

Article

# Hybrid CUSUM Change Point Test for Time Series with Time-Varying Volatilities Based on Support Vector Regression

Sangyeol Lee \*, Chang Kyeom Kim and Sangjo Lee

Department of Statistics, Seoul National University, Seoul 08826, Korea; ckdk95@snu.ac.kr (C.K.K.); lscho13@snu.ac.kr (S.L.)

\* Correspondence: sylee@stats.snu.ac.kr; Tel.: +82-2-880-8814

Received: 9 March 2020; Accepted: 19 May 2020; Published: 20 May 2020



**Abstract:** This study considers the problem of detecting a change in the conditional variance of time series with time-varying volatilities based on the cumulative sum (CUSUM) of squares test using the residuals from support vector regression (SVR)-generalized autoregressive conditional heteroscedastic (GARCH) models. To compute the residuals, we first fit SVR-GARCH models with different tuning parameters utilizing a time series of training set. We then obtain the best SVR-GARCH model with the optimal tuning parameters via a time series of the validation set. Subsequently, based on the selected model, we obtain the residuals, as well as the estimates of the conditional volatility and employ these to construct the residual CUSUM of squares test. We conduct Monte Carlo simulation experiments to illustrate its validity with various linear and nonlinear GARCH models. A real data analysis with the S&P 500 index, Korea Composite Stock Price Index (KOSPI), and Korean won/U.S. dollar (KRW/USD) exchange rate datasets is provided to exhibit its scope of application.

**Keywords:** GARCH time series; change point detection; CUSUM of squares test; support vector regression; machine learning

## 1. Introduction

In this study, we aim to develop a method to detect a significant change in the conditional variance of time series with time-varying volatility using the cumulative sum (CUSUM) of squares test based on the support vector regression (SVR)-generalized autoregressive conditional heteroscedastic (GARCH) residuals. In this task, obtaining accurate predictions of volatilities is crucial because the residuals are obtained as the ratios of the observations and the forecasts of volatility. Traditionally, the prediction has been regarded as an important and challenging task in financial time series analysis owing to its non-stationary and nonlinear nature and, in particular, its volatility. The GARCH model, proposed by Engle [1] and Bollerslev [2], is the most popular model for measuring the volatility of financial time series. This model has advantages for characterizing the properties of time series well, such as persistency, heavy tails, and volatility clustering. Subsequently, several GARCH variants have been developed to better capture these properties, e.g., the asymmetric GARCH model proposed by Engle [3], the exponential GARCH (EGARCH) model proposed by Nelson [4], the threshold GARCH (TGARCH) model proposed by Zakoian [5], the GJR-GARCH model proposed by Glosten, Jagannathan and Runkle [6], and the asymmetric power ARCH (APARCH) model proposed by Ding, Granger and Engle [7]. Refer to Carrasco and Chen [8] for the properties of these models.

Although these GARCH models perform well in general, empirical studies show that they have low forecasting performance in the presence of model misspecification. As the underlying model

of time series is often difficult to identify, new analytical tools such as neural networks (NN) and support vector machines (SVM) have been adopted as a replacement of the classical GARCH models for forecasting time series volatilities, owing to their excellent capability of approximating nonlinearity without knowing the underlying dynamic structure of time series a priori. Refer to Fernandez-Rodriguez, Gonzalez-Martel and Sosvilla-Rivero [9], Cao and Tay [10], Pérez-Cruz, Afonso-Rodriguez and Giner [11], Cherkassky and Ma [12], Chen, Härdle and Jeong [13], Shim, Kim, Lee and Hwang [14], Shim, Hwang and Seok [15], Bezerra and Albuquerque [16], and the papers cited therein. In this study, the SVM method is adopted because it is subordinate to the “structural risk minimization principle” (Vapnik [17]) and therefore seeks a balance between the model complexity and the empirical risk (Smola and Schölkopf [18]) and performs properly for relatively small datasets. In contrast, as mentioned in Tay and Cao [19] and Chen, Härdle and Jeong [13], the NN might suffer from a number of tuning parameters, difficulty in obtaining a global solution, and over-fitting.

Since Page [20], the parameter change detection problem has been a crucial issue in economics, engineering, and medicine, and numerous articles have been published in various research areas. As financial time series often suffer from structural changes owing to changes in governmental policy and critical social events and ignoring them leads to a false conclusion, the change point test has been viewed as a core research topic in time series analysis. See Csörgő and Horváth [21] and Chen and Gupta [22] for a general review. In particular, detecting the volatility change is important in risk management since it affects the calculation of risk measurements such as value-at-risk (VaR) and expected shortfall (ES) (Kim and Lee [23]). The CUSUM test has long been used as a functional tool for detecting a change point owing to its handiness in practice. For earlier works, we refer to Inclán and Tiao [24], Kim, Cho and Lee [25], Lee, Ha, Na and Na [26], Berkes, Horvath and Kokoszka [27], Hillebrand [28], Gombay [29], and Tahmasbi and Rezaei [30]. Subsequently, numerous articles have appeared in the literature. For further developments, we refer to Ross [31], Oh and Lee [32], and the papers cited therein.

Among the CUSUM tests, the estimate-based CUSUM test has been a conventional method for detecting a change point (Lee, Ha, Na and Na [26]) for decades because it directly compares the discrepancy among sequentially obtained estimators and, as such, is quite appealing to the intuition of practitioners. This method performs well in general, but often suffers from severe instability and produces low powers, particularly when the underlying model is a GARCH type. See Kang and Lee [33] and Oh and Lee [34]. To rectify this problem, the residual-based CUSUM test has been proposed because the residuals can effectively discard the correlations of time series and highly enhance the performance of the CUSUM test in terms of stability and power. This method has been firmly advocated by Lee, Tokutsu and Maekawa [35], De Pooter and Van Dijk [36], and Lee and Lee [37]. Oh and Lee [32] recently proposed a modified residual-based CUSUM test to cope with more general location-scale time series models including GARCH models. Despite its merits, the method relies on presumed parametric models and bears the potential risk of leading to a false conclusion when the assumed models are misspecified. To resolve this problem, Lee, Lee and Moon [38] recently designated a hybrid method combining the SVR and CUSUM methods in autoregressive and moving average (ARMA) models and demonstrated its superiority over the classical ARMA-based CUSUM test when the time series has significant nonlinear characteristics. However, as the SVR-ARMA models cannot capture volatility in financial time series and the CUSUM test based on them is highly affected by volatility and easily misidentifies the change point (Lee, Lee and Moon [38]), we aim to develop a CUSUM test based on the SVR-GARCH models.

The rest of this paper is organized as follows. Section 2 introduces the SVR-GARCH model for forecasting volatilities and describes the SVR method in a general framework. Section 3 presents the principle of the CUSUM of squares test for the GARCH models and explains how to apply the SVR-GARCH method for constructing the CUSUM of squares test. Section 4 conducts Monte Carlo simulations to evaluate the performance of the proposed method. Section 5 performs a real data

analysis using the S&P 500 index, KOSPI, and KRW/USD exchange rate datasets. Finally, Section 6 provides the concluding remarks.

## 2. Support Vector Regression for the GARCH Model

Let us consider the GARCH model:

$$\begin{aligned} y_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= g(y_{t-1}^2, \dots, y_{t-p}^2, \sigma_{t-1}^2, \dots, \sigma_{t-q}^2), \end{aligned} \quad (1)$$

where  $g$  is an unknown function,  $p, q$  are nonnegative integers, and  $\{\epsilon_t\}$  is a sequence of independent and identically distributed (i.i.d.) random variables with zero mean and unit variance. When  $g$  is linear, the inferences for stationary GARCH( $p, q$ ) models are well developed in the literature (Francq and Zakoian [39] and Francq and Zakoian [40]). Ideally, this approach provides an excellent analytic tool for prediction, but the accuracy of prediction can be deteriorated owing to a violation of the assumptions such as stationarity and linearity. As such, in this study, we use the SVR-GARCH method for prediction.

SVR is a revision of the SVM, initially proposed by Cortes and Vapnik [41], which seeks an optimal hyperplane that separates the inputs by maximizing the margins between the support vectors and the hyperplane. Equipped with a high generalization ability (Abe [42]), a notable strength of the SVM is that its usage can be extended to nonlinear prediction and classification methods because of its ability to map the inputs to a higher dimensional feature space via nonlinear kernel functions (Cortes and Vapnik [41]).

SVR aims to identify a nonlinear function of the form:

$$f(\mathbf{x}) = w^T \phi(\mathbf{x}) + b,$$

where  $\mathbf{x}$  denotes a vector of inputs,  $w$  and  $b$  are vectors of the regression parameter, and  $\phi$  is a known kernel function. In the context of this paper,  $\mathbf{x}$  is comprised of squared inputs up to lag  $p$  ( $y_{t-1}^2, \dots, y_{t-p}^2$ ) and conditional variance up to lag  $q$  ( $\sigma_{t-1}^2, \dots, \sigma_{t-q}^2$ ). Notice that  $w$  and  $b$  have the same dimension as  $\mathbf{x}$ . The optimal  $w$  and  $b$  that yield the best approximation of  $f$  are determined by exploiting the structure of the  $\epsilon$ -insensitive loss function (Vapnik [17]):

$$\ell_\epsilon(y, f(x)) = \begin{cases} |y - f(x)| - \epsilon & \text{if } |y - f(x)| \geq \epsilon \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

which neglects the errors lying inside the  $\epsilon$ -band surrounding the estimated function  $f$ . The SVR aims to seek a function that keeps the data inside such a band and dismisses excessive complexity (Smola and Schölkopf [18]).

Given the input vectors  $x_i$ , scalar output  $y_i$ ,  $i = 1, \dots, n$ , and a constant  $C > 0$ , we construct the objective function of the SVR using the  $\epsilon$ -insensitive loss function (Vapnik [17]):

$$\begin{aligned} &\text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_{1,i} + \xi_{2,i}), \\ &\text{subject to } \begin{cases} y_i - w^T \phi(x_i) - b \leq \epsilon + \xi_{2,i} \\ w^T \phi(x_i) + b - y_i \leq \epsilon + \xi_{1,i} \\ \xi_{1,i} \geq 0, \xi_{2,i} \geq 0, \end{cases} \end{aligned} \quad (3)$$

where  $\xi_{1,i}, \xi_{2,i} > 0$  denote slack variables that allow some points to lie outside the  $\epsilon$ -band with a penalty and  $C$  denotes a trade-off between the function complexity and the training error. Notice that the former two constraints designate the upper and lower bound, respectively.

To obtain the optimal  $w$  and  $b$ , we formulate an unconstrained optimization problem using Lagrange multipliers (Smola and Schölkopf [18]):

$$\begin{aligned} \text{minimize } L := & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_{1,i} + \xi_{2,i}) \\ & - \sum_{i=1}^n (\beta_{1,i} \xi_{1,i} + \beta_{2,i} \xi_{2,i}) \\ & - \sum_{i=1}^n \alpha_{1,i} (\epsilon + \xi_{1,i} - y_i + w^T \phi(x_i) + b) \\ & - \sum_{i=1}^n \alpha_{2,i} (\epsilon + \xi_{2,i} + y_i - w^T \phi(x_i) - b). \end{aligned} \quad (4)$$

Then, the optimal solution must satisfy the following Karush–Kuhn–Tucker conditions (Vapnik [17]):

$$\begin{cases} w - \sum_{i=1}^n (\alpha_{1,i} - \alpha_{2,i}) \phi(x_i) = 0 \\ \sum_{i=1}^n (\alpha_{1,i} - \alpha_{2,i}) = 0 \\ C - \alpha_{1,i} - \beta_{1,i} = 0 & \forall i = 1, \dots, n \\ C - \alpha_{2,i} - \beta_{2,i} = 0 & \forall i = 1, \dots, n \\ \alpha_{1,i} \geq 0, \alpha_{2,i} \geq 0, \xi_{1,i} \geq 0, \xi_{2,i} \geq 0 & \forall i = 1, \dots, n, \end{cases} \quad (5)$$

and:

$$\begin{cases} \alpha_{1,i} (\epsilon + \xi_{1,i} - y_i + w^T \phi(x_i) + b) = 0 & \forall i = 1, \dots, n \\ \alpha_{2,i} (\epsilon + \xi_{2,i} + y_i - w^T \phi(x_i) - b) = 0 & \forall i = 1, \dots, n \\ (C - \alpha_{1,i}) \xi_{1,i} = 0 & \forall i = 1, \dots, n \\ (C - \alpha_{2,i}) \xi_{2,i} = 0 & \forall i = 1, \dots, n. \end{cases} \quad (6)$$

Substituting Equations (5) and (6) into (4), we obtain the following dual form:

$$\begin{aligned} \text{maximize } & -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_{1,i} - \alpha_{2,i}) (\alpha_{1,j} - \alpha_{2,j}) \phi(x_i)^T \phi(x_j) \\ & - \epsilon \sum_{i=1}^n (\alpha_{1,i} + \alpha_{2,i}) \\ & + \sum_{i=1}^n (\alpha_{1,i} - \alpha_{2,i}) y_i, \end{aligned} \quad (7)$$

$$\text{subject to } \begin{cases} \sum_{i=1}^n (\alpha_{1,i} - \alpha_{2,i}) = 0 \\ 0 \leq \alpha_{1,i} \leq C, 0 \leq \alpha_{2,i} \leq C, \end{cases}$$

where  $\alpha_{1,i}$  and  $\alpha_{2,i}$  denote dual variables (Vapnik [17]). Consequently, the optimization problem in (7) yields the solutions  $\hat{w}, \hat{b}, \hat{f}$  of  $w, b, f$  of the following form (Smola and Schölkopf [18]):

$$\begin{aligned} \hat{w} &= \sum_{i=1}^n (\alpha_{1,i} - \alpha_{2,i}) \phi(x_i), \\ \hat{b} &= \begin{cases} y_i - \hat{w}^T \phi(x_i) - \epsilon, & 0 < \alpha_{1,i} < C \\ y_i - \hat{w}^T \phi(x_i) + \epsilon, & 0 < \alpha_{2,i} < C, \end{cases} \end{aligned}$$

$$\begin{aligned} \hat{f}(x) &= \sum_{i=1}^n (\alpha_{1,i} - \alpha_{2,i}) \phi(x_i)^T \phi(x) + \hat{b} \\ &= \sum_{i=1}^n (\alpha_{1,i} - \alpha_{2,i}) K(x_i, x) + \hat{b}, \end{aligned} \tag{8}$$

with  $K(x, y) = \phi(x)^T \phi(y)$ . Note that  $\hat{w}$  is determined uniquely, whereas  $\hat{b}$  is not, although the two cases rarely coincide. One way to determine  $\hat{b}$  is to obtain the average of the above two values (Abe [42]).

Furthermore, when constructing a model with the SVR here, we employ the Gaussian kernel for  $K$  in (8):

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\gamma^2}\right).$$

As such, we have to determine the tuning parameters  $\gamma^2$ ,  $C$  in (3), and  $\epsilon$  in the loss function (2) appropriately. Accordingly, we consider a cube of  $(C, \gamma^2, \epsilon)$  with  $1 \leq C \leq 100$ ,  $0.1 \leq \gamma^2 \leq 1$ , and  $0.1 \leq \epsilon \leq 1$ . We then employ a grid search method on this cube.

An analogous approach can be adopted for the GARCH time series (Chen, Härdle and Jeong [13]) by constructing the design matrix  $X$  with each row  $x_s$  comprising the lagged time series of lag  $p$  and the lagged conditional variance of lag  $q$ . Specifically, for the GARCH( $p, q$ ) model,

$$\begin{aligned} x_s &= (\check{y}_s^T, \check{\sigma}_s^T)^T, \\ \check{y}_s &= (y_{s-p}^2, y_{s-p+1}^2, \dots, y_{s-1}^2)^T, \\ \check{\sigma}_s &= (\sigma_{s-q}^2, \sigma_{s-q+1}^2, \dots, \sigma_{s-1}^2)^T. \end{aligned} \tag{9}$$

To reflect the structure above, we present a five-step procedure to obtain the predicted SVR-GARCH function.

Given the time series of length  $n + n'$  and the space of tuning parameters, the procedure is illustrated as follows:

- Step 1. Prescribe the points to be evaluated within this space, then divide the given time series into training and validation time series of size  $n$  and  $n'$ , respectively. This preliminary procedure is required for the subsequent task of validating the fitted SVR-GARCH model, which determines the best tuning parameter sets.
- Step 2. Note that the conditional variance  $\sigma_t^2$  of (9) is unknown. As a remedy, replace  $\sigma_t^2$  with the initial estimates  $\check{\sigma}_t^2$ , which plays the role of a proxy of  $\sigma_t^2$ . The estimate  $\check{\sigma}_t^2$  is based on the training time series using a moving average method (Niemira [43]):

$$\check{\sigma}_t^2 = \frac{1}{m} \sum_{j=1}^m y_{t-j+1}^2,$$

where  $m$  is a positive integer. When  $t$  is smaller than  $m$ ,  $\check{\sigma}_t^2$  is computed as an average of the first to the  $t$ th squares of observations, i.e.,  $\check{\sigma}_t^2 = \sum_{j=1}^t y_{t-j+1}^2 / t$  when  $t < m$ .

- Step 3. Given a set of tuning parameters, we estimate  $g$  in (1) with  $\hat{g}$  using the SVR with  $\sigma_t^2$  replaced by  $\check{\sigma}_t^2$ . Then, the estimate  $\hat{\sigma}_t^2$  of  $\sigma_t^2$  is obtained as:

$$\hat{\sigma}_t^2 = \hat{g}(y_{t-1}^2, \dots, y_{t-p}^2, \check{\sigma}_{t-1}^2, \dots, \check{\sigma}_{t-q}^2).$$

- Step 4. Applying the estimated SVR-GARCH model and using the same proxy formula as in Step 2 for the validation time series, the mean absolute error (MAE) is computed as follows:

$$\text{MAE} = \frac{1}{n'} \sum_{t=1}^{n'} |\hat{\sigma}_t^2 - \sigma_t^2|.$$

The MAE escalates the robustness of the model against outliers and therefore provides more flexibility in a model fitting than the root mean squared error.

- Step 5. Repeat Steps 2 to 4 for all the tuning parameter sets selected in Step 1 and choose the combination that minimizes the MAE. Then, perform Steps 2 and 3 using the training and validation time series together to determine the final model, which is used in obtaining the residuals.

**Remark 1.** One might contemplate utilizing different methods regarding the selection of tuning parameters, such as the random search method proposed by Bergstra and Bengio [44]. However, we employ the grid search method in our simulation study because it outperforms the random search method. Furthermore, the choice of  $m$  in the construction of proxy volatilities is critical in constructing a stable test. In this study, we report the case of  $m = 5$  because it provides reasonably satisfactory results the most consistently.

**Remark 2.** One may consider iteratively updating  $\hat{\sigma}_t^2$  until a specific convergent criterion is satisfied, as observed in Chen, Härdle and Jeong [13] and Lee, Lee and Moon [38]. However, the  $\hat{\sigma}_t$  obtained from this iterative approach makes a function either “flatten out” after each iteration, ultimately yielding a constant function, or drastically follow the peak outliers among the initial  $\tilde{\sigma}_t^2$ 's. Thus, this approach is inadequate and disregarded in our study.

**Remark 3.** Instead of MAE, other loss functions, such as the mean squared error (MSE), the root mean squared error (RMSE), and the mean absolute percentage error (MAPE), could be employed. These loss functions, however, are likely to yield suboptimal results. To illustrate, when either the MSE or the RMSE is used, the result deteriorates as these amplify or shrink the losses according to their values. On the other hand, MAPE yields inconsistent results because it uses unobservable  $\tilde{\sigma}_t^2$  rather than  $\hat{\sigma}_t^2$ .

### 3. Hybrid CUSUM Test via the SVR-GARCH Model

Based on the work of Lee, Tokutsu and Maekawa [35], Oh and Lee [34] developed the CUSUM of squares test for the GARCH(1,1) model as follows:

$$\begin{aligned} y_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \quad t = 1, \dots, n, \end{aligned}$$

where  $\omega > 0, \alpha \geq 0, \beta \geq 0, \alpha + \beta < 1$ , and  $\epsilon_t$  are i.i.d. random variables with zero mean and unit variance, to perform a test for a parameter change in  $\omega, \alpha, \beta$ .

More precisely, the null and alternative hypotheses are formulated as follows:

$$\begin{aligned} H_0 &: \omega, \alpha, \beta \text{ remain the same over } t = 1, \dots, n. \text{ vs.} \\ H_1 &: \text{not } H_0. \end{aligned}$$

To test these,  $\omega, \alpha, \beta$  are estimated using their consistent estimators  $\hat{\omega}, \hat{\alpha}, \hat{\beta}$ , such as Gaussian quasi-maximum likelihood estimates (QMLEs) in Francq and Zakoian (2004), and the residuals  $\hat{\epsilon}_t$  are recursively obtained via the equation:

$$\begin{aligned} \hat{\epsilon}_t &= \frac{y_t}{\hat{\sigma}_t}, \\ \hat{\sigma}_t^2 &= \hat{\omega} + \hat{\alpha} y_{t-1}^2 + \hat{\beta} \hat{\sigma}_{t-1}^2, \quad t = 1, \dots, n, \end{aligned}$$

with the initials  $y_0 = 0, \sigma_0^2 = y_1$ . Subsequently, the CUSUM of squares test is given by

$$\hat{T}_n = \max_{1 \leq k \leq n} \frac{1}{\sqrt{n \hat{\tau}_n}} \left| \sum_{t=1}^k \hat{\epsilon}_t^2 - \frac{k}{n} \sum_{t=1}^n \hat{\epsilon}_t^2 \right|$$

with

$$\hat{\tau}_n^2 = \frac{1}{n} \sum_{t=1}^n \hat{\epsilon}_t^4 - \left( \frac{1}{n} \sum_{t=1}^n \hat{\epsilon}_t^2 \right)^2.$$

Letting

$$T_n = \max_{1 \leq k \leq n} \frac{1}{\sqrt{n\tau}} \left| \sum_{t=1}^k \epsilon_t^2 - \frac{k}{n} \sum_{t=1}^n \epsilon_t^2 \right|,$$

with  $\tau^2 = \text{Var}(\epsilon_1^2)$ , Oh and Lee [34] verified that under regularity conditions,  $\hat{T}_n$  behaves asymptotically the same as  $T_n$  under  $H_0$  as  $n$  tends to  $\infty$ , and thus, the limiting null distribution of  $\hat{T}_n$  is the same as  $\sup_{0 \leq s \leq 1} |B^\circ(s)|$ , where  $B^\circ$  denotes a Brownian bridge on the unit interval because  $T_n$  converges to the supremum of a Brownian bridge in distribution due to Donsker's invariance principle; see Billingsley [45]. The critical values then can be obtained asymptotically. For instance, the null hypothesis  $H_0$  of no changes is rejected if  $\hat{T}_n \geq 1.3397$  at the level of 0.05, which can be obtained through Monte Carlo simulations. Furthermore, provided that a change point exists, the location of change is identified as:

$$\hat{k}_n = \operatorname{argmax}_{1 \leq k \leq n} \left| \sum_{t=1}^k \hat{\epsilon}_t^2 - \frac{k}{n} \sum_{t=1}^n \hat{\epsilon}_t^2 \right|.$$

Multiple change points can be detected by following the scheme of Inclán and Tiao [24]. The same approach can be adopted for more general nonlinear location-scale time series models as seen in Oh and Lee [32].

This CUSUM procedure can be applied to the residuals obtained through the SVR-GARCH models, as can be seen in our simulation study, since they will quite likely behave like the true errors  $\epsilon_t$ . Lee, Lee and Moon [38] recently adopted the SVR-ARMA scheme to calculate the residuals in the change point detection problem on ARMA models and affirmed its validity.

In the next section, we evaluate the SVR-GARCH model in the previous section for various settings, wherein we only consider the case of  $p = q = 1$  as most financial time series can be sufficiently described with GARCH(1,1) models (Hansen and Lunde [46]).

#### 4. Simulation Results

In this section, we evaluate the performance of the SVR-GARCH model through simulations using the linear, asymmetric, threshold, GJR, and exponential GARCH models. For this task, we generate a time series of length  $n = 500, 1000, \text{ and } 2000$  to evaluate the empirical sizes and powers at the level of 0.05. The experiment is comprised of the following four steps.

- Step 1. Generate a time series of length  $2n$  from a prescribed GARCH model.
- Step 2. Follow the estimation scheme described in Section 3 with  $m = 5$ . In this procedure, the first  $0.7n$  number of time series constitute the training set, and the following  $0.3n$  number of time series constitute the validation set.
- Step 3. Conduct the CUSUM of squares test described in Section 3. We utilize the remaining  $n$  number of time series as a testing set.
- Step 4. Repeat Steps 1 to 3 1000 times iteratively, and then, compute the empirical sizes and powers.

In the power evaluation, we assume that the change point existed in the middle of the time series. Here, we utilize R 3.5.1 running on Windows 10 and the packages "e1071" and "fGarch".

For this task, various GARCH models are considered as listed below, wherein  $\epsilon_t, t = 1, \dots, 2n$ , denote i.i.d. normal errors with zero mean and unit variance:

- GARCH(1,1) model:

$$\begin{aligned}y_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \\ \alpha &\geq 0, \beta \geq 0.\end{aligned}$$

- AGARCH(1,1) model:

$$\begin{aligned}y_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \omega + \alpha (y_{t-1} - b)^2 + \beta \sigma_{t-1}^2, \\ \alpha &\geq 0, \beta \geq 0.\end{aligned}$$

- GJR-GARCH(1,1) model:

$$\begin{aligned}y_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \omega + \alpha_1 y_{t-1}^{+2} + \alpha_2 y_{t-1}^{-2} + \beta \sigma_{t-1}^2, \\ y_t^+ &= \max(y_t, 0), y_t^- = -\min(y_t, 0), \\ \alpha_1 &\geq 0, \alpha_2 \geq 0, \beta \geq 0.\end{aligned}$$

- TGARCH(1,1) model:

$$\begin{aligned}y_t &= \sigma_t \epsilon_t, \\ \sigma_t &= \omega + \alpha |y_{t-1}| + \beta \sigma_{t-1}, \\ \alpha &\geq 0, \beta \geq 0.\end{aligned}$$

- Log-linear GARCH(1,1) model (a specific variation of the EGARCH( $p, q$ ) model):

$$\begin{aligned}y_t &= \sigma_t \epsilon_t, \\ \log \sigma_t^2 &= \omega + \alpha \log y_{t-1}^2 + \beta \log \sigma_{t-1}^2.\end{aligned}$$

When examining empirical power, we first consider the change in parameters. To illustrate this, we consider the change in  $\omega$ , the sum of  $\alpha$  and  $\beta$  for the GARCH, TGARCH, and log-linear GARCH models, and the change in  $b$  for the AGARCH model. Similarly, for the GJR-GARCH model, we consider a change in the sum of  $\alpha_1$ ,  $\alpha_2$ , and  $\beta$  instead. Under the null hypothesis, the parameters are set as follows:

1. GARCH model:  $\omega = 0.3$ ,  $\alpha = 0.3$ ,  $\beta = 0.3$ ;
2. AGARCH model:  $\omega = 0.3$ ,  $\alpha = 0.3$ ,  $\beta = 0.4$ ,  $b = 1$ ;
3. GJR-GARCH model:  $\omega = 0.3$ ,  $\alpha_1 = 0.3$ ,  $\alpha_2 = 0.4$ ,  $\beta = 0.3$ ;
4. TGARCH model:  $\omega = 0.3$ ,  $\alpha = 0.3$ ,  $\beta = 0.3$ ;
5. log-linear GARCH model:  $\omega = 0.3$ ,  $\alpha = 0.3$ ,  $\beta = 0.3$ .

Tables 1 and 2 indicate the results for the GARCH and AGARCH models, respectively. In both models, no size distortions are observed in most cases. Furthermore, the results confirm that the test performs well in general, in terms of power. Furthermore, as anticipated, the power increases as the sample size increases. Tables 3–5 depict the results for the GJR-GARCH, TGARCH, and log-linear GARCH models, respectively. Although some mild distortions of size can be noticed in the GJR-GARCH model, the results are mostly similar to those of the GARCH and AGARCH models.

For the log-linear GARCH model, the truncated residuals are employed in the construction of the CUSUM test, that is,

$$\tilde{\epsilon}_t := \hat{\epsilon}_t I(|\hat{\epsilon}_t| < N) + NI(|\hat{\epsilon}_t| \geq N)$$

with  $N = 20$  to mitigate the influence of extreme outliers. This modification does not hamper the asymptotic behavior of the CUSUM of squares test, whereas it significantly improves the performance of the CUSUM test in terms of stability and power, as demonstrated in Table 5.

**Table 1.** Empirical sizes and powers for the GARCH(1,1) model.

		$\omega = 0.3$ $\alpha = 0.3$ $\beta = 0.3$	$n = 500$	$n = 1000$	$n = 2000$
size			0.023	0.038	0.055
change of $\omega$	$\rightarrow \omega = 1$		0.761	0.826	0.956
	$\rightarrow \omega = 0.1$		0.612	0.792	0.97
change of $\alpha, \beta$	$\rightarrow \alpha = 0.1, \beta = 0.2$		0.355	0.651	0.949
	$\rightarrow \alpha = 0.4, \beta = 0.5$		0.649	0.802	0.952
change of mixed parameters	$\rightarrow \omega = 0.7, \alpha = 0.1$		0.871	0.969	0.981
	$\rightarrow \omega = 0.1, \beta = 0.1$		0.848	0.952	0.964

**Table 2.** Empirical sizes and powers for the AGARCH(1,1) model.

		$\omega = 0.3$ $\alpha = 0.3, \beta = 0.4$ $b = 1$	$n = 500$	$n = 1000$	$n = 2000$
size			0.031	0.04	0.03
change of $\omega$	$\rightarrow \omega = 1$		0.881	0.951	0.975
	$\rightarrow \omega = 0.1$		0.26	0.613	0.904
change of $\alpha, \beta$	$\rightarrow \alpha = 0.1, \beta = 0.1$		0.783	0.939	0.975
	$\rightarrow \alpha = 0.6$		0.762	0.863	0.941
change of $b$	$\rightarrow b = 0$		0.591	0.897	0.976
	$\rightarrow b = 3$		0.898	0.936	0.957
change of mixed parameters	$\rightarrow \omega = 0.1, \alpha = 0.5, \beta = 0.4, b = 0$		0.565	0.726	0.846
	$\rightarrow \omega = 0.1, \beta = 0.6, b = 2$		0.879	0.926	0.966

**Table 3.** Empirical sizes and powers for the GJR-GARCH(1,1) model.

		$\omega = 0.3$ $\alpha_1 = 0.3, \alpha_2 = 0.4$ $\beta = 0.3$	$n = 500$	$n = 1000$	$n = 2000$
size			0.021	0.029	0.028
change of $\omega$	$\rightarrow \omega = 1$		0.851	0.912	0.947
	$\rightarrow \omega = 0.1$		0.644	0.787	0.866
change of $\alpha_i, \beta (i = 1, 2)$	$\rightarrow \alpha_1 = 0.5, \alpha_2 = 0.8$		0.418	0.695	0.879
	$\rightarrow \alpha_2 = 0.1, \beta = 0.1$		0.421	0.751	0.888
change of mixed parameters	$\rightarrow \omega = 0.5, \alpha_2 = 0.6$		0.657	0.863	0.928
	$\rightarrow \omega = 0.1, \beta = 0.7$		0.717	0.857	0.925

**Table 4.** Empirical sizes and powers for the TGARCH(1,1) model.

		$\omega = 0.3$ $\alpha = 0.3$ $\beta = 0.3$	$n = 500$	$n = 1000$	$n = 2000$
size			0.037	0.049	0.059
change of $\omega$	$\rightarrow \omega = 1$		0.718	0.795	0.879
	$\rightarrow \omega = 0.1$		0.647	0.84	0.902
change of $\alpha, \beta$	$\rightarrow \alpha = 0.1, \beta = 0.2$		0.805	0.886	0.913
	$\rightarrow \alpha = 0.4, \beta = 0.5$		0.735	0.838	0.897
change of mixed parameters	$\rightarrow \omega = 0.7, \alpha = 0.1$		0.907	0.97	0.994
	$\rightarrow \omega = 0.1, \beta = 0.1$		0.499	0.674	0.772

**Table 5.** Empirical sizes and powers for the log-linear GARCH(1,1) model.

		$\omega = 0.3$ $\alpha = 0.3$ $\beta = 0.3$	$n = 500$	$n = 1000$	$n = 2000$
size			0.047	0.039	0.037
change of $\omega$	$\rightarrow \omega = 1$		0.906	0.984	0.997
	$\rightarrow \omega = 0.1$		0.228	0.382	0.507
change of $\alpha, \beta$	$\rightarrow \alpha = -0.1, \beta = -0.2$		0.868	0.946	0.976
	$\rightarrow \alpha = -0.4, \beta = 0.5$		0.917	0.985	1
change of mixed parameters	$\rightarrow \omega = 0.7, \beta = 0.5$		0.82	0.973	0.998
	$\rightarrow \omega = 0.1, \alpha = -0.1$		0.862	0.944	0.971

Finally, to simulate the situation in which the underlying model is unknown, we examine the case where both the parameters and the underlying model itself change. Analogously to the simulation study regarding the above log-linear GARCH model, we truncate the residuals and implement  $\tilde{\epsilon}_t$  to formulate the CUSUM statistic in Cases 1 and 2:

1. GARCH(1,1) changes to log-linear GARCH(1,1);
2. log-linear GARCH(1,1) changes to GARCH(1,1);
3. TGARCH(1,1) changes to AGARCH(1,1);
4. AGARCH(1,1) changes to GJR-GARCH(1,1).

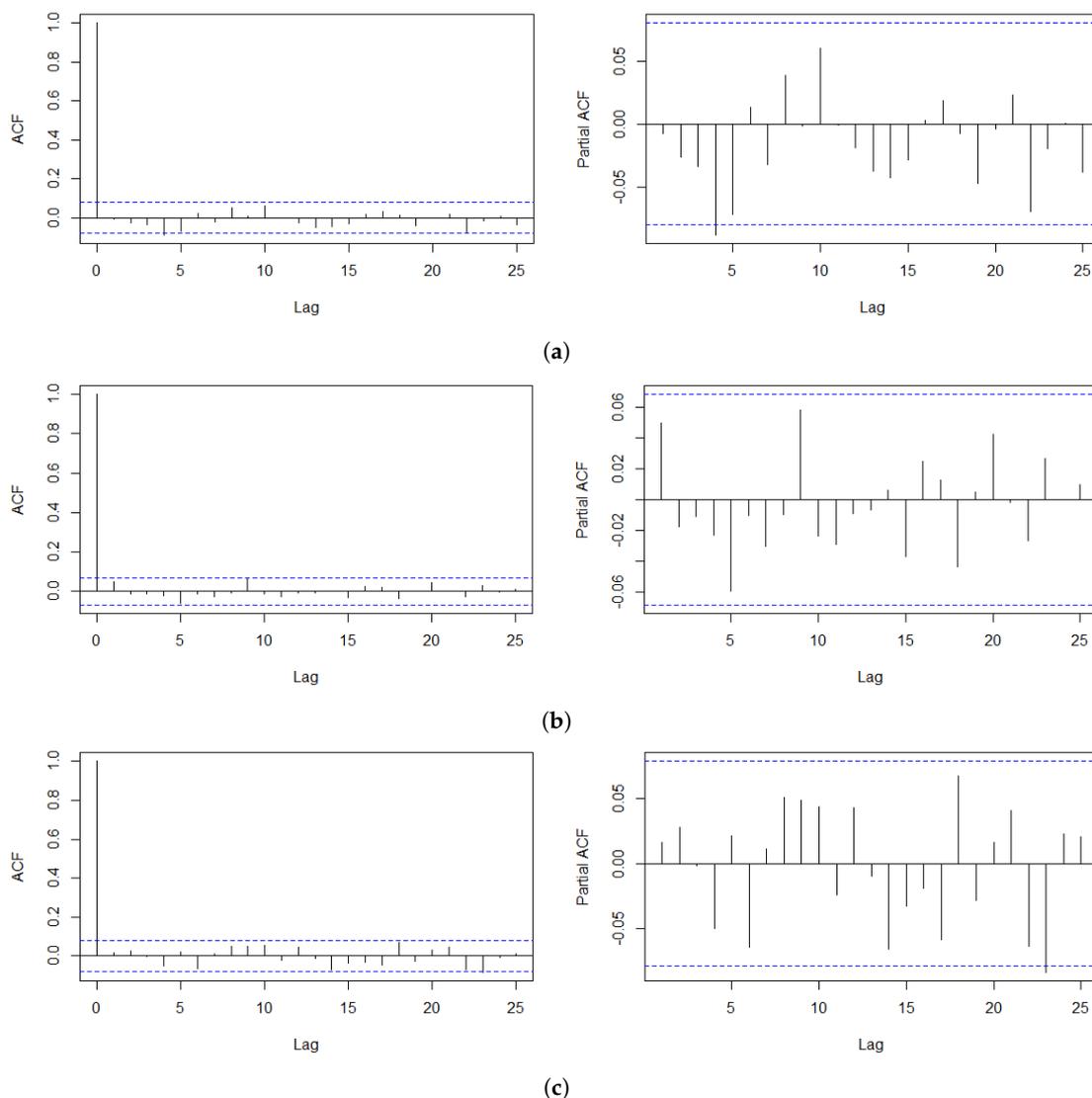
Table 6 shows that the test produces powers comparable to those with only parameter changes. Overall, our findings strongly support the validity of the SVR-GARCH model for detecting a change.

**Table 6.** Empirical powers for model changes (log-linear GARCH = log-GARCH).

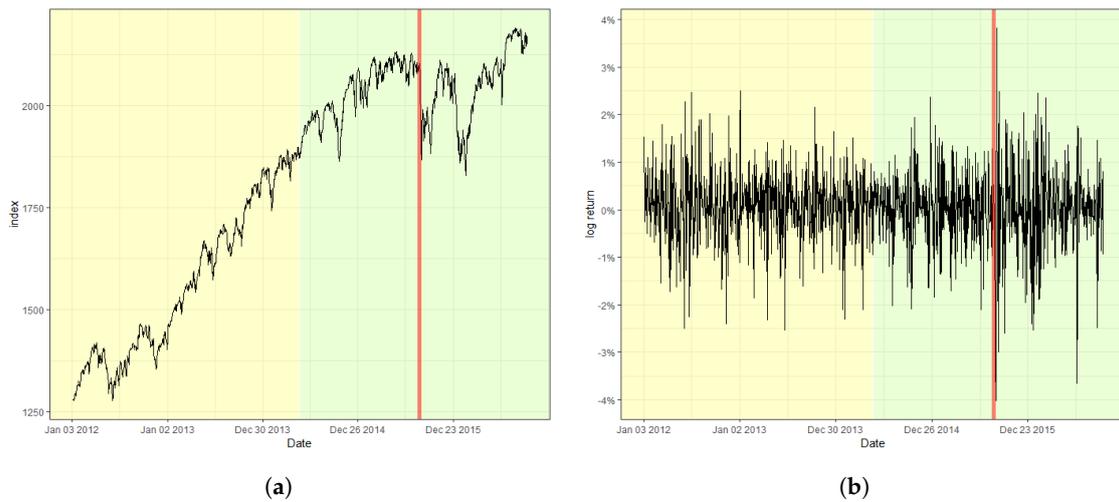
		$\omega = 0.3$ $\alpha = 0.3$ $\beta = 0.3$	$n = 500$	$n = 1000$	$n = 2000$
GARCH $\rightarrow$ log-GARCH	$\rightarrow \omega = 1$		0.879	0.946	0.956
	$\rightarrow \alpha = 0.1, \beta = 0.1$		0.668	0.918	0.969
log-GARCH $\rightarrow$ GARCH	$\rightarrow \omega = 1$		0.907	0.97	0.994
	$\rightarrow \alpha = 0.1, \beta = 0.1$		0.499	0.674	0.772
TGARCH $\rightarrow$ AGARCH	$\rightarrow \omega = 1, b = 1$		0.728	0.795	0.879
	$\rightarrow \alpha = 0.1, \beta = 0.1, b = 1$		0.851	0.891	0.919
		$\omega = 0.3$ $\alpha = 0.3, \beta = 0.3$ $b = 1$	$n = 500$	$n = 1000$	$n = 2000$
AGARCH $\rightarrow$ GJR-GARCH	$\rightarrow \omega = 0.7, \alpha_1 = 0.3, \alpha_2 = 0.7, b = 0$		0.549	0.861	0.958
	$\rightarrow \alpha_1 = 0.1, \alpha_2 = 0.3, \beta = 0.1, b = 0$		0.802	0.931	0.976

## 5. Real Data Analysis

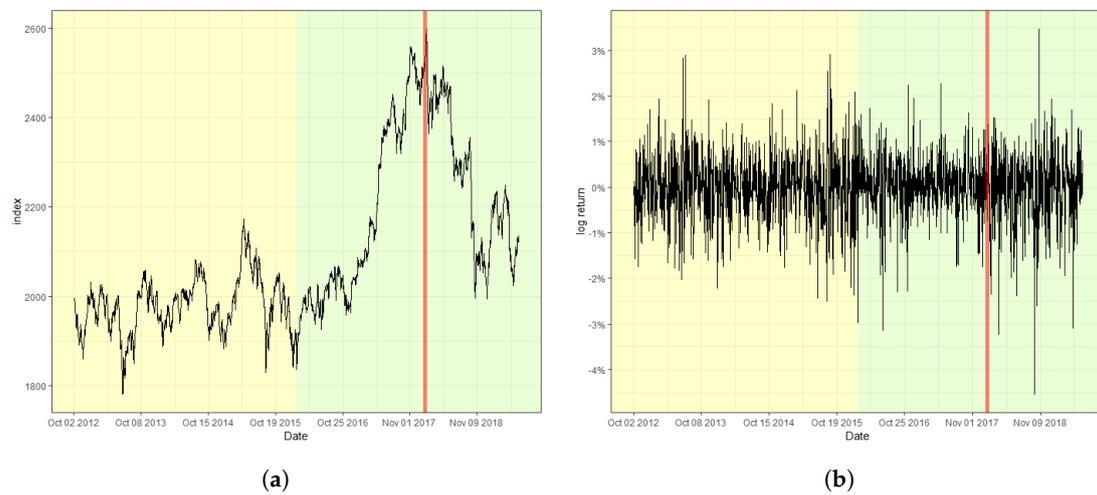
In this section, we analyze the log-returns of the daily index of the S&P 500 from 3 January 2012 to 30 September 2016, the Korea Composite Stock Price Index (KOSPI) from 2 October 2012 to 28 June 2019, and the exchange rate between U.S. dollars (USD) and South Korean Won (KRW), denoted as KRW/USD, from 2 January 2012 to 30 September 2016. We obtain these time series from the website “investing.com”. The two South Korean economic indices are selected because they are well appreciated to be susceptible to various international affairs of the country owing to its geographical location and heavy dependence on the exports of the South Korean economy. Prior to fitting the SVR-GARCH model, we first inspect the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the log-returns of each training time series, consisting of the first half of the entire time series, to examine the presence of irregular patterns of autocorrelations. This step is needed to verify the adequacy of the training time series. Figure 1 shows that the ACFs and PACFs for all three training datasets support stationarity to a great extent, indicating that estimation via the SVR-GARCH model with these training time series would not undermine the outcomes. The selected training datasets are highlighted with yellow shaded areas in Figures 2–4. For example, the training period of the S&P 500 index is from 2 April 2012 to 16 April 2014.



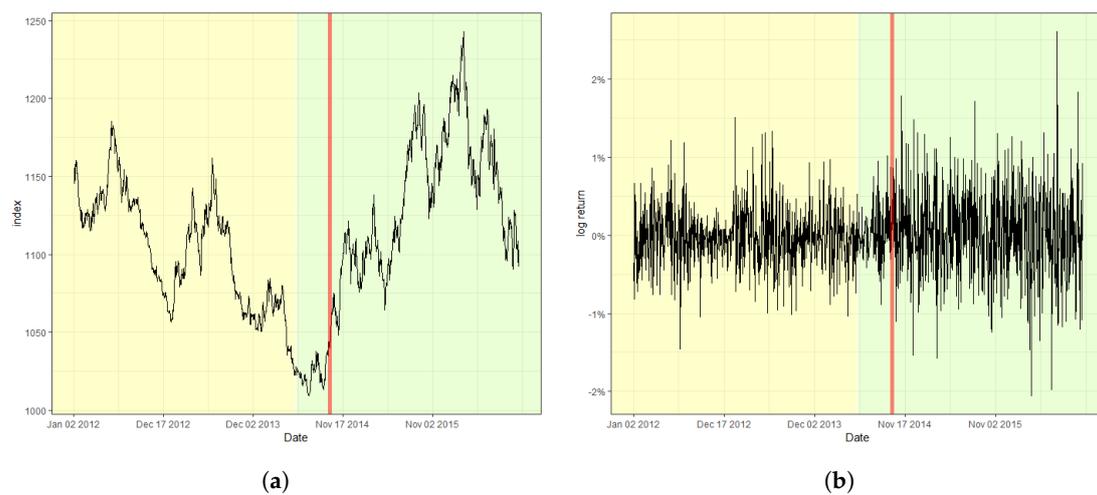
**Figure 1.** Plot of ACF and partial autocorrelation function (PACF) up to lag 25 of log-returns of the (a) S&P 500, (b) KOSPI, and (c) KRW/USD indices.



**Figure 2.** Plot of (a) the raw index of S&P 500 and (b) its log-returns with the detected change point.



**Figure 3.** Plot of (a) the raw index of KOSPI and (b) its log-returns with the detected change point.



**Figure 4.** Plot of (a) the KRW/USD index and (b) its log-returns with the detected change point.

Figure 2 displays the S&P 500 index and its log-returns with the detected change point, 18 August 2015, indicated by a red vertical line. The obtained value of the CUSUM statistic in this case is 2.4751, which rejects the null hypothesis of a lack of changes at the nominal level of 0.05. The result signifies that the index fluctuates more heavily after the change point in comparison with that before

the change point. Moreover, (a) of Figure 2, the original index, shows a relatively consistent trend of steady increase before the change point. However, a double-dip is observed after the change point, which coincides with the fact that a change in trend is often accompanied by a change in variation in stock markets. Furthermore, Figures 3 and 4 plot the result of the change point detection process for the KOSPI and KRW/USD indices, respectively. Here, we obtain the CUSUM statistic values of 1.5081 and 4.2754, respectively, which indicates the detection of a change in both cases at the level of 0.05. The identified location of a change point for the KOSPI index appears to be 22 January 2018, whereas that of KRW/USD is 26 September 2014. Similar to the case of the S&P 500 index, a significant change in trend is also observed in the original datasets, as shown in the plot (a) of Figures 3 and 4. Our finding also illuminates that the log-return of the KRW/USD index experiences a more significant change in volatility, in contrast to the other two cases.

In general, it is not feasible to find the matching incidents that cause change points in economic indices either because they can be obscured from media or they may affect the market with a significant delay. Nevertheless, for the KOSPI example, we could deduce that unstable international affairs could be the culprit behind such changes. These affairs include the nuclear weapon test of North Korea in September 2017, and the participation of North Korea in the Pyeongchang Winter Olympic Games in February 2018, which caused massive turmoil in the Korean peninsula. However, one could argue that the raising of the interest rate by the FRB twice in March and June 2018, was a much more significant factor affecting both the changes in volatility and trend, considering that the impact of the international affairs of the Korean peninsula has been limited on many occasions. In contrast, in the case of the KRW/USD index, the report from the Bank of Korea in January 2015, reasoned that the high volatility after the change point was due to a byproduct of ending quantitative easing and the accompanying improvement of the U.S. economy.

## 6. Conclusions

In this study, we proposed the CUSUM of squares test based on the residuals obtained with the SVR-GARCH model in order to detect a parameter change in the volatility of time series. Monte Carlo simulations were conducted with various linear and nonlinear GARCH models, including the GARCH, GJR-GARCH, AGARCH, TGARCH, and log-linear GARCH models, and the obtained results confirmed the validity of the SVR-GARCH method. Our method was then applied to the analysis of financial datasets such as the S&P 500, KOSPI, and KRW/USD indices and detected one change in all cases. Overall, our findings supported the validity of our method and the practicality in financial time series analysis. Here, we only considered a plain SVR method for emphasizing the hybrid of the CUSUM and SVR methods and for easy access to general readers. However, more sophisticated methods could be employed for refinement concerning the selection of tuning parameters and kernel functions, although they do not necessarily guarantee a better performance, possibly due to over-fitting; see Remark 2 of Lee, Lee and Moon [38]. Furthermore, in this study, we only investigated the retrospective change point test. However, a similar method could be applied to an on-line monitoring process (Huh, Oh and Lee [47]), which aims at an early detection of an anomaly in sequentially observed time series. In this case, the training data could be redesigned via a rolling window procedure. As these issues go beyond the scope of the current study, we leave them to our future project.

**Author Contributions:** Conceptualization, S.L. (Sangyeol Lee); data curation, C.K.K.; formal analysis, C.K.K. and S.L. (Sangjo Lee); funding acquisition, S.L. (Sangyeol Lee); investigation, S.L. (Sangjo Lee); methodology, S.L. (Sangyeol Lee); project administration, S.L. (Sangyeol Lee); validation, C.K.K. and S.L. (Sangjo Lee); visualization, C.K.K.; writing, original draft, S.L. (Sangyeol Lee) and C.K.K. All authors read and agreed to the published version of the manuscript.

**Acknowledgments:** This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning (Grant 2018R1A2A2A05019433).

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

CUSUM	cumulative sum
SVR	support vector regression
SVM	support vector machine
GARCH	generalized autoregressive conditionally heteroscedastic
EGARCH	exponential GARCH
GJR-GARCH	Glosten, Jagannathan, and Runkle-GARCH
TGARCH	threshold GARCH
APARCH	asymmetric power ARCH
NN	neural network
ARMA	autoregressive and moving average
QMLE	quasi-maximum likelihood estimator
KOSPI	Korea Composite Stock Price Index
KRW	Korean Won
VaR	value at risk
ES	expected shortfall

## References

- Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **1982**, *50*, 987–1007. [\[CrossRef\]](#)
- Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [\[CrossRef\]](#)
- Engle, R.F. Stock volatility and the crash of '87: discussion. *Rev. Financ. Stud.* **1990**, *3*, 103–106. [\[CrossRef\]](#)
- Nelson, D.B. Stationarity and persistence in the GARCH (1, 1) model. *Econom. Theory* **1990**, *6*, 318–334. [\[CrossRef\]](#)
- Zakoïan, J.M. Threshold heteroskedastic models. *J. Econ. Dynamic. Control* **1994**, *18*, 931–955. [\[CrossRef\]](#)
- Glosten, L.R.; Jagannathan, R.; Runkle, D.E. On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Financ.* **1993**, *48*, 1779–1801. [\[CrossRef\]](#)
- Ding, Z.; Granger, C.W.; Engle, R.F. A long memory property of stock market returns and a new model. *J. Emp. Financ.* **1993**, *1*, 83–106. [\[CrossRef\]](#)
- Carrasco, M.; Chen, X. Mixing and moment properties of various GARCH and stochastic volatility models. *Econ. Theory* **2002**, *18*, 17–39. [\[CrossRef\]](#)
- Fernandez-Rodriguez, F.; Gonzalez-Martel, C.; Sosvilla-Rivero, S. On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market. *Econ. Lett.* **2000**, *69*, 89–94. [\[CrossRef\]](#)
- Cao, L.; Tay, F. Financial forecasting using support vector machines. *Neural Comp. Appl.* **2001**, *10*, 184–192. [\[CrossRef\]](#)
- Pérez-Cruz, F.; Afonso-Rodriguez, J.; Giner, J. Estimating GARCH models using SVM. *Quant. Financ.* **2003**, *3*, 163–172. [\[CrossRef\]](#)
- Cherkassky, V.; Ma, Y. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Net.* **2004**, *17*, 113–126. [\[CrossRef\]](#)
- Chen, S.; Härdle, W.K.; Jeong, K. Forecasting volatility with support vector machine-based GARCH model. *J. Forecast.* **2010**, *433*, 406–433. [\[CrossRef\]](#)
- Shim, J.; Kim, Y.; Lee, J.; Hwang, C. Estimating value at risk with semiparametric support vector quantile regression. *Comp. Stat.* **2012**, *27*, 685–700. [\[CrossRef\]](#)
- Shim, J.; Hwang, C.; Seok, K. Support vector quantile regression with varying coefficients. *Comp. Stat.* **2016**, *31*, 1015–1030. [\[CrossRef\]](#)
- Bezerra, P.C.S.; Albuquerque, P.H.M. Volatility forecasting via SVR–GARCH with mixture of Gaussian kernels. *Comp. Manag. Sci.* **2017**, *14*, 179–196. [\[CrossRef\]](#)
- Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 2000.

18. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comp.* **2004**, *14*, 199–222. [[CrossRef](#)]
19. Tay, F.; Cao, L. Application of support vector machines in financial time series forecasting. *Omega* **2001**, *29*, 309–317. [[CrossRef](#)]
20. Page, E.S. A test for a change in a parameter occurring at an unknown point. *Biometrika* **1955**, *42*, 523–527. [[CrossRef](#)]
21. Csörgő, M.; Horváth, L. *Limit Theorems in Change-Point Analysis*; John Wiley & Sons Inc.: New York, NY, USA, 1997.
22. Chen, J.; Gupta, A.K. *Parametric Statistical Change Point Analysis with Applications to Genetics, Medicine, and Finance*; Wiley: New York, NY, USA, 2012.
23. Kim, M.; Lee, S. Nonlinear expectile regression with application to value-at-risk and expected shortfall estimation. *Comp. Stat. Data Anal.* **2016**, *94*, 1–19. [[CrossRef](#)]
24. Inclán, C.; Tiao, G.C. Use of cumulative sums of squares for retrospective detection of changes of variance. *J. Am. Stat. Assoc.* **1994**, *89*, 913–923.
25. Kim, S.; Cho, S.; Lee, S. On the cusum test for parameter changes in GARCH (1, 1) models. *Comm. Stat. Theory Meth.* **2000**, *29*, 445–462. [[CrossRef](#)]
26. Lee, S.; Ha, J.; Na, O.; Na, S. The CUSUM test for parameter change in time series models. *Scand. J. Stat.* **2003**, *30*, 781–796. [[CrossRef](#)]
27. Berkes, I.; Horváth, L.; Kokoszka, P. Testing for parameter constancy in GARCH(p,q) models. *Stat. Prob. Lett.* **2004**, *70*, 263–273. [[CrossRef](#)]
28. Hillebrand, E. Neglecting parameter changes in GARCH models. *J. Econometrics* **2005**, *129*, 121–138. [[CrossRef](#)]
29. Gombay, E. Change detection in autoregressive time series. *J. Multi. Anal.* **2008**, *99*, 451–464. [[CrossRef](#)]
30. Tahmasbi, R.; Rezaei, S. Change point detection in garch models for voice activity detection. *IEEE Trans. Audio Speech Lang. Proc.* **2008**, *16*, 1038–1046. [[CrossRef](#)]
31. Ross, G.J. Modelling financial volatility in the presence of abrupt changes. *Phys. A Stat. Mech. Appl.* **2013**, *392*, 350–369. [[CrossRef](#)]
32. Oh, H.; Lee, S. Modified residual CUSUM test for location-scale time series models with heteroscedasticity. *Ann. Inst. Stat. Math.* **2019**, *71*, 1059–1091. [[CrossRef](#)]
33. Kang, J.; Lee, S. Parameter change test for Poisson autoregressive models. *Scand. J. Stat.* **2014**, *41*, 1136–1152. [[CrossRef](#)]
34. Oh, H.; Lee, S. On score vector- and residual-based CUSUM tests in ARMA-GARCH models. *Stat. Methods Appl.* **2018**, *27*, 385–406. [[CrossRef](#)]
35. Lee, S.; Tokutsu, Y.; Maekawa, K. The cusum test for parameter change in regression models with ARCH errors. *J. Japan Stat. Soc.* **2004**, *34*, 173–188. [[CrossRef](#)]
36. De Pooter, M.; Van Dijk, D. *Testing for Changes in Volatility in Heteroskedastic Time Series—A Further Examination*; Econometric Institute Research Papers EI 2004-38; Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute: Rotterdam, the Netherlands, 2004.
37. Lee, S.; Lee, J. Parameter change test for nonlinear time series models with GARCH type errors. *J. Korean Math. Soc.* **2015**, *52*, 503–553. [[CrossRef](#)]
38. Lee, S.; Lee, S.; Moon, M. Hybrid change point detection for time series via support vector regression and CUSUM method. *Appl. Soft Comput.* **2020**, *89*, 106101. [[CrossRef](#)]
39. Francq, C.; Zakoïan, J.M. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **2004**, *10*, 605–637. [[CrossRef](#)]
40. Francq, C.; Zakoïan, J. *GARCH Models: Structure, statistical Inference, and Financial Applications*; Wiley: New York, NY, USA, 2010.
41. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
42. Abe, S. *Support Vector Machines for Pattern Classification*; Springer: London, UK, 2005.
43. Niemira, M.P. *Forecasting Financial and Economic Cycles*; John Wiley & Sons Inc.: New York, NY, USA, 1994; Volume 49.
44. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
45. Billingsley, P. *Convergence of Probability Measure*; Wiley: New York, NY, USA, 1968.

46. Hansen, P.R.; Lunde, A. A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *J. Appl. Econ.* **2005**, *20*, 873–889. [[CrossRef](#)]
47. Huh, J.; Oh, H.; Lee, S. Monitoring parameter change for time series models with conditional heteroscedasticity. *Econ. Lett.* **2017**, *152*, 66–70. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).