

**PARAMETER ESTIMATION OF A CLASS OF HIDDEN MARKOV MODELS
USING SEQUENTIAL MONTE CARLO EXPECTATION-MAXIMIZATION
ALGORITHM**

BY

**AHANI, EDESIRI BRIDGET
(039075017)**

B Sc., MATHEMATICS/EDUCATION, DELTA STATE UNIVERSITY, ABRAKA

M Sc., STATISTICS, UNIVERSITY OF LAGOS, LAGOS

**A THESIS SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES
UNIVERSITY OF LAGOS**

**IN PARTIAL FULLFILMENT OF THE REQUIREMENT FOR THE AWARD OF
THE DEGREE OF DOCTOR OF PHILOSOPHY (Ph. D) IN STATISTICS**

**DEPARMENT OF MATHEMATICS
UNIVERSITY OF LAGOS**

JULY, 2012

**SCHOOL OF POSTGRADUATE STUDIES
UNIVERSITY OF LAGOS**

CERTIFICATION

This is to certify that the thesis

**PARAMETER ESTIMATION OF A CLASS OF HIDDEN MARKOV MODELS
USING SEQUENTIAL MONTE CARLO EXPECTATION-MAXIMIZATION
ALGORITHM**

**Submitted to the
SCHOOL OF POSTGRADUATE STUDIES
UNIVERSITY OF LAGOS**

FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY (Ph. D)

IS A RECORD OF ORIGINAL RESEARCH CARRIED OUT

BY:

**AHANI, EDESIRI BRIDGET
IN THE DEPARTMENT OF MATHEMATICS**

_____	_____	
Author's Name	Signature	Date
_____	_____	
First Supervisor's Name	Signature	Date
_____	_____	
Second Supervisor's Name	Signature	Date
_____	_____	
First Internal Examiner	Signature	Date

_____	_____	
Second Internal Examiner	Signature	Date
_____	_____	
External Examiner	Signature	Date
_____	_____	
SPGS Representative	Signature	Date

DEDICATION

This research is dedicated to God Almighty and to the blessed memory of my late parent,
Mr. and Mrs. Palmer Ahani.

ACKNOWLEDGEMENT

I lack fitting words to thank and praise God, for the strength, and opportunity He extended to me that culminated in the production of this work.

I am particularly indebted to my supervisor, Prof. Olaide Abass, for taking the time and having the patience to guide me through the course of this work. Sir, I will ever remain grateful for your corrections and contributions which helped me to give this work shape.

Special thanks go to my second supervisor, Dr. R. A. Kasumu, for his encouragement, support and contributions. Thank you so much for everything.

My appreciation goes to staff in the Department of Mathematics, especially Prof. R. O. Okafor, Prof. S. O. Ajala, Prof. J. O. Olaleru, Prof. S. A. Okunuga, Dr. M. O. Adamu, A. Adeniyi, Dr Mogbademu and Dr. H. Akewe, for their constructive criticisms and useful contributions. I would like to thank Oyebanjo Wasiu who gave me the much needed instructions that led to the carrying out of this research effectively.

Special thanks to my dear husband, Mr. Nkemnole Stanley, for being there for me. He has provided me with unconditional love and ceaseless support, in joy and in sorrow.

I will not fail to thank my siblings, Uriri, Hossana, Sabbath, Blessed and Mathew for their encouragements. I will not fail to cherish the memory of my late brother, Jordan, who left this world at the prime of his youth. You will ever remain green in our memory.

I am also grateful to all my friends and colleagues whom I am unable to name here, but

who supported me in one way or another during the course of this work. I owe all of you a great deal.

TABLE OF CONTENTS		Pages
Title pages	i	
Certification	ii	
Dedication	iii	
Acknowledgement	iv	
Table of Contents	v	
List of Figures	viii	
List of Tables	x	
Abstract	xi	

CHAPTER ONE: INTRODUCTION

1.1 Background to the Study	1
1.2 Statement of Problem	7
1.3 Objectives of the Study	8
1.4 Significance of Study	9

1.5 Research Questions

9

1.6 Scope of Study

10

1.7 Operational Definition of Terms

11

CHAPTER TWO: LITERATURE REVIEW

2.1 Markov Model

14

2.1.1 Introduction to Stochastic Processes

14

2.2 Hidden Markov Model (HMM)

17

2.2.1 The Basic HMM

18

2.2.2 The Inference Problems

21

2.2.3 State-Space

Representation

22

2.3 Examples

23

2.3.1 Finite HMM

23

2.3.2 Normal Hidden Markov Model

24

2.3.3 Gaussian Linear State Space Model

24

2.3.4	General State Space HMM	25
2.4	Parameter Estimation in Stochastic Volatility model	27
2.4.1	Review of Methods of Parameter Estimation of SV Model	29
2.4.2	Review of Multivariate Stochastic Volatility Model	37
2.4.3	Review of Methods in Bearings-Only Target	40
2.5	Review of Expectation Maximization (EM) Algorithm	44
2.5.1	Monte Carlo EM Algorithm	46
2.6	Sequential Monte Carlo- Particle Filters and Smoothers	47
2.6.1	Simple Monte Carlo Sampling	50
2.6.2	Importance Sampling	50
2.6.3	Sequential Importance Sampling	51
2.6.4	Sequential Importance Resampling	52
2.6.5	Algorithm Description	54
2.7	Particle Filter Algorithm	55
2.7.1	Monte Carlo Filtering for General State Space Models	56

2.8 Particle Smoothing Algorithm

56

2.8.1 Particle Smoother using Backwards Simulation

57

CHAPTER THREE: THEORETICAL FRAMEWORK AND METHODOLOGY

3.1 Stochastic Volatility Model

58

3.2 Stochastic Volatility Model with Heavy Tail

59

3.2.1 Stochastic Volatility Model Normal Mixture

60

3.2.2 The Proposed Model

61

3.2.2.1 Student-t as an Observation Noise

62

3.2.2.2 Generalized Error Distribution as an Observation noise

63

3.3 Sequential Monte Carlo Expectation Maximization Analysis

64

3.3.1 Filtering Step

64

3.3.2 Smoothing Step

65

3.3.3 Estimation Step

66

3.4 Other Considerations

71

3.4.1 Initial Parameter Selection

71

3.4.2 Relative Likelihood	
72	
3.4.3 Stopping Rule and Selection of Particle Size	
72	
3.5 Proof of Convergence of the Proposed technique	
74	
3.5.1	Theorem
74	
3.5.2 Consistency and Asymptotic	Normality
79	
3.6 Multivariate Stochastic Volatility Model based on Student-t distribution	
82	

CHAPTER FOUR: SIMULATION RESULTS AND DATA ANALYSIS

4.1 Simulation Study	
85	
4.2 Estimation Results	
88	
4.3 Application to Real Life Financial Data	
94	
4.3.1 Estimation Results for Real Life Financial Data	
96	
4.4 Akaike Information Criteria (AIC) – Model Selection Criteria	
98	
4.4.1 Estimation Results for the MSV model	
102	
4.5 Application to Bearing-Only Tracking	
104	
4.5.1 SMCEM for Bearings-Only Tracking	
106	

4.5.2 Tracking Performance

108

**CHAPTER FIVE: CONCLUSION AND CONTRIBUTIONS TO
KNOWLEDGE**

5.1 Summary of Findings

112

5.2 Conclusion

113

5.2 Contributions to knowledge

114

5.3 Suggestions for Further Research

115

REFERENCES

116

APPENDIXES

128

LIST OF FIGURES

2.1	visualization of the concept of a Markov Model.	16
2.2	Graphical representation of the dependence structure of a HMM, where y_1, y_2, y_3 are the observable states and x_1, x_2, x_3 are the hidden states	19
2.3	Sequential Importance Sampling particle filter	53
2.4	problem of the SIS particle filter: degeneracy	54
2.5	Resampling	55
4.1:	Time plot and histogram of simulation 1	86
4.2:	Time plot and histogram of simulation 2	86
4.3:	Time plot and histogram of simulation 3	87
4.4:	Parameter estimation results and the relative likelihood for technique based on student-t: for data generated from normal mixture model	91
4.5:	Parameter estimation results and the relative likelihood for technique based on student: for data generated from student-t model	91
4.6:	Parameter estimation results and the relative likelihood for technique based on normal mixture: for data generated from student-t model	91
4.7	Parameter estimation results and the relative likelihood for technique based on normal mixture: for data generated from normal mixture model	91
4.8	Parameter estimation results and the relative likelihood for technique based on GED:	

	for data generated from GED model	
	93	
4.9	Parameter estimation results and the relative likelihood for technique based on GED: for data generated from normal mixture model	
	93	
4.10	Parameter estimation results and the relative likelihood for technique based on normal mixture: for data generated from GED model	
	93	
4.11	Parameter estimation results and the relative likelihood for technique based on normal: for data generated from normal mixture model	
	93	
4.12	Naira/dollar daily exchange rate and the log returns	
	95	
4.13	Cedi/dollar exchange rate and the log returns	
	95	
4.14	Euro/dollar daily exchange rate and the log returns	
	95	
4.15	Pound/dollar daily exchange rate and the log returns	
	96	
4.16	Estimation results for the naira/dollar exchange rates	
	98	
4.17	Estimation results for the cedi/dollar exchange rates	
	98	
4.18	Estimation results for the Euro/dollar exchange rates	98
4.19	Estimation results for the pound/dollar exchange rates	98
4.20	The Bearings-Only Tracking (BOT) problem.	105
4.21	Representation of the trajectories of the true target path, observer, the SMCEM estimate and position MSE	
	109	
4.22.	Estimates of the vehicle being tracked	
	111	

LIST OF TABLES

4.1	Parameter estimation on technique based on the normal mixture and Student-t on data generated from the normal mixture model	88
4.2	Parameter estimation on technique based on the normal mixture and Student-t on data generated from the Student-t model	89
4.3	Parameter estimation on technique based on the normal mixture and GED on data generated from the normal mixture model	92
4.4	Parameter estimation on technique based on the normal mixture and GED on data generated from the GED model	93
4.5	Descriptive statistics of Daily Returns for the Exchange Rate	94
4.6	Estimation Results - Distributions Comparison for the returns	97
4.7	Evaluation Statistics - Distributions Comparison for the returns	101
4.8	SMCEM estimation results for the MSV model	102
4.9	unconditional variance–covariance matrix of the returns	103
4.10	sample variance– covariance matrix of the returns	103

ABSTRACT

Much research has been advanced in the development of Monte Carlo methods for stochastic processes. A particular focus is on sequential Monte Carlo methods (particle filters and particle smoothers) and the Expectation-Maximization (EM) algorithm which allows the estimation of a class of Hidden Markov Models (HMMs) with nonlinear, non-Gaussian state-space models. The Stochastic Volatility (SV) model can be regarded as a nonlinear state space model. SV model has become increasingly popular for explaining the behaviour of financial variables (e.g. stock prices and exchange rates). This has resulted in several different proposed approaches to estimating the parameters of the model. This thesis proposes a Sequential Monte Carlo Expectation Maximization (SMCEM) approximation method for the nonlinear state space representation and applies it for estimating the SV model. The basic idea of our approach is to combine the Expectation-Maximization (EM) algorithm with particle filters and smoothers in order to estimate parameters of the model. In addition to mixture-of-normal distributions of Kim & Stoffer (2008), the scope of application of SV models is expanded by adopting a student-t and the Generalized Error Distribution (GED), for the observational error term. To establish the viability of the extended volatility models, simulation studies as well as real life data analysis results are presented. Furthermore, the research establishes the convergence properties of the proposed technique. The results obtained from the models indicate that the student-t and the GED are comparable to the normal mixture SV model but empirically more successful. The proposed model allows for a more robust fit, giving

us a new tool to explore the tail fit. In the same vein, there are theoretical as well as empirical reasons to study multivariate volatility models. The application of the SMCEM approach to a multivariate factor model with stochastic volatility using the student-t distribution indicates that it performs quite well in explaining the joint dynamics in the volatility of a number of asset returns.

In the same vein, this work applies the proposed procedure to nonlinear problems in signal processing such as bearings-only tracking; again the procedure is successful in accommodating nonlinear model for a target tracking scenario.

Key words: Hidden Markov model, Stochastic Volatility model, State-space model, Sequential

Monte Carlo, EM algorithm, Student-t distribution

CHAPTER ONE

1. INTRODUCTION

1.1 BACKGROUND OF STUDY

Hidden Markov Models (HMMs), originally introduced in 1957, [see MacDonald and Zucchini (1997) and Cappe et al., (2005)] have found many applications in most contemporary fields like signal processing, medicine, engineering, and management applications. It is a doubly stochastic process $(X_t, Y_t)_{t \geq 0}$, with an underlying stochastic process, X_t , that is not directly observable but can be observed only through another process, Y_t , that produces the sequence of independent random observations. MacKay (2003) has observed that it is used for two purposes, namely, to make inferences about an

unobserved process based on the observed one, and to explain variation in the observed process based on variation in a postulated hidden process. Rabiner (1989) put it thus:

First the model is very rich in mathematical structure and hence can form the theoretical basis for use in a wide range of applications. Second the models, when applied properly, works very well in practice for several important applications.

A number of representative classes of hidden Markov models in applications are taken from a variety of fields:

- finite-hidden Markov model [DNA or protein sequencing, Churchill, (1992)]
- Normal hidden Markov model [ion channel modeling, Ball and Rice (1992)]
- Gaussian Linear State-Space Model [signal processing, Juang and Rabiner (1991)]
- Conditionally Gaussian Linear State-Space Models [Digital communication, Punskey et al., (2002) and Fearnhead (2003)].
- non-linear Gaussian state-space model [stochastic volatility, Shumway & Stoffer (2000)].

In HMMs, the output probabilities impose a veil (Ferguson, 1980) between the state sequence and the observed time series. In an effort to lift the veil, a substantial body of theories has been developed over the past years. The initial work dealt with finite probability spaces and addressed the problems of tractability of probability computation, iterative maximum-likelihood estimation of model parameters from observed time series and the proof of consistency of the estimates: [see Baum and Eagon (1967), Baum and Petrie (1966)]. In most cases, however, exact maximum likelihood estimation is impossible as the likelihood function of the model cannot be evaluated explicitly. Consequently, a lot of papers have been concerned with approximations by means of numerical and simulation techniques (see Del Moral et al., 2001). There are many

methods of identifying the parameters of a nonlinear stochastic system among which are maximum likelihood estimators and Bayes estimators (Casella and Berger, 2001). The latter especially, is related to the sequential Monte Carlo (SMC) methods (Kitagawa, 1996) also known as Particle filter methods introduced by Gordon et al., (1993).

HMMs are equivalently defined through a functional representation known as state space model. State-space model of a HMM certainly is one of the concepts of statistical model processing that has had profound practical impact in the recent years. Recent advances in this field include SMC filtering approximation techniques which makes it possible to perform inference in models that are more general than linear Gaussian models [Cappe et al., (2005) and Doucet et al., (2001)]. Estimation of parameters in general state-space models is a crucial but often difficult task (Cappe and Moulines, 2009). An online Expectation-Maximization (EM) algorithm proposed for HMM (Cappe, 2009) can be extended to more general settings, including non-linear non-Gaussian state-space models that necessitate the use of SMC filtering approximations. These advances notwithstanding, the estimation of the model parameters remains challenging.

The state space model (Doucet and Johansen, 2009) of a HMM is represented by the following two equations:

$$\text{(State equation)} \quad x_t = f(x_{t-1}) + w_t \quad (1.1)$$

$$\text{(Observation equation)} \quad y_t = g(x_t) + v_t \quad (1.2)$$

where f and g are either linear or nonlinear functions, while w_t and v_t are white noise processes. Models represented by (1.1) and (1.2) are referred to as state space model and this includes a class of HMMs with non-linear Gaussian state-space model such as stochastic volatility model and the bearings-only tracking model.

Stochastic Volatility (SV) model is prominent in the formulation of volatility by an unobservable process that allows the volatilities to vary stochastically (Taylor, 1986, 1994). The SV model introduced by Taylor (1982) accounts for the time-varying and persistent volatility, and the leptokurtosis in financial return series. The model has been commonly and successfully used to explain the behavior of financial variables such as stock prices and exchange rates [Durbin and Koopman (2000) and Doucet and Tadic (2003)].

In the early stage of time series data analysis, researchers [Box and Jenkins (1976) and Granger (1981)] were interested mainly in mean behaviour of the data and tried to find a model which could explain it effectively.

Recently, concern about volatility in the data has been raised as changes in volatility are observed in real data, especially in financial data, and knowledge of volatility can be a good piece of information for decision-making process. Furthermore, if it is possible to predict the future variance of financial variables, it would be very useful to control the risk. For example, if there are two stocks which have the same mean but different variances, people would prefer the stock with smaller variance in that it is less risky. SV models, which are time series models concerned with volatility, would provide this type of information.

SV model has become increasingly popular for explaining the behaviour of financial variables, and its popularity has resulted in several different proposed approaches regarding the problem of estimating the parameters of the model. It models the variance as an unobserved component that follows a particular stochastic process. Though

theoretically attractive, it is empirically challenging owing to the fact that the unobserved volatility process enters the model in a non-linear fashion, leading to the likelihood function depending upon high-dimensional integrals. In the standard form of the model, the SV is modeled as an autoregressive (AR) process,

$$x_t = \phi x_{t-1} + w_t \quad (1.3)$$

$$y_t = \beta \exp\left(\frac{x_t}{2}\right) v_t \quad (1.4)$$

where $w_t \sim N(0, \tau)$, $x_0 \sim N(\mu_0, \sigma_0^2)$, $v_t \sim N(0, 1)$, $\{y_t\}_{t \geq 0}$ is the log-returns on day t , we call β the constant scaling factor, so that $\{x_t\}_{t \geq 0}$ represents the log of volatility of y_t . In order to ensure stationarity of y_t , it is assumed that $|\phi| < 1$. Squaring (1.4) and taking the logarithm of it results in a linear equation (1.5),

$$y_t = \alpha + x_t + z_t. \quad (1.5)$$

Equations (1.3) and (1.5) form the version of the SV model which can be modified in many ways; together they form a linear, non-Gaussian, state-space model for which (1.5) is the observation equation and (1.3) is the state equation.

SV model usually assumes that the distribution of asset returns conditional on the latent volatility is normal. Financial data often have heavier tails than can be captured by the standard SV model. This has naturally led to the use of non normal distributions to “better-model” and to address the problem of heavy tails. [see Shephard (1996), Kim et al., (1998), Bai et al., (2003), Sadorsky (2005) and Kim and Stoffer (2008)]. Liesenfeld and Jung (2000) fit a student-t distribution to the error distribution in the SV model using the simulated maximum likelihood method developed by Danielsson and Richard (1993)

and Danielsson (1994). A promising distribution that models both the skewness and the kurtosis is the Skewed Student-t of Fernandez and Steel (1998). Hence, there is the utmost necessity to determine the best-fitted model out of a potentially huge class of candidates. In other words, it has become very pertinent to develop efficient model selection criteria. The above background shows the ever-growing literature on time-varying financial market volatility which abounds in empirical studies in which competing models are evaluated and compared on the basis of their forecast performance , Andersen et al., (2005).

In addition, there are theoretical as well as empirical reasons to study multivariate volatility models. This multivariate specification accounts not only for the volatility dynamics of the individual assets but also, due to the common factor, for time varying correlations across assets returns which require simultaneous multivariate estimation for full efficiency.

Supposedly, solutions to nonlinear problems such as bearings-only tracking, assume that the state transition and measurement models are known a priori [see Gordon et al., (1993), Carpenter et al., (1999) and Bergman (1999)]. Nonetheless, there are situations where the model parameters are not known a priori or are known only partially and can be considered as a special case of maximum likelihood estimation with incomplete data. In such situations, standard estimation algorithms like the Kalman filter and the extended Kalman filter (EKF) which assume perfect knowledge of the model parameters are not accurate. Recently, the growth in computational power has made computer intensive

statistical methods feasible. Both Markov Chain Monte Carlo methods and sequential Monte Carlo methods (particle filters) are now implemented in several applications.

From a Bayesian perspective, the tracking problem can be solved by recursively calculating some degree of belief in the target state, taking different values, given available observations. Since the target state uncertainty and the measurement-originated uncertainty are the two major unavoidable obstacles for target tracking, a good model of the target motion will effectively facilitate the design of the required tracking algorithm.

1.2 STATEMENT OF PROBLEM

Estimation of parameters of a class of HMMs such as SV model has been the focus of much academic research, in recent years. [see Sandman and Koopman (1998), Danielson (1994), Chib et al., (2002), Gallant et al., (1995), Djuric et al., (2002) Shumway and Stoffer (2000) and Doucet and Tadic (2003)]. The reason is not far-fetched: parameter estimation for the SV-type model is intractable when the evaluation of the likelihood function can not be expressed in a closed form. As a result, likelihood-based inferences are computationally demanding. A number of specifications of the SV model has been suggested since its introduction by Taylor (1982).

Recent works extend the basic SV model to include heavy-tailed error distributions as financial data often have heavier tails than can be captured by the standard SV model. [see Liesenfeld and Richard (2003), Bai et al., (2003), Sadorsky (2005), Shimada and Tsukuda (2005) and Kim and Stoffer (2008)]. The recent progress in the incorporation of the Expectation-Maximization algorithm computation and Sequential Monte Carlo

methods (particle filters and smoothers), as developed in Kim and Stoffer (2008), has made it possible to fit increasingly complex statistical models. It is however necessary to determine the best-fitting model out of a potentially huge class of candidates.

Consequently, the focus of this research is to develop a method that can handle the estimation problem of SV model formulated as a state space form of a HMM and in particular to address the problem of heavy tails, by analyzing the model using non normal distributions as the error distribution. Ultimately, a comparison will be drawn to determine the distribution that is most adequate in the analysis of SV models. The method is also applied to solve the nonlinear bearing-only tracking problem in order to verify whether the procedure is successful in accommodating other nonlinear model state space model.

1.3 OBJECTIVES OF THE STUDY

The aim of this study is to develop a method for estimating parameters of a class of HMM with non-linear Gaussian state-space model such as stochastic volatility model using sequential Monte Carlo Expectation Maximization (SMCEM) technique and to extend results in the study and analysis of SV model with the parameter estimated in the procedure. The specific objectives are to:

- i) extend the SV model by modelling the observational error using a student-t distribution and the generalized error distribution (GED), and juxtapose them with the normal mixture distribution of Kim and Stoffer and then apply model diagnostic tools of Ibrahim et al., (2008), to determine the most empirically successful;

- ii) improve on the results of Kim and Stoffer parameter estimation by techniques based on the student-t and GED SV model and a strategy for fitting the model that combines the EM algorithm and SMC;
- iii) establish the viability of the proposed model by simulation studies and real data;
- iv) establish the convergence properties of the proposed technique using the theorem of Chan and Ledolter, (1995);
- v) build multivariate SV models for asset returns in financial economics and apply the proposed procedure to other state-space model such as the bearings-only tracking model, in other to show that it is successful in accommodating nonlinear model.

1.4 SIGNIFICANCE OF THE STUDY

The Sequential Monte Carlo Expectation Maximization (SMCEM) algorithm for estimating parameter for the extended model provides a highly generic procedure for an accurate Monte Carlo (MC) evaluation of the marginal likelihood which depends upon high-dimensional interdependent integrals. Real data can be very complex, typically involving elements of non-Gaussianity, high dimensionality and non linearity, which conditions usually preclude analytic solution. This is a problem of fundamental importance that permeates most discipline. The proposed technique can provide a convenient and attractive approach to computing the complex distributions. Also, By estimating parameters, the dynamics of the process can be fully specified, and future values can be estimated from them. In addition, the proposed technique can be used to

analyse data from many areas, such as economics, medicine and engineering, thus widening the scope of its application to other state-space models.

1.5 RESEARCH QUESTIONS

This research was carried out in an attempt to answer the following questions:

- i. Using SMCEM algorithm, how can we develop a method to estimate the parameters for the extended volatility model?
- ii. How do we employ the use of the student-t distribution and the GED, while juxtaposing both with the normal mixture distribution of Kim and Stoffer (2008) and the method proposed by Ibrahim et al., (2008), to extend the SV model?
- iii. How can the much needed improvement of the known outcomes of parameter estimation be effected through techniques based on the student-t and the GED SV model?
- iv. By simulation studies and real data, does the student-t or GED SV model exhibit viability?
- v. How do we establish the convergence properties of the proposed technique by using the theorem of Chan and Ledolter, (1995)?
- vi. How can we build multivariate SV model for asset returns in financial economics, and apply the proposed technique to other state-space models such as the bearings only tracking model?

1.6 SCOPE OF THE STUDY

This research work is limited to a class of HMMs with non-linear Gaussian state-space model such as stochastic volatility model with a student-t distribution and the GED. A comparison is drawn between these two models and the mixture-of-normal distributions of Kim and Stoffer. A Sequential Monte Carlo with Expectation Maximization (SMCEM) algorithm techniques is used to estimate the parameters of the extended

volatility model; this is extended to the multivariate SV models. Model diagnostic tools are employed to compare the fit of the distributions. The performance of the SV models is assessed using three simulated data sets and daily exchange rate series of the Nigerian Naira, Ghana Cedi, British Pound and Euro, all against the U. S. Dollars, from March 3rd, 2009 to March 3rd, 2011. The proposed procedure is also applied to solve a nonlinear problem in signal processing such as bearing-only tracking model.

1.7 OPERATIONAL DEFINITION OF TERMS

The following definitions are useful in this research work.

Filtering distribution: a filtering distribution is a distribution function which is used to estimate the sequence of hidden parameters, x_t , for $t \geq 0$, based only on the observed data y_t , for $t \geq 0$. It corresponds to estimating the distribution of the current state of an HMM based upon the observations received up until the current time.

General state-space models: are represented by the following two equations

$$x_t = F_t(x_{t-1}, w_t) \quad (1.6)$$

$$y_t = H_t(x_t, v_t) \quad (1.7)$$

Kurtosis: is the quality property of a distribution which is based on the peakedness and tails of the distribution. The degree of non-Gaussianity of a data set of most interest is measured by the kurtosis of the distribution.

Leptokurtic distribution: In probability theory, a leptokurtic distribution is a probability distribution with a kurtosis greater than that of a normal distribution.

Likelihood function: Let $f(x, \theta)$ represent a probability density function indexed by a parameter θ which can be a scalar or vector. Let x_1, x_2, \dots, x_n be a random sample drawn from the distribution. The likelihood of θ given observations x_1, x_2, \dots, x_n is given by

$$L(\theta, x_1, x_2, \dots, x_n) = \prod f(x_i; \theta).$$

Linear State-Space Model: is when the state equation and the observation equation are as (1.8) and (1.9).

$$x_t = \phi x_{t-1} + w_t \quad (1.8)$$

$$y_t = B_t x_t + v_t \quad (1.9)$$

Linear Gaussian State-Space Model: is when both w_t and v_t are normal distributions in (1.8) and (1.9).

Simulation: is an attempt to model a real-life or hypothetical situation so that it can be studied to see how the system works.

Smoothing distribution: smoothing distribution is a distribution that corresponds to estimating the distribution of the state at a particular time given all of the observations up to some later time.

State space model: is in principle any model that includes an observation process y_t and a state process x_t . The equations may be nonlinear or non-Gaussian.

Particle: is a realization of the stochastic process.

Volatility: is the relative rate at which the price of financial data moves up and down. If the price of a stock moves up and down rapidly over short time periods, it has high volatility. If it almost never changes, it has low volatility

$f_t^{(j)}$: j^{th} particle filter at time t which approximates $f(x_t|Y_t), .$

$s_t^{(j)}$: j^{th} particle smoother at time t which approximates $f(x_t|Y_n), (n > t) .$

$p_t^{(j)}$: j^{th} particle predictor at time t which approximates $f(x_t|Y_s), (s < t) .$

$w_t^{(j)}$:	a weight associated with $f_t^{(j)}$ or $s_t^{(j)}$.
χ_1^2	chi-square distribution with one degree of freedom.

CHAPTER TWO

2. LITERATURE REVIEW

The likelihood of the SV model formulated as a state-space form of a Hidden Markov Model is intractable as the evaluation of the likelihood function cannot be expressed in a closed form. This has stimulated the development of various inference techniques for their parameters and so either an approximation of the likelihoods or numerical methods has been considered. Lately, the incorporation of the EM algorithm and SMC (particle filters and smoothers) forms a basic idea in handling the estimation problem. Consequently, this chapter takes a look at the Markov Model, Hidden Markov Model, and a review of literature of other methods by different researchers in SV model and the

bearings only tracking problem, the EM algorithm and SMC (particle filters and smoothers).

2.1 THE MARKOV MODEL

2.1.1 INTRODUCTION TO STOCHASTIC PROCESSES

Stochastic modelling is an interesting and challenging area of probability and statistics.

This introductory section, while explaining what a stochastic process is, also strives to describe what is meant by the Markov property, with examples.

Suppose that each performance of an experiment yields a real-valued function on an interval. If the experiment is performed at times t_0, t_1, \dots, t_n ($t_0 < t_1, \dots, t_n$), $X_{t_0}, X_{t_1}, \dots, X_{t_n}$ are random variables. A stochastic process is a sequence $X_{t_0}, X_{t_1}, \dots, X_{t_n}$ of random variables based on the same sample space Ω . The sequence of outcomes when repeatedly casting the die is a stochastic process. The possible outcomes of the random variables are the set of possible states of the process.

Basically a stochastic process is an indexed set of random variables $\{X_t, t \in I\}$. The index set I is either a finite set or an interval in R and is usually thought of as the time variable.

Assume that for any finite set of indices $\{t_1, \dots, t_n\} \subset I$, the joint density $\phi(x_{t_1}, x_{t_2}, \dots, x_{t_n})$ exists. For an ordered set of indices $t_1 \geq t_2 \geq \dots \geq t_k \geq t_{k+1} \geq t_{k+2} \geq \dots \geq t_{k+n}$, we define the conditional density

$$\phi(x_{t_1}, \dots, x_{t_k} \mid x_{t_{k+1}}, \dots, x_{t_n}) = \phi(x_{t_1}, x_{t_2}, \dots, x_{t_n}) / \phi(x_{t_{k+1}}, x_{t_{k+2}}, \dots, x_{t_n}) \text{ (see Lawler, 1995).}$$

For Manning and Schutze (1999), it is a family of random variables $\{X_\theta\}$, indexed by a parameter θ , where θ belongs to some index set Θ . If Θ is a set of integers,

representing specific time points, we have a stochastic process in discrete time and we replace the general subscript θ by n . If Θ is the real line (or some interval of the real line), we have a stochastic process in continuous time and we replace the general subscript θ by t . In general, for a discrete time process, the random variable X_n will depend on earlier values of the process, X_{n-1}, X_{n-2}, \dots . Similarly, in continuous time, X_t will generally depend on values X_u for $u < t$. Therefore, we are often interested in conditional distributions of the form

$$p(X_{t_k} = s_{t_k} \mid X_{t_{k-1}} = s_{t_{k-1}}, X_{t_{k-2}} = s_{t_{k-2}}, \dots, X_{t_1} = s_{t_1}) , \quad (2.1)$$

for some set of times $t_k > t_{k-1} > \dots > t_1$. where s_t denotes a state at time t . In general, this conditional distribution will depend upon values of $X_{t_{k-1}}, X_{t_{k-2}}, \dots, X_{t_1}$.

However, emphasis will be laid on the processes that satisfy the Markov property, which says that

$$p(X_{t_k} = s_{t_k} \mid X_{t_{k-1}} = s_{t_{k-1}}, X_{t_{k-2}} = s_{t_{k-2}}, \dots, X_{t_1} = s_{t_1}) = p(X_{t_k} = s_{t_k} \mid X_{t_{k-1}} = s_{t_{k-1}}) \quad (2.2)$$

The Markov process is a special type of stochastic process. Historically, the Markov property is named after the Russian probabilist, Andrei Andreyevich Markov (1856-1922) who began the study of an important new type of dynamic stochastic chance process. In this class of processes, the outcome of a given experiment can influence the outcome of the next experiment. An informal mnemonic for remembering the Markov property is this. ‘Given the present X_{k-1} , the future X_k is independent of the past $X_{k-2}, X_{k-3}, \dots, X_1$.

These are the Markov properties described in (Manning and Schutze, 1999):

$$p(X_{t+1} = s_{t+1} \mid X_0 = s_0, X_1 = s_1, \dots, X_t = s_t) = p(X_{t+1} = s_{t+1} \mid X_t = s_t) \quad (2.2)$$

As time passes the process remain or change to another state.

Markov models are mathematical models of stochastic processes that generate random sequences of outcomes according to certain probabilities. The figure below helps to visualize the concept of a Markov Model.

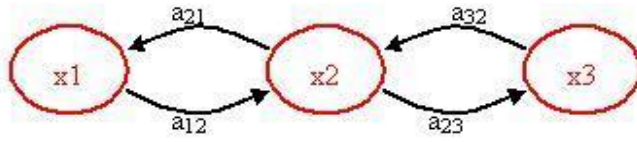


Figure 2.1 *visualization of the concept of a Markov Model*

$x_1, x_2, x_3 = \text{states}$, $a_{12}, a_{21}, a_{23}, a_{32} = \text{transition probabilities}$

Symbolically, x_1, x_2 and x_3 in the diagram represent the possible states of the process and the arrows represent transitions between states. The label on each arrow a_{ij} 's represents the transition probability. These numbers constitute the transition matrix.

A Markov model consists of:

- a finite set of states $\Omega = \{x_1, \dots, x_n\}$.
- an $n \times n$ state transition matrix $P = \{p_{ij}\}$, where $p_{ij} = P(X_{t+1} = j \mid X_t = i)$.
- and an initial vector $\pi_i = P(X_1 = s_i)$.

In this context, the Markov property suggests that the distribution for this variable depends only on the distribution of the previous state. In many cases, the prediction of the next state and its associated observation only depends on the current state, meaning that the state transition probabilities do not depend on the whole history of the past process.

However, in some cases, the patterns that we wish to find are not described sufficiently by Markov process, as the observed sequence of states is probabilistically related to the hidden process. Such processes are modelled using a hidden Markov model.

2.2 HIDDEN MARKOV MODEL (HMM)

An HMM is an extension of Markov model which models the dynamics of hidden states each of which produces a value of an observable (Visser et al., 2000). It assumes that there is a set of states generating our data, and that the identity of each successive state still depends only on the state(s) before it, but now we cannot observe these states directly. Instead, each state is associated with an emission probability density function that generates our observed data.

Li et al., (2000) define HMM as a set of states, each of which is associated with a probability distribution. Transitions among the states are governed by a set of probabilities called transition probabilities. In a particular state an outcome or observation can be generated, according to the associated probability distribution. It is only the outcome that is visible to an external observer and not the state. Therefore, states are hidden to the outside, hence the name HMM (Rabiner and Juang 1986).

Grundy et al., (1999) describes HMM as a mathematical framework which models a series of observations based upon a hypothesized, underlying but hidden process. The model consists of a set of states and transitions between these states. Each state emits a signal based upon a set of emission probabilities and then stochastically transits to some other state, based upon a set of transition probabilities. These two probability distributions, when combined with the initial distribution characterized an HMM.

Bickel et al., (1998) have defined HMM as a discrete-time stochastic process $\{(X_t, Y_t)\}$ such that (i) $\{X_t\}$ is a finite-state Markov chain, and (ii) given $\{X_t\}$, $\{Y_t\}$ is a sequence of conditionally independent random variables with the conditional distribution of Y_n depending on $\{X_t\}$ only through X_n .

2.2.1 THE BASIC HIDDEN MARKOV MODEL

Let $\{Y_t\} = \{Y_t, t \geq 0\}$ denote a sequence of observations and $\{X_t\} = \{X_t, t \geq 0\}$ a Markov chain defined on the state space $\{0, 1, \dots, J-1\}$. This Markov chain is assumed to be discrete, homogeneous, and irreducible with transition matrix. For better readability, we use the notation

$$Y_{t_0}^{t_1} = \{Y_{t_0}, \dots, Y_{t_1}\}$$

with $t_0 < t_1$; $X_{t_0}^{t_1}$ is defined similarly.

Consider a stochastic process consisting of two parts: Firstly, the unobserved parameter process $\{X_t\}$, which fulfils the Markov property

$$p(X_t = x_t | X_1^{t-1} = s_1^{t-1}) = p(X_t = x_t | X_{t-1} = x_{t-1}), \quad (2.3)$$

and secondly the state dependent observation process $\{Y_t\}$, for which the conditional independence property

$$p(Y_t = y_t | Y_0^{t-1} = y_0^{t-1}, X_0^t = x_0^t) = p(Y_t = y_t | X_t = x_t) \quad (2.4)$$

holds. Then, the pair of stochastic processes $\{X_t, Y_t\}$ is Hidden Markov Model. The basic structure of the HMM component is illustrated below:

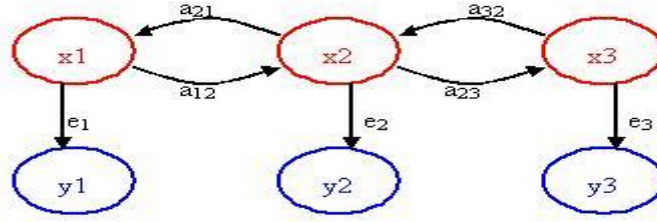


Figure 2.2 : Graphical representation of the dependence structure of a HMM, where

y_1, y_2, y_3 are the observable states and x_1, x_2, x_3 are the hidden states

- A set of states (x 's).
- A set of possible output symbols (y 's).
- A state transition matrix (a 's), probability of making transition from one state to another:

$P = \{p_{ij}\}$, $p_{ij} = p(X_{t+1} = j \mid X_t = i)$, $p_{ij} \geq 0$, $1 \leq i, j \leq N$, X_t , denotes the current state.

- Output emission matrix (e 's), probability of emitting or observing a symbol at a particular state:

$B = [b_i(k)]$, $b_i(k) = p(y_t = v_k \mid x_t = j)$, $1 \leq i \leq N$, $1 \leq k \leq M$, v_k , denotes the k^{th} observation symbol per state.

- Initial probability vector, probability of starting at a particular state: $\pi_i, i \in S$,
 $\pi_i = p(X_1 = s_i)$.

An application of HMM requires specification of two model parameters (N and M) - the number of set of states an set of possible output - and of the three probability measures $\{P, B, \pi\}$. For convenience, we use the compact notation

$$\lambda = \{P, B, \pi\},$$

to indicate the complete parameter for HMMs.

Two assumptions are made by the model. The first, is the Markov assumption, which states that the current state is dependent only on the previous state,

$$p(X_{t+1} \mid X_1^t) = p(X_{t+1} \mid x_t). \quad (2.5)$$

The independence assumption states that the output observation at time t is dependent only on the current state, it is independent of previous observations and states:

$$p(Y_t \mid Y_1^{t-1}, X_1^t) = p(Y_t \mid X_t). \quad (2.6)$$

2.2.2 THE INFERENCE PROBLEMS

Once a system can be described as a HMM, some inference problems must be solved for the model to be useful in real-world applications (Rabiner, 1989).

- **Model evaluation:** What is $p(y_1, \dots, y_t)$ under the current model?
- **State prediction (“filtering”):** What is $p(x_{t+1} \mid y_1, \dots, y_t)$? That is, what do we think the current state is, given the history of observations?
- **State estimation (“smoothing”):** What is $p(x_t \mid y_1, \dots, y_t)$? That is, what do we think the true state was at a certain time, given all the observations including those after that time?
- **Trajectory estimation:** what is the most likely trajectory through states, given $p(x_t \mid y_1, \dots, y_t)$?
- **Learning:** How can the factors $p(x_1)$, $p(x_{t+1} \mid x_t)$ and $p(y_t \mid x_t)$ be improved in the light of data?

One of the most important problems in HMM theory is that of parameter estimation. The relatively long history of this problem started with the works of Baum and Petrie (1966) and Petrie (1969), which demonstrated the consistency and asymptotic normality of the maximum likelihood estimator (MLE) for some classes of finite HMM. These results

were extended to other classes of HMM [see Bickel et al., (1998), Jensen and Peterson, (1999) and Leroux (1992)]. Another direction of research in the area of parameter estimation for HMM is the development of computationally efficient methods of maximizing the likelihood function. The Baum-Welch algorithm, which is based on the expectation-maximization (EM) algorithm for HMMs, is one of the most popular methods [Archer and Titterton (2002), Baum et al., (1970) and Rabiner (1989)]. In practice, the MLE is often computed using EM algorithm (Dempster et al., 1977). However the EM algorithm is a deterministic algorithm, which is sensitive to initialization and can become trapped in severe local maxima. To avoid this and to deal with cases where the E-step cannot be performed in closed-form, some Monte Carlo variants of the EM algorithm have been proposed [see Wei and Tanner (1991) and Chan and Ledolter (1995)]. The more challenging issue of HMM with continuous state space was much studied throughout the 1990s, mostly using simulation-based approaches made possible by recent advances in Monte Carlo methods [Doucet et al., (2002) and Jacquier et al., (2007)].

2.2.3. STATE-SPACE REPRESENTATION

In great generality, an HMM is equivalently defined through a functional representation known as a (general) state-space model

$$x_t = f_t(x_{t-1}, w_t) \quad (2.7)$$

$$y_t = g_t(x_t, v_t) \quad (2.8)$$

where f and g are either linear or nonlinear functions, w_t and v_t are white noise processes. Practically, the x 's are the unseen true signals in signal processing (Liu and Chen, 1995), the actual words in speech recognition (Rabiner, 1989), the target features in a multitarget tracking problem (Avitzour, 1995), the image characteristics in computer vision (Isard and Blake, 1996), the gene indicator in a DNA sequence analysis (Churchill, 1992) or the underlying volatility in an economical time series (Pitt and Shephard, 1997).

Equations (2.7) and (2.8) correspond to a recursive, generative form of the model, as opposed to our initial exposition, which focused on the specification of the joint probability distribution of the variables.

2.3 EXAMPLES

HMMs are used in many different areas. The bibliography by Cappe (2005) gives an idea of the rich literature in this domain, with different applications. Some applications are in areas such as some results on the data on fetal lamb movements (Leroux, 1989) which were also used in Leroux (1992) who found that the data set is best modeled by a HMM with two states with Poisson rates. Hughes and Guttorp (1994) presented a multivariate HMM for data consisting of daily binary rainfall observations (rain or no rain) at four different stations. Others include speech processing (Juang and Rabiner, 1991), modeling and predicting corporate default frequencies (Banachewicz et al., 2007), Biology (Felsenstein and Churchill, 1996).

HMM has been successfully applied to financial data. Engel and Hamilton (1990) modeled segmented time-trends in the US dollar exchange rates via HMMs. HMMs

reproduce most of the stylized facts about daily series of returns (Ryden et al., 1998) and accurately estimate stochastic volatility (Rossi and Gallo, 2006). The HMM has also been utilized in the simulation of the Nigeria Stock Exchange (Ahani et al., 2010).

It should be stressed that the idea one has about the nature of the HMM may be quite different from one case to another. These differences are illustrated in the examples below.

Example 2.3.1 - Finite HMM.

In this case, both the state space X_t of the HMM and the set Y_t on which the output lies are finite. The HMM is characterized by the transition probabilities

$$p(X_{t_k} = s_{t_k} \mid X_{t_{k-1}} = s_{t_{k-1}}) , \quad (2.9)$$

of the Markov chain, the conditional probabilities

$$p(Y_t = y_t \mid X_{t_{k-1}} = s_{t_{k-1}}) , \quad (2.10)$$

and the initial probability vector, probability of starting at a particular state:

$$\pi_i, i \in S, \pi_i = p(X_1 = s_i) .$$

Examples of this type of model are found in areas such as Capture-Recapture in the study of population with unknown sizes as in census undercount, survey, animal abundance evaluation and software debugging. Dupuis (1995) employed the model to lizard population. Here, the lizards move between three spatially connected zones denoted as 1, 2, and 3. The main focus is basically the modeling of these moves. The model has also been employed in genetics, signal processing, and Computer Sciences (see Koski, 2001) in which they can describe, amongst many other things, arbitrary finite-state machines.

Example 2.3.2 Normal Hidden Markov Model

A normal HMM is an HMM in which the conditional distribution of Y_k given X_k is Gaussian. In this case, given $X_k = i$, $Y_k \sim N(\mu_i, \sigma_i^2)$ the marginal distribution of Y_k is a finite mixture of normals. Example of normal HMM includes application to speech recognition, (Rabiner and Juang, 1993). The famous tutorial by Rabiner (1989) gives condensed description of the model.

Example 2.3.4 Gaussian Linear State-Space Model.

The standard state-space model takes the form

$$X_{t+1} = AX_t + BV_t, \quad (2.11)$$

$$Y_t = CX_t + DW_t, \quad (2.12)$$

where $V_t \sim N(0, I_{t_v})$, $W_t \sim N(0, I_{t_w})$. The initial condition X_0 is Gaussian with mean μ_w and covariances I_w and is uncorrelated with the processes $\{V_t\}$ and $\{W_t\}$. The state transition matrix A , the measurement transition matrix C , the square-root of the state noise covariance B , and the square-root of the measurement noise covariance D are known with appropriate dimensions. Ever since the pioneering work by Kalman and Bucy (1961), the study of the above model has been a favorite both in the engineering (automatic control, signal processing, target tracking) and time series literature, (Ristic et al., 2004). If the data are modeled by the Gaussian linear state-space model, it is possible to derive an exact analytical expression. This recursion is known as the Kalman filter. The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process. It is reliable and efficient, but it is limited to a relatively restricted class of linear Gaussian problems. To solve problems beyond this restricted class, particle filters are proving to be dependable methods for stochastic dynamic estimation.

Example 2.3.5 - General state space HMMs - Stochastic Volatility model.

The state space model (Doucet and Johansen, 2009) of a HMM is represented by the following two equations:

$$\text{(State equation)} \quad x_t = f(x_{t-1}) + w_t, \quad (2.13)$$

$$\text{(Observation equation)} \quad y_t = g(x_t) + v_t, \quad (2.14)$$

where f and g can be either linear or nonlinear functions, w_t and v_t are white noise processes. Models represented above are referred to as general state space models. These include such models as the Stochastic Volatility (SV) model.

This type of model and its generalizations have been used in various areas of economics and mathematical finance [Shumway and Stoffer (2000) and Chan (2002)].

Stochastic Volatility models belong to state-space form of a HMM; they take the volatility of the data into account. In a general SV model framework, the data, y , are generated from a probability model $f(y|x)$, where x is a vector of volatilities, and this unobserved vector x has a probabilistic structure $f(x|\theta)$ where θ is a vector of parameters (see Jacquier et al., 1994). This work focused on the SV problem formulated as a state space form of a HMM. The SV model as in Taylor (1982) is expressed thus:

$$x_t = \phi x_{t-1} + w_t \quad (2.15)$$

$$r_t = \beta \exp\left(\frac{x_t}{2}\right) v_t \quad (2.16)$$

Here, r_t is the return of the asset and x_t is the log-volatility at time t and follows an autoregressive (AR) process. Both w_t and v_t are independent normal error. The parameter β is the constant scaling factor, ϕ is the persistence in the volatility.

By taking logarithms of the squared returns,

$$r_t^2 = \left[\beta \exp\left(\frac{x_t}{2}\right) v_t \right]^2$$

$$r_t^2 = \beta^2 \left[\exp \frac{x_t}{2} \right]^2 v_t^2$$

$$\log (r_t^2) = \log \left[\beta^2 \exp\left(\frac{x_t}{2}\right) \right]^2 v_t^2$$

$$\log (r_t^2) = \log (\beta^2) + \log \exp(x_t) + \log (v_t^2) + E (\log (v_t^2)) - E (v_t^2)$$

one obtains,

$$y_t = \alpha + x_t + z_t$$

$$y_t = \log (r_t^2),$$

$$\text{where } \alpha = \log (\beta^2) + E (\log (v_t^2)),$$

$$z_t = \log (v_t^2) - E (v_t^2).$$

$$v_t \sim \log (\chi_1^2) E (\log (\chi_1^2))$$

$$\text{since } v_t^2 \sim \chi_1^2, \log (v_t^2) \sim \log \chi_1^2$$

$$E (\log (v_t^2)) \sim E (\log (\chi_1^2)) \quad \text{and} \quad E (\log (v_t^2)) = E (v_t^2)$$

2.4 PARAMETER ESTIMATION

SV model have long attracted researchers in finance and econometricians since the 1970's for various theoretical reasons. One use of the SV model is to explain the random behavior of financial markets. In the simplest framework, the series of returns, y_t , is modeled as the product of two stochastic processes. The early work on SV model to explain the random behavior of the market includes those of Clark (1973) and Tauchen and Pitts (1983). Clark used SV model to represent the random and uneven flow of new information to the financial market. Tauchen and Pitts (1983) refined this work and proposed a mixture of distributions of asset returns with temporal dependence in information arrival.

Some detailed reviews of ARSV models are given by Taylor (1994), Shephard (1996), Capobianco (1996) and Barndorff-Nielsen and Shephard (2001).

Carnero et al. (2001) showed that the ARSV (1) model is more adequate than the more popular GARCH (1,1) model to represent the empirical regularities often observed in financial time series. Helena (2004) proposed three extensions to the simple Log-AR(1) SV model. In the first extension the researcher extended the model fitting procedure to higher order AR(p) for the volatility to examine the fit of a higher order Log-AR(p) SV model to the data. The second extension to the model is to allow the correlation between the two error processes in the return and volatility process. The third extension combines the historical and implied volatility in one model to provide a unified approach to forecast volatility.

However, one of the most serious limitations of SV models is that the distribution of y_t conditional on past observations up to time $t-1$ is unknown. Consequently, the exact likelihood function is difficult to evaluate. In order to derive it, the vector of the unobserved volatilities has to be integrated out of the joint probability distribution. If we denote $y = \{y_1, \dots, y_t\}$ the vector of observations, $x = \{x_1, \dots, x_t\}$ the vector of log volatilities, and $\theta = \{\phi, \tau, \alpha\}$, the corresponding parameter vector, the likelihood is given by,

$$f(y|\theta) = \int f(y|x, \theta) f(x|\theta) dx \quad (2.17)$$

The dimension of the integral in (2.17) is equal to the sample size, T , and, in order to evaluate it, numerical methods must be used. Therefore, Maximum Likelihood (ML) estimation of the parameters of SV models is not straightforward.

The empirical application of Stochastic Volatility (SV) models has been limited due to the difficulties involved in the evaluation of the likelihood function. However, recently there has been fundamental progress in this area due to several proposed estimation methods that try to overcome this problem. As a consequence, several extensions of the SV models have been proposed and their empirical implementation is increasing. Recent developments in estimating non-linear latent variable models have made efficient estimation of SV models a reality. This makes extensive research on studying the empirical aspect of SV model feasible.

We review a number of methods proposed in the literature for fitting a SV model in the next section.

2.4.1 REVIEW OF METHODS OF PARAMETER ESTIMATION OF SV MODEL

Recent developments in statistical technology have made the estimation of nonlinear latent variable models possible with the increasing computing power. Among them are Generalized Method of Moments (GMM) and Quasi-Maximum-Likelihood Estimation (QML). Other computationally extensive procedures include simulated maximum likelihood (Daniellson, 1994), Efficient Method of Moment (Gallant et al., 1995).

Method of moments is a tool in econometrics. Applications of this method to the SV model include the work of Melino and Turnbull (1990) that used the GMM to estimate the SV model. The method uses model moments obtained either through simulations or analytically. Those estimators, however, can also be highly inefficient, depending on the choice of moment conditions. The main idea of the Method of Moments is to exploit the stationary properties of the process which yield the convergence of sample moments to their unconditional expectations. For a fully specified parameter model like SV model, the Generalized Method of Moments method is inefficient (Jacquier et al., (1994).

Furthermore, the GMM approach is an unattractive approach for the SV model as it does not deliver an estimate of the volatility process x_t , either filtered or smoothed. Since the main use of the model is to forecast volatility, another estimation procedure has to be developed. Jacquier et al., (1994) provided the accuracy of the GMM estimates and demonstrated that it is a less efficient estimation procedure for the SV model.

Harvey et al., (1994) proposed a Quasi-Maximum Likelihood (QML) approach to estimate the SV model which relies on a transformation of the model to a state-space form by taking the logarithm of the squares of the observation. By using a normal approximation to the $\log \chi_1^2$, they approximated the model to a Gaussian state-space model and employed Kalman filter and smoother technique to estimate the latent volatility process. This method is easy to perform and extend to more general model. It provides filtered and smoothed estimates of the latent volatility process but might not be efficient as the $\log \chi_1^2$ distribution is poorly approximated by the normal distribution. Jacquier et al., (1994) provided empirical evidence on the performance of QML and concluded that the procedure gave poor sampling properties.

Gallant et al., (1995) used the Efficient Method of Moments (EMM) to fit the standard SV model and various extensions to several daily financial time series. EMM is a simulation-based moment matching procedure. The moments that get matched are the scores of an auxiliary model called the score generator. If the score generator is able to approximate the probability distribution of return data reasonably well, then estimates of the parameters of the structural model are as efficient as maximum likelihood estimates. The EMM work has tried to maximize efficiency through the optimal choice of moment conditions, empirical implementation remains challenging

In the simulated maximum likelihood (SML) approach, Danielsson (1994) approximated the marginal likelihood of the observable process by simulating the latent volatility conditional on available information, using the importance-sampling simulation tool.

Shephard (1994) developed a partial non-Gaussian state-space framework under which one assumes that the state-space model is only conditionally Gaussian.

Jacquier et al., (1994) undertook a Bayesian approach based on the Markov chain Monte Carlo (MCMC) methodology. They were able to draw samples from the joint posterior distribution of the model parameters and the latent volatility realizations by constructing a simulation tool using a cyclic Metropolis algorithm.

Kim et al., (1998) demonstrated how to perform either maximum likelihood (ML) or Bayesian estimation for a linearized (log-squared transformed) SV model using a mixture of normals to approximate the error-term distribution. Bayesian analysis was carried out

using Markov chain Monte Carlo (MCMC) techniques, while a simulated EM (SEM) algorithm facilitated ML estimation. In both cases, Monte Carlo (MC) simulation is carried out through a multimove Gibbs sampler that enables one to sample from the density of the volatility process. The MC simulation based SML, partial non-Gaussian state-space using SEM or Bayesian analysis, and Bayesian MCMC procedures have the major disadvantage of being computationally demanding and much harder to implement, using nonconventional software that is not widely available among researchers and practitioners in the field. Successful application of the MC integration component in these techniques requires careful consideration of the simulation convergence and simulation error rates in the implementation. Furthermore, the Bayesian MCMC technique requires the choice of an appropriate conjugate prior distribution (Fridman and Harris, 1998).

Sandmann and Koopman (1998) used a Monte Carlo maximum likelihood (MCL) method of estimating stochastic volatility (SV) model. The basic SV model is expressed as a linear state space model with log chi-square disturbances. The likelihood function is approximated arbitrarily and accurately by decomposing it into a Gaussian part, constructed by the Kalman filter, and a remainder function, whose expectation is evaluated by simulation. However, due to the log chi-square disturbances in the measurement equation of the SV model, the Gaussian likelihood will only make up a part of the true likelihood function. The MCL estimator proposed here approximates the term via Monte Carlo simulation.

For Fridman and Harris (1998) a maximum likelihood approach for the analysis of stochastic volatility models is developed. They further explored this time series by fitting to it five variations of the simple SV model. The first variation is an SV in mean (SV-M) type model, where w_t and v_t are independent and Gaussian. The second model fitted is an extension that includes an autoregressive conditional heteroscedasticity (ARCH) term in the log-volatility equation. The third variation is one in which v_t is t-distributed. In the fourth variation, they let v_t follow a kurtotic distribution obtained by a signed power transformation of the normal distribution. Finally, the fifth model fitted is a combination of the last two models, where both w_t and v_t have kurtotic distributions. The method uses a recursive numerical integration procedure that directly calculates the marginal likelihood. In principle, both numerical and Monte Carlo integration deliver accurate approximations to the exact maximum likelihood estimator, but practical considerations have impeded their widespread use. In particular, the methods are computationally intensive and rely on assumptions that are hard to check in practice, such as the accuracy of numerical integrals and the convergence of simulated Markov chains to their steady state.

Durbin and Koopman (2000) used the idea of linearization of general state-space models and matched terms in the likelihood of a linearized model to those of a linear Gaussian model. As a result, the usual estimation skill for the linear Gaussian model can be applied to general (non-Gaussian and non-linear) state space models.

Shumway and Stoffer (2000) used an approximate likelihood in the normal mixtures setting. Their basic idea is to approximate the observation error, z_t , in equation (1.5) to two normal mixtures and then get the normal approximation of the conditional density of y_t given Y_{t-1} and I_t , where Y_{t-1} represents the previous observation and I_t is the indicator variable representing which of two normal's the observation, y_t , comes from.

Chib et al., (2002) applied the Metropolis-Hastings scheme for the analysis of the SV. Their main idea is to get a random sample from the posterior density of parameters given the data. In their approach, parameters are assumed to have some prior density, while in the classical analysis (non-Bayesian analysis), parameters are fixed and unknown, as in this study. Usually, their algorithms are iterative algorithms consisting of the following steps:

- (1). To sample from $p(\theta|Y_n, X_n)$
- (2). To sample from $p(X_n|\theta, X_n)$
- (3). By repeating above, a random sample $\{\theta, X_n\}$ from the joint posterior density $p(\theta, X_n|Y_n)$ can be obtained and this sample makes it possible to estimate parameters and to get filters and smoothers.

It is not obvious how to sample from the desired densities; so there are some techniques, which vary by author, needed to actually get random samples.

Andrieu et al., (2005) used the fixed-memory pseudo likelihoods, otherwise known as split likelihoods, to approximate the exact likelihood in SV model. This approach only requires fixed memory smoothing which is provably easier to achieve in the context of

Sequential Monte Carlo. Likewise the approach of Cappe, et al., (2005) used fixed-lag approximation ideas also inspired by an algorithm previously considered for HMMs (Krishnamurthy, 1993).

Cappe and Moulines (2009) used on-line expectation algorithm for latent data model. Here, they proposed a generic online (also sometimes called adaptive or recursive) version of the Expectation-Maximization (EM) algorithm applicable to latent variable models of observations. Compared to the algorithm of Titterton (1984), this approach is directly connected to the usual EM algorithm and does not rely on integration with respect to the complete data distribution. The resulting algorithm is usually simpler and is shown to achieve convergence to the stationary points between the marginal distribution of the observation and the model distribution at the optimal rate, i.e., that of the maximum likelihood estimator. The main advantages of the proposed approach to online parameter estimation in latent data models are its analogy with the standard batch EM algorithm, which makes the online algorithm easy to implement, and its optimal convergence behaviour.

Standard methods for maximum likelihood parameter estimation in latent variable models rely on the Expectation-Maximization algorithm and its Monte Carlo variants. Johansen et al., (2006) proposed an approach that reminds us of simulated annealing but also capitalizes on Expectation Maximization (EM) related ideas. That is, they built a sequence of artificial distributions whose support concentrates itself on the set of maximum likelihood estimates. They sampled from these distributions using a sequential Monte Carlo approach and demonstrated state-of-the-art performance for a non-linear

non-Gaussian state space model which is used in financial modeling. The intention of this example is to show how the algorithm can be applied in more complex settings. The results shown do not provide rigorous evidence that the algorithm is performing well..

Mongillo and Deneve (2008) proposed an online version of the EM algorithm. Their main tool is a recursion which allows for data recursive computation of smoothing functional required by the EM algorithm. However, this recursion appears to be very specific, that is, HMMs for which both the states and observations take a finite number of values. In Cappe and Moulines (2009), it is shown that this is indeed not the case and that this idea can be seen as a generalization of the online EM algorithm for mixture combined with a scheme for recursive implementation of smoothing for sum functionals of the state variables.

Cappe (2009) showed that the algorithm of Mongillo and Deneve (2008) can be extended to provide online EM estimation in more general settings, including non-linear non-Gaussian state-space models that necessitate the use of sequential Monte Carlo filtering approximations. The performance of the proposed online sequential Monte Carlo EM algorithm is illustrated using the simple noisy Gaussian AR(1) model observed in noise. The resulting algorithm is generic, and its performance in simulated scenarios is promising.

Recently, particle methods (Sequential Monte Carlo) have been applied to the SV model. As particle filters are designed to get the samples of hidden states given parameters and data, it is necessary to either adopt other methods for the parameter estimation, or extend

the particle filtering method to embrace the parameters as a part of the hidden states, in order to solve the estimation problem. Kim and Stoffer (2008) incorporated the EM algorithm and SMC (particle filters and smoothers). Their attempt forms a basic idea to handle the parameter estimation problem in the SV model. The SMCEM estimation procedure offers a unified simulation approach which works on actual likelihood distribution rather than asymptotic approximations for parameter estimation and smoothing.

In order to expand the scope of application of SV models, this research extends the SMC techniques with EM algorithm of Kim and Stoffer (2008), to estimate the parameters of SV model with student-t distribution and the GED.

2.4.2 REVIEW OF MULTIVARIATE STOCHASTIC VOLATILITY MODEL

We review a number of methods proposed in the literature for fitting a univariate SV model. Although univariate SV models have undergone a lot of research, the literature on multivariate models is less abundant with the works of Harvey et al. (1994), Danielsson (1998), Aguilar and West (2000), Lopes and Migon (2003) and Chib et al., (2006) forming the basis for the developments we consider here. They modeled the levels of a set of financial time-series by a standard normal factor model in which both the common factor variances and the specific (or idiosyncratic) time-series variances are modeled as univariate SV processes.. An ultimate stylized fact which cannot be captured by univariate descriptions is the covariation effect. That is what necessitates the extension of

the univariate model to the multivariate case is the covariation effect. Volatilities of different financial series are often seen to move together and large changes in one asset are matched by large movements in another. This is the result of either a formal linkage between a set of assets, e.g. currency exchange rates or due to the assets being subject to the same overall environment and as a result having factors in common which influence their behavior.

As Aydemir (1998) noted, we often observe related movements between markets, sectors, stocks, or exchange rates. This phenomenon is usually due to their being influenced by common unobserved factors. Nevertheless, there are theoretical as well as empirical reasons to study multivariate volatility models. In the first place, portfolio allocation and asset pricing can only be meaningfully discussed within a multivariate framework. Secondly, correlation across asset returns requires simultaneous multivariate estimation for full efficiency. Thirdly, multivariate structural volatility models can provide useful information about the factors driving the volatility process.

Apart from the well-established constant conditional correlation (see Harvey et. al., (1994)) and multivariate factor SV models (see Pitt and Shephard (1997)), other possible alternatives seem suitable to describe the time evolution of the joint distribution of different assets. However, a thorough understanding of the estimation procedures developed for the univariate SV model is imperative before we can extend the ideas to a multi-dimensional equivalent.

In the multivariate context, when one is dealing with a collection of financial time series denoted by $y_t = (y_1, \dots, y_t)$, the major objective is to model the time-varying conditional covariance matrix of y_t and this can be done in several ways within the SV context (Asai et al., 2006).

In the context of stochastic volatility models, Harvey et al., (1994) propose the following multivariate model which allow the variances and covariances to evolve through time with possibly common trends

$$\begin{aligned} x_t &= \phi_i x_{t-1} + w_t^i \\ y_t^i &= \beta_i \exp\left(\frac{x_t}{2}\right) v_t^i \quad i=1, 2, \dots, N; \quad t=1, 2, \dots, T, \end{aligned} \quad (2.18)$$

where $w_t \sim N(0, \tau)$, $v_t^i = (v_{1,t}, \dots, v_{n,t})'$ has a multivariate normal distribution with covariance matrix \sum_v , and $w_t^i = (w_{1,t}, \dots, w_{n,t})'$ is distributed independently from v_t following a multivariate normal distribution with covariance matrix \sum_w . Harvey et al., (1994) estimate the parameters by the QML method. The main limitation of this model is that it restricts the correlations to be constant over time. Danielsson (1998) extends this model to allow for leverage effects and time-varying correlations. Later, Ray and Tsay (2000) used the same model to study common long memory components in daily stock volatilities of groups of companies. The parameters of this model are estimated by the simulated maximum likelihood (SML) method.

Alternatively, Jacquier et al., (1999) and Shephard (1996) propose a factor model for returns, where the factors are SV processes.

$$R_t = D_{x_t} + e_t \quad (2.19)$$

$$y_t = \beta \exp\left(\frac{x_t}{2}\right) v_t, v_t \sim N(0,1) \quad (2.20)$$

$$x_t = \phi x_{t-1} + w_t, w_t \sim N(0,1) \quad (2.21)$$

According to this model a set of asset returns are driven by latent factors which are specified as SV processes. Such a multivariate factor SV specification, originally proposed by Shephard (1996) and Jacquier et al., (1999), has several attractive features. Foremost, the dimension of the parameter space remains operational as it only increases linearly with the number of assets being modeled. Moreover, this multivariate specification accounts, not only for the volatility dynamics of the individual assets but also, due to the common factor, for time varying correlations across assets returns which require simultaneous multivariate estimation for full efficiency.

Chib et al., (2006) generalized model (2.19) allowing series-specific jumps at each time and a fat tailed t-distribution for the errors, e_t , and Lopes and Migon (2003) allowed for factor loadings to evolve over time. Recently, Tims and Mahieu (2006) have proposed to fit a multivariate SV model to the logarithmic range. The logarithmic range is a proxy for volatility that is approximately Gaussian, allowing for the application of standard Kalman filter techniques to estimate the corresponding models.

Kim, et al., (1998) put forward the basic model structure we suggest in this paper. They allow them to follow independent SV processes - although this model was not fitted in practice. In a recent paper, Aguilar and West (2000) have implemented this model using the Kim, et al., (1998) mixture MCMC approach. Estimation of parameters in MSV models is difficult as there is a sequence of latent variables and parameters needed to be

estimated. Thus, estimation can be based on approximations (quasi maximum likelihood), numerical methods for evaluating the likelihood (numerical integration) or simulation methods.

2.4.3 TARGET TRACKING ESTIMATION - REVIEW OF METHODS IN BEARINGS-ONLY TARGET

The problem of bearings-only target estimation and tracking has a history which dates back to the 1960's and has attracted a variety of important practical applications, such as submarine tracking (using passive sonar) or aircraft surveillance (using a radar in a passive mode or an electronic warfare device) . The problem is sometimes referred to as target motion analysis (TMA), and its objective is to track the kinematics (position and velocity) of a moving target using noise-corrupted bearing measurements. The observation platform needs to manoeuvre in order to estimate the target range Nardone, et al., (1984).

The problem of target tracking using bearings-only measurements is a difficult task. The filtering algorithms involve a nonlinear measurement process, which, when linearized, can lead to time-varying parameters, biases as explained by Aidala, (1979). The common estimation algorithms used for bearings-only target tracking are: Least Squares (batch and recursive forms), Maximum Likelihood Estimator, Extended Kalman Filter (EKF), and Particle Filters or Bayesian Methods.

Most researchers in the field of bearings-only tracking have concentrated on tracking a nonmanoeuvring target. Due to inherent nonlinearity and observability issues, it is difficult to construct a finite-dimensional optimal filter even for this relatively simple

problem. As for the bearings-only tracking of a manoeuvring target, the problem is much more difficult.

Early research focuses mainly on analytical derivations for the observability criteria of the estimation process, and comparisons of the convergence properties and performance of the different types of method used for target tracking. Since bearings-only target estimation involves a non-linear measurement process, several filtering and observability complications arise. Lindgren and Gong (1978) analyze the observability associated with a least-squares estimation approach and show that, for a constant velocity target and a constant velocity vehicle moving in a 2-D plane, the target estimation is unobservable until the vehicle executes a maneuver (change in heading). Kalman Filtering techniques are used by Aidala, (1979) and Nardone, et al., (1984). Since the bearings-only estimation problem involves nonlinear measurements, an Extended Kalmer filter (EKF) approach needs to be used instead of the normal Kalman Filter. The traditional EKF however is sensitive to initialization techniques and measurement errors which can cause early covariance collapse and other filter instabilities Aidala, (1979). The moving vehicle trajectory affects the observability and convergence of the target estimation, suggesting that a good trajectory design can reduce filter instability and estimation errors.

A pseudolinear filter formulation is proposed by Aidala and Nardone (1982), which attempts to linearize the dynamics and measurement models. However by linearizing the dynamics the noise becomes non-Gaussian which, when propagated through the filter, causes estimation bias. For the bearings-only tracking problem, the bias is introduced

only in the position estimate, and is highly dependent on the geometry of the vehicle maneuvers, once again suggesting that the estimation performance can be improved by proper design of the vehicle trajectory. Comparisons of the properties and performance between several different filtering algorithms are explored by Nardone, et al.,(1984). Aidala and Hammel (1983) proposed the modified polar coordinates filter. The filter uses an EKF algorithm with a state vector choice, based on polar coordinates, that attempts to separate the observable and unobservable components of the estimated state by using a different coordinate system. The resulting filter is stable and asymptotically unbiased. The modified polar coordinate filter shows the dependence of the target estimation on the vehicle maneuvers, once again suggesting that the estimation can be improved by designing a good trajectory.

De Vlieger (1992) used a piecewise linear model of the target motion and a Maximum Likelihood Estimator approach for target tracking. He uses numerical methods to condition the measurement model to increase the observability of the estimation. Goshen-Meskin and Bar-Itzhack, (1992) derive the observability requirements for piecewise

constant linear systems. Tao et al., (1996) shows that for a MLE approach it is important to consider the correlation of the noise, and that ignoring it degrades the performance of the estimation.

Several modifications to the classical estimation algorithms have been explored. Some attempt to smooth the trajectory, within the constraints of a known target behavior model. Others consist of designing multiple filters for different known target scenario and using statistical properties of the innovation to switch between the algorithms. Another approach has been to support multiple Kalman filters simultaneously and develop an estimate by combining all the filters. Later research by Bar-Shalom et al. (2002) has focused on using interacting multiple models (IMM). These algorithms employ a constant velocity (CV) model along with manoeuvre models to capture the dynamic behaviour of a manoeuvring target scenario. Le Cadre and Tremois, (1998) modelled the manoeuvring target using the CV model with Gaussian noise and developed a tracking filter in the hidden Markov model framework.

Particle filtering or Sequential Monte Carlo techniques have also been explored by Liu et al., (2002), Bar-Shalom et al., (2001), Ristic et al., (2004). Particle filters have the advantage of being able to deal with nonlinear systems and non-Gaussian noise models making them particularly well suited to bearings-only tracking. They can also accommodate unknown and stochastic target models making them more versatile than classical filters. However, they require increased computational resources and, for fast convergence, need a fairly accurate description of the measurement likelihood function and a good initial distribution on the estimated target location.

2.5 MAXIMUM LIKELIHOOD AND THE EXPECTATION-MAXIMIZATION ALGORITHM

When we have a density function $p(X|\theta)$ that depends on a set of parameters θ and a data set $X = \{x_1, \dots, x_n\}$ that are independent and identically distributed, the likelihood is

$$p(X|\theta) = \prod_{i=1}^n p(x_i|\theta) = L(\theta|X). \quad (2.22)$$

The resulting $L(\theta|X)$ is the likelihood function that we will maximize. We usually maximize $l(\theta) = \log(L(\theta|X))$ because it is easier. However, there are many cases where $l(\theta)$ has no analytic solution and so we resort to the Expectation-Maximization (EM) algorithm as one possible solution. The EM algorithm is a parameter estimation tool for finding the maximum likelihood estimates of parameters in probabilistic model, where the model depends on some unobserved variables.

The method in this work is basically based on the EM algorithm. When the EM algorithm is applied to the SV model, it is not possible to get the exact expected likelihood due to the complex dependent structure of the SV model. In order to solve this problem, we bring out particle methods and calculate an approximate expected likelihood by using the output of the particle methods.

In practice, the transition and observation distributions of state-space models usually depend on an unknown parameter vector θ which needs to be estimated. A very useful approach is to apply the Expectation-Maximization (EM) algorithm. A considerable collection of literature exists on the EM algorithm and its application (See Dempster et al., 1977). The EM algorithm is an iterative algorithm for obtaining Maximum likelihood

parameter estimates. Generally its application has been seen much in the cases where the data set is considered to be incomplete in the sense that it is not fully observable. The EM algorithm as proposed by Dempster et al., (1977) consists in iteratively building a sequence $\{\theta^k\}_{k>1}$ of parameter estimates given an initial guess θ^0 . Each iteration is broken into two steps. Below is a brief description of the EM algorithm applied to the data:

Let X and Y be the two sample spaces where there is a many-one mapping, $y^* : x \rightarrow y^*(x) = y \in Y$ from X to Y . y is a realization in Y . The corresponding x in X is not observed and is termed the complete data. If there is a family of sampling densities $f(x|\theta)$ depending on parameters θ , then the density of observed data Y is

$$g(y|\theta) = \int_{X(y)} f(x|\theta) \quad (2.23)$$

$X(y)$ is the pre-image of y in X . In this framework, the problem that calls for a solution is to get MLE of θ by maximizing $l(\theta) = \log(y|\theta)$. In several instances $l(\theta)$ has no analytic solution. The EM algorithm is an iterative algorithm to find MLE by using $f(x|\theta)$ instead of $g(y|\theta)$. One iteration of the EM algorithm is formed by the combination of Expectation step (E-step) and Maximization step (M-step). At k^{th} iteration, the updated parameter $\theta^{(k)}$ is obtained from $\theta^{(k-1)}$ as follows: In the E-step, the conditional expectation:

E-step: the expected likelihood, $Q(\theta|\theta^{(k)})$

$$Q(\theta|\theta^{(k)}) = E_{\theta^{(k)}} (\log f(x|\theta') | y, \theta), \quad (2.24)$$

is computed, where $\theta^{(k)}$ is the current parameter estimate.

M-step: In the M-step, a new parameter estimate $\theta^{(k+1)}$ is obtained by maximizing $Q(\theta | \theta^{(k)})$.

The E-step and M-step are repeated until some stopping criterion is met, such as $|\hat{\theta}^{n+1} - \hat{\theta}^n| < Q$, for some specified Q , obtaining of suitable initial parameters inclusive.

2.5.1 MONTE CARLO EXPECTATION-MAXIMIZATION ALGORITHM

In some applications of the EM algorithm, the E-step is difficult to perform; and is sometimes impossible. Wei and Tanner (1991) introduced the Monte Carlo Expectation-Maximization (MCEM) algorithm where the E-step is executed by Monte Carlo methods. The MCEM algorithm returns the expectation of M realizations of the likelihood by random sampling in its E-step. Chan and Ledolter (1995) applied the MCEM algorithm to a time series model with count data, where the E-step is intractable even by numerical integration. The following is a brief summary of the MCEM algorithm.

Monte Carlo expectation step (Monte Carlo E-step) and M-step form one iteration of the MCEM algorithm. At the k^{th} iteration, the updated parameter, $\theta^{(k)}$ is obtained from $\theta^{(k-1)}$ as follows:

Monte Carlo E-step: Compute $Q^*(\theta | \theta^{(k-1)})$, where

$$Q^*(\theta | \theta^{(k-1)}) = \frac{1}{M} \sum_{j=1}^M \log f(x_j^*, y | \theta) \quad \text{and} \quad \{x_j^*\}_{j=1}^M \text{ is random sample from } f(x | y, \theta^{(k-1)})$$

M-step: Choose $\theta^{(k)}$ which maximizes $Q^*(\theta | \theta^{(k-1)})$.

In MCEM algorithms, it is not as obvious how to get samples from desired distributions, and is sometimes difficult. To obtain an approximation of the conditional expectations, particles smoothing distributions are used. The method in this work is similar to the MCEM algorithm through the use of the particle smoother and filter method to execute the E-step while the MCEM algorithm uses the Monte Carlo method. The particle filter method takes advantage of the sequential structure of the state-space models, thus it is easier to produce samples. The clue is that it is often possible to carry out the M-step analytically. If this is the case, a closed-form expression for the estimator $\theta^{(k+1)}$ is obtained which depends on the smoothing particles which are generated with respect to the old parameter estimate $\theta^{(k)}$.

2.6 SEQUENTIAL MONTE CARLO: PARTICLE FILTERS AND SMOOTHERS

Since the introduction of SMC in 1993 (Gordon et al., 1993), it has become a well known class of numerical methods for the solution of optimal estimation problems in non-linear non-Gaussian state-space models. Chen and Lai (2003) applied SMC methodology to tackle the problems of optimal filtering and smoothing in HMM. SMC has also stirred great interest in the engineering and statistical literature (see Doucet et al., (2001) for a summary of the state of the art). Lately, by Johansen et al., (2008), SMC has been applied for resolving a marginal Maximum Likelihood problem. In Gordon et al., (1993), the application of SMC to optimal filtering was first offered. Here, SMC methods is developed for prediction of state by estimating the probability $p(x_t | y_{1:t-1})$. Zhang et al., (2007), applied SMC method for predicting actual data of monthly national air

passengers in America. Ahani and Abass (2010) applied a SMC approach for stock market prediction.

The SMC approach shows a great promise in providing such a fundamental technology. The interest has arisen with the proposal of the bootstrap filter (Gordon et al., 1993), simultaneously developed by (Kitagawa, 1993). Since then, Sequential Monte Carlo (SMC) algorithms, under the names of particle filters, sequential importance resampling (SIR) has been applied to a wide range of problems in the fields of engineering, financial data analyses, genetics, medicine, biology, to name a few (see Doucet et al., 2001).

SMC methods are an efficient class of simulation techniques to approximate sequences of complex probability distributions. These probability distributions are approximated by a large number of random samples which are propagated over time using a combination of importance sampling and resampling steps. SMC methods allow us to carry out on-line approximation of probability distributions using samples (particles). They are very useful in scenarios involving real-time signal processing, where data arrival is inherently sequential. Furthermore, one might wish to adopt a sequential processing strategy to deal with non-stationarity in signals, so that information from the recent past is given greater weighting than information from the distant past.

The key idea is to represent the concerned density function $p(x_{0:t-1} | y_{0:t-1})$ at time $t-1$ by a set of random samples with associated weights, $\{x_{0:t-1}^{(i)}, w_{0:t-1}^{(i)} | i = 1, \dots, N\}$ and compute estimates based on these samples and associated weights. As the number of samples becomes very large, this Monte Carlo characterization develops into an

equivalent representation to the functional description of the density function (Zhang et al., 2007).

If we let $\{x_{0:t-1}^{(i)}, w_{0:t-1}^{(i)} \mid i=1, \dots, N\}$ be respectively samples and associated weights approximating the density function $p(x_{0:t-1} \mid y_{0:t-1})$, $\{x_{0:t-1}^{(i)}\}_{i=1}^N$ is a set of particles with associated weights $\{w_{0:t-1}^{(i)}\}_{i=1}^N$ with $\sum_{i=1}^N w_{t-1}^{(i)} = 1$, then the density function are approximated by

$$p(x_{0:t-1} \mid y_{0:t-1}) \approx \sum_{i=1}^N w_{t-1}^{(i)} \delta(x_{t-1} - x_{t-1}^{(i)}) \quad (2.25)$$

$\delta(x)$ signifies the Dirac delta role. The particle approximation $\{w_t^{(i)}, x_t^{(i)}\}_{i=1}^N$ are transformed into an equally weighted random sample from $p(x_{0:t-1} \mid y_{0:t-1})$ by sampling, with replacement, from the distribution. This procedure, otherwise called resampling, produces a new sample with uniformly distributed weights.

Particle filters and smoothers are SMC methods grounded in particle representations, and are considered as generalizations of well-known Kalman filters and smoothers for general state-space models.

The fundamental approach used to get particles from the desired density is based on sequential importance sampling (SIS) and resampling. SIS, a Monte Carlo method, forms the basis for most particle filtering methods. To approximate the conditional density of x_t given the previous states, x_{t-1} , and the past and present data, y_t , $p(x_t \mid x_{t-1}, y_t)$, SIS introduces a importance sampling density, $\pi(x_t \mid x_{t-1}, y_t)$ where it is easier to sample from the distribution, $p(x_t \mid x_{t-1}, y_t)$.

2.6.1 SIMPLE MONTE CARLO SAMPLING

Assume we are able to simulate N independent and identically distributed random samples $\{X_{0:t}^{(i)}, i=1, \dots, N\}$ according to $p(x_{0:t} \mid y_{1:t})$. An empirical estimate of this distribution is given by

$$P_N(dx_{0:t} \mid y_{0:t}) = \frac{1}{N} \sum_{i=1}^N \delta_{x_{0:t}^{(i)}}(dx_{0:t}) \quad (2.26)$$

or the expectation of a function

$$E(f_t(x_{0:t})) = \int f_t(x_{0:t}) P_N(dx_{0:t} \mid y_{1:t}) = \frac{1}{N} \sum_{i=1}^N f_t(x_{0:t}^{(i)}) . \quad (2.27)$$

This estimate given in equation (2.27) is unbiased with the variance proportional to $\frac{1}{N}$

for the finite variance of $f_t(x_{0:t})$ (Doucet et al., 2001).

It is seldom possible to sample efficiently from the distribution $p(x_{0:t} \mid y_{1:t})$ at any time t .

2.6.2 IMPORTANCE SAMPLING

An alternative solution to estimate $p(x_{0:t} \mid y_{1:t})$ and $E(f_t(x_{0:t}))$ consists of using importance sampling (IS) method (Geweke, 1989). This is a fundamental Monte Carlo method and the basis of all the algorithms developed later on. IS relies on the introduction of an importance sampling distribution $\pi(x_{0:t} \mid y_{1:t})$, from which one can easily sample and such that $p(x_{0:t} \mid y_{1:t}) > 0$ implies $\pi(x_{0:t} \mid y_{1:t}) > 0$, then

$$E(f_t(x_{0:t})) = \frac{\int f_t(x_{0:t}) w(x_{0:t}) \pi(x_{0:t} \mid y_{1:t}) dx_{0:t}}{\int w(x_{0:t}) \pi(x_{0:t} \mid y_{1:t}) dx_{0:t}} . \quad (2.28)$$

where $w(x_{0:t}) = \frac{p(x_{0:t} \mid y_{1:t})}{\pi(x_{0:t} \mid y_{1:t})}$ is the importance weight.

If we have *N i.i.d.* random samples $\{X_{0:t}^{(i)}, i=1, \dots, N\}$ distributed according to $\pi(x_{0:t} \mid y_{1:t})$ then a Monte Carlo estimate of $E(f_t(x_{0:t}))$ is

$$E(f_t(x_{0:t})) = \frac{\frac{1}{N} \sum_{i=1}^N f_t(X_{0:t}^{(i)}) w(X_{0:t}^{(i)})}{\frac{1}{N} \sum_{j=1}^N w(X_{0:t}^{(j)})} = \sum_{i=1}^N f_t(X_{0:t}^{(i)}) \tilde{w}_t^{(i)}, \quad (2.29)$$

where the normalized importance weights $\tilde{w}_t^{(i)}$ are given by

$$\tilde{w}_t^{(i)} = \frac{w(X_{0:t}^{(i)})}{\sum_{j=1}^N w(X_{0:t}^{(j)})}. \quad (2.30)$$

The importance sampling is a general Monte Carlo integration method. However, in its simplest form, it is not adequate for recursive estimation.

2.6.3 SEQUENTIAL IMPORTANCE SAMPLING

The importance sampling method can be modified so that it becomes possible to obtain at any time t an estimate $\hat{P}_N(dx_{0:t} \mid y_{1:t})$ of the distribution $p(x_{0:t} \mid y_{1:t})$ without modifying subsequently the past simulated trajectories $\{X_{0:t-1}^{(i)}, i=1, \dots, N\}$. This means that the importance function $\pi(x_{0:t} \mid y_{1:t})$ at time t admits as marginal distribution at time $t-1$ the importance function $\pi(x_{0:t-1} \mid y_{1:t-1})$ that is,

$$\pi(x_{0:t} \mid y_{1:t}) = \pi(x_{0:t-1} \mid y_{1:t-1}) \pi(x_t \mid x_{0:t-1}, y_{1:t}). \quad (2.31)$$

Iterating, one obtains

$$\pi(x_{0:t} \mid y_{1:t}) = \pi(x_0) \prod_{k=1}^t \pi(x_k \mid x_{0:k-1}, y_{1:k}), \quad (2.32)$$

where $\pi(x_k \mid x_{0:k-1}, y_{1:k})$ is the probability density function of X_k conditional upon $x_{0:k-1}$ and $y_{1:k}$.

This importance function allows us to evaluate recursively in time the importance weights (2.30). Thus, one has

$$\tilde{w}_t^{(i)} \propto \tilde{w}_{t-1}^{(i)} \frac{p(Y_t \mid X_t^{(i)})p(X_t^{(i)} \mid X_{t-1}^{(i)})}{\pi(X_t^{(i)} \mid X_{0:t-1}^{(i)}, Y_{1:t})}. \quad (2.33)$$

2.6.4 SEQUENTIAL IMPORTANCE RESAMPLING

The problem with sequential importance resampling (SIS) is that the variance of the importance weights $\tilde{w}_t^{(i)}$ increases stochastically over time [Kong et al., (1994) and Doucet et al., (2001)]. In practice, after a few iterations of the algorithm, only one particle has a non-zero importance weight. The algorithm fails to represent the distributions of interest. To avoid the degeneracy of the SIS simulation method, a resampling stage (which is referred to as bootstrap filter) are used to eliminate samples with low importance weights $\tilde{w}_t^{(i)}$ and multiply samples with high importance weights (Doucet, et al. 2001). Many of the ideas on resampling have stemmed from the work of Efron (1982), Rubin (1988), and Smith and Gelfand (1992).

Formally, the weighted empirical distribution

$$\hat{P}_N(dx_{0:t} \mid y_{0:t}) = \sum_{i=1}^N \tilde{w}_t^{(i)} \delta_{x_{0:t}^{(i)}}(dx_{0:t}). \quad (2.34)$$

is replaced by the unweighted measure

$$\hat{P}_N(dx_{0:t} \mid y_{0:t}) = \frac{1}{N} \sum_{i=1}^N N_t^{(i)} \delta_{x_{0:t}^{(i)}}(dx_{0:t}), \quad (2.35)$$

where $N_t^{(i)}$ is the number of offspring associated to particle $X_{0:t}^{(i)}$, it is an integer number

such that $\sum_{i=1}^N N_t^{(i)} = N$. If $N_t^{(j)} = 0$, then the particle $X_{0:t}^{(j)}$ dies. The $N_t^{(i)}$ are chosen such

that $P_N(dx_{0:t} | y_{1:t})$ is close to $\hat{P}_N(dx_{0:t} | y_{1:t})$ in the sense that, for any function f_t ,

$$\int f_t(x_{0:t}) P_N(dx_{0:t} | y_{1:t}) \approx \int f_t(x_{0:t}) \hat{P}_N(dx_{0:t} | y_{1:t}) . \quad (2.36)$$

After the selection step, the surviving particles $X_{0:t}^{(i)}$, that is the ones with $N_t^{(i)} > 0$, are thus approximating distributed according to $p(x_{0:t} | y_{1:t})$.

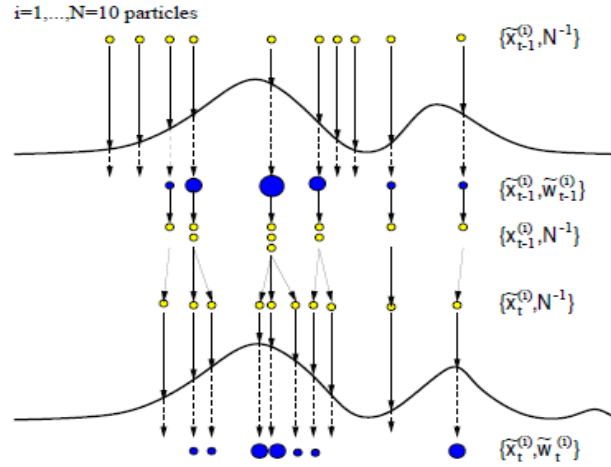


Figure 2.3: SIS particle filter.

Figures 2.3 - 2.5 are adopted from Maskell and Gordon (2002): In this example, a particle filter starts at time $(t-1)$ with an unweighted measure $\{\tilde{X}_{t-1}^{(i)}, N^{-1}\}$, which provides an approximation of $p(x_{t-1} | y_{1:t-2})$. For each particle we compute the importance weights using the information at time $(t-1)$. This results in the weighted measure $\{\tilde{X}_{t-1}^{(i)}, \tilde{w}_{t-1}^{(i)}\}$, which yields an approximation $p(x_{t-1} | y_{1:t-1})$. Subsequently, a resampling step selects only the fittest particles to obtain the unweighted measure $\{\tilde{X}_{t-1}^{(i)}, N^{-1}\}$, which is still an

approximation of $p(x_{t-1} | y_{1:t-1})$. Finally, the sampling step introduces variety resulting in the measure $\{\tilde{X}_{t-1}^{(i)}, N^{-1}\}$.

2.6.5 Algorithm description

We specify the algorithm as follows:

Bootstrap Filter

1) *Initialisation*, $t = 0$.

* For $i = 1, \dots, N$ sample $X_{0:t}^{(i)} \sim p(x_0)$ and set $t = 1$.

2). *Importance Sampling Step*

* For $i = 1, \dots, N$ sample $\tilde{X}_t^{(i)} \sim p(X_t | X_{t-1}^{(i)})$ and set $\tilde{X}_{0:t}^{(i)} = (X_{0:t-1}^{(i)}, \tilde{X}_t^{(i)})$

3). *Sequential Importance Sampling Step*

* For $i = 1, \dots, N$ sample $X_t^{(i)} \sim \pi(X_t | x_{0:t-1}^{(i)}, y_{1:t})$ and set $\tilde{X}_{0:t}^{(i)} = (X_{0:t-1}^{(i)}, \tilde{X}_t^{(i)})$

* For $i = 1, \dots, N$ evaluate the importance weights up to a normalizing constant:

$$\tilde{w}_t^{(i)} \propto \tilde{w}_{t-1}^{(i)} \frac{p(Y_t | X_t^{(i)}) p(X_t^{(i)} | X_{t-1}^{(i)})}{\pi(X_t^{(i)} | X_{0:t-1}^{(i)}, Y_{1:t})}.$$

* For $i = 1, \dots, N$ normalize the importance weights:

$$\tilde{w}_t^{(i)} = \frac{w(X_{0:t}^{(i)})}{\sum_{j=1}^N w(X_{0:t}^{(j)})}.$$

4). *Selection Step** Multiply / eliminate particles $\{\tilde{X}_{0:t}^{(i)}\}_{i=1}^N$ with respect to high/low normalized importance weights $\tilde{w}_t^{(i)}$ to obtain N particles $\{\tilde{X}_{0:t}^{(i)}\}_{i=1}^N$ with $\tilde{w}_t^{(i)} = \frac{1}{N}$

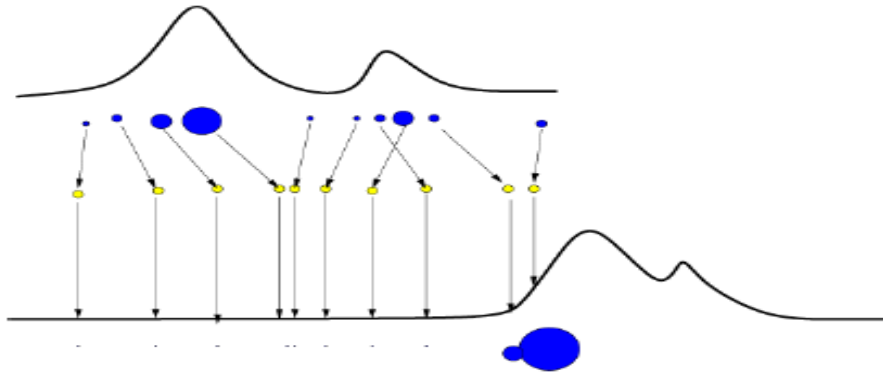


Fig 2.4: problem of the SIS particle filter: degeneracy.

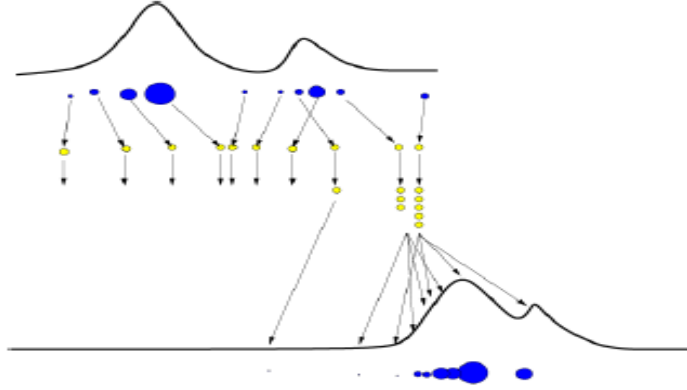


Fig 2.5: Resampling

In Figure 2.4, two particles dominate the distribution and the others are of no use since their weights are almost zero. Resampling is designed to solve this problem by removing particles with small weights, concentrating, instead, on particles with large weights. This resampling step involves generating a new set, $\{x_t^{(i*)}\}_{i=1}^M$, by resampling with replacement M times from $\{x_t^{(i)}\}_{i=1}^M$ so that $p(x_t^{(i*)} = x_t^{(i)}) = w_t^{(i)}$. Figure 2.5 shows how resampling works. A generic particle filter draws $\{x_t^{(i)}\}_{i=1}^M$ using a SIS particle filter, and resamples $\{x_t^{(i*)}\}_{i=1}^M$ when degeneracy has occurred. For more details, see Zhang et al., (2007) or Doucet et al., (2001).

2.7 PARTICLE FILTER ALGORITHM

Suppose that we have at time t weighted particles $\{f_t^{(i)}, w_t^{(i)}\}$ drawn from $f(x_t | y_t)$, $f_t^{(i)}$ is a set of particle filter with associated weight $w_t^{(i)}$. This is considered as an empirical approximation for the density,

$$f(x_t | y_t) \approx \sum_{i=1}^M w_t^{(i)} \delta(x_t - f_t^{(i)}). \quad (2.37)$$

Kitagawa and Sato (2001) and Kitagawa (1996) gave an algorithm for filtering in general state space model thus:

2.7.1 Monte Carlo filtering for general state-space models

- 1) For $i=1, \dots, N$, generate a random number $f_0^{(i)} \sim p(x_0)$.
- 2) Repeat the following steps for $t=1, \dots, T$.
 - a. For $i=1, \dots, N$, generate a random number $w_t^{(i)} \sim q(w)$.
 - b. For $i=1, \dots, N$, Compute $p_t^{(i)} = F(f_{t-1}^{(i)}, w_t^{(i)})$.
 - c. For $i=1, \dots, N$, Compute $w_t^{(i)} = p(y_t | p_t^{(i)})$.
 - d. Generate $f_t^{(i)}, i=1, \dots, N$ by resampling $p_t^{(i)}, \dots, p_t^{(N)}$.
- 3) This Monte Carlo filter returns

$$\{f_t^{(i)}, i=1, \dots, N, t=1, \dots, m\} \text{ so that } \sum_{i=1}^N \frac{1}{N} \delta(x_t - f_t^{(i)}) \approx f(x_t | Y_t).$$

2.8 PARTICLE SMOOTHING ALGORITHM

If we let $\{s_t^{(i)}, w_t^{(i)}\}_{i=1}^M$ be set of particle smoothers and associated weights approximating the density function $f(x_t | Y_n)$, then the density function are approximated by

$$f(x_t | Y_n) \approx \sum_{j=1}^M w_t^{(j)} \delta(x_t - s_t^{(j)}). \quad (2.38)$$

The problem with smoothed estimates is degeneracy. Godsill et al., (2004) suggested a new smoothing method (particle smoother using backwards simulation). The method assumes that the filtering has already been performed. Thus, the particles and associated weights, $\{f_t^{(i)}\}_{i=1}^M$, $\{w_t^{(i)}\}_{i=1}^M$ can approximate the filtering density, $f(x_t | Y_t)$, by

$$= \frac{\sum w_t^{(i)} \delta(x_t - f_t^{(i)})}{\sum_{i=1}^N w_t^{(i)}}. \text{ Godsill et al., (2004) gave the following algorithm :}$$

2.8.1 Particle smoother using backwards simulation

Suppose weighted particles $\{f_t^{(i)}, w_t^{(i)}; i = 1, 2, \dots, M\}$ are available for $t = 1, 2, \dots, n$. For $i = 1, 2, \dots, M$,

- 1) Choose $s_n^{(i)} = f_n^{(j)}$ with probability $w_n^{(j)}$
- 2) For $n-1$ to 1
 - a. Calculate $w_{t|t+1}^{(j)} \propto w_t^{(j)} f(s_{t+1}^{(i)} | f_t^{(j)})$ for each j .
 - b. Choose $s_t^{(i)} = f_t^{(j)}$ with probability $w_{t|t+1}^{(j)}$.
- 3) $s_{1:n}^{(i)} = (s_1^{(i)}, \dots, s_n^{(i)})$ is an approximate realization from $p(X_n | Y_n)$.

CHAPTER THREE

3. THEORETICAL FRAMEWORK AND METHODOLOGY

This work is cardinally hinged in its analysis and extension, on the essential findings of Kim and Stoffer (2008) and Cappe (2009). As online estimation of model parameters of the state space model of a HMM has proved to be a crucial but often a difficult task, it has been demonstrated through simulations that the estimates converge in a distribution that enables exact finite-sample inference. This research adopts methods from Kim and Stoffer (2008), Cappe (2009) and Cappe and Moulines (2009) in the literature of SMCEM algorithm and applications. Kim and Stoffer (2008) proposed a version of the particle methods and EM algorithm to estimate parameters of the SV model using chi-square which was later modified to normal mixture as its observation noise. This is what endeared it to this work as a framework.

In this section, a description of the SV model followed by SV model with a normal mixture distribution is discussed. The student-t SV model as well as the Generalized Error Distribution (GED) is also introduced, which represent the extension. An overview of the estimation procedure is presented, as well as theoretical results of the proposed technique. .

3.1 STOCHASTIC VOLATILITY MODEL

Stochastic Volatility models belong to state-space form of a Hidden Markov model which take the volatility of the data into account. The idea behind the family of SV models is that the volatility is driven by a latent process representing the flow of price relevant

information. However, this feature makes the SV model difficult to implement. Several specifications of the model have been suggested since its introduction by Taylor (1982). To fix notation and set the stage for our discussion, if we let y_t denote the return from an asset at time t , the SV model due to Taylor (1982) can be expressed as

$$x_t = \phi x_{t-1} + w_t \quad (3.1)$$

$$r_t = \beta \exp\left(\frac{x_t}{2}\right) v_t \quad (3.2)$$

where $w_t \sim N(0, \tau)$, $x_0 \sim N(\mu_0, \sigma_0^2)$, $v_t \sim N(0, 1)$, $\{y_t\}_{t \geq 0}$ is the log-returns on day t , we call β as the constant scaling factor, so that $\{x_t\}_{t \geq 0}$ represents the log of squared volatility. We assume that $|\phi| < 1$. The model are linearized by taking logarithms of the squared returns,

$$y_t = \alpha + x_t + z_t \quad (3.3)$$

Equations (3.1) and (3.2) form the version of the SV model which can be extended in many ways.

3.2 STOCHASTIC VOLATILITY WITH HEAVY –TAILED DISTRIBUTION

The standard form of the SV model is given in equations (3.1) and (3.3). In equation (3.3) z_t follows a normal distribution. Various authors have argued that real data may have heavier tails than can be captured by the standard SV model. It is a known fact that daily asset returns have leptokurtic distributions. The kurtosis of x_t following the above SV model is given by

$$k = \frac{E(x_t^4)}{(E(x_t^2))^2} = E(v_t^4) \exp(\tau_x), \quad (3.4)$$

where τ represents the unconditional variance of x_t .

The standard normal distribution is usually assumed for the distribution of v_t . If so,

$E(v_t^4) = 3$ and hence the kurtosis x_t is

$$k = 3 \exp(\tau) \geq 3 \quad (3.5)$$

where $k = 3$ only if $\tau = 0$. This distribution indicates that, as long as the volatility changes over time, the unconditional distribution of x_t is leptokurtic even if v_t follows the standard normal.

3.2.1 The stochastic volatility model with Normal mixture

The observational noise process of Kim and Stoffer (2008) is a mixture of two normal distributions with unknown parameters given as

$$y_k = x_t + z_t \quad (3.6)$$

with

$$z_t = I_t z_{t-1} + (1 - I_t) z_{t0}, \quad z_{t0} \sim N(m_0, R_0), \quad z_{t1} \sim N(m_1, R_1), \quad m_0 = \alpha - \mu\pi,$$

$m_1 = \alpha + (1 - \pi)\mu$ and $I_t \sim \text{Bernuolli}(\pi)$. I_t is an indicator variable, where π is an unknown mixing probability, i.e. $p(I_t = 1) = \pi = 1 - p(I_t \sim \text{Bernuolli}(\pi))$. The simplest extension of the normal distribution exhibiting higher kurtosis is a mixture of two normals.

The kurtosis of the normal mixture distribution Asai (2009) is

$$\frac{3(I_t R_0^4 + (1 - I_t) R_1^4)}{(I_t R_0^2 + (1 - I_t) R_1^2)^2}. \quad (3.7)$$

The likelihood of $\{x_0, \dots, x_n, y_1, \dots, y_n, I_1, \dots, I_n\}$ is

$$f(X, Y, I) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(x_0 - \mu_0)^2}{2\sigma_0^2}\right) \prod_{t=1}^n \frac{1}{\sqrt{2\pi\tau}} \exp\left(-\frac{(x_t - \phi x_{t-1})^2}{2\tau}\right) \prod_{t=1}^n \pi^{I_t} (1-\pi)^{1-I_t} \\ \prod_{t=1}^n \frac{1}{\sqrt{2\pi R_t^*}} \exp\left(-\frac{(y_t - x_t - \mu_t^*)^2}{2R_t^*}\right) \quad (3.8)$$

where $R_t^* = I_t R_1 + (1 - I_t) R_0$, $\mu_t^* = I_t q_1 + (1 - I_t) q_0$.

$\{q_0, q_1, R_0, R_1, \pi\}$ are estimated in the estimation step along with the other parameters, $\{\phi, \tau\}$.

3.2.2 The Proposed model - technique based on the student-t SV model

One of the main assumptions of the SV model is that v_t in equation (3.2) has normal distribution. It is generally known that the densities of many financial time series exhibit larger kurtosis than can be explained by the standard SV models with normal error distribution. The SV model with a fat tail distribution has been proposed to deal with this problem.

However, the extension from normal to heavy-tailed distributional assumption increases the computational load dramatically. In this study, the impact of the distribution assumption on the v_t variable is analyzed. In particular, this study extends the observation equation by allowing the v_t variable to follow the student-t distribution and a Generalized Error distribution (GED). The student-t distribution and the GED allow for fatter tails compared to the normal distribution. A Sequential Monte Carlo with Expectation-Maximization algorithm technique is developed for the extended volatility model.

3.2.2.1 Student-t as an observation noise

We now consider an extension of the linearized version of the SV model (see equation (3.1) and (3.2), wherein it is assumed that the observational noise process, z_t , follow a student-t distribution.

The model, first presented in Shumway and Stoffer (2006), retains the state equation for the volatility as:

$$x_t = \phi x_{t-1} + w_t$$

but the proposed student-t distribution with degrees of freedom, ν , for the observation error term, z_t , effects a change in the observation equation:

$$y_t = \alpha + x_t + z_t \quad z_t \sim t_{\nu}, t = 1, \dots, n. \quad (3.9)$$

The distribution of the error term for this specification according to Shimada and Tsukuda (2005) takes the form

$$f(y_t | x_t) = \frac{1}{\sqrt{\pi(\nu-2)}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{x_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{\nu-2}\right)^{-\frac{\nu+1}{2}}, \quad (3.10)$$

where ν represents a parameter of degree of freedom and Γ stands for the Gamma function.

The likelihood function of $\{x_0, x_1, \dots, x_n, y_1, \dots, y_n\}$ is

$$f(X, Y) = f(x_0) \prod_{t=1}^n f(x_t | x_{t-1}) \prod_{t=1}^n f(y_t | x_t)$$

$$x_t \sim N(\mu_0, \sigma_0^2), x_t | x_{t-1} \sim N(\phi x_{t-1}, \tau), y_t | x_t \sim t \text{ distribution}(\nu)$$

$$f(X, Y) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{1}{2}\left(\frac{x_0 - \mu_0}{\sigma_0}\right)^2\right) \prod_{t=1}^n \frac{1}{\sqrt{2\pi\tau}} \exp\left(-\frac{1}{2}\left(\frac{x_t - \phi x_{t-1}}{\tau}\right)^2\right) \\ \prod_{t=1}^n \frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} e^{-\frac{x_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{v-2}\right)^{-\frac{v+1}{2}} \quad (3.11)$$

Here, the target parameters are $\{\alpha, \phi, \tau, v\}$. The kurtosis of the student-t distribution is

$$E(v_t^4) = \frac{3(v-2)}{v-4} \quad ,$$

which is greater than three if $v < \infty$.

This change to the proposed model allows for a more robust fit, giving us a new tool to explore the tail fit.

3.2.2.2 Generalized error distribution as an observation noise

The distribution of the GED according to Bao et al., (2006) takes the form

$$f(y_t | x_t) = \frac{v \exp\left[-\frac{1}{2}((y_t - \alpha - v_t)/\psi)^v\right]}{\psi \Gamma\left(\frac{1}{v}\right) 2^{\left(1+\frac{1}{v}\right)}}, \quad (3.12)$$

where

$$\psi = \left[2^{-\frac{2}{v}} \frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)} \right]^{\frac{1}{2}}.$$

The Kurtosis of the GED is given by

$$E(\varepsilon_t^4) = \Gamma\left(\frac{1}{v}\right) \Gamma\left(\frac{5}{v}\right) \left[\Gamma\left(\frac{3}{v}\right) \right]^{-2},$$

which is greater than three if $v < 2$.

The log-likelihood function for the GED model is given by

$$\log f(y_t | x_t) = n \log v - \frac{1}{2} \sum_{t=1}^n \left(\frac{y_t - \alpha - v_t}{\psi} \right)^v - n \log \psi - n \log \Gamma\left(\frac{1}{v}\right) - n \log 2^{\left(1+\frac{1}{v}\right)}. \quad (3.13)$$

3.3 SEQUENTIAL MONTE CARLO EXPECTATION-MAXIMIZATION ANALYSIS

In this section, the method used in this work for the analysis of the SV model with the student-t distribution and the GED is explained. As is well known, it is difficult to estimate the parameters in the SV model using the maximum likelihood method. Several alternative methods have been proposed. Among such method, is the extension of the SMCEM techniques.

Parameter estimation

The entire estimation procedure consists of three main steps: filtering, smoothing, and estimation. We consider parameter estimation for the student-t and GED SV model. A basic approach for the student-t SV model, equation (3.2), is to apply the EM algorithm. With the output of filtering and smoothing step an approximate expected likelihood is calculated.

3.3.1 Filtering Step:

The algorithm below for the filtering and smoothing steps shows an extension of Godsill et al., (2004) and Kim and Stoffer (2008) results. From here, M samples from $f(x_t, | Y_t)$ for each t were obtained.

- 1) Generate $f_0^{(i)} \sim N(\mu_0, \sigma_0^2)$
- 2) For $t = 1, \dots, n$
 - a. Generate a random number $w_t^{(i)} \sim N(0, \tau)$, $i = 1, \dots, M$
 - b. Compute $p_t^{(i)} = \phi f_{t-1}^{(i)} + w_t^{(i)}$
 - c. Compute $w_t^{(i)} = p(y_t | p_t^{(i)},) \propto e^{-\frac{x_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{v-2} \right)^{-\frac{v+1}{2}}$
 - d. Generate $f_t^{(i)}$ by resampling with weights, $w_t^{(i)}$

3.3.2 Smoothing step

In the smoothing step, particle smoothers that are needed to get the expected likelihood in the expectation step of the EM algorithm were obtained:

Suppose that equally weighted particles $\{f_t^{(i)}\}, i = 1, \dots, M$ from $f(x_t, | Y_t)$ are available for $t = 1, \dots, n$ from the filtering step.

1) Choose $[s_n^{(i)}] = [f_n^{(j)}]$ with probability $\frac{1}{M}$.

2) For $n-1$ to 0

a. Calculate

$$w_{t|t+1}^{(i)} \propto f(s_{t+1}^{(i)} | f_t^{(j)}) \propto \exp\left(-\frac{(s_{t+1}^{(i)} - \phi f_t^{(j)})^2}{2\tau}\right) \frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left[\frac{v+1}{2}\right]}{\Gamma\left(\frac{v}{2}\right)} \exp^{-\frac{\tilde{s}_{t+1}}{2}} \left(1 + \frac{y_t^2 e^{-\tilde{s}_{t+1}^{(j)}}}{v-2}\right)^{-\frac{v+1}{2}}$$

for each j

b. Choose $[s_t^{(i)}] = [f_t^{(j)}]$ with probability $w_{t|t+1}^j$.

3) $(s_{0:n}^{(i)}) = \{(s_0^{(i)}, \dots, s_n^{(i)})\}$ is the random sample from $f(x_0, \dots, x_n | Y_n)$.

4) Repeat steps (1) – (3), for $i = 1, \dots, M$ and calculate

$$\hat{x}_t^n = \frac{\sum_{i=1}^M s_t^{(i)}}{M}, \hat{p}_t^n = \frac{\sum_{i=1}^M (s_t^{(i)} - \hat{x}_t^n)^2}{M-1}, \hat{p}_{t,t-1}^n = \frac{\sum_{i=1}^M (s_t^{(i)} - \hat{x}_t^n)(s_{t-1}^{(i)} - \hat{x}_{t-1}^n)}{M},$$

$$E\left[1 + \frac{y_t^2 e^{x_t}}{v-2}\right]^{-\frac{v+1}{2}} = \frac{n(v-2)}{(v+1) \sum_{t=1}^n y_t^2 e^{-y_t + v_t} \left[1 + \frac{y_t^2 e^{x_t}}{v-2}\right]^{-1}}.$$

3.3.3 Estimation Step

This step consists of obtaining parameter estimates by setting the derivative of the expected likelihood, of the complete data $\{x_0, \dots, x_n, y_1, \dots, y_n\}$ given $\{x_0, \dots, x_n\}$, with respect to each parameter to zero and solving for $\hat{\phi}$, $\hat{\tau}$, and $\hat{\alpha}$.

We view $\{x_0, \dots, x_n\}$ as unobserved and apply the EM algorithm. However, problems arise when the model is not Gaussian because it can be difficult to compute such expected likelihood. A simulation-based particle filters and smoothers are therefore used to obtain the values needed for calculating the expected likelihood. The procedure of this algorithm consists of running a filtering and smoothing step for the given parameters, and then running an estimation step to get updated parameter estimates.

To apply the EM algorithm, there is need to calculate the complete data likelihood and the expected likelihood given the data. The followings are the details.

The complete likelihood of $\{x_0, x_1, \dots, x_n, y_1, \dots, y_n\}$ is

$$f(X, Y) = f(x_0) \prod_{t=1}^n f(x_t | x_{t-1}) \prod_{t=1}^n f(y_t | x_t)$$

$$x_t \sim N(\mu_0, \sigma_0^2), x_t | x_{t-1} \sim N(\phi x_{t-1}, \tau), y_t | x_t \sim t \text{ distribution}(v)$$

$$f(x_0) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{1}{2}\left(\frac{x_0 - \mu_0}{\sigma_0}\right)^2\right), \quad f(x_t | x_{t-1}) = \frac{1}{\sqrt{2\pi\tau}} \exp\left(-\frac{1}{2}\left(\frac{x_t - \phi x_{t-1}}{\tau}\right)^2\right)$$

$$f(y_t | x_t) = \frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} e^{-\frac{x_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{v-2}\right)^{-\frac{v+1}{2}}$$

$$\begin{aligned}\log f(X, Y) &= \log \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_0} + \log \exp\left(-\frac{(x_0 - \mu_0)^2}{2\sigma_0^2}\right) + \log \prod_{t=1}^n \frac{1}{\sqrt{2\pi\tau}} \exp\left(-\frac{(x_t - \phi x_{t-1})^2}{2\tau}\right) \\ &\quad + \log \prod_{t=1}^n \frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} e^{-\frac{(y_t - \alpha - v_t)^2}{2}} \left(1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)^2}}{v-2}\right)^{-\frac{v+1}{2}}\end{aligned}$$

$$\begin{aligned}\log f(X, Y) &= -\frac{1}{2} \log \sigma_0^2 - \frac{(x_0 - \mu_0)^2}{2\sigma_0^2} - \frac{n}{2} \log \tau - \frac{1}{2} \sum_{t=1}^n \frac{(x_t - \phi x_{t-1})^2}{\tau} \\ &\quad - \frac{1}{2} \sum_{t=1}^n (y_t - \alpha - v_t) + \sum_{t=1}^n \log \left[1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)^2}}{v-2}\right]^{-\frac{v+1}{2}} + C\end{aligned}$$

where C is constant. Multiply through by -2

$$\begin{aligned}-2 \log f(X, Y) &= \log \sigma_0^2 + \frac{(x_0 - \mu_0)^2}{\sigma_0^2} + n \log \tau - \sum_{t=1}^n \frac{(x_t - \phi x_{t-1})^2}{\tau} \\ &\quad + \sum_{t=1}^n (y_t - \alpha - v_t) - 2 \sum_{t=1}^n \log \left[1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)^2}}{v-2}\right]^{-\frac{v+1}{2}} + C\end{aligned}$$

Hence, the expected likelihood of $\{x_0, x_1, \dots, x_n, y_1, \dots, y_n\}$ given $\{y_1, \dots, y_n\}$ is

$$\begin{aligned}P(\theta) &= E(-2 \log f(X, Y) | Y) = E \left[\begin{aligned} &\log \sigma_0^2 + \frac{(x_0 - \mu_0)^2}{\sigma_0^2} + n \log \tau + \sum_{t=1}^n \frac{(x_t - \phi x_{t-1})^2}{\tau} \\ &+ \sum_{t=1}^n (y_t - \alpha - v_t) - 2 \sum_{t=1}^n \log \left[1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)^2}}{v-2}\right]^{-\frac{v+1}{2}} + C \end{aligned} \right] \\ P(\theta) &= \log \sigma_0^2 + \frac{(x_0^n - \mu_0)^2}{\sigma_0^2} + p_0^n + n \log \tau + \sum_{t=1}^n \frac{(x_t^n - \phi x_{t-1}^n)^2}{\tau} + \frac{p_t^n + \phi^2 p_{t-1}^n - 2\phi p_{t,t+1}^n}{\tau} \\ &\quad + \sum_{t=1}^n (y_t - \alpha - v_t) + \sum_{t=1}^n \log \left[1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)^2}}{v-2}\right]^{-\frac{v+1}{2}} + C\end{aligned}$$

where $x_t^n = E(x_t | Y_n)$, $p_t^n = E[(x_t - x_t^n)^2 | Y_n]$ and $p_{t,t-1}^n = E[(x_t - x_t^n)(x_{t-1} - x_{t-1}^n) | Y_n]$

By maximizing $P(\theta)$, we get the estimates $\hat{\phi}$, $\hat{\tau}$, $\hat{\alpha}$

$$\begin{aligned}\frac{\partial P(\theta)}{\partial \phi} &= \sum_{t=1}^n \frac{2(x_t^n - \phi x_{t-1}^n)(-x_{t-1}^n) + 2\phi p_{t-1}^n - 2p_{t,t+1}^n}{\tau} \\ &= \sum_{t=1}^n \frac{2}{\tau} [\phi[(x_{t-1}^n)^2 + p_{t-1}^n] - (x_t^n x_{t-1}^n + p_{t,t+1}^n)]\end{aligned}$$

$$\frac{\partial P(\theta)}{\partial \phi} = \frac{2}{\tau} \phi (S_{00} - S_{10}) = 0$$

where

$$S_{00} = \sum_{t=1}^n (x_{t-1}^n)^2 + p_{t-1}^n ,$$

$$S_{11} = \sum_{t=1}^n (x_t^n)^2 - p_t^n ,$$

$$S_{10} = \sum_{t=1}^n x_t^n x_{t-1}^n + p_{t,t-1}^n$$

$$\hat{\phi} = \frac{S_{10}}{S_{00}}$$

$$\frac{\partial P(\theta)}{\partial \tau} = \frac{n}{\tau} + \sum_{t=1}^n \frac{-(x_t^n - \phi x_{t-1}^n)^2 + (-x_{t-1}^n) + p_t^n + \phi^2 p_{t-1}^n - 2\phi p_{t,t-1}^n}{\tau^2}$$

$$= \frac{1}{\tau^2} \left[n\tau - \sum_{t=1}^n [(x_t^n)^2 + p_t^n] - \phi^2 \sum_{t=1}^n [(x_{t-1}^n)^2 + p_{t-1}^n] + 2\phi \sum_{t=1}^n [x_t^n x_{t-1}^n + p_{t,t-1}^n] \right]$$

$$\tau = \frac{1}{n} [S_{11} + \phi^2 S_{00} - 2\phi S_{10}]$$

Putting $\hat{\phi} = \frac{S_{10}}{S_{00}}$

$$\hat{\tau} = \frac{1}{n} \left[S_{11} - \frac{S_{10}^2}{S_{00}} \right]$$

$$\begin{aligned} \frac{\partial P(\theta)}{\partial \alpha} &= \sum_{t=1}^n (-1) - 2 \sum_{t=1}^n \frac{-\left(\frac{v+1}{2}\right) \left[1 + \frac{y_t^2 e^\alpha e^{-y_t+v_t}}{v-2} \right]^{-\left(\frac{v+1}{2}\right)-1} \frac{y_t^2 e^\alpha e^{-y_t+v_t}}{v-2}}{\left[1 + \frac{y_t^2 e^{-y_t+v_t}}{v-2} \right]^{-\frac{v+1}{2}}} \\ &= -n + 2 \sum_{t=1}^n \left(\frac{v+1}{2} \right) \left[1 + \frac{y_t^2 e^\alpha e^{-y_t+v_t}}{v-2} \right]^{-1} \frac{y_t^2 e^\alpha e^{-y_t+v_t}}{v-2} = 0 \\ &= \frac{e^\alpha}{v-2} (v+1) \sum_{t=1}^n y_t^2 e^{-y_t+v_t} \left[1 + \frac{y_t^2 e^\alpha e^{-y_t+v_t}}{v-2} \right]^{-1} = n \\ &= \frac{e^\alpha}{v-2} (v+1) \sum_{t=1}^n y_t^2 e^{-y_t+v_t} \left[1 + \frac{y_t^2 e^{x_t}}{v-2} \right]^{-1} = n \end{aligned}$$

$$e^\alpha = \frac{n(v-2)}{(v+1) \sum_{t=1}^n y_t^2 e^{-y_t+v_t} \left[1 + \frac{y_t^2 e^{x_t}}{v-2} \right]^{-1}}$$

$$\hat{\alpha} = \log \left[\frac{n(v-2)}{(v+1) \sum_{t=1}^n y_t^2 e^{-y_t+v_t} \left[1 + \frac{y_t^2 e^{x_t}}{v-2} \right]^{-1}} \right].$$

When z_t follows the GED, it is not possible to represent it as equation (3.2). Hence, for the SV-GED model, we sample v as well as the other parameters and x_t from their full conditional distributions using the SMCCEM techniques.

The log likelihood function for the GED model is given by

$$\log f(X, Y) = \log \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_0} + \log \exp \left(-\frac{(x_0 - \mu_0)^2}{2\sigma_0^2} \right) + \log \prod_{t=1}^n \frac{1}{\sqrt{2\pi\tau}} \exp \left(-\frac{(x_t - \phi x_{t-1})^2}{2\tau} \right)$$

$$+ \log \prod_{t=1}^n \frac{v \exp \left[-\frac{1}{2} \left((y_t - \alpha - v_t) / \psi \right)^v \right]}{\psi \Gamma \left(\frac{1}{v} \right) 2^{\left(1 + \frac{1}{v} \right)}}$$

$$\log f(y_t | x_t) = n \log v - \frac{1}{2} \sum_{t=1}^n \left(\frac{y_t - \alpha - v_t}{\psi} \right)^v - n \log \psi - n \log \Gamma \left(\frac{1}{v} \right) - n \log 2^{\left(1 + \frac{1}{v} \right)}$$

$$\log f(y_t | x_t) = C_1 - \frac{1}{2} \sum_{t=1}^n \left(\frac{y_t - \alpha - v_t}{\psi} \right)^v$$

$$\text{where } C_1 = n \log v - n \log \psi - n \log \Gamma \left(\frac{1}{v} \right) - n \log 2^{\left(1 + \frac{1}{v} \right)} \text{ and } \psi = \left[2^{-\frac{2}{v}} \frac{\Gamma \left(\frac{1}{v} \right)}{\Gamma \left(\frac{3}{v} \right)} \right]^{\frac{1}{2}}$$

$$P(\theta) = E(\log f(X, Y) | Y) = E \left[C_1 - \frac{1}{2} \sum_{t=1}^n \left(\frac{y_t - \alpha - v_t}{\psi} \right)^v \right]$$

$$\frac{\partial \log f(y_t|x_t)}{\partial \alpha} = -\frac{1}{2} \sum_{t=1}^n v \left(\frac{y_t - \alpha - v_t}{\psi} \right)^{v-1} - \frac{1}{\psi}$$

$$\text{for } \frac{\partial \log f(y_t|x_t)}{\partial \alpha} = 0$$

$$\frac{v}{2\psi} \sum_{t=1}^n \left(\frac{y_t - \alpha - v_t}{\psi} \right)^{v-1} = 0$$

$$\frac{v}{2\psi} \sum_{t=1}^n \left(\frac{y_t - v_t}{\psi} - \frac{\alpha}{\psi} \right)^{v-1} = 0$$

$$\frac{v}{2\psi} \sum_{t=1}^n \left[\left(\frac{y_t - v_t}{\psi} \right)^{v-1} - \left(\frac{\alpha}{\psi} \right)^{v-1} \right] = 0$$

$$\frac{v}{2\psi} \left[\sum_{t=1}^n \left(\frac{y_t - v_t}{\psi} \right)^{v-1} - \sum_{t=1}^n \left(\frac{\alpha}{\psi} \right)^{v-1} \right] = 0$$

$$\sum_{t=1}^n \left(\frac{y_t - v_t}{\psi} \right)^{v-1} = \sum_{t=1}^n \left(\frac{\alpha}{\psi} \right)^{v-1}$$

$$n \frac{\alpha^{v-1}}{\psi^{v-1}} = \frac{1}{\psi^{v-1}} \sum_{t=1}^n (y_t - v_t)^{v-1}$$

$$\alpha^{v-1} = n \sum_{t=1}^n (y_t - v_t)^{v-1}$$

$$\hat{\alpha} = \left[n \sum_{t=1}^n (y_t - v_t)^{v-1} \right]^{\frac{1}{v-1}} .$$

3.4 OTHER CONSIDERATIONS

3.4.1 Initial Parameter Selection

To save computing time, it is essential to start with good initial parameters. Anderson et al., (1969) suggested consistent estimates of the parameters of a linear system based on the idea of the method of moments. Their idea is applied to get initial parameters.

1) $[\alpha]$

$$E(y_t) = \alpha + E(x_t) + E(z_t) = \alpha$$

$$\alpha^{(0)} = \frac{\sum_{t=0}^n y_t}{n} \quad . \quad (3.14)$$

2) $[\phi]$

$$\begin{aligned} E\{(y_t - E(y_t))(y_{t-2} - E(y_{t-2}))\} &= E\{(\alpha + x_t + z_t - \alpha)(y_{t-2} - E(y_{t-2}))\} \\ &= E\{(\phi x_{t-1} + w_t + z_t)(y_{t-2} - E(y_{t-2}))\} \\ &= E\{(\phi(y_{t-1} - \alpha - z_{t-1}) + w_t + v_t)(y_{t-2} - E(y_{t-2}))\} \\ &= \phi\{(y_{t-1} - E(y_{t-1}))(y_{t-2} - E(y_{t-2}))\} \\ &= \phi^{(0)} = \frac{\sum_{t=3}^n (y_t - \bar{y}_t)(y_{t-2} - \bar{y}_{t-2})}{\sum_{t=3}^n (y_{t-1} - \bar{y}_{t-1})(y_{t-2} - \bar{y}_{t-2})} \quad . \end{aligned} \quad (3.15)$$

3) $[\tau]$

$$\begin{aligned} &E[\{y_t - E(y_t) - \phi(y_{t-1} - E(y_{t-1}))\}^2] \\ &= E\{\alpha + x_t + z_t - \alpha - \phi(\alpha + x_{t-1} + z_{t-1} - \alpha)\}^2 \\ &= E\{w_t + z_t - \phi z_{t-1}\}^2 \\ &= \tau + \text{var}(z_t) + \phi^2 \text{var}(z_{t-1}) \\ &= \tau^{(0)} = \frac{\sum_{t=2}^n (y_t - \bar{y}_t - \phi(y_{t-1} - \bar{y}_{t-1}))^2}{n-1} - \hat{\text{var}}(z_t) - \hat{\phi}^2 \text{var}(z_{t-1}) \quad . \end{aligned} \quad (3.16)$$

3.4.2 Relative Likelihood

In the EM algorithm, the likelihood of the observed data increases at every iteration.

Although the E-step, which uses particles to calculate the expected likelihood, does not guarantee the monotone likelihood property, the behavior of the relative likelihood is worth studying. The relative likelihood is the ratio of the likelihoods at two adjacent

iterations, and the relative likelihood at i th iteration, $\frac{f_{\theta^{(i)}}(y)}{f_{\theta^{(i-1)}}(y)}$, are calculated by using

the complete likelihood as follows:

$$\frac{f_{\theta^{(i)}}(y)}{f_{\theta^{(i-1)}}(y)} = \frac{f_{\theta^{(i)}}(x, y)}{f_{\theta^{(i)}}(x|y)} \cdot \frac{f_{\theta^{(i-1)}}(x|y)}{f_{\theta^{(i-1)}}(x, y)},$$

where y is observed data and x is complete data.

On Multiplying $f_{\theta^{(i)}}(x|y)$ and integrating out x , we get

$$\frac{f_{\theta^{(i)}}(y)}{f_{\theta^{(i-1)}}(y)} = E_{\theta^{(i-1)}} \left[\frac{f_{\theta^{(i)}}(x, y)}{f_{\theta^{(i)}}(x|y)} | y \right].$$

The change in the log-likelihood is

$$\begin{aligned} l(\theta^{(i-1)}, \theta^{(i)}) &= \log f_{\theta^{(i)}}(y) - \log f_{\theta^{(i-1)}}(y) \\ &= \log \frac{f_{\theta^{(i)}}(y)}{f_{\theta^{(i-1)}}(y)} \\ &= \log E_{\theta^{(i-1)}} \left[\frac{f_{\theta^{(i)}}(x, y)}{f_{\theta^{(i)}}(x|y)} | y \right]. \end{aligned} \tag{3.17}$$

3.4.3 Stopping Rule and Selection of Particle Size

Theoretically, the suggested algorithm converges when the particle size, M , and the number of iterations, N , are large. Practically, it is not possible to use infinitely large M

and N . The choice of N is important as the estimates from the procedure when stopped too early may not be reliable. Moreover, it is a waste of time and resources to run the procedure longer than necessary. Many numerical procedures involving iteration compares estimates from contiguous iteration; if the two are close enough, the process is considered to have converged, and is stopped. Equivalently, the relative likelihood, which is the difference between two likelihood values at two adjacent iterations are considered as a measure in assessing convergence. In particular, a small relative likelihood signifies that a process approaches convergence. Here, the relative likelihood is used to assess convergence. More specifically, we conclude that the process is converged if the relative likelihood is less than some pre-determined tolerance, ϵ .

Regarding the selection of M , Tanner (1996) mentioned that it is wise to start with a small M , increasing it as the current approximation moves closer to the maximum likelihood estimate in the Markov chain EM setting. We apply Tanner's method to save computing time.

On the other hand, a key element in statistical analysis of HMM is a strong law of large numbers for the log-likelihood function. In previous works, stability theory of Markov chains and the subadditive ergodic theorem were used; see Baum and Petrie (1966), Douc and Matais (2001), Legland and Mevel (2000) and Leroux (1992). Early reviews of the theoretical properties of particle filters are found in Doucet et al., (Chapter 2 and 3, 2001) and Crisan and Doucet (2002) while Chapter 9 of Cappe et al., (2005) includes an introduction to consistency and asymptotic normality for several particle filtering algorithms.

3.5 PROOF OF CONVERGENCE OF THE PROPOSED SMCEM ALGORITHM

In this section, the convergence of the proposed method presented is proved using the following theorem:

The Theorem states that for suitable starting values close to an isolated maximizer of the log-likelihood of the observations and with sufficiently large particle sample, an SMCEM sequence will get close to the maximizer with high probability.

Theorem 3.5.1 (Chan and Ledolter, 1995)

Let θ^* be an isolated local maximum of $l_y(\theta)$, the observed likelihood, then there exists a neighbourhood, say V_1 , of θ^* such that for starting values of the algorithm presented inside V_1 and for all $\varepsilon_1 > 0$, there exists a k_0 such that $p(\|\theta_k^{(M)} - \theta^*\| < \varepsilon_1 \text{ for some } k \leq k_0) \rightarrow 1 \text{ as } M \rightarrow \infty$.

Propositions

(i): If for all θ' , $l_x(\theta') = r(W, \theta')$, where W is a measurable vector function of X , then

r is linear in W and $l_x(\theta')$ is the complete likelihood at the parameter θ' .

(ii): If $r(W, \theta')$ attains the unique global maximum at $N(W)$ then N is continuous in W .

(iii): Let θ^* be an isolated local maximum of $l_y(\theta)$ then $l_y(\theta)$ is continuous in θ .

(iv): If $\bar{W}_n(\theta)$ is the sample mean of W values computed from the particle sample

$X_1(\theta), \dots, X_n(\theta)$ then the convergence, in probability, of $\bar{W}_n(\theta) \rightarrow E_\theta(W)$ as

$N \rightarrow \infty$ is uniform over compact subsets of Ω where Ω is a parameter space.

Proof:

(i): The log-likelihood of the complete data for the student-t SV model is

$$\begin{aligned}
 \log f(X, Y) &= -\frac{1}{2} \log \sigma_0^2 - \frac{(x_0 - \mu_0)^2}{2\sigma_0^2} - \frac{n}{2} \log \tau - \frac{1}{2} \sum_{t=1}^n \frac{(x_t - \phi x_{t-1})^2}{\tau} \\
 &\quad - \frac{1}{2} \sum_{t=1}^n (y_t - \alpha - v_t) + \sum_{t=1}^n \log \left[1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)}}{v - 2} \right]^{-\frac{v+1}{2}} \\
 \log f(X, Y) &= -\frac{1}{2} \sum_{t=1}^n \frac{(x_t - \phi x_{t-1})^2}{\tau} - \frac{1}{2} \sum_{t=1}^n (y_t - \alpha - v_t) + \sum_{t=1}^n \log \left[1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)}}{v - 2} \right]^{-\frac{v+1}{2}} \\
 &= -\frac{1}{2} \left[\sum_{t=1}^n \frac{(x_t^2 - 2\phi x_t x_{t-1} + \phi^2 x_{t-1}^2)}{\tau} + \sum_{t=1}^n y_t - n\alpha - \sum_{t=1}^n v_t + (v+1) \sum_{t=1}^n \log \left[1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)}}{v - 2} \right] \right]
 \end{aligned}$$

But, we recall that

$$\begin{aligned}
 \log(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \\
 \Rightarrow \log \left(1 + \frac{y_t^2 e^{-(y_t - \alpha - v_t)}}{v - 2} \right) &= \frac{y_t^2 e^{-(y_t - \alpha - v_t)}}{v - 2} - \frac{\left[\frac{y_t^2 e^{-(y_t - \alpha - v_t)}}{v - 2} \right]^2}{2} + \dots
 \end{aligned}$$

$$\begin{aligned}
 \log f(X, Y) = l_X(\theta) &= -\frac{1}{2} \left[\frac{W_1}{\tau} \right] - \frac{2\phi}{\tau} W_2 + \frac{\phi^2}{\tau} W_3 + W_4 - n\alpha - W_5 + \frac{e^\alpha (v+1)}{v-2} W_6 - \frac{e^{2\alpha} (v+1)}{(v-2)^2} W_7 + c \\
 &= r(W, \theta)
 \end{aligned}$$

where $W = (W_1, \dots, W_7)$

The dimension of W is not assumed constant and may grow with that of X .

$$W_1 = \sum_{t=1}^n x_t^2$$

$$W_2 = \sum_{t=1}^n x_t x_{t-1}$$

$$W_3 = \sum_{t=1}^n x_{t-1}^2$$

$$W_4 = \sum_{t=1}^n y_t,$$

$$W_5 = \sum_{t=1}^n v_t$$

$$W_6 = \sum_{t=1}^n y_t^2 e^{-(y_t - v_t)}$$

$$W_7 = \sum_{t=1}^n y_t^4 e^{-2(y_t - v_t)}$$

$$\text{and } C = \log \sigma_0^2 + \frac{(x_0 - \mu_0)^2}{\sigma_0^2} + n \log \tau - n\alpha.$$

Since W is a measurable vector function of X and r is linear in W , then proposition (i) is satisfied.

(ii): Also, $r(W, \theta')$ attains the unique global maximum at $N(W)$ where

$$N(W) = \left(\arg \max_{\phi} q, \arg \max_{\tau} q, \arg \max_{\alpha} q \right)$$

$$l_x(\theta) = -\frac{1}{2} \left[\frac{W_1}{\tau} \right] - \frac{2\phi}{\tau} W_2 + \frac{\phi^2}{\tau} W_3 + W_4 - n\alpha - W_5 + \frac{e^\alpha (v+1)}{v-2} W_6 - \frac{e^{2\alpha} (v+1)}{(v-2)^2} W_7 + c$$

$$\frac{\partial l_x(\theta)}{\partial \phi} = \frac{W_2}{\tau} - \frac{\phi}{\tau} W_3 = 0$$

$$\phi = \frac{W_2}{W_3}$$

$$\frac{\partial l_x(\theta)}{\partial \tau} = -\frac{1}{2} \left[-\frac{1}{\tau^2} \left(W_1 - 2\phi W_2 + \phi^2 W_3 - \frac{n}{\tau} \right) \right] = 0$$

$$\tau = \frac{W_1 - 2\phi W_2 + \phi^2 W_3}{n}$$

$$\begin{aligned} \frac{\partial l_x(\theta)}{\partial \tau} &= -n + \frac{e^\alpha (v+1)}{v-2} W_6 - \frac{2e^{2\alpha} (v+1)}{(v-2)^2} W_7 = 0 \\ &= W_6 e^\alpha (v+1)^{(v-2)} - 2e^{2\alpha} (v+1) W_7 - n(v-2)^2 = 0 \end{aligned}$$

$$e^\alpha = \frac{-(v+1)(v-2) W_6 \pm \sqrt{(v+1)^2 (v-2)^2 W_6^2 - 4 - 2(v+1) W_7 - n(v-2)^2}}{2[-2(v+1) W_7]}$$

$$\alpha = \log \left(\frac{-(v+1)(v-2) W_6 \pm \sqrt{(v+1)^2 (v-2)^2 W_6^2 - 8n(v+1)(v-2)^2 W_7}}{-4(v+1) W_7} \right)$$

$$N(W) = \left(\arg \max_{\phi} q, \arg \max_{\tau} q, \arg \max_{\alpha} q \right)$$

$$= \left(\frac{W_2}{W_3}, \frac{W_1 - 2\phi W_2 + \phi^2 W_3}{n}, \log \left(\frac{-(v+1)(v-2) W_6 \pm \sqrt{(v+1)^2 (v-2)^2 W_6^2 - 8n(v+1)(v-2)^2 W_7}}{-4(v+1) W_7} \right) \right)$$

and $N(W)$ is continuous in W .

(iii). Furthermore, proposition (iii) is satisfied since W is a measurable vector function of X and $l_x(\theta') = r(W, \theta')$, which is linear in W and $l_x(\theta')$ is the complete likelihood at the parameter θ' where $r(W, \theta')$ attains the unique global maximum at $N(W)$, and it has been shown that $N(W)$ is continuous in W . Therefore, letting θ^* be an isolated local maximum of $l_y(\theta)$, it follows that $l_y(\theta)$ is continuous in θ .

(iv): we use the convergence result of the particle smoothing algorithm in Godsill et al., (2004). By Theorem A2 of Godsill et al., (2004), for all $t \in (1, \dots, n)$, there exists $c_{t:n}(\theta)$ independent of N such that for any bounded, Borel-measurable function ϕ on R^n ,

$$E \left[\left(\int \phi(x_1, \dots, x_n) d\pi_{1:n:n}^M(x_1, \dots, x_n; \theta) - \int \phi(x_1, \dots, x_n) d\pi_{1:n:n}(x_1, \dots, x_n; \theta) \right)^2 \right] \leq c_{1:n}(\theta) \frac{\|\phi\|^2}{N}$$

$$\int \phi(x_1, \dots, x_n) d\pi_{1:n:n}^N(x_1, \dots, x_n; \theta) \xrightarrow{P} \int \phi(x_1, \dots, x_n) d\pi_{1:n:n}^N(x_1, \dots, x_n; \theta)$$

We can assume squared returns since both equations above are equivalent, and hence the y_t , are bounded by a large number, then since W is bounded, we obtain

$$\overline{W}_N(\theta) \xrightarrow{P} E_\theta(W) \text{ as } N \rightarrow \infty.$$

Thus, we need to show that for every $\varepsilon > 0$ there is an integer M such that $N \geq M$ implies $P(|\overline{W}_N(\theta) - E_\theta(W)| \geq \varepsilon) < \varepsilon$ for all $\theta \in E$ where E is a compact subset of Ω to establish the uniform convergence.

Since \overline{W}_N and $E_\theta(W)$ are continuous in θ , for any $\varepsilon_1 > 0$ there exists $\delta_1 > 0$ and $\delta_2 > 0$ such that $|\theta_1 - \theta_2| < \delta_1$ implies $|\overline{W}_N(\theta_1) - \overline{W}_N(\theta_2)| < \varepsilon_1$ and $|\theta_1 - \theta_2| < \delta_2$ implies $|E_{\theta_1}(W) - E_{\theta_2}(W)| < \varepsilon_1$.

Let us consider an open covering $\{O_\alpha\}$ of E such that $O_\alpha = \{\theta : |\theta - \theta_\alpha| < \delta\}$ where δ is the minimum of δ_1 and δ_2 . Then by compactness of E we can choose the finite subset of $\{O_\alpha\}, \{\alpha\}, \alpha = 1, \dots, K$ which covers E .

Then for any θ in E we can find O_β such that $\theta \in O_\beta$ and $\beta \in \{1, \dots, K\}$.

Since

$$|\theta - \theta_\beta| < \delta, \text{ we get } |\bar{W}_N(\theta) - \bar{W}_N(\theta_\beta)| < \varepsilon_1 \text{ and } |E_\theta(W) - E_{\theta_\beta}(W)| < \varepsilon_1.$$

Therefore,

$$\begin{aligned} & p(|\bar{W}_N(\theta) - E_\theta(W)| \geq \varepsilon) \\ & \leq p(|\bar{W}_N(\theta) - \bar{W}_N(\theta_{\alpha'})| + |\bar{W}_N(\theta_{\alpha'}) - E_{\theta_{\alpha'}}(W)| + |E_{\theta_{\alpha'}}(W) - E_\theta(W)| \geq \varepsilon) \\ & \leq p\left(|\bar{W}_N(\theta) - \bar{W}_N(\theta_{\alpha'})| \geq \frac{\varepsilon}{3}\right) + p\left(|\bar{W}_N(\theta_{\alpha'}) - E_{\theta_{\alpha'}}(W)| \geq \frac{\varepsilon}{3}\right) \\ & \quad + p\left(|E_{\theta_{\alpha'}}(W) - E_\theta(W)| \geq \frac{\varepsilon}{3}\right) \end{aligned}$$

The first and the third terms of the above expression are less than $\frac{\varepsilon}{3}$ if we let $\varepsilon_1 = \frac{\varepsilon}{3}$,

and the second term is less than $\frac{\varepsilon}{3}$ by pointwise convergence for $N \geq M_{\alpha'}$. Thus, letting

$M = \max M_{\alpha'}$, we get that for every $\theta \in E$, $p(|\bar{W}_N - E_\theta(W)| \geq \varepsilon) \leq \varepsilon$ for $n \geq M$.

Remark: In propositions (i) and (ii), the propositions that $l_X(\theta') = r(W, \theta')$ is linear in W and that, $\bar{W}_n(\theta) \rightarrow E_\theta(W)$, uniformly for θ over compact sets in probability, are needed only to guarantee that $N_m(\theta)$ converges in probability to $E_\theta(W)$ uniformly for θ over compact sets.

Hence, as the algorithm satisfies all four propositions of the theorem, the convergence follows.

3.5.2 Consistency and asymptotic normality

In the context of i.i.d.-observations convergence properties of recursive EM-type algorithms have been studied in Titterton (1984), Sato (2000), Wang and Zhao (2006),

and Cappe and Moulines (2009) where also proofs of consistency and asymptotic normality are provided.

Early reviews of the theoretical properties of particle filters can be found in Chapters 2-3 of Doucet et al., (2001) and Crisan and Doucet (2002) while Chapter 9 of Cappe et al., (2005) includes introduction to consistency and asymptotic normality for several particle filtering algorithms. More recent papers on consistency and asymptotic normality are Douc et al., (2007) and Douc and Moulines (2008). Douc and Moulines (2008) also Cappe et al., (2005)) prove preliminary theorems showing that if one starts with a sample $\{x_{0:n-1}^{(i)}, w_{n-1}^{(i)}\}_{i=1}^N$ that produces a consistent and asymptotically normal estimator for a function f , then one iteration of the sampling and resampling operations produce a new sample $\{x_{0:n}^{(i)}, w_n^{(i)}\}_{i=1}^N$ whose estimator is also consistent and asymptotically normal for the function f . These theorems also govern consistency and asymptotic normality for the SIR algorithm of Rubin (1988). Therefore, an additional assumption is that estimates produced by the initial particles $\{x_0^{(i)}, w_0^{(i)}\}_{i=1}^N$ are consistent and asymptotically normal for a function f .

Given the propositions (i) - (iv) stated above, then Theorem 10 of Douc and Moulines (2008) states that the estimator which is the sample mean of W computed from the particle samples $X_1(\theta), \dots, X_n(\theta)$ produced by the particle filter and smoother is consistent at iteration n meaning that as $N \rightarrow \infty$

$$\overline{W}_N(\theta) \rightarrow E_\theta(W).$$

Proposition v.

Let $\bar{W}_N(\theta)$ be an estimator (based on sample size N) of θ and let $E_\theta(W) = W_\theta$, so that

$\bar{W}_N(\theta) \rightarrow W_\theta$. If for each $\varepsilon > 0$, we have that

$$P\left[\left|\bar{W}_N(\theta) - W_\theta\right| < \varepsilon\right] \rightarrow 1 \quad \text{as } N \rightarrow \infty$$

then $\bar{W}_N(\theta)$ is a consistent estimator.

Proof:

It is sufficient to prove that for any $\varepsilon > 0$ the probability is one that all limit points \bar{W} of the Sequence $\{\bar{W}_N\}$ satisfy the inequality $|\bar{W} - W_\theta| \leq \varepsilon$. The event that there exists a

limit point \bar{W} of the sequence $\{\bar{W}_N\}$ such that $|\bar{W} - W_\theta| > \varepsilon$ implies that

$\sup_{|\bar{W} - W_\theta| \geq \varepsilon} \{\phi(x_1, \theta) \cdots \phi(x_n, \theta)\} \geq \phi(x_1, \bar{w}_N) \cdots \phi(x_n, \bar{w}_N)$ for infinitely many N . But then

$$\frac{\sup_{|\bar{W} - W_\theta| \geq \varepsilon} \phi(x_1, \theta) \cdots \phi(x_n, \theta)}{\phi(x_1, w_\theta) \cdots \phi(x_n, w_\theta)} \geq c > 0$$

for infinitely many n . By theorem 1 of Wald (1949),

$$prob \left\{ \lim_{n \rightarrow \infty} \frac{\sup_{\theta \in W} f(x_1, \theta) f(x_2, \theta) \cdots f(x_n, \theta)}{f(x_1, \theta_0) f(x_2, \theta_0) \cdots f(x_n, \theta_0)} = 0 \right\} = 1$$

We have shown that the probability is one that all limit points \bar{W} of $\{\bar{W}_N\}$ satisfy the inequality $|\bar{W} - W_\theta| > \varepsilon$.

They are asymptotically normal meaning that as $N \rightarrow \infty$

$$\sqrt{N} (\bar{W}_N(\theta) - W_\theta) \rightarrow N(0, \sigma_N^2)$$

Due to the recursive nature of the algorithm, both the asymptotic variance σ_N^2 and the sets of functions for which these results will hold are determined recursively.

3.6 Multivariate Stochastic Volatility model based on student-t

The contribution of this work to the literature, on multivariate modeling of volatilities and correlations, lies in the introduction of observation-driven time-varying parameter models with heavy tailed distributions. Here, we extend SMCEM analysis to multivariate SV models with student-t distribution. Let $r_t = (r_{1,t}, \dots, r_{n,t})'$ denote a vector of n asset returns. A multivariate stochastic volatility (MSV) model with Student-t for r_t is

$$r_t = D_{x_t} + e_t \quad e_t \sim t(0, \sum_e, v), \quad (3.18)$$

$$y_t = \beta \exp\left(\frac{x_t}{2}\right) v_t, \quad v_t \sim t_v, \quad t = 1, \dots, n, \quad (3.19)$$

$$x_t = \phi x_{t-1} + w_t, \quad w_t \sim N(0, 1). \quad (3.20)$$

Let observation vector $y_t \in \mathfrak{R}^k$ follow a standardized Student's t distribution with v degrees of freedom. The variance covariance matrix of y_t is denoted by \sum_t . We assume that $v > 2$, such that the variance covariance matrix exists.

The observation density of y_t is given by

$$p(y_t | \sum_t; v) = \frac{\Gamma\left(\frac{v+k}{2}\right)}{[(v-2)\pi]^{\frac{k}{2}} \Gamma\left(\frac{v}{2}\right) |\sum_t|^{\frac{1}{2}}} e^{-\frac{\sum_{x_t}}{2} \left[1 + \frac{y_t' \sum_t^{-1} y_t}{v-2}\right]^{-\frac{v+k}{2}}} \quad (3.21)$$

Under this model, the conditional joint distribution of the returns r_t given x_t is multivariate student-t given by $r_t|x_t \sim t(0, DD' \exp(x_t) + \sum_t)$, where the time-series behavior of the conditional variance–covariance is driven by the volatility process of the common factor y_t . The relative importance of the factor for each of the returns considered is shown by considering the unconditional variance estimated from the model. This is compared with the corresponding sample variance [Engle (2002), Engle and Shephard (2001), and Tse and Tsui (2002)]. The unconditional variance–covariance matrix of r_t is given by $\text{var}(r_t) = DD' E[\exp(x_t)] + \text{diag}(\sigma_{t1}^2, \dots, \sigma_{tN}^2)$. Hence, the overall variance–covariance is decomposed into a component which is due to the variation in the common factor and a component reflecting the variation in the vector of noises. Following an interpretation offered by Diebold and Nerlove (1989), the common factor reflects the flow of new information relevant to the pricing of all assets, upon which asset specific shocks represented by the vector of noises are superimposed.

Again, the implementation of SMCem for the MSV model requires extending the conditional density of y_t given x_t in the baseline algorithm described in Section 3.3.

Given a sample of vector observations y_t with mean zero, $t = 1, \dots, n$, the log likelihood function for the multivariate Student's t model is given by

$$L = \sum_{t=1}^n \left\{ \log \left[\Gamma \left(\frac{v+k}{2} \right) \right] - \log \left[\Gamma \left(\frac{v}{2} \right) \right] - \frac{1}{2} \log |\sum_t| - \frac{k}{2} \log [(v-2)\pi] - \frac{(v+k)}{2} \log \left[1 + \frac{y_t' \sum_t^{-1} y_t}{(v-2)} \right] \right\}$$

In principle, estimation can be done using the same methods suggested for univariate models, although not each method may be applicable to every model. Still, SMCem

estimation technique appears to be flexible and efficient estimation technique for MSV models.

The EM algorithm maximizes the likelihood by iteratively carrying out an E-step and an M-step. In the E-step, the expectation

$$Q(\sum | \sum^{(k)}) = E_{\sum^{(k)}} (\log f(x | \theta') | y, \theta) \quad (3.22)$$

needs to be approximated, where $\sum^{(k)}$ is the current estimator. In the M-step, a new parameter estimate $\sum^{(k+1)}$ is obtained by maximizing $Q(\sum | \sum^{(k)})$. An approximation of $Q(\sum | \sum^{(k)})$ are computed based on the smoothing particles $\{\tilde{X}_{t-1}^{(i)}, \tilde{w}_{t-1}^{(i)}\}$ from our particle filter or from existing particle smoothing algorithms (Godsill et al., 2004 and Briers et al., 2010).

CHAPTER FOUR

4 DATA ANALYSIS AND RESULTS

4.1 SIMULATION STUDY

The models which pertain to parameter estimation for the student-t and GED SV model are illustrated using three simulated data sets and daily exchange rate series of the Nigerian Naira, Ghana Cedi, British Pound and Euro, all against the U. S. Dollars, from March 3, 2009 to March 3, 2011, to compare the fit of the distributions. In the simulation studies, one data set is generated from the SV model which has the normal mixtures distribution as its observation noise. The second data set is generated from the student-t SV model. The third data set is generated from the SV model with GED as its observation error. In each case, we carry out the SMCEM simulation with 500 particles and increase the number of the particles until we have achieved the tolerance $\varepsilon = 0.001$. The processes were stopped when the value of relative likelihood was less than tolerance, ε , which assesses convergence. To guarantee the processes had reached stability, we generated more than 11000 observations and discarded the initial 10000 samples.

Simulation 1: Data were generated from the normal mixture SV model

$$\begin{aligned}x_t &= 0.7x_{t-1} + w_t \\ y_t &= -2.75 + x_t + v_t\end{aligned}$$

where $w_t \sim N(0, 0.96)$, $v_t \sim I_t N(-2, 6) + (1 - I_t) N(-3.5, 4)$ and $I_t \sim \text{Bernoulli}(0.5)$. with true parameter set of $(\phi, \tau, q_0, q_1, R_0, R_1, \pi) = (0.7, 0.96, -3.5, -2, 4, 6, 0.5)$. We applied the techniques based on mixture and student-t SV to this data to examine the performance of the proposed model. To make this process stationary, we generated 11000

samples and discarded the first 10000 values. Figure 4.1 shows the plot and the histogram of simulation 1.

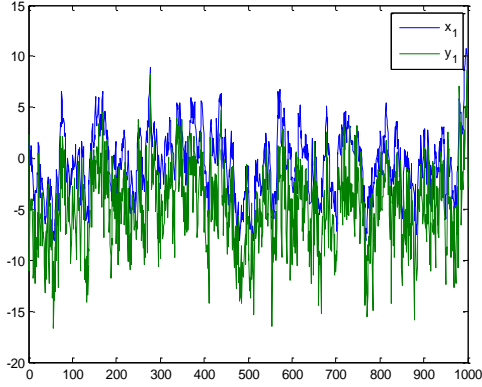
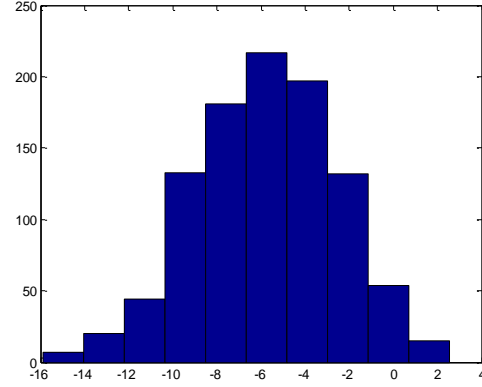


Fig. 4.1a: Representation of SMCEM sequence simulated from the normal mixture SV model.



4.1b Histogram of the final values of the parameters of the normal mixture SV model

Simulation 2: Data were generated from the student-t SV model with true parameter set of $(\phi, \tau, \alpha, \nu) = (0.81, 1.45, -3.01, 8)$. The techniques based on mixture and student-t SV model were applied to this data to see the merit of the student-t idea. Again, the length of the data, $\{y_t\}$, is 1000. Figure 4.2 below shows the plot and the histogram of simulation 2.

We used the second data set to observe the behavior of the estimation procedure when there is a departure from the normal mixture observational error assumption.

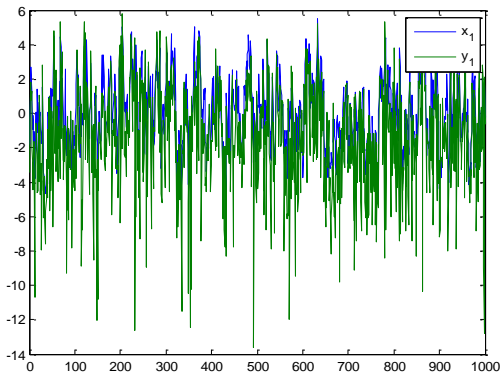
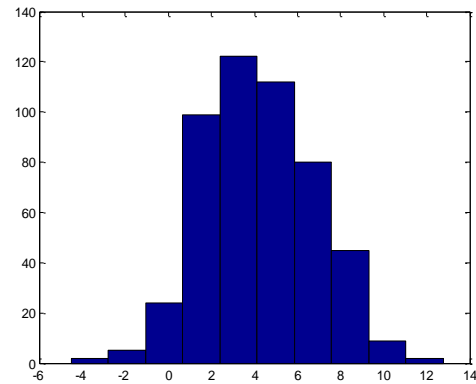


Fig. 4.2a: Representation of SMCEM sequence simulated from the student-t SV model.



4.2b Histogram of the final values of the parameters of the student-t SV model.

Simulation 3: Data were generated from the GED SV model with true parameter set of $(\phi, \tau, \alpha, \nu) = (0.9, 1.6, 0.7)$. We applied techniques based on mixture and the GED SV model to this data to see the merit of GED idea. Again, the length of the data, $\{y_t\}$, is 1000.

Figure 4.3 below shows the plot and the histogram of simulation 3.

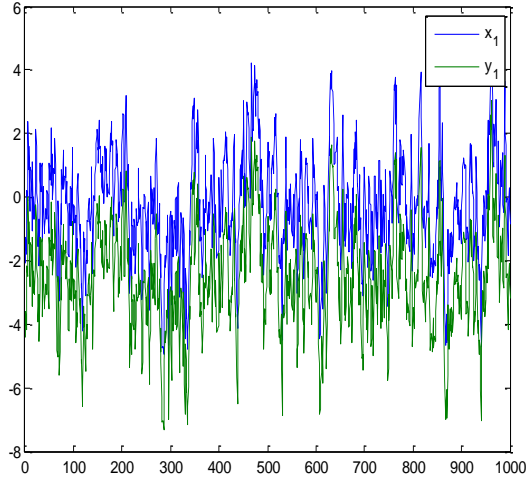
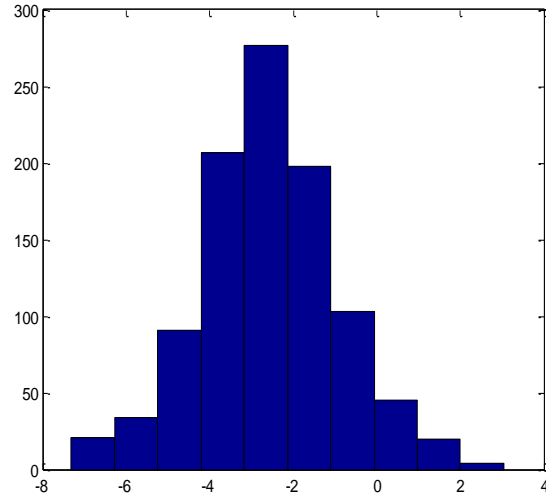


Fig. 4.3a: Representation of SMCEM sequence simulated from the normal mixture SV model.



4.3b Histogram of the final values of the parameters of the normal mixture SV model.

Figures 4.1b – 4.3b shows the SV for the three distributions. The results confirm that the student-t and the GED SV models, when compared with the normal mixture SV model, have fatter tails. The emphasis of the GED SV model is on the sharpness around the mean rather than the tail fatness, making the student-t to have fatter tail than that of the GED. As a result, student-t SV model attributes a larger proportion of extreme return values to ν_t instead of w_t than the SV normal and the SV GED. Figures 4.1a – 4.3a plot the volatilities jointly with the squared returns. No major difference in the volatilities among the three models.

4.2 ESTIMATION RESULTS

Utilizing the functions provided by MATLAB, Tables 4.1, 4.2, 4.3 and 4.4 show the results of the estimation for simulation 1, 2 and 3. At the beginning of each iteration, parameters are assumed to be known. For given parameters, the filtering step and a smoothing step were run. With the output of these two steps, updated parameter estimates were obtained by running the estimation step. In each simulation case, we started with 500 particles ($M = 500$) and $\varepsilon = 0.1$. The process was stopped as the relative likelihood at the next iteration was less the 0.1. We increased M and decreased ε , and repeated the procedure until we achieved the tolerance 0.001.

TABLE 4.1:
Parameter estimation on technique based on the normal mixture and Student-t on data generated from the normal mixture model

	$\phi^{(i)}$	$\tau^{(i)}$	$q_0^{(i)}$	$q_1^{(i)}$	$R_0^{(i)}$	$R_1^{(i)}$	$\pi^{(i)}$	$\alpha^{(i)}$	Rel. Lik	M	ε
True parameter	0.7	1.06	-3.5	-2	4	6	0.5	-2.75			
Mixture SV	0.7368 (0.028438)	0.9408 (0.014796)	-0.9826 (0.085676)	-3.5081 (0.037916)	2.2144 (0.38686)	7.5758 (1.0062)	0.4330 (0.017748)	-2.0761	0.0865		
Student-t SV	0.6976 (0.01475)	1.3654 (0.0029492)						-2.1465 (0.0038743)	0.0250	500	0.1
Mixture SV	0.8677 (0.048671)	0.1186 (0.0059242)	-1.8746 (0.084237)	-3.4744 (0.00879)	1.8883 (0.27135)	7.9490 (0.59001)	0.4267 (0.014455)	-1.983935	0.0025		
Student-t SV	0.7308 (0.049383)	0.7625 (0.20677)						-1.6703 (0.10626)	-	0.0021	1000 0.01
Mixture SV	0.7568 (0.027826)	0.3466 (0.00931)	-1.9486 (0.10989)	-3.7620 (0.08690)	2.3169 (0.20936)	7.5643 (0.67241)	0.3854 (0.02635)	-2.094344	0.0004		
Student-t SV	0.6913 (0.037981)	1.0336 (0.14839)						-1.9009 (0.024501)	-0.000	1000	0.001

Where ε is tolerance which assesses convergence, M is the number of particles, standard errors are shown in parenthesis, τ is the unconditional variance of x_t , ϕ is the estimated value between -1 and + to make the process stationary, $(q_0, q_1, R_0, R_1, \pi)$ are parameters in the normal mixture model which are estimated along with the other parameters (ϕ, τ)

By the procedure described in Section 3, [0.9500, 1.0729, -0.6794, -3.6794, 4.000, 4.000, and 0.5000] are selected for the initial parameters for $(\phi, \tau, q_0, q_1, R_0, R_1, \pi)$. Table 4.1 shows the final estimates alongside their standard error (in parenthesis) for simulation 1. The final estimates, along with their standard deviations (in parentheses), were $\hat{\phi} = 0.7568$ (0.027826),

$\hat{\tau} = 0.3466(0.00931)$, $\hat{q}_0 = -1.9486(0.10989)$, $\hat{q}_1 = -3.7620(0.08690)$, $\hat{R}_0 = 2.3169(0.20936)$, $\hat{R}_1 = 7.5643(0.67241)$, $\hat{\pi} = 0.3854(0.02635)$ where the true parameters are (0.7, 1.06, -3.5, -2, 4, 6, 0.5). In this approach, $\hat{\alpha} = \hat{\pi} \hat{q}_1 + (1 - \hat{\pi}) \hat{q}_0 = -2.6475$; $(\hat{\phi}, \hat{\tau}, \hat{\alpha}) = (0.7568, 0.3466, -2.6475)$. It can be said that the estimation procedure based on the normal mixture model works well in the sense that the estimates are close to the true parameters.

Based on the technique of student-t, we use (0.9500, 1.0729, -2.1496) as the initial values for the parameters, (ϕ, τ, α) . The process was stopped when the value of relative likelihood was less than 0.001. The final estimates, along with their standard deviations (in parentheses), were $\hat{\phi} = 0.6913(0.037981)$, $\hat{\tau} = 1.0336(0.14839)$, $\hat{\alpha} = -2.9009(0.024501)$. These results show that the model gives good estimates despite the fact that the true observation noise is not a normal mixture distribution.

A similar simulation study was performed using the data from simulation 2, and the results are presented in the table below:

TABLE 4.2:
Parameter estimation on technique based on the normal mixture and Student-t on data generated from the Student-t model

	$\phi^{(i)}$	$\tau^{(i)}$	$q_0^{(i)}$	$q_1^{(i)}$	$R_0^{(i)}$	$R_1^{(i)}$	$\pi^{(i)}$	$\alpha^{(i)}$	Rel. Lik	M	ε
True parameter	0.81	1.45						-3.01			
Mixture SV	0.6388 (0.021863)	1.2585 (0.029249)	-2.8256 (0.035245)	-5.7964 (0.011406)	3.8225 (0.075907)	4.6134 (0.14959)	0.4911 (0.0025628)	-4.28456	0.0643		
Student-t SV	0.8439 (0.005971)	1.2696 (0.02815)						-3.1243 (0.000596)	0.0045	500	0.1
Mixture SV	0.5601 (0.052847)	1.5805 (0.092768)	-2.8979 (0.02815)	-5.7544 (0.002826)	3.8111 (0.057768)	4.2066 (0.060831)	0.4911 (0.001607)	-4.300727	0.0057		
Student-t SV	0.8693 (0.002036)	1.4500 (0.037812)						-3.1645 (0.005662)	0.0042	1000	0.01
Mixture SV	0.6547 (0.005272)	1.2930 (0.002473)	-3.0180 (0.035241)	-5.8536 (0.012445)	3.3275 (0.15)	5.2663 (0.32564)	0.4806 (0.004338)	-4.3808	0.0009		
Student-t SV	0.8383 (0.008552)	1.5357 (0.12403)						-3.0912 (0.005302)	0.0010	1000	0.001

Where ε is tolerance which assesses convergence, M is the number of particles, standard errors are shown in parenthesis, τ is the unconditional variance of x_t , ϕ is the estimated value between -1 and + to make the process stationary, $(q_0, q_1, R_0, R_1, \pi)$ are parameters in the normal mixture model which are estimated along with the other parameters (ϕ, τ) .

The initial parameter set $[0.8214, 1.3359, -2.7823, -5.7823, 4.000, 4.000, \text{ and } 0.5000]$ are selected for the parameters $(\phi, \tau, q_0, q_1, R_0, R_1, \pi)$. Table 4.2 shows the results of the parameter estimation procedure based on the normal mixture. The final estimates, along with their standard deviations (in parentheses), were $\hat{\phi} = 0.6547$ (0.005272), $\hat{\tau} = 1.2930$ (0.002473), $\hat{q}_0 = -3.0180$ (0.035241), $\hat{q}_1 = -5.8536$ (0.012445), $\hat{R}_0 = 3.3275$ (0.15), $\hat{R}_1 = 5.2663$ (0.32564), $\hat{\pi} = 0.4806$ (0.004338) where the true parameters are (0.81, 1.45, -3.01) for the parameters, $(\hat{\phi}, \hat{\tau}, \hat{\alpha})$; where $\hat{\alpha} = \hat{\pi} \hat{q}_1 + (1 - \hat{\pi}) \hat{q}_0 = -4.3808$.

The following is the result, when we fit the data from simulation 2 with the techniques based on the student-t. (0.8214, 1.3359, -2.2823) were used as initial parameters of (ϕ, τ, α) . At 11th iteration, the relative likelihood was less than 0.001, and the process was considered converged. The final estimates, along with their standard deviations were $\hat{\phi} = 0.8383$ (0.008552), $\hat{\tau} = 1.5357$ (0.12403), $\hat{\alpha} = -3.0912$ (0.005302). These estimates are pretty similar to the true parameters (0.81, 1.45, -3.01), while Normal mixture returns (0.6547, 1.2930, -4.3808) as $(\hat{\phi}, \hat{\tau}, \hat{\alpha})$. The method based on the student-t SV model worked well in both cases. When the estimation procedure based on the normal mixture SV model was applied, the estimates were distant to the true parameter. On the other hand, the application of the technique based on student-t model indicated a better proximity to the true parameters. Therefore, extension of the SV model by adopting student-t is meaningful. Figure 4.4 - 4.7 shows the trajectory of the parameter estimates and the relative likelihoods.

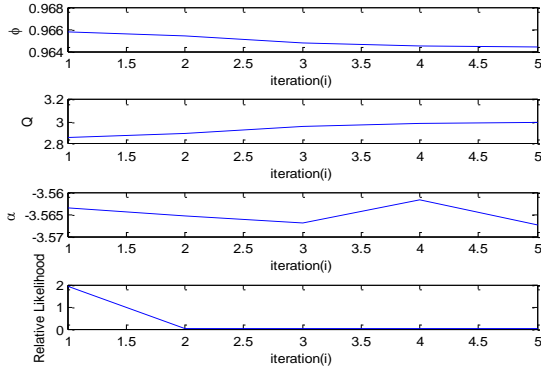


Figure 4.4: Parameter estimation results (1st to 3rd), and the relative likelihood (bottom) for technique based on student-t: for data generated from student-t model

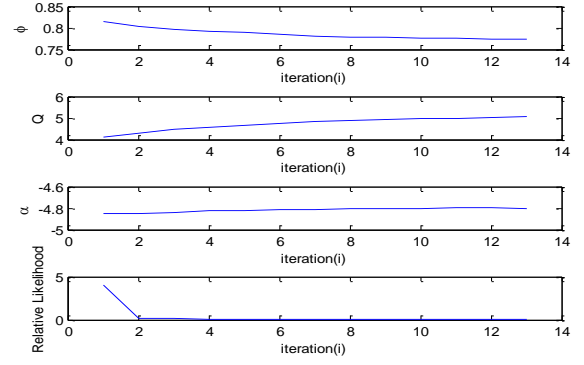


Figure 4.5: Parameter estimation results (1st to 3rd), and the relative likelihood (bottom) for technique based on student: for data generated from normal mixture model

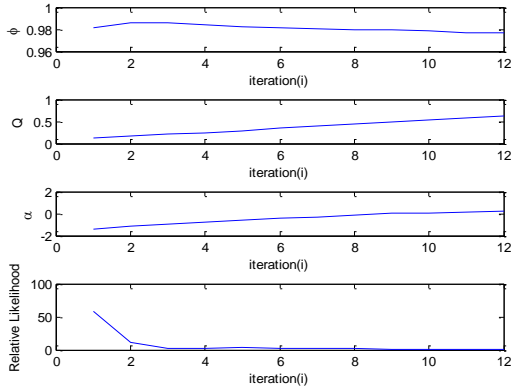


Figure 4.6: Parameter estimation results (1st to 3rd), and the relative likelihood (bottom) for technique based on normal mixture: for data generated from student-t model

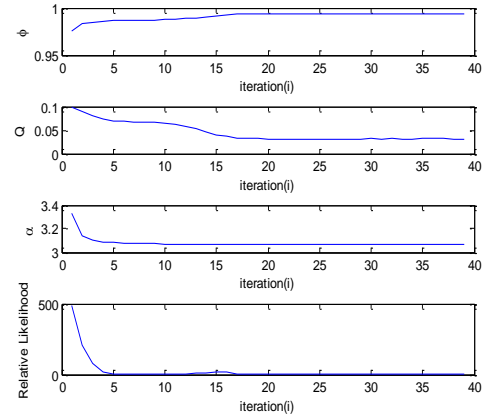


Figure 4.7: Parameter estimation results (1st to 3rd), and the relative likelihood (bottom) for technique based on normal mixture: for data generated from normal mixture model

Table 4.3 shows the results of the parameter estimation procedure on technique based on the normal mixture SV and GED on data generated from the normal mixture model. Figures 4.8 - 4.11 show the trajectory of the parameter estimates and the relative likelihoods. $[0.8699, 3.6899, -4.8897, -7.8897, 4.000, 4.000, \text{ and } 0.5000]$ are selected for the initial parameter for the parameters $(\phi, \tau, q_0, q_1, R_0, R_1, \pi)$. The final estimates, along with their standard deviations (in parentheses), were $\hat{\phi} = 0.9869$ (0.0013341), $\hat{\tau} = 4.1936(0.51468)$, $\hat{q}_0 = -4.3900(0.021891)$, $\hat{q}_1 = -6.4554(0.026309)$, $\hat{R}_0 = 3.7342(0.069579)$, $\hat{R}_1 = 4.3517(0.16268)$, $\hat{\pi} = 0.4895(0.0032833)$ where the true parameters are $(0.8, 3.5, -5, -8, 3, 4.1, 0.5)$. In this approach, $\hat{\alpha} = \hat{\pi} \hat{q}_1 + (1 - \hat{\pi}) \hat{q}_0 = -5.4010$; $(\hat{\phi}, \hat{\tau}, \hat{\alpha}) = (0.9869, 4.1936, -5.4010)$. It can be said that the estimation procedure based on the normal mixture model works well in the sense that the estimates are close to the true parameters.

Based on the technique of GED, we used (0.8699, 3.6899, -4.4365,) as the initial values for the parameters, (ϕ, τ, α) . Table 4.3 shows the results of the estimation procedure. The process was stopped when the value of relative likelihood was less than 0.001. The final estimates, along with their standard deviations (in parentheses), were $\hat{\phi} = 0.8127(0.0027854)$, $\hat{\tau} = 4.2368(0.020652)$, $\hat{\alpha} = -4.7144(0.0083309)$. This results show that the GED model gives good estimates even though the true observation noise is not a normal mixture distribution.

TABLE 4.3:
Parameter estimation on technique based on the normal mixture SV and GED on data generated from the normal mixture model

	$\phi^{(i)}$	$\tau^{(i)}$	$q_0^{(i)}$	$q_1^{(i)}$	$R_0^{(i)}$	$R_1^{(i)}$	$\pi^{(i)}$	$\alpha^{(i)}$	Rel. Lik	M	\mathcal{E}
True parameter	0.8	3.5	-5	-8	3	4.1	0.5	-5.55			
Mixture SV	0.7875 (0.0045855)	3.4771 (0.17068)	-5.5322 (0.0059081)	-8.5269 (0.00932)	3.6937 (0.10267)	4.4847 (0.16612)	0.4895 (0.003056)	-6.9980	0.0248		
GED SV	0.8485 (0.0059856)	4.0273 (0.13077)						-4.3205 (0.035379)	0.0853	500	0.1
Mixture SV	0.8425 (0.0046651)	2.9591 (0.17631)	-4.8080 (0.026411)	-7.9658 (0.032108)	3.7054 (0.081339)	4.1121 (0.057281)	0.4930 (0.0025453)	-6.3648	-0.0148		
GED SV	0.8788 (0.0037027)	3.0502 (0.052579)						-4.2576 (0.003807)	0.0064	1000	0.01
Mixture SV	0.9869 (0.0013341)	4.1936 (0.51468)	-4.3900 (0.021891)	-6.4554 (0.026309)	3.7342 (0.069579)	4.3517 (0.16268)	0.4805 (0.0032833)	-5.3824	-0.0684		
GED SV	0.8127 (0.0027854)	4.2368 (0.020652)						-4.7144 (0.0083309)	0.0007	1000	0.001

Where \mathcal{E} is tolerance which assesses convergence, M is the number of particles, standard errors are shown in parenthesis, τ is the unconditional variance of x_t , ϕ is the estimated value between -1 and + to make the process stationary, $(q_0, q_1, R_0, R_1, \pi)$ are parameters in the normal mixture model which are estimated along with the other parameters (ϕ, τ)

Table 4.4 shows the results of the parameter estimation procedure on technique based on the normal mixture SV and GED on data generated from GED SV model. The method based on the GED model works well in both cases. When the estimation procedure based on the normal mixture SV model was applied, the estimates were quite far from the true parameters.

On the other hand, the application of the technique based on GED model indicated a better proximity to the true parameters. (0.9500, 1.3288, 0.6309) were used as initial parameters of (ϕ, τ, α) . The final estimates, along with their standard deviations were $\hat{\phi} = 0.9749(0.0026845)$, $\hat{\tau} = 2.3496(0.2678)$, $\hat{\alpha} = 0.6821(0.014247)$. These estimates are pretty similar to the true

parameters (0.9, 1.6, 0.7), while normal mixture returns (0.7515, 2.3496, 1.3259732) as (ϕ, τ, α) .

Therefore, the method based on the GED idea works well in both cases.

TABLE 4.4:

Parameter estimation on technique based on the normal mixture SV and GED on data generated from the GED model

	$\phi^{(i)}$	$\tau^{(i)}$	$q_0^{(i)}$	$q_1^{(i)}$	$R_0^{(i)}$	$R_1^{(i)}$	$\pi^{(i)}$	$\alpha^{(i)}$	Rel. Lik	M	ε
True parameter	0.9	1.6						0.7			
Mixture SV	0.9050 (0.015127)	0.3136 (0.01557)	0.1491 (0.08715)	-2.4081 (0.1269)	2.9473 (0.2282)	7.8467 (0.84311)	0.4448 (0.015076)	-0.9883	0.0052		
GED SV	0.9770 (0.0021463)	2.1311 (0.2265)						0.7654 (0.018455)	0.0248	500	0.1
Mixture SV	0.8871 (0.026021)	0.4605 (0.025703)	-0.3568 (0.11521)	-2.6609 (0.13263)	2.5108 (0.28846)	9.4387 (1.1097)	0.3916 (0.028448)	1.2590	0.0125		
GED SV	0.9754 (0.0025267)	2.3108 (0.26119)						0.7627 (0.021224)	0.0057	1000	0.01
Mixture SV	0.7515 (0.015215)	1.2777 (0.069801)	-0.3193 (0.088286)	-3.1988 (0.15449)	1.6365 (0.47054)	8.7981 (1.1994)	0.3496 (0.041351)	-1.3259732	0.0003		
GED SV	0.9749 (0.0026845)	2.3496 (0.2678)						0.6821 (0.014247)	0.0007	1000	0.001

Where ε is tolerance which assesses convergence, M is the number of particles, standard errors are shown in parenthesis, τ is the unconditional variance of x_t , ϕ is the estimated value between -1 and +1 to make the process stationary, $(q_0, q_1, R_0, R_1, \pi)$ are parameters in the normal mixture model which are estimated along with the other parameters (ϕ, τ)

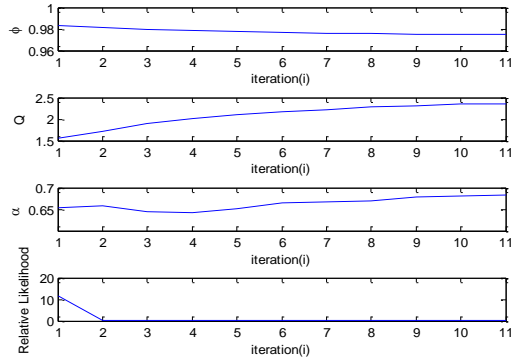


Figure 4.8: Parameter estimation results (1st to 3rd), and the relative likelihood (bottom) for technique based on GED: for data generated from GED model

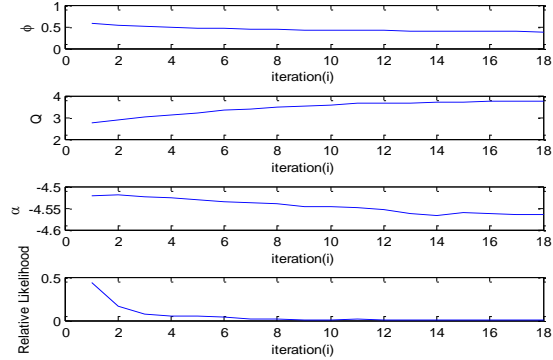


Figure 4.9: Parameter estimation results (1st to 3rd), and the relative likelihood (bottom) for technique based on GED: for data generated from normal mixture model

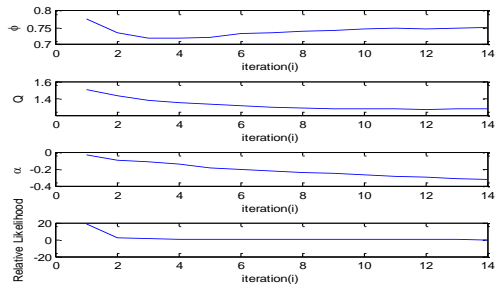


Figure 4.10: Parameter estimation results (1st to 3rd), and the relative likelihood (bottom) for technique based on normal mixture: for data generated from GED model

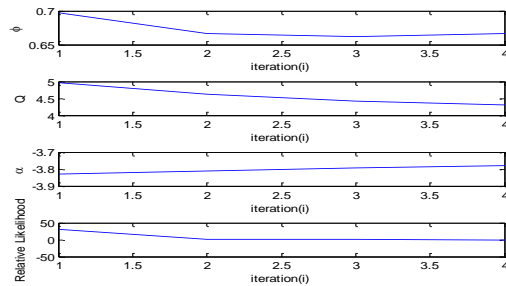


Figure 4.11: Parameter estimation results (1st to 3rd), and the relative likelihood (bottom) for technique based on normal: for data generated from normal mixture model.

4.3 APPLICATION TO REAL LIFE FINANCIAL DATA

Volatility which is one of the most important motivations for the development of SV model is most obvious in exchange rate data as exchange rate typically exhibit high degree of volatility. We apply the normal mixture, student-t and GED SV model to analyze daily rates on the Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar exchange rates from March 3rd, 2009 to March 3rd, 2011. Figure 4.12 - 4.15 shows the plots of the daily exchange rates and log returns of the data.

A log return of exchange rates is

$$r_t = 100 \ln \frac{p_t}{p_{t-1}} \quad (4.1)$$

where p_t is the exchange rate at time index t .

The descriptive statistics are summarized in Table 4.5. Skewness and kurtosis are observed, leading to a high valued Jarque and Bera (1987) test which indicates non-normality of the distribution. The kurtosis of the returns of the series is significantly above three, indicating leptokurtic return distributions.

Table 4.5 Descriptive statistics of Daily Returns for the Exchange Rate

Statistics	Naira/Dollar rate	Cedi/Dollar	Pound/Dollar	Euro/Dollar
Mean	-0.001385	-0.006258	0.020803	0.001676
Std. Dev.	0.708650	0.536507	0.506541	0.488392
Skewness	-0.074139	0.966923	0.022958	0.434943
Kurtosis	8.805879	13.11769	4.262290	7.993814
Jarque-Bera	735.0376	2312.255	34.76827	559.9343

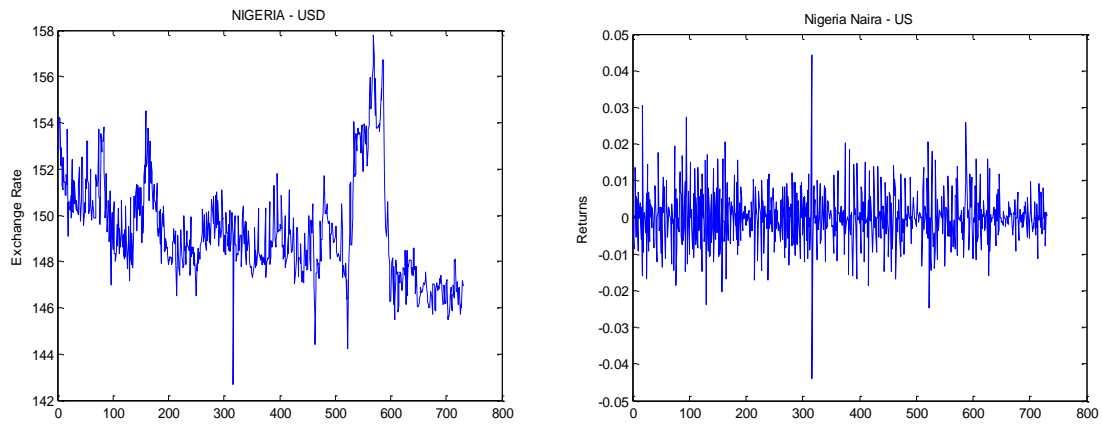


Fig. 4.12 Naira/dollar daily exchange rate and the log returns

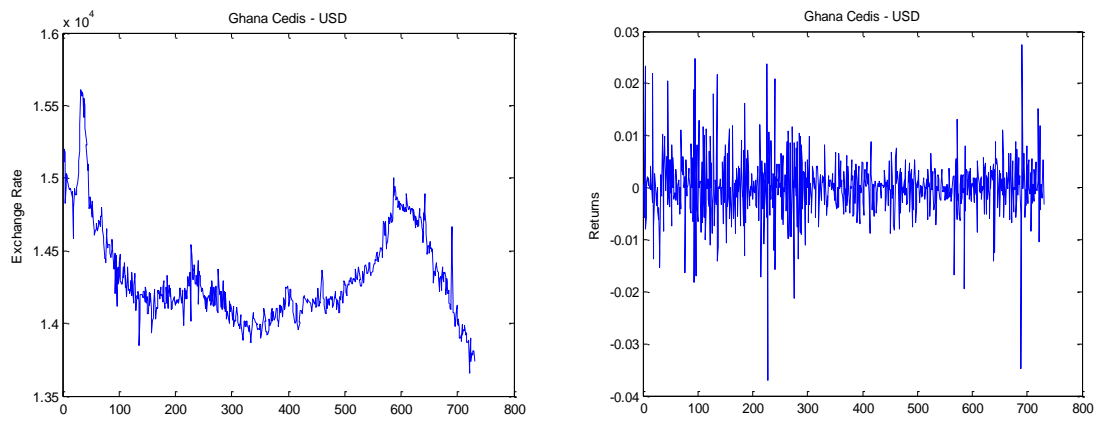


Fig. 4.13 Cedi/dollar exchange rate and the log returns

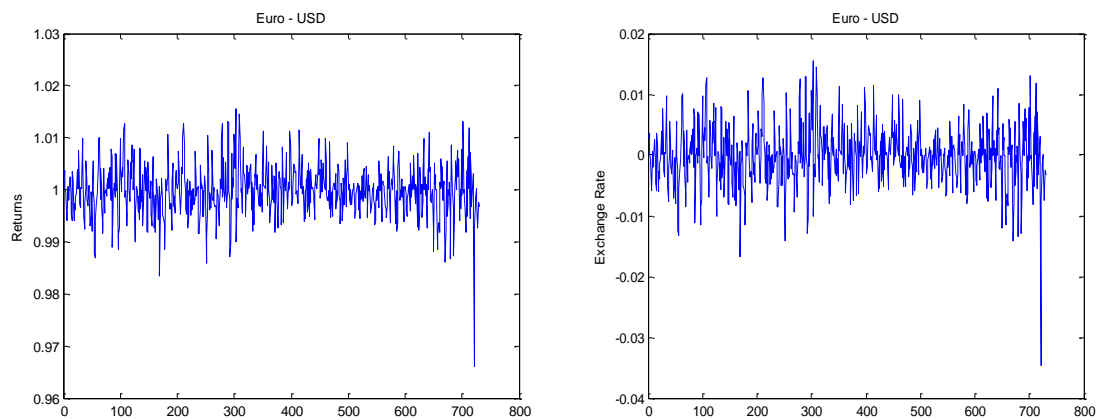


Fig. 4.14 Euro/dollar daily exchange rate and the log returns

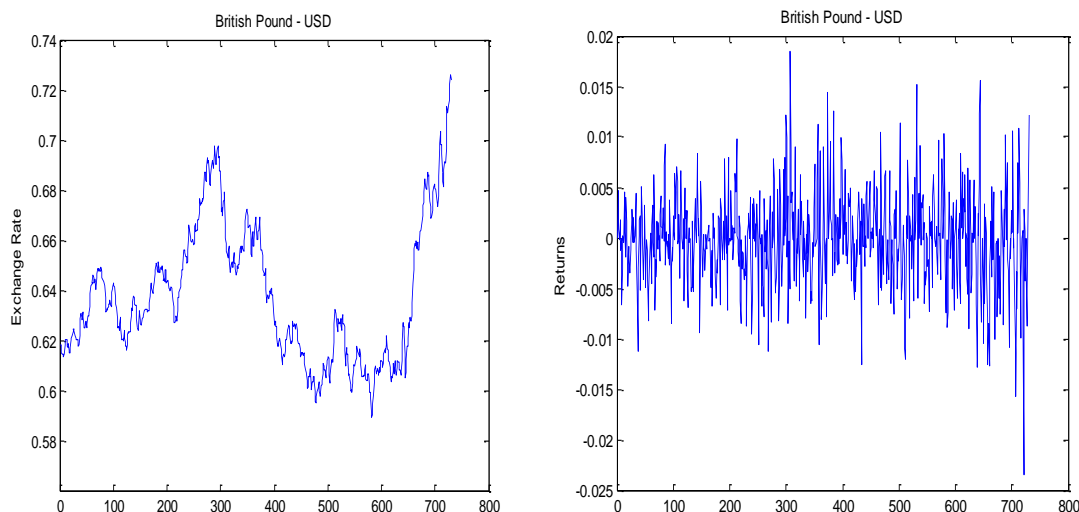


Fig. 4.15 Pound/dollar daily exchange rate and the log returns

Some patterns of behavior are evident in the second plot of each of Figure 4.12-4.15: the data experience a small variance for some periods of time, and for other periods, they show a large variance. Therefore, it cannot be assumed that the data have a constant variance.

4.3.1 Estimation Results

Utilizing the functions and the means of various routines written in Matlab, the SMCEM were applied to the data, parameters are estimated using the techniques described in the previous sections. Table 4.6 presents the estimation results along with their standard deviations in parenthesis, for the student-t, normal mixture and the GED SV model. These distributions produce comparable maximum likelihood values, indicating an acceptable overall fit. The values (ranging from 0.927 to 0.988) suggest high persistence of the volatility of the series indicating that volatility clustering is observed in all the exchange rates return series. Figure 4.16 - 4.19 shows the trajectory of parameter estimates and the history of the relative likelihood.

Table 4.6: Estimation Results - Distributions Comparison

Naira/Dollar	ϕ	τ	q_0	q_1	R_0	R_1	π	α
Normal								
Mixture SV	0.9759 (0.0041766)	0.0988 (0.019774)	3.0673 (0.043267)	1.6633 (0.65851)	0.6954 (0.11387)	1.9713 (0.88689)	0.0546 (0.017691)	-2.0731
Student-t SV	0.9869 (0.0029595)	0.1300 (0.16604)						-2.1245 (0.5369)
GED SV	0.9684 (0.015815)	0.3227 (0.030638)						-2.8836 (0.35001)
Cedi/Dollar	ϕ	τ	q_0	q_1	R_0	R_1	π	α
Normal								
Mixture SV	0.9830 (0.0049939)	0.0902 (0.033484)	3.4686 (0.015039)	0.9826 (0.22583)	0.1865 (0.015231)	2.3618 (0.10685)	0.0705 (0.006188)	-4.28356
Student-t SV	0.9887 (0.00078342)	0.0854 (0.0057923)						-3.2243 (0.0019277)
GED SV	0.9741 (0.01001)	0.09831 (0.048961)						-4.9834 (0.0022972)
Pound/Dollar	ϕ	τ	q_0	q_1	R_0	R_1	π	α
Normal								
Mixture SV	0.9895 (0.0013354)	0.7114 (0.36943)	3.3388 (0.022015)	0.4351 (0.023938)	0.7835 (0.056033)	4.4129 (0.14603)	0.4941 (0.0028666)	-4.28356
Student-t SV	0.9754 (0.0025267)	0.7627 (0.26119)						2.3108 (0.021224)
GED SV	0.9697 (0.0084966)	0.9018 (0.183815)						1.9588 0.0082487
Euro/Dollar	ϕ	τ	q_0	q_1	R_0	R_1	π	α
Normal								
Mixture SV	0.9279 (0.0090241)	0.4170 (0.24553)	-1.9638 (0.21412)	-4.3710 (0.54791)	4.1677 (1.1257)	8.1096 (0.4471)	0.4404 (0.03083)	-4.28356
Student-t SV	0.9513 (0.022502)	0.4731 (0.45638)						-2.2692 (0.17681)
GED SV	0.9283 (0.0021063)	0.7258 (0.361172)						-3.5033 (0.058033)

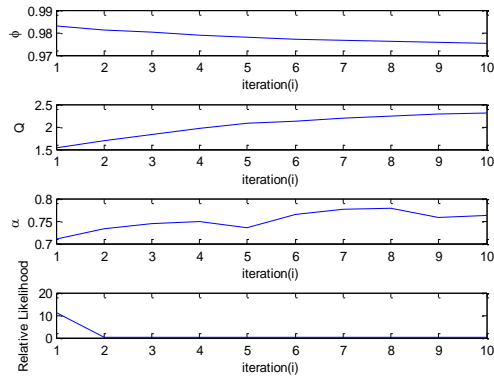


Fig. 4.16: Estimation results for the naira/dollar exchange rates

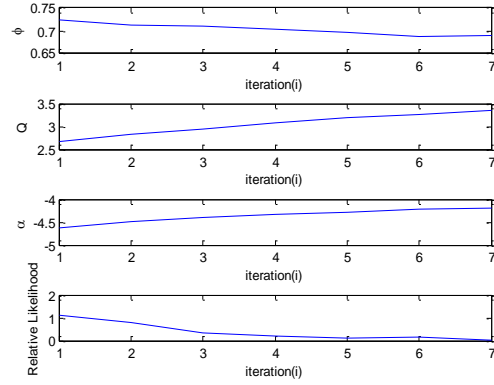


Fig. 4.17: Estimation results for the cedi/dollar exchange rates

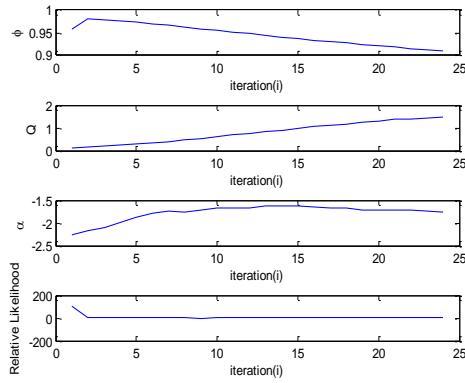


Fig. 4.18: Estimation results for the Euro/dollar exchange rates

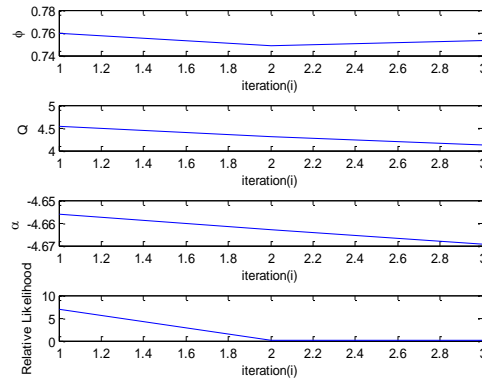


Fig. 4.19: Estimation results for the pound/dollar exchange rates

4.4 AKAIKE INFORMATION CRITERION BASED MODEL SELECTION CRITERIA

Model comparison in SMCEM framework can be performed using the Akaike information Criterion (AIC). In general, the AIC is designed to select a model that best describes a set of observed data from a collection of potential probability models. If a potential probability model is characterized by a parameter vector θ of length k which has log-likelihood $l(\theta)$, the AIC is defined by

$$AIC = 2k - 2l(\hat{\theta}), \quad (4.2)$$

where $\hat{\theta}$ is the maximum likelihood estimate of θ . The AIC uses the observed data to simultaneously estimate the parameters of the model by the method of maximum likelihood and calculate the model selection criteria. The model with the smallest value of the AIC is the most desirable. It suffices to note that while the method of maximum likelihood gives an unbiased estimate $\hat{\theta}$ of the parameters θ , the value of the loglikelihood when evaluated at the estimated parameter vector $\hat{\theta}$ produces a biased estimate of $l(\theta)$ (Ibrahim et al., 2008). The term $2k$ in the definition of the AIC is used to correct the bias in the estimated maximum log likelihood $l(\hat{\theta})$. To evaluate the AIC using the EM algorithm, a different definition of the AIC is used, as the EM algorithm actually estimates the expected log likelihood $Q(\theta|\theta')$; this expected log-likelihood is used in defining the AIC. Furthermore, using the expected log likelihood in the new definition of the AIC implies that the bias correction required is different from that in equation (4.2).

In equation (4.2), the term $2k$ is added to make this statement true of the AIC. However, the expected log likelihood $Q(\theta|\theta')$ when evaluated at the maximum likelihood estimate $\hat{\theta}$ provides an unbiased estimate of the mean expected log likelihood. The bias of the expected log likelihood is zero.

Thus for the EM algorithm

$$AIC = -2Q(\theta, \theta') \quad (4.3)$$

On the other hand, comparing normal density with non-normal ones has also been explored in a number of cases [Cairns et al., (2009), Rossi and Gallo (2006), Amisano and Giacomini (2007), Maheu et al., (2010)]. The evaluation of volatility forecasting models is complicated by the fact that the volatility is not observable, even ex post. However, there exist conditionally unbiased

proxies of the unobserved variance process. In general, the proxies used are of three types: daily squared returns, intra daily range and realized volatility.

Basically, statistical measures evaluate the difference between forecasts at time t and realized values at time $t + k$. However, asset price volatility is not directly observable and measuring the realized values of volatility is quite tasking. The daily squared return has been extensively used in the literature as ex-post volatility. Nonetheless, Andersen and Bollerslev (1998) show that it is a very noisy volatility estimator and does not present reliable inferences regarding the underlying latent volatility in daily samples. A new volatility measure, known as realized volatility is introduced, which estimates volatility by summing squared intraday returns. Hence, the volatility of a price process is treated as an observable process. Until recently, the SV models have focused on using either squared returns or absolute returns as a proxy to the true but unobservable volatility.

Here, the forecasting performance for the student-t, mixture and the GED.SV model are presented. The forecasts are made on the four exchange rate series. The forecast summary statistics include well known measures like mean squared error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE).

$$\text{MSE} = \frac{\sum_{t=1}^n (\hat{\sigma}_t^2 - \sigma^2)^2}{n}, \text{MAE} = \frac{1}{n} \sum_{t=1}^n |\hat{\sigma}_t^2 - \sigma^2|, \text{MAPE} = \frac{\sum_{t=1}^n |\hat{\sigma}_t^2 - \sigma^2| / \sigma^2}{n}$$

where $\hat{\sigma}^2$ is the forecasted variance and σ^2 the actual variance time period t and n is the number of forecasts.

The Akaike values and the evaluation statistics using the entire data material are presented in Table 4.7. The Akaike information criteria and the log-likelihood values highlight the fact that (GED) student-t distribution better estimates the series than the normal mixture distribution for the SV model. Indeed, the log-likelihood function increases, leading to AIC criteria of 2.805, 3.4593, 3.9989 and 9.6632 with the normal mixture versus (2.7814304) 2.736433, (3.4223827) 3.321374, (3.9749741) 3.869968 and (9.6513786) 9.446376 with the non normal densities, for the Naira/Dollar, Cedi/Dollar, Pound/Dollar and Euro/Dollar rate respectively. The evaluation statistics from the volatility forecasts, Sadorsky (2005), are presented. In terms of MSE, student-t performs better than the normal mixture for the Naira/Dollars and the Euro/Dollar exchange rate while the opposite is true for the Cedi/Dollars and Pound/Dollar exchange rate. Generally, the MAE results are not very different from the MSE results. In terms of MAPE, the student-t SV model is preferred in three cases and the GED SV model once.

Table 4.7: Evaluation Statistics - Distributions Comparison

Naira/Dollar	AIC	Log-like	MSE	MAE	MAPE
Normal					
Mixture SV	2.805	2537.403	0.2601914	0.210278	0.233428
Student-t SV	2.736433	2565.97	0.1111914	0.075379	0.353437
GED SV					
	2.781430	2563.16	0.1131914	0.081256	0.264035
Cedi/Dollar	AIC	Log-like	MSE	MAE	MAPE
Normal					
Mixture SV	3.4593	2652.836	0.037074	0.059524	0.147437
Student-t SV	3.321374	2769.66	0.121086	0.171621	0.226563
GED SV	3.4223827	2760.56	0.16377	0.188346	0.131424

Pound/Dollar	AIC	Log-like	MSE	MAE	MAPE
Normal					
Mixture SV	3.9989	2801.144	0.151623	0.072823	0.192232
Student-t SV	3.869968	2869.07	0.178377	0.098361	0.345245
GED SV	3.974974	2857.27	0.184677	0.103932	1.1434675
Euro/Dollar	AIC	Log-like	MSE	MAE	MAPE
Normal					
Mixture SV	9.6632	2874.968	0.170612	0.260176	0.3988
Student-t SV	9.446376	2903.9	0.095612	0.194364	0.5578
GED SV	9.6513786	2897.8	0.101414	0.196355	0.2634

The AIC picks the model with the highest likelihood. We could select our model using log-likelihood instead of the AIC. The model with the smallest AIC corresponds to the model with highest log likelihood, and so a model with a high likelihood (good fit) will make the AIC small (more negative). The MSE, MAE, and MAPE measure the magnitude of the forecast errors.

4.4.1

Table 4.8 SMCEM estimation results for the MSV model

	NN	Cedi	BP	Euro
d_j	0.8754(0.03)	1.5175 (0.0201)	0.5463(0.0349)	0.7815(0.03811)
$\sigma_{e,j}$	0.7750 (0.01)	0.3050 (0.0070)	0.8737 (0.0069)	0.4893 (0.011)
	ϕ	τ	α	log-likelihood
	0.9777 (0.0065)	0.1587 (0.0176)	0.6857 (0.0236)	18,271.6

The data to which we fit this multivariate model consists of the daily exchange rates: Nigeria Naira/US-Dollar, Ghana Cedi/Dollar, British Pound/US-Dollar (BP), and Euro/Dollar.

The SMCEM results for the multivariate SV model (3.18) - (3.20) are summarized in Table 4.8.

All parameter estimates are numerically reasonable as indicated by the small standard deviations (in parentheses). The estimates of the factor loadings d_j indicate that the Cedi currencies load more on the common factor than the others. Moreover, the estimated volatility parameters of the

factor are similar in magnitude to the those obtained under univariate SV models, and the estimate of ϕ , which is close to one, implies that the common factor exhibits a strongly persistent volatility process. The common factor explains 76%, 90%, 52% and 61% of the overall variation in the returns of Naira, Cedi, Pound and Euro, respectively. In addition, the log-likelihood of the multivariate model is 18,271.6, which is considerably larger than the sum of the likelihood values obtained under the four independent SV models which equals 11,108.6. This significant difference reflects the fact that, in contrast to the univariate specifications, the multivariate model can account for the correlation between the returns. Using the parameter estimates, we computed the corresponding estimate of the unconditional variance–covariance matrix,

Table 4.9 unconditional variance–covariance matrix of the returns

Variables	NGN/USD	BP/USD	GHC/USD	EUR/USD
NGN/USD	0.428222	0.0259	0.01118	0.01035
BP/USD		0.119209	0.03821	0.0509
GHC/USD			0.2043	0.07523
EUR/USD				0.215

to be compared with the sample variance– covariance matrix of the returns.

Table 4.10 sample variance– covariance matrix of the returns

Variables	NGN/USD	BP/USD	GHC/USD	EUR/USD
NGN/USD	0.558222	0.0433	0.06892	0.02525
BP/USD		0.216509	0.03539	0.0659
GHC/USD			0.342	0.02187
EUR/USD				0.39

respectively. The two matrices are quiet similar. However, the diagonal elements from our model are smaller in each case than those of the sample variance indicating that there is more volatility in the data than the model accounts for.

4.5. APPLICATION TO BEARINGS - ONLY TRACKING

In this section another application of the proposed estimation technique developed in section 3 is considered. The scenario considered is the nonlinear bearing-only tracking (BOT) problem. In the BOT problem, the objective is to track the movements of an object using measurements of the relative angle to the object. Target tracking applications have proved to be an active area of research among scholars [see Liu et al., (2002), Bar-Shalom et al., (2001), Bergman (1999), Bar-Shalom and Fortmann (1988) and Mazor et al., (1996)]. It is an important component of many modern applications, such as robots localization, visual tracking, radar tracking, and satellite navigation. The key to a successful target tracking depends on an effective extraction of the useful information about the target state from available observations.

The commonly studied estimation technique tracking models relies upon the Kalman filter or, for non-linear systems the extended Kalman filter. Recently, the growth in computational power has made computer intensive statistical methods feasible. The main breakthrough came with the seminar paper of Gordon et al., (1993). Both Markov Chain Monte Carlo methods and sequential Monte Carlo methods (particle filters) are now implemented in several applications. Several writers including Gordon et al., (1993) and Pitt and Shephard (1997) have implemented particle filters for simple tracking models.

Bearings-only tracking involves estimating the target states based on angle measurements at a antenna node. The target is assumed to move in the x - y plane and to follow a constant-velocity motion model, Bar-Shalom and Fortmann (1988), with a state update period of 1s.

The BOT problem is illustrated below:

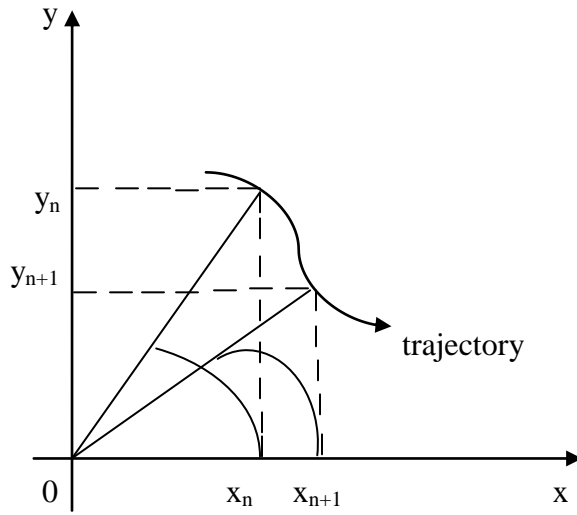


Figure 4.20: The Bearings-Only Tracking (BOT) problem.

In addition, BOT involves tracking of an object (typical examples include ships, planes and other moving vehicles) based upon its measured bearing (angle) with respect to a fixed sensor. The system state in BOT is 4 dimensional and includes the positions and velocities along the Cartesian x and y directions. The positions are expressed with reference to the fixed sensor location. The state equation in the BOT problem describes the system kinematics modelled with a Constant Velocity model. The measurement is the bearing or angle of the object with respect to the sensor axis (y axis). The system model that we describe here is similar to that used in Gordon et al., (1993) and Chapter 6 of Ristic et al., (2004) [see also Pitt and Shephard (1997) and Carpenter et al., (1999)]. The state vector X_t contains the positions and velocities of the object in the x - y directions respectively : $X_t = (x_t, y_t, \dot{x}_t, \dot{y}_t)^T$, where $T = 1\text{sec}$ denotes the sampling period. One possible discretization of this model, is given by Gordon et al., (1993)

$$X_t = \phi x_{t-1} + \Gamma u_t \quad t = 1, \dots, N \quad . \quad (4.4)$$

The equation of the observed bearing, z_t , is

$$z_t = \tan^{-1}\left(\frac{y_t}{x_t}\right) + w_t \quad (4.5)$$

$$\text{where } u_t = (u_x, u_y)^T, \phi = \begin{pmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{pmatrix}, \Gamma = \begin{pmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \\ 0 & T \end{pmatrix}$$

Parameter u_t represents the system noise and is Gaussian distributed with covariance $\sum_u = \sigma_u^2 I_2$ where I_2 is a 2×2 identity matrix. This vector can be thought of as representing the acceleration in the x and y directions. z_t represents the observed bearing of the object measured by the sensor at time t . w_t represents a Gaussian measurement noise with mean zero and variance σ_w^2 . Before measurements are taken, the particle filter recursion is started with initial state vector in the form of a 4 dimensional Gaussian variable with known mean and covariance matrix. As can be seen, this model is 4 dimensional and nonlinear due to a transcendental function in the observation equation. The particle filters handle these situations efficiently. It has been shown through intense simulations in Gordon et al., (1993), that particle filters are much more efficient for this problem than the traditional EKF.

4.5.1 SMCEM for Bearings-Only Tracking

Using equation (4.5) and (4.5), the proposed estimation technique for target tracking is applied. The state update is used to propose new particles. This provides a sub-optimal recursive estimate of the target position in the $x - y$ plane.

Previous approaches to density estimation have mostly focused on Gaussian nonlinear measurement dynamics in practice. The SMCEM algorithm is developed based on non-Gaussian, which are heavier tailed than Gaussians and hence more robust. Simulations have demonstrated not only the effectiveness but also the improved performance of the non-Gaussian distribution over Gaussian.

As we observed the target in its motion, new data z_t accrue, along with new parameters (\dot{x}_t, \dot{y}_t) .

The vector of unknown at time t is

$$\theta_t = (x_1, y_1, \dot{x}_1, \dot{y}_1, \dots, \dot{x}_t, \dot{y}_t) ,$$

and the data are $z_{1:t} = (z_1, \dots, z_t)$. Therefore, the target distribution evolves in an expanding space, Ω_t . As t increases, the aim here is to maintain a set of sampled particles in Ω_t which can be used to estimate aspect of the distribution of interest. In particular, these particles can be used at any given time point t to approximate the conditional distribution for the current state of the object, given the data, $z_{1:t}$ accumulated up to that point. The procedure for the BOT problem is summarized below.

Given the observed data z_t at t

For $i = 1, 2, \dots, N$ sample particles, $X_t^{(i)}$ are drawn from the density

$$X_t^{(i)} \sim p(X_t^{(i)} | X_{t-1}^{(i)}) \text{ using (4.4).}$$

$$x_t^{(i)} = x_{t-1}^{(i)} + \dot{x}_{t-1}^{(i)} + u_{x_t}^{(i)}$$

$$\dot{x}_t^{(i)} = \dot{x}_{t-1}^{(i)} + u_{\dot{x}_t}^{(i)}$$

$$y_t^{(i)} = y_{t-1}^{(i)} + \dot{y}_{t-1}^{(i)} + u_{y_t}^{(i)}$$

$$\dot{y}_t^{(i)} = \dot{y}_{t-1}^{(i)} + u_{\dot{y}_t}^{(i)}$$

the weights are updated recursively using

$$\hat{w}_t^{(i)} = w_{t-1}^{(i)} p(z_t | x_t^{(i)}, \dot{x}_t^{(i)}, \dot{y}_t^{(i)}) \quad \text{where } z_t | x_t^{(i)} \sim t \text{ distribution } (v)$$

we evaluate this distribution at time t for the parameters estimation by using the EM algorithm and SMC,
and then calculate the output:

$$\begin{aligned}\hat{x}_t &= \sum_{i=1}^N w_t^{(i)} x_t^{(i)} , \\ \hat{\dot{x}}_t^{(i)} &= \sum_{i=1}^N w_t^{(i)} \dot{x}_t^{(i)} , \\ \hat{y}_t &= \sum_{i=1}^N w_t^{(i)} y_t^{(i)} , \\ \hat{\dot{y}}_t^{(i)} &= \sum_{i=1}^N w_t^{(i)} \dot{y}_t^{(i)} ,\end{aligned}$$

The mean estimate of the target state and the covariance matrix of the estimate error are approximated by

$$\begin{aligned}\mu_t &= \sum_{i=1}^N w_t^{(i)} x_t^{(i)} , \\ \sum_u &= \sum_{i=1}^N w_t^{(i)} (x_t^{(i)} - \mu_t)(x_t^{(i)} - \mu_t)^T .\end{aligned}$$

4.5.2 Tracking Performance

An in-depth analysis of the tracking performance for the BOT problem is investigated in various works including Gordon et al., (1993) and Doucet et al., (2001). A bearings-only tracking problem is simulated to present implementation of the proposed estimation technique. A target trajectory and associated measurements is generated according to equations (4.4) and (4.5) with the parameter values

$$\sqrt{\sigma_u^2} = 0.001, \sqrt{\sigma_w^2} = 0.005$$

the initial state of the target

$$X_0 = (5, 5, 1, 0)^T$$

and covariance = $diag(7.5, 2.55, 5.5, 2.7)$.

The time between successive measurement is $T = 1s$ and a single bearing measurement is obtained in each time step.

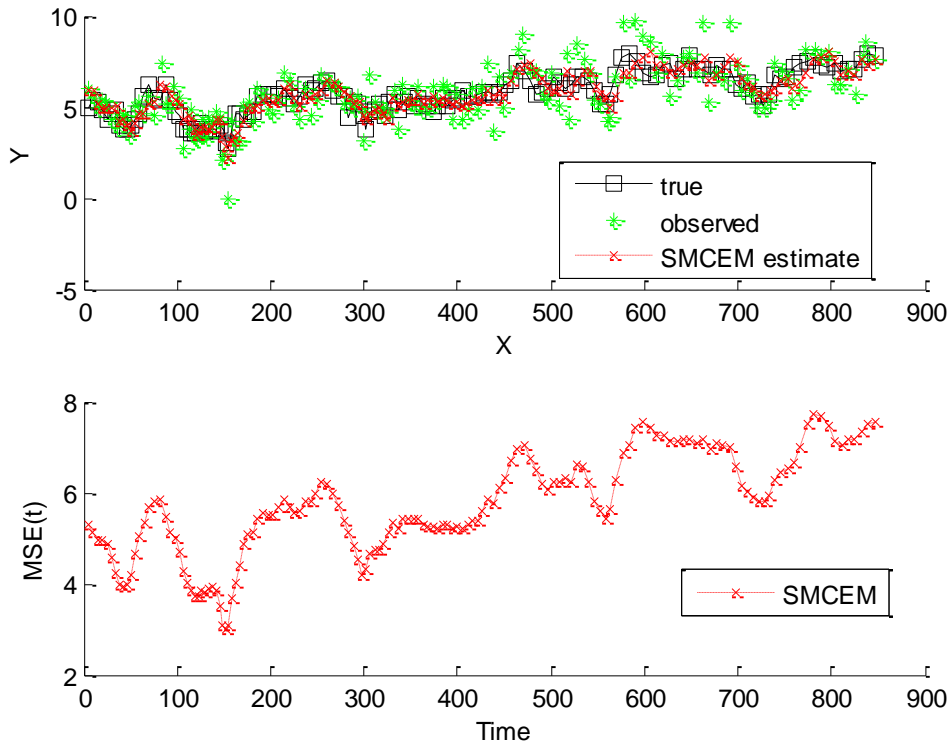


Fig. 4.21. Three scenarios for the BOT. Representation of the trajectories of the true target path (shown by a square), observer (shown by an asterisk) the SMCEM estimate and position MSE (shown by a cross)

Figure 4.21 gives the true target path in the $x - y$ plane, with the position of the target at each time being shown by a square and the observation by an asterisk. The result of applying the SMCEM with $N = 2000$ particles is shown in the figure. The number of particles was chosen such that further increase in N does not bring a significant improvement in the tracking performance. The cross symbol gives the SMCEM estimate such that the estimate moves towards the true target path. The performance is evaluated using the mean square error (MSE) for each time, according to Sanjeev et al., (2004)

$$MSE(t) = \frac{1}{N} \sum_{n=1}^N (x_t^{true} - x_t)^2 \quad (4.6)$$

where x_t denotes the estimate at time t , N is the total number of realizations over which the MSE is averaged. Each of these realizations used observations from the same generated true state. Obviously, with increasing particles, the performance in terms of mean square error (MSE) improves. The MSE is obtained independently for each element of the state in the BOT problem. In Figure 4.21, the MSE value (7.2197) for different times are presented. For the SMCEM based method, no tracks diverged. Thus, the performance of the proposed estimation technique is successful for the target tracking problem.

The proposed estimation technique was also applied to the problem of tracking a moving vehicle. Data were taken from the digital GSM real-time data logging tracking system, (see www.trackingtheworld.com). It models the dynamic properties of the tracked vehicle and estimates it using the proposed technique.

From the data collected, we estimate the vehicle's position $X_t = (x_t, y_t)$ at time t , and its velocity v_t . Also at each time step we obtain a new measurement z_t . the velocity evolves over time according to

$$p(v_t | v_{t-1}) .$$

The vehicle moves based on the evolved velocity according to a dynamics model:

$$p(X_t | X_{t-1}, v_t) .$$

The measurements are governed by a measurement model:

$$p(Z_t | X_t) .$$

The measurement likelihood factor is

$$p(Z_t | X_t) = \prod_{i=1}^N p(z_i | X) .$$

At each time step t , we produce an estimate of the proposed technique about the tracked vehicle trajectory and velocity based on a set of measurements:

$$M_t = p(X_t, v_t | Z_t) . \quad (4.7)$$

Equation (4.7) encodes the vehicle motion and is approximated using the student-distribution. The motion update of the particle filter is carried out using the vehicle dynamic model. The measurement update is carried out by computing the importance weights w_t for all particles:

$$\hat{w}_t^{(i)} = w_{t-1}^{(i)} [p(Z_t | X_t)]$$

where $z_t | x_t \sim t \text{ distribution}(v)$.

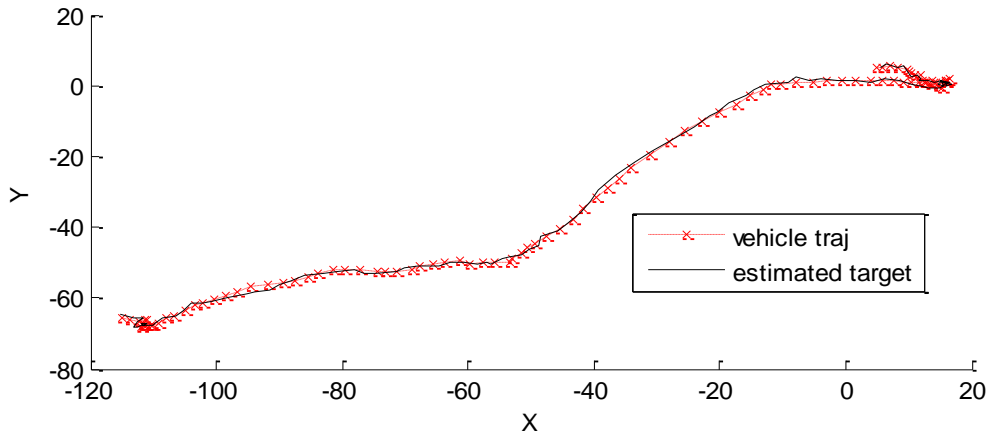


Fig 4.22. Estimates of the vehicle being tracked

The results for this case are presented in Figure 4.22. The SMCCEM based method is able to track the true path of the vehicle being tracked and remains stable and converges, showing that increasing the information provided by the measurements improves the accuracy and convergence of the estimation.

CHAPTER FIVE

5. SUMMARY OF FINDINGS, CONCLUSION AND CONTRIBUTION TO KNOWLEDGE

5.1 SUMMARY OF FINDINGS

In Sections 4.1 and 4.2, from the simulation results presented in Table 4.1 – 4.4, the new parameter estimation technique which combines the EM algorithm along with particle filters and smoothers yields acceptable results when the normal assumption about the observation error is violated as well as when the normal assumption holds, thus exhibiting the viability of the proposed model. Furthermore, the research establishes the proof of convergence of the proposed technique. The convergence results extend the results of Chan and Ledolter (1995) and Godsill et al., (2004).

In Section 4.3, the SMCEM results on the real life data presented in Table 4.6, show that the distributions produced comparable maximum likelihood estimates, indicating an acceptable overall fit. The values (ranging from 0.927 to 0.988) suggest high persistence of the volatility of the series indicating that volatility clustering is observed in all the exchange rates return series.

In Section 4.4, from the results based on the akaike values presented in Table 4.7, , it is evident that the student-t and the GED SV model are comparable to the normal mixture SV model but empirically more successful. This change to the proposed model allows for a more robust fit, giving us a new tool to explore the tail fit.

In Section 4.4.1, the SMCEM results for the multivariate SV model in Table 4.8 show that all parameter estimates are numerically accurate as indicated by the small standard deviations.

Moreover, the estimated volatility parameters of the factor are similar in magnitude to those obtained under univariate SV models. Using the parameter estimates, we computed the corresponding estimate of the unconditional variance–covariance matrix, to be compared with the sample variance– covariance matrix of the returns. The diagonal elements from our model are smaller in each case than those of the sample variance. This indicates that there is more volatility in the data than the model accounts for.

In section 4.5, the proposed technique when applied to solve the nonlinear bearing-only tracking problem showed that the algorithm is capable of modeling the observations and accurately tracking the state vector, thus showing the vast applicability of the proposed technique to other state space models.

5.2 CONCLUSION

This work presents an extension of the observation error in the SV model from normal mixture to student-t and GED distribution. A Sequential Monte-Carlo Expectation Maximization experiment is used to estimate the parameters for the extended SV model. The functions provided by MATLAB enabled us to develop the techniques based on the student-t and GED SV model and a strategy for fitting the model that combines the EM algorithm & SMC. This change to the proposed model allows for a more robust fit, giving us a new tool to explore the tail fit. The student-t and GED SV model were compared and evaluated with the normal mixture. The EM algorithm makes it possible to get maximum likelihood estimators. the estimation algorithm were completed by applying the particle smoothing algorithm of Godsill et al., (2004), to the SV model with (1.3) and (1.5) as an observation equation and a state equation. The experimental

outcome of the simulation and real data analyses confirms the viability of the proposed method. The estimation results show that this proposed estimation algorithm yields acceptable results when the normal assumption is violated, as well as when the normal assumption holds, thus widening the range of application of the SV model.

The empirical studies for the real data reveal that the technique based on the student-t and GED yields similar estimates of coefficients to those of Kim and Stoffer (2008). The evaluation statistics are calculated to compare the fit of distributions. The results, based on daily data from the naira/dollar, cedi/dollar exchange rate, pound/dollar and Euro/dollar reveal that the student-t is comparable to the normal mixture SV model but empirically more successful. A multivariate one-factor SV model performs quite well in explaining the joint dynamics in the volatility of a set of four exchange rates series. Additionally, the work considered another application of the proposed estimation technique. The application of proposed technique to solve a nonlinear bearing-only tracking problem showed that the algorithm is capable of modeling the observations and accurately tracking the state vector. Thus, the procedure can be used for general non-linear non-Gaussian state-space models.

5.3 CONTRIBUTIONS TO KNOWLEDGE

This research has

- i) introduced the techniques based on the student-t and GED SV models, and a strategy for the estimation of parameters of the above models by using the EM algorithm and SMC, thereby expanding the scope of application of the SV models

- ii) generalized and extended previously known results (results of Kim and Stoffer, 2008, Cappe, 2009) for dealing with the problem of heavy tails using techniques based on the student-t and GED SV model.
- iii) established the convergence properties of the proposed technique using the theorem of Chan and Ledolter, (1995) .
- iv) shown that the proposed technique can be used to explain data from many areas, such as economics, engineering and medicine as indicated in its successful application in signal processing (bearing-only tracking).
- v) built multivariate SV models for asset returns in financial economics.

5.4 SUGGESTION FOR FURTHER RESEARCH

This research estimated and tested extensions of the standard SV model to multivariate SV model, while a multivariate one-factor SV model performs quite well in explaining the joint dynamics in the volatility of a set of four exchange rates series. Thus, adding additional dynamic factors in the future would be useful.

The focus here was on a parameter estimation of a class of HMMs with non-linear non-Gaussian state-space models, specifically the SV models in explaining data in financial time series as well as signal processing - bearings-only tracking. However the application of the proposed technique can be applied to other state-space models.

REFERENCES

- Aguilar, O. and West, M., (2000). Bayesian Dynamic Factor Models and Portfolio Allocation, *Journal of Business and Economic Statistics*, 18: 338-357.
- Ahani, E., Abass, O. and Okunoye, O. B. (2010). Simulation of the Nigerian Stock Exchange using Hidden Markov Model, *International Journal of Statistics and Systems*, 5 (1): 29-41.
- Ahani, E. and Abass O. (2010). A Sequential Monte Carlo Approach for Online Stock Market Prediction using Hidden Markov Model, *Journal of Modern Mathematics and Statistics*, 4(2): 73-77.
- Ahani, E., Abass O. and Kasumu R. A. (2010). Online Sequential Monte Carlo Expectation Maximization for estimating parameters of a Hidden Markov Model, Accepted in *Journal of Association for Advancement of Modelling and Simulation (AMSE)*.
- Aidala, V. J. (1979). Kalman filter behavior in bearings-only tracking applications. *Aerospace and Electronic Systems*, 15 (1): 29–39.
- Aidala V. J. and S. C. Nardone, S. C. (1982). Biased estimation properties of the pseudolinear tracking filter. *Aerospace and Electronic Systems*, 18(4):432–441.
- Aidala, V. and Hammel, S. (1983). Utilization of modified polar coordinates for bearings-only tracking. *Automatic Control*, 28 (3) : 283–294.
- Alizadeh, S. Brandt, M. W. and Diebold, F. X. (2002). Range-based Estimation of Stochastic Volatility Models, *Journal of Finance*, 507: 1047–1091.
- Amisano, G. and Giacomini, R. (2007). Comparing Density Forecasts via Weighted Likelihood Ratio tests, *Journal of Business and Economic Statistics*, 25 (2): 177–190.
- Anderson, W. N., Kleindorfer, G. B., Kleindorfer, P. R., and Woodroffe, M. B. (1969). Consistent estimates of the parameters of a linear system. *The Annals of Mathematical Statistics*, 40 (6): 2064–2075
- Andersen, T. G. and Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review*, 39: 885-905.
- Andersen, T.G., Bollerslev, T., and Diebold, F.X. (2005). Parametric and Nonparametric *Financial Econometrics*. Amsterdam: North-Holland.
- Andrieu, C., Doucet, A. and Tadic, V. B. (2005). Online Simulation based Methods for Parameter Estimation in non-linear non Gaussian State-Space Models, *In proceedings IEEE Conference Decision Control*.42: 332-337.

- Archer, G. E. and Titterton, D.M. (2002). Parameter Estimation for Hidden Markov Chains. *Journal of Statistical Planning and Inference* 108: 365–390.
- Asai, M., McAleer, M., and Yu, J. (2006). Multivariate Stochastic Volatility: *Econometric Reviews*, 25: 145-175.
- Asai M. (2009). Bayesian Analysis of Stochastic Volatility Models with mixture-of- normal distributions, *Mathematics and Computers in Simulation*, 79: 2579-2596.
- Avitzour, D. (1995). Stochastic Simulation Bayesian Approach to Multitarget Tracking, *In Proceedings of the IEEE on Radar, Sonar and Navigation*, 142: 41-44.
- Aydemir, A. B. (1998). Volatility Modelling in Finance, in J. Knight & S. Satchell, eds, *Forecasting Volatility in the Financial Markets*, Butterworth-Heinemann, Oxford, 1-46
- Bai, X., Russell, J. R. and Tiao, G. C. (2003). Kurtosis of GARCH and Stochastic Volatility Models with Non-normal Innovations, *Journal of Econometrics*, 114: 349–60.
- Ball, F. G. and Rice, J. H. (1992). Stochastic models for ion channels: Introduction and bibliography, *Mathematical Bioscience*, 112, 189-206.
- Banachewicz, K., Lucas, A. and Vaart, A. (2007). Modeling Portfolio Defaults Using Hidden Markov Models with Covariates, *Econometrics Journal*, 10: 1-18.
- Bao, Y., Tae-HwyLee and Burak, S. (2006). Comparing Density Forecast Models, *Journal of Forecasting*, 26 (3): 203-225.
- Barndorff -Nielsen, O.E. and N.G. Shephard, N. G. (2001). Non-Gaussian Ornstein-Uhlenbeck based Models and Some of their Uses in Financial Econometrics, *Journal of the Royal Statistical Society*, 63: 167-241.
- Baum, L.E. and Petrie, T. (1966). Statistical Inference for Probabilistic Functions of Finite State Markov Chains, *Annals of Mathematical Statistics*, 37: 1554-1563.
- Baum, L. E. and Eagon J. A. (1967). An Inequality with Applications to Statistical Estimation for Probabilistic Functions of Markov Processes, *Bulletin American Mathematical Society*, 73: 360-363.
- Baum, L.E., Petrie, T., Souled, G. and Weiss, N. (1970). A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains *Annals of Mathematical Statistics*, 41 (1): 164-171.
- Bar-Shalom, Y. and T. E. Fortmann, T. E. (1988). *Tracking and Data Association*, Academic-Press, New York.

- Bar-Shalom, Y., Li, X. R. and Kirubarajan, T. (2001). Estimation with Application to Tracking and Navigation, New York: John Wiley and Sons, Inc.
- Bergman, N. (1999). Recursive Bayesian Estimation: Navigation and Tracking Applications, Dissertation 579. Department. of Electrical Engineering, Linkoping University, Linkoping.
- Bickel P. J., Ritov Y. and Ryden T. (1998). Asymptotic Normality of the Maximum Likelihood Estimator for General Hidden Markov models, *The Annals of Statistics*, 26 (4):1614-1635.
- Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.
- Briers, M, Doucet, A. and Maskell, S. (2010). Smoothing Algorithms for state-space models, *Annals of the Institute of Statistical Mathematics*, 62: 61-89.
- Cairns, A. J., Blake, K., Dowd, G. D., Coughlan, D., Epstein, A., and Balevich, I. (2009). A Quantitative Comparison of Stochastic Mortality Models Using Data from England & Wales and the United States, *North American Actuarial Journal* 13 (1): 1–35.
- Capobianco, E. (1996). State-Space Stochastic Volatility Models: A Review of Estimation Algorithms, *Applied Stochastic Models and Data Analysis*, 12: 265-279.
- Cappe, O., Moulines, E. and Ryden, T. (2005). Inference in Hidden Markov Models. New York: Springer -Verlag.
- Cappe, O. (2009). Online Sequential Monte Carlo EM algorithm, *Statistical Signal Processing IEEE/SP 15th Workshop*, 37-40.
- Cappe, O., and Moulines, E. (2009). On-line Expectation-Maximization Algorithm for Latent Data Models, *Journal of the Royal Statistical Society*, 71: 593-613.
- Carnero, A., Pena D. and E. Ruiz (2001). Is Stochastic Volatility more Flexible than GARCH?, Working Paper, 01-08, Universidad Carlos III de Madrid.
- Carpenter, J., Clifford, P. and Fernhead, P. (1999). Improved particle filter for non linear Problems, *IEEE proceedings on radar and sonar navigation*, 146 (1): 2-7.
- Casella, G. and Berger, J. O. (2001). Statistical Inference, 2nd edition, Duxbury Press, California.
- Chan K. S. and Ledolter, J. (1995). Monte Carlo EM Estimation for Time Series models involving Counts, *Journal of the American Statistical Association*, 90: 242–52.

- Chan, N. H. (2002). Time series: application to Finance, John Wiley & Sons, New York.
- Chen, Y., and Lai, T. L. (2003). Sequential Monte Carlo Methods for Filtering and Smoothing in Hidden Markov Models, Technical Report, Department of Statistics, Stanford University.
- Chib, S., Nardari, F. and Shephard, N. (2002). Markov Chain Monte Carlo methods for Stochastic Volatility models, *Journal of Econometrics*, 108: 281–316.
- Chib, S., Nardari, F., and Shephard N. (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics* 134, 341-371.
- Churchill, G. A. (1992). Stochastic Models for Heterogeneous DNA Sequences, *Bulletin of Mathematical Biology*, 51: 79-94.
- Clark, P. K., (1973). A Subordinated Stochastic Process Model with Finite Variances for Speculative Prices. *Econometrica*, 41: 135-156.
- Crisan, D. and Doucet, A. (2002). A Survey of Convergence Results on Particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50: 736-746.
- Danielsson, J. and Richard, J. F. (1993). Accelerated Gaussian Importance Sampler with Application to Dynamic Latent variable models, *Journal of Applied Econometrics*, 8: 153–173.
- Danielson, J. (1994). Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood, *Journal of Econometrics*, 61: 375-400.
- Danielsson, J. (1998). Multivariate Stochastic Volatility Models: Estimation and a Comparison with VGARCH models, *Journal of Empirical Finance*, 5: 125– 173.
- De Vlieger, J. H. and R. H. J. G. Meyling, H. J. G. (1992). Maximum likelihood estimation for long-range target tracking using passive sonar measurements. *Acoustics, Speech, and Signal Processing*, 40(5):1216–1225.
- Del Moral, Jacod, J., and Protter, P. (2001). The Monte-Carlo Method for Filtering with Discrete time Observations, *Probability Theory Related Fields*, 120: 346-368.
- Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society*, 39 (1), 1-38.
- Diebold, F.X., Nerlove, M. (1989). The dynamics of exchange rate volatility: a Multivariate latent-factor ARCH model, *Journal of Applied Econometrics* 4: 1 –22.

- Djuric, P. M., Kotecha, J., Esteve, F. and Perret, E. (2002). Sequential Parameter Estimation of Time-Varying non-Gaussian Autoregressive Processes, *Journal on Applied Signal Processing*, 8: 865-875.
- Douc, R. and Matias, C. (2001). Asymptotics of the Maximum Likelihood Estimator for general Hidden Markov Models, *Bernoulli*, 7: 381-420.
- Douc, R., A., Guillin, J. M. and Robert, C. P. (2007). Convergence of Adaptive Sampling Schemes, *The Annals of Statistics* 35 (1): 1–35.
- Douc, R. and E. Moulines (2008). Limit theorems for weighted samples with applications to sequential Monte Carlo methods, *The Annals of Statistics* 36 (5): 2344–2376.
- Doucet, A., De Freitas, J. F. G. and Gordon, N. (2001). Sequential Monte Carlo Methods in Practice, Springer-Verlag: New York.
- Doucet, A., Godsill, S. J. and Robert, C. P. (2002). Marginal Maximum a Posteriori Estimation using Markov Chain Monte Carlo, *Statistics and Computing*, 12: 77–84.
- Doucet, A. and Tadic, B. B. (2003). Parameter Estimation in General state-space models using particle methods, *Annals of Institute of Statistical Mathematics*, 55: 409–22.
- Doucet A., and Johansen A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. In Oxford Handbook of Nonlinear Filtering, D. Crisan and B. Rozovsky (eds.). Oxford University Press.
- Dupuis, J. A. (1995). Bayesian Estimation of Movement Probabilites in Open Populations using Hidden Markov Model, *Biometrika*, 82: 761-772.
- Durbin, J. and Koopman, S. J. (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives, *Journal of the Royal Statistical Society*, 62: 3–56.
- Efron, B. (1982). The bootstrap, jackknife and other resampling plans. Philadelphia: *Society of Industrial and Applied Mathematics* (SIAM) Monographs, 38.
- Engel C. and Hamilton J.D. (1990). Long swings in the dollar: Are they in the data and do markets know it?, *American Economic Review*, 80: 689–713.
- Engle, R. F. and Shepherd, K. (2001). Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. University of Carlifornia, San Diego. Working Paper.

- Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics*, 20: 339–350.
- Felsenstein, J. and Churchill, G. A. (1996). A hidden Markov Model approach to variation among sites in rate of evolution, *Molecular Biology and Evolution*, 13: 93-104.
- Fernandez, C. and Steel, M. (1998). On Bayesian modelling of fat tails and skewness, *Journal of the American Statistical Association*, 93: 359–371.
- Fearnhead, P. and Clifford, P. (2003). On-line inference for hidden Markov models via particle Filters, *Journal of Royal Statistical Society*, 65: 887-899.
- Ferguson, J. D. (1980). Hidden Markov Analysis: An Introduction, in *Proceedings of the Symposium on the Applications of Hidden Markov Models to Text and Speech*, Princeton, 143-179.
- Fridman, M. and Harris, L. (1998). A Maximum Likelihood Approach for Non-Gaussian Stochastic Volatility Model, *Journal of Business and Economic Statistics*, 16 (3): 284-291.
- Gallant, A. R., Hsieh, D. and Tauchen, G. (1995). Estimation of Stochastic Volatility Model with Diagnostics, *Journal of Econometrics*, 81: 159-192.
- Goshen-Meskin, D. and I. Y. Bar-Itzhack, I. Y. (1992). Observability analysis of piece-wise constant systems. *Aerospace and Electronic Systems*, 28 (4) :1056–1067.
- Geweke, J. (1989). Bayesian Inference in Econometric Models using Monte Carlo integration, *Econometrica* 57: 1317-1339.
- Godsill, S., Doucet, A. and West, M. (2004). Monte Carlo smoothing for non-linear time series. *Journal of the American Statistical Association* , 199: 156–168.
- Gordon, N. J., Salmond, D. J. and Smith, A. (1993). Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation, In *IEEE proceedings on Radar signal process*, 140 (2): 107–113.
- Granger, C.W.J. (1981). Some properties of time series data and their use in econo-metric model specification, *Journal of Econometrics*, 16: 121-130.

- Grundy, W. N., Bailey T. L., Elkan C. P. and Baker M. E. (1999). Motif-based Hidden Markov Models of Protein Families. In *Computer Applications in the Biosciences*, 13 (4): 397-406.
- Harvey, A. C., Ruiz, E. and Shephard, N. (1994). Multivariate Stochastic Variance models, *Review of Economic Studies*, 61: 247–264.
- Hughes, J. P. and Guttorp, P. (1994). A Class of Stochastic Models for Relating Synoptic Atmospheric Patterns to Regional Hydrologic Phenomena, *Water Resources Research*, 30: 1535–1546.
- Ibrahim, J. G., Zhu H. T, and Tang N.S. (2008). Model selection criterion for missing data problems via the EM algorithm, *Journal of the American Statistical Association Theory and Method*, 103: 1648-1658.
- Isard, M., and Blake, A. (1996). Contour Tracking by Stochastic Propagation of Conditional Density. In *Computer Vision – Buxton and R. Cipolla*, New York: Springer.
- Jacquier, E., Polson, N. G. and Rossi, P. E. (1994). Bayesian analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics*, 12 (4): 371–417.
- Jacquier, E., Polson, N. G., and Rossi, P. E. (1999). Stochastic volatility: Univariate and multivariate extensions. Working paper 99s-26, CIRANO, Montreal.
- Jacquier, E., Johannes, M. and Polson, N. (2007). MCMC Maximum Likelihood for Latent State Models, *Journals of Econometrics*, 137 (2): 615-640.
- Jarque, C., and Bera, A. (1987). A Test for Normality of Observations and Regression Residuals, *International Statistical Review*, 55: 163–172.
- Jensen, J. L. and Petersen, N. V. (1999). Asymptotic Normality of the Maximum Likelihood estimator in state space models, *Annals of Statistics*, 27, 514–535.
- Johansen, A., Doucet, A. and Davy M. (2006). Maximum Likelihood Parameter Estimation for Latent Variable Models using Sequential Monte Carlo, *Statistics and Computing*, 18: 47-57.
- Johansen A. M., Doucet A. and Davy M. (2008). Particle methods for Maximum Likelihood Parameter Estimation in Latent Variable Models, *Statistics and Computing*, 18 (1): 47-57.
- Juang, B. H. and Rabiner, L. R. (1991). Hidden Markov Models for Speech Recognition, *Technometrics*, 33 (3): 251-272.

- Kalman, R. E. and Bucy, R. (1961). New Results in Linear Filtering and Prediction Theory, *Journal of Basic Engineering Transaction*, 83: 95-108.
- Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models, *The review of economic studies*, 65 (3): 361–393.
- Kim, J. and Stoffer, D. S. (2008). Fitting Stochastic Volatility Models in the Presence of Irregular Sampling via Particle Methods and the EM Algorithm, *Journal of Time Series Analysis*, 29 (5): 811-833.
- Kitagawa, G. (1993). A Monte Carlo Filtering and Smoothing Method for Non-Gaussian Nonlinear State Space Models. In *Proceedings of the 2nd US-Japan Joint Seminar on Statistical Time Series Analysis*, 110-131, Honolulu, Hawaii.
- Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *Journal of Computational and Graphical Statistics*, 5:1–25.
- Kitagawa, G. and Sato, S. (2001). Monte Carlo smoothing and self-organising state space model. In Doucet, A., de Freitas, N., and Gordon, N., editors, *Sequential Monte Carlo Methods in Practice*, pages 177–195. Springer-Verlag, New York.
- Kong, A., Liu, J. S. and. Wong, W. H. (1994). Sequential Imputations and Bayesian Missing Data Problems, *Journal of the American Statistical Association.*, 89 (425): 278–288.
- Koski, T. (2001). Hidden Markov Models for Bioinformatics. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Krishnamurthy V. and Moore, J. B. (1993). On-line estimation of hidden Markov model parameters based on the kullbackleibler information measure. *IEEE Transactions on Signal Processing*, 41(8): 2557–2573.
- Lawler G. F. (1995). Intorduction to Stochastic Processes, Second Edition: Chapman & Hall Publishers, New York.
- LeGland, F. and Mevel, L. (2000). Exponential forgetting and geometric ergodicity in hidden Markov models, *Mathematics of Control, Signals and Systems*, 13: 63-93.
- Leroux, B. G. (1989). Maximum Likelihood Estimation for Mixture Distributions and hidden Markov Models. PhD Thesis. University of British Columbia, Vancouver.
- Leroux, B. G. (1992). Maximum-Likelihood Estimation for Hidden Markov Models, *Stochastic Processes and their Applications*, 40: 127-143.

- Le Cadre J. P. and Tremois, O. (1998) Bearings-only tracking for maneuvering sources,” *IEEE Transactions on Aerospace and Electronic Systems*, 34 (1) : 179–193.
- Li, J., Najmi, A. and Gray, R. M. (2000). Image Classification by a Two Dimensional HMM, *IEEE Transactions on Signal Processing*, 48 (2): 517-333.
- Liesenfeld, R. and Jung, R. C. (2000). Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions, *Journal of Applied Econometrics*, 15: 137-160.
- Liesenfeld, R. and Richard, J. F. (2003). Univariate and multivariate stochastic volatility models: estimation and diagnostics, *Journal of Empirical Finance*, 10: 505-531.
- Lindgren A. G. and Kai F. Gong. (1978). Position and velocity estimation via bearing observations. *Aerospace and Electronic Systems*, 14 (4): 564–577.
- Liu, J. S. and Chen, R. (1995). Blind Deconvolution via Sequential Imputations. *Journal of the American Statistical Association*, 90: 567-576.
- Liu, X., Kirubarajan, T., Bar-Shalom, Y. and Maskell, S. (2002). Comparison of EKF, Pseudomeasurement and particle filters for bearing-only target tracking problem. In *Proc. SPIE Conference on Image and Signal Processing for Small Targets*, 4728: 240–250.
- Lopes, H. F. and Migon, H. S. (2003). Comovements and Contagion in Emergent Markets: Stock Indexes Volatilities, In *Case Studies in Bayesian Statistics*. New York: Springer-Verlag, pp. 285-300.
- MacDonald I. L. and Zucchini W. (1997). Hidden Markov and other Models for Discrete-Valued Time Series, *Vol., 70 of Monographs on Statistics and Applied Probability*, London: Chapman and Hall.
- MacKay, R. J. (2003). Hidden Markov Models: Multiple Processes and Model Selection. Ph.D Thesis - Dept. of Statistics, The University of British Columbia.
- Maheu, J. M., Thomas, H. and McCurdy, T. H., (2010). Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics*, 160: 69-76.
- Manning C. D. and Schutze H. (1999). Markov Models, In *Foundations of Statistical Natural Language Processing*, Papers in Textlinguistics, 317-379.
- Maskell, S. and Gordon, N. (2002). A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking, *IEEE Transactions on Signal Processing*, 50: 174-188

- Mazor, E., Averbuch, A., Bar-Shalom, Y. and J. Dayan, J. (1996). Interacting multiple model methods in target tracking: A survey. *IEEE Transactions on Aerospace and Electronic Systems*, 34: 103-123.
- Mellino, A. and Turnbull, S. (1990). Pricing foreign currency options with stochastic volatility, *Journal of Econometrics*, 45: 7–39.
- Mongillo, G. and Deneve, S. (2008). Online Learning with Hidden Markov Models. *Neural Computation*, 20 (7): 1706-1716.
- Nardone, S., Lindgren, A. and Kai Gong (1984). Fundamental properties and performance of conventional bearings-only target motion analysis. *Automatic Control*, 29 (9) :775–787.
- Petrie, T. (1969). Probabilistic functions of finite state Markov chains, *Annals of Mathematical Statistics*, 40: 97–115.
- Pitt, M. K. and Shephard, N. (1997). Filtering via Simulation: Auxiliary Particle Filter, *Journal of America Statistical Association*, 94: 590-599.
- Punskeya, E., Doucet, A. and Fitzgerald, W. (2002). On the use of particle filtering in digital communications, in *Proceeding of European Signal Processing Conference*, 2: 173-176.
- Rabiner, L. R. and Juang, B. H. (1986). An Introduction to Hidden Markov Models, *IEEE ASSP Magazine*, 3 (1): 4-16.
- Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, in *Proceedings of the IEEE*, 77 (2): 257-286.
- Rabiner, L. R. and Juang, B. H. (1993). Fundamentals of Speech Recognition. Prentice-Hall, USA.
- Ray, B. K. and Tsay, R. S. (2000). Long-range dependence in daily stock volatilities, *Journal of Business and Economic Statistics*, 18: 254–262.
- Ristic, B., Arulampalam, M. and Gordon, A. (2004). Beyond Kalman Filters: Particle Filters for Target Tracking Application. Artech House, London, UK.
- Rossi A. and Gallo G. M. (2006), Volatility estimation via hidden Markov models, *Journal of Empirical Finance*, 13 (2): 203-230.

- Rubin, D. B. (1988). Using the SIR algorithm to Simulate Posterior Distributions, *Bayesian Statistics*, 3: 395–402.
- Ryden T., Teräsvirta T. and Asbrink S. (1998). Stylized facts of daily return series and the Hidden Markov model, *Journal of Applied Econometrics*, 13: 217–244.
- Sadorsky P. (2005). Stochastic volatility forecasting and risk management, *Applied Financial Economics*, 15: 121–135.
- Sandmann, G. and Koopman, S. J. (1998). Estimation of Stochastic volatility models via Monte Carlo maximum likelihood, *Journal of Econometrics*, 87: 271–301.
- Sanjeev, A. M., Ristic, B., Gordon, N. and Mansell, T. (2004). Bearings-only tracking for Manoeuvring Targets using Particle filters, *Journal on Applied Signal Processing*, 15: 2351–2365
- Sato, M. (2000). Convergence of on-line EM algorithm, *In Proceedings of Int. Conf. on Neural Information Processing*, 1: 476–481.
- Shephard, N. (1994). Partial Non-Gaussian State Space, *Biometrika*, 81: 115–131.
- Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In Time Series Models: In Econometrics, Finance and Other Fields (eds D. R. Cox, D. V. Hinkley and O. E. Barndorff-Nielsen). London: Chapman and Hall, 1–67.
- Shimada J. and Tsukuda Y. (2005). Estimation of Stochastic Volatility Models: An Approximation to the Nonlinear State Space Representation, *Communications in Statistics - Simulation and Computation*, 34: 429 – 450.
- Shumway, R. H. and Stoffer, D. S. (2006). Time Series Analysis and its Applications: Springer, New York.
- Shumway, R. H. and Stoffer, D. S. (2000). Time Series Analysis and its Applications. Springer, New York.
- Smith, A. F. M and Gelfand, A. E. (1992). Bayesian Statistics without Tears: A Sampling-Resampling Perspective, *The American Statistician*, 46: 84–88.
- Tanner, M. A. (1996). Tools for Statistical Inference. Springer-Verlag, New York.
- Tauchen, G. and M. Pitt, (1983). The price variability-volume relationship on speculative markets, *Econometrica*, 51: 485–505.

- Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic Processes—a study of daily sugar prices, 1961–79. In *Time Series Analysis: theory and Practice*, 1, (ed. O. D. Anderson). New York: Elsevier Science Publishing Co., Amsterdam: North-Holland Publishing Co., pp. 203–26.
- Taylor, S. (1986). *Modeling Financial Time Series*. John Wiley and Sons, New York.
- Taylor, S. J. (1994). Modeling Stochastic Volatility, *Mathematical Finance*, 4: 183-204.
- Tims, B., and Mahieu R. (2006). A range-based multivariate stochastic volatility model for exchange rates. *Econometric Reviews* 25, 409-424.
- Titterton, D. M. (1984). Recursive parameter estimation using incomplete data, *Journal of Royal Statistical Society*, 46 (2): 257–267.
- Tse, Y., Tusi, A. (2002). Multivariate model with time-varying correlations. *Journal of Business and Economic Statistics*, 17: 351–362.
- Visser, I., Rijamakers, M. E. J and Molenaar, P. C. M. (2000). Confidence Intervals for Hidden Markov Model Parameters, *British Journal of Mathematical and Statistical Psychology*, 53: 317-327.
- Wald, A. (1949). Note on the Consistency of Maximum Likelihood Estimate, *The Annals of Statistics*, 20,: 595-601.
- Wang, S., and Zhao, Y. (2006). Almost sure convergence of Titterton's recursive estimator for mixture models, *In Statistics and Probability Letters*, 76: 2001-2006.
- Wei, G. and Tanner, M. A. (1991). A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms, *Journal of the American Statistical Association*, 85: 699–704.
- Yaakov Bar-Shalom, Thiagalingam Kirubarajan, and X.-Rong Li. (2002) *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., New York, USA.
- Zhang, D, Ning X., Liu, X., and Hongwei M. (2007). Prediction in Hidden Markov Models Using Sequential Monte Carlo Methods, *In Proceedings of IEEE International Conference on Grey Systems and Intelligent Services*, 718-722.

Appendix A

Parameter estimation results for normal mixture and student-t SV model

A1. Estimation results from data generated from normal mixture SV model: techniques based on normal mixture

iteration	$\phi^{(i)}$	$\tau^{(i)}$	$q_0^{(i)}$	$q_1^{(i)}$	$R_0^{(i)}$	$R_1^{(i)}$	$\pi^{(i)}$	Rel.Lik	M	ε
0	0.9500	1.0729	-0.6794	-3.6794	4.0000	4.0000	0.5000			
1	0.8093	0.9699	-0.7390	-3.6016	3.2444	4.9142	0.4792	14.3001		
2	0.7629	0.9357	-0.8163	-3.5716	2.8007	5.7570	0.4690	2.0087		
3	0.7445	0.9301	-0.8563	-3.5277	2.5575	6.4485	0.4569	0.6524	500	0.1
4	0.7384	0.9345	-0.8918	-3.5236	2.3984	6.9496	0.4469	0.2407		
5	0.7371	0.9326	-0.9276	-3.5090	2.2782	7.2959	0.4393	0.1087		
6	0.7368	0.9408	-0.9826	-3.5081	2.2144	7.5758	0.4330	0.0865		
7	0.8226	0.1103	-0.8358	-3.4686	1.9520	7.8731	0.4379	0.0565		
8	0.8429	0.1125	-0.8457	-3.4818	1.9178	7.8919	0.4346	0.1314		
9	0.8552	0.1144	-0.8508	-3.4750	1.9108	7.9290	0.4314	0.0128	500	0.01
10	0.8643	0.1164	-0.8696	-3.4745	1.8972	7.9594	0.4294	0.0363		
11	0.8677	0.1186	-0.8746	-3.4744	1.8883	7.9490	0.4267	0.0025		
12	0.7631	0.3385	-1.0200	-3.7210	2.2676	7.5133	0.3981	0.0319		
13	0.7622	0.3409	-1.0325	-3.7205	2.2993	7.5590	0.3949	0.0019	1000	0.001
14	0.7609	0.3456	-1.0421	-3.7425	2.2926	7.5224	0.3894	0.0089		
15	0.7568	0.3466	-1.0486	-3.7620	2.3169	7.5643	0.3854	-0.0004		

A2. Estimation results for normal mixture model: techniques based on student-t

Iteration	$\phi^{(i)}$	$\tau^{(i)}$	$\alpha^{(i)}$	Rel.Lik	M	ε
0	0.9500	1.0729	-2.1496			
1	0.7266	1.3670	-2.1504	1.0773	500	0.1
2	0.7070	1.3613	-2.1543	0.1175		
3	0.6976	1.3654	-2.1465	0.0250		
4	0.8891	0.1578	-1.3467	0.1126	500	0.01
5	0.8850	0.1712	-1.3326	0.1747		
6	0.8840	0.1876	-1.3451	0.1881		
7	0.8795	0.2044	-1.3414	0.0870		
8	0.8785	0.2254	-1.3696	0.0893		

9	0.8661	0.2468	-1.3726	0.0576		
10	0.8600	0.2699	-1.3740	0.0672		
11	0.8522	0.2925	-1.3873	0.0686		
12	0.8439	0.3152	-1.3917	0.0702		
13	0.8363	0.3397	-1.4063	0.1875		
14	0.8385	0.3740	-1.4191	0.2537		
15	0.8279	0.4036	-1.4413	0.0477		
16	0.8192	0.4368	-1.4748	0.0995		
17	0.8082	0.4683	-1.4786	0.1374		
18	0.8047	0.5017	-1.5146	0.1622		
19	0.7970	0.5322	-1.5281	-0.0112		
20	0.7878	0.5668	-1.5468	0.0800		
21	0.7748	0.6010	-1.5680	0.0653		
22	0.7698	0.6345	-1.5958	0.0760		
23	0.7636	0.6710	-1.6062	0.0428		
24	0.7522	0.7009	-1.6193	0.0517		
25	0.7413	0.7308	-1.6474	0.0297		
26	0.7308	0.7625	-1.6703	-0.0021		
27	0.7225	0.9135	-1.8566	0.0040	1000	0.001
28	0.7245	0.9333	-1.8619	0.0250		
29	0.7158	0.9377	-1.8710	0.0072		
30	0.7160	0.9557	-1.8700	0.0146		
31	0.7081	0.9670	-1.8778	0.0076		
32	0.6998	0.9664	-1.8686	0.0145		
33	0.7006	0.9667	-1.8933	0.0105		
34	0.7068	0.9930	-1.8915	0.0351		
35	0.7034	1.0124	-1.8962	0.0093		
36	0.6962	1.0165	-1.8958	0.0027		
37	0.6898	1.0192	-1.8935	0.0028		
38	0.6955	1.0332	-1.8950	0.0224		
39	0.6913	1.0336	-1.9009	-0		

A3. Estimation results for student-t SV model: techniques based on normal mixture

Iteration	$\phi^{(i)}$	$\tau^{(i)}$	$q_0^{(i)}$	$q_1^{(i)}$	$R_0^{(i)}$	$R_1^{(i)}$	$\pi^{(i)}$	Rel.Lik	M	ε
0	0.5710	1.1606	-2.7823	-5.7823	4.0000	4.0000	0.5000			
1	0.5952	1.2001	-2.7557	-5.8142	3.9742	4.3193	0.4961	0.4400		
2	0.6200	1.2322	-2.7986	-5.8177	3.8930	4.5138	0.4925	0.1333	500	0.1

3	0.6388	1.2585	-2.8256	-5.7964	3.8225	4.6134	0.4911	0.0643		
4	0.5800	1.6199	-2.8816	-5.7510	3.7019	4.1275	0.4932	0.0826	500	0.01
5	0.5704	1.5957	-2.8931	-5.7492	3.7727	4.1449	0.4914	0.0141		
6	0.5601	1.5805	-2.8979	-5.7544	3.8111	4.2066	0.4911	0.0057		
7	0.6549	1.2925	-2.9987	-5.8696	3.3282	5.1130	0.4843	0.0085	1000	0.001
8	0.6541	1.2921	-3.0092	-5.8682	3.3253	5.1832	0.4820	0.0088		
9	0.6542	1.2913	-3.0130	-5.8643	3.3292	5.2233	0.4810	0.0015		
10	0.6547	1.2930	-3.0180	-5.8536	3.3275	5.2663	0.4806	0.0009		

A4. Estimation results for student-t SV model: techniques based on student-t

Iteration	$\phi^{(i)}$	$\tau^{(i)}$	$\alpha^{(i)}$	Rel. Like	M	ε
0	0.8214	1.3359	-3.1199		500	0.1
1	0.8354	1.3094	-3.1252	0.0401		
2	0.8439	1.2696	-3.1243	0.0045		
3	0.8693	1.4500	-3.1645	0.0042	500	0.01
4	0.8228	1.7098	-2.1057	0.0056	1000	0.001
5	0.8247	1.6772	-2.1068	0.0016		
6	0.8275	1.6367	-2.1016	0.0062		
7	0.8324	1.5988	-2.1002	0.0199		
8	0.8365	1.5766	-2.0951	0.0084		
9	0.8353	1.5506	-2.0891	0.0037		
10	0.8389	1.5519	-2.0862	0.0157		
11	0.8383	1.5357	-2.0912	-0.0010		

Appendix B

Parameter estimation results for normal mixture and GED SV model

B1 Estimation results from data generated from GED SV model: techniques based on GED

Variance Estimate

0.0001 -0.0005 -0.0003
-0.0005 0.0688 0.0197
-0.0003 0.0197 -0.5506

No of Iteration = 6

Total time = 279.9397

Standard deviation of final estimate = 0.0021463 0.2265 0.018455

iteration(i) phi(i) Q(i) alpha(i) Relative Likelihood

0 0.9500 1.3288 0.6309

1.0000 0.9828 1.5411 0.7105 11.1953

2.0000	0.9814	1.6990	0.7327	0.3987
3.0000	0.9802	1.8264	0.7443	0.1294
4.0000	0.9789	1.9643	0.7497	0.2142
5.0000	0.9780	2.0703	0.7352	0.1113
6.0000	0.9770	2.1311	0.7654	0.0248

0.01 500 sample

Variance Estimate

0.0001	-0.0006	0.0000
-0.0006	0.0768	-0.0111
0.0000	-0.0111	-0.1236

No of Iteration = 10

Total time = 455.1017

Standard deviation of final estimate = 0.0025267 0.26119 0.021224

iteration(i) phi(i) Q(i) alpha(i) Relative Likelihood

0	0.9500	1.3288	0.6309	
---	--------	--------	--------	--

1.0000	0.9828	1.5411	0.7105	11.1953
2.0000	0.9814	1.6990	0.7327	0.3987
3.0000	0.9802	1.8264	0.7443	0.1294
4.0000	0.9789	1.9643	0.7497	0.2142
5.0000	0.9780	2.0703	0.7352	0.1113
6.0000	0.9770	2.1311	0.7654	0.0248
7.0000	0.9766	2.1884	0.7755	0.0181
8.0000	0.9762	2.2325	0.7790	0.0121
9.0000	0.9757	2.2812	0.7585	0.0166
10.0000	0.9754	2.3108	0.7627	0.0057

0.001 1000 sample

Variance Estimate

0.0001	-0.0005	0.0003
-0.0005	0.0742	-0.0369
0.0003	-0.0369	-0.1461

No of Iteration = 11

Total time = 5003.1666

Standard deviation of final estimate = 0.0026845 0.2678 0.014247

iteration(i) phi(i) Q(i) alpha(i) Relative Likelihood

0	0.9500	1.3288	0.6309	
---	--------	--------	--------	--

1.0000	0.9830	1.5482	0.6532	11.4154
2.0000	0.9813	1.7226	0.6576	0.4121
3.0000	0.9797	1.8869	0.6437	0.2577
4.0000	0.9782	2.0188	0.6426	0.1740
5.0000	0.9775	2.1014	0.6503	0.0460
6.0000	0.9769	2.1617	0.6645	0.0233

7.0000	0.9761	2.2268	0.6660	0.0323
8.0000	0.9758	2.2788	0.6691	0.0158
9.0000	0.9754	2.3150	0.6790	0.0109
10.0000	0.9750	2.3445	0.6800	0.0058
11.0000	0.9749	2.3496	0.6821	0.0007

B2 Estimation results from data generated from normal mixture SV model: techniques based on GED

Variance Estimate

0.0007	-0.0054	-0.0023
-0.0054	0.2457	0.1535
-0.0023	0.1535	0.3758

No of Iteration = 2

Total time = 88.6583

Standard deviation of final estimate = 0.0059856 0.13077 0.035379

iteration(i)	phi(i)	Q(i)	alpha(i)	Relative Likelihood
0	0.8699	3.6899	-4.4365	

1.0000	0.8570	3.8424	-4.3706	0.8560
2.0000	0.8485	4.0273	-4.3205	0.0853

0.01 500 sample

Variance Estimate

0.0008	-0.0093	0.0020
-0.0093	0.2671	-0.0493
0.0020	-0.0493	0.0394

No of Iteration = 3

Total time = 133.336

Standard deviation of final estimate = 0.0037027 0.052579 0.003807

iteration(i)	phi(i)	Q(i)	alpha(i)	Relative Likelihood
0	0.9500	3.0366	-4.2449	

1.0000	0.8861	2.9451	-4.2504	4.6311
2.0000	0.8814	3.0011	-4.2560	0.0192
3.0000	0.8788	3.0502	-4.2576	0.0064

0.001 1000 sample

Variance Estimate

0.0011	-0.0077	0.0017
-0.0077	0.2698	-0.0356
0.0017	-0.0356	0.0933

No of Iteration = 6

Total time = 786.1738

Standard deviation of final estimate = 0.0027854 0.020652 0.0083309

iteration(i) phi(i) Q(i) alpha(i) Relative Likelihood

0 0.9075 4.6686 -4.7389

1.0000 0.8187 4.2881 -4.7372 6.1792

2.0000 0.8113 4.2367 -4.7367 0.0578

3.0000 0.8126 4.2523 -4.7294 0.0012

4.0000 0.8117 4.2510 -4.7284 0.0013

5.0000 0.8117 4.2319 -4.7265 0.0016

6.0000 0.8127 4.2368 -4.7144 0.0007

B3 Estimation results from data generated from GED SV model: techniques based on normal mixture

No of Iteration = 5

Total time = 259.3772

Standard deviation of final estimate = 0.015127 0.01557 0.08715 0.1269 0.2282

0.84311 0.015076

iteration(i) phi(i) Q(i) m_0(i) m_1(i) R0(i) R1(i) pi(i) Relative Likelihood

Columns 1 through 6

0 0.8250 0.2572 0.3484 -2.6516 4.0000

Columns 7 through 8

4.0000 0.5000

Columns 1 through 6

1.0000 0.8695 0.2745 0.3719 -2.7158 3.5192

2.0000 0.8889 0.2846 0.3099 -2.6360 3.2317

3.0000 0.9038 0.2974 0.2563 -2.5317 3.0711

4.0000 0.9026 0.3045 0.2047 -2.4549 3.0149

5.0000 0.9050 0.3136 0.1491 -2.4081 2.9473

Columns 7 through 9

5.7045 0.4819 8.2085

6.5096 0.4718 1.0975

7.0490 0.4587 0.4812

7.4874 0.4514 0.1234

7.8467 0.4448 0.0052

0.01 500 sample

No of Iteration = 11

Total time = 539.7878

Standard deviation of final estimate = 0.026021 0.025703 0.11521 0.13263 0.28846
1.1097 0.028448

iteration(i) phi(i) Q(i) m_0(i) m_1(i) R0(i) R1(i) pi(i) Relative Likelihood
Columns 1 through 6

0 0.7418 0.3584 0.0012 -2.9988 4.0000

Columns 7 through 8

4.0000 0.5000

Columns 1 through 6

1.0000	0.8067	0.3808	0.0083	-3.0392	3.4186
2.0000	0.8477	0.3970	-0.0912	-2.9562	3.0021
3.0000	0.8711	0.4072	-0.1585	-2.8640	2.7730
4.0000	0.8807	0.4124	-0.2109	-2.7880	2.6284
5.0000	0.8859	0.4204	-0.2399	-2.7236	2.5635
6.0000	0.8871	0.4271	-0.2747	-2.7165	2.5550
7.0000	0.8882	0.4347	-0.2989	-2.7031	2.5307
8.0000	0.8897	0.4432	-0.3209	-2.6769	2.5183
9.0000	0.8909	0.4524	-0.3295	-2.6580	2.5098
10.0000	0.8910	0.4560	-0.3333	-2.6388	2.4895
11.0000	0.8871	0.4605	-0.3568	-2.6609	2.5108

Columns 7 through 9

5.9123	0.4795	11.1725
6.8847	0.4648	2.3854
7.4628	0.4531	0.9593
7.8875	0.4420	0.4326
8.2703	0.4323	0.2220
8.5237	0.4247	0.0821
8.7384	0.4186	0.0574
8.9514	0.4112	0.0573
9.1711	0.4036	0.0664
9.3319	0.3972	0.0276
9.4387	0.3916	-0.0125

>> No of Iteration = 11

Total time = 539.7878

Standard deviation of final estimate = 0.026021 0.025703 0.11521 0.13263 0.28846
1.1097 0.028448

iteration(i) phi(i) Q(i) m_0(i) m_1(i) R0(i) R1(i) pi(i) Relative Likelihood
Columns 1 through 6

0 0.7418 0.3584 0.0012 -2.9988 4.0000

Columns 7 through 8

4.0000 0.5000

Columns 1 through 6

1.0000	0.8067	0.3808	0.0083	-3.0392	3.4186
2.0000	0.8477	0.3970	-0.0912	-2.9562	3.0021
3.0000	0.8711	0.4072	-0.1585	-2.8640	2.7730
4.0000	0.8807	0.4124	-0.2109	-2.7880	2.6284
5.0000	0.8859	0.4204	-0.2399	-2.7236	2.5635
6.0000	0.8871	0.4271	-0.2747	-2.7165	2.5550
7.0000	0.8882	0.4347	-0.2989	-2.7031	2.5307
8.0000	0.8897	0.4432	-0.3209	-2.6769	2.5183
9.0000	0.8909	0.4524	-0.3295	-2.6580	2.5098
10.0000	0.8910	0.4560	-0.3333	-2.6388	2.4895
11.0000	0.8871	0.4605	-0.3568	-2.6609	2.5108

Columns 7 through 9

5.9123	0.4795	11.1725
6.8847	0.4648	2.3854
7.4628	0.4531	0.9593
7.8875	0.4420	0.4326
8.2703	0.4323	0.2220
8.5237	0.4247	0.0821
8.7384	0.4186	0.0574
8.9514	0.4112	0.0573
9.1711	0.4036	0.0664
9.3319	0.3972	0.0276
9.4387	0.3916	

0.001 1000 sample

Variance Estimate

Columns 1 through 6

0.0000	0.0110	0.0004	-0.0013	0.0036	-0.0131
0.0110	-0.0684	-0.0067	0.0131	-0.0369	0.0889
0.0004	-0.0067	0.0013	-0.0089	0.0031	0.0077
-0.0013	0.0131	-0.0089	-0.0141	-0.0398	0.0218
0.0036	-0.0369	0.0031	-0.0398	0.1317	0.0705
-0.0131	0.0889	0.0077	0.0218	0.0705	-0.0310
0.0024	0.0268	0.0246	0.1122	0.2717	-0.1239

Column 7

0.0024
0.0268
0.0246
0.1122
0.2717
-0.1239
1.0414

No of Iteration = 14

Total time = 646.1515

Standard deviation of final estimate = 0.015215 0.069801 0.088286 0.15449 0.47054

1.1994 0.041351

iteration(i) phi(i) Q(i) m_0(i) m_1(i) R0(i) R1(i) pi(i) Relative Likelihood

Columns 1 through 6

0 0.9500 1.7658 0.0472 -2.9528 4.0000

Columns 7 through 8

4.0000 0.5000

Columns 1 through 6

1.0000	0.7759	1.5043	-0.0330	-2.8441	3.2101
2.0000	0.7357	1.4275	-0.0921	-2.7778	2.6811
3.0000	0.7185	1.3735	-0.1115	-2.7611	2.3335
4.0000	0.7193	1.3473	-0.1431	-2.7671	2.1048
5.0000	0.7219	1.3303	-0.1870	-2.7968	1.9461
6.0000	0.7326	1.3140	-0.2014	-2.8313	1.8516
7.0000	0.7362	1.2962	-0.2243	-2.8647	1.7928
8.0000	0.7402	1.2823	-0.2385	-2.9185	1.7313
9.0000	0.7425	1.2729	-0.2502	-2.9853	1.6991
10.0000	0.7473	1.2782	-0.2709	-3.0331	1.6821
11.0000	0.7487	1.2731	-0.2855	-3.0712	1.6670
12.0000	0.7472	1.2687	-0.2967	-3.1117	1.6375
13.0000	0.7499	1.2796	-0.3091	-3.1637	1.6357
14.0000	0.7515	1.2777	-0.3193	-3.1988	1.6365

Columns 7 through 9

4.9086	0.4810	18.7099
5.8650	0.4639	2.2298
6.7320	0.4465	1.1780
7.3280	0.4346	0.2285
7.7537	0.4227	0.1868

8.0216	0.4112	0.0831
8.2753	0.4005	0.0603
8.4366	0.3911	0.0370
8.4710	0.3853	0.0231
8.5269	0.3762	0.0090
8.6863	0.3691	0.0094
8.7283	0.3615	0.0327
8.7644	0.3562	0.0106
8.7981	0.3496	-0.0003

B4 Estimation results from data generated from normal mixture SV model: techniques based on normal mixture

No of Iteration = 3

Total time = 137.4582

Standard deviation of final estimate = 0.0045855 0.17068 0.0059081 0.00932 0.10267
0.16612 0.003056

iteration(i) phi(i) Q(i) m_0(i) m_1(i) R0(i) R1(i) pi(i) Relative Likelihood

Columns 1 through 6

0	0.8987	4.4300	-5.5210	-8.5210	4.0000
---	--------	--------	---------	---------	--------

Columns 7 through 8

4.0000	0.5000
--------	--------

Columns 1 through 6

1.0000	0.7923	3.8149	-5.5214	-8.5094	3.8962
2.0000	0.7831	3.6039	-5.5225	-8.5125	3.7653
3.0000	0.7875	3.4771	-5.5322	-8.5269	3.6937

Columns 7 through 9

4.1539	0.4955	8.2747
4.2923	0.4914	0.1702
4.4847	0.4895	0.0248

0.01 500 sample

No of Iteration = 4

Total time = 199.6295

Standard deviation of final estimate = 0.0046651 0.17631 0.026411 0.032108 0.081339
0.057281 0.0025453

iteration(i) phi(i) Q(i) m_0(i) m_1(i) R0(i) R1(i) pi(i) Relative Likelihood

Columns 1 through 6

0 0.8699 3.6899 -4.8897 -7.8897 4.0000

Columns 7 through 8

4.0000 0.5000

Columns 1 through 6

1.0000	0.8324	3.3688	-4.8714	-7.8968	3.8962
2.0000	0.8330	3.2129	-4.8508	-7.9105	3.8038
3.0000	0.8371	3.0787	-4.8429	-7.9483	3.7550
4.0000	0.8425	2.9591	-4.8080	-7.9658	3.7054

Columns 7 through 9

3.9752	0.4979	1.6141
4.0192	0.4941	0.0760
4.0460	0.4921	0.0338
4.1121	0.4930	-0.0148

0.001 1000 sample

No of Iteration = 3

Total time = 464.2807

Standard deviation of final estimate = 0.0013341 0.51468 0.021891 0.026309 0.069579
0.16268 0.0032833

iteration(i) phi(i) Q(i) m_0(i) m_1(i) R0(i) R1(i) pi(i) Relative Likelihood

Columns 1 through 6

0 0.9500 5.8502 -1.4214 -4.4214 4.0000

Columns 7 through 8

4.0000 0.5000

Columns 1 through 6

1.0000	0.9843	5.2198	-1.4331	-4.4032	3.8701
2.0000	0.9858	4.6363	-1.4184	-4.4234	3.7759
3.0000	0.9869	4.1936	-1.3900	-4.4554	3.7342

Columns 7 through 9

4.0264	0.4960	9.0561
4.1807	0.4921	0.0746
4.3517	0.4895	-0.0684

Appendix C

C1 Parameter estimation for the GED SV model

Filtering Step:

The algorithm for the filtering and smoothing steps below shows a slight modification of Godsill et al. (2004) and Kim & Stoffer (2008). From here M samples from $f(x_t, | Y_t)$ for each t were obtained.

- i) Generate $f_0^{(i)} \sim N(\mu_0, \sigma_0^2)$
- ii) For $t = 1, \dots, n$
- iii) Generate a random number $w_t^{(i)} \sim N(0, \tau)$, $j = 1, \dots, M$
- iv) Compute $p_t^{(i)} = \phi f_{t-1}^{(i)} + w_t^{(i)}$
- v) Compute $w_t^{(i)} = p(y_t | p_t^{(i)},) \propto \frac{v \exp\left[-\frac{1}{2}((y_t - \alpha - v_t)/\psi)^v\right]}{\psi \Gamma\left(\frac{1}{v}\right) 2^{\left(1+\frac{1}{v}\right)}}$
- vi) Generate $f_t^{(i)}$ by resampling with weights, $w_t^{(j)}$

C2 Smoothing step

In the smoothing step, particle smoothers that are needed to get the expected likelihood in the expectation step of the EM algorithm were gotten:

Suppose that equally weighted particles $\{f_t^{(i)}\}$, $i = 1, \dots, M$ from $f(x_t, | Y_t)$ are available for $t = 1, \dots, n$ from the filtering step.

- 1. Choose $[s_n^{(i)}] = [f_n^{(j)}]$ with probability $\frac{1}{M}$.
- 2. For $n-1$ to 0
- c. Calculate

$$w_{t|t+1}^{(i)} \propto f(s_{t+1}^{(i)} | f_t^{(j)}) \propto \exp\left(-\frac{(s_{t+1}^{(i)} - \phi f_t^{(j)})^2}{2\tau}\right) \frac{v \exp\left[-\frac{1}{2}((y_t - \alpha - v_t)/\psi)^v\right]}{\psi \Gamma\left(\frac{1}{v}\right) 2^{\left(1+\frac{1}{v}\right)}} \text{ for each } j$$

- d. Choose $[s_t^{(i)}] = [f_t^{(j)}]$ with probability $w_{t|t+1}^j$.
3. $(s_{0:n}^{(i)}) = \{(s_0^{(i)}, \dots, s_n^{(i)})\}$ is the random sample from $f(x_0, \dots, x_n | Y_n)$
4. Repeat 1-3, for $i = 1, \dots, M$ and calculate

$$\hat{x}_t^n = \frac{\sum_{i=1}^M s_t^{(i)}}{M}, \hat{p}_t^n = \frac{\sum_{i=1}^M (s_t^{(i)} - \hat{x}_t^n)^2}{M-1}, \hat{p}_{t,t-1}^n = \frac{\sum_{i=1}^M (s_t^{(i)} - \hat{x}_t^n)(s_{t-1}^{(i)} - \hat{x}_{t-1}^n)}{M},$$

$$E \left[-\frac{1}{2} \sum_{t=1}^n \left(\frac{y_t - \alpha - v_t}{\psi} \right)^v \right] = \left[n \sum_{t=1}^n (y_t - v_t)^{v-1} \right]^{\frac{1}{v-1}}$$

Appendix D

The table below shows historical exchange rates between the Nigerian naira (NGN) and the US Dollar (USD), British Pound (GBP) and the US Dollar, Ghanaian Cedi (GHS) and the US Dollar, Euro (EUR) and the US Dollar.

Date	USD/NGN	USD/GBP	USD/GHC	USD/EUR
3/9/2009	146.897	0.7149	13,784.30	0.7907
3/10/2009	146.952	0.7237	13,739.10	0.7882
3/11/2009	147.148	0.7263	13,811.20	0.7862
3/12/2009	146.008	0.7226	13,807.60	0.7804
3/13/2009	146.207	0.7163	13,779.40	0.7747
3/14/2009	145.716	0.7139	13,788.90	0.7731
3/15/2009	145.716	0.7139	13,794.50	0.7732
3/16/2009	146.882	0.7109	13,735.50	0.772
3/17/2009	146.873	0.7113	13,900.10	0.77
3/18/2009	146.799	0.7133	13,755.70	0.7645
3/19/2009	146.151	0.6967	13,654.40	0.7385
3/20/2009	147.144	0.6906	13,863.80	0.7338
3/21/2009	146.72	0.6911	13,880.40	0.736
3/22/2009	146.72	0.6911	13,871.60	0.736
3/23/2009	148.089	0.6883	13,868.20	0.7336
3/24/2009	148	0.6815	13,937.80	0.7355
3/25/2009	148.058	0.6833	13,920.40	0.7408
3/26/2009	146.417	0.687	13,935.90	0.7367
3/27/2009	146.961	0.6945	13,984.90	0.7433
3/28/2009	146.56	0.6981	13,947.00	0.7522
3/29/2009	146.56	0.6984	13,955.90	0.7523
3/30/2009	145.846	0.7036	13,960.40	0.7571
3/31/2009	146.48	0.6987	13,917.00	0.7532
4/1/2009	146.903	0.6964	13,995.20	0.756

4/2/2009	146.448	0.6855	13,879.30	0.75
4/3/2009	146.195	0.6777	13,918.60	0.7439
4/4/2009	145.626	0.6736	13,901.40	0.7413
4/5/2009	145.626	0.6736	13,898.30	0.7413
4/6/2009	145.49	0.6729	13,973.50	0.7408
4/7/2009	146.927	0.6801	14,031.30	0.7505
4/8/2009	147.208	0.6805	13,988.20	0.7557
4/9/2009	147.033	0.6808	14,012.70	0.755
4/10/2009	146.186	0.6826	14,124.70	0.7609
4/11/2009	146.186	0.6812	14,077.00	0.7581
4/12/2009	146.186	0.6812	14,072.50	0.7581
4/13/2009	147	0.6794	13,991.90	0.7558
4/14/2009	146.2	0.6721	13,982.00	0.7514
4/15/2009	146.156	0.6699	14,074.60	0.7559
4/16/2009	146.783	0.6686	14,071.70	0.7577
4/17/2009	146.944	0.6736	14,151.70	0.7635
4/18/2009	147.4	0.6756	14,546.50	0.7663
4/19/2009	147.4	0.6755	14,667.00	0.7663
4/20/2009	147.042	0.6824	14,166.20	0.7705
4/21/2009	147.243	0.686	14,094.30	0.7728
4/22/2009	146.814	0.6854	14,101.10	0.7718
4/23/2009	146.942	0.6873	14,133.70	0.7672
4/24/2009	146.933	0.6814	14,141.50	0.7574
4/25/2009	146.756	0.68	14,223.90	0.755
4/26/2009	147.08	0.681	14,204.30	0.7549
4/27/2009	146.806	0.6842	14,308.60	0.7604
4/28/2009	145.879	0.6846	14,279.50	0.7669
4/29/2009	145.919	0.6794	14,198.90	0.7567
4/30/2009	146.937	0.6751	14,326.00	0.753
5/1/2009	147.049	0.6739	14,353.30	0.7541
5/2/2009	147.049	0.6699	14,347.90	0.7533
5/3/2009	147.049	0.6699	14,349.50	0.7533
5/4/2009	145.72	0.6694	14,283.20	0.7512
5/5/2009	146.427	0.6642	14,314.90	0.7474
5/6/2009	146.523	0.664	14,389.00	0.7516
5/7/2009	146.697	0.6636	14,325.30	0.7492
5/8/2009	146.696	0.6629	14,228.80	0.7436
5/9/2009	145.995	0.6563	14,323.20	0.7332
5/10/2009	145.995	0.6563	14,308.20	0.7332
5/11/2009	146.158	0.6593	14,221.90	0.7344
5/12/2009	146.249	0.6579	14,315.70	0.7341
5/13/2009	146.499	0.6567	14,355.10	0.7326
5/14/2009	146.896	0.66	14,361.20	0.7359
5/15/2009	146.909	0.6574	14,359.10	0.7363
5/16/2009	147.541	0.6585	14,401.50	0.7407
5/17/2009	147.25	0.6586	14,421.00	0.7408
5/18/2009	147.002	0.6567	14,372.60	0.7418
5/19/2009	147.035	0.6484	14,331.40	0.7357
5/20/2009	147.082	0.6429	14,358.40	0.731
5/21/2009	146.704	0.6349	14,275.30	0.7246

5/22/2009	146.651	0.6299	14,352.00	0.7163
5/23/2009	146.651	0.6275	14,512.20	0.7141
5/24/2009	146.651	0.6276	14,492.30	0.7141
5/25/2009	146.88	0.6289	14,516.70	0.7141
5/26/2009	146.558	0.6293	14,509.10	0.7162
5/27/2009	146.364	0.6252	14,528.30	0.7169
5/28/2009	146.23	0.6273	14,565.80	0.7206
5/29/2009	146.054	0.6208	14,486.00	0.712
5/30/2009	146.054	0.6175	14,533.00	0.7065
5/31/2009	146.054	0.6175	14,533.00	0.7065
6/1/2009	147.799	0.6124	14,475.80	0.7056
6/2/2009	147.312	0.6073	14,532.30	0.7039
6/3/2009	146.622	0.6051	14,614.10	0.7024
6/4/2009	147.781	0.6146	14,681.30	0.7055
6/5/2009	147.646	0.6222	14,764.10	0.7079
6/6/2009	148.571	0.6255	14,895.90	0.7157
6/7/2009	148.571	0.6256	14,712.40	0.7157
6/8/2009	147.405	0.6271	14,764.40	0.7186
6/9/2009	147.717	0.6191	14,556.90	0.7168
6/10/2009	147.396	0.6117	14,621.80	0.7115
6/11/2009	147.663	0.6077	14,663.90	0.7123
6/12/2009	148.073	0.606	14,684.00	0.7115
6/13/2009	148.073	0.6079	14,733.00	0.7132
6/14/2009	148.073	0.6079	14,731.70	0.7132
6/15/2009	147.942	0.6114	14,792.10	0.7202
6/16/2009	147.766	0.6111	14,706.20	0.7225
6/17/2009	147.88	0.6113	14,670.30	0.7209
6/18/2009	146.534	0.6116	14,669.40	0.7171
6/19/2009	148.496	0.6094	14,712.00	0.7179
6/20/2009	146.154	0.6061	14,678.90	0.7173
6/21/2009	146.154	0.606	14,703.00	0.7172
6/22/2009	148.484	0.6086	14,735.60	0.7203
6/23/2009	147.388	0.6121	14,660.10	0.7172
6/24/2009	147.68	0.6066	14,729.60	0.7117
6/25/2009	147.711	0.6108	14,785.30	0.7167
6/26/2009	147.307	0.6074	14,708.90	0.7116
6/27/2009	146.75	0.605	14,793.20	0.7111
6/28/2009	146.75	0.605	14,793.70	0.7111
6/29/2009	147.396	0.6053	14,785.50	0.7118
6/30/2009	147.427	0.6036	14,789.70	0.7099
7/1/2009	147.546	0.6074	14,756.20	0.71
7/2/2009	147.089	0.609	14,786.40	0.7103
7/3/2009	147.417	0.6114	14,807.00	0.7149
7/4/2009	146.11	0.6121	14,790.60	0.7151
7/5/2009	145.79	0.6121	14,781.00	0.7151
7/6/2009	147.547	0.6161	14,789.40	0.717
7/7/2009	147.448	0.6167	14,820.30	0.716
7/8/2009	148.04	0.6219	14,895.60	0.7197
7/9/2009	147.221	0.6179	14,768.40	0.7167
7/10/2009	146.251	0.6152	14,796.20	0.7166

7/11/2009	145.469	0.6167	14,776.20	0.7173
7/12/2009	147.8	0.6167	14,767.50	0.7173
7/13/2009	146.117	0.6172	14,793.00	0.7168
7/14/2009	147.404	0.6143	14,819.10	0.7156
7/15/2009	148.153	0.6106	14,777.40	0.712
7/16/2009	147.956	0.6092	14,804.30	0.7091
7/17/2009	147.956	0.6115	14,804.30	0.7085
7/18/2009	146.29	0.612	14,817.80	0.7089
7/19/2009	146.29	0.612	14,825.30	0.7089
7/20/2009	147.92	0.6076	14,785.20	0.7046
7/21/2009	148.337	0.6071	14,839.10	0.7036
7/22/2009	148.798	0.6092	14,831.30	0.704
7/23/2009	150.577	0.6061	14,798.50	0.703
7/24/2009	150.431	0.6072	14,845.20	0.7045
7/25/2009	149.1	0.6085	14,868.20	0.7039
7/26/2009	149.1	0.6084	14,871.60	0.7038
7/27/2009	149.436	0.6072	14,849.40	0.7024
7/28/2009	149.436	0.6068	14,896.20	0.7029
7/29/2009	151.299	0.6096	14,947.00	0.7081
7/30/2009	155.266	0.6079	14,879.60	0.7112
7/31/2009	156.7	0.6034	15,000.00	0.7068
8/1/2009	156.7	0.5981	14,711.00	0.7012
8/2/2009	156.7	0.598	14,709.20	0.7011
8/3/2009	154.999	0.5946	14,695.80	0.6989
8/4/2009	155.446	0.5902	14,733.30	0.6943
8/5/2009	154.511	0.5893	14,761.10	0.6942
8/6/2009	154.316	0.5912	14,733.50	0.6951
8/7/2009	153.648	0.5973	14,688.30	0.6989
8/8/2009	153.95	0.5992	14,723.70	0.7048
8/9/2009	153.8	0.5992	14,729.30	0.7048
8/10/2009	153.849	0.6019	14,635.80	0.7053
8/11/2009	153.754	0.6066	14,561.80	0.7068
8/12/2009	153.744	0.6066	14,544.90	0.706
8/13/2009	153.773	0.6041	14,509.40	0.7016
8/14/2009	155.941	0.604	14,700.00	0.7014
8/15/2009	154.34	0.6043	14,609.20	0.7038
8/16/2009	156	0.6044	14,609.80	0.7039
8/17/2009	156.871	0.6103	14,690.00	0.7084
8/18/2009	157.806	0.6083	14,700.00	0.7086
8/19/2009	156.466	0.6061	14,700.00	0.7064
8/20/2009	155.978	0.6053	14,455.60	0.7026
8/21/2009	155.859	0.606	14,488.80	0.6999
8/22/2009	154.62	0.6055	14,518.20	0.6978
8/23/2009	154.619	0.6055	14,516.30	0.6978
8/24/2009	155.974	0.6068	14,505.90	0.6982
8/25/2009	154.503	0.6102	14,466.10	0.6991
8/26/2009	153.871	0.614	14,464.20	0.7003
8/27/2009	153.72	0.6165	14,449.40	0.7011
8/28/2009	152.919	0.6135	14,422.00	0.6968
8/29/2009	152.13	0.6145	14,476.00	0.6989

8/30/2009	152.13	0.6145	14,489.90	0.6989
8/31/2009	153.868	0.6154	14,461.30	0.6991
9/1/2009	153.081	0.6155	14,431.40	0.6989
9/2/2009	153.198	0.6174	14,440.60	0.7028
9/3/2009	153.517	0.6129	14,407.20	0.7004
9/4/2009	153.94	0.6116	14,410.70	0.7009
9/5/2009	151.9	0.6099	14,374.20	0.6995
9/6/2009	151.9	0.6099	14,360.70	0.6995
9/7/2009	153.204	0.6102	14,362.50	0.6976
9/8/2009	153.917	0.6079	14,351.90	0.6934
9/9/2009	153.202	0.6051	14,343.70	0.6887
9/10/2009	152.706	0.6031	14,330.90	0.6866
9/11/2009	153.56	0.5992	14,364.20	0.6853
9/12/2009	153.56	0.6001	14,415.00	0.6861
9/13/2009	153.56	0.6	14,414.00	0.686
9/14/2009	153.487	0.6026	14,378.10	0.6861
9/15/2009	153.779	0.6044	14,346.70	0.6839
9/16/2009	153.214	0.6065	14,342.20	0.6809
9/17/2009	152.799	0.6061	14,366.00	0.679
9/18/2009	153.537	0.6117	14,400.30	0.6796
9/19/2009	151.76	0.6144	14,392.20	0.6795
9/20/2009	151.771	0.6144	14,393.90	0.6795
9/21/2009	151.676	0.6171	14,383.20	0.6814
9/22/2009	154.053	0.6136	14,279.70	0.6778
9/23/2009	152.672	0.6099	14,306.40	0.6761
9/24/2009	150.558	0.6162	14,358.40	0.6791
9/25/2009	149.934	0.6256	14,315.20	0.6816
9/26/2009	148.72	0.6267	14,255.00	0.6805
9/27/2009	148.72	0.6266	14,254.00	0.6805
9/28/2009	151.429	0.6304	14,298.60	0.6833
9/29/2009	151.336	0.6281	14,300.10	0.6853
9/30/2009	149.165	0.6242	14,293.50	0.6838
10/1/2009	147.827	0.626	14,387.10	0.6856
10/2/2009	147.827	0.6287	14,387.10	0.687
10/3/2009	144.214	0.6268	14,309.30	0.6858
10/4/2009	144.21	0.6269	14,303.50	0.6857
10/5/2009	147.195	0.6267	14,318.50	0.6836
10/6/2009	146.36	0.6272	14,321.00	0.6798
10/7/2009	147.938	0.6286	14,370.60	0.6803
10/8/2009	147.29	0.6236	14,282.70	0.6776
10/9/2009	147.384	0.6261	14,306.10	0.6785
10/10/2009	147.84	0.6309	14,303.00	0.6786
10/11/2009	147.84	0.6309	14,304.30	0.6787
10/12/2009	147.33	0.6323	14,272.60	0.6782
10/13/2009	149.354	0.6323	14,270.00	0.6754
10/14/2009	150.48	0.6265	14,288.70	0.6717
10/15/2009	150.01	0.619	14,315.50	0.6697
10/16/2009	148.03	0.6123	14,311.40	0.6705
10/17/2009	148.05	0.6112	14,288.30	0.6706
10/18/2009	148.05	0.6114	14,290.70	0.6706

10/19/2009	148.17	0.6122	14,265.60	0.6705
10/20/2009	148.36	0.6097	14,272.50	0.6686
10/21/2009	148.18	0.606	14,233.40	0.6686
10/22/2009	148.21	0.6027	14,193.40	0.6667
10/23/2009	148.425	0.6062	14,216.00	0.6655
10/24/2009	149.5	0.6131	14,262.00	0.6662
10/25/2009	149.5	0.6132	14,263.60	0.6662
10/26/2009	149.11	0.613	14,280.90	0.6671
10/27/2009	148.92	0.6116	14,290.60	0.6731
10/28/2009	148.843	0.6109	14,235.20	0.6762
10/29/2009	148.785	0.6078	14,139.40	0.6776
10/30/2009	148.84	0.6051	14,144.70	0.6754
10/31/2009	149.63	0.6079	14,221.00	0.6793
11/1/2009	149.63	0.6079	14,221.00	0.6793
11/2/2009	149.17	0.6096	14,160.80	0.6774
11/3/2009	149.186	0.611	14,184.40	0.6789
11/4/2009	148.12	0.6064	14,077.00	0.6771
11/5/2009	149.74	0.6042	14,075.20	0.6731
11/6/2009	149.66	0.6026	14,123.70	0.6725
11/7/2009	149.66	0.6017	14,158.00	0.6733
11/8/2009	149.66	0.6016	14,158.20	0.6731
11/9/2009	149.99	0.5977	14,092.90	0.6687
11/10/2009	150.47	0.5984	14,160.90	0.6674
11/11/2009	150.09	0.5995	14,174.40	0.6667
11/12/2009	150.2	0.6035	14,217.10	0.669
11/13/2009	150.071	0.601	14,173.20	0.6723
11/14/2009	151.7	0.5996	14,128.00	0.6708
11/15/2009	151.7	0.5995	14,125.90	0.6707
11/16/2009	150.16	0.5975	14,106.90	0.6683
11/17/2009	150.109	0.5948	14,183.20	0.6703
11/18/2009	149.59	0.5955	14,163.40	0.6703
11/19/2009	148.95	0.5995	14,183.50	0.6709
11/20/2009	148.95	0.6032	14,183.50	0.6718
11/21/2009	147.99	0.6056	14,169.50	0.6726
11/22/2009	149.75	0.6056	14,169.10	0.6726
11/23/2009	148.18	0.6034	14,128.30	0.6691
11/24/2009	147.24	0.6035	14,167.90	0.6689
11/25/2009	147.939	0.6004	14,099.70	0.6654
11/26/2009	147.861	0.6024	14,152.00	0.6633
11/27/2009	148.1	0.6087	14,182.80	0.6694
11/28/2009	147.52	0.6057	14,069.70	0.6669
11/29/2009	147.52	0.6055	14,074.00	0.6668
11/30/2009	146.86	0.606	14,084.20	0.6654
12/1/2009	144.417	0.6053	14,125.00	0.6644
12/2/2009	145.519	0.6013	14,167.50	0.6631
12/3/2009	147.02	0.601	14,179.60	0.6626
12/4/2009	147.92	0.6039	14,287.10	0.6663
12/5/2009	150.03	0.607	14,366.90	0.6729
12/6/2009	150.5	0.6071	14,362.50	0.6729
12/7/2009	149.216	0.6086	14,244.30	0.6736

12/8/2009	149.8	0.6109	14,142.00	0.6758
12/9/2009	147.61	0.615	14,149.30	0.6792
12/10/2009	148.12	0.6144	14,119.20	0.679
12/11/2009	148.975	0.6142	14,149.40	0.68
12/12/2009	149.989	0.6148	14,245.20	0.6839
12/13/2009	149.9	0.6148	14,243.30	0.6839
12/14/2009	148.3	0.6152	14,119.50	0.6828
12/15/2009	147.67	0.6146	14,113.50	0.6854
12/16/2009	147.66	0.6134	14,107.30	0.6874
12/17/2009	148.49	0.6168	14,179.10	0.6942
12/18/2009	147.83	0.6186	14,130.30	0.6968
12/19/2009	149.25	0.6187	14,150.00	0.6972
12/20/2009	149.25	0.6187	14,151.90	0.6973
12/21/2009	147.08	0.6205	14,167.60	0.698
12/22/2009	147.126	0.624	14,142.80	0.7
12/23/2009	147.571	0.6266	14,078.70	0.7003
12/24/2009	147.13	0.626	14,092.50	0.6963
12/25/2009	147.13	0.6261	14,137.50	0.695
12/26/2009	147.13	0.6243	14,142.00	0.6945
12/27/2009	148.75	0.6264	14,142.00	0.6948
12/28/2009	147.9	0.6259	14,148.60	0.6949
12/29/2009	148.13	0.6254	14,156.00	0.695
12/30/2009	148.162	0.6278	14,168.60	0.6977
12/31/2009	147.38	0.62	14,138.40	0.696
1/1/2010	149.45	0.6188	14,151.60	0.6985
1/2/2010	149.45	0.6198	14,154.00	0.6949
1/3/2010	149.45	0.6188	14,171.20	0.6984
1/4/2010	148.047	0.6199	14,174.80	0.6966
1/5/2010	147.44	0.6227	14,177.60	0.6936
1/6/2010	147.457	0.6253	14,189.50	0.6961
1/7/2010	147.036	0.6263	14,175.00	0.6961
1/8/2010	147.84	0.6257	14,143.90	0.6976
1/9/2010	149.9	0.6239	14,074.20	0.6937
1/10/2010	149.9	0.6238	14,069.00	0.6936
1/11/2010	148.48	0.6205	14,087.90	0.6894
1/12/2010	148.37	0.6201	14,086.10	0.6898
1/13/2010	148.5	0.6163	14,010.10	0.6895
1/14/2010	149.349	0.6136	14,007.20	0.689
1/15/2010	149.21	0.6134	14,016.60	0.6936
1/16/2010	149.17	0.6147	13,958.00	0.6949
1/17/2010	151.1	0.6147	13,964.80	0.695
1/18/2010	148.33	0.613	13,999.00	0.6957
1/19/2010	149.02	0.6104	14,123.20	0.697
1/20/2010	148.2	0.6134	14,142.00	0.705
1/21/2010	148.193	0.6161	14,018.00	0.7097
1/22/2010	148.004	0.618	14,000.60	0.7078
1/23/2010	147.75	0.6207	14,038.00	0.7072
1/24/2010	150	0.6206	14,047.70	0.707
1/25/2010	148.178	0.619	14,083.70	0.7064
1/26/2010	148.71	0.6177	14,152.80	0.7089

1/27/2010	148.12	0.6186	14,160.00	0.7112
1/28/2010	148.73	0.6175	14,163.30	0.7141
1/29/2010	148.62	0.6211	14,183.10	0.7176
1/30/2010	149.622	0.6253	14,231.80	0.7211
1/31/2010	150.88	0.6255	14,233.80	0.7211
2/1/2010	149.4	0.6275	14,158.90	0.7198
2/2/2010	149.585	0.6269	14,118.50	0.7176
2/3/2010	149.339	0.6259	14,185.50	0.7164
2/4/2010	149.35	0.6307	14,254.90	0.7224
2/5/2010	149.26	0.637	14,241.60	0.7305
2/6/2010	149.6	0.6383	14,240.20	0.7308
2/7/2010	151.82	0.6383	14,247.00	0.7306
2/8/2010	149.27	0.6408	14,149.60	0.7317
2/9/2010	148.93	0.6401	14,092.20	0.729
2/10/2010	148.61	0.6385	14,151.10	0.7267
2/11/2010	149	0.6398	14,173.30	0.7283
2/12/2010	149.11	0.6381	14,140.50	0.7335
2/13/2010	151.3	0.6367	14,046.40	0.7333
2/14/2010	151.3	0.6368	14,054.00	0.7334
2/15/2010	148.92	0.6383	14,118.10	0.735
2/16/2010	148.64	0.6361	14,074.80	0.7303
2/17/2010	148.83	0.6352	14,122.40	0.7301
2/18/2010	148.91	0.6404	14,127.00	0.7361
2/19/2010	148.18	0.6485	14,055.80	0.7394
2/20/2010	147.85	0.6461	13,978.10	0.7344
2/21/2010	150.6	0.6461	14,008.40	0.7342
2/22/2010	148.28	0.646	14,028.70	0.7348
2/23/2010	148.133	0.6468	14,102.90	0.7364
2/24/2010	148.124	0.648	14,100.00	0.7381
2/25/2010	148.15	0.6542	14,103.20	0.7407
2/26/2010	147.96	0.6561	14,035.10	0.7352
2/27/2010	147.291	0.6561	14,030.90	0.7335
2/28/2010	150.3	0.6563	14,024.10	0.7333
3/1/2010	148.785	0.6658	14,080.70	0.7371
3/2/2010	147.86	0.6692	14,063.70	0.7378
3/3/2010	148.14	0.664	14,015.70	0.7318
3/4/2010	147.98	0.664	14,101.40	0.7336
3/5/2010	148.18	0.6634	14,069.00	0.7355
3/6/2010	147.85	0.6604	14,004.00	0.7337
3/7/2010	147.85	0.6604	14,001.70	0.7337
3/8/2010	148.3	0.6613	13,984.40	0.7323
3/9/2010	148.244	0.6673	14,028.00	0.7358
3/10/2010	148.07	0.669	13,955.70	0.7344
3/11/2010	148.25	0.6657	13,930.60	0.7321
3/12/2010	148.42	0.6603	13,919.90	0.7275
3/13/2010	150.3	0.6575	13,939.00	0.726
3/14/2010	148.54	0.6575	13,943.00	0.726
3/15/2010	148.33	0.6632	14,014.60	0.7298
3/16/2010	147.81	0.6604	14,012.00	0.7286
3/17/2010	148.08	0.6534	14,010.30	0.7264

3/18/2010	148.501	0.6547	14,103.30	0.7325
3/19/2010	147.73	0.6621	14,051.90	0.7374
3/20/2010	149.45	0.6658	13,998.00	0.7387
3/21/2010	147.98	0.6658	13,997.50	0.7388
3/22/2010	147.66	0.6654	13,939.40	0.7392
3/23/2010	147.307	0.6648	13,898.70	0.7398
3/24/2010	147.29	0.6697	13,993.60	0.7482
3/25/2010	147.556	0.6716	13,990.80	0.7504
3/26/2010	147.78	0.6723	13,955.90	0.7472
3/27/2010	147.73	0.671	13,967.00	0.7468
3/28/2010	147.73	0.6707	13,958.20	0.745
3/29/2010	148.01	0.6677	13,957.10	0.7425
3/30/2010	148.617	0.6635	13,995.80	0.7432
3/31/2010	149.303	0.66	13,992.70	0.7414
4/1/2010	148.924	0.6557	13,987.50	0.7391
4/2/2010	148.485	0.6559	14,007.10	0.7385
4/3/2010	149.7	0.6574	14,027.00	0.7406
4/4/2010	147.88	0.6572	14,029.10	0.7406
4/5/2010	148.462	0.655	14,040.50	0.7413
4/6/2010	149.75	0.6575	14,071.50	0.7463
4/7/2010	149.255	0.6565	14,018.60	0.7481
4/8/2010	148.747	0.6567	13,994.80	0.7502
4/9/2010	149.554	0.6515	13,921.30	0.7451
4/10/2010	150.4	0.6503	13,871.00	0.7405
4/11/2010	149.75	0.6503	13,872.00	0.7408
4/12/2010	148.866	0.6489	13,875.70	0.7349
4/13/2010	147.789	0.65	13,989.30	0.7359
4/14/2010	149.154	0.6475	13,944.80	0.733
4/15/2010	149.142	0.6462	13,973.70	0.7366
4/16/2010	148.633	0.6485	13,988.40	0.7395
4/17/2010	148.477	0.6507	14,019.00	0.7404
4/18/2010	150	0.6509	14,014.80	0.7406
4/19/2010	148.28	0.655	13,981.70	0.7431
4/20/2010	148.36	0.6509	13,978.50	0.7425
4/21/2010	148.27	0.6494	14,011.80	0.7459
4/22/2010	148.03	0.6493	14,016.80	0.7494
4/23/2010	148.729	0.6511	13,992.10	0.751
4/24/2010	148.729	0.6501	13,886.00	0.7472
4/25/2010	150	0.6501	13,887.50	0.7473
4/26/2010	149.106	0.647	13,944.80	0.7495
4/27/2010	142.679	0.6517	13,999.00	0.7519
4/28/2010	149.12	0.6576	14,056.70	0.758
4/29/2010	149.881	0.6555	13,975.80	0.7554
4/30/2010	150.138	0.6527	13,965.10	0.752
5/1/2010	148.8	0.6544	13,992.30	0.7521
5/2/2010	148.8	0.6543	13,989.20	0.7517
5/3/2010	148.42	0.6557	14,007.00	0.7565
5/4/2010	150.194	0.6587	14,047.70	0.7638
5/5/2010	148.5	0.661	14,106.20	0.7749
5/6/2010	149.618	0.6668	14,113.50	0.7852

5/7/2010	148.67	0.6792	14,043.90	0.7864
5/8/2010	149.249	0.6758	14,011.70	0.7836
5/9/2010	149.25	0.6756	14,022.10	0.7824
5/10/2010	150.099	0.6699	13,976.60	0.7746
5/11/2010	148.29	0.6727	14,115.20	0.7867
5/12/2010	148.36	0.6714	14,021.80	0.79
5/13/2010	149.154	0.6781	14,028.40	0.7939
5/14/2010	148.44	0.6864	14,104.20	0.8024
5/15/2010	148.44	0.6877	14,104.20	0.8088
5/16/2010	149.86	0.6878	14,131.00	0.8091
5/17/2010	149.843	0.6933	14,083.60	0.811
5/18/2010	151.12	0.6928	14,154.10	0.8095
5/19/2010	149.43	0.6975	13,998.00	0.8151
5/20/2010	149.05	0.6967	14,039.80	0.8067
5/21/2010	149.745	0.6934	14,084.10	0.7963
5/22/2010	149.745	0.6904	14,084.10	0.795
5/23/2010	148.65	0.6905	14,027.00	0.7948
5/24/2010	149.646	0.6927	14,093.80	0.8051
5/25/2010	150.641	0.6974	14,210.70	0.8156
5/26/2010	149.58	0.6945	14,083.40	0.8153
5/27/2010	151.009	0.6888	14,223.10	0.813
5/28/2010	149.452	0.6885	14,101.90	0.8101
5/29/2010	150.86	0.6918	14,251.00	0.8145
5/30/2010	149.75	0.6919	14,290.00	0.8146
5/31/2010	149.01	0.6892	14,096.00	0.8128
6/1/2010	150.835	0.6857	14,219.70	0.8178
6/2/2010	149.666	0.682	14,112.20	0.8181
6/3/2010	150.21	0.6822	14,220.80	0.817
6/4/2010	150.249	0.6862	14,197.30	0.8273
6/5/2010	150.655	0.6916	14,252.50	0.8352
6/6/2010	150.65	0.6917	14,252.00	0.8354
6/7/2010	148.71	0.691	14,376.50	0.8372
6/8/2010	148.81	0.6928	14,075.00	0.837
6/9/2010	149.654	0.6891	14,091.40	0.8339
6/10/2010	148.79	0.6834	14,092.00	0.8281
6/11/2010	150.152	0.6842	14,258.30	0.8263
6/12/2010	150.152	0.687	14,258.30	0.8253
6/13/2010	148.9	0.687	14,124.00	0.8253
6/14/2010	150.138	0.6793	14,247.00	0.8176
6/15/2010	148.48	0.677	14,114.30	0.815
6/16/2010	148.8	0.6756	14,121.90	0.8125
6/17/2010	148.81	0.6768	14,138.40	0.8099
6/18/2010	150.044	0.6747	14,293.80	0.8078
6/19/2010	149.128	0.6752	14,154.00	0.8082
6/20/2010	148.85	0.6743	14,127.00	0.8065
6/21/2010	149.565	0.674	14,250.70	0.8073
6/22/2010	150.249	0.6766	14,187.90	0.8135
6/23/2010	149.26	0.6713	14,155.00	0.8147
6/24/2010	148.87	0.6681	14,190.40	0.8118
6/25/2010	148.683	0.6683	14,178.10	0.8113

6/26/2010	148.12	0.6638	14,178.10	0.8081
6/27/2010	148.12	0.6638	14,078.00	0.8081
6/28/2010	148.4	0.6634	14,165.30	0.8107
6/29/2010	147.94	0.6635	14,186.80	0.8191
6/30/2010	148.555	0.6666	14,101.50	0.8163
7/1/2010	147.59	0.6655	14,193.80	0.8089
7/2/2010	147.914	0.6585	14,168.60	0.7975
7/3/2010	146.53	0.658	14,180.00	0.7955
7/4/2010	146.53	0.6581	14,180.00	0.7955
7/5/2010	147.73	0.6603	14,243.00	0.7977
7/6/2010	148.369	0.659	14,222.80	0.7937
7/7/2010	147.802	0.6599	14,245.20	0.7935
7/8/2010	148.601	0.6597	14,262.90	0.7895
7/9/2010	148.33	0.661	14,327.00	0.7898
7/10/2010	148.33	0.6636	14,327.00	0.7908
7/11/2010	148.63	0.6636	14,356.00	0.7909
7/12/2010	148.19	0.6658	14,133.00	0.7947
7/13/2010	149.949	0.6621	14,429.90	0.7917
7/14/2010	147.419	0.6558	14,255.80	0.7858
7/15/2010	148.961	0.6515	14,287.70	0.7792
7/16/2010	148.493	0.6508	14,326.70	0.7726
7/17/2010	148.597	0.6534	14,271.50	0.7731
7/18/2010	148.6	0.6534	14,272.00	0.7732
7/19/2010	148.869	0.655	14,401.90	0.7723
7/20/2010	149.342	0.6559	14,369.50	0.7738
7/21/2010	147.88	0.6562	14,375.00	0.7795
7/22/2010	148	0.6559	14,299.00	0.778
7/23/2010	148.671	0.6502	14,452.80	0.7758
7/24/2010	148.476	0.6481	14,464.60	0.7743
7/25/2010	150	0.6482	14,545.00	0.7744
7/26/2010	147.83	0.6458	14,016.00	0.7723
7/27/2010	147.75	0.6438	14,354.00	0.7693
7/28/2010	147.88	0.6411	14,271.50	0.7691
7/29/2010	148.307	0.6401	14,253.90	0.7654
7/30/2010	149.385	0.6392	14,321.40	0.7667
7/31/2010	149.385	0.6372	14,321.40	0.766
8/1/2010	147.63	0.6371	14,166.00	0.7659
8/2/2010	147.41	0.6317	14,184.00	0.7622
8/3/2010	147.73	0.6278	14,225.90	0.7566
8/4/2010	147.89	0.6284	14,168.00	0.7582
8/5/2010	149.137	0.6298	14,240.30	0.7591
8/6/2010	149.045	0.628	14,230.70	0.7556
8/7/2010	146.51	0.6271	13,987.00	0.7528
8/8/2010	146.51	0.627	13,987.00	0.7527
8/9/2010	148.05	0.6271	14,157.00	0.7543
8/10/2010	149.325	0.6333	14,226.80	0.7606
8/11/2010	148.779	0.6362	14,137.20	0.7703
8/12/2010	149.398	0.6403	14,194.80	0.7781
8/13/2010	148.842	0.6406	14,242.90	0.7805
8/14/2010	148.842	0.6412	14,242.90	0.7838

8/15/2010	148.02	0.6411	14,189.00	0.7837
8/16/2010	147.87	0.6397	14,141.00	0.7801
8/17/2010	148.13	0.6401	14,162.00	0.7768
8/18/2010	148.2	0.6409	14,161.00	0.7771
8/19/2010	148.19	0.6408	14,148.00	0.7795
8/20/2010	148.04	0.6438	14,140.00	0.7852
8/21/2010	148.21	0.6436	14,158.00	0.7864
8/22/2010	147.88	0.6436	14,126.00	0.7865
8/23/2010	148.34	0.6431	14,163.00	0.7878
8/24/2010	148.615	0.6482	14,185.00	0.7907
8/25/2010	148.57	0.6479	14,163.00	0.7902
8/26/2010	148.83	0.6433	14,165.00	0.7867
8/27/2010	149	0.6447	14,150.00	0.7857
8/28/2010	148.47	0.6437	14,100.00	0.7832
8/29/2010	148.43	0.6437	14,096.00	0.7832
8/30/2010	149.19	0.6446	14,138.00	0.7872
8/31/2010	150.42	0.6497	14,240.90	0.7885
9/1/2010	149.491	0.6483	14,182.60	0.7824
9/2/2010	150.086	0.6494	14,262.40	0.7803
9/3/2010	149.075	0.648	14,196.60	0.778
9/4/2010	147.92	0.6469	14,101.00	0.7752
9/5/2010	150.25	0.647	14,330.00	0.7752
9/6/2010	148.75	0.6488	14,143.00	0.776
9/7/2010	150.16	0.6513	14,246.10	0.7843
9/8/2010	149.26	0.647	14,141.00	0.7864
9/9/2010	149.4	0.648	14,129.00	0.7867
9/10/2010	149.08	0.6492	14,106.00	0.7866
9/11/2010	149.67	0.6506	14,162.00	0.7882
9/12/2010	151.4	0.6504	14,290.00	0.7881
9/13/2010	150.445	0.6481	14,252.50	0.7791
9/14/2010	150.164	0.6467	14,148.80	0.7739
9/15/2010	150.605	0.6428	14,200.20	0.7696
9/16/2010	150.701	0.6405	14,187.60	0.7658
9/17/2010	151.336	0.6389	14,247.10	0.7642
9/18/2010	150.25	0.6395	14,168.00	0.766
9/19/2010	150.35	0.6395	14,177.00	0.7659
9/20/2010	151.18	0.6411	14,229.70	0.7646
9/21/2010	151.882	0.6427	14,205.00	0.7611
9/22/2010	152.339	0.6384	14,224.40	0.7485
9/23/2010	151.39	0.6376	14,101.00	0.7489
9/24/2010	151.732	0.6348	14,112.00	0.7451
9/25/2010	153.2	0.6317	14,112.00	0.7409
9/26/2010	150.63	0.6317	14,029.00	0.7409
9/27/2010	153.787	0.6315	14,196.70	0.7424
9/28/2010	153.406	0.6323	14,233.60	0.7408
9/29/2010	152.38	0.6326	14,127.50	0.7347
9/30/2010	152.22	0.6329	14,035.00	0.7339
10/1/2010	152.07	0.6326	14,021.40	0.7285
10/2/2010	154.5	0.6319	14,021.40	0.7248
10/3/2010	151.37	0.6321	13,938.00	0.7248

10/4/2010	152.16	0.632	14,096.70	0.7294
10/5/2010	151.692	0.63	14,204.50	0.7255
10/6/2010	150.768	0.6291	14,174.90	0.7205
10/7/2010	149.24	0.6285	14,123.00	0.7172
10/8/2010	150.418	0.6287	14,215.20	0.7184
10/9/2010	148.829	0.6263	14,063.90	0.7172
10/10/2010	151.2	0.6265	14,063.90	0.717
10/11/2010	150.617	0.6281	14,179.90	0.7181
10/12/2010	149.99	0.6316	14,138.40	0.722
10/13/2010	151.206	0.6306	14,222.80	0.7161
10/14/2010	151.161	0.6247	14,246.70	0.7104
10/15/2010	150.24	0.6239	14,149.00	0.7121
10/16/2010	151.28	0.6251	14,218.00	0.7152
10/17/2010	151.161	0.6252	14,219.40	0.7151
10/18/2010	149.344	0.6287	14,220.80	0.7181
10/19/2010	151.326	0.634	14,173.30	0.7221
10/20/2010	150.937	0.6337	14,198.10	0.7208
10/21/2010	149.89	0.6346	14,195.40	0.7158
10/22/2010	150.829	0.6371	14,199.40	0.7179
10/23/2010	150.829	0.6375	14,199.40	0.7163
10/24/2010	149.29	0.6375	14,041.00	0.7162
10/25/2010	147.716	0.6352	13,845.10	0.7135
10/26/2010	148.05	0.6318	14,149.40	0.719
10/27/2010	149.234	0.6328	14,217.50	0.7248
10/28/2010	147.86	0.6297	14,171.70	0.7207
10/29/2010	148.21	0.6266	14,192.30	0.7198
10/30/2010	150.75	0.6233	14,192.30	0.7167
10/31/2010	147.18	0.6233	14,087.00	0.7167
11/1/2010	149.518	0.623	14,342.40	0.7171
11/2/2010	149.048	0.6236	14,324.20	0.7161
11/3/2010	148.09	0.6223	14,267.30	0.7124
11/4/2010	147.68	0.618	14,142.20	0.7052
11/5/2010	148.632	0.616	14,210.30	0.7069
11/6/2010	148.632	0.6177	14,210.30	0.7124
11/7/2010	150.4	0.6177	14,350.00	0.7124
11/8/2010	148.217	0.6194	14,192.80	0.7164
11/9/2010	149.354	0.6206	14,292.60	0.7202
11/10/2010	148.09	0.6237	14,241.70	0.7264
11/11/2010	148.31	0.6198	14,208.70	0.7279
11/12/2010	149.173	0.621	14,233.20	0.7318
11/13/2010	148.63	0.6203	14,230.00	0.7302
11/14/2010	150.65	0.6203	14,390.00	0.7302
11/15/2010	149.405	0.6212	14,326.40	0.7323
11/16/2010	148.6	0.6249	14,183.00	0.7365
11/17/2010	148.627	0.6291	14,186.20	0.7404
11/18/2010	149.436	0.6266	14,328.80	0.7353
11/19/2010	148.33	0.6242	14,247.50	0.7319
11/20/2010	150.3	0.6256	14,415.00	0.7311
11/21/2010	150.3	0.6256	14,415.00	0.7309
11/22/2010	148.656	0.6252	14,318.10	0.7299

11/23/2010	149.62	0.6287	14,390.20	0.7392
11/24/2010	148.35	0.6331	14,268.20	0.7479
11/25/2010	148.02	0.6343	14,272.00	0.7497
11/26/2010	148.33	0.6372	14,289.10	0.753
11/27/2010	148.33	0.6413	14,289.10	0.7549
11/28/2010	150.6	0.6412	14,475.00	0.7546
11/29/2010	149.994	0.6416	14,414.80	0.758
11/30/2010	148.51	0.643	14,321.90	0.7654
12/1/2010	149.573	0.6415	14,269.70	0.7659
12/2/2010	150.112	0.6409	14,387.40	0.7602
12/3/2010	148.84	0.639	14,355.40	0.7537
12/4/2010	146.97	0.6336	14,115.00	0.7451
12/5/2010	151.05	0.6337	14,470.00	0.7451
12/6/2010	148.84	0.6361	14,208.10	0.7502
12/7/2010	150.308	0.6344	14,479.80	0.7501
12/8/2010	150.627	0.6346	14,385.10	0.7554
12/9/2010	150.212	0.6333	14,374.90	0.7544
12/10/2010	149.254	0.6331	14,457.80	0.7553
12/11/2010	149.254	0.6326	14,457.80	0.7558
12/12/2010	151.78	0.6326	14,580.00	0.7558
12/13/2010	150.666	0.633	14,509.50	0.754
12/14/2010	149.9	0.6313	14,453.50	0.7457
12/15/2010	150.85	0.6372	14,456.50	0.7504
12/16/2010	151.94	0.6418	14,485.30	0.7562
12/17/2010	153.836	0.6416	14,539.40	0.7547
12/18/2010	153.21	0.6435	14,545.00	0.7579
12/19/2010	153.21	0.6435	14,545.00	0.758
12/20/2010	152.744	0.6441	14,520.50	0.7605
12/21/2010	153.525	0.6448	14,598.10	0.7607
12/22/2010	152.416	0.6475	14,652.10	0.7624
12/23/2010	149.62	0.649	14,415.30	0.763
12/24/2010	150.782	0.6475	14,444.50	0.762
12/25/2010	150.782	0.6476	14,444.50	0.7618
12/26/2010	153.73	0.6475	14,470.00	0.7618
12/27/2010	151.523	0.6485	14,556.00	0.7612
12/28/2010	151.47	0.6478	14,590.80	0.7578
12/29/2010	151.22	0.6489	14,639.60	0.7609
12/30/2010	151.774	0.6464	14,801.10	0.7545
12/31/2010	149.86	0.645	14,741.30	0.7495
1/1/2011	149.86	0.6404	14,741.30	0.747
1/2/2011	149.84	0.6405	14,666.00	0.7473
1/3/2011	149.89	0.6445	14,685.80	0.7503
1/4/2011	150.13	0.6433	14,705.70	0.7489
1/5/2011	150.541	0.643	14,713.30	0.7554
1/6/2011	150.616	0.6451	14,664.70	0.763
1/7/2011	149.81	0.6459	14,638.80	0.7705
1/8/2011	149.81	0.643	14,638.80	0.7744
1/9/2011	150.47	0.6431	14,691.00	0.7746
1/10/2011	151.978	0.6434	14,654.10	0.7743
1/11/2011	150.278	0.6421	14,640.20	0.7722

1/12/2011	150.077	0.6387	14,606.60	0.768
1/13/2011	151.34	0.6335	14,634.00	0.7579
1/14/2011	151.52	0.6309	14,719.20	0.7486
1/15/2011	153.2	0.63	14,719.20	0.7468
1/16/2011	153.2	0.6299	14,840.00	0.7467
1/17/2011	151.15	0.6298	14,731.40	0.7509
1/18/2011	151.59	0.6267	14,808.90	0.7491
1/19/2011	150.13	0.6252	14,793.70	0.7432
1/20/2011	150.53	0.6272	14,874.00	0.7429
1/21/2011	150.094	0.6274	14,820.20	0.7393
1/22/2011	149.564	0.6248	14,792.40	0.7341
1/23/2011	152.25	0.6249	15,100.00	0.7342
1/24/2011	152.5	0.626	15,052.10	0.7348
1/25/2011	152.5	0.6292	15,140.80	0.733
1/26/2011	150.648	0.6313	15,229.40	0.7307
1/27/2011	149.76	0.628	15,230.00	0.7294
1/28/2011	149.82	0.6294	15,220.00	0.7304
1/29/2011	151.1	0.6305	15,349.00	0.7344
1/30/2011	152.06	0.6306	15,500.00	0.7346
1/31/2011	151.281	0.6287	15,550.30	0.733
2/1/2011	150.04	0.6217	15,421.90	0.7274
2/2/2011	150.2	0.6184	15,581.00	0.7235
2/3/2011	150.634	0.6177	15,593.50	0.7274
2/4/2011	150.23	0.6204	15,567.00	0.7345
2/5/2011	150.725	0.6205	15,605.50	0.736
2/6/2011	150.74	0.6206	15,607.00	0.7362
2/7/2011	151.489	0.6203	15,368.50	0.7365
2/8/2011	150.49	0.6207	15,181.80	0.7339
2/9/2011	151.535	0.622	15,122.40	0.732
2/10/2011	150.13	0.6217	15,029.00	0.732
2/11/2011	150.19	0.6235	14,962.40	0.7376
2/12/2011	152.4	0.6246	14,962.40	0.7377
2/13/2011	149.87	0.6246	14,934.00	0.7379
2/14/2011	150.876	0.6239	14,883.10	0.7408
2/15/2011	150.23	0.6217	14,887.00	0.7403
2/16/2011	150.395	0.6207	14,927.90	0.7393
2/17/2011	150.58	0.6202	14,899.00	0.7364
2/18/2011	150.52	0.6172	14,786.90	0.7344
2/19/2011	149.09	0.6151	14,587.00	0.73
2/20/2011	153.73	0.615	14,910.00	0.7299
2/21/2011	151.32	0.616	14,867.00	0.7309
2/22/2011	151.749	0.6185	14,872.00	0.7335
2/23/2011	150.851	0.6173	14,866.70	0.7292
2/24/2011	150.86	0.6177	14,927.50	0.726
2/25/2011	151.446	0.6205	14,921.80	0.725
2/26/2011	151.446	0.6202	14,921.80	0.7268
2/27/2011	152.5	0.6203	14,950.00	0.7268
2/28/2011	151.1	0.6184	14,957.50	0.7254
3/1/2011	151.3	0.6143	14,983.60	0.7239
3/2/2011	152.047	0.6143	15,014.20	0.7243

3/3/2011	152.933	0.6135	15,029.20	0.7201
3/4/2011	152.11	0.6146	14,944.50	0.7159
3/5/2011	154.2	0.6146	14,825.00	0.715
3/6/2011	154.2	0.6145	15,175.00	0.7146
3/7/2011	154.25	0.6151	15,200.00	0.7148
3/8/2011	152.924	0.618	15,113.70	0.7174
3/9/2011	152.638	0.618	15,122.30	0.7195

Appendix E

Simulation of data from the SV model

Simulation 1

```
function [y,x]=sim_model1(phi,Q,alpha,q1,R0,R1,p,n)
% sim_model1 : give stationary random sample from the sv model(2): normal
mixtures
% y(t)=alpha+x(t)+v(t)
% x(t)=phi*x(t-1)+w(t)
% w(t) Gaussian (0,Q)
% v(t) mixtures of two normals N(mu1,R1),N(mu0,R0) mixing prob=p
w=randn(1,n+10000)*sqrt(Q);
Ind=(rand(1,n+10000)<p);
nor1=randn(1,n+10000);
v1=Ind.*(nor1*sqrt(R1)+mu1)+(1-Ind).*(nor1*sqrt(R0));
x0=randn(1);
x(1)=phi*x0+w(1);
for t=2:n+10000
x(t)=phi*x(t-1)+w(t);
end
y=alpha+x+v1;
y=y(10001:n+10000);
x=x(10001:n+10000);
t=1:n;
figure(3),plot(t,x,t,y)
legend('x_1','y_1')
figure(4),hist(y)
```

Simulation 2

```
function [y,x]=sim_model2(phi,Q,alpha,n)
% randsp : give random sample from the sv model
% y(t)=alpha+a(t)*x(t)+v(t)
% x(t)=phi*x(t-1)+w(t)
% w(t) Gaussian (0,Q)
% n=1000; number of observation
% v(t) student-t distribution

w=randn(1,n)*sqrt(Q);
v1=abs(trnd(3,1,n));
x0=randn(1);
x(1)=phi*x0+ w(1);
for t=2:n
    x(t)=phi*x(t-1)+w(t);
```

```

end
y=alpha+x+v1;

y=y(1:n);
x=x(1:n);

nn=1:n;
figure(1),plot(nn,x,nn,y)

legend('x_1','y_1')
figure(2),hist(y)

Simulation 3

function [y,x]=sim_model3(phi,Q,alpha,n,v)
% randsp : give random sample from the sv model
% y(t)=alpha+a(t)*x(t)+v(t)
% x(t)=phi*x(t-1)+w(t)
% w(t) Gaussian (0,Q)
% n=1000; number of observation
% v(t) generalized error distribution

w=randn(1,n+10000)*sqrt(Q);
v1=grnd(v,1,n+10000);
x0=randn(1);
x(1)=phi*x0+w(1);
for t=2:n+10000
    x(t)=phi*x(t-1)+w(t);
end
y=alpha+x+v1;
y=y(10001:n+10000);
x=x(10001:n+10000);
nn=1:n;
figure(1),plot(nn,x,nn,y)

legend('x_1','y_1')
figure(2),hist(y)

```

Parameter Estimation for on technique based on the normal mixture and Student-t on data generated from the normal mixture model

```

clc
clear
n=1000;
sim_model(0.9500, 1.0729,-0.6794, -3.6794, 4.000, 4.000, 0.5000, n);

Iparm=Iparmnorm(y);
a=~isnan(y);
Tol=0.001;
Miter=100;
Npar=1000;

tt=1;

```

```

[Eparm,Vparm,Vparmt,RelLike,Tcal,Niter]=mainnorm(y,a,Iparm,Tol,Miter,Npar,n,t
t)

disp('USING TECHNIQUE BASED ON THE NORMAL MIXTURE ON NORMAL SV MODEL')

disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i)  Q(i)  q_0(i)  q_1(i)  R0(i)  R1(i)  pi(i)  Relative
Likelihood')
disp([IniP])
disp([Results])
diary off


figure(1),subplot(4,1,1),plot(1:Niter,Eparm(:,1)),ylabel('\phi'),xlabel('iter
ation(i)')
subplot(4,1,2),plot(1:Niter,Eparm(:,2)),ylabel('Q'),xlabel('iteration(i)')
subplot(4,1,3),plot(1:Niter,Eparm(:,3)),ylabel('\alpha'),xlabel('iteration(i)
')
subplot(4,1,4),plot(1:Niter,RelLike),ylabel('Relative
Likelihood'),xlabel('iteration(i)')


clc
clear
n=1000;
sim_model1(0.9500,1.0729,-2.1496, n);

Iparm=Iparmnorm(y);
a=~isnan(y);
Tol=0.001;
Miter=100;
Npar=1000;

tt=1;
[Eparm,Vparm,Vparmt,RelLike,Tcal,Niter]=mainnorm(y,a,Iparm,Tol,Miter,Npar,n,t
t)
disp('USING TECHNIQUE BASED ON THE NORMAL MIXTURE ON STUDENT-T MODEL')

disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];

```

```

IniP=[0 Iparm];
disp('iteration(i)  phi(i)  Q(i)  q_0(i)  q_1(i)  R0(i)  R1(i)  pi(i)  Relative
Likelihood')
disp([IniP])
disp([Results])
diary off

figure(2),subplot(4,1,1),plot(1:Niter,Eparm(:,1)),ylabel('\phi'),xlabel('iteration(i)')
subplot(4,1,2),plot(1:Niter,Eparm(:,2)),ylabel('Q'),xlabel('iteration(i)')
subplot(4,1,3),plot(1:Niter,Eparm(:,3)),ylabel('\alpha'),xlabel('iteration(i)')
subplot(4,1,4),plot(1:Niter,RelLike),ylabel('Relative
Likelihood'),xlabel('iteration(i)')

clc
clear
n=1000;
sim_model2(0.8214,1.3359,-2.7823,-5.7823,4.000,4.000,0.5000, n,v);

Iparm=Iparm_t(y);
a=~isnan(y);
Tol=0.001;
Miter=100;
Npar=1000;

tt=1;
[Eparm,Vparm,RelLike,Tcal,Niter]=mainchis(y,a,Iparm,Tol,Miter,Npar,n,tt)
disp(' USING TECHNIQUE BASED ON STUDENT-T MODEL ON NORMAL SV MODEL')
disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i)  Q(i)  alpha(i)  Relative Likelihood')
disp([IniP])
disp([Results])
diary off
figure(3),subplot(4,1,1),plot(1:Niter,Eparm(:,1)),ylabel('\phi'),xlabel('iteration(i)')
subplot(4,1,2),plot(1:Niter,Eparm(:,2)),ylabel('Q'),xlabel('iteration(i)')
subplot(4,1,3),plot(1:Niter,Eparm(:,3)),ylabel('\alpha'),xlabel('iteration(i)')
subplot(4,1,4),plot(1:Niter,RelLike),ylabel('Relative
Likelihood'),xlabel('iteration(i)')

```

```

clc
clear
n=1000; v=12;
sim_model1(0.8214, 1.3359, -2.2823, n, v);

Iparm=Iparm_t(y);
a=~isnan(y);
Tol=0.001;
Miter=100;
Npar=1000;
tt=1;
[Eparm, Vparm, RelLike, Tcal, Niter]=mainchis(y, a, Iparm, Tol, Miter, Npar, n, tt)
diary Result_Alg_A_A.txt
disp(' USING TECHNIQUE BASED ON STUDENT-T MODEL ON STUDENT-T SV MODEL ')
disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ', num2str(Niter)])
disp(['Total time = ', num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ', num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i)  Q(i)  alpha(i)  Relative Likelihood')
disp([IniP])
disp([Results])
diary off
figure(4), subplot(4,1,1), plot(1:Niter, Eparm(:,1)), ylabel('\phi'), xlabel('iteration(i)')
subplot(4,1,2), plot(1:Niter, Eparm(:,2)), ylabel('Q'), xlabel('iteration(i)')
subplot(4,1,3), plot(1:Niter, Eparm(:,3)), ylabel('\alpha'), xlabel('iteration(i)')
subplot(4,1,4), plot(1:Niter, RelLike), ylabel('Relative Likelihood'), xlabel('iteration(i)')

```

Parameter estimation on technique based on the normal mixture SV and GED on data generated from the normal mixture model

```

clc
clear
n=1000;
sim_model2(0.8699, 3.6899, -4.8897, -7.8897, 4.000, 4.000, 0.5000, n);

Iparm=Iparmnorm(y);
a=~isnan(y);
Tol=0.001;
Miter=100;
Npar=1000;

tt=1;
[Eparm, Vparm, Vparmt, RelLike, Tcal, Niter]=mainnorm(y, a, Iparm, Tol, Miter, Npar, n, t)

```

```

disp(' USING TECHNIQUE BASED ON NORMAL MIXTURE ON GED SV MODEL ')

disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i)  Q(i)  m_0(i)  m_1(i)  R0(i)  R1(i)  pi(i)  Relative
Likelihood')
disp([IniP])
disp([Results])
diary off


figure(5),subplot(4,1,1),plot(1:Niter,Eparm(:,1)),ylabel('\phi'),xlabel('iteration(i)')
subplot(4,1,2),plot(1:Niter,Eparm(:,2)),ylabel('Q'),xlabel('iteration(i)')
subplot(4,1,3),plot(1:Niter,Eparm(:,3)),ylabel('\alpha'),xlabel('iteration(i)')
subplot(4,1,4),plot(1:Niter,RelLike),ylabel('Relative
Likelihood'),xlabel('iteration(i)')


clc
clear
n=1000; v=12;
[y_C, x_C] = sim_model3(0.8699, 3.6899, -4.4365, n,v);

Iparm=Iparm_g(y);
a=~isnan(y);
Tol=0.1;
Miter=100;
Npar=500;
tt=1;
[Eparm,Vparm,RelLike,Tcal,Niter]=maintndist(y,a,Iparm,Tol,Miter,Npar,n,tt,v)
diary Result_Algo_C_tdist.txt
disp(' USING TECHNIQUE BASED ON GED ON NORMAL MIXTURE SV MODEL ')
disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i)  Q(i)  alpha(i)  Relative Likelihood')
disp([IniP])
disp([Results])
diary off

```



```

figure(6),subplot(4,1,1),plot(1:Niter,Eparm(:,1)),ylabel('\phi'),xlabel('iteration(i)')
subplot(4,1,2),plot(1:Niter,Eparm(:,2)),ylabel('Q'),xlabel('iteration(i)')
subplot(4,1,3),plot(1:Niter,Eparm(:,3)),ylabel('\alpha'),xlabel('iteration(i)')
subplot(4,1,4),plot(1:Niter,RelLike),ylabel('Relative Likelihood'),xlabel('iteration(i)')

clc
clear
n=1000; v=12;
[y_A, x_A] = sim_model3(0.9500, 1.3288, 0.6309, n,v);

Iparm=Iparm_g(y_A);
a=~isnan(y_A);
Tol=0.1;
Miter=100;
Npar=500;
tt=1;
[Eparm,Vparm,RelLike,Tcal,Niter]=maintndist(y_A,a,Iparm,Tol,Miter,Npar,n,tt,v)
diary Result_Algo_C_tdist.txt
disp(' USING TECHNIQUE BASED ON GED ON GED SV MODEL ')
disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i)  Q(i)  alpha(i)  Relative Likelihood')
disp([IniP])
disp([Results])
diary off
figure(7),subplot(4,1,1),plot(1:Niter,Eparm(:,1)),ylabel('\phi'),xlabel('iteration(i)')
subplot(4,1,2),plot(1:Niter,Eparm(:,2)),ylabel('Q'),xlabel('iteration(i)')
subplot(4,1,3),plot(1:Niter,Eparm(:,3)),ylabel('\alpha'),xlabel('iteration(i)')
subplot(4,1,4),plot(1:Niter,RelLike),ylabel('Relative Likelihood'),xlabel('iteration(i)')

```

Parameter estimation for real life financial data

```

clc
clear all

yn=xlsread('exchangeratedata.xls','NGN_US','H5:H183');

y=yn';
n=length(y);

```

```

Iparm=Iparmnorm(y);
a=~isnan(y);
Tol=0.1;
Miter=100;

tt=1;
[Eparm,Vparm,Vparmt,RelLike,Tcal,Niter]=mainnorm(y,a,Iparm,Tol,Miter,Npar,n,t
t)

disp('USING NORMAL MIXTURE')

disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i) Q(i)  m_0(i) m_1(i) R0(i) R1(i) pi(i) Relative
Likelihood')
disp([IniP])
disp([Results])
diary off


clc
clear all
%yn=xlsread('Bankdata.xlsx');
yn=xlsread('exchangeratedata.xls','NGN_US','H5:H183');

y=yn';
n=length(y);


Iparm=Iparm_c(y);
a=~isnan(y);
Tol=0.001;
Miter=100;
tt=1;
[Eparm,Vparm,RelLike,Tcal,Niter]=mainchis(y,a,Iparm,Tol,Miter,Npar,n,tt)
disp('USING STUDENT-T')
disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i) Q(i)  alpha(i)  Relative Likelihood')
disp([IniP])

```

```

disp([Results])
diary off
figure(8),subplot(4,1,1),plot(1:Niter,Eparm(:,1)),ylabel('\phi'),xlabel('iteration(i)')
subplot(4,1,2),plot(1:Niter,Eparm(:,2)),ylabel('Q'),xlabel('iteration(i)')
subplot(4,1,3),plot(1:Niter,Eparm(:,3)),ylabel('\alpha'),xlabel('iteration(i)')
subplot(4,1,4),plot(1:Niter,RelLike),ylabel('Relative Likelihood'),xlabel('iteration(i)')

clc
clear all
%yn=xlsread('Bankdata.xlsx');
yn=xlsread('exchangeratedata.xls','NGN_US','I5:I183');

y=yn';
n=length(y);

Iparm=Iparm_c(y);
a=~isnan(y);
Tol=0.001;
Miter=100;
Npar=1000;
tt=1;
[Eparm,Vparm,RelLike,Tcal,Niter]=mainchis(y,a,Iparm,Tol,Miter,Npar,n,tt)
disp('USING GENERALIZED ERROR DISTRIBUTION')
disp('Variance Estimate')
disp(Vparm)
disp(['No of Iteration = ',num2str(Niter)])
disp(['Total time = ',num2str(Tcal)]);
stdev=std(Eparm);
disp(['Standard deviation of final estimate = ',num2str(stdev)])
nn1=1:Niter;
Results=[nn1' Eparm RelLike];
IniP=[0 Iparm];
disp('iteration(i)  phi(i)  Q(i)  alpha(i)  Relative Likelihood')
disp([IniP])
disp([Results])
diary off
figure(9),subplot(4,1,1),plot(1:Niter,Eparm(:,1)),ylabel('\phi'),xlabel('iteration(i)')
subplot(4,1,2),plot(1:Niter,Eparm(:,2)),ylabel('Q'),xlabel('iteration(i)')
subplot(4,1,3),plot(1:Niter,Eparm(:,3)),ylabel('\alpha'),xlabel('iteration(i)')
subplot(4,1,4),plot(1:Niter,RelLike),ylabel('Relative Likelihood'),xlabel('iteration(i)')

```

Summary statistics for real life financial data

```

clc
clear
diary ResultsNew.txt
X=xlsread('exchangeratenew1.xls','exchangeratenew','B5:M734');
NN=X(:,2);

```

```

BP=X(:,5);
GHC=X(:,8);
EUR=X(:,11);

disp(' Results for Nigeria Naira - USD')
[Coeff,Error,LLF,Innov,Sigmas,Summary]=st_svfit(NN)
disp(' ')
disp(' Results for British Pound - USD ')
[Coeff2,Error2,LLF2,Innov2,Sigmas2,Summary2]=st_svfit(BP)

disp(' ')
disp(' Results for Ghana Cedi - USD ')
[Coeff3,Error3,LLF3,Innov3,Sigmas3,Summary3]=st_tfit(GHC)

disp(' ')
disp(' Results for EURO - USD ')
[Coeff4,Error4,LLF4,Innov4,Sigmas4,Summary4]=st_tfit(EUR)

disp('AIC')
a=aic([LLF LLF2 LLF3 LLF4],[4 4 4 4])
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%#
st_tfit (NN)
st_tfit (BP)
st_tfit (GHC)
st_tfit (EUR)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%FACTOR ANALYSIS%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp('FACTOR ANALYSIS')
XX=[NN BP GHC EUR];
disp('COVARIANCE')
disp(cov(XX))
[Lambda,Psi,T,stats,F] = factoran(XX,1,'scores','regression')

diary off

```

Scenarios for the BOT.

```

clc
clear
% State = (x y xdot ydot). We only
% observe (x y).
ss = 4; % state size
os = 2; % observation size

F = [1 1 0 0; 0 1 0 0; 0 0 1 1; 0 0 0 1];
H = [1 0 0 0; 0 1 0 0];
Q=[0.0667 0.1 0 0;0.1 0.2 0 0;0 0 0.0667 0.1;0 0 0.2 0.8];

R = 1*eye(os);
initx = [5 5 1 0]';
initV = 10*eye(ss);
seed = 9;
rand('state', seed);

```

```

randn('state', seed);
T =150;
%[x,y] = sample_lds(F, H, Q, R, initx, T);
[x,y] = sample_lds(F, H, Q, R, initx, T);

[xfilt, Vfilt, VVfilt, loglik] = pfilter(y, F, H, Q, R, initx, initV);
[xsmooth, Vsmooth] = kalman_smoother(y, F, H, Q, R, initx, initV);
dfilt = x([1 2],:) - xfilt([1 2],:);
mse_filt = sqrt(sum(sum(dfilt.^2)))
dsmooth = x([1 2],:) - xsmooth([1 2],:);
mse_smooth = sqrt(sum(sum(dsmooth.^2)))
subplot(2,1,1)
hold on
plot(x(1,:), x(2,:), 'ks-');
plot(y(1,:), y(2,:), 'g*');
plot(xfilt(1,:), xfilt(2,:), 'rx:');
% for t=1:T
%     plotstudent_t2d(xfilt(1:2,t), Vfilt(1:2, 1:2, t));
% end
hold off
legend('true', 'observed', 'SMCEM estimate', 0)
xlabel('X')
ylabel('Y')

subplot(2,1,2)
hold on
%plot(x(1,:), x(2,:));
%plot(y(1,:), y(2,:));
plot(xsmooth(1,:), xsmooth(2,:), 'rx:');

% for t=1:T
%     plotstudent-t2d(xsmooth(1:2,t),Vsmooth(1:2, 1:2, t));
% end
hold off
legend('SMCEM', 0)
xlabel('Time')
ylabel('MSE(t)')

```