Forecasting electric energy consumption under recession: an application of boosting to the case of the Brazilian state Rio Grande do Sul

Guilherme Schultz Lindenmeyer * Pedro Pablo Skorin†

Hudson da Silva Torrent ‡

Abstract

This paper seeks to test the component-wise boosting validity as an instrument of forecasting regional series in economic recessions. We use 822 predictors to forecast the monthly electricity consumption of the Brazilian state Rio Grande do Sul. The time series has 190 observations and occurs during the 2015 political and economic crisis of Brazil, the biggest economic crisis in the country's history until then, which significantly impacted the electricity consumption behavior. Boosting manages to select predictors associated with the crisis and is capable of understanding faster the trend change compared to a standard SARIMA benchmark. By filtering a high-dimensional data set, the machine learning algorithm appear as an useful instrument when forecasting short series with uncertainty.

Keywords: Boosting. Electric energy consumption. Forecast. Regional. Recession.

Resumo

Este artigo procura testar a validade do *component-wise boosting* como um instrumento de previsão de séries curtas em recessões econômicas. Utilizamos 822 variáveis preditoras para prever o consumo mensal de eletricidade do estado brasileiro do Rio Grande do Sul. A série temporal tem 190 observações e ocorre durante a crise política e econômica de 2015 do Brasil, a maior crise econômica da história do país até então, que impactou significativamente o comportamento do consumo de eletricidade. O Boosting consegue selecionar os preditores associados à crise e é capaz de entender mais rapidamente a mudança de tendência em relação a um padrão de referência SARIMA. Ao filtrar um conjunto de dados de alta dimensão, o algoritmo de *machine learning* aparece como um instrumento útil na previsão de séries curtas com incerteza.

Palavras-chave: *Boosting*. Consumo de energia elétrica. Previsão. Regional. Recessão.

Classificação JEL: C53 · Q47 · R11

Área de Submissão: 8 - Econometria

*Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, E-mail: gslindenmeyer@gmail.com

†Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, E-mail: pedro.skorin@hotmail.com

‡Departamento de Estatística, Universidade Federal do Rio Grande do Sul, E-mail: hudson.torrent@ufrgs.br

1 Introduction

Forecasting time series under unstable times requires a special look. The task is particularly hard when there are few observations and not much can be inferred about the future dynamics of the series. This is the scenario of policy-makers and investors when making decisions in times of economic recession. Our paper is motivated by the struggle uncertainty creates in forecasting short series at the regional level. As a case study, we use the monthly electric energy consumption of the Brazilian state Rio Grande do Sul between 2002 and 2017. From 2002 trough 2014 this series had a growing trend dynamic (see section 2), however from 2015 to 2017 Brazil suffered the greatest recession in its history until then, where the electric sector was particularly affected. Between the start of the recession and 2016, Brazil lost 11% of its GDP per capita and the former president Dilma Rouseff was removed from office.

Having good forecasts of the electric energy demand is highly desirable in an economy. Since electricity can't be stored in an efficient way, deviations of the supply from the demand create unwanted inefficiencies. The authors in (PARAJULI et al., 2014) assess short-run econometric models of primary energy consumption in Nepal. They discuss about the relation between energy consumption and economic performance. The article of (ASSIS CABRAL; LEGEY; FREITAS CABRAL, 2017) focus in understanding the spatial dependence in regional electricity consumption in Brazil. The authors show a Spatial ARIMA model with a better predictive performance than the standard ARIMA model. Machine learning methods have also been used in the literature. The authors in (GONZÁLEZ-ROMERA; JARAMILLO-MORÁN; CARMONA-FERNÁNDEZ, 2008) combine two instruments to forecast the spanish monthly electric consumption. A Fourier series is adjusted to reproduce the fluctuation behavior of the consumption and a neural network is used to forecast the trend of the data.

In this paper, the data set that will be used to perform the forecast is monthly and the series vary from 2002 to 2017. The forecasting exercise happens in the last 20% of our date range, from 2014 to 2017. Therefore, we are only concerned about the crisis period. 822 predictors will be used, thus, qualifying our data set as a high-dimension environment, time series filled. To deal with this scenario, the conventional and classical econometric methods, such as regression, does not work, since by dealing with a large amount of predictors, others methods may be more appropriate. As (VARIAN, 2014) points out, in a high-dimensional environment, bringing machine learning models can be interesting for econometrics. To see this, the article from (MEDEIROS et al., 2019) is a good current example of the use of several machine learning models to forecast US inflation with a high-dimensional data set.

Among the diverse world of machine learning, we have selected the boosting method. To visualize the application and performance of these other different methods, the (MEDEIROS et al., 2019) uses a vast collection of machine learning models and compares them. The boosting method was initially introduced by (SCHAPIRE, 1990; FREUND, 1995; FREUND; SCHAPIRE, et al., 1996) and it consists of an algorithm that can be used to significantly reduce the error of another model. In order to understand the different types of boosting, the recent article (CHU et al., 2020) provides details

and models the different approaches of boosting, as well as their different economic applications. The method to be used in this article is the L_2 boosting, which is the application of the algorithm in linear regression models. This method was first conceived and developed by (BÜHLMANN, 2003) and its consistency[1] was proven in (BÜHLMANN et al., 2006). Its efficiency in dealing with macroeconomic data can be seen in (BAI; NG, 2009), where the L_2 component-wise boosting method is used to select variables as predictors of the model. As pointed out by (SCHMID; HOTHORN, 2008) "when the number of covariates p in a data set is large (and when selecting a small number of relevant covariates is desirable), boosting is usually superior to standard estimation techniques for regression models (such as backward stepwise linear regression, which, e.g., cannot be applied if p is larger than the number of observations n)". In (ROBINZONOV; TUTZ; HOTHORN, 2012) the method is used in models for forecasting the german industrial production and concludes that the boosting method performs better than auto-regressive models in that scenario. The study of (LEHMANN; WOHLRABE, 2016) and the one of (ZENG, 2014) use boosting to select variables and to perform forecasting at national level, and both conclude that the method is competitive. After realizing the lack of regional studies, the article by (LEHMANN; WOHLRABE, 2017) brings this same method when dealing with the forecast of the GDP of two states and one region of Germany. The study concludes that the boosted model brings more precise forecasts and gives relevance to the variable selecting, giving special attention to local variables. To implement the boosting algorithm we use the "mboost" package (HOTHORN et al., 2020) with the R language. To see in details the math and theory used behind the code, the paper (BÜHLMANN; HOTHORN, 2007) provides the background for the authors.

We seek to test the L_2 boosting validity under the circumstance of a regional series with uncertainty. A good algorithm should anticipate future changes of the time series by looking at associated variables. We compare our results with what already exists in the literature of electricity consumption and boosting forecasting. Calculating the relative importance of each variable, we find three important predictors groups: weather parameters which model seasonality, the lagged electricity consumption around Brazil which update recent changes of the consumption behavior and unemployment rates which indicate the recession. There is a change in some predictors as h increases, bringing possible ideas about how the decisions in electricity consumption change in different horizons. Boosting performance is highly superior in $h = 1$ compared to the SARIMA benchmark, however this superiority does not stand with larger values of h.

The paper is organized as follows: in Section 2 we present the data-set, followed by the exposition of the boosting and the discussion of our forecast approach. In Section 3 we discuss our results by performance measures and by most important predictors. We close this article with some concluding remarks in Section 4.

[1] The reasons that make the chosen method efficient and consistent is shown in (BÜHLMANN; YU, 2003a).

2 Data and Forecasting Strategy

2.1 Data

Between January 2002 and November 2017 the electric energy in Rio Grande do Sul was distributed by three companies: AES Sul (today named RGE SUL), CEEE and RGE. The three companies divided their product in seven categories: residential, commercial, industrial, rural, public services and others. Our target variable Electric Energy Consumption in Rio Grande do Sul consists of the sum of all companies categories of consumption. We collected this data from the former agency Fundação de Economia e Estatística (FEE).

From 2002 through to 2013, Brazil observed a steady average GDP growth of 3.7%. Compared to other countries, Brazil had a good performance under the 2008 financial recession. However, in 2014 there was a decrease in the GDP of two consecutive quarters of the year, making experts believe in a possible new crisis. 2014 still finished with a positive GDP growth value of 0.5%, but the economy contracted 3.5% in 2015 and 3.6% in 2016, resulting in a 11% GDP per capita loss in the period. In addition, Dilma Rouseff, ex-president of Brazil re-elected in 2014, suffered an impeachment. The former president was removed from office in August 2016, consolidating the Brazilian economic and political crisis.

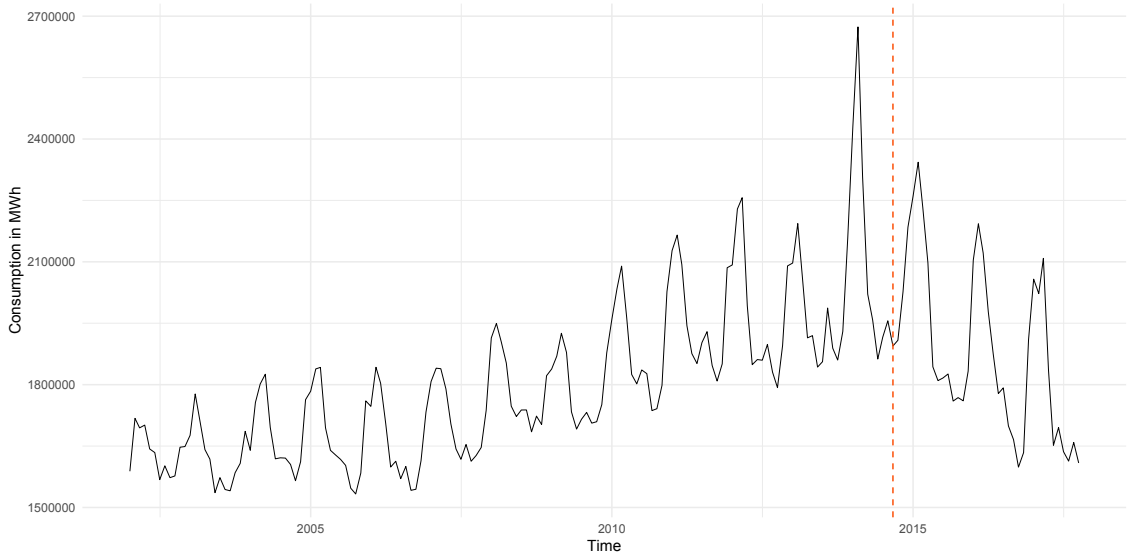
It is important to point out that in 2012 the Brazilian government created an agreement with the electricity distribution companies. The companies could anticipate the renewal of concessions without having to go through bids. In exchange, the companies would have to reduce the price charged for their service. Government control of the price resulted in increased consumption in the years following the agreement, however when electricity production costs rose in Brazil the distributors were unable to adjust prices. In 2015 the government had to change its conduct to ensure the profitability of the distributors and the price of electricity increased. The political uncertainty is a factor that adds difficulty for the forecasting process.

We chose September 2014 to be the starting date of the forecast exercise. The observations until August 2014 are used as the training set. Then, the algorithm is tested in the 38 observations from September 2014 to October 2017, coinciding with Brazilian economic recession. Figure 1 shows the electric energy consumption in Rio Grande do Sul in the period. The red line indicates the September 2014 split date.

The forecasting exercise is based on a monthly time series filled data set. We have a total of 822 parameters, which are grouped into four categories: international, national, meteorological and regional. The study from (KOPOIN; MORAN; PARÉ, 2013) shows that regional forecast with high-dimensional data sets containing national and international data help improve the medium-term forecasting performance. At international level, we collected data from Federal Reserve Economic Data (FRED) and Círculo de Estudos Latinoamericanos (CESLA) according to the main commercial partners of Rio Grande do Sul in agreement with (MDIC, 2019). Therefore from Chile, Peru, Argentina, United States, Russia, China, OECD and Euro Area.

At the national level, we selected the Ipeadata macroeconomics data set to describe Brazil in the period. 511 monthly predictors were chosen from the following

Figure 1 – Electricity Consumption in Rio Grande do Sul



topics: exchange rates, sales and consumption, employment, market expectations, financial market indicators, currency and credit indicators, prices, production, wages and income.

The electrical energy consumption is also affected by the behavior of the weather, as described in (ABDEL-AAL, 2008) and (GONZÁLEZ-ROMERA; JARAMILLO-MORÁN; CARMONA-FERNÁNDEZ, 2008). To model that, we gather monthly meteorological parameters of 12 cities around the state of Rio Grande do Sul. This parameters include variables such as average maximum temperature, average minimum temperature, relative humidity, average wind velocity, total precipitation and others. Because of the lack of data, some cities don't have all the weather parameters. All of the meteorological variables were collected from Instituto Nacional de Metereologia (INMET). For regional economic data, we collected indicators from Federação das Indústrias do estado do Rio Grande do Sul (FIERGS), Instituto Brasileiro de Geografia e Estatística (IBGE) and the former Fundação de Economia e Estatística (FEE). We present the division of the predictors in Table 1.

Table 1 – Breakdown of the Data Set

Variables	**Category**	**Source**
22	International	FRED
8	International	Cesla
511	National	Ipeadata
90	Regional	Fiergs
1	Regional	IBGE
91	Regional	FEE
99	Meteorological	INMET

Originally the data starts in January 2002 and runs through October 2017.

We have checked the stationarity of each series by using the Augmented Dickey–Fuller test (PFAFF, 2008). For those that were not stationary, we transformed the series either by logging it or by doing discrete growth[2]. Then the test was reapplied. Finally, for the series not yet stationary, we applied a difference and rechecked the test. Since we had to make only one difference in the non-stationary series, our stationary data starts with all predictors in February 2002. Our train set loses one observation, thus beginning in February 2002 and ending in August 2015, summing to 151 train observations, roughly 80% of the entire data set.

2.2 Boosting Algorithm

We are dealing with a high-dimensional data set. As exposed in (BÜHLMANN et al., 2006), the linear boosting method is consistent in such situations. In this section we will present the boosting model and how it was implemented in the "mboost" R package. Generally, boosting algorithms construct a linear or non linear model iteratively. We use the linear approach to our data. For that, we will estimate a function $\hat{f}(x_t) = \hat{y}_t$ that is a sum of M boosting components plus a constant:

$$\hat{f}\,(x_t) = \hat{f}^{(0)} + \nu \sum_{m=1}^{M} \hat{g}^{(m)}, \tag{1}$$

where x_t is a vector containing our N predictor variables. $\hat{f}(0)$ is a constant, v is a shrinkage parameter such that $0 < v < 1$ that reduces the learner variance and improve prediction performance (need citation). Finally, $\sum_{m=1}^{M} \hat{g}^{(m)}$ is the sum of $m = 1$ to M, where M can be chosen arbitrarily or by risk control measures (e.g. AIC or cross-validation). $\hat{g}^{(m)}$ is the estimated learner. That is, boosting adds its estimated learner which minimizes the sum of squared residual (SSR), within each m iteration. This will be explained in detail below.

The model begins with a temporary fitted value $\hat{f}_t^{(0)} = \bar{y}$, and then the first residual is calculated: $u_t^{(0)} = y_t - \hat{f}_t^{(0)}$. Now we introduce the set of potential predictors $z_{(k)}$, which contains p_y lags of the target variable y and p_x lags of all N potential exogenous parameters. In our case we are using one lag of the target variable y as well as one lag of all the 822 exogenous potential predictors, thus making the length of our $z_{(k)}$ to be 823. The algorithm proceeds by regressing the residual on each variable in in $z_{(k)}$ and selecting the one variable $z_{(k^*)}$ which minimizes the SSR. After that, the following temporary fitted value $\hat{f}_t^{(1)}$ is updated by the fraction v of the best regression made before ($\hat{f}_t^{(1)} = \hat{f}_t^{(0)} + v\hat{g}^{(m)}$ with $\hat{g}^{(m)} = \hat{\beta}_{(k^*)} z_{(k^*)}$), thus reducing the learner's variance, as it does not completely embrace $z_{(k^*)}$. This procedure happens until we have $\hat{f}_t^{(m)}$ with m = M:

1. Set $m = 0$ and begin with $\hat{f}_t^{(0)} = \bar{y}$.

2. From each iteration step $m = 1$ to M repeat:

- Calculate the residuals $u_t = y_t - \hat{f}_t^{(m-1)}$.

[2] Sometimes it was necessary to add a constant to each element before transforming.

- Regress the residuals u_t on each predictor $z_{(k)}$, with $k = 1, 2, ..., k, ..., p_y + p_x N$, and compute $SSR_{(k)} = \sum_{t=1}^{T}(u_t - z_{t,(k)}\hat{\beta}_{(k)})^2$.
- Select the predictor $z_{(k^*)}$ which has the smallest SRR.
- Set $\hat{g}^{(m)} = \hat{\beta}_{(k^*)} z_{(k^*)}$.
- Update $\hat{f}_t^{(m)} = \hat{f}_t^{(m-1)} + v\hat{g}^{(m)}$.

As declared in (BÜHLMANN; YU, 2003b) "L2-Boosting is nothing else than repeated least squares fitting of residuals". The algorithm manages to deal with a large set of predictors by exploring little by little their contribution. This way boosting is capable of selecting the best parameters in the pool of possibilities.

2.3 Forecasting

Here we describe the forecasting model. As mentioned before, we have restricted our analysis to only one-period lag, thus $p_x = p_y = 1$. Also recall $\hat{g}^{(m)} = \hat{\beta}_{(k^*)} z_{(k^*)}$ and z as the vector of all predictors. Let h be the time horizon and suppose we want to forecast the variable y_t. For example, when we define h = 1, we consider until z_{t-1} as predictors. In general, the equation have the following form:

$$\hat{y}_t = \bar{y} + v \sum_{m=1}^{M} \widehat{\beta}_{k_m^*} z_{k_m^*|t-h}, \tag{2}$$

$$y_t = \bar{y} + v \sum_{m=1}^{M} \widehat{\alpha}_{k_m^*} y_{k_m^*|t-h} + v \sum_{m=1}^{M} \widehat{\delta}_{k_m^*} x_{k_m^*|t-h} + \varepsilon_t, \tag{3}$$

$$\widehat{\alpha}_{k_m^*}, \widehat{\delta}_{k_m^*} \in \widehat{\beta}_{k_m^*}. \tag{4}$$

The values of v and M must be selected before the forecast exercise. How they are selected matters because this two key parameters influence the trade-off between variance and bias. A large value of v means in each iteration the algorithm is making larger steps, as a larger part of the selected parameter is being embraced. In the literature $v = 0.1$ is the standard choice, therefore we are going to use it here. In the case of M, a large value means the algorithm is going to have a lot of iterations, thus possibly over-fitting. On the other hand, a small value of iterations may not give enough time for boosting to produce a good fitting. To select M, we use the conventional strategy for modeling time series with linear models, which is the corrected Akaike Information Criteria (AICc)[3].

$$AICc(m) = \log(\hat{\sigma}_m^2) + \frac{1 + df(m)/T}{[1 - df(m) + 2]/T}, \tag{5}$$

[3] Cross-validation may also be used, however since the study case consists of a short time series, we choose to use AICc.

where $df(m)$ denote the degrees of freedom at m and $\hat{\sigma}^2_m$ the estimate of the residual variance of the additive model[4]. There are two approaches to calculate the AICc when dealing with linear boosting algorithms. The standard approach defines degrees of freedom by the trace of the boosting hat matrix, thus $df(m) = tr(\mathcal{B}_m)$, where $\mathcal{B}_m$ is the constructed hat matrix. However, as discussed in (HASTIE, 2007), this approach tends to overshoot the value of M. The suggestion given by the author is the second approach to calculate the AICc: computing degrees of freedom using the number of non-zero coefficients. We choose the second approach. Therefore, each time the algorithm produces a forecast to a new observation, M is selected by the AICc, where the degrees of freedom are computed as the sum of the non-zero coefficients.

To evaluate the boosting performance we include a SARIMA$(1,1,0)[1,1,0]_{12}$ model as benchmark, choosed via Box-Jenkins methodology (BOX; JENKINS, 1976). Auto-regressive models are a standard strategy when dealing with time series. The seasonal component comes from the fact the electricity consumption in Rio Grande do Sul varies seasonally (e.g. increasing in the summer), that can be seen in Figure 1. We use six measures of forecast accuracy, where the first one is the root mean squared forecast error (RMSFE). It is described as:

$$RMSFE_h^{Model} = \sqrt{T^{-1}\sum_{k=1}^{T}(FE_{t+h,k}^{Model})^2}, \tag{6}$$

where $FE_{t+h,k}^{Model} = y_{t+h,k} - \hat{y}_{t+h,k}^{Model}$ represents the h-step-ahead forecast error of the observation k. To make the comparison between SARIMA and boosting RMSFE performance we use the ratio between the two values (rRMSFE):

$$rRMSFE_h = \frac{RMSFE_h^{Boost}}{RMSFE_h^{SARIMA}}. \tag{7}$$

If the ratio is larger than one, $RMSFE_h^{Boost} > RMSFE_h^{SARIMA}$, thus the boosting model produces on average greater quadratic forecast errors. On the other hand, if $rRMSFE_h < 1$ we have the opposite. Since the electrical energy forecast literature conventionally uses mean absolute percentage error (MAPE) to measure performance, we shall embrace it. MAPE is defined as following:

$$\text{MAPE} = T^{-1}\sum_{k=1}^{T}\left|\frac{y_{t+h,k} - \hat{y}_{t+h,k}^{Model}}{y_{t+h,k}}\right| 100. \tag{8}$$

Besides MAPE, we calculate percentile 95 of the absolute percentage errors (P95), percentile 90 of the absolute percentage errors (P90) and maximum negative and positive errors (MNE and MPE). Finally, we verify the boosting accuracy superiority with the Diebold-Mariano test (DIEBOLD; MARIANO, 1995). For this reason, we go according to Diebold in (DIEBOLD, 2015), that the Diebold-Mariano

[4] For more details about the AICc calculation, we suggest the read of (BÜHLMANN; HOTHORN, 2007)

test is useful in pseudo-out-of-sample model comparisons such as historical episodes. With all these values, it is possible to have a sense of model behavior and quality.

3 Results

3.1 Performance Measures

In this section we present the results of the model and compare them with the benchmark. We use $v = 0.1$, a 0.80 ratio between training set and total set, and we select M through AICc. According to Table 2 and Table 3, the boosting MAPE value is better than the benchmark when h is 1 or 2[5]. Specifically in the case where $h = 1$, we see that the model outperforms the benchmark in all indexes. Considering the fact that we are dealing with an unstable series, the 3% MAPE shows that boosting can be considered a competitive electrical energy forecasting model. It is worth noting that we are using boosting to capture both trend and seasonality. We analyze the selected variables from the distinct time horizons in Section 3.2. The rRMSFE of 0.568 shows that we are closer to the true values compared to the benchmark. Worth to note the MPE of 7.649% compared to a MPE of 12.124% indicates that boosting is remarkably less tendentious than SARIMA. As for $h = 2$, boosting outperforms SARIMA only through MAPE.

Table 2 – Boosting Measures Table

	MAPE (%)	**P95 (%)**	**P90 (%)**	**MPE (%)**	**MNE (%)**	**rRMSFE**
$h = 1$	3.005	7.290	6.293	7.649	-8.307	0.568
$h = 2$	4.928	13.068	10.454	18.550	-12.975	1.229
$h = 3$	5.607	12.571	10.706	13.163	-13.455	1.173

Table 3 – Benchmark Measures Table

	MAPE (%)	**P95 (%)**	**P90 (%)**	**MPE (%)**	**MNE (%)**
$h = 1$	3.879	7.918	7.377	12.124	-8.508
$h = 2$	4.958	10.787	8.531	12.082	-11.824
$h = 3$	4.928	10.906	10.363	11.181	-15.245

The forecast plot is presented in Figure 3.1. By increasing h, both SARIMA and boosting loose predictive power. The main point to be seen here is, in the start of the recession, SARIMA model tends to overshoot the summer peak consumption. Meanwhile boosting is capable of recognizing a historical peak fall. The 2015 summer in Brazil is the first summer with recession, thus uncertainty hinders the forecast. Also, as mentioned before, electricity prices increased in this period, adding problems to the exercise. With $h = 3$ the SARIMA overshoot is greater than with $h = 1$, which may indicates a positive relation between forecasting horizon and first peak value error. On the other hand, boosting continues to maintain a conservative summer peak value as h increases. From the second

[5] When $h \geq 3$ the benchmark starts to perform better than our algorithm.

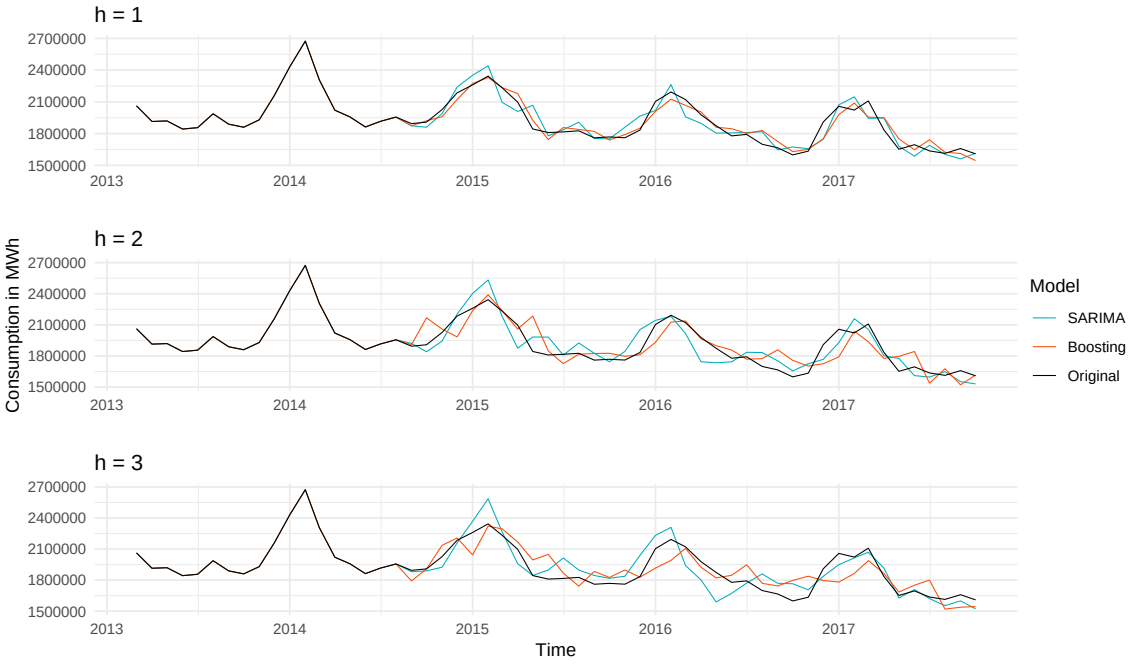
Table 4 – Diebold-Mariano Test: p-value

H1	**Superiority of boosting accuracy**	**Superiority of benchmark accuracy**
$h = 1$	0.026	0.974
$h = 2$	0.438	0.562
$h = 3$	0.717	0.283

to the third summer into the recession, both models show a similar performance. After some time, SARIMA understands the new trend. Boosting capacity to understand how the recession hit affects the electricity consumption comes from looking at associated variables. We look at what predictors boosting choose in the next section. Worth mentioning the capacity boosting have to recognize seasonal patterns decreases with a larger h value in our model. SARIMA does not have many mechanisms to forecast never seen dynamics, which raises doubts about its usefulness at uncertain times.

Finally, Table 4 brings the p-values of the Diebold-Mariano test with different alternative hypothesis[6]. The null hypothesis in each test is that the two models have the same forecast accuracy. For $h = 1$, the p-value rejects the null hypothesis within a 95% confidence level, thus supporting the superiority of boosting accuracy. For both $h = 2$ and $h = 3$, neither the boosting nor the benchmark are superior to each other when looking to the p-values. Thus, the Diebold-Mariano test only shows a significant accuracy superiority for boosting in $h = 1$.

Figure 2 – Forecast



[6] We used loss function power equals to 1, since the electrical energy forecasting literature commonly use MAPE, however the main results do not change using a loss function power of 2.

Table 5 – Top 10 Variables for Different Time Horizons

$h = 1$	
Average Maximum Temperature in Encruzilhada do Sul	15.14%
Electricity Consumption in the Central-West Region of Brazil	8.718%
Importation of Apparel Goods in Rio Grando do Sul	5.080%
Industry Electricity Consumption in Rio Grande do Sul by CEEE distributor	4.514%
Apparent Oil Product Consumption in Brazil	4.188%
Importation of Textile Goods in Rio Grande do Sul	3.969%
Apparent Consumption in the Metallurgy Sector in Brazil	3.436%
Physical Food Production Index in Brazil	3.063%
Open Unemployment Rate in São Paulo Metropolitan Region	2.215%
Hours Paid in Industry in the State of São Paulo	1.946%
$h = 2$	
Average Maximum Temperature in Encruzilhada do Sul	7.547%
Unemployment Rate in São Paulo Metropolitan Region	7.482%
Average Atmospheric Pressure in Bom Jesus	6.360%
Open Unemployment Rate in São Paulo Metropolitan Region	4.887%
Electricity Consumption in the Central-West Region of Brazil	4.337%
Wind Direction in Torres	4.095%
Industry Electricity Consumption in Rio Grande do Sul by CEEE distributor	4.031%
Importation of Apparel Goods in Rio Grando do Sul	2.583%
Rural Electricity Consumption in Rio Grande do Sul by CEEE distributor	2.299%
Apparent Oil Product Consumption in Brazil	2.114%
$h = 3$	
Rural Electricity Consumption in Rio Grande do Sul by CEEE distributor	13.71%
Unemployment Rate in São Paulo Metropolitan Region	5.189%
Average Maximum Temperature in Encruzilhada do Sul	5.019%
Average Atmospheric Pressure in Bom Jesus	4.341%
Open Unemployment Rate in São Paulo Metropolitan Region	4.213%
Electricity Consumption in the Central-West Region of Brazil	3.590%
Wind Direction in Torres	3.422%
Export of Other Transport Equipment from Rio Grande do Sul	1.762%
Importation of Apparel Goods in Rio Grando do Sul	1.738%
Industry Electricity Consumption in Rio Grande do Sul by CEEE distributor	1.586%

3.2 Variable Importance

In Table 5 we seek to understand which predictors matter more for the boosting algorithm. In Section 2.2 we showed boosting model works as a selector: in each iteration M the model choose the best parameter candidate to embrace. Here we discuss about the most important variables for the model. According to (HOTHORN et al., 2020), variable importance is calculated by a function that extracts risk reductions, AICc reduction in our case, per boosting iteration. It constructs a vector that accumulates the reduction contribution of each predictor contained in the model. This allows us to visualize and compare the different importance of the variables.

By looking at Table 5, one can assert weather variables have an important role in the forecast. They model the seasonality of the series. The Average Maximum Temperature in Encruzilhada do Sul is the most selected proxy for the season changes in Rio Grande do Sul, what makes sense since the city is located at the center of the state. Other important kind of parameter is the lagged electricity consumption. The electricity consumption in the Central-West Region of Brazil and the industrial and rural electricity consumption in Rio Grande do Sul appear often in the table. They serve to update the model the recent status of the electricity consumption in general.

However, the variables that are more capable of explaining the sudden change of direction of the series are the ones somehow related to the crisis. Boosting select the unemployment rates for this role. The Unemployment Rate in São Paulo Metropolitan Region parameter appears in the three horizons and is the second most important for $h = 2$ and $h = 3$. It may be argued that for larger values of h, the variables more important for the general decisions in the economy, such as the unemployment rate, have a greater importance. On the other hand, for $h = 1$ looking at the recent consumption of electricity, other energy parameters (e.g. oil consumption) and the production sector (e.g. metallurgy consumption and food production) may be more effective. In general, boosting succeeds in selecting parameters associated with the recession and use them to predict the falling trend in the series.

4 Concluding Remarks

Our goal with this research was to validate the use of boosting in an exceptional situation, where there is uncertainty and time series are short. To do this, we created a data set with regional variables as well as national and international variables to forecast monthly electrical energy consumption. It is also worth noting that we intended to test the boosting in an environment where there was no removal of seasonality. We found that boosting was able to capture seasonality in its own way as well as being a competitive model in $h = 1$. In addition, it was possible to bring the analysis of the results and understand which variables were key to plot the model behavior. We found three predictor clusters: the weather variables, the lagged electricity consumption around Brazil and the unemployment rates.

It is important to note the conclusions in the paper are partial, in the sense that future research is necessary to understand more how boosting deals with short unstable regional time series. In this paper we analysed one specific scenario. We suggest performance researches of the algorithm in other series affected by exogenous shocks. The emerging countries series are a good suggestion because the features desired are more usual in there.

Bibliography

ABDEL-AAL, RE. Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks. *Computers & Industrial Engineering*, Elsevier, v. 54, n. 4, p. 903–917, 2008.

ASSIS CABRAL, Joilson de; LEGEY, Luiz Fernando Loureiro; FREITAS CABRAL, Maria Viviana de. Electricity consumption forecasting in Brazil: A spatial econometrics approach. *Energy*, Elsevier, v. 126, p. 124–131, 2017.

BAI, Jushan; NG, Serena. Boosting diffusion indices. *Journal of Applied Econometrics*, Wiley Online Library, v. 24, n. 4, p. 607–629, 2009.

BOX, George.E.P.; JENKINS, Gwilym M. *Time Series Analysis: Forecasting and Control*. Holden-Day, 1976.

BÜHLMANN, Peter Lukas. Bagging, subagging and bragging for improving some prediction algorithms. In: SEMINAR für statistik, eidgenössische technische hochschule (eth), zürich. RESEARCH report/Seminar für Statistik, Eidgenössische Technische Hochschule (ETH). 2003. v. 113.

BÜHLMANN, Peter et al. Boosting for high-dimensional linear models. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 34, n. 2, p. 559–583, 2006.

BÜHLMANN, Peter; HOTHORN, Torsten. Boosting Algorithms: Regularization, Prediction and Model Fitting (with Discussion). *Statistical Science*, v. 22, n. 4, p. 477–505, 2007.

BÜHLMANN, Peter; YU, Bin. Boosting with the L 2 loss: regression and classification. *Journal of the American Statistical Association*, Taylor & Francis, v. 98, n. 462, p. 324–339, 2003.

_____. Boosting With the L2 Loss. *Journal of the American Statistical Association*, Taylor Francis, v. 98, n. 462, p. 324–339, 2003. DOI: 10.1198/016214503000125. eprint: https://doi.org/10.1198/016214503000125. Available from: <https://doi.org/10.1198/016214503000125>.

CHU, Jianghao et al. Boosting. In: MACROECONOMIC Forecasting in the Era of Big Data. Springer, 2020. P. 431–463.

DIEBOLD, Francis X. Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 33, n. 1, p. 1–1, 2015.

DIEBOLD, Francis X; MARIANO, Robert S. Comparing predictive accuracy. *Journal of Business & economic statistics*, American Statistical Association, v. 13, n. 3, p. 253–263, 1995. Available from: <http://www.jstor.org/stable/1392185>.

FREUND, Yoav. Boosting a weak learning algorithm by majority. *Information and computation*, Elsevier, v. 121, n. 2, p. 256–285, 1995.

FREUND, Yoav; SCHAPIRE, Robert E, et al. Experiments with a new boosting algorithm. In: CITESEER. ICML. 1996. v. 96, p. 148–156.

GONZÁLEZ-ROMERA, E; JARAMILLO-MORÁN, MA; CARMONA-FERNÁNDEZ, D. Monthly electric energy demand forecasting with neural networks and Fourier series. *Energy Conversion and Management*, Elsevier, v. 49, n. 11, p. 3135–3142, 2008.

HASTIE, Trevor. Comment: Boosting Algorithms: Regularization, Prediction and Model Fitting. *Statistical Science*, Institute of Mathematical Statistics, v. 22, n. 4, p. 513–515, 2007.

HOTHORN, Torsten et al. *mboost: Model-Based Boosting*. 2020. R package version 2.9-2. Available from: <https://CRAN.R-project.org/package=mboost>.

KOPOIN, Alexandre; MORAN, Kevin; PARÉ, Jean-Pierre. Forecasting regional GDP with factor models: How useful are national and international data? *Economics Letters*, Elsevier, v. 121, n. 2, p. 267–270, 2013.

LEHMANN, Robert; WOHLRABE, Klaus. Boosting and regional economic forecasting: the case of Germany. *Letters in Spatial and Resource Sciences*, Springer, v. 10, n. 2, p. 161–175, 2017.

_____. Looking into the black box of boosting: the case of Germany. *Applied Economics Letters*, Taylor & Francis, v. 23, n. 17, p. 1229–1233, 2016.

MDIC. Comex Vis, 2019. Available from: <http://comexstat.mdic.gov.br/pt/home>.

MEDEIROS, Marcelo C et al. Forecasting Inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics*, Taylor & Francis, p. 1–22, 2019.

OECD. Main Economic Indicators - complete database, 2015. DOI: https://doi.org/https://doi.org/10.1787/data-00052-en. Available from: <https://www.oecd-ilibrary.org/content/data/data-00052-en>.

PARAJULI, Ranjan et al. Energy consumption projection of Nepal: An econometric approach. *Renewable Energy*, Elsevier, v. 63, p. 432–444, 2014.

PFAFF, B. *Analysis of Integrated and Cointegrated Time Series with R*. Second. New York: Springer, 2008. ISBN 0-387-27960-1. Available from: <http://www.pfaffikus.de>.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2019. Available from: <https://www.R-project.org/>.

ROBINZONOV, Nikolay; TUTZ, Gerhard; HOTHORN, Torsten. Boosting techniques for nonlinear time series models. *AStA Advances in Statistical Analysis*, Springer, v. 96, n. 1, p. 99–122, 2012.

SCHAPIRE, Robert E. The strength of weak learnability. *Machine learning*, Springer, v. 5, n. 2, p. 197–227, 1990.

SCHMID, Matthias; HOTHORN, Torsten. Boosting additive models using component-wise P-splines. *Computational Statistics & Data Analysis*, Elsevier, v. 53, n. 2, p. 298–311, 2008.

VARIAN, Hal R. Big data: New tricks for econometrics. *Journal of Economic Perspectives*, v. 28, n. 2, p. 3–28, 2014.

ZENG, Jing. Forecasting Aggregates with Disaggregate Variables: Does boosting help to select the most informative predictors? Kiel und Hamburg: ZBW-Deutsche Zentralbibliothek für ..., 2014.