



# Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting



Tao Xiong, Yukun Bao\*, Zhongyi Hu

School of Management, Huazhong University of Science and Technology, Wuhan 430074, PR China

## ARTICLE INFO

### Article history:

Received 25 January 2013  
Received in revised form 9 September 2013  
Accepted 9 October 2013  
Available online 16 October 2013

### Keywords:

Stock price forecasting  
Interval-valued data  
Multiple-output support vector regression  
Firefly algorithm  
Trading strategy

## ABSTRACT

Highly accurate interval forecasting of a stock price index is fundamental to successfully making a profit when making investment decisions, by providing a range of values rather than a point estimate. In this study, we investigate the possibility of forecasting an interval-valued stock price index series over short and long horizons using multi-output support vector regression (MSVR). Furthermore, this study proposes a firefly algorithm (FA)-based approach, built on the established MSVR, for determining the parameters of MSVR (abbreviated as FA-MSVR). Three globally traded broad market indices are used to compare the performance of the proposed FA-MSVR method with selected counterparts. The quantitative and comprehensive assessments are performed on the basis of statistical criteria, economic criteria, and computational cost. In terms of statistical criteria, we compare the out-of-sample forecasting using goodness-of-forecast measures and testing approaches. In terms of economic criteria, we assess the relative forecast performance with a simple trading strategy. The results obtained in this study indicate that the proposed FA-MSVR method is a promising alternative for forecasting interval-valued financial time series.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Forecasting stock prices is a fascinating issue in financial market research. Accurately forecasting stock prices, which forms the basis for decision making regarding financial investments, is most likely the greatest challenge for the capital investment industry and is thus of great interest to academic researchers and practitioners.

According to an extensive literature investigation, it is not difficult to find a wide variety of methodologies and techniques that have been used for stock price forecasting with various degrees of success, such as the Box–Jenkins method [1], general autoregressive conditional heteroskedasticity [2], stochastic volatility model [3], fuzzy logic approach [4], gray-based approaches [5], wavelet transforms and adaptive models [6], neural networks [7], support vector regression [8,9], hybrid models [10,11], and decision support systems [12]. However, it is important to note that the studies above considered point forecasting rather than interval forecasting.

Interval forecasts of stock prices have the advantage of taking into account variability and/or uncertainty, reducing the amount of random variation relative to that found in classic single-valued stock price time series (e.g., stock closing price). As Hu and He [13] noted, the interval forecasts of stock price are superior to the traditional point forecasts in terms of the overall lower mean error and higher average accuracy ratio. Moreover, intervals of

stock prices have been widely used in the construction of a variety of technical trading rules [14]. To date, there has been a great deal of research focused on exploring the underlying dynamics of interval-valued stock prices and developing suitable models for forecasting them [13,15–18]. For example, Maia et al. [15] proposed a hybrid methodology that combines the ARIMA and ANN models for interval-valued stock price forecasting. Cheung et al. [18] found evidence of cointegration between the daily log highs and log lows of several major stock indices and further forecasted the daily log highs and log lows using a vector error correction model (VECM). The reader is referred to Arroyo et al. [16] for a recent survey of the present methodologies and techniques employed for interval-valued stock price forecasting. It should be noted that the interval-valued data in this study do not come from noise assumptions as in [19], but rather from the expression of variation or the aggregation of huge databases into a reduced number of groups as in [13,15–18].

Our study focuses on extending the multi-output support vector regression (MSVR) to adapt to the scenario of interval forecasting of a stock price index. As a well-known intelligent algorithm, support vector regression (SVR) [20] has attracted particular attention from both practitioners and academics for use in time series forecasting during the last decade. SVR algorithms have been found to be a viable contender among various time-series models [21,22] and have been successfully applied to different areas [23]. Despite the promising SVR demonstrated in [21–24], the applications of SVR in interval-valued time series (ITS) have not been widely

\* Corresponding author. Tel.: +86 27 87558579; fax: +86 27 87556437.  
E-mail addresses: [yukunbao@hust.edu.cn](mailto:yukunbao@hust.edu.cn), [y.bao@ieee.org](mailto:y.bao@ieee.org) (Y. Bao).

explored. This is because the standard formulation of SVR can only be used as a univariate modeling technique for ITS forecasting due to its inherent single-output structure. The univariate technique fits and forecasts the interval bounds of ITS independently, without considering the possible interrelations that are present among them, which has been criticized in [16]. To generalize the SVR from a regression estimation to multi-dimensional problems, Pérez-Cruz et al. [25] proposed a multi-dimensional SVR that uses a cost function with a hyperspherical intensive zone, which is capable of obtaining better predictions than using an SVR model independently for each dimension. Recently, Tuia et al. [26] proposed a multi-output SVR model (MSVR), based on the previous contribution in [27], for simultaneously estimating different biophysical parameters from remote sensing images. In the work of [25,27], the MSVRs are established and justified in a variety of disciplines [26–28]. Although past studies have clarified the capability of the MSVR, there have been very few, if any, efforts to evaluate the performance of the MSVR for time series forecasting, particularly interval-valued time series forecasting. As such, we set out to investigate the possibility of forecasting the lower and upper bounds of stock index series simultaneously by making use of an MSVR. In this model, the inputs of the MSVR are the lagged intervals, while the two outputs of the MSVR correspond to the forecasts of the bounds.

Parameter selection for the MSVR is another issue addressed in this study. The generalization ability of the MSVR depends on adequately setting parameters, such as the penalty coefficient  $C$  and kernel parameters. Therefore, the selection of optimal parameters is crucial to obtain good performance in handling forecasting tasks with MSVR. To date, a large number of evolutionary algorithms, such as the genetic algorithm (GA) and particle swarm optimization (PSO), have been employed to optimize the parameters of SVR. The firefly algorithm (FA), a novel swarm-based intelligent metaheuristic, was recently introduced by Yang [29]. The FA mimics the social behavior of fireflies, which move and communicate with each other based on characteristics such as the brightness, frequency and time period of their flashing. The superiority of the FA against the GA and PSO in existing studies [29–31] motivates us to use the FA for selecting parameters for the MSVR. By doing so, this study proposes a FA-based approach to appropriately determine the parameters of MSVR for ITS forecasting (abbreviated as FA-MSVR). For comparison purposes, a univariate technique (fitting the interval bounds independently), standard SVR, and three well-established interval-valued forecasting methods (fitting the interval bounds simultaneously), namely Holt's exponential smoothing method for intervals (Holt<sup>t</sup>) [17], the vector error correction model (VECM) [18], and the interval multilayer perceptron (iMLP) [32], are selected as benchmarks. Three globally traded broad market indices, the S&P 500, FTSE 100, and Nikkei 225, are chosen as experimental datasets.

To examine the performance of the proposed FA-MSVR method for interval forecasts of a stock price index, we analyze the out-of-sample one- and multi-step-ahead forecasts from the FA-MSVR and selected benchmarks in two ways. First, we examine whether the out-of-sample forecasts generated by the FA-MSVR are more accurate than and preferable to those generated by the benchmark methods for an interval-valued stock index series, employing statistical criteria such as the goodness of forecast measure (e.g., the interval's average relative variance) and the accuracy compared to competing forecasts (e.g., the analysis of variance test and Tukey's HSD test). Second, we analyze whether the FA-MSVR is superior to the selected benchmarks in practice, assessing its relative forecast performance with economic criteria. We use the forecasts of lower and upper bounds from the different methods in a simple trading strategy and compare the returns to determine whether the FA-MSVR is a useful forecasting approach for an investor.

In summary, our contributions could be outlined as follows. First, we extend the MSVR in a novel manner to adapt to the scenario of interval-valued time series forecasting. Second, the possibility of forecasting the lower and upper bounds of interval-valued stock index series simultaneously by the established MSVR is examined. Third, to address the determination of parameters for the MSVR, the parameters of MSVR are tuned using a recently proposed FA. Finally, not only statistical accuracy but also economic criteria are used to assess the practicability of the FA-MSVR for interval-valued stock index forecasting.

This paper is structured as follows. In Section 2, we provide a brief introduction to the MSVR and illustrate the data representation of an interval-valued stock index series analysis. Afterwards, the proposed FA-MSVR method is discussed in detail in Section 3. Section 4 details the research design of the data descriptions, statistical and economic criteria, input selection, and implementation of the methodologies. Following that, in Section 5, the experimental results are discussed. Section 6 concludes this work.

## 2. MSVR with an interval-valued stock index series

This section presents the overall formulation process of the MSVR for interval-valued stock index series forecasting. First, the data representation of interval-valued stock index series is illustrated. Then, the MSVR for the obtained ITS is formulated in detail.

### 2.1. Construction of an interval-valued stock index series

An interval-valued variable,  $[X]$ , is a variable defined for all of the elements  $i$  of a set  $E$ , where  $[X_i] = \{[X_i^L, X_i^U]^T : X_i^L, X_i^U \in \mathbb{R}, X_i^L \leq X_i^U\} \forall i \in E$ . Table 1 shows the daily interval values of the S&P 500 index.

The particular value of  $[X]$  for the  $i$ th element can be denoted either by the interval's lower and upper bounds,  $[X_i] = [X_i^L, X_i^U]^T$ , or the center (mid-point) and radius (half-range),  $[X_i] = [X_i^C, X_i^R]^T$ , where  $X_i^C = (X_i^L + X_i^U)/2$  and  $X_i^R = (X_i^U - X_i^L)/2$ . Fig. 1 illustrates the structure of an interval.

An interval-valued time series (ITS) is a chronological sequence of interval-valued variables. The value of a variable at each instant in time  $t$  ( $t = 1, \dots, n$ ) is expressed as a two-dimensional vector  $[X_t^L, X_t^U]^T$  with the elements in  $\mathbb{R}$  representing the lower bound  $X_t^L$  and upper bound  $X_t^U$ , with  $X_t^L \leq X_t^U$ . Thus, an ITS is  $[X_t] = [X_t^L, X_t^U]^T$  for  $t = 1, \dots, n$ , where  $n$  denotes the number of intervals of the time series (sample size). Fig. 2 illustrates the stock market in which a daily interval-valued S&P 500 index series arises. Fig. 2(a) illustrates a 10-min S&P 500 index from December 11, 2012 to December 27, 2012. Fig. 2(b) depicts the corresponding daily S&P 500 index intervals.

**Table 1**  
Interval-valued variables.

Year 2012	S&P 500 index [Lower, Upper]
December 11	[1418.55, 1434.27]
December 12	[1426.76, 1438.59]
December 13	[1416, 1431.36]
December 14	[1411.88, 1419.45]
December 17	[1413.54, 1430.67]
...	...
...	...

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات