

REFERENCE EVAPOTRANSPIRATION PREDICTION FOR PRECISION AGRICULTURE USING FUZZY TIME SERIES

Patrícia de Oliveira e Lucas

Graduate Program in Electrical Engineering - Universidade Federal de Minas Gerais
Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brazil
patricia.lucas@ifnmg.edu.br

Marcos Antônio Alves

Graduate Program in Electrical Engineering - Universidade Federal de Minas Gerais
Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brazil
marcosalves@ufmg.br

Petrônio Cândido Lima e Silva

Instituto Federal do Norte de Minas Gerais (IFNMG), Brazil. Grupo de Pesquisa em
Ciência de Dados e Inteligência Computacional - ciðic, Brazil
Fazenda São Geraldo, S/N Km 06, 39480-000, Januária, MG, Brazil
petronio.candido@ifnmg.edu.br

Frederico Gadelha Guimarães

Department of Electrical Engineering - Machine Intelligence and Data Science (MINDS)
Laboratory, Universidade Federal de Minas Gerais, Brazil
Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brazil
fredericoguimaraes@ufmg.br

RESUMO

Predição de evapotranspiração de referência (ET_o) é um processo importante na agricultura de precisão, que impacta na conservação da água para a gestão da irrigação. Como as soluções atuais possuem alto custo computacional e baixa explicabilidade, os métodos de séries temporais nebulosas (FTS) se encaixam como uma alternativa capaz de produzir previsões pontuais e intervalares de um a muitos passos à frente. Neste artigo, modelos FTS foram aplicados com diferentes ordens, variáveis e tipo de previsão. Como métodos competidores, foram utilizados os métodos ARIMA, KNN e QAR. Para avaliar as previsões de pontos e intervalares foram utilizados RMSE e Score Winkler. Os resultados mostraram que o método FTS Multivariado Ponderado superou as abordagens tradicionais e outros modelos FTS na previsão por ponto, mas tem desempenho inferior na previsão intervalar. Conclui-se que os modelos FTS são viáveis para predição de séries temporais ET_o proporcionando boa precisão e bom tempo computacional.

PALAVRAS CHAVE. Evapotranspiração de referência. Séries Temporais Nebulosas. Predição de séries temporais.

IC – Inteligência Computacional. **ESTMP** – Estatística e Modelos Probabilísticos. **AG&MA** – PO na Agricultura e Meio Ambiente.

ABSTRACT

The reference evapotranspiration (ET_o) forecast is an important process of precision agriculture, which has impact on water conservation for irrigation management. Since the current solutions have high computational cost and low explainability, the fuzzy time series (FTS) methods fit as an alternative capable of produce point and interval forecasts from one to many steps ahead. In this paper, we apply FTS models with different order, variables and type of prediction. As a baseline, it was used the known ARIMA, KNN and QAR methods. To evaluate the point and interval predictions, it was used RMSE and Score Winkler. The results showed that the Weighted Multivariate FTS method outperformed the traditional approaches and other FTS models in the point forecast, but has lower performance of interval forecasting. We conclude that FTS models are viable for ET_o time series prediction providing good accuracy and good computational time consuming.

KEYWORDS. Reference evapotranspiration. Fuzzy time series. Time series prediction.

IC – Inteligência Computacional. **ESTMP** – Estatística e Modelos Probabilísticos. **AGMA** – PO na Agricultura e Meio Ambiente.

1. Introduction

The agricultural sector is increasingly seeking for accurate, efficient and fast computational applications to ensure an adequate and regular supply of food to society with social responsibility. In the Agenda for Sustainable Development proposed for 2030 [Nations, 2015], strongly related to Agriculture 4.0, one of the goals for the governments is to provide best practices to agriculture. A real example is the irrigation management, an essential issue that deals with reference evapotranspiration (ET_o) forecast to know in advance the needs of crops to planning and manage water resources.

Agriculture 4.0 [Zarco-Tejada et al., 2014] and of smart farming [Liakos et al., 2018] come at this juncture of resource scarcity as management that seeks through approaches such as Artificial Intelligence (AI) to supply plants and animals their needs, increasing productivity and reducing environmental impacts. At the field of irrigation management, the prediction of the ET_o defines the volume of water to be applied to an irrigation event based on weather data [Patil e Deka, 2016; Chia et al., 2020]. Basically, ET_o defines the amount of water evaporated and transpired by a surface of soil with vegetable. Thus, the good prediction of ET_o allows reduce the water consumption and the cost of electricity by allowing it to be used at lower cost periods [Perea et al., 2015], to offer possibility to analyze possible stress in the cultivation [Zhao et al., 2019] and operation of reservoirs, is possible to predict the amount of water to be delivered [Bachour et al., 2016].

In general, the decision making in irrigation systems depends on forecasts of 1 to 10 steps ahead. The shorter time, up to 7 days, for example, can be used in planning the use of irrigation systems as well as in optimizing electricity. Longer forecast horizons, from 7 to 10 days, help in the water management of irrigation channels in irrigated perimeters and in reservoirs for irrigation use [Lucas, 2019].

There is a vast literature in order to monitor the climatic conditions of cultivated areas for the purpose of estimating and forecasting ET_o . Recently, Chia et al. [2020] presented the advances in evapotranspiration estimation using AI techniques. Manikumari N et al. [2017] proposed a daily reference estimation by using neural networks ensemble learning. Tawegoum et al. [2015] apply the parametric SARIMA model to predict hourly ET_o , while Lucas [2019] applied more complex models such as Convolutional Neural Networks (CNN), Temporal CNN and ensemble with probabilistic predictions.

Despite these researches reported good accuracy and provide advances to the industries and agriculture, these approaches have problems with the stochastic nature of the time series, interpretability and are not able to consider one or more variables. Based on this, we propose the application of Fuzzy Time Series (FTS) models [Song e Chissom, 1993b,a, 1994] in order to predict the ET_o . FTS methods produce data-driven, white-box and non-parametric models and have become attractive due to their simplicity, versatility, high accuracy and low computational cost. Other key advantages include the deterministic nature of its training procedures and the human-readable representation of its models making its results easy to audit and reproduce.

In this paper, we evaluate different FTS models in order to cover the existing gap in the literature on reference evapotranspiration predictions. For this, we used a dataset with data collected at a Brazilian station. It was tested univariate and multivariate FTS methods, with different orders and evaluating point and interval predictions. It was also compared the results achieved with the parametric methods Autoregressive Integrated Moving Average (ARIMA), K-Nearest Neighbors (KNN) and Quantile Autoregression (QAR).

2. Preliminaries

Evapotranspiration can be defined as the amount of water evaporated and transpired by a vegetable surface over a given period of time. Basically, it is the equivalent height of evaporated water in $mm\ t^{-1}$ where t is the time unit (hours, day, month, growing seasons, or years [Frizzzone et al., 2012]. Since the ET is a continuous process, the water lost to the atmosphere is replenished through precipitation and moist soil. In the context of agricultural production, when the available water supply is naturally not sufficient for a crop, it must be supplied by irrigation.

Thus, properly managing irrigation involves determining when and how much water is required for a crop to produce well. The amount of water required by a crop is calculated from the reference evapotranspiration (ET_o) and the crop coefficient (K_c) [Allen et al., 2006]. The characteristics of the ET_o were described by Allen et al. [2006].

The methods to determine the ET_o can be direct or indirect. The former is the lysimeters and the water balance methods in the soil. The latter includes the micro and meteorological, theoretical and empirical models based on the use of climatic data, such as parametric and non-parametric time series models [Frizzzone et al., 2012].

The most precise method to measure the ET_o is the lysimeter, however, this equipment is huge and expensive. The measurement takes a lot of time, hard execution and depends on the farm conditions, justifying its use aimed to develop and calibrate estimation methods [Pereira et al., 2013].

Among the available models to estimate the ET_o by indirect methods, the Penman-Monteith was selected by the Food and Agriculture Organization of the United Nations (FAO) because of the good results, similar to the direct methods [Frizzzone et al., 2012]. The standard ET_o (ET_{FAO}) estimated by the Penman-Monteith equation (1) is parameterized according to with the reference surface (PM-FAO), became a reference in substitution to the lysimeter measurements [Pereira et al., 2013].

$$ET_{FAO} = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

where ET_{FAO} represents the standard ET_o ($mm d^{-1}$), R_n is the balance of radiation on the crop surface ($MJ m^{-2} d^{-1}$), G is the ground heat flow ($MJ m^{-2} d^{-1}$), T is the average daily temperature ($^{\circ}C$), u_2 the wind speed at 2m high (ms^{-1}), e_s is the saturation value pressure (kPa), e_a is the steam pressure (kPa), $e_s - e_a$ is the deficit of the saturation value pressure (kPa), Δ is the slope of the vapor pressure curve ($kPa C^{-1}$) and γ is the psychrometric constant ($kPa C^{-1}$). This equation utilizes the parameters: air temperature, air humidity, heat stroke and wind speed.

3. Fuzzy Time Series

The Fuzzy Time series was first introduced by Song e Chissom [1993a, 1994]; Chen [1996] as a method that employs Fuzzy Sets [Zadeh, 1965] to transform numerical time series into linguistic time series and then model its behavior (generally using membership matrices or rules). The trained model is later used to forecast unseen values of the time series.

Given a univariate time series $Y \in \mathbb{R}$, and its individual values $y(t) \in Y$, for $t = 0, 1, \dots, T$, Each membership function $\mu_{A_i} : \mathbb{R} \rightarrow [0, 1]$ is responsible to assess the membership of a value $y(t) \in Y$ to a fuzzy set $A_i \in \tilde{A}$. There are many possible shapes for

Membership functions but the most common are the triangular, trapezoidal, sigmoidal and Gaussian functions.

There are several categories of FTS methods, varying mainly by their order Ω 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 Y , First Order models need just $y(t-1)$ data to forecast $y(t)$, while Higher Order models require $y(t-1), \dots, y(t-\Omega)$ data to forecast $y(t)$. The time variance indicates if an FTS model $\mathcal{M}$ changes along the time. If the model is static, it is called Time Invariant, otherwise Time-Variant. This work focuses only on Time Invariant methods.

The pioneer FTS work in Song e Chissom [1993a] used weight matrices to store the knowledge learned from the data. In Chen [1996] rule models were introduced as a way to improve the readability and computational cost of the FTS approaches. Following the approach proposed by Chen [1996], an FTS model $\mathcal{M}$ is trained with the Training Procedure, which contains three main steps:

- A) *Partitioning*: the Universe of Discourse (UoD) U is delimited by the known bounds of Y , such that $U = [\min(Y), \max(Y)]$. U is then divided in k overlapping intervals and for each interval a fuzzy set A_i is defined with its membership function μ_{A_i} and midpoint c_i . A linguistic variable $\tilde{A}$ is defined upon U , where the linguistic terms are fuzzy sets $A_i \in \tilde{A}$, for $i = 1 \dots k$.
- B) *Fuzzyfication*: transform the original numeric time series Y into a fuzzy time series F , where each data point $f(t) \in F$ is an $1 \times k$ array with the fuzzyfied values of $y(t) \in Y$ with respect to the linguistic terms $A_i \in \tilde{A}$.
- C) *Knowledge Extraction*: analyze the sequential terms $f(t-1), f(t)$ in F , grouping the patterns in rules $A_i \rightarrow A_j, A_k, \dots$ which can be read as "IF $f(t)$ IS A_i THEN $f(t+1)$ IS $A_j, A_k, \dots$ ".

The trained FTS model $\mathcal{M}$ contains both the linguistic variable $\tilde{A}$ and the rules. Given a sample $y(t)$ model can be used to forecast $y(t+1)$ with the forecasting procedure:

- A) *Fuzzyfication*: transform the numeric sample $y(t) \in Y$ into a fuzzy sample $f(t)$, where $f(t) \in \tilde{A}$.
- B) *Rule Matching*: Find the rules that contain the fuzzy set $f(t)$ in its precedent, the set of selected rules is R . The activation $\mu_r \in R$ of the rules are given by the membership values in $f(t)$.
- C) *Defuzzyfication*: For each rule $r \in R$, first find the midpoint of each rule by summing the midpoints of each fuzzy set of the consequent:

$$mp_r = \sum_{i \in \text{consequent}} c_i \quad (2)$$

Compute the forecast as the weighted sum of the rule mid-points mp by their membership grades μ for each selected rule r :

$$y(t+1) = \frac{\sum_{r \in R} \mu_r \cdot mp_r}{\sum_{r \in R} \mu_r} \quad (3)$$

There are several models available in the literature that extended the classical Song and Chissom method [Song e Chissom, 1993a, 1994], such as High Order FTS (HOFTS) [Severiano et al., 2017]. Weighted FTS (WFTS) [Yu, 2005] which add weights to the rules and modify the way that the rule midpoint is calculated. Other weighted methods include Improved Weight FTS (IWFTS) [Efendi et al., 2013], Trend-Weighted FTS (TWFTS) [Cheng et al., 2006] and Exponentially Weighted FTS (EWFTS) [Sadaei et al., 2014]. In Silva [2019] is proposed the Weighted High Order FTS (WHOFTS), which extend the HOFTS model including weights in its rules. In Silva et al. [2019d] a new kind of weighted method is proposed, the Probabilistic Weighted FTS (PWFTS), including weights in the precedent and the consequent of the rules. These weights represent conditionals probabilities and the PWFTS was the first approach to integrate point, interval and probabilistic forecasting for one to many steps ahead. The most important hyperparameters of an FTS model are summarized in Table 1.

Alias	Parameter
$k \in \mathbb{N}^+$	Number of partitions (fuzzy sets)
$\mu : U \rightarrow [0, 1]$	Measures the membership of a value $y \in U$ to a fuzzy set
$\Omega \in \mathbb{N}^+$	Order, the number of past lags used by the model

Table 1: FTS hyperparameters

The most common FTS approaches are univariate methods. However many multivariate approaches were presented. These approaches can be divided into Multiple Input Single Output (MISO) and Multiple Input Multiple Output (MIMO) approaches. The multivariate time series are matrices $Y \in \mathbb{R}^n$ where $n = |\mathcal{V}|$ and $\mathcal{V}$ is the set of attributes of Y . Each vector $y(t) \in Y$ contains all attributes $\mathcal{V}_i \in \mathcal{V}$. There are some FTS methods for multivariate time series, such as Chen e Tanuwijaya [2011].

In the MISO approaches, for the set of variables $\mathcal{V}$ one of them is chosen as the target (or endogenous) variable $\ast \mathcal{V}$ and the others are referred to as the explanatory (or exogenous) variables. For sake of simplification, from now on the target variable will be distinguished from the others by an asterisk. That is the case of the Weighted Multivariate FTS (WMVFTS) method [Silva et al., 2019a], which consists of a set of variables $\mathcal{V}$, a fuzzy linguistic variable $\tilde{\mathcal{V}}$ for each $\mathcal{V}_i \in \mathcal{V}$ and a set of weighted fuzzy rules over the linguistic variables $\tilde{\mathcal{V}}$. The inputs of the training procedure are the crisp time series training data Y and the set of hyperparameters (k and μ) for each $\mathcal{V}_i \in \mathcal{V}$. This method is first order but allows point and interval forecasting for one to many steps ahead.

A MIMO approach has to be able to forecast all the input variables (or a subset of interest), i. e., all variables are target variables. That is the case of the Fuzzy Information Granule-Fuzzy Time Series (FIG-FTS) method [Silva et al., 2019c], which proposes a new kind of multivariate fuzzy set based on Fuzzy Information Granules (FIG) [Zadeh, 1996]. The FIG transforms each multivariate point $y(t)$ in a single fuzzy set, enabling the use of common univariate models to forecast multivariate time series. The FIG-FTS method is high order and capable to forecast points, intervals and probability distributions for one to many steps ahead.

In the literature, several extensions are embracing other aspects of FTS and a review of the methods can be found in Silva [2019] while the review of concepts, applications and future directions was recently provided in Bose e Mali [2019].

Interval forecasting approaches were first proposed in Silva et al. [2016], providing prediction intervals $\mathbb{I} = [\underline{l}, \bar{u}]$ and it is expected the the true value will fall between the lower bound ($\underline{l}$) and the upper bound ($\bar{u}$) of the interval. Prediction intervals are a way to represent the inherent uncertainty of the forecasting, usually not well understood when dealing only with point forecasts. The Interval FTS (IFTS) [Silva et al., 2016] calculates the interval bounds using the bounds of the fuzzy sets of the matched rules in the defuzzification step. After the IFTS, Ensemble FTS [Silva et al., 2017] and PWFTS [Silva et al., 2019d] also proposed new interval forecasting methods.

4. Computational Experiments

4.1. Dataset

The climatic data used to generate the time series of ET_o come from station 83586 of National Institute of Meteorology (*Instituto Nacional de Meteorologia* (INMET), in Portuguese), located in Sete Lagoas city, Brazil. The historical data used covers the period from 1993 to 2015, constituting 22 years of daily data, referring to the following meteorological variables: T_{max} and T_{min} , maximum and minimum temperature, respectively, in Celsius; U_2 the wind speed at 2 meters high (ms^{-1}); RH the average relative humidity (%); and I the insolation (h).

The ET_o was calculated for each day using Equation (1), generating the time series illustrated in Fig 1.

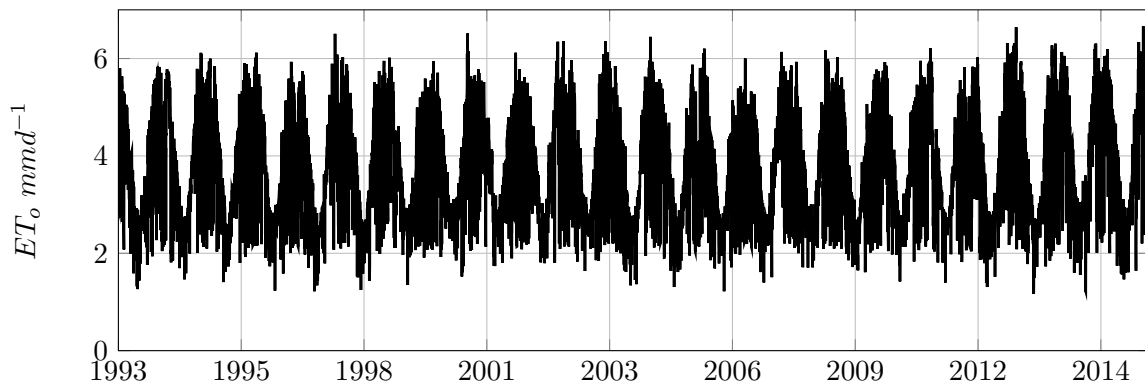


Figure 1: Time series of daily reference evapotranspiration from 1993 to 2015.

It was used 8000 instances with a sliding window of 2000 instances, being 80% training data and 20% test.

4.2. Methodological procedures

Four candidate FTS methods were evaluated: WHOFTS [Silva, 2019], PWFTS [Silva et al., 2019d], WMVFTS [Silva et al., 2019b] and FIG-FTS [Silva et al., 2019c]. The competing methods are the known ARIMA, KNN and QAR. The idea was to evaluate these methods applied to reference evapotranspiration. The approach aims to help the agricultural sector in better irrigation management, especially in sustainable water consumption.

The FTS methods were tested with different orders (1, 2 and 3) and considering point and interval forecasts. Regarding the number of partitions, it was employed a recent optimization technique provided in Lucas et al. [2019], and returned that the best value is

50. For point and also for interval forecast, we tested different steps ahead (1 to 7). As mentioned before, shorter time, i.e. 1-7 days, are frequently used by the experts. A brief notation about the methods, order e variables tested are described in Table 2. The best model is also shown.

Method	Order	# variables	Best Model
WHOFTS	1, 2*, 3*	1	3
PWFTS	1, 2, 3	1	3
WMVFTS	1	N	1
FIG-FTS	1, 2, 3	N	3
ARIMA	1, 0, 0; 1, 0, 1; 2, 0, 0	1	(2,0,0)**
KNN	1, 2, 3	1	2
QAR	1, 2	1	2

Table 2: Information about the methods, order and variables. * WHOFTS for order 2 and 3 we used the mechanism from Interval FTS. ** The parameters (2,0,0) for ARIMA was chosen based on “auto.arima” function.

The metric used to evaluate the point predictions was the Root Mean Squared Error (RMSE) (4), where y are the real data and $\hat{y}$ are the predict values. For the interval predictions, we used the Winkler Score (WS) metric [Winkler, 1972], where the target value y and an interval prediction $I = [l, u]$ with nominal probability $(1 - \alpha)$. WS score is defined by the Equation (5), where $\delta = u - l$. The average score is defined by the Equation (6), where T is the sample size.

$$RMSE = \sqrt{\sum_{t=0}^n (y(t) - \hat{y}(t))^2} \quad (4)$$

$$WS(\alpha, y(t), \Pi(t)) = \begin{cases} \delta & \text{if } l \leq y \leq u \\ \delta + 2(\bar{l} - y)/\alpha & \text{if } y < \bar{l} \\ \delta + 2(y - \underline{u})/\alpha & \text{if } \underline{u} < y \end{cases} \quad (5)$$

$$\overline{WS(\alpha, Y, \Pi)} = T^{-1} \sum_{t=1}^T S(\alpha, y(t), \Pi(t)) \quad (6)$$

5. Results and Discussion

The average RMSE for point forecast for each model presented in the Table 2 and the time horizon (H) for 1-7 days are described in Table 3.

The Friedman Aligned Ranks [Friedman, 1940] of these models are (in order): WMVFTS: 3.00000, knn: 3.14286, QAR: 3.28571, WHOFTS: 3.71429, FIG-FTS: 3.85714, PWFTS: 4.00000 and ARIMA: 7.00000. The test statistic for those results is $Q = 4.08824$, where the p-value is equal to 0.00315. For this, H_0 is reject with a confidence interval of $\alpha = 0.05$, indicating that there is a statistical difference between the averages of competing models.

The average results for interval forecasts by using WS metric for each model considering $\alpha = 0.05$, for the time horizons (H) from 1 to 7 are presented in Table 4.

H	ARIMA	QAR	kNN	WHOFTS	PWFTS	WMVFTS	FIG-FTS
1	0.957 $\pm$ 0.48	0.744 $\pm$ 0.79	0.797 $\pm$ 0.57	0.506 $\pm$ 0.4	0.408 $\pm$ 0.49	0.554 $\pm$ 0.45	0.391 $\pm$ 0.5
2	1.042 $\pm$ 0.48	0.525 $\pm$ 0.67	0.559 $\pm$ 0.5	0.708 $\pm$ 0.48	0.598 $\pm$ 0.5	0.745 $\pm$ 0.66	0.585 $\pm$ 0.51
3	1.413 $\pm$ 0.64	0.594 $\pm$ 0.69	0.582 $\pm$ 0.57	0.669 $\pm$ 0.71	0.736 $\pm$ 0.8	0.659 $\pm$ 0.65	0.699 $\pm$ 0.81
4	1.659 $\pm$ 1.29	0.899 $\pm$ 0.99	1.102 $\pm$ 0.83	0.722 $\pm$ 0.72	0.699 $\pm$ 0.78	0.498 $\pm$ 0.56	0.724 $\pm$ 0.7
5	2.174 $\pm$ 0.93	0.61 $\pm$ 0.79	0.742 $\pm$ 0.77	0.855 $\pm$ 0.69	0.914 $\pm$ 0.77	0.969 $\pm$ 0.69	0.979 $\pm$ 0.76
6	2.2 $\pm$ 0.93	0.806 $\pm$ 0.86	0.668 $\pm$ 0.72	0.602 $\pm$ 0.53	0.911 $\pm$ 0.57	0.569 $\pm$ 0.54	0.861 $\pm$ 0.65
7	2.027 $\pm$ 0.75	0.755 $\pm$ 0.66	0.55 $\pm$ 0.36	0.784 $\pm$ 0.7	0.665 $\pm$ 0.34	0.495 $\pm$ 0.46	0.647 $\pm$ 0.43

Table 3: Average RMSE for point forecast

H	ARIMA	QAR	kNN	WHOFTS	PWFTS	WMVFTS	FIG-FTS
1	0.606 $\pm$ 0.42	0.847 $\pm$ 0.54	0.739 $\pm$ 0.71	15.795 $\pm$ 16.06	114.942 $\pm$ 68.98	17.934 $\pm$ 17.58	121.268 $\pm$ 63.22
2	1.165 $\pm$ 0.66	0.846 $\pm$ 0.78	0.729 $\pm$ 0.65	22.285 $\pm$ 17.42	122.664 $\pm$ 66.06	23.644 $\pm$ 25.85	133.608 $\pm$ 52.21
3	1.461 $\pm$ 0.7	0.961 $\pm$ 0.76	0.699 $\pm$ 0.74	22.275 $\pm$ 31.49	126.839 $\pm$ 77.27	17.98 $\pm$ 24.86	135.02 $\pm$ 68.51
4	1.369 $\pm$ 1.05	1.065 $\pm$ 0.9	0.891 $\pm$ 1.04	24.714 $\pm$ 29.77	133.742 $\pm$ 75.28	13.784 $\pm$ 22.52	146.089 $\pm$ 58.22
5	1.534 $\pm$ 0.9	1.23 $\pm$ 1.13	1.121 $\pm$ 1.15	27.464 $\pm$ 23.97	127.498 $\pm$ 67.93	29.962 $\pm$ 26.45	144.876 $\pm$ 40.62
6	1.126 $\pm$ 1.11	1.045 $\pm$ 1.23	1.161 $\pm$ 1.26	14.162 $\pm$ 16.22	113.634 $\pm$ 61.56	14.151 $\pm$ 19.28	130.154 $\pm$ 36.76
7	0.971 $\pm$ 0.94	0.94 $\pm$ 1.0	1.009 $\pm$ 0.88	23.892 $\pm$ 24.01	118.8 $\pm$ 67.78	11.231 $\pm$ 16.44	135.211 $\pm$ 45.4

Table 4: Average WS metric considering $\alpha = 0.05$

The Friedman Aligned Ranks of the models are (in order): kNN: 1.71429, QAR: 1.85714, ARIMA: 2.42857, MVFTS: 4.42857, WHOFTS: 4.57143, PWFTS: 6.00000 and FIG-FTS: 7.00000. The statistical test for these results is $Q = 70.22222$ with p-value equal to 0.00000. For this statistic, H_0 is reject to $\alpha = 0.05$, indicating that there is a statistical difference between the averages of the competing methods.

To visualize the predicted values by the best models, Figure 2 illustrates the behaviour of WMVFTS for point and interval 1 step ahead, while Figure 3 shows this method predicting 40 steps ahead.

Among the FTS approaches evaluated the WMVFTS had the best performance for point forecasting, being also superior to the traditional forecasting methods as ARIMA, KNN and QAR. This is mainly because of its multivariate structure that aggregates more information than that univariate ones. The result of the interval prediction of the traditional methods had a better performance than the WMVFTS. This is because the interval calculation used by WMVFTS considers the width of the fuzzy sets in its rules as the measure of uncertainty, such that the width of the interval predicted by the WMVFTS is the weighted sum of the widths of the fuzzy sets by their degree of activation. Therefore, the

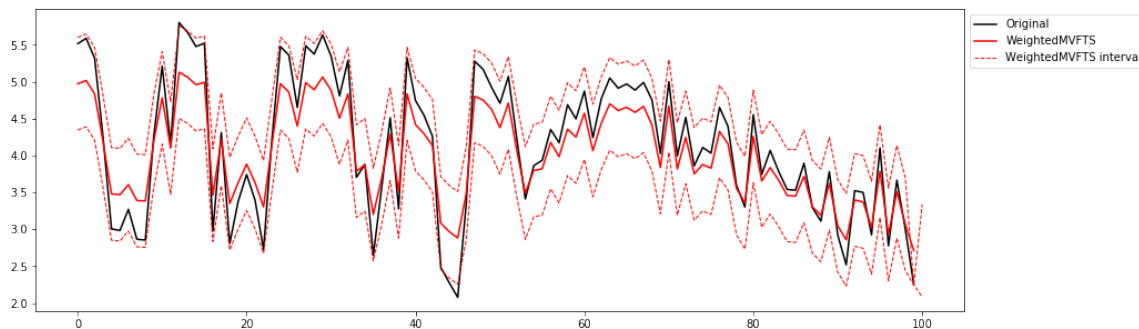


Figure 2: Point and interval forecast by the WMVFTS for 1 step ahead.

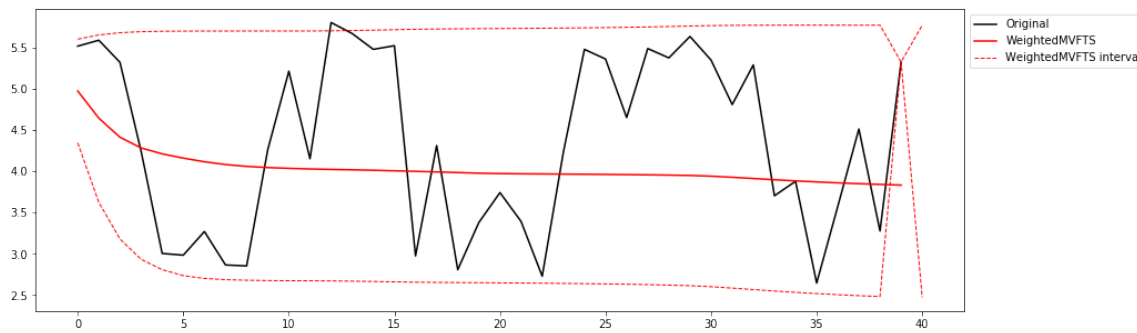


Figure 3: Point and interval forecast by the WMVFTS for 40 steps ahead.

number of fuzzy sets in the model (i.e. the partitioning) directly reflects the width of the predicted range.

6. Conclusions

Smart farming is quickly becoming a major trend in the agricultural sector, where the forecasting methods play one of the most important roles. In this context, the irrigation management is a process that has great environmental impact and is directly responsible for the conservation of water, saving important amounts of this natural resource. In the irrigation management, the ET_o is the most important variable and is used to identify the amount of water will be demanded by the culture for irrigation. However, direct methods is expensive and AI-based alternative approaches have been helping. FTS models applied to ET_o prediction was introduced in this paper considering different order, number of variables and point and interval forecasts.

For the experiments were considered univariate methods (WHOFTS and PWFTS) and multivariate methods (WMVFTS and FIG-FTS), beyond traditional forecasting approaches (ARIMA, QAR and KNN). The experiments show that WMVFTS outperformed the traditional approaches and the other FTS models in the point forecast but has lower performance of interval forecasting. As an integrated tool it offers explainability, fast computational processing and the possibility to be applied in other areas, such precise agriculture and in the context of smart farming. Point and interval forecasting were considered using the forecasting interval of 7 days, the usual term for irrigation planning by experts. The application cover an existing gap in the literature regarding to the application of FTS to ET_o prediction. Future research challenge involves implementing new strategies for ET_o prediction, evaluate new methods and aggregate more concepts of smart farms, such as the generation of renewable energy and best hidrology practices.

7. Acknowledgements

Patrícia de Oliveira e Lucas would like to thank the support given by the PBQS program of IFNMG. This work has been supported by the Brazilian agencies (i) National Council for Scientific and Technological Development (CNPq), Grant no. 312991/2020-7; (ii) Coordination for the Improvement of Higher Education Personnel (CAPES) and (iii) Foundation for Research of the State of Minas Gerais (FAPEMIG, in Portuguese), Grant no. APQ-01779-21. MINDS Laboratory – <https://minds.eng.ufmg.br/>.

References

- Allen, R. G., Pereira, L. S., e Raes, D. (2006). Evapotranspiración del cultivo. Guías para la determinación de los requerimientos de agua de los cultivos. Technical report, FAO, Roma, Itália. URL <http://academic.uprm.edu/abe/backup2/tomas/fao\%2056.pdf>.
- Bachour, R., Maslova, I., Ticlavilca, A. M., Walker, W. R., e McKee, M. (2016). Wavelet-multivariate relevance vector machine hybrid model for forecasting daily evapotranspiration. *Stochastic environmental research and risk assessment*, 30(1):103–117.
- Bose, M. e Mali, K. (2019). Designing fuzzy time series forecasting models: A survey. *International Journal of Approximate Reasoning*, 111:78–99.
- Chen, S.-M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy sets and systems*, 81(3):311–319.
- Chen, S. M. e Tanuwijaya, K. (2011). Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques. *Expert Systems with Applications*, 38.
- Cheng, C.-H., Chen, T.-L., e Chiang, C.-H. (2006). Trend-weighted fuzzy time-series model for taiex forecasting. In *International Conference on Neural Information Processing*, p. 469–477. Springer.
- Chia, M. Y., Huang, Y. F., Koo, C. H., e Fung, K. F. (2020). Recent advances in evapotranspiration estimation using artificial intelligence approaches with a focus on hybridization techniques—a review. *Agronomy*, 10(1):101.
- Efendi, R., Ismail, Z., e Deris, M. M. (2013). Improved weight fuzzy time series as used in the exchange rates forecasting of us dollar to ringgit malaysia. *International Journal of Computational Intelligence and Applications*, 12(01):1350005.
- Friedman, M. (1940). A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, 11(1):86–92.
- Frizzzone, J. A., Freitas, P. S. L., Rezende, R., e Faria, M. A. (2012). *Microirrigação - Gotejamento e microaspersão*. Editora Eduem, 1^a edition. ISBN 978-85-7628-460-4. URL <http://www.eduem.uem.br/novapagina/?q=node/524>.
- Liakos, K., Busato, P., Moshou, D., Pearson, S., e Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8):2674.
- Lucas, P. O., Silva, P. C. L., e Guimarães, F. G. (2019). Otimização evolutiva e hiper-parâmetros para modelos de séries temporais nebulosas. In *14 Simpósio Brasileiro de Automação Inteligente (SBAI 2019)*. SBA.
- Lucas, P. O. (2019). *Previsão de Séries Temporais de Evapotranspiração de Referência com Redes Neurais Convolucionais*. PhD thesis, Universidade Federal de Minas Gerais. URL <https://10.5281/zenodo.3629616>.
- Manikumari N, Murugappan A, e Vinodhini G (2017). Time Series Forecasting of Daily Reference Evapotranspiration by Neural Network Ensemble Learning for Irrigation System. *IOP Conference Series: Earth and Environmental Science*, 80(1):012069. URL <https://iopscience.iop.org/article/10.1088/1755-1315/80/1/012069/pdf>.

- Nations, U. (2015). Transforming our World: The 2030 Agenda for Sustainable Development. Technical report.
- Patil, A. P. e Deka, P. C. (2016). An extreme learning machine approach for modeling evapotranspiration using extrinsic inputs. *Computers and Electronics in Agriculture*, 121:385–392.
- Perea, R. G., Poyato, E. C., Montesinos, P., e Díaz, J. R. (2015). Irrigation demand forecasting using artificial neuro-genetic networks. *Water Resources Management*, 29 (15):5551–5567.
- Pereira, A. R., Sedyama, G. C., e Villa Nova, N. A. (2013). *Evapotranspiração*. Fundag.
- Sadaei, H. J., Enayatifar, R., Abdullah, A. H., e Gani, A. (2014). Short-term load forecasting using a hybrid model with a refined exponentially weighted fuzzy time series and an improved harmony search. *International Journal of Electrical Power & Energy Systems*, 62:118–129.
- Severiano, C. A., Silva, P. C. L., Sadaei, H. J., e Guimarães, F. G. (2017). Very short-term solar forecasting using fuzzy time series. In *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, p. 1–6.
- Silva, P., Sadaei, H. J., e Guimarães, F. (2016). Interval Forecasting with Fuzzy Time Series. In *IEEE Symposium Series on Computational Intelligence (IEEE SSCI 2016)*, Athens, Greece. ISBN 9781509042401.
- Silva, P. C. L., Severiano Jr., C. A., Alves, M. A., Cohen, M. W., e Guimarães, F. G. (2019a). A new granular approach for multivariate forecasting. In Cota, V. R., Barone, D. A. C., Dias, D. R. C., e Damázio, L. C. M., editors, *Computational Neuroscience*, p. 41–58, Cham. Springer International Publishing. ISBN 978-3-030-36636-0.
- Silva, P. C. d. L., e Lucas, P. d. O., e Guimarães, F. G. (2019b). A distributed algorithm for scalable fuzzy time series. In *International Conference on Green, Pervasive, and Cloud Computing*, p. 42–56. Springer.
- Silva, P. C. d. L., Severiano Jr, C. A., Alves, M. A., Cohen, M. W., e Guimarães, F. G. (2019c). A new granular approach for multivariate forecasting. In *Latin American Workshop on Computational Neuroscience*, p. 41–58. Springer.
- Silva, P. C., Sadaei, H. J., Ballini, R., e Guimarães, F. G. (2019d). Probabilistic forecasting with fuzzy time series. *IEEE Transactions on Fuzzy Systems*.
- Silva, P. C. d. L. (2019). *Scalable Models For Probabilistic Forecasting With Fuzzy Time Series*. PhD thesis, Universidade Federal de Minas Gerais. URL <https://doi.org/10.5281/zenodo.3374641>.
- Silva, P. C. d. L., Alves, M. A., Severiano Jr, C. A., Vieira, G. L., Guimarães, F. G., e Sadaei, H. J. (2017). Probabilistic Forecasting with Seasonal Ensemble Fuzzy Time-Series. In *XIII Brazilian Congress on Computational Intelligence*, Rio de Janeiro.

- Song, Q. e Chissom, B. S. (1993a). Forecasting enrollments with fuzzy time series—part i. *Fuzzy sets and systems*, 54(1):1–9.
- Song, Q. e Chissom, B. S. (1993b). Fuzzy time series and its models. *Fuzzy Sets and Systems*, 54(3):269–277.
- Song, Q. e Chissom, B. S. (1994). Forecasting enrollments with fuzzy time series—part ii. *Fuzzy sets and systems*, 62(1):1–8.
- Tawegoum, R., Leroy, F., Sintès, G., e Chassériaux, G. (2015). Forecasting hourly evapotranspiration for triggering irrigation in nurseries. *Biosystems Engineering*, 129:237–247. URL <https://doi.org/10.1016/j.biosystemseng.2014.10.011>.
- Winkler, R. L. (1972). A decision-theoretic approach to interval estimation. *Journal of the American Statistical Association*, 67(337):187–191.
- Yu, H.-K. (2005). Weighted fuzzy time series models for TAIEX forecasting. *Physica A: Statistical Mechanics and its Applications*, 349(3):609–624. ISSN 03784371.
- Zadeh, L. A. (1996). Fuzzy sets and information granularity. In *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*, p. 433–448. ISBN 981-02-2422-2.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353.
- Zarco-Tejada, P. J., Hubbard, N., e Loudjani, P. (2014). Precision agriculture: An opportunity for eu farmers—potential support with the cap 2014-2020. *Joint Research Centre (JRC) of the European Commission*.
- Zhao, T., Wang, Q. J., Schepen, A., e Griffiths, M. (2019). Ensemble forecasting of monthly and seasonal reference crop evapotranspiration based on global climate model outputs. *Agricultural and forest meteorology*, 264:114–124.