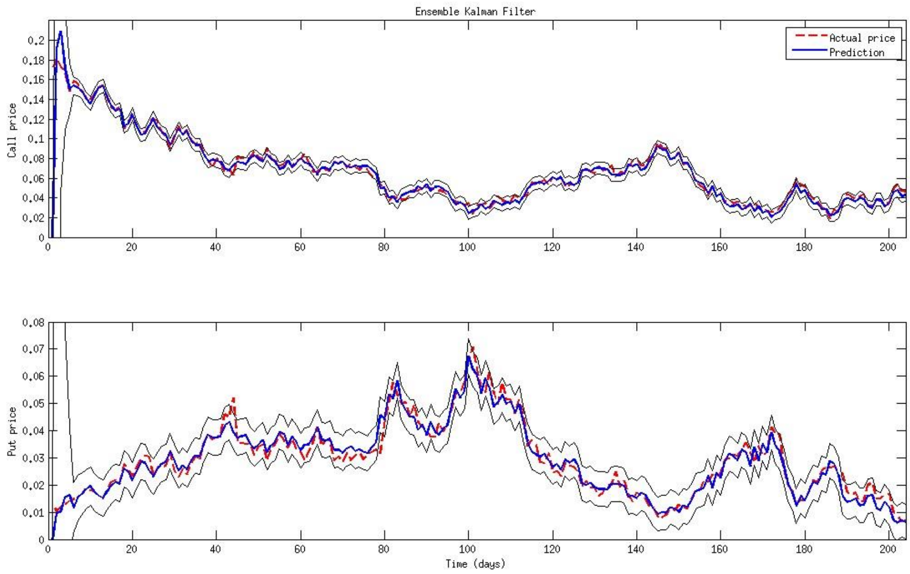

Figure 4.5: Plot of estimated call and put option prices from the ensemble Kalman filter

4.3.3 Estimates for Contract with Strike Price of 3125

Results for contract with strike price of 3125, shows that the estimate of the volatility has the highest values approaching the time to maturity but interestingly, Figure 4.7 still indicated that the value of the options still worth less as they approach the time to maturity.

Table 4.3: Mean of the normalized RMSE with strike price of 3125

Option type	**Algorithm**	**Mean NSE**
Call	Trivial	0.052427
	Ensemble Kalman filter (EnKF)	0.036825
Put	Trivial	0.049033
	Ensemble Kalman filter (EnKF)	0.031532

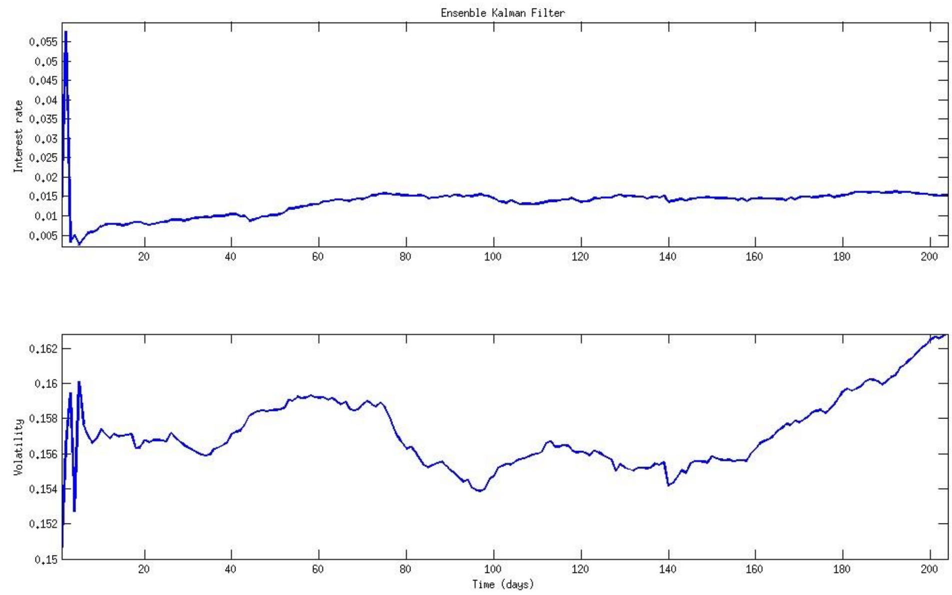

Figure 4.6: EnKF Estimate of interest rate and volatility for a contract with strike price 3125

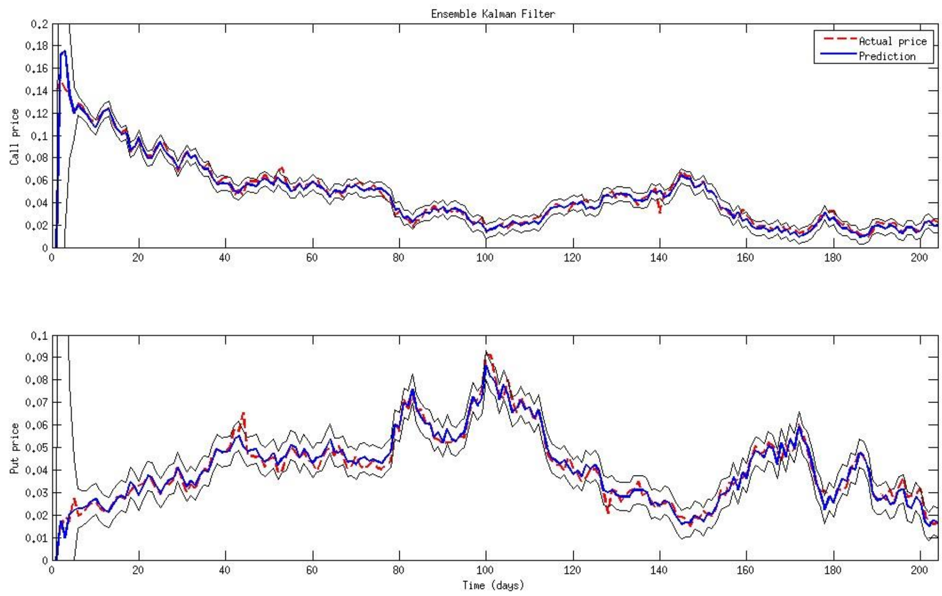

Figure 4.7: Plot of estimated call and put option prices from the ensemble Kalman filter

4.3.4 Estimates for Option Contract with Strike Price of 3225

The prediction obtianed for the varying strike prices were very close to the measured data as evident in the results. That of the strike price of 3225 is shown in Figure 4.9. Figure 4.8 shows the estimated volatility and interest rate for contract with a strike price of 3225. The one-step-ahead normalized square errors obtained for the EnKF on a pair of options with strike price 3225 csn be seen in Table 4.4

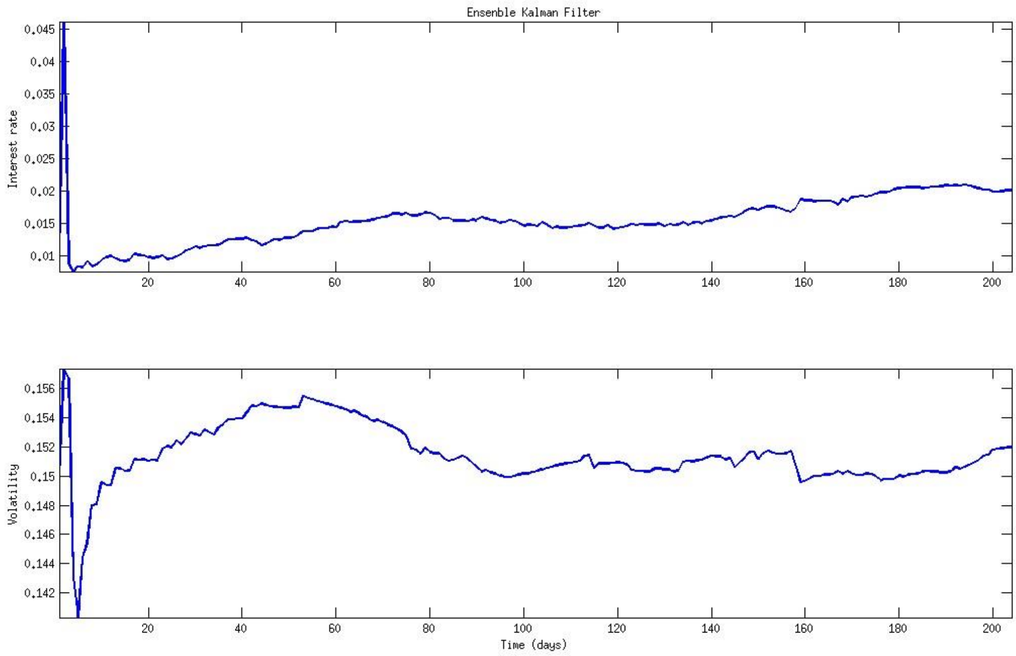

Figure 4.8: EnKF Estimate of interest rate and volatility for a contract with strike price 3225

Table 4.4: Mean of the normalized RMSE with strike price of 3225

Option type	Algorithm	Mean NSE
Call	Trivial	0.033906
	Ensemble Kalman filter (EnKF)	0.024048
Put	Trivial	0.068802

Ensemble Kalman filter (EnKF)	0.040229

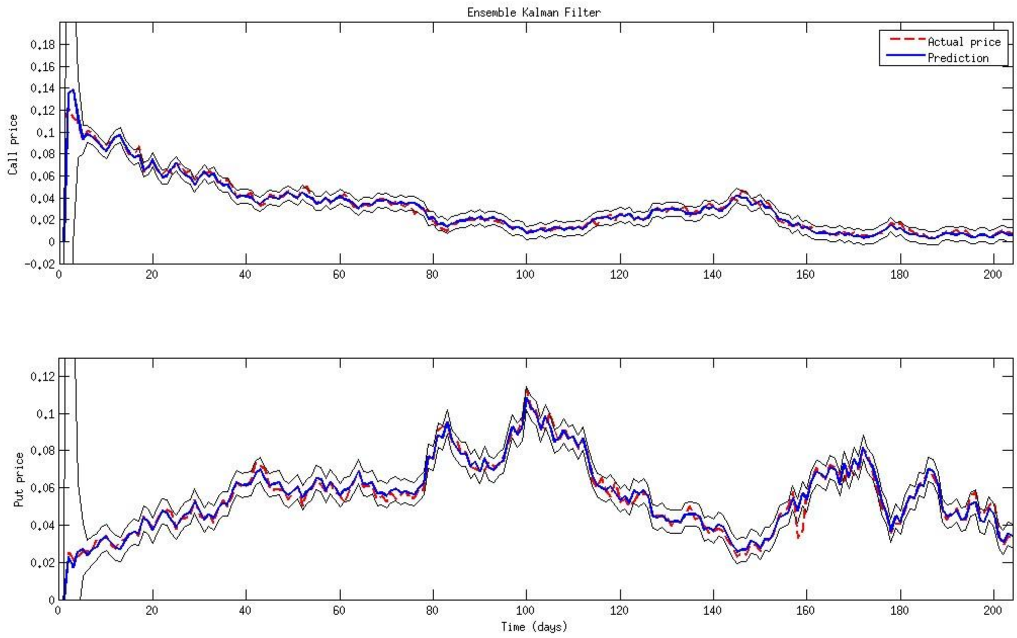

Figure 4.9: Plot of estimated call and put option prices from the ensemble Kalman filter

4.3.5 Estimates for Option Contract with Strike Price of 3325

It is observed that for the different strike prices considered the estimation of the interest rate has been consistent with slight variations in the peak values. Figure 4.11 holds estimates of interest rate and volatility and compared to estimates from the rest of the stock prices, the volatility on the other hand has the dynamics changing consistently.

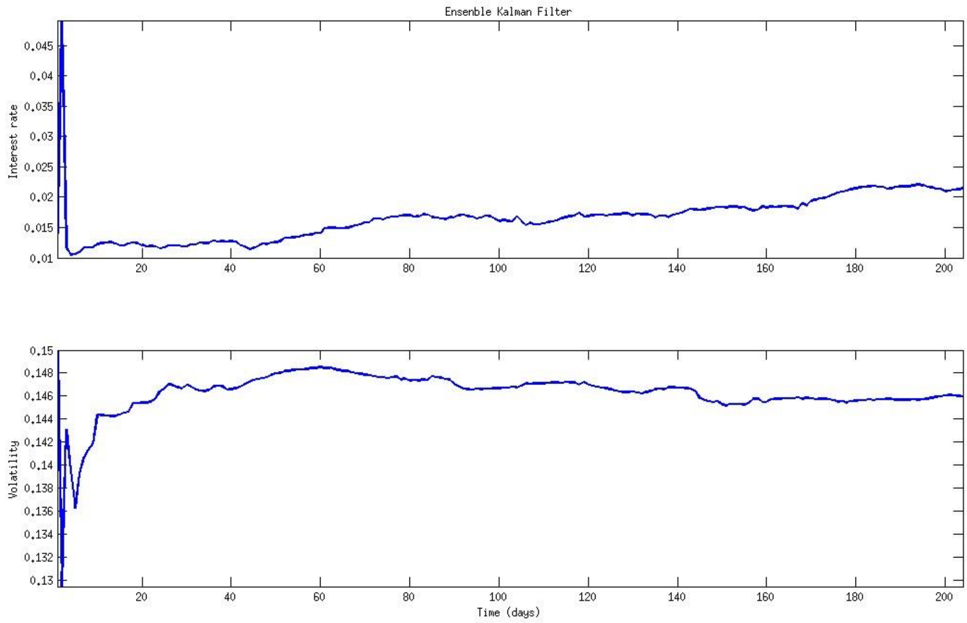

Figure 4.10: EnKF Estimate of interest rate and volatility for a contract with strike price 3325

The worth of the call option seems not to vary much as it approaches the maturity time from day 90 to 204 while there are significant variations in the worth of the put option as in Figure 4.11. Consistently as the call option worths more, the put option worth less with time to maturity.

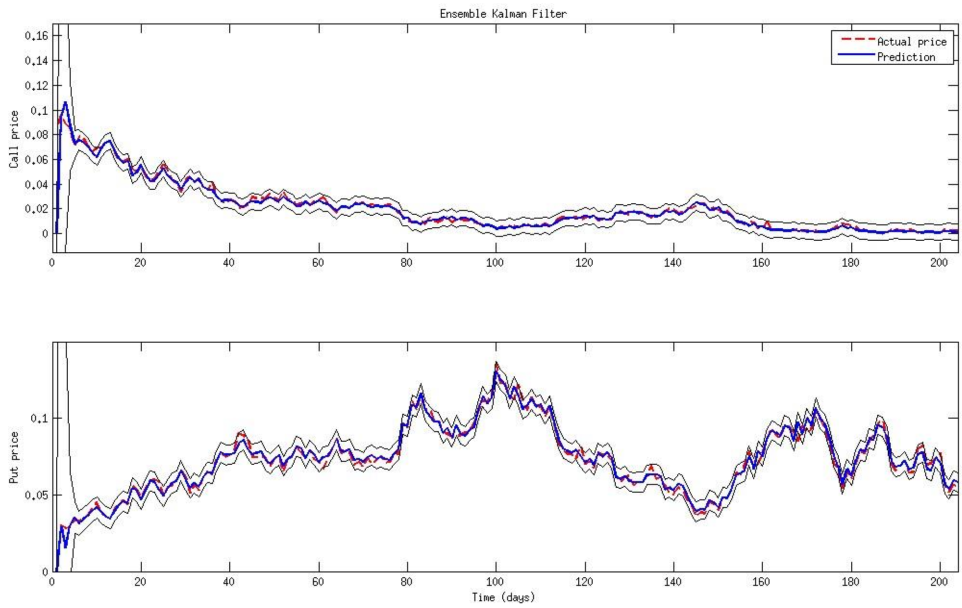

Figure 4.11: Plot of estimated call and put option prices from the ensemble Kalman filter

Table 4.5: Mean of the normalized RMSE with strike price of 3325

Option type	**Algorithm**	**Mean NSE**
Call	Trivial	0.020488
	Ensemble Kalman filter (EnKF)	0.013225
Put	Trivial	0.071863
	Ensemble Kalman filter (EnKF)	0.034004

4.4 Comparing the Performance of the EnKF and EKF

Comparing the results from both filters, Figures 4.12 and 4.14 show the estimated volatility and interest rate from both the ensemble Kalman filter and the extended Kalman filter for strike prices 2925 and 3125 respectively. The red plot is the estimate from the EKF whilst the blue is EnKF prediction.

The predictions from the two filter of the interest rate and volatility are so close to each

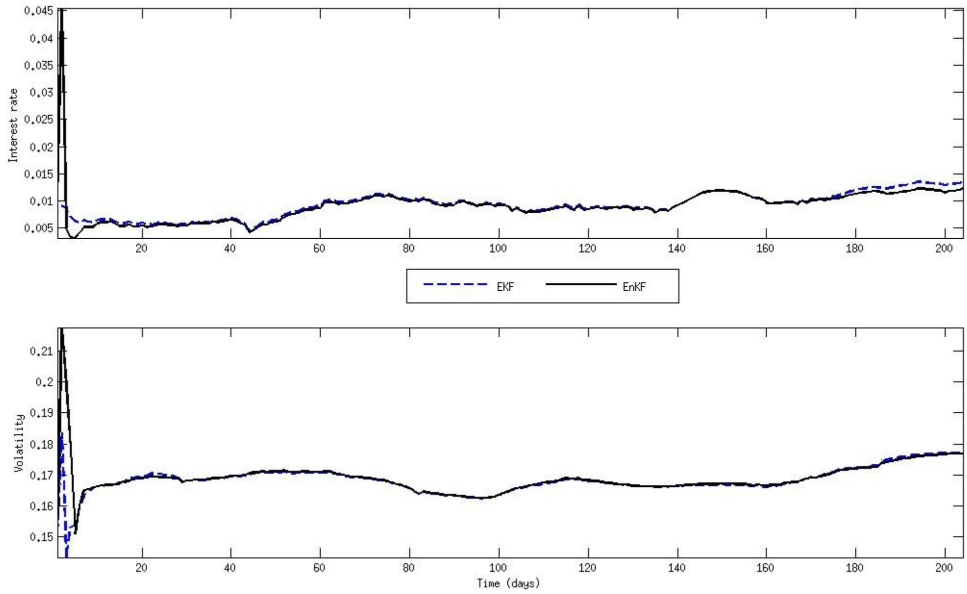

Figure 4.12: Estimate of interest rate and volatility for a contract with strike price 2925

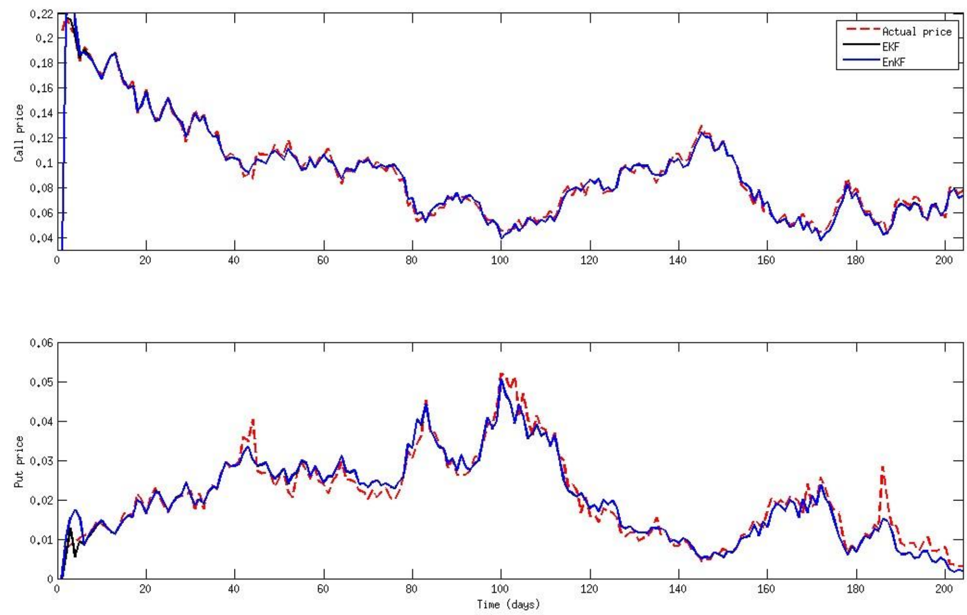

Figure 4.13: Estimate of call and put options for strike price of 2925

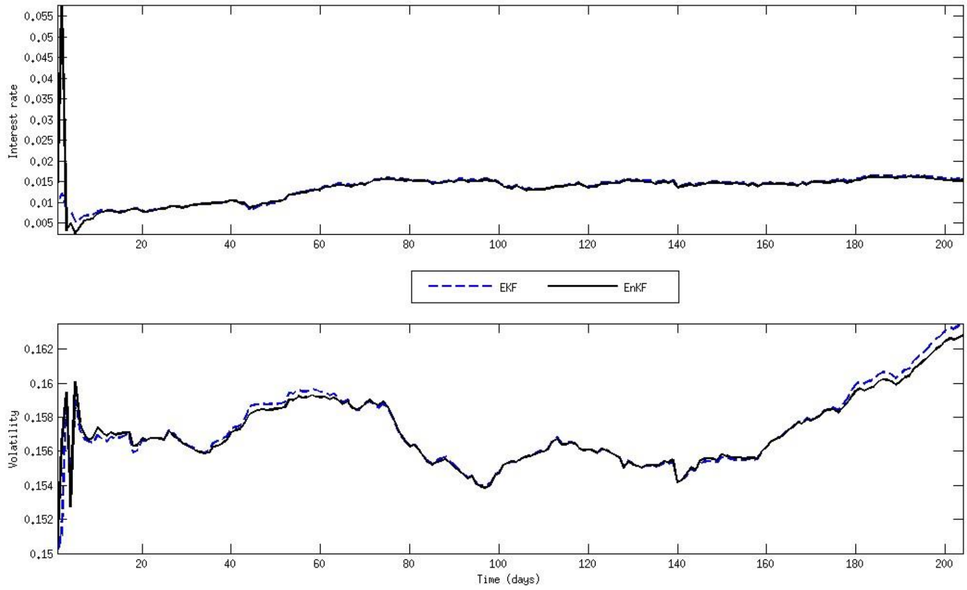

Figure 4.14: Estimate of interest rate and volatility for a contract with strike price 3125

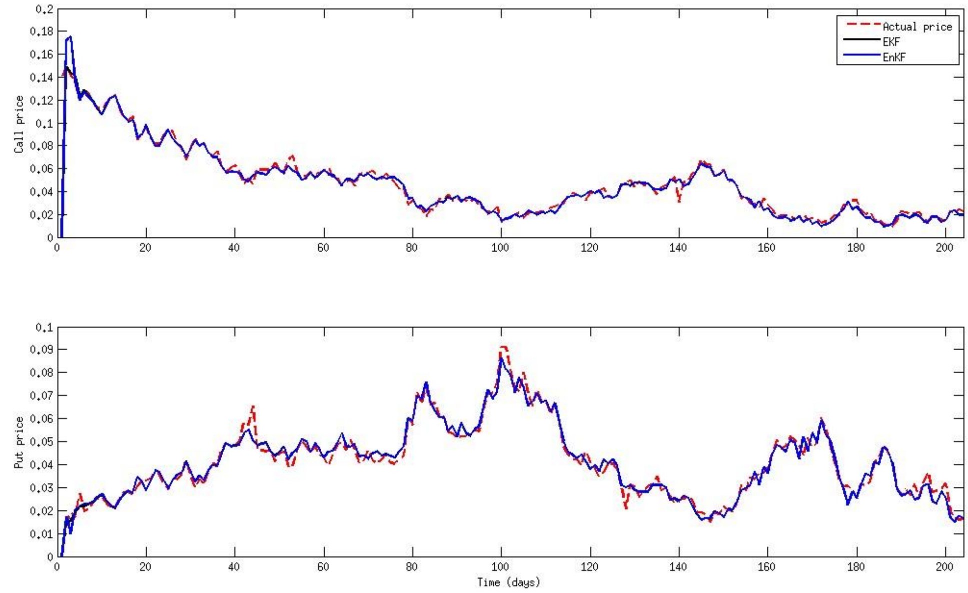

Figure 4.15: Estimate of call and put options with strike price 3125

other. The ensemble Kalman filter estimates for the interest rate flatuate sharply for the first couple of days which is not the case for the extended Kalman filter. The EKF predicts relatively low values for the interest rate at these instances.

Table 4.6, contains the mean of the root mean square errors. These error rates are computed over the last 100 days. The root mean square error in the filters are so close indicating that the two methods approach the same accuracy in using the Black-Scholes model for predicting the options. Notable is the constant decrease in the root mean square values for the value of the call options in all the filters while the reverse occurs in the case of the put option when the strike price increases.

Table 4.6: Mean of the normalized root mean squared errors for the specified strike prices

Option type	**Algorithm**	**X=2925**	**X=3125**	**X=3325**
Call	Trivial	0.0783	0.0524	0.0205
	Extended Kalman filter (EKF)	0.0417	0.0368	0.0132
	Ensemble Kalman filter (EnKF)	0.0418	0.0368	0.0132
Put	Trivial	0.0354	0.0490	0.0719
	Extended Kalman filter (EKF)	0.0269	0.0315	0.0339
	Ensemble Kalman filter (EnKF)	0.0269	0.0315	0.0340

Chapter 5

Conclusion and Recommendation

5.1 Introduction

In this study, we investigated the performance of the Ensemble Kalman filter considering the effect of different process noises, different measurement noises, different initial error covariances and the ensemble size through a Monte Carlo simulation. The ensemble Kalman filter was implemented on two estimation problems, the first being a synthetic experiment using the Van der Pol equation and the second is a real world problem using the Black-

Scholes model in the pricing of financial options considering on the British FTSE-100 index. Estimation for the pricing of options were obtained for five different strike prices 2925, 3025,

3125, 3225 and 3325. The performance of the ensemble Kalman filter was compared to the Extended Kalman filter and the ensemble open loop, judging from the estimate of the mean and variance of the root mean squared errors. In this chapter, we present the conclusion and make some recommendations.

5.2 Conclusion

The extended Kalman filter is widely used data assimilation schemes for estimating the state of nonlinear dynamical systems. In this study we presented the Ensemble Kalman filter as an alternative to the EKF. The EnKF addresses many of the estimation issues in using the EKF and EnKF attains the same or better level of performance. The EnKF algorithm have been demonstrated on a number estimation problems since it introduction by Evensen (1994). In this study we implemented the ensemble Kalman filter on two estimation problems.

In summary, the scenarios considered are

1. The effect of different process noise on the performance of the EnKF
2. The effect of different measurement noise on the performance of the EnKF
3. The effect of different initial error covariance on the performance of the EnKF
4. The effect of different ensemble size on the performance of the EnKF

Comparing the performance of the EnKF to that of the EKF and the EnOL. On the basis of the comparison and the advantages of the EnKF over the EKF, the EnKF is proposed as an alternative to EKF in the pricing of financial instruments.

For the optimally determined parameters for the estimation problems considered, the EnKF was found to be the best performing filter among the three filters considered. In the pricing of financial options the EnKF closely predicted the observed especially in the case of the call option and the estimate of the interest rate and volatility are also obtained appropriately.

It was found that for the different process noises, the performance of the filters improved with decrease in the process noise. Same results were obtained for the different measurement noises and initial error covariances. In all the case the EnKF happened to be the best performing of the filters except in some few case where the EKF was performing better that the EnKF. The experiment showed that for high values of the noises, the filters break down as in the case of $Q = I_2$.

There was a decline in the mean and variance of the RMSE for the EnKF as the ensemble size grows or increases. That is to say that the performance of the EnKF improves with increases in the ensemble size.

The EnKF is a Monte Carlo approximation of the Kalman filter, representing the distribution of the system state by using random samples (ensembles) and computes the covariance from the ensemble. A major advantage of the EnKF is, it can be easily implemented and no analytical derivative need to be computed as in the case of the EKF. The EnKF relies on functional evaluations through the use of sample means and sample variances. The EnKF also has some limitations despite its clear advantage over the EKF.

5.3 Recommendation

Based on our findings, the following recommendations are made to further examine the performance of the EnKF. We recommend that the EnKF should be used as an alternative to the EKF and in using these filter the process noise, measurement noise and initial error covariance should be minimal. Large ensemble sizes are recommended for the use of EnKF.

It is also recommended for further research; the performance of the EnKF be verified for other factors that may affect the performance of the filter which were not considered in this research. Some of these may be to investigate the impact of varying observation frequency or sparse data on the performance of the filter, and so on.

Bibliography

J. D. Annan and J. C. Hargreaves. Efficient parameter estimation for a highly chaoticsystem. 2004.

J.D. Annan, J.C. Hargreaves, N.R. Edwards, and R. Marsh. Parameter estimation in an intermediate complexity earth system model using an ensemble kalman filter. *Frontier Research System for Global Change*, 2003.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, Vol. 81(No.3):637 – 654, May - June 1973.

N. Bolia and S. Huneja. Monte carlo methods for pricing financial options. Vol. 5, 2005.

F. Bouttier and P. Courtier. *Data assimilation concepts and methods*, March 1999. Meteorological Training Course Lecture Series.

G. Burgers, P. Jan Van Leeuwen, and G. Evensen. Analysis scheme in the ensemble kalman filter. 1998.

J. F. G. de Freitas, M. Niranjan, A. H. Gee, and A. Doucet. Sequential monte carlo methods to train neural network models. *Cambridge University Engineering Department, Cambridge CB2 1PZ, England, U.K*, 2000.

H. Etienne and E. Dombrowsky. Estimation of the optimal interpolation parameters in a quasi-geostrophic model of the northeast atlantic using ensemble methods. *Journal of Marine Systems*, 2002.

G. Evensen. Using the extended kalman filter with a multilayer quasi-geostrophic ocean model. *Geophysical Research*, 1992.

G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Geophysical Research*, 99(C5), May 1994.

G. Evensen. The ensemble kalman filter for combined state and parameter estimation. *IEEE Control Systems Magazine*, 2009.

G. Evensen, D. Dee, and J. Schroter. Parameter estimation in dynamical models. 1998.

A. R. Gallant, P. E. Rossi, and G. Tauchen. Stock prices and volume. *The Review of Financial Studies*, Vol. 5(No. 2):199–242, 1992.

A. Gupta and C. Reisinger. Robust calibration of financial models using baysian estimation. *Mathematical Institute, University of Oxford*, February 7 2012.

S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, Vol 6(No. 2), 1993.

J. C. Hull. *Options, Futures and Other Derivatives: 6th edition*. International Edition, Prentice Hall, New Jersey, 2006.

G. Indragoby and O. Pironneau. Calibration by automatic differentiation. 2004.

A. Javaheri, D. Lautier, and A. Galli. Filtering in finance. December 20 2002.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Transaction of the ASMEJournal of Basic Engineering*, 82 (Series D), 1960.

S. Kumar, A. Yildirim, Y. Khan, H. Jafari, K. Sayevand, and L. Wei. Analytical solutionof fractional black-scholes european option pricing equation by using laplace transform. *Journal of Fractional Calculus and Applications*, Vol. 2(No. 8):1–9, 2012.

R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, Vol. 4(No. 1):141–183, 1973.

H. T. Nguyen and I. T. Nabney. Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. *Non-linearity and Complexity Research Group, School of Engineering and Applied Science, Aston University, Aston Triangle, Birmingham, B4 7ET, UK*, July 15 2010.

N. Nikolaev and E. Smirnov. Unscented grid filtering and smoothing for nonlinear time series analysis. *World Academy of Science, Engineering and Technology*, 2009.

E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D.J. Patil, and J. A. Yorke. A local ensemble kalman filter for atmospheric data assimilation. *Institute for Research in Electronics and Applied Physics*, 2004.

S. Pastorello, E. Renault, and N. Touzi. Statistical inference for random-variance option pricing. *Journal of Business & Economic Statistics*, Vol. 18, No. 3:pp. 358–367, June 2000.

M. Rubinstein. Displayed diffusion option pricing. *Journal of Finance*, pages 213–217, 1983.

H. W. Sorenson. Least squares estimation: from gauss to kalman. *IEEE*, 1970.

E. A. Wan and R. van der Merwe. *Kalman Filtering and Neural Networks*. John Wiley & Sons, Inc., New York, 2001.

G. Welch and G. Bishop. An introduction to kalman filter. *Siggraph 2001, Los Angeles, CA*, August 2001.