International Journal of Statistics and Applications 2015, 5(6): 317-337
DOI: 10.5923/j.statistics.20150506.07

Comparison of Forecasting Methods for Interval-Valued Time Series

Ebrucan Islamoglu[1,*], Alı Islamoglu[2], Hasan Bulut[1]

[1]Department of Statistics, Faculty of Arts and Sciences, Ondokuz Mayıs University, Samsun, Turkey
[2]Strategy development authority, Ondokuz Mayıs University, Samsun, Turkey

Abstract This study examines the interval-valued time series which is an ongoingness issue in time series. The study aims to obtain new time series forecasting methods using different combination of several analysis methods and modeling techniques and to determine the methods and models that provide the optimal accuracy by comparing the forecasting accuracy of the proposed methods. Different approaches and appropriate modeling techniques are used for analyzing interval-valued time series. The data of experimental study is obtained by analyzing the interval-valued time series via forecasting methods. These values represent variables in the variance analysis of our study. The obtained results are analyzed by using statistical analysis. Kruskal-Wallis H test and Mann-Whitney U-test are applied to evaluate the performance. Which one of the methods provided better forecasting results and their pros and cons are examined. Four real time series are used in the implementation and the forecasting performance of the methods are compared and evaluated.

Keywords Interval-Valued Time Series, Kruskal-Wallis H Test, Mann-Whitney U-Test, Time Series, Variance Analysis

1. Introduction

Forecasting in time series analysis is extremity important both national economy and preproduction. Time series are commonly used in economic magnitude analysis, population forecasting and all other branches of science. So that it comes into prominence day after day. The criterion that increases the performance comes into prominence in various fields. Prediction modelings are used in many scientific fields. Interval-valued data represent uncertainty (for instance, confidence intervals), variability (minimum and maximum of daily temperature), etc. The field of interval analysis assumes that observations and estimations in the real world are usually incomplete and, consequently, do not precisely represent the real data. Interval-valued time series are interval-valued data collected in a chronological sequence. Maia and De Carvalho [1] stated that interval-valued time series could be seen in various different fields, e.g. economics (daily stock price of a company, which can be expressed as the lowest and highest trading prices over the day); engineering (the variation in an electric current, expressed by the lowest and highest intensities over a given day); medicine (diastolic and systolic blood pressure in a given day); weather (maximum and minimum rainfall in a month for a particular place); and relative humidity (also measured as the highest and lowest values in a given month).

Billard and Diday [2], Bock and Diday [3] and Diday and Noirhomme Fraiture [4] stated that interval-valued data have also been considered in the field of symbolic data analysis (SDA). This field is related to multivariate analysis, pattern recognition and artificial intelligence, and aims to extend classical exploratory data analysis and statistical methods to symbolic data. Symbolic data allow multiple (sometimes weighted) values for each variable and new variable types (set-valued, interval-valued and histogram valued variables) have been introduced. These new variables make it possible to take into account the variability and/or uncertainty present in the data. Symbolic Data Analysis (SDA) has been introduced by Billard and Diday [5] as a domain related to multivariate analysis, pattern recognition and artificial intelligence in order to introduce new methods and to extend classical data analysis techniques and statistical methods to symbolic data.

Various different approaches have been introduced for analyzing interval-valued data. A number of authors have considered neural network models in order to manage interval-valued data. For example, Beheshti et al. [6] propose a three-layer perceptron in which inputs, weights, biases and outputs are intervals, band show how to obtain the optimal weights and biases for a given training data set by means of interval computation algorithms. Patinŏ-Escarcina et al. [7] describe a one-layer perceptron for classification tasks in which inputs, weights and biases are represented by intervals. Roque et al. [8] propose and analyze a new multilayer perceptron model based on

* Corresponding author:
ebrucan.tiring@omu.edu.tr (Ebrucan Islamoglu)
Published online at http://journal.sapub.org/statistics

interval arithmetic which facilitates the handling of input and output interval data, but which has single valued rather than interval valued weights and biases.

In the field of SDA, several applications for managing interval-valued data have been provided: Ichino et al. [9] introduced a symbolic classifier as a region-oriented approach for symbolic interval data. Rasson and Lissoir [10] presented a symbolic kernel classifier based on dissimilarity functions suitable for symbolic interval data. Périnel and Lechevallier [11] proposed a tree-growing algorithm for classifying symbolic interval data. Bock [12] proposed several clustering algorithms for symbolic data described by interval variables and presented a sequential clustering and updating strategy for constructing a self-organizing map (SOM) to visualize symbolic interval data. Chavent and Lechevallier [13] proposed a dynamic clustering algorithm for interval data where the class representatives are defined by an optimality criterion based on a modified Hausdorff distance. Central tendency and dispersion measures are extended by Billard and Diday [5] and Chavent and Saracco [14]; factorial methods by Irpino [15]; multidimensional scaling by Groenen et al. [16]; hierarchical clustering by Gowda and Diday [17], Ichino and Yaguchi [18], Gowda and Ravi [19, 20], Chavent [21], and Guru et al. [22]; fuzzy by El-Sonbaty and Ismail [23], De Carvalho [24] and Yang et al. [25]; hard partition clustering by De Carvalho et al. [26], De Carvalho and Lechevallier [27,28], De Carvalho et al. [29], De Souza and De Carvalho [30], and Irpino and Verde [31]; regression by Lima Neto and De Carvalho [32]; and forecasting by Arroyo and Mat´e [33].

Souza and De Carvalho [34] presented partitioning clustering methods for interval data based on (adaptive and non-adaptive) city-block distances. More recently, De Carvalho et al. [35] have proposed an algorithm using an adequacy criterion based on adaptive Hausdorff distances. De Carvalho [36] proposed histograms for interval-valued data. Concerning factorial methods, Cazes et el. [37] and Lauro and Palumbo [38] proposed principal component analysis methods suitable for interval-valued data. According to Diday [39] statistical units described by interval data can be assumed as special cases of Symbolic Objects (SO). Currently, different approaches have been introduced to analyze symbolic interval data. Regarding univariate statistics, Bertrand and Goupil [40] and Billard and Diday [5] introduced central tendency and dispersion measures suitable for symbolic interval data. Billard and Diday [41] presented the first approach to fit a linear regression model on interval-valued data sets. Their approach consisted of fitting a linear regression model on the midpoint of the interval values assumed by the variables in the learning set and to apply this model on the lower and upper boundaries of the interval values of the explanatory variables to predict, respectively, the lower and upper boundaries of the interval values of the dependent variable. Lima Neto and De Carvalho [42] improved the former approach presenting a new method based on two linear regression models, the first regression model over the midpoints of the intervals and the second one over the ranges, which reconstruct the boundaries of the interval-values of the dependent variable in a more efficient way when compared with the Billard and Diday's method. Palumbo and Verde [43] and Lauro et al. [44] generalized factorial discriminant analysis (FDA) to interval-valued data. Concerning interval-valued time series, Maia et al. [45] have introduced autoregressive integrated moving average (ARIMA), artificial neural network (ANN) as well as a hybrid methodology that combines both ARIMA and ANN models in order to forecast interval-valued time series. Silva [46] used copulas approach to the regression models for interval-valued data attack the problem from an optimization point of view. Maia and De Carvalho [1] introduced three approaches to forecasting interval-valued time series. The first two approaches are based on multilayer perceptron (MLP) neural networks and Holt's exponential smoothing methods, respectively. In Holt's method for interval valued time series, the smoothing parameters are estimated by using techniques for non-linear optimization problems with bound constraints. The third approach is based on a hybrid methodology that combines the MLP and Holt models. The practicality of the methods is demonstrated through simulation studies and applications using real interval-valued stock market time series.

This study examines the interval-valued time series which has been an ongoingness issue in time series. The study aims to obtain new time series forecasting methods using different combination of several analysis methods and modeling techniques and to determine the methods and models that provide the optimal accuracy by comparing the forecasting accuracy of the proposed methods. To do this, four real time series are used. These time series are Usage of Gold Exchange I-Istanbul (Weekday, Turkish Liras/US dollar), Usage of Gold Exchange II-Istanbul (Weekday, Turkish Liras/US dollar), Cost of Living Index (Wage Earners) (1995=100) and Selling Rate of Exchange-Euro (Exchange Selling). Obtained results represent the dependent variables in the variance analysis of our study. Independent variables involve forecasting methods. In the next section, brief information about interval-valued time series processing in time series analysis is given. By utilizing tables, Section 3 presents the results obtained from the implementation in which four real time series are analyzed. In the last section, the obtained results are summarized, interpreted and discussed.

2. Interval-Valued Time Series Processing

Maia and De Carvalho [1] reported that classical statistics and data analysis deal with individuals who are described by usual variables that assume a single value for a given individual: either a real value (for a quantitative variable) or a category from a set of alternatives (for either a nominal or an ordinal qualitative variable). Symbolic variables generalize this classical paradigm, as the value of a symbolic variable may be a subset of a set of categories, an interval of R, or a histogram for an underlying classical variable. In this paper, we are concerned with interval-valued variables. In the context of SDA, Bock and Diday [3] used an interval-valued variable X is a correspondence defined from Ω (the set of individuals) into R, such that, for each $k \in \Omega$, $X(k) = [a, b] \in \Im$, where $\Im = \{[a, b] : a, b \in R, a \leq b\}$ is the set of closed intervals defined from R. Maia et al.[45] indicated that when interval-valued variables are collected in an ordered sequence over time, we say that we have an interval-valued time series. In the development of this work, we consider that, at each point in time (t = 1, 2, . . . , n), an interval is described as a two-dimensional vector $[XL_t, XU_t]$ with components in R representing the upper boundary XU_t and the lower boundary XL_t, with $XL_t \leq XU_t$. Thus, an ITS (Interval-valued time series) is

$$[X_{L_1}; X_{U_1}],\ [X_{L_2}; X_{U_2}],\ \cdots, [X_{L_n}; X_{U_n}],$$

where n denotes the number of intervals of the time series (sample size). Specifically, an observed interval at time t is it, and is represented as

$$\mathbf{I_t} = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix}.$$

Thus, an ITS is a set of intervals $\mathbf{I_t}$, each recorded at a specific time *t* and generally time equidistant.

We often come across interval-valued variables in practice. For instance, an interval may describe the smallest and largest values of a measurement for an individual over a day, or the range of monthly income values in a city. Fig. 1 illustrates two typical situation which an ITS arises. At the top of Fig. 1, the ITS on the right is the series of temperature intervals, which are obtained at each instant in time from the minimal and maximal values of the time series of temperatures on the left. At the bottom of Fig. 1, the ITS on the right is the series of intervals of stock prices, which are obtained at each instant in time from the highest and lowest trading values of the stock options recorded monthly.

2.1. Constructing Models for Interval-Valued Time Series Forecasting

2.1.1. Approach 1

Maia et al. [45] suggest that the upper and lower bounds (XU_t - the upper bounds of the interval, XL_t –the lower bounds of the interval) of interval-valued time series predicted by neural networks separately.

In approach 1, forecasts for the lower and upper bounds obtain from these models, respectively.

$$XU_t = f_1(XU_{t-1},\ XU_{t-2}, \ldots\ XU_{t-m})$$
$$XL_t = f_2(XL_{t-1},\ XL_{t-2}, \ldots, XL_{t-m})$$

In the precedent expression, f_1 ve f_2 are a non-linear function determined by the neural networks, respectively.

2.1.2. Approach 2

Maia et al. [45] suggest that the upper and lower bounds of interval-valued time series predicted by neural networks together. In approach 2 forecasts for the lower and upper bounds obtain from this model. In this approach only one function determined by the neural networks and obtained the lower and upper bounds.

$$[XL_t, XU_t] = f_3(XU_{t-1},\ XU_{t-2}, \ldots, XU_{t-m}, XL_{t-1}, XL_{t-2}, \ldots, XL_{t-m})$$

In the precedent expression f_3 is a non-linear function determined by the neural networks. In this approach, the inputs of neural networks the lower and the upper bounds of lagged variables of time series

$$(XU_{t-1}, XU_{t-2}, \ldots, XU_{t-m}, XL_{t-1}, XL_{t-2}, \ldots,\ XL_{t-m}),$$

outputs are $[XU_t, XL_t]$ values.

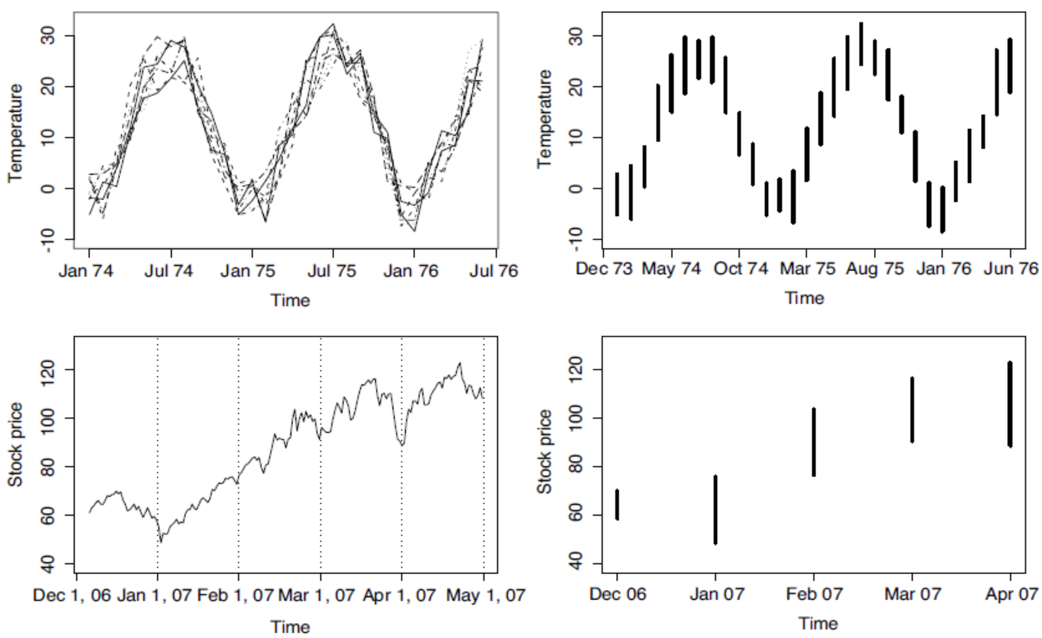

Figure 1. Two interval-valued time series (right hand side) obtained from a set of usual time series (left hand side)

2.1.3. Approach 3

Lima Neto and De Carvalho [47] presented two time series. These are the interval centre series XC_t and the range interval series XR_t, respectively, represent the centre and the range interval series by,

$$XC_t=\frac{XU_t\text{-}XL_t}{2} \quad \text{and} \quad XR_t=\frac{XU_t\text{-}XL_t}{2}$$

XC_t and XR_t time series are analyzed by using feed forward neural networks, separately. In this circumstance, the forecasts of XC_t and XR_t time series obtained from two models given below.

$$XC_t= f_4(XC_{t-1}, XC_{t-2}, \ldots, XC_{t-m})$$

$$XR_t=f_5(XR_{t-1}, XR_{t-2}, \ldots, XR_{t-m})$$

In the methods presented here, f_4 and f_5 are respectively a non-linear function determined by the neural network of lagged variables of mid-point and range series. Thus, the values predicted by these models for the lower and upper bounds of the interval, given by.

$$XL_t=XC_t\text{-}XR_t$$

$$XU_t=XC_t+XR_t$$

2.2. Modeling Techniques Used in Approaches

2.2.1. Box-Jenkins (ARIMA) Models

An often-used methodology in handling and predicting time series is known as the Box-Jenkins method or simply ARIMA. An autoregressive (AR) model is simply a model used to find an estimation of a variable based on previous input values of the variable. The actual equation for the AR model is as follows:

$$y_t= \theta_0 +\sum_{i=1}^{p} \emptyset_i y_{t-i}+\varepsilon_t$$

where y_t is the current value of the time series at time t. The $\emptyset_i$ (i=1,2,…,p) are the model parameters to be estimated. The model consists of three parts: a constant part θ_o, a random error part ε_t (white noise) and the AR summation. The parameter p represents the order of the model AR(p).

Autoregressive moving average (ARMA) models are created from a finite, linear combination of past values of the series and a finite linear combination of past errors. The particular model that will be used in the present paper, known as ARMA(p; q), is represented as follows:

$$y_t=\theta_0+\sum_{i=1}^{p} \emptyset_i y_{t-i} +\varepsilon_t+\sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

where ε_t is the random error at time t; $\emptyset_i$ (i=1; 2; . . . ; p) and θ_j (j=1,2,…,q) are the model parameters to be estimated; p and q refer to the order of the model; the random errors ε_t are assumed to be independent and identically distributed with a zero mean and σ^2 constant variance. In practice, one applies the ARMA process not to the original time series, but to the transformed time series. Often, the time series of differences is stationary despite the non-stationary of the underlying process. Stationary time series can be well estimated by the ARMA model. This leads to the definition of the ARIMA model:

$\Delta^d y_t$ is ARMA(p; q) process $\rightarrow$ y_t is ARIMA(p; d; q) process,

where $\Delta^d y_t$ is the d order differencing operator. The differencing operator is applied to the time series until it becomes stationary. Thus, an ARIMA (p; d; q) process models the stationary differences of the order d of the time series y_t using the ARMA (p; q) process.

2.2.2. Neural Networks for Interval-Valued Time Series

Zhang et al. [48] mention that there is a huge range of different types of artificial neural networks, but the most popular type is the multilayer perceptron (MLP). In particular, feed-forward MLP networks with two layers (one hidden layer and one output layer) are often used for modeling and forecasting time series. Approaches based on artificial neural networks have also been proposed for the non-linear modeling of time series.

Kaastra and Boyd [49] proposed the relationship between the output, y_t, and the inputs y_{t-1}, y_{t-2},…,y_{t-p}, is as follows:

$$\hat{y_t}=\alpha_o+\sum_{j=1}^{q} \alpha_j \cdot g(\beta_{oj} + \sum_{i=1}^{p} \beta_{ij} y_{t-i})$$

in which α_o and β_{oj} denote the weights of the connection between the constant input (bias) and the output, and between the bias and hidden nodes, respectively; and where α_j and β_{ij} are the weights associated with each node; p is the number of inputs; q is the number of hidden nodes; and g denotes the transfer function used in the hidden layer. Transfer functions such as the logistic g(u) = 1/{1+exp(-u)} are commonly used for time series data, as they are non-linear and continuously differentiable, which are desirable properties for network learning.

In this paper, the neural network used is a MLP network for ITS processing (MLP^1), and is based on typical MLP networks for time series data. For ITS processing and forecasting, we use a MLP network with two feed-forward layers with 2p inputs (lagged intervals at *t-1,…, t-p),* q nodes in the hidden layer, and two output nodes, with each output corresponding to the forecasting of the boundaries, X_t^L and X_t^U.

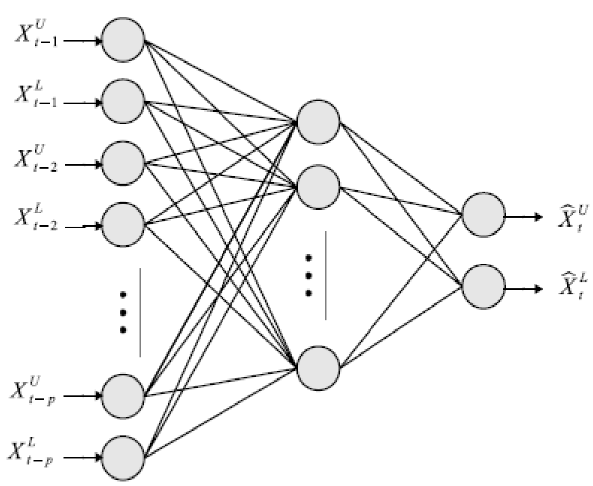

Figure 2. Neural network for interval-valued time series processing, with 2p inputs and one hidden layer with q nodes and two outputs

Figure 2 displays the MLP^1 structure suitable for ITS processing. Note that the $2p$ inputs are used as a lagged ITS, i.e.,

$$\{\mathbf{I}_{t-1},\mathbf{I}_{t-2},\ldots,\mathbf{I}_{t-p}\}=\left\{(X^U_{t-1},X^L_{t-1})^T,(X^U_{t-2},X^L_{t-2})^T,\ldots,(X^U_{t-p},X^L_{t-p})^T\right\}$$

There is a constant input unit, called a bias node, connected to every node in the hidden layer, as well as to the output nodes.

Let φ_{ij} and ω_j be two-dimensional vectors,

$$\varphi_{ij}=\begin{bmatrix}\varphi^U_{ij}\\ \varphi^L_{ij}\end{bmatrix} \quad \text{and} \quad \omega_j=\begin{bmatrix}\omega^U_j\\ \omega^L_j\end{bmatrix}$$

for a MLP^1 network with one hidden layer of q nodes. The general prediction equation for computing forecasts of X^L_t and X^U_t (two outputs) using selected past intervals $I_{t-1},\ldots, I_{t-p}$ as the inputs is written in the following form:

$$\widehat{I_t}=\begin{bmatrix}\widehat{X_t}^U\\ \widehat{X_t}^L\end{bmatrix}$$

$$=\omega_0+\sum_{j=1}^{q}\omega_j\cdot\begin{bmatrix}g(\varphi^U_{0_j}+\sum_{i=1}^{p}(\varphi_{ij})^T I_{t-i})\\ g(\varphi^L_{0_j}+\sum_{i=1}^{p}(\varphi_{ij})^T I_{t-i})\end{bmatrix}$$

$$=\begin{bmatrix}\omega^U_0+\sum_{j=1}^{q}\omega^U_j\cdot g(\varphi^U_{O_J}+\sum_{i=1}^{p}(\varphi^U_{ij}X^U_{t-i}+\varphi^L_{ij}X^L_{t-i}))\\ \omega^L_0+\sum_{j=1}^{q}\omega^L_j\cdot g(\varphi^L_{O_J}+\sum_{i=1}^{p}(\varphi^U_{ij}X^U_{t-i}+\varphi^L_{ij}X^L_{t-i}))\end{bmatrix}$$

where φ denotes the weights for the connections between the inputs (lagged intervals) and the hidden nodes; ω denotes the weights between the hidden nodes and the output nodes; and is the Hadamard product.

We train the MLP using conjugate gradient error minimization. The conjugate gradient approach finds the optimal weight vector along the current gradient by doing a line-search. Details of the conjugate gradient algorithm are given by Bishop [50].

2.2.3. Holt's Method for Interval-Valued Time Series

The exponential smoothing method has become very popular due to its simplicity and good overall performance. Common applications range from business tasks (e.g., forecasting sales or stock fluctuations) to environmental studies (e.g., measurements of atmospheric components or rainfall data)—with typically no more a priori knowledge than the possible existence of trends in seasonal patterns. Such methods are sometimes also called naive, because no covariates are used in the models, i.e., the data are assumed to be self-explanatory. Their success is rooted in the fact that they belong to a class of local models that adapt their parameters to the data automatically during the estimation procedure, and therefore implicitly account for (slow) structural changes in the training data. Moreover, as the influence of new data is controlled by hyperparameters, this has the effect of smoothing the original time series.

Holt's method proposed by Holt [51] is viewed as a common extension of the simple ES with an additional trend component. Two components must be updated each period—*the level* and *the trend.* The level is a smoothed estimate of the value of the data at the end of each period, while the trend is a smoothed estimate of the average growth at the end of each period. The specific formulae for Holt's method are:

$$\widehat{L_t}=\alpha y_t+(1-\alpha)(\widehat{L_{t-1}}+\widehat{T_{t-1}})$$

$$\widehat{T_t}=\beta(\widehat{L_t}-\widehat{L_{t-1}})+(1-\beta)\widehat{T_{t-1}}$$

where $0<\alpha,\ \beta<1$ are the smoothing parameters. In this model, α controls the length of the average for the estimation of the level and β controls the smoothing of the trend.

A straightforward approach to finding the optimal values of both α and β-constrained to be in the range (0, 1)—is to look for the parameter combination that minimizes the sum of squared errors of the one-stepahead predictions,

$$\sum_{t=3}^{n}(y_t-\widehat{y_t})^2$$

where $\widehat{y_{t+1}}=\widehat{L_t}+\widehat{T_t}$. The starting values for $\widehat{L_2}$ and $\widehat{T_2}$ are typically taken as being y_2 and y_2-y_1, respectively. More details are given by Gardner [52].

Numerous variations of the original exponential smoothing methods have been proposed. For example, Williams and Miller [53] proposed modifications for dealing with discontinuities. In this paper, the interval Holt's exponential smoothing method (Holt') follows a similar representation for usual quantitative data, and has the following form:

$$\widehat{L_t}'=AI_t+(I-A)(\widehat{L_{t-1}'}+\widehat{T_{t-1}'})$$

$$\widehat{T_t}'=B(\widehat{L_t}'-\widehat{L_{t-1}'})+(1-B)\widehat{T_{t-1}}'$$

where

$$\widehat{L_t}'=\begin{bmatrix}\widehat{L^u_t}\\ \widehat{L^L_t}\end{bmatrix} \quad \text{and} \quad \widehat{T_t}'=\begin{bmatrix}\widehat{T^u_t}\\ \widehat{T^L_t}\end{bmatrix} \quad \text{and} \quad \widehat{I_{t+1}}=\widehat{L_t}'+\widehat{T_t}'$$

A and B denote the (2 × 2) smoothing parameter matrices,

$$A=\begin{bmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{bmatrix} \quad \text{and} \quad B=\begin{bmatrix}\beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{bmatrix}$$

and I is a (2 × 2) identity matrix.

The Holt' method is given by

$$\widehat{L_t'}=\begin{bmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{bmatrix}\begin{bmatrix}X^u_1\\ X^L_1\end{bmatrix}+\begin{bmatrix}1-\alpha_{11} & -\alpha_{12}\\ -\alpha_{21} & 1-\alpha_{22}\end{bmatrix}\begin{bmatrix}\widehat{L^U_{t-1}}+\widehat{T^U_{t-1}}\\ \widehat{L^L_{t-1}}+\widehat{T^L_{t-1}}\end{bmatrix}$$

$$=\begin{bmatrix}\alpha_{11}X^u_t+(1-\alpha_{11})\left(\widehat{L^U_{t-1}}+\widehat{T^U_{t-1}}\right)+\alpha_{12}\left(X^L_t-\widehat{L^L_{t-1}}-\widehat{T^L_{t-1}}\right)\\ \alpha_{22}X^L_t+(1-\alpha_{22})\left(\widehat{L^L_{t-1}}+\widehat{T^L_{t-1}}\right)+\alpha_{21}(X^U_t-\widehat{L^U_{t-1}}-\widehat{T^U_{t-1}})\end{bmatrix}$$

and

$$\widehat{T_t'} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} \widehat{L_t^U} - \widehat{L_{t-1}^U} \\ \widehat{L_t^L} - \widehat{L_{t-1}^L} \end{bmatrix} + \begin{bmatrix} 1-\beta_{11} & -\beta_{12} \\ -\beta_{21} & 1-\beta_{22} \end{bmatrix} \begin{bmatrix} \widehat{T_{t-1}^U} \\ \widehat{T_{t-1}^L} \end{bmatrix}$$

$$= \begin{bmatrix} \beta_{11}(\widehat{L_t^U} - \widehat{L_{t-1}^U}) + (1-\beta_{11})(\widehat{T_{t-1}^u}) + \beta_{12}(\widehat{L_t^L} - \widehat{L_{t-1}^L} - \widehat{T_{t-1}^L}) \\ \beta_{22}(\widehat{L_t^L} - \widehat{L_{t-1}^L}) + (1-\beta_{22})\widehat{T_{t-1}^L} + \beta_{21}(\widehat{L_t^U} - \widehat{L_{t-1}^U} - \widehat{T_{t-1}^U}) \end{bmatrix}$$

In the Holt' method, the interpretation of the A and B parameter matrices is that they regulate the degree of smoothing. Thus, if A is an identity matrix (i.e. $\propto_{11} = \propto_{22} = 1$ ve $\alpha_{12} = \propto_{21} = 0$), the curves are not smoothed at all, and if A is a null matrix (i.e. $\propto_{11} = \propto_{22} = \alpha_{12} = \propto_{21} = 0$) , the curves are absolutely smoothed; in fact, they show no variation at all. The parameter matrix B controls the smoothing of the trend. Moreover, the α_{12} and β_{12} parameters introduce information on the lower boundaries regarding the fitting of the upper boundaries, whereas the α_{21} and β_{21} parameters introduce information on the upper boundaries regarding the fitting of the lower boundaries.

In this method, the smoothing parameter matrices A and B, with elements constrained to the range (0, 1), can be estimated by minimizing the interval sum of squared one-step-ahead forecast errors:

$$R(A, B) = \sum_{t=3}^{n} (I_t - \widehat{I_t})^T (I_t - \widehat{I_t})$$

$$= \sum_{t=3}^{n} \begin{bmatrix} X_t^U - \widehat{L_{t-1}^U} - \widehat{T_{t-1}^U} \\ X_t^L - \widehat{L_{t-1}^L} - \widehat{T_{t-1}^L} \end{bmatrix}^T * \begin{bmatrix} X_t^U - \widehat{L_{t-1}^U} - \widehat{T_{t-1}^U} \\ X_t^L - \widehat{L_{t-1}^L} - \widehat{T_{t-1}^L} \end{bmatrix}$$

$$= \sum_{t=3}^{n} (X_t^U - \widehat{L_{t-1}^U} - \widehat{T_{t-1}^U})^2 + \sum_{t=3}^{n} (X_t^L - \widehat{L_{t-1}^L} - \widehat{T_{t-1}^L})^2$$

The start vectors for $\widehat{L_2'}$ and $\widehat{T_2'}$ are typically taken to be I_2 and I_2-I_1, respectively. Based on the above objective function, we may write the estimation of the smoothing parameter matrices as a constrained nonlinear programming problem, formulated as:

$$\min_{\propto_{ij},\, \beta_{ij}} R(A, B) \qquad 0 \le \propto_{ij},\, \beta_{ij} \le 1$$

This is the method presented for estimating the optimal weight matrix for the Holt model intervals and for the prediction of ITS. The solution of this problem can be obtained using the *limited* memory *BFGS method for bound constrained optimization* (L-BFGS-B), developed by Byrd et al. [54] This method allows box constraints, that is, each parameter can be given lower and upper boundaries. It is based on the gradient projection method, and a quasi-Newton approach is used: the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. The main goal of the BFGS algorithm is to use an approximate Hessian rather than the true Hessian. Nocedal and Wright [55] provided a comprehensive reference for the previous algorithms. The L-BFGS-B algorithm is implemented in the R software package by R Development Core Team [56].

2.2.4. Vector Autoregressive Models (VAR)

According to Füss [57] in ARIMA models we only derive the actual value from past values for an endogenous variable. However, there is often no theoretical background available. Therefore, we can use Vector Autoregressive (VAR) Models. The single equation approach explains an endogenous variable (e.g. private consumption) by a range of other variables (e.g., disposal income, property, interest rate) Assumption: Explanatory variables are exogenous. But macroeconomic variables are often not exogenous (endogenity problem). In theory exists no assumptions about the dynamic adjustment. Dynamic modeling of the consumption function, e.g. adjustment of consumption behavior to substantial income, occurs only slowly. Multivariate (linear) time series models eliminate both problems. Development of a variable is explained by the development of potential explanatory variables. Value is explained by its own history and simultaneously by considering various variables and their history. E.g. dependency of the long-term interest rate from the short run interest rate: is difficult in univariate time series analysis, because one would need estimations of the development of short run interest rate. Therefore, it is not possible to make good forecast out of such a model. In contrast to a bivariate model, where both, the history of long and short term interest rate are taken into consideration.

With vector autoregressive models it is possible to approximate the actual process by arbitrarily choosing lagged variables. Thereby, one can form economic variables into a time series model without an explicit theoretical idea of the dynamic relations.

2.2.4.1. Modeling of a VAR(1)-Models

Füss [57] stated that the most easy multivariate time series model is the bivariate vector autoregressive model with two dependent variables $y_{1,t}$ and $y_{2,t}$ where t = 1, ..., T. The development of the series should be explained by the common past of these variables. That means, the explanatory variables in the simplest model are $y_{1,t-1}$ and $y_{2,t-1}$. The VAR(1) with lagged values for every variable is determined by:

$$y_{1,t} = \alpha_{11} y_{1,t-1} + \alpha_{12} y_{2,t-1} + \varepsilon_{1,t}$$

$$y_{2,t} = \alpha_{21} y_{1,t-1} + \alpha_{22} y_{2,t-1} + \varepsilon_{2,t}$$

Matrix Notation:

$$y_t = A_1 y_{t-1} + \varepsilon_t$$

$$A_1 = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$$

Assumptions about the Error Terms:

1. The expected residuals are zero:
$E[\varepsilon_{i,t}] = 0$ with i=1,2
2. The error terms are not autocorrelated:
$E[\varepsilon_{i,t}, \varepsilon_{j,t}] = 0$ with $t \neq r$

Interpretation of VAR Models:

VAR-Models themselves do not allow us to make statements about causal relationships. This holds especially when VAR-Models are only approximately adjusted to an unknown time series process, while a causal interpretation requires an underlying economic model. However, VAR-Models allow interpretations about the dynamic relationship between the indicated variables.

2.2.4.2. VAR(p)-Models with more than two Variables

An VAR(p)-Model, with p variables, is given as:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \varepsilon_t$$

If one wants to expand the equation with a trend, intercept or seasonal adjustment, it will be necessary to augment the vector x_t, which includes all the deterministic components, and the matrix B (VARX-Model):

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + Bx_t + \varepsilon_t$$

Estimation of VAR-Models Specifications:

- Determination of endogenous variables according to economic theory, empirical evidence and experience.
- Transformation of time series (take logs or log-returns).
- Insert seasonal component, especially for macro data.
- Control for deterministic terms.

Determination of Lag Length:

The determination of lag length is a trade-off between the curse of dimensionality and abbreviate models, which are not appropriate to indicate the dynamic adjustment. If the lag length is to short, autocorrelation of the error terms could lead to apparently significant and inefficient estimators. Therefore, one would receive wrong results. With the so called curse of dimensionality we understand, that even with a relatively small lag length a large number of parameters if required. On the other hand, with increasing number of parameters, the degrees of freedom decrease, which could possibly result in significant of inefficient estimators.

Information Criteria:

The idea of information criteria is similar to the trade-off discussed above. On the one hand, the model should be able to reflect the observed process as precise as possible (error terms should be as small as possible) and on the other hand, to many variables lead to inefficient estimators. Therefore, the information criteria are combined out of the squared sum of residuals and a penalty term for the number of lags. In detail, for T observations we chose the lag length p in a way that the reduction of the squared residuals after augmenting lag p+1, is smaller than the according boost in the penalty term.

Squared residuals:

$$\text{In}\left(\frac{\varepsilon'\varepsilon}{T}\right)$$

Penalty terms:

AIC : $\frac{2p}{T}$

SIC : $\frac{p.\text{In}(T)}{T}$

HQ : $\frac{2\cdot c\cdot p\cdot \text{In}(T)}{T}$

with constant c > 1.

3. Implementation

The time series which are used in implementation obtained from internet site of Central Bank of the Turkish Republic. These data sets are described below.

Set1: Usage of Gold Exchange I-Istanbul (Weekday, TL/US dollar) between 03/01/2000 and 31/12/2011 dates.

Set2: Usage of Gold Exchange II-Istanbul (Weekday, TL/US dollar) between 03/01/2000 and 31/12/2011 dates.

Set3: Cost of Living Index (Wage Earners) (1995=100) between 03/01/2000 and 31/12/2011 dates.

Set4: Selling Rate of Exchange-Euro (Exchange Selling) between 03/01/2000 and 31/12/2011 dates.

The data of this time series are used in the implementation. This series have **the lowest (US/ounce), the highest (US/ounce), centre (US/ounce)** and **range (US/ounce)** values. To compare the methods, Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Theil's interval statistics (U') and interval average relative variance (ARV') values are calculated by using the following formulas.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

$$\text{MAPE} = \frac{\%100}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right|$$

$$U' = \sqrt{\frac{\sum_{j=1}^{m}\left(Y_{j+1}^{u} - \widehat{Y_{j+1}^{U}}\right)^2 + \sum_{J=1}^{m}\left(Y_{j+1}^{L} - \widehat{Y_{j+1}^{L}}\right)^2}{\sum_{J=1}^{m}\left(Y_{J+1}^{U} - \widehat{Y_{j}^{U}}\right)^2 + \sum_{j=1}^{m}\left(Y_{J+1}^{L} - \widehat{Y_{J}^{L}}\right)^2}}$$

$$\text{ARV}' = \frac{\sum_{J=1}^{m}\left(Y_{J+1}^{U} - \widehat{Y_{J+1}^{U}}\right)^2 + \sum_{j=1}^{m}\left(Y_{J+1}^{L} - \widehat{Y_{J+1}^{L}}\right)^2}{\sum_{J=1}^{m}\left(Y_{J+1}^{U} - \overline{Y^{U}}\right)^2 + \sum_{J=1}^{m}\left(Y_{j+1}^{L} - \overline{Y^{L}}\right)^2}$$

where y_t is the actual value; $\hat{y}_t$ is the predicted value; n is the number of data; m is the number of fitted intervals; $\hat{I}_t = (\widehat{X_t^U}, \widehat{X_t^L})$ is the t-th fitted interval; $\bar{I} = (\overline{X^U}, \overline{X^L})^T$ is the sample average interval; $\overline{X^U}$ is upper boundary averages and $\overline{X^L}$ lower boundary averages.

Forecasting methods compose of different combinations of several analysis methods and modeling techniques. Experimental study has done by using these different combinations. The data of experimental study are obtained by analyzing the interval-valued time series via forecasting methods. These values represent the dependent variables in the variance analysis of our study. The dependent variables measures accuracy of forecast outcomes and these are performance criterions. These performance criterions are Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Theil's Interval Statistics (U') and Interval Average Relative Variance (ARV') values. Independent variables involved eight forecasting methods. Kruskal-Wallis H test, multivariate statistical hypothesis test, was applied to evaluate the performance. Pairwise comparisons of the methods were performed using Mann-Whitney U-test. The study aimed to determine the methods and models that provide the optimal accuracy by comparing the forecasting accuracy of the proposed methods.

Additionally, the answer of the question that "which methods providing better forecasting results?" and their pros and cons were examined. In analysis stage, MATLAP, SPSS, EVIEWS and EXCEL are used. Programs are written in MATLAB programming language. The use of the methods and the contents of the methods are summarized in Table 1.

These methods are applied to four real interval-valued time series. The data of our experimental study were obtained by analyzing the interval-valued time series via forecasting methods. The obtained results are examined by using Kruskal-Wallis H and Mann-Whitney U nonparametric hypothesis tests. The reason of using nonparametric hypothesis test is that samples do not have Normal Distribution. And, median is preferred as a location parameter.

Table 1. The use of the methods and the contents of the methods

FORECASTİNG METHODS	ANALYSİS METHODS	MODELİNG TEKNİQUES
Method 1	Approach 1	ARIMA
Method 2	Approach 2	Vector Autoregressive Models (VAR)
Method 3	Approach 1	Holt's Exponential Smoothing Methods
Method 4	Approach 3	Holt's Exponential Smoothing Methods
Method 5	Approach 1	Artificial Neural Networks (ANN)
Method 6	Approach 2	Artificial Neural Networks (ANN)

4. Emprical Results

4.1. Method 1

In this forecasting method, we are used **Approach 1** as analysis method and **ARIMA** as modeling technique. We analyzed four interval-valued time series via forecasting methods. The obtained results are below.

Table 2. The obtained results for Usage of Gold Exchange I-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE I-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 1				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	1, 4782	**0, 0032**	0, 9553	0, 0453
2001	**1, 2482**	0, 0033	0, 8500	0, 2269
2002	2, 7777	0, 0052	0, 8506	0, 0155
2003	3, 1633	0, 0057	0, 7882	0, 0057
2004	4, 2641	0, 0053	0, 7867	0, 0136
2005	5, 3302	0, 0075	**0, 6415**	0, 0084
2006	4, 5610	0, 0052	0, 7543	0, 0272
2007	9, 3468	0, 0093	0, 6569	0, 0050
2008	15, 8394	0, 0098	0, 7225	0, 0182
2009	13, 5067	0, 0088	0, 6939	0, 0052
2010	12, 9907	0, 0072	0, 7147	**0, 0048**
2011	21, 7562	0, 0100	0, 6902	0, 0187

Table 3. The obtained results for Usage of Gold Exchange II-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE II-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 1				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	**46, 5046**	0, 0049	0, 8307	0, 0218
2001	201, 5830	0, 0118	0, 8876	**0, 0035**
2002	230, 6398	0, 0102	0, 7241	0, 0104
2003	141, 3513	0, 0048	0, 8004	0, 0099
2004	140, 0281	**0, 0046**	0, 7739	0, 0064
2005	185, 6846	**0, 0046**	**0, 4994**	0, 0048
2006	257, 8999	0, 0062	0, 8156	0, 0319
2007	331, 8900	0, 0082	0, 7609	0, 0366
2008	711, 6193	0, 0127	0, 8197	0, 0263
2009	3963, 7144	0, 0574	0, 7678	0, 3521
2010	4407, 2150	0, 0475	0, 7929	0, 2625
2011	2071, 0900	0, 0166	0, 9673	0, 0187

Table 4. The obtained results for Cost of Living Index (Wage Earners) (1995=100)

COST OF LIVING INDEX (WAGE EARNERS) (1995=100) METHOD 1				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	3, 2409	0, 0004	0, 7071	**0, 0000**
2001	4, 7504	0, 0003	0, 7071	**0, 0000**
2002	9, 6402	0, 0005	0, 7071	**0, 0000**
2003	6, 1781	0, 0002	0, 7071	0, 0001
2004	21, 6285	0, 0235	0, 7071	0, 0039
2005	5, 1502	0, 0002	0, 7071	0, 0002
2006	6, 6886	0, 0002	0, 7071	**0, 0000**
2007	10, 9924	0, 0004	0, 7071	0, 0002
2008	104, 4083	0, 0036	0, 7071	0, 0658
2009	8, 8804	0, 0003	0, 7071	0, 0002
2010	**2, 9160**	**0, 0001**	0, 7071	**0, 0000**
2011	6, 6363	0, 0002	0, 7071	**0, 0000**

Table 5. The obtained results for Selling Rate of Exchange-Euro (Exchange Selling)

SELLING RATE OF EXCHANGE-EURO (EXCHANGE SELLING) METHOD 1				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	**0, 0052**	0, 0057	0, 7746	0, 0522
2001	0, 0165	0, 0095	0, 6962	**0, 0028**
2002	0, 0189	0, 0094	0, 7035	0, 0066
2003	0, 0138	0, 0059	**0, 6723**	0, 0435
2004	0, 0090	0, 0035	0, 7262	0, 0050
2005	0, 0084	0, 0038	0, 8388	0, 0094
2006	0, 0128	0, 0048	0, 7479	0, 0197
2007	0, 0172	0, 0075	0, 7059	0, 0466
2008	0, 0414	0, 0146	0, 6755	0, 0399
2009	0, 0100	**0, 0028**	0, 7061	0, 0252
2010	0, 0089	0, 0030	0, 6802	0, 0780

4.2. Method 2

In this forecasting method, we are used **Approach 2** as analysis method and **Vector Autoregressive Models (VAR)** as modeling technique. We analyzed four interval-valued time series via forecasting methods. The obtained results are below.

Table 6. The obtained results for Usage of Gold Exchange I-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE I-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 2				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	1, 0810	0, 2337	0, 5731	0, 2124
2001	**0, 3628**	**0, 0937**	0, 3334	0, 0266
2002	1, 0210	0, 2001	0, 4230	0, 0100
2003	1, 6492	0, 3104	0, 5122	0, 0249
2004	0, 8213	0, 1088	0, 2049	0, 0103
2005	0, 9072	0, 1368	**0, 1845**	**0, 0022**
2006	1, 1856	0, 1410	0, 2690	0, 0104
2007	2, 0508	0, 2026	0, 2162	0, 0130
2008	4, 7426	0, 4483	0, 2807	0, 0103
2009	3, 5322	0, 2530	0, 2149	0, 0062
2010	3, 1477	0, 1748	0, 2407	0, 0211
2011	5, 0830	0, 2303	0, 2630	0, 0061

Table 7. The obtained results for Usage of Gold Exchange II - Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE II-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 2				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	15, 4716	0, 1846	0, 3741	0, 0795
2001	71, 6620	0, 4309	0, 5001	0, 0163
2002	33, 7503	0, 1560	0, 1459	0, 0017
2003	23, 5690	0, 0860	0, 2061	0, 0139
2004	33, 4015	0, 1221	0, 2380	0, 0061
2005	**13, 2449**	**0, 0393**	**0, 0734**	**0, 0002**
2006	65, 1071	0, 1738	0, 2626	0, 0133
2007	86, 1811	0, 2080	0, 2868	0, 0272
2008	135, 0478	0, 2579	0, 2130	0, 0055
2009	943, 9929	1, 1476	0, 2335	0, 0639
2010	2450, 9928	2, 2059	0, 4388	0, 1557
2011	301, 13375	0, 2344	0, 3418	0, 0191

Table 8. The obtained results for Cost of Living Index (Wage Earners) (1995=100)

COST OF LIVING INDEX (WAGE EARNERS)(1995=100) METHOD 2				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	**2, 4138**	0, 1450	0, 6528	0, 0919
2001	5, 9985	0, 1694	0, 7715	0, 1901
2002	8, 1025	0, 1864	0, 6365	0, 1273
2003	5, 9383	0, 1208	0, 7105	0, 1964
2004	6, 8748	**0, 0737**	**0, 3373**	**0, 0214**
2005	3, 7423	0, 0827	0, 5364	0, 0421
2006	4, 6892	0, 1013	0, 5501	0, 0858
2007	5, 2530	0, 1259	0, 5358	0, 0615
2008	44, 6645	0, 3114	0, 4461	0, 0269
2009	5, 8511	0, 1033	0, 5359	0, 0336
2010	5, 2130	0, 0884	0, 8301	0, 2743
2011	9, 2932	0, 1467	0, 7652	0, 1655

Table 9. The obtained results for Selling Rate of Exchange-Euro (Exchange Selling)

SELLING RATE OF EXCHANGE-EURO (EXCHANGE SELLING) METHOD 2				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	0, 0019	0, 2464	0, 3290	0, 0205
2001	0, 0078	0, 4905	0, 3699	0, 0158
2002	0, 0032	0, 1640	**0, 0596**	0, 0035
2003	0, 0026	0, 1231	0, 1026	0, 0105
2004	**0, 0016**	**0, 0549**	0, 2073	0, 0132
2005	0, 0021	0, 1059	0, 2373	0, 0191
2006	0, 0026	0, 0997	0, 2032	0, 0082
2007	0, 0035	0, 1486	0, 1880	0, 0183
2008	0, 0039	0, 1504	0, 0961	**0, 0031**
2009	0, 0030	0, 0890	0, 2871	0, 0185
2010	0, 0019	0, 0818	0, 1896	0, 0043

4.3. Method 3

In this forecasting method, we are used **Approach 1** as analysis method and **Holt's Exponential Smoothing Methods** as modeling technique. We analyzed four interval-valued time series via forecasting methods. The obtained results are below.

Table 10. The obtained results for Usage of Gold Exchange I-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE I-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 3						
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**	**α**	**γ**
2000	3, 2615	0, 0050	0, 4202	0, 0115	0, 9060	0, 0000
2001	**2, 8967**	**0, 0046**	0, 4073	0, 0294	0, 9210	0, 0000
2002	4, 0620	0, 0063	0, 5724	0, 0150	0, 9030	0, 0000
2003	6, 9835	0, 0082	0, 4717	0, 0069	0, 8010	0, 0000
2004	5, 3599	0, 0068	0, 6820	0, 0132	0, 9000	0, 0000
2005	8, 2954	0, 0096	0, 5150	0, 0087	1, 0000	0, 0010
2006	5, 9374	0, 0059	0, 6271	0, 0263	0, 9980	0, 0000
2007	19, 7802	0, 0121	0, 4131	0, 0051	1, 0000	0, 0010
2008	36, 8241	0, 0223	0, 3920	0, 0185	1, 0000	0, 0010
2009	30, 0968	0, 0120	0, 4106	0, 0052	1, 0000	0, 0010
2010	37, 1100	0, 0108	**0, 3208**	**0, 0047**	1, 0000	0, 0000
2011	51, 8705	0, 0142	0, 3871	0, 0200	0, 9550	0, 0010

Table 11. The obtained results for Usage of Gold Exchange II-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE II-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 3						
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**	**α**	**γ**
2000	**126, 9286**	**0, 0080**	0, 4279	0, 0319	0, 6880	0, 0000
2001	1325, 3921	0, 0242	**0, 2139**	0, 0068	0, 8000	0, 0000
2002	738, 2861	0, 0160	0, 3029	0, 0107	0, 9410	0, 0020
2003	162, 4803	0, 0055	0, 7724	0, 0091	0, 9540	0, 0020
2004	437, 2877	0, 0088	0, 3672	0, 0091	0, 7560	0, 0010
2005	301, 7586	0, 0067	0, 5252	0, 0047	0, 8950	0, 0000
2006	617, 4893	0, 0096	0, 4015	0, 0320	0, 8950	0, 0000
2007	313, 0417	0, 0077	0, 7289	0, 0319	0, 6730	0, 0020
2008	783, 3112	0, 0137	0, 7203	0, 0215	0, 8150	0, 0020
2009	3451, 5656	0, 0311	0, 7332	0, 2617	1, 0000	0, 0010
2010	4547, 3233	0, 0492	0, 8012	0, 2624	0, 4000	0, 0000
2011	5024, 5754	0, 0141	0, 1797	**0, 0026**	0, 9880	0, 0010

Table 12. The obtained results for Cost of Living Index (Wage Earners) (1995=100)

COST OF LIVING INDEX (WAGE EARNERS) (1995=100) METHOD 3						
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**	**α**	**γ**
2000	96, 4534	0, 0101	0, 0397	0, 0001	0, 9990	0, 0000
2001	234, 9875	0, 0130	0, 0292	**0, 0000**	0, 9990	0, 0000
2002	240, 6683	0, 0098	0, 0453	0, 0001	0, 9990	0, 0000
2003	148, 9961	0, 0056	0, 0519	0, 0002	0, 9990	0, 0000
2004	**67, 8746**	0, 0031	0, 3045	0, 0039	0, 9990	0, 0000
2005	76, 7581	**0, 0029**	0, 0716	0, 0002	1, 0000	0, 0000
2006	130, 4467	0, 0045	0, 0543	0, 0001	1, 0000	0, 0000
2007	178, 7869	0, 0060	0, 0684	0, 0002	0, 9990	0, 0000
2008	145, 8654	0, 0069	0, 5830	0, 0661	0, 9990	0, 0000
2009	140, 5379	0, 0046	0, 0702	0, 0003	0, 9990	0, 0000
2010	164, 0313	0, 0044	**0, 0243**	**0, 0000**	0, 9990	0, 0000
2011	220, 4769	0, 0059	0, 0409	0, 0001	0, 9990	0, 0000

Table 13. The obtained results for Selling Rate of Exchange-Euro (Exchange Selling)

SELLING RATE OF EXCHANGE-EURO (EXCHANGE SELLING) METHOD 3						
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**	**α**	**γ**
2000	**0, 0067**	0, 0073	0, 6923	0, 0542	0, 5010	0, 0000
2001	0, 1217	0, 0218	**0, 1371**	**0, 0029**	1, 0000	0, 0000
2002	0, 0766	0, 0150	0, 2141	0, 0053	1, 0000	0, 0010
2003	0, 0287	0, 0081	0, 4689	0, 0423	0, 8300	0, 0020
2004	0, 0298	0, 0057	0, 2698	0, 0050	0, 9430	0, 0010
2005	0, 0207	0, 0049	0, 3396	0, 0077	0, 9000	0, 0000
2006	0, 0415	0, 0081	0, 2965	0, 0193	1, 0000	0, 0010
2007	0, 0273	0, 0081	0, 5260	0, 0460	0, 9000	0, 0000
2008	0, 0424	0, 0153	0, 6670	0, 0411	0, 9080	0, 0000
2009	0, 0108	**0, 0028**	0, 6856	0, 0237	0, 9990	0, 0000
2010	0, 0181	0, 0040	0, 4264	0, 0760	1, 0000	0, 0010

4.4. Method 4

In this forecasting method, we are used **Approach 3** as analysis method and **Holt's Exponential Smoothing Methods** as modeling technique. We analyzed four interval-valued time series via forecasting methods. The obtained results are below.

Table 14. The obtained results for Usage of Gold Exchange I-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE I-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 4						
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**	**α**	**γ**
2000	3, 2615	0, 0050	**0, 1766**	0, 0115	0, 8070	0, 0000
2001	**2, 8967**	**0, 0046**	0, 4073	0, 0294	0, 8790	0, 0000
2002	4, 0782	0, 0062	0, 5784	0, 0156	0, 9990	0, 0000
2003	6, 9994	0, 0081	0, 4729	0, 0070	0, 7000	0, 0000
2004	5, 3599	0, 0068	0, 6820	0, 0132	0, 9010	0, 0000
2005	8, 2954	0, 0096	0, 5150	0, 0087	0, 9020	0, 0000
2006	5, 9374	0, 0059	0, 6271	0, 0263	0, 9980	0, 0000
2007	20, 0901	0, 0124	0, 4128	0, 9975	0, 9600	0, 0010
2008	35, 3572	0, 0215	0, 4072	0, 0191	1,0000	0, 0010
2009	30, 7076	0, 0124	0, 4165	0, 0055	1,0000	0, 0010
2010	38, 4667	0, 0106	0, 3012	**0, 0042**	1,0000	0, 0010
2011	52, 0653	0, 0138	0, 3551	0, 0174	0, 9650	0, 0020

Table 15. The obtained results for Usage of Gold Exchange II-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE II-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 4						
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**	**α**	**γ**
2000	**126, 9314**	0, 0084	0, 4279	0, 0319	0, 7050	0, 0000
2001	1325, 3993	0, 0242	0, 2139	0, 0068	0, 7980	0, 0000
2002	738, 2967	0, 0168	0, 3029	0, 0107	0, 9260	0, 0010
2003	162, 6850	**0, 0053**	0, 7726	0, 0091	0, 7770	0, 0030
2004	437, 2951	0, 0084	0, 3672	0, 0091	0, 8440	0, 0010
2005	301, 7564	0, 0067	0, 5252	0, 0047	0, 9140	0, 0000
2006	617, 4889	0, 0095	0, 4015	0, 0320	0, 8960	0, 0000
2007	313, 0432	0, 0076	0, 7289	0, 0319	0, 6810	0, 0010
2008	783, 3168	0, 0132	0, 7203	0, 0215	0, 8280	0, 0020
2009	3444, 4981	0, 0309	0, 7267	0, 2607	1, 0000	0, 0010
2010	4544, 8054	0, 0489	0, 8005	0, 2622	0, 3990	0, 0000
2011	5024, 5759	0, 0139	**0, 1797**	**0, 0026**	0, 9890	0, 0010

Table 16. The obtained results for Cost of Living Index (Wage Earners) (1995=100)

COST OF LIVING INDEX (WAGE EARNERS) (1995=100) METHOD 4						
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**	**α**	**γ**
2000	96, 4534	0, 0101	0, 0397	0, 0001	0, 9990	0, 0000
2001	234, 9875	0, 0130	0, 0292	**0, 0000**	0, 9990	0, 0000
2002	240, 6683	0, 0098	0, 0453	0, 0001	0, 9990	0, 0000
2003	148, 9961	0, 0056	0, 0519	0, 0002	0, 9990	0, 0000
2004	**67, 8746**	0, 0031	0, 3045	0, 0039	0, 9990	0, 0000
2005	76, 7581	**0, 0029**	0, 0716	0, 0002	1, 0000	0, 0000
2006	130, 4467	0, 0045	0, 0543	0, 0001	1, 0000	0, 0000
2007	178, 7869	0, 0060	0, 0684	0, 0002	0, 9990	0, 0000
2008	145, 8654	0, 0069	0, 5830	0, 0661	0, 9990	0, 0000
2009	140, 5379	0, 0046	0, 0702	0, 0003	0, 9990	0, 0000
2010	164, 0313	0, 0044	**0, 0243**	**0, 0000**	0, 9990	0, 0000
2011	220, 4769	0, 0059	0, 0409	0, 0001	0, 9990	0, 0000

Table 17. The obtained results for Selling Rate of Exchange-Euro (Exchange Selling)

SELLING RATE OF EXCHANGE-EURO (EXCHANGE SELLING) METHOD 4						
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**	**α**	**γ**
2000	**0, 0084**	0, 0057	0, 6923	0, 0542	0, 5010	0, 0000
2001	0, 0681	0, 0144	**0, 1358**	**0, 0028**	1, 0000	0, 0000
2002	0, 0482	0, 0096	0, 2134	0, 0053	1, 0000	0, 0010
2003	0, 0202	0, 0049	0, 4689	0, 0423	0, 8300	0, 0020
2004	0, 0204	0, 0034	0, 2698	0, 0052	0, 9430	0, 0010
2005	0, 0145	0, 0041	0, 3396	0, 0077	0, 9000	0, 0000
2006	0, 0269	0, 0050	0, 2965	0, 0193	1, 0000	0, 0010
2007	0, 0215	0, 0047	0, 5260	0, 0460	0, 9000	0, 0000
2008	0, 0418	0, 0127	0, 6670	0, 0411	0, 9080	0, 0000
2009	0, 0102	**0, 0019**	0, 6856	0, 0237	0, 9990	0, 0000
2010	0, 0136	0, 0030	0, 4264	0, 0760	1, 0000	0, 0010

4.5. Method 5

In this forecasting method, we are used **Approach 1** as analysis method and **Artificial Neural Networks (ANN)** as modeling technique. We analyzed four interval-valued time series via forecasting methods. The obtained results are below.

Table 18. The obtained results for Usage of Gold Exchange I-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE I-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 5				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	1, 3977	0, 0009	0, 8834	0, 6069
2001	1, 2104	0, 0010	0, 8277	0, 2772
2002	3, 1983	0, 0055	0, 8740	**0, 0847**
2003	4, 4039	0, 0079	0, 8421	0, 1579
2004	5, 2515	0, 0077	0, 8936	0, 3677
2005	**0, 0817**	0, 0116	0, 8054	0, 1755
2006	4, 2679	0, 0013	0, 7691	0, 1240
2007	10, 0521	0, 0106	0, 6924	0, 3392
2008	16, 7496	0, 0021	0, 7045	0, 1176
2009	13, 8036	**0, 0008**	0, 7084	0, 0952
2010	12, 5628	0, 0043	0, 7023	0, 3259
2011	19, 7296	0, 0025	**0, 6544**	0, 0931

Table 19. The obtained results for Usage of Gold Exchange II-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE I-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 5				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	**39, 7079**	**0, 0001**	0, 7973	0, 5109
2001	164, 2257	0, 0071	0, 7751	0, 0816
2002	236, 9452	0, 0032	0, 7363	0, 0877
2003	127, 0014	0, 0015	0, 7975	0, 3334
2004	140, 4388	0, 0007	0, 7912	0, 0992
2005	195, 7668	0, 0086	**0, 7301**	**0, 0426**
2006	245, 7378	0, 0002	0, 8133	0, 1874
2007	360, 2266	0, 0120	0, 7993	0, 4756
2008	749, 4441	0, 0132	0, 8039	0, 1638
2009	3063, 9500	0, 0443	0, 7473	0, 6741
2010	3586, 4000	0, 0180	0, 7388	0, 3340
2011	876, 3881	0, 0034	0, 7388	0, 1536

Table 20. The obtained results for Cost of Living Index (Wage Earners) (1995=100)

COST OF LIVING INDEX (WAGE EARNERS)(1995=100) METHOD 5				
YEARS	**RMSE**	**MAPE**	**U'**	**ARV'**
2000	**3, 1839**	0, 0002	0, 7136	0, 1548
2001	4, 7024	**0, 0000**	0, 7105	0, 1121
2002	9, 5938	0, 0002	**0, 7082**	0, 1814
2003	6, 3328	0, 0008	0, 7235	0, 2279
2004	21, 6970	0, 0010	0, 7274	0, 1502
2005	6, 8718	0, 0006	0, 7681	0, 1432
2006	11, 9746	0, 0040	0, 9354	0, 5556
2007	17, 3665	0, 0050	0, 9210	0, 1920
2008	102, 8203	0, 0025	0, 7090	0, 0781
2009	8, 7230	0, 0003	0, 7353	0, 0752
2010	30, 2264	0, 0090	1, 0044	1, 1821
2011	7, 5302	0, 0004	0, 7707	**0, 1108**

Table 21. The obtained results for Selling Rate of Exchange-Euro (Exchange Selling)

SELLING RATE OF EXCHANGE-EURO (EXCHANGE SELLING) METHOD 5				
YEARS	RMSE	MAPE	U'	ARV'
2000	**0, 0060**	0, 0022	0, 8812	0, 1985
2001	0, 0151	0, 0070	**0, 6387**	**0, 0592**
2002	0, 0190	0, 0030	0, 6913	0, 1056
2003	0, 0132	0, 0015	0, 6661	0, 2696
2004	0, 0104	0, 0059	0, 7660	0, 4379
2005	0, 0080	0, 0035	0, 8224	0, 2791
2006	0, 0120	0, 0015	0, 7444	0, 1631
2007	0, 0157	**0, 0007**	0, 7680	0, 3698
2008	0, 0410	0, 0100	0, 6778	0, 3551
2009	0, 0095	0, 0009	0, 7307	0, 1793
2010	0, 0092	0, 0020	0, 6790	0, 0932

4.6. Method 6

In this forecasting method, we are used **Approach 2** as analysis method and **Artificial Neural Networks (ANN)** as modeling technique. We analyzed four interval-valued time series via forecasting methods. The obtained results are below.

Table 22. The obtained results for Usage of Gold Exchange I-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE I-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 6				
YEARS	RMSE	MAPE	U'	ARV'
2000	3, 7514	0, 0231	0, 9710	0, 0551
2001	**1, 6425**	0, 0068	0, 8596	0, 0460
2002	27, 5468	0, 1550	0, 9992	2, 6946
2003	10, 6184	0, 0424	0, 9697	0, 0706
2004	14, 0890	0, 0555	0, 9888	0, 1462
2005	17, 7728	0, 0550	0, 9657	0, 0938
2006	4, 6745	**0, 0010**	0, 7648	0, 0289
2007	27, 1624	0, 0331	0, 9632	0, 0447
2008	27, 9611	0, 0131	0, 8768	0, 0570
2009	40, 0183	0, 0558	0, 9724	0, 0488
2010	33, 9382	0, 0325	0, 9410	0, 0341
2011	22, 8020	0, 0092	**0, 6981**	**0, 0206**

Table 23. The obtained results for Usage of Gold Exchange II-Istanbul (Weekday, Turkish Liras/US Dollar)

USAGE OF GOLD EXCHANGE II-ISTANBUL (WEEKDAY, TURKISH LIRAS/US DOLLAR) METHOD 6				
YEARS	RMSE	MAPE	U'	ARV'
2000	43, 1528	0, 0019	0, 8706	0, 0184
2001	267, 8614	0, 0324	0, 8618	**0, 0061**
2002	468, 9186	0, 0225	0, 9051	0, 0438
2003	201, 0252	0, 0087	0, 8966	0, 0217
2004	249, 8072	0, 0138	0, 9510	0, 0205
2005	1835, 2000	0, 1350	1, 0090	0, 4570
2006	305, 3547	0, 0022	0, 8579	0, 0444
2007	2211, 7000	0, 1346	1, 0016	1, 8799
2008	1454, 9000	0, 0045	0, 9597	0, 1152
2009	3878, 3500	0, 0038	**0, 8122**	0, 3449
2010	5040, 2000	0, 0352	0, 8669	0, 3450
2011	2307, 0000	0, 3256	0, 9997	1, 7111

Table 24. The obtained results for Cost of Living Index (Wage Earners) (1995=100)

COST OF LIVING INDEX (WAGE EARNERS) (1995=100) METHOD 6				
YEARS	RMSE	MAPE	U'	ARV'
2000	11, 3080	0, 0098	0, 9995	0, 0006
2001	46, 3924	0, 0267	1, 0003	0, 0024
2002	145, 4691	0, 0692	0, 9988	0, 0156
2003	**8, 9487**	0, 0021	**0, 8143**	0, 0003
2004	59, 4780	0, 0236	0, 9401	0, 0295
2005	9, 1602	**0, 0010**	0, 8596	0, 0006
2006	66, 0852	0, 0283	1, 0055	0, 0084
2007	60,7977	0, 0219	1, 0046	0, 0061
2008	147,3349	0, 0256	0, 8322	0, 1352
2009	41,0712	0, 0136	0, 9817	0, 0062
2010	67,8322	0, 0229	1, 0011	0, 0057
2011	14,1902	0, 0021	0, 9418	**0, 0002**

Table 25. The obtained results for Selling Rate of Exchange-Euro (Exchange Selling)

SELLING RATE OF EXCHANGE-EURO (EXCHANGE SELLING) METHOD 6				
YEARS	RMSE	MAPE	U'	ARV'
2000	**0, 0089**	0, 0151	0, 9419	0, 1964
2001	0, 0204	0, 0156	0, 7265	**0, 0042**
2002	0, 0254	0, 0021	0, 8230	0, 0121
2003	0, 0165	0, 0043	0, 7105	0, 0640
2004	0, 0192	0, 0167	0, 9477	0, 0225
2005	0, 0110	0, 0089	0, 9294	0, 0161
2006	0, 0147	0, 0026	0, 7685	0, 0257
2007	0, 0167	**0, 0019**	0, 7523	0, 0495
2008	0, 0531	0, 0263	0, 8276	0, 0668
2009	0, 0281	0, 0068	0, 9438	0, 2844
2010	0, 0098	0, 0046	**0, 6939**	0, 0928

These values represent variables in the variance analysis of our study. The obtained results are analyzed by using statistical analysis. Kruskal-Wallis H test and Mann-Whitney U-test are applied to evaluate the performance.

Table 26. The best methods of performance criterias

TİME SERİES	U'	ARV'	RMSE	MAPE
USAGE OF GOLD EXCHANGE-I	Method2	Method4	Method2	Method6
USAGE OF GOLD EXCHANGE-II	Method2	Method2	Method2	Method6
COST OF LIVING INDEX	Mtd4 and Mtd5	Mtd1	Method2	Method1
SELLING RATE OF EXCHANGE	Method2	Method2	Method2	Method6

5. Conclusions and Discussion

This study examined the interval-valued time series which is an ongoingness issue in time series. Daily time series data are used as a data set. Different forecasting methods are proposed. The data of experimental study are obtained by analyzing the interval-valued time series via forecasting methods. These values represent the dependent variables in the variance analysis of our study. The obtained results are examined by using Kruskal-Wallis H and Mann-Whitney U nonparametric hypothesis tests. As a result of study, it can be said that, VAR model is more suitable for forecasting. The evaluation of the forecasting performance of the methods revealed that Method 2 which uses Approach 2 as analysis method and Vector Autoregressive Models (VAR) as modeling technique provide the highest forecast accuracy and closest results to the actual values. The superiority of VAR models is that variables do not pose a problem.

REFERENCES

[1] A.L.S. Maia and F.A.T. De Carvalho, "Holt's exponential smoothing and neural network models for forecasting interval-valued time series" *International Journal of Forecasting,* vol. 27,no. 3, pp. 740-759, 2011.

[2] L. Billard and E. Diday, *Symbolic data analysis. Conceptual statistics and data mining,* Chichester: Wiley, 2006.

[3] H.-H. Bock and E. Diday, *Analysis of symbolic data: exploratory methods for extracting statistical information from complex data,* Springer-Verlag, Berlin, 2000.

[4] E. Diday and M. Noirhomme-Fraiture, *Symbolic data analysis and the SODAS Software,* Chichester: Wiley, 2008.

[5] L. Billard and E. Diday, "From the statistics of data to the statistics of knowledge: symbolic data analysis." *Journal of the American Statistical Association,* vol. 98, no. 462, pp. 470-487, 2003.

[6] M. Beheshti, A. Berrached, A. De Korvin, C. Hu and O. Sirisaengtaksin, "On interval weighted three-layer neural networks." in *Proceedings of the 31st annual simulation symposium,* pp.188-194, Los Alamitos: IEEE Computer Society Press, 1998.

[7] R. E. Patinǒ-Escarcina, B. R. C. Bedregal and A. Lyra, "Interval computing in neural networks: one layer interval neuralnetworks." İn *Proceedings of the 7th international conference on information technology,* pp. 68-75, 2004.

[8] A. M. Roque, C. Maté, J. Arroya and A. Sarabia, "İMLP: applying multi-layer perceptrons to interval-valued data." *Neural Processing Letters,* vol.25, pp.157-169, 2007.

[9] M. Ichino, H. Yaguchi and E. Diday, "A fuzzy symbolic pattern classifier." *in: Diday, E. et al. (Eds.), Ordinal and Symbolic Data Analysis*, pp.92-102, Springer, Berlin, 1996.

[10] J.P. Rasson and S. Lissoir, "Symbolic kernel discriminant analysis." *in: Bock, H.-H., Diday, E. (Eds.), Symbolic Data Analysis,* pp. 244-265, Springer, Heidelberg, 2000.

[11] E. Périnel and Y. Lechevallier, "Symbolic Discriminant Rules.", *In: Bock, H.-H., Diday, E. (Eds.), Analysis of Symbolic Data,* pp. 244-265, Springer, Heidelberg, 2000.

[12] H.-H. Bock, "Clustering algorithms and Kohonen maps for symbolic data." *J. Jpn. Soc. Comput. Statist*., vol.15, pp.1-13, 2002.

[13] M. Chavent and Y. Lechevallier, "Dynamical clustering algorithm of interval data: optimization of an adequacy criterion based on Hausdorff distance." *In: Sokolowski, A., Bock, H.-H. (Eds.), Classification, Clustering and Data Analysis,* pp.53-59, Springer, Heidelberg, 2002.

[14] M. Chavent and J. Sarraco, "On central tendency and dispersion measures for intervals and hypercubes." *Communications in Statistics-Theory and Methods,* vol.37, pp.1471-1482, 2008.

[15] A. Irpino, "“Spaghetti” PCA analysis: an extension of principal components analysis to time dependent interval data." *Pattern Recognition Letters,* vol.27, pp.504-513, 2006.

[16] P. J. F. Groenen, S. Winsberg, O. Rodrigues and E. Diday, "I -scal: multidimensional scaling of interval dissimilarities." *Computational Statistics and Data Analysis,* vol.51, pp. 360-378, 2006.

[17] K. C. Gowda, E. Diday, "Symbolic clustering using a new dissimilarity measure." *Pattern Recognition,* vol. 24, pp. 567-578, 1991.

[18] M. Ichino and H. Yaguchi, "Generalized Minkowski metrics for mixed feature type data analysis." *IEEE Transactions on Systems, Man and Cybernetics,* vol. 24, pp.698-708, 1994.

[19] K. C. Gowda and T. R. Ravi, "Agglomerative clustering of symbolic objects using the concepts of both similarity and dissimilarity." *Pattern Recognition Letters,* vol.16, pp.647-652,1995.

[20] K. C. Gowda and T. R. Ravi, "Clustering of symbolic objects using gravitational approach." *IEEETransactions on Systems, Man and Cybernetics,* vol.29, pp. 888–894, 1999.

[21] M. Chavent, "Criterion-based divisive clustering for symbolic objects." In *H.-H. Bock, & E. Diday (Eds.), Analysis of symbolic data, exploratory methods for extracting statistical information from complex data,* pp. 299-311, Springer-Verlag, Berlin, 2000.

[22] D.S. Guru, B. B. Kiranagi and P. Nagabhushan, "Multivalued type proximity measure and concept of mutual similarity value useful for clustering symbolic patterns." *Pattern Recognition Letters,* vol.25, pp.1203–1213, 2004.

[23] Y. El-Sonbaty and M.A. Ismail, "Fuzzy clustering for symbolic data." *IEEE Transactions on Fuzzy Systems,* vol.6, pp.195-204, 1998.

[24] F. A. T. De Carvalho, "Fuzzy c-means clustering methods for interval-valued data." *Pattern Recognition Letters,* vol.28, pp.423-437, 2007.

[25] M.-S. Yang, P.-Y. Hwang and D.-H. Chen, "Fuzzy clustering algorithms for mixed feature variables." *Fuzzy Sets and Systems,* vol.141, pp. 301–317, 2004.

[26] F. A. T. De Carvalho, M. Csernel and Y. Lechevallier, "Clustering constrained symbolic data." *Pattern Recognition Letters,* vol.30, pp. 1037–1045, 2009.

[27] F. A. T. De Carvalho and Y. Lechevallier, "Partitionalclustering algorithms symbolic interval data based on single adaptive distances." *Pattern Recognition,* vol.42, pp.1223–1236, 2009a.

[28] F. A. T. De Carvalho and Y. Lechevallier, "Dynamic clustering of interval-valued data based on adaptive quadratic distances." *IEEE Transactions on System, Man and Cybernetics—Part A: Systems and Humans,* vol.39, pp. 1295–1306, 2009b.

[29] F. A. T. De Carvalho, R. M. C. R. De Souza, M. Chavent and Y. Lechevallier, "Adaptive Hausdorff distances and dynamic clustering of symbolic data." *Pattern Recognition Letters,* vol. 27, pp. 167–179, 2006.

[30] R. M. C. R. De Souza and F. A. T. De Carvalho, "Clustering of interval data based on city-block distances." *Pattern Recognition Letters,* vol. 25, pp. 353-365, 2004.

[31] A. Irpino and R. Verde, "Dynamic clustering of interval data using a Wassertein-based distance." *Pattern Recognition Letters,* vol. 29, pp. 1648–1658, 2008.

[32] E. A. Lima Neto and F. A. T. De Carvalho, "Constrained linear regression models for symbolic interval-valued variables." *Computational Statistics and Data Analysis,* vol.54, pp. 333–347, 2010.

[33] J. Arroyo and C. Mat´e, "Forecasting histogram time series with k-nearest neighbours methods." *International Journal of Forecasting,* vol. 25, pp. 192–207, 2009.

[34] R. M. C. R. Souza and F. A. T. De Carvalho, "Clustering of interval data based on city-block distances." *Pattern Recognition Lett.,* vol. 25, no.3, pp.353-365, 2004.

[35] F. A. De Carvalho, R.M.C.R. Souza, M. Chavent and Y. Lechevallier, "Adaptive Hausdorff distances and dynamic clustering of symbolic data." *Pattern Recognition Lett.,* vol.27, no.3, pp.167-179, 2006.

[36] F.A.T. De Carvalho, "Histograms in symbolic data analysis." *Annals of Operations Research,* vol.55, pp.229-322, 1995.

[37] P. Cazes, A. Chouakria, E. Diday and Y. Schektman, "Extension de l'analyse en composantes principales `a des donn´ees de type intervalle" Revue de Statistique Appliquee, vol.45, no.3, pp.5-24, 1997.

[38] C.N. Lauro and F. Palumbo, "Principal component analysis of interval data: a symbolic data analysis approach." *Computational Statistics,* vol.15, no.1, pp. 73-87, 2000.

[39] E. Diday, "Introduction ´a l'approche symbolique en analyse des donn´ees." *Journ´ees Symbolique-Numerique,* vol. 23, no. 2, pp. 193-236, 1987.

[40] P. Bertrand and F. Goupil, "Descriptive statistic for symbolic data", in: *Bock, H.-H., Diday, E. (Eds.), Analysis of Symbolic Data,* pp.106-124, Springer, Heidelberg, 2000.

[41] L. Billard and E. Diday, "Regression analysis for interval-valued data, Data Analysis, Classification and Related Methods", in: *Proceedings of the Seventh Conference of the International Federation of Classification Societies (IFCS'00),* pp. 369–374, Springer, Belgium, 2000.

[42] E.A. Lima Neto and F.A.T. De Carvalho, "Centre and range method for fitting a linear regression model to symbolic interval data." *Comput. Stat. Data Anal*. vol.52, pp. 1500–1515, 2008.

[43] F. Palumbo and R. Verde, "Non-symmetrical factorial discriminant analysis for symbolic objects." *Applıed Stochastıc Models ın Busıness and Industry,* vol.15, no.4, pp.419-427, 1999.

[44] N.C. Lauro, R. Verde and F. Palumbo, "Factorial Discriminant Analysis on Symbolic Objects." İn *Bock, H.-H., Diday, E. (Eds.), Analysis of Symbolic Data,* pp.212-233, Springer, Heidelberg, 2000.

[45] A.L.S. Maia, F.A.T. De Carvalho and T.B. Ludermir, "Forecasting models for interval- valued time series." *Neurocomputing,* vol.71, no.16-18, pp.3344-3352, 2008.

[46] A. D. Silva, E.D. Lima Neto and U.U. Anjos, "A regression model to interval-valued variables based on Copula approach." İn Proceedings of the 58th World Statistical Congress, Dublin, p.6481,21-26 Augst 2011.

[47] E.A. Lima Neto, F.A.T. De Carvalho, "*Centre and Range Method For Fitting A Linear Regression Model To Symbolic Interval Data." Computational Statistics& Data Analysis,* vol.53, no.3, pp.1500-1515, 2008.

[48] G. Zhang, B.E. Patuwo and M. Y. Hu, "Forecasting with artificial neural networks: the state of the art", *International Journal of Forecasting,* vol.14, pp.35-62, 1998.

[49] L. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic time series." *Neurocomputing,* vol.10, pp.215-236,1996.

[50] C. Bishop, *Neural networks for pattern recognition,* Oxford: Oxford University Press, 1996.

[51] C.C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages.", *International Journal of Forecasting,* vol.20, pp.5-13, 2004.

[52] E.S. Gardner, "Exponential smoothing: the state of the art.", *Journal of Forecasting,* vol.4, no.1, pp.1- 28, 1985.

[53] D.W. Williams and D. Miller, "Level-adjusted exponential smoothing for modeling planned discontinuities", *International Journal of Forecasting*, vol.15, pp.273-289, 1999.

[54] R.H. Byrd, P. Lu, J. Nocedal and C. Zhu, "A limited memory algorithm for bound constrained optimization", *SIAM Journal on Scientific Computing,* vol.16, pp.1190-1208, 1995.

[55] J. Nocedal and S.J. Wright, *Numerical Optimization,* Springer, 1999.

[56] R Development Core Team, *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.Rproject.org, 2008

[57] R. Füss, *Department of empirical research and econometrics,* Financial Data Analysis, Winter Term 2007/08.