

# Stochastic Volatility Models with Endogenous Breaks in Volatility Forecasting



Akram S. Hasanov  and Salokhiddin S. Avazkhodjaev 

**Abstract** The need for research on modelling and forecasting financial volatility has increased noticeably due to its essential role in portfolio and risk management, option pricing, and dynamic hedging. This paper contributes to the ongoing discussion of how researchers use regime shifts or structural breaks information to improve forecast accuracy. To accomplish this, we use the data on renewable energy markets. Thus, this study examines several models that accommodate regime shifts and investigates their forecasting performance. First, a subset of competing models (GARCH-class and stochastic volatility) employ the modified iterative cumulative sum of squares method to determine the estimation windows. This paper's novel aspect is that it studies the forecasting performance of various specifications of stochastic volatility models under this procedure. Second, we employ Markov switching GARCH models under alternative distribution assumptions. The rolling window-based forecast analysis reveals that Markov switching models offer more accurate volatility forecast results for most cases. Regarding distribution functions' relevance, the normal distribution followed by Student's  $t$ , skew Student  $t$ , and generalized hyperbolic distribution commonly dominates the series under investigation in the superior sets under all considered loss metrics.

**Keywords** Volatility modelling and forecasting · Regime shifts · Renewable energy · The rolling window

---

A. S. Hasanov (✉)

Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, Subang Jaya, 47500 Kuala Lumpur, Selangor, Malaysia

e-mail: [akram.hasanov@monash.edu](mailto:akram.hasanov@monash.edu)

S. S. Avazkhodjaev

Tashkent Institute of Finance, 60A, A. Temur Street, 100000 Tashkent, Republic of Uzbekistan

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

M. K. Terzioğlu (ed.), *Advances in Econometrics, Operational Research,*

*Data Science and Actuarial Studies*, Contributions to Economics,

[https://doi.org/10.1007/978-3-030-85254-2\\_6](https://doi.org/10.1007/978-3-030-85254-2_6)

## 1 Introduction

The studies on modelling and forecasting financial volatility occupy a considerable portion of empirical finance literature due to its essential applications in portfolio and risk management, option pricing, and dynamic hedging. This study mainly explores the role of regime changes in forecasting the volatility in renewable energy markets through the moving window approach. Also, there have been somewhat little research works on modelling and predicting the volatility of renewable energy markets despite the substantial necessity resulting from renewable energy supporting policies worldwide. Thus, this investigation also aims to contribute to the limited literature on renewable energy markets' volatility prediction analysis. We employ the three commonly used indices in the renewable energy sector, namely, the European Renewable Energy Index (ERIX), the S&P Global Clean Energy Index (S&P GCE), and the Wilder Hill Clean Energy Index (ECO).

Past literature documents that ignoring regime changes in volatility models may overestimate the model coefficients (i.e., persistence), resulting in inaccurate volatility forecasts (see, among others, Lamoureux and Lastrapes 1990; Nomikos and Pouliasis 2011). Thus, the models that take structural breaks into account can be more suitable for predicting renewable energy market volatility.

We employ several models to accommodate structural breaks. First, some models under consideration employ the modified iterative cumulative sum of squares (ICSS) algorithm developed by Sansó et al. (2004) to determine the volatility models' estimation windows. We employ the ICSS method to determine the samples for estimations. This strategy of computing the volatility forecasts is, to some extent, similar to the approach employed in Rapach and Strauss (2008). However, while Rapach and Strauss (2008) assume Gaussian distribution for the maximum likelihood function, we consider, together with Gaussian, some other conditional densities, such as the asymmetric and symmetric fat-tailed density functions. Another novel aspect is that we examine the forecasting performance of stochastic volatility models under this procedure. Second, we employ the Markov switching GARCH models (i.e., MS-GARCH, MS-EGARCH, and MS-GJRGARCH) proposed by Haas et al. (2004) with different distributional assumptions. Besides regime shifts, we have attempted to accommodate various stylized facts in renewable energy markets such as skewness, non-normality, asymmetry, and excess kurtosis.

As usually documented, the proper conditional distribution for asset returns is essential for options valuation and asset pricing (e.g., Hasanov et al. 2018). However, the role of various conditional densities in out-of-sample forecasting analysis is still an open question in the literature. One may compute the volatility forecasts by employing various methods, depending on the variance models and conditional density functions.

Chuang et al. (2007) mentioned that a distribution function needs to comprise some essential features of asset returns, such as shapes, skewness, and kurtosis. In the literature, the statisticians have developed several complex distribution functions (e.g., Fernandez and Steel 1998; Theodossiou 1998). However, a limited number of works

focus on stochastic volatility (SV) and GARCH-type and models' volatility forecasting performance using various distribution functions (with the notable exception of Harvey et al. (1994); Giot and Laurent (2003); Chuang et al. (2007); Hasanov et al. (2018)). This paper employed the four types of conditional density functions in the empirical estimations of GARCH-type models. These are the Student  $t$  (STD), Gaussian (N), skew Student  $t$  (SSTD), and generalized hyperbolic distribution (GHYP). In comparison, we rely on two distributions for SV models, such as Gaussian and Student  $t$ .

In this study, we employ the model confidence set (MCS) test technique proposed in a study by Hansen et al. (2011) to assess the predicting performance of models under consideration following the comparatively recent research works in this context (see, among others, Charles and Darné (2017); Laporta et al. (2018); Zhang et al. (2019); and Hasanov et al. (2020)).

We contribute to a few research questions. First, we examine how one needs to consider the structural breaks to improve the forecast accuracy for the markets under consideration. Second, we investigate whether stochastic volatility models' forecasting performance is improved compared with GARCH-type models when the breaks in log-returns are considered in both models. Third, we analyze the role of distributions in volatility prediction performance. Finally, we also study whether asymmetric models perform better than symmetric models.

We organize the remainder of this study as follows. In Sect. 2, we provide some information about the data employed in this study. Section 3 comprises the methodology, including model specifications, break test, and out-of-sample forecasting analysis procedure. Section 4 includes forecasting results and provides some discussion. Finally, Sect. 5 concludes the paper and highlights some practical implications.

## 2 Data

We rely on the three most widely used stock indices in the renewable energy sector to represent the renewable energy market. First, in the analysis, we use the Wilder Hill Clean Energy Index (ECO). These series are constructed as the weighted mean of the corporate stocks of publicly traded entities whose business operations may have benefited noticeably from a conventional societal position towards cleaner energy use and conservation. Second, we selected the European Renewable Energy Index (ERIX). This index series comprises 10 renewable energy companies' corporate stocks in biomass, water, wind, and solar energy in Europe. Finally, we took the S&P Global Clean Energy Index (S&P GCE). These index series are computed as the weighted mean of the corporate shares of 30 companies devoted to developing renewable energy technologies worldwide. The sample periods for all three indices end on November 30, 2020, and start on January 3, 2005. We retrieve the data on these renewable energy indices from the Bloomberg database.

### 3 Empirical Models

We use some univariate models to predict renewable energy market volatility. We employ the historical mean (HM), exponentially smoothing (ES), GARCH-class, Markov-switching GARCH (MS-GARCH), and stochastic volatility models.<sup>1</sup> We take the structural breaks or regime shifts into account by using two ways. First, we rely on the adjusted ICSS method to select the estimation windows for different specification of SV and GARCH-type conditional variance models. Second, we employ Markov-switching GARCH models under various distribution assumptions to model and forecast conditional volatility.

#### 3.1 The Stochastic Volatility Models

The main characteristics of the SV model are its stochastic and time-varying features of the variance evolution. Specifically, one assumes that the log-variance of the model follows an AR(1) process. Hence, we may specify the following models. Let vector  $y = (y_1, \dots, y_n)^T$  comprises the demeaned log-return observations of an asset. The usual SV model assuming normal distribution can be specified as:

$$\begin{aligned} y_t &= \mathbf{x}_t \boldsymbol{\beta} + \exp\left(\frac{h_t}{2}\right) \varepsilon_t \\ h_{t+1} &= \mu + \varphi(h_t - \mu) + \sigma \eta_t \\ \varepsilon_t &\sim N(0, 1) \\ \eta_t &\sim N(0, 1) \end{aligned} \tag{1}$$

We have also considered the SV model with the conditional Student's  $t$  distribution suggested by Harvey et al. (1994).

$$\begin{aligned} y_t &= \mathbf{x}_t \boldsymbol{\beta} + \exp\left(\frac{h_t}{2}\right) \varepsilon_t \\ h_{t+1} &= \mu + \varphi(h_t - \mu) + \sigma \eta_t \\ \varepsilon_t &\sim t_v(0, 1) \\ \eta_t &\sim N(0, 1) \end{aligned} \tag{2}$$

where  $t_v$  is the Student's  $t$  distribution with  $v$  degrees of freedom, unit variance and zero mean.

We also employ the SV model that accounts for the leverage effect.

---

<sup>1</sup> For a detailed description of the HM and ES models, we refer interested readers to Sadorsky (2006) and Hasanov et al. (2020).

$$\begin{aligned}
 y_t &= \mathbf{x}_t \boldsymbol{\beta} + \exp\left(\frac{h_t}{2}\right) \varepsilon_t \\
 h_{t+1} &= \mu + \varphi(h_t - \mu) + \sigma \eta_t \\
 \varepsilon_t &\sim \mathbf{N}(0, 1) \\
 \eta_t &\sim \mathbf{N}(0, 1)
 \end{aligned}
 \tag{3}$$

where the correlation matrix of  $(\varepsilon_t, \eta_t)$  is defined as

$$\Sigma^\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

In the above SV models,  $\varepsilon_t$  and  $\eta_t$  are independent. The  $h$  denotes the log-variance process,  $\mathbf{x}_t$  is a vector of regressors, and  $\boldsymbol{\beta}$  is a vector of coefficients. We use the Greek letters  $\mu$ ,  $\varphi$ , and  $\sigma$  to denote the parameters of models.

### 3.2 The GARCH-Type Models

This study’s first deterministic conditional volatility model is the simple GARCH(1, 1) model proposed by Bollerslev (1986). The GARCH(1, 1) model for a given log-return series (i.e.,  $r_t$ ) with a mean model, which follows ARMA( $p, q$ ) process can be written as:

$$r_t = \mu + \sum_{i=1}^p \phi_i r_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}
 \tag{4}$$

$$\varepsilon_t = \sigma_t \varepsilon_t, \sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2
 \tag{5}$$

where  $\varepsilon_t$  is a series of identically and independently distributed (i.i.d.) random variables with unit variance and zero mean;  $\sigma_t^2$  denotes the conditional variance, and the parameters of mean and variance models are as follows:  $|\delta| < 1$ ,  $\omega > 0$ ,  $\alpha_1 \geq 0$ ,  $\beta_1 \geq 0$ , and  $\alpha_1 + \beta_1 < 1$ .

As commonly noted, the financial markets frequently demonstrate evidence of the asymmetric volatility phenomenon. This implies that research works should utilize the models that encompass the data’s asymmetry features to examine financial markets’ volatility. The GJR-GARCH specification suggested in Glosten et al. (1993) is another GARCH-class of specification that accounts for the phenomenon of asymmetry. The mean and variance models can be specified as:

$$r_t = \mu + \sum_{i=1}^p \phi_i r_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}
 \tag{6}$$

$$\sigma_t^2 = \omega + \{\alpha_1 + \gamma_1 I(\varepsilon_{t-1} > 0)\} \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (7)$$

where  $\omega > 0$ ,  $\alpha_1 \geq 0$ ,  $\alpha_1 + \gamma_1 \geq 0$ ,  $\beta_1 \geq 0$ , and  $I(\varepsilon_{t-1} > 0)$  is an indicator variable that takes one when  $\varepsilon_{t-1} < 0$  and obtains zero if the argument is false. One may capture an asymmetric impact in the log-return series by analyzing the coefficient estimate of  $\gamma_1$ .

To incorporate conditional variances' asymmetric responses to negative and positive shocks with a similar absolute value, Nelson (1991) developed the exponential GARCH (EGARCH) model specified as:

$$\ln(\sigma_t^2) = \omega + \alpha |z_{t-1}| + \gamma z_{t-1} + \beta \ln(\sigma_{t-1}^2) \quad (8)$$

where  $z_{t-1} = \varepsilon_{t-1} \sigma_{t-1}^{-1}$ .

The GARCH and GJR-GARCH models impose non-negative constraints on the variance equation parameters, while there are no restrictions imposed on the variance coefficients in the EGARCH model.

This paper considers four conditional distributions in the empirical estimations of GARCH-type models. These distributions are as follows: the standardized Student  $t$  (STD), standard Gaussian (N), skew standardized Student  $t$  (SSTD), and generalized hyperbolic distribution (GHYP). For a detailed description of the functional forms of the aforementioned conditional distribution functions, we refer interested readers to Hasanov et al. (2018) and their paper's references.

### 3.3 *The SV and GARCH-Type Models with Endogenously Determined Breaks*

To accommodate the possible structural changes in the log-return series, we rely on the adjusted ICSS procedure developed in the paper by Sansó et al. (2004) to select the estimation windows for the SV and GARCH-type conditional variance model specifications. In the initial step, we apply the  $\kappa_2$  test to all existing observations one through  $T$ . Assume we detected one or more breaks through the modified ICSS procedure. The last break has been found to happen at the time  $T_B$ . Then all models considered in this study are estimated using log-return series starting from  $T_B + 1$  to  $T$  to get an estimate of  $\sigma_{T+1}^2$ . Then we again apply the  $\kappa_2$  test to observations from two through  $T + 1$ . The last breakpoint is now found to happen at the time  $T_{1B}$ . In the next step, we estimate all models for log-returns from  $T_{1B} + 1$  to  $T + 1$  to compute the following estimate of  $\sigma_{T+1}^2$ . We continue this process until all out-of-sample forecast period is exhausted. In sum, we rely on the  $\kappa_2$  break detection test to find the estimation sample so-called the "structural break" model.<sup>2</sup>

<sup>2</sup> In Tables 1, 2 and 3, we use the prefix "SB" to indicate the models with structural breaks.

**Table 1** Model confidence set for ECO ( $s = 1, R = 500$ )

AE	HMAE		HMSE		QLIKE	
	1.000	1.000	1.000	1.000	1.000	1.000
SB-HM		MS-GARCH-STD	MS-GARCH-STD	MS-GARCH-STD	MS-GJR-GARCH-SSTD	1.000
	0.805	MS-GARCH-N	MS-GARCH-N	MS-GARCH-N	MS-GJR-GARCH-N	0.448
	0.805	MS-GARCH-SSTD	MS-GARCH-SSTD	MS-GARCH-SSTD	MS-GARCH-STD	0.963
	0.582	MS-GJR-GARCH-N	MS-GJR-GARCH-N	MS-GJR-GARCH-N	MS-EGARCH-SSTD	0.937
	0.582	MS-GJR-GARCH-SSTD	MS-GJR-GARCH-SSTD	MS-GJR-GARCH-SSTD	MS-GARCH-N	0.937
	0.488	MS-EGARCH-SSTD	MS-EGARCH-SSTD	MS-EGARCH-SSTD	MS-GARCH-SSTD	0.931
		MS-GJR-GARCH-STD	MS-GJR-GARCH-STD	MS-GJR-GARCH-STD	MS-GJR-GARCH-STD	0.519
		MS-EGARCH-N	MS-EGARCH-STD	MS-EGARCH-STD	MS-EGARCH-N	0.324
			MS-EGARCH-N	MS-EGARCH-N	MS-EGARCH-STD	0.324
			SB-ES	SB-ES	SB-ES	0.324
			SB-GJR-GARCH-SSTD	SB-GJR-GARCH-SSTD	SB-GJR-GARCH-STD	0.324
			SB-GJR-GARCH-GHYP	SB-GJR-GARCH-GHYP	SB-GJR-GARCH-SSTD	0.324
			SB-GJR-GARCH-STD	SB-GJR-GARCH-STD	SB-HM	0.324
			SB-GJR-GARCH-N	SB-GJR-GARCH-N	SB-GJR-GARCH-GHYP	0.324
			SB-GARCH-SSTD	SB-GARCH-SSTD	SB-GJR-GARCH-N	0.324
			SB-GARCH-GHYP	SB-GARCH-GHYP	SB-GARCH-STD	0.324
			SB-GARCH-STD	SB-GARCH-STD	SB-GARCH-SSTD	0.324
			SB-HM	SB-HM	SB-GARCH-GHYP	0.324
			SB-GARCH-N	SB-GARCH-N	SB-GARCH-N	0.324

(continued)

Table 1 (continued)

AE	HMAE		HMSE		QLIKE		
SB-HM	1.000	MS-GARCH-STD	1.000	MS-GARCH-STD	1.000	MS-GJR-GARCH-SSTD	1.000
				SB-SV-leverage	0.169	SB-SV-leverage	0.302
				SB-SV	0.169	SB-SV	0.154
				SB-SV-t	0.169	SB-EGARCH-N	0.154
				SB-EGARCH-N	0.169	SB-SV-t	0.137

Notes: The table includes the Superior set of models and the corresponding  $p$ -values calculated through  $T_R$  statistics under the considered loss metrics with  $R = 500$  and  $s = 1$ . The prefix "SB" shows the models with structural changes (endogenously detected). The distributions employed in the maximum likelihood estimation process are (i) a standardized Student  $t$  (STD), (ii) a Gaussian (N), (iii) a skew Student  $t$  (SSTD), and (iv) a generalized hyperbolic (GHYP). The adjusted ICSS regime change test is employed to detect an estimation sample in the models with structural breaks. The prefix "MS" indicates the Markov switching GARCH models, which are estimated by using three conditional distributions: (i) a standardized Student  $t$  (STD), (ii) a Gaussian (N), (iii) and a skew Student  $t$  distribution (SSTD)



**Table 2** Model confidence set for ERIX ( $s = 1, R = 500$ )

AE	HMAE		HMSE		QLIKE	
	1.000	MS-GARCH-N	1.000	MS-GARCH-N	1.000	MS-EGARCH-SSTD
MS-EGARCH-SSTD	1.000	MS-GARCH-N	1.000	MS-GARCH-N	1.000	MS-EGARCH-SSTD
MS-EGARCH-N	0.629	MS-GARCH-SSTD	0.832	MS-GARCH-STD	0.646	MS-GARCH-N
MS-GJR-GARCH-SSTD	0.629	MS-GARCH-STD	0.829	MS-EGARCH-N	0.646	MS-EGARCH-STD
MS-GARCH-N	0.629	MS-EGARCH-SSTD	0.829	MS-GARCH-SSTD	0.646	MS-GARCH-SSTD
MS-EGARCH-STD	0.629	MS-GJR-GARCH-SSTD	0.820	MS-EGARCH-SSTD	0.646	MS-EGARCH-N
MS-GARCH-STD	0.629	MS-GJR-GARCH-N	0.439	MS-GJR-GARCH-SSTD	0.608	MS-GJR-GARCH-SSTD
SB-SV-leverage	0.629	MS-EGARCH-N	0.439	MS-EGARCH-STD	0.435	MS-GARCH-STD
MS-GARCH-SSTD	0.507	MS-GJR-GARCH-STD	0.439	MS-GJR-GARCH-N	0.435	MS-GJR-GARCH-STD
MS-GJR-GARCH-N	0.446	MS-EGARCH-STD	0.439	MS-GJR-GARCH-STD	0.435	MS-GJR-GARCH-N
SB-GJR-GARCH-SSTD	0.446	SB-ES	0.439	SB-ES	0.435	SB-SV-leverage
SB-GJR-GARCH-STD	0.332	SB-SV	0.439	SB-GARCH-N	0.435	SB-HM
		SB-GARCH-N	0.439	SB-GARCH-GHYP	0.435	SB-GJR-GARCH-STD
		SB-SV-t	0.439	SB-GARCH-STD	0.435	SB-GJR-GARCH-SSTD
		SB-GARCH-STD	0.439	SB-GARCH-SSTD	0.435	SB-EGARCH-SSTD
		SB-GARCH-SSTD	0.439	SB-SV	0.435	SB-GJR-GARCH-GHYP
		SB-GARCH-GHYP	0.439	SB-SV-t	0.435	SB-EGARCH-GHYP
		SB-HM	0.439	SB-EGARCH-N	0.435	SB-SV-t
		MS-EGARCH-N	0.439	SB-HM	0.435	SB-SV
		SB-SV-leverage	0.439	SB-EGARCH-SSTD	0.435	SB-GJR-GARCH-N
				SB-EGARCH-GHYP	0.364	SB-EGARCH-N
				SB-GJR-GARCH-N	0.364	SB-GARCH-SSTD

(continued)

**Table 2** (continued)

AE	HMAE	HMSE	QLIKE
MS-EGARCH-SSTD	1.000	MS-GARCH-N	MS-EGARCH-SSTD
		1.000	1.000
		SB-GJR-GARCH-STD	SB-GARCH-STD
		0.253	0.217
		SB-GJR-GARCH-GHYP	SB-GARCH-GHYP
		0.253	0.217
		SB-GJR-GARCH-SSTD	SB-GARCH-N
		0.149	0.217
		SB-SV-leverage	SB-ES
		0.149	0.122

*Notes:* The table includes the Superior set of models and the corresponding  $p$ -values calculated through  $T_R$  statistics under the considered loss metrics with  $R = 500$  and  $s = 1$ . The prefix “SB” shows the models with structural changes (endogenously detected). The distributions employed in the maximum likelihood estimation process are (i) a standardized Student  $t$  (STD), (ii) a Gaussian (N), (iii) a skew Student  $t$  (SSTD), and (iv) a generalized hyperbolic (GHYP). The adjusted ICSS regime change test is employed to detect an estimation sample in the models with structural breaks. The prefix “MS” indicates the Markov switching GARCH models, which are estimated by using three conditional distributions: (i) a standardized Student  $t$  (STD), (ii) a Gaussian (N), (iii) and a skew Student  $t$  distribution (SSTD)

**Table 3** Model confidence set for SPGCE ( $s = 1, R = 500$ )

AE	HMAE		HMSE		QLIKE	
	1.000	MS-GARCH-N	1.000	MS-GARCH-N	1.000	MS-GARCH-N
SB-SV-t	0.629	MS-GARCH-STD	0.832	MS-GARCH-SSTD	0.786	MS-GJR-GARCH-N
	0.629	MS-GARCH-SSTD	0.829	MS-GARCH-STD	0.704	MS-EGARCH-SSTD
	0.629	MS-GJR-GARCH-N	0.829	MS-GJR-GARCH-N	0.704	MS-EGARCH-STD
	0.629		0.820	MS-EGARCH-SSTD	0.226	MS-GARCH-STD
	0.629		0.439	MS-GJR-GARCH-STD	0.226	MS-GARCH-SSTD
	0.629		0.439	MS-EGARCH-N	0.226	MS-GJR-GARCH-STD
	0.507		0.439	MS-GJR-GARCH-SSTD	0.226	MS-EGARCH-N
	0.446		0.439	MS-EGARCH-STD	0.226	MS-GJR-GARCH-SSTD
	0.446		0.439	SB-HM	0.226	SB-HM
	0.332		0.439	SB-EGARCH-N	0.226	SB-EGARCH-N
			0.439	SB-EGARCH-STD	0.226	SB-EGARCH-STD
			0.439	SB-EGARCH-SSTD	0.226	SB-EGARCH-GHYP
			0.439	SB-EGARCH-GHYP	0.226	SB-EGARCH-SSTD
			0.439	SB-GJR-GARCH-STD	0.226	SB-GJR-GARCH-N
			0.439	SB-GJR-GARCH-GHYP	0.226	SB-SV-leverage
			0.439	SB-GJR-GARCH-SSTD	0.226	SB-GJR-GARCH-GHYP
			0.439	SB-GJR-GARCH-N	0.226	SB-GARCH-GHYP
			0.439	SB-SV-leverage	0.226	SB-GJR-GARCH-STD
				SB-SV	0.226	SB-GARCH-SSTD
				SB-SV-t	0.226	SB-GARCH-N

(continued)

**Table 3** (continued)

AE	HMAE		HMSE		QLIKE		
SB-SV-t	<b>1.000</b>	MS-GARCH-N	<b>1.000</b>	MS-GARCH-N	<b>1.000</b>	MS-GARCH-N	<b>1.000</b>
				SB-GARCH-N	0.226	SB-GJR-GARCH-SSTD	0.180
				SB-GARCH-GHYP	0.226	SB-GARCH-STD	0.180
				SB-GARCH-STD	0.226	SB-SV	0.154
				SB-GARCH-SSTD	0.226	SB-SV-t	0.154

*Notes:* The table includes the Superior set of models and the corresponding  $p$ -values calculated through  $T_R$  statistics under the considered loss metrics with  $R = 500$  and  $s = 1$ . The prefix “SB” shows the models with structural changes (endogenously detected). The distributions employed in the maximum likelihood estimation process are (i) a standardized Student  $t$  (STD), (ii) a Gaussian (N), (iii) a skew Student  $t$  (SSTD), and (iv) a generalized hyperbolic (GHYP). The adjusted ICSS regime change test is employed to detect an estimation sample in the models with structural breaks. The prefix “MS” indicates the Markov switching GARCH models, which are estimated by using three conditional distributions: (i) a standardized Student  $t$  (STD), (ii) a Gaussian (N), (iii) and a skew Student  $t$  distribution (SSTD)

If regime shifts or structural changes occur in the log-return series, employing the whole available log-return series to estimate a model may generate inaccurate forecasts, despite having a lesser variance (Pesaran and Timmermann 2007). The existing theory in this context advocates that neglecting the structural breaks or regime shifts may induce upward biases in volatility persistence estimates (for example, Lamoureux and Lastrapes 1990; Mikosch and Stărică 2004). Hence, the observations only over after the break period have been used (i.e., from  $T_B + 1$  to  $T$ ) to estimate the models with breaks, given after-break period is sufficient to run the estimations. Rapach and Strauss (2008) mentioned that a possible shortcoming of this method is that the observations might be insufficient to estimate the model coefficients.

### 3.4 Markov-Switching GARCH Models

In addition to the models described in the previous sections, we also estimate the MS-GARCH model proposed by Haas et al. (2004), assuming different distribution assumptions (Student  $t$ , Gaussian, and skew Student  $t$ ). Ardia et al. (2018) specify the MS-GARCH model as:

$$r_t | (s_t = k, \Psi_{t-1}) \sim D(0, h_{k,t}, \vartheta_k) \quad (9)$$

where  $D(0, h_{k,t}, \vartheta_k)$  is a distribution (i.e., continuous) function with a mean equal to zero, and a time-varying conditional volatility  $h_{k,t}$  in regime  $k$ . Here, a vector,  $\vartheta_k$ , includes additional parameters like skew or tail parameter. The symbol,  $\Psi_{t-1}$ , denotes the available information set. Here, the state variable,  $s_t$ , changes in line with a first-order homogeneous Markov chain with  $k$  states. We rely on three conditional variance models: the GARCH, the EGARCH, and the GJR-GARCH.

## 4 Results and Discussion

The coefficient estimates of all considered models<sup>3</sup> outlined in Sect. 3 have been used to produce daily single-step ahead predictions for conditional variances. We move forward the beginning as well as end dates of the estimation period one day. We re-estimate the model coefficients, and finally, obtained new parameter estimates are employed to predict daily single-step variance (i.e., conditional) over the pre-set out-of-sample period. This process continues until the out-of-sample period has been completed. The numerous studies in the literature applied a rolling window

---

<sup>3</sup> Since HM estimation relies on non-parametric way of estimating the volatility, we do not have any parameters to estimate in HM.

approach of computing out-of-sample predictions. (e.g., Sadorsky 2006; Wen et al. 2016; Charles and Darné 2017; Hasanov et al. 2018; Hasanov et al. 2020).

Following many previous studies, we separate the entire sample for the log-return series under consideration into two parts: in-sample and out-of-sample. While the out-of-sample period comprises the last  $R$  observations, the in-sample period covers the initial  $T$  observations. In this study, the out-of-sample period  $R$  is set to include 500 observations.

As mentioned earlier, we rely on the MCS testing procedure in forecast comparison analysis. We predict the volatility employing the models defined in the previous section estimated on each of the three renewable energy market returns to apply this technique. The MCS testing procedure has been shown to offer a robust comparison for the predictions generated from many models simultaneously and generate a superior set that includes the superior models in forecasting performance with the given pre-specified confidence level. In this study, we have set the confidence level for the MCS test to  $\alpha = 0.90$ . The software code we have written for the analysis relies on the R (version 4.0.2) software and the MCS (Catania and Bernardi 2015), the rugarch (Ghalanos 2016), the stochvol (Hosszejni and Kastner 2016), and the MSGARCH (Ardia et al. 2019) packages.

In this paper, the MCS test is based on the following loss metrics:

$$\begin{aligned}
 AE_{t+1} &= \left| \tilde{V}_{t+1} - \hat{V}_{t+1} \right| \\
 HMAE_{t+1} &= \left| 1 - \frac{\tilde{V}_{t+1}}{\hat{V}_{t+1}} \right| \\
 HMSE_{t+1} &= \left( 1 - \frac{\tilde{V}_{t+1}}{\hat{V}_{t+1}} \right)^2 \\
 QLIKE_{t+1} &= \log\left(\hat{V}_{t+1}^2\right) + \tilde{V}_{t+1}^2 \hat{V}_{t+1}^{-2}
 \end{aligned}$$

where  $\hat{V}_{t+1}$  is the volatility forecasts computed by the estimated models at time  $t$ . And,  $\tilde{V}_{t+1}$  is an actual volatility's proxy at time  $t$ . The squared returns serve as a proxy (see Sadorsky 2006; Hasanov et al. 2020 among many others).

Table 1 shows that the majority of Markov switching models are in the superior set  $\hat{M}_{0.9}$  in terms of three loss metrics (i.e., HMAE, HMSE, and QLIKE). Thus, these three Markov switching models happen to be the most accurate single-period forecasting models for the ECO series. Moreover, the SB-HM model is also in the superior set of models  $\hat{M}_{0.9}$  under most of the loss metrics, and, therefore, this is also considered as a promising model. Under  $AE_1$  loss function, the MCS procedure excludes all other competing models except SB-HM from a superior set of models (SSM). The model confidence set comprises the GARCH-type models and SV models

with endogenously determined structural breaks under two out of four loss criteria (i.e., HMSE and QLIKE).

As one can see from Table 2, under the AE and QLIKE, the MS-EGARCH model with SSTD distribution for the ERIX series shows the highest forecasting performance according to the  $p$ -values computed by the MCS algorithm. Meanwhile, the MS-GARCH with Gaussian distribution provides the most accurate forecasts under HMAE and HMSE loss criteria. The SB-SV model with leverage effects survives in all SSMs, indicating that this is also a favourable model. The MCS procedure selects SB-GJR-GARCH with Student  $t$ , and skew Student  $t$  distributions in the SSM under AE, HMSE, and QLIKE loss functions.

The three loss criteria (HMAE, HMSE, and QLIKE) choose the MS-GARCH-N model as the most promising forecasting model for the SPGCE series. Meanwhile, the MS-GARCH with SSTD, MS-GARCH with STD, and MS-GJR-GARCH with N survive in three out of four SSMs, showing that they are also favourable models. Moreover, the SB-SV-t model for the SPGCE series turned out to be the model with the highest forecasting performance under the AE criterion, based on the  $p$ -values computed by the MCS test (see Table 3). This model also appears in the superior set of models  $\hat{M}_{0,9}$  under HMSE and QLIKE loss criteria. The rest of the models under consideration for the SPGCE index survive when the MCS test relies on HMSE and QLIKE forecast summary statistics.

We also looked into the importance of distributions in forecasting analysis for the models under consideration. The findings for the relevance of distributions in single-step-ahead forecasting are rather mixed. The results show that distribution specifications employed in rolling estimations of models with the most accurate forecasting performance are not consistent with the true underlying distribution of returns. Besides, no dominant distribution is found for the markets under study, which increases forecasting performance. In general, the normal distribution followed by Student  $t$ , skew Student  $t$ , and generalized hyperbolic distribution commonly dominates for the series under investigation in the model confidence sets under all considered accuracy criteria. This result is consistent with those found by Chuang et al. (2007) and Hasanov et al. (2018), who find that the distribution function's complexity does not consistently outperform the less complex one because of the over-fitting problem.

It is worth mentioning that most models in SSMs are asymmetric models across all series and loss metrics. Thus, the renewable energy market participants must not neglect the stylized facts like regime shifts or structural changes and asymmetry when they perform risk management. Moreover, we analyze whether GARCH-based and SV models' forecasting performance, which considers the log-return series' regime changes, is improved. The results suggest that SB-GARCH and SB-SV models show a similar forecasting performance across all the markets under consideration.

## 5 Conclusion and Implications

In this study, we have addressed several research questions. We investigate how one needs to use the structural breaks information to improve the forecast accuracy for the markets under consideration. We find that the Markov switching GARCH models are the superior one-period forecasting models for all markets under investigation. Moreover, we analyze whether stochastic volatility models' forecasting performance is improved compared with GARCH-type models when the endogenously determined breaks in log-returns are considered in both models. The results suggest that both GARCH-type and SV models show a similar forecasting performance across all the markets under consideration.

Also, we analyze the relevance of several distribution specifications in volatility forecasting accuracy analysis. The results indicate no dominant asymmetric and fat-tailed distribution for the markets under study, which increases forecasting performance. In general, the normal distribution followed by Student  $t$ , skew Student  $t$ , and generalized hyperbolic distribution commonly dominates for the series under investigation in the model confidence sets under all considered accuracy criteria. Therefore, renewable energy investors and policymakers must be cautious when employing the SV and GARCH-type models to forecast market volatility. The models with complex distribution functions do not necessarily lead to better forecasting results. Finally, we also study the relevance of typical renewable energy markets like asymmetry. The results suggest that most models in superior sets are asymmetric models across all series and loss measures. Thus, renewable energy markets' participants must not neglect the essential stylized facts when they perform risk management.

This paper's modelling approach helps investors or market participants identify future renewable energy market fluctuations. The market participants, such as portfolio managers and international investors, can predict the future renewable energy market dynamics and design proper portfolio selection and risk management. They intend to generate a more accurate return and volatility predictions to evaluate the portfolio risk exposure and update the hedge ratio according to computed predictions. As Kilian and Park (2009) note, the hedge ratio's continuous adjustment confirms the proper dynamic hedging strategies.

**Acknowledgements** The first author acknowledges the financial support provided by Monash University Malaysia under the Seed grant (B-2-2020).

## References

- Ardia D, Bluteau K, Boudt K, Catania L (2018) Forecasting risk with Markov-switching GARCH models: a large-scale performance study. *Int J Forecast* 34:733–747
- Ardia D, Bluteau K, Boudt K, Catania L, Ghalanos A, Peterson B, Trottier D-A (2019) Package 'MSGARCH'. <https://cran.r-project.org>
- Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. *J Econ* 31:307–327



- Catania L, Bernardi M (2015) Package 'MCS'. <https://cran.r-project.org>
- Charles A, Darné O (2017) Forecasting crude-oil market volatility: further evidence with jumps. *Energy Econ* 67:508–519
- Chuang I, Lu J, Lee P (2007) Forecasting volatility in the financial markets: a comparison of alternative distributional assumptions. *Appl Financ Econ* 17:1051–1060
- Fernandez C, Steel MF (1998) On Bayesian modeling of fat tails and skewness. *J Am Stat Assoc* 93:359–371
- Ghalanos A (2016) Rugarch: univariate GARCH models. R package version 1.4-4. <https://cran.r-project.org>
- Giot P, Laurent S (2003) Market risk in commodity markets: a VaR approach. *Energy Econ* 25:435–457
- Glosten LR, Jaganathan R, Runkle DE (1993) On the relation between the expected value and the volatility of the nominal excess returns on stocks. *J Financ* 48:1779–1801
- Haas M, Mittnik S, Paolella MS (2004) A new approach to Markov-switching GARCH models. *J Financ Economet* 2(4):493–530
- Hansen PR, Lunde A, Nason JM (2011) The model confidence set. *Econometrica* 79:453–497
- Harvey AC, Ruiz E, Shephard N (1994) Multivariate stochastic variance models. *Rev Econ Stud* 61(2):247–264
- Hasanov AS, Poon WC, Al-Freedi A, Heng ZY (2018) Forecasting volatility in the biofuel feedstock markets in the presence of structural breaks: a comparison of alternative distribution functions. *Energy Econ* 70:307–333
- Hasanov AS, Shaiban MS, Al-Freedi A (2020) Forecasting volatility in the petroleum futures markets: a re-examination and extension. *Energy Econ* 86:104626
- Hosszejni D, Kastner G (2016) Package 'stochvol'. <https://cran.r-project.org>
- Kilian L, Park C (2009) The impact of oil price shocks on the US stock market. *Int Econ Rev* 50:1267–1287
- Lamoureux CG, Lastrapes WD (1990) Persistence in variance, structural change, and the GARCH model. *J Bus Econ Stat* 8(2):225–234
- Laporta AG, Merlo L, Petrella L (2018) Selection of value at risk models for energy commodities. *Energy Econ* 74:628–643
- Mikosch T, Střaričá C (2004) Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Rev Econ Stat* 86:378–390
- Nelson D (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* 59:347–370
- Nomikos NK, Pouliasis PK (2011) Forecasting petroleum futures markets volatility: the role of regimes and market conditions. *Energy Econ* 33:321–337
- Pesaran MH, Timmermann A (2007) Selection of estimation window in the presence of structural breaks. *J Econ* 137:134–161
- Rapach DE, Strauss JK (2008) Structural breaks and GARCH models of exchange rate volatility. *J Appl Econ* 23:65–90
- Sadorsky P (2006) Modeling and forecasting petroleum futures volatility. *Energy Econ* 28:467–488
- Sansó A, Arragó V, Carrion JL (2004) Testing for change in the unconditional variance of financial time series. *Rev Econ Financ* 4:32–53
- Theodossiou P (1998) Financial data and the skewed generalised-t distribution. *Manag Sci* 44:1650–1661
- Wen F, Gong X, Cai S (2016) Forecasting the volatility of crude oil futures using HARtype models with structural breaks. *Energy Econ* 59:400–413
- Zhang Y-J, Yao T, Ling-Yun He L-Y, Ripple R (2019) Volatility forecasting of crude oil market: can the regime switching GARCH model beat the single-regime GARCH models? *Int Rev Econ Financ* 59:302–317