Engineering Applications of Artificial Intelligence 64 (2017) 272–285

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A novel intelligent approach for state space evolving forecasting of seasonal time series

Selmo Eduardo Rodrigues Júnior [a], Ginalber Luiz de Oliveira Serra [b,*]

[a] Federal University of Maranhão, Av. dos Portugueses, 1966, Bacanga, CEP: 65080-805, São Luís-MA, Brazil
[b] Federal Institute of Education, Sciences and Technology, Av. Getúlio Vargas, 04, Monte Castelo, CEP: 65030-005, São Luís-MA, Brazil

ARTICLE INFO

Keywords:
Seasonal time series
Forecasting
Unobservable components
Evolving systems
Neuro-fuzzy modelling
Singular spectral analysis

ABSTRACT

This paper proposes a new methodology for modelling based on an evolving Neuro-Fuzzy Takagi–Sugeno (NF-TS) network used for seasonal time series forecasting. The NF-TS considers the unobservable components extracted from the time series data to evolve, that is, to adapt and to adjust its structure, where the fuzzy rules number of this network can increase or decrease according to components behaviour. The method used to extract these components is a recursive version, proposed in this paper, based on the Spectral Singular Analysis (SSA) technique. The NF-TS network adopts the principle divide to conquer, where it divides a complex problem into subproblems easier to deal, forecasting separately each unobservable component, because they present dynamic behaviours that are simpler to forecast. The consequent propositions of fuzzy rules are linear state space models, where the states are the unobservable components data. When there are available observations from the time series, the training stage of NF-TS is performed, i.e., the NF-TS evolves its structure and adapts its parameters to carry out the mapping between the components data and the available sample of original time series. On the other hand, if this observation is not available, the network considers the forecasting stage, keeping its structure fixed and using the states of consequent fuzzy rules to feedback the unobservable components data to NF-TS. The NF-TS was evaluated and compared with other recent and traditional techniques for seasonal time series forecasting, obtaining competitive and advantageous results in relation to other papers. This paper also presents a case study about real-time detection of anomalies based on a patient electrocardiogram data.

1. Introduction

Time series forecasting is an important study theme in several fields, such as physics, economics, medicine, and engineering. The methods of time series forecasting are important tools to help the decision making and planning of preventive actions by experts. These forecasting methods seek to identify patterns in historical time series data such that a model is developed to build the future temporal patterns of considered series (Brockwell and Davis, 2002; Abdollahzade et al., 2015).

Besides the forecasting, the experts may also be interested in the characterization of a time series, i.e., identify its unobservable components or patterns in data. These components can have features that are important to understand the hidden behaviour of time series and to improve the forecasting results (Abdollahzade et al., 2015). The amount and type of unobservable components depend on the method used to extract them from the time series data (Brockwell and Davis, 2002; Young, 2011; Harvey and Koopman, 2009). Examples of unobservable components are trend, seasonality, cyclic pattern, level, impulse, and other types. In particular, this paper is interested in the multi-step ahead forecasting and characterization of seasonal time series. Empirically, seasonal series have patterns that occur regularly over a period of time that can be monthly, quarterly, annual, and so on Brockwell and Davis (2002) and Štěpnička et al. (2013).

Currently, there is a large volume of data in various formats, where these data are generated in a high speed. Thus, the experts need to deal with an increasing amount of information efficiently, extracting useful knowledge for modelling complex systems. In addition, the real data are usually complex, nonlinear and nonstationary. Therefore, researchers are challenged to develop algorithms that consider large amounts of data, be online and running in real time, adapt to changes in the environment, be computationally efficient, and preserve the transparency of the model (Angelov, 2013). These challenges also involve the time series analysis and the proposed methodology is included in this context.

* Corresponding author.
E-mail addresses: selmoeduardo@gmail.com (S.E. Rodrigues Júnior), ginalber@ifma.edu.br (G.L.O. Serra).

http://dx.doi.org/10.1016/j.engappai.2017.06.016
Received 26 December 2016; Received in revised form 2 May 2017; Accepted 21 June 2017
0952-1976/

Because of these current requirements, the paradigm of evolving systems has recently emerged. This paradigm is based on the concept of system structure modification, i.e., the system is able to adapt according to the changes in the environment, modifying its internal architecture and adjusting its parameters. The evolving models are applied in situations where purely adaptive models are not sufficient to represent systems (Lughofer, 2011). It is important for evolving systems algorithms that are online and work recursively. In addition, these systems are able to extract knowledge from data samples for each instant (Abdollahzade et al., 2015; Angelov, 2013; Lughofer, 2011). Therefore, there is the necessity to develop a methodology to satisfy these current requirements and challenges in time series area.

In this context, the main motivation of this paper is the needing of a methodology that deals with the strict requirements of time and with the amount of data arriving at all instants, using the promising theory of evolving systems, and considering seasonal time series that are very important in practice. Another important motivation is the gap that exists in time series domain in relation to methods that work with both objectives recursively: forecasting and characterization. Hence, this paper focuses on the development of a methodology that provides important knowledge to support the decision-making by experts.

1.1. State-of-the-art

To deal with the nonstationary behaviour, uncertainties, unobservable components, and other complexities of real time series, methods for time series forecasting based on computational intelligence techniques has been widely investigated in scientific community, mainly fuzzy, neuro-fuzzy, and neural networks approaches (Siminski, 2017; Papageorgiou and Poczetac, 2017; Bodyanskiy et al., 2017; Bodyanskiy and Vynokurova, 2013; Li et al., 2013; Chai and Lim, 2016; Li and Hu, 2012; Singh and Borah, 2013; Peng et al., 2015; Xiong et al., 2015). A modified Takagi–Sugeno–Kang fuzzy neural system, entitled Locally Recurrent Neuro-Fuzzy Forecasting System (LR-NFFS), is presented in Mastorocostas and Hilas (2012), where the consequent parts of the fuzzy rules are neural networks with an internal recurrence, introducing the dynamics to the overall system. The LR-NFFS is used for telecommunications time series forecasting. In Abdollahzade et al. (2015), a hybrid method for nonlinear and chaotic time series forecasting based on a local linear neuro-fuzzy model and optimized singular spectrum analysis is developed. This hybrid method is applied to forecast several well-known time series with different structures and characteristics. A hybrid evolutionary system composed by a simple exponential smoothing filter, ARIMA and autoregressive (AR) linear models and a Support Vector Regression (SVR) model is applied for time series forecasting in de Oliveira and Ludermir (2016). Particle swarm optimization is employed to optimize the order of AR model, SVR parameters, and number of time series lags. In Cui et al. (2015), the authors propose a novel hybrid networks model Complex Rotation Quantum Dynamic Neural Networks (CRQDNN) applied to application studies: the chaotic time series prediction and electronic remaining useful life prediction.

Some new methodologies is also being developed for seasonal time series forecasting (Zhang and Qi, 2005). In Štěpnička et al. (2013), the researchers introduce novel methods based on computational intelligence techniques for multi-step seasonal time series forecasting, for example, evolutionary artificial neural networks, support vector machines, and linguistic fuzzy rules. The most important contribution of Štěpnička et al. (2013) is the introduction of a new hybrid combination using linguistic fuzzy rules and others computational intelligence methods presented. In Gan et al. (2014), a quasi-linear autoregressive model, that belongs to a class of varying coefficient models in which its autoregressive coefficients are constructed by radial basis function networks, is applied for seasonal and trend time series forecasting.

Many researches involving the time series forecasting are considering the machine learning concepts in their proposals, highlighting the unsupervised learning method called clustering (Lapidot et al., 2017; Homenda and Jastrzebska, 2017; Rodriguez-Fernandez et al., 2017; Huang et al., 2016). A forecasting model based on a modified fuzzy c-means and information granulation which does not require that data have the same dimensionality is proposed in Wang et al. (2015) for time series long-term prediction. In Wu and Lee (2015), a local modelling strategy and the investigation of effectiveness of local modelling with three popular machine learning based on forecasting methods, Neural Network, Adaptive Neuro-Fuzzy Inference System, and Least Squares Support Vector Machine, applied for time series prediction is studied. An improved extreme learning machine for online sequential prediction of multivariate time series is presented in Wang and Han (2015). On the basis of the specific network function of extreme learning machine, an improved Levenberg–Marquardt algorithm, in which Hessian matrix and gradient vector are calculated iteratively, is developed to implement online sequential prediction of multivariate time series in Wang and Han (2015).

We can also highlight, in state-of-art, the use of evolving neuro-fuzzy networks, adapting its structure to the data (Angelov, 2013; Lughofer, 2011; Lughofer and Angelov, 2011; Lughofer, 2013; Rocha and Serra, 2017; Lemos et al., 2011). In Komijani et al. (2012), an eTS-LS-SVM algorithm (evolving Takagi–Sugeno Least Square Support Vector Machine) applied for time series forecasting, is presented. The main contribution of Komijani et al. (2012) is that it addresses nonlinear local models based on least squares support vector machines as consequence of fuzzy rules. In Birek et al. (2014), a modified evolving fuzzy Takagi–Sugeno (eTS) algorithm applied to a leakage forecasting problem is described. The eTS groups the data into several clusters based on Euclidean distance between the relevant independent variables. So, the Mod eTS algorithm, which incorporates a modified dynamic update of cluster radii while accommodating new available data is proposed. The created clusters serve as a base for fuzzy If-Then rules with Gaussian membership functions which are defined using the cluster centres and have linear functions in the consequent. An evolving fuzzy neural network (eFNN) predictor is proposed in Li et al. (2014) to extract representative information from multi-dimensional data sets for more accurate system state forecasting. The proposed predictor possesses online learning capability and can address nonstationary properties of data sets. In Júnior and Serra (2016), an algorithm for nonstationary and seasonal time series forecasting, with an evolving Neuro-Fuzzy Takagi–Sugeno (NF-TS) structure, is proposed. For this algorithm, the NF-TS inputs are unobservable patterns extracted from the time series by the Holt–Winters method optimized by Particle Swarm Optimization.

1.2. Originality and contributions

In this paper, an evolving Neuro-Fuzzy Takagi–Sugeno (NF-TS) modelling approach applied for multi-step forecasting and characterization of seasonal time series is proposed. The unobservable components are extracted from the time series through of a recursive version, proposed in this paper, based on Singular Spectrum Analysis (SSA) method (Golyandina et al., 2001). This new approach was named of Recursive SSA (RSSA). The NF-TS uses the unobservable components data to adapt and evolve its structure based on fuzzy rules (adding or removing) to forecast future observations for the time series. Firstly, the forecasting is performed for each unobservable component separately. Next, the forecasting results for each component are considered to compose the future value for the original time series. The consequent proposition of NF-TS fuzzy rules are equations in the state space, whose the states are formed by the extracted components, representing the dynamic behaviour of time series components.

So, the originality and the main contributions of the proposed methodology are presented in the following aspects:

- Evolving forecasting methodology of seasonal time series multi-steps ahead;
- Use of unobservable components data to forecast the time series;