

Non-linear state-space models for estimating and forecasting financial volatility evolution: disentangling some knots

Antonio A. F. Santos

Faculty of Economics, University of Coimbra, Portugal

`aasantos@fe.uc.pt`

An extended abstract

In this paper, we present the implementation of a simplified algorithm for the estimation and forecasting of the stochastic volatility (SV) model, which is adapted to the statistical structure of the model, and can be easily extended to different versions. It is basically defined using second-order approximations for the distributions of interest, which imply essentially gaussian-type of approximations. The algorithm is easy to implement, allowing that the results can be compared with others in the literature, which have been revealed as being difficult to interpret (Meyer and Yu 2000, Yu 2005, Djegn e and McCausland 2015). The non-linear state-space SV model is commonly used to estimate and forecast the evolution of the financial volatility. As a highly-dimensional non-linear model, analytical results are scarce, and the estimates and forecasts are based on numerical approaches like Markov chain Monte Carlo (MCMC) simulations, used within a Bayesian estimating and forecasting framework. It is certain that the non-linear and non-gaussian

characteristics of the observations equation, coupled with the volatility persistence, represented by an autoregressive linear form, may be treated by very sophisticated MCMC algorithms. However, they are usually not easy to implement, and probably more error prone when they are compared with simpler algorithms. Despite that they can be reliable, they may be difficult to consider for different versions of the basic model, which can include the Student- t innovations in the observations equation, leverage effects with an Asymmetric Stochastic Volatility (ASV), and bivariate versions, which are obtained when additional observations equations are added, for example, to take into account the evolution of the volume (Watanabe 2000), or more recently, to include the evolution of the realised volatility (Koopman and Scharth 2013, Takahashi et al. 2009, Omori and Watanabe 2015).

We begin by giving an example, where in the literature, for the ASV model (Yu 2005), trying to find an analytical formula for the expected future value of the volatility giving the current state of knowledge, the return at t , incorrect results were presented. No analytical solution exists, even less a linear relationship between the financial return at t , and the future expected volatility.

Next, we proceed to the estimation of the model's parameters. Due to the influence of the linear autoregressive form with high persistence of the states equation, in a Bayesian framework, the conditional posterior densities of the states can be very well approximated by a set of univariate gaussian densities. This is true for the basic model, as well as for the several extensions that have been considered. The main task of the algorithm is to find the set of means and variances that will be used to perform the gaussian approximations. Due to the log-concavity of the posterior densities, it is straightforward to find such means and variances by applying successively the Newton's method, or even in the simpler case of the basic model, find them through an analytical formula. Our estimates are compared with some presented in the literature,

which in some cases are distinct from ours (Djegnéné and McCausland 2015), and others clearly incorrect (Meyer and Yu 2000), which highlights the fact that the simpler algorithms allow that coding errors can be detected at earlier stages, avoiding the presentation of incorrect results.

Almost as a by-product of the algorithm used in the estimation process is the one developed for the forecasting of the volatility evolution. Given the values for the parameters, the main aim is to update the filter distribution of the states as new information arrive. The filter distribution is approximated through Particle Filter methods. It is approximated through a set of points, “particles”, and their respective weights. Assuming that a similar set is available to approximate the prior, the posterior (filter) is obtained by incorporating the new information at t . This can be done by using the same kind of gaussian approximations considered in the estimation process, coupled with a Sampling Importance Resampling (SIR) algorithm to define the “particles” that approximate the filter distribution for the states. Algorithms highly elaborated proposed in the literature recently (Li et al. 2016) were used to approximate the filter distribution of the states for a linear-gaussian state-space model, which is essentially useless as they are simply gaussian. In a more realistic scenario, as in the case of the SV model, particle filter methods are necessary because of the lack of analytically tractable results. However, unnecessary complex algorithms are more error prone, and may divert the attention from the main characteristics of the model. We show how with a SIR algorithm, based on the gaussian approximations, for the filter distribution for the states gives reliable results when applied to the SV model.

Our results are compared with the ones presented in the literature, and to allow the reproducibility of the results, code in R was developed. Two series were used to make the relevant comparisons. The first series is the daily Pound/Dollar exchange rates from 01/10/81 to 28/6/85 considered in

Shephard and Pitt (1997), Kim et al. (1998) and Meyer and Yu (2000). A second series is associated with the returns of the S&P500 from January 1980 to December 1987, also considered by Yu (2005) and Djegné and McCausland (2015). We present some dissimilar results from the ones presented in the literature, most of them associated with the ASV model. We show that the structure of the algorithm implemented, allows an easier extension for generalisations of the basic model, and the results obtained are reliable.

References

- Djegné, B. and McCausland, W. J.: 2015, The hessian method for models with leverage-like effects, *Journal of Financial Econometrics* **13**(3), 722–755.
- 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.
- Koopman, S. J. and Scharth, M.: 2013, The analysis of stochastic volatility in the presence of daily realized measures, *Journal of Financial Econometrics* **11**(1), 76–115.
- Li, W., Chen, R. and Tan, Z.: 2016, Efficient sequential monte carlo with multiple proposals and control variates, *Journal of the American Statistical Association* **111**(513), 298–313.
- Meyer, R. and Yu, J.: 2000, Bugs for a bayesian analysis of stochastic volatility models, *The Econometrics Journal* **3**(2), 198–215.
- Omori, Y. and Watanabe, T.: 2015, *Current Trends in Bayesian Methodology with Applications*, CRC Press, chapter Stochastic Volatility and Realized Stochastic Volatility Models, p. 435.

- Shephard, N. and Pitt, M. K.: 1997, Likelihood analysis of non-gaussian measurement time series, *Biometrika* **84**(3), 653–667.
- Takahashi, M., Omori, Y. and Watanabe, T.: 2009, Estimating stochastic volatility models using daily returns and realized volatility simultaneously, *Computational Statistics & Data Analysis* **53**(6), 2404–2426.
- Watanabe, T.: 2000, Bayesian analysis of dynamic bivariate mixture models: Can they explain the behavior of returns and trading volume?, *Journal of Business & Economic Statistics* **18**(2), 199–210.
- Yu, J.: 2005, On leverage in a stochastic volatility model, *Journal of Econometrics* **127**(2), 165–178.

Non-linear state-space models for estimating and forecasting financial volatility evolution: disentangling some knots

Antonio A. F. Santos

Faculty of Economics, University of Coimbra, Portugal

`aasantos@fe.uc.pt`

Abstract

We present two integrated algorithms tailored to the estimation and forecasting associated with the stochastic volatility model. We consider the characteristics of these models and highlight what we think are still some incorrect results presented in the literature. The financial volatility evolution is an important subject that is addressed by many financial econometric models. As it is not directly observable, state-space models have assumed an important role for such analysis. Very sophisticated methods have been presented to estimate the parameters and forecast through these models. Some layers of extra complexity have led to the development of algorithms for estimation and forecasting that are more error prone when compared with simpler versions, which are more robust than the ones aforementioned.

Keywords: Bayesian estimation; Financial Volatility; Markov chain Monte Carlo; Particle Filter; Simulation; Stochastic Volatility

1 Introduction

The level of risk is a key factor that economic agents must take into account. It is even more important to agents that take decisions in financial markets. Measuring correctly the level of risk/volatility associated with the returns of a given asset is important for decision-making in this environment. Professionals and academics have been using different statistical models to measure the level of risk in financial markets. The first family of models was the Autoregressive Conditional Heteroskedastic (ARCH) model proposed by Engle (1982). Since then, several generalisations were proposed as the Generalised ARCH Bollerslev (1986) and the Exponential GARCH Nelson (1991). Another family of models is the stochastic volatility, first proposed by Taylor (1986). At the time, some of the recent econometric developments were not available and the application of the model was limited due to the difficulties in estimating the parameters.

The stochastic volatility (SV) model has been the object of widespread research. The model is used to characterise the volatility evolution associated with financial returns. The volatility (variance of returns) is not directly observable, but it exists empirical evidence that varies over time. Some of the properties found in the empirical distributions of the returns are replicated by the SV model. Extensions of the SV model were considered and some incorrect results associated with the main characteristics of the models, and also with the estimation of the parameters have happened with some frequency.

The estimation of the parameters of the SV model through Bayesian estimation methods has led to the development of very sophisticated algorithms, but also, to the use of some general-use software packages for Bayesian estimation with Markov chain Monte Carlo (MCMC) simulations. However some incorrect results still persist. Finally, forecasting is still a field of ac-

tive research. In this article is presented a simple and unified (estimation and forecasting) version for the algorithms, which are used in the characterisation of the volatility evolution.

A major breakthrough happened with the seminal paper of Jacquier et al. (1994). Since its publication the developments associated with the stochastic volatility model have been enormous. With the computational power available nowadays and the statistical techniques based on MCMC simulations, the main focus is not if we are able to estimate the model, but instead, look for new generalisations able to accommodate different features found in the financial returns.

Several generalisations were proposed, the asymmetric stochastic volatility (ASV) model, fat-tails in the distribution of the innovations, and the multivariate stochastic volatility model. Unfortunately some misleading results have been published. One case is related to the results in Yu (2005), where the ASV model is presented and compared with the version given by Jacquier et al. (2004). The comparison is related to the innovations of the observation and system equations, and how their dependence is modelled. The argument is that the only way to accommodate the efficient market hypothesis, where the returns follow a martingale difference stochastic process, is assuming that the innovation of the measurement equation at t is correlated with the system innovation at $t + 1$, and not for the same period, as was considered in Jacquier et al. (2004). It is also argued that with the formulation proposed, there is a clear and intuitive interpretation for the leverage effect, which can be expressed using an analytical formula. This establishes a linear relationship between returns and expected future volatilities, but within the SV model this result is difficult to understand.

For the estimation of the parameters in this kind of models, which are non-linear state space models, since Jacquier et al. (1994), Bayesian estimation with MCMC is one of the preferred methods. General-purpose software have

been developed to estimate statistical models through Bayesian estimation using MCMC. Perhaps the best known is the BUGS package Spiegelhalter et al. (2003). More recently a more sophisticated package gained substantial attention, the STAN package Stan Development Team (2014). These software packages can be very useful, but sometimes they are used as black-boxes, and incorrect results may not be detected. As an example, we can refer the results in Meyer and Yu (2000), which were obtained using the package BUGS to estimate the ASV model for a series of exchange rates between the British pound and the US dollar. The estimate for the correlation coefficient that characterises the asymmetric effect is incorrect for the aforementioned series.

The estimation of the parameters of the SV model is still an area of active research McCausland (2012), Kastner and Frühwirth-Schnatter (2014) and Djegnéé and McCausland (2015). Very complex algorithms to estimate the parameters of the models were proposed as in McCausland (2012) and Djegnéé and McCausland (2015). Simpler structures give identical results and are less error prone.

The rest of the paper proceeds as follows. In Section 2 we present the two best known SV models, the basic and one assuming the leverage effect, highlighting the incorrect results still found in the literature. In Section 3 a set of algorithms are presented to simulate the states in the SV model, algorithms based on gaussian approximations within the MCMC setting. In Section 4 the results are extended, and a particle filter algorithm is developed, which is used to update the filter distributions for the states for new arriving information. The filter distributions are used to forecast the volatility. In Section 5 the methods proposed in previous sections are tested with two series of financial returns. Finally, Section 6 presents some concluding remarks.

2 Stochastic volatility models

The SV model can be expressed as

$$y_t = \exp\left(\frac{\alpha_t}{2}\right) \varepsilon_t \quad (1)$$

$$\alpha_{t+1} = \mu + \phi \alpha_t + \sigma_\eta \eta_{t+1} \quad (2)$$

where y_t and α_t represent the return and the state, respectively, for $t = 1, 2, \dots, n$. It is the state at t that defines the evolution of the volatility. The vector $(\varepsilon_t, \eta_{t+1})$ follows a standard bivariate normal distribution, $N(\mathbf{0}, \mathbf{I}_2)$. The parameters of the model are given in the vector $\theta = (\mu, \phi, \sigma_\eta)$, where μ represents the mean of the volatility, ϕ the persistence parameter, with $|\phi| < 1$, assuming the stationarity of the volatility evolution, and finally, σ_η represents the volatility of the volatility.

The leverage effect leads to an asymmetric stochastic volatility model. The main characteristic is that the volatility responds with greater intensity to negative than to positive shocks. This is modelled in the ASV model by establishing a dependence between the innovation terms of the observation and system equations, $\text{corr}(\varepsilon_t, \eta_{t+1}) = \rho$, which makes the parameter vector $\theta = (\mu, \phi, \sigma_\eta, \rho)$. The vector of the innovations assumes the distribution $(\varepsilon_t, \eta_{t+1}) \sim N(\mathbf{0}, \Sigma)$, with $\Sigma = \{\{1, \rho \sigma_\eta\}, \{\rho \sigma_\eta, \sigma_\eta^2\}\}$. The model can be rewritten through independent innovations as

$$y_t = \exp\left(\frac{\alpha_t}{2}\right) \varepsilon_t \quad (3)$$

$$\alpha_{t+1} = \mu + \phi \alpha_t + \rho \sigma_\eta \exp\left(-\frac{\alpha_t}{2}\right) y_t + \sigma_\eta \sqrt{1 - \rho^2} \zeta_{t+1}, \quad (4)$$

where $(\varepsilon_t, \zeta_{t+1}) \sim N(\mathbf{0}, \mathbf{I}_2)$.

This kind of model was first proposed by Harvey and Shephard (1996). However, recently, new developments have been considered, mainly due to new estimation techniques and computational advances that have become available. Yu (2005) has made a comparison between different forms of modelling the leverage effect. The comparisons are essentially with the model

presented in Jacquier et al. (2004). There is a straightforward interpretation of the leverage effect, an extreme negative shock, modelled by ε_t , at t , daily shock, implies an increase in the expected volatility in the next period, represented by α_{t+1} , modelled by a negative correlation between ε_t and η_{t+1} , $\text{corr}(\varepsilon_t, \eta_{t+1}) = \rho < 0$. However, Jacquier et al. (2004) assumed that the negative correlation is between ε_t and η_t , which is difficult to interpret, and leads to results that are not compatible with the martingale difference hypothesis for the daily returns.

Due to the assumption of the efficiency of financial markets, it is generally accepted that stock returns follow a martingale difference stochastic process, $E(y_{t+1}|y_t) = 0$. Assuming that $\text{corr}(\varepsilon_t, \eta_t) = \rho$ as in Jacquier et al. (2004), the expression to $E(y_{t+1}|y_t)$, obtained by Yu (2005), is given by

$$\begin{aligned} E(y_{t+1}|y_t) &= E(E(y_{t+1}|y_t, \alpha_t)) \\ &= \frac{1}{2} \rho \sigma_\eta \exp\left(\frac{2 - \phi}{2 - 2\phi} \mu + \frac{2 - \phi^2}{8 - 8\phi^2} \sigma_\eta^2\right), \end{aligned} \quad (5)$$

which is not compatible with the martingale difference hypothesis. Even if it is true that with $\text{corr}(\varepsilon_t, \eta_t) = \rho$, returns do not follow a martingale difference process, the formula presented assumes an incorrect form.

A more difficult to understand claim made by Yu (2005) is that exists a linear relationship between stock returns at t and expected log-variance at $t + 1$, which is formalised by

$$E(\alpha_{t+1}|y_t) = \mu + \frac{\mu \phi}{1 - \phi^2} + \rho \sigma_\eta \exp\left(-\frac{\sigma_\eta^4}{4(1 - \phi^2)^2} + \frac{\sigma_\eta^2 \mu}{(1 - \phi^2)(1 - \phi)}\right) y_t. \quad (6)$$

Assuming that $\rho < 0$, it means that as y_t increases, the expected future volatility decreases. The claim that there is a linear relationship between returns and expected future volatility is a false claim.

2.1 Results reformulation

The claim that in the ASV model $E(y_{t+1}|y_t) = 0$ is correct, assuming that $\text{corr}(\varepsilon_t, \eta_{t+1}) = \rho$. The same result cannot be obtained with $\text{corr}(\varepsilon_t, \eta_t) = \rho$. For the correlation structure as in Jacquier et al. (2004), $\text{corr}(\varepsilon_{t+1}, \eta_{t+1}) = \rho$, it is straightforward to show that

$$E\left(\exp\left(\frac{\sigma_\eta \eta_{t+1}}{2}\right) \varepsilon_{t+1}\right) = \frac{1}{2} \rho \sigma_\eta \exp\left(\frac{\sigma_\eta^2}{8}\right). \quad (7)$$

Considering the conditional expectation given by

$$E(y_{t+1}|y_t, \alpha_t) = \exp\left(\frac{\mu + \phi \alpha_t}{2}\right) E\left(\exp\left(\frac{\sigma_\eta \eta_{t+1}}{2}\right) \varepsilon_{t+1}|y_t, \alpha_t\right), \quad (8)$$

with the expectation in (7), and

$$E\left(\exp\left(\frac{\mu + \phi \alpha_t}{2}\right)\right) = \exp\left(\frac{\phi^2 \sigma_\eta^2 + 4\mu(1 + \phi)}{8(1 - \phi^2)}\right) \quad (9)$$

the correct formulation to $E(y_{t+1}|y_t)$ is given by

$$E(y_{t+1}|y_t) = \frac{1}{2} \rho \sigma_\eta \exp\left(\frac{\sigma_\eta^2 + 4\mu(1 + \phi)}{8(1 - \phi^2)}\right) \quad (10)$$

instead of (5). The conclusion is the same, the returns do not follow a martingale difference process, but the formula presented in Yu (2005) is incorrect. It is straightforward to verify that with $\text{corr}(\varepsilon_t, \eta_{t+1}) = \rho$ the martingale difference hypothesis represented by $E(y_{t+1}|y_t) = 0$ is obtained.

A more serious problem is the one related to the interpretation of the leverage effect as a linear relationship between returns and expected future volatility. To establish the linear relationship between returns and future volatility the conditional expectation $E(\alpha_{t+1}|y_t)$ in Yu (2005) is given by (6), which can be expressed as $E(\alpha_{t+1}|y_t) = a + b y_t$, a linear function of y_t , where a and b are two fixed coefficients. With $\rho < 0$, a negative shock will increase the volatility and a positive one will decrease it, which cannot be true. The

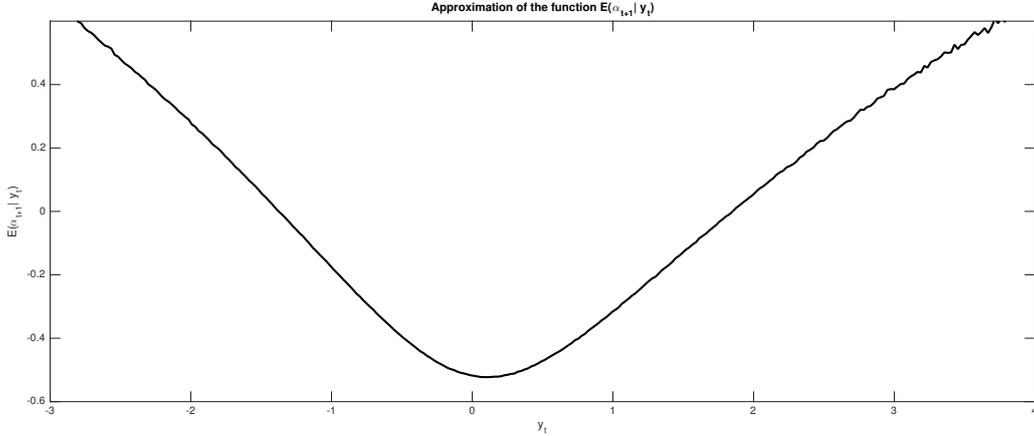


Figure 1: The figure depicts the approximation to the function $E(\alpha_{t+1}|y_t)$ for the ASV model, obtained through simulations with the parameters $\mu = -0.01$, $\phi = 0.97$, $\sigma_\eta = 0.15$ and $\rho = -0.4$.

previous formula, using the model's formulation in (3)-(4) must be obtained from

$$E(\alpha_{t+1}|y_t, \alpha_t) = \mu + \phi \alpha_t + \rho \sigma_\eta \exp\left(-\frac{\alpha_t}{2}\right) y_t, \quad (11)$$

which through the law of iterated expectations gives

$$E(\alpha_{t+1}|y_t) = E(E(\alpha_{t+1}|y_t, \alpha_t)) \quad (12)$$

$$= \mu + \frac{\phi \mu}{1 - \phi} + \rho \sigma_\eta E\left(\exp\left(-\frac{\alpha_t}{2}\right) y_t\right). \quad (13)$$

However, because y_t depends on α_t , $E\left(\exp\left(-\frac{\alpha_t}{2}\right) y_t\right)$ cannot be equal to

$$\exp\left(-\frac{\sigma_\eta^4}{4(1 - \phi^2)^2} + \frac{\sigma_\eta^2 \mu}{(1 - \phi^2)(1 - \phi)}\right) y_t. \quad (14)$$

With a linear relationship, and $\rho < 0$, if y_t assumes a positive extreme value, it would imply a decrease in the expected future volatility, which would be counterintuitive.

There is no analytical formula that relates returns and expected future volatility. The relationship is nonlinear, and even with $\rho < 0$, an increase

in the returns implies an increase of the expected future volatility. As there is no analytical formula, here with an illustrative purpose, the function that relates $E(\alpha_{t+1}|y_t)$ to y_t is approximated numerically for a given parameter vector $\theta = (\mu, \phi, \sigma_\eta, \rho) = (-0.01, 0.97, 0.15, -0.4)$. For the calculations made, depicted in Figure 1, the linear relationship between returns and the expected future log-volatilities can be seen to be incorrect.

This issue is addressed in the literature subordinated to the name News Impact Curve (NIC) or Function (NIF), and in the multivariate case, News Impact Surface (NIS). Except in Takahashi et al. (2013), where is recognised the U-shaped configuration for the function that relates returns and future volatilities, there are still errors in the literature. Takahashi et al. (2013) refer to the results in Yu (2005) and Asai and McAleer (2009a) but do not question their validity. On the other hand, the authors use a more than needed sophisticated way to approximate the NIC for the SV and ASV models. From the model, we simulated returns and states, obtained in a straightforward manner, and by approximating the joint distribution of α_{t+1} and y_t , the conditional distribution and respective moments are readily obtained.

This article goes beyond the results of Takahashi et al. (2013) and explains why in the SV model the news impact measured by the expectation $E(\alpha_{t+1}|y_t)$ has to have the U-shaped form. In Asai et al. (2006) the formula to $E(\alpha_{t+1}|y_t)$ given in the page 169 of Yu (2005) is assumed to be correct. On the other hand, in Asai and McAleer (2009b), it is recognised as an incorrect formula, and an alternative formula is presented, which is also incorrect in this context.

3 Stochastic volatility models estimation

The article of Jacquier et al. (1994) began a series of important research on how to estimate the SV model, and the most cited are Shephard and Pitt

(1997), Kim et al. (1998), Chib et al. (2002), Chib et al. (2006) and Omori et al. (2007). Different approaches have been proposed, which are associated with different algorithms to obtain samples for the vector of states. Here, we develop a gaussian approximation, which gives high acceptance rates within MCMC simulations.

To estimate the parameters of the model the marginal posterior distribution for the parameters is approximated, which is needed for a Bayesian estimation procedure. This can be done by simulating from the distribution of $\alpha_{1:n}, \theta | y_{1:n}$, where $\alpha_{1:n} = (\alpha_1, \dots, \alpha_n)$ and $y_{1:n} = (y_1, \dots, y_n)$. Using Gibbs sampling, the parameter vector is sampled conditional on the states, $\theta | \alpha_{1:n}, y_{1:n}$, and the states conditional on the parameter vector, $\alpha_{1:n} | \theta, y_{1:n}$. We develop a single-move sampler to simulate from $\alpha_t | \alpha_{\setminus t}, \theta, y_{1:n}$, where $\alpha_{\setminus t} = (\alpha_1, \dots, \alpha_{t-1}, \alpha_{t+1}, \dots, \alpha_n)$. Through a second order Taylor approximation to the target density, a gaussian density is constructed. Assuming that at iteration k the sampled elements are given in $\theta^{(k)} = (\mu^{(k)}, \phi^{(k)}, \sigma_\eta^{(k)})$ and $\alpha^{(k)} = (\alpha_1^{(k)}, \dots, \alpha_n^{(k)})$, at iteration $k + 1$ the algorithm proceeds as

1. Sample from $\alpha_t | \alpha_{t-1}^{(k+1)}, \alpha_{t+1}^{(k)}, y_t, \theta^{(k)}$; $t = 1, \dots, n$.
2. Sample from $\mu, \phi | \sigma_\eta^{(k)}, \alpha^{(k+1)}, y_{1:n}$.
3. Sample from $\sigma_\eta | \mu^{(k+1)}, \phi^{(k+1)}, \alpha^{(k+1)}$.

Here the single-move sampler contrasts with the algorithms based on multi-move samplers, from which one of the more robust and cited is the one used in Omori et al. (2007). The algorithm has been used by other authors with different adaptations and also serving as a benchmark to the results for other algorithms. Recent examples can be found in Kastner and Frühwirth-Schnatter (2014) and Djegnéné and McCausland (2015). We found that these algorithms are more error prone and in some circumstances do not represent an additional improvement in terms of estimation efficiency.

The single-move sampler has been used by Jacquier et al. (1994, 2004) and the main difference here is the approximating density used to sample the states. Due to the characteristics of the model, the chains obtained to approximate the marginal posterior distribution of the parameters can present high levels of autocorrelation. However, due to the high levels of acceptance associated with the draws in the vector of states, the convergence of the chains does not usually represent a major concern.

Finally, we have to distinguish between sampling the states and sampling the parameters. The first task is harder to implement, essentially, due to the dimension issue. It is in this part that the results in this article can be differentiated from the others. Given the vector of states, it is straightforward to simulate from the marginal posterior distribution of the parameters, and we are going to consider very similar procedures from the ones already presented in the literature when sampling the parameter vector.

3.1 Sampling the states

Most of the research associated with the estimation of the parameters on SV models has been related to the development of algorithms to simulate the vector of states within a MCMC setting. The reason is that this constitutes the main challenge. Jacquier et al. (1994) have used approximations based on the gamma distribution. Major modifications were proposed by Shephard and Pitt (1997), Kim et al. (1998), Chib et al. (2002) and Omori et al. (2007) using gaussian state-space approximation with multi-move samplers. The most recent approaches to this problem can be found in Kastner and Frühwirth-Schnatter (2014) and Djegnéné and McCausland (2015). The complexity of the algorithms has led to some incorrect implementations and misleading values for the estimates of the parameters were reported in the literature. It is argued that with a simpler algorithm to sample the vector of states, based on an independent and single-move sampler, defined through

gaussian approximations, gives reliable and robust estimates for the parameters, and also, presenting the non-negligible characteristics of being very easy to translate as software code.

The aim is to simulate from the vector of states given the observations and the parameter vector, $\alpha_{1:n}|y_{1:n}, \theta$. Instead of considering a gaussian state-space approximation and a multi-move sampler, we consider only a gaussian approximation to the distribution of $\alpha_t|\alpha_{\setminus t}, y_{1:n}, \theta$, for $t = 1, \dots, n$, with $\alpha_{\setminus t} = (\alpha_1, \dots, \alpha_{t-1}, \alpha_{t+1}, \dots, \alpha_n)$. Due to the assumed structure of the states evolution, the gaussian approximation constitutes an extremely good approximation. The main point is to find the mean and variance of such approximation. For the basic SV model an analytical formula can be obtained to calculate such quantities. For most generalisations, assuming the log-concavity of the target density, readily-to-use numerical estimates for the mean and variance can be obtained through numerical procedures.

By using a single-move sampler we are breaking the n -dimensional problem of sampling from the distribution of $\alpha_{1:n}|y_{1:n}, \theta$, into n problems of dimension one, by simulating from $\alpha_t|\alpha_{\setminus t}, y_{1:n}, \theta$, for $t = 1, \dots, n$, which can be simplified to sampling from $\alpha_t|\alpha_{t-1}, \alpha_{t+1}, y_t, \theta$. The logarithm of the target density function assumes the form

$$\ell(\alpha_t) \propto -\frac{\alpha_t}{2} - \frac{y_t^2}{2e^{\alpha_t}} - \frac{(\alpha_{t+1} - \mu - \phi \alpha_t)^2}{2\sigma_\eta^2} - \frac{(\alpha_t - \mu - \phi \alpha_{t-1})^2}{2\sigma_\eta^2}. \quad (15)$$

The independent sampler is defined through a gaussian approximation to such density, which is given by $g(\cdot)$. The next step is standard, and a candidate to α_t is obtained, denoted by α_t^+ , sampled from $g(\alpha_t^+)$, a gaussian density with mean m_t and variance s_t^2 . The candidate is accepted with probability

$$p(\alpha_t^c, \alpha_t^+) = \min \left\{ \frac{f(\alpha_t^+) g(\alpha_t^c)}{f(\alpha_t^c) g(\alpha_t^+)}, 1 \right\}, \quad (16)$$

where α_t^c represents the current value of the state at t . As the approximation is very good the acceptance rates are very high.

The added value of this algorithm is related to its simplicity. We have only to determine at each sub-iteration the values for m_t and s_t^2 and simulate from the respective gaussian distribution. For the basic SV model, such quantities can be obtained through an analytical formula, and the maximiser of (15), which defines the mean of the gaussian approximation is given by

$$m_t = W \left(\frac{y_t^2 \sigma_\eta^2 e^{-\varphi_t}}{2(1 + \phi^2)} \right) + \varphi_t \quad (17)$$

where

$$\varphi_t = \frac{\phi (\alpha_{t+1} + \alpha_{t-1}) + \mu (1 - \phi)}{1 + \phi^2} - \frac{\sigma_\eta^2}{2(1 + \phi^2)}, \quad (18)$$

and W is the Lambert function. For the variance, using the the second order Taylor approximation to $\ell(\alpha_t)$ around m_t , the log-kernel of a gaussian density with variance

$$s_t^2 = -\frac{1}{\ell''(m_t)} = \frac{2 \sigma_\eta^2 e^{m_t}}{y_t^2 \sigma_\eta^2 + 2 (1 + \phi^2) e^{m_t}} \quad (19)$$

is obtained.

The approximation is very good and the acceptance rates are very high. However, even with chains that always move, sometimes they move slowly, and high levels of autocorrelation are obtained. Due to the simplicity of the sampler, several strategies may be considered to reduce the levels of autocorrelation, and to define more efficient estimation procedures. The main point to highlight is that with SV models, gaussian approximations are straightforward to implement.

The basic ideas for the standard SV model can now be developed for the ASV model, and with the formulation in (3)-(4), an equivalent expression to (15) is developed

$$\begin{aligned} l(\alpha_t) &\propto -\frac{\alpha_t}{2} - \frac{y_t^2}{2 \exp(\alpha_t)} \\ &\quad - \frac{(\alpha_t - \mu - \phi \alpha_{t-1} - \rho \sigma_\eta \exp(-\frac{\alpha_{t-1}}{2}) y_{t-1})^2}{2 (1 - \rho^2) \sigma_\eta^2} \\ &\quad - \frac{(\alpha_{t+1} - \mu - \phi \alpha_t - \rho \sigma_\eta \exp(-\frac{\alpha_t}{2}) y_t)^2}{2 (1 - \rho^2) \sigma_\eta^2}. \end{aligned} \quad (20)$$

Unfortunately, no analytical expression exists to define the value that maximizes the function as in the case of the standard SV model, but numerical procedures can be used to approximate it. The main insight is that locally the gaussian approximation is very good. The mode can be obtained easily using the Newton's method after evaluating the first and second derivatives of $\ell(\alpha_t)$. The mode will constitute the mean of the gaussian approximation, and after evaluating the second derivative at the mode the variance is also defined.

3.2 Sampling the parameters

It is the marginal posterior distribution of the parameters that will serve in a Bayesian analysis to estimate the parameters of the models. As the distribution does not present an analytical tractable form, it is approximated by simulating from $\theta|\alpha_{1:n}, y_{1:n}$, where $\theta = (\mu, \phi, \sigma_\eta^2)$ for the SV model and $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$ for the ASV model, conditional on the vector of observations and vector of states. The critical element to simulate in these models is the vector of states, mainly due to its dimension. Given the vector of states, it is relatively straightforward to simulate from the marginal posterior distribution of the parameters.

To define the posterior, a prior distribution for the vector of parameters needs to be specified. It is common to assume the independence of the parameters within the prior and for the ASV model $p(\theta) = p(\mu)p(\phi)p(\sigma_\eta^2)p(\rho)$ Kastner and Frühwirth-Schnatter (2014), Omori et al. (2007), Chib et al. (2002), Kim et al. (1998). For μ , ϕ and σ_η^2 taken individually, conditional on the observations, states, and remaining parameters, conjugate prior distributions can be defined, gaussian for μ and ϕ , and inverse-gamma for σ_η^2 .

There is a well established set of results on the sensitivity of the marginal posterior distribution to the different forms that the prior distributions may assume. Naturally the influence of the priors depend on the information con-

tained in the likelihood. With the amount of information usually available, the sensitivity of the marginal posterior distribution to different forms for the priors can be small.

For the ASV the same priors are used, and to distinguish σ_η and ρ , the results presented in Jacquier et al. (2004) are considered. With $\psi = \rho \sigma_\eta$ and $\Omega = \sigma_\eta^2 (1 - \rho^2)$, estimated through the linear model, the original parameters are obtained as $\sigma_\eta^2 = \Omega + \psi^2$ and $\rho = \psi / \sigma_\eta$.

4 Volatility forecasting

The SV model is a state-space model where the evolution of the states defines the evolution of the volatility. Forecasts for the evolution of the states in this setting require the development of Sequential Monte Carlo (SMC) simulation techniques, also referred as particle filter methods Gordon et al. (1993), Carpenter et al. (1999), Pitt and Shephard (1999), Doucet et al. (2000), Godsill and Clapp (2001), Del Moral et al. (2006), Andrieu et al. (2010), Fearnhead et al. (2010). The aim is to update the filter distribution for the states when new information arrive.

The quantities of interest are the values of the states governing the evolution of the volatility, which are propagated to define the predictive density of the returns, defined here as $f(y_{t+1}|y_{1:t})$. However, essential to the definition of this distribution is the filter density associated with the states, $f(\alpha_t|y_{1:t})$. Bayes's rule allows us to assert that the posterior density $f(\alpha_t|y_{1:t})$ of states is related to the density $f(\alpha_t|y_{1:t-1})$ prior to y_t , and the density $f(y_t|\alpha_t)$ of y_t given α_t by $f(\alpha_t|y_{1:t}) \propto f(y_t|\alpha_t)f(\alpha_t|y_{1:t-1})$. The predictive density of y_{t+1} given $y_{1:t}$ is defined by $f(y_{t+1}|y_{1:t}) = \int f(y_{t+1}|\alpha_{t+1})f(\alpha_{t+1}|y_{1:t}) d\alpha_{t+1}$.

Particle filters approximate the posterior density of interest, $f(\alpha_t|y_{1:t})$, through a set of k "particles" $\{\alpha_{t,1}, \dots, \alpha_{t,k}\}$ and their respective weights $\{\pi_{t,1}, \dots, \pi_{t,k}\}$, where $\pi_{t,j} \geq 0$ and $\sum_{j=1}^k \pi_{t,j} = 1$. This procedure must be

implemented sequentially with the states evolving over time to accommodate new arriving information. It is difficult to obtain samples from the target density, and an approximating density is used instead, afterwards the particles are resampled to better approximate the target density. This is known as the sample importance resampling (SIR) algorithm. A possible approximating density is given by $f(\alpha_t|\alpha_{t-1})$, however, Pitt and Shephard (1999, 2001) pointed out that as a density to approximate $f(\alpha_t|y_{1:t})$ is not generally efficient, because it constitutes a *blind* proposal that does not take into account the information contained in y_t .

4.1 Particle Filter for the SV Model

Through SMC with SIR the aim is to update sequentially the filter density for the states. The optimal importance density is given by $f(y_t|\alpha_t)f(\alpha_t|\alpha_{t-1}, y_t)$, which induces importance weights with zero variance. Usually it is not possible to obtain samples from this density, and an importance density $g(\alpha_t)$, different from the optimal one, is used to approximate the target density.

To approximate the filter densities associated with the SV model, Pitt and Shephard (1999) considered the same kind of approximations used to obtain the samples from the vector of the states in a static MCMC setting. However, the approximations were based on a first order Taylor approximation, and it was demonstrated by Smith and Santos (2006) that they are not robust when information contained in more extreme observations need to be updated (also called very informative observations). In Smith and Santos (2006) a second order Taylor approximation for the likelihood combined with the predictive density for the states leads to improvements in the particle filter algorithm. As the auxiliary particle filter in Pitt and Shephard (1999), it avoids *blind* proposals like the ones in Gordon et al. (1993), takes into account the information in y_t , and defines a robust approximation for the target density without the problem represented by the sample impoverishment.

Here we develop the aforementioned results using a robusiter approximation for the importance density. The logarithm of the density $f(y_t|\alpha_t)f(\alpha_t|\alpha_{t-1})$, $\ell(\alpha_t)$, is concave on α_t , and to maximize the function in order to α_t , let us consider the first derivative equal to zero, $\ell'(\alpha_t) = 0$. Solving in order to α_t the solution is

$$\alpha_t^* = W \left(\frac{y_t^2 \sigma_\eta^2 e^{-\gamma_t}}{2} \right) + \gamma_t, \text{ with } \gamma_t = \mu + \phi \alpha_{t-1} - \frac{\sigma_\eta^2}{2}. \quad (21)$$

The second derivative is given by $\ell''(\alpha_t) = -(2e^{\alpha_t} + \sigma_\eta^2 y_t^2)/(2\sigma_\eta^2 e^{\alpha_t})$, which is strictly negative for all α_t , so α_t^* maximizes the function $\ell(\alpha_t)$ defining a global maximum. The second order Taylor expansion of $\ell(\alpha_t)$ around α_t^* defines the log-kernel of a gaussian density with mean $m_t = \alpha_t^*$ and variance

$$s_t^2 = \frac{2\sigma_\eta^2 e^{m_t}}{2e^{m_t} + \sigma_\eta^2 y_t^2}. \quad (22)$$

This gaussian density will be used as the importance density in the SIR algorithm.

In the procedures implemented, the densities are approximated using particles with equal weights, which means that a resampling step is performed. Assuming at $t - 1$ a set of k particles $\alpha_{t-1}^k = \{\alpha_{t-1,1}, \dots, \alpha_{t-1,k}\}$ with associated weights $1/k$, which approximate the density $f(\alpha_{t-1}|y_{1:t-1})$, the algorithm proceeds as follows

1. For each element of the set α_{t-1}^k , sample a value from a gaussian distribution with mean and variance defined by (21) and (22), respectively, obtaining the set $\{\alpha_{t,1}^*, \dots, \alpha_{t,k}^*\}$.
2. Calculate the weights,

$$w_i = \frac{f(y_t|\alpha_{t,i}^*)f(\alpha_{t,i}^*|\alpha_{t-1,i})}{g(\alpha_{t,i}^*|m_t, s_t^2)}, \quad \pi_i = \frac{w_i}{\sum_{i=1}^k w_i}, \quad i = 1, \dots, k. \quad (23)$$

3. Resample from the set $\{\alpha_{t,1}^*, \dots, \alpha_{t,k}^*\}$ using the set of weights $\{\pi_1, \dots, \pi_k\}$ obtaining a sample $\{\alpha_{t|1:t,1}, \dots, \alpha_{t|1:t,k}\}$, where to each particle a weight of $1/k$ is associated.

For the one step-ahead volatility forecast, having the approximation to the density $f(\alpha_t|y_{1:t})$, and due to the structure of the system equation in the SV model, an AR(1), is easy to sample from $f(\alpha_{t+1}|y_{1:t})$, the predictive density for the states.

5 An empirical demonstration

In this section we use two time series that are frequently considered when the aim is to apply estimation and forecasting procedures associated with the SV model. The first series is the daily Pound/Dollar exchange rates from 01/10/81 to 28/6/85 considered in Shephard and Pitt (1997), Kim et al. (1998) and Meyer and Yu (2000). A second series is associated with the returns of the S&P500 commonly used in the literature Chib et al. (2002), Jacquier et al. (2004), Yu (2005), McCausland (2012), Djegn  n   and McCausland (2015). Here is used the series from January 1980 to December 1987, also considered by Yu (2005) and Djegn  n   and McCausland (2015). The utilisation of these series serves mainly to compare with the results presented in the literature.

The SV model is a nonlinear state-space model, and in contrast with models used for the same end like the GARCH, is not compatible with the estimation of the parameters through standard maximum likelihood estimation methods, and forecasting through a simple recursive formula. Estimation and forecasting procedures need numerical methods. However, the model incorporates a simple structure given by the Markovian nature for the evolution of the states. For this reason gaussian proposals to obtain samples from the vector of states give intuitive and reliable results.

The development of too complex algorithms to estimate the parameters and to forecast through the model, in some circumstances may be reliable, but surely can be cumbersome to implement and error prone, and in some

Table 1: Estimation of the parameters in the ASV model. The used series is the S&P500 returns (January 1980 to December 1987). Left-side: results obtained using the algorithm proposed to sample the states in this article. Right-side: Estimates obtained using the STAN package.

	mean	sd	inef	mean	sd	inef
μ	-0.008	0.006	62.19	-0.007	0.005	45.44
ϕ	0.956	0.013	177.28	0.957	0.013	132.47
σ_η	0.195	0.027	336.67	0.193	0.029	209.93
ρ	-0.322	0.082	77.68	-0.318	0.080	3.24

cases can lead to results that differ from the obtained through simpler algorithms. Examples can be found in Djegn   and McCausland (2015) to estimate the parameters, and Li et al. (2016) to forecast the evolution of the states.

For the estimates of the parameters, we present only the results associated with the ASV model. Compared with the basic model, it adds a parameter, and our aim is to compare with the results previously found in the literature. To compare the forecasting procedures, the basic SV model is used. It is a model commonly used when the aim is to test particle filter algorithms associated with state-space models. A recent example is found in Li et al. (2016), where we think that the extra layers of complexity in the design of the algorithm are not strictly necessary, especially, if the examples considered are the AR(1) with gaussian noise, and the SV model.

5.1 Parameter's estimation

In Table 1, we present the results associated with the estimation of the ASV model for the series S&P500 from January 1980 to December 1987. The results are obtained through Bayesian estimation using MCMC with chains

of 10 000 observations for a burn-in period of 2 000. The results are obtained using the algorithm presented in this paper. To demonstrate the reliability of the algorithms proposed, the results are compared with the ones obtained through the package STAN, and the results obtained using both procedures are very similar. The main difference is related to the inefficiency factors that are lower in the STAN version, especially for the parameter ρ . Here the inefficiency (inef) factors are calculated as the ratio between the number of observations used in the chain and the effective sample size (ESS). These statistics are standard and the library CODA through the R software package was used to perform such computations. Comparing the results with the ones given in Yu (2005), the inefficiency measures are similar, and we find less persistence and more variance associated with the innovations of the system equation.

Our results differ substantially from the ones in Djegn   and McCausland (2015), mainly in the estimated value for the parameter ρ . The results are not shown here, but we estimated the parameters for the other two series in Djegn   and McCausland (2015), and our results are similar with the ones presented by the authors. A possible explanation is that the algorithm presented by Djegn   and McCausland (2015) is not sufficiently robust, and is not able to cope with the abnormal returns associated with the crash of October 87 and observations that have followed.

In Table 2, we present the results associated with the estimation of the ASV model for the series of the Pound/Dollar exchange rates from 01/10/81 to 28/6/85, and the results are in full contrast with the results in Meyer and Yu (2000). A significative leverage effect was not found for the exchange rates, which seems more intuitive than the other way around. Due to the structure of the exchange rates market, how the negotiation is made, the players involved and the fundamentals driving the exchange rates, it seems more plausible that the leverage effect can be small or even inexistent. However,

Table 2: Estimation of the parameters in the ASV model. The used series is the £/\$ exchange rates (01/10/81 to 28/6/85). Left-side: results obtained using the algorithm proposed to sample the states in this article. Right-side: Estimates obtained using the STAN package.

	mean	sd	inef	mean	sd	inef
μ	-0.035	0.021	236.62	-0.036	0.018	98.77
ϕ	0.965	0.016	233.14	0.963	0.016	104.54
σ_η	0.192	0.049	504.87	0.201	0.039	173.07
ρ	-0.014	0.155	123.43	0.005	0.135	6.69

in Meyer and Yu (2000) a significative leverage effect is reported measured by the estimate for the parameter ρ , assuming a value around -0.2 . This is another example of the conflicting results still present in the literature that may justify the need for more research associated with these models.

5.2 Forecasting with the SV

We combine the insights of the developed algorithm to obtain samples from the states in a MCMC setting to estimate the parameters, and we develop also an algorithm within a particle filter approach, used to update the distributions of the states when new arriving information happens sequentially. Assuming that the parameters are fixed or known, it might be of interest updating the distributions for the states. It is a procedure considered in different fields like signal processing, object tracking, etc, but in this case the state represents the volatility. As the state is not directly observable, the information obtained through the returns is used to update the filter distributions. Assuming that at $t - 1$ the filter distribution is characterised by the density $f(\alpha_{t-1}|y_{1:t-1})$, the aim is to use the information in y_t to update the filter density $f(\alpha_t|y_{1:t})$.

Here the basic SV model is considered with the series of the S&P500 returns. The aim is to approximate the mean of the filter distribution, which is done after the estimation for the values of the parameters. Using one third of the sample, the filter distributions are updated for each observation, mimicking the arrival of new information. Each time the filter distribution is updated the respective mean is calculated. The mean is obtained in two ways. First, applying a MCMC static procedure to obtain samples from the vector of states each time a new observation is considered. Chains with a very large number of observations allow to obtain the *true* value for the means of the filter distributions. Second, the sequential particle filter approach is used to approximate the means of the filter distributions. We show that the particle filter approach gives very precise results. We have to note that the particle filter uses only partial information, the previous approximated filter distribution and the return observed at t . The first is used as the prior distribution, and the second is combined through the likelihood function.

The results show that adding extra layers of complexity in the algorithms, when in fact due to structure of the model those extra layers are not needed can lead to results more difficult to interpret. One example is related to the results presented by Li et al. (2016), where two models are used to test the effectiveness of the algorithms proposed. The first is the linear gaussian state-space, described as the AR(1) with gaussian noise. The second is the basic SV model. In the first case, the true value for the means associated with the filter and smooth distributions can be obtained analytically, and due to the structure of the model, any deviation from the true value that do not correspond to errors comparable with rounding errors, makes the particle filter algorithm essentially useless.

To show the reliability of the particle filter algorithm proposed in this paper, an experiment was considered to calculate the evolution for the means of the filter distributions and compare such evolution with the *true* evolu-

tion. Using the 2022 observations for the S&P500 series, the estimates for the parameters are defined, and in this case $\mu = -0.007$, $\phi = 0.968$ and $\sigma_\eta = 0.165$. Using two thirds of the series, with fixed parameters, the filter distribution for the states is approximated (observation 1348), which will constitute the numerical approximation for the prior distribution in the first iteration of the particle filter. The particle filter is run from the observation 1349 to the observation 2022, simulating the arrival of new information and the updating of the filter distribution. The means and variances associated with the filter distributions are compared with the ones obtained by running the MCMC where sequentially a new observation is added in each iteration. These are considered here the *true* values of the means and variances. As it is depicted in Figure 2, the evolution for the means of the filter distributions is well approximated using the particle filter approach. The particle filter uses only partial information, but due to the robustness of the algorithm tailored to the model, the approximation is good and remains good through the successive application of the particle filter procedure.

Li et al. (2016) try to establish the comparison for the adequacy of different algorithms in the implementation of different approaches in the particle filter methods. The linear state-space model does not need numerical algorithms to approximate the distributions, filter and smoothing. The nonlinear state-space models are models that justify the use of such numerical techniques, which can be used to approximate the quantities of interest. But as we show in this paper, a well adapted algorithm for the SV model gives very reliable results, and puts into perspective the need of more complex algorithms, which are also more error prone, as in Li et al. (2016).

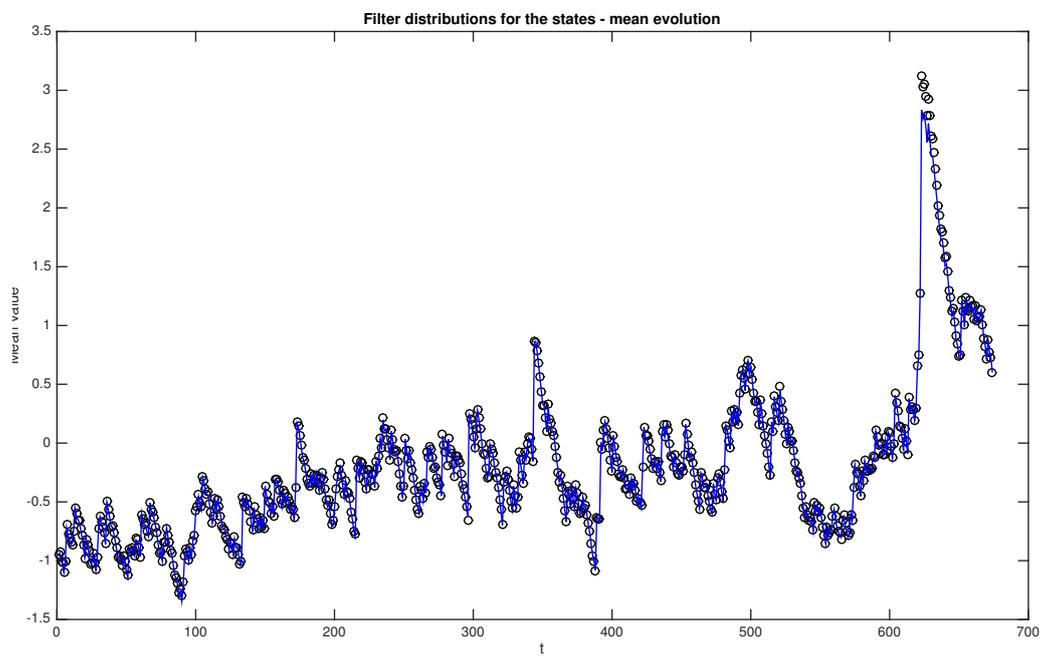


Figure 2: Evolution of the means of the filter distributions for the states with the SV model. The *true* mean (o) is compared with the approximated mean (—) obtained through particle filter methods.

6 Conclusions

In this paper we show how to estimate and forecast through SV models using a robust algorithm tailored to the model. A comparison is made with other methods presented in the literature and an explanation for the conflicting results is sought. Some algorithms proposed in the literature are too complicated. They are more error prone than simpler algorithms, which give more reliable results. These are not the most efficient in terms of the estimators produced, but they can avoid grossly wrong estimates for the parameters, and consequently to the forecasts.

References

- Andrieu, C., Doucet, A. and Holenstein, R.: 2010, Particle markov chain monte carlo methods, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **72**(3), 269–342.
- Asai, M. and McAleer, M.: 2009a, Multivariate stochastic volatility, leverage and news impact surfaces, *The Econometrics Journal* **12**(2), 292–309.
- Asai, M. and McAleer, M.: 2009b, The structure of dynamic correlations in multivariate stochastic volatility models, *Journal of Econometrics* **150**(2), 182–192.
- Asai, M., McAleer, M. and Yu, J.: 2006, Multivariate stochastic volatility: a review, *Econometric Reviews* **25**(2-3), 145–175.
- Bollerslev, T.: 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of econometrics* **31**(3), 307–327.
- Carpenter, J., Clifford, P. and Fearnhead, P.: 1999, Improved particle filter for nonlinear problems, *IEE Proceedings-Radar, Sonar and Navigation* **146**(1), 2–7.

- Chib, S., Nardari, F. and Shephard, N.: 2002, Markov chain monte carlo methods for stochastic volatility models, *Journal of Econometrics* **108**(2), 281–316.
- Chib, S., Nardari, F. and Shephard, N.: 2006, Analysis of high dimensional multivariate stochastic volatility models, *Journal of Econometrics* **134**(2), 341–371.
- Del Moral, P., Doucet, A. and Jasra, A.: 2006, Sequential monte carlo samplers, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68**(3), 411–436.
- Djegnéné, B. and McCausland, W. J.: 2015, The hessian method for models with leverage-like effects, *Journal of Financial Econometrics* **13**(3), 722–755.
- Doucet, A., Godsill, S. and Andrieu, C.: 2000, On sequential monte carlo sampling methods for bayesian filtering, *Statistics and computing* **10**(3), 197–208.
- Engle, R. F.: 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, *Econometrica: Journal of the Econometric Society* pp. 987–1007.
- Fearnhead, P., Wyncoll, D. and Tawn, J.: 2010, A sequential smoothing algorithm with linear computational cost, *Biometrika* **97**(2), 447–464.
- Godsill, S. and Clapp, T.: 2001, Improvement strategies for monte carlo particle filters, *Sequential Monte Carlo Methods in Practice*, Springer, pp. 139–158.
- Gordon, N. J., Salmond, D. J. and Smith, A. F.: 1993, Novel approach to nonlinear/non-gaussian bayesian state estimation, *IEE Proceedings F (Radar and Signal Processing)*, Vol. 140, IET, pp. 107–113.

- Harvey, A. C. and Shephard, N.: 1996, Estimation of an asymmetric stochastic volatility model for asset returns, *Journal of Business & Economic Statistics* **14**(4), 429–434.
- Jacquier, E., Polson, N. G. and Rossi, P. E.: 1994, Bayesian analysis of stochastic volatility models, *Journal of Business & Economic Statistics* **12**(4), 371–89.
- Jacquier, E., Polson, N. G. and Rossi, P. E.: 2004, Bayesian analysis of stochastic volatility models with fat-tails and correlated errors, *Journal of Econometrics* **122**(1), 185–212.
- Kastner, G. and Frühwirth-Schnatter, S.: 2014, Ancillarity-sufficiency interweaving strategy (asis) for boosting mcmc estimation of stochastic volatility models, *Computational Statistics & Data Analysis* **76**, 408–423.
- 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.
- Li, W., Chen, R. and Tan, Z.: 2016, Efficient sequential monte carlo with multiple proposals and control variates, *Journal of the American Statistical Association* **111**(513), 298–313.
- McCausland, W. J.: 2012, The hessian method: Highly efficient simulation smoothing, in a nutshell, *Journal of Econometrics* **168**(2), 189–206.
- Meyer, R. and Yu, J.: 2000, Bugs for a bayesian analysis of stochastic volatility models, *The Econometrics Journal* **3**(2), 198–215.
- Nelson, D. B.: 1991, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica: Journal of the Econometric Society* pp. 347–370.

- Omori, Y., Chib, S., Shephard, N. and Nakajima, J.: 2007, Stochastic volatility with leverage: Fast and efficient likelihood inference, *Journal of Econometrics* **140**(2), 425–449.
- Pitt, M. K. and Shephard, N.: 1999, Filtering via simulation: Auxiliary particle filters, *Journal of the American statistical association* **94**(446), 590–599.
- Pitt, M. K. and Shephard, N.: 2001, Auxiliary variable based particle filters, *Sequential Monte Carlo methods in practice*, Springer, pp. 273–293.
- Shephard, N. and Pitt, M. K.: 1997, Likelihood analysis of non-gaussian measurement time series, *Biometrika* **84**(3), 653–667.
- Smith, J. and Santos, A. A. F.: 2006, Second-order filter distribution approximations for financial time series with extreme outliers, *Journal of Business & Economic Statistics* **24**(3), 329–337.
- Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D.: 2003, Winbugs user manual.
- Stan Development Team: 2014, *Stan Modeling Language Users Guide and Reference Manual, Version 2.5.0*.
URL: <http://mc-stan.org/>
- Takahashi, M., Omori, Y. and Watanabe, T.: 2013, News impact curve for stochastic volatility models, *Economics Letters* **120**(1), 130–134.
- Taylor, S. J.: 1986, *Modelling Financial Time Series*, Wiley: Chichester.
- Yu, J.: 2005, On leverage in a stochastic volatility model, *Journal of Econometrics* **127**(2), 165–178.