Simulating Brazilian Electricity Demand Under Climate Change Scenarios

Ian Michael Trotter, José Gustavo Féres, Torjus Folsland Bolkesjø and Lavínia Rocha de Hollanda

WP – 01/2015
Simulating Brazilian Electricity Demand Under Climate Change Scenarios

Ian M. Trotter∗ José Gustavo Féres† Torjus Folsland Bolkesjø‡
Lavínia Rocha de Hollanda§

4th September 2015

Abstract

Long-term load forecasts are important for planning the development of the electric power infrastructure. We present a methodology for simulating ensembles of daily long-term load forecasts for Brazil under climate change scenarios. For certain applications, it is important to choose an ensemble approach in order to estimate the (conditional) probability distribution of the load. High temporal resolution is necessary in order to preserve key features of the electricity demand that are particularly important in the face of increasing penetration of intermittent renewable power generation.

Keywords: long-term load forecast, climate change.

∗(Corresponding Author) Universidade Federal de Viçosa, Department of Agricultural Economics. E-mail: ian.trotter@ufv.br.
†Institute for Applied Economic Research (IPEA), Department of Macroeconomic Policies and Studies (DIMAC). E-mail: jose.feres@ipea.gov.br.
‡Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management. E-mail: torjus.bolkesjo@nmbu.no.
§Fundação Getulio Vargas, Centro de Estudos de Energia. E-mail: lavinia.hollanda@fgv.br.
1 Introduction

The electric power sector in Brazil is particularly exposed to weather risk, since hydroelectric power generation accounts for about 80% of total power generation (U.S. Energy Information Administration, 2013) and the electricity consumption for cooling purposes is significant and growing rapidly. The Brazilian electric power sector will therefore be particularly severely impacted by the changes that are expected to occur in the Earth’s climatic system over the next few decades (Intergovernmental Panel on Climate Change, 2013).

In this study, we create electric load simulations with high temporal resolution for the distant future. Such simulations can be useful for studying the properties of the electric load under a variety of future conditions, and are essential for the purposes of long-term generation capacity planning, especially in the presence of an increasing share of intermittent renewable power generation. We will use the approach to characterise the likely impacts of climate change on the distribution of the system load of the electric system in Brazil, which will be compared to results from other studies.

The connection between weather and consumption of electric power is well established, and weather variables have been extensively used for forecasting electricity demand for many decades already. Dryar (1944) noted that electric system load was influenced by weather conditions and radio programs of “unusual interest”. In the following decades, the use of weather variables for short-term load forecasting gained popularity and became important in the day-to-day operations of the power system and in production scheduling (Gillies et al., 1956; Davies, 1959; Clair and Einwechter, 1961; Heinemann et al., 1966; Matthewman and Nicholson, 1968a,b; Lijesen and Arvanitidis, 1970). Edwin and Elsayed (1980) showed how the errors of short-term load forecasts affected the operating costs and that higher accuracy of load forecasts could lead to more cost-efficient operation of the electric power system. Although primarily used in the daily operations of utilities, the relationship between weather conditions and electric system load also started receiving more attention in the context of long-term load forecasting and system planning (Stanton et al., 1969; Davey et al., 1973; Clayton et al., 1973; Asbury, 1975; Thompson, 1976; Rao

In the following period, timeseries methods with integrated weather variables gradually became common for short-term load forecasting, as they were able to account for inertia in the load response and complex serial correlation in the load (Christiaanse, 1971; Gupta and Yamada, 1972; Keyhani and Daniels, 1976; Keyhani and Rad, 1978; Phi et al., 1978; Abou-Hussien et al., 1981; Nakamura and Miyano, 1983; Bolzern and Fronza, 1986; Vemuri et al., 1986; Costin et al., 1987; Campo and Ruiz, 1987; Matthews et al., 1988). Barakat and Eissa (1989) studied the dynamic properties of electricity demand in a fast growing utility, and used a logistic model to represent a trend component with a saturation point. Timeseries approaches were also adapted to longer-term load forecasting, where the importance economic and demographic factors was also recognised and included (Uri, 1977, 1978a,b, 1979; Cardelli et al., 1980).

As personal computers became more widespread, they naturally became very popular in the field of load forecasting, and eventually became deeply integrated into the daily power system operations (Takenawa et al., 1980; Lajda and Reichert, 1981; Keyhani and Miri, 1983; Laing and Metcalfe, 1986; Jabbour et al., 1988; Wen and Jiang, 1988; Rahman and Baba, 1989). Multiple linear regression was originally the most common method, but with the increased availability of computational power, researchers started exploring the application of machine learning techniques to load forecasting (Shu-Ti, 1981; Nemeth and Nagy, 1981; Lajda, 1981; Dehdashti et al., 1982; Rahman and Bhatnagar, 1988; Lebby, 1990; Chen et al., 1991). Machine learning techniques, particularly neural networks applied to short-term load forecasting, gained tremendous popularity for load forecasting purposes in the following period. Bansal and Pandey (2005) performed a more thorough review of 265 papers on the topic, and research in this area has remained very active since the review was written. Although machine learning techniques usually perform well within the range of their training sample, they have been applied much less frequently to the problem of long-term load forecasting due to concerns that they may be poor for extrapolation. However, several studies have successfully applied machine learning techniques to long-term load forecasting (Kermanshahi, 1998; Kermanshahi and Iwamiya,
In addition to neural networks, researchers have also applied many other machine learning techniques to the problem, such as Kalman filters, data mining techniques, fuzzy logic, self-organizing maps and support vector machines (Al-Hamadi and Soliman, 2004; Liu and Lin, 2004; Wang et al., 2005; Bao et al., 2005; Fan et al., 2005; Al-Hamadi and Soliman, 2006; Pan et al., 2007; ul Asar and Amjad, 2008; Chang-chun and Min, 2008; Guo, 2009). Decomposition techniques, primarily based on wavelet theory and Fourier transforms, have also been widely applied to load forecasting (Kim et al., 2000; Zheng et al., 2000; Yu, 2000; Degaudenzi and Arizmendi, 2000; Huang and Yang, 2001; Zhang et al., 2007; Zheng et al., 2008; Truong et al., 2008; Bashir and El-Hawary, 2009; Kelo and Dudul, 2010; Chauhan and Hanmandlu, 2010; Pandey et al., 2010; Bahrani et al., 2014).

A number of studies have focused on predicting the load for “special events”, such as holidays, unusual weather patterns or extraordinary political events, for which the statistical models have traditionally performed badly due to low number of observations on which they can be calibrated (Lambert-Torres et al., 1991; Liu et al., 1992; Moharari and Debs, 1993; Park et al., 1993; Rahman et al., 1993; Srinivasan et al., 1995; Song et al., 2002; Yong, 2003; Ding et al., 2005; Qza et al., 2005; Li et al., 2006; Hor et al., 2008; Quansheng et al., 2009; Li and Gao, 2011; Akdemir and Çetinkaya, 2012; Wi et al., 2012; Li et al., 2013; Raza et al., 2014). In a particularly curious study, Yu et al. (2007) shows that text mining and sentiment analysis of news articles can be used to improve energy demand forecasts during abnormal events.

Some research has also been dedicated to distributions of load forecasts rather than simple point forecasts, mainly in order to more clearly understand the uncertainty of the forecasts, the effect of uncertainty in the parameters on the forecasts and the consequences of uncertainty in the forecasts (Adams et al., 1991; Mumford et al., 1991; Belzer and Kellogg, 1993; Ranaweera et al., 1995; Miyake et al., 1995; Ranaweera et al., 1996; Charytoniuk and Niebrzydowski, 1998; Douglas et al., 1998a,b; Charytoniuk et al., 1999; Zhengling and Kongyuan, 2002; Taylor and Buizza, 2002, 2003; Teisberg et al., 2005; Yang et al., 2007; Chaoyun and Ran, 2007; Hor et al., 2008; Hong et al., 2010; Fay and
Another modeling approach that has grown increasingly popular – in particular in the context of so-called *smart grids* – is attempting to create the load forecast bottom-up by modeling the end-use of the consumers (Noureddine et al., 1992; Bartels et al., 1992; Wu and Wong, 1993; Pratt et al., 1993; Harris and Liu, 1993; Rastogi and Roulet, 1994; Levi, 1994; Yan, 1998; Tao and Shen, 2006; Frąckowiak and Tomczykowski, 2007; Tiedemann, 2008; Beccali et al., 2008; Ali et al., 2011; Penya et al., 2011; Mao et al., 2011; Peng and Wei, 2011; Nejat and Mohsenian-Rad, 2012; Shen et al., 2012; Marinescu et al., 2013; Bacher et al., 2013; Sun et al., 2013; Hovgaard et al., 2013; Palchak et al., 2013; Bagnasco et al., 2014; Powell et al., 2014; Sandels et al., 2014; Horowitz et al., 2014; Lü et al., 2015). Some recent studies have also discussed how the proliferation of smart grid or microgrid technologies may influence consumer behaviour (Schachter and Mancarella, 2014; Jin et al., 2014; Mirowski et al., 2014).

Recognising that low-quality inputs can be detrimental to the forecast accuracy, some researchers focused on improving the accuracy of the inputs to the load forecasts. Khotanzad et al. (1996) and Seerig and Sagerschnig (2009) discuss how to upsample weather data in order to improve high-resolution load forecasts. Savelieva et al. (2000) have treated the problem of missing weather data and varying economic conditions. Challa et al. (2005) and Cao et al. (2012) discuss how to incorporate real-time weather data in the forecasting process, and Jiao et al. (2012) discusses strategies for constructing training samples for support vector machines with the purpose of forecasting electric system loads.

discussed the use of seasonal climate forecasts for medium-term electricity demand forecasting. Although one year forecasts may be sufficient for many operational tasks, such as production scheduling, it is insufficient capacity expansion planning – in particular at the time horizon in which climate change becomes important, several decades.

Since electricity demand is affected by weather conditions, it will naturally be influenced by changing climatic conditions. Mideksa and Kallbekken (2010) performed a review of the literature on the impacts of climate change on the electricity market, noting that most of the reviewed studies predict a net increase in electricity demand due higher temperatures and more use of air conditioning and further that more research was needed on the demand-side impacts in Latin America. Schaeffer et al. (2012) also summarise a number studies on the impact of climate change on energy demand. Furthermore, Zachariadis (2010) and Zachariadis and Hadjinicolaou (2014) studied the impacts of climate change on electricity use on Cyprus. Kaufmann et al. (2013) note that utilising hourly temperature and a flexible cut-off point for calculating heating degrees allows for improving the accuracy of models for electricity consumption.

In the specific case of Brazil, most of the published studies ignore the effects of weather on electric system load. Zebulum et al. (1995) and Carpinteiro et al. (2004) studied the use of neural networks for short- and medium-term load forecasting, respectively, for small regions of Brazil, although no weather data included. Several studies estimate the price- and income-elasticity of electricity in Brazil, and demonstrate how this may be used to forecast electricity demand, but do not consider the influence of weather conditions (Andrade and Lobão, 1997; Schmidt and Lima, 2004; Mattos and Lima, 2005; Camargo, 2007). Further studies have experimented with load forecasting based on various time-series methods for various regions of Brazil, but again weather conditions have not been included in the analysis (Soares and Souza, 2006; Soares and Medeiros, 2008; Castro and Montini, 2010; Neto et al., 2006; Viana and Silva, 2014). Siqueira et al. (2006) and Irffi et al. (2009) made forecasts for the North-East region of Brazil, but fail to take weather conditions into account. Campos (2008) studied several techniques for long-term load forecasting for the state of Minas Gerais, but did not consider weather conditions. On
the other hand, Schaeffer et al. (2008) conducted a comprehensive study of the impacts of climate change on the Brazilian electric energy sector, including long-term electricity demand projections. Caldas and Santos (2012) developed a simple multiple regression model for forecasting long-term residential consumption in the city of Petrolina, which included temperature. Migon and Alves (2013) experimented with a multivariate dynamic regression model on the hourly electricity loads of the Brazilian South-East region, considering trends, calendar effects, short-term dynamics and non-linear cooling effects. Using a dynamic panel approach, dos Anjos Rodrigues et al. (2014) also used weather variables to show that climate change is likely to cause a significant increase in residential electricity consumption in Brazil in the long term.

The strategic planning and expansion of the Brazilian electricity supply infrastructure is largely guided by the research conducted by Empresa de Pesquisa Energética (EPE). Some of the main reports created by EPE include an annual ten-year expansion plan (EPE-PDE, 2014), and an occasional longer-term plan – the latest of which is currently under development and will treat the period until 2050 (EPE-PNE, 2014). These reports are instrumental to the development of the Brazilian electricity supply infrastructure and naturally contain demand projections.

However, in the context of climate change, a somewhat longer planning horizon is required. The IPCC generally develops scenarios whose span reaches up to year 2100 and forecasts the most severe weather effects to occur beyond year 2050 (Intergovernmental Panel on Climate Change, 2013). Therefore, a time horizon up to about year 2100 appears more adequate for analysing climate risk. Traditionally, the reports published by EPE have also not explicitytly treated climate uncertainty, although the climate uncertainty and weather risk are of great interest the context of climate change.

This study will make several important contributions. Firstly, our research will contribute to the understanding of how weather conditions affect electricity demand in Brazil, which has received little attention in the existing literature. Secondly, realistic long-term load simulations with high temporal resolution for Brazil do not appear to have been created before, and are essential for long-term system planning. Thirdly, we will provide
a desperately needed updated assessment of the impacts of climate change on electricity demand in Brazil, using the latest available results and data. And finally, the chosen approach will enable us to discuss the probability distribution of the forecasts for electric system load, in contrast to the existing literature which only provides point forecasts.

Ferreira et al. (2015) calls for more research on the stochastic modelling of the Brazilian Electric Power Sector, and this study provides a probabilistic long-term electricity demand forecast through simulation.

This paper is structured as follows: the following section discusses and justifies the selection and treatment of input variables, the appropriate data sources and applicable calibration methods for building the load model. Section 3 presents the methodology framework, whereas section 4 present the load model and the results of calibrating the model on historical data, and compares some of the fitted model parameters with results from earlier studies. Section 5 presents the approach that will be used for performing load simulations for the future, as well as results and insights gained from performing these simulations. Finally, section 6 briefly summarises the main findings of this study.

2 Data Description and Analysis

2.1 Electricity Demand

Daily aggregated electricity demand on the Brazilian Interconnected Power System (SIN) is available from the Electric System National Operator (ONS), as far back as August 3, 2005 (ONS, 2015). Although data with a higher temporal resolution – say, hourly or half-hourly resolution – would have been preferable, only daily data is available to the public. In the ten years of daily aggregated electricity demand data between August 3, 2005 and August 2, 2015, there are only two missing observations. Figure 1 shows a graph of the daily load data.

Although there is some power generation and consumption which occurs outside of the interconnected grid, the SIN accounts for an overwhelming share of Brazilian electricity demand. Only 1.7% of electricity demand occurs outside this grid, mainly small isolated
systems in the Amazon region. Our study therefore only considers the electricity demand served by the SIN.

In principle, there may be some advantages to separating the demand between various sectors and modelling them individually, since for example the residential, commercial and industrial sectors can have very different demand patterns. In some cases, it might also be interesting to divide even further, for example distinguish between various industries (steel and pulp, for example) or different classes of residential customers (for example based on income). We do not, however, have access to daily data at this level of detail. Essentially, this means that our modelling approach will be in large part a top-down approach where we focus on aggregate behaviour, rather than a bottom-up approach where one would concentrate on smaller units. Although daily demand data is available for four geographically separated subsystems of the SIN, our main focus remains on the daily total demand of the system. It might be possible to improve on our results by modelling each geographical region separately, although our approach – simply modelling the aggregate – hopefully benefits somewhat from the law of large numbers.
2.2 Weather Data

The main purpose of this study is to determine the behaviour of electricity demand under changing climatic conditions. Therefore, it is important that we first and foremost consider the impact of weather variables on the electricity demand.

The Integrated Surface Database (ISD) from the National Oceanic and Atmospheric Administration (NOAA) contains quality-controlled weather observations from 490 weather stations on Brazilian territory (NOAA, 2015). However, only 28 of these weather stations have less than 10% missing hourly observations in the ten-year period between August 3, 2005 and August 2, 2015. We initially focus on this subset of weather stations.

For simplicity, we consider only the temperature rather than a wider set of weather variables. Although humidity, cloud coverage and wind most likely also affect electricity demand, temperature is generally considered to be by far the most important single weather variable.

Since the temperature observations contain missing values, we replace the missing values with artificial values imputed by the procedure described by Josse and Husson (2011). The temperature observations contain few missing values in addition to a substantial amount of both temporal and geographical regularity, which is why we believe that a PCA-based imputation method is well suited for our purposes.

We assume that when average daily temperature (that is, daily maximum temperature plus daily minimum temperature, divided by two) is below a certain base temperature, $T_H$, then electricity demand increases for heating purposes. When the temperature is above a certain base temperature, $T_C$, electricity demand increases for cooling purposes, and when the temperature is above an even higher base temperature $T_{CA}$, then electricity demand for cooling purposes increases at a higher rate. Altogether, this means that we estimate a piecewise linear temperature response curve with knots at $T_H$, $T_C$ and $T_{CA}$, and which is flat between $T_H$ and $T_C$. The base temperatures are chosen by simply scanning over the base temperatures and selecting the base temperatures that minimise an out-of-model error measure, that is, we selected the base temperatures that appeared to provide the best forecasting performance. This procedure lead to a base temperature for heating.
at 18°C, a base temperature for cooling at 25°C, and a base temperature for accelerated cooling at 28°C. Each of these three base temperatures will give rise to a linear term in the model.

In order to select a smaller subset of stations with the highest impact on electricity demand, we employ a variation of the procedure suggested by Hong et al. (2015) for selecting weather stations. The framework proposed by Hong et al. (2015) first ranks individual weather stations based on a goodness-of-fit measure when including each station individually in a simple model, then selects a composite weather variable containing the highest ranking weather stations such that a given model error measure is minimised. We depart slightly from the proposed framework by avoiding the construction of a composite weather variable and instead including all the highest ranking weather variables individually in the model. This procedure results in a selection of 18 weather terms pertaining to 13 different weather stations.

For the selected weather variables, autoregressive and lagged terms are also included in order to account for inertia (that is, serial correlation) and accumulative weather effects, as discussed by Li et al. (2009). By scanning over various lagged and autoregressive terms and calculating out-of-model error measures, we chose to include the 9-day moving average of the weather variables (average of ten days earlier up until and including yesterday) and the value of yesterday’s weather variable.

Since this can give rise to a large number of weather-related variables in the model, we also perform a backwards stepwise regression at the end using the Bayesian Information Criterion to select a more parsimonious model (Hyndman and Athanasopoulos, 2014). After this pruning, only nine weather variables from eight weather stations is included in the model: heating degree days from a single weather station (number of degrees below 18°C), accelerated cooling degree days from a single weather station (number of degrees above 28°C), and regular cooling degree days from seven weather stations (number of degrees above 25°C).

Figure 2 shows the location of the final eight weather stations used in the model, plotted on a population density map of Brazil. Despite the appearance of being some-
what unbalanced, the resulting selection of weather stations appears reasonable: a larger number of weather stations were selected in the south of the country where population density is much higher, the seasonal weather variations are much greater, and the population in general is more affluent. These three reasons would contribute to making the electricity demand in the southern part of Brazil more sensitive to weather, and would explain why the selection procedure would choose a greater number of weather stations from the southern parts of the country.

2.3 Demographic and Economic Data

Electricity demand is known to be influenced by demographic conditions, such as the size and structure of the population, and economic factors, such as industrial production. Since the economic and demographic conditions can change substantially in the period in consideration, it is important to explicitly include such factors in the analysis. Our
efforts are, however, restricted by the availability of data: not only with respect to the historical information available for calibrating a model, but also with respect to forecasts for the main drivers. It makes little sense to calibrate a model using factors for which we will not be able to obtain credible forecasts for the period of interest. We therefore only consider the population number and the gross domestic product, variables whose impact on energy demand are relatively well understood and for which scenarios can be obtained with relative ease.

The Brazilian Institute of Geography and Statistics (IBGE) provides both historical GDP figures and population numbers. We use the quarterly GDP figures (that is, not seasonally adjusted), and copy the quarterly value to each day in the quarter (IBGE, 2015). For the population numbers, we choose to use projected population numbers rather than numbers directly from the censuses, since the numbers from the censuses appear to contain numerous statistical artifacts (IBGE, 2013). The projected population numbers estimate the Brazilian population on the first day of each month, and we obtain a value for every day simply by linear interpolation. The GDP figures and the projected population number are shown in figures 3 and 4, respectively.

---

**Figure 3:** Quarterly GDP in 1995 prices, seasonally unadjusted
Source: The Brazilian Institute of Geography and Statistics (IBGE, 2015)
2.4 Calendar Effects

The pattern of electricity demand changes between weekdays, holidays and throughout the year. Since we are interested in constructing daily resolution forecasts, we must take into account recurring factors that affect the electricity demand on specific days. We include these so-called calendar effects mainly as dummy variables in the model. We include a dummy variable for each day of the week, for major national and regional holidays, and for bridge days. In some cases, we also include a number of dummies that distinguish between whether a holiday occurs on a weekday or during the weekend, since a holiday that occurs on a weekday usually impacts demand differently than a holiday that occurs during a weekend. Table 1 shows an overview of the dummy variables used to capture calendar effects that remained after pruning the model using a backwards stepwise regression procedure employing the Bayesian Information Criterion to select a parsimonious set of predictors.

In addition to dummies for capturing weather variables, we include the number of daylight hours per day in the model. The number of daylight hours presumably affects electricity demand by directly affecting usage of electrical lighting, but may also capture
Table 1: Dummy variables used for capturing calendar effects

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday, Saturday, Sunday</td>
<td>Day-of-the-week effects</td>
</tr>
</tbody>
</table>

**Fixed dates**

- January 1: New Year’s Day
- January 2: Knock-on effects from New Year’s Day
- First week of January: Knock-on-effects from earlier holidays
- April 21: Tiradente’s Day
- May 1: Labour Day
- May 2: Knock-on effects from Labour Day
- September 7: Independence Day
- September 8: Knock-on effects from Independence Day
- October 12: Our Lady of Aparecida
- November 2: All Soul’s Day
- November 15: Republic Day
- November 16: Knock-on effects from Republic Day
- December 23: Two days before Christmas
- December 24: Christmas Eve
- December 25: Christmas Day
- December 31: New Year’s Eve

**Fixed dates combined conditions**

- January 1 and Sunday: Moderates the New Year’s Day effect if it falls on a Sunday
- April 20 and Monday: Bridge day for Tiradente’s Day
- April 21 and weekend: Moderates the effect of Tiradente’s Day if it falls on a weekend
- May 1 and weekend: Moderates the effect of Labour Day if it falls on a weekend
- July 9 and not Sunday: São Paulo State Day, although the effect is indistinguishable when it falls on a Sunday
- September 6 and Monday: Bridge day for Independence Day
- September 7 and weekend: Moderates the effect of Independence Day when it falls on a weekend
- October 11 and Monday: Bridge day for Our Lady of Aparecida
- October 12 and weekend: Moderates the effect of Our Lady of Aparecida when it falls on a weekend
- November 1 and Monday: Bridge day for All Soul’s Day
- November 2 and weekend: Moderates the effect of All Soul’s Day when it falls on a weekend
- November 15 and weekend: Moderates the effect of Republic Day when it falls on a weekend
- November 20 and weekday: Zumbi dos Palmares Day

**Moving Dates**

- Saturday-Tuesday before Ash Wednesday: Carneval
- Ash Wednesday: Religious holiday
- Good Friday: Religious holiday
- Easter Sunday: Religious holiday
- Corpus Christi: Religious holiday
- Friday after Corpus Christi: Bridge day for Corpus Christi

Source: Own elaboration
more profound influences in daylight variations and changes in the diurnal routine of consumers over the year, as well as having some correlation with electricity demand for heating and cooling purposes. Since Brazil is a very large country, latitudinally speaking, there are large regional differences in how the number of daylight hours varies over the year: there is hardly any seasonal variation in daylight hours close to the equator in the north, whereas the extreme south experiences significant change in the number of daylight hours over the year. Therefore, we initially include the number of daylight hours from four locations, relatively evenly spread throughout the latitudinal extent of the country: Porto Alegre, São Paulo, Salvador and Fortaleza. The location of these cities is shown in figure 5.

**Figure 5:** Location of the cities whose number of daylight hours were considered in the model
Source: Own elaboration
2.5 Special events

Special rare or irregular events can also affect electricity consumption. Although it would be difficult to obtain forecasts for these events in the remote future, it is necessary to include known historical events for the calibration of the model. When creating forecasts with the model, there is some comfort in knowing that these types of events normally infrequent. In this respect, in our demand model we include a dummy variable to mark the weekdays on which the Brazilian national team played matches in the FIFA World Cups in 2006, 2010 and 2014.

2.6 Error Term

The error of the model is likely to contain residual amounts of serial correlation. We model the error term as an ARMA process, and use a Ljung-Box test to show that the residuals from the ARMA process are uncorrelated. Modelling the error term thusly will help us create probabilistic forecasts that recreate the serial correlation characteristics of the original series during the simulation step.

2.7 Intentionally Omitted Drivers

Electricity demand is affected by a greater number of factors than those mentioned until now. Although the following factors have not been included in the model for practical reasons, they have not been forgotten.

Technological change, for instance, is an important determinant of electricity demand. However, new technology may both increase and decrease electricity demand through creating new ways of using electricity or replacing current uses of electricity. Also the proliferation of important existing technologies, such as air conditioning, is something that could have been explicitly included in the model, especially considering that many technologies in Brazil are far from their saturation point. Some of these developments, however, are somewhat accounted for by including the GDP.

Intermittent renewable electricity generation, such as solar or wind power, is frequently modelled as “negative demand”, since electricity from these sources could substitute elec-
tricity from the national grid with local sources of electricity. Although this might increase in importance in the future, we do not include this in the model. Although these are weather-dependent, thus apparently of great interest in this study, we consider these as sources of supply in this context, rather than sources of “negative demand”.

Perhaps the most ominous omission is the price of electricity. Firstly, however, several studies on the household sector in Brazil assert that at least price elasticity in that particular sector is low, much lower than income elasticity (Andrade and Lobão, 1997; Schmidt and Lima, 2004; Mattos and Lima, 2005). Secondly, demand is a crucial component to price formation, which means that it would be necessary to forecast demand in order to forecast price, but we would also need a price forecast to construct a demand forecast. As such, this would lead to a circular problem, or at best a simultaneous problem, which is beyond the scope of this study.

These factors have been omitted partially in an effort to keep the assumptions necessary for simulation step as simple as possible: including even a small number of additional factors can dramatically increase the number of possible scenarios that need to be considered, as well as create consistency problems, since creating scenarios that accurately reflect the appropriate conditional probabilities can be exceptionally complicated. In many cases, there is also little current or historical information to base forecasts of the factors on.

The factors that have been omitted are also presumably of lower importance than the variables already included. If the selection of predictors already included provides sufficiently pleasing results, there is little incentive to complicate the model further. At best, including these factors would provide marginal benefits in return for a substantial increase in complexity, typically showing diminishing returns in response to the effort.

3 Methodology Framework

The methodology follows the same three basic steps as Hyndman and Fan (2010): first a statistical demand model is calibrated using the available historical data described above, then the demand model is used to forecast future demand by combining temperature and
residual simulation with future assumed demographic and economic scenarios, and finally the results and the forecasting performance will be evaluated critically.

The statistical model of demand is constructed using multiple linear regression. Although more complex approaches have been explored in the literature, such as machine learning, Hong et al. (2014a) argue that multiple linear regression approaches often prove superior to machine learning approaches, in addition to being considerably simpler to operationalise. The resulting models are also more defensible, in the sense that the influence of each factor is directly and explicitly quantified, which is not the case for most machine learning based approaches. The weather variables of the model are selected using a slightly modified version of the approach developed by Hong et al. (2015). Final predictor selection is performed by backwards stepwise regression, using the Bayesian Information Criterion (BIC) in order to indicate a preference for parsimonious models (Hyndman and Athanasopoulos, 2014).

Future demand scenarios are constructed through repeatedly creating point forecasts with the calibrated model by applying it to simulated residuals, simulated future temperatures, and assumed future economic and demographic scenarios. After choosing the correct timeseries model for the residual, residual simulation can be performed by a simple bootstrapping method. Future temperatures are provided by global circulation models MIROC5 and HadGem3 that simulate the climate. However, the climate simulations do not provide a sufficient number of weather scenarios for assessing the entire probability distribution, and because of this we will apply a bootstrapping method to the climate simulation results in order to generate a much greater number of realistic weather scenarios. For assumed future values of demographic and economic drivers, we will use selected scenarios from the Shared Socioeconomic Pathways (SSP) database (O’Neill et al., 2014a).

In the final step, the forecasting performance is evaluated and the forecasts are compared to forecasts provided in the existing literature. An idea of the performance can be obtained by checking the forecasting performance of the calibrated model on an out-of-sample validation set, paying close attention to the differences between the ex-ante and ex-post forecasts. We also compare results with electricity demand forecasts developed in
other reports, such as the Brazilian official ten-year plan (EPE-PDE, 2014), the Brazilian official long-term plan (EPE-PNE, 2014), and the demand projections developed by Schaeffer et al. (2008).

4 Model Establishment

4.1 Model Description

The model of the daily electric system load takes the following form:

\[
\ln(L_t) = \sum_{i=1}^{p} \alpha_i \text{HDD}_t^{TH} + \sum_{i=1}^{q} \beta_i \text{CDD}_t^{TC} + \sum_{i=1}^{r} \gamma_i \text{CDD}_t^{TCA} + \theta_1 \ln(\text{POP}_t) + \theta_2 \ln(\text{GDP}_t) \\
+ \sum_{i=1}^{n} \kappa_i \text{CAL}_t^{i} \sum_{i=1}^{s} \lambda_i \text{DH}_t^{i} + \eta_t
\]

where

- \( L_t \) denotes the electric system load on day \( t \);
- \( \text{HDD}_t^{TH} \) denotes how many degrees below the base temperature \( T_H \) the daily average temperature (that is, daily maximum plus daily minimum divided by two) is at weather station \( i \in \{1, \ldots, p\} \) on day \( t \);
- \( \text{CDD}_t^{TC} \) denotes by how many degrees the daily average temperature at weather station \( i \in \{1, \ldots, q\} \) exceeds the base temperature \( T_C \) on day \( t \);
- \( \text{CDD}_t^{TCA} \) denotes by how many degrees the daily average temperature at weather station \( i \in \{1, \ldots, r\} \) exceeds the base temperature \( T_{CA} \) on day \( t \);
- \( \text{POP}_t \) denotes the estimated population on day \( t \);
- \( \text{GDP}_t \) denotes the gross domestic product in the quarter to which day \( t \) belongs;
- CAL_{t,i} is a set of dummies marking calendar effects and the occurrence of special events;
- DH_{t,i} is the number of hours of daylight on day \( t \) at location \( i \in \{1, \ldots, s\} \);
- \( \eta_t \) denotes the model error on day \( t \), which can be serially correlated and is modelled by an ARMA process.

The model is formulated in terms of the natural logarithm of demand, as such the effects of weather and calendar dummies are multiplicative in this model. The demographic and economic variables, population and gross domestic product, also appear as natural logarithms, so they are modelled as having a constant elasticity rather than a constant marginal effect. In addition to being a reasonable modelling choice (a change in the GDP of one percent increases the electricity demand by a determined percentage), an advantage of this choice is that the elasticities can be easily compared to estimates of elasticities found in the existing literature.

### 4.2 Variable Selection

The set of weather variables was chosen using a method similar to the one developed by Hong et al. (2015). The final set of predictors were selected using backwards stepwise regression (Hyndman and Athanasopoulos, 2014).

### 4.3 Model Fitting

We divided our sample into two parts: a training sample and a test sample. The model parameters are calibrated using the training sample, and an ex-post forecast on the validation sample gives an impression of how the model performs for forecasting purposes, disconsidering uncertainty in the predictor variables.

The results of the model calibration are shown in appendix A. The mean absolute percentage error (MAPE) of the training sample is 1.64\%, whereas the MAPE of the ex-post forecast (i.e. given the observed historical value of all the predictors) on the test sample is 1.93\%. To give an idea of the fit of the model and the forecasting performance,
Figure 6: In-sample forecast, January 2014
Source: Own elaboration, Brazilian Independent System Operator (ONS, 2015)

gure 6 and 7 show an in-sample forecast for January 2014 and an ex-post out-of-sample forecast for January 2015, respectively.

The estimated coefficient of the term $\ln(GDP_t)$, $\theta_2 = 0.475$, corresponds to the income elasticity of electricity demand. Earlier studies have estimated the income elasticity of electricity demand for various sectors or geographical regions of Brazil, but no study has reported the elasticity of the aggregate electricity demand. Even so, the value obtained here conforms well to those found in earlier studies: Andrade and Lobão (1997) estimated the income elasticity of the residential sector in Brazil at 0.2132, Schmidt and Lima (2004) estimated the income elasticities of the residential, commercial and industrial sectors to 0.539, 0.636 and 1.92, respectively, and Mattos and Lima (2005) placed the income elasticity of the residential sector in Minas Gerais at 0.532. Although our estimated income elasticity differs from earlier attempts, it appears to be well within reason despite the radically different approach of the present study.
5 Forecasting and Evaluation

5.1 Forecasting Procedure

Repeatedly calculate the daily load from January 1, 2016 through December 31, 2100, using the calibrated model together with:

- Assumed values for population and GDP;
- Temperature data from weather simulations;
- Simulated residual from a SARIMA model calibrated on the residuals from the load model.

Here we explain in greater detail how the different input variables are treated.

5.1.1 Assumed Population and GDP: Shared Socioeconomic Pathways

The Shared Socioeconomic Pathways are a set of five demographic and socioeconomic scenarios developed to serve as a common starting point for climate change researchers (Van Vuuren et al., 2014; O’Neill et al., 2014b). Each of the five scenarios is accompanied
by a narrated storyline (O’Neill et al., 2012), and several teams of researchers have performed simulations to quantify key economic and demographic variables for each of the scenarios (IIASA, 2015).

For our work, we have selected three of the five scenarios:

**SSP1:** Global Sustainable Development

**SSP2:** Business As Usual

**SSP5:** Conventional Development/Economic Optimism

We have chosen to use the quantifications that have been created by the OECD and are available in the SSP Database, since these are considered illustrative (IIASA, 2015). Since the SSP quantifications only provide data for one year each decade until year 2100, we perform an exponential interpolation (that is, a linear interpolation in log-space) on the population data to generate daily figures necessary for the model. We perform an exponential interpolation on the GDP data to generate annual GDP figures, to which we apply a simple seasonal profile in-line with historical observations (Q1: 23.98%, Q2: 24.91%, Q3: 25.70%, Q4: 25.40%). The population series based on the SSP is then adjusted by a constant factor such that the population on the first of January 2015 matches the observed data, and the GDP series based on the SSP is similarly adjusted by a constant factor such that it matches the observed GDP for 2014. Figure 8 shows the resulting population of Brazil for the three chosen scenarios, and figure 9 shows the annual GDP used in the three scenarios.

### 5.1.2 Weather Simulation

The IPCC has chosen four representative global climate pathways, called Representative Concentration Pathways (RCP), each of which is largely defined by the level of radiative forcing reached in year 2100: 2.6 W/m$^2$, 4.5 W/m$^2$, 6 W/m$^2$ and 8 W/m$^2$ (Van Vuuren et al., 2011). These pathways have been selected in order to facilitate the comparison between climate simulations performed by different researchers.
**Figure 8:** Population of the Shared Socioeconomic Pathway scenarios  

**Figure 9:** GDP of the Shared Socioeconomic Pathway scenarios  
The National Aeronautics and Space Administration (NASA) has created global daily downscaled climate projections for two of the RPCs, RCP4.5 and RCP8.5, using 21 different climate models (NASA-GDDP, 2015). With a spatial resolution of 0.25 degrees, these projections are considered appropriate for climate change impact studies on local and regional scales. From the NASA-GDDP dataset, we obtained the daily maximum and minimum temperature projections for each weather station used in our calibrated model from the MIROC5 global circulation model for the two available RCPs (Watanabe et al., 2010).

This temperature data is then used to simulate a large number of weather paths, which enables us to estimate the impact of weather uncertainty on the electricity demand. The simulated weather paths are created using the following procedure:

1. A trend spline and a seasonal spline are fitted to the data for each variable from each weather station, as illustrated for a single weather variable in figures 10 and 11;

2. A trend spline and a seasonal spline are fitted to the annual and daily standard deviations of the detrended and deseasonalised temperature data for each variable from each weather station, as illustrated for a single weather variable in figures 12 and 13;

3. A normalised, detrended and deseasonalised residual is created for each variable from each weather station by subtracting the trend and seasonal splines, then dividing by the trend and seasonal splines that were fitted to the standard deviation, as illustrated for a single weather variable in figure 14;

4. The normalised residual is resampled (with replacement) in blocks containing a random number of days between 180 and 545, and containing all weather variables. This procedure is intended to capture trends in the weather variables, as well as preserve the temporal and spatial correlation present in the dataset. The last step of the procedure is repeated 500 times in order to generate 500 weather simulations for each of the two selected RCPs. Figures 15 and 16 illustrate the results of this procedure for the maximum
**Figure 10:** Maximum daily temperature from a single weather station (820980), and a spline fitted to the data
Source: Own elaboration, National Aeronautics and Space Administration (NASA-GDDP, 2015)

**Figure 11:** Detrended maximum daily temperature from a single weather station (820980) at fractions of the year, and a spline fitted to the data
Source: Own elaboration
Figure 12: Standard deviation of the detrended and deseasonalised maximum daily temperature from a single weather station (820980), and a spline fitted to the data. Source: Own elaboration.

Figure 13: Standard deviation of the detrended and deseasonalised maximum daily temperature from a single weather station (820980) throughout the year, and a spline fitted to the data. Source: Own elaboration.
daily temperature of a single weather station, for RCP4.5 and RCP8.5 respectively. The figures show the mean annual temperature of the 500 simulations as a thick black line, and the 25-75 and 10-90 percentiles of the temperature simulations as shaded areas, together with the annual mean temperature in the original scenario as a dotted line.

The temperature simulations appear to reflect important characteristics of the original scenario quite well: the mean of the simulations closely follows the overall trend of the original scenario, and the percentiles of the simulations appear reasonable when compared to the original. The block resampling step ensures that temporal and spatial correlation is preserved.

5.1.3 Residual Simulation

In order to capture serial correlation that was not captured by the load model, we pair each simulation of the weather variables with a simulation of the error term. The simulation of the error term is created by calibrating a SARIMA model on the residual from the load model in the calibration period August 2005 to August 2014. The lag structure is
Figure 15: Annual mean of the daily maximum temperature of the weather scenarios, based on the MIROC5 global circulation model run with the representative concentration pathway 4.5
Source: Own elaboration, National Aeronautics and Space Administration (NASA-GDDP, 2015)
Figure 16: Annual mean of the daily maximum temperature of the weather scenarios, based on the MIROC5 global circulation model run with the representative concentration pathway 8.5.
Source: Own elaboration, National Aeronautics and Space Administration (NASA-GDDP, 2015)
determined by the Akaika Criterion, which indicates that a SARIMA\((3,1,4)(1,0,2)\) model is appropriate. Figure 17 and 18 show the autocorrelation function and partial autocorrelation function calculated at different lags on the residuals of the SARIMA\((3,1,4)(1,0,2)\) model. The plots indicate that the residual of this process behaves almost like white noise, thus we are satisfied that this process is sufficiently well-suited for our purposes. We generate 500 simulations using this process, each of which is paired with a simulation of the weather variables.

5.2 Forecasting Results

We consider three main scenarios: SSP1 paired with RCP4.5, SSP2 paired with RCP4.5, and SSP5 paired with RCP8.5. Although no official recommendations on the pairing of RCPs and SSPs is currently available, the pairings of socioeconomic pathways and representative concentration pathways used in this paper are considered to represent plausible combinations, and have been suggested as possible reference scenarios (Van Vuuren et al., 2014).

For each of the three scenarios, load forecasts from 2016 to 2100 were created by
calculating the load model on the 500 weather simulations created for the respective RCP and the assumed values for population and GDP given by the selected SSP, and then adding a simulated error term. The distribution of the estimated annual demand for the three scenarios are illustrated in figure 19, figure 20 and figure 21. A summary of the distribution of estimated annual demand for key years is also given in table 2, and a graphical comparison between the mean of the simulations for each of the three scenarios is given in figure 22.

The first scenario, RCP4.5 paired with SSP1 as shown in figure 19, shows rising annual electricity demand until approximately year 2060, which rapidly drops again by year 2100 to a level only slightly above today’s level. This trajectory clearly reflects the assumed low population growth, combined with the relatively modest increase in GDP and temperature over the period. The wide range between the percentiles of the distribution shows the tremendous impact of weather uncertainties on the distribution.

The second scenario, RCP4.5 paired with SSP2 as shown in figure 20, shows an increase in annual electricity demand until year 2060, and a very slow decline from then until year 2100. This is the population with the highest population growth and the lowest
Table 2: The mean and quartiles of the estimated probability distribution of annual electricity demand (TWh) for selected years in the three scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Year</th>
<th>Mean</th>
<th>1st Quartile</th>
<th>2nd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSP1-RCP4.5</td>
<td>2020</td>
<td>626.19</td>
<td>606.32</td>
<td>642.12</td>
</tr>
<tr>
<td></td>
<td>2030</td>
<td>789.34</td>
<td>752.37</td>
<td>827.05</td>
</tr>
<tr>
<td></td>
<td>2040</td>
<td>939.92</td>
<td>876.11</td>
<td>995.58</td>
</tr>
<tr>
<td></td>
<td>2050</td>
<td>1036.97</td>
<td>952.03</td>
<td>1111.22</td>
</tr>
<tr>
<td></td>
<td>2060</td>
<td>1071.18</td>
<td>971.65</td>
<td>1160.23</td>
</tr>
<tr>
<td></td>
<td>2070</td>
<td>1028.56</td>
<td>928.95</td>
<td>1122.11</td>
</tr>
<tr>
<td></td>
<td>2080</td>
<td>936.04</td>
<td>833.80</td>
<td>1027.20</td>
</tr>
<tr>
<td></td>
<td>2090</td>
<td>799.39</td>
<td>698.15</td>
<td>880.53</td>
</tr>
<tr>
<td></td>
<td>2099</td>
<td>657.84</td>
<td>569.06</td>
<td>734.49</td>
</tr>
<tr>
<td>SSP2-RCP4.5</td>
<td>2020</td>
<td>634.80</td>
<td>614.68</td>
<td>650.96</td>
</tr>
<tr>
<td></td>
<td>2030</td>
<td>803.61</td>
<td>765.98</td>
<td>842.00</td>
</tr>
<tr>
<td></td>
<td>2040</td>
<td>939.72</td>
<td>875.90</td>
<td>995.35</td>
</tr>
<tr>
<td></td>
<td>2050</td>
<td>1038.44</td>
<td>953.39</td>
<td>1112.80</td>
</tr>
<tr>
<td></td>
<td>2060</td>
<td>1099.15</td>
<td>997.03</td>
<td>1190.53</td>
</tr>
<tr>
<td></td>
<td>2070</td>
<td>1109.56</td>
<td>1002.11</td>
<td>1210.47</td>
</tr>
<tr>
<td></td>
<td>2080</td>
<td>1087.29</td>
<td>968.53</td>
<td>1193.18</td>
</tr>
<tr>
<td></td>
<td>2090</td>
<td>1036.20</td>
<td>905.00</td>
<td>1141.36</td>
</tr>
<tr>
<td></td>
<td>2099</td>
<td>988.65</td>
<td>855.23</td>
<td>1103.90</td>
</tr>
<tr>
<td>SSP5-RCP8.5</td>
<td>2020</td>
<td>625.60</td>
<td>606.82</td>
<td>642.52</td>
</tr>
<tr>
<td></td>
<td>2030</td>
<td>808.86</td>
<td>771.40</td>
<td>847.31</td>
</tr>
<tr>
<td></td>
<td>2040</td>
<td>999.74</td>
<td>934.05</td>
<td>1058.49</td>
</tr>
<tr>
<td></td>
<td>2050</td>
<td>1134.25</td>
<td>1040.90</td>
<td>1215.05</td>
</tr>
<tr>
<td></td>
<td>2060</td>
<td>1200.14</td>
<td>1088.05</td>
<td>1298.01</td>
</tr>
<tr>
<td></td>
<td>2070</td>
<td>1181.04</td>
<td>1065.74</td>
<td>1282.36</td>
</tr>
<tr>
<td></td>
<td>2080</td>
<td>1105.18</td>
<td>986.24</td>
<td>1214.25</td>
</tr>
<tr>
<td></td>
<td>2090</td>
<td>973.55</td>
<td>851.78</td>
<td>1068.97</td>
</tr>
<tr>
<td></td>
<td>2099</td>
<td>826.43</td>
<td>712.22</td>
<td>924.20</td>
</tr>
</tbody>
</table>

Source: Own elaboration
Figure 19: Annual electricity demand
GDP growth over the period, and these forces affect the demand in opposite directions, although the final result is a relatively high electricity demand in year 2100 due to a large population. Again, the wide range between the percentiles of the distribution shows that the climatic risk to electricity demand is enormous.

The third scenario, RCP8.5 paired with SSP5 as shown in figure 21, also shows a strong increase in annual electricity demand until year 2060, followed by a sharp decline. The assumed GDP growth and temperature increase are both strong throughout the period, and the decrease in electricity demand is therefore a result of declining population. The range spanned by the percentiles shows again that the climatic risk to electricity demand is considerable.

Figure 22 shows the mean annual electricity demand over all the simulations for all the scenarios. The scenarios are almost indistinguishable until approximately year 2030. After
Figure 21: Annual electricity demand
Figure 22: Annual electricity demand, mean of simulations for the three scenarios
Source: Own elaboration

In figures 19, 20 and 21, we also compare our results to the official Brazilian projections found in the ten-year expansion plan PDE-2023 (EPE-PDE, 2014) and the long-term projections from PNE-2050 (EPE-PNE, 2014). The official projections are well above even our estimate 10% percentile. The annual electricity demand projections for the years 2040 and 2050 found in PNE-2050 are completely out of line with the projections created by our framework. The aforementioned figures also show the demand projections for years 2030, 2080, 2090 and 2100 presented by Schaeffer et al. (2008). A quick comparison shows
that the projections for year 2030 are significantly above all three scenarios considered here. The projections for year 2080 for both the A2 and B2 scenarios are more or less in line with all three of our scenarios, and their projections for years 2090 and 2100 are quite close to our results for both the RCP4.5-SSP2 and RCP8.5-SSP5 scenarios. Their projections for year 2090 and 2100 for both the A2 and B2 scenarios are considerably above the results of our RCP4.5-SSP1 scenario.

5.3 Model Evaluation

The framework we have used to create electricity demand projections contains a few critical limitations that must be taken into consideration.

Firstly, the procedure made no explicit allowance for structural and/or technological change, for example widespread adoption of electrical vehicles or deployment of electricity storage technology that would radically change the determinants of electricity demand. Calibrating a model on a decade of historical data, then subsequently using the model to make demand projections for the next 85 years, may seem like little more than a shot in the dark to some and pure madness to others. We recognise this limitation, and even though the next 85 years certainly will be very different from the last ten, we still believe that the demand projections provide some guidance. It is a very brave assumption, but every attempt to avoid it would perhaps be equally brave. In essence, all forecasts are wrong, but they can still be useful.

Secondly, we did not allow for uncertainty in the GDP and population numbers. Although this could be incorporated in principle, our main focus was on climatic risks and therefore we concentrated on weather uncertainty.

Thirdly, we have also disregarded the uncertainty in the model parameters. That is, the calibration of the model was considered to be absolutely accurate. This is unrealistic, but again this assumption allowed us to focus on climate risk and weather uncertainty.

These shortcomings discussed here are clear, and although they are cause to be cautious, they do not invalidate the results. Taking into consideration the radically different framework in this study and the different datasets used during both calibration and fore-
casting, it is comforting that our projections share some characteristics with the results presented in the existing literature. All in all, the projections seem quite reasonable in general.

6 Conclusion

We have provided electricity demand projections for Brazil for three climate change scenarios, with probability distributions illustrating the climatic risk that the electricity demand is exposed to. We have improved on the existing literature in three important ways: firstly, we have applied a probabilistic approach to Brazilian electricity demand forecasting. Since forecasting is essentially a stochastic problem, this is very appropriate, although no such approach appears to have been considered yet in the existing literature. Secondly, we provide updated demand projections for climate change scenarios in light of the new climatic, demographic and macroeconomic simulations accompanying the IPCC AR5 report, using the latest available data. This is important, because much has happened since the last comprehensive electricity demand projections that explicitly considered climate change scenarios were published in 2008. And finally, by creating the electricity demand projections using a model for daily electricity demand, our framework gives easy access to an unprecedented amount of detail about the Brazilian electricity demand. Such a level of detail, for instance being able to retrieve the probability distribution of the maximum daily demand of a year, can be very valuable for planning purposes.

Acknowledgments

The author would like to thank the Brazilian agency CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for the scholarship BEX 10893-14.
References


URL http://www.ibge.gov.br/home/estatistica/populacao/projecao_da_populacao/2013/default.shtm


URL http://www.ibge.gov.br/home/estatistica/indicadores/pib/defaultcnt.shtm


URL https://tntcat.iiasa.ac.at/SspDb


energia elétrica para classes de consumo na região nordeste, usando ols dinâmico e mudança de regime. Economia Aplicada 13 (1), 69–98.


URL https://cds.nccs.nasa.gov/nex-gddp/


URL https://www.ncdc.noaa.gov/isd


URL http://www.ons.org.br/resultados_operacao/boletim_diario/index.htm


Sun, Y.-y., Zhao, Y.-h., Wang, W., Jia, H., 2013. Application study of air conditioning system operational load prediction based on grnn. Journal of Beijing University of Technology 7, 020.


URL http://www.eia.gov/countries/analysisbriefs/brazil/brazil.pdf


A Calibrated Model

t test of coefficients:

<p>| Estimate  | Std. Error | t value | Pr(&gt;|t|) |
|-----------|------------|---------|-----------|
| (Intercept) | -2.9902e+00 | 7.4830e+00 | -0.3996 | 0.6894750 |
| daylight_hours_pa | -3.2164e+00 | 1.3053e+00 | -2.4641 | 0.0137860 * |
| daylight_hours_sp | 8.0754e+00 | 2.3586e+00 | 3.4238 | 0.0006252 *** |
| daylight_hours_sa | -7.2021e+00 | 1.4636e+00 | -4.9208 | 9.053e-07 *** |
| D_Mon | -2.7672e-02 | 6.8651e-04 | -40.3077 | &lt; 2.2e-16 *** |
| D_Sat | -8.9385e-02 | 1.0246e-03 | -87.2412 | &lt; 2.2e-16 *** |
| D_Sun | -2.0298e-01 | 9.0462e-04 | -224.3757 | &lt; 2.2e-16 *** |
| D_Jan01 | -2.5794e-01 | 1.3363e-02 | -19.3035 | &lt; 2.2e-16 *** |
| D_Jan02 | -7.7532e-02 | 9.4730e-03 | -8.1846 | 3.897e-16 *** |
| D_Apr21 | -1.4522e-01 | 8.7693e-03 | -16.5597 | &lt; 2.2e-16 *** |
| D_May01 | -1.7870e-01 | 8.2543e-03 | -21.6497 | &lt; 2.2e-16 *** |
| D_Jul09 | -5.2531e-02 | 6.4678e-03 | -8.1220 | 6.471e-16 *** |
| D_Sep07 | -1.4756e-01 | 9.5850e-03 | -15.3948 | &lt; 2.2e-16 *** |
| D_Oct12 | -1.5613e-01 | 9.2497e-03 | -16.8800 | &lt; 2.2e-16 *** |
| D_Nov02 | -1.7726e-01 | 8.1625e-03 | -21.7165 | &lt; 2.2e-16 *** |
| D_Nov15 | -1.3064e-01 | 1.0719e-02 | -12.1872 | &lt; 2.2e-16 *** |
| D_Nov20 | -3.3166e-02 | 6.4912e-03 | -5.1093 | 3.421e-07 *** |
| D_Dec23 | -4.4558e-02 | 1.3079e-02 | -3.4068 | 0.0006653 *** |
| D_Dec24 | -1.6632e-01 | 8.4632e-03 | -19.6519 | &lt; 2.2e-16 *** |
| D_Dec25 | -2.8454e-01 | 1.4692e-02 | -19.3667 | &lt; 2.2e-16 *** |
| D_Dec31 | -8.0852e-02 | 1.3305e-02 | -6.0769 | 1.369e-09 *** |
| D_Dec25Sun | 1.8908e-01 | 2.2226e-02 | -8.5072 | &lt; 2.2e-16 *** |
| D_Dec24SatSun | 1.0909e-01 | 2.1944e-02 | 4.9712 | 7.002e-07 *** |
| D_Jan01Sun | 1.9088e-01 | 1.5722e-02 | 12.1406 | &lt; 2.2e-16 *** |
| D_ChristmasWeek | -9.2825e-02 | 8.8881e-03 | -10.4437 | &lt; 2.2e-16 *** |</p>
<table>
<thead>
<tr>
<th>Date</th>
<th>Value1</th>
<th>Value2</th>
<th>Value3</th>
<th>Value4</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_Apr20Mon</td>
<td>-9.4232e-02</td>
<td>6.2052e-03</td>
<td>-15.1858</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_Apr21SatSun</td>
<td>1.2173e-01</td>
<td>9.4486e-03</td>
<td>12.8831</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_May01SatSun</td>
<td>1.0173e-01</td>
<td>3.0693e-02</td>
<td>3.3143</td>
<td>0.0009289</td>
</tr>
<tr>
<td>D_May02</td>
<td>-4.1787e-02</td>
<td>9.9591e-03</td>
<td>-4.1959</td>
<td>2.792e-05</td>
</tr>
<tr>
<td>D_Sep06Mon</td>
<td>-6.6169e-02</td>
<td>4.1756e-03</td>
<td>-15.8465</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_Sep07SatSun</td>
<td>1.1748e-01</td>
<td>2.4673e-02</td>
<td>4.7614</td>
<td>2.007e-06</td>
</tr>
<tr>
<td>D_Sep08</td>
<td>-3.6633e-02</td>
<td>8.8685e-03</td>
<td>-4.1307</td>
<td>3.709e-05</td>
</tr>
<tr>
<td>D_Oct11Mon</td>
<td>-6.9830e-02</td>
<td>4.8233e-03</td>
<td>-14.4777</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_Oct12SatSun</td>
<td>1.2955e-01</td>
<td>3.2596e-02</td>
<td>3.9746</td>
<td>7.206e-05</td>
</tr>
<tr>
<td>D_Nov01Mon</td>
<td>-8.3570e-02</td>
<td>5.0037e-03</td>
<td>-16.7016</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_Nov02SatSun</td>
<td>1.5202e-01</td>
<td>3.1126e-02</td>
<td>4.8841</td>
<td>1.090e-06</td>
</tr>
<tr>
<td>D_Nov15SatSun</td>
<td>9.0162e-02</td>
<td>1.5185e-02</td>
<td>5.9375</td>
<td>3.204e-09</td>
</tr>
<tr>
<td>D_Nov16</td>
<td>-3.5443e-02</td>
<td>8.5976e-03</td>
<td>-4.1225</td>
<td>3.843e-05</td>
</tr>
<tr>
<td>D_1stJanWeek</td>
<td>-4.0138e-02</td>
<td>7.6555e-03</td>
<td>-5.2431</td>
<td>1.681e-07</td>
</tr>
<tr>
<td>D_Carneval1</td>
<td>-1.6225e-01</td>
<td>6.4132e-03</td>
<td>-25.2994</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_Carneval2</td>
<td>-1.3515e-01</td>
<td>1.1346e-02</td>
<td>-11.9116</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_Carneval3</td>
<td>-2.9211e-02</td>
<td>5.2968e-03</td>
<td>-5.5148</td>
<td>3.769e-08</td>
</tr>
<tr>
<td>D_Carneval4</td>
<td>-2.6619e-02</td>
<td>4.0956e-03</td>
<td>-6.4994</td>
<td>9.312e-11</td>
</tr>
<tr>
<td>D_AshWednesday</td>
<td>-5.7668e-02</td>
<td>5.5411e-03</td>
<td>-10.4072</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_GoodFriday</td>
<td>-2.1129e-01</td>
<td>9.1480e-03</td>
<td>-23.0967</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_EasterSunday</td>
<td>-2.8532e-02</td>
<td>6.8778e-03</td>
<td>-4.1484</td>
<td>3.435e-05</td>
</tr>
<tr>
<td>D_CorpusChristi</td>
<td>-1.1208e-01</td>
<td>4.3595e-03</td>
<td>-25.7093</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>D_CorpusChristiFri</td>
<td>-4.9414e-02</td>
<td>6.7415e-03</td>
<td>-7.3298</td>
<td>2.903e-13</td>
</tr>
<tr>
<td>D_WorldCup_Matches</td>
<td>-4.1775e-02</td>
<td>6.4568e-03</td>
<td>-6.4698</td>
<td>1.130e-10</td>
</tr>
<tr>
<td>logpib</td>
<td>4.7509e-01</td>
<td>9.6663e-02</td>
<td>4.9150</td>
<td>9.325e-07</td>
</tr>
<tr>
<td>logpop</td>
<td>1.7086e+00</td>
<td>3.4013e-01</td>
<td>5.0235</td>
<td>5.350e-07</td>
</tr>
<tr>
<td>C25.837460</td>
<td>5.8294e-04</td>
<td>6.5874e-05</td>
<td>8.8493</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>C25.837800</td>
<td>3.1173e-04</td>
<td>1.1217e-04</td>
<td>2.7791</td>
<td>0.0054831</td>
</tr>
<tr>
<td>Variable</td>
<td>Coefficient</td>
<td>Standard Error</td>
<td>t value</td>
<td>Pr(&gt;</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>----------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>C25.837680</td>
<td>3.3852e-04</td>
<td>1.1772e-04</td>
<td>2.8757</td>
<td>0.0040577 **</td>
</tr>
<tr>
<td>C25.838270</td>
<td>2.8912e-04</td>
<td>7.6326e-05</td>
<td>3.7880</td>
<td>0.0001547 ***</td>
</tr>
<tr>
<td>HDD.838400</td>
<td>-1.3928e-04</td>
<td>2.4900e-05</td>
<td>-5.5936</td>
<td>2.410e-08 ***</td>
</tr>
<tr>
<td>C25.839710</td>
<td>2.0234e-04</td>
<td>7.5715e-05</td>
<td>2.6723</td>
<td>0.0075703 **</td>
</tr>
<tr>
<td>C25.825790</td>
<td>1.4577e-04</td>
<td>7.9643e-05</td>
<td>1.8303</td>
<td>0.0673019 .</td>
</tr>
<tr>
<td>C25.832480</td>
<td>1.2992e-04</td>
<td>7.8748e-05</td>
<td>1.6498</td>
<td>0.0990745 .</td>
</tr>
<tr>
<td>L1.HDD.838400</td>
<td>-1.5552e-04</td>
<td>2.0663e-05</td>
<td>-7.5264</td>
<td>6.735e-14 ***</td>
</tr>
<tr>
<td>MA9.C25.837460</td>
<td>3.2237e-04</td>
<td>1.5410e-04</td>
<td>2.0920</td>
<td>0.0365167 *</td>
</tr>
<tr>
<td>MA9.C25.837800</td>
<td>2.0306e-03</td>
<td>7.5656e-04</td>
<td>2.6840</td>
<td>0.0073130 **</td>
</tr>
<tr>
<td>MA9.HDD.838400</td>
<td>-2.4583e-04</td>
<td>7.5619e-05</td>
<td>-3.2509</td>
<td>0.0011623 **</td>
</tr>
<tr>
<td>MA9.C25.839710</td>
<td>-7.9136e-04</td>
<td>3.9880e-04</td>
<td>-1.9843</td>
<td>0.0473028 *</td>
</tr>
<tr>
<td>MA9.C28.820980</td>
<td>-6.1332e-04</td>
<td>2.8309e-04</td>
<td>-2.1665</td>
<td>0.0303442 *</td>
</tr>
<tr>
<td>MA9.C25.825990</td>
<td>5.8429e-04</td>
<td>1.7966e-04</td>
<td>3.2521</td>
<td>0.0011573 **</td>
</tr>
<tr>
<td>MA9.C28.838270</td>
<td>-2.3624e-03</td>
<td>1.1997e-03</td>
<td>-1.9692</td>
<td>0.0490176 *</td>
</tr>
<tr>
<td>MA9.C28.839710</td>
<td>1.9521e-03</td>
<td>9.9293e-04</td>
<td>1.9660</td>
<td>0.0493883 *</td>
</tr>
</tbody>
</table>

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1