Journal of Econometrics and Statistics
Vol. 5, Issue 1, 2025, pp. 1-13

Stochastic Volatility Model via Gaussian Process

Lucas Marques Oliveira[a], Fabio A. Tavares Ramos[b] and Ralph dos Santos Silva[b,c]

[a] *Semantix AI, Brazil*
[b] *Instituto de Matemática, Universidade Federal de Rio de Janeiro, Brazil*

ABSTRACT
This paper aims to model the returns on a financial asset with volatilities expressed through a Gaussian process. This approach avoids a rigid structure for the functional form (over time) of volatility. Inference about the unknown quantities of the models is made under a Bayesian approach with the application of numerical methods such as Gibbs sampling and particle filter. Although the model is defined with a Markovian structure, with the introduction of the Gaussian process, there is a dependence between the states in the unconditional distributions. The state prior is complicated, so to approximate it and speed up the estimation process, a data window is proposed. In the applications, a small Monte Carlo study is presented to evaluate the estimation process and compare performance with other models, as well as with real data.

KEYWORDS
Gibbs sampling, Sequential Monte Carlo methods, State space model.

1. Introduction

In the finance literature, the importance of quantifying volatility has progressively increased since Engle [7]. The term volatility has been used in different contexts such as current, historical, realized, implied, and future volatilities. The estimation of historical volatility is the focus of this paper.

The difference between volatility and other variables, such as the price of an asset or a country's inflation, is that this variable is generally not observed. Therefore, in this paper, its estimation is done through statistical models, exploring its relationship with other variables, specifically the asset return.

Engle [7] explores the relationship between volatility and return through an autoregressive form, where the return has its variance (normal) distribution varying over time, and the variance depends on lagged terms of the return itself. Since then, other models have been developed, such as the classic GARCH described in Bollerslev [3], where volatility depends on lagged values of the return and its own lagged volatility values.

Subsequently, the use of probability distributions to model volatility was researched, resulting in stochastic volatility models. In the class of state-space models, these models involve a latent stochastic variable (unobserved) that changes over time and influ-

CONTACT Author[c]. Email: ralph@im.ufrj.br

Article History
Received : 14 October 2024; Revised : 13 November 2024; Accepted : 08 December 2024; Published : 08 January 2025

To cite this paper

Lucas Marques Oliveira, Fabio A. Tavares Ramos & Ralph dos Santos Silva (2025). Stochastic Volatility Model via Gaussian Process. *Journal of Econometrics and Statistics.* 5(1), 1-13.

ences the observed variable. A common approach in the use of stochastic volatility is when the log of variance follows an autoregressive model [13].

The existence of a temporal relationship between volatility values is largely due to the empirically observed phenomenon known as volatility clustering. This term was first discussed by Mandelbrot [15]. The use of dependence on current volatility values with past volatility and past squared log returns is largely due to this phenomenon.

Another widely discussed phenomenon in quantitative finance literature is the so-called leverage effect, first discussed by Black [2], as the negative correlation between past return values and future volatility values. Thus, during a crisis, it is unclear in the subsequent moment what the fair price of the asset should be.

An established fact in stochastic volatility models is that logarithmic returns generally exhibit heavy-tailed distributions. Nevertheless, many models use the normal distribution as part of their specification. However, this does not diminish the estimation process because this normal distribution is typically combined with other structures that account for the heavy-tail effect in log returns [5].

Wu et al. [23] propose a volatility model through a Gaussian process [19] in which the normal covariance matrix is set through a covariance function, also known as a kernel. This approach has usually the advantage of being a smooth function from the past to the new volatility values. This paper builds on this idea and introduces a novel approach using a distribution in the function space for volatility. This approach offers great flexibility in estimating this variable.

Moreover, we employ a particle Markov chain Monte Carlo (PMCMC) algorithm [1] called particle Gibbs with ancestor sampling (PGAS), proposed by Lindsten et al. [14]. These algorithms handle the posterior distribution of static parameters through standard MCMC methods and manage the posterior of latent variables using sequential Monte Carlo sampling, also known as particle filtering. Particle filtering involves simulating a certain number N of samples for each time t of the state variable. As in Pitt et al. [17], notice that increasing the number of particles enhances the acceptance rate of PMCMC sampling, albeit at a higher computational cost.

The remainder of this paper is structured as follows: Section 2 briefly discusses a few volatility models and introduces our proposed approach to modelling volatility. The development of inference and algorithm to deal with the stochastic volatility model via Gaussian process is given in Section 3 by applying the Bayesian inference through particle Gibbs with ancestor sampling. Section 4 brings a small Monte Carlo study and real data applications. Section 5 concludes.

2. Modelling volatility

Most financial studies concentrate on analysing return series rather than using asset price series. The preference for return series stems from two factors: returns provide information relevant to investors' interests and exhibit more statistically intriguing properties compared to price series. Therefore, if P_t represents the price of an asset at time t, the log-return at time t is defined as: $y_t = \log(P_t) - \log(P_{t-1})$.

2.1. *GARCH model*

A generalized autoregressive conditionally heteroscedastic (GARCH) model [3] is used - for comparisons - to model the variance of a time series using values of the past squared means of the observations and past variances. The GARCH model is given

by

$$y_t \sim \mathcal{N}(0, \sigma_t^2) \quad \text{and} \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

where $\mathcal{N}$ stands for normal distribution, $\alpha_0 > 0$, $\alpha_i \geqslant 0$, $\beta_j \geqslant 0$, $\forall i, j$, respectively. Other restrictions may be applied in the classical maximum likelihood estimation process used in this paper. See Francq and Zakoian [8] for more details. The model can be generalized with an autoregressive moving average structure on the mean.

2.2. *Stochastic volatility model*

First, a state space model can be described with an observation distribution given by $p(y_t|x_t;\theta)$ and a system distribution given by $p(x_t|x_{t-1};\theta)$, both for $t = 1, \ldots, n$, where θ and $p(.)$ denote a parameter vector and general probability (density) functions, respectively. The initial state x_0 is distributed according to $p(x_0|\theta)$. See West and Harrison [22] for more details.

The observation and system equations of the stochastic volatility (SV) model is given by [13]:

$$\begin{aligned} y_t &= e^{x_t/2}\epsilon_t \\ x_t &= \alpha + \phi(x_{t-1} - \alpha) + \omega_t. \end{aligned}$$

We consider that ϵ_t and ω_t are independent, with $\epsilon_t \sim \mathcal{N}(0,1)$ and $\omega_t \sim \mathcal{N}(0, \tau^2)$. The SV model is used for comparisons and as the base for our main model in the following.

To complete our Bayesian model specification, we assume the following prior distribution: $\tau^2 \sim \mathcal{IG}(a_1, b_1)$, $\phi \sim Beta(a_2, b_2)$, $\alpha \sim \mathcal{N}(a_3, b_3^2)$, and $x_0 \sim \mathcal{N}(\alpha, \tau^2/(1+\phi^2))$, where $\mathcal{IG}(a_1, b_1)$ is the inverse-gamma distribution with a_1 and b_1 as shape and scale parameters; and $Beta(a_2, b_2)$ is the beta distribution with mean $a_2/(a_2 + b_2)$.

2.3. *Stochastic volatility model via Gaussian process*

Gaussian process is a generalization of the Gaussian probability distribution. Whilst a probability distribution is associated with a random variable, which can be a scalar or a vector, the Gaussian process is associated with an object that resides in the space of functions. Formally, Gaussian process is a collection of random variables, such that any finite number of these variables has a joint multivariate normal distribution [19].

A Gaussian process is completely determined by its mean and covariance functions. If $f(x)$ follows a Gaussian process ($\mathcal{GP}$), then

$$f(x) \sim \mathcal{GP}(m(x), \kappa(x, x')),$$

where $m(x)$ is the mean function and $\kappa(x, x')$ is the covariance function, with x and x' being input data. There are several covariance functions such as the squared exponential, the γ-exponential, the rational quadratic and the periodic [19].

Our proposed stochastic volatility model via Gaussian process is given by

$$y_t = e^{x_t/2}\epsilon_t, \quad (1)$$
$$x_t = f(x_{t-1}|y_{t-1}) + \omega_t, \quad (2)$$

where ϵ_t and ω_t are again independent, with $\epsilon_t \sim \mathcal{N}(0,1)$ and $\omega_t \sim \mathcal{N}(0,\tau^2)$. The state variable of the model x_t represents the log volatility. A Gaussian process distribution is proposed for the function $f(\cdot|y_{t-1})$ to enhance its modelling flexibility: $f(x|y_{t-1}) \sim GP(m(x_t), \kappa(x_t, x_{t-s}))$. In contrast to Wu et al. [23], we marginalize $f(x|y_{t-1})$ at the posterior distribution to deal better with the inference process through numerical methods.

We consider the mean function as $m(x_t) = ax_{t-1} + by_{t-1}$, where a and b are parameters to be estimated. Note that despite the rigid structure of the mean function, the covariance function has the potential to capture complex patterns according to the observed data. The parameter b captures the leverage effect of the time series.

Moreover, we model the covariance structure through a quadratic kernel as a function of the temporal distance between volatilities. Therefore, let $\kappa(x_t, x_{t-s})$, where $t > s$, be the covariance function evaluated between variables x_t and x_{t-s}. This function is given by

$$\kappa(x_t, x_{t-s}) = h^2 \exp\left(\frac{s^2}{\ell^2}\right), \quad (3)$$

where h^2 and ℓ^2 are parameters that determine the scale of the log volatility and how much one point influences another, respectively.

Hence, the model has a set of (static) parameters $\theta = (a, b, h^2, \ell^2, \tau^2)$. These parameters, together with the log volatility values x_t, $t = 1, 2, \ldots, n$, form the set of variables under which Bayesian inference is developed. Next, we deal with the development of inference and the algorithm dealing with the stochastic volatility model via Gaussian process.

3. Bayesian inference through particle Gibbs with ancestor sampling

In the stochastic volatility model via Gaussian process, directly learning from the posterior of all unknown variables, $(f, x_{1:n}, \theta)$, is a highly complex task since $x_{1:n} = (x_1, x_2, \ldots, x_n)$ are unobserved. Moreover, the posterior distribution has a known but non-standard form. Thus, we resort to numerical algorithms such as particle Gibbs with ancestor sampling [14] to obtain a sample from the posterior distribution of $(f, x_{1:n}, \theta)$ and, consequently, facilitates the inference.

From Equations (1) and (2), the posterior distribution can be represented as

$$p(x_{1:n}, \theta|y_{1:n}) \propto p(y_{1:n}|x_{1:n}, \theta)p(x_{1:n}|\theta)p(\theta),$$

where $p(\theta)$ is the prior distribution (given in the applications) and the likelihood is given by

$$p(y_{1:n}|x_{1:n}, \theta) = \prod_{t=1}^{n} p(y_t|x_t) \quad \text{with} \quad (y_t|x_t) \sim \mathcal{N}(0, \exp(x_t)),$$

and $y_{1:n} = (y_1, y_2, \ldots, y_n)$.

The latent state prior - with θ known and f marginalized - $p(x_{1:n}|\theta)$ is obtained through the product of predictive distributions of a Gaussian process, as discussed in Frigola et al. [9]. Thus, defining $\Sigma_{t,s} = k(x_t, x_{t-s})$ and $m_t = m(x_t)$, the prior $p(x_{1:n}|\theta)$ takes the following form:

$$\begin{aligned} p(x_{1:n}|\theta) &= p(x_1|\theta)\prod_{t=2}^{n} p(x_t|\theta, x_{1:t-1}, y_{1:t-1}) \quad \text{with} \qquad (4)\\ (x_t|\theta, x_{1:t-1}, y_{1:t-1}) &\sim \mathcal{N}\left(x_t|\mu_t(x_{1:t-1}, y_{1:t-1}), \Omega_t(x_{1:t-1})\right)\\ \mu_t(x_{1:t-1}, y_{1:t-1}) &= m_{t-1} + \Sigma_{t-1,1:t-2}\tilde{\Sigma}^{-1}_{1:t-2,1:t-2}(x_{1:t-1} - m_{1:t-2}),\\ \Omega_t(x_{1:t-1}) &= \tilde{\Sigma}_{t-1,t-1} - \Sigma_{t-1,1:t-2}\tilde{\Sigma}^{-1}_{1:t-2,1:t-2}\Sigma^{\top}_{t-1,t-2}, \end{aligned}$$

with $m_{1:t} = (m_1, m_2, \ldots, m_t)$. Note that the prior distribution of x_t depends on all preceding states $x_{1:t-1}$ and preceding observations $y_{1:t-1}$ [23].

The particle Gibbs with ancestor sampling is an algorithm described by Lindsten et al. [14] which is in the class of particle Markov chain Monte Carlo methods [1]. It combines concepts from Markov chain Monte Carlo and particle filtering. Consequently, it enables inference for both the parameters and the states of a state-space model.

3.1. *Standard Particle Filters*

In the state space model described in Section 2.2 - with $p(y_t|x_t;\theta)$, $p(x_t|x_{t-1};\theta)$, and $p(x_0|\theta)$ being the observation, the system and the prior distributions, we need to solve the following integral at time t:

$$p(x_t|y_{1:t-1};\theta) = \int p(x_t|x_{t-1};\theta)p(x_{t-1}|y_{1:t-1};\theta)dx_{t-1}, \qquad (5)$$

and use the result to update the posterior distribution at time t, that is,

$$\begin{aligned} p(x_t|y_{1:t};\theta) &= \frac{p(y_t|x_t;\theta)p(x_t|y_{1:t-1};\theta)}{p(y_t|y_{1:t-1};\theta)}, \quad \text{and} \qquad (6)\\ p(y_t|y_{1:t-1};\theta) &= \int p(y_t|x_t;\theta)p(x_t|y_{1:t-1};\theta)dx_t. \qquad (7) \end{aligned}$$

Except in a few cases such as linear Gaussian models [22], the integrals given in Equations (5)–(7) are in general hard to solve, as is in our model. To address this problem, the standard particle filter can be used to approximate these distributions.

The standard particle filter, also known as the sequential importance resampling (SIR) method, was proposed by Gordon et al. [10]. Suppose that we have a sample $x^{(\ell)}_{t-1}$, $\ell = 1, \ldots, L$ with probabilities $\pi^{(\ell)}_{t-1}$ from $p(x_{t-1}|y_{1:t-1};\theta)$. It is easy to notice that the simplest values of $\pi^{(\ell)}_{t-1}$ are $1/L$. An approximation to Equation (5) is given by:

$$p(x_t|y_{1:t-1};\theta) \approx \sum_{\ell=1}^{L} p(x_t|x^{(\ell)}_{t-1};\theta)\pi^{(\ell)}_{t-1}. \qquad (8)$$

For that reason, $p(x_t|y_{1:t-1};\theta)$ can be viewed as a mixture density with L components where $p(x_t|x_{t-1}^{(\ell)};\theta)$ represents the system equation conditional at each particle $x_{t-1}^{(\ell)}$. That would give us a sample $\widetilde{x}_t^{(\ell)}$, $\ell = 1,\ldots,L$, from the density $p(x_t|y_{1:t-1};\theta)$. Now, we can update the posterior distribution using Equation (6). We obtain a sample $\widetilde{x}_t^{(\ell)}$, $\ell = 1,\ldots,L$ from $p(x_t|y_{1:t};\theta)$ by assigning a probability of

$$\widetilde{\pi}_t^{(\ell)} = \frac{p(y_t|\widetilde{x}_t^{(\ell)};\theta)\pi_{t-1}^{(\ell)}}{\sum_{j=1}^{L} p(y_t|\widetilde{x}_t^{(j)};\theta)\pi_{t-1}^{(j)}} \tag{9}$$

to $\widetilde{x}_t^{(\ell)}$. Thus, we have a sample $\widetilde{x}_t^{(\ell)}$, $\ell = 1,\ldots,L$ with probabilities $\widetilde{\pi}_t^{(\ell)}$ from $p(x_t|y_{1:t};\theta)$. Finally, we resample L values (with replacement) from the particles $\widetilde{x}_t^{(\ell)}$ with weights $\widetilde{\pi}_t^{(\ell)}$ to obtain a sample from $p(x_t|y_{1:t};\theta)$, then restarts the procedure for time $t+1$.

3.2. *Particle Gibbs with ancestor sampling*

Now, particle filter and Gibbs sampling can be combined into an appropriate algorithm called the particle Gibbs with ancestor sampling [14]. Nevertheless, the computational cost of the particle Gibbs with ancestor sampling (PGAS) is $\mathcal{O}(NMn^5)$, where N is the number of particles, M is the number of PGAS iterations, and the term n^5 comes from the inversion of matrices of size t by t varying between 1 and n, which generates a computational cost of $\mathcal{O}(n^4)$. Additionally, the product of these terms is computed, which has a computational cost of n. This computational cost could limit its use in stochastic volatility models via Gaussian process, where the sample size n required for parameter estimation is usually high. That being so, a technique described in Doucet et al. [6] can be used, in which a data window is utilized to reduce the computational cost of matrix inversion. In this way, the factor n^3 can be reduced to J^3, where J is the window size and $J \ll n$. Consequently, the cost is reduced to $\mathcal{O}(NMn^2J^3)$.

To sample from the posterior distribution of the parameters $p(\theta|x_{1:n},y_{1:n})$, which is Step (a) of PGAS, a Metropolis-Hastings [12, 16] algorithm is used. The proposed distribution is a random walk where the order of magnitude of the variance is calculated following the idea of Braak [4] with $\widetilde{\Sigma} = 2.38^2/d \times \widetilde{\text{var}}(\theta)$. Here, $d = 5$ is the dimension of the parameter vector and $\widetilde{\text{var}}(\theta)$ is an estimate of the posterior variance of θ. It follows that Step (b) of PGAS consists of sampling from the states, $p(x_{1:n}|\theta,y_{1:n})$, via SIR given in Section 3.1 by selecting one of the trajectories based on π_n^j, for $j = 1,2,\ldots,L$.

4. Applications

In this section, we present a small Monte Carlo study and two real data applications with Brazilian companies in the BOVESPA stock exchange. All computations were carried out using `R` software [18] for fitting the GARCH model and running the PGAS of the stochastic volatility model via Gaussian process, whilst OpenBUGS [20] for running the MCMC of the classical stochastic volatility model.

4.1. *Monte Carlo study*

In this section, we conduct a small Monte Carlo study using 20 replicas for two sets of parameters and two different sample sizes, $n = \{100, 200\}$. For each possibility, the PGAS algorithm was executed twice (in parallel) for 5,000 iterations, with a burn-in period of 1,000 iterations and kept 1 from every 4 draws (a posterior sample size of 2,000 observations). A window size (J) of 10 data points was utilized. The objective of the Monte Carlo study presented in this section is to empirically evaluate the methodology adopted from the literature and outlined in Section 3. The stochastic volatility model via Gaussian process has its parameters and state variables inferred through the PGAS algorithm described in Sections 3.1 and 3.2.

The evaluation focus on the consistency of point estimates and the level of credibility of interval estimates for the model parameters. Thus, our results include the mean of means and the median of medians, accompanied by their respective standard deviations, as well as the 2.5% and 97.5% percentiles based on the 20 replicas.

The covariance function used is the squared exponential for two reasons: firstly, as noted by Rasmussen and Williams [19], this kernel adapts well to functions with complex shapes; secondly, it offers an intuitive interpretation of its parameters, (h^2,ℓ^2). Whilst other kernels with long-term structures could be considered for future work, they are not explored here.

To complete our Bayesian model specification, we set the following prior distributions: $a \sim \mathcal{N}(0;\, 3)$, $b \sim \mathcal{N}(0;\, 3)$, $h^2 \sim \mathcal{G}(1/40;\, 1/20)$, $\ell^2 \sim \mathcal{G}(256/100;\, 16/100)$ and $\tau^2 \sim \mathcal{G}(1/40;\, 1/20)$, where $\mathcal{G}(a, b)$ is the gamma distribution with mean a/b.

Table 1. Summary results of a Monte Carlo study with 20 replicas for two sets of parameters and two different sample sizes, $n = \{100,\, 200\}$. PGAS was run for 5,000 iterations and the burn-in period used was 1,000 iterations. A window size of 10 data points was employed. We report the mean of means and the median of medians, along with standard deviations and the 2.5% and 97.5% percentiles bases on the 20 replicas.

Parameter	True	Mean	Median	Std.Dev	$2,5\%$	$97,5\%$
			$n = 100$			
a	0.2	0.080	0.099	0.151	-0.208	0.390
b	-0.2	-0.155	-0.181	0.057	-0.268	-0.041
h^2	10	12.802	11.988	3.972	6.380	22.094
ℓ^2	30	23.101	21.115	8.768	9.900	43.675
τ^2	5	8.985	7.579	4.128	3.585	19.290
a	0.5	0.291	0.342	0.173	-0.070	0.601
b	0.3	0.254	0.269	0.078	0.097	0.404
h^2	5	8.033	7.432	3.276	3.653	16.377
ℓ^2	13	15.139	14.739	8.611	3.043	36.271
τ^2	10	16.774	13.791	10.895	3.603	45.736
			$n = 200$			
a	0.2	0.080	0.099	0.115	-0.208	0.390
b	-0.2	-0.155	-0.181	0.044	-0.268	-0.041
h^2	10	12.802	11.988	2.911	6.380	22.094
ℓ^2	30	23.101	21.115	8.091	9.900	43.675
τ^2	5	8.985	7.579	3.626	3.585	19.290
a	0.5	0.285	0.309	0.132	0.017	0.532
b	0.3	0.243	0.252	0.061	0.121	0.361
h^2	5	8.256	7.088	2.401	4.552	13.737
ℓ^2	13	16.523	14.739	7.946	5.268	36.037
τ^2	10	19.138	13.791	9.571	6.270	42.246

Table 1 shows the summary results of our Monte Carlo study. By visual inspection of trace plots, convergence was reached quickly for posteriors of each dataset, although they showed high correlated samples for some parameters. Given the standard deviation, the mean of means and the median of medians are relatively close to the true values of the parameters. Moreover, the 2.5% and 97.5% percentiles include the true values of the parameters, nevertheless the intervals may seem large due to the small sample sizes of our artificial data. Furthermore, when the sample size increases from 100 to 200, the overall standard deviations decrease, showing consistency of the proposed model. Despite of being a small Monte Carlo, we believe that the results are satisfactory and encouraging. Thus, we proceed with our real data applications in the following.

4.2. *Real data applications*

The objective of this section is to compare our proposed stochastic volatility model via Gaussian process given in Section 2.3 with the GARCH model and with the stochastic volatility model presented in Sections 2.1 and 2.2, respectively. Our comparisons are based on the L_1 and L_2 losses [11] defined as follows:

$$L_1 = \sum_{t=1}^{n} \left| \hat{\sigma}_t^2 - y_t^2 \right| \quad \text{and} \quad L_2 = \sum_{t=1}^{n} \left(\hat{\sigma}_t^2 - y_t^2 \right)^2. \tag{10}$$

Note that y_t^2 is used as a proxy of the conditional variance whilst $\hat{\sigma}_t^2$ comes from each fitted model. An approximation of $\mathrm{E}(\exp(x_t)|y_{1:n})$, from a sample from the posterior distribution, is used for $\hat{\sigma}_t^2$ for the stochastic volatility models.

The data consists of the returns from Eletrobras and Petrobras stocks - two Brazilian companies - traded on BOVESPA[1]. Both series were collected from the Yahoo Finance website[2], containing data on the assets for the period between 2012/09/04 and 2016/08/24. Figures 1 and 2 show the price and return series for Eletrobras and Petrobras from September 4, 2012, to August 24, 2016, respectively.

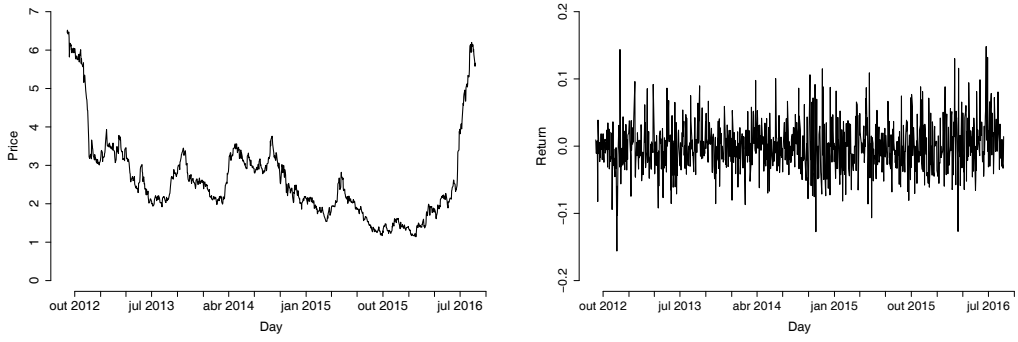


Figure 1. Price (left) and return (right) of Eletrobras stock traded on BOVESPA from 2012/09/04 to 2016/08/24.

The Ljung-Box test [see for example 21] applied to Eletrobras' returns resulted in p-values greater than or equal to 0.1 for degrees of freedom between 2 and 10. However,

[1] `https://www.b3.com.br/en_us/`

[2] `https://finance.yahoo.com/`

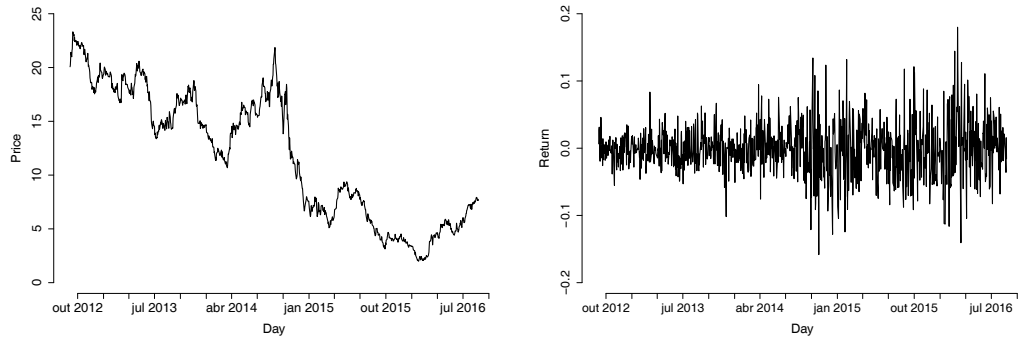


Figure 2. Price (left) and return (right) of Petrobras stock traded on BOVESPA from 2012/09/04 to 2016/08/24.

the same test applied to squared returns yielded values below 0.01, indicating the presence of heteroskedasticity and highlighting the need to model volatility.

The same test was conducted for Petrobras' returns and their squares, with p-values exceeding 0.2 for the former and very low values (below 10^{-5}) for the latter. This also suggests a requirement to model volatility.

In Table 2, it can be seen that both kurtoses of the unconditional log return distributions are greater than 3, which is a characteristic of a distribution with heavy tails. Another relevant fact is that the skewness coefficients are positives and both close to zero, indicating a slight right skew. Notwithstanding, this does not diminish our modelling and fitting of the data because the hierarchical structure of our model accounts for the heavy-tail effect in log returns [5].

Table 2. Summary statistics of Eletrobras and Petrobras stocks.

Stock	Mean	Standard Deviation	Skewness	Excess kurtosis
Eletrobras	-0.00013	0.0367	0.2652	1.264
Petrobras	-0.00097	0.0381	0.1585	1.870

For both datasets - Eletrobras and Petrobras - we ran the PGAS algorithm (see Sections 3.1 and 3.2) twice in parallel for 20,000 iterations and discarded the first 10,000 draws from each chain. A window size (J) of 10 data points was used to estimate the states. The convergence of the chains was again checked by visually inspecting the trace plots of all parameters. Table 3 brings model comparison by means of L_1 and L_2 losses. For theses datasets, our proposed stochastic volatility model via Gaussian process outperforms both GARCH and stochastic volatility models. As

Table 3. Model comparisons by means of L_1 and L_2 losses.

Model	Eletrobras		Petrobras	
	L_1	L_2	L_1	L_2
GP Vol.	**0.90**	**0.002**	**0.87**	**0.002**
GARCH	1.40	0.006	1.50	0.007
SV	1.30	0.005	1.20	0.007

said before, our proposed model seems more flexible than the other two models in this paper even with just using the quadratic kernel. Figures 3 and 4 depict the marginal posterior distributions for a, b, ℓ^2, h^2 and τ^2 for the Eletrobras and Petrobras datasets, respectively. Note that a is close to 1 for both datasets, indicating persistent volatilities, whilst b is around zero, suggesting almost no leverage effect. Additionally, both ℓ^2 and h^2 are in the same range for both datasets, respectively.

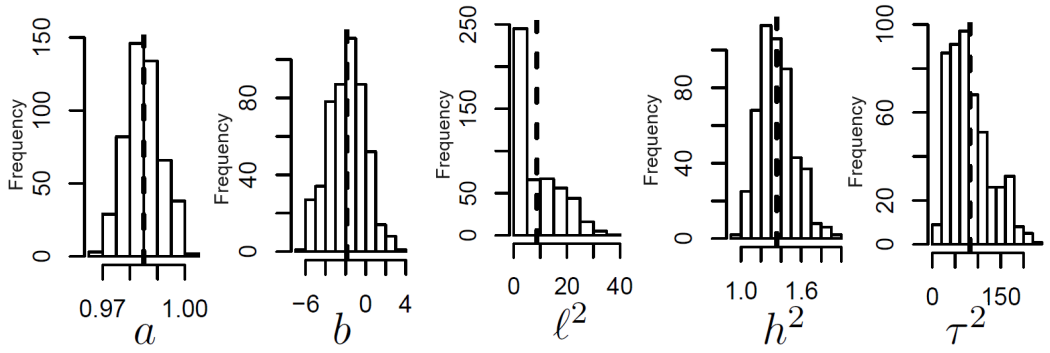


Figure 3. Summary of the posterior distribution for the stochastic volatility model via Gaussian process applied to Eletrobras returns

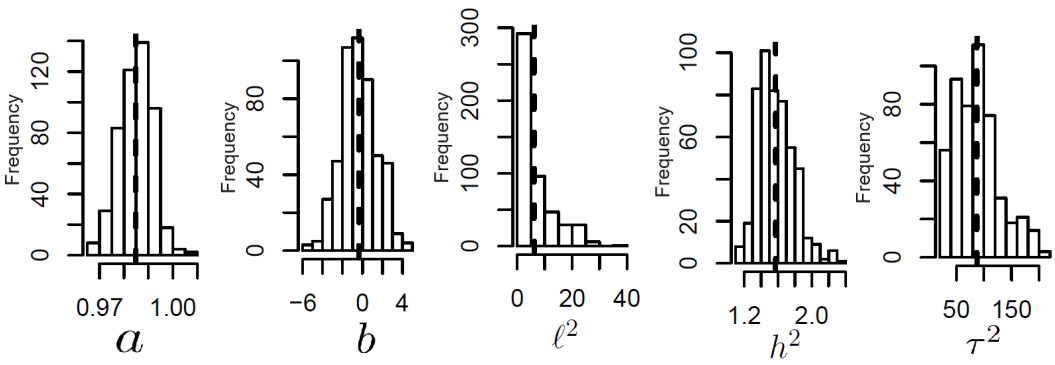
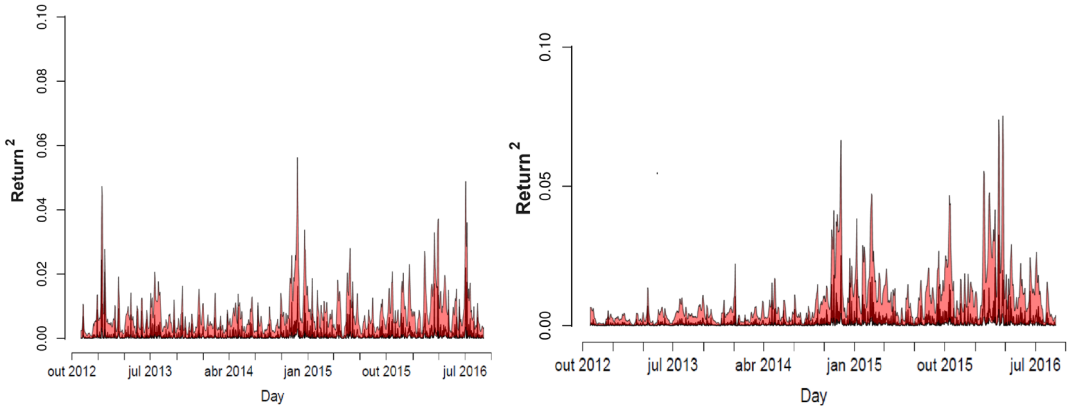


Figure 4. Summary of the posterior distribution for the stochastic volatility model via Gaussian process applied to Petrobras returns

Furthermore, Figure 5 illustrates the estimated posterior distributions of volatilities for both Eletrobras and Petrobras. When analyzed jointly with Figures 1 and 2, it can be observed that our proposed stochastic volatility model via Gaussian process effectively captures the key features of market volatility over time.

5. Concluding remarks

This paper develops Bayesian inference for the stochastic volatility model via Gaussian process through a hybrid algorithm of particle filters and Gibbs sampling. By using techniques to reduce computational complexity, the model achieves better results than conventional volatility models.



(a) ELETROBRAS (b) PETROBRAS

Figure 5. The estimated posterior distributions of volatilities for both Eletrobras and Petrobras. The black lines represent squared returns, whilst the red shaded region indicates the 0.95 credible interval.

This paper provides a broader understanding of stochastic volatility models, Gaussian processes, and the statistical modelling of financial assets. To reduce computational complexity, the first technique involved using a data window in both the likelihood function and the particle filter. The second technique was to start with a smaller sample to more quickly converge on the region with the highest probability mass. This paper also demonstrates the effectiveness of the PGAS algorithm through a Monte Carlo study, providing robust parameter estimations. It illustrates the efficiency of the stochastic volatility model using Gaussian processes by comparing it with other widely used models.

The proposed model was tested in restricted scenarios, but this allows for a more objective presentation of the idea. In the future, further comparisons and modifications to the model may be made based on this paper.

For example, other covariance functions such as the quadratic rational function or the Matérn function could be considered. Combinations of two or more covariance functions could also be considered, such as their product or sum.

Finally, a modification to the model structure could be made by replacing the stochastic volatility distribution. The most plausible modification would be to replace the normal distribution with the Student's t-distribution. This extension could, in principle, be done via a mixture of scale (precision) of the normal distribution with the gamma distribution. In this case, a heavy-tailed distribution for volatility would be admitted.

References

[1] Andrieu, C., Doucet, A., Holenstein, R., 2010. Particle Markov chain Monte Carlo methods. Journal of the Royal Statistical Society: Series B 72, 269–342. doi: 10.1111/j.1467-9868.2009.00736.x.

[2] Black, F., 1976. Studies of stock price volatility changes, in: Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association. pp.

177–181.

[3] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327. doi: 10.1016/0304-4076(86)90063-1.

[4] Braak, C.J.F.T., 2006. A Markov chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter spaces. Statistics and Computing 16, 239–249. doi: 10.1007/s11222-006-8769-1.

[5] Carmona, R., 2014. Statistical Analysis of Financial Data in R. second ed., Springer, New York, USA. doi: 10.1007/978-1-4614-8788-3.

[6] Doucet, A., Briers, M., Sénécal, S., 2006. Efficient block sampling strategies for sequential Monte Carlo methods. Journal of Computational and Graphical Statistics 15, 693–711. doi: 10.1198/106186006X142744.

[7] Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1008. doi: 10.2307/1912773.

[8] Francq, C., Zakoian, J.M., 2019. GARCH Models: Structure, Statistical Inference and Financial Applications. second ed., Wiley, New Jersey, USA. doi: 10.1002/9781119313472.

[9] Frigola, R., Lindsten, F., Schön, T.B., Rasmussen, C.E., 2013. Bayesian inference and learning in Gaussian process state-space models with particle MCMC, in: Burges, C.J., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems, Curran Associates, Inc.

[10] Gordon, N.J., Salmond, D.J., Smith, A.F.M., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proceedings F. 140, 107–113. doi: 10.1049/ip-f-2.1993.0015.

[11] Hansen, P.R., Lunde, A., 2005. A comparison of volatility models: Does anything beat a GARCH(1,1)? Journal of Applied Econometrics 20, 873–889. doi: 10.1002/jae.800.

[12] Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109. doi: 10.1093/biomet/57.1.97.

[13] Jacquier, E., G.Polson, N., Rossi, P.E., 2002. Bayesian analysis of stochastic volatility models. Journal of Business & Economic Statistics 12, 371–389. doi: 10.1080/07350015 .1994.10524553.

[14] Lindsten, F., Jordan, M.I., Schön, T.B., 2014. Particle Gibbs with ancestor sampling. Journal of Machine Learning Research 15, 2145–2184. URL: `http://jmlr.org/papers/v15/lindsten14a.html`.

[15] Mandelbrot, B., 1963. The variation of certain speculative prices. The Journal of Business 36, 394–419. doi: 10.1086/294632.

[16] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. The Journal of Chemical Physics

21, 1087–1092. doi: 10.1063/1.1699114.

[17] Pitt, M.K., Silva, R.S., Giordani, P., Kohn, R., 2012. On some proprieties of Markov chain Monte Carlo simulation methods based on the particle filter. Journal of Econometrics 171, 134–151. doi: 10.1016/j.jeconom.2012.06.004.

[18] R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.r-project.org/`.

[19] Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. MIT Press, California, USA. URL: `https://gaussianprocess.org/gpml/`.

[20] Thomas, A., 2014. OpenBUGS. OpenBUGS Foundation. Finland.

[21] Tsay, R.S., 2010. Analysis of Financial Time Series. third ed., Wiley, New Jersey, USA.

[22] West, M., Harrison, J., 1997. Bayesian Forecasting and Dynamic Models. second ed., Springer, New York, USA.

[23] Wu, Y., Lobato, J.M.H., Ghahramani, Z., 2014. Gaussian process volatility model. arXiv. doi: 10.48550/arXiv.1402.3085.