Mathematics and Computers in Simulation 225 (2024) 430–445

Contents lists available at ScienceDirect

Mathematics and Computers in Simulation

journal homepage: www.elsevier.com/locate/matcom

Original articles

Efficient likelihood estimation of Heston model for novel climate-related financial contracts valuation

Augusto Blanc-Blocquel [a,b], Luis Ortiz-Gracia [c,d,*], Rodolfo Oviedo [e]

[a] Departamento de Finanzas, Facultad de Ciencias Empresariales, Universidad Austral, Paraguay 1950, 2000 Rosario, Santa Fe, Argentina
[b] Department of Statistics and Operations Research, Universitat Politècnica de Catalunya, Barcelona, Spain
[c] Departament d'Econometria, Estadística i Economia Aplicada, Universitat de Barcelona (UB), Av. Diagonal, 690, 08034 Barcelona, Spain
[d] RISKcenter, Institut de Recerca en Economia Aplicada (IREA), Universitat de Barcelona (UB), Av. Diagonal, 690, 08034 Barcelona, Spain
[e] Independent Financial Advisor, Rosario, Argentina

ARTICLE INFO

Keywords:
Climate management
Derivatives
Heston model
Bitcoin
Cryptocurrencies
Likelihood evaluation
Shannon wavelets

ABSTRACT

We propose novel Bitcoin-denominated derivatives contracts on carbon bonds. We consider a futures contract on carbon bonds where its price is expressed in terms of bitcoins. Then, we put forward options on a futures contract of the former type. Governments can use such contracts to hedge climate change and influence the prices of carbon bonds and cryptocurrencies. We show how these derivatives transfer volatility to the bitcoin market without a negative effect in the carbon bonds market. Since the aforementioned options are not yet traded in the market, we price them by assuming that the underlying is driven by the well-known Heston model, where the model parameters are estimated by a novel method based on Shannon wavelets. Heston model belongs to the class of stochastic volatility (SV) models. The discrete observations from the SV model can be seen as a state-space model, that is, a stochastic model in discrete-time which contains two sets of equations, the state equation and the observation equation. While the first describes the transition of a latent process in time, the second shows how an observer measures the latent process at each time period. We infer the properties of the latent variable by means of a filtering algorithm, and we estimate the parameters of the model via maximum likelihood. The evaluation of the likelihood function is a time-consuming task that involves updating and prediction steps of the state variable, leading to the computation of complicated integrals. We calculate these integrals by means of an integration method based on Shannon wavelets, and compare the root mean square error (RMSE) of the estimation with state-of-the-art methods. The results show that the RSME is dramatically reduced in a short CPU time with the use of wavelets.

1. Introduction

Climate change has been one of the main concerns for policymakers for the last 20 years. One of the main effects of climate change, as stated in [1], is its impact on agriculture and food production. In addition, climate change has clear consequences for human health. As detailed in [2], the direct and indirect health consequences include excessive heat-related illnesses, vector and waterborne diseases, increased exposure to environmental toxins, exacerbation of cardiovascular and respiratory diseases due to

* Corresponding author at: Departament d'Econometria, Estadística i Economia Aplicada, Universitat de Barcelona (UB), Av. Diagonal, 690, 08034 Barcelona, Spain.

E-mail address: luis.ortiz-gracia@ub.edu (L. Ortiz-Gracia).

https://doi.org/10.1016/j.matcom.2024.05.024
Received 30 November 2023; Received in revised form 22 May 2024; Accepted 24 May 2024
Available online 1 June 2024

declining air quality, and mental health stress, among others. Another consequence of climate change, as explained in [3], is the impact on dryland productivity and dynamics.

In this work, we develop novel financial products to mitigate climate risk. To be more precise, we consider futures contracts on one unit of carbon bond (CB) with price given in bitcoins (BTC), as well as options on such futures. An important task associated to the launch of new options contracts is the modeling of the dynamics of the underlying asset and the pricing of the derivative. We consider the well-known Heston model (see [4]) for the underlying due to its mean reverting and stochastic volatility features. Since these contracts are not yet quoted in the market, we must estimate the parameters of the model directly from the underlying asset data, instead of performing the classical calibration based on minimizing the sum of the squared differences between the quoted option prices and the model prices.

We propose a maximum likelihood approach to estimate the parameters of the Heston model. Generally speaking, the SV model can be seen as a state-space model, that is, a stochastic model in discrete-time which contains two sets of equations, the state equation and the observation equation. While the first describes the transition of a latent process in time, the second shows how an observer measures the latent process at each time period. We infer the properties of the latent variable by means of a filtering algorithm. The evaluation of the likelihood function is a very time-consuming task that involves updating and prediction steps of the state variable, leading to the computation of complicated integrals, which poses a serious computational challenge. There are several attempts in the literature for estimating the parameters of SV models, like for instance [5–10] who work on the autoregressive model of order one (among others). The Heston model is considered for instance in [11–13] within the framework of likelihood estimation, as well as in [14] where the authors use option price data. Importance sampling techniques are used in [15] for likelihood evaluation. Our contributions can be summarized as follows.

- We design novel financial contracts to help governmental authorities to manage climate risk, being one of the main utilities of these novel contracts the increase in the volatility of BTC.
- We carry out an econometric analysis with real data to show that the volatility transfer from the derivatives to the BTC spot price does not have a negative side effect in the form of volatility spillover from BTC to CB.
- We use a filtering algorithm to estimate the parameters of SV models via computation of the log-likelihood function, where the main challenge here is the efficient calculation of integrals on the entire real line. We propose a novel method for computing them by means of Shannon wavelets.
- We test our methodology for estimating the three parameters of an autoregressive SV model of order one. The parameters of our numerical method can be fixed a priori based on the error analysis provided.
- We perform numerical experiments and compare the results with the literature. We show that the RMSE is dramatically reduced in a short computation time.
- We apply our method to the estimation of parameters of Heston model and we perform an error analysis. We present numerical experiments and compare the results with the extended Kalman filter (EKF) method, which is well-known and used in the literature. We show that the wavelet method can estimate particularly well the correlation parameter.
- We estimate the parameters of Heston model with real data, and compare again with EKF estimation. A Kolmogorov–Smirnov test shows the ability of Shannon wavelets for the estimation.

This work is organized as follows. Section 2 is devoted to the design of novel futures and option contracts along with the volatility spillover analysis. We develop in Section 3 the methodology to estimate the log-likelihood function by means of Shannon wavelets. We estimate the parameters of two SV models in Section 4, accompanied with an error analysis. A wide variety of numerical experiments either with simulated or with real data is presented in Section 5. Finally, Section 6 concludes.

2. Definition of climate-related derivatives and their rationale

As detailed in [16], city governments are increasingly developing policies and programs designed to reduce greenhouse gas (GHG) emissions and adapt to the consequences of climate change. As stated in [17], energy production accounts for almost three-quarters of GHG emissions. A large part is devoted to producing computational power for different industries. Over the last decade, the blockchain sector has become a prominent consumer of computational power and energy. This has given rise to concerns about the sustainability of this technology and the hazards it could impose on society in terms of climate change and global warming. With this concern in mind, we design novel financial instruments aiming to help the regulatory authorities impact the climate change by causing a volatility transfer from derivatives to bitcoin's spot markets. A high volatility in these markets would disincentive investments and potentially diminish the amount of energy devoted to the mining of these coins. In what follows, we describe the structure and potential utility of the novel contracts for regulators and financial authorities.

We define a novel futures contract with one unit of CB as underlying and its price expressed in terms of BTC. The mechanics of the contract is as follows. Suppose that at inception, a futures contract without delivery is traded at 0.40 BTC/CB. At maturity, if the spot BTC price is 200,000 dollars and the spot CB price is 50,000 dollars then, these prices imply that the last settlement price is 0.25 BTC/CB. Sovereign authorities may have short positions in BTC (since they prefer its price to fall) and long positions in CB (since they prefer its price to increase so that pollution is more expensive for industries). In contrast, miners have a long position in BTC as they receive bitcoins for their mining activities, and they may also have a short position in CB if they need to pay for pollution. Sovereign authorities could buy the future on CB, which would increase the price of CB and decrease the price of BTC, since CB has its price in BTC.

Furthermore, we propose the listing of European calls and puts on futures contracts. The underlying of each option is a futures contract of the type described in the former paragraph. A call/put holder exercising an option gets a long/short futures contract at maturity. The writer will receive the opposite futures position. The utility of the option for sovereign authorities is similar to the utility of its underlying futures contract. By buying a call, the governmental authority gets the right to have a long position on the underlying futures contract for climate management, while limiting the maximum loss to the premium.

By influencing CB and BTC prices, these derivatives will transfer volatility to BTC markets, which is consistent with the objectives of regulatory authorities. However, while the volatility transfer from derivatives to the spot markets is in line with the objectives of climate risk management, it is an unwanted effect the transfer of volatility from BTC markets to CB markets. We show in Section 2.1 that there is no relation between the volatility of CB and the volatility of BTC. Therefore, such an undesirable scenario is unlikely to occur. For this analysis we perform a vector autoregression with the two volatility series (BTC and CB). The volatility modeling is carried out with GARCH models for which a previous and necessary step is the modeling of returns, which is performed with ARMA models. This type of models are routinely employed for time series volatility modeling.

After the econometric analysis of Section 2.1 about the consistency of these novel derivatives on the underlying CB with price in BTC, we tackle the problem of their pricing. A popular model for driving the dynamics of the underlying is the well-known Heston stochastic volatility model put forward in Section 4. In this case, we recall that the underlying is the CB (with price in BTC). Since there are no market prices for these novel derivatives on the underlying considered, we cannot calibrate the parameters of the Heston model to those prices as usual. Instead, we propose a calibration method in Section 3 based on maximum likelihood estimation where the likelihood function is evaluated via a filtering algorithm that can be potentially used to estimate any other stochastic volatility model. Our contribution to the literature is the use of Shannon wavelets to efficiently compute the integrals appearing in the filtering steps. This new methodology is developed in Section 3.3 and tested with a well-known AR-(1) model in Section 4 for which there are many references in the literature to compare with.

It is worth remarking that this is not a fair price for the options on CB with price in BTC, since there is no market trading that underlying and therefore the classical risk-neutral pricing theory does not apply. The computation of the price by means of the expected discounted payoff can be used to propose a price for this derivative within the framework of climate-related risk assessment. A similar argument is underlined in [18] in the context of pricing derivatives on temperature.

2.1. Volatility spillover analysis in the spot markets

One of the intended consequences of both the futures and the options contracts is the increase in the volatility of BTC. Now we want to check the absence of negative side effects. One potential side effect would be an increase in the volatility of CB. A disturbance of the CB market would be an unintended consequence. In what follows we check whether there is further volatility transfer between BTC and the CB spot prices. To check this, we obtain historical prices of BTC and CB (both expressed in dollars), and we consider their logarithmic returns. The BTC prices are obtained from `www.coingecko.com`, and the CB prices correspond to EUA daily futures from ENDEX exchange, and they are obtained from www.barchart.com The period ranges from 16/12/2019 to 17/12/2021. Since this period of data contains the COVID-19 outbreak, the lockdown measures could have potentially affected the carbon bond market due to a reduction in the consumption during the year 2020, as stated in [19–21]. Therefore, the period considered in this work includes the data from the lockdown measures in 2020 plus the additional year 2021, which is not affected by an abnormal use of carbon. It would be advisable to assess again the volatility spillover when new data becomes available to see whether these conclusions are still valid.

We choose a model for the returns, denoted by y_t. We restrict ourselves to an ARMA(p,q) process of the form,

$$y_t = \delta + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t, \tag{1}$$

with p autoregressive terms and q moving-average terms. As for the volatility process of each series, we carry out the estimation by fitting a GARCH model. In the case of the BTC volatility, we use an exponential GARCH $(1,1)$ model of the form,

$$\ln(\sigma_t^2) = \omega + \sum_{j=1}^{p} \left(\alpha_j z_{t-j} + \gamma_j \left(|z_{t-j}| - \mathbb{E}\left(|z_{t-j}|\right)\right)\right) + \sum_{j=1}^{q} \beta_j \ln\left(\sigma_{t-j}^2\right), \tag{2}$$

where σ_t is the conditional volatility at time t, the coefficient α_j captures the sign effect, γ_j the size effect and β_j the persistence of volatility. A standard normal innovation is given by z_j. For the CB volatility, we use a standard GARCH $(1,1)$ model given by,

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \tag{3}$$

For the BTC returns, we found the best fit with an ARMA$(2,2)$, while for the CB returns, we chose an ARMA$(2,0)$. We display in Table 1 the values of the parameters estimated for the BTC series, while parameters corresponding to CB series are shown in Table 2. Note that α_1 of BTC volatility estimate is negative, allowing asymmetry in the variance. We believe that this is congruent with the nature of the asset.

We take the daily volatilities time series produced by the GARCH models of BTC and CB and set a vector autoregression (VAR) scheme to analyze whether there is a statistically significant relation between the two. Volatility spillovers between assets are

Table 1
BTC parameters corresponding to the returns of expression (1) and volatility of expression (2).

ARMA	Estimate	p-value
δ	0.00307	0.154
ϕ_1	−0.16459	< 0.001
ϕ_2	−0.93476	< 0.001
θ_1	0.08388	< 0.001
θ_2	0.94757	< 0.001
GARCH	Estimate	p-value
ω	−0.57865	0.016
α_1	−0.06490	0.033
β_1	0.90285	< 0.001
γ_1	0.19474	< 0.001

Table 2
CB parameters corresponding to the returns of expression (1) and volatility of expression (3).

ARMA	Estimate	p-value
δ	0.00275	0.008
ϕ_1	−0.14449	0.003
ϕ_2	0.06744	0.151
GARCH	Estimate	p-value
ω	0.00009	0.008
α_1	0.14617	< 0.001
β_1	0.75515	< 0.001

Table 3
VAR of BTC volatility expression.

Coefficient	Estimate	p-value
a^1_{11}	0.80278	< 0.001
a^1_{12}	−0.04304	0.247
a^2_{11}	0.07311	0.200
a^2_{12}	0.06423	0.190
a^3_{11}	0.05093	0.258
a^3_{12}	−0.07047	0.056
c_1	−0.40399	< 0.001

commonly studied by means of VAR, like for instance in [22]. A VAR(p) in two variables, where p is the maximum lag, can be written in matrix form as,

$$\begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} a^1_{11} & a^1_{12} \\ a^1_{21} & a^1_{22} \end{bmatrix} \begin{bmatrix} v_{1,t-1} \\ v_{2,t-1} \end{bmatrix} + \cdots + \begin{bmatrix} a^p_{11} & a^p_{12} \\ a^p_{21} & a^p_{22} \end{bmatrix} \begin{bmatrix} v_{1,t-p} \\ v_{2,t-p} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}, \tag{4}$$

where $v_{1,t}$ is the volatility of BTC at t and $v_{2,t}$ is the volatility of CB at t. In our analysis we select the optimal number of lags by means of four information criteria: the Akaike information criterion, the Schwartz criterion, the Hannan and Quinn criterion, and the final prediction criterion. We select the mode of these 4 values, which in our analysis is $p = 3$. The estimation results for the parameters of the BTC and CB volatility equations are given in Tables 3 and 4, respectively. The covariance and correlation matrices of residuals are given in Tables 5 and 6, respectively. Overall, results shown in Table 3, 4, 5 and 6 do not support a significant transfer of volatility between BTC and CB, since the coefficients $a^1_{1,2}, a^2_{1,2}, a^1_{2,1}, a^3_{2,1}$ are all small, have high p-values and the residuals show very low covariance and correlation, meaning that there is no pattern left to explain in the VAR model. These econometric results suggest that the volatility transfer from the derivatives to the BTC spot price does not have a negative side effect in the form of volatility spillover from BTC to CB.

3. Maximum likelihood estimation of stochastic volatility models

In this section, we consider stochastic volatility models, in which the price process $\{y_t\}$ is governed by a stochastic differential equation depending on the log-volatility process $\{h_t\}$, which, in turn, is considered stochastic. We present the filtering problem and likelihood evaluation in Section 3.1. The methodology to estimate the likelihood by means of Shannon wavelets is detailed in Section 3.3. For sake of completeness, a brief description of multi-resolution analysis and Shannon wavelets is put forward in Section 3.2.

Table 4
VAR of CB volatility expression.

Coefficient	Estimate	p-value
a^1_{21}	−0.05814	0.272
a^1_{22}	0.86375	< 0.001
a^2_{21}	0.19117	0.005
a^2_{22}	0.09143	0.117
a^3_{21}	−0.05521	0.304
a^3_{22}	−0.07591	0.085
c_2	−0.19429	0.040

Table 5
Covariance matrix of residuals.

	BTC	CB
BTC	0.00628	0.00050
CB	0.00050	0.00895

Table 6
Correlation matrix of residuals.

	BTC	CB
BTC	1	0.06693
CB	0.06693	1

3.1. Filtering and evaluation of the likelihood

In order to estimate the likelihood, we carry out the so-called filtering algorithm, summarized as follows. Given $t = 1, \dots, T$, the likelihood contribution at time t of return y_t, conditional on the past returns Y_{t-1} reads,

$$f(y_t|Y_{t-1}) = \int_{\mathbb{R}} f(y_t, h_t|Y_{t-1}) dh_t = \int_{\mathbb{R}} f_1(y_t|h_t) f_3(h_t|Y_{t-1}) dh_t, \tag{5}$$

where $Y_t = (y_t, y_{t-1}, \dots, y_1)$, Y_0 is empty, $\mathbb{R}$ is the domain of the latent state variable h_t, and $f_1(y_t|h_t)$ is the density of the return y_t given h_t. The posterior distribution of the state variables at time t, conditional on Y_t, can be obtained as follows,

$$f(h_t|Y_t) = f(h_t|y_t, Y_{t-1}) = \frac{f(y_t, h_t|Y_{t-1})}{f(y_t|Y_{t-1})} = \frac{f_1(y_t|h_t) f_3(h_t|Y_{t-1})}{f(y_t|Y_{t-1})}, \tag{6}$$

and finally, the one-step ahead prediction of h_t, conditional on the past returns, is given by,

$$f_3(h_{t+1}|Y_t) = \int_{\mathbb{R}} f(h_{t+1}, h_t|Y_t) dh_t = \int_{\mathbb{R}} f_2(h_{t+1}|h_t) f(h_t|Y_t) dh_t, \tag{7}$$

where $f_2(h_{t+1}|h_t)$ is the transition probability distribution of h_{t+1} given h_t and the past returns. If we have $f(y_t|Y_{t-1})$, $t = 1, \dots, T$, we can calculate the log-likelihood,

$$L(\boldsymbol{\Theta}|Y_T) = \sum_{t=1}^{T} \log\ f(y_t|Y_{t-1}), \tag{8}$$

where $\boldsymbol{\Theta}$ stands for the set of model parameters.

3.2. Multi-resolution analysis and Shannon wavelets

Consider the space,

$$L^2(\mathbb{R}) = \{ f \ : \int_{-\infty}^{+\infty} |f(x)|^2 \, dx < \infty \},$$

of square integrable functions. A general structure for wavelets in $L^2(\mathbb{R})$ is called a *multi-resolution analysis*. We start with a family of closed nested subspaces,

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots,$$

in $L^2(\mathbb{R})$ where,

$$\bigcap_{j\in\mathbb{Z}} V_j = \{0\}, \qquad \overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R}),$$

and,

$$f(x) \in V_j \iff f(2x) \in V_{j+1}.$$

If these conditions are met, then a function $\phi \in V_0$ exists such that $\{\phi_{j,k}\}_{k\in\mathbb{Z}}$ forms an orthonormal basis of V_j, where,

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k).$$

In other words, the function ϕ, called the *scaling function* or *father wavelet*, generates an orthonormal basis for each V_j subspace.

For any $f \in L^2(\mathbb{R})$, a projection map of $L^2(\mathbb{R})$ onto V_m, $\mathcal{P}_m : L^2(\mathbb{R}) \to V_m$, is defined by means of,

$$\mathcal{P}_m f(x) = \sum_{k\in\mathbb{Z}} c_{m,k}\phi_{m,k}(x), \tag{9}$$

where $c_{m,k} = \int_{-\infty}^{+\infty} f(x)\phi_{m,k}(x)\,dx$ are the *scaling coefficients* (see [23] for general theory on multi-resolution analysis and wavelets).

We here use Shannon wavelets (see [24]). The *sinc* function or Shannon scaling function is the starting point for the definition of the Shannon wavelet family. A set of Shannon scaling functions or father wavelets in the subspace V_m is defined as,

$$\phi_{m,k}(x) = 2^{m/2}\frac{\sin(\pi(2^m x - k))}{\pi(2^m x - k)}, \quad k \in \mathbb{Z}. \tag{10}$$

It is clear that for $m = k = 0$, we have the basic scaling function or father wavelet,

$$\phi(x) = \mathrm{sinc}(x),$$

where,

$$\mathrm{sinc}(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x}, & \text{if } x \neq 0, \\ 1, & \text{if } x = 0. \end{cases} \tag{11}$$

The following theorem will be used in Section 3.3 to compute some integrals in the entire real line in a very efficient way. A generalization of this theorem can be found in [25].

Theorem 1 (*Theorem 1.3.2 of [26]*). *Let f be defined on $\mathbb{R}$, and let its Fourier transform, denoted by $\hat{f}$, be such that, for some positive constant d,*

$$|\hat{f}(\xi)| = \mathcal{O}\left(e^{-d|\xi|}\right), \quad \xi \to \pm\infty . \tag{12}$$

Then, as $a \to 0$,

$$\frac{1}{a}\int_{\mathbb{R}} f(y)\mathcal{S}(k,a)(y)dy - f(ka) = \mathcal{O}\left(e^{-\frac{\pi d}{a}}\right),$$

where $\mathcal{S}(k,a)(y) := \mathrm{sinc}\left(\frac{y}{a} - k\right)$.

3.3. Likelihood estimation with Shannon wavelets

We start by computing,

$$f(y_1|Y_0) = \int_{\mathbb{R}} f_1(y_1|h_1)f_3(h_1|Y_0)dh_1, \tag{13}$$

and approximate $f_3(h_1|Y_0)$ by means of a finite combination of Shannon wavelets,

$$f_3(h_1|Y_0) \approx \sum_{k=k_1}^{k_2} c^1_{m,k}\phi_{m,k}(h_1), \tag{14}$$

with,

$$c^1_{m,k} = \int_{\mathbb{R}} f_3(h_1|Y_0)\phi_{m,k}(h_1)dh_1 \approx \frac{1}{2^{m/2}} f_3\left(\frac{k}{2^m}|Y_0\right), \quad k = k_1, \ldots, k_2, \tag{15}$$

where the approximation in expression (15) is justified by the application of Theorem 1. Finally,

$$f(y_1|Y_0) \approx \sum_{k=k_1}^{k_2} c^1_{m,k}\int_{\mathbb{R}} f_1(y_1|h_1)\phi_{m,k}(h_1)dh_1 \approx \frac{1}{2^{m/2}}\sum_{k=k_1}^{k_2} c^1_{m,k} f_1\left(y_1|\frac{k}{2^m}\right), \tag{16}$$

where the last approximation in expression (16) comes again from the application of Theorem 1.

Now, we compute,

$$f(y_2|Y_1) = \int_{\mathbb{R}} f_1(y_2|h_2)f_3(h_2|Y_1)dh_2, \tag{17}$$

and approximate $f_3(h_2|Y_1)$ by means of a finite combination of Shannon wavelets,

$$f_3(h_2|Y_1) \approx \sum_{k=k_1}^{k_2} c^2_{m,k}\phi_{m,k}(h_2), \tag{18}$$

with,

$$c^2_{m,k} = \int_{\mathbb{R}} f_3(h_2|Y_1)\phi_{m,k}(h_2)dh_2, \quad k = k_1,\dots,k_2. \tag{19}$$

If we take into account expression (7), we have,

$$c^2_{m,k} = \int_{\mathbb{R}} f_3(h_2|Y_1)\phi_{m,k}(h_2)dh_2 = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f_2(h_2|h_1)f(h_1|Y_1)dh_1\right)\phi_{m,k}(h_2)dh_2. \tag{20}$$

We apply Fubini's theorem to interchange the order of integration, and apply Theorem 1 to the inner integral,

$$c^2_{m,k} = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f_2(h_2|h_1)\phi_{m,k}(h_2)dh_2\right) f(h_1|Y_1)dh_1 \approx \frac{1}{2^{m/2}}\int_{\mathbb{R}} f_2\left(\frac{k}{2^m}|h_1\right) f(h_1|Y_1)dh_1. \tag{21}$$

Finally, by means of expression (6), expression (14), and Theorem 1, we find that,

$$c^2_{m,k} \approx \frac{1}{2^{m/2}}\frac{1}{f(y_1|Y_0)}\int_{\mathbb{R}} f_2\left(\frac{k}{2^m}|h_1\right) f_1(y_1|h_1)f_3(h_1|Y_0)dh_1 \approx \frac{1}{2^m}\frac{1}{f(y_1|Y_0)}\sum_{j=j_1}^{j_2} c^1_{m,j} f_1\left(y_1|\frac{j}{2^m}\right) f_2\left(\frac{k}{2^m}|\frac{j}{2^m}\right), \tag{22}$$

and,

$$f(y_2|Y_1) \approx \frac{1}{2^{m/2}}\sum_{k=k_1}^{k_2} c^2_{m,k} f_1\left(y_2|\frac{k}{2^m}\right). \tag{23}$$

We iterate this process following the steps summarized in Algorithm 1. It is worth remarking that, although different m, k_1, k_2 can be estimated at each step of Algorithm 1, that is, for each contribution of the likelihood function, we consider them constant after the first choice.

Algorithm 1: Likelihood estimation

Data: $Y_T = (y_T, y_{T-1},\dots,y_1)$
Determine the scale of approximation m and the truncation range k_1, k_2
Compute $f(y_1|Y_0)$ by means of expression (15) and expression (16)
Initialize $L(\Theta|Y_T) = \log f(y_1|Y_0)$
for $t = 2,\dots,T$ **do**
  Compute $\log f(y_t|Y_{t-1})$
  Update $L(\Theta|Y_T)$ with $\log f(y_t|Y_{t-1})$
return $L(\Theta|Y_T)$

4. Stochastic volatility models

In this section, we estimate the parameters of two well-known stochastic volatility models, the standard AR(1)-SV model of Section 4.1 and the Heston model of Section 4.2. The estimation of parameters is a complex procedure that involves two main tasks: the computation of the likelihood function and the maximization of the likelihood function. The optimization part is an iterative process that starts with an initial seed of parameters and evaluates the likelihood function at each step. The selection of the initial values of these parameters might determine whether the optimal set of parameters is reached. On top of that, we deal with models for which the existence of a global optimal value for their parameters is yet an open problem. Furthermore, since we used an iterative optimization method that needs to evaluate many times the likelihood function, and we also typically deal with a large data vector, the likelihood computation is a crucial step that must be carried out very efficiently. For these reasons, we focus on the likelihood estimation and select as initial seed for the optimization step the values of parameters used to simulate the data (or, in some cases, small perturbations of these parameters) as it is a common practice in the references cited in this work.

4.1. The standard AR(1)-SV model

We consider the popular stochastic volatility model studied, for instance, in [6–10]. We carry out an error analysis in Section 4.1.1, and a wide range of numerical experiments[1] in Section 5.1. The model is formulated as follows,

$$\begin{aligned} y_t &= \exp(h_t/2)\epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,1), \\ h_{t+1} &= \alpha + \beta h_t + \eta_{t+1}, \quad \eta_{t+1} \sim \mathcal{N}(0,\sigma^2_\eta), \end{aligned} \tag{24}$$

[1] Computations were performed in R code on a personal computer with a 3.40 GHz Intel Core i7-6700 processor and 32.0 GB of RAM. Function `optim` with Nelder–Mead method of the R package `nloptr` is used for optimization.

where $\{\epsilon_t\}, \{\eta_{t+1}\}$ are independent processes, $\sigma_t^2 = e^{h_t}$ is the volatility of y_t, and the distribution of $h_1 \sim \mathcal{N}\left(\frac{\alpha}{1-\beta}, \frac{\sigma_\eta^2}{1-\beta^2}\right)$, typically known as an initial condition, with probability density function,[2]

$$f_3(h_1|Y_0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(h_1-\mu)^2}{2\sigma^2}},$$

where $\mu = \frac{\alpha}{1-\beta}, \sigma^2 = \frac{\sigma_\eta^2}{1-\beta^2}$, denotes the unconditional distribution of the process $\{h_t\}$ for $t = 1$. The density function of y_t given h_t and the density function of h_t given h_{t-1} are, respectively,

$$f_1(y_t|h_t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y_t^2}{2} e^{-h_t} - \frac{h_t}{2}}, \quad f_2(h_t|h_{t-1}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(h_t-\alpha-\beta h_{t-1}\right)^2}{2\sigma_\eta^2}}.$$

In this case, the set of parameters to be estimated is $\boldsymbol{\Theta} = (\alpha, \beta, \sigma_\eta)$.

4.1.1. Error analysis

We compute the corresponding Fourier transforms[3] to assess whether the hypothesis of Theorem 1 is satisfied. Let $\hat{f}_1(y_t|\xi)$ be the Fourier transform of $f_1(y_t|h_t)$,

$$\hat{f}_1(y_t|\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\xi h_t} e^{-\frac{y_t^2}{2} e^{-h_t} - \frac{h_t}{2}} dh_t, \tag{25}$$

and let us approximate $f_1(y_t|h_t)$ by $f_1^a(y_t|h_t)$,

$$f_1(y_t|h_t) \approx f_1^a(y_t|h_t) := \frac{1}{\sqrt{2\pi}} e^{-\frac{y_t^2}{2}\left(\frac{h_t^2}{2} - h_t + 1\right) - \frac{h_t}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{y_t^2}{4} h_t^2 + \frac{1}{2}(y_t^2-1)h_t - \frac{y_t^2}{2}}, \tag{26}$$

where we have replaced the term e^{-h_t} of expression (25) by its second order Taylor expansion $\frac{h_t^2}{2} - h_t + 1$ in order to facilitate the Fourier transform computation. Then, the Fourier transform $\hat{f}_1(y_t|\xi)$ of $f_1(y_t|h_t)$ is approximated by the Fourier transform $\hat{f}_1^a(y_t|\xi)$ of $f_1^a(y_t|h_t)$,

$$\hat{f}_1^a(y_t|\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\xi h_t} e^{-\frac{y_t^2}{4} h_t^2 + \frac{1}{2}(y_t^2-1)h_t - \frac{y_t^2}{2}} dh_t = \frac{1}{y_t} e^{-\frac{y_t^2}{2} + \left(\frac{\frac{1}{2}(y_t^2-1) - i\xi}{y_t}\right)^2}. \tag{27}$$

We observe that $|\hat{f}_1^a(y_t|\xi)| = \mathcal{O}\left(e^{-\frac{1}{y_t^2}\xi^2}\right)$, from which the hypothesis of Theorem 1 is satisfied for f_1^a. This gives us the intuition that Theorem 1 might work out well for f_1 although this fact has not been proved theoretically, since we have employed the approximation f_1^a of f_1 to ease the computation of the Fourier transform. Let us perform some numerical experiments to show that the integral in expression (16) is computed very efficiently when it is calculated following Theorem 1.

We define the integral,

$$I = \int_{\mathbb{R}} G_{m,k}(x) dx, \tag{28}$$

where,

$$G_{m,k}(x) := g(x)\phi_{m,k}(x), \quad g(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}e^{-x} - \frac{x}{2}}. \tag{29}$$

In our experiments, we approximate the integral of expression (28) for values of scale $m = 0, 2$ and translation parameters $k = 0, 1$, respectively. We plot $G_{m,k}$ in Fig. 1 for $m = k = 0$ as well as in Fig. 2 for $m = 2, k = 1$. We observe that $G_{2,1}$ is a highly oscillatory function and it is more peaked than $G_{0,0}$. While increasing the scale m represents a challenge for numerical quadrature methods, we will show that integral I is very efficiently computed following Theorem 1.

We use the R function called `integrate` to solve the integral of expression (28). If one or both limits are infinite (as it is our case) the infinite range is mapped onto a finite interval. For a finite interval, globally adaptive interval subdivision is used in connection with extrapolation by Wynn's Epsilon algorithm, with the basic step being Gauss–Kronrod quadrature (see [27] for details). We refer to this method with the acronym GK.

Since the integral I cannot be calculated exactly, we also perform Monte Carlo simulation to have an estimation for comparison, since the order of approximation error in that case is known to be $\mathcal{O}(1/\sqrt{n})$, whe n is the sample size. If we truncate the entire real line and consider the integration interval $[a, b]$, then,

$$I \approx \int_a^b G_{m,k}(x)dx = (b-a)\int_0^1 G_{m,k}((b-a)y + a)dy, \tag{30}$$

[2] Note that we use the same letter for the random variable of the stochastic volatility model and for the variable of the density function.

[3] The Fourier transforms have been calculated with WolframAlpha software.

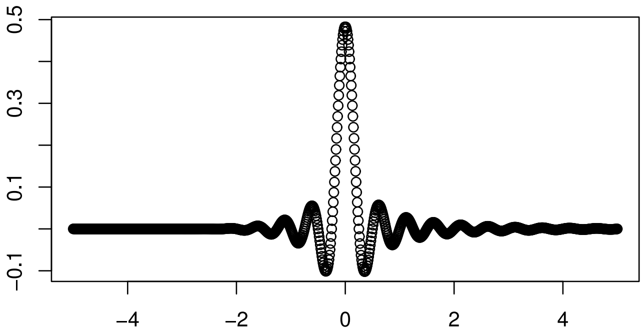

Fig. 1. Function $G_{0,0}(x)$.

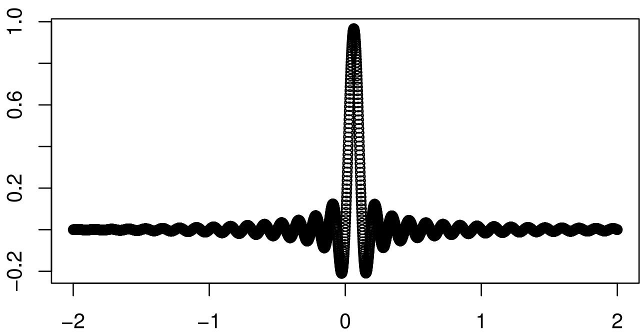

Fig. 2. Function $G_{2,1}(x)$.

Table 7
Values of the approximated integral I. The truncation interval to perform Monte Carlo simulation is $[a,b] = [-10,10]$. Parameter tol is defined as the absolute error in R documentation for routine `integrate` without further details. Given a value of tol, `integrate` returns the number of subintervals used.

Method	$m = 0, k = 0$	Subintervals	$m = 2, k = 1$	Subintervals
GK, tol=10^{-1}	0.1151907	2	0.0543461	3
GK, tol=10^{-3}	0.1211707	10	0.0605648	25
GK, tol=10^{-6}	0.1209852	21	0.0604349	75
MC, $n = 10^6$	0.1217226	–	0.0607478	–
MC, $n = 10^8$	0.1208844	–	0.0605098	–
TH	0.1209854	–	0.0604349	–

and the integral in the right hand side of expression (30) can be approximated by the Monte Carlo sum,

$$\int_0^1 G_{m,k}((b-a)y + a)dy \approx \frac{1}{n}\sum_{i=1}^{n} G_{m,k}((b-a)U_i + a), \tag{31}$$

with U_i independent and identically distributed uniform random variables in $[0,1]$. We refer to this method with the acronym MC.

Finally, we compute the integral of expression (28) by means of the one point approximation of Theorem 1,

$$I = \int_{\mathbb{R}} G_{m,k}(x)dx \approx \frac{1}{2^{m/2}} g\left(\frac{k}{2^m}\right). \tag{32}$$

We refer to this method with the acronym TH.

Results are presented in Table 7. We can observe that TH and GK give very similar results when many subintervals are considered for GK integration. In this case, MC results with $n = 10^8$ agree with GK and TH up to the third decimal place. These outcomes confirm the impressive accuracy and efficiency of TH method, where only one evaluation function is required without domain truncation.

Following similar steps for $f_2(\cdot|h_{t-1})$, $f_2(h_t|\cdot)$ and $f_3(h_1|Y_0)$, we obtain,

$$\hat{f}_2(\xi|h_{t-1}) = e^{-\frac{\sigma_\eta^2}{2}\xi^2 - i(\alpha+\beta h_{t-1})\xi}, \quad \hat{f}_2(h_t|\xi) = \frac{1}{\beta} e^{-\frac{\sigma_\eta^2}{2\beta^2}\xi^2 - i\frac{h_t-\alpha}{\beta}\xi}, \quad \hat{f}_3(\xi|Y_0) = e^{-\frac{\sigma^2}{2}\xi^2 - i\mu\xi}. \tag{33}$$

Then,

$$|\hat{f}_2(\xi|h_{t-1})| = \mathcal{O}\left(e^{-\frac{\sigma_\eta^2}{2}\xi^2}\right), \quad |\hat{f}_2(h_t|\xi)| = \mathcal{O}\left(e^{-\frac{\sigma_\eta^2}{2\beta^2}\xi^2}\right), \quad |\hat{f}_3(\xi|Y_0)| = e^{-\frac{\sigma^2}{2}\xi^2}, \tag{34}$$

and the hypothesis of Theorem 1 is satisfied for all three functions.

It remains to compute the scale of approximation m and the truncation parameters k_1 and k_2.

Proposition 1. *Let ϵ_m be a given tolerance error. Then, for all $m \in \mathbb{Z}$, $m > \log_2\left(-\frac{1}{\pi\sigma}\Phi^{-1}\left(\sigma\sqrt{\frac{\pi}{2}}\epsilon_m\right)\right)$, the projection error $f_3(h_1|Y_0) - \mathcal{P}_m f_3(h_1)$ satisfies $|f_3(h_1|Y_0) - \mathcal{P}_m f_3(h_1)| < \epsilon_m$, where Φ^{-1} represents the inverse of the cumulative distribution function of the standard normal distribution.*

Proof. Lemma 3 of [28] states that,

$$|f_3(h_1|Y_0) - \mathcal{P}_m f_3(h_1)| \leq \frac{1}{2\pi}\int_{|\xi|>2^m\pi} |\hat{f}_3(\xi|Y_0)| d\xi. \tag{35}$$

Then, the result follows by integrating the Fourier transform $\hat{f}_3(\xi|Y_0)$ of $f_3(h_1|Y_0)$ detailed in expression (34). □

Proposition 2. *Let $\epsilon_{k_1}, \epsilon_{k_2}$ be given tolerance errors. Then for all $k_1, k_2 \in \mathbb{Z}$, $k_1 < 2^m(\mu + \sigma\Phi^{-1}(\epsilon_{k_1}))$, $k_2 > 2^m(\mu + \sigma\Phi^{-1}(1-\epsilon_{k_2}))$, the cumulative distribution function $F_3(h_1|Y_0)$ of h_1 satisfies $F_3(k_1/2^m|Y_0) < \epsilon_{k_1}$ and $1 - F_3(k_2/2^m|Y_0) < \epsilon_{k_2}$.*

Proof. The result follows straightforwardly since h_1 is normally distributed. □

Based on Propositions 1 and 2, in what follows, we select $m = \lceil \log_2\left(-\frac{1}{\pi\sigma}\Phi^{-1}\left(\sigma\sqrt{\frac{\pi}{2}}\epsilon_m\right)\right)\rceil$, $k_1 = \lfloor 2^m(\mu + \sigma\Phi^{-1}(\epsilon_{k_1}))\rfloor$, and $k_2 = \lceil 2^m(\mu + \sigma\Phi^{-1}(1-\epsilon_{k_2}))\rceil$, where $\lceil x\rceil := \min\{k \in \mathbb{Z} : k \geq x\}$ and $\lfloor x\rfloor := \max\{k \in \mathbb{Z} : k \leq x\}$.

4.2. Heston model

We consider the celebrated Heston stochastic volatility model studied in [4]. The parameters estimation problem via the likelihood function has been studied, for instance, in [11,13]. We carry out an error analysis in Section 4.2.1, and we show some numerical results in Section 5.2. The model is formulated as follows,

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dB_t, \\ dv_t &= (\omega - \theta v_t)dt + \xi\sqrt{v_t} dZ_t, \end{aligned} \tag{36}$$

where, B_t and Z_t are two correlated Brownian motions with correlation parameter ρ,

$$\rho B_t + \sqrt{1-\rho^2}\tilde{Z}_t = Z_t, \tag{37}$$

where $\tilde{Z}_t$ is a Brownian motion independent of B_t. In this case, the set of parameters to be estimated is, $\boldsymbol{\Theta} = (\omega, \theta, \xi, \rho)$, where $\frac{\omega}{\theta}$ represents the long-run average variance of the asset, θ the rate at which v_t reverts to the mean, and ξ is the volatility of volatility. For sake of convenience, we transform the model in (36) into the log-asset and log-variance form by means of the Itô lemma, and we end up with,

$$\begin{aligned} d\left(\ln S_t\right) &= \left(\mu - \frac{1}{2}v_t\right)dt + \sqrt{v_t}dB_t, \\ d\left(\ln v_t\right) &= \frac{1}{v_t}\left((\omega - \theta v_t) - \frac{1}{2}\xi^2\right)dt + \xi\frac{1}{\sqrt{v_t}}dZ_t. \end{aligned} \tag{38}$$

This representation of the model fits with the form of the integration domain of volatility in the filtering algorithm of Section 3.1. In order to be able to compute the likelihood function, we discretize the model in expression (38),

$$\begin{aligned} \ln S_{t+\Delta t} &= \ln S_t + \left(\mu - \frac{1}{2}v_t\right)\Delta t + \sqrt{v_t}\sqrt{\Delta t}B_t, \\ \ln v_{t+\Delta t} &= \ln v_t + \frac{1}{v_t}\left((\omega - \theta v_t) - \frac{1}{2}\xi^2\right)\Delta t + \xi\frac{1}{\sqrt{v_t}}\sqrt{\Delta t}Z_t, \end{aligned} \tag{39}$$

where Δt is the time step (for daily data we will consider $\Delta t = \frac{1}{252}$ as in [29]). If we define,

$$y_t := \ln S_{t+\Delta t} - \ln S_t, \quad \text{and,} \quad h_t = \ln v_t,$$

then the system in expression (39) can be written as,

$$\begin{aligned} y_t &= \left(\mu - \frac{1}{2}e^{h_t}\right)\Delta t + e^{\frac{h_t}{2}}\sqrt{\Delta t}B_t, \\ h_{t+1} &= h_t + e^{-h_t}\left(\omega - \theta e^{h_t} - \frac{1}{2}\xi^2\right)\Delta t + \xi e^{-\frac{h_t}{2}}\sqrt{\Delta t}Z_t. \end{aligned} \tag{40}$$

We eliminate the correlation in the system of expression (40). To do this, we start by replacing Z_t of expression (37) in the second equation of the system (40),

$$h_{t+1} = h_t + e^{-h_t}\left(\omega - \theta e^{h_t} - \frac{1}{2}\xi^2\right)\Delta t + \xi e^{-\frac{h_t}{2}}\sqrt{\Delta t}\left(\rho B_t + \sqrt{1-\rho^2}\tilde{Z}_t\right), \tag{41}$$

and then, we replace B_t of expression (41) by using the first equation of the system (40). Finally, after some algebraic manipulation we obtain the system,

$$\begin{aligned} y_t &= \left(\mu - \frac{1}{2}e^{h_t}\right)\Delta t + e^{\frac{h_t}{2}}\sqrt{\Delta t}B_t, \\ h_{t+1} &= h_t + \left(e^{-h_t}\left(\omega - \theta e^{h_t} - \frac{1}{2}\xi^2\right) - \xi\rho e^{-h_t}\left(\mu - \frac{1}{2}e^{h_t}\right)\right)\Delta t + \xi\rho e^{-\frac{h_t}{2}} y_t\sqrt{\Delta t} + \xi e^{-\frac{h_t}{2}}\sqrt{1-\rho^2}\sqrt{\Delta t}\tilde{Z}_t. \end{aligned} \tag{42}$$

The density function of y_t given h_t and the density function of h_t given h_{t-1} are, respectively,

$$f_1(y_t|h_t) = \frac{1}{\sqrt{2\pi}s}e^{-\frac{(y_t-m)^2}{2s^2}}, \quad f_2(h_t|h_{t-1}) = \frac{1}{\sqrt{2\pi}\bar{s}}e^{-\frac{(h_t-\bar{m})^2}{2\bar{s}^2}},$$

where,

$$m := \left(\mu - \frac{1}{2}e^{h_t}\right)\Delta t, \quad s := e^{\frac{h_t}{2}}\sqrt{\Delta t},$$

and,

$$\bar{m} := h_{t-1} + \left(e^{-h_{t-1}}\left(\omega - \theta e^{h_{t-1}} - \frac{1}{2}\xi^2\right) - \xi\rho e^{-h_{t-1}}\left(\mu - \frac{1}{2}e^{h_{t-1}}\right)\right)\Delta t + \xi\rho e^{-\frac{h_{t-1}}{2}} y_{t-1}\sqrt{\Delta t},$$

$$\bar{s} := \xi e^{-\frac{h_{t-1}}{2}}\sqrt{1-\rho^2}\sqrt{\Delta t}.$$

It remains to define the distribution of h_1. As the initial variance, we select the exponential distribution of parameter 1, $v_1 \sim \exp(1)$ as in [11]. The initial density function for the log-variance reads,

$$f_3(h_1|Y_0) = e^{-e^{h_1}+h_1}, \quad h_1 \in \mathbb{R}.$$

4.2.1. Error analysis

Following similar steps as those detailed in Section 4.1.1 for $f_1(y_t|h_t)$, $f_2(\cdot|h_{t-1})$, $f_2(h_t|\cdot)$ and $f_3(h_1|Y_0)$, we can see that the hypothesis of Theorem 1 is again satisfied, and it remains to compute the scale of approximation m and the truncation parameters k_1 and k_2.

Proposition 3. *Let ϵ_m be a given tolerance error. Then, for all $m \in \mathbb{Z}$, $m > \log_2\left(-\frac{1}{\pi}\Phi^{-1}\left(\frac{e}{2}\epsilon_m\right)\right)$, the projection error $f_3(h_1|Y_0) - \mathcal{P}_m f_3(h_1)$, satisfies $|f_3(h_1|Y_0) - \mathcal{P}_m f_3(h_1)| \lesssim \epsilon_m$, where Φ^{-1} represents the inverse of the cumulative distribution function of the standard normal distribution.*

Proof. Lemma 3 of [28] states that,

$$|f_3(h_1|Y_0) - \mathcal{P}_m f_3(h_1)| \le \frac{1}{2\pi}\int_{|\xi|>2^m\pi} |\hat{f}_3(\xi|Y_0)| d\xi. \tag{43}$$

We use the approximation $e^{h_1} \approx \frac{h_1^2}{2} + h_1 + 1$ to obtain $f_3(h_1|Y_0) \approx e^{-\frac{h_1^2}{2}-1}$ and an approximation to its Fourier transform,

$$|f_3(h_1|Y_0) - \mathcal{P}_m f_3(h_1)| \le \frac{1}{2\pi}\int_{|\xi|>2^m\pi} |\hat{f}_3(\xi|Y_0)| d\xi \approx \frac{1}{\sqrt{2\pi}}\int_{|\xi|>2^m\pi} e^{-\frac{1}{2}\xi^2-1} d\xi = \frac{2}{e}\Phi\left(-2^m\pi\right). \tag{44}$$

Then, the result follows immediately from expression (44). □

Proposition 4. *Let $\epsilon_{k_1}, \epsilon_{k_2}$ be given tolerance errors. Then for all $k_1, k_2 \in \mathbb{Z}$, $k_1 < 2^m \ln\left(\ln\left(\frac{1}{1-\epsilon_{k_1}}\right)\right)$, and $k_2 > 2^m \ln\left(\ln\left(\frac{1}{\epsilon_{k_2}}\right)\right)$, the cumulative distribution function $F_3(h_1|Y_0)$ of h_1 satisfies $F_3(k_1/2^m|Y_0) < \epsilon_{k_1}$ and $1 - F_3(k_2/2^m|Y_0) < \epsilon_{k_2}$.*

Proof. The result follows straightforwardly since the distribution function of h_1 is calculated in closed form. □

Based on Propositions 3 and 4, in what follows, we select $m = \lceil \log_2\left(-\frac{1}{\pi}\Phi^{-1}\left(\frac{e}{2}\epsilon_m\right)\right)\rceil$, $k_1 = \lfloor 2^m \ln\left(\ln\left(\frac{1}{1-\epsilon_{k_1}}\right)\right)\rfloor$ and $k_2 = \lceil 2^m \ln\left(\ln\left(\frac{1}{\epsilon_{k_2}}\right)\right)\rceil$, where $\lceil x\rceil := \min\{k \in \mathbb{Z} : k \ge x\}$ and $\lfloor x\rfloor := \max\{k \in \mathbb{Z} : k \le x\}$.

5. Numerical experiments

We devote Section 5.1 to numerical experiments on the standard AR(1)-SV model, while Section 5.2 is dedicated to Heston model. Finally, parameters estimation with real data is performed in Section 5.3.

Table 8
RMSE corresponding to the estimation of parameters $\boldsymbol{\Theta} = (\alpha, \beta, \sigma_\eta)$ with $\boldsymbol{\Theta^0} = \boldsymbol{\Theta}$, projection error $\epsilon_m = 10^{-6}$ and truncation errors $\epsilon_{k_1} = \epsilon_{k_2} = 10^{-6}$.

T	N	M	α	β	σ_η	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_\eta$	$\mathcal{E}(\alpha)$	$\mathcal{E}(\beta)$	$\mathcal{E}(\sigma_\eta)$
500	500	10	−0.821	0.9	0.675	−0.814	0.901	0.696	0.0227	0.0034	0.0416
500	500	10	−0.736	0.9	0.363	−0.734	0.901	0.381	0.0083	0.0026	0.0391
500	50	20	−0.706	0.9	0.135	−0.695	0.902	0.154	0.0172	0.0031	0.0456
2000	500	10	−0.736	0.9	0.363	−0.736	0.900	0.370	0.0000	0.0000	0.0238

Table 9
RMSE corresponding to the estimation of parameters $\boldsymbol{\Theta} = (-0.736, 0.9, 0.363)$ with $\boldsymbol{\Theta^0} = (-0.746, 0.89, 0.353)$, projection error $\epsilon_m = 10^{-6}$, truncation errors $\epsilon_{k_1} = \epsilon_{k_2} = 10^{-6}$, $T = 2000$, $N = 500$ and $M = 10$.

$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_\eta$	$\mathcal{E}(\alpha)$	$\mathcal{E}(\beta)$	$\mathcal{E}(\sigma_\eta)$
−0.720	0.902	0.379	0.0158	0.0026	0.0158

Table 10
RMSE corresponding to the estimation of parameters $\boldsymbol{\Theta} = (\alpha, \beta, \sigma_\eta)$ for sample size $T = 500$.
Source: All these results were taken from [6].

Method	α	β	σ_η	α	β	σ_η	α	β	σ_η
TRUE	−0.821	0.9	0.675	−0.736	0.90	0.363	−0.706	0.90	0.135
GLQ [7]	0.28	0.03	0.08	0.43	0.05	0.08	1.72	0.24	0.12
RQ [5]	0.25	0.03	0.08	0.45	0.06	0.09	1.24	0.17	0.10
LA [9]	0.28	0.04	0.10	0.42	0.06	0.11	1.55	0.22	0.14
AGH [6]	0.23	0.03	0.15	0.17	0.02	0.04	0.21	0.03	0.08

5.1. The standard AR(1)-SV model

In what follows, the true parameter set used for data simulation is denoted by $\boldsymbol{\Theta} = (\alpha, \beta, \sigma_\eta)$, while the estimated values are represented by $\hat{\boldsymbol{\Theta}} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}_\eta)$. The initial parameter set used for the optimization process of the likelihood function is denoted by $\boldsymbol{\Theta^0} = (\alpha^0, \beta^0, \sigma_\eta^0)$. The number of samples is N, the size of each sample is T, and the number of iterations of the optimization method is M.

The parameters estimated for $\epsilon_m = \epsilon_{k_1} = \epsilon_{k_2} = 10^{-4}$, when $\boldsymbol{\Theta} = \boldsymbol{\Theta^0} = (-0.821, 0.9, 0.675)$, $N = 1$, $T = 2000$, and $M = 30$ are $\hat{\boldsymbol{\Theta}} = (-0.801, 0.901, 0.656)$. If we consider $\epsilon_m = \epsilon_{k_1} = \epsilon_{k_2} = 10^{-6}$ then $\hat{\boldsymbol{\Theta}} = (-0.803, 0.900, 0.662)$ and no differences are observed if we further reduce the tolerance errors to 10^{-8}. For these reasons, we finally set $\epsilon_m = \epsilon_{k_1} = \epsilon_{k_2} = 10^{-6}$ for the next experiments.

Our next experiment consist of the estimation of three different sets of values corresponding to parameters $(\alpha, \beta, \sigma_\eta)$ taken from [9]. The estimated values $\hat{\boldsymbol{\Theta}} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}_\eta)$ represent the mean of the estimations for N simulations. We consider the values $N = 50{,}500$, $T = 500, 2000$ and $M = 10, 20$, and compute the root mean square error (RMSE) in each case associated with $(\alpha, \beta, \sigma_\eta)$ and denoted by $\mathcal{E}(\alpha), \mathcal{E}(\beta), \mathcal{E}(\sigma_\eta)$, respectively. The RMSE reported ranges from orders 10^{-3} to 10^{-2}, except for the sample size $T = 2000$, where $\mathcal{E}(\alpha) = \mathcal{E}(\beta) = 0$.

This zero error, is due to an extremely accurate computation of the likelihood function by means of Shannon wavelets, since the first iteration of the optimization method (the initial seed equals the set of parameters used for data simulation) gives us the maximum value of the likelihood. We illustrate this fact in Fig. 3, where we compute and plot the likelihood in terms of one changing parameter while the others two remain fixed (the vertical line represents the optimal value of the changing parameter).

If we repeat the experiment for $T = 2000$ but this time with a small perturbation of the initial seed, then the RSME lies within the range 10^{-3} to 10^{-2}. This is illustrated in Table 9.

For sake of comparison, we collect some results on RMSE stated in [6] and present them in Table 10. The first row of the table shows the real values of the parameters used for simulating the data, while the second, third, fourth and fifth rows refer to the methods GLQ (Gauss–Legendre quadrature), RQ (rectangular quadrature), LA (Laplace approximation) and AGH (adaptive Gaussian Hermite quadrature), which were developed by [5–7,9], respectively. In particular, the AGH method is implemented with 21 quadrature points, as stated by the authors of [6]. If we compare those figures with the results of Table 8, we realize that our method is capable of recovering the parameter values with a reduction of the RSME by a factor ranging from 10 to 100. Further, the execution time is illustrated in Fig. 4 for different sample sizes. For $T = 500$ and $M = 10$, the parameters are estimated in about 2 s.

5.2. Heston model

Non-zero correlation is the element that most distinguishes the Heston model from the previous stochastic volatility model that assumed zero correlation between the asset and volatility. Heston model tries to reproduce the so-called leverage effect, that is, the increase in volatility with the decrease of the returns of the underlying, which results in negative correlation between the returns of

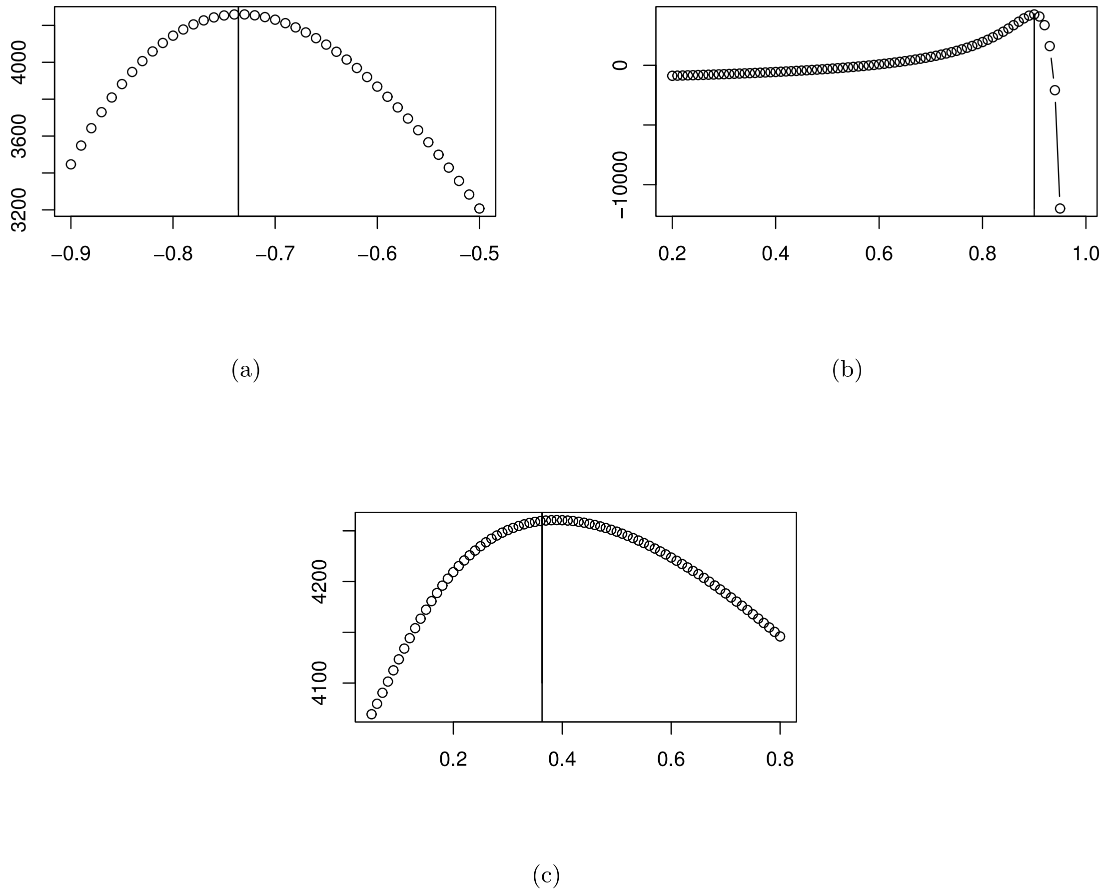

Fig. 3. Likelihood function $-L(\boldsymbol{\Theta}|Y_T)$ corresponding to: (a) $\boldsymbol{\Theta} = (\alpha, 0.9, 0.363)$, (b) $\boldsymbol{\Theta} = (-0.736, \beta, 0.363)$, (c) $\boldsymbol{\Theta} = (-0.736, 0.9, \sigma_\eta)$. Vertical lines denote the true values of the parameters $\alpha, \beta, \sigma_\eta$, respectively. The sample size is $T = 2000$.

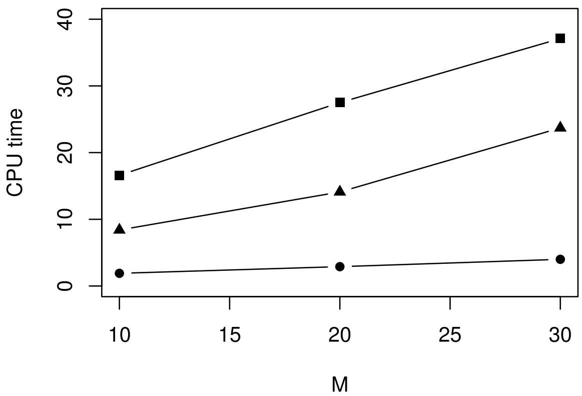

Fig. 4. CPU time in seconds employed in the estimation of parameters with $N = 1$, varying sample size (circles for $T = 500$, triangles for $T = 2500$, squares for $T = 5000$) and different number of iterations $M = 10, 20, 30$ in the optimization method. The tolerance of projection and truncation errors are $\epsilon_m = \epsilon_{k_1} = \epsilon_{k_2} = 10^{-6}$.

the underlying and their volatility. This negative correlation, among other factors, drives the skewness of the distribution of returns in discrete time. This is especially pronounced in stock indexes. Correlation is much more difficulty to estimate than volatility. With just a few observations, we can get a decent estimation of volatility. To estimate volatility, we work with one dimension, the returns of one asset. However, to estimate correlation, we deal with two dimensions of risk. In the Heston model and its simulations, the correlation between the returns of the asset and its volatility is driven by the correlation between two Brownian motions driving each.

Table 11
Results corresponding to the estimation of parameters $\boldsymbol{\Theta} = (\omega, \theta, \xi, \rho)$ with the initial parameter set $\boldsymbol{\Theta^0} = (\omega = 0.15, \theta = 15.0, \xi = 0.02, \rho = -0.40)$, projection error $\epsilon_m = 10^{-6}$, truncation errors $\epsilon_{k_1} = \epsilon_{k_2} = 10^{-6}$, $N = 1$, and $M = 50$. The parameter set used for simulation is $\boldsymbol{\Theta} = (0.10, 10.0, 0.03, -0.50)$. For the sake of comparison, results estimated by means of EKF taken from [29] are presented (rounded to two decimal digits).

T	Shannon wavelets				EKF from [29]			
	$\hat{\omega}$	$\hat{\theta}$	$\hat{\xi}$	$\hat{\rho}$	$\hat{\omega}$	$\hat{\theta}$	$\hat{\xi}$	$\hat{\rho}$
5000	0.132396	14.638018	0.024106	−0.610392	0.150854	15.294576	0.266175	−0.128835
50000	0.105109	15.218574	0.011538	−0.386075	0.126387	12.748852	0.020521	−1.000000
100000	0.182143	15.096761	0.009195	−0.521110	0.136023	13.700906	0.044353	−0.439961

We assess the performance of our method by comparing the outcomes with the results reported in [29], where the authors use the EKF. We estimate the optimal values for the Heston model parameters by simulating large data series and choosing the same perturbation as [29] to the true value of the parameters for the initial seed $\boldsymbol{\Theta^0}$. Results are presented in Table 11. We can observe that Shannon wavelets method is capable of accurately estimating the correlation parameter, while EKF is accurate only for the largest T. Further, when $T = 5000$, the EKF method gives a very bad estimation of the volatility of volatility parameter ξ, while Shannon wavelets estimation is again accurate.

5.3. Estimation of parameters with real market data and option valuation

We devote this section to the valuation of an option of the type defined in Section 2. To carry out this task, we need to select the dynamics of the underlying asset. In the present work, we choose the Heston model described in Section 4.2 because it accounts for stochastic and mean-reverting behavior of the volatility of the underlying asset and the correlation of its changes with the returns of the underlying. We recall that, as it was defined in Section 2, the underlying asset is the price of CB in terms of BTC. As the volatility and returns of the BTC are negatively correlated, we expect the volatility and return of the price of the CB in terms of BTC to positively correlate. The ability of the Heston model to capture the correlation is extremely important. Since the options of interest do not exist in the market, we cannot obtain the parameters of the model by means of calibration with quoted option prices. We therefore turn to estimate them from the time series of the underlying by EKF (see [29] for details) as well as by means of the wavelet method proposed in Section 3.3.

The time series used for the filtering are the same described in Section 2.1, and we consider as the underlying the quotient between the price in dollars of CB and the price in dollars of BTC and then the differences of the logarithm of the resulting data series. This ratio synthesizes the price of CB in terns of BTC. The results obtained with EKF method are,

$$\hat{\omega} = 0.828, \hat{\theta} = 0.707, \hat{\xi} = 0.982, \hat{\rho} = 0.573, \hat{\mu} = 1.180205e - 05. \tag{45}$$

It is natural for μ to be close to zero since it represents the risk free rate in BTC. As BTC has a limited supply one would expect a return on a risk free bond denominated in BTC to be close to zero, as there are no drawbacks of delaying consumption and purchase power in BTC (due to the limited supply).

Note the positive correlation ρ between the returns and changes in volatility. As it happens with many risky assets, drops in the price of BTC increases its volatility. But a drop in the value of the BTC produces an increase in the price of carbon bonds in terms of BTC. Therefore, it is expected to find a positive correlation between the price of the CB in terms of BTC and its volatility. Further, we note the high mean reversion θ of the volatility, also explained mostly by that of the BTC. Once the parameters have been estimated, we proceed to price a call option with 1 day to maturity and strike $K = 0.0015$ BTC/CB. The value obtained is 0.0001625 BTC/CB. The pricing of the option under the Heston model is carried out by means of the method described in [30] with the function `callHestoncf` of R package `nmof`.

This derivative can be used as a hedge against a position in the underlying. Assume that with a drop in the BTC price of the CB (0.00005 BTC/CB), the pollution produces losses of 10,000,000 USD, equivalent to 500 BTC. The delta of this position is 10,000,000 CB, that is, 500 BTC/(0.00005 BTC/CB). Therefore, it can be hedged by selling 10,000,000 CB futures. Alternatively, a put can be used as a hedge. As an example, we consider a put with a strike of 0.0015 BTC/CB and three months to expiration while the futures price is 0.001660029 BTC/CB. The premium of this put is 0.0002497738 BTC/CB or a total of 2497.738 BTC (0.0002497738 BTC/CB × 10,000,000 CB). To compute the premiums, we used the parameters of expression (45).

Next, we estimate the parameters using the wavelet method put forward in this work. We use as the initial seed the values obtained by EKF method (we set $\mu = 0$ and $M = 50$), the values of the estimated parameters are,

$$\hat{\omega} = 0.801, \hat{\theta} = 1.089, \hat{\xi} = 0.304, \hat{\rho} = 0.979,$$

and the call option price is 0.0001613 BTC/CB. We include in Table 12 option prices for different strikes.

As we can observe, there are significant differences between the two methods for the volatility of volatility parameter as well as for the correlation parameter, which is the same behavior observed in the simulation experiments given in Table 11.

In order to check whether the wavelets parameters accurately reproduce the data we perform a Kolmogorov–Smirnov test against the real returns and compare the results with the same test applied to the EKF simulation against the real returns. The null hypothesis of this test is that the data comes from the same distribution. The p-value for the returns produced with the wavelets parameters is 0.3682 and, hence, we cannot reject the null hypothesis that the data comes from the same distribution. The p-value for the returns produced with the EKF parameters is $7.139e - 14$ so we can reject the hypothesis that the data comes from the same distribution.

Table 12
Prices of call options on CB.

Strike	Price
0.0012	0.000460003
0.0013	0.000360003
0.0014	0.000260031
0.0015	0.000161283
0.0016	7.44e−05
0.0017	2.15e−05
0.0018	3.54e−06

6. Conclusions

In this work, we have developed novel financial instruments that could be a helpful tool for governments in designing and implementing climate management policies. The success of this type of instruments relies not only on the financial implementation but also on the body of legislation to support it. We believe that financial climate management by means of derivatives trading by regulatory agencies is a promising strategy for improving social well-being. In regard to the option designed appropriate dynamics for the underlying are needed when it comes to the pricing of the contract. We choose the Heston model because its volatility is stochastic, mean reverting and correlated with price of the underlying, and we estimate its parameters by means of a novel method based on Shannon wavelets. The numerical experiments show the high efficiency of the method in terms of speed and accuracy. Future research might be devoted to finding an initial seed for the maximum likelihood optimization step, which is the most challenging part when dealing with real data, regardless of the method used for computing the likelihood function. A possible solution can be the method of moments (see for instance [31]). As stated in [32], these estimators are not optimal but they are often easy to compute, and they are also useful as starting values for other methods that require iterative numerical routines.

Funding

The research leading to these results received funding from the Spanish Ministry of Economy and Competitiveness, Spain under grants PID2019-105986GB-C21 and PID2020-118339GB-I00.

CRediT authorship contribution statement

Augusto Blanc-Blocquel: Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Luis Ortiz-Gracia:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Rodolfo Oviedo:** Methodology, Conceptualization.

References

[1] J.R. Porter, L. Xie, A. Challinor, K. Cochrane, M. Howden, M.M. Iqbal, D. Lobell, M.I. Travasso, Food security and food production systems, in: Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 2014, pp. 1–82.
[2] G. Luber, N. Prudent, Climate change and human health, Trans. Am. Clin. Climatol. Assoc. 120 (2009) 113–117.
[3] L. Wang, W. Jiao, N. MacBean, M.C. Rulli, S. Manzoni, G. Vico, P. D'Odorico, Dryland productivity under a changing climate, Nature Clim. Change 12 (2022) 981–994.
[4] S. Heston, A closed-form solution for options with stochastic volatility with applications to bond and currency options, Rev. Financ. Stud. 6 (1993) 327–343.
[5] F. Bartolucci, G. De Luca, Maximum likelihood estimation for a latent variable time series model, Appl. Stoch. Models Bus. Ind. 17 (2001) 5–17.
[6] S. Cagnone, F. Bartolucci, Adaptive quadrature for maximum likelihood estimation of a class of dynamic latent variable models, Comput. Econ. 49 (2017) 599–622.
[7] M. Fridman, L. Harris, A maximum likelihood approach for non-Gaussian stochastic volatility models, J. Bus. Econom. Statist. 16 (3) (1998) 284–291.
[8] W. Lee, J. Lim, Y. Lee, J. Castillo, The hierarchical-likelihood approach to autoregressive stochastic volatility models, Comput. Statist. Data Anal. 55 (2011) 248–260.
[9] J. Shimada, Y. Tsukuda, Estimation of stochastic volatility models: An approximation to the nonlinear state space representation, Commun. Stat. - Simul. Comput. 34 (2005) 429–450.
[10] T. Watanabe, A non-linear filtering approach to stochastic volatility models with an application to daily stock returns, J. Appl. Econometrics 14 (1999) 101–121.
[11] A.F. Atiya, S. Wall, An analytic approximation of the likelihood function for the heston model volatility estimation problem, Quant. Finance 9 (3) (2009) 289–296.
[12] J.F. Bégin, M. Boudreault, Likelihood evaluation of jump-diffusion models using deterministic nonlinear filters, J. Comput. Graph. Statist. 30 (2) (2021) 452–466.
[13] F. Cacace, A. Germani, M. Papi, On parameter estimation of heston's stochastic volatility model: a polynomial filtering method, Decis. Econ. Finance 42 (2019) 503–525.
[14] A.S. Hurn, K.A. Lindsay, A.J. and McClell, Estimating the parameters of stochastic volatility models using option price data, J. Bus. Econom. Statist. 33 (4) (2015) 579–594.
[15] S.J. Koopman, A. Lucas, M. Scharth, Numerically accelerated importance sampling for nonlinear non-Gaussian state-space models, J. Bus. Econom. Statist. 33 (1) (2015) 114–127.
[16] S. Hughes, S. Giest, L. Tozer, Accountability and data-driven urban climate governance, Nature Clim. Change 10 (2020) 108–1090.

[17] H. Ritchie, M. Roser, P. Rosado, CO_2 and greenhouse gas emissions, 2020, Available at https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions.
[18] A. Alfonsi, N. Vadillo, A stochastic volatility model for the valuation of temperature derivatives, 2022, Available at www.arxiv.org.
[19] I. Dafnomilis, M. den Elzen, H. van Soest, F. Hans, T. Kuramochi, N. Höhne, Exploring the impact of the COVID-19 pandemic on global emission projections, 2020, Available at www.newclimate.org.
[20] R.I. Ray, V.P. Singh, S.K. Singh, B.S. Acharya, Y. He, What is the impact of COVID-19 pandemic on global carbon emissions? Sci. Total Environ. 816 (2022) 151503.
[21] M. Ronaghi, E. Scorsone, The impact of COVID-19 outbreak on CO_2 emissions in the ten countries with the highest carbon dioxide emissions, J. Environ. Pub. Health (2023) article ID 4605206, 10 pages.
[22] L. Barbaglia, C. Croux, I. Wilms, Volatility spillovers in commodity markets: A large *t*-vector autoregressive approach, Energy Econ. 85 (2020) 104555.
[23] I. Daubechies, Ten lectures on wavelets, in: CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, Pa, USA, 1992.
[24] C. Cattani, Shannon wavelets theory, Math. Probl. Eng. 2008 (2008) article ID 164808, 24 pages.
[25] A. Leitao, L. Ortiz-Gracia, E.I. Wagner, SWIFT valuation of discretely monitored arithmetic Asian options, J. Comput. Sci. 28 (2018) 120–139.
[26] F. Stenger, Handbook of Sinc Numerical Methods, CRC Press, Inc., Boca Raton, FL, USA, 2011.
[27] R. Piessens, E. deDoncker Kapenga, C. Uberhuber, D. Kahaner, Quadpack: A Subroutine for Automatic Integration, Springer Verlag, 1983.
[28] S.C. Maree, L. Ortiz-Gracia, C.W. Oosterlee, Pricing early-exercise and discrete barrier options by Shannon wavelet expansions, Numer. Math. 136 (4) (2017) 1035–1070.
[29] A. Javaheri, Inside Volatility Arbitrage, the Secrets of Skewness, Wiley Finance, 2005.
[30] J. Gatheral, The Volatility Surface: A Practitioner's Guide, in: Wiley Finance Series, 2006.
[31] R. Azencott, P. Ren, I. Timofeyev, Realised volatility and parametric estimation of Heston SDEs, Finance Stoch. 24 (2020) 723–755.
[32] L. Wasserman, All of Statistics: A Concise Course in Statistical Inference, in: Springer Texts in Statistics, 2004.