ESANN 2018 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2018, i6doc.com publ., ISBN 978-287587047-6. Available from http://www.i6doc.com/en/.

An extension of nonstationary fuzzy sets to heteroskedastic fuzzy time series

Marcos Antônio Alves[1], Petrônio Cândido de Lima e Silva[1,3], Carlos Alberto Severiano Junior[1,4], Gustavo Linhares Vieira[1], Frederico Gadelha Guimaraes[2], Hossein Javedani Sadaei[2] *

1- Graduate Program in Electrical Engineering - Universidade Federal de Minas Gerais
Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brazil
2- Universidade Federal de Minas Gerais (UFMG) - Dept. of Electrical Engineering
Av. Antonio Carlos 6627, 31270-901, Belo Horizonte, MG - Brazil
3- Instituto Federal do Norte de Minas Gerais (IFNMG)
Rua Coronel Luiz Pires, 202, Centro, Montes Claros, MG - Brazil
4- Instituto Federal de Minas Gerais (IFMG) - Campus Sabará
Av. Serra da Piedade, 299, Conj. Morada da Serra, Sabará, MG - Brazil

Abstract. Most applications deal with unconditional variance of the time series. Fuzzy time series allow an inexpensive computation to forecasting dynamic processes and uncertainties. In this paper we have extended the concept of nonstationary fuzzy sets to Fuzzy Time Series, termed Non-stationary Fuzzy Time Series (NSFTS). While some models require new data before adapting, the NSFTS is capable of adapting to heteroskedastic time series. In the experiments, NSFTS outperformed other known FTS methods with box-cox transformations available. Statistical tests in three different datasets indicate that the results achieved by the proposed model are either superior or non-inferior to other FTS models.

1 Introduction

In order to deal with vagueness and imprecise knowledge in time series data, Fuzzy Time Series (FTS) were introduced by Song and Chissom [1]. FTS are based on Fuzzy Set Theory first proposed by Zadeh [2]. Some of the data allow parameter variability, suggesting a lack of stationarity, i.e. unconditional instants where time series may vary over time [3, 4]. Unlike homeoskedastic time series, the data variance changes over time and, because the information that is inherent to the data generating system is not in the model, the magnitude of the prediction errors increases with time.

Since most of real-world applications are concerned with dynamic processes and uncertainties, the problem has motivated reviews [5, 6], forecasting applications [3, 4, 7] and new design methodologies [8, 9]. These articles and research works consider the nonstationarity of time series but because of the nature of accumulated uncertainty in FTS modeling, the literature focused on investigating nonstationary data: no paper ever addressed extensions of nonstationary

*This work has been supported by the Brazilian agencies FAPEMIG, CAPES and CNPq.

ESANN 2018 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2018, i6doc.com publ., ISBN 978-287587047-6. Available from http://www.i6doc.com/en/.

fuzzy sets (NSFS) presented by Garibaldi et al. [10] to FTS in order to investigate heteroskedasticity i.e. temporal variability of variance [4] in time series. Therefore, those analyses may be over-simplifying.

We propose to term this new approach Nonstationary Fuzzy Time Series (NSFTS). Its purpose is predicting trends and scale changes in time series. Because it relies on NSFS the model should be able to keep itself updated without requiring any new training, as long as there is no break in the trend. Some models ask for new data before adapting. For instance, in offline learning, when the data is known and has a definite trend, one can try to model the trend and remove it from data so that one can apply usual methods. Regarding online learning or streaming, NSFTS appears to be best suited because it automatically deals with changes in trend and variance.

2 Preliminaries

2.1 Fuzzy Time Series

There are several categories of FTS methods, varying mainly by their *order* and time-variance. The order is the number of time-delays (lags) that are used in modeling the time series. Given the time series data F, First Order models need $F(t-1)$ data to predict $F(t)$ ones while Higher Order models require $F(t-1), \ldots, F(t-k)$ data to predict $F(t)$. According to Song and Chissom [1], FTS starts with an univariate time series $Y(t) \in \mathbb{R}^1$, for $t = 0, 1, \ldots, T$, where the Universe of Discourse (UoD) U is delimited by the known bounds of $Y(t)$, such that $U = [\min(Y), \max(Y)]$. Fuzzy sets A_i are defined upon U, for $i = 1 \ldots q$, each one with its own membership function (MF) μ_{A_i}. $F(t)$ is a Fuzzy Time Series over $Y(t)$ if $F(t) = \mu_{A_i}(Y(t))$ is the collection of fuzzyfied values of $Y(t)$ for $i = 1 \ldots q$ and $t = 0, 1, \ldots, T$. Fuzzy time series $F(t)$ can also be seen as a Linguistic Variable and the fuzzy sets A_i as their linguistic values.

Suppose $F(t-1) = A_{i1}$, $F(t-2) = A_{i2}$, $F(t-3) = A_{i3}$, …, $F(t-p) = A_{ip}$ and $F(t) = A_j$. The high order fuzzy logical relashionship (FLR) can be defined as $A_{i1}, A_{i2}, A_{i3}, \cdots, A_{ip} \to A_j$ where $A_{i1}, A_{i2}, A_{i3}, \cdots, A_{ip}$ is the left-hand side (LHS) and A_j is the right-hand side (RHS) of the FLR. Based on Chen's model [11], these FLRs can be grouped into the FLRG since they have the same fuzzy sets on the LHS.

2.2 Nonstationary Fuzzy Sets

Based on the assumption by Garibaldi et al. [10], "*the ability of FST to model and/or minimize the effects of uncertainty is restricted*", since the data can vary with time. Thus, through a slight variation in the MF, the NSFTS could be able to deal with the incorporation of variability into decisions over time. NSFS are based on two functions: $\mu(x)$: the membership function, $\mu \in (0, 1)$; and $p(\mu, t)$ the perturbation function that alters the parameters of the function μ by a temporal parameter t.

ESANN 2018 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2018, i6doc.com publ., ISBN 978-287587047-6. Available from http://www.i6doc.com/en/.

NSFS shape changes along time, given the parameter t. This change can take any form, including at random. According to Garibaldi et al. [10] these forms of the nonstationarity can be formalized as: a) variation in location: translate the parameters of μ function along the UoD, without changing its shape; b) variation in width: Change the shape of μ, stretching or contracting its bounds; and c) noise variation: Adds random noise (for instance White Gaussian Noise (WGN), $x \sim N(0, 1)$) to the membership grade.

3 Nonstationary Fuzzy Time Series

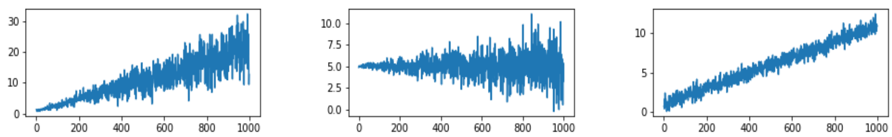
Nonstationary Fuzzy Time Series provides FTS with an incremental component capable of representing mean and time variations observed over time. This component is identified during the training process and embedded in the perturbation function mentioned previously. Then, it is applied to the model during the forecasting step. The next subsections describe the training and forecasting steps.

3.1 Training Procedure

In order to identify the perturbation function, a segmentation of the time series is performed. The observations are split into sliding windows, which are intervals where mean and variance parameters present smaller variations. Given a training set D with n instances, consider sliding windows of size w and a parameter d for the order of perturbation functions. The training process is described below.

1. Define the Universe of Discourse (UoD) based on D. Define a partition of the UoD into K even-length fuzzy sets A_i with triangular MF $\mu_i(\cdot)$ where $i \in \{1, 2, \cdots, K\}$.
2. Learn the fuzzy logical relationships and the FLRGs using D;
3. Split D into d_j sliding windows where $j \in \{1, 2, \cdots, (n/w)\}$;
4. For each subset $d_j \in D$
   (a) Find $Y_{max,j} = \max(Y(t) \in d_j)$ and $t_{max,j}$ = time at $Y_{max,j}$;
   (b) Find $Y_{min,j} = \min(Y(t) \in d_j)$ and $t_{min,j}$ = time at $Y_{min,j}$;.
5. Interpolate polynomials for location perturbation functions:
   (a) The lower polynomial p^l_1, order d, is interpolated with $(Y_{min,j};\ t_{min,j})$;
   (b) The upper polynomial p^l_K with order d is interpolated with data $(Y_{max,j};\ t_{max,j})$;
   (c) The intermediate polynomials p^l_s, $s \in \{2, 3, \ldots (K-1)\}$, are determined by the linear interpolation between p^l_1 and p^l_K parameters.
6. Interpolate polynomials for width perturbation functions:

ESANN 2018 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence
and Machine Learning. Bruges (Belgium), 25-27 April 2018, i6doc.com publ., ISBN 978-287587047-6.
Available from http://www.i6doc.com/en/.



(a) (b) (c)

Fig. 1: Synthetic datasets for model evaluation

   (a) Compute the values of the location polynomial p_i^l, $i \in 1 \ldots K$, for $t \in 0 \ldots n$;
   (b) Compute the distance w_t between each pair of location polynomials p_i^l and p_{i+1}^l for $t \in 0 \ldots n$;
   (c) The width polynomials p_i^w, $\forall i \in 1 \ldots K$ with order d are interpolated with $(w_t;\ t)$ data.

7. For each fuzzy set A_i with triangular MF $\mu_i(\cdot)$, add the location and width perturbation functions p_i^l and p_i^w, where $i \in \{1, 2, \cdots, K\}$.

3.2 Forecasting Procedure

Given a time series $Y(t)$, for $t = 1, 2, \ldots, L$, where L is the number of lags (order)

1. Compute the membership grade $\mu_i(Y(t))$ for each nonstationary fuzzy set A_i, perturbed by the width and location functions with parameter t;
2. Select FLRG's where the fuzzy sets A_i in the LHS have $\mu_i > 0$;
3. For each selected FLRG k, compute the mean point $mpfl_k(t)$:

$$mpfl_k(t) = \frac{\sum_{j \in RHS} mp_j(t)}{|RHS|} \tag{1}$$

   where $mp_j(t)$ is the mean point of a fuzzy set A_j

4. Compute the forecast as:

$$Y(t+1) = \sum_{i \in K} \mu_i(Y(t)) \cdot mpfl_i(t) \tag{2}$$

4 Empirical results

To measure the performance of the proposed method, three synthetic datasets with unconditional heteroskedastic behavior were generated: Figure 1a, dataset A: increasing mean and variance. Figure 1b, dataset B: constant mean and increasing variance. Figure 1c, dataset C: increasing mean and constant variance.

ESANN 2018 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2018, i6doc.com publ., ISBN 978-287587047-6. Available from http://www.i6doc.com/en/.

Each dataset had 1000 instances and a sliding window cross validation methodology was applied, using a working set of 300 instances, divided in 70% for training (in-sample) and 30% for testing with a sliding increment of 30 instances, totalizing 25 experiments for each tested method. The accuracy metric to evaluate the models was the Symmetrical Mean Average Percent Error (SMAPE), defined as $SMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - \hat{Y}_i|}{|\hat{Y}_i|+|Y_i|}$, where Y is the correct value, $\hat{Y}$ is the forecasted value and n is the number of instances.

NSFTS was tested against known FTS methods in the literature: the traditional Fuzzy Time Series (FTS) [1], Conventional FTS (CFTS) [11], Weighted FTS (WFTS) [12], Improved Weighted FTS (IWFTS) [13], Trend Weighted FTS (TWFTS) [14], Exponentially Weighted FTS (EWFTS) [15] and High Order FTS (HOFTS) [16]. For these methods a Box-Cox transformation [17] was applied to normalize the data and allow these methods to deal with this kind of nonstationary time series. All methods were trained with a grid partitioning scheme, the number of partitions varying in the interval $[5, 100]$ and the high order models were tested with orders 2 and 3. The best performing models where chosen for each method.

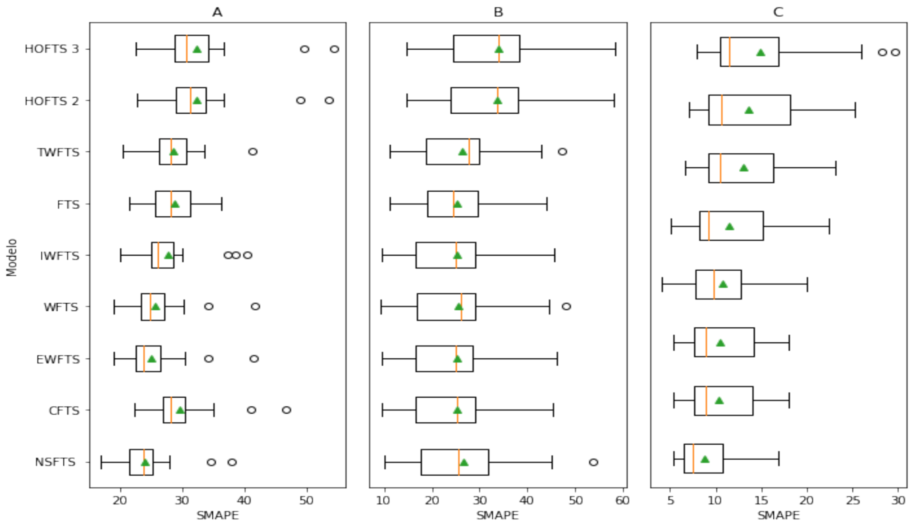


Fig. 2: Experiments results

It can be seen in Figure 2 that the NSFTS outperformed the other FTS models on average for datasets A and C and tied with CFTS, EWFTS, WFTS and IWFTS on dataset B. Statistical significance tests show that for all datasets, the proposed NSFTS model SMAPE mean was either superior or non-inferior to the other FTS models, considering a 95% significance interval. For dataset A, NSFTS mean was statistically superior to all methods but EWFTS, WFTS and IWFTS. For dataset B, it was superior to HOFTS2 and HOFTS3. Finally, it was superior to FTS, IWFTS and CFTS for dataset C.

It is interesting to note that NSFTS performed best (in comparison to other methods) on dataset A, which presents both increasing mean and variance. This might be an indication that the proposed method's performance is less affected by these characteristics than the other FTS methods.

References

[1] Qiang Song and Brad S Chissom. Fuzzy time series and its models. *Fuzzy Sets and Systems*, 54(3):269–277, 1993.

[2] Lotfi A Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.

[3] Sébastien Van Bellegem and Rainer Von Sachs. Forecasting economic time series with unconditional time-varying variance. *International Journal of Forecasting*, 20(4):611–627, 2004.

[4] Y Liu, B Wang, H Zhan, Y Fan, Y Zha, and Y Hao. Simulation of nonstationary spring discharge using time series models. *Water Resources Management*, 31(15):4875–4890, 2017.

[5] Pritpal Singh. Fuzzy time series modeling approaches: A review. In *Applications of Soft Computing in Time Series Forecasting*, pages 11–39. Springer, 2016.

[6] Changqing Cheng, Akkarapol Sa-Ngasoongsong, Omer Beyca, Trung Le, Hui Yang, Zhenyu Kong, and Satish TS Bukkapatnam. Time series forecasting for nonlinear and non-stationary processes: a review and comparative study. *IIE Transactions*, 47(10):1053–1071, 2015.

[7] Xiao-Ying Wang, Jonathan M Garibaldi, Shang-Ming Zhou, and Robert I John. Methods of interpretation of a non-stationary fuzzy system for the treatment of breast cancer. In *IEEE International Conference on Fuzzy Systems, 2009. FUZZ-IEEE 2009.*, pages 1187–1192. IEEE, 2009.

[8] Zhengbing Hu, Yevgeniy V Bodyanskiy, Oleksii K Tyshchenko, and Olena O Boiko. Adaptive forecasting of non-stationary nonlinear time series based on the evolving weighted neuro-neo-fuzzy-anarx-model. *arXiv preprint arXiv:1610.06486*, 2016.

[9] Riswan Efendi, Mustafa Mat Deris, and Zuhaimy Ismail. Implementation of fuzzy time series in forecasting of the non-stationary data. *International Journal of Computational Intelligence and Applications*, 15(02):1650009, 2016.

[10] Jonathan M Garibaldi, Marcin Jaroszewski, and Salang Musikasuwan. Nonstationary fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 16(4):1072–1086, 2008.

[11] Shyi-Ming Chen. Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems*, 81(3):311–319, 1996.

[12] Hui-Kuang Yu. Weighted fuzzy time series models for TAIEX forecasting. *Physica A: Statistical Mechanics and its Applications*, 349(3):609–624, 2005.

[13] Zuhaimy Ismail and R. Efendi. Enrollment forecasting based on modified weight fuzzy time series. *Journal of Artificial Intelligence*, 4(1):110–118, 2011.

[14] Ching-Hsue Cheng, Tai Liang Chen, Hia Jong Teoh, and Chen Han Chiang. Fuzzy time-series based on adaptive expectation model for TAIEX forecasting. *Expert Systems with Applications*, 34:1126–1132, 2008.

[15] Hossein Javedani Sadaei. *Improved models in Fuzzy Time Series for forecasting.* PhD thesis, Universiti Teknologi Malaysia, 2013.

[16] Shyi Ming Chen and Chao Dian Chen. TAIEX forecasting based on fuzzy time series and fuzzy variation groups. *IEEE Transactions on Fuzzy Systems*, 19(1):1–12, 2011.

[17] M. H. Lee, Hossein Javedani Sadaei, and Suhartono. Improving TAIEX forecasting using fuzzy time series with Box–Cox power transformation. *Journal of Applied Statistics*, 40(11):2407–2422, 11 2013.