Hindawi
Discrete Dynamics in Nature and Society
Volume 2023, Article ID 7657430, 12 pages
https://doi.org/10.1155/2023/7657430

Research Article

Testing Data Cloning as the Basis of an Estimator for the Stochastic Volatility in Mean Model

E. Romero [1] and E. Ropero-Moriones [2]

[1] *Department of Statistics and OR, Complutense University, Madrid, Spain*
[2] *School of Applied Social Sciences and Communication, UNIE University, Madrid, Spain*

Correspondence should be addressed to E. Romero; evarom03@ucm.es

Received 17 February 2023; Revised 12 August 2023; Accepted 24 August 2023; Published 12 September 2023

Academic Editor: Tien Van Do

Developed as a refinement of stochastic volatility (SV) models, the stochastic volatility in mean (SVM) model incorporates the latent volatility as an explanatory variable in both the mean and variance equations. It, therefore, provides a way of assessing the relationship between returns and volatility, albeit at the expense of complicating the estimation process. This study introduces a Bayesian methodology that leverages data-cloning algorithms to obtain maximum likelihood estimates for SV and SVM model parameters. Adopting this Bayesian framework allows approximate maximum likelihood estimates to be attained without the need to maximize pseudo likelihood functions. The key contribution this paper makes is that it proposes an estimator for the SVM model, one that uses Bayesian algorithms to approximate the maximum likelihood estimate with great effect. Notably, the estimates it provides yield superior outcomes than those derived from the Markov chain Monte Carlo (MCMC) method in terms of standard errors, all while being unaffected by the selection of prior distributions.

1. Introduction

When analyzing time series of returns on financial assets, we need to consider their specific properties, particularly their volatility, i.e., whether they exhibit behaviors such as heteroskedasticity, volatility clustering [1, 2], and excess leptokurtosis [3]. Models such as GARCH [4, 5] and stochastic volatility (SV) [6, 7] have previously been developed to account for these properties.

GARCH models define the conditional variance as a function of the past squared innovations and lagged conditional variances [8, 9]. In contrast, variance in SV models is characterized as an unobserved component that follows a stochastic process [3, 10, 11].

Furthermore, whereas the SV model captures the deviation of returns from the mean using a function of two disturbance terms, the GARCH model relies on a single disturbance term [3]. This added complexity in the SV model allows for more flexibility [12, 13] and improved accuracy in capturing the volatility clustering of financial series [8, 14–18]. The SV model is also better at handling the negative relationship between volatility and returns. Furthermore, SV models are more robust against misspecification and radical changes in the data [13, 17] and are better at estimating the properties of financial series.

Recently, the stochastic volatility in mean (SVM) models has emerged as a refinement of the SV model. It allows the mean and variance of financial time series data to be modeled simultaneously, enabling the relationship between volatility and returns to be analyzed at the same time—an important aspect of financial modeling [3]. Other models, such as ARCH-M and GARCH-M, also attempt to estimate this relationship but do not provide a simultaneous estimation of the ex-ante relationship between volatility and returns. We can therefore expect SVM models to provide more accurate estimates of the behavior of financial time series data when analyzing leverage effects [19] and the effect of volatility feedback [3].

While SV models have been shown to be superior to GARCH models in the literature, they are not as widely used due to the complexity involved in their estimation. This is because it is difficult to evaluate the likelihood function directly and because they require us to estimate both return and volatility at the same time.

Various techniques have been used to estimate SV models, including methods based on the method of moments [20–22] and likelihood-based methods. The estimators of moments do have the advantage of not requiring a likelihood assessment, but their effectiveness is known to be suboptimal compared to likelihood-based inference methods [23]. However, likelihood-based methods have limitations as well, including that they are computationally intensive, require excessive simulation efforts, and call for assumptions to be made that can be difficult to satisfy. Financial markets often require real-time decision-making, which means the estimators need to be computationally fast and robust and are limited in the amount of sampling they can involve [24].

Bayesian methods, such as Monte Carlo Markov Chains (MCMC) [16, 25–31] and Integrated nested Laplace approximations [32], are good solutions for estimating the parameters of SV models because they allow for efficient evaluation of the posterior distribution of parameters and volatility. However, these methods also have limitations such as requiring a prior distribution for the parameters and a numerical evaluation of the likelihood function, and they can also experience problems getting the simulated chains to converge [33].

This paper proposes using a different approach called "data cloning" [34] to estimate the parameters, utilizing the computational simplicity of MCMC algorithms while also enabling frequentist inferences, such as maximum likelihood estimates and standard errors, to be made. The method involves applying a Bayesian methodology to a dataset constructed by cloning the original dataset as many times as necessary for the solution to approximate the maximum likelihood estimate [35, 36]. The main advantage of using data cloning over other Bayesian methods is that the inferences are invariant to the choice of the prior distributions, and no likelihood estimation is required. Overall, data cloning is a powerful method for estimating and studying complex models, especially when analyzing volatility.

We propose using this methodology to estimate the parameters of SV and SVM models as it has been shown to be particularly useful for complex models, as discussed in studies by authors in [34–37]. Recently, this method has been successfully used to estimate the parameters of other complex financial models in [38, 39]. Although it is beyond the scope of this article, models have recently been developed to estimate volatility in the valuation of financial options using two volatility components [40, 41]. These models are strong potential candidates for using an algorithm like the one we have constructed in this paper to estimate their parameters.

This paper makes three important contributions to the literature. First, it provides an algorithm to estimate SV and SVM model parameters based on the data-cloning method. This is a simpler way of estimating SVM that allows frequentist inferences to be obtained without having to estimate likelihood. Second, by performing an analysis of simulated data using the proposed algorithm, we show that its estimates are more accurate than those obtained using the MCMC method. Third, in order to evaluate the predictive ability of the model over a real financial series, the methodology is applied to model Bitcoin returns, allowing us to draw new conclusions about the relationship between volatility and profitability in cryptocurrencies—conclusions that can only be obtained with the SVM method.

The structure of the article is as follows. In Section 2, we specify the SV and SVM models that will be used, and in Section 3, we explain the data-cloning method in general terms. In Section 4, we lay out the algorithms required to apply this method to the SV and SVM models, then obtain the results and compare them with the MCMC methodology, demonstrating that the data-cloning method is superior. In Section 5, we apply SVM to a real example of a financial series (Bitcoin) and analyze the relationship between return and volatility to test the hypotheses of the leverage effect and volatility feedback. Finally, in Section 6, we present the main conclusions of the paper.

2. Definition and Specification of the SV and SVM Models

Definition 1. The stochastic volatility model defines the returns of process Y_t in discrete time t as

$$Y_t = \mu_t + \sigma_t\epsilon_t, \epsilon_t \sim \mathrm{NID}(0,1), \quad (1)$$

$$\mu_t = a + \sum_{i=1}^{k} b_i x_{i,t}, \quad \text{for } t = 1,2,\ldots. \quad (2)$$

Here, $x_{i,t}$ can be independent variables or lags of the dependent variable. The mean μ_t also depends on constants a and b_i for $i = 1,\ldots, k$ regression coefficients. In the volatility process, σ_t^2 is defined as

$$\sigma_t^2 = \sigma^{*2} e^{h_t}, \quad (3)$$

where σ^{*2} is a positive scaling factor and $h_t t$ is a stochastic process defined as

$$h_t = \phi h_{t-1} + \sigma_\eta \eta_t, \eta_t \sim \mathrm{NID}(0, 1). \quad (4)$$

In (4), ϕ and σ_η are model parameters. Parameter σ_η is the variance of the independent and identically distributed normal variables η_t, while ϕ is the volatility persistence parameter. It is important for ϕ to be positive and smaller than 1 ($\phi \in (0, 1)$) to ensure stationarity.

It can be assumed that in (3), σ_t^2 is specified in the logarithmic form, considering that $h_t = \ln(\sigma_t^2/\sigma^{*2})$.

The SV model has two sources of variability in the form of two independent and mutually uncorrelated disturbance terms, ϵ_t and ηt. This is the main difference between SV models and GARCH models [3, 4].

The unconditional variance implied in the SV model is $\sigma^{*2} e^{\sigma_\eta/2(1-\phi^2)}$.

One important characteristic of SV models is that they capture part of the excess of kurtosis that financial series present. The kurtosis of an SV series is defined by

$$k_y = \frac{k_\epsilon E\left(\sigma_t^4\right)}{\left(E\left(\sigma_t^2\right)^2\right)} = 3e^{\sigma_\eta^2/1-\phi^2}. \quad (5)$$

Definition 2. The return of the stochastic volatility in mean (SVM) model is defined as (1), and its mean is defined as

$$\mu_t = a + \sum_{i=1}^{k} b_i x_{i,t} + d\sigma_t^2, \quad (6)$$

where parameter d measures the effect of volatility in the mean of the process.

The variance of the SVM model is defined by equations (3) and (4).

The inclusion of variance in the mean equation allows for a better understanding of the relationship between returns and volatility. It enables studies, such as [42], to be performed and the returns' partial dependence on volatility to be analyzed, as featured in all financial theory [3].

3. Data-Cloning Estimation

The estimation of these models, particularly SVM, is not straightforward. To resolve this, this paper proposes a technique based on data cloning to obtain approximates of the maximum likelihood estimates through Bayesian algorithms. The main idea is to clone the series k times and assume that each series represents an independent sample path of the process. We consider all sample paths to be equal because the sample path with the highest probability is the one obtained. Although the heuristic explanation implies that the cloned trajectories are independent, the mathematical proof of the algorithm does not rely on this assumption, and it does not assume that the k clones are independent.

This method was introduced by the authors in [34, 43], as a means of obtaining maximum likelihood estimates for parameters of complex models where direct maximization of the likelihood is infeasible.

The data-cloning method offers an effective solution for estimating the parameters of SV and SVM models as it avoids the need for direct maximization of the likelihood function. Instead, it utilizes Bayesian algorithms to approximate the likelihood. Moreover, this methodology is not reliant on the specific prior distributions chosen, resulting in improved solutions compared to those provided by MCMC estimators.

Previous studies by the authors in [39, 44] have successfully applied this method to estimate the SV model, albeit using a less general model. Their findings demonstrate the enhanced accuracy of the parameter estimation compared to the standard Bayesian approach. Therefore, we aim to assess the effectiveness of this method in the context of a more general SV model and the SVM model.

The data-cloning method begins with an observed dataset $y = (y_1, y_2, \ldots, y_n)$ and the prior distributions for the parameters. It utilizes the posterior distribution of the parameter set θ, denoted as $\pi(\theta|y)$, which is proportional to the likelihood function $L(\theta|y)$ multiplied by the prior distribution $\pi(\theta)$. This posterior distribution is then used to generate samples using an MCMC method. In the data-cloning method, samples are drawn from the posterior distribution $\pi^{(k)}(\theta|y)$, which is proportional to the k-th power of the likelihood function $[L(\theta \mid y)]^{(k)}$ multiplied by the prior distribution $\pi(\theta)$.

The data-cloning method is based on the principle that when k is sufficiently large, $\pi^{(k)}(\theta \mid y)$ converges to a multivariate normal distribution that has the maximum likelihood estimate of the model parameters as its mean. In addition, the covariance of this multivariate normal distribution is equal to $1/k$ times the inverse of the Fisher information matrix for the maximum likelihood estimate [34]. The data-cloning algorithm can therefore be summarized in the following steps:

Step 1: Create k-cloned dataset $\mathbf{y}^{(k)} = (\mathbf{y}, \mathbf{y}, \ldots, \mathbf{y})$ by cloning the observed dataset k times.

Each copy of y is treated as an independent sample path of the same process.

Step 2: Use an MCMC method to generate random values from the posterior distribution. Start the algorithm with the prior distribution $\pi(\theta)$ and the cloned data vector $\mathbf{y}^{(k)} = (\mathbf{y}, \mathbf{y}, \ldots, \mathbf{y})$.

Step 3: After running the MCMC method for a total of B iterations, compute the sample means and variances of the values obtained for the marginal posterior distribution, denoted as $(\theta)_j$, where $j = 1, \ldots, B$. The sample means correspond to the maximum likelihood estimates, while the approximate variances of the maximum likelihood estimates are k times the posterior variances.

4. Data-Cloning Algorithms to Estimate the SV and SVM Models

To simplify the estimation algorithms for both models, the estimation of the constant parameter was excluded. Although it is possible to include this parameter in the algorithms, doing so significantly increases the computation time as it requires a higher number of clones. After conducting several empirical tests, we observed that excluding the constant parameter does not significantly affect the results, so we decided to omit it in the simulations and work with variables in differences.

4.1. Data-Cloning Estimator for the SV Model. The algorithm based on the data-cloning method was used to estimate the model parameters for the SV model described in Section 2 by equations (1), (3), and (4) and simplifying equation (2) to

$$\mu_t = b\tilde{y}_{t-1}, \tag{7}$$

being $\tilde{y} = y_t - \overline{y}$ (the returns in differences).

The model included just one autoregressive term. More autoregressive terms, or other kinds of terms, could be easily included if necessary, but each additional term would probably increase the required number of clones to achieve convergence, and consequently the computation time.

This model was characterized by four parameters: ϕ, σ_η, σ^{*2}, and b.

To apply the data-cloning method, an MCMC procedure needed to be designed, and prior distributions therefore needed to be chosen, even though it has been proven that they do not affect the final results [34]. In light of this, the following vaguely informative distributions were chosen as prior distributions: $\phi \sim U(0,1)$, $\sigma_\eta \sim U(0,10)$, $\sigma^{*2} \sim U(0,10)$, and $b \sim U(-10,10)$.

The joint posterior distribution was obtained by assuming that $y_i \sim N(\mu_t, \sigma_t^2)$, with μ_t defined in (7) and σ_t^2 defined in (3). The likelihood function of the SV model was therefore

$$L\left(b, \sigma^{*2}, \phi, \sigma_\eta | \tilde{y}\right) = \left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^{*2}e^{h_i}}}\right) \exp\left\{-\frac{1}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1}\right)^2}{\exp\left(h_i\right)}\right\}, \tag{8}$$

being $h_i i$ defined by (4). With this likelihood function, the joint posterior was

$$\begin{aligned}\pi^{(k)}\left(\phi, \sigma_\eta, \sigma^{*2}, b\right) &\propto \left[L\left(\phi, \sigma_\eta, \sigma^{*2}, b | \tilde{y}\right)\right]^k \pi(\phi)\pi\left(\sigma_\eta\right)\pi\left(\sigma^{*2}\right)\pi(b) \\ &\propto \left(\prod_{i=1}^{n} 2\pi\sigma^{*2}e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1}\right)^2}{\exp\left(h_i\right)}\right\} \\ &\quad \cdot I_{(0,1)}(\phi) I_{(0,10)}\left(\sigma_\eta\right) I_{(0,10)}\left(\sigma^{*2}\right) I_{(-10,10)}(b).\end{aligned} \tag{9}$$

The conditional posterior distributions for the parameters were

$$\begin{aligned}\pi^{(k)}\left(\phi | \sigma_\eta, \sigma^{*2}, b, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2}e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1}\right)^2}{\exp\left(h_i\right)}\right\} I_{(0,1)}(\phi), \\ \pi^{(k)}\left(\sigma_\eta \middle|, \phi, \sigma^{*2}, b, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2}e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1}\right)^2}{\exp\left(h_i\right)}\right\} I_{(0,10)}\left(\sigma_\eta\right), \\ \pi^{(k)}\left(\sigma^{*2} | \phi, \sigma_\eta, b, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2}e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1}\right)^2}{\exp\left(h_i\right)}\right\} I_{(0,10)}\left(\sigma^{*2}\right), \\ \pi^{(k)}\left(b | \phi, \sigma_\eta, \sigma^{*2}, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2}e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1}\right)^2}{\exp\left(h_i\right)}\right\} I_{(-10,10)}(b).\end{aligned} \tag{10}$$

The data-cloning algorithm started from an initial solution $\phi^{(0)}$, $\sigma_\eta^{(0)}$, $\sigma^{*2(0)}$, and $b^{(0)}$. From the conditional posterior distributions, it generated values for $\phi^{(m)}$, $\sigma_\eta^{(m)}$, $\sigma^{*2(m)}$, and $b^{(m)}$ in each iteration m. The initial values were simulated directly from the prior distributions since there was no need to use specific values to achieve convergence within a reasonable time.

After a sufficient number of iterations, a sample was obtained to constitute the posteriors whose means formed the basis of the maximum likelihood estimates of the model parameters. The steps of this algorithm can be summarized as follows:

Step 1: set initial solution at $m = 0$ as $\phi^{(0)}$, $\sigma_\eta^{(0)}$, $\sigma^{*2(0)}$, and $b^{(0)}$.

Step 2: generate $\phi^{(m+1)}$ from its conditional posterior distribution

$$\sigma_\eta^{(m)} \sim \pi^{(k)}\left(\sigma_\eta \middle| \phi, \sigma^{*2}, b, \tilde{y}\right). \quad (11)$$

Step 3: generate $\sigma_\eta^{(m)}$ from its conditional posterior distribution

$$\sigma_\eta^{(m)} \sim \pi^{(k)}\left(\sigma_\eta \middle| \phi, \sigma^{*2}, b, \tilde{y}\right). \quad (12)$$

Step 4: generate $\sigma^{*2(m)}$ from its conditional posterior distribution

$$\sigma^{*2(m)} \sim \pi^{(k)}\left(\sigma^{*2} | \phi, \sigma_\eta, b, \tilde{y}\right). \quad (13)$$

Step 5: generate $b^{(m)}$ from its conditional posterior distribution

$$b^{(m)} \sim \pi^{(k)}\left(b | \phi, \sigma_\eta, \sigma^{*2}, \tilde{y}\right). \quad (14)$$

Step 6: set $m = m + 1$ and go to Step 2.

This algorithm was implemented using the dclone package [37] from the R project [45].

To test the performance of the algorithm in estimating the parameters of the SV model, a sample path of this model was simulated. This allowed us to compare the real parameters against the estimated ones. A simulator for this model was developed using R to generate the series, which consisted of 245 values, representing approximately the number of business days in a year. The purpose of this was to assess the performance of the algorithm over the annual evolution of the daily returns of a financial asset. The selected parameter values for simulating the model were $\phi = 0.97$, $\sigma_\eta = 0.12$, $\sigma^{*2} = 0.2$, and $b = 0.2$.

The data-cloning algorithm requires the optimal number of clones to be determined. This is achieved by evaluating the maximum eigenvalue of the posterior variance, the minimum squared error, the R^2 statistic, and the $\widehat{R}$ criterion [43, 46]. All these metrics can be computed using the dclone package. Based on these results, no significant improvements were found by using more than 20 clones, so the optimal number of clones is fixed at 20.

The results obtained by applying the algorithm to a single sample path are presented in Table 1. It displays the real values for all parameters, the estimated parameters, the standard errors, and the 95% confidence intervals. In addition, the last two columns include the parameter estimates using an MCMC estimator and the corresponding standard errors of estimates. This allows a comparison to be made with the results obtained using data cloning.

Bearing in mind that only one sample path was simulated, we can observe that the estimator produces values that closely match the real values used in generating the path. In addition, the standard errors of estimates are very small across the board, indicating that the estimator yields good results, based on a single sample path. Moreover, all real values fall within the 95% confidence intervals, as expected.

Comparing these results with those obtained using a traditional MCMC estimator, data cloning demonstrates superior performance in almost all cases. It provides estimates with smaller standard errors that do not depend on the selected priors.

Figure 1 shows the posterior distributions obtained by the data-cloning algorithm, providing a better understanding of the estimates' behavior. We can see a slight tendency to underestimate the value of ϕ, but for the remaining parameters, the higher probabilities of the posterior density function closely are aligned with the true parameter values.

It should be noted that data-cloning estimates are approximations of maximum likelihood estimates, so they will have the same analytical properties.

4.2. Data-Cloning Estimator for the SVM Model. The estimator for the SVM model based on the data-cloning method also required the mean equation (6) to be simplified in order to work with the returns in differences and fix the variables to be used. Thus, the equation of the mean is defined by

$$\mu_t = b\tilde{y}_{t-1} + d\sigma^{*2} e^{h_t}. \quad (15)$$

Again, a single autoregressive term was included to simplify the algorithm execution. Thus, the model had five parameters: ϕ, σ_η, σ^{*2}, b, and, d i.e., one more than the SV model. The prior distributions used in the algorithm were $\phi \sim U(0, 1)$, $\sigma_\eta \sim U(0, 10)$, $\sigma^{*2} \sim U(0, 10)$, $b \sim U(-10, 10)$, and $d \sim U(-10, 10)$.

The joint posterior distribution was obtained considering that $y_i \sim N(\mu_t, \sigma_t^2)$, with μ_t defined in (7) and σ^2 defined in (3), so the likelihood function of the SVM model was

$$L\left(\phi, \sigma_\eta, \sigma^{*2}, b, d | \tilde{y}\right) = \left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^{*2} e^{h_i}}}\right) \exp\left\{-\frac{1}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1} - d\sigma^{*2} \exp\left(h_i\right)\right)^2}{\exp\left(h_i\right)}\right\}, \quad (16)$$

being h_i defined by (4).

Based on this likelihood function, the joint posterior was

Table 1: Estimates of the stochastic volatility model parameters using the data-cloning method.

Parameter	Real value	Data-cloning estimates	SD	95% confidence intervals	MCMC estimates	SD (MCMC)
ϕ	0.97	0.8879	0.03931	(0.5433 and 1.2324)	0.8335	0.2077
σ_η	0.12	0.1478	0.03758	(−0.1816 and 0.4771)	0.1910	0.1220
σ^{*2}	0.2	0.2130	0.06577	(−0.3635 and 0.7895)	0.2036	0.0392
b	0.2	0.2192	0.01462	(0.0911 and 0.3474)	0.1199	0.0676

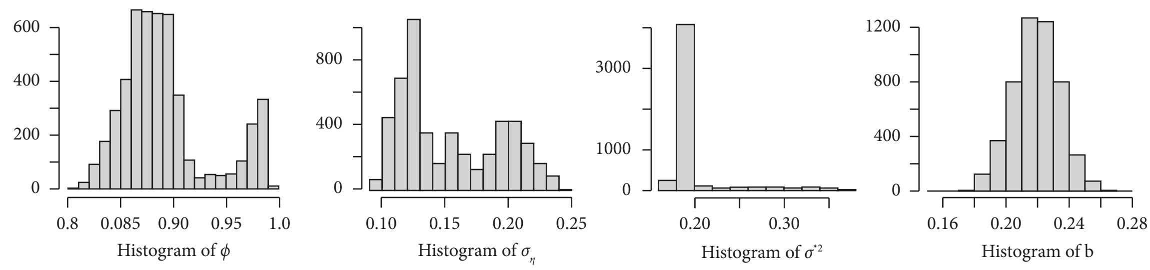

Figure 1: Histograms of the posterior distributions of the stochastic volatility model parameters.

$$\begin{aligned}\pi^{(k)}\left(\phi, \sigma_\eta, \sigma^{*2}, b, d\right) &\propto \left[L\left(\phi, \sigma_\eta, \sigma^{*2}, b, d \mid \tilde{y}\right)\right]^k \pi(\phi)\pi\left(\sigma_\eta\right)\pi\left(\sigma^{*2}\right)\pi(b)\pi(d) \\ &\propto \left(\prod_{i=1}^{n} 2\pi\sigma^{*2} e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1} - d\sigma^{*2}\exp\left(h_i\right)\right)^2}{\exp\left(h_i\right)}\right\} \\ &\quad \cdot I_{(0,1)}(\phi) I_{(0,10)}\left(\sigma_\eta\right) I_{(0,10)}\left(\sigma^{*2}\right) I_{(-10,10)}(b) I_{(-10,10)}(d).\end{aligned} \tag{17}$$

From this likelihood, the conditional posteriors were

$$\begin{aligned}\pi^{(k)}\left(\phi \mid \sigma_\eta, \sigma^{*2}, b, d, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2} e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1} - d\sigma^{*2}\exp\left(h_i\right)\right)^2}{\exp\left(h_i\right)}\right\} I_{(0,1)}(\phi), \\ \pi^{(k)}\left(\sigma_\eta \mid \phi, \sigma^{*2}, b, d, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2} e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1} - d\sigma^{*2}\exp\left(h_i\right)\right)^2}{\exp\left(h_i\right)}\right\} I_{(0,10)}\left(\sigma_\eta\right), \\ \pi^{(k)}\left(\sigma^{*2} \mid \phi, \sigma_\eta, b, d, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2} e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1} - d\sigma^{*2}\exp\left(h_i\right)\right)^2}{\exp\left(h_i\right)}\right\} I_{(0,10)}\left(\sigma^{*2}\right), \\ \pi^{(k)}\left(b \mid \phi, \sigma_\eta, \sigma^{*2}, d, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2} e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1} - d\sigma^{*2}\exp\left(h_i\right)\right)^2}{\exp\left(h_i\right)}\right\} I_{(-10,10)}(b), \\ \pi^{(k)}\left(d \mid \phi, \sigma_\eta, \sigma^{*2}, b, \tilde{y}\right) &\propto \left(\prod_{i=1}^{N} \sigma^{*2} e^{h_i}\right)^{-k/2} \exp\left\{-\frac{k}{2\sigma^{*2}} \sum_{i=1}^{n} \frac{\left(\tilde{y}_i - b\tilde{y}_{i-1} - d\sigma^{*2}\exp\left(h_i\right)\right)^2}{\exp\left(h_i\right)}\right\} I_{(-10,10)}(d).\end{aligned} \tag{18}$$

The algorithm started from an initial solution $\phi^{(0)}$, $\sigma_\eta^{(0)}$, $\sigma^{*2(0)}$, $b^{(0)}$, and $d^{(0)}$, and consider these values generated the intermediate solutions ($\phi^{(m)}$, $\sigma_\eta^{(m)}$, $\sigma^{*2(m)}$, $b^{(m)}$, and $d^{(m)}$) in each iteration (m) from the conditional posterior distributions.

They were then used to obtain the posterior sample, and its arithmetic means constituted the maximum likelihood estimates. The algorithm steps can be summarized as follows:

Step 1: set initial solution at $m = 0$ as: $\phi^{(0)}$, $\sigma_\eta^{(0)}$, $\sigma^{*2(0)}$, $b^{(0)}$, and $d^{(0)}$.

Step 2: generate $\phi^{(m+1)}$ from its conditional posterior distribution

$$\phi^{(m+1)} \sim \pi^{(k)}\left(\phi | \sigma_\eta, \sigma^{*2}, b, d, \tilde{y}\right). \tag{19}$$

Step 3: generate $\sigma_\eta^{(m)}$ from its conditional posterior distribution

$$\sigma_\eta^{(m)} \sim \pi^{(k)}\left(\sigma_\eta \middle| \phi, \sigma^{*2}, b, d, \tilde{y}\right). \tag{20}$$

Step 4: generate $\sigma^{*2(m)}$ from its conditional posterior distribution

$$\sigma^{*2(m)} \sim \pi^{(k)}\left(\sigma^{*2} | \phi, \sigma_\eta, b, d, \tilde{y}\right). \tag{21}$$

Step 5: generate $b^{(m)}$ from its conditional posterior distribution

$$b^{(m)} \sim \pi^{(k)}\left(b | \phi, \sigma_\eta, \sigma^{*2}, d, \tilde{y}\right). \tag{22}$$

Step 6: generate $d^{(m)}$ from its conditional posterior distribution

$$d^{(m)} \sim \pi^{(k)}\left(d | \phi, \sigma_\eta, \sigma^{*2}, b, \tilde{y}\right). \tag{23}$$

Step 7: set $m = m + 1$ and go to Step 2.

Again, the dclone package [37] from the R project [45] was used to program the algorithm analogously to how the data-cloning algorithm was programmed to estimate the SV model. The initial values were simulated directly from the prior distribution.

The same procedure as for the SV model was followed to analyze the quality of the estimates. Therefore, a series with 245 observations was simulated using the following parameters for the model: $\phi = 0.97$, $\sigma_\eta = 0.12$, $\sigma^{*2} = 0.2$, $b = 0.2$, and $d = 0.1$. The parameters of the model were then estimated using the series data, and the proximity of the estimated values to the real values and the standard errors of estimate were examined. Confidence intervals were also obtained for the parameters to test whether they include the true values.

To determine the optimal number of clones, the following criteria from the dclone package were employed: maximum eigenvalue of the posterior variance, minimum squared error, R^2, and $\widehat{R}$ [43, 46]. As this model had one more parameter than the previous one, a considerably higher number of clones needed to be used in order to achieve convergence. After testing several estimates, we concluded that 40 clones are sufficient to make high-quality estimates that are not substantially improved by including a larger number of clones. Hence, the optimal number of clones is set at 40.

Table 2 shows the real data used to estimate the series, the estimates obtained, the standard errors of estimate, and the confidence intervals for each parameter. It also includes the estimates obtained using the MCMC method, with their respective standard errors, for the purpose of comparing the quality of the estimates yielded by the two methodologies. Figure 2 displays the posterior distributions of the parameters obtained by the algorithm.

Only one trajectory was considered, but we can see that the estimation algorithm provides values very close to the real values of the parameters used in simulating it. The standard errors of estimates are also small enough to prove that the estimates obtained are of high quality. Finally, we can see that the 95% confidence intervals include the real values of the parameters and that the estimates are better than those obtained by an MCMC procedure in terms of the standard errors of estimates. The histograms demonstrate the close correspondence between the estimated SVM parameters and the real values. From the results obtained, we can observe that the values of ϕ, σ^{*2}, and d are overestimated and the values of σ_η and b are underestimated.

Although we can see that the estimates based on a single trajectory are good enough, different trajectories were also estimated from the same parameters, obtaining the average of all the estimates as a result. As expected, this method provided values that are even closer to the true values of the parameters. We do not include further details of this option as it may not be applicable to real data, where only a single trajectory is available. However, it is worth mentioning that this approach enhances the quality of the estimator by reducing variance and improving the accuracy of the mean value.

The algorithm was also evaluated with different size sample paths, showing good performance in all of them. We observed that when the sample paths were small in size, the estimation results depended to a greater extent on the path considered. In contrast, convergence was achieved with a number of clones even smaller than the 40 clones proposed. When the size of the sample paths was moderately large, the estimates were more stable and depended less on the path considered, but in some cases, more than 40 clones needed to be used in order to reach convergence. The results are summarized in Table 3.

5. Applying the Estimators to Real Data: Bitcoin

Cryptocurrencies have undoubtedly become hugely important in the economy since the initial introduction of Bitcoin to the markets in 2008 [5, 47, 48]. Cryptocurrencies exhibit higher volatility and are more susceptible to bubbles than traditional currencies [49]. In addition, the volatility of Bitcoin returns is subject to long memory, resulting in their being analyzed as financial assets rather than traditional currencies. They are increasingly being included in financial portfolios, which make

Table 2: Estimates of the stochastic volatility in mean model parameters using the data-cloning method.

Parameter	Real value	Data-cloning estimates	SD	Confidence intervals	MCMC estimates	SD (MCMC)
ϕ	0.97	0.9717	0.0053	(0.9055 and 1.0379)	0.9368	0.0810
σ_η	0.12	0.1386	0.0171	(−0.0736 and 0.3509)	0.1878	0.0807
σ^{*2}	0.2	0.1831	0.0671	(−0.6493 and 1.0155)	0.1717	0.0511
b	0.2	0.2548	0.0103	(0.1267 and 0.3829)	0.2527	0.0654
d	0.1	0.1386	0.0171	(−0.1454 and 0.4145)	0.1402	0.1408

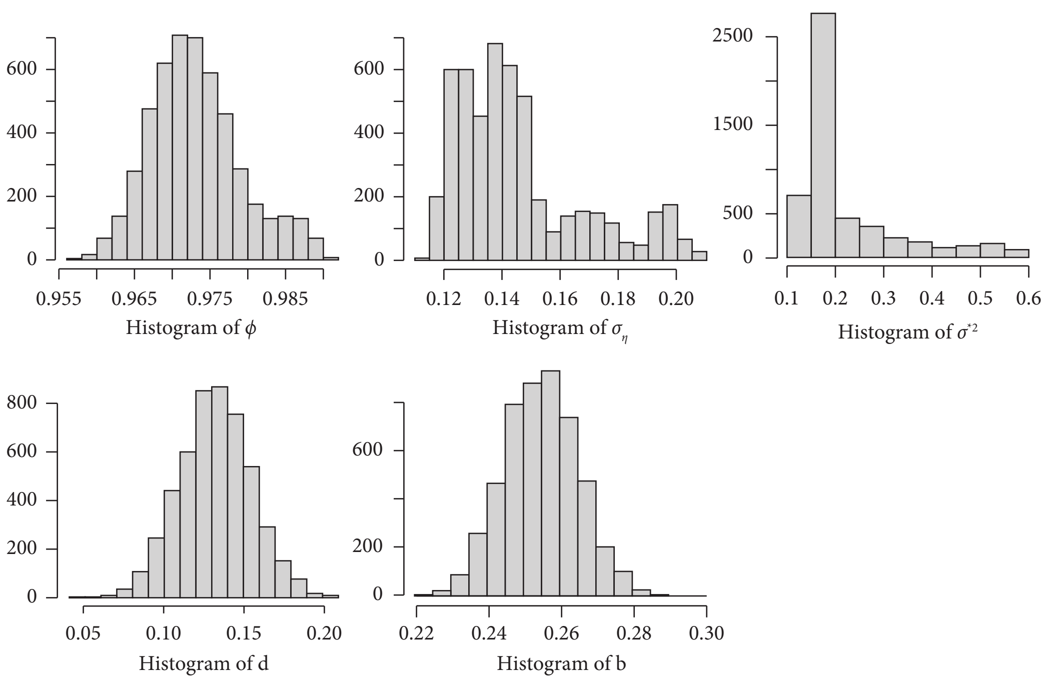

Figure 2: Histograms of the posterior distributions of the stochastic volatility in mean model parameters.

Table 3: Estimates of the stochastic volatility in mean model parameters using the data-cloning method in sample paths of different sizes.

Parameter	Real value	Estimates (SD) $n = 100$	Estimates (SD) $n = 245$	Estimates (SD) $n = 500$	Estimates (SD) $n = 1000$
Φ	0.97	0.9635 (0.0054)	0.9717 (0.0053)	0.9893 (0.0039)	0.9631 (0.0024)
σ_η	0.12	0.1143 (0.0143)	0.1386 (0.0171)	0.1401 (0.0086)	0.1465 (0.0047)
σ^{*2}	0.2	0.4505 (0.0261)	0.1831 (0.0671)	0.5086 (0.1575)	0.2571 (0.0049)
b	0.2	0.1335 (0.0190)	0.2548 (0.0103)	0.1341 (0.0072)	0.2042 (0.0051)
d	0.1	−0.061 (0.0252)	0.1386 (0.0171)	0.0656 (0.0133)	0.1512 (0.0092)

modeling their volatility and its relationship to returns very important in portfolio optimization, hedging, and the valuation of derivative securities. Bitcoin remains the largest cryptocurrency in terms of market capitalization [17, 47], hence it is why we chose it as an example. Tiwari et al. [17] found that in general, SV models consistently outperform the GARCH models when it comes to analyzing cryptocurrencies (particularly in the case of Bitcoin and, to a lesser extent, Litecoin). Moreover, they show that in general using t-distributed innovations greatly improves the results of standard GARCH models, but this result is not significant for SV models. Considering this, in this paper, we use innovations that follow a normal distribution. Nevertheless, the analysis can be easily extended to incorporate Student's t distribution.

The data considered were the daily returns of the cryptocurrency from October 1, 2020, to March 1, 2021. The dataset was obtained from the Spanish financial news website https://es.investing.com/.

5.1. Modeling Bitcoin Returns Using the SV Model Estimated by the Data-Cloning Method. Given that the estimation algorithm excludes the intercept term, we used the deviations from the mean of the data to model the real data within an SV model. Furthermore, the five most recent data values were excluded to be used later to test the predictions. The estimated model parameters, the standard errors of estimate, and credible intervals are shown in Table 4. It also

Table 4: Estimates of the SV model parameters to estimate Bitcoin, using the data-cloning and MCMC methods.

Parameter	Data-cloning estimates	SD DC	HPD 0.95	MCMC estimates	SD MCMC
ϕ	0.4722044	0.2779	(0.0170559 and 0.9464303)	0.4165	0.26151
σ_η	0.1012176	0.06509	(0.0196413 and 0.3175769)	0.4825	0.3382
σ^{*2}	0.0001425	$7.516e-6$	(0.0001297 and 0.0001603)	0.0001489	$4.6174e-5$
b	−0.2081462	0.02999	(−0.2671039 and −0.1502172)	−0.1194	0.1324

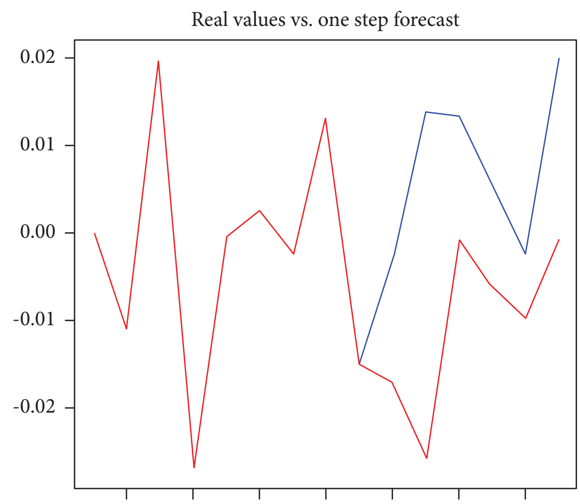

Figure 3: Bitcoin returns vs. estimates of Bitcoin returns generated using the SV model estimated by the data-cloning algorithm.

includes the estimates of the model parameters and corresponding standard errors using the MCMC method in order to compare the two methodologies. Bayesian confidence intervals are included because they can be used to analyze the significance of the parameters from a Bayesian point of view. However, as shown above, if a frequentist approach to the study is desired, confidence intervals can readily be calculated. This is one of the advantages of the data-cloning method.

As expected, the data-cloning and MCMC algorithms provided close values for all parameters except for σ_η. This is probably due to a high standard error in the MCMC method. Note that all parameters except ϕ have lower estimation errors in the estimates obtained through data cloning.

All the parameters are significant at 5%, according to the credible intervals. The parameter b represents the effect of the lagged return in the expected value of the return, and in this case, a negative value was obtained. ϕ is the first-order coefficient of the log-volatility equation (4), while σ_η moderates the effect of disturbance in the log-volatility equation (4). Finally, σ^{*2} is the constant coefficient of variance and represents a small part of the total volatility to which ϕ and σ_η are added.

The value of ϕ is significant, providing evidence of volatility clustering. However, its value is relatively low, suggesting that there is not a substantial persistence of volatility across consecutive periods. At the same time, the value of σ_η is quite high and significant, which means that the volatility of a period is strongly affected by the shocks within that period, increasing the value of the variance. That implies that the course of the volatility is less easily predictable. Finally, b is negative, indicating that the differences in profitability within one period negatively affect the profitability of the following period. Therefore, we can conclude the following:

(i) The negative value of b implies that returns from one period have a negative impact on the returns of the subsequent period

(ii) The variance is generally high, showing little dependence on the variance of the previous period but significant sensitivity to shocks occurring in the current period

These parameters allow equations to be constructed for predicting subsequent values using a one-step prediction method. This involves using the actual values from the previous period to generate predictions for returns. The true value of the required lag (in this case, 1) is used to construct the next values in the series. Similarly, a lag is required for volatility, but since volatility is unobservable, the estimated value is used here.

Figure 3 displays the predicted Bitcoin returns obtained through SV modeling compared to the actual Bitcoin returns. It demonstrates the model's ability to generate accurate one-step predictions for future values in this series.

5.2. Modeling Bitcoin Returns Using the SVM Model Estimated by the Data-Cloning Method. The same dataset was also modeled using the SVM model estimated through the data-cloning algorithm introduced earlier. This model was expected to better incorporate the unobservable behavior of

Table 5: Estimates of the stochastic volatility in mean model parameters to estimate Bitcoin, using the data-cloning method.

Parameter	Data-cloning estimates	SD	HPD 0.95	MCMC estimates	SD MCMC
ϕ	0.4918337	0.3506	(0.011611 and 0.9855730)	0.4290	0.2661
σ_η	0.1158941	0.06447	(0.048043 and 0.2844250)	0.4556	0.3318
σ^{*2}	0.0001422	$7.998e-6$	(0.000131 and 0.0001663)	$1.49e-4$	$5.28e-5$
b	−0.214432	0.02091	(−0.25487 and −0.173575)	−0.1212	0.1314
d	7.1425561	1.558	(3.851365 and 9.7414312)	2.3354	5.0931

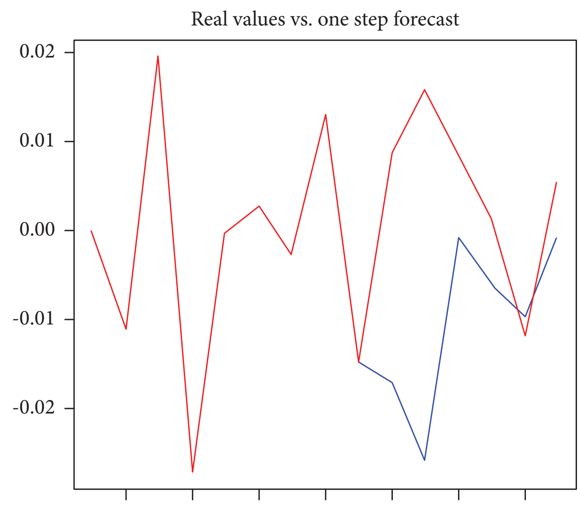

Figure 4: Bitcoin returns vs. estimates of Bitcoin returns generated using the SVM model estimated by the data-cloning algorithm.

volatility by considering its effects on both the returns and their mean simultaneously. Table 5 presents the estimated parameter values, their standard errors of estimate, and the credible intervals. It also includes the parameter values estimated through the MCMC method and their corresponding standard errors.

Both estimation methods yielded similar parameter values, except for σ_η and d, where the MCMC method exhibited higher standard errors, resulting in less agreement with the data-cloning estimate.

All parameters are statistically significant at a 5% significance level, as indicated by the credible intervals. The significance of ϕ once again supports the presence of volatility clustering, although its magnitude is not particularly high. Similar to the SV model, parameter b takes a negative value, indicating a negative impact of lagged returns on current returns.

In the SVM model, a new parameter d is estimated, which represents the effect of volatility on the mean returns. Its significance suggests that the variance has a substantial influence on the expected returns, and the positive value indicates a feedback effect of volatility on returns, aligning with our expectation when analyzing returns in differences.

Figure 4 presents the predicted values of the last observations obtained from the SVM model compared to the actual values. It demonstrates the effectiveness of the one-step prediction method in capturing the future behavior of the series. The close alignment between the predicted values and the actual observations highlights the accuracy of the SVM model in forecasting future values.

6. Final Conclusions

The main goal of this paper is to introduce an estimator for the SVM model parameters based on the data-cloning algorithm, which provides an approximation of the maximum likelihood estimates of the model parameters. The main findings of this study are as follows:

(i) The data-cloning algorithm is a good solution for estimating the parameters of SV and SVM models, whose complexity makes it difficult to use other estimation methods.

(ii) Data cloning is especially useful for estimating the SVM model because it allows the return and the volatility to be estimated at the same time.

(iii) The estimates obtained by the data-cloning method to estimate the parameters of SV and SVM models are shown to be better in terms of their standard errors than those obtained by the conventional MCMC algorithms in the simulation study.

(iv) The SVM data-cloning estimation algorithm demonstrates consistent performance regardless of the sample path size. However, the estimates are observed to be more stable and less path-dependent when we increase its size.

(v) The hybrid nature of the data-cloning method proves to be a very suitable solution when estimating parameters by the maximum likelihood method using Bayesian algorithms.

(vi) SV and SVM models are suitable for modeling financial data with volatility jumps, and they provide a means of understanding the behavior of these series.

(vii) SV and SVM models are empirically shown to be highly capable of providing one-step predictions for cryptocurrencies like Bitcoin. They show that Bitcoin volatility is strongly related to the return in the same period.

Data Availability

This study has employed simulated data from the stochastic volatility and stochastic volatility in mean models, both of which are accessible to interested parties. Furthermore, Bitcoin closing data spanning the period between October 1, 2020, and March 1, 2021, have been utilized. These datasets are publicly available and have been sourced from the website https://es.investing.com/.

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge the open access funding, enabled and organized by CRUE-UNIRIS Gold.

References

[1] B. Mandelbrot, "The variation of certain speculative prices," *Journal of Business*, vol. 36, no. 4, pp. 394–419, 1963.
[2] J.-J. Tseng and S.-P. Li, "Asset returns and volatility clustering in financial time series," *Physica A: Statistical Mechanics and Its Applications*, vol. 390, no. 7, pp. 1300–1314, 2011.
[3] S. Koopman and E. Hol Uspensky, "The stochastic volatility in mean model: empirical evidence from international stock markets," *Journal of Applied Econometrics*, vol. 17, no. 6, pp. 667–689, 2002.
[4] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
[5] P. Katsiampa, "Volatility estimation for Bitcoin: a comparison of GARCH models," *Economics Letters*, vol. 158, pp. 3–6, 2017.
[6] S. Taylor, "Financial returns modelled by the product of two stochastic processes: a study of daily sugar prices 1961-79," in *Time Series Analysis: Theory and Practice*, O. Anderson, Ed., North-Holland, Amsterdam, The Netherlands, 1982.
[7] S. Taylor, "Modelling stochastic volatility: a review and comparative study," *Mathematical Finance*, vol. 4, no. 2, pp. 183–204, 1994.
[8] J. Chan and A. Grant, "Modeling energy price dynamics: GARCH versus stochastic volatility," *Energy Economics*, vol. 54, pp. 182–189, 2016.
[9] M. Manera, M. Nicolini, and I. Vignati, "Modelling futures price volatility in energy markets: is there a role for financial speculation?" *Energy Economics*, vol. 53, pp. 220–229, 2016.
[10] C. Brooks and M. Prokopczuk, "The dynamics of commodity prices," *Quantitative Finance*, vol. 13, no. 4, pp. 527–542, 2013.
[11] A. B. Trolle and E. S. Schwartz, "Unspanned stochastic volatility and the pricing of commodity derivatives," *Review of Financial Studies*, vol. 22, no. 11, pp. 4423–4461, 2009.
[12] M. Asai, M. McAleer, and J. Yu, "Multivariate stochastic volatility: a review," *Econometric Reviews*, vol. 25, no. 2-3, pp. 145–175, 2006.
[13] M. Balcilar and Z. Ozdemir, "The volatility effect on precious metals price returns in a stochastic volatility in mean model with time-varying parameters," *Physica A: Statistical Mechanics and Its Applications*, vol. 534, pp. 122329–122414, 2019.
[14] T. Agbeyegbe, "Modeling JSE stock returns dynamics: GARCH versus stochastic volatility," *The Journal of Developing Areas*, vol. 56, no. 1, pp. 175–191, 2022.
[15] P. Carr, H. Geman, D. Madan, and M. Yor, "Stochastic volatility for Lévy processes," *Mathematical Finance*, vol. 13, no. 3, pp. 345–382, 2003.
[16] S. Kim, N. Shepherd, and S. Chib, "Stochastic volatility: likelihood inference and comparison with ARCH models," *The Review of Economic Studies*, vol. 65, no. 3, pp. 361–393, 1998.
[17] A. Tiwari, S. Kumar, and R. Pathak, "Modelling the dynamics of Bitcoin and Litecoin: GARCH versus stochastic volatility models," *Applied Economics*, vol. 51, no. 37, pp. 4073–4082, 2019.
[18] J. Yu, "Forecasting volatility in the New Zealand stock market," *Applied Financial Economics*, vol. 12, no. 3, pp. 193–202, 2002.
[19] J. Bouchaud, A. Matacz, and M. Potters, "Leverage effect in financial markets: the retarded volatility model," *Physical Review Letters*, vol. 87, no. 22, Article ID 228701, 2001.
[20] A. Melino and S. Turnbull, "Pricing foreign currency options with stochastic volatility," *Journal of Econometrics*, vol. 45, no. 1-2, pp. 239–265, 1990.
[21] E. Renault, "Moment-based estimation of stochastic volatility models," in *Handbook of Financial Time Series*, T. Andersen, R. Davis, J. Kreib, and T. Mikosch, Eds., Springer, New York, NY, USA, pp. 269–311, 2009.
[22] S. Taylor, *Modelling Financial Time Series*, John Wiley and Sons, New York, NY, USA, 1986.
[23] G. Sandmann and S. Koopman, "Estimation of stochastic volatility models via Monte Carlo maximum likelihood," *Journal of Econometrics*, vol. 87, no. 2, pp. 271–301, 1998.
[24] X. Yang, Y. Wu, Z. Zheng, and J.-Q. Hu, "Method of moments estimation for Lévy- driven Ornstein-Uhlenbeck Stochastic Volatility models," *Probability in the Engineering and Informational Sciences*, vol. 35, no. 4, pp. 975–1004, 2021.
[25] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society-Series B: Statistical Methodology*, vol. 72, no. 3, pp. 269–342, 2010.
[26] A. Beskos, K. Kalogeropoulos, and E. Pazos, "Advanced MCMC methods for sampling on diffusion pathspace," *Stochastic Processes and Their Applications*, vol. 123, no. 4, pp. 1415–1453, 2013.
[27] C. Broto and E. Ruiz, "Estimation methods for stochastic volatility models: a survey," *Journal of Economic Surveys*, vol. 18, no. 5, pp. 613–649, 2004.
[28] E. Jacquier, N. Polson, and P. Rossi, "Bayesian analysis of stochastic volatility models," *Journal of Business & Economic Statistics*, vol. 12, no. 4, pp. 371–389, 1994.
[29] G. Kastner, S. Frühwirth-Schnatter, and H. Lopes, "Efficient Bayesian inference for multivariate factor stochastic volatility models," *Journal of Computational & Graphical Statistics*, vol. 26, no. 4, pp. 905–917, 2017.
[30] H. Li, K. Yang, and D. Wang, "A threshold stochastic volatility model with explanatory variables," *Statistica Neerlandica*, vol. 73, no. 1, pp. 118–138, 2019.
[31] N. Shephard, "Fitting non-linear time series models, with applications to stochastic variance models," *Journal of Applied Econometrics*, vol. 8, no. S1, pp. 135–152, 1993.
[32] S. Martino, K. Aas, O. Lindqvist, L. Neef, and H. Rue, "Estimating stochastic volatility models using integrated nested Laplace approximations," *The European Journal of Finance*, vol. 17, no. 7, pp. 487–503, 2011.
[33] H. Rue, S. Martino, and N. Chopin, "Approximate bayesian inference for latent Gaussian models by using integrated nested Laplace approximations," *Journal of the Royal Statistical Society-Series B: Statistical Methodology*, vol. 71, no. 2, pp. 319–392, 2009.
[34] S. Lele, B. Dennis, and F. Lutscher, "Data cloning: easy maximum likelihood estimation for complex ecological models using bayesian Markov chain Monte Carlo methods," *Ecology Letters*, vol. 10, no. 7, pp. 551–563, 2007.

[35] J. Ponciano, M. Taper, B. Dennis, and S. Lele, "Hierarchical models in ecology: confidence intervals, hypothesis testing, and model selection using data cloning," *Ecology*, vol. 90, no. 2, pp. 356–362, 2009.

[36] P. Chaim and M. P. Laurini, "Data cloning estimation and identification of a medium- scale DSGE model," *Stat*, vol. 6, no. 1, pp. 17–29, 2022.

[37] P. Sólymos, "Dclone: data cloning in R," *RELC Journal*, vol. 2, no. 2, pp. 29–37, 2010.

[38] J. M. Marín, M. T. Rodríguez-Bernal, and E. Romero, "Data cloning estimation of GARCH and COGARCH models," *Journal of Statistical Computation and Simulation*, vol. 85, no. 9, pp. 1818–1831, 2015.

[39] P. de Zea Bermudez, J. M. Marín, and H. Veiga, "Data cloning estimation for asymmetric stochastic volatility models," *Econometric Reviews*, vol. 39, no. 10, pp. 1057–1074, 2020.

[40] P. Pasricha and X.-J. He, "Exchange options with stochastic liquidity risk," *Expert Systems with Applications*, vol. 223, Article ID 119915, 2023.

[41] S. Lin and X.-J. He, "Analytically pricing variance and volatility swaps with stochastic volatility, stochastic equilibrium level and regime switching," *Expert Systems with Applications*, vol. 217, Article ID 119592, 2023.

[42] K. R. French, G. Schwert, and R. F. Stambaugh, "Expected stock returns and volatility," *Journal of Financial Economics*, vol. 19, no. 1, pp. 3–29, 1987.

[43] S. Lele, K. Nadeem, and B. Schmuland, "Estimability and likelihood inference for generalized linear mixed models using data cloning," *Journal of the American Statistical Association*, vol. 105, no. 492, pp. 1617–1625, 2010.

[44] M. Laurini, "A hybrid data cloning maximum likelihood estimator for stochastic volatility models," *Journal of Time Series Econometrics*, vol. 5, no. 2, pp. 193–229, 2013.

[45] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012.

[46] S. Brooks and A. Gelman, "General methods for monitoring convergence of iterative simulations," *Journal of Computational & Graphical Statistics*, vol. 7, no. 4, pp. 434–455, 1998.

[47] H. Akkus and I. Çeli̇k, "Kri̇ptopara pi̇yasa volati̇li̇tesi̇ni̇n modellenmesi̇, tahmi̇ni̇ ve Bİ̇TCOİ̇N'İ̇N ri̇ske maruz değer di̇nami̇kleri̇," *Muhasebe Bilim Dünyası Dergisi*, vol. 22, no. 2, pp. 296–312, 2020.

[48] A. Urquhart, "The inefficiency of Bitcoin," *Economics Letters*, vol. 148, pp. 80–82, 2016.

[49] E. Cheah and J. Fry, "Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin," *Economics Letters*, vol. 130, pp. 32–36, 2015.