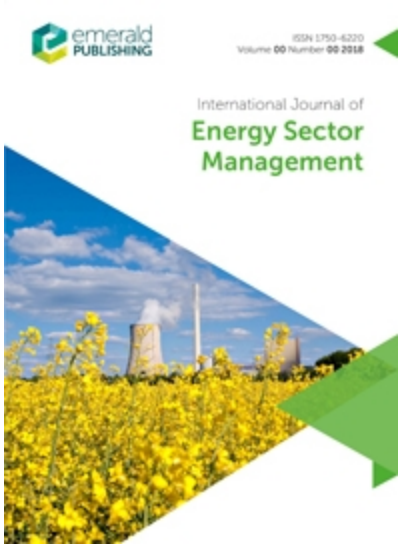

FORECASTING THE BRAZILIAN LIGHT FUEL CONSUMPTION USING TIME SERIES AND MACHINE LEARNING TECHNIQUES

Journal:	*International Journal of Energy Sector Management*
Manuscript ID	Draft
Manuscript Type:	Research Papers
Subject Area:	Demand-side management, Demand forecasting, Forecasting, Gasoline, Liquid fuels, Biofuels
Methods/ Tools / Analytical Frameworks:	Autoregressive, Co-integration, Econometric, Fuzzy-logic model, Neural networks, Time series analysis

FORECASTING THE BRAZILIAN LIGHT FUEL CONSUMPTION USING TIME SERIES AND MACHINE LEARNING TECHNIQUES

Abstract

Purpose: Fuel demand forecast is a fundamental tool to guide private planning actions and public policies aim to guarantee energy supply. In this paper, different forecasting methods were evaluated to project the consumption of light fuels in Brazil (fuels used by vehicles with an internal combustion engine).

Design: Eight different methods were implemented, besides of ensemble learning technics that combine the different models. The evaluation was carried out based on the forecast error for a forecast horizon of 3, 6 and 12 months.

Findings: The statistical tests performed indicated the superiority of the evaluated models compared to a naive forecasting method. Furthermore, for 12 months forecast, it was found methods that outperform, with statistical significance, the SARIMA method, that is widely used. The results indicate, for all forecast horizon, that is possible to estimate a model whose mean absolute percentage error is less than 3%.

Practical implications: The level of accuracy reached allows the use of these models as tools to assist public and private agents that operate in this market.

Originality: The study seeks to fill a gap in the literature on the Brazilian light fuel market. In addition, the methodological strategy adopted assesses projection models from different areas of knowledge using a robust evaluation procedure.

Keywords: Fuel Demand; Forecasting Methods; Time Series; Machine Learning; Forecast Evaluation

1 INTRODUCTION

Energy regulation and policy aim, among other elements, to ensure predictability of domestic supply. From an economic point of view, this element is critical for better allocation of resources. In the social sphere, greater predictability is highlighted when the society's well-being is impacted by the effects of a supply crisis.

In fact, the management of energy demand has become a mandatory issue for public and private agents working in this area, aiming to plan the need for future resources, promote environmental conservation measures, optimize the use of energy sources, assist on decision making, and ensure regularity of supply (Suganthi and Samuel, 2012).

In this context, the use of quantitative models has become recurrent to improve the forecasts associated with this market and is being consolidated as a fundamental tool to meet the aforementioned objectives.

Among the various sources of energy, liquid fuels used in spark-ignited vehicles (light fuels or Otto cycle [1] fuels) are highlighted, which are mostly used to meet the needs of private transportation around the globe.

In Brazil, as it will be detailed in the following topics, the liquid fuels sector has undergone technological, economic, and regulatory changes that have significantly altered its functioning. Since 2003, the country has witnessed the growth of the flex-fuel vehicle fleet and the possibility of choosing the fuel to be used in these vehicles at each filling. In the regulatory sphere, the beginning of the 2000s was marked by the release of fuel prices and the beginning of free market competition.

Additionally, over the past two decades, the fuel sector has had different pricing strategies for oil products internally, with impacts on the average price of light fuels, and also, significant changes in the economic condition of the country, with recurrent reflections on the dynamics of the Otto cycle fuel consumption.

These changes have made the Brazilian light fuel market a unique case in the world and have substantially increased its complexity. Several factors that are difficult to predict started to influence the formation of fuel prices, the dynamics of domestic production, the behavior of consumption and, consequently, the guarantee of internal supply.

Therefore, the present study seeks to evaluate different methods for making predictions about the consumption of light fuels in Brazil. The analysis recognizes the need to anticipate the behavior of demand, considered as one of the main elements associated with greater predictability to the fuel market.

The study seeks to answer questions such as: what is the best method to forecast the consumption of light fuels in the country? Is the performance of the models satisfactory for use as a decision-making support tool for public and private agents?

To answer these questions, the performance of the most common models in the literature on the subject will be evaluated: Sarima (SAR) and Arima with binary variables for seasonality (AR.Bin), Vector Error Correction Model (VECM), univariate and multivariate Artificial Neural Networks (ANN and MANN), Support Vector Machine (SVM), Hybrid Neural Fuzzy Inference System (HyFIS), Random Forest (RF) and, finally, combined models (Ensemble Learning).

As in other nations, the proposed assessment is justified by offering grounds to expand the guarantee of domestic supply. Specifically for the Brazilian market, the results also provide

useful information for achieving various public policy standards and measures associated with the future behavior of demand for fuels.

Finally, the proposed study seeks to fill a gap in the literature on the Brazilian market. This is because great part of the papers available on the theme did not address consumption forecast. In fact, these studies aimed to understand the effect of prices, income, and other economic variables on the behavior of demand for hydrous ethanol, gasoline, or natural gas (NGV).

2 THE BRAZILIAN LIGHT FUEL MARKET

The Brazilian light fuel market has peculiarities derived from technological, productive, and institutional complexity that affect, on the one hand, consumption decisions and, on the other, the dynamics of production and supply.

Regarding technology, the introduction of flex-fuel vehicles in early 2003 gave the Brazilian consumer the opportunity to decide between ethanol and gasoline or any mixture of both products at each filling. This change affected the dynamics of price formation and, consequently, of the country's consumption of fuels by making the demand for individual fuels more elastic (Freitas and Kaneko, 2011).

In 2020, the light fleet with flex-fuel engines represented 80,1% of automobiles and 41,1% of motorcycles in circulation in the country (UNICA, 2020). Unlike other countries where gasoline is the main fuel and, in the Brazilian market the configuration established by the flex-fuel fleet requires that the demand for light fuels incorporate the aggregate consumption of gasoline, hydrous ethanol and, also, natural gas ou NGV (Rodrigues and Bacchi, 2017).

In the productive sphere, since sugarcane is the raw material for Brazilian ethanol, fuel production is subject to climatic conditions and to the fundamentals observed in the sugar market. International cycles and fluctuations in the sweetener's value affect domestic production decisions, since most sugar plants may, albeit in a restricted way, choose to use this material shared for one or another product (Drabik *et al*., 2014).

In the internal supply of gasoline, Petrobrás [2] has a near-monopoly of the activity, being responsible for more than 90% of the domestic production of the product. This condition allows the oil company to define the domestic prices regardless of market conditions.

In the institutional sphere, from 2008 to 2016 there was constant government intervention in the oil sector aiming at inflation control. The main policies were the freezing of

domestic gasoline prices and the exemption from federal taxes levied on the oil products (Rodrigues and Bacchi, 2016).

After being immersed in an economic crisis, as of 2015 the federal government began to tax fuels again to contain the public deficit. At the same time, Petrobras's pricing policy was changed, and the price of gasoline began to be adjusted periodically after the company underwent major financial losses (Rodrigues and Rodrigues, 2018).

In this context, the dimensioning of domestic consumption of fuels becomes essential for the management of resources both by the public sector (need for inspection, infrastructure for the flow of production, etc.), and by the private sector, to define the increase in production capacity, decision of the plants' production strategy, dimensioning of investments in logistics by distributors, among others.

In fact, this complexity of the national market has stimulated countless studies on the subject. Specifically in the case of demand, most studies conducted aimed at assessing gasoline, hydrous ethanol, or NGV consumption behavior individually in the face of changes in economic variables. This is the case of those elaborated by Cardoso *et al*. (2019), Santos *et al*. (2018), Isabella *et al*. (2017), Gomez and Legey (2015), Barros, Gil-Alana and Wanke (2014), Du and Carriquiry (2013), Salvo and Huse (2013), Santos (2013), Freitas and Kaneko (2013), Pacini and Silveira (2011) and Alves and Bueno (2003), Burnquist and Bacchi (2002), among others.

Among the authors who assessed the total demand for light fuels in Brazil (gasoline, hydrous ethanol, and NGV together) are those developed by Rodrigues and Bacchi (2017), Costa *et al*. (2017), Figueira *et al*. (2014) and Rodrigues and Bacchi (2016). Although the proposed models adopt the same dependent variable of this study, these papers focused on the estimation of demand equations, identifying the price and income elasticities for fuel consumption in the Otto cycle.

Finally, it is worth mentioning the studies that used quantitative methods to forecast fuel consumption in Brazil. Castro and Santiago (2012) and Mattos and Perobelli (2011) used auto-regressive vectors with error correction (VECM) to forecast the Brazilian consumption of each of the fuels individually. Santiago, Mattos and Perobelli (2011) also incorporated input-product techniques and macroeconomic variables into the analysis.

Zanini *et al.* (2001), in their turn, used a hybrid model of neural networks and dynamic regression to forecast gasoline consumption in Brazil. Figueira *et al.* (2010) used classic time series models to forecast ethanol consumption.

In general, the studies carried out for the Brazilian market are limited to the following aspects: i) they are essentially concentrated on the use of traditional time series methods; ii) they did not evaluate models to forecast the consumption of light fuels - most studies aim to verify the fit of the model in the training data and the significance of the elasticities; iii) they did not dynamically test several methods in the same database and with the same error metric.

In fact, despite the small number of studies focused on the fuel consumption forecast in the national market, there are numerous studies proposing the forecast of the demand for energy in the international literature, especially for electricity.

That said, when reviewing these works, the following techniques were most frequently identified: Fuzzy Logic models (Chang *et al.*, 2011; Kucukali and Baris, 2010; Padmakumari *et al.*, 1999; Pai, 2006), Support Vector Machine (Ahmad *et al*., 2014; Che and Wang, 2014; Chen *et al*., 2017; Debnath and Mourshed, 2018; Fan *et al*., 2008; García Nieto *et al*., 2018; Hong, 2009) and Random Forest (Dudek, 2015; Ibarra-Berastegi *et al*., 2015; Lahouar and Ben Hadj Slama, 2015; Kane *et al*., 2014). Although little explored to forecast fuel consumption, the Artificial Neural Network models (Ebrahimpour *et al*., 2011; Kavaklioglu *et al*., 2009; Kandananond, 2011; Khashei and Bijari, 2010; Ringwood *et al*., 2001; Sözen *et al*., 2005) have been recurrently adopted to forecast the demand for different energy sources and other applications.

Other methods applied to forecasting the demand and production of different energy sources can be consulted in Deb *et al*., (2017), Debnath and Mourshed (2018) and Suganthi and Samuel (2012).

3 METHODOLOGY

3.1 Data source and pre-treatment

The database used to represent the total demand for light fuels is composed by the sum of the consumption of gasoline, hydrous ethanol, and NGV, converted into liters of equivalent gasoline.

According to the model proposed by Rodrigues and Bacchi (2017), this conversion is necessary to consider the distinct efficiency of each fuel in the demand composition. Therefore, the consumption of light fuels (C_t) will be given by Eq. (1).

$$C_t = \sum_{i=1}^{3} C_{it} \times Re_i \quad (1)$$

where C_{it} representing the amount consumed of each fuel $i = \{\text{hydrous ethanol, gasoline, NGV}\}$ in month t; Re_i, is the coefficient of energy equivalence for the conversion into equivalent liters of gasoline (0.70 in case of hydrous ethanol, 1.23 for NGV, and 1.00 for gasoline).

The monthly data on gasoline and hydrous ethanol consumption were obtained from information by the Brazilian Agency of Petroleum, Natural Gas, and Biofuels (ANP, 2018). Additionally, the amount of natural gas commercialized in the country was compiled by the Brazilian Association of Piped Gas Distributing Companies (ABEGÁS, 2018).

The consumption series, which stars in January 2002 and finishes in August 2018 [3], can be viewed in Figure 1.

Insert Figure 1

For multivariate forecast models, two other series were forecast together with the Otto cycle consumption. The first consists of the GDP per capita estimated from the monthly GDP released by the Central Bank of Brazil (BACEN, 2018) and the resident population by the Brazilian Institute of Geography and Statistics (IBGE, 2018). The second, in its turn, refers to the average price of light fuels to the final consumer (ANP, 2018).

3.2 Forecast models evaluated

3.2.1 Seasonal Autoregressive Integrated Moving Average Model (Sarima)

Sarima model $(p,d,q)(P,D,Q)$ (SAR) aims at explaining a stationary time series through the past values of the series itself and the model errors (Box *et al*., 2008).

The estimated model can be represented according to Eq. (2).

$$\Delta^D\left(\Delta^{\mathrm{d}} N_t\right) = \phi_0 + \sum_{i=1}^{p} \phi_i \Delta^D\left(\Delta^{\mathrm{d}} N_{t-i}\right) + \sum_{i=1}^{q} \theta_i e_{t-i} + \sum_{j=1}^{P} \Phi_j \Delta^D\left(\Delta^{\mathrm{d}} N_{t-f\times j}\right) + \sum_{j=1}^{Q} \Theta_j e_{t-f\times j} + e_t, t = 1,\ldots,n \quad (2)$$

where Δ^{d} and Δ^D are the difference operator of order d and D respectively, ϕ_0 is the intercept of the stationary series and ϕ_i, θ_i, Φ_j and Θ_j are coefficients that multiply the autoregressive

terms and the moving averages, f is the series frequency that characterizes its seasonality (for a monthly series $f = 12$), and e_t is a normally distributed error with zero mean and variance σ_e^2.

Along with the traditional approach, which uses the series and errors lags to incorporate seasonality effects, this study proposes an alternative approach using binary regressors to control seasonality (AR.Bin). Thus, the model starts being represented by Eq. (3).

$$\Delta^D\left(\Delta^{\mathrm{d}} N_t\right) = \phi_0 + \sum_{i=1}^{p} \phi_i \Delta^D\left(\Delta^{\mathrm{d}} N_{t-i}\right) + \sum_{i=1}^{q} \theta_i e_{t-i} + \sum_{m=1}^{f-1} \beta_m \mathrm{b}_m + e_t \,, t = 1, \quad (3)$$

Where b_m is a binary variable representing one month m and β_m is the coefficient representing the differential effect of this month related to the basal month. The seasonality effect on the basal month is given by ϕ_0, while for any $m = 1,...,f-1$ this effect will be $\phi_0 + \beta_m$.

To determine the number of autoregressive components and moving averages, the autocorrelation and partial autocorrelation functions were evaluated. The order of integration of the series was determined from the KPSS (Kwiatkowski *et al*., 1992) and DF-GLS (Elliott *et al.*, 1996) tests for a unit root. The DP test (Dickey and Pantula, 1987) was implemented to check for the presence of two roots and the presence of seasonal unit root was evaluated according to the procedure described by Osbom *et al*. (1988). The final selection was based on the AIC information criterion.

3.2.2 Autoregressive Vector with Error Correction Mechanism (VECM)

The Autoregressive Vector with Error Correction Mechanism (VECM) consists of an alternative approach to forecast non-stationary time series. The structure starts from an autoregressive vector (VAR) with endogenous variables m with multivariate autoregressive order behavior, as described by Eq. (4).

$$\mathbf{Y}_t = \mathbf{\Phi}_0 + \sum_{i=1}^{p} \mathbf{\Phi}_i \mathbf{Y}_{t-i} + \mathbf{G}\mathbf{Z}_t + \mathbf{e}_t \,, t = 1,...,n \quad (4)$$

Where $\mathbf{Y}_t$ and $\mathbf{e}_t$ are vectors of m endogenous variables and white noise, $\mathbf{Z}_t$ is a vector of g exogenous variables, the vectors of coefficients $\mathbf{\Phi}_0$ and $\mathbf{\Phi}_i$, $i = 1,...,p$ of length m and $\mathbf{G}$ is the matrix of exogenous variable coefficients with dimension $m \times g$.

When the variables are non-stationary, but have a common long-term trajectory, it is still possible to use a VAR by introducing an error correction model.

Once the presence of long-term equilibrium is identified, the variables contained in $\mathbf{Y}_t$ are said to be cointegrated in order (d,b) or $C(d,b)$. Thus, it is possible to estimate up to $m-1$ cointegration vectors. In this case, Eq. (4) can be changed and $\mathbf{Y}_t$ can be expressed by a VECM model according to Eq. (5).

$$\Delta\mathbf{Y}_t = \boldsymbol{\alpha}\mathbf{B}^{\mathrm{T}}\mathbf{Y}_{t-1} + \sum_{j=1}^{p-1}\Pi_j\Delta\mathbf{Y}_{t-j} + \mathbf{G}\mathbf{Z}_t + \mathbf{e}_t \quad (5)$$

Where $\mathbf{B} = [\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...,\boldsymbol{\beta}_r]$ and $\boldsymbol{\alpha}$ are matrices $m \times r$ such that $r = posto(\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}})$, $r < m$ is the number of non-zero cointegration vectors and, consequently, error correction vectors. The latter are expressed by $\mathbf{B}^{\mathrm{T}}\mathbf{Y}_{t-1} = \mathbf{u}_{t-1} = [u_{t-1}^1,u_{t-1}^2,...,u_{t-1}^r]$ and $\Pi_j = -\sum_{i=1+j}^{p-1}\boldsymbol{\Phi}_i$, $j = 1,...,p-1$ are vectors of coefficients that multiply $\Delta\mathbf{Y}_{t-j}$.

The cointegration test was conducted as proposed by Johansen (1988, 1991). The selection of the number of autoregressive terms was based on the AIC information criterion. In the verification process, the Portmanteau (Lütkepohl, 2007) and Edgerton and Shukur (1999) autocorrelation tests, the Jarque-Bera (Lütkepohl, 2007) and Shapiro-Wilk (Shapiro and Wilk, 1965) multivariate normality test and the residual multivariate heteroscedasticity test ARCH-LM (Engle, 1982) have been implemented.

3.2.3 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) can be represented by a network of oriented links that are interconnected at certain points called nodes or neurons (Haykin, 2001).

Algebraically, in an ANN where all neurons of a layer are interconnected with all neurons of the next layer (referred as totally connected ANN), the output of a neuron k of layer l, represented by f_k^l, is given by the application of an activation function $g(.)$ on the linear combination of the outputs of neurons belonging to layer $l-1$, i.e, $f_k^l = g\left(\sum_{k' \in l-1}\omega_{k'}^l f_{k'}^{l-1}\right)$, where $\omega_{k'}^l$ is the vector of weights that weigh the synaptic links of neurons k' belonging to layer $l-1$ to neuron k of layer l.

Let the parameters vector of the model given by $\boldsymbol{\omega} = \left\{\left\{\omega_{k'}^l\right\}_{k' \in l-1}\right\}_{l=2}^{L}$, $\{\mathbf{Y}_t,\mathbf{x}_t\}_{t=1}^{n}$, a set of known data where $\mathbf{Y}_t$ is vector m-varied of the optimal response to a vector p–varied of input $\mathbf{x}_t$. The estimated value of $\boldsymbol{\omega}$ is given by Eq. (6).

$$\boldsymbol{\omega} = \underset{\boldsymbol{\omega}}{\operatorname{argmin}}\left\{\frac{1}{m}\sum_{j=1}^{m}\sum_{t=1}^{n}\left(Y_{t,j} - \hat{f}_j^L(\boldsymbol{\omega},\mathbf{x}_t)\right)^2\right\} \quad (6)$$

Where $\hat{f}^L(\boldsymbol{\omega},\mathbf{x}_t) = \left\{\hat{f}_j^L(\boldsymbol{\omega},\mathbf{x}_t)\right\}_{j=1}^{m}$ are nonlinear mappings of $\mathbf{x}_t \rightarrow \mathbf{Y}_t$ resulting in layer L, called the ANN output layer. In the present study $L = 3$ and the number of neurons in the output layer is equal to the number of endogenous variables (m) in the model.

To optimize the function, the BFGS algorithm was used because it presents better performance in smaller dimension problems (Haykin, 2001).

Two ANN structures were estimated. The first is a univariate model $(m = 1)$ with Otto cycle fuel consumption as the output variable and p autoregressive terms as input variables. The second is a multivariate model $(m = 3)$ that jointly forecasts the Otto cycle consumption, the average fuel price and GDP per capita through p autoregressive terms for each of these variables as an input vector.

3.2.4 Support Vector Machines (SVM)

Support Vector Machine (SVM) is a non-parametric model capable of estimating nonlinear relations from an input vector to an output variable, using the so-called kernel trick (Hastie *et al.*, 2009).

It consists of applying a nonlinear transformation $\left\{\left\{\varphi_j(\mathbf{x}_t)\right\}_{j=1}^{m_1}\right\}_{t=1}^{n}$ on the input variables $\left\{\left\{\mathrm{x}_{i,t}\right\}_{i=1}^{m_0}\right\}_{t=1}^{n}$ so as the dimension increases from m_0 to m_1. The objective of the method is to transform a nonlinear relation of $\mathbf{x}_t \rightarrow Y_t$ in a linear relation of $\boldsymbol{\varphi}(\mathbf{x}_t) \rightarrow Y_t$.

It is necessary to estimate the vector of weights $\boldsymbol{\omega}$ that allows the construction of a hyperplane $L\left(Y_t,\hat{f}(\boldsymbol{\omega},\mathbf{x}_t)\right) = 0$, so as $\hat{f}(\boldsymbol{\omega},\mathbf{x}_t) = \boldsymbol{\omega}^T\boldsymbol{\varphi}(\mathbf{x}_t) = \sum_{j=0}^{m_1} w_j\varphi_j(\mathbf{x}_t)$. The estimate of coefficients vector takes place indirectly, from the resolution of the problem of optimization given by Eq. (7).

$$\max_{\boldsymbol{\alpha},\boldsymbol{\alpha}'} Q(\boldsymbol{\alpha},\boldsymbol{\alpha}') = \sum_{t=1}^{n} Y_t(\alpha_t - \alpha_t') - \epsilon\sum_{t=1}^{n}(\alpha_t - \alpha_t') - \frac{1}{2}\sum_{t=1}^{n}\sum_{j=1}^{n}(\alpha_t - \alpha_t')(\alpha_j - \alpha_j')K(\mathbf{x}_t,\mathbf{x}$$

$$s.t$$

$$\sum_{t=1}^{n}(\alpha_t - \alpha_t') = 0 \qquad (7)$$

$$0 \le \alpha_t \le C\,,\ i = 1,2,\ldots,n$$

$$0 \le \alpha_t' \le C\,,\ i = 1,2,\ldots,n$$

In Eq. (7) $K(\mathbf{x}_t,\mathbf{x}_j) = \boldsymbol{\varphi}^{\boldsymbol{T}}(\mathbf{x}_t)\boldsymbol{\varphi}(\mathbf{x}_j) = \sum_{j=0}^{m_1}\varphi_j(\mathbf{x}_t)\varphi_j(\mathbf{x}_j)$ is called kernel function. In practice, to estimate an SVM it is not necessary to know the function $\boldsymbol{\varphi}(\mathbf{x})$, only the kernel function and that it is in accordance with Mercer's theorem (Haykin, 2001). All input vectors $\mathbf{x}_t$ whose $(\alpha_t,\alpha_t') \neq 0$ are called support vectors. In fact, these observations are the ones that define the estimated hyperplane since the optimal weight vector will be given by $\boldsymbol{\omega} = \sum_{t=1}^{n}(\alpha_t - \alpha_t')\boldsymbol{\varphi}(\mathbf{x}_t)$ and the forecast value for an input $\mathbf{x}_h$ will be $Y_h = \sum_{t=1}^{n}(\alpha_t - \alpha_t')K(\mathbf{x}_h,\mathbf{x}_t)$

.

3.2.5 Hybrid Neural Fuzzy Inference system (HyFIS)

The fuzzy controls are ways to establish relations among linguistic variables from modeling of rules connecting a condition (If) and a consequence (Then). The systems mixing the fuzzy systems and the neural networks paradigms can be called "neural *fuzzy*" (Jang *et al.*, 1997).

This study implemented the method Hybrid Neural Fuzzy Inference System - HyFIS (Kim and Kasabov, 1999). The step of rules learning was performed using the procedure proposed by Wang and Mendel (1992).

3.2.6 Random Forest

Random Forest model uses decision trees to make forecasts of y_t in function of an input vector $\mathbf{x}_t$. In this case, several regression trees are estimated, and the final forecast of the algorithm is given by the average of trees forecasts (Breiman, 2001; Hastie *et al.*, 2009).

Assuming that a regression tree set $\{T_b\}_{b=1}^{B}$ will be estimated to map the relation $\mathbf{x}\rightarrow Y$, then the algorithm performs the following steps:

1. Obtaining a subsample randomically of size n_b from the training sample.
2. Generating a regression tree T_b with the simulated sample and, using the stop criterion that each terminal node reaches the minimum size n_{min}^{b}, or the variation of the training sample residual sum of squares is below a tolerance limit:
3. The remaining observations on the terminal nodes are used in the forecast of an individually tree given by $T(\mathbf{x},\Theta_b)$. Where Θ_b represents the variables selected, the optimal division points and the observations in each terminal node.
4. The forecast value for an input vector $\mathbf{x}$ is given by $\hat{f} = \frac{1}{B}\sum_{b=1}^{B}T(\mathbf{x},\Theta_b)$.

Knowing that there is a set of S possible divisions for the nodes of a tree T_b, the selection of the optimal node break point was made according to the algorithm CART for decision trees (Breiman *et al*., 1984).

3.2.7 Ensemble Learning Models

The ensemble learning models aim at combining the results of different algorithms to reduce forecast error variance and to increase its accuracy. The most usual form is based on the linear combination of forecasts according to Eq. (8) and (9).

$$\hat{\mathbf{y}} = \hat{\mathbf{Y}}\mathbf{W} \tag{8}$$

$$\sum_{j}^{J} w_j = 1 \tag{9}$$

where $\hat{\mathbf{y}}$ is a combined forecast vector with dimension $T \times 1$, $\hat{\mathbf{Y}}$ is a matrix $T \times J$ containing the forecasts of the J different algorithms for T periods and $\mathbf{W}$ is the dimension vector $J \times 1$ which contains the weights w_j for the linear combination of forecasts.

The study results proposing a combination of forecasts have been unanimous, favoring the improvement of forecast accuracy (Hyndman and Athanasopoulos, 2018). Two forms of combination of predictions were evaluated from an empirical point of view. The first estimates the weights by means of restricted least squares (RLS) imposing that the sum of the weights is equal to one. In the second method, a simple average (AVG) of the predictions of the different models is taken as the final projection of the data.

3.3 Determination of hyperparameters and evaluation metrics

3.3.1 Determination of hyperparameters of the models

Table I shows the hyperparameters of each estimated model.

Insert Table I

The determination of the set of hyperparameters was obtained by minimizing the estimate of the generalization error or forecast error (Hastie *et al.*, 2009).

For linear models, a simple way to assess the generalization capacity is to compare the information criteria of Akaike or AIC (Hastie *et al.*, 2009; Hyndman and Athanasopoulos,

2018) after estimating the models with all available data. These are the cases of the Sarima and VECM methods.

For non-linear models, the cross-validation process was used. Cross-validation involving time series is different from problems involving panel data due to the series autoregressive dynamics (Bergmeir *et al.*, 2018). In the present study the procedure was structured as follows:

1. A set of parameters Θ_i is initially selected for the model.
2. An initial cut is performed in the series of consumption of the Otto cycle, with only the data from January 2002 to December 2011 remaining, that is, 120 observations of a total of 200. The selection of this cut point is explained by the change of trend in the series that was observed from 2013 on.
3. This subseries is used to estimate the models and obtained the h step ahead forecasts ($\hat{Y}$), and the forecast error vector is calculated and stored.
4. A new observation is added to the training base and step 3 is repeated.
5. Steps 3 and 4 are repeated until the h-step ahead forecast has as its last forecast period the last observation known of the series. The root mean square error (RMSE) is calculated for the model with parameters Θ_i according to Eq. (10).

$$RMSE_i = \sqrt{\frac{1}{(81-h)h}\sum_{j=1}^{(81-h)h}\left(Y_j - \hat{Y}_j\right)^2} \quad (10)$$

Obtaining the forecast errors of the subseries, it was possible to estimate RMSE as metrics of the model generalization error, and to select the set of hyperparameter with better performance.

3.3.2 Evaluation metrics and statistical tests for comparing algorithms

After defining the hyperparameters, one can proceed to compare the different forecasting methods. An interactive process was conducted to generate a robust estimate of the accuracy of the models. The following steps were performed:

1. An initial cut is made in the Otto cycle consumption series leaving only the data from January 2002 to December 2011.
2. This subseries is used to estimate the models and carry out the h-step ahead forecast. The RMSE is calculated, according to Eq. (13), and stored as a measure of accuracy in

the period. In this case, each cutoff period has an RMSE value for the h-step ahead forecast.

3. A new observation is added to the training base and step 3 is repeated.

Step 3 is repeated until the h-step ahead forecast has the last known observation of the data as the last predicted period. Thus, a series with the RMSE is obtained for each point in time and for each model. Figure 2 illustrates the stages of the interactive process based on a 3-month horizon.

Insert Figure 2

In this study, two nonparametric tests were implemented (Shingala, 2016). Friedman's test was performed to check for a difference between the average ranking of an algorithm i, given by $\bar{r}_i$, and the average performance of all models $\bar{r}$. The test hypotheses were shown on Eq. (11).

$$\text{Friedman} \quad \begin{matrix} \text{H0}: \bar{r}_i = \bar{r}, \text{ there is no difference in the ranking of methods in general.} \\ \text{HA}: \bar{r}_i \neq \bar{r}, \text{ there is difference in the ranking of methods in general.} \end{matrix} \tag{11}$$

If Friedman's test rejects H0, Nemenyi test can be performed to compare the pairs of methods. In this case, these are compared pair-to-pair through its average ranking. Eq. (12) shows the null hypotheses and test alternatives.

$$\text{Nemenyi} \quad \begin{matrix} \text{H0}: \bar{r}_i = \bar{r}_j \; \forall \; i \neq j, \text{ there is no difference in the ranking of methods } i \text{ and } j. \\ \text{HA}: \bar{r}_i \neq \bar{r}_j \; \forall \; i \neq j, \text{ there is no difference in the ranking of methods } i \text{ and } j. \end{matrix} \tag{12}$$

The execution of multiple tests for pairwise comparison of multiple methods leads to the inflation of type I error. To avoid this issue, the Bonferroni control method (Shingala, 2016; Demsar, 2006) was performed.

A naive forecasting method was introduced between the models to verify the superiority of the methods in relation to a model with low cost of implementation. In this case, the months of the last year will be used as a forecast for a subsequent year.

The RMSE was used as a criterion for ranking the models. However, the Mean Absolute Square Error (MAPE) will be displayed for each point in time (that is, for the h-step ahead

forecast from that date), calculated according to Eq. (14), in order to facilitate the interpretation of results by excluding the dimensional issue in errors.

Additionally, the maximum Percentage Error (PE) considering the h step ahead forecast horizon is also presented (Eq. 15).

$$RMSE_t = \sqrt{\frac{1}{h}\sum_{j=1}^{h}\left(Y_j - \hat{Y}_j\right)^2} \quad (13)$$

$$MAPE_t = \frac{100\%}{h}\sum_{j=1}^{h}\frac{\left|Y_j - \hat{Y}_j\right|}{Y_j} \quad (14)$$

$$\max PE_t = \max\left\{\frac{\left|Y_j - \hat{Y}_j\right|}{Y_j}\right\}_{j=1}^{h} \quad (15)$$

4 RESULTS AND DISCUSSION

4.1 Cross-validation results

Table II shows the results of the selection of hyperparameters. Models whose parameters are chosen according to the forecast horizon show a break for 3, 6 and 12 months.

Insert Table II

The use of non-seasonally adjusted series in the ANN, Multi ANN, SVM, HyFIS and Random Forest models comprises the alternative with the smaller forecast error.

The additional diagnostic tests performed for the Sarima and VECM models indicate good adherence of the models to the data, except for the normality tests of the residues. The Arima model with binary variables, in its turn, presented problems of non-normality and heteroscedasticity.

Despite the results of the normality tests, it is worth noting that they are highly sensitive to the presence of outliers and, in such cases, tend to reject the hypothesis of normally distributed data with ease (Coin, 2008). Thus, these approaches were also considered in the model comparison procedures.

4.2 Accuracy evaluation and model comparison

At first, the comparison with the combined models was disregarded in order to choose the algorithms with superior individual performances.

Table III shows the average rankings for each method and the result of the Friedman test.

Insert Table III

The results of the Nemenyi test, in its turn, are presented in Figures 3 to 5. The best algorithm identified by this procedure was SVM, followed by AR.Bin and VECM for the 12- and 6-month forecast. Only in the case of the 3-month horizon did these models exceed SVM's performance.

However, it is not possible to affirm that the first 6 best ranked models have a statistically lower performance than SVM in the 12-month forecast. The same is valid for the forecast of 6 and 3 months ahead. The confidence interval of the test is represented in Figures 3 to 5 by the gray area.

Insert Figures 3 to 5

The performances of the best-ranked methods were statistically superior to the naive model. This fact suggests greater robustness of these procedures, in contrast to the higher computational cost necessary to implement them.

Tables IV to VI show the mean and standard deviation of the MAPE of the forecast, the MAPE average of the training sample, the maximum and minimum MAPE value and the maximum percentage error of each model considering all periods used in the validation.

Insert Table IV to VI

Tables IV to VI end up reflecting the results obtained in the Nemenyi test. SVM, AR.Bin and VECM lead with the best accuracy.

Except for MANN for 12 months forecast, apparently no method has an over-fitting problem, since the training and validation errors do not differ significantly.

The longer forecast horizon implies an increase in the average MAPE of the models, which have greater accuracy in the 3-month forecast, followed by the 6- and then 12-month forecast.

In addition, this increase in the forecast horizon does not lead to generalized degradation of the forecasting capacity of the methods and, on the other hand, allows to discriminate them more precisely. Consequently, there is a greater heterogeneity of rankings and MAPE in the 12-month forecast. The SAR model is no longer part of the significant region of Nemenyi test.

Therefore, there is a strong indication that the SVM showed to be superior to the Sarima method for the data used.

In order to assess the stability of the forecast error, one can analyze the MAPE standard deviation and the maximum percentage error within a point in time. SVM, whose performance ranking was the best among all methods in the 12- and 6-month forecast, also obtained the smallest standard deviation of MAPE considering all horizons. Considering the longer forecast horizon, the maximum percentage error of this model was 11.6% and the maximum MAPE was 5.7%. On the other hand, the AR.Bin model, whose ranking was the second best, has the fifth largest standard deviation, maximum percentage error of 16.2% and maximum MAPE of 8.4%.

After analyzing the individual methods, the joint learning models were implemented considering the 3 methods with the lowest MAPE in each forecast horizon, these being SVM, AR.Bin and VECM.

4.3 Results of the Ensemble Learning Models

The Friedman and Nemenyi tests were reevaluated including the combined methods. Table VII presents the average ranking of the models and Figures 6 to 8 the result of the Nemenyi test.

Insert Table VII

Insert Figures 6 to 8

Tables VIII to X show the metrics previously mentioned, in the individual models, for the combined models.

Insert Table VIII to X

Although the combined models have lower forecasting errors, the performed tests do not reject the null hypothesis that their accuracy is as good as that generated by the individual methods. Also, the combination using the mean of the forecasts generates a MAPE as good as the combination based on the inference of the weights using the least squares method.

4.4 General evaluation of results and considerations on the application of the models

In general terms, model forecasts for three months of consumption showed an average MAPE below 2.3%, with maximum variations (Max MAPE) below 6%. These values indicate that the structure developed can be useful for commercial and logistical planning carried out routinely by agents of the private sector, in the dimensioning of their operations in the short term. In the case of fuel supply contracts, for example, the negotiations cover permitted variations of up to 5% in the withdrawal of the volume initially contracted.

Short-term projections can be used by the public sector in the process of monitoring domestic fuel supply. In the case of the Otto cycle, this activity is carried out periodically by the Supply Monitoring Committee, coordinated by the Ministry of Mines and Energy. This committee was created by Resolution No. 14/2017 of the National Energy Policy Council (CNPE, 2017) and, among other aspects, aims to monitor the conditions of supply and demand for light fuels, in addition to discussing strategies to ensure adequate supply of the light fuels.

In the case of annual forecast, the results obtained show options for models with an average MAPE below 3%. These parameters also reveal the possibility of using the evaluated models in the coordination and planning of agents operating in this sector.

To cite an example, the projection of fuel consumption of the Otto cycle for 12 months is one of the parameters to be evaluated by agents in the process of contracting anhydrous ethanol. This is because ANP Resolution No. 67/2011 (ANP, 2011) established incentives for the advance contracting of the supply of this biofuel between producers and distributors. Currently, the minimum volume to be contracted is defined based on the consumption of gasoline observed in the previous year.

Still in the regulatory sphere, the approval of Law No. 13,576 of December 2017 and the institution of the National Biofuels Policy in Brazil (RenovaBio [4]) started to require regular forecasts for the definition of the ten-year goals of reducing the carbon intensity of the Brazilian fuel matrix (CRBIO, 2020; CNPE, 2020). This definition is annually approved after a wide public consultation process. Among the parameters used in that definition, there is the consumption of light fuels in the country. Thus, the tools developed here can also be useful for the calibration of the decarbonization goals mentioned above.

5 CONCLUSIONS

The present study compared different methods for predicting the consumption of Otto cycle fuels in Brazil. The cross-validation procedure allowed estimate the generalization error of the models and the average MAPE obtained was less than 3% in all forecast horizons.

Additionally, the methods with the best performance proved to be significantly superior to the naive forecasting model.

In addition, for the 12-month forecast, the SVM and the combined RLS method were found to be statistically superior to Sarima. This fact is relevant given the wide use of this method, which makes it an interesting benchmarking for forecasting.

Even in periods of great uncertainty, the models' performance proved to be stable, with model options whose maximum monthly error remained at 11.61% for the 12-month forecast and was less than 10% for the 6- and 3-month periods.

Thus, it is possible to conclude that, despite the technological, market, and institutional changes observed in the Brazilian light fuel sector in the last 10 years, the forecast models evaluated indicate that it is possible to make forecasts for the consumption of light fuels with a satisfactory level of accuracy.

Total consumption forecast for light fuels is useful for private agents in the definition of operational and investment actions related to the production, distribution, transport, and resale of fuels in the country.

In the case of the public sector, the models implemented can be used in the dimensioning of resources to adjust domestic production capacity, as well as the logistical infrastructure necessary to supply the demand for transport. In addition, in the Brazilian case, consumption forecast is critical for the more efficient operation of marketing rules and established public policies.

Finally, the analysis conducted here does not cover all the need for further studies related to the topic. Models for carrying out regional forecasts and forecasting the individual demand for each fuel are important for a deeper understanding of the topic. Finally, studies to assess the structural impact of the recent COVID-19 pandemic on fuel consumption in the future are essential for a thorough analysis on the topic.

Notes

1. The term "Otto cycle" refers to the thermodynamic cycle associated with the operation of internal combustion engines with spark ignition.

2. The company Petróleo Brasileiro SA (Petrobrás) is a Brazilian mixed capital company whose majority shareholder is the federal government.

3. The analysis period was limited to the availability of data up to the research conclusion.

4. The program aims to establish emission targets for fuel distributors in view of the level of pollution related to each fuel.

6 REFERENCES

ABEGÁS – Brazilian Association of Pipeline Gas Distribution Companies. (2018), "Consumption Statistics", available at: https://www.abegas.org.br/estatisticas-de-consumo (accessed 19 September 2018).

Ahmad, A.S., Hassan, M.Y., Abdullah, M.P., Rahman, H.A., Hussin, F., Abdullah, H. and Saidur, R. (2014), "A review on applications of ANN and SVM for building electrical energy consumption forecasting", *Renewable and Sustainable Energy Reviews*. Vol.33, pp.102–109. doi.org/10.1016/j.rser.2014.01.069

Alves, D.C.O. and Bueno, R.L.S. (2003), "Short-run, long-run and cross elasticities of gasoline demand in Brazil", *Energy Economics*, Vol.25 No.2, pp.191–199. doi.org/10.1016/S0140-9883(02)00108-1

ANP – National Agency of Petroleum, Natural Gas and Biofuels. (2011), "Resolution No. 67, of 9.12.2011", available at: http://legislacao.anp.gov.br/?path=legislacao-anp/resol-anp/2011/dezembro&item=ranp-67-2011 (accessed 31 October 2020).

ANP – National Agency of Petroleum, Natural Gas and Biofuels. (2018), "Statistical data", available at: https://www.gov.br/anp/pt-br/centrais-de-conteudo/dados-estatisticos (accessed 19 September 2018).

BACEN – Central Bank of Brazil. (2018), "Time Series Management System", available at: https://www3.bcb.gov.br/sgspub (accessed 19 September 2018].

Barros, C.P., Gil-Alana, L.A. and Wanke, P. (2014), "Ethanol consumption in Brazil: Empirical facts based on persistence, seasonality and breaks", *Biomass and Bioenergy*, Vol.63, pp.313–320. doi.org/10.1016/j.biombioe.2014.02.012

Bergmeir, C., Hyndman, R.J. and Koo, B. (2018), "A note on the validity of cross-validation for evaluating autoregressive time series prediction", *Computational Statistics & Data Analysis*, Vol.120, pp.70–83. doi.org/10.1016/j.csda.2017.11.003

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (2008), *Time Series Analysis: Forecasting and Control*. 4th ed., John Wiley & Sons, New Jersey.

Breiman, L., Friedman, J., Stone, C.J. and Olshen, R.A. (1984), *Classification and Regression Trees*, CRC Press, New York.

Breiman, L. (2001), "Random Forests", *Machine Learning*, Vol.45, pp.5–32.

Bueno, R.L.S. (2018), "Econometria de Séries Temporais", 2nd ed., Cengage Learning, São Paulo.

Burnquist, H.L. and Bacchi, M.R.P. (2002), "A demanda por gasolina no Brasil: uma análise utilizando técnicas de co-integração", paper presented at XL Congresso Brasileiro de Economia e Sociologia Rural, Passo Fundo, Rio Grande do Sul.

Cardoso, L.C.B., Bittencourt, M.V.L., Litt, W.H. and Irwin, E.G. (2019), "Biofuels policies and fuel demand elasticities in Brazil", *Energy Policy*, Vol.128, pp.296–305. doi.org/10.1016/j.enpol.2018.12.035

Castro, G.R. (2012), "Modelo de Previsão de Demanda por Combustíveis Automotivos no Brasil". Graduation conclusion work. *Universidade Federal do Rio de Janeiro*. http://hdl.handle.net/11422/1867

Chang, P.C., Fan, C.Y. and Lin, J.J. (2011), "Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach", *International Journal of Electrical Power and Energy Systems*, Vol.33, pp.17–27. doi.org/10.1016/j.ijepes.2010.08.008

Che, J. and Wang, J. (2014), "Short-term load forecasting using a kernel-based support vector regression combination model", *Applied Energy*, Vol.132, pp.602–609. doi.org/10.1016/j.apenergy.2014.07.064

Chen, Y., Xu, P., Chu, Y., Li, W., Wu, Y., Ni, L., Bao, Y. and Wang, K. (2017), "Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings", *Applied Energy*, Vol.195, pp.659–670. doi.org/10.1016/j.apenergy.2017.03.034

CNPE – National Energy Policy Council. (2017), "Resolution No. 17, of 8.6.2017", available at: http://legislacao.anp.gov.br/?path=legislacao-federal/resolucoes/resol-cnpe/2017 (accessed 19 August 2020).

CNPE – National Energy Policy Council. (2020), "Resolution No. 8, of 18.8.2020", available at: http://www.mme.gov.br/documents (accessed 19 August 2020).

Coin, D. (2008), "Testing normality in the presence of outliers. Statistical Methods and Applications", Vol.17, pp.3–12. doi.org/10.1007/s10260-007-0046-8

Costa, C.C., Burnquist, H.L., Souza, M.J.P. and Valdez, C. (2017), "Modeling fuel demand in Brazil: the importance of the flex-fuel vehicle", *Revista Econômica do Nordeste*, Vol.48 No.3, pp. 113-127.

CRBIO – RenovaBio Commitment. (2020), "Public Consultation No. 94 of June 5, 2020", available at: http://www.mme.gov.br/web/guest/conselhos-e-comites/crbio (accessed 19 August 2020).

Deb, C., Zhang, F., Yang, J., Lee, S.E. and Shah, K.W. (2017), "A review on time series forecasting techniques for building energy consumption", *Renewable and Sustainable Energy Reviews*, Vol.74, pp.902–924. doi.org/10.1016/j.rser.2017.02.085

Debnath, K.B. and Mourshed, M. (2018), "Forecasting methods in energy planning models", *Renewable and Sustainable Energy Reviews*, Vol.88, pp.297–325. doi.org/10.1016/j.rser.2018.02.002

Demsar, J. (2006), "Statistical Comparisons of Classifiers over Multiple Data Sets", *Journal of Machine Learning Research,* Vol.7, pp.1–30.

Dickey, D.A. and Pantula, S.G. (1987), "Determining the Order of Differencing in Autoregressive Processes", *Journal of Business & Economic Statistics*, Vol.5 No.4, pp.455–461. doi.org/10.1080/07350015.1987.10509614

Drabik, D., Gorter, H., Just, D.R. and Timilsina, GR. (2015), "The economics of Brazil's ethanol-sugar markets, mandates, and tax exemptions", *American Journal of Agricultural Economics,* Vol.97 No.5, pp.1433–1450. doi.org/10.1093/ajae/aau109

Du, X. and Carriquiry, M.A. (2013), "Flex-fuel vehicle adoption and dynamics of ethanol prices: Lessons from Brazil", *Energy Policy*, Vol.59, pp.507–512. doi.org/10.1016/j.enpol.2013.04.008

Dudek, G. (2015), "Short-Term Load Forecasting using Random Forests. Intelligent Systems'2014", Filev, D. *et al*. (eds), *Advances in Intelligent Systems and Computing*, Springer, Cham, pp.821–828. doi.org/10.1007/978-3-319-11310-4_71

Ebrahimpour, R., Nikoo, H., Masoudnia, S., Yousefi. M.R. and Ghaemi, M.S. (2011), "Mixture of mlp-experts for trend forecasting of time series: A case study of the tehran stock exchange", *International Journal of Forecasting*, Vol.27 No.3, pp.804–816. doi.org/10.1016/j.ijforecast.2010.02.015

Edgerton, D. and Shukur, G. (1999), "Testing autocorrelation in a system perspective testing autocorrelation", *Econometric Reviews*, Vol.18 No.4, pp.343–386. doi.org/10.1080/07474939908800351

Elliott, G., Rothenberg, T.J. and Stock, J.H. (1996), "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, Vol.64 No.4, pp.813–836. doi.org/10.2307/2171846

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, Vol.50 No.4, pp.987–1007. doi.org/10.2307/1912773

Fan, S., Chen, L. and Lee, W.J. (2008), "Machine learning based switching model for electricity load forecasting", *Energy Conversion and Management*, Vol.49 No.6, pp.1331–1344. doi.org/10.1016/j.enconman.2008.01.008

Figueira, S.R., Burnquist, H.L. and Bacchi, M.R.P. (2010), "Forecasting fuel ethanol consumption in Brazil by time series models: 2006–2012", *Applied Economics*, Vol.42 No.7, pp.865-874. doi.org/10.1080/00036840701720978

Figueira, S.R.F., Santos, D.F.L, Borges, A.C.G., Sanches, A. and Campos, A.L.S. (2014), "Demand for Automotive Fuel in Brazil: An Empirical Analysis Using Cointegration Techniques", *Business and Management Review*, Vol. 4 No.3, pp.343-356.

Freitas, L.C. and Kaneko, S. (2011), "Ethanol demand under the flex-fuel technology regime in Brazil", *Energy Economics*, Vol.33 No.6, pp.1146–1154. doi.org/10.1016/j.eneco.2011.03.011

García Nieto, P.J., Lasheras, F.S., García-Gonzalo, E. and Cos Juez, F.J. (2018), "PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study", *Science of the Total Environment*, Vol.621, pp.753–761. doi.org/10.1016/j.scitotenv.2017.11.291

Gomez, J.M.A. and Legey, L.F.L. (2015), "An analysis of the impact of Flex-Fuel vehicles on fuel consumption inBrazil, applying Cointegration and the Kalman Filter", *Energy*, Vol.81, pp.696–705. doi.org/10.1016/j.energy.2015.01.015

Hastie, T., Tibshirani, R. and Frieman, J. (2009), *The Elements of Statistical Learning*, 2nd ed., Springer, New York.

Haykin, S. (2001), *Redes Neurais: Princípios e Prática*, 2nd ed., Bookman, Porto Alegre.

Hong, W-C. (2009), "Electric load forecasting by support vector model", *Applied Mathematical Modelling*, Vol.33 No.5, pp.2444–2454. doi.org/10.1016/j.apm.2008.07.010

Hyndman, R.J. and Athanasopoulos, G. (2018), *Forecasting: Principles and Practice*, 2nd ed., Otexts, Melbourne.

Ibarra-Berastegi, G., Saénz, J., Esnaola, G., Ezcurra, A. and Ulazia A. (2015), "Short-term forecasting of the wave energy flux: Analogues, random forests, and physics-based models", *Ocean Engineering*, Vol.104, pp.530–539. doi.org/10.1016/j.oceaneng.2015.05.038

IBGE - Brazilian Institute of Geography and Statistics. (2018), "Population Statistics", available at: https://www.ibge.gov.br/estatisticas/sociais/populacao (accessed 19 September 2018).

Isabella, G., Yu, A.S.O., Silva, A.M. and Pegetti, A.L. (2017), "Another driver of the Brazilian fuel ethanol supply chain: the consumers' preferences", *Revista de Administração*, Vol.52 No.3, pp.304–316. doi.org/10.1016/j.rausp.2017.05.003

Jang, J-SR., Sun, C-R. and Mizutani, E. (1997), *Neuro-Fuzzy and Soft Computing: A computational Approach to Learning and Machine Intelligence*, Prentice Hall, New Jersey.

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector", *Econometrica*, Vol.59 No.6. pp.1551–1580. doi.org/10.2307/2938278

Johansen, S. (1988), "Statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, Vol.12 No.2-3, pp.231–254. doi.org/10.1016/0165-1889(88)90041-3

Kandananond, K. (2011), "Forecasting electricity demand in Thailand with an artificial neural network approach", *Energies*, Vol.4 No.8, pp.1246–1257. doi.org/10.3390/en4081246

Kane, M.J., Price, N., Scotch, M. and Rabinowitz, P. (2014), "Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks", *BMC Bioinformatics*, Vol.15. doi.org/10.1186/1471-2105-15-276

Kavaklioglu, K., Ceylan, H., Ozturk, H.K. and Canyurt, O.E. (2009), "Modeling and prediction of Turkey's electricity consumption using Artificial Neural Networks", *Energy Conversion and Management*, Vol.50 No.11, pp;2719–2727. doi.org/10.1016/j.enconman.2009.06.016

Khashei, M. and Bijari, M. (2010), "An artificial neural network (p, d, q) model for timeseries forecasting", *Expert Systems with Applications*, Vol.37, pp.479–489. doi.org/10.1016/j.eswa.2009.05.044

Kim, J. and Kasabov, N. (1999), "HyFIS: adaptive neuro-fuzzy inference systems e their application to nonlinear dynamical systems", *Neural Networks*, Vol.12 No.9, pp.1301–1319. doi.org/10.1016/S0893-6080(99)00067-2

Kucukali, S. and Baris, K. (2010), "Turkey's short-term gross annual electricity demand forecast by fuzzy logic approach", *Energy Policy*, Vol.38 No.5, pp.2438–2445. doi.org/10.1016/j.enpol.2009.12.037

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992), "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?", *Journal of Econometrics*, Vol.54 No.1-3, pp.159–178. doi.org/10.1016/0304-4076(92)90104-Y

Lahouar, A. and Ben Hadj Slama J. (2015), "Day-ahead load forecast using random forest and expert input selection", *Energy Conversion and Management* Vol.103, pp.1040–1051. doi.org/10.1016/j.enconman.2015.07.041

Lütkepohl, H. (2007), *New Introduction to Multiple Time Series Analysis*. 2nd ed., Springer, Berlin.

Osbom, D.R., Chui, A.P.L., Smith, J.P. and Birchenhall, C.R. (1988), "Seasonality and the Order of Integration for Consumption", *Oxford Bulletin of Economics and Statistics*, Vol.50 No.4, pp.361–377. https://doi.org/10.1111/j.1468-0084.1988.mp50004002.x

Pacini, H. and Silveira S. (2011), "Consumer choice between ethanol and gasoline: Lessons from Brazil and Sweden", *Energy Policy*, Vol.39 No.11, pp.6936–6942. doi.org/10.1016/j.enpol.2010.09.024

Padmakumari, K., Mohandas, K.P. and Thiruvengadam, S. (1999), "Long term distribution demand forecasting using neuro fuzzy computations", *International Journal of Electrical Power and Energy System*, Vol.21 No.5, pp.315–322. doi.org/10.1016/S0142-0615(98)00056-8

Pai, P.F. (2006), "Hybrid ellipsoidal fuzzy systems in forecasting regional electricity loads", *Energy Conversion and Management*, Vol.47 No.15-16, pp.2283–2289. doi.org/10.1016/j.enconman.2005.11.017.

Ringwood, J.V., Bofelli, D. and Murray, F.T. (2001), "Forecasting Electricity Demand on Short, Medium and Long Time Scales Using Neural Networks", *Journal of Intelligent and Robotic Systems*, Vol.31, pp.129–147. doi.org/10.1023/A:1012046824237

Rodrigues, L. and Bacchi, M.R.P. (2017), "Analyzing light fuel demand elasticities in Brazil using cointegration techniques", *Energy Economics*, Vol.63, pp.322–331. dx.doi.org/10.1016/j.eneco.2017.02.012

Rodrigues, L. and Bacchi, M.R.P. (2016), "Light fuel demand and public policies in Brazil, 2003-2013", *Applied Economics*, Vol.48, pp.5300–5313. dx.doi.org/10.1080/00036846.2016.1176115

Rodrigues, L. and Rodrigues, L. (2018), "Economic-financial performance of the Brazilian sugarcane energy industry: An empirical evaluation using financial ratio, cluster and discriminant analysis", *Biomass and Bioenergy*, Vol.108, pp.289–296. doi.org/10.1016/j.biombioe.2017.11.013

Salvo, A. and Huse, C. (2013), "Build it, but will they come? Evidence from consumer choice between gasoline and sugarcane ethanol", *Journal of Environmental Economics and Management*, Vol. 66 No.2, pp.251–279. doi.org/10.1016/j.jeem.2013.04.001.

Santiago, F.S., Mattos, R.S. and Perobello, F.S. (2011), "Um modelo integrado econométrico+insumo-produto para previsão de longo prazo da demanda de combustíveis no Brasil", *Nova Economia*, Vol.21 No.3, pp.423–455. doi.org/10.1590/S0103-63512011000300005.

Santos, A.S., Gilio, L., Halmenschlager, V., Diniza, T.B. and Almeida, A.N. (2018), "Flexible-fuel automobiles and CO2 emissions in Brazil: Parametric and semiparametric analysis using panel data", *Habitat International*, Vol.71, pp.147–155. doi.org/10.1016/j.habitatint.2017.11.014.

Santos, G.F. (2013), "Fuel demand in Brazil in a dynamic panel data approach", *Energy Economics*, Vol.36, pp.229–240. doi.org/10.1016/j.eneco.2012.08.012

Shapiro, S.S. and Wilk, M.B. (1965), "An Analysis of Variance Test for Normality (Complete Samples)", *Biometrika*, Vol.52 No.3-4, pp.591–611. doi.org/10.2307/2333709

Shingala, M.C. (2016), *Comparison of different post HOC tests*, Thesis for the degree of Doctor of Philosophy. http://hdl.handle.net/10603/171827

Sözen, A., Arcaklioğlu, E. and Özkaymak, M. (2005) "Turkey's net energy consumption", *Applied Energy*, Vol.81 No.2, pp.209–221. doi.org/10.1016/j.apenergy.2004.07.001

UNICA – Brazilian Sugarcane Industry Association. (2020), "Brazilian automobile and light fuel fleet (Otto cycle)", available at: https://observatoriodacana.com.br/listagem.php?idMn=55&idioma=2 (accessed 31 January 2021).

Wang, L.X. and Mendel, J.M. (1992), "Generating Fuzzy Rules by Learning from Examples", *IEEE Transactions on Systems, Man and Cybernetics*, Vol.22 No.6, pp.1414–1427. doi.org/10.1109/21.199466

Zanini, A., Souza, R.C. and Pedreira, C.E. (2001), "Um modelo híbrido para previsão de curto prazo da demanda de gasolina automotiva no Brasil", paper presented at: Proceedings of the V

Brazilian Conference on Neural Networks, April 2-5, Rio de Janeiro. doi.org/10.21528/CBRN2001-086.

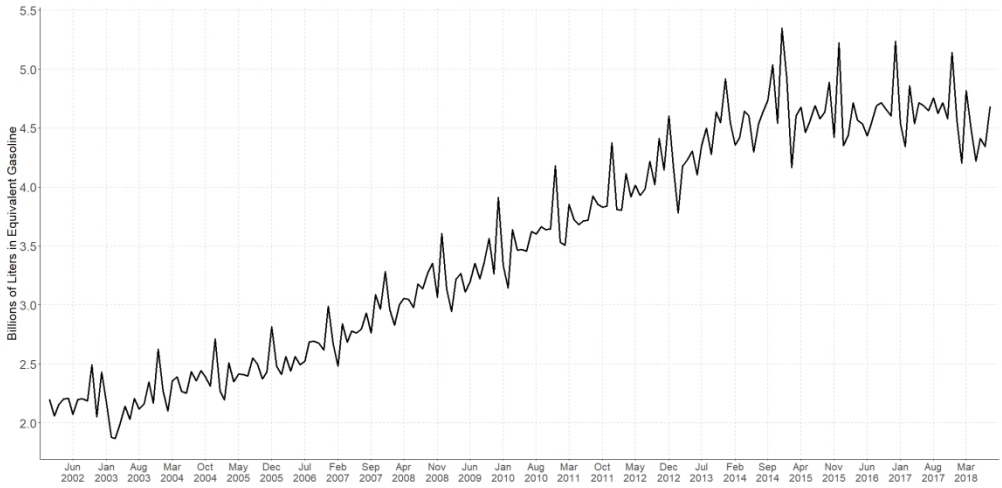

Figure 1. Consumption of Hydrous Ethanol, Gasoline and NGV in liters of equivalent gasoline monthly from August 2002 to August 2018.

542x261mm (96 x 96 DPI)

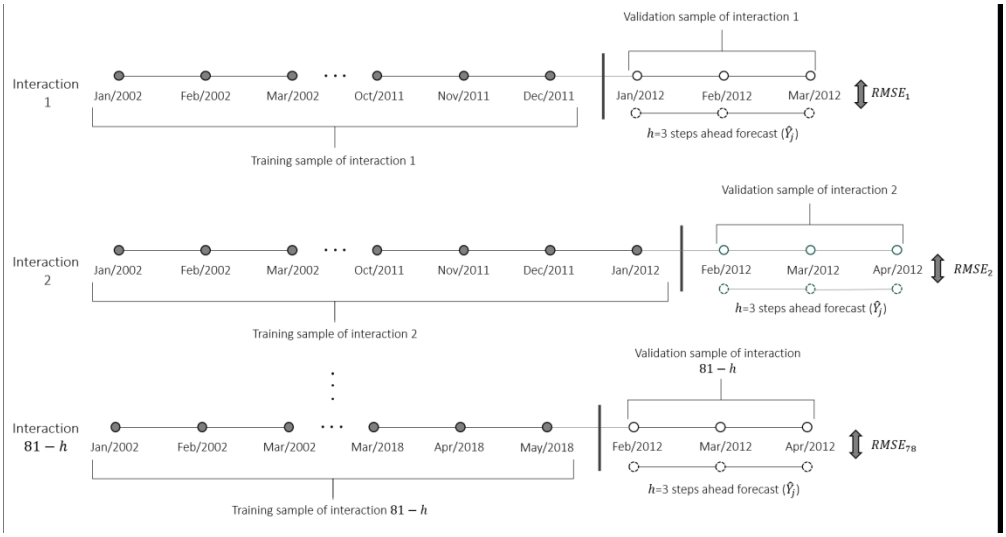

Figure 2. Procedure for estimating the prediction error of the models.

324x171mm (150 x 150 DPI)

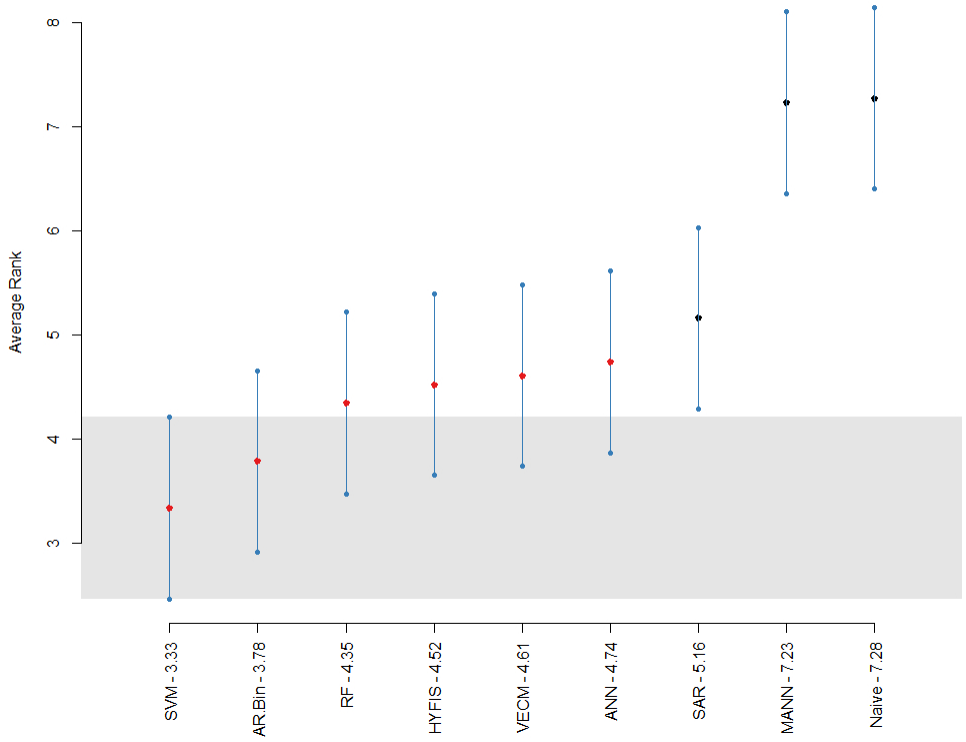

Figure 3. Result and confidence interval of the Nemenyi test for forecast 12 months ahead.

264x223mm (96 x 96 DPI)

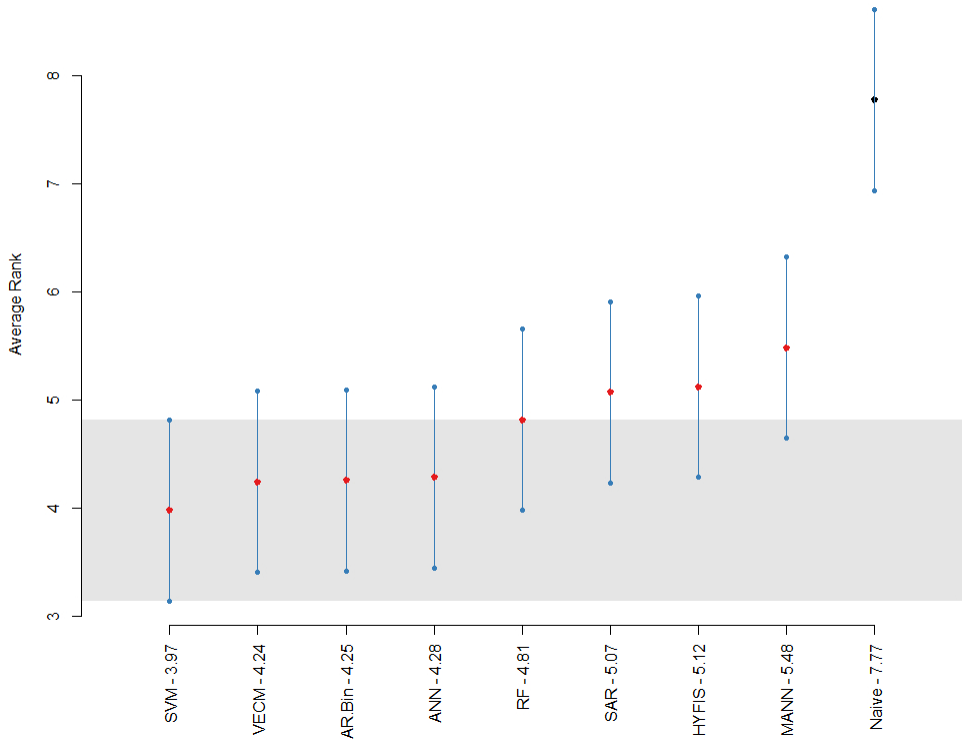

Figure 4. Result and confidence interval of the Nemenyi test for forecast 6 months ahead.

264x223mm (96 x 96 DPI)

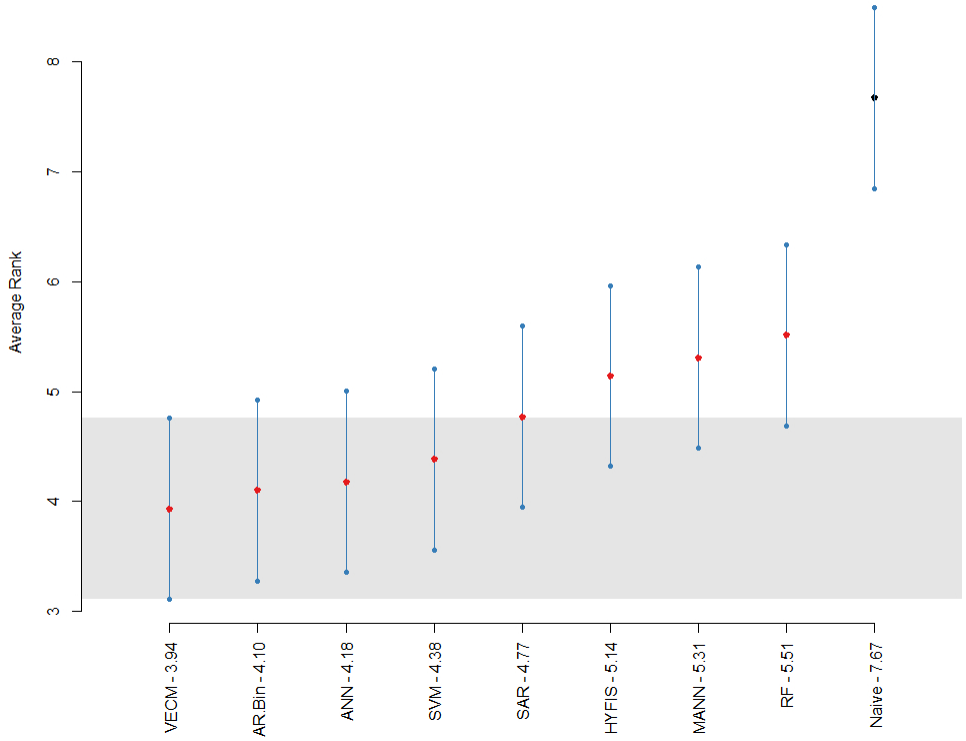

Figure 5. Result and confidence interval of the Nemenyi test for forecast 3 months ahead.

264x223mm (96 x 96 DPI)

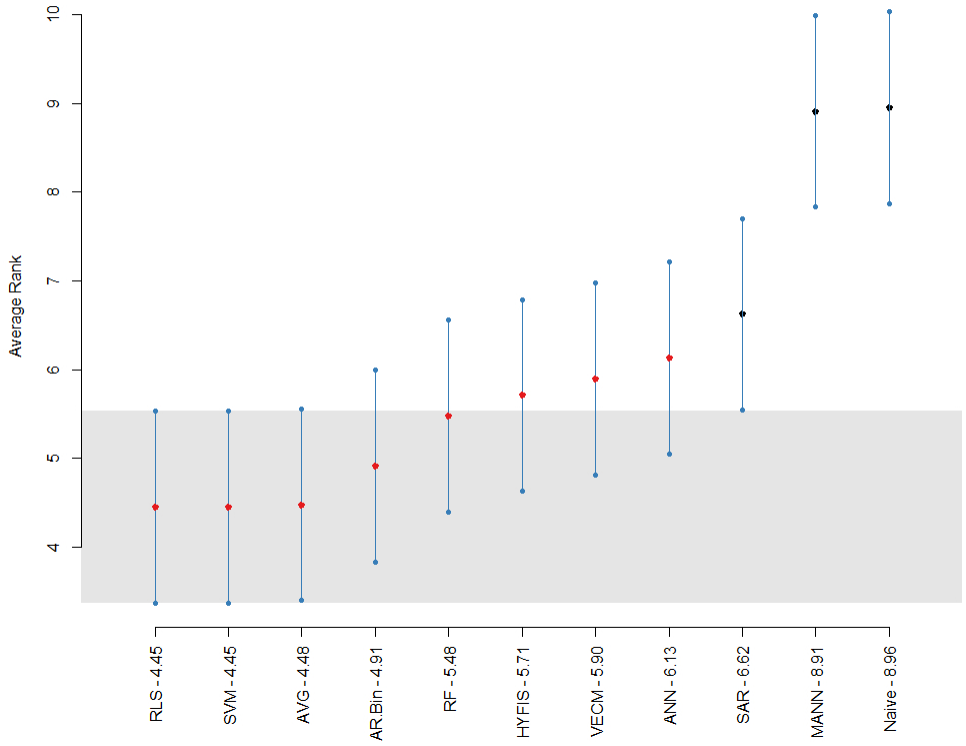

Figure 6. Result and confidence interval of the Nemenyi test for forecast 12 months ahead including the ensemble learning models.

264x223mm (96 x 96 DPI)

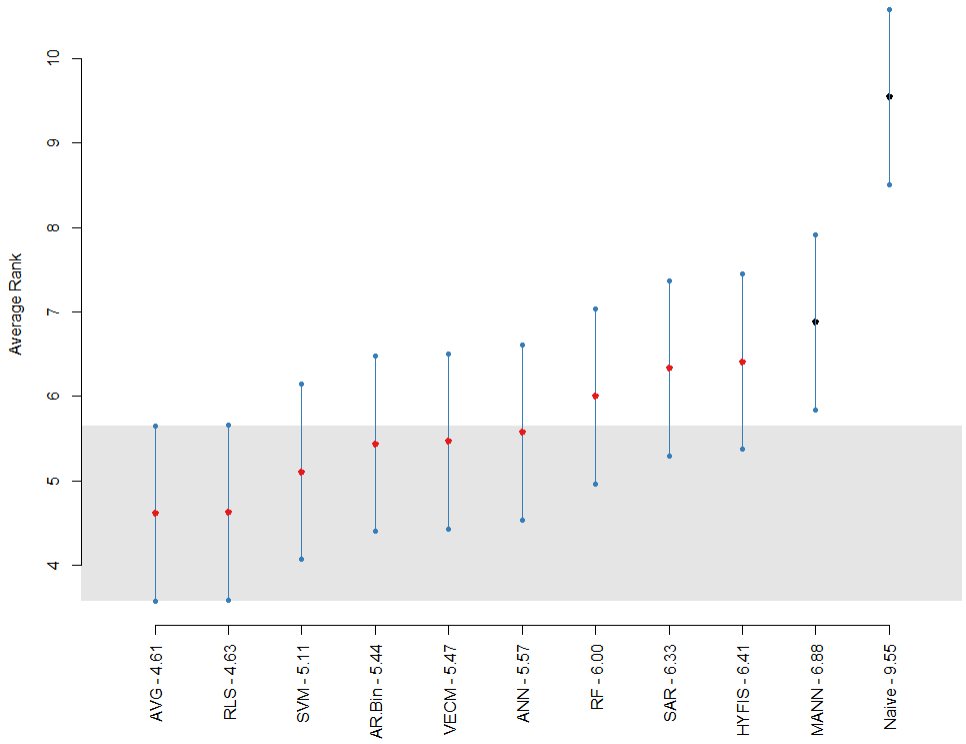

Figure 7. Result and confidence interval of the Nemenyi test for forecast 6 months ahead including the ensemble learning models.

264x223mm (96 x 96 DPI)

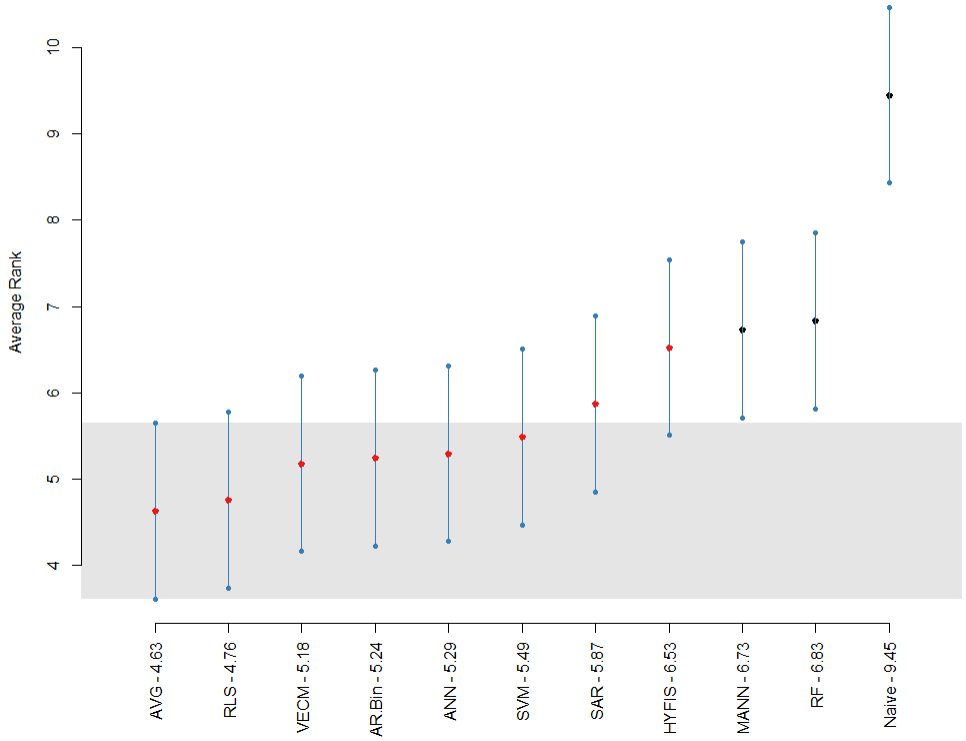

Figure 8. Result and confidence interval of the Nemenyi test for forecast 3 months ahead including the ensemble learning models.

264x223mm (96 x 96 DPI)

Table I: Hyperparameters to be priori determined in each model.

Model	Hyperparameters
Sarima	Number of autoregressive terms (p), moving average (q), integration order (d), number of seasonal autoregressive terms (P), seasonal moving average (Q) and seasonal integration order (D).
Arima w/ Binarys	Number of autoregressive terms (p), moving average (q), integration order (d),
VECM	Number of autoregressive terms (p) and cointegration vectors (r).
ANN	Number of autoregressive terms (p), number of neurons in the first and second hidden layer (k_1, k_2).
Multi ANN	Number of autoregressive terms (p) of the endogenous variables, i.e, Otto cycle consumption, average price of light fuels and GDP per capita, number of neurons in the first and second hidden layer (k_1,k_2).
SVM	Number of autoregressive terms (p), error tolerance threshold (ϵ), constant C and the kernel function for input space transformation.
Hyfis	Number of autoregressive terms (p), number of labels that define the input and output speech universe (L).
Random Forest	Number of autoregressive terms (p), number of trees (B), number of explanatory variables used by a tree (m) and the minimum size of a terminal node (n_{min}^{B}).

Table II: Hyperparameters selected from cross-validation procedures and information criteria.

Model	Treatment of Seasonality	Data pre-processing	Optimal Hyperparameters	
Sarima	Seasonal autoregressive terms	Logarithm	$p = 2, d = 1, q = 3$ $P = 1, D = 0, Q = 1$	
Arima w/ Binarys	Binary variables	Logarithm	$p = 5, d = 1, q = 5$	
VECM	Binary variables	Logarithm	$p = 5, r = 1$	
ANN	Moving Averages Filter	Standardization by mean and standard deviation	3 months 6 months 12 months	$p = 3, k_1 = 1, k_2 = 2$ $p = 5, k_1 = 1, k_2 = 5$ $p = 4, k_1 = 1, k_2 = 2$
Multi ANN	3 meses 6 meses 12 meses Moving Averages Filter Binary variables	Standardization by mean and standard deviation	3 months 6 months 12 months	$p = 4, k_1 = 10, k_2 = 10$ $p = 3, k_1 = 4, k_2 = 0$ $p = 3, k_1 = 10, k_2 = 9$
SVM	Moving Averages Filter	Standardization by mean and standard deviation	3 months 6 months 12 months	$p = 3, \epsilon = 0{,}1, C = 6,$ kernel=gaussiano $p = 3, \epsilon = 0{,}1, C = 10,$ kernel=gaussiano $p = 3, \epsilon = 0{,}1, C = 10,$ kernel=gaussiano
HyFIS	Moving Averages Filter	Standardization by mean and standard deviation	3 months 6 months 12 months	$p = 3, L = 22$ $p = 3, L = 22$ $p = 3, L = 16$
Random Forest	Moving Averages Filter	No transformation	3 months 6 months 12 months	$p = 6, B = 100, m = 5,$ $n_{min}^{B} = 6$ $p = 6, B = 100, m = 4,$ $n_{min}^{B} = 3$ $p = 6, B = 100, m = 5,$ $n_{min}^{B} = 5$

Table III: Average ranks of the models and result of the Friedman's test.

Model	Average Rank (12 months)	Average Rank (6 months)	Average Rank (3 months)
SVM	3.33	3.97	4.38
AR.Bin	3.78	4.25	4.10
RF	4.35	4.81	5.51
HYFIS	4.52	5.12	5.14
VECM	4.61	4.24	3.94
ANN	4.74	4.28	4.18
SAR	5.16	5.01	4.77
MANN	7.23	5.48	5.31
Naive	7.28	7.77	7.67
General average	5.00	5.00	5.00
p-value Friedman test (Ha: Different performances)	0.00	0.00	0.00

Table IV: Summary of accuracy performance of methods for forecasting 12 months ahead.

Model	Average MAPE	Average MAPE (Training)	MAPE standard deviation	Max MAPE	Min MAPE	Max EP
SVM	2.78%	2.12%	1.11%	5.67%	1.47%	11.61%
AR.Bin	2.89%	1.92%	1.44%	8.40%	1.22%	16.21%
VECM	2.91%	1.99%	1.21%	7.31%	1.34%	16.78%
RF	3.05%	2.66%	1.33%	6.43%	1.71%	12.03%
ANN	3.25%	2.21%	1.48%	7.24%	1.52%	13.24%
HYFIS	3.26%	2.98%	1.50%	7.14%	1.21%	12.81%
SAR	3.32%	2.15%	1.67%	8.58%	1.25%	19.71%
Naive	4.62%	-	1.71%	7.37%	1.98%	13.23%
MANN	4.71%	2.31%	1.87%	9.34%	1.35%	21.58%

Table V: Summary of accuracy performance of methods for forecasting 6 months ahead.

Model	Average MAPE	Average MAPE (Training)	MAPE standard deviation	Max MAPE	Min MAPE	Max EP
AR.Bin	2.45%	1.92%	0.99%	6.07%	0.97%	11.43%
SVM	2.47%	2.12%	0.81%	4.48%	1.11%	8.94%
VECM	2.53%	1.99%	1.01%	5.17%	0.79%	10.10%
ANN	2.54%	2.20%	0.93%	5.85%	0.69%	9.45%
RF	2.66%	2.64%	0.96%	5.34%	1.20%	9.73%
SAR	2.68%	2.15%	1.20%	6.29%	1.06%	13.81%
HYFIS	2.74%	2.67%	1.02%	5.34%	0.90%	10.22%
MANN	3.37%	3.25%	1.99%	12.56%	2.00 %	18.42%
Naive	4.65%	-	1.94%	8.64%	1.24%	13.23%

Table VI: Summary of accuracy performance of methods for forecasting 3 months ahead.

Model	Average MAPE	Average MAPE (Training)	MAPE standard deviation	Max MAPE	Min MAPE	Max EP
AR.Bin	2.28%	1.92%	1.17%	6.06%	0.49%	11.79%
VECM	2.28%	1.99%	1.23%	5.63%	0.64%	9.19%
SVM	2.34%	2.12%	1.03%	5.54%	0.68%	7.97%
ANN	2.40%	2.20%	1.09%	5.11%	0.51%	8.05%
SAR	2.42%	2.15%	1.29%	7.31%	0.67%	13.20%
HYFIS	2.56%	2.66%	1.18%	5.45%	0.70%	9.46%
RF	2.58%	2.65%	1.17%	5.68%	0.54%	9.44%
MANN	2.77%	3.26%	1.29%	6.88%	0.79%	10.93%
Naive	4.72%	-	2.26%	9.54%	0.95%	13.23%

Table VII: Average ranks of individual and combined models and Friedman's test result.

Model	Average Rank (12 months)	Average Rank (6 months)	Average Rank (3 months)
RLS	6.45	6.69	4.76
AVG	6.52	6.72	4.63
SVM	6.96	7.57	5.49
AR.Bin	7.00	7.65	5.24
RF	8.33	8.75	6.83
VECM	8.41	8.11	5.18
HYFIS	8.68	9.52	6.53
ANN	9.04	8.21	5.29
SAR	9.46	8.99	5.87
MANN	13.51	10.20	6.73
Seas Naive	13.38	14.32	9.45
General average	9.00	9.00	9.00
p-value Friedman test (Ha: Different performances)	0.00	0.00	0.00

Tabela VIII: Summary of accuracy performance of combined models for forecasting 12 months ahead.

Model	Average MAPE	Average MAPE (Training)	MAPE standard deviation	Max MAPE	Min MAPE	Max EP
AVG	2.65%	1.90%	0.95%	6.75%	1.23%	14.01%
RLS	2.68%	1.88%	1.21%	7.89%	1.15%	15.33%

Tabela IX: Summary of accuracy performance of combined models for forecasting 12 months ahead.

Model	Average MAPE	Average MAPE (Training)	MAPE standard deviation	Max MAPE	Min MAPE	Max EP
AVG	2.33%	1.90%	0.82%	4.98%	0.88%	9.85%
RLS	2.34%	1.88%	0.91%	5.68%	0.85%	10.23%

Tabela X: Summary of accuracy performance of combined models for forecasting 12 months ahead.

Model	Average MAPE	Average MAPE (Training)	MAPE standard deviation	Max MAPE	Min MAPE	Max EP
RLS	2.15%	1.88%	1.14%	5.54%	0.54%	9.05%
AVG	2.18%	1.90%	1.08%	5.54%	0.54%	8.83%

Appendice A: Statistical tests of the models Sarima and Arima with binary variables

Table A.1. Results of Dickey-Pantula test for two-unit roots for the Otto cycle consumption series.

Dickey-Pantula test conclusion	Type of test	Test statistics	Critical Value (5%)
Number of unit roots < 2	Without deterministic terms	-3.8142	-3.4333
Number of unit roots < 2	With constant	-5.2328	-3.4334
Number of unit roots < 2	With constant and trend	-5.5911	-3.4335

Table A.2. Results of DF-GLS unit root test for the Otto cycle consumption series.

DF-GLS test conclusion	Type of test	Test statistics	Critical Value (5%)
Not stationary	With trend	-0.4756	-2.93
Not stationary	With constant	1.2701	-1.94

Table A.3. Results of KPSS unit root test for the Otto cycle consumption series.

KPSS test conclusion	Type of test	Test statistics	Critical Value (5%)
Not stationary	With constant	1.3934	0.463
Not stationary	With constant and trend	0.2629	0.146

Table A.4. Results of OCSB test for seasonal unit root for the Otto cycle consumption series.

OCSB test conclusion	Test statistics	Critical Value (5%)
Not have seasonal unit root	-5.3775	-1.803

Table A.5. Autocorrelation test of the residuals of the Sarima model for the Otto cycle consumption series.

Ljung-Box test conclusion (CL=5%)	p-value	Number of lags used in the test
There is no autocorrelation	0.9104	2
There is no autocorrelation	0.6804	4
There is no autocorrelation	0.8003	6
There is no autocorrelation	0.6484	12
There is no autocorrelation	0.7556	18

Table A.6. Normality test of the residuals of the Sarima model for the Otto cycle consumption series.

Normality test conclusion (CL=5%)	Test	p-value
Not normally distributed	Jarque-Bera	0.001785
Not normally distributed	Shapiro-Wilk	0.0121

Table A.7. Heteroscedasticity test of the residuals of the Sarima model for the Otto cycle consumption series.

Heteroscedasticity test conclusion (CL=5%)	Test	Number of lags used in the test
Not heteroscedastic	LM	4, 8, 12, 16

Table A.8. Autocorrelation test of the residuals of the Arima with binary variables model for the Otto cycle consumption series.

Ljung-Box test conclusion (CL=5%)	p-value	Number of lags used in the test
There is no autocorrelation	0.8326	2
There is no autocorrelation	0.9756	4

There is no autocorrelation	0.9874	6
There is no autocorrelation	0.9896	12
There is no autocorrelation	0.9971	18

Table A.9. Normality test of the residuals of the Arima with binary variables model for the Otto cycle consumption series.

Normality test conclusion (CL=5%)	**Test**	**p-value**
Not normally distributed	Jarque-Bera	2.2×10^{-16}
Not normally distributed	Shapiro-Wilk	5.2×10^{-6}

Table A.10. Heteroscedasticity test of the residuals of the Arima with binary variables model for the Otto cycle consumption series.

Heteroscedasticity test conclusion (CL=5%)	**Test**	**Number of lags used in the test**
Not homoscedastic	LM	4, 8, 12, 16

Appendice B: Statistical tests of the VECM model

Table B.1. Results of Dickey-Pantula test for two-unit roots for the average price of light fuels series.

Dickey-Pantula test conclusion	Type of test	Test statistics	Critical Value (5%)
Number of unit roots < 2	Without deterministic terms	-9.1702	-3.4333
Number of unit roots < 2	With constant	-9.1500	-3.4334
Number of unit roots < 2	With constant and trend	-9.1287	-3.4335

Table B.2. Results of DF-GLS unit root test for the average price of light fuels series.

DF-GLS test conclusion	Type of test	Test statistics	Critical Value (5%)
Not stationary	With trend	-1.9387	-2.93
Not stationary	With constant	-1.111	-1.94

Table B.3. Results of KPSS unit root test for the average price of light fuels series.

KPSS test conclusion	Type of test	Test statistics	Critical Value (5%)
Not stationary	With constant	1.029	0.463
Not stationary	With constant and trend	0.1976	0.146

Table B.4. Results of OCSB test for seasonal unit root for the average price of light fuels series.

OCSB test conclusion	Test statistics	Critical Value (5%)
Not have seasonal unit root	-14.1002	-1.803

Table B.5. Results of Dickey-Pantula test for two-unit roots for the GDP per capita series.

Dickey-Pantula test conclusion	Type of test	Test statistics	Critical Value (5%)
Number of unit roots < 2	Without deterministic terms	-3.1210	-3.4333
Number of unit roots < 2	With constant	-3.6393	-3.4334
Number of unit roots < 2	With constant and trend	-4.2924	-3.4335

Table B.6. Results of DF-GLS unit root test for the GDP per capita series.

DF-GLS test conclusion	Type of test	Test statistics	Critical Value (5%)
Not stationary	With trend	-0.5627	-2.93
Not stationary	With constant	0.4032	-1.94

Table B.7. Results of KPSS unit root test for the GDP per capita series.

KPSS test conclusion	Type of test	Test statistics	Critical Value (5%)
Not stationary	With constant	1.2468	0.463
Not stationary	With constant and trend	0.3007	0.146

Table B.8. Results of OCSB test for seasonal unit root for the GDP per capita series.

OCSB test conclusion	Test statistics	Critical Value (5%)
Not have seasonal unit root	-4.1836	-1.803

Table B.9. Johansen maximum eigenvalue cointegration test.

Test null hypothesis	Test value	Critical Value (5%)
$r \leq 2$	5.53	9.24
$r \leq 1$	13.75	15.67
$r = 0$	186.92	22.00

Table B.10. Johansen trace cointegration test.

Test null hypothesis	Test value	Critical Value (5%)
$r \leq 2$	5.53	9.24
$r \leq 1$	19.28	19.96
$r = 0$	206.20	34.91

Table B.11. Edgerton-Shukur multivariate autocorrelation test for the residuals.

Edgerton-Shukur test conclusion (CL=5%)	p-value	Number of lags used in the test
There is no autocorrelation	0.9652	2
There is no autocorrelation	0.8031	4
There is no autocorrelation	0.258	6

Table B.12. Portmanteau multivariate autocorrelation test for the residuals.

Portmanteau test conclusion (CL=5%)	p-value	Number of lags used in the test
There is autocorrelation	0.0022	6
There is no autocorrelation	0.0629	12
There is no autocorrelation	0.1542	18

Table B.13. Jarque-Bera multivariate normality test for the residuals.

Jarque-Bera test conclusion (CL=5%)	p-value	Tested series
Not normally distributed	2.138e-06	Otto cycle consumption
Not normally distributed	8.151e-13	Average price of light fuels
Not normally distributed	1.109e-06	GDP per capita
Not normally distributed	2.200e-16	All series together

Table B.13. Shapiro-Wilk multivariate normality test for the residuals.

Shapiro-Wilk test conclusion (CL=5%)	p-value	Tested series
Not normally distributed	1.24e-2	Otto cycle consumption
Not normally distributed	1.15e-6	Average price of light fuels
Not normally distributed	8.17e-3	GDP per capita
Not normally distributed	9.58e-4	All series together

Table B.14. LM multivariate heteroscedasticity test for the residuals.

Heteroscedasticity test conclusion (CL=5%)	Tested series	Number of lags used in the test
Not heteroscedastic	Otto cycle consumption	2, 5, 12, 18
Not heteroscedastic	Average price of light fuels	2, 5, 12, 18
Not heteroscedastic	GDP per capita	2, 4, 6, 12, 18
Not homoscedastic	All series together	2, 4, 6, 12, 18

Appendice C: MAPE, maximum and minimum percentage error series for all models

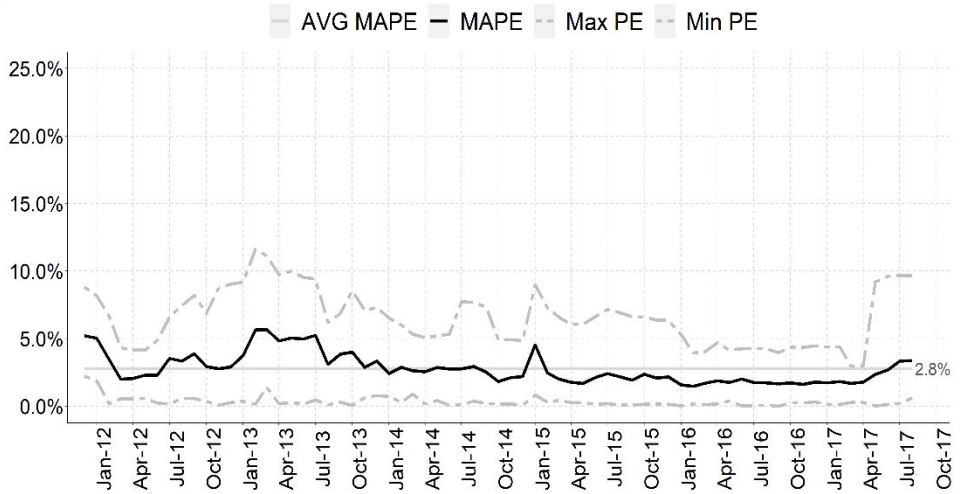

Figure C.1. MAPE, average MAPE and maximum and minimum absolute percentage error of the SVM model for 12-month forecast ahead.

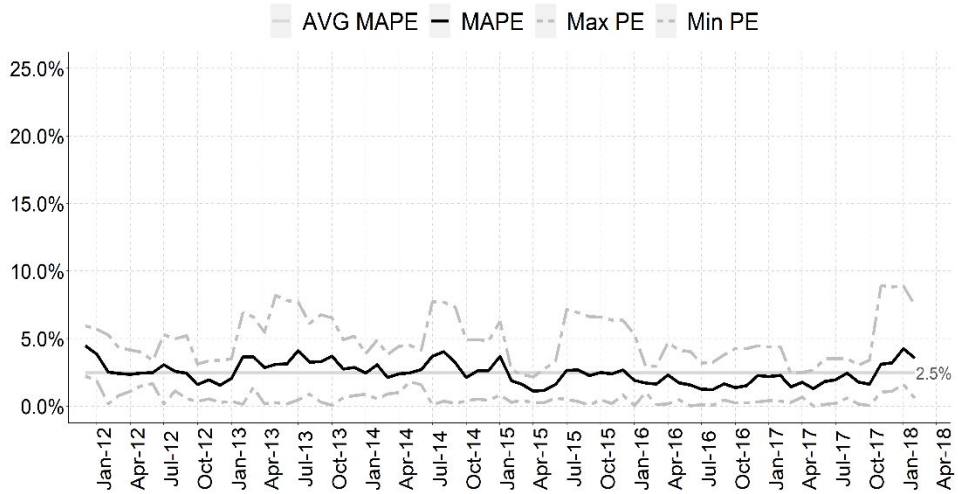

Figure C.2. MAPE, average MAPE and maximum and minimum absolute percentage error of the SVM model for 6-month forecast ahead.

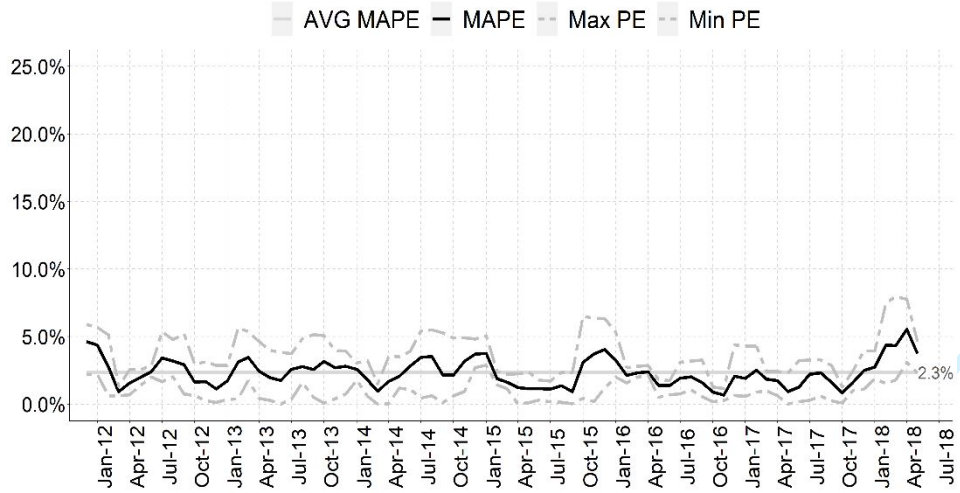

Figure C.3. MAPE, average MAPE and maximum and minimum absolute percentage error of the SVM model for 3-month forecast ahead.

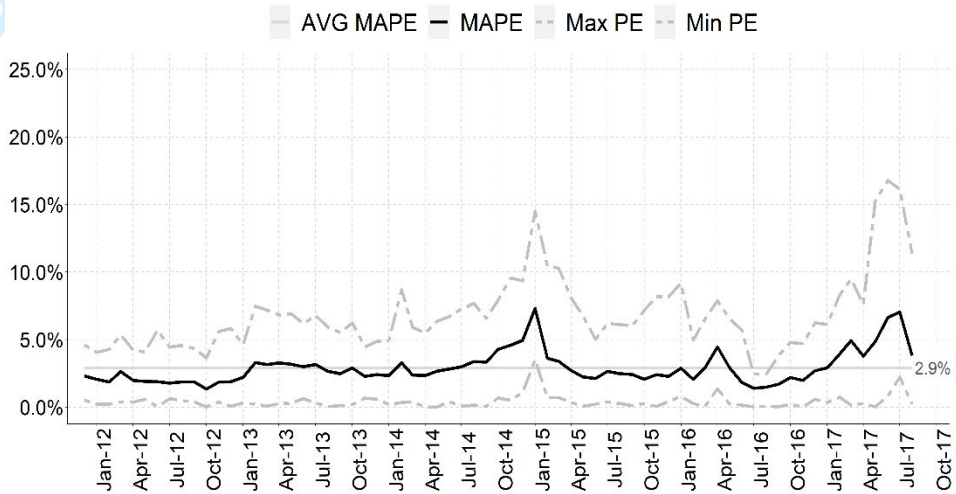

Figure C.4. MAPE, average MAPE and maximum and minimum absolute percentage error of the VECM model for 12-month forecast ahead.

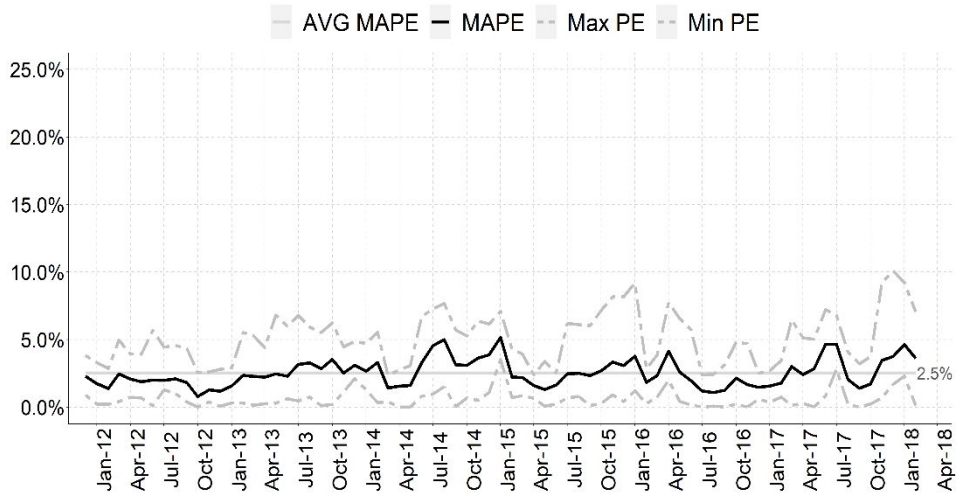

Figure C.5. MAPE, average MAPE and maximum and minimum absolute percentage error of the VECM model for 6-month forecast ahead.

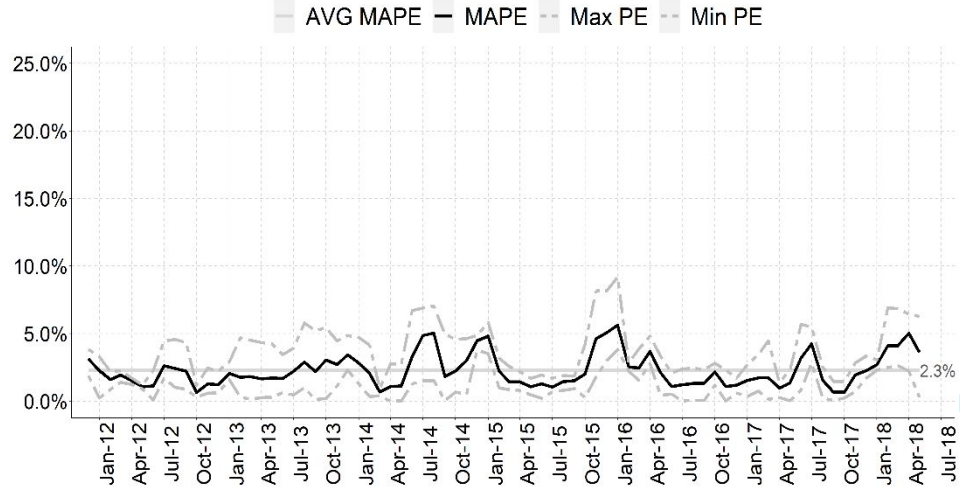

Figure C.6. MAPE, average MAPE and maximum and minimum absolute percentage error of the VECM model for 3-month forecast ahead.

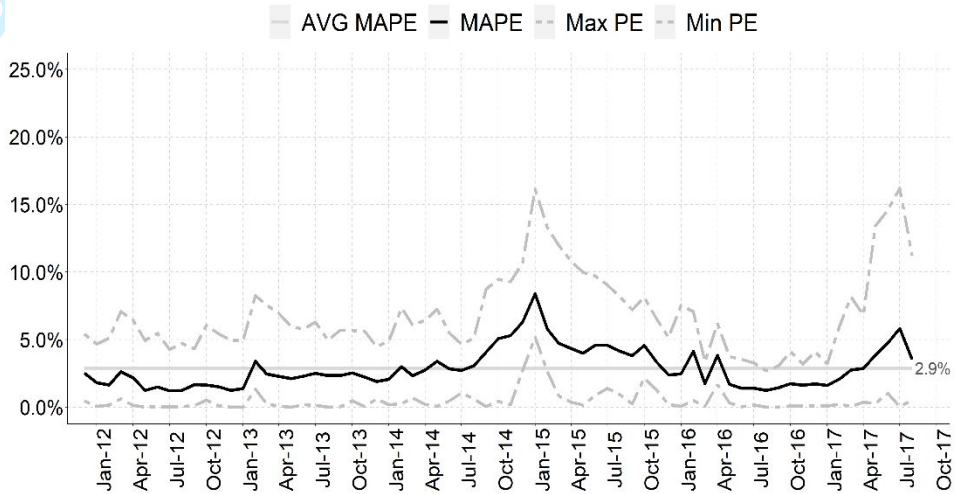

Figure C.7. MAPE, average MAPE and maximum and minimum absolute percentage error of the AR.Bin model for 12-month forecast ahead.

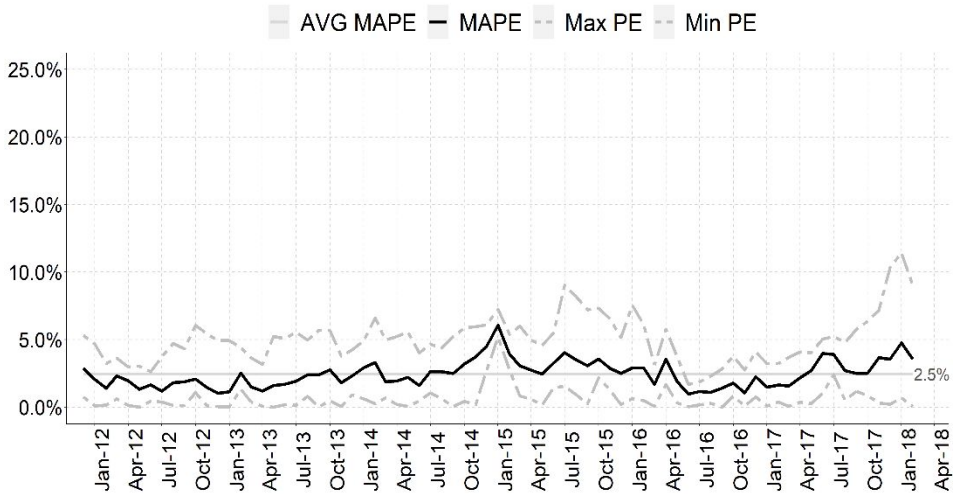

Figure C.8. MAPE, average MAPE and maximum and minimum absolute percentage error of the AR.Bin model for 6-month forecast ahead.

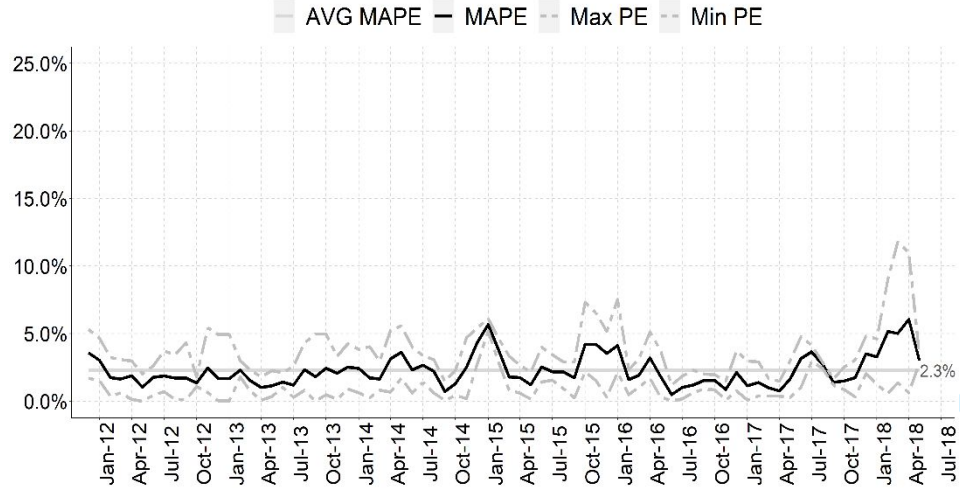

Figure C.9. MAPE, average MAPE and maximum and minimum absolute percentage error of the AR.Bin model for 3-month forecast ahead.

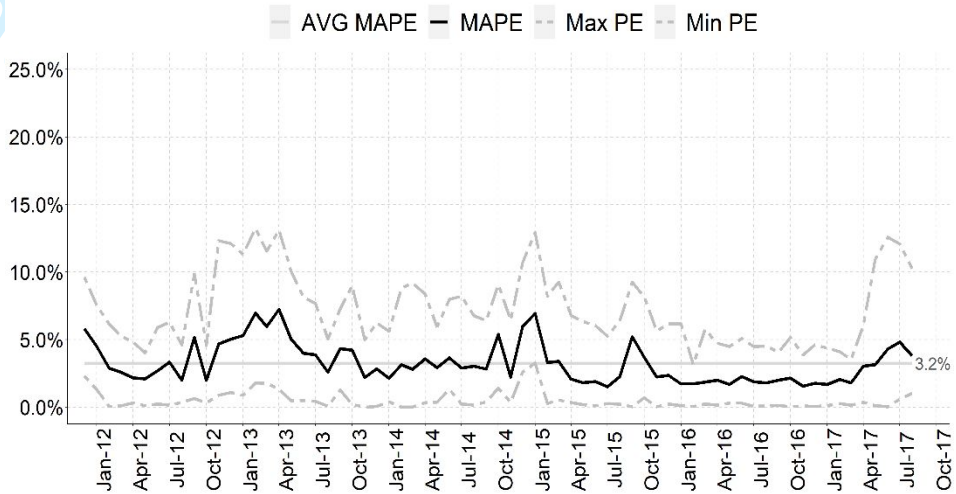

Figure C.10. MAPE, average MAPE and maximum and minimum absolute percentage error of the ANN model for 12-month forecast ahead.

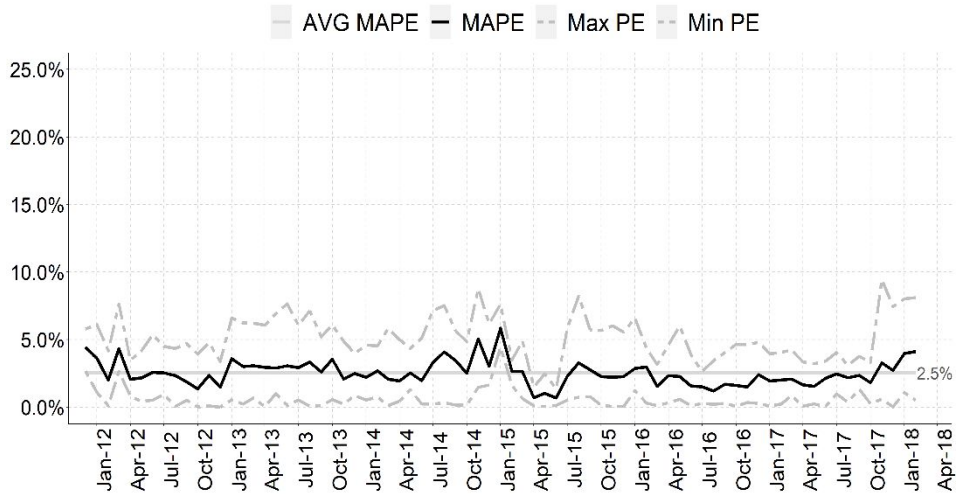

Figure C.11. MAPE, average MAPE and maximum and minimum absolute percentage error of the ANN model for 6-month forecast ahead.

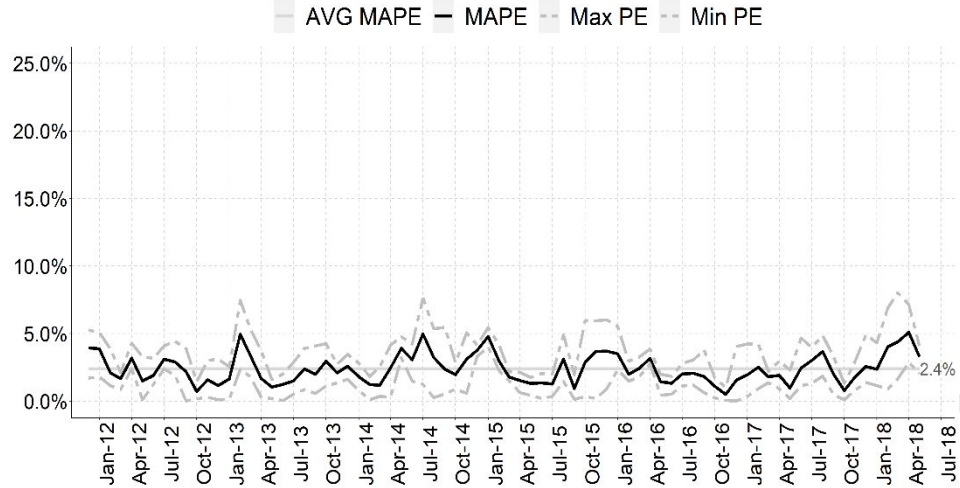

Figure C.12. MAPE, average MAPE and maximum and minimum absolute percentage error of the ANN model for 3-month forecast ahead.

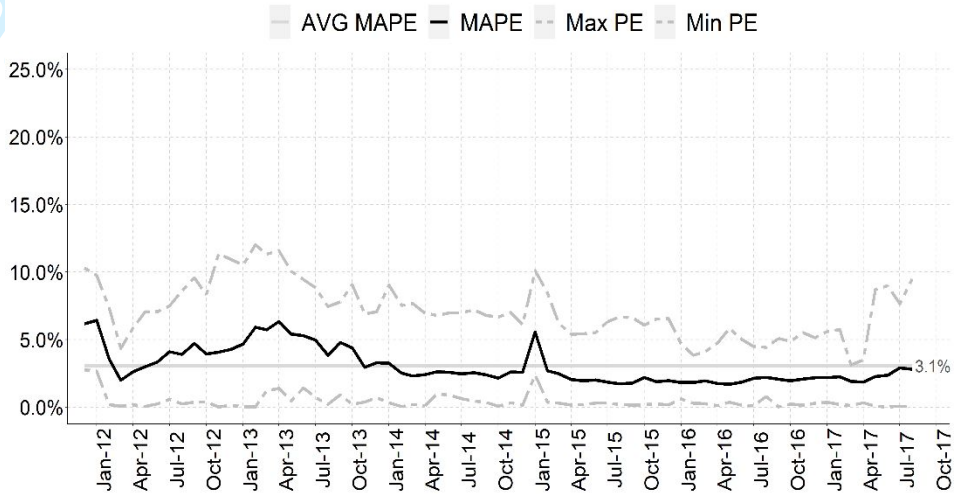

Figure C.13. MAPE, average MAPE and maximum and minimum absolute percentage error of the RF model for 12-month forecast ahead.

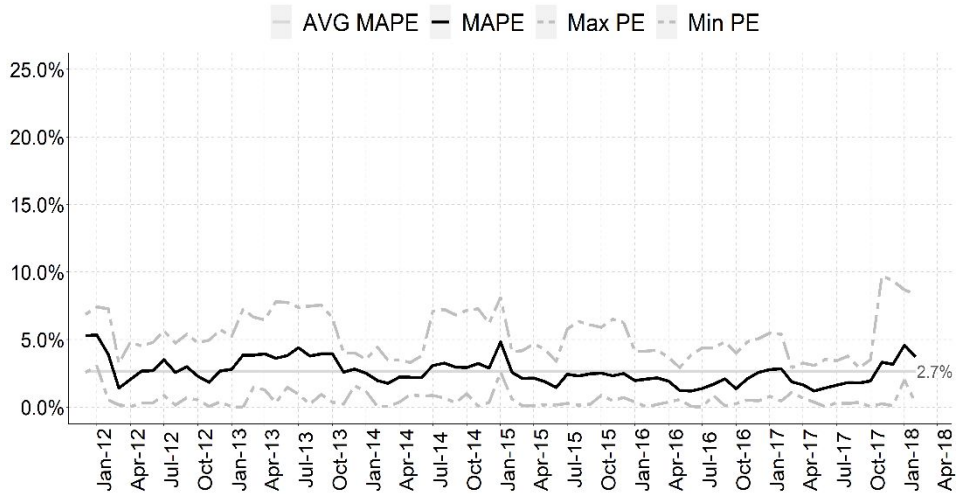

Figure C.14. MAPE, average MAPE and maximum and minimum absolute percentage error of the RF model for 6-month forecast ahead.

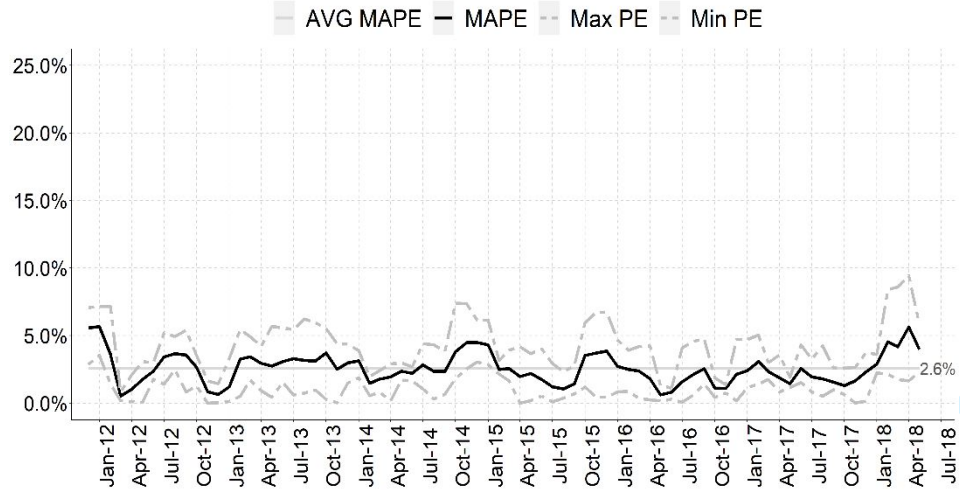

Figure C.15. MAPE, average MAPE and maximum and minimum absolute percentage error of the RF model for 3-month forecast ahead.

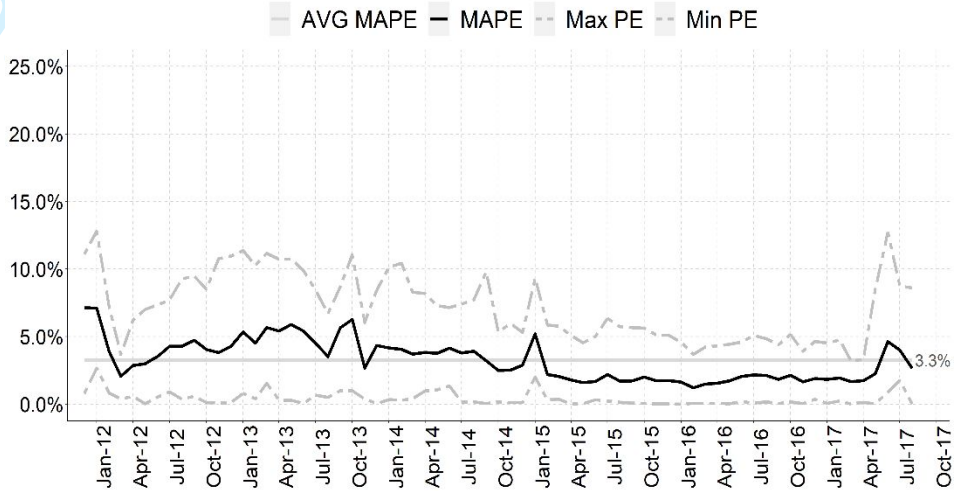

Figure C.16. MAPE, average MAPE and maximum and minimum absolute percentage error of the HyFIS model for 12-month forecast ahead.

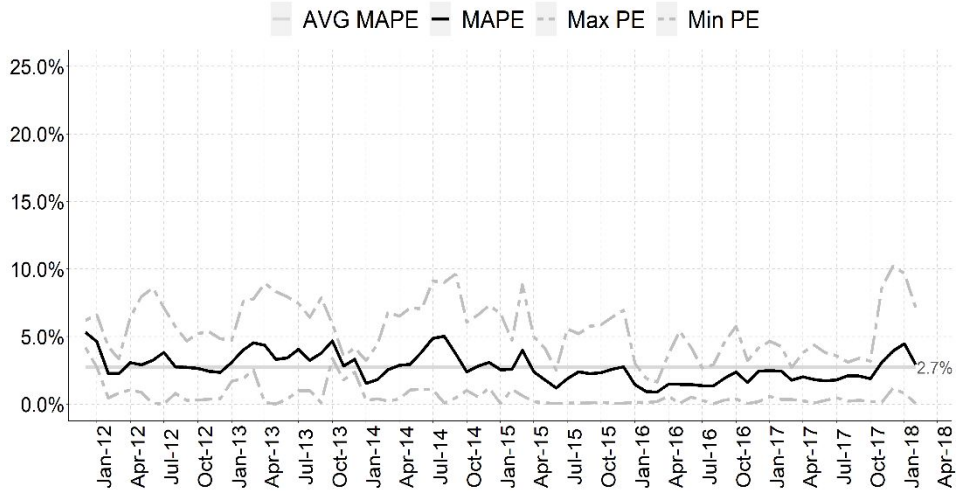

Figure C.17. MAPE, average MAPE and maximum and minimum absolute percentage error of the HyFIS model for 6-month forecast ahead.

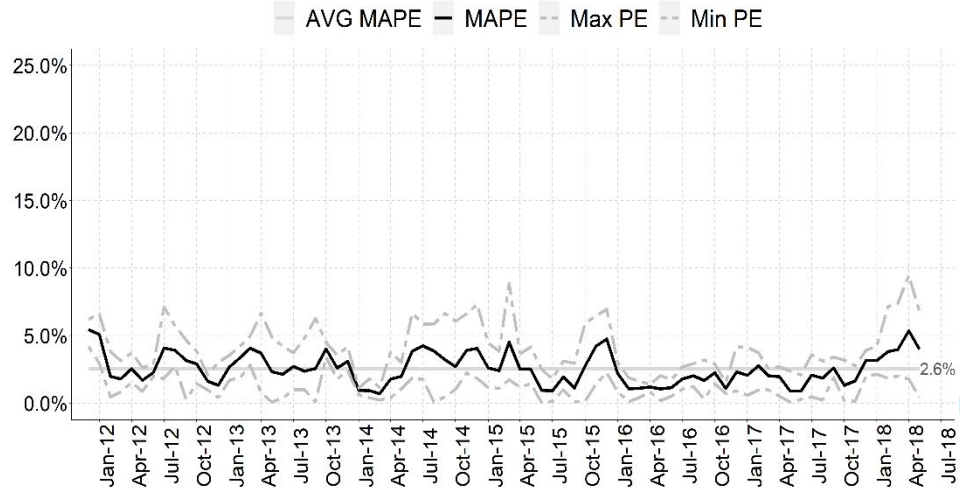

Figure C.18. MAPE, average MAPE and maximum and minimum absolute percentage error of the HyFIS model for 3-month forecast ahead.

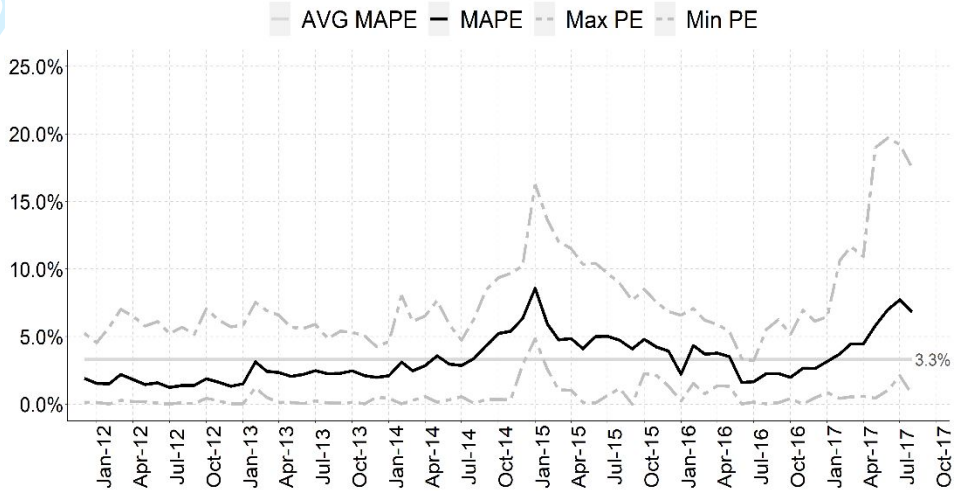

Figure C.19. MAPE, average MAPE and maximum and minimum absolute percentage error of the SAR model for 12-month forecast ahead.

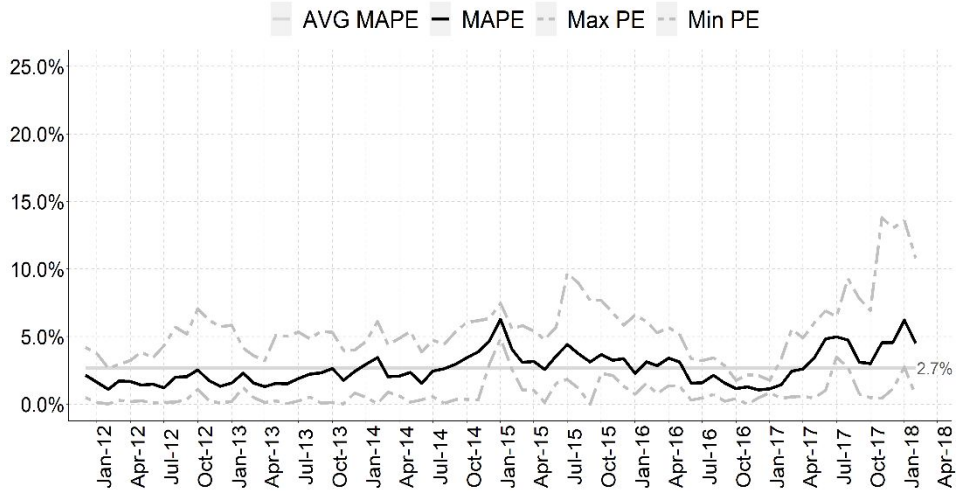

Figure C.20. MAPE, average MAPE and maximum and minimum absolute percentage error of the SAR model for 6-month forecast ahead.

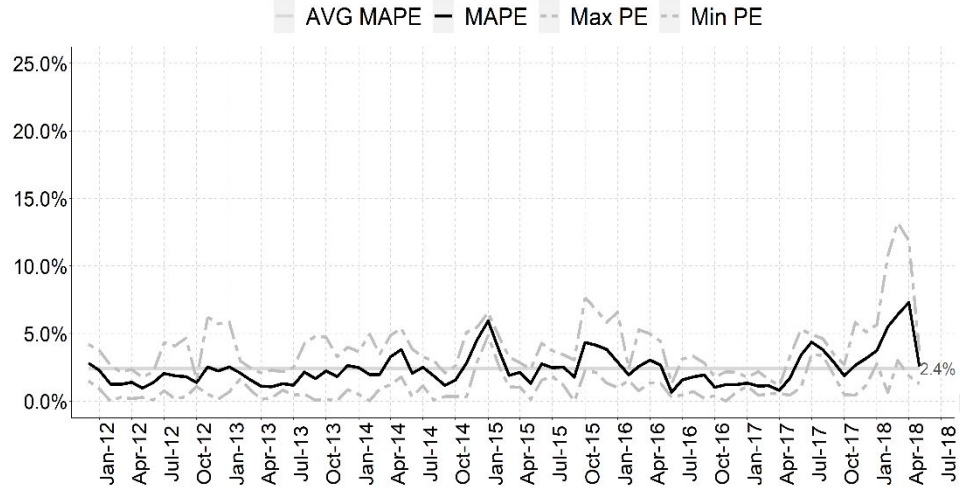

Figure C.21. MAPE, average MAPE and maximum and minimum absolute percentage error of the SAR model for 3-month forecast ahead.

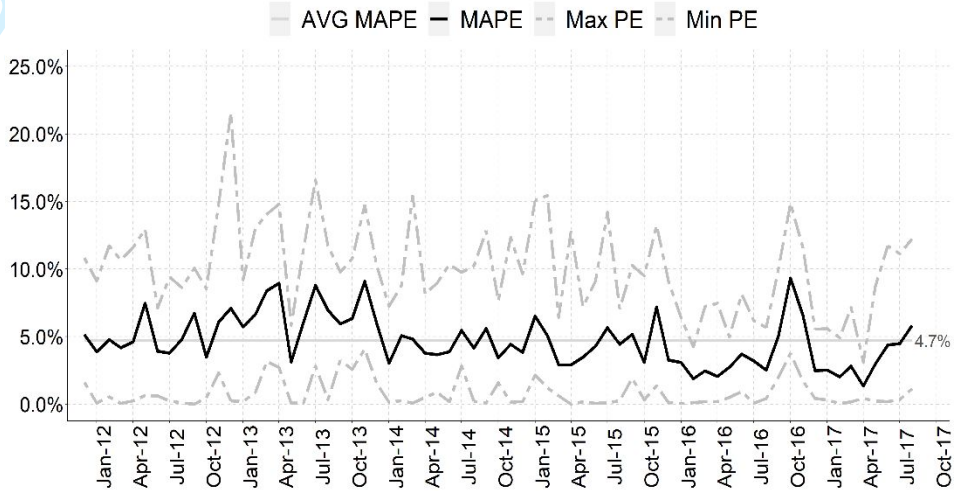

Figure C.22. MAPE, average MAPE and maximum and minimum absolute percentage error of the MANN model for 12-month forecast ahead.

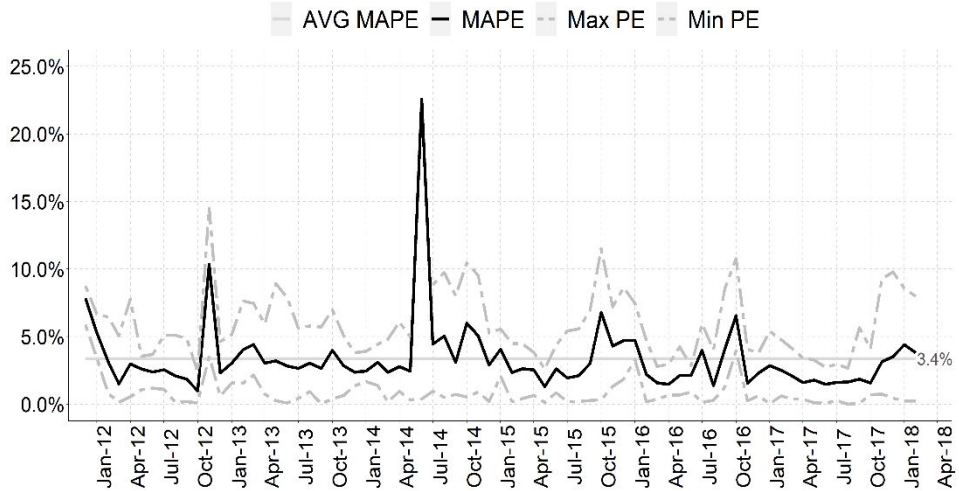

Figure C.23. MAPE, average MAPE and maximum and minimum absolute percentage error of the MANN model for 6-month forecast ahead.

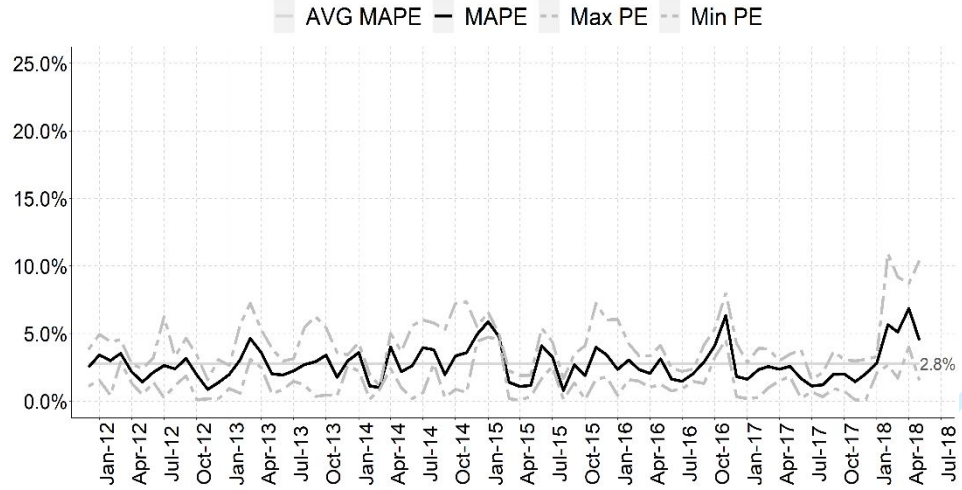

Figure C.24. MAPE, average MAPE and maximum and minimum absolute percentage error of the MANN model for 3-month forecast ahead.

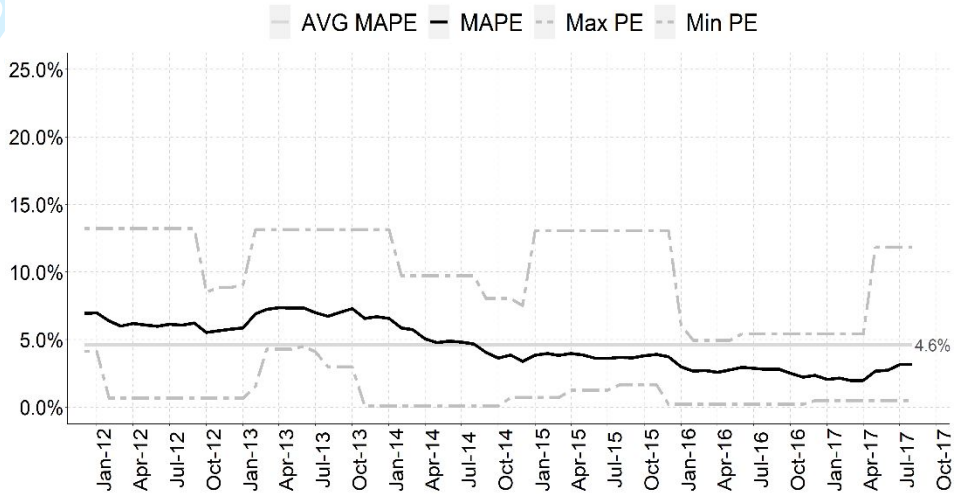

Figure C.25. MAPE, average MAPE and maximum and minimum absolute percentage error of the Naive model for 12-month forecast ahead.

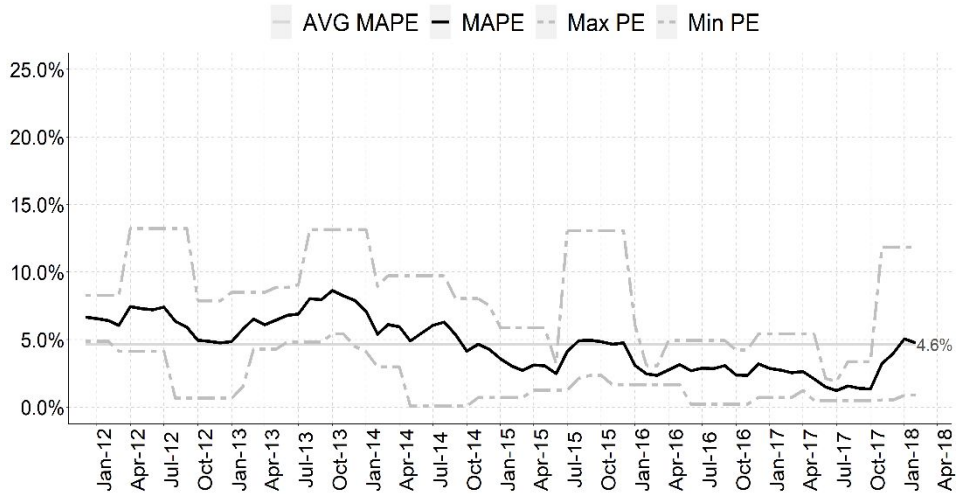

Figure C.26. MAPE, average MAPE and maximum and minimum absolute percentage error of the Naive model for 6-month forecast ahead.

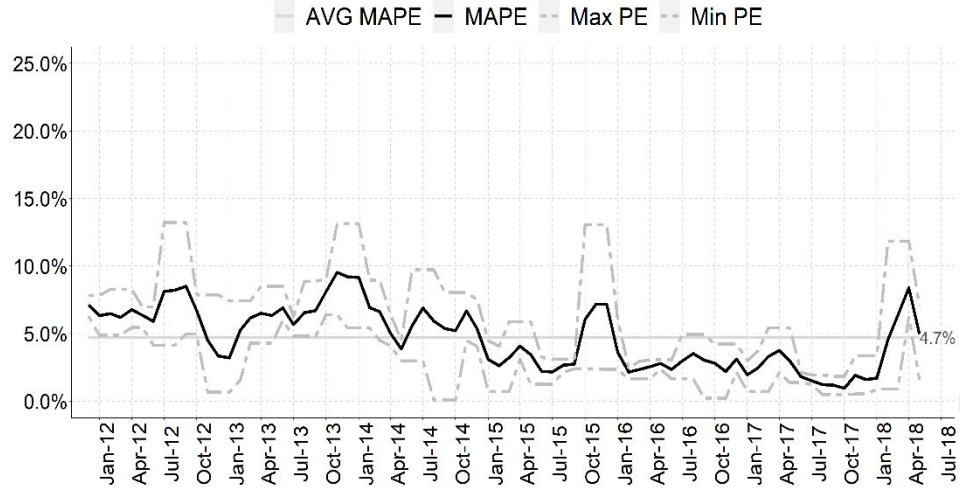

Figure C.27. MAPE, average MAPE and maximum and minimum absolute percentage error of the Naive model for 3-month forecast ahead.

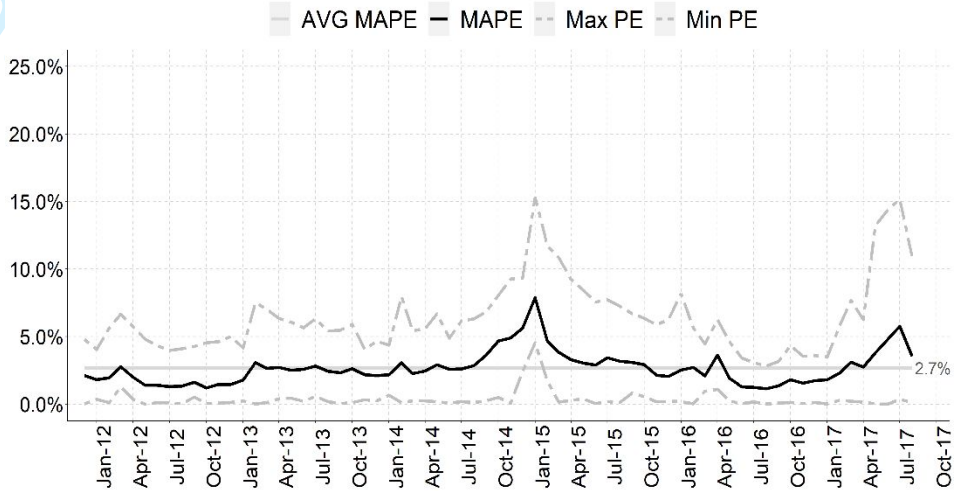

Figure C.28. MAPE, average MAPE and maximum and minimum absolute percentage error of the RLS model for 12-month forecast ahead.

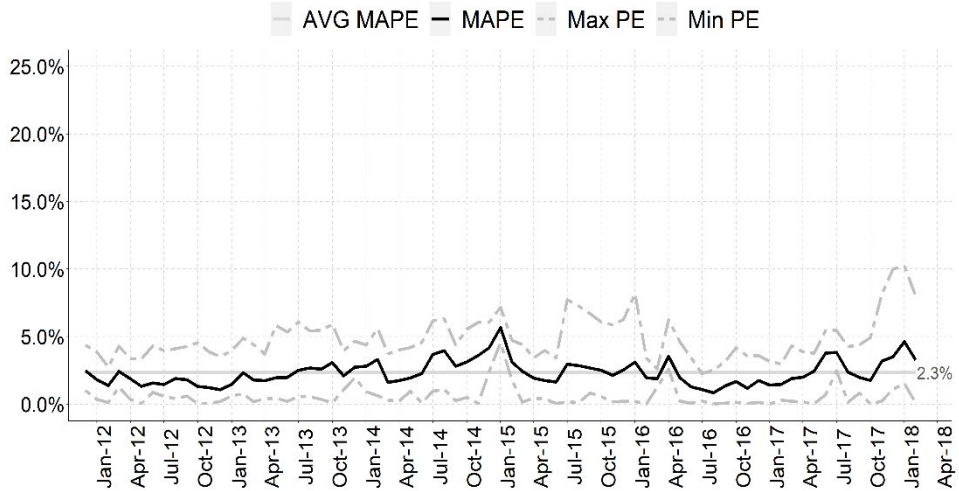

Figure C.29. MAPE, average MAPE and maximum and minimum absolute percentage error of the RLS model for 6-month forecast ahead.

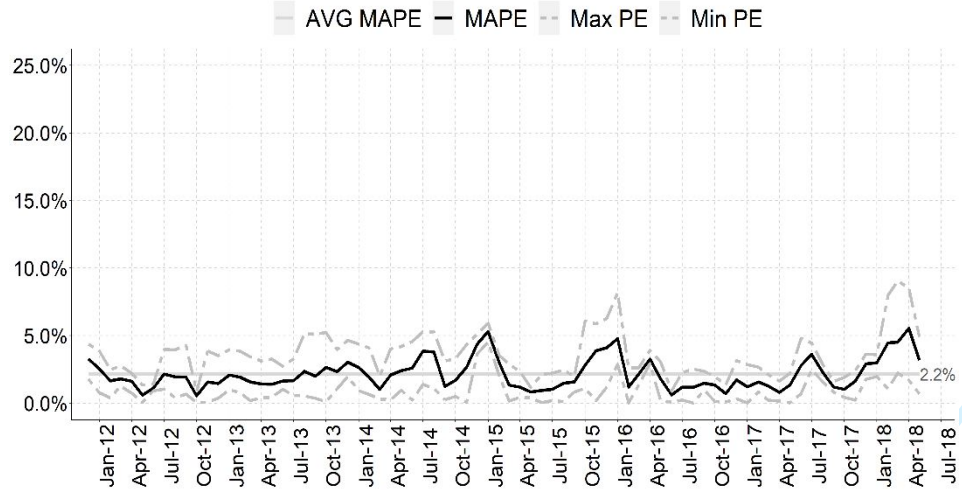

Figure C.30. MAPE, average MAPE and maximum and minimum absolute percentage error of the RLS model for 3-month forecast ahead.

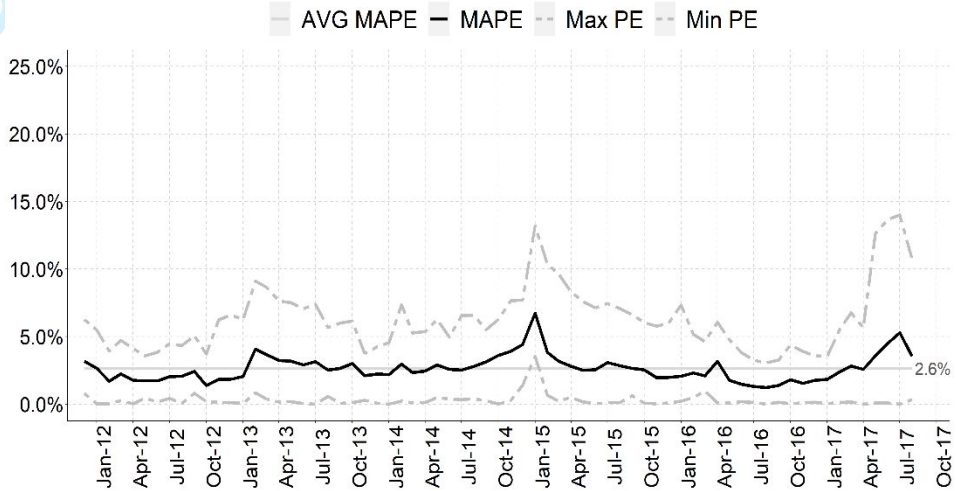

Figure C.31. MAPE, average MAPE and maximum and minimum absolute percentage error of the AVG model for 12-month forecast ahead.

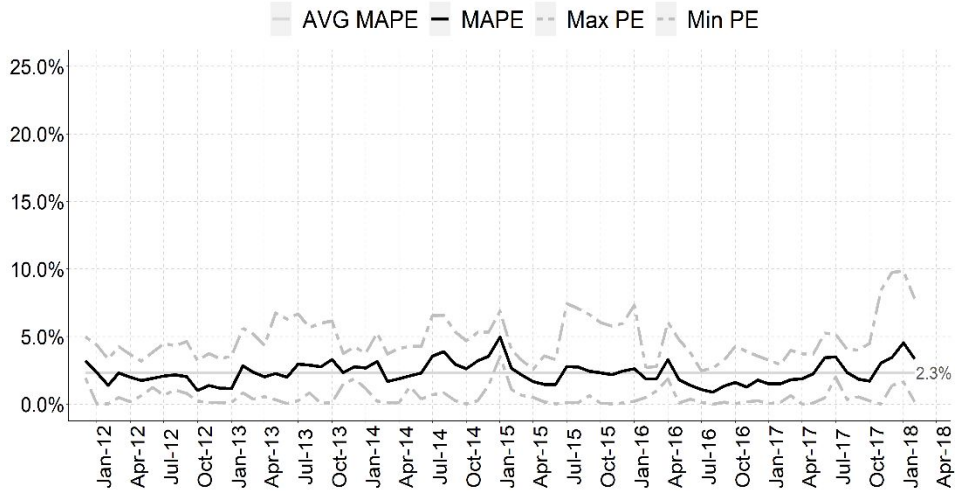

Figure C.32. MAPE, average MAPE and maximum and minimum absolute percentage error of the AVG model for 6-month forecast ahead.

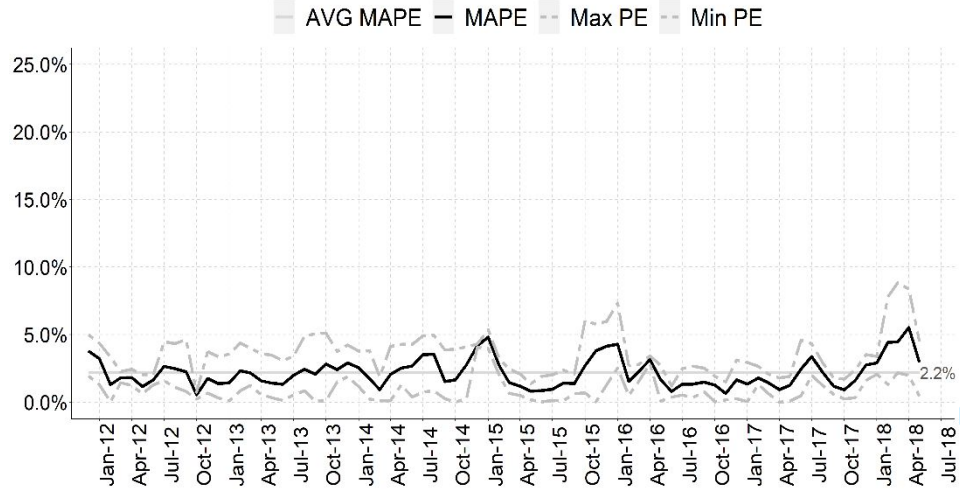

Figure C.33. MAPE, average MAPE and maximum and minimum absolute percentage error of the AVG model for 3-month forecast ahead.