



Adaptive fuzzy modeling of interval-valued stream data and application in cryptocurrencies prediction

Leandro Maciel¹ · Rosangela Ballini² · Fernando Gomide³

Received: 23 February 2021 / Accepted: 26 June 2021 / Published online: 7 July 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

This paper introduces an adaptive interval fuzzy modeling method using participatory learning and interval-valued stream data. The model is a collection of fuzzy functional rules in which the rule base structure and the parameters of the rules evolve simultaneously as data are input. The evolving nature of the method allows continuous model adaptation using the stream interval input data. The method employs participatory learning to cluster the interval input data recursively, constructs a fuzzy rule for each cluster, uses the weighted recursive least squares to update the parameters of the rule consequent intervals, and returns an interval-valued output. The method is evaluated using actual data to model and forecast the daily lowest and highest prices of the four most traded cryptocurrencies, BitCoin, Ethereum, XRP, and Litecoin. The performance of the adaptive interval fuzzy modeling is compared with the adaptive neuro-fuzzy inference system, long short-term memory neural network, autoregressive integrated moving average, exponential smoothing state model, and the naïve random walk methods. Results show that the suggested interval fuzzy model outperforms all these methods in predicting prices in the digital coin market, especially when considering directional accuracy measure.

Keywords Adaptive machine learning · Fuzzy modeling · Interval-valued stream data · Forecasting · Cryptocurrencies

1 Introduction

Conventionally, time series analysis uses a succession of single-valued data samples. This may be restrictive in situations in which complex data analysis is needed to comprehend the inherent variability and uncertainty of a

phenomenon. For instance, in economics, the daily stock price of a corporation is expressed as the daily minimum and maximum trading prices. If only the lowest (or the highest, or closing) price of each day is considered, then the resulting time series is single-valued, neglecting the inherent intra-daily price variability. If both, the daily lowest and the highest prices are accounted for, then an interval-valued time series (ITS) is formed with its intrinsic trend (or level) and volatility information (the range between the boundaries) [29]. ITS is encountered in many fields such as finance [3], energy [30], environment and climate [8], and agriculture [32] to mention but a few.

This paper introduces a novel adaptive interval fuzzy modeling method to model and forecast interval-valued time series. The method processes stream interval-valued input data, employs participatory learning to cluster interval-valued data, develops a fuzzy rule for each cluster, uses weighted recursive least squares to update the parameters of the rule consequents, and outputs interval-valued forecasts. The adaptive interval fuzzy method is evaluated in modeling and forecasting the daily lowest and highest prices of the most traded cryptocurrencies, respectively,

✉ Leandro Maciel
leandromaciel@usp.br

Rosangela Ballini
ballini@unicamp.br

Fernando Gomide
gomide@unicamp.br

¹ Department of Business, Faculty of Economics, Business and Accounting, University of São Paulo, Av. Prof. Luciano Gualberto, 908, São Paulo 05508-010, Brazil

² Department of Economics Theory, Institute of Economics, University of Campinas, R. Pitágoras, 353, Campinas 13083-857, Brazil

³ Department of Computer Engineering and Automation, School of Electrical and Computer Engineering, University of Campinas, Av. Albert Einstein, 400, Campinas, São Paulo 13083-852, Brazil

Bitcoin, Ethereum, XRP and Litecoin. Its performance is compared against the adaptive neuro-fuzzy inference system (ANFIS), long short-term memory (LSTM) neural network, autoregressive integrated moving average (ARIMA), exponential smoothing state model (ETS), and the naïve random walk benchmarks. These methods do not rely on interval arithmetic, and their interval forecasts are produced from the individual lower and upper bound forecasts.

The rest of this paper proceeds as follows. Section 2 overviews the current literature and state of the art of the area. Section 3 details the nature of the data, the structure of the fuzzy models, and suggests the adaptive method to develop fuzzy models from interval stream data. Section 4 concerns cryptocurrencies modeling and forecasting using actual daily lowest and highest prices data. Section 5 concludes the paper, summarizing its contributions and suggesting issues for future investigation.

2 Literature review

Different approaches have been developed to model and forecast interval-valued data. The first linear regression model for interval-valued data was investigated in [4] using the center of the interval method. A constrained center and interval range method with nonnegative constraints on the coefficients of the range regression model is addressed in [16]. The use of ARIMA and neural networks models to forecast the center and range of intervals is pursued in [21]. An autoregressive time series modeling approach is addressed in [33], and a threshold autoregressive interval model in [29]. Nonparametric alternatives such as the interval kernel regression method [7] and the nonparametric additive approach [15] have also been developed.

Machine learning has also been used in interval-valued data modeling and forecasting. Examples include the interval multilayer perceptron (iMLP) model [26], a multilayer perceptron (MLP) neural network and Holt's exponential smoothing [20], a multiple-output support vector regression method for interval-valued stock price index forecasting [34].

The possibilistic fuzzy, evolving method for interval time series modeling and forecasting was proposed in [19]. In this case, the model employs memberships and typicalities to recursively cluster data, uses participatory learning to update the forecasting model structure as stream data are input, and processes interval-valued data.

Intuitionistic Fuzzy Grey Cognitive Map (IFGCM) for interval-valued data [9] has also been explored in modeling and forecasting. IFGCM was evaluated using stock market data and the results show the high efficiency of the IFGCM, especially when compared with state-of-the-art models.

Participatory learning in fuzzy clustering [27] was extended for clustering interval data in [18]. The clustering method uses interval arithmetic, and the computational experiments are reported using synthetic interval data sets with linearly non-separable clusters of different shapes and sizes.

This paper introduces an adaptive interval fuzzy method to model stream interval-valued data. The method employs participatory learning to cluster interval-valued data using the Hausdorff–Pompeiu [11, 14] distance to compute the (dis)similarity between intervals. Moore and Hukuhara subtraction are studied to verify and evaluate the impact of the distinct subtraction operations on ITS modeling and forecasting performance. Moreover, the paper also contributes with an alternative forecasting technique for cryptocurrencies based on highest and lowest prices, which serve as a measure of volatility, an essential information in decision making and risk management in high price variation markets such as digital coins.

3 Adaptive fuzzy modeling of interval-valued stream data

This section addresses the nature of data, the structure of interval fuzzy models, and develops an adaptive method to construct the models within the framework of evolving participatory learning from interval-valued stream data [2, 18]. First, a brief reminder of interval-valued time series and interval arithmetic is given. Next, the structure of the rule-based model is shown. Finally, the interval adaptive fuzzy modeling method is explained, and the computational steps of the corresponding procedure are summarized.

3.1 Interval time series and interval arithmetic

An interval-valued time series (ITS) is a sequence of interval-valued data indexed by successive time steps $t = 1, 2, \dots, N$. An interval datum is expressed as $[x] = [x^L, x^U] \in \mathcal{K}_c(\mathbb{R})$, where $\mathcal{K}_c(\mathbb{R}) = \{[x^L, x^U] : x^L, x^U \in \mathbb{R}, x^L \leq x^U\}$ is the set of closed intervals of the real line \mathbb{R} , and x^L and x^U are the lower and upper bounds of the interval $[x]$. An interval $[x]$ may also be expressed by a two-dimensional vector $[x] = [x^L, x^U]^T$.

Modeling ITS requires interval arithmetic, an extension of traditional arithmetic to operate on intervals. This paper uses the arithmetic operations introduced by Moore [23] summarized below:

$$\begin{aligned}
 [x] + [y] &= [x^L + y^L, x^U + y^U] \\
 [x] - [y] &= [x^L - y^U, x^U - y^L] \\
 [x] \cdot [y] &= [\min\{x^L y^L, x^L y^U, x^U y^L, x^U y^U\}, \\
 &\quad \max\{x^L y^L, x^L y^U, x^U y^L, x^U y^U\}] \\
 [x]/[y] &= [x] \cdot (1/[y]) \quad \text{with } 1/[y] = [1/y^U, 1/y^L].
 \end{aligned}
 \tag{1}$$

It is well known that $[x] - [x] \neq [0]$ for Moore subtraction, with $[0] = [0, 0]$. An alternative subtraction operation for which $[x] - [x] = [0]$ is recovered is the generalized Hukuhara difference [28] defined as follows:

$$\begin{aligned}
 [x] - [y] &= [\min\{x^L - y^L, x^U - y^U\}, \\
 &\quad \max\{x^L - y^L, x^U - y^U\}].
 \end{aligned}
 \tag{2}$$

This paper considers both Moore and Hukuhara subtraction to verify and evaluate their impact on ITS modeling and forecasting performance.

The forecasting ability of the models must be verified and tested using interval-based accuracy metrics, which requires the notions of union and intersection of intervals $[x]$ and $[y]$ [23]. They are defined as follows:

$$[x] \cup [y] = [\min\{x^L, y^L\}, \max\{x^U, y^U\}], \tag{3}$$

$$[x] \cap [y] = [\max\{x^L, y^L\}, \min\{x^U, y^U\}]. \tag{4}$$

The intersection of $[x]$ and $[y]$ is empty if $\max\{x^L, y^L\} > \min\{x^U, y^U\}$. Real numbers are considered as intervals of zero length, that is, intervals whose lower bounds are equal to the upper bounds.

The concept of distance between two intervals is important to measure their (dis)similarity. The Hausdorff–Pompeiu distance measures how far two intervals $[x]$ and $[y]$ are from each other [11, 14]. Denoted $dH([x], [y])$, it is computed using:

$$dH([x], [y]) = \max\{|x^L - y^L|, |x^U - y^U|\}. \tag{5}$$

3.2 Fuzzy interval model structure

A fuzzy interval model is a collection of interval fuzzy rules, a fuzzy rule base in which each rule has two parts: an antecedent part identifying the state of the interval-valued input variable, and a consequent part specifying the corresponding interval-valued output variable. Adaptive fuzzy interval modeling (aFIM) uses functional interval fuzzy rules, an interval extension of functional fuzzy rules of the form introduced in [31]. An interval

fuzzy rule base is a collection of functional interval fuzzy rules of the form:

$$\begin{aligned}
 \mathcal{R}_1: \quad &\text{IF } [x] \text{ is } \mathcal{M}_1 \text{ THEN } [y_1] = [\beta_{1,0}] + [\beta_{1,1}][x_1] + \dots + [\beta_{1,n}][x_n], \\
 &\quad \vdots \\
 \mathcal{R}_i: \quad &\text{IF } [x] \text{ is } \mathcal{M}_i \text{ THEN } [y_i] = [\beta_{i,0}] + [\beta_{i,1}][x_1] + \dots + [\beta_{i,n}][x_n], \\
 &\quad \vdots \\
 \mathcal{R}_c: \quad &\text{IF } [x] \text{ is } \mathcal{M}_c \text{ THEN } [y_c] = [\beta_{c,0}] + [\beta_{c,1}][x_1] + \dots + [\beta_{c,n}][x_n],
 \end{aligned}
 \tag{6}$$

where \mathcal{R}_i is the i th fuzzy rule, $i = 1, 2, \dots, c$, c is the number of fuzzy rules in the rule base, $[x] = ([x_1], [x_2], \dots, [x_n])^T$ the vector of inputs, $[x_j] = [x_j^L, x_j^U] \in \mathcal{K}_c(\mathbb{R})$, $j = 1, \dots, n$, and $[\beta_{i,l}] = [\beta_{i,l}^L, \beta_{i,l}^U] \in \mathcal{K}_c(\mathbb{R})$ are interval-valued parameters of the rule consequent, $l = 0, 1, \dots, n$. \mathcal{M}_i is the fuzzy set of the antecedent whose membership function is $\mu_i([x]) : \mathcal{K}_c(\mathbb{R}^n) \rightarrow [0, 1]$, and $[y_i] = [y_i^L, y_i^U] \in \mathcal{K}_c(\mathbb{R})$ is the output of the i th rule. Fuzzy inference with interval functional rules (6) produces outputs $[y]$ as the weighted average:

$$\begin{aligned}
 [y] &= \sum_{i=1}^c \left(\frac{\mu_i([x])[y_i]}{\sum_{j=1}^c \mu_j([x])} \right) \\
 &= \sum_{i=1}^c \lambda_i [y_i],
 \end{aligned}
 \tag{7}$$

where $\lambda_i = \frac{\mu_i([x])}{\sum_{j=1}^c \mu_j([x])}$ is the normalized degree of activation of the i th rule. The membership degree $\mu_i([x])$ of datum $[x]$ is given by:

$$\mu_i([x]) = \left[\frac{\sum_{j=1}^n \left(\max\{|x_j^L - v_{ij}^L|, |x_j^U - v_{ij}^U|\} \right)}{\sum_{j=1}^n \left(\max\{|x_j^L - v_{hj}^L|, |x_j^U - v_{hj}^U|\} \right)} \right]^{\frac{2}{(m-1)}}, \tag{8}$$

where m is a fuzzification parameter (usually $m = 2$), and $[v_{i,j}] = [v_{i,j}^L, v_{i,j}^U] \in \mathcal{K}_c(\mathbb{R})$, $j = 1, \dots, n$ and $i = 1, \dots, c$ is the cluster center of the i th cluster/rule.

Functional interval fuzzy modeling uses parameterized fuzzy regions of the data space, and associates each region with a local affine interval-valued model. The nature of the rule-based model emerges from the fuzzy weighted combination of the collection of the multiple local interval models. The contribution of a local model to the model output is proportional to the normalized activation degree of the corresponding rule. The construction of interval-valued fuzzy models needs two steps: (1) to learn the antecedent part of the rules using fuzzy clustering to granulate the input interval-valued data space into

parameterized regions; and (2) to estimate the parameters of the local models of the rules consequents. These two steps are detailed in the next section.

3.3 Learning the antecedents of interval fuzzy rules

Identification of the fuzzy rules of an interval fuzzy model is done using the participatory learning fuzzy clustering algorithm extended to handle interval-valued data [18]. The aim of interval-valued data clustering is to partition an interval data set $\mathbf{X} = \{[\mathbf{x}_1], \dots, [\mathbf{x}_N]\}$ in c fuzzy subsets, $2 \leq c \leq N$, where N is the number of samples. Intervals $[\mathbf{x}_j]$ bounds are assumed to be normalized using min–max operator:

$$x_{\text{norm}}^B = \frac{x^B - \min \{x^L, x^U\}}{\max \{x^L, x^U\} - \min \{x^L, x^U\}}, \tag{9}$$

where B denotes either the lower bound L, or the upper bound U of the interval.

Participatory learning clustering is based on the idea that the current cluster structure influences the cluster learning process whenever a new data sample is an input. Model learning depends on what the system already knows about the model itself. The input data and the cluster structure affect self-organization depending on the compatibility of the data with the current cluster structure [27]. The impact of new data in inducing model revision depends on its compatibility with the current rule base structure or, equivalently, its compatibility with the current cluster structure.

The cluster structure is defined by the cluster centers. If $[\mathbf{V}] = [[\mathbf{v}_1], \dots, [\mathbf{v}_c]]$, $[\mathbf{v}_i] = ([v_{i,1}], \dots, [v_{i,n}])^T$, $[v_{i,j}] = [v_{i,j}^L, v_{i,j}^U] \in \mathcal{K}_c([0, 1])$, $j = 1, \dots, n$ and $i = 1, \dots, c$, then $[\mathbf{V}]$ represents a cluster structure. The aim of the antecedent learning is to determine $[\mathbf{V}]$ from inputs $[\mathbf{x}]^t \in [0, 1]^n$, $t = 1, \dots$, i.e., $[\mathbf{x}]^t$ is used as a vehicle to learn about $[\mathbf{V}]$.

The learning process is participatory if the contribution of $[\mathbf{x}]^t$ to the learning process depends upon its acceptance by the cluster structure $[\mathbf{V}]^t$ at t being valid. In other words, data $[\mathbf{x}]^t$ must be compatible with current cluster structure $[\mathbf{V}]^t$. The compatibility $\rho_i^t \in [0, 1]$ of input $[\mathbf{x}]^t$ with the cluster center $[\mathbf{v}_i]^t$ of $[\mathbf{V}]^t$, $i = 1, \dots, c$ is calculated as:

$$\rho_i^t = 1 - \frac{1}{n} \sum_{j=1}^n \left(\max \left\{ |x_j^{L,t} - v_{i,j}^{L,t-1}|, |x_j^{U,t} - v_{i,j}^{U,t-1}| \right\} \right). \tag{10}$$

Participatory clustering updates only the cluster center whose compatibility with input $[\mathbf{x}]^t$ is the highest. Thus, if

cluster center $[\mathbf{v}_i]^{t-1}$ is the most compatible with $[\mathbf{x}]^t$, that is, $i = \arg \max_{j=1, \dots, c} \{\rho_j^t\}$, then it is updated as follows:

$$[\mathbf{v}_i]^t = [\mathbf{v}_i]^{t-1} + G_i^t([\mathbf{x}]^t - [\mathbf{v}_i]^{t-1}), \quad G_i^t = \alpha \rho_i^t \tag{11}$$

where $\alpha \in [0, 1]$ is the basic learning rate. Notice that (11) needs the value of $([\mathbf{x}]^t - [\mathbf{v}_i]^{t-1})$. Both Moore’s (1) and generalized Hukuhara subtraction (2) are used in (11) to verify if they make a difference in interval time series modeling and forecasting performance.

Adaptive modeling is particularly relevant in time-varying domains. A sequence of input data with low compatibility with the current cluster structure indicates that the current model should be revised in front of new information. Participatory learning uses an arousal mechanism to monitor how the compatibility measure values progress, i.e., the arousal mechanism acts as a reminder of when current cluster structure should be revised, given the new information carried in the data. A high arousal value indicates less confidence in how the current model fits recent input data. Thus, arousal can be seen as the complement of confidence of [18]. A way to express the arousal $a_i^t \in [0, 1]$ at step t is:

$$a_i^t = a_i^{t-1} + \beta(1 - \rho_i^t - a_i^{t-1}), \tag{12}$$

where $\beta \in [0, 1]$ controls the rate of change of arousal. The closer β is to one, the faster the learning process senses compatibility variations. If the values of arousal a_i^t are greater than or equal a threshold $\tau \in [0, 1]$ for $i = 1, \dots, c$, then a new cluster should be created, assigning the current data as its cluster center, that is, $[\mathbf{v}_{c+1}]^t = [\mathbf{x}]^t$. Otherwise, the center with the highest compatibility is updated to accommodate input the data using (11). The arousal mechanism (12) becomes part of the learning process by converting G_i^t of (11) in an effective learning rate:

$$G_i^t = \alpha(\rho_i^t)^{1-a_i^t}. \tag{13}$$

Updates of cluster centers (11) can be viewed as a form of exponential smoothing modulated by the compatibility of data with the model structure, the very nature of participatory learning. Participatory clustering also accounts for redundant clusters because updating a cluster center using (11) can cause the center to get closer to another one, and a redundant cluster may be formed. A cluster i is redundant if its similarity with any other cluster h , $\rho_{i,h}^t$, is greater than or equal to a threshold $\lambda \in [0, 1]$. The similarity between cluster centers i and h is found using:

$$\begin{aligned} \rho_{i,h}^t &= 1 - \frac{1}{n} dH([\mathbf{v}_i]^t, [\mathbf{v}_h]^t) \\ &= 1 - \frac{1}{n} \sum_{j=1}^n \left(\max \left\{ |v_{i,j}^{L,t} - v_{h,j}^{L,t}|, |v_{i,j}^{U,t} - v_{h,j}^{U,t}| \right\} \right). \end{aligned} \tag{14}$$

If clusters i and h are declared redundant, then they are replaced by a cluster whose center is the average of their

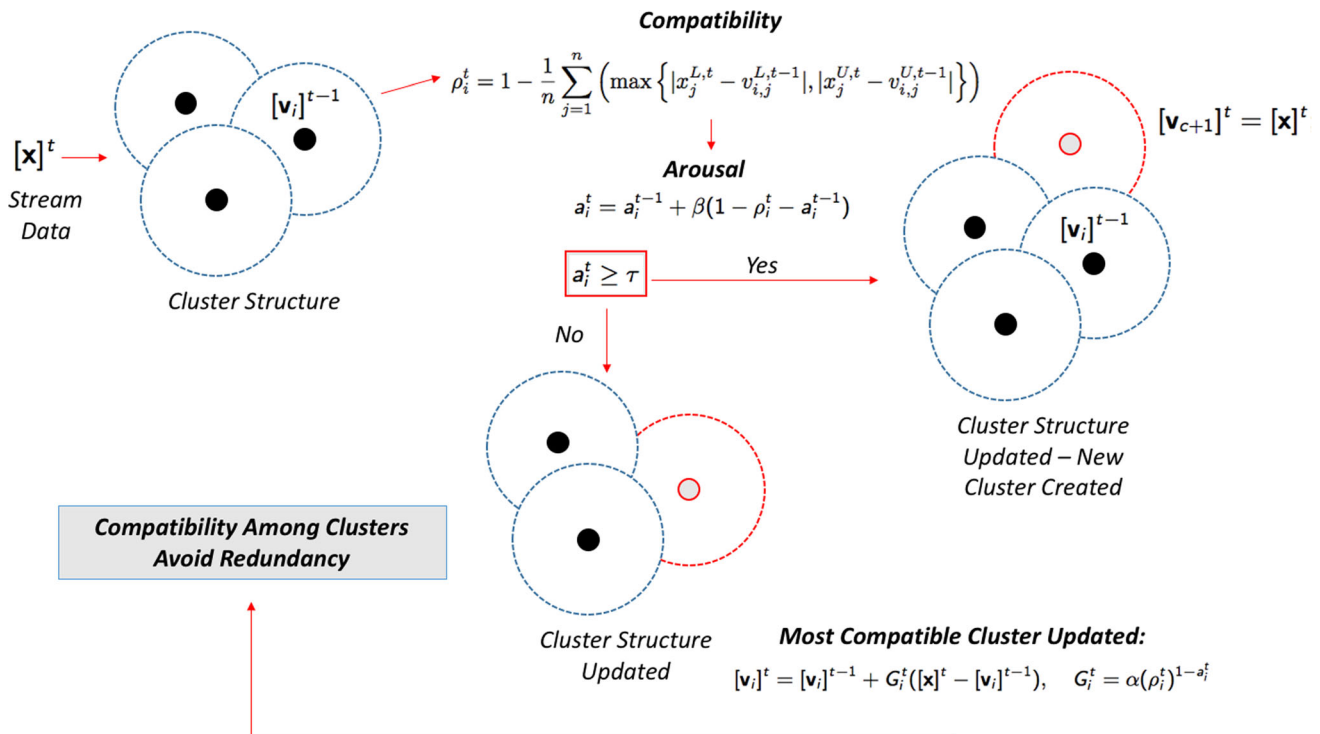


Fig. 1 Participatory learning in clustering interval stream data and antecedents learning

centers. Figure 1 summarizes the rule antecedent learning process of aFIM.

3.4 Parameter estimation of interval fuzzy rule consequents

To complete the interval rule-based model construction procedure, the interval-valued parameters $[\beta_{i,0}], [\beta_{i,1}], \dots, [\beta_{i,n}]$ of the rule consequent are estimated using the classic weighted recursive least squares algorithm (wRLS) [1, 17] to compute the lower bounds $\beta_{i,l}^L$ and the upper bounds $\beta_{i,l}^U$, $i = 1, \dots, c$ and $l = 0, 1, \dots, n$ separately. To do this, expression (7) is expressed in terms of the input interval data bounds individually, namely $y^B = \Lambda^T \Theta$ where $\Lambda = [\lambda_1 \mathbf{x}_e^T, \lambda_2 \mathbf{x}_e^T, \dots, \lambda_c \mathbf{x}_e^T]^T$ is the fuzzily weighted input data, $\mathbf{x}_e = [1, x_1^B, x_2^B, \dots, x_n^B]^T$ is the extended input, $\Theta = [\beta_1^T, \beta_2^T, \dots, \beta_c^T]^T$, and $\beta_i = [\beta_{i,0}^B, \beta_{i,1}^B, \dots, \beta_{i,n}^B]^T$. Recall that superscript B denotes either the lower L, or the upper bound U of the interval. The locally optimal error criterion wRLS is considered:

$$\min E_i^t = \min \sum_{k=1}^t \lambda_i (y^{B,k} - (\mathbf{x}_e^k)^T \beta_i^k)^2, \tag{15}$$

whose solution can be expressed recursively by [17]:

$$\beta_i^{t+1} = \beta_i^t + \Sigma_i^t \mathbf{x}_e^t \lambda_i^t (y^{B,t} (\mathbf{x}_e^t)^T \beta_i^t), \quad \beta_i^0 = 0, \tag{16}$$

$$\Sigma_i^{t+1} = \Sigma_i^t - \frac{\lambda_i^t \Sigma_i^t \mathbf{x}_e^t (\mathbf{x}_e^t)^T \Sigma_i^t}{1 + \lambda_i^t (\mathbf{x}_e^t)^T \Sigma_i^t \mathbf{x}_e^t}, \quad \Sigma_i^0 = \Omega I, \tag{17}$$

where Ω is a large number (typically $\Omega = 1000$), and Σ is the dispersion matrix.

More precisely, parameter estimation of interval fuzzy rule consequents is done in two steps. The first step uses expressions (16)–(17) to individually estimate the lower bounds $\beta_{i,l}^L$ and the upper bounds $\beta_{i,l}^U$ of the interval-valued parameters $[\beta_{i,l}] = [\beta_{i,l}^L, \beta_{i,l}^U]$, $l = 0, 1, \dots, n$. The second step computes the outputs of the fuzzy rules (6) at $t + 1$ using:

$$[y_i]^{t+1} = [\beta_{i,0}]^t + [\beta_{i,1}]^t [x_1]^t + \dots + [\beta_{i,n}]^t [x_n]^t. \tag{18}$$

$i = 0, 1, \dots, c$. The model output $[y]^{t+1}$ at $t + 1$ is the weighted average of the outputs $[y_i]^{t+1}$ computed using (7).

3.5 Adaptive interval fuzzy modeling method

Adaptive interval fuzzy modeling is inherently recursive, which means that it is memory efficient in continuously and endlessly learning and adaptation using stream data. This is of major importance, especially in online and real-time application situations. The procedure to construct interval fuzzy models is as follows.

Adaptive interval fuzzy modeling procedure

1. choose parameters α, β, τ , and λ .
 2. set $a_1^1 = 0, \Sigma_i^0 = \Omega I, [\beta_0^1] = [0, 0]$.
 3. start cluster structure with the first input datum: $c = 1, [\mathbf{v}_1]^1 = [\mathbf{x}]^1$.
 4. **for** $t = 2, 3, \dots$ **do**
 5. read $[\mathbf{x}]^t$
 6. **for** $i = 1, \dots, c$
 7. compute $\rho_i^t = 1 - \frac{1}{n} \sum_{j=1}^n \left(\max \left\{ |x_j^{L,t} - v_{i,j}^{L,t-1}|, |x_j^{U,t} - v_{i,j}^{U,t-1}| \right\} \right)$
 8. update $a_i^t = a_i^{t-1} + \beta(1 - \rho_i^t - a_i^{t-1})$
 9. **end for**
 10. **if** $a_i^t \geq \tau$ for $i = 1, \dots, c$
 11. create a new cluster: $c = c + 1$
 12. $[\mathbf{v}_c]^t = [\mathbf{x}]^t$
 13. reset $a_i^t = 0$
 14. **else**
 15. find the most compatible cluster $s: s = \arg \max_{j=1, \dots, c} \{\rho_j^t\}$
 16. update $[\mathbf{v}_s]^t = [\mathbf{v}_s]^{t-1} + G_s^t([\mathbf{x}]^t - [\mathbf{v}_s]^{t-1})$,
 17. **end if**
 18. **for** $i = 1, \dots, c - 1$ and $h = i + 1, \dots, c$
 19. compute cluster compatibility $\rho_{i,h}^t = 1 - \frac{1}{n} dH([\mathbf{v}_i]^t, [\mathbf{v}_h]^t)$
 20. **if** $\rho_{i,h}^t \geq \lambda$
 21. redefine $[\mathbf{v}_i]^t$ using centers i and h , and delete $[\mathbf{v}_h]^t$
 22. $c = c - 1$
 23. **end if**
 24. **end for**
 25. compute rule consequent parameters using the wRLS
 26. compute model output $[y]^{t+1}$
 27. **end for**
-

4 Computational experiments

This section illustrates the adaptive fuzzy interval modeling method (aFIM) introduced in this paper in modeling and forecasting the daily lowest and highest prices of the currently most traded cryptocurrencies, respectively, Bitcoin, Ethereum, XRP and Litecoin. Moreover, the performance of aFIM is compared against ANFIS, LSTM, ARIMA, ETS, and the naïve random walk.

4.1 Cryptocurrency trade

Current interest in cryptocurrencies has offered to investors and speculators a diversity of electronic crypto assets to trade due to benefits such as anonymity, transparency, lower transaction costs, and diversification [6]. Because of their high volatile dynamics, risk management of cryptocurrencies investment positions based on volatility modeling and forecasting is of key interest. When risk management models are built based on intervals forecasts, more accurate analyses can be done. Intervals of daily highest and lowest digital currency prices convey a better idea of portfolio risk perception (volatility) at a time period than a single value such as closing prices. Interval-valued

modeling and forecasting inherently embed price variability information contained in the data samples. This section addresses the performance evaluation of the adaptive fuzzy interval modeling (aFIM) using real-world ITS data from the cryptocurrency market.

4.2 Data

Adaptive fuzzy interval modeling and forecasting use daily highest and lowest prices as interval bounds of four leading cryptocurrencies: Bitcoin, XRP (Ripple), Ethereum, and Litecoin, respectively. Data cover January 1, 2016 to December 31, 2019, with a total of 1461 observations.¹ The data set was divided into in-sample and out-of-sample sets. The in-sample data set contains data from January 1, 2016 to December 31, 2018, and the out-of-sample data contains the 2019 data. Thus, 2019 data are used to evaluate the modeling power and forecasting performance of aFIM and its counterparts. For the in-sample set, data from January 1, 2016 to December 31, 2017 were used for models learning, and the year of 2018 was used as a validation set on

¹ Selection is made one choosing cryptocurrencies with the highest liquidity and market capitalization. The data are available at <https://coinmarketcap.com/>.

choosing control parameters. aFIM processes data as a stream and learning may be done endlessly, which means that there is no need to split data in training, validation, and testing data as usual in machine learning experiments. In spite of that, a data set split was done to keep evaluations and comparisons with the competing methods fair. Forecasting is one-step-ahead with an iterated strategy. The forecasting techniques considered for comparison are the autoregressive moving average (ARIMA), the exponential smoothing state-space model (ETS) [12], the adaptive neuro-fuzzy inference system (ANFIS), a long short-term memory (LSTM) neural network, and the naïve random walk (RW). These methods produce interval forecasts from the individual lowest and highest daily currency data.

4.3 Performance measures

Performace evaluation of the methods is accessed considering the mean absolute percentage error (MAPE) and the root mean squared scaled error (RMSSE) [3], respectively:

$$MAPE^B = \frac{100}{N} \sum_{t=1}^N \frac{|y^{B,t} - \hat{y}^{B,t}|}{y^{B,t}}, \tag{19}$$

$$RMSSE^B = \sqrt{\text{mean} \left(\frac{(y^{B,t} - \hat{y}^{B,t})^2}{\frac{1}{N-1} \sum_{i=2}^N (y^{B,i} - y^{B,i-1})^2} \right)}, \tag{20}$$

where the superscript B denotes either the lower bound L, or the upper bound U of the interval prices, $[y]^t = [y^{L,t}, y^{U,t}]$ is the actual price interval, $[\hat{y}]^t = [\hat{y}^{L,t}, \hat{y}^{U,t}]$ is the forecast price interval at t , and N is the size of the out-of-sample data set.

Additionally, in practice, the direction of price change is as important as, sometimes more important than, the magnitude of the forecasting error [5]. A measure of direction accuracy (DA) [10] is as follows:

$$DA^B = \frac{1}{N} \sum_{t=1}^N Z^{B,t},$$

$$Z^{B,t} = \begin{cases} 1, & \text{if } (y^{B,t+1} - y^{B,t})(\hat{y}^{B,t+1} - \hat{y}^{B,t}) > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

Because MAPE, RMSSE, and DA are computed individually for each bound of an interval, they neglect the inherent interval nature of the data. Performance evaluation of forecasting methods of interval time series data uses the normalized symmetric difference of intervals (NSD) [25] defined by:

$$NSD([y]^t, [\hat{y}]^t) = \frac{1}{N} \sum_{t=1}^N \frac{w([y]^t \cup [\hat{y}]^t) - w([y]^t \cap [\hat{y}]^t)}{w([y]^t \cup [\hat{y}]^t)}, \tag{22}$$

as well as the average coverage (R^C) rate, and the efficiency rate (R^E) [25]:

$$R^C = \frac{1}{N} \sum_{t=1}^N \frac{w([y]^t \cap [\hat{y}]^t)}{w([y]^t)},$$

$$R^E = \frac{1}{N} \sum_{t=N}^n \frac{w([y]^t \cap [\hat{y}]^t)}{w([\hat{y}]^t)}. \tag{23}$$

R^C and R^E give information on what part of ITS data are covered by their forecasts (coverage), and what part of the forecasts cover ITS data (efficiency). If the observed intervals are fully enclosed in the predicted intervals, then the coverage rate is 100%, but the efficiency could be less than 100% and reveal the fact that the forecasted ITS is in average wider than the actual ITS data samples. Hence, R^C and R^E values must be considered jointly. As the literature indicates [25], good forecast performance is expected when the average coverage and efficiency rates are reasonably high, and the difference between them is small.

4.4 Results

The aFIM is compared with ARIMA, ETS, ANFIS, LSTM and RW in cryptocurrencies lowest and highest price forecast. ARIMA, ETS, ANFIS, LSTM and RW are univariate techniques, and their price forecasts are developed for the interval bounds individually. Contrary, aFIM is an interval-based method and develops interval-valued price forecasts. aFIM assumes that forecasts are produced assuming:

$$[\hat{y}]^{t+1} = f([y]^t, [y]^{t-1}, \dots, [y]^{t-l}), \tag{24}$$

that is, the inputs are lagged values of the interval time series $[y]^t = [y^{L,t}, y^{U,t}]$, the output is the one-step-ahead forecast $[\hat{y}]^{t+1} = [\hat{y}^{L,t+1}, \hat{y}^{U,t+1}]$, and f encodes the aFIM output in (7).

Performance evaluation was done using the out-of-sample data, i.e., from January 1, 2019 to December 31, 2019. aFIM was implemented in MATLAB, while ARIMA, ETS, ANFIS and LSTM were developed using software R. Table 1 shows the values of the control parameters for each method. aFIM parameters were chosen by doing simulations to select the ones that give the best MAPE and RMSSE values. Structures of ARIMA, ETS, ANFIS and LSTM are different for lower (L) and upper (U) interval bounds. In ARIMA(p, d, q), p, d and q are the number of autoregressive, difference, and moving average

Table 1 Structure of the forecasting models and their respective control parameters

Currency	ARIMA	ETS	ANFIS	LSTM	aFIM _M	aFIM _H
Bitcoin	L(2, 1, 1)	L(M, A, N)	L(3) U(4)	L(2, 150) U(3, 180)	$l = 4, \beta = 0.10, \tau = 0.46, \alpha = 0.01, \lambda = 0.87$	$l = 6, \beta = 0.11, \tau = 0.55, \alpha = 0.03, \lambda = 0.83$
	U(2, 1, 2)	U(M, A, N)				
Ethereum	L(2, 1, 2)	L(M, A, N)	L(5) U(5)	L(4, 120) U(3, 150)	$l = 4, \beta = 0.11, \tau = 0.48, \alpha = 0.02, \lambda = 0.83$	$l = 7, \beta = 0.09, \tau = 0.44, \alpha = 0.06, \lambda = 0.88$
	U(0, 1, 1)	U(M, A, N)				
XRP	L(4, 1, 5)	L(M, A, N)	L(2) U(5)	L(3, 150) U(2, 150)	$l = 3, \beta = 0.13, \tau = 0.44, \alpha = 0.03, \lambda = 0.85$	$l = 5, \beta = 0.11, \tau = 0.48, \alpha = 0.02, \lambda = 0.85$
	U(5, 1, 5)	U(M, A, N)				
LiteCoin	L(3, 1, 1)	L(M, A, N)	L(3) U(2)	L(4, 180) U(4, 200)	$l = 4, \beta = 0.11, \tau = 0.48, \alpha = 0.03, \lambda = 0.83$	$l = 5, \beta = 0.12, \tau = 0.49, \alpha = 0.03, \lambda = 0.89$
	U(2, 1, 2)	U(M, A, N)				

terms, respectively. In ETS(er, tr, sea), er, tr, and sea mean error, trend, and season components, respectively, such that each component could be of the type *A*, *M* or *N*, denoting additive, multiplicative, and none, respectively. In ANFIS(*l*), *l* is the number of lagged time series values (inputs). In LSTM(*l*, *h*), *l* is the number of lagged time series values (inputs), and *h* is the number of hidden units. The structures of ARIMA, ETS, ANFIS and LSTM were selected to achieve their highest accuracy. aFIM parameters were chosen by doing simulations to select the ones that give the best MAPE and RMSSE values. Recall that aFIM may use either Moore (aFIM_M) and generalized Hukuhara (aFIM_H) subtractions. ANFIS is as detailed in [13]. LSTM has one layer, 800 as the maximum number of epochs, and uses tangent hyperbolic and sigmoid as state and gate activation functions, respectively.

Table 2 summarizes the forecasting performance in terms of MAPE, RMSSE, and DA. Best results are highlighted in bold. aFIM outperforms all competitors from the point of view of MAPE and RMSSE, except for Ethereum high prices when MAPE is the lowest. LSTM is the second best approach, from the point of view of lowest values of MAPE and RMSSE, in general. It is worth to note that naïve random walk performs worst than aFIM, but better than ARIMA, ETS and ANFIS, in line with the well-known “Meese–Rogoff puzzle” [22] which states that exchange rate forecasting models are not able to outperform random walk. The simulations show that this is also the case for cryptocurrency forecasts, but when an adaptive nonlinear model such as aFIM or LSTM is considered, the RW can be outperformed. Concerning the use of the two distinct

interval subtraction operations in aFIM, the generalized Hukuhara generally produces, except for Ethereum, slightly higher performance for all currencies lowest and highest prices forecasts.

It is interesting to recall that the Meese–Rogoff puzzle [24] also suggests that dynamic forecasting models can outperform the random walk in out-of-sample forecasting if performance is measured by the direction of change and profitability. Indeed, this is the case in cryptocurrency price forecasting as well. Table 2 shows that aFIM outperforms random walk, ARIMA, ETS, ANFIS and LSTM from the point of view of the direction accuracy (DA). These results agree with those of [5, 24], in which the random walk is decisively beaten when the direction is used as a comparison measure. The high performance that aFIM achieves in predicting directions of price changes is due to its evolving, continuous adaptation ability to capture price changes more accurately in time-varying environments such as digital currency markets. When trading strategies use direction, the potential to anticipate price change is crucial.

To further illustrate the forecasting efficiency of aFIM, Fig. 2 shows an example of the actual and forecast values developed for Bitcoin (BTC) daily lowest and highest prices from July to December 2019. Prices are in logarithmic values to give better visualization. aFIM is able to accurately predict both the lowest and highest price dynamics of the BTC during the period evaluated. Because the differences between lowest and highest values give an efficient measure of volatility, aFIM appears as a potentially strategic tool to manage risk in digital coin markets, especially in real-time situations.

Table 2 MAPE, RMSSE and DA performance for one-step-ahead forecasting of the daily cryptocurrencies lowest and highest prices

Error measure	ARIMA	ETS	ANFIS	LSTM	RW	aFIM _M	aFIM _H
<i>Bitcoin</i>							
MAPE ^L	0.0205	0.0197	0.0199	0.0186	0.0200	0.0182	0.0177
MAPE ^U	0.0212	0.0200	0.0198	0.0195	0.0199	0.0188	0.0182
RMSSE ^L	1.0068	0.9997	1.0056	1.0023	0.9986	0.9961	0.9957
RMSSE ^U	1.0071	0.9986	1.0016	0.9980	0.9986	0.9969	0.9952
DA ^L	0.5110	0.5852	0.5860	0.5987	–	0.6019	0.6166
DA ^U	0.5385	0.5440	0.5490	0.5492	–	0.5671	0.5542
<i>Ethereum</i>							
MAPE ^L	0.0249	0.0244	0.0239	0.0239	0.0245	0.0231	0.0234
MAPE ^U	0.0243	0.0246	0.0244	0.0240	0.0239	0.0241	0.0240
RMSSE ^L	1.0066	1.0026	1.0022	0.9980	0.9986	0.9970	0.9965
RMSSE ^U	0.9997	1.0041	0.9855	0.9987	0.9989	0.9973	0.9970
DA ^L	0.5000	0.5247	0.5265	0.5563	–	0.5798	0.5911
DA ^U	0.5719	0.5440	0.5610	0.5700	–	0.5800	0.6009
<i>XRP</i>							
MAPE ^L	0.0232	0.0196	0.0197	0.0192	0.0192	0.0184	0.0187
MAPE ^U	0.0244	0.0250	0.0246	0.0230	0.0215	0.0198	0.0207
RMSSE ^L	1.0782	1.0053	1.0032	0.9987	0.9986	0.9953	0.9951
RMSSE ^U	1.0014	1.1244	1.1230	1.0928	0.9994	0.9960	0.9968
DA ^L	0.4753	0.4863	0.4980	0.5462	–	0.5911	0.5877
DA ^U	0.5385	0.5055	0.5198	0.5672	–	0.6166	0.6013
<i>LiteCoin</i>							
MAPE ^L	0.0278	0.0271	0.0269	0.0252	0.0271	0.0253	0.0249
MAPE ^U	0.0304	0.0290	0.0287	0.0277	0.0283	0.0275	0.0271
RMSSE ^L	1.0049	0.9996	0.1008	0.9970	0.9984	0.9966	0.9962
RMSSE ^U	1.0104	1.0087	1.0074	0.9988	0.9983	0.9969	0.9964
DA ^L	0.4863	0.5604	0.5701	0.5799	–	0.5811	0.5890
DA ^U	0.5549	0.5110	0.5230	0.5722	–	0.5880	0.5913

Fig. 2 BitCoin actual and forecast values of aFIM lowest and highest daily log-prices

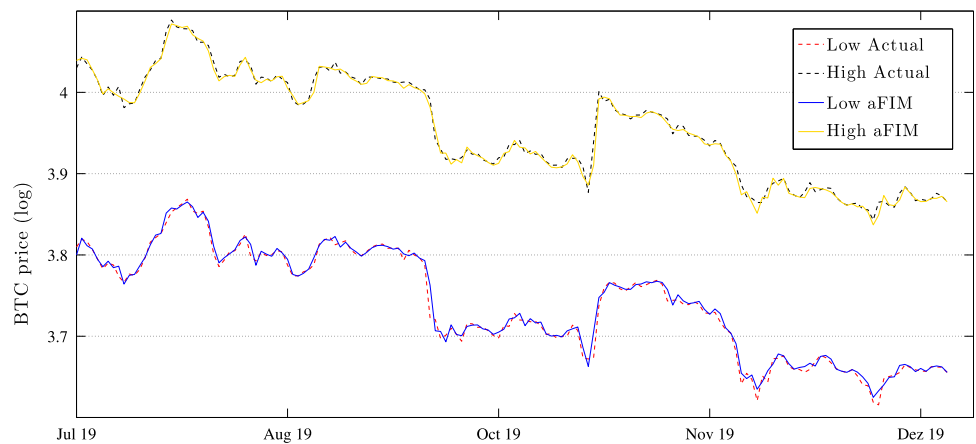


Table 3 NSD, R^C and R^E for one-step-ahead forecast of cryptocurrencies lowest and highest prices

Error measure	ARIMA	ETS	ANFIS	LSTM	RW	aFIM _M	aFIM _H
<i>Bitcoin</i>							
NSD	0.6507	0.6114	0.6018	0.5487	0.5926	0.5419	0.5013
R^C	0.5418	0.5677	0.5770	0.6013	0.6092	0.6218	0.6330
R^E	0.6812	0.5971	0.5955	0.6419	0.5805	0.7116	0.7382
<i>Ethereum</i>							
NSD	0.6086	0.6029	0.6009	0.5877	0.5840	0.5573	0.5607
R^C	0.5945	0.5899	0.5911	0.6123	0.6165	0.6450	0.6388
R^E	0.5761	0.5808	0.5917	0.6099	0.5781	0.6098	0.6213
<i>XRP</i>							
NSD	0.6899	0.6184	0.5811	0.5228	0.5541	0.5019	0.4881
R^C	0.5098	0.5718	0.5884	0.5910	0.6483	0.6981	0.7013
R^E	0.5448	0.5514	0.5619	0.6275	0.6052	0.6125	0.6441
<i>LiteCoin</i>							
NSD	0.6228	0.5976	0.6018	0.5317	0.5777	0.5211	0.5008
R^C	0.5827	0.5916	0.6012	0.6344	0.6258	0.6775	0.6810
R^E	0.5447	0.5978	0.6009	0.6177	0.5926	0.6309	0.6566

Better results are highlighted in bold

Forecasting performance, considering the interval nature of the data, measured by the normalized symmetric divergence (NSD), coverage rate (R^C), and the efficiency rate (R^E), are shown in Table 3. Again, aFIM outperforms all the remaining methods. Table 3 also indicates that when forecasts are evaluated by interval-oriented metrics, aFIM achieves superior performance, especially when it uses the generalized Hukuhara subtraction (aFIM_H).

The average number of fuzzy rules that aFIM developed using the out-of-sample data (the year of 2019) when using Moore/Hukuhara subtraction operator were 4/4 (Bitcoin), 4/3 (Ethereum), 3/3 (XRP) and 3/2 (LiteCoin). It shows that model complexity varies slightly depending on the dynamic nature of the the digital coin, and that generally model complexity is insensitive to the choice of either Moore or Hukuhara subtraction.²

5 Conclusion

This paper has introduced a novel adaptive fuzzy interval modeling and forecasting method for interval-valued time series. The model collects functional fuzzy rules with a learning mechanism that continuously updates the rule base structure and its parameters using interval stream input data. The structure of the rule base is found employing participatory learning clustering, and parameters of the

fuzzy rules consequents are estimated using the weighted recursive least squares algorithm. Experimental evaluation was done using an actual interval time series of cryptocurrency prices whose lowest and highest prices are the low and high interval bounds. One-step-ahead forecasts of the interval-valued prices of Bitcoin, Ethereum, XRP, and Litecoin for the period from January 2016 to December 2019 were produced. Comparison of the adaptive fuzzy interval model aFIM with ARIMA, ETS, ANFIS, LSTM and naïve random walk indicate that aFIM outperforms all of them. The performance of aFIM also is the highest when performance is measured by the direction of price changes. This is a key feature, especially when trading with the price and direction-based strategies. Future work shall consider autonomous mechanisms to select and adjust control parameters such as thresholds in real-time, the evaluation of the results through a trading strategy based on forecasts, and model and forecast interval time series from commodities markets, energy, environment, and transportation systems.

Acknowledgements The authors are grateful to the Brazilian National Council for Scientific and Technological Development (CNPq) for Grants 304456/2020-9 (Leandro Maciel) and 302467/2019-0 (Fernando Gomide). The comments and suggestions of the reviewers are also kindly acknowledged.

Declaration

Conflict of interest The authors declare that they have no conflict of interest.

² Examples of the fuzzy rules are not shown. They are available upon request.

References

- Angelov P (2010) Evolving Takagi–Sugeno fuzzy systems from data streams (eTS+). In: Angelov P, Filev D, Kasabov N (eds) *Evolving intelligent systems: methodology and applications*. Wiley and IEEE Press, Hoboken, pp 21–50
- Angelov P, Filev D (2004) An approach to online identification of Takagi–Sugeno fuzzy models. *IEEE Trans Syst Man Cybern B* 34(1):484–498
- Arroyo J, Espínola R, Maté C (2011) Different approaches to forecast interval time series: a comparison in finance. *Comput Econ* 27(2):169–191
- Billard L, Diday E (2000) Regression analysis for interval-valued data. In: *Data analysis, classification and related methods: proceedings of the 7th conference of the IFCS*, pp 369–374
- Burns K, Moosa I (2015) Enhancing the forecasting power of exchange rate models by introducing nonlinearity: does it work? *Econ Model* 50:27–39
- Corbet S, Meegan A, Larkin C, Lucey B, Yarovaya L (2018) Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Econ Lett* 165:28–34
- Fagundes R, De Souza R, Cysneiros F (2014) Interval kernel regression. *Neurocomputing* 128:371–388
- Froelich W, Salmeron J (2014) Evolutionary learning of fuzzy grey cognitive maps for the forecasting of multivariate, interval-valued time series. *Int J Approx Reason* 55:1319–1335
- Hajek P, Froelich W, Prochazka F (2020) Intuitionistic fuzzy grey cognitive maps for forecasting interval-valued time series. *Neurocomputing* 400:173–185
- Hamilton J (1994) *Time series analysis*. Princeton University Press, Princeton
- Huttenlocher D, Klanderman G, Rucklidge W (1993) Comparing images using the Hausdorff distance. *IEEE Trans Pattern Anal Mach Intell* 15(9):850–863
- Hyndman R, Koehler A, Ord J, Snyder R (2008) *Forecasting with exponential smoothing: the state space approach*. Springer, Berlin
- Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23(3):665–685
- Li B, Shen Y, Li B (2008) A new algorithm for computing the minimum Hausdorff distance between two point sets on a line under translation. *Inf Process Lett* 106(2):52–58
- Lim C (2016) Interval-valued data regression using nonparametric additive models. *J Korean Stat Soc* 45(3):358–370
- Lima Neto E, de Carvalho F (2010) Constrained linear regression models for symbolic interval-valued variables. *Comput Stat Data Anal* 54(2):333–347
- Ljung L (1988) *System identification: theory for the user*. Prentice-Hall, Englewood Cliffs
- Maciel L, Ballini R, Gomide F, Yager RR (2016) Participatory learning fuzzy clustering for interval-valued data. In: *Proceedings of the 16th international conference on information processing and management of uncertainty in knowledge-based systems (IPMU 2016)*. Eindhoven, pp 1–8
- Maciel L, Gomide F, Ballini R (2015) Evolving possibilistic fuzzy modeling for financial interval time series forecasting. In: *Proceedings of annual conference on North American fuzzy information processing society (NAFIPS) held jointly 5th world conference on soft computing (WConSC)*, pp 1–6
- Maia A, de Carvalho F (2011) Holt’s exponential smoothing and neural network models for forecasting interval-valued time series. *Int J Forecast* 27(3):740–759
- Maia A, de Carvalho F, Ludermir T (2008) Forecasting models for interval-valued time series. *Neurocomputing* 71(16–18):3344–3352
- Meese R, Rogoff K (1983) Empirical exchange rate models of the seventies: do they fit out of sample? *J Int Econ* 14(1–2):3–24
- Moore R, Kearfott R, Cloud M (2009) *Introduction to interval analysis*. SIAM Press, Philadelphia
- Moosa I, Burns K (2014) The unbeatable random walk in exchange rate forecasting: reality or myth? *J Macroecon* 40:69–81
- Rodrigues P, Salish N (2015) Modeling and forecasting interval time series with threshold models. *Adv Data Anal Class* 9(1):41–57
- Roque A, Maté C, Arroyo J, Sarabia A (2007) iMLP: applying multi-layer perceptrons to interval-valued data. *Neural Process Lett* 25(2):157–169
- Silva L, Gomide F, Yager R (2005) Participatory learning in fuzzy clustering. In: *IEEE International conference on fuzzy systems*, Reno, pp 857–861
- Stefanini L (2010) A generalization of Hukuhara difference and division for interval and fuzzy arithmetic. *Fuzzy Set Syst* 161:1564–1584
- Sun Y, Han A, Hong Y, Wang S (2018) Threshold autoregressive models for interval-valued time series data. *J Econom* 206:414–446
- Sun Y, Zhang X, Hong Y, Wang S (2018) Asymmetric pass-through of oil prices to gasoline prices with interval time series modelling. *Energy Econ* 78:165–173
- Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans Syst Man Cybern SMC–15(1):116–132*
- Tao X, Li C, Bao Y, Hu Z, Lu Z (2015) A combination method for interval forecasting of agricultural commodity futures prices. *Knowl Based Syst* 77:92–102
- Wang X, Li S (2011) The interval autoregressive time series model. In: *Proceedings IEEE international conference fuzzy system (FUZZ-IEEE)*, pp 2528–2533
- Xiong T, Bao Y, Hu Z (2014) Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting. *Knowl Based Syst* 55:87–100

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.