Gest. Prod., São Carlos, v. 23, n. 4, p. 787-797, 2016
http://dx.doi.org/10.1590/0104-530X-1868-15

Proposal of an empirical method to adjust time series for calendar and temperature effects

Proposta de um método empírico de ajuste de séries temporais influenciadas pelos ofensores calendário e temperatura

Rodrigo Flora Calili[1]
Fernando Luiz Cyrino Oliveira[2]
Reinaldo Castro Souza[3]
Thiago Gomes de Araujo[3]
Pedro Guilherme Costa Ferreira[4]

Abstract: In Brazil, the behavior of electrical load, particularly in energy consumption, has been widely investigated over the past years. In general, this interest is due to the great financial and social importance of this input, as its failure or shortage can have a variety of damaging impacts to the country. This paper proposes a method to generate monthly load series freed from variations arising from two sources: calendar and temperature. To find the best fitting approach to removing these effects, we considered a totally empirical method and one with hybrid features, as it uses both empirical procedures and time series models. The data set used comes from daily observations from each one of the four subsystems that form the Brazilian Electricity Grid. However, the final task is to obtain unique monthly series for the entire grid, and not only the four subsystems. The quarterly GDP series was used to check the performance of the two proposed methods. It was noted that the adjustment difference is minimal in the two approaches studied, and that both series had a great explanatory power when compared with the time series without removing calendar and temperature effects.

Keywords: Electrical load; Time series; Empirical method.

Resumo: *No Brasil, o comportamento da carga, em especial do consumo de energia, tem sido amplamente investigado nos últimos anos. Esse interesse, em geral, é devido à grande importância financeira e social desse insumo, pois sua falta pode causar todo tipo de dano ao país. O objetivo do presente trabalho é a geração de uma série mensal de carga elétrica livre das variações de ofensores não econômicos, no caso, calendário e temperatura. Foram comparadas duas abordagens com vistas à seleção da mais eficiente na remoção dos efeitos dos referidos ofensores: a primeira de natureza empírica e a segunda com características híbridas, utilizando métodos empíricos e modelos de Séries Temporais. Os dados utilizados são provenientes de observações diárias de cada um dos quatro subsistemas que integram o Sistema Interligado Nacional (SIN), porém a ideia é produzir séries mensais do SIN e não apenas de cada um dos subsistemas. A série trimestral do PIB foi utilizada para decidir qual abordagem melhor ajustou os dados de carga. Verificou-se que a diferença dos ajustes é mínima entre os métodos propostos, apresentando alto poder de explicação quando comparadas à série sem a retirada dos ofensores calendário e temperatura.*

Palavras-chave: *Carga elétrica; Séries temporais; Método empírico.*

[1] Programa de Pós-graduação em Metrologia, Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio, Rua Marquês de São Vicente, 225, Gávea, CEP 22453-900, Rio de Janeiro, RJ, Brazil, e-mail: rcalili@esp.puc-rio.br

[2] Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio, Rua Marquês de São Vicente, 225, Gávea, CEP 22453-900, Rio de Janeiro, RJ, Brazil, e-mail: cyrino@puc-rio.br

[3] Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio, Rua Marquês de São Vicente, 225, Gávea, CEP 22453-900, Rio de Janeiro, RJ, Brazil, e-mail: reinaldo@ele.puc-rio.br; thiago_gda@hotmail.com

[4] Instituto Brasileiro de Economia, Fundação Getúlio Vargas – FGV, Rua Barão de Itambi, 60, Botafogo, CEP 22231-000, Rio de Janeiro, RJ, Brazil, e-mail: pedro.guilherme@fgv.br

Received Nov. 11, 2015 - Accepted June 8, 2016

Financial support: The authors would like to thank the financial support of CAPES, CNPq (Research Project 443595/2014-2) and FAPERJ (Research Project E-26/202.806/2015).

1 Introduction

In Brazil, the load pattern, in particular in energy consumption has been widely investigated in recent years. This interest is generally due to the major financial and social importance of this input, since the lack of reliable energy services may cause serious harm to the country (Maçaira et al., 2014).

The separation of the three stages that make up the electricity supply chain (generation, transmission, and distribution), known as vertical de-integration, or unbundling, is a key feature of electricity sector reform. The prevention of predatory behavior and the increase of competitors in the generation segment, given free access to the network, was the purpose of unbundling (Leite & Santana, 2006). However, the reform proposal presented significant failures both in planning and execution, leading to energy rationing in 2001. It is important to note that in the post-rationing period, utilities faced serious liquidity crises and operational losses (Pires et al., 2002). In 2003 Brazil started a process of re-orienting and restructuring the electrical power sector, resulting in a new power structure.

Under the Ministry of Mines and Energy (MME), a series of studies was initiated to formulate and implement a new model for the Brazilian electricity sector, whose institutional and legal bases were approved by Congress through Laws 10.438 and 10.848 of 2004. This new model has two main objectives: to ensure security of the electricity supply, and reasonable tariffs (Brasil, 2002, 2004b).

With regard to supply security, the "new" and the current Brazilian model incorporated: (a) a shift of the focus of electricity contracts from short- to long-term, in order to reduce price volatility and create a market of long-term contracts that can be used as a firm guarantee for financing; (B) mandatory contract coverage, by the distributors and free consumers, of 100% of their electricity consumption; (C) creation of the Monitoring Committee for the Power Sector (CMSE), whose purpose is to improve the balance between supply and demand over short, medium, and long terms; (D) preliminary environmental permits for prospective bidders in new auctions; and (E) resumption of centralized, cross-sectoral planning by the State, through the creation of the Energy Research Enterprise (EPE), by Decree 5184 of 2004 (Brasil, 2004c).

The Brazilian electricity market was divided into two trading environments, with clearly distinct logics of action and structuring. The first, the Regulated Trading Floor (*Ambiente de Contratação Regulada, ACR)*, meets the demands of captive consumers, primarily residential, service and industry consumers, with lower consumption levels. The second is the Free Trading Floor (*Ambiente de Contratação Livre, ACL)*, focused exclusively on companies with the highest volume of consumption and strategic needs for a higher volume of energy in the short term, known as free consumers, and seeks to ensure competition and the effective freedom of the participants.

An important feature of this model is that the vertical integration of the companies is no longer allowed. Distributors are not granted the right to own generation assets or market directly with free consumers. In other words, they can act only in the ACR, buying through auctions and selling to captive consumers. This model feature reduces the likely and undesirable cross-subsidies, whereby the tariffs of captive consumers could subsidize lower prices for free customers, as reported by Castro & Leite (2008).

Thus, according to Law No. 10.848/2004 (further regulated by Decree No. 5.163/2004) companies with a concession, permission, or authorization for rendering public services in the National Interconnected System (*Sistema Interligado Nacional*, *SIN*) purchase electric energy through public energy auctions to ensure that their market needs are met (Brasil, 2004a, b).

The winners of an auction will be those that offered electricity at the lowest price per megawatt-hour (MWh) to meet the demand expected by the distributors; that is, the lowest tariff criterion (item VII of art. 20 of Decree 5.163 / 2004) will be used to define the winners of the auction. Thus, the Electricity Trading Contracts in a Regulated Environment *(Comercialização de Energia Elétrica em Ambiente Regulado, CCEAR)* will be concluded between the winners and the distributors who declared the need to purchase, to initiate the energy supply contracted in the auction (Reis, 2003).

In order to have 100% of its contracted market, one of the regulatory frameworks of the "new model", the distributors have to participate in the regulated auctions, proposed and organized by the government, via the Energy Research Company (*Empresa de Pesquisa Energética, EPE*). Therefore, distributors have to provide their load and declare it in the auction. It is known, *a priori*, that the variables related to calendar and temperature have a significant impact on energy consumption (Apadula et al., 2012), and therefore on the load.

For the distribution companies to make accurate forecasts of the electricity load to be offered, so that the country's supply is guaranteed, the load series must be well treated and free of non-economic offenders, such as the above mentioned calendar and temperature. Thus, given the importance of generating electricity load series free from non-economic offenders, the main objective of this work is to produce monthly and quarterly series of load, free of such offenders.

This adjustment is carried out in some countries. In Spain, Rede Eléctrica de España (2012) discloses in its *Boletín Mensual* data about physical measurements of the load, and the data are corrected to compensate for effects of atypical temperatures and differences in the number of working days. The data series are available on a monthly frequency and separate the load variation caused by temperature, by differences in the calendar, and by economic activity. Likewise, in Portugal the Rede Eléctrica Nacional (2012) disseminates monthly, through the producer information system, physical measurements of the load and adjustments for the effects of temperature and the number of working days.

In Brazil, Castro et al. (2012) developed a methodology appropriate for the specificities of the country to produce adjusted load data, able to isolate the influence of non-economic factors. These authors listed the non-economic factors that influence the electricity load in Brazil, and that can be estimated based on available data, as:

- Measurement discontinuities: these occur when, for example, a system isolated from the National Interconnected System (SIN) is incorporated. In October 2009, the former isolated system Acre-Rondônia began to be connected to the SIN. Since mid-October, part of the load of these two states, at the time amounting to approximately 400 aMW, started to be computed in SIN. This load increase, corresponding to 0.75% of the SIN load from one month to the other, is not caused by any economic growth, but is the result of the simple fact that the National System Operator (ONS) began measuring the load of a larger portion of the country.
- Power Transmission network losses: significant fluctuations can occur over the year for reasons closely linked to system operation. Although transmission losses have as their basis electrical, not economic, reasons, in most countries transmission losses tend to be small and to vary slightly—and this is probably why the adjustment of the losses is not made through load data treatment in other countries. But in Brazil, the transmission system is very extensive and carries large volumes of energy, since it was built to take advantage of differences in hydrological regimes among the different basins with hydroelectric exploitations. SIN's controlling agency, the Electric System National Operator (Operador Nacional do Sistema Elétrico, ONS), uses the transmission system to transport energy from places where it is abundant to where it is scarce. The heavy use of large interconnections significantly affects the level of losses. The losses increase, for example, due to greater use of energy transport over long distances and decrease when consumption is met primarily by the nearest generator.
- Occurrence of atypical temperatures: these can also strongly influence the load. Whereas industrial consumption tends not to be sensitive to temperature variations, commercial, and especially residential, consumption respond strongly to temperature variations. Part of the effect of temperature on the load is seasonal, and a large increase is expected in residential consumption during the summer. However, the occurrence of atypical temperatures can also cause the load to escape the normal seasonal pattern. For example, a summer with relatively mild temperatures can result in an abnormally low load. This reduction of the load has little or no economic significance and can be estimated based on available data.
- Variations on the calendar: the international literature on the subject usually sticks to the influence of the number of calendar days in a month. To properly compare an indicator of a given month with the reading of the same indicator for the same month last year, an adjustment is necessary to reflect the difference between the number of working days.

This paper proposes a novel approach to the treatment of the temperature effect. For this purpose, it will establish a time series model that will enable the use of the parameters and estimate the effects on the variable load, differing from the existing literature in that it provides a statistical basis predicated on the withdrawal of these offenders.

Besides this introduction, this paper is organized as follows: Section 2 is devoted to the methodology for withdrawing calendar and temperature effects from the load series; Section 3 describes the results obtained on real data from ONS; and finally in Section 4 the final considerations are exposed.

2 Methodology

This section will describe our methodology, first using an empirical method, aimed at the removal of the calendar and temperature effects from the load series. Next, we will propose an alternative approach, using models of time series with explanatory variables. Importantly, this model is applied only to the removal of the temperature effect on load series. For purposes of better understanding, the notation of the load and temperature variables are written as follows (see Equation 1):

$$LOAD_{t|s}^{r,d,m,a} \qquad TEMP_{t|s}^{r,d,m,a} \tag{1}$$

Where:

t is the index that references the day in question;

s is the index that indicates to which subsystem the observation belongs;

a indicates the year of observation;

m indicates the month of observation, $1 \leq m \leq 12$;

d indicates the day of observation, $1 \leq d \leq 31$; and

r indicates the day of the week, ranging from Sunday to Saturday plus holidays and semi-holidays (days before holidays), using $1 \leq r \leq 9$. Sunday is encoded with the value of r = 1, Monday r = 2, and so on until Saturday, which will receive the value of r = 7. Holidays will be coded r = 9, and semiholidays r = 8.

2.1 Empirical method

In this case, the adjusted load series are built to exclude the influence of non-economic and / or fortuitous factors influencing the load, with two advantages over traditional series. First, they are a good basis for predicting the load in the medium and long term, in that they reflect not the load which effectively occurred, but that which would have been observed in the absence of random factors. Second, the adjusted load series allow for a more detailed analysis of the relationship between economy and load, thereby being an indicator of good quality economic activity.

2.1.1 Adjustment of daily load series for the calendar factor

The calendar adjustment procedure for the load of the subsystems entails re-estimating the daily load series to obtain a series that does not reflect the specifics of the verified calendar. This methodology consists of recalculating the daily data by multiplying it by the inverse ratio normally observed between the daily load in question and the weekly load.

2.1.1.1 Weights for each day

First, it is necessary to calculate the weight of each observation for the week to which this observation belongs, that is, the weight of day r of a specific week. Once the calculation for every day of the series is made, it is possible to calculate the typical weight of a day in any given week.

Also, to perform the calculation of weight, it should be noted whether the week has a holiday or semiholiday. Knowing that the week starts on a Sunday and ends on a Saturday, a week without a holiday or semiholiday within this range will be considered week type 1. If the week contains a holiday or semiholiday, it will be considered week type 2.

The formula to calculate the weight of the day t within a week Type 1 is given by Equation 2:

$$Weight_{t|s}^{r,d,m,a} = \frac{LOAD_{t|s}^{r,d,m,a}}{\sum_{i=1-r}^{7-r} LOAD_{t+i|s}^{r+i,d,m,a} \Big/ 7} \tag{2}$$

As can be seen in (2), the weight of day t is calculated by dividing the load of day t by the average of the loads of the week to which day t belongs. It is understood that this weight is distorted for a week with holiday or semiholiday. Thus, the load of the day is compared with the average between the average load of the previous week and that of the following week, if neither of them have holidays. If one of the two weeks in question has a holiday or semiholiday, we calculate the average of both weeks without holidays or semiholidays equidistant from the week to which t belongs, and as close as possible to it. The typical weight of each weekday (Equation 3) will be the average of the weights for all the days r:

$$Typical\_Weight^{r} = \frac{\sum_{i=1}^{n} Weight_{m|s}^{r,d,m,a}}{n} \tag{3}$$

Wherein

n is the number of occurrences of day

r at the base.

2.1.1.2 Daily load adjusted for calendar effects

The daily load value is determined by Equation 4:

$$adj\_LOAD_{t|s}^{r,d,m,a} = \frac{LOAD_{t|s}^{r,d,m,a}}{Typical\_Weight^{r}} \tag{4}$$

2.1.2 Adjustment of daily load series for the temperature factor

The task of temperature adjustment is addressed by compensating for the effect of atypical temperatures on the load. This implies re-estimating the daily load series using the typical daily temperatures instead of the observed temperatures. If the checked temperature is greater than the normal value for the season, the load, which reacts positively to heat, must also have been greater than that which would normally occur. The temperature adjustment, in this case, consists of calculating an adjusted load smaller than the one observed, with the adjustment proportional to the difference between the detected and the typical temperature.

2.1.2.1 Verification of the relationship between load and temperature daily series

It is known that the influence of the temperature variable on the load variable is different in each season. Thus, the data series will be divided into twelve sets of data, each of which collects information of the same month; in other words, twelve linear regression models will be built, taking into account every day of a particular month. For example, the 1st regression includes the data for the month of January. The standardization of the series will also be considered, and there will be seven differentiations in the variables *Adj_load* and temperature. Mathematically, we have, in Equations 5 and 6:

$$DTemp_{t-i|s}^{r,d,m,a} = Temp_{t|s}^{r,d,m,a} - Temp_{t-i|s}^{r,d,m,a} \quad (5)$$

$$DLOAD_{t-i|s}^{r,d,m,a} = \frac{LOAD\_adj_{t|s}^{r,d,m,a}}{Adj\_LOAD_{t-i|s}^{r,d,m,a}} - 1 \quad (6)$$

Where $1 \leq i \leq 7$.

From this point on, the new databases possess the variables $DLoad_{k/s}^{m}$ and $DTemp_{k/s}^{m}$, where k is equal to $(t-i)$, i is the nth lag of t, and m is the month that characterizes the database.

2.1.2.2 Analysis of the gradient of the regressions

The regressions are developed with the intention of using the gradient to adjust the load value through the temperature effect. If this coefficient is significant, the load will be adjusted by the temperature, and if not this will not occur. For this work, the variable $DLoad_{k/s}^{m}$ is used as the dependent variable in the regression, and $DTemp_{k/s}^{m}$ as the regressor variable.

Importantly, given the nature of the linear regression model used, error lags were included, ensuring the absence of serial correlation in the residuals. This was tested for all adjusted models using the Ljung Box test. In all cases, there was no evidence to reject the null hypothesis (no serial correlation in the first k lags, considering k = 20) at a significance level of 5%.

After calculating the regression coefficient for the observations of the new data sets that contain information about each month, it is necessary to remove the atypical days from the analysis. Thus, confidence intervals are calculated, considering the standard error and the estimated value of the load, to assess whether the variation of the load of a given day is an outlier or not. If the load variation value is within the calculated confidence interval, this means that this value is not atypical. Otherwise, this value will be considered an outlier and excluded from the analysis. With the removal of outliers, the regression coefficient is recalculated and used to calculate the Temperature Factor, which will be shown in section 2.1.2.5.

2.1.2.3 Typical daily temperature

The typical daily temperature, which takes into account the observations of the daily series of temperature, will be denoted by , Equation 7, and is defined as the temperature average of the same day and same month of every year, that is:

$$TTA_{L|s}^{d,m} = \frac{\sum_{i=w}^{v} Temp_{t|s}^{r,d,m,i}}{v-w+1} \quad (7)$$

Where:

L is the day of the year. January 1st, regardless of the year, will be referred to as $L = 1$, and December 31st as L = 365, i.e. $1 \leq L \leq 365$;

v is the last year of the series that contains the day / month under analysis;

w is the first year of the series that contains the day / month under analysis.

2.1.2.4 Typical daily temperature II

This typical day temperature is calculated from the TTAs and will be denoted by *TTM*, which is the average of the 3TTTAs of the nearest days, that is (see Equation 8),

$$TTM_{L|s}^{d,m} = \frac{\sum_{i=L-15}^{L+15} TTA_{i|s}^{d,m,a}}{31} \quad (8)$$

2.1.2.5 Temperature factor

This factor is used to adjust the load value. It takes into account the gradient estimated by the linear regression described in 2.1.2.2. Equation 9 describes this procedure, wherein $b_{l/m}^{m}$ is the gradient estimated by regression in item 2.1.2.2

$$Fact\_Temp_{t|s}^{r,d,m,a} = 1 + \left[\frac{\left(TTM_{L|s}^{d,m} - Temp_{t|s}^{r,d,m,a} \right) \cdot b_{l|s}^{m}}{100} \right] \quad (9)$$

2.1.2.6 Temperature-adjusted load

The effect of temperature is corrected by Equation 10:

$$LOAD\_adjT_{t|s}^{r,d,m,a} = LOAD\_adj_{t|s}^{r,d,m,a} \cdot Fat\_Temp_{t|s}^{r,d,m,a} \quad (10)$$

2.1.3 Daily series adjustment for the dead week factor

The week between Christmas and New Year is a period of low electricity consumption because many industries take advantage of the celebrations to promote collective holidays, stopping or greatly reducing production. During this week there is load reduction with respect to the typical one, both on informal and formal semiholidays and holidays. Notably, December 25th tends to have a reduced load, much smaller than that of a regular holiday. Thus, the days belonging to the weeks of the year-end festivities receive special treatment. For each observation of the load series, a deviation from the Weekly Weight is calculated (as shown in Equation 11), Equation 2, for the Typical Weight of the day, Equation 3:

$$Deviation_{t|s}^{r,d,m,a} = \frac{Weight_{t|s}^{r,d,m,a}}{Typical\_Weight_{t|s}^{r,d,m,a}} - 1 \quad (11)$$

The factor to correct a day that belongs to the dead week is calculated as follows (Equation 12),

$$Dead\_Week\_Fact_s = \frac{\sum_{t=1}^{k} Deviation_{t|s}^{r,d,m,a}}{k} \quad (12)$$

where k is the total number of days belonging to the dead week in the load series. Hence, the sum contains only the dead week's days.

The adjustment factor for the Dead Week is not applied to ordinary days; hence, the daily load series adjusted by the Dead Week factor is given as Equation 13:

$$DAILY\_load_{t|s}^{r,d,m,a} = \begin{cases} adj\_LOADT_{t|s}^{r,d,m,a}, & ordinary\ day \\ \dfrac{adj\_LOADT_{t|s}^{r,d,m,a}}{1 + Fact\_Dead\_Week_s}, & dead\ week\ day \end{cases} \quad (13)$$

2.1.4 Monthly load adjustment

The adjusted monthly series takes into account both the data available on daily basis, such as: monthly generation data of small power plants, measurement discontinuities, and losses in the core grid. The adjustment in this case is direct, and the Monthly Net Load Loss is exemplified by the following steps:

i) The monthly load is the mean of the daily load values, as Equation 14.

$$Monthly\_LOAD_s^{m,a} = \frac{\sum_{t=1}^{v} Load_{t|s}^{r,d,m,a}}{v} \quad (14)$$

where v is the number of days of month m in a particular year a.

ii) The monthly adjusted load is the average of the daily load values, and can be calculated using the following equation 15:

$$Monthly\_LOAD\_adj_s^{m,a} = \frac{\sum_{t=1}^{v} \backslash Daily\_LOAD_{t|s}^{r,d,m,a}}{v} \quad (15)$$

iii) The monthly adjustment factor is the ratio of the series *Monthly_load* and *Adj-Monthly-Load*, as Equation 16.

$$Monthly\_Adjustment_s^{m,a} = \frac{Adj - Monthly - Load_s^{m,a}}{Monthly\_load_s^{m,a}} \quad (16)$$

iv) The Adjusted Monthly Load in the previous equation is obtained by adding the Generation of Small Power Plants, and applying the monthly correction factor (Equation 17).

$$Adjusted\_Load_s^{m,a} = \begin{pmatrix} Monthly\_LOAD_s^{m,a} + \\ Gener\_SMALL\_Plants_s^{m,a} \end{pmatrix} \cdot Monthly\_Adj_s^{m,a} \quad (17)$$

v) Finally, the Net Monthly Adjusted Load Loss is calculated by discounting the losses in the core grid from the Adjusted Monthly Load, as in the following equation (18):

$$Net\_Adjusted\_LOAD_s^{m,a} = Adjusted\_LOAD_s^{m,a} \cdot \left(1 - Losses_s^{m,a}\right) \quad (18)$$

Where $Losses_s^{m,a}$ is a percentage value of lost load in the core grid.

2.1.5 Quarterly adjusted load

This is the average of the Adjusted Net Load Loss of three consecutive months, amounting to four per year. The average between the months of January, February, and March will form the first observation of the year; the second will be formed by the average monthly values of April, May, and June, and so on. This general relation is presented in Equation 19:

$$Quart\_LOAD_s^{q,a} = \frac{\sum_{i=3\cdot q-2}^{3\cdot q} Net\_Adj\_LOAD_s^{i,a}}{3} \quad (19)$$

Where $1 \le q \le 4$.

2.1.6 Quarterly Variation of Adjusted Load in relation to the same quarter of the previous year

This indicator, presented in Equation 20, which is dimensionless, will show how much a quarter's percentage varied compared to the same quarter of the previous year.

$$Quart_{Var} - LOAD_s^{q,a} = \frac{Quart\_LOAD_s^{q,a}}{Quart\_LOAD_s^{q,a-1}} - 1 \quad (20)$$

2.2 Times series approach

The purpose of this section is to eliminate the temperature effect from the variable Load, using a time series model with the inclusion of explanatory variables. This class of models is described in detail by Hyndman & Athanasopoulos (2013). A model will be adjusted for the daily series of *Adj_Load*, using the daily series of temperature as the explanatory variable. The adjustment model is as follows (Equation 21):

$$Adj\_LOAD_{t|s}^{r,d,m,a} = \beta_0 + \sum_{i=1}^{k_1} \alpha_i \cdot Adj\_LOAD_{t-i|s}^{r,d,m,a} + \sum_{i=1}^{k_2} \beta_i \cdot Temp_{t-i+1|s}^{r,d,m,a} + \varepsilon_t \quad (21)$$

As one of the objectives of this method is to eliminate the Temperature effect from the load variable, we simply subtract the third term on the right side of the Equation 21 from *Adj_Load*. Thus, *Adj-Load* adjusted for the temperature effect is represented as follows (Equation 22):

$$Adj\_LOADT_{(t|s)}^{r,d,m,a} = Adj\_LOAD_{(t|s)}^{r,d,m,a} - \sum_{i=1}^{k_2} \hat{\beta}_i \cdot Temp_{(t-i+1|s)}^{r,d,m,a} \quad (22)$$

3 Results

This section presents the results of the proposed methodology applied to daily data from July 2003 to July 2012 for the South East / Central West subsystem, because this is the most representative one in terms of load in Brazil. It may be noted that by removing the calendar and temperature effects, as described in paragraphs 2.1.1 and 2.1.2 of the method, the load series became less volatile (Figures 1 and 2). The first

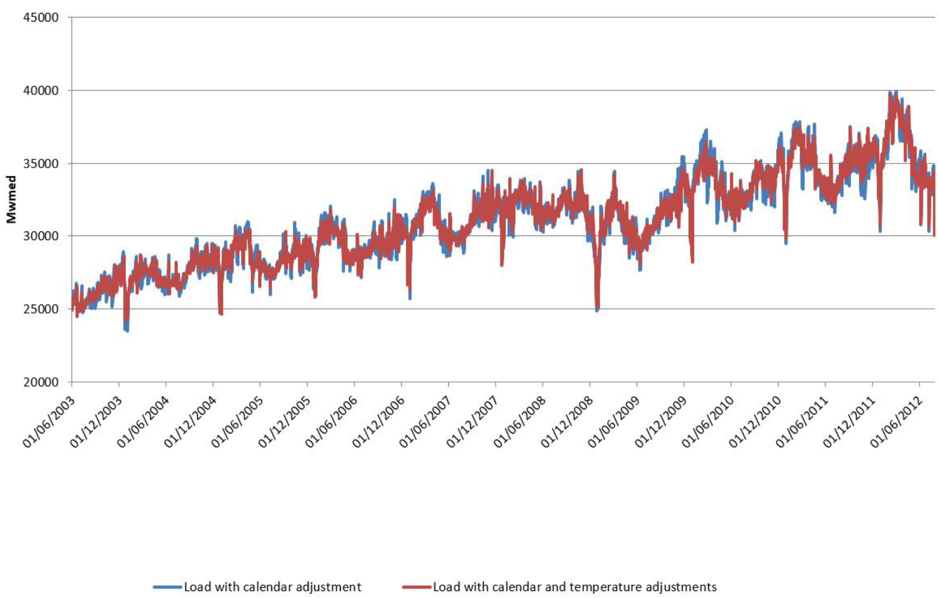

Figure 1. Load Series with Calendar adjustment and Load series with Calendar and Temperature adjustments (empirical method) for the Southeast / Midwest subsystem.

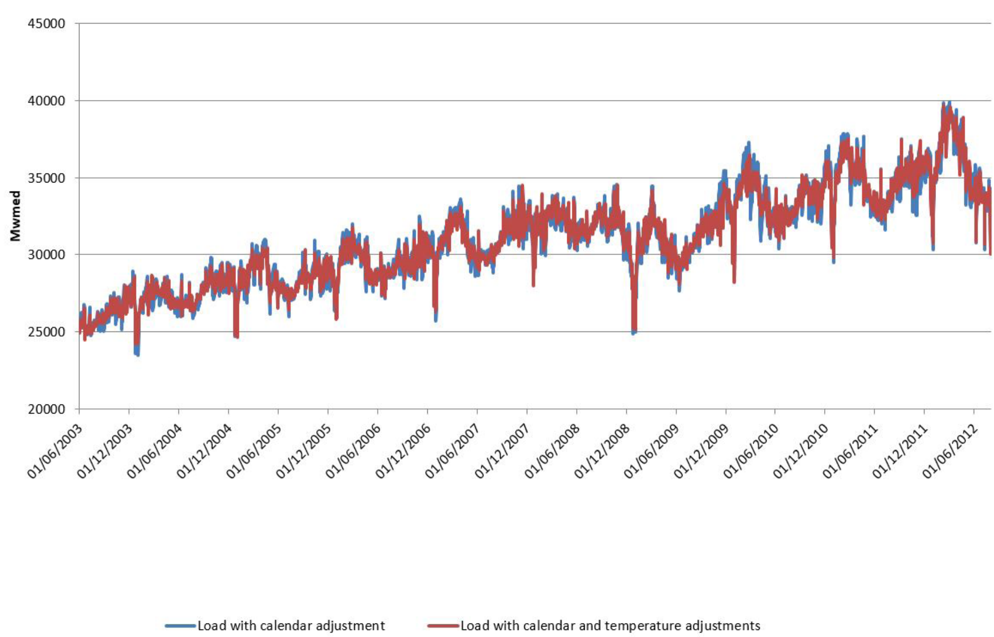

Figure 2. Load Series with Calendar adjustment and Load series with Calendar and Temperature adjustments (Time Series model) for the Southeast / Midwest subsystem.

figure shows the empirical method to remove the temperature effect, and the second shows the method using regression analysis to remove this same effect.

The monthly adjusted monthly series takes into account both the data available on a daily basis and the monthly data generated by small plants, measurement discontinuities, and losses in the core grid. These series were obtained using the methodology described in the previous section.

Next, the SIN series was obtained by adding the Adjusted Monthly Load series of the four subsystems. Figure 3 shows the SIN Monthly series, with two methods used to remove the temperature effect from the Daily Temperature series.

The quarterly series of the SIN were obtained in order to compare them with the quarterly GDP series, considering both methodologies for removal of the temperature effect, as shown in Figure 4.

At this point of the work, with the exception of economic variables, all variables that influence the load have been removed. Thus, the previously generated Quarterly series are clean regarding the Calendar, Temperature, and Dead Week effects. Hence, the Quarterly Load series that better matches GDP variations and/or better explains them will be considered better adjusted than the other. The change from one quarter to the same quarter of the previous year of the GDP series and the unadjusted Load series is shown in Figure 5.

In Figure 5, one can see that the pattern of the quarterly series of adjusted Load variation in relation to the same quarter of the previous year is similar to that of the quarterly series of GDP variation over the same period. We calculated the correlation coefficients between the quarter's load variations series with the same quarter of the previous year, as well as the series of quarterly GDP variations in relation to the same quarter of the previous year (Table 1). The quarterly load series adjusted by the Time series model showed a 92.0% correlation coefficient, while the one adjusted by the empirical method had a 90.7% correlation coefficient. However, the unadjusted quarterly Load series showed a 83.6% correlation coefficient with the quarterly GDP series. The calendar, temperature, and Dead Week adjustments added significant gains in relation to the coefficient of correlation with the quarterly GDP series, showing that these adjustments have been successfully performed to exclude random factors that influence the load series.

Another comparison device consists of explaining the quarterly GDP series by the quarterly series of adjusted Load. That is, the GDP series will be modeled, via Time Series models, by the adjusted load and its lags. The model will not include GDP and error lags. The chosen selection criteria will be the value of the coefficient of determination (R^2), and the adjusted Quarterly Load series that presents the greatest value of R^2 in modeling is the one that

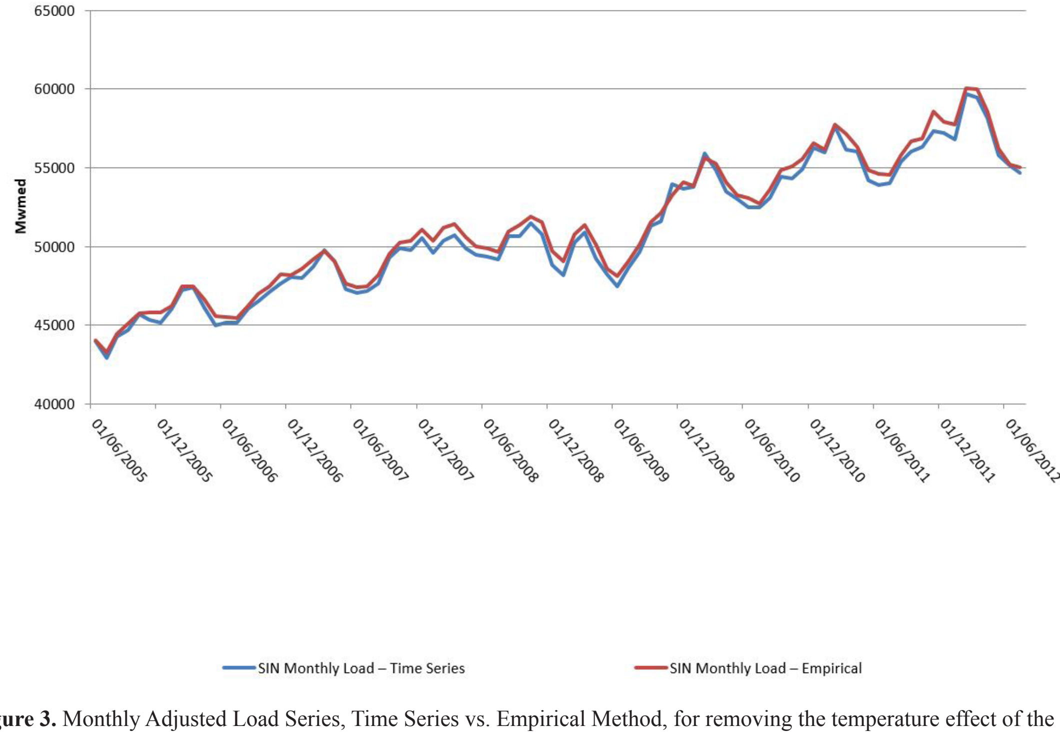

Figure 3. Monthly Adjusted Load Series, Time Series vs. Empirical Method, for removing the temperature effect of the SIN

Table 1. Correlation between the quarterly series of load variation and the quarterly series of GDP variation.

Method	Correlation
Load with adjustment – Empirical	90.7%
Load with adjustment – Time series	92.0%
Unadjusted Load series	83.6%

best fits the GDP values, and will be considered the best fitting series.

R^2 adjusted by the time series model was 89.58%, while in the empirical method it was 89.55%. Regarding the fit, it is noted that both methods are capable of filtering the load variation to remove non-economic

Figure 4. Quarterly adjusted load series vs. Empirical Method for removal of the SIN temperature effect.

Figure 5. Quarterly variance series of adjusted Load and GDP in the relation to the same quarter of the previous year.

offenders, and that there is no significant difference between these methodologies.

4 Final considerations

Knowing that the load series are influenced by several factors, including calendar, temperature, and economic effects, the series of monthly Load Adjusted and quarterly Adjusted Load Variation compared to the same quarter of the previous year—presented above—are very important to give foundation and support to the decisions of electricity distribution companies that are required to participate in regulated auctions proposed by the government. To make adjustments, daily data about load and temperature were used, provided by the SIN.

For these adjustments, we used two different approaches: a totally empirical method for removal of the Temperature and Calendar offenders from the daily and monthly series, and another approach with the same adjustments to the monthly and daily series for the Calendar effect, the exception being that the temperature effect was removed from the daily series. The temperature adjustment in the daily series in the second approach was based on a statistical modeling of time series.

With the application of the methods, it is understood that only the economic offender remains as an influence on the load series. To compare the approaches, monthly load series were consolidated on a quarterly basis and compared with the quarterly GDP series. Thus, the quarterly load series that better fits the GDP series and presents the higher correlation coefficient between these series will be better for calendar and temperature effects. To adjust the GDP series we used the Time Series model only, with the load variable and its lags as regressors.

It was noted that the difference between the adjustments is minimal, and that both series have a great explanatory power for the GDP series (R^2 values close to 90%). It is noteworthy that lagged values of errors and GDP series were not included in the modelling, showing that the load and adjusted GDP series are strongly correlated. The correlation coefficient between the load series adjusted by the second approach and the GDP series was slightly higher than that between the load series adjusted by the first approach and the GDP series, of 92.0% and 90.7%, respectively. Moreover, the correlation coefficient between the unadjusted load series and the GDP series was also calculated, being 83.6%. This shows that both approaches led to clean series calendar and temperature variations.

It is suggested, as future research, an investigation into other ways to evaluate the performance of the two methods described in this paper, applied to other series.

References

Apadula, F., Bassini, A., Elli, A., & Scapin, S. (2012). Relationships between meteorological variables and monthly electricity demand. *Applied Energy*, 98, 346-356. http://dx.doi.org/10.1016/j.apenergy.2012.03.053.

Brasil. (2002, 29 de abril). *Lei nº 10.438, de 26 de abril 2002. Dispõe sobre a expansão da oferta de energia elétrica emergencial, recomposição tarifária extraordinária, cria o Programa de Incentivo às Fontes Alternativas de Energia Elétrica (Proinfa), a Conta de Desenvolvimento Energético (CDE), dispõe sobre a universalização do serviço público de energia elétrica*. Brasília, DF: Diário Oficial da União.

Brasil. (2004a, 30 de julho). *Decreto nº 5.163, de 29 de junho de 2004. Regulamenta a comercialização de energia elétrica, o processo de outorga de concessões e de autorizações de geração de energia elétrica, e dá outras providências* (edição extra). Brasília, DF: Diário Oficial da União.

Brasil. (2004b, 16 de março). *Lei nº 10.848 de 15 de março de 2004. Dispõe sobre a comercialização de energia elétrica, altera as Leis nos 5.655, de 20 de maio de 1971, 8.631, de 4 de março de 1993, 9.074, de 7 de julho de 1995, 9.427, de 26 de dezembro de 1996, 9.478, de 6 de agosto de 1997, 9.648, de 27 de maio de 1998, 9.991, de 24 de julho de 2000, 10.438, de 26 de abril de 2002, e dá outras providências* (pp. 2). Brasília, DF: Diário Oficial da União.

Brasil. (2004c, 17 de agosto). *Decreto nº 5.184, de 16 de agosto de 2004. Cria a Empresa de Pesquisa Energética - EPE, aprova seu Estatuto Social e dá outras providências*. Brasília, DF: Diário Oficial da União.

Castro, N. J., Brandão, R., Mathias, J. F., Souza, R. C., Lobão, V., Ferreira, P. G. C., & Cyrino Oliveira, F. L. (2012). *Pesquisa sobre previsão de carga com base em cenários macro econômicos*. Rio de Janeiro: Grupo de Estudos do Setor Elétrico, UFRJ.

Castro, N. J., & Leite, A. L. S. (2008). Estrutura de governança e a formação de holdings no setor elétrico brasileiro. *Estratégia e Negócios*, 1(2), 1-15.

Hyndman, R. J., & Athanasopoulos, G. (2013). *Forecasting: principles and practice*. Recuperado em 07 de julho de 2015, de http://otexts.org/fpp/

Leite, A. L. S., & Santana, E. A. (2006). Mercado de capacidade: uma alternativa para o setor elétrico brasileiro. *Revista de Desenvolvimento Econômico*, 8(14), 23-33.

Maçaira, P. M., Silva, F.L.C., Oliveira, F.L.C., Calili, R.F., Lourenço, P. (2014). Statistical Analysis of the Brazilian Electricity Sector: a Top-down Long Range Energy Consumption and Supply Forecast Model. In *Anais do XLVI Simpósio Brasileiro de Pesquisa Operacional*. Salvador: SBPO.

Pires, J. C. L., Giambiagi, F., & Sales, A. F. (2002). As perspectivas do setor elétrico após o racionamento. *Revista do BNDES*, 9(18), 163-204.

Rede Eléctrica de España – REE. (2012). *Boletín mensual*. Recuperado em 01 de janeiro de 2012, de http://www.ree.es

Rede Eléctrica Nacional – REN. (2012). *Informação mensal do sistema eletroprodutor*. Recuperado em 01 de janeiro de 2012, de http://www.ren.pt

Reis, L. B. (2003). *Geração de energia elétrica: tecnologia, inserção ambiental, planejamento, operação e análise de viabilidade* (3. ed.). São Paulo: Manole.