



Covariance matrix forecasting using support vector regression

Piotr Fiszeder^{1,2} · Witold Orzeszko³

Accepted: 13 January 2021 / Published online: 23 February 2021
© The Author(s) 2021

Abstract

Support vector regression is a promising method for time-series prediction, as it has good generalisability and an overall stable behaviour. Recent studies have shown that it can describe the dynamic characteristics of financial processes and make more accurate forecasts than other machine learning techniques. The first main contribution of this paper is to propose a methodology for dynamic modelling and forecasting covariance matrices based on support vector regression using the Cholesky decomposition. The procedure is applied to range-based covariance matrices of returns, which are estimated on the basis of low and high prices. Such prices are most often available with closing prices for many financial series and contain more information about volatility and relationships between returns. The methodology guarantees the positive definiteness of the forecasted covariance matrices and is flexible, as it can be applied to different dependence patterns. The second contribution of the paper is to show with an example of the exchange rates from the forex market that the covariance matrix forecasts calculated using the proposed approach are more accurate than the forecasts from the benchmark dynamic conditional correlation model. The advantage of the suggested procedure is higher during turbulent periods, i.e., when forecasting is the most difficult and accurate forecasts matter most.

Keywords Support vector regression · Machine learning · Multivariate volatility models · High and low prices · Range-based models · Covariance forecasting

1 Introduction

Artificial intelligence (AI) offers new approaches to modelling and forecasting real-time data. In the context of financial data analysis, one of the most relevant AI methods is machine learning (ML). Generally, machine learning includes methods that help computer systems automatically improve their performance with experience. These data-driven, self-adaptive techniques require very few assumptions about the models

used for the investigated data. In recent years, several ML methods have been successfully used for forecasting purposes. One of them is the support vector machine (SVM) method proposed by Vapnik [1]. This method, applied to solve both classification and regression problems, is designed to have good generalisability and an overall stable behaviour, implying good out-of-sample performance.

The literature on SVM has been systematically expanding, both in the area of methodology and practical applications. In particular, new methodological approaches, including some modifications of the original SVM models or specific SVM-based hybrid models, have been proposed (e.g., [2–7]).

Originally, the SVM method was developed to solve classification problems; later, however, it was extended to the domain of regression problems [8]. In the literature, the term SVM is typically applied in the context of classification problems, while the term support vector regression (SVR) is used to describe regression with support vector methods.

The econometric literature extensively discusses the empirical properties of financial time series, which include volatility clustering, weak autocorrelation of returns, occurrence of an asymmetric impact of positive and negative shocks on conditional volatility (the so-called leverage effect), long memory, existence of strong dependencies between returns of various

✉ Piotr Fiszeder
piotr.fiszeder@umk.pl

Witold Orzeszko
witold.orzeszko@umk.pl

¹ Department of Econometrics and Statistics, Faculty of Economic Sciences and Management, Nicolaus Copernicus University in Torun, ul. Gagarina 13a, 87-100 Torun, Poland

² Faculty of Finance and Accounting, University of Economics, Prague, W. Churchill Sq. 1938/4, 130 67 Prague 3 – Žižkov, Czech Republic

³ Department of Applied Informatics and Mathematics in Economics, Faculty of Economic Sciences and Management, Nicolaus Copernicus University in Torun, ul. Gagarina 13a, 87-100 Torun, Poland

financial instruments, and some characteristics of return distributions, such as fat tails, leptokurtosis and asymmetry. Many empirical studies have shown that the dynamics of financial processes can be nonlinear (see, e.g., [9–14]; and references therein). Support vector methods, such as other nonparametric statistical analysis techniques, tend to be useful tools of nonlinear forecasting, as they do not presume the linearity of the data-generating process but let the data speak for themselves. They have been successfully applied to forecast financial time series, such as stock indices [3, 15–19], stock prices [20–23], volatility indices [24], derivatives [25–27], exchange rates [15, 28–32], exchange-traded funds [33, 34] and corporate bonds [35].

In the literature, volatility models are usually constructed on the basis of only closing price data. However, databases also usually contain daily low and high prices. These values come from intraday prices and are very important for the measurement of price changes during day. It has been shown that the use of low and high daily prices leads to more accurate estimates and forecasts of variances (see, e.g., [36–40]), covariances (see, e.g., [41–44]) and value-at-risk measures (see, e.g., [45, 46]). Moreover, in contrast to very-high-frequency data, the application of low and high prices does not suffer from a large computational burden. For these reasons, the use of these prices in forecasting is very important from a practical viewpoint.

The main motivation of our paper is to combine two gaining importance and popularity approaches, namely, the application of SVR and low and high prices, to create a new forecasting procedure for the covariance matrix of returns based on daily prices. Modelling and forecasting covariance matrices are vital because financial institutions and investors usually possess portfolios of assets. In the process of construction, valuation and management of financial instruments portfolios, knowledge about the relationships between assets is as important as knowledge about the dynamics of returns and volatilities. Forecasting the covariance matrix is also crucial in applications such as risk management, option pricing and hedging strategies. Such applications require a multivariate approach, whereas most of the SVR studies in finance are univariate. In contrast, in this study, SVR was applied to forecast not only variances of financial returns but also covariances. Modelling and forecasting the covariance matrix are much more demanding tasks because the matrix constructed from the forecasts of variances and covariances obtained by the disjoint models is not guaranteed to be positive definite.

We apply range-based variances and covariances of returns, which are formulated on the basis of low and high prices. Our approach is nonparametric, which makes it more general than the papers mentioned above. Chou et al. [41] combined the conditional autoregressive range (CARR) model by Chou [36] with the dynamic conditional correlation (DCC) model by Engle [47] to propose the range-based

DCC model. Fiszeder et al. [44] proposed the DCC model constructed using the range generalised autoregressive conditional heteroskedasticity (R-GARCH) model by Molnár [39]. Fiszeder and Faldziński [43] suggested the DCC model formulated using the CARR model and the range-based estimator of covariance of returns. The model introduced by Fiszeder [42] was based on the BEKK model by Engle and Kroner [48] and the use of range-based estimators of variances and covariances of returns. All the above papers are methodologically different from the approach considered in this paper because we do not use any parametric range-based volatility models.

The paper has four contributions:

- First, we propose a new method for dynamic modelling and forecasting covariance matrices based on SVR. This approach guarantees the positive definiteness of the forecasted covariance matrices and is flexible, as it can be applied to different dependence patterns. At the beginning, we decompose the range-based covariance matrices of returns into the Cholesky factors, and then we forecast the univariate series of the entries of the Cholesky factors using the SVR model. Afterwards, we reconstruct the covariance matrix from these forecasts as a result of the reverse operation of the Cholesky decomposition. We use the range-based variances and covariances; the proposed approach, however, is quite general and can be applied to other proxies of covariance matrices formulated on the basis of daily data (e.g., squared returns and products of returns) or intraday data (e.g., realised variances and covariances).
- Second, we provide empirical evidence that the forecasts of both the whole covariance matrix and each single covariance obtained by the proposed procedure are more accurate than those obtained by the DCC model. This model is a natural benchmark because it is one of the most popular multivariate volatility models and can even be applied to very large portfolios. To the best of our knowledge, this is the first attempt in the literature to use SVR to forecast covariance matrices.
- Third, we demonstrate that the variance forecasts based on the proposed procedure are more precise than the forecasts from the univariate GARCH model [49]. It has already been shown in the literature that variance forecasts based on the SVR model can be more accurate than the forecasts calculated from the GARCH model (see, e.g., [15, 28, 31, 50]); however, range data have not been used in such applications so far.
- Fourth, we show that the forecasting advantage of the proposed method over the DCC model is higher during high market volatility and dependence between assets. This conclusion is important since such periods are often associated with market turmoil and high market uncertainty, i.e., when forecasting is the most difficult and accurate

forecasts matter most. Such a finding for the range-based estimators has not been formulated in the literature so far.

The rest of the paper is organised as follows. Section 2 provides an outline of the SVR model, describes the range-based covariance estimator and, most importantly, introduces the proposed method for covariance matrix forecasting. Section 3 presents the empirical research aimed at assessing the performance of the proposed procedure, the data applied, and a detailed description of the study and its results. Section 4 provides the conclusions.

2 Theoretical background

2.1 SVR model

Let us assume the following regression model:

$$y = r(\mathbf{x}) + \delta \quad (1)$$

where $r(\mathbf{x})$ is the regression function, y is the dependent variable, \mathbf{x} is the vector of regressors and δ is additive zero-mean noise with variance σ^2 . On the basis of a training dataset $\{(\mathbf{x}_t, y_t)\}_{t=1, \dots, T}$, we want to approximate the unknown regression function by a function $f(\mathbf{x})$ that has a deviation of at most ε from the outputs y_t and is as flat as possible [51]. In SVR, the input \mathbf{x} is first mapped onto a high-dimensional feature space using fixed (nonlinear) mapping, and then a linear model is constructed in this feature space:

$$f(\mathbf{x}) = \sum_{i=1}^d \omega_i \varphi_i(\mathbf{x}) + b \quad (2)$$

where d is the dimension of the feature space, $\varphi_i(\mathbf{x})$ denotes (nonlinear) transformations, ω_i are the coefficients and b is the bias term [52]. It should be noted that the dimension of the feature space determines the capacity of the SVR model to approximate a smooth input-output mapping; higher values of the dimension d lead to a more accurate approximation [15].

According to Eq. (2), to derive the function $f(\mathbf{x})$, one must estimate $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_d)'$ and b . To measure the estimation quality, Vapnik [1] proposed the ε -insensitive loss function:

$$L_\varepsilon(y, f(\mathbf{x})) = \begin{cases} 0, & |y - f(\mathbf{x})| \leq \varepsilon, \\ |y - f(\mathbf{x})| - \varepsilon, & \text{otherwise} \end{cases} \quad (3)$$

which means that errors below ε are not penalised. SVR performs linear regression in the d -dimensional feature space using the ε -insensitive loss function and, at the same time, tries to reduce model complexity by minimising $\|\boldsymbol{\omega}\|^2 = \boldsymbol{\omega}'\boldsymbol{\omega}$. The optimal regression function is given by the minimum of the functional:

$$\Phi(\boldsymbol{\omega}, \boldsymbol{\xi}) = \frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \sum_{t=1}^n (\xi_t + \xi_t^*) \quad (4)$$

where C is a pre-specified positive value and ξ_t and ξ_t^* are nonnegative slack variables representing the upper and lower constraints, respectively, on the outputs of the system; i.e.,

$$y_t - f(\mathbf{x}_t) \leq \varepsilon + \xi_t^* \quad (5)$$

$$f(\mathbf{x}_t) - y_t \leq \varepsilon + \xi_t \quad (6)$$

for all $t = 1, 2, \dots, T$. The parameter C controls the penalty imposed on observations that lie outside the ε -margin and, consequently, helps to prevent overfitting. Both the ε and the C parameters of SVR must be determined by the user.

The optimisation problem described above can be transformed into a dual problem SVR for which the solution is given by:

$$f(\mathbf{x}) = \sum_{t=1}^{T_{SV}} (\alpha_t - \alpha_t^*) K(\mathbf{x}_t, \mathbf{x}) \quad \text{s.t.} \quad 0 \leq \alpha_t \leq C, 0 \leq \alpha_t^* \leq C \quad (7)$$

where α_t and α_t^* are Lagrange multipliers, T_{SV} is the number of support vectors and K is the kernel function of the form:

$$K(\mathbf{x}_t, \mathbf{x}) = \sum_{i=1}^d \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}_t) \quad (8)$$

The dual problem can be solved more easily than the primal problem. The use of the kernel function prevents the need to explicitly compute the functional form of φ_i , which greatly reduces the computational complexity of the high-dimensional hidden space. Instead, the kernel function K computes the inner product of the vector $\boldsymbol{\varphi}(\mathbf{x}) = (\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x}), \dots, \varphi_d(\mathbf{x}))'$ [31].

In practice, the most popular kernel functions are the following:

- Linear (dot product): $K(\mathbf{x}_t, \mathbf{x}) = \mathbf{x}_t' \mathbf{x}$,
- Gaussian: $K(\mathbf{x}_t, \mathbf{x}) = \exp(-\|\mathbf{x}_t - \mathbf{x}\|^2)$,
- Polynomial: $K(\mathbf{x}_t, \mathbf{x}) = (1 + \mathbf{x}_t' \mathbf{x})^p$; $p = 2, 3, \dots$

The application of the linear kernel leads to linear SVR, while the Gaussian and polynomial kernels allow nonlinear SVR to be performed.

2.2 Range-based covariance estimator

We apply the estimator of covariance of returns calculated on the basis of low and high prices (see [53–55]). This estimator has an advantage over that based on only the closing prices because it uses information about the price changes during the day. It is given by:

$$\text{cov}(X, Y) = 0.5[\text{var}(X + Y) - \text{var}(X) - \text{var}(Y)] \quad (9)$$

where variances $\text{var}(X + Y)$, $\text{var}(X)$, $\text{var}(Y)$ are estimated using low and high prices.

The Parkinson [56] estimator of variance can be used to calculate all the variances in Eq. (9). It is expressed as:

$$\text{var}_{P_t} = [\ln(H_t/L_t)]^2 / (4\ln 2) \quad (10)$$

where H_t and L_t are the daily high and low prices, respectively. The Parkinson estimator was advocated by Brunetti and Lildholdt [53] and Brandt and Diebold [55]; however, other range-based variance estimators can also be applied.

Equation (9) can be applied when the range of the sum of variables X and Y is known, although, in practice, this range is not easy to calculate. It can be computed from tick-by-tick prices; however, such data are difficult to obtain for many financial assets. However, in the case of foreign exchange rates, the aforementioned range can be easily calculated. Let us consider two exchange rates of currencies x and y in terms of currency z , denoted by x/z and y/z , respectively. In the absence of triangular arbitrage opportunities, the return of the cross rate can be written as:

$$\Delta \ln x/y = \Delta \ln x/z - \Delta \ln y/z \quad (11)$$

Then, the range-based estimator of covariance for the currency pairs can be represented as:

$$\begin{aligned} \text{cov}(\Delta \ln x/z, \Delta \ln y/z) \\ = 0.5[\text{var}(\Delta \ln x/z) + \text{var}(\Delta \ln y/z) - \text{var}(\Delta \ln x/y)] \end{aligned} \quad (12)$$

This approach was used by some authors to analyse covariance of returns (see, e.g., [53, 57, 58]). Such an estimator was also employed for the construction of multivariate GARCH models by Fiszeder [42] for the BEKK model and by Fiszeder and Faldziński [43] for the DCC model.

The estimator of covariance based on low and high prices is less efficient than the most common estimator based on intraday prices, i.e., realised covariance, although it is more robust to microstructure noise, which makes the estimators biased [55]. Chou et al. [41] and Martens and van Dijk [59] showed how this bias of the range-based covariance estimator can be eliminated. Compared with the estimator calculated on the basis of closing prices, the estimator calculated on the basis of low and high prices is highly efficient. Monte Carlo simulations have indicated that when the Parkinson estimator is applied to all the variances in Eq. (9), the range-based covariance estimator is approximately five times more efficient (see, e.g., [53, 55]).

2.3 Forecasting the range-based covariance matrix using SVR

In this subsection, we introduce a methodology for dynamic modelling and forecasting of covariance matrices based on

SVR models using the Cholesky decomposition. The approach guarantees the positive definiteness of the forecasted covariance matrices and is flexible, as it can be applied to different dependence patterns.

It should be noted that the matrix constructed from the variance and covariance forecasts obtained from the disjoint application of the forecasting models is not guaranteed to be positive definite. In this paper, we apply the Cholesky decomposition to preserve the positive definiteness of the forecasted covariance matrices. The Cholesky decomposition, also known as the Cholesky factorisation, is a method of decomposing a symmetric positive-definite matrix A into the product of a unique upper triangular matrix U with real and positive diagonal entries and its conjugate transpose U' ; i.e., $A = U'U$. The matrix U is known as the Cholesky factor of A and can be interpreted as the square root of A . The motivation for modelling and forecasting the Cholesky factors instead of the elements of the range-based covariance matrix of returns directly is that we do not need to impose any restrictions. The idea of using the Cholesky factorisation in financial modelling is not new. For example, Tsay [60] applied it to re-parameterise the conditional covariance matrix of returns in the multivariate GARCH model. The idea of using the Cholesky decomposition of the realised covariance matrix in modelling and forecasting was put forward by Andersen et al. [58] and initially implemented in empirical studies by Chiriac and Voev [61].

To forecast the covariance matrix for the forecast horizon τ , we follow five steps:

- Step 1. We calculate the $N \times N$ range-based covariance matrices of returns G_t , $t = 1, 2, \dots, T$, where T is the time-series length. In general, the range-based variances of the returns are the diagonal entries of these matrices, while the range-based covariances based on Eq. (9) are the other entries. To estimate the range-based covariance matrices for the currency pairs, we use the estimator of the covariance of the returns given by Eq. (12) and the Parkinson estimator of the variance expressed in Eq. (10).
- Step 2. The matrices G_t ($t = 1, 2, \dots, T$) are decomposed using the Cholesky decomposition into the form $G_t = P'_t P_t$.
- Step 3. For each entry p_t^{ij} ($1 \leq i \leq j \leq N$) of the Cholesky factor P_t , we construct and train the autoregressive SVR model of the form:

$$p_{t+l}^{ij} = f(p_t^{ij}, p_{t-1}^{ij}, \dots, p_{t-l+1}^{ij}) \quad (13)$$

where l is the lag length. The SVR model (13) is estimated separately for each (i, j) based on univariate series p_t^{ij} ($t = 1, 2, \dots, T$).

- Step 4. We forecast the Cholesky factor $\mathbf{P}_{T+\tau} = [p_{T+\tau}^{ij}]$. To achieve this aim, the forecasts $\hat{p}_{T+\tau}^{ij}$ are calculated using the models trained in Step 3.
- Step 5. The forecast of the covariance matrix is reconstructed using the reverse operation of the Cholesky decomposition; i.e., $\hat{\mathbf{G}}_{T+\tau} = \hat{\mathbf{P}}'_{T+\tau} \hat{\mathbf{P}}_{T+\tau}$, where $\hat{\mathbf{P}}_{T+\tau}$ is the forecast of the Cholesky factor.

The outline of the proposed algorithm is depicted in Fig. 1.

2.4 Alternative forecasting methods

Modelling a covariance matrix is a challenging task for two reasons. First, the chosen model must guarantee the positive definiteness of the estimated and forecasted covariance matrices. Second, to limit the inflation of the number of estimated parameters and computational difficulties, severe restrictions on the model dynamics must be imposed. To model conditional covariance matrices, many methods have been

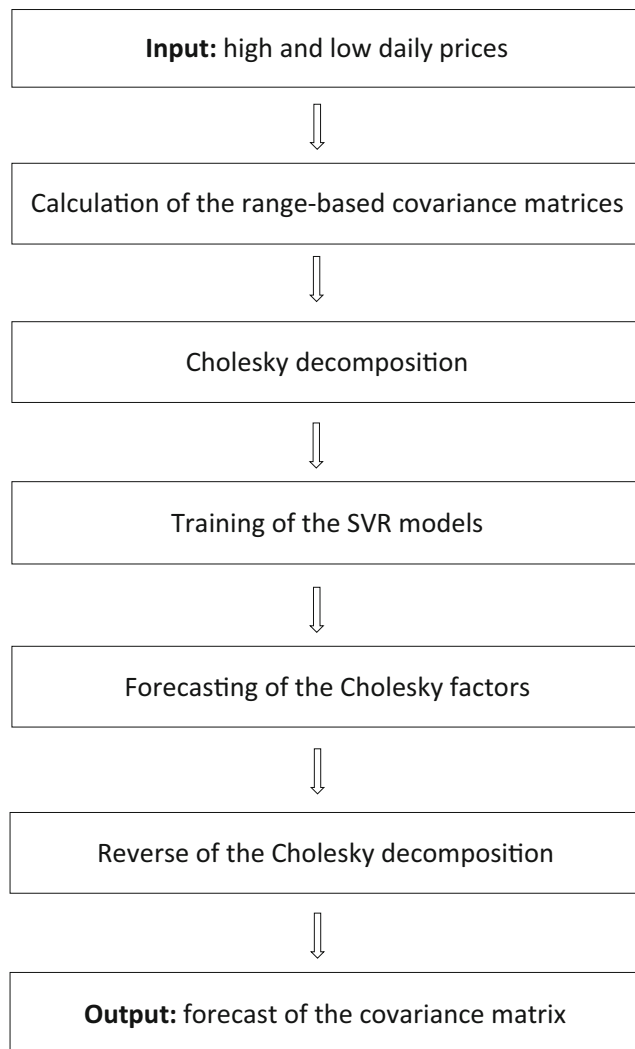


Fig. 1 Outline of the algorithm for covariance matrix forecasting

proposed in the literature. Two of the most popular approaches are (1) modelling and forecasting realised covariance matrices and (2) applying multivariate GARCH models (see, e.g., [62]). The first method relies on the usage of intraday data to calculate nonparametric measures of variances and covariances, such as realised variances and covariances. Models based on these measures provide more precise estimates and forecasts than models based on daily closing prices; however, high-frequency data are not commonly available and are significantly more expensive than daily data. In this paper, we apply daily low and high prices, which are usually available with closing prices but contain much more information about volatility and relationships between returns. The use of such prices also has more advantages than intraday data, such as wider availability, lower acquisition costs, considerably lower database requirements, and greater robustness to some micro-structure effects. Furthermore, the direct application of intraday data means some problems, such as the existence of daily cyclical fluctuations, the existence of strong autocorrelation or a significant impact of the publication of macroeconomic information on quoted prices. The goal of this study is to create a forecasting procedure based on daily data, which is why we do not consider models for realised covariance matrices.

The second most popular approach is to apply multivariate GARCH models. These are parametric models where, by definition, the structure of dependencies between variables is restricted to a specific analytical form. When dealing with large portfolios, many multivariate GARCH models present unsatisfactory performance or problems with estimation. Therefore, we choose for our forecasting study the dynamic conditional correlation (DCC) model by Engle [47], which has important advantages, such as the positive definiteness of the conditional covariance matrices and the ability to describe time-varying conditional correlations and covariances in a parsimonious way. Furthermore, the parameters of the DCC model can be estimated in two stages by the quasi-maximum likelihood method, which makes this approach relatively simple and possible to apply even to very large portfolios. The DCC model is one of the most popular multivariate GARCH models used to describe financial time series. Moreover, many papers, such as [63–66], show that it is very difficult for other multivariate GARCH models to outperform the DCC model.

There are also alternative methods to forecast conditional covariance matrices that are not included in the two above approaches. The method based on an SVR model that we propose in this paper is one of them. There are several benefits of applying SVR models to forecast time series. First, it preserves the common advantages of other machine learning methods. It is widely claimed that machine learning offers a more general approach than parametric models (cf. [67, 68]). It is capable of approximating nonlinear functions based on noisy and nonstationary data. Machine learning concentrates

on prediction by using general-purpose learning algorithms to find patterns in often rich and unwieldy data. This approach makes minimal assumptions about the data-generating systems; it can be effective even when the data are gathered without a carefully controlled experimental design and in the presence of complicated nonlinear interactions. Moreover, it can be particularly helpful when the number of input variables exceeds the number of subjects. It should also be noted that empirical studies in the literature are promising since they confirm that specific ML methods can outperform econometric models (cf. [69]). In particular, one can find first attempts to forecast conditional covariance matrices by applying ML methods [70, 71]. In both cited papers, artificial neural networks (ANNs) were used. However, in contrast to our method, both studies can be included in the first mentioned approach since realised covariance matrices are modelled there.

The most important ML forecast methods are artificial neural networks, support vector machines, random forests, the nearest neighbours algorithm, Bayesian regression, kernel ridge regression and generalised linear models. We decided to apply the SVR model due to its promising properties. It has been shown that SVR combines the training efficiency and simplicity of linear algorithms with the accuracy of the best nonlinear techniques. In many practical applications, this approach can tolerate high-dimensional or incomplete data and is robust to outliers [28, 72]. One of the main advantages of SVR is that its computational complexity does not depend on the dimensionality of the input space. Additionally, it has excellent generalisation capability and high prediction accuracy [73].

Many of the conducted studies have confirmed that SVR can make more accurate forecasts than other machine learning techniques, including ANNs (see, e.g., [28, 74, 75]; however, some authors have claimed that the superiority of SVR over ANN is not obvious and can depend on the type of neural network [76, 77]. Moreover, unlike neural network training, which requires nonlinear optimisation with the risk of falling into the local minima, the SVR solution is always unique and globally optimal [75]. It has also been shown that SVR has advantages over artificial neural networks for real-world data of limited size: (i) fewer calibration samples are required to obtain a desired model performance, (ii) SVR is less sensitive to sampling variations in small datasets and (iii) cross-validation is an approximately unbiased option for evaluating the true support vector regression model performance even for small datasets [78].

3 Application of the SVR models for forecasting exchange rates

In the empirical study, we applied the proposed methodology for dynamic modelling and forecasting covariance matrices

based on SVR models to exchange rates from the forex market. We assessed the accuracy of several SVR models with different lags and kernels and compared this approach with the DCC model.

3.1 Data applied

The three most heavily traded currency pairs in the forex market, namely, EUR/USD, USD/JPY and GBP/USD, were investigated. The analysis of currencies uses triangular arbitrage to calculate the covariance of the returns.

First, we evaluated the data for the 11-year period from 2 January 2006 to 30 December 2016 (2853 returns). The descriptive statistics for the returns, squared returns and products of the returns are presented in Table 1. The daily returns were calculated as $r_t = 100 \ln(p_t/p_{t-1})$, where p_t is the closing price at time t .

The variability of the returns, measured by the standard deviation, was quite similar for all the currency pairs; however, there were significant differences in the skewness and kurtosis of the distributions. Owing to perturbations caused by the 2016 Brexit referendum, the distribution of returns was more leptokurtic, and the minimum return was significantly lower for GBP/USD than for the remaining pairs. A weak autocorrelation was present in the returns of the GBP/USD rate. The autocorrelation of the squared returns and products of the returns was much stronger than the autocorrelation of the returns. Moreover, there were considerably higher deviations from the normal distribution for the squared returns and the products of the returns. This finding means that modelling the covariance matrices of the returns is a much more demanding task than modelling the returns.

3.2 Description of the models and procedures

We applied autoregressive SVR models (cf. Eq. (13)) with different lags and two kernels: a linear kernel (which leads to linear SVR) and a Gaussian kernel (which leads to nonlinear SVR). We also applied several lag values; however, we present only the results for lags $l = 1$ and $l = 15$; lag $l = 1$ leads to the simplest specification, in which only one lagged variable, i.e., p_T^j , is used as the regressor. Our calculations showed that larger lags may lead to more accurate forecasts; however, this effect ceased to be visible for $l > 15$. According to the results, lag $l = 15$ seemed to be optimal when considering the accuracy of the forecasts and the computation time.

Finally, we considered four specifications of the SVR models:

- 1) SVR with a linear kernel and $l = 1$ (hereinafter SVR_lin_1),

Table 1 Summary statistics of the daily currency pairs

Currency pairs	Mean $\times 10^2$	Min	Max	SD	Skew	Kurt	LB(10)
Returns							
EUR/USD	-0.417	-2.554	3.503	0.623	0.088	2.035	6.655
JPY/USD	0.026	-5.448	3.779	0.670	0.051	4.549	12.592
GBP/USD	-1.173	-8.322	2.870	0.612	-1.198	14.041	25.432*
Squared returns							
EUR/USD	38.796	0.000	12.272	0.779	5.351	44.756	542.085
JPY/USD	44.811	0.000	29.675	1.147	10.059	179.258	263.955
GBP/USD	37.461	0.000	69.252	1.504	34.598	1543.373	190.737
Products of returns							
EUR/USD-JPY/USD	8.886	-9.653	8.545	0.654	-1.154	64.028	201.952
EUR/USD-GBP/USD	23.235	-2.157	21.255	0.677	12.777	336.127	245.532
JPY/USD-GBP/USD	1.815	-31.449	8.555	0.866	-18.536	636.843	99.220

Mean is the arithmetic mean, Min is the minimum, Max is the maximum, SD is the standard deviation, Skew is the skewness, Kurt is the excess kurtosis, and LB(10) is the Ljung-Box statistic for 10 lags

- 2) SVR with a linear kernel and $l = 15$ (hereinafter SVR_lin_15),
- 3) SVR with a Gaussian kernel and $l = 1$ (hereinafter SVR_Gauss_1),
- 4) SVR with a Gaussian kernel and $l = 15$ (hereinafter SVR_Gauss_15).

All the above models were repeatedly constructed on the basis of a rolling sample with a fixed size of $T = 527$ (i.e., the number of observations in the first 2-year period, from 3 January 2006 to 31 December 2007, which was used as the initial sample). In the case when the range-based covariance matrices of the returns \mathbf{G}_t were not positive definite, a simple method based on eigenvalues was applied (see, e.g., [79, 80]); however, this procedure did not significantly affect the dynamic dependencies between the covariance matrices. According to the methodology described in Subsection 2.3, the SVR models were constructed for the series p_t^{ij} ($t = 1, 2, \dots, T$) obtained from the Cholesky decomposition. The decomposed series were standardised, i.e., centred by subtracting the mean and divided by the standard deviation.

As described in Subsection 2.1, the values of the ε and C parameters (also called meta-parameters) must be determined to create the SVR models. There are competitive propositions in the literature on how to tune these parameters (see, e.g., [50, 52, 81–83]; and references therein); however, previous studies did not demonstrate a clear superiority for any of them. Therefore, we applied two tuning methods in our study. In the first method, we determined the parameters using the default settings in MATLAB (in our study, we used the function *fitrsvm* in MATLAB (R2015b) to perform SVR); i.e., $C = 1$ for

the linear kernel and $C = \frac{lqr(Y)}{1.349}$ for the Gaussian kernel, where $lqr(Y)$ is the interquartile range of the response variable Y . For both kernels, the default value of the ε parameter was $\varepsilon = \frac{lqr(Y)}{13.49}$. The main advantage of this method is its simplicity and time effectiveness. The second method we applied was the grid search technique. This method consists of constructing many SVR models for different values of the parameters and selecting the optimal model on the basis of a validation set. We performed a grid search for the C and ε parameters by considering consecutive values of $C = 2^{-5}, 2^{-4}, \dots, 2^4, 2^5$ and $\varepsilon = 2^{-5}, 2^{-4}, \dots, 2^4, 2^5$. To select the optimal parameters, we applied a 10-fold cross-validation procedure. According to this approach, the investigated sample was randomly partitioned into 10 equal-sized subsamples. Nine of these subsamples were used to construct the SVR model, while the remaining one was used to validate the model. To this end, the mean squared error (MSE) was computed on the observations in the validation subsample. This procedure was repeated 10 times (for each of the 10 subsamples used as the test set), and the average of the 10 obtained MSEs was calculated. Finally, the parameters that led to the smallest MSE were considered optimal. It should be noted that this approach can be very time consuming. However, this problem can be avoided because it is reasonable to assume that the optimal parameters for consecutive rolling samples should be very similar (as these samples differ only in 1 of 527 observations), which means that there is no need to determine these parameters for each sample. In our study, we decided to perform the grid search technique to tune the parameters every 100 days.

Our study results showed that the grid search method produced better values than the default values in MATLAB

(according to the in-sample *MSE* calculated by the cross-validation and to the out-of-sample prediction errors). Therefore, in the next subsection, we will present only the results from the SVR models with parameters tuned using the grid search technique.

3.3 Forecasting performance

Based on all the considered models, 1-day-ahead forecasts of covariance matrices – $\hat{\mathbf{G}}_{T+\tau}$ (i.e., $\tau = 1$) for the 9-year period from 2 January 2008 to 30 December 2016 were calculated (i.e., 2336 forecasts). The considered period was relatively long, and it covered both turbulent periods, such as the global financial crisis of 2008, European sovereign debt crisis and 2016 Brexit referendum, and tranquil periods; therefore, the results should be robust to the state of the global economy.

As a proxy of the daily covariance for the evaluation of forecasts, the sum of products of intraday returns (the realised covariance) was employed, while as a proxy for the daily variance, the sum of squared intraday returns (the realised variance) was used. This is a commonly accepted approach in the literature (see, e.g., [84–86]). One major problem of using such data is the choice of the appropriate frequency of observations (see, e.g., [87]). In this study, 15-min returns were applied; however, the main results did not change for the 5- or 30-min returns. It should be noted that we used intraday data only to evaluate the forecasts; we did not use them to construct the models or to calculate the forecasts.

First, following Laurent et al. [88], we evaluated the forecasts of the whole conditional covariance matrix using the squared Frobenius loss function given by:

$$LF = (1/m) \sum_{t=1}^m \text{Tr} \left[\left(\hat{\mathbf{G}}_t - \mathbf{G}_{tR} \right)' \left(\hat{\mathbf{G}}_t - \mathbf{G}_{tR} \right) \right] \quad (14)$$

where m is the number of forecasts, Tr is the trace of a matrix, $\hat{\mathbf{G}}_t$ is a forecast of the conditional covariance matrix and \mathbf{G}_{tR} is the realised covariance matrix (with the realised variances on the diagonal and the realised covariances on the off-diagonal).

The proposed method of covariance matrix forecasting using SVR models utilises the Cholesky decomposition. Unfortunately, the forecasting performance of the procedure may be sensitive to the order of the variables in the covariance matrix because each permutation of the elements in the original matrix yields a different decomposition and different factors. Therefore, we considered all possible permutations of the analysed currency pairs (permutation 1: EUR/USD, USD/JPY and GBP/USD; permutation 2: EUR/USD, GBP/USD and USD/JPY; permutation 3: USD/JPY, EUR/USD and GBP/USD; permutation 4: USD/JPY, GBP/USD and EUR/USD; permutation 5: GBP/USD, EUR/USD and USD/JPY; and permutation 6: GBP/USD, USD/JPY and EUR/USD).

To evaluate the statistical significance of the results, we applied the model confidence set (MCS) of Hansen et al. [89]. The objective of the MCS procedure is to determine the set of models that consists of the best model(s) from a given collection of models. The best models are selected with a given level of confidence in terms of a user-specified criterion. In our analysis, a criterion based on the squared Frobenius loss function was applied. The values of the squared Frobenius loss function and the corresponding p -values (calculated by the bootstrap method) of the MCS test are presented in Table 2. It can be seen in the table that the forecasts from the SVR_lin_15 model were significantly more accurate than those from the other SVR models and the DCC model for all permutations.¹ This finding means that the order of the variables in the covariance matrix did not affect the forecasting superiority of the SVR_lin_15 model.

The results of the analysis of the whole covariance matrix did not show whether the superiority of one model was due to more accurate forecasts of the variances, covariances or both. This question is important since, in some applications, only the volatility of financial processes is used, and in other cases, the relationship between processes plays a key role. Therefore, from a practical point of view, it is advisable to analyse the forecasts of the variances and covariances separately. To this end, the mean squared error (*MSE*), mean absolute error (*MAE*) and coefficient of determination from the Mincer-Zarnowitz regression were calculated. These criteria are often used to evaluate volatility forecasts in empirical studies (see, e.g., [90, 91]). We also tried other loss functions, and they yielded similar results. The statistical significance of the results was verified again by the MCS test. To save space, we present only the results for permutation 2, which was the worst permutation according to the squared Frobenius loss function for the SVR_lin_15 model. The other permutations also led to more accurate forecasts for the separate series of the variances and covariances. The results for the forecasts of the covariances are presented in Table 3.

According to the results of the MCS test, only the SVR_lin_15 model belonged to the MCS. The forecasting superiority of the SVR_lin_15 model did not depend on the type of loss function; all the considered criteria indicated that the covariance forecasts based on this model were the most accurate.

We also compared the forecasts of the variances. To this end, we applied the GARCH models, which were previously used in the DCC model. The obtained results are presented in Table 4.

The results for variance forecasting were not unequivocal but also indicated the advantage of the SVR_lin_15 model.

¹ This model was also the best in terms of the in-sample *MSE* calculated by the cross-validation procedure. This finding proves that the cross-validation method can be effectively used to construct the best SVR model.

Table 2 Evaluation results of the covariance matrix forecasts based on the squared Frobenius loss function

Method of forecasting	Permutation 6		Permutation 4		Permutation 1		Permutation 3		Permutation 5		Permutation 2	
	<i>LF</i>	<i>p</i> value	<i>LF</i>	<i>p</i> value	<i>LF</i>	<i>p</i> value	<i>LF</i>	<i>p</i> value	<i>LF</i>	<i>p</i> value	<i>LF</i>	<i>p</i> value
DCC model	2.870	0.041	2.870	0.048	2.870	0.061	2.870	0.064	2.870	0.094	2.870	0.094
SVR_lin_1	2.921	0.000	2.918	0.000	2.932	0.000	2.933	0.000	2.948	0.000	2.952	0.000
SVR_lin_15	2.615	1.000*	2.619	1.000*	2.633	1.000*	2.636	1.000*	2.658	1.000*	2.659	1.000*
SVR_Gauss_1	2.964	0.000	3.058	0.000	3.039	0.000	3.057	0.000	2.960	0.000	3.043	0.000
SVR_Gauss_15	3.282	0.000	3.337	0.000	3.306	0.000	3.341	0.000	2.960	0.000	3.335	0.000

LF is the squared Frobenius loss function for the conditional covariance matrix, the *p* values are for the MCS test and ‘*’ indicates that the model belongs to the MCS, with a confidence level of 0.90. The evaluation period was from 2 January 2008 to 30 December 2016; the realised covariance and the realised variance, which were used in the evaluation of the forecasts, were based on 15-min returns. The permutations were ordered according to the *LF* for the SVR_lin_15 model

Based on the *MSE* measure, the forecasts from this model were the most accurate for all the series, with the single exception of the JPY/USD pair (when only the GARCH model was included in the MCS). Additionally, for the EUR/USD pair, two models (the GARCH model and the SVR_lin_15 model) belonged to the MCS, and there was no evidence to

reject the null hypothesis of equal predictive ability for these models. However, under the *MAE* criterion, which is less sensitive to outliers than the *MSE* measure, the SVR_lin_15 model was the best forecasting model for all the currency pairs. The superiority of the SVR_lin_15 model was also confirmed by the R^2 criterion.

Table 3 Evaluation results of the covariance forecasts based on the *MSE*, *MAE* and R^2 criteria

Method of forecasting	Forecast evaluation criteria				
	<i>MSE</i>	<i>p</i> value	<i>MAE</i>	<i>p</i> value	R^2
EUR/USD-GBP/USD					
DCC model	0.123	0.096	0.131	0.000	0.344
SVR_lin_1	0.150	0.000	0.149	0.000	0.231
SVR_lin_15	0.116	1.000*	0.121	1.000*	0.398
SVR_Gauss_1	0.144	0.000	0.145	0.000	0.247
SVR_Gauss_15	0.180	0.000	0.167	0.000	0.054
EUR/USD-JPY/USD					
DCC model	0.097	0.000	0.140	0.000	0.163
SVR_lin_1	0.096	0.001	0.135	0.000	0.165
SVR_lin_15	0.084	1.000*	0.119	1.000*	0.275
SVR_Gauss_1	0.106	0.000	0.141	0.000	0.077
SVR_Gauss_15	0.113	0.000	0.148	0.000	0.013
JPY/USD-GBP/USD					
DCC model	0.220	0.003	0.119	0.000	0.058
SVR_lin_1	0.206	0.007	0.113	0.000	0.104
SVR_lin_15	0.195	1.000*	0.102	1.000*	0.140
SVR_Gauss_1	0.212	0.007	0.116	0.000	0.063
SVR_Gauss_15	0.225	0.002	0.126	0.000	0.003

The *p* values are for the MCS test; ‘*’ indicates that the model belongs to the MCS, with a confidence level of 0.90; and R^2 is the coefficient of determination from the Mincer-Zarnowitz regression. The evaluation period was 2 January 2008 to 30 December 2016; the realised covariance and the realised variance, which were used in the evaluation of the forecasts, were based on 15-min returns

Table 4 Evaluation results of the variance forecasts based on the *MSE*, *MAE* and R^2 criteria

Method of forecasting	Forecast evaluation criteria				
	<i>MSE</i>	<i>p</i> value	<i>MAE</i>	<i>p</i> value	R^2
EUR/USD					
GARCH model	0.168	0.412*	0.206	0.000	0.367
SVR_lin_1	0.203	0.000	0.222	0.000	0.269
SVR_lin_15	0.164	1.000*	0.185	1.000*	0.410
SVR_Gauss_1	0.209	0.000	0.226	0.000	0.241
SVR_Gauss_15	0.264	0.000	0.258	0.000	0.047
GBP/USD					
GARCH model	1.190	0.013	0.230	0.010	0.097
SVR_lin_1	1.091	0.013	0.245	0.000	0.133
SVR_lin_15	0.999	1.000*	0.198	1.000*	0.192
SVR_Gauss_1	1.087	0.013	0.239	0.000	0.120
SVR_Gauss_15	1.195	0.003	0.271	0.000	0.028
JPY/USD					
GARCH model	0.631	1.000*	0.313	0.019	0.193
SVR_lin_1	0.751	0.003	0.318	0.000	0.180
SVR_lin_15	0.706	0.010	0.295	1.000*	0.226
SVR_Gauss_1	0.822	0.003	0.327	0.000	0.043
SVR_Gauss_15	0.839	0.002	0.337	0.000	0.031

The *p* values are for the MCS test; ‘*’ indicates that the model belongs to the MCS, with a confidence level of 0.90; and R^2 is the coefficient of determination from the Mincer-Zarnowitz regression. The evaluation period was from 2 January 2008 to 30 December 2016; the realised covariance and the realised variance, which were used in the evaluation of the forecasts, were based on 15-min returns

For both the covariances and variances, the differences between the values of the *MAE* amongst the different currency pairs were much smaller than those in the case of the *MSE* and *R*² criteria. These differences were associated with the existence of numerous outliers in the currency pairs returns, which have a much stronger impact on the latter measures.

Our research showed the superiority of the linear kernel over the nonlinear (Gaussian) kernel, which means that the autoregressive relations in each forecasted series were linear or almost linear. It can be concluded that the applied linear SVR models have a form similar to that of the ARCH model; however, it should be noted that they are not applied directly to the raw series of the variances and covariances but to the series transformed by the Cholesky decomposition. By the same analogy, one can easily explain the superiority of the SVR_lin_15 model over the SVR_lin_1 model. Many empirical studies have shown that the conditional variance (covariance) usually appears to be a function of many lagged past squared errors (products of errors), which is why a more parsimonious parametrisation in the form of the GARCH model is frequently used. Additionally, it should be noted that our calculations showed that, in the case of linear SVR, longer lags led to more accurate forecasts; however, this effect ceased to be clearly visible for $l > 15$.

3.4 Influence of market conditions on the superiority of forecasts

Two recent studies suggest that the application of low and high prices in volatility models leads to the largest improvement in the estimation and forecasting of volatility during turbulent periods [38, 39]. For this reason, in this section, we examine whether market conditions, i.e., market volatility and dependencies between assets, can affect the accuracy of the proposed forecasting procedure based on the SVR model.

For this aim, we applied a quantile regression model and tested whether extreme improvements in forecasts can be explained by the level of market volatility (for variance forecasts) or dependence (for covariance forecasts) on the previous day. Let $d_{\text{var},T}$ and $d_{\text{cov},T}$ denote loss differentials defined for the variance and covariance forecasts, respectively, as:

$$d_{\text{var},T} = (\text{var}_{DCC,T} - \text{var}_{R,T})^2 - (\text{var}_{SVR_lin_15,T} - \text{var}_{R,T})^2 \quad (15)$$

$$d_{\text{cov},T} = (|\text{cov}_{DCC,T}| - |\text{cov}_{R,T}|)^2 - (|\text{cov}_{SVR_lin_15,T}| - |\text{cov}_{R,T}|)^2 \quad (16)$$

where $\text{var}_{DCC,T}$ and $\text{var}_{SVR_lin_15,T}$ are the forecasts of the conditional variances based on the DCC and SVR_lin_15 models, respectively, $\text{var}_{R,T}$ is the realised variance; $\text{cov}_{DCC,T}$ and $\text{cov}_{SVR_lin_15,T}$ are the forecasts of the conditional covariances based on the DCC and SVR_lin_15 models, respectively, and $\text{cov}_{R,T}$ is the realised covariance.

When $d_{\text{var},T}$ and $d_{\text{cov},T}$ are positive, then the forecasts based on the SVR_lin_15 model are more accurate than the forecasts from the DCC model. The loss differentials described in Eqs. (15)–(16) are based on the MSE loss function, but very similar formulas can be written for the MAE loss function.

The τ -th linear quantile regression model can be specified for the variance and covariance as follows:

$$d_{\text{var},T} = \beta_0(\tau) + \beta_1(\tau) \text{var}_{R,T-1} + \varepsilon_{\text{var},T}(\tau) \quad (17)$$

$$d_{\text{cov},T} = \alpha_0(\tau) + \alpha_1(\tau) |\text{cov}_{R,T-1}| + \varepsilon_{\text{cov},T}(\tau) \quad (18)$$

The parameter estimation results for the above quantile regression models are presented in Tables 5 and 6. We report the results based on the 90th percentile because we are interested in analysing large forecast improvements; however, very similar results were achieved for other high quantiles (e.g., the 75th and 95th percentiles).

The estimates of the coefficient β_1 were positive and significant for all the currency pairs, which means that higher forecast improvement of the SVR_lin_15 model over the DCC model was observed when the realised variance was large. This conclusion is important since high market volatility is associated with turbulent periods and high market uncertainty, i.e., when forecasting is the most difficult and accurate forecasts matter most.

The results reported in Table 5 show that the estimates of the coefficient of interest α_1 were positive and highly significant for all the analysed relations, which means that the forecasting advantage of the SVR_lin_15 model over the DCC model was higher when the dependence between currency pairs was large. This conclusion is also important and, in particular, confirms the previous results concerning the impact of the realised variance, since strong relations between assets often exist during turbulent periods (see, e.g., [92]). Such a finding for the range-based estimators has not been formulated in the literature so far.

We have presented the results only for the loss differentials based on the MSE loss function; however, the main results did not change for the MAE loss function.

Table 5 The parameter estimation results for the 90th percentile regression of the loss differential $d_{\text{var},T}$ on the lagged realised variance

Currency pairs	β_0	$s(\beta_0)$	p value	β_1	$s(\beta_1)$	p value
EUR/USD	0.002	0.006	0.778	0.187	0.026	0.000
GBP/USD	−0.035	0.029	0.226	0.499	0.211	0.018
JPY/USD	0.011	0.013	0.411	0.405	0.049	0.000

White's heteroskedastic-consistent standard errors are used. The evaluation period was 2 January 2008 to 30 December 2016; the realised variance was calculated based on 15-min returns

Table 6 The parameter estimation results for the 90th percentile regression of the loss differential $d_{\text{cov},T}$ on the lagged realised covariance

Currency pairs relations	α_0	$s(\alpha_0)$	p value	α_1	$s(\alpha_1)$	p value
EUR/USD-JPY/USD	0.002	0.001	0.213	0.279	0.057	0.000
EUR/USD-GBP/USD	-0.001	0.001	0.382	0.108	0.017	0.000
JPY/USD-GBP/USD	-0.003	0.003	0.277	0.589	0.086	0.000

White's heteroskedastic-consistent standard errors are used. The evaluation period was 2 January 2008 to 30 December 2016; the realised covariance was calculated based on 15-min returns

4 Conclusions

We have proposed a methodology for dynamic modelling and forecasting covariance matrices based on SVR, which is our main contribution. The procedure guarantees the positive definiteness of the forecasted covariance matrices and is flexible, as it can be applied to different dependence patterns. The range-based covariance matrices of returns are decomposed into the Cholesky factors, and then SVR models are applied to forecast the elements of these factors. Afterwards, the forecast of the covariance matrix of returns is reconstructed from the forecasts of these elements as a result of the reverse operation of the Cholesky decomposition.

The procedure is based on the decomposed range-based covariance matrices; however, the proposed approach is quite general and can be applied to other proxies of the covariance matrices formulated on the basis of daily data (e.g., squared returns and products of returns) or intraday data (e.g., realised variances and covariances).

The proposed procedure was applied to analyse the most heavily traded currency pairs in the forex market: EUR/USD, USD/JPY and GBP/USD. Our second primary contribution was to show that the forecasts of each separate covariance and the whole covariance matrix obtained by the SVR models were more accurate than those obtained by the competing benchmark multivariate GARCH model. Moreover, the variance forecasts based on the proposed procedure were more precise than those from the univariate GARCH model. It should be emphasised that the advantage of the suggested procedure was higher during turbulent periods, i.e., when forecasting is the most difficult and accurate forecasts matter most. Furthermore, we showed that the order of the variables in the covariance matrix, which yields different Cholesky decomposition results, did not affect the forecasting superiority of the SVR model. The main conclusions of the study were also robust to the forecast evaluation criterion employed.

We applied the DCC model as a benchmark because it is one of the most popular multivariate volatility models; moreover, the estimation of its parameters is relatively simple, and it is possible to apply it even to very large portfolios. The comparison can be performed with other more or less complex models. However, the search for such models was not the purpose of this investigation. Similarly, other variants of the

SVR models or even other machine learning techniques can also be considered in future research. The procedure proposed in this paper was an effective approach to forecast the covariance matrices; however, there are some potential directions on how to improve it. For example, one can apply other kernel functions or other methods for tuning the meta-parameters. These issues were not the primary objective of this work but can be investigated in future studies.

Acknowledgements The authors received support from the National Science Centre (project no. 2016/21/B/HS4/00662 (Piotr Fiszeder) and project no. 2019/35/B/HS4/00642 (Witold Orzeszko)).

The authors would like to thank three anonymous reviewers for helpful and constructive comments.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Vapnik VN (1995) The nature of statistical learning theory. Springer, Verlag
2. Jayadeva KR, Chandra S (2007) Twin support vector machines for pattern classification. *IEEE Trans Pattern Anal Mach Intell* 29:905–910
3. Kao L-J, Chiu C-C, Lu C-J, Yang J-L (2013) Integration of nonlinear independent component analysis and support vector regression for stock price forecasting. *Neurocomputing* 99:534–542
4. Shao Y, Chen W, Deng N (2014) Nonparallel hyperplane support vector machine for binary classification problems. *Inf Sci* 263:22–35
5. Zhan L, Li C (2017) A hybrid PSO-SVM-based method for predicting the friction coefficient between aircraft tire and coating. *Meas Sci Technol* 28(2):025004
6. Wang Q, Tian Y, Liu D (2019) Adaptive FH-SVM for imbalanced classification. *IEEE Access* 7:130410–130422
7. Zhao J, Xu Y, Fujita H (2019) An improved non-parallel Universum support vector machine and its safe sample screening rule. *Knowl-Based Syst* 170:79–88

8. Vapnik VN, Golowich S, Smola A (1997) Support vector method for function approximation, regression estimation, and signal processing. In: Mozer M, Jordan M, Petsche T (eds) *Advances in neural information processing systems 9*. MIT Press, Cambridge, pp 281–287
9. Matilla-García M (2007) Nonlinear dynamics in energy futures. *Energy J* 28:7–29
10. Lim K-P, Brooks RD, Hinich MJ (2008) Nonlinear serial dependence and the weak-form efficiency of Asian emerging stock markets. *J Int Finan Markets Inst Money* 18:527–544
11. Orzeszko W (2008) The new method of measuring the effects of noise reduction in chaotic data. *Chaos, Solitons Fractals* 38:1355–1368
12. Sadique MS (2011) Testing for neglected nonlinearity in weekly foreign exchange rates. *Rev Econ Finan* 3:77–88
13. Wey MA (2018) Nonlinear dynamics of U.S. equity factor portfolios. *Chaos* 28:113109
14. Aliyev F (2019) Testing market efficiency with nonlinear methods: evidence from Borsa Istanbul. *Int J Finan Stud* 7:27–38
15. Chen S, Härdle WK, Jeong K (2010) Forecasting volatility with support vector machine-based GARCH model. *J Forecast* 29:406–433
16. Chen Y-S, Cheng C-H, Tsai W-L (2014) Modeling fitting-function-based fuzzy time series patterns for evolving stock index forecasting. *Appl Intell* 41:327–347
17. Patel J, Shah S, Thakkar P, Kotecha K (2015) Predicting stock market index using fusion of machine learning techniques. *Expert Syst Appl* 42:2162–2172
18. Qu H, Zhang Y (2016) A new kernel of support vector regression for forecasting high-frequency stock returns. *Math Probl Eng* 1:1–9
19. Simian D, Stoica F, Bărbulescu A (2020) Automatic optimized support vector regression for financial data prediction. *Neural Comput & Applic* 32:2383–2396
20. Khemchandani R, Chandra S (2009) Regularized least squares fuzzy support vector regression for financial time series forecasting. *Expert Syst Appl* 36:132–138
21. Henrique BM, Sobreiro VA, Kimura H (2018) Stock price prediction using support vector regression on daily and up to the minute prices. *J Finan Data Sci* 4:183–201
22. Zhang J, Teng Y-F, Chen W (2019) Support vector regression with modified firefly algorithm for stock price forecasting. *Appl Intell* 49:1658–1674
23. Xu Y, Yang C, Peng S, Nojima Y (2020) A hybrid two-stage financial stock forecasting algorithm based on clustering and ensemble learning. *Appl Intell* 50:3852–3867
24. Psaradellis I, Sermpinis G (2016) Modelling and trading the U.S. implied volatility indices. Evidence from the VIX, VXN and VXD indices. *Int J Forecast* 32:1268–1283
25. Cao LJ, Tay FEH (2003) Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Trans Neural Netw* 14:1506–1518
26. Huang CL, Tsai CY (2009) A hybrid SOM-SVR with a filterbased feature selection for stock market forecasting. *Expert Syst Appl* 36:1529–1539
27. Faldziński M, Fiszeder P, Orzeszko W (2021) Forecasting volatility of energy commodities: comparison of GARCH models with support vector regression. *Energies* 14:1–18
28. Gavrishchaka VV, Ganguli SB (2003) Volatility forecasting from multiscale and high-dimensional market data. *Neurocomputing* 55:285–305
29. Ni H, Yin H (2009) Exchange rate prediction using hybrid neural networks and trading indicators. *Neurocomputing* 72:2815–2823
30. Sermpinis G, Stasinakis C, Theofilatos K (2015) Modeling, forecasting and trading the EUR exchange rates with hybrid rolling genetic algorithms – support vector regression forecast combinations. *Eur J Oper Res* 247:831–846
31. Peng Y, Albuquerque PH, Camboim de Sá JM, Padula AJA, Montenegro MR (2018) The best of two worlds: forecasting high frequency volatility for cryptocurrencies and traditional currencies with support vector regression. *Expert Syst Appl* 97:177–192
32. Nayak RK, Mishra D, Rath AK (2019) An optimized SVM-k-NN currency exchange forecasting model for Indian currency market. *Neural Comput & Applic* 31:2995–3021
33. Sermpinis G, Stasinakis C, Rosillo R, de la Fuente D (2017) European exchange trading funds trading with locally weighted support vector regression. *Eur J Oper Res* 258:372–384
34. Sermpinis G, Stasinakis C, Hassanniakalager A (2017) Reverse adaptive krill herd locally weighted support vector regression for forecasting and trading exchange traded funds. *Eur J Oper Res* 263:540–558
35. Nazemi A, Heidenreich K, Fabozzi FJ (2018) Improving corporate bond recovery rate prediction using multi-factor support vector regressions. *Eur J Oper Res* 271:664–675
36. Chou RY (2005) Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. *J Money Credit Bank* 37:561–582
37. Chan JSK, Lam CPY, Yu PLH, Choy STB, Chen CWS (2012) A Bayesian conditional autoregressive geometric process model for range data. *Comput Stat Data Anal* 56:3006–3019
38. Fiszeder P, Perczak G (2016) Low and high prices can improve volatility forecasts during the turmoil period. *Int J Forecast* 32:398–410
39. Molnár P (2016) High-low range in GARCH models of stock return volatility. *Appl Econ* 48:4977–4991
40. Wu X, Hou X (2020) Forecasting volatility with component conditional autoregressive range model. *N Am J Econ Financ* 51:101078
41. Chou RY, Wu CC, Liu N (2009) Forecasting time-varying covariance with a range-based dynamic conditional correlation model. *Rev Quant Finan Acc* 33:327–345
42. Fiszeder P (2018) Low and high prices can improve covariance forecasts: the evidence based on currency rates. *J Forecast* 37:641–649
43. Fiszeder P, Faldziński M (2019) Improving forecasts with the co-range dynamic conditional correlation model. *J Econ Dyn Control* 108:103736
44. Fiszeder P, Faldziński M, Molnár P (2019) Range-based DCC models for covariance and value-at-risk forecasting. *J Empir Financ* 54:58–76
45. Chen CWS, Gerlach R, Hwang BBK, McAleer M (2012) Forecasting value-at-risk using nonlinear regression quantiles and the intra-day range. *Int J Forecast* 28:557–574
46. Meng X, Taylor JW (2020) Estimating value-at-risk and expected shortfall using the intraday low and range data. *Eur J Oper Res* 280:191–202
47. Engle RF (2002) Dynamic conditional correlation – a simple class of multivariate GARCH models. *J Bus Econ Stat* 20:339–350
48. Engle RF, Kroner KF (1995) Multivariate simultaneous generalized ARCH. *Economic Theory* 11:122–150
49. Bollerslev T (1986) Generalised autoregressive conditional heteroscedasticity. *J Econ* 31:307–327

50. Santamaría-Bonfil G, Frausto-Solís J, Vázquez-Rodarte I (2015) Volatility forecasting using support vector regression and a hybrid genetic algorithm. *Comput Econ* 45:111–133
51. Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14:199–222
52. Cherkassky V, Ma Y (2004) Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Netw* 17:113–126
53. Brunetti C, Lildholdt PM (2002) Return-based and range-based (co)variance estimation, with an application to foreign exchange markets. SSRN. <https://doi.org/10.2139/ssrn.296875>
54. Fernandes M, Mota B, Rocha G (2005) A multivariate conditional autoregressive range model. *Econ Lett* 86:435–440
55. Brandt MW, Diebold FX (2006) A no-arbitrage approach to range-based estimation of return covariances and correlations. *J Bus* 79: 61–73
56. Parkinson M (1980) The extreme value method for estimating the variance of the rate of return. *J Bus* 53:61–65
57. Lopez JA, Walter CA (2001) Evaluating covariance matrix forecasts in a value-at-risk framework. *J Risk* 3:69–97
58. Andersen TG, Bollerslev T, Diebold FX, Labys P (2003) Modeling and forecasting realized volatility. *Econometrica* 71:579–625
59. Martens M, van Dijk D (2007) Measuring volatility with the realized range. *J Econ* 138:181–207
60. Tsay RS (2002) Analysis of financial time series. Wiley, New York
61. Chiriac R, Voev V (2011) Modelling and forecasting multivariate realized volatility. *J Appl Econ* 26:922–947
62. Bauwens L, Hafner CM, Laurent S (eds) (2012) Handbook of volatility models and their applications. John Wiley & Sons, Hoboken
63. Bauwens L, Hafner CM, Pierret D (2013) Multivariate volatility modeling of electricity futures. *J Appl Econ* 28:743–761
64. Noureldin D, Shephard N, Sheppard K (2014) Multivariate rotated ARCH models. *J Econ* 179:16–30
65. de Almeida D, Hotta LK, Ruiz E (2018) MGARCH models: trade-off between feasibility and flexibility. *Int J Forecast* 34:45–63
66. Trucios C, Zevallos M, Hotta L, Santos AAP (2019) Covariance prediction in large portfolio allocation. *Econometrics* 7:19
67. Bzdok D, Altman N, Krzywinski M (2018) Statistics versus machine learning. *Nat Methods* 15:233–234
68. Makridakis S, Spiliotis E, Assimakopoulos V (2018) Statistical and machine learning forecasting methods: concerns and ways forward. *PLoS One* 13:1–26
69. Hsu M-W, Lessmann S, Sung M-C, Ma T, Johnson JEV (2016) Bridging the divide in financial market forecasting: machine learners vs. financial economists. *Expert Syst Appl* 61:215–234
70. Cai X, Lai G, Lin X (2013) Forecasting large scale conditional volatility and covariance using neural network on GPU. *J Supercomput* 63:490–507
71. Bucci A (2020) Cholesky-ANN models for predicting multivariate realized volatility. *J Forecast* 39:865–876
72. Yu J, Weng Y, Rajagopal R (2017) Mapping rule estimation for power flow analysis in distribution grids. *arXiv:1702.07948*. <https://arxiv.org/abs/1702.07948>. Accessed 20 Nov 2020
73. Awad M, Khanna R (2015) Support vector regression. In: Awad M, Khanna R (eds) Efficient learning machines. Apress, Berkeley, pp 67–80
74. Cristianini N, Shawe-Taylor J (2000) Introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, Cambridge
75. Yumlu MS, Gurgen FS (2011) SVR for time series prediction. In: Boyle BH (ed) Support vector machines: data analysis, machine learning and applications. Nova Science Publishers, New York, pp 117–130
76. Shen G, Tan Q, Zhang H, Zeng P, Xu J (2018) Deep learning with gated recurrent unit networks for financial sequence predictions. *Procedia Comput Sci* 131:895–903
77. Ryll L, Seidens S (2019) Evaluating the performance of machine learning algorithms in financial market forecasting: a comprehensive survey. *arXiv:1906.07786 [q-fin.CP]*. <https://arxiv.org/abs/1906.07786>. Accessed 21 Nov 2020
78. Tange RI, Rasmussen MA, Taira E, Bro R (2017) Benchmarking support vector regression against partial least squares regression and artificial neural network: effect of sample size on model performance. *J Near Infrared Spectrosc* 25(6):381–390
79. Rousseeuw PJ, Molenberghs G (1993) Transformation of non positive semidefinite correlation matrices. *Commun Stat Theory Methods* 22:965–984
80. Zhao T, Roeder K, Liu H (2014) Positive semidefinite rank-based correlation matrix estimation with application to semiparametric graph estimation. *J Comput Graph Stat* 23:895–922
81. Chalimourda A, Scholkopf B, Smola AJ (2004) Experimentally optimal γ in support vector regression for different noise models and parameter settings. *Neural Netw* 17:127–141
82. Phan AV, Nguyen ML, Bui LT (2017) Feature weighting and SVM parameters optimization based on genetic algorithms for classification problems. *Appl Intell* 46:455–469
83. Wang H, Xu D (2017) Parameter selection method for support vector regression based on adaptive fusion of the mixed kernel function. *J Control Sci Eng* 2017:1–12
84. Andersen TG, Benzoni L (2009) Realized volatility. In: Andersen TG, Davis RA, Kreiss JP, Mikosch T (eds) Handbook of Financial Time Series. Springer Verlag, New York, pp 555–575
85. Andersen TG, Bollerslev T, Diebold FX (2010) Parametric and nonparametric volatility measurement. In: Hansen LP, Ait-Sahalia Y (eds) Handbook of financial econometrics, vol 1. Tools and techniques. North-Holland, Amsterdam, pp 67–138
86. McAleer M, Medeiros MC (2008) Realized volatility: a review. *Econ Rev* 27:10–45
87. Pigorsch C, Pigorsch U, Popov I (2012) Volatility estimation based on high-frequency data. In: Duan JC, Härdle WK, Gentle JE (eds) Handbook of computational finance. Springer-Verlag, Berlin, pp 335–369
88. Laurent S, Rombouts JVK, Violante F (2013) On loss functions and ranking forecasting performances of multivariate volatility models. *J Econ* 173:1–10
89. Hansen PR, Lunde A, Nason JM (2011) The model confidence set. *Econometrica* 79:453–497
90. Patton AJ, Sheppard K (2009) Evaluating volatility and correlation forecasts. In: Andersen TG, Davis RA, Kreiss J-P, Mikosch TV (eds) Handbook of financial time series. Springer, Berlin, pp 801–838
91. Violante F, Laurent S (2012) Volatility forecasts evaluation and comparison. In: Bauwens L, Hafner C, Laurent S (eds) Handbook of volatility models and their applications. John Wiley & Sons, Hoboken, pp 465–486
92. Forbes K, Rigobon R (2002) No contagion, only interdependence: measuring stock market co-movements. *J Financ* 57:2223–2261

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Piotr Fiszeder Professor at the Faculty of Economic Sciences and Management, Nicolaus Copernicus University in Toruń, Poland. Member of the Editorial Board of several scientific journals. Author of over 60 academic publications in financial econometrics and empirical finance in prestigious journals like *International Journal of Forecasting*, *Journal of Economic Dynamics and Control*, *Journal of Empirical Finance*, *Empirical Economics*.

He participated in many research projects with support from the Ministry of Science and Higher Education in Poland, National Science Centre and the European Community's Phare ACE Programme. At present, he leads a research project entitled "Multivariate volatility models - the application of low and high prices" financed by the National Science Centre.



Witold Orzeszko received a Ph.D. degree in Economics from the Nicolaus Copernicus University in Toruń, Poland, in 2005. He is currently a professor at the Faculty of Economic Sciences and Management, Nicolaus Copernicus University and head of the Department of Applied Informatics and Mathematics in Economics. His principal areas of research are financial econometrics, nonparametric statistics and machine learning techniques.