Probabilistic Forecasting with Seasonal Ensemble Fuzzy Time-Series

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

[1] Instituto Federal do Norte de Minas Gerais, IFNMG, Januaria, Brazil
petronio.candido@ifnmg.edu.br
[2] Instituto Federal de Minas Gerais, IFMG, Sabara, Brazil
carlossjr@ufmg.br
[3] Department of Electrical Engineering, Universidade Federal de Minas Gerais, UFMG, Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brazil
Machine Intelligence and Data Science Laboratory (MINDS Lab)
marcosalves@ufmg.br, gustavolvieira@ufmg.br, fredericoguimaraes@ufmg.br, h.javedani@gmail.com

Abstract. The aim of this paper is to propose a method for probabilistic forecasting based on the aggregation of seasonal Fuzzy Time-Series techniques with ensemble learning. The proposed method generates different seasonal FTS models and the best ones are combined into an ensemble learning. The forecasting procedure consists in evaluating individual models and combining their outputs into a continuous probability distribution using Kernel density estimation. The method was applied to SONDA dataset considering three seasonal indexes on solar radiation data. The best Ensemble models were those with 15 minutes interval index and Entropy partitioning in their different parameters. The built ensemble forecasts were then compared with ARIMA and Quantile Auto-Regression models using Continuous Ranked Probability Score (CRPS) metric. The Ensemble FTS method presented a slightly larger CRPS, especially for the Epanechnikov, Tophat and Triangular kernels, which suggests a better model.

Keywords: Fuzzy forecasting; Time-series; Fuzzy seasonality; Ensemble Learning.

1 Introduction

Time-series can be seen as a set of data observed at a discrete point of time. Essentially, information can be inferred from the patterns of past observations and can be used to forecast future values of the series [15]. In order to deal with vague and incomplete data, Fuzzy Time-Series (FTS), introduced by [30] and based on Fuzzy Set Theory (FST) [31], appear as computationally inexpensive, efficient and simple to implement forecasting methods. FTS have been used in several application fields, such as tourism [16], electric load [12, 24], stock markets [6, 23] and seasonal time-series [3, 29]. There are a variety of models to forecast future events. When the data repeats at fixed intervals it becomes a Seasonal Fuzzy Time-Series (SFTS) [15]. A wide variety of seasonalities can be found in a time-series and the current prediction methods have not been able to provide satisfactory accuracy rates for forecasts [22]. For this reason, Probabilistic forecasts assign probabilities to different outcomes and have received good attention [9, 28].

However, probabilistic forecasts are not equally accurate and different metrics are necessary to assess the respective accuracy of distinct probabilistic forecasts. The tradeoff between bias and variance in FTS is controlled by the number of fuzzy sets and their distribution. For improving performance, some approaches of preprocessing data with transformations are applicable, such as differentiation [23], box-cox [17], adaptive forecasts [8] etc. These transformations can help predict a future point as it improves accuracy over the time-series under analysis.

A measure of uncertainty is lacking for point forecasts, which means that this type of forecasting is not enough to capture all the uncertainty of an estimate. More complex forecasting tasks involving dynamic and nonlinear processes, such as climatological and economic forecasts, contain many uncertainty sources [24]. Moreover, combinations of different learning methods have been used to improve predictive performance, since it is expected that ensemble learning for probabilistic forecasts is able to result in better accuracy, on average, than any individual prediction. The main advantage of ensemble learning is the flexibility and ease implementation, since the individual models can be replaced (or added) for any other point forecast, such as any FTS method [18]. In this paper we provide a new probabilistic forecasting approach using seasonal FTS, ensemble learning and kernel density estimation (KDE). The proposed method generates different seasonal FTS models and the best ones are combined into an ensemble. The forecasting procedure consists in evaluating individual models and combining their outputs into a continuous probability distribution using KDE. As numerical example, it was applied to solar forecasting data, using SONDA dataset.

2 Background

2.1 Fuzzy Time Series

Fuzzy Time-Series (FTS) are non-parametric forecasting methods introduced by [30] based on Fuzzy Set theory [31]. FTS provide a different representation of a time series data. While conventional time series are composed by sequential observations represented by real numbers indexed by a time index t, FTS are composed by fuzzy sets. These fuzzy sets form the Universe of Discourse (UoD) for the forecasting problem. The UoD is obtained from the range of values observed in the conventional time series. For example, consider a crisp time series $Y(t) \in \mathbb{R}$, for $t \in [0, n]$. The UoD can be divided into overlapping sub intervals such as $U = u_1, u_2, \cdots, u_n$. The fuzzy sets A_i are then defined over each sub interval with a corresponding membership function $\mu_{A_i} : u_i \longmapsto [0, 1]$. Therefore, if $F(t)$ consists of $\mu_{A_i}(t)$, then $F(t)$ is considered an FTS over $Y(t)$, Fuzzy Logical Relationships (FLR) represent the causal relationship between the observations at time t and previous observations. Establishing the FLR is one of the main steps for an FTS algorithm. If there exists a fuzzy relationship $R(t-p,t)$, such that $F(t) = F(t-p) \circ R(t-p,t)$, where $\circ$ is an arithmetic operator, then $F(t)$ is said to be caused by $F(t-p)$. The relationship between $F(t)$ and $F(t-p)$ can be denoted by $F(t-p) \rightarrow F(t)$.

Consider $F(t-1) = A_i$ and $F(t) = A_j$. The FLR can be defined as $A_i \rightarrow A_j$ where A_i and A_j are called the left-hand side (LHS) and the right-hand side (RHS) of the FLR, respectively.

The main steps of all FTS methods were proposed in [30], but its computation demanded many matrix operations for each forecasting, making the process computationally expensive. Then, Chen [7] simplified Song and Chissom's algorithm by creating the Fuzzy Logical Relationship Groups (FLRG), making the forecasting process cheaper by avoiding the use of matrix manipulations. The FLRG represent the knowledge base (rule base) of the model and are human readable and easy to interpret.

The FLRs with the same LHS are gathered into FLRGs. LHS of groups indicate input value (the point which prediction is performed) and RHS corresponds to the outputs that were experienced in the estimation period.

If $F(t)$ is caused by $F(t-1), F(t-2), F(t-3), \cdots, F(t-p)$, then the corresponding FLR is $F(t-1), F(t-2), F(t-3), \cdots, F(t-p) \rightarrow F(t)$.

Additionally, *Seasonal Fuzzy Time Series* (SFTS) models were proposed in [29] basically by defining a seasonal period i, where $F(t) = F(t-i)$. Before, Chang [3] proposed a method for capturing fuzzy trend and fuzzy seasonal indexes using Fuzzy Regression.

2.2 Probabilistic Forecasting

Probabilistic Forecasting assigns probabilities to different outcomes, instead of returning a single-valued forecast. The set of probabilities represents the probability forecasting. The set of outcomes delimited by these probabilities is known as prediction interval.

A simple method for creating prediction intervals for generic forecasting models was proposed by [4], namely mean-variance model. From the point forecast $\mu = \mathbb{E}[Y_{t+1}|Y_t, Y_{t-1}, ...]$ with the variance of the residuals $\sigma_\epsilon = \sqrt{VAR[\epsilon]}$ by assuming that these residuals as $\epsilon \sim \mathcal{N}(0,1)$. The prediction interval is calculated by $I = [\mu - z_{\alpha/2}\sigma_\epsilon\ ,\ \mu + z_{\alpha/2}\sigma_\epsilon]$ and $z_{\alpha/2} = \Phi((1-\alpha)/2)$ is the standard normal distribution function.

The main probabilistic approach for interval forecasting is the Quantile Auto Regression (QAR) [14], based on the Quantile Regression [13]. The QAR estimates a conditional quantile function $Q_{Y_t}(\tau|Y_{t-1}, ...) = \min_\theta \sum_{i=1}^{n} \rho_\tau(y_t - y_i\theta)$ where $\hat{Y}_t$ is the estimated quantile value, τ is the quantile level, θ are the fitted coefficients for the Y_i lagged values and $\rho_\tau(u) = u(\tau - \mathbb{I}(u < 0))$ is the Pinball Loss Function, where $\mathbb{I}(x) = \{1 \text{ if } x \geq 0 \text{ or } 0 \text{ if } x < 0\}$. QAR approaches have been used in many application fields, for instance energy load forecasting [11] and wind forecasting [20]. Each QAR model is fitted for a specific τ, so for a certain α two QAR models are needed. The independence of quantiles also allows to create asymmetric inter-quantile intervals, if needed.

2.3 Ensemble Learning

An ensemble prediction consists of multiple runs of numerical weather prediction models, which differ in the initial conditions [1, 9]. This concept was exploited in [1] and [18], which proposed an ensemble learning approach for solar power probabilistic forecasting based on k-Nearest Neighbors, Regression Trees, Random Forests and regression methods. Given an ensemble with k models and taken the ordered set of the k individual forecasts, the probabilistic forecast is constructed as an empirical cumulative distribution F, calculated with the percentiles of the individual forecasted values. F can be made with three approaches: Quantile Linear Interpolation, Normal Distribution and Normal Distribution with initial different conditions. The linear interpolation approach calculates the τ quantile position r_τ on the individual forecasts as $r_\tau = \frac{k\tau}{100} + 0.5$. With the set of individual forecasts, the mean μ and variance σ are calculated and the τ quantile is given by $\tau = \mu + z_\tau \cdot \sigma$. The third approach is specific for the application domain of solar power. Ensemble Learning for time series forecasting is also proposed in [5].

2.4 Kernel Density Estimation

Distribution generating techniques for ensemble forecasts exist as Kernel Density Estimation [11] and Kernel Dressing [2, 21]. Both approaches smooth the discrete values in a continuous function that approximates the empirical distribution of data, as $P(x) = (nh)^{-1} \sum_{i \in Y} K((x-i)/h)$ where Y is the set of individual forecasts, K is the kernel function and h is a smoothing parameter also known as bandwidth. A kernel function K has to be non-negative, real-valued, symmetric, integrable and normalized. A review of density estimation methods can be found in [26] and a specific study on estimation of h parameter can be found in [25].

3 Proposed Method

3.1 Training Procedure

The aim of the training procedure is to build an ensemble $\mathcal{M}$ with k individual models m_i, given a crisp training set $Y(t)$. The overall training procedure is shown in Figure 1 and it is composed as follows:

Step A) Data Preprocessing The first step necessary in any forecasting problem is to verify the dataset beforehand and preprocess it if necessary. Very often the data contain outliers, missing fields and wrong info that, if ignored, can worsen the prediction results. This step includes data selection, cleaning,

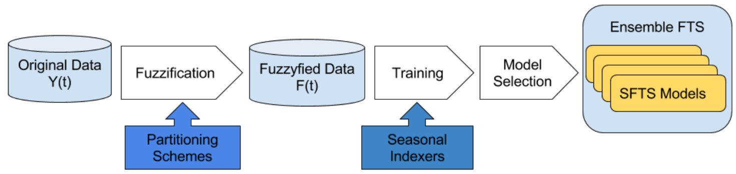


Fig. 1: Ensemble FTS training procedure

dimensionality reduction, removing seasonality and trend. Besides, there are transformations that can be applied in order to facilitate the predictions and improve the results. On this method, all models were built using differentiated time series (with a lag of one sample) or simply the original data.

Step B) Universe of Discourse Partitioning Two methods of partition were applied: a) Grid Partitioning [3, 7, 29], which divides the UoD in n overlapping intervals of equal length and b) Entropy Partition which creates n partitions of different lengths based on the distribution of observations in the UoD. The fuzzy sets were created with triangular membership functions in both cases. For each of these methods of partition, the number of partitions still needs to be supplied as a parameter, but its definition is empirical and data dependent.

Step C) Data Fuzzification Once defined the partitioning schemes and number of partitions, each data point on crisp time series $Y(t)$ is replaced by the fuzzy set with maximum membership value, i. e, $F(t) = arg\max_{A_i} \mu_{A_i}(Y(t))$. The fuzzified time series $F(t)$ is different for each partitioning scheme and number of partitions.

Step D) Individual Seasonal FTS model training In this step the Multi-Seasonal FTS method and its training procedure are defined. Given a fuzzified time series $F(t)$, the traditional Seasonal FTS defines a seasonal period i such that $F(t) = F(t-i)$. The Multi-Seasonal FTS deals with the definition of multiple and nested seasonal indexes $i_0, .., i_k$ with different time granularities $g_0, ..., g_k$, such that $F(t) = F(t \mod g_0 = i_0 \wedge ... \wedge t \mod g_k = i_k)$. It allows us to construct seasonal indexes by combining months, hours, minutes etc.

For each data point $F(t) = A_k$, the time index t is decomposed such that $m_i = t \mod g_i$, where m_i is the seasonal index, generating the seasonal pattern $[i_0, ..., i_k] \rightarrow A_k$. For example consider the $F(t)$ on (1), using time granularities month $(0-11)$ and hour $(0-23)$ the seasonal indexing generates the patterns presented on (2). These seasonal patterns are equivalent to the FLR's on conventional FTS models.

$$\begin{array}{ll}
(\text{2015-01-03 00:00am}, A_0) & (\text{2015-01-03 06:00am}, A_1) \\
(\text{2015-01-03 12:00pm}, A_2) & (\text{2015-01-03 06:00pm}, A_1) \\
(\text{2015-01-04 00:00am}, A_0) & (\text{2015-01-04 06:00am}, A_0) \\
(\text{2015-01-04 12:00pm}, A_1) & (\text{2015-01-04 06:00pm}, A_1)
\end{array} \quad (1)$$

$$
\begin{array}{lll}
(0,0) \rightarrow A_0 & (0,6) \rightarrow A_1 & (0,12) \rightarrow A_2 \\
(0,18) \rightarrow A_1 & (0,0) \rightarrow A_0 & (0,6) \rightarrow A_0 \\
(0,12) \rightarrow A_1 & (0,18) \rightarrow A_1 &
\end{array} \tag{2}
$$

Then all seasonal patterns with the same seasonal indexes are grouped, representing all fuzzy possibilities that may happen on that seasonality. The final model is the set of all possible seasonal indexes and their consequent fuzzy sets. By instance, consider the seasonal patterns on (2) will generate the seasonal pattern groups on (3).

$$
\begin{array}{ll}
(0,0) \rightarrow A_0 & (0,6) \rightarrow A_0, A_1 \\
(0,12) \rightarrow A_1, A_2 & (0,18) \rightarrow A_1
\end{array} \tag{3}
$$

Step E) Model selection Based on the previous steps, many different models can be created, varying the transformations applied to the series (differential or none), the type of partitioning of the UoD (grid or entropy), the number of partitions and the seasonal indexes defined. Given all the possible combinations, all distinct models m_j are benchmarked on a validation dataset and ranked by their Root Mean Squared Error (RMSE). After that, the best 20 models are selected and used to create the ensemble $\mathcal{M}$.

3.2 Forecasting Procedure

With the ensemble built as in previous section, it is desired to forecast a full probability distribution for $Y(t+1)$. The overall process is described in Figure 2 and is composed by: a) the forecasting of individual models; b) forecast selection and c) distribution smoothing with the KDE.

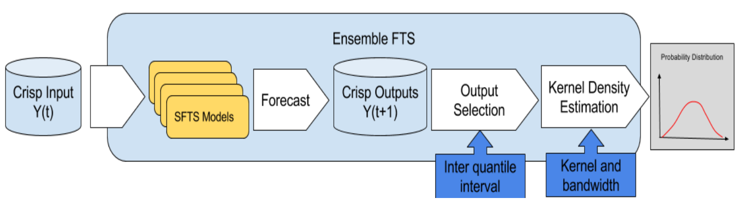


Fig. 2: Ensemble FTS forecasting procedure

Step A) Individual Seasonal FTS forecasts On Multi-Seasonal FTS the forecast is defined solely by extracting the seasonal indexes from t, using the same method of model training. With the seasonal indexes $i_0, ..., i_k$ extracted from t it is necessary to find the seasonal pattern group with this antecedent or premise and the fuzzy forecast, $F(t+1)$ will be the consequent.
The crisp forecast is calculated by defuzzifying $F(t+1)$ using the center of mass of each fuzzy set $A_i \in F(t+1)$, such that $Y(t+1) = n^{-1} \sum_1^n A_i$ where

n is the number of fuzzy sets in $F(t+1)$. The output set of this step is composed by the individual forecasts $Y_j(t+1)$ of each model $m_j \in \mathcal{M}$.

Step B) Forecast selection In order to control the total forecast variance and eliminate the effect of possible outliers the forecasted output is limited by an inter quantile interval $(\alpha, 1-\alpha)$ where $\alpha \in (0,1)$ is the confidence level. By varying α parameter it is possible to fine tune the final distribution accuracy by eliminating forecasts that are too distant from the mean.

Step C) Kernel density estimation The final step uses KDE methods to smooth the individual forecasts and provide a continuous probability density distribution over the UoD. Two parameters are necessary on this step: the type of kernel and the bandwidth h. Both parameters are domain specific and need to be empirically evaluated for each application. Common kernel functions are Epanechnikov, Histogram, Tophat and Triangular or Retangular kernels are used to create different models [27]. The bandwidth h was selected based on computational experiments, with values varying in the interval [0.1, 0.9].

4 Computational Experiments

In order to evaluate the proposed method[4] among the environmental datasets available in the literature, the SONDA - System of National Organization of Environmental Data (Sistema de Organização Nacional de Dados Ambientais - in Portuguese) stands out for providing reliable and valuable information to be used by energy sector in planning and deployment of energy resources for electricity generation and distribution in Brazil [19]. For this work, three different seasonal indexes were used in order to experiment with the best period as follows:

1. 15 minutes interval index, which groups the data in 15 minutes intervals during the day (0-95) and allows 96 FLRG;
2. Month (1-12) and hour (0-23) index, which allows 288 FLRG;
3. Month (1-12), hour (0-23) and 15 minutes interval (0-3) allowing 1152 FLRG.

The instances[5] were split as Training data (from January 2012 to December 2014. 105.312 instances) and Test data (from January 2015 to November 2015. 24.225 instances).

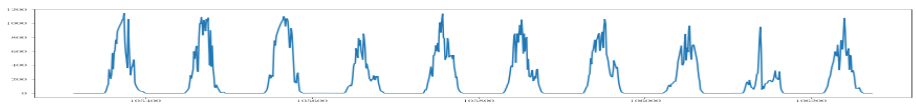
The CRPS (Continuous Ranked Probability Score) is used to evaluate the probabilistic predictions of the models [10] and can be seen as a ranked probability score.

4.1 Ensemble FTS parameter tunning

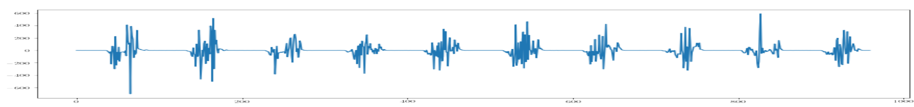
It was adopted an empirical method to select the models and parameters of the Ensemble FTS, based on exploring all the parameter space by building models and evaluating their performances on validation dataset.

[4] Source code is available in: http://bit.ly/pyFTS_ensemble

[5] Dataset is available in: a) http://bit.ly/sonda_bsb_hourly and b) http://bit.ly/sonda_bsb_15

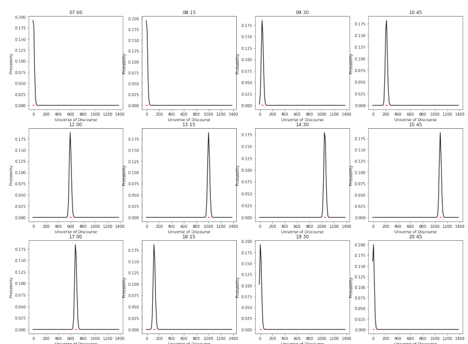


(a) Original data

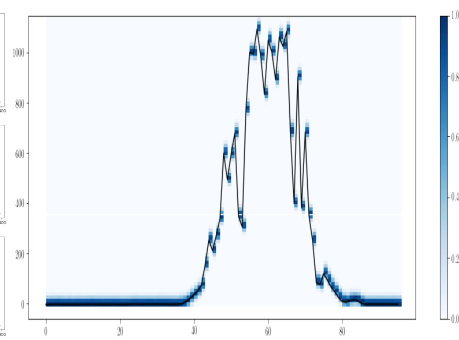


(b) Defferentiated data

Fig. 3: 10 day sample of SONDA dataset



(a) Probability density along the time



(b) Probability density over the SONDA time series

Fig. 4: Ensemble FTS sample results for SONDA dataset

The models were trained both with original data and differentiated data and also varying the number of partitions of the UoD in the interval from 10 to 90 (with a step size of 10). For each one, Grid and Entropy partitioning were tested, both with triangular fuzzy membership functions. In total 108 models were trained and evaluated, of which the best 20 ones were selected to compose the ensemble. The inter-quantile parameter $\alpha \in [0.05, 0.95]$ was evaluated under its CRPS score. It was chosen $\alpha = 0.1$. The kernel function and bandwidth h parameter were also tested with the built ensemble. We tested the kernels Epanechnikov, Histogram, Tophat and Triangular considering the bandwidth in the interval $h \in [0, 1]$. The best CRPS scores were achieved by the Epanechnikov kernel with $h = 0.55$, the values used for benchmarks.

The forecast of the trained Ensemble FTS can be seen in Figure 4a, where one can see the probability distributions along a day, and in Figure 4b, where the probability density is plotted over the time series data.

4.2 Model Evaluation

Following the parameter tuning for the Ensemble FTS models, their predictions for the test dataset were evaluated using the CRPS metric. ARIMA and QAR models were also evaluated for comparison. The results are shown in Figure 5.

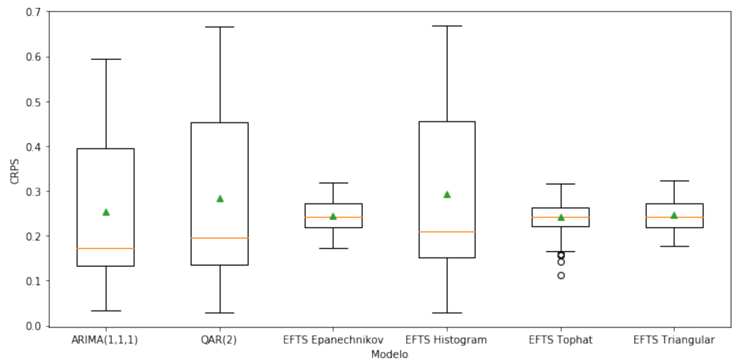


Fig. 5: CRPS values boxplot

In order to check for a statistically significant difference between the mean CRPS for each model, an ANOVA test was performed, corrected for the heteroscedasticity present in the data. The test result refuted the null hypothesis that all the models are equivalent in terms of mean CRPS ($p\_value = 5.708e-07$).

A multiple comparison test was then performed between each pair of models, considering a significance level of 95% ($\alpha = 0.05$). The comparisons are summarized in Figure 6. The results show non-inferiority between the Tophat, Epanechnikov and Triangular ensemble models proposed and the ARIMA model, for the effect sizes $\delta^\star$ greater than 0.02. These three models also have a statistically significant lower mean than the QAR model, and there is no statistically significant difference between them. Finally, the Histogram ensemble model has the worse mean than all models (with statistical significance for all except the QAR model)

4.3 Discussion

The best models as candidates for the Ensemble were those with 15-minute seasonal indexes. The transformation of the data using Differentiation had a positive result, as well as other pre-processing, such as horizontal and vertical reduction.

It was observed that 18 of the 20 ensemble models formed from the combination of different partitioning models of the UoD and seasonality indexes, the ones with the best performance in terms of RMSE were those of 15-minute

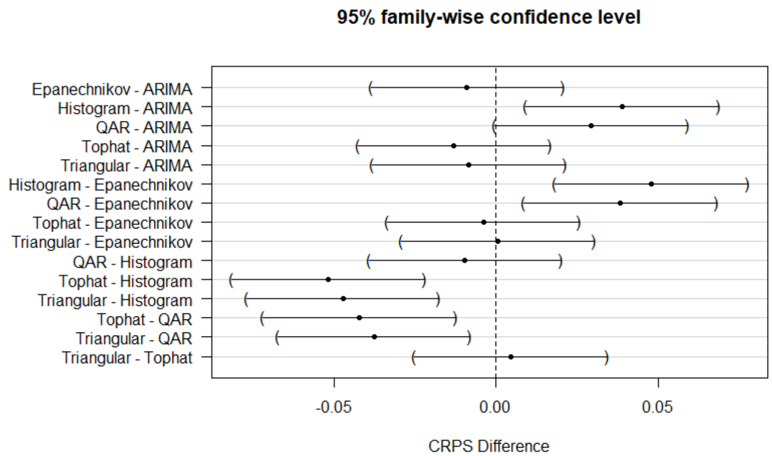


Fig. 6: Pairwise CRPS comparison between models

seasonality indexes and the Entropy partitioning model in their different parameters. It is known that grouping by minutes provides much more data available and this allows better learning of the proposed model. The monthly evaluation, considering the range of available data (2 years for training and 1 for testing), may not provide as much information as possible to allow for a better accuracy of the model.

Three of the proposed ensemble FTS models (Tophat, Epanechnikov and Triangular) had better CRPS results than QAR model, and were non-inferior to the ARIMA model. They also had much smaller variance, which indicates more stability in a model.

It is noted that for some seasonal time series, especially those with environmental data, data can be indexed in several ways, for example by a minute, hour, month, quarter etc, and combinations between them allowing to capture of different seasonality effects.

The number of partitions of the UoD has a great impact on FTS accuracy, and different partitioning schemes generate different FTS representations. So, on Ensemble FTS the models were trained with both schemes and several numbers of partitions with the objective that the model diversity, and their consequent forecasting variance, would cover different scenarios and fluctuations that an isolated model cannot cover.

5 Conclusions

This paper deals with a new method for probabilistic forecast which aggregates traditional seasonal Fuzzy Time-Series techniques with ensemble learning and KDE. The proposed method has as main feature the capability of capturing a wide spectrum of seasonal effects and time series trends by the variation of several parameters of its internal models. A parameter tuning method is also presented.

We tested the proposed method in SONDA dataset, comparing the CRPS values of the proposed models and ARIMA and QAR models. Statistical tests show that the proposed method resulted in models with non-inferior mean CRPS and much lower variance for this dataset. These results show that the method could bring better results for some probabilistic predictions problems, and suggest further investigation with other datasets.

Acknowledgments

This work has been supported by the Brazilian agency CAPES and CNPq.

References

[1] Ahmed Mohammed, A., Aung, Z.: Ensemble learning approach for probabilistic forecasting of solar power generation. Energies 9(12), 1017 (2016)
[2] Bröcker, J., Smith, L.A.: From ensemble forecasts to predictive distribution functions. Tellus, Series A: Dynamic Meteorology and Oceanography 60 A(4), 663–678 (2008)
[3] Chang, P.T.: Fuzzy seasonality forecasting. Fuzzy Sets and Systems 90(1), 1–10 (1997)
[4] Chatfield, C.: Prediction Intervals for Time-Series Forecasting. In: Armstrong, J. (ed.) Principles of forecasting, pp. 475–494. Springer US, 1 edn. (2001)
[5] Chen, D.W., Zhang, J.P.P.: Time Series Prediction Based On Ensemble ANFIS. Proceedings of the Fourth International Conference on Machine Learning and Cybernetics pp. 18–21 (2005)
[6] Chen, M.Y.: A high-order fuzzy time series forecasting model for internet stock trading. Future Generation Computer Systems 37, 461–467 (2014)
[7] Chen, S.M.: Forecasting enrollments based on fuzzy time series. Fuzzy Sets and Systems 81(3), 311–319 (1996)
[8] Cheng, C.H., Chen, T.L., Teoh, H.J., Chiang, C.H.: Fuzzy time-series based on adaptive expectation model for TAIEX forecasting. Expert Systems with Applications 34, 1126–1132 (2008)
[9] Gneiting, T.: Editorial: Probabilistic forecasting. Journal of the Royal Statistical Society 171(2), 319–321 (2008)
[10] Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather and Forecasting 15(5), 559–570 (2000)
[11] Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., Hyndman, R.J.: Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. International Journal of Forecasting 32, 896–913 (2016)
[12] Ismail, Z., Efendi, R., Deris, M.M.: Application of Fuzzy Time Series Approach in Electric Load Forecasting. New Mathematics and Natural Computation 11(3), 229–248 (2015)
[13] Koenker, R., Hallock, K.F.: Quantile Regression. Journal of Economic Perspectives—Volume 15(4), 143–156 (2001)

[14] Koenker, R., Xiao, Z.: Quantile Autoregression. Journal of the American Statistical Association 101(405), 980–990 (2006)
[15] Kumar, A.: Convexity , Loss functions and Gradient pp. 1–15 (2011)
[16] Lee, M.H., Javedani, H.: A weighted fuzzy integrated time series for forecasting tourist arrivals. Informatics Engineering and Information Science pp. 206–217 (2011)
[17] Lee, M.H., Sadaei, H.J., Suhartono: Improving taiex forecasting using fuzzy time series with box–cox power transformation. Journal of Applied Statistics 40(11), 2407–2422 (2013)
[18] Mohammed, A.A., Yaqub, W., Aung, Z.: Probabilistic forecasting of solar power: An ensemble learning approach. In: Intelligent Decision Technologies, pp. 449–458. Springer (2015)
[19] INPE Instituto Nacional de Pesquisas Espaciais, .: Sonda - sistema de organização nacional de dados ambientais (2017), http://sonda.ccst.inpe.br/
[20] Pinson, P., Kariniotakis, G.: Conditional prediction intervals of wind power generation. IEEE Transactions on Power Systems 25(4), 1845–1856 (2010)
[21] Pinson, P., Madsen, H.: Ensemble-based probabilistic forecasting at horns rev. Wind energy 12(2), 137–155 (2009)
[22] Poulsen, J.: Fuzzy time series forecasting–developing a new forecasting model based on high order fuzzy time series. AAUE: CIS 4, 67 (2009)
[23] Sadaei, H.J., Enayatifar, R., Lee, M.H., Mahmud, M.: A hybrid model based on differential fuzzy logic relationships and imperialist competitive algorithm for stock market forecasting. Applied Soft Computing 40, 132–149 (2016)
[24] Sadaei, H.J., Guimarães, F.G., da Silva, C.J., Lee, M.H., Eslami, T.: Short-term load forecasting method based on fuzzy time series, seasonality and long memory process. International Journal of Approximate Reasoning 83, 196–217 (2017)
[25] Sheather, S.J., Jones, M.C.: A reliable data-based bandwidth selection method for kernel density estimation. Journal of the Royal Statistical Society. Series B (Methodological) pp. 683–690 (1991)
[26] Silverman, B.W.: Density estimation for statistics and data analysis, vol. 26. CRC press (1986)
[27] Silverman, B.W.: Density estimation for statistics and data analysis, vol. 26. CRC press (1986)
[28] Sloughter, J.M., Gneiting, T., Raftery, A.E.: Probabilistic wind speed forecasting using ensembles and bayesian model averaging. Journal of the american statistical association 105(489), 25–35 (2010)
[29] Song, Q.: Seasonal forecasting in fuzzy time series. Fuzzy Sets and Systems 107(2), 235–236 (1999)
[30] Song, Q., Chissom, B.S.: Fuzzy time series and its models. Fuzzy Sets and Systems 54(3), 269–277 (1993)
[31] Zadeh, L.A.: Fuzzy sets. Information and control 8(3), 338–353 (1965)