Modeling and Forecasting Intraday Electricity Load

Helio S. Migon Larissa C. Alves
Universidade Federal do Rio de Janeiro

February 29, 2012

Abstract

This paper aims models electricity load curves for short-term forecasting purposes. A broad class of multivariate dynamic regression model is proposed to model hourly electricity load. Alternative forecasting models, special cases of our general model, include separate time series regressions for each hour and week day. All the models developed include components that represent trends, seasons at different levels (yearly, weekly etc.), dummies to take into account weekends/holidays and other special days, short-term dynamics and weather regression effects, discussing the necessity of nonlinear functions for cooling effects. Our developments explore the facilities of dynamic linear models such as the use of discount factors, subjective intervention, variance learning and smoothing/filtering. The elicitation of the load curve is considered in the context of subjective intervention analysis, which is especially useful to take into account the adjustments for daylight savings time. The theme of combination of probabilistic forecasting is also briefly addressed.

Keywords: Electricity load curve, Dynamic multivariate discount regression models, Factor models, Generalized Diebold and Li factor models.

1 Introduction

Short-term load forecasting has long been an issue of major interest for the electricity industry. Traditionally, hourly forecasts with a lead time of between one hour and seven days are required for the scheduling and control of power systems. From the perspective of the system operators and regulatory agencies, these forecasts are a primary input for the safe and reliable operation of the system.

Load curve forecasting is very important for the electricity industry, especially in a deregulated economy. It has multiple applications including energy purchasing and generation. For optimally operating the huge Brazilian electrical system, mainly composed of hydroelectric plants with complementary thermoelectric generation, the authorities first need to decide daily how many megawatts to produce from each alternative. In order to optimize the generation of power from a vast complex of hydroelectric plants, various types of time series models have been used to describe and simulate flows into reservoirs. For instance, a new run-off model is described in Fernandes *et al.* (2009), where applications to some Brazilian basins are presented, including comparison with alternative models.

This paper models electricity load curves for short-term forecasting purposes. Since load curves usually exhibit long-term trends, which are caused by economic and demographic factors, they are not taken into account in this application. It has long been known that electricity load has a large predictable component due to its very strong daily, weekly and yearly periodic behavior, along with meteorological-based variations.

A multivariate dynamic regression model is introduced and some of its inferential aspects are discussed, including variance learning procedure. Particular cases included in this broad class of models are versions to deal with hourly electricity load forecasting based on separate time series regressions for each hour and joint modeling of the weekly data by hour. This full multivariate dynamic regression model based on the concept of discount factor and with variance learning is a novelty in the area of electricity load forecasting. All the models developed include components that represent trends, seasons at different levels (yearly, weekly etc.), dummies to take into account weekends/holidays and other special days, short-term dynamics and weather regression effects, including nonlinear functions for cooling effects.

Special care is taken with the role played by weekends/holidays and their effect on the estimated coefficients. In particular, we explore the similarities between the shapes of load curves occurring on the day just before and just after weekends/holidays. Our developments explore the facilities of dynamic linear models such as the use of discount factors, subjective intervention, variance learning and smoothing. The elicitation of the load curve is considered in the context of subjective intervention analysis, which is especially useful to take into account the adjustments for daylight savings time. More precisely, we model a suitable transformation of the electricity load at hour τ and day t as a linear combination of the transformed load at the same hour on the previous day and appropriate functions of the prevailing temperatures on day t. The coefficients depend on the pair (τ, t) in a continuous nonlinear manner. The paper ends by briefly addressing the theme of combination of probabilistic forecasting.

The remainder of the paper is organized as follows. A review of alternative models is presented in Section 2. In Section 3 we explore the dataset analyzed. Some stylized facts always presented in the load electricity data are also reviewed. Our application to

Brazilian southeast hourly electricity loads is also presented. In Section 4, our general multivariate dynamic regression model is introduced. Some aspects of the inference and some special cases are mentioned. Our main findings are presented in Section 5 and Section 6 concludes and mentions some extensions.

2 A Review of Alternative Models

A review and categorization of electric load forecasting techniques can be seen in Alfares and Nazeeruddin (2002), where a wide range of methodologies and models often used in the current literature are discussed. The categories of load forecasting techniques considered by them, roughly in chronological order, include at least multiple regression, exponential smoothing, iterative reweighted least-squares, stochastic time series: ARMA type models, and some fashionable techniques like fuzzy logic and neural networks.

Most papers deal with 24 hours-ahead load forecasting or next-day peak load forecasting. These methods forecast power demand by using predicted temperature as forecast information. But, when the temperature curves change rapidly on the forecast day, loads change greatly and the forecasting error increases (Badran *et al.* (2008)). Typically, load forecasting can be long-term, medium-term, short-term or very short-term. This paper concentrates on short-term load forecasting dynamic regression models.

The experiences of some countries are described in the recent literature. Cancelo *et al.* (2008) present the building process and models used by Red Eléctrica de España (REE), the Spanish system operator, in short-term electricity load forecasting. The methods developed in Cottet and Smith (2003) are applied to several multiequation models of half-hourly total system load in New South Wales, Australia. An hourly periodic state space model for modeling French national electricity load is presented in Dordonnat *et al.* (2008).

A wide variety of models, varying in complexity of functional form and estimation procedures, have been proposed to improve load forecasting accuracy. Double seasonal exponential smoothing models are introduced by Taylor (2003) to make univariate online electricity demand forecasting for lead times from a half hour ahead to a day ahead. Since the time series of demand recorded at half-hourly intervals contains more than one seasonal pattern, they adapt the Holt-Winters exponential smoothing formulation to accommodate two periodicities. A comparative study including the recently proposed exponential smoothing method for double seasonality and a new method based on principal component analysis is presented in Taylor *et al.* (2006). Both time series of hourly demand for Rio de Janeiro and half-hourly demand for England and Wales were used to compare the alternative models. The new double seasonal Holt-Winters method outperforms those from traditional Holt-Winters and from a well-specified multiplicative double seasonal ARIMA model.

More advanced methods for load forecasting are developed by Cottet and Smith (2003), who adopt Bayesian procedures for forecasting high-dimensional vectors of time series. A multi-equation regression model with a diagonal first-order stationary vector autoregression for modeling and forecasting intraday electricity load is proposed. The covariance structures in such multivariate time series are of key importance for an effective forecasting strategy. They take account of the correlation between hourly loads when computing their forecasts.

A dynamic multivariate periodic regression model for hourly electricity load forecasting based on stochastically time-varying processes was developed by Dordonnat *et al.* (2008). The model consists of different equations and parameters for each hour of the day and the dependence between the equations is introduced by covariances between disturbances that drive the time-varying processes. The implementation of the forecasting procedure relies on the multivariate linear Gaussian state space framework. Although the analyses are mainly illustrated for only two hours (9 and 12), the forecasting results are presented for all twenty-four hours. Dordonnat *et al.* recognize that the unrestricted dynamic multivariate periodic regression model contains many unknown parameters and developed an effective methodology within the state-space framework that imposes common dynamic factors for the parameters that drive the dynamics across different equations. The factor model approach leads to more precise estimates of the coefficients.

Many other methodologies have been applied to forecast the electricity load curve. Special attention has been devoted to the functional data analysis. A hierarchical model for aggregated functional data is introduced by Dias *et al.*, with an application to the distribution of energy among different type of consumers. Other fashionable techniques often used are neural networks (see Alfares and Nazeeruddin (2002)) and the Gaussian process (Lourenço and Santos).

3 Preliminary Data Analysis - Some Stylized Facts

The data set available used in this paper pertains to the Brazilian southeastern submarket power consumption from 01 Jun 2001 to 20 Jan 2011. This hourly time series covers almost ten years and consists of 3521 daily or 84504 hourly observations. The models were implemented for a subset of data, starting on 01 Jan 2002 and going to 20 Jan 2011 (so 3307 days and 79368 hours). This region has 94 million inhabitants and an area of 2.530 million square kilometers, corresponding to a demographic density of

37 inhabitants per square kilometer, living predominantly in urban areas (89%), with an average daily consumption of 35000 Mw and with a large number of plants. The illustrations of the models developed are mainly based on the 2010 data only (from 03 Jan 2010 to 01 Jan 2011) to keep the figures clearer.

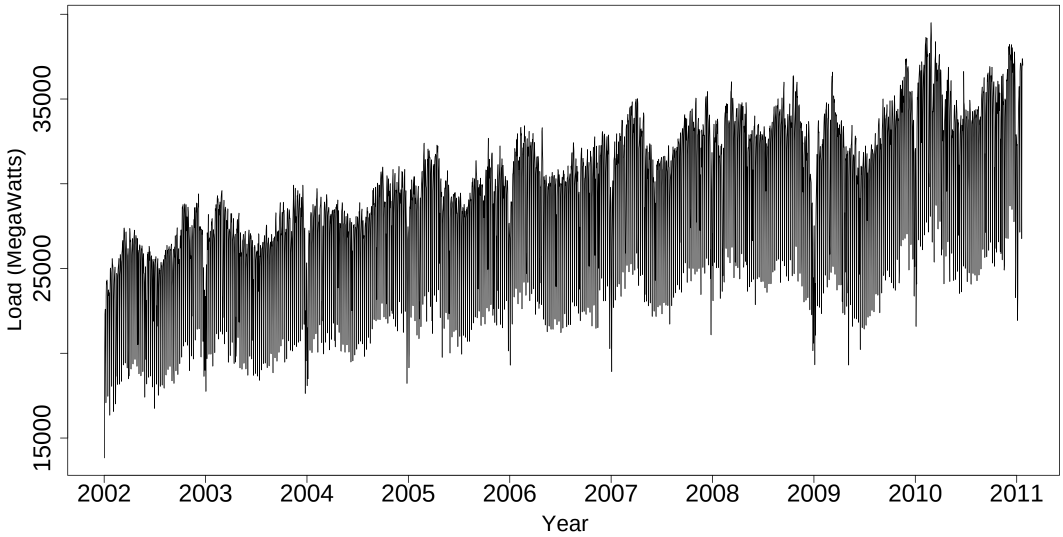
The main features of load series have been extensively reported in the literature: trend, superimposed levels of seasonality, short-term dynamics, special days, nonlinear effects of meteorological variables, possible nonlinear time dependence, etc. The system load is actually a random non-stationary process composed of thousands of individual components, whose behavior is influenced by a number of factors, as describe before, and also includes random effects.

Electricity load has a long predictable component due to its very strong daily, weekly, and yearly periodic behavior, along with meteorological-based variation. Although the meteorological variables that affect load can differ according to region, temperature appears to be by far the most important meteorological factor in most locations. Apart from the strong daily and weekly periodicity, the load profiles also vary substantially across seasons. The overall forecasting model consists of one daily model for forecasting the daily load up to ten days ahead, and 24 hourly models for computing hourly predictions for horizons up to three days. The daily model is aimed at producing forecasts for network outage planning, while the hourly models are used to derive forecasts for the next-day hourly dispatch.

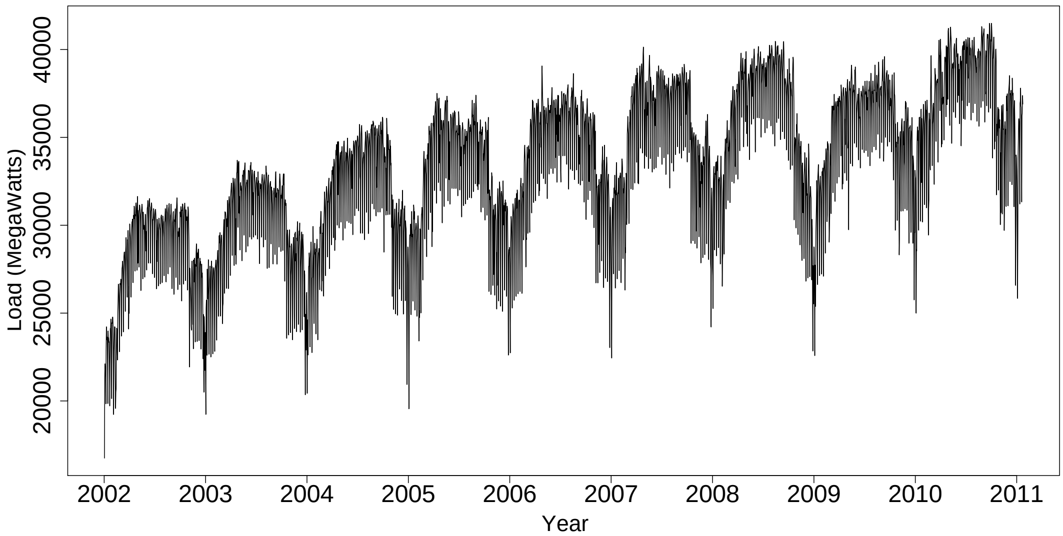
The figures below provide some evidence about the main characteristics of the electricity demand data. Figure 1 depicts the daily electricity load at 9 a.m. and at 7 p.m., for all the data available in our data set. At 7 p.m., we clearly note a smooth trend component, possibly due to the vegetative population growth and also to the economic performance, superimposed on an annual seasonal cycle. For instance, the effect of the 2008 economic crisis impacts the level of the daily electricity load only in the beginning of 2009. The same sort of trend can be seen in the 9 a.m. daily time series. The seasonal cycle is not so clear due to large volatility of this time series. Those characteristics are also present in the other hourly time series.

Figure 2 shows the daily electricity load at 9 a.m. and 7 p.m. only for the year 2010, allowing more detailed examination of the components describing the data generation process. The effect of the seasons and also some weekly seasonality are evident.

Finally, in Figure 3 we illustrate the time evolution of the electricity load curve in a three-week period, starting on Monday, 09 Aug 2008 and ending on Sunday, 29 Aug 2010, making the intraday movements clearer. It is easy to note the lower load on weekends (see at the beginning and at end of the graph) and its evolution throughout the days, with a dip around 4 p.m. on weekdays.

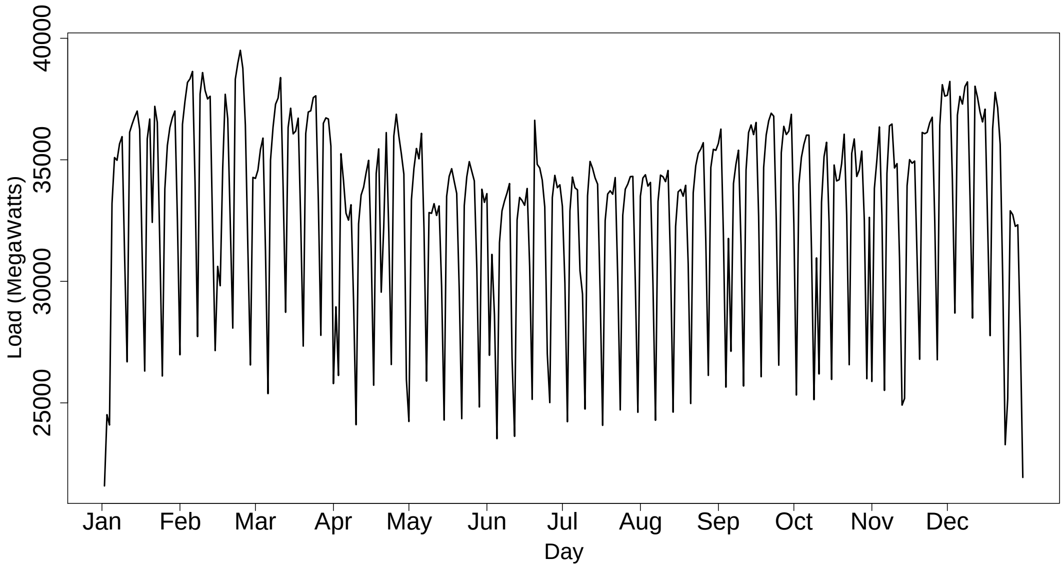


9 a.m.

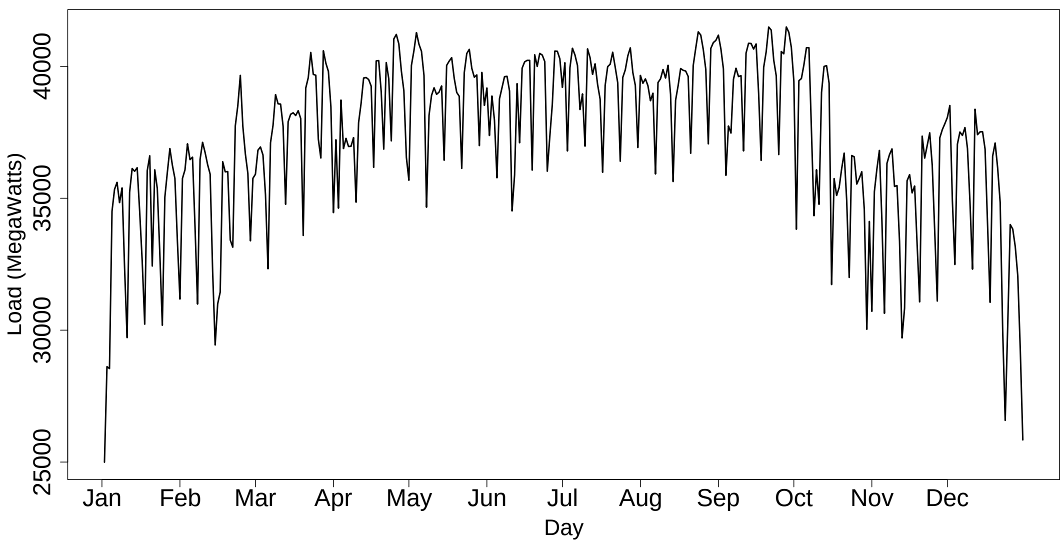


7 p.m

Figure 1: *The daily electricity load consumption, from 01 Jan 2002 to 20 Jan 2011, at 9 a.m. and at 7 p.m.*



9 a.m.



7 p.m.

Figure 2: *The daily electricity load consumption, at 9 a.m. and 7 p.m., in 2010.*

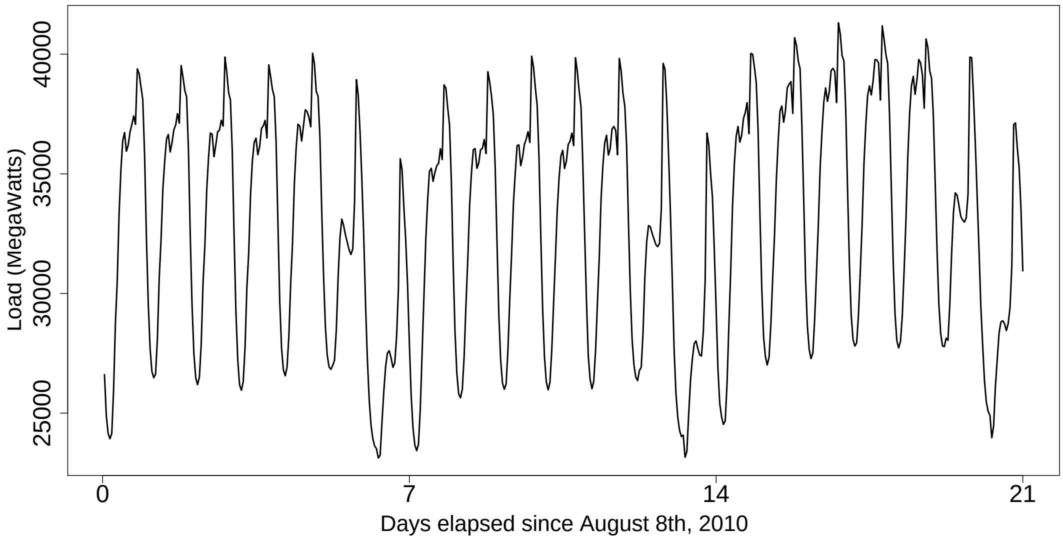


Figure 3: *The daily electricity load consumption, in three weeks starting on 9 Aug 2010.*

Temperature has two well defined impacts on load. Firstly, the daily load responds significantly to variations on temperature. On the other hand, the seasonality of electricity load and temperature is almost in phase (see Figure 4).

The effect of temperature on load is very well documented in the literature. Dordonnat *et al.* (2008) claim that a non-linear effect is present in the French data. The smoothed-heating / cooling-degrees temperature variables were built up to approximate the nonlinear relationship between electricity load and temperature in a linear relationship (see Fig. 2, page 570, Dordonnat *et al.* (2008)). The relationship between consumption and maximum temperature for the daily data during 2005 in Spain has different behavior. The relationship is U-shaped and slightly asymmetric. There is also some evidence of an exhaustion effect, especially for low temperatures (see Cancelo *et al.* (2008), Fig. 5, page 581).

Figure 5 shows the relationship between the logarithm of the daily average load versus the average temperature for the year 2010. The symbol (+) represents the consumption in the daylight saving period and the broken line shows the fitted straight line describing the relationship between the log of the mean of daily electricity consumption and the mean temperature.

Some cross-section regressions fitted to the intraday data have clearly shown that the models implemented must have time varying parameters. From those analysis we can conclude also that some dummy variables are needed to control for the type of day and the effect of holidays and daylight savings. We conclude also that there are some

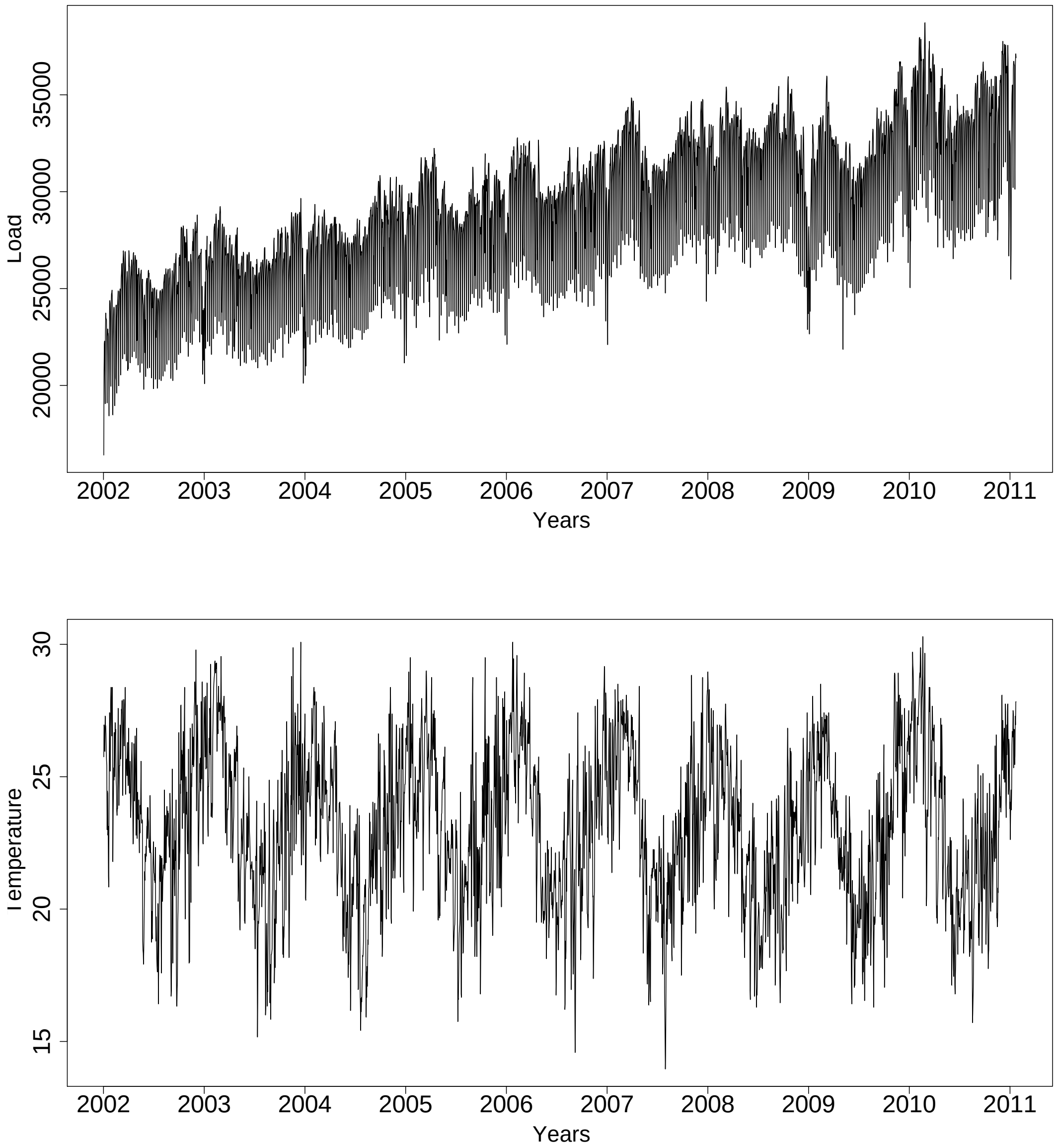


Figure 4: *The average electricity load and the average temperature, from 01 Jan 2002 to 20 Jan 2011*

seasonal effects. This way, our models are in the class of dynamic regression models. In order to keep the computational burden under control, we decided to use the principles of discount factor and some other approximations.

Separate univariate models for daily time series, at each our of the day, were developed including components for trend (linear growth models), seasonal effect (via harmonics), the effect of temperature and some dummies describing the nature of the weekday.

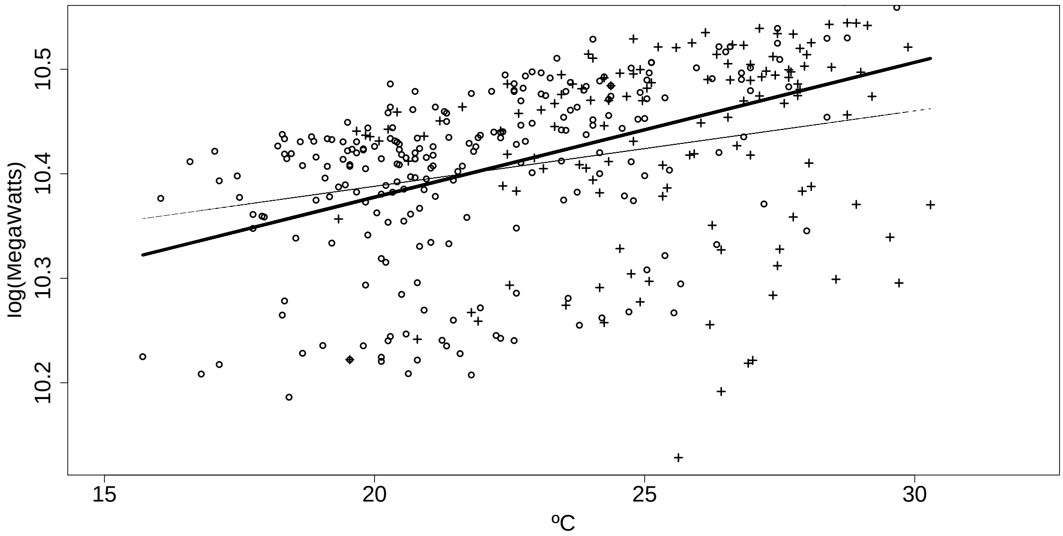
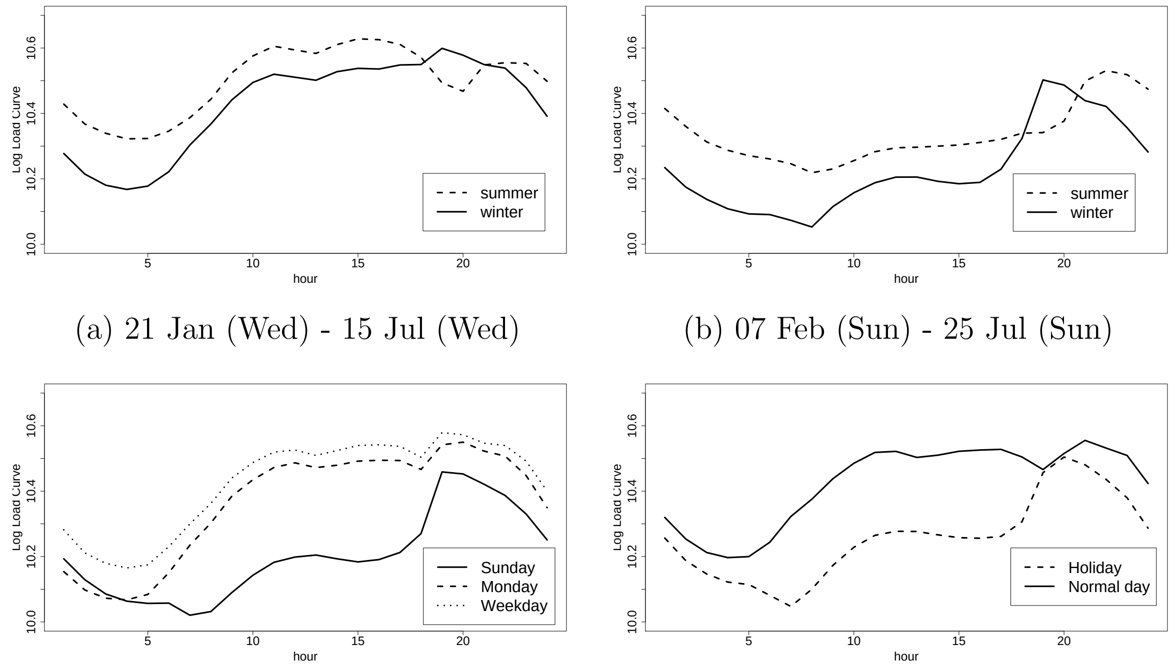


Figure 5: *Dispersion Diagram: average load versus average temperature.*

Two classes of factorial models will be discussed. One of them, the structural factor model, is based on some extensions of Diebold and Li (2006), often used in financial econometrics to describe the term structure of the interest rate. Therefore, besides the VAR component, the model introduced in this paper includes four dynamic factors based on two different decay factors (λ_1, λ_2). The effect of temperature and dummies to take care of the nature of the weekdays are also included in the model.

To conclude this section, some plots of the daily electricity log curve are presented in Figure 6. Panels (a) and (b) show how the load curve changes during the seasons on a weekday (Wednesday) and weekend day (Sunday), respectively. It is clear that the consumption is higher on weekdays than on the weekend, as can be observed in panel (c), where the figures for the 15th week of 2010 are plotted. Since the 15th week of the year corresponds to Autumn, the load curves for the corresponding weekdays are quite different from those in panels (a) and (b), which correspond to Summer and Winter. Finally, the effect of a holiday on consumption can be seen in panel (d), where both load curves refer to a Tuesday.

The class of models we are introducing in this paper encompasses many of the recent developments reviewed above. We show that our Discount Multivariate Dynamic Regression Model includes as special cases the developments of Cottet and Smith (2003) and also the recent models introduced by Dordonnat *et al.* (2008) and Dordonnat *et al.*. It is worth pointing out that our model includes two sub-classes of factor models: stochastic dynamic factor models and determinist factor models. The latter are strongly inspired



(a) 21 Jan (Wed) - 15 Jul (Wed)

(b) 07 Feb (Sun) - 25 Jul (Sun)

(c) 15th week of 2010

(d) 12 Oct (Holiday) - 19 Oct (Weekday)

Figure 6: *The electricity load curve for different days and seasons in 2010.*

by various extensions of the models developed by Nelson and Siegel (1987) in the context of financial econometrics (Diebold and Li (2006)) as discussed by De Pooter (2007), who has shown that using more flexible models leads to a better in-sample fit of the term structure of the interest rate as well as improving the out-of-sample predictability. The four-factor model, which adds a second slope factor to the three-factor Nelson-Siegel model, forecasts particularly well.

4 Multivariate Dynamic Regression Models

The general linear model structures assume that the equally spaced observations, y_t are described over time by the observation and the evolution equations:

$$y_t = F_t'\theta_t + \epsilon_t, \qquad \epsilon_t \sim N(0, \phi_t^{-1}\Sigma) \tag{1a}$$

$$\theta_t = G_t\theta_{t-1} + \omega_t, \qquad \omega_t \sim N(0, \phi_t^{-1}\Sigma \otimes W_t) \tag{1b}$$

where:

- y_t is the $m \times 1$ observation conditionally independent vector.

- θ_t is the $mp \times 1$ state vector at time t.
- F_t is an $m \times mp$ matrix of known constants or regressors at time t accommodating level, trend, seasonality, etc.
- G_t is an $mp \times mp$ matrix of known quantities describing the state time evolution.
- Σ, an $m \times m$ matrix, and W_t, a $p \times p$ matrix, are, respectively, the observational and the parameter evolution error covariance matrices.
- $\theta_0|y^0 \sim N(m_0, \phi^{-1}C_0)$ is the initial information.

For future reference, this class of models is defined by the quadruple $\{F_t, G_t, \Sigma, W_t\}$, where F_t is the design matrix, G_t is the state evolution and Σ and W_t, respectively are the observational covariance and the state evolution matrix. The observational variance is composed of a scale factor, ϕ_t, common to all the m time series, and a joint cross-sectional structure Σ. The scale factor, often assumed known, represents common measurement errors, common sample variance etc.

In our application, the observations are m dimensional for each time t, $y_t = (y_1, \cdots, y_m)'_t$ and the time varying design matrix and the p-dimensional vector of regression coefficients, for each "separate" hourly model, are defined as: $\boldsymbol{F}_t = diag(\cdots, F_{jt}, \cdots)$ and $\boldsymbol{\theta}_t = (\cdots, \theta_{jt}, \cdots)$, for $j = 1, \cdots m$, where m denotes the daily hours, so $m = 24$. Then, in terms of scalar time series elements we have m univariate dynamic models:

$$\begin{aligned} y_{t,\tau} &= F'_{t,\tau}\theta_{t,\tau} + \epsilon_{t,\tau} \\ \theta_{t,\tau} &= \theta_{t-1,\tau} + \omega_{t,\tau} \end{aligned}$$

where $\epsilon_{t,\tau} \sim N[0, \phi_{t,\tau}^{-1}\sigma_\tau^2]$, $\omega_{t,\tau} \sim N[0, \phi_{t,\tau}^{-1}\sigma_\tau^2 W_{t,\tau}]$, $\tau = 1, \cdots, m$.

Assume also that the observational variance is given by $\phi_t^{-1}\Sigma$ and the evolution variance is $\boldsymbol{W}_t = \phi_t^{-1}\Sigma \otimes W_t$, where the $p \times p$ matrix W_t is obtained using the discount factors.

We also consider the alternative model where the observational error has an autoregressive behavior. The time varying p-dimensional vectorial autoregressive process ($VAR(p)$), defined as $y_t = \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \epsilon_t$, can be stated as a dynamic linear model as follows:

$$\begin{aligned} y_t &= E'\xi_t + \epsilon_t \\ \xi_t &= G\xi_{t-1} + \omega_t \end{aligned}$$

where: $\xi_t = (y_{t-1}, \cdots, y_{t-p})'$, $\omega_t = (\epsilon_t, 0, \cdots, 0)'$, $E = (I_m\, 0 \cdots 0)'$ and $G = \begin{pmatrix} \Phi_1 & \cdots & \Phi_{p-1} & \Phi_p \\ I_m & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & I_m & 0 \end{pmatrix}$.

In the present application we limit ourselves to the case of $p = 1$. Then in our model the error term is given as $\epsilon_t = \Phi\epsilon_{t-1} + a_t$, where $a_t \sim N[0, \Sigma]$, which is equivalent to modeling $y_t - \Phi y_{t-1}$.

As is well established in the econometric literature, the VAR(p) provides a flexible description of dynamic interaction in a model system, while using initial information very sparsely - only the variables that enter the system and the lag length need to specified. One problem consists just of select the adequate order to capture the dynamic structure. The Litterman prior (Litterman (1986)) is one of the earliest priors proposed. It is appealing as an easy and accurate expression of prior beliefs about the nature of the dynamics involved in the data generation process. It is specified via some hyperparameters describing the tightness of beliefs on the own variables lag, on the others variables and a rate of beliefs tightness with increasing lag. In the modern literature the model selection is approached via some special hierarchical priors. Comparing the performance of the Litterman prior and Bayesian lasso (Kyung *et al.* (2010)) is part of our research agenda.

4.1 Inference in Multivariate Dynamic Linear Models

The Bayesian inferential procedure is sequential by nature, so it combines well with data ordered by time. The main steps involved in the inference include the computation and summarization of posterior distributions (forward filtering) as well as predictive inferences for future observations (forecast distribution).

At any time t we have various distributions available and the analysis allows specific ways to change these distributions over time. Let the information available at time t be denoted by $y^t = (y^0, y_1, \ldots, y_t)$, where y^0 is the initial information. The following distributions are of interest:

- $p(\theta_{t+h}|y^t))$: the prior density, if $h > 0$, for the state vector at time $t + s$ given the information up to time t; the smoothing density if $h < 0$; and the posterior density at time t if $h = 0$.

- $p(y_{t+h}|y^t)$: the h steps ahead forecasting distribution for the future observation.

Assuming a conjugate normal-inverse Wishart prior, $\theta_0, \Sigma \sim N - IW(m_0, C_0, n_0, D_0)$, where $n_0 > 0$ is the initial degree of freedom and D_0 is the $m \times m$ sum of squares matrix, that is $\Sigma|y^0 \sim IW(n_0, D_0)$, with harmonic mean given by $E[\Sigma^{-1}|y^0]^{-1} = \frac{D_0}{n_0+m-1}$ and $E[\Sigma|y^0] = \frac{D_0}{n_0-2}, n_0 > 2$. The one-step-ahead predictive distribution and the posterior distributions, for all t, are given by:

(a) Posterior at $t-1$:

$$\theta_{t-1}, \Sigma|y^{t-1} \sim NIW(m_{t-1}, C_{t-1}, n_{t-1}, D_{t-1})$$

where m_{t-1} is an $mp \times 1$ mean vector, C_{t-1} is an $mp \times mp$ covariance matrix and d_t is an $mp \times mp$ sum of squares matrix. Define $S_{t-1} = \frac{D_{t-1}}{n_{t-1}}$.

(b) Priori at t:

$$\theta_t, \Sigma|y^{t-1} \sim NIW(a_t, R_t, n_{t-1}, D_{t-1})$$

where $a_t = G_t m_{t-1}$, $R_t = G_t C_{t-1} G_t' + W_t$.

(c) one-step-ahead forecast distribution:

$$y_t|y^{t-1} \sim T_{n_{t-1}}(f_t, Q_{t-1})$$

where $f_t = F_t' a_t$ and $Q_t = F_t' R_t F_t + S_{t-1}$

(d) Posteriori in t:

$$\theta_t, \Sigma|y^t \sim NIW(m_t, C_t, n_t, D_t)$$

where $m_t = a_t + A_t e_t$, $C_t = R_t - A_t Q_t A_t'$, $n_t = n_{t-1} + 1$ and $D_t = D_{t-1} + e_t Q_t^{-1} e_t'$, with $A_t = R_t F_t Q_t^{-1}$ and $e_t = y_t - f_t$

Some special cases

The above model can accommodate many different sub-models by some careful specification of the components.

- The state θ_t evolution covariance component matrix, W_t, can be indirectly specified through the use of the discount factor principles. The discount factor is typically a number in $(0, 1)$ which can be interpreted as the information percentage that passes from time $t-1$ to t. Then $W_t = (\frac{1}{\delta} - 1)C_{t-1}$, with $R_t = (G_t C_{t-1} G_t')/\delta$. Without loss of generality, the same discount could be applied to each separate regression component, keeping the covariance among regression blocks fixed.

- All the results presented below are based on the special case of separate discounted dynamic regression models for each hour of the day. This corresponds to setting $\Sigma_t = I_m$ and defining $W_{jt} = (\frac{1}{\delta} - 1)C_{j,t-1}$, where $C_{j,t-1}$ is the block of the covariance matrix corresponding to the j^{th} hour.

- ◦ The model described in Cottet and Smith (2003) is a special case of the above model with $\epsilon_t = \Phi\epsilon_{t-1} + a_t$, where $a_t \sim N(0, I)$ and the regression coefficients are time invariant, which corresponds to setting the discount factor equal to 1.
- ◦ The Dordonnat *et al.* (2008) model is also cast in state space form, but it differs from ours since $\Sigma_t = I_m, \forall t$, which corresponds merely to separate regressions. Besides that, it does not include any AR component and the estimation is done through the maximum likelihood approach.
- ◦ The model in Cancelo *et al.* (2008) follows the traditional ARIMA model and of course can be viewed as a particular case of our general model.

4.2 Model Components

In this section we introduce the main components of our full multivariate dynamic regression model. We start with the class of factorial models. We distinguish two alternative factorial models: the structural and the latent factor models. The first is mainly described by observable factors, which can be interpreted as a base of functions approximating the "smoothing" load curve for each day. From the financial econometric literature we borrow the structural factor model (De Pooter (2007)), which will be specified below. This is a base of functions parameterized by some exponential decay factors which follows as a natural extension of the classical Nelson-Siegel three-factor model, as discussed in Diebold and Li (2006).

To model the intraday electricity load, $y_t(\tau)$, where t denotes a weekday and $\tau \in \{1, \cdots, m\}$ an hour, we define the $m \times 4$ regression design matrix $F_1 = (f_1(1) \cdots f_1(m))'$, to represent the factorial load, with components:

$$f_1(\tau) \quad = \quad (g_1(\tau, \lambda_1), g_2(\tau, \lambda_1), g_3(\tau, \lambda_1), g_4(\tau, \lambda_2))$$

where: $g_1(\tau) = 1$, $g_2(\tau, \lambda_1) = \frac{1-e^{-\tau\lambda_1}}{\tau\lambda_1}$, $g_{j+2}(\tau, \lambda_j) = \frac{1-e^{-\tau\lambda_j}}{\tau\lambda_j} - e^{-j\tau\lambda_j}$, $j = 1, 2$, with $\lambda_1 > \lambda_2$.

The time varying regression coefficients, $\underline{\theta}_1 = \begin{pmatrix}\theta_{1t} & \theta_{2t} & \theta_{3t} & \theta_{4t}\end{pmatrix}'$, are interpreted as the factors and represent the level, growth and curvature. As usual in dynamic regression models, the G_1 matrix is the identity. This generalization is introduced in De Pooter (2007) to increase the flexibility and fit of the Nelson-Siegel model by adding a second hump-shape factor with a separate decay. The introduction of a second medium-term component to the model makes it easier to fit load curve shapes, which typically have more than one local maximum or minimum along the daily hours. In this representation, we have factors interpreted as the level, growth factors and two describing curvatures. The curvatures differ on the decay factor parameters λ_1 and λ_2. The decay parameters are kept as fixed throughout time. The values of these decay factors were obtained after a sensitivity study and plugged in to keep the model easy to implement.

Now we describe the vector autoregressive component as specified in section 4. The matrix of VAR coefficients, Φ, of dimension $m \times m$, is part of the state vector, and corresponds to the block (F_{2t}, G_2), with parametrization $\underline{\theta}_2 = vec(\Phi)'$ defined as:

$$F_{2t} = \begin{pmatrix} y'_{t-1} & & 0 \\ & \ddots & \\ 0 & & y'_{t-1} \end{pmatrix} \quad \text{and} \quad G_2 = I_{m^2}$$

where F_{2t} is an $m \times m^2$ matrix.

Additionally, we include a linear growth component, the covariate temperature, which varies with t and τ, and dummy variables to take into account weekends, holidays etc.

The full model, with parametrization $\theta_t = \begin{pmatrix} \underline{\theta}_1 & \theta_5 & \theta_6 & \dots & \theta_{13} & \underline{\theta}_2 \end{pmatrix}'_t$ is defined as:

$$F_t = (F_1, X_t, \mathbf{0}, \dots, \mathbf{0}, F_{2t}) \quad \text{and} \quad G_t = \begin{pmatrix} G_1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & I_3 & Du_t & 0 \\ 0 & 0 & 0 & I_8 & 0 \\ 0 & 0 & 0 & 0 & G_2 \end{pmatrix}$$

where X_t is the temperature on day t, $\mathbf{0}$ is a vector of zeros, with dimension m, and $Du_t = diag[d_{1t}, d_{2t}]$ denotes the block diagonal of the dummies matrix, with $d_{1,t} = I_{Weekend}(t) \quad d_{2,t} = I_{Weekday}(t)$, where $I_x(A) = 1$ iff $x \in A$, and where Weekday is composed by the days of the week except Monday. Note that the state vector is expanded to include the effect of the dummy variables and the matrix of the autoregressive process. Those covariates affect the load curve, modifying the factors or the regression coefficients. The effect of the dummy variables is to modify the factors in the presence of those special days. This is another novelty introduced in the analysis of electricity load data.

4.2.1 *Univariate Models*

Although univariate models can be viewed as particular cases of the former model with $m = 1$, some modifications should be detailed. Since now we are taking into account hourly time series, the polynomials do not apply anymore. On the other hand, in the preliminary data analysis we saw the presence of strong seasonal components in the data generating process. Using the principle of model superposition, our general univariate model is parameterized as: $\theta = (\theta_1, ...\theta_6)'_t$ where $\theta_{1t} = (\mu_{1t}, \mu_{2t})'$, $\theta_{2t} = (\beta_{1t}, \beta_{2t})'$, $\theta_{3t} = (a_{1t}, b_{1t})'$, $\theta_{4t} = (a_{2t}, b_{2t})'$, $\theta_{5t} = (a_{3t}, b_{3t})'$ and $\theta_{6t} = (\gamma_{1t}, \gamma_{2t})'$. $F_t = (F_{1t}, F_{2t}, F_{3t}, F_{4t}, F_{5t}, F_{6t})$ and $G_t = diag(G_1, G_2, G_3, G_4, G_5, G_6)$.

The first component represents the linear growth and is defined by the pair:

$$F_{1t} = (1, 0) \quad \text{and} \quad G_{1t} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

The second component includes two regressors: one representing the $AR(1)$ term and the other the coefficient of the variable temperature. This block is defined as: $F_{2t} = (T_t, y_{t-1})$ and $G_{2t} = I_2$. All the models include the linear growth component, the AR term and also the temperature covariate.

The third and fourth components are seasonal terms and so are defined by

$$F_{jt} = (1, 0) \quad \text{and} \quad G_{jt} = \begin{pmatrix} cos(2\pi s/p) & sen(2\pi s/p) \\ -sen(2\pi s/p) & cos(2\pi s/p) \end{pmatrix}$$

with $j = 3$, 4 or 5. Each of them has a different seasonal period. The first is a daily seasonal component ($p = 365.25$) and the other a weekly component ($p = 7$). For the weekly component, we considerer the first two harmonics ($s = 1$ *or* 2), while for the daily component we only consider the first harmonic ($s = 1$).

To describe the effect of the weekday we use the dummies, $d_{1,t}$ and $d_{2,t}$, previously defined.

Then, representing it as a DLM we have: $F_{6t} = (d_{1t}, d_{2t})$ and $G_{6t} = I_2$.

The univariate DLM model is used to describe the time series of electricity load at each hour of the day, throughout the days. After running 24 separate DLMs (one for each hour), the one-step-ahead predictions are collapsed together to illustrate the daily load curve. Of course a drawback is that the intraday correlations are not directly considered. However, the predictions are obtained almost instantaneously and their precision is compatible with that of similar models in the literature.

4.3 Practical Aspects of DLM Modeling

Some special aspects involved in dynamic Bayesian modeling include variance law, discount factor, smoothing, intervention and monitoring (Migon *et al.*). Model average or model selection are also extensively discussed in the literature (see West and Harrison (1997), chapter 12).

The use of a discount factor is often recommended to avoid the difficult task of directly setting the state parameters evolution matrix. These are fixed numbers between zero and one describing subjectively the loss of information through time.

Other relevant aspects of dynamic linear models are to easily deal with missing observations and to automatically implement subjective interventions. In the first case, it suffices not to use the updating equations when the observations are missing. In this way, the uncertainties increase with the evaluation of the new prior distribution and the recurrence equation continues to be valid without any additional problem. From the intervention point of view, the simplest proposal is to use a small discount factor, close to zero, at the time of announced structural changes in the data generation process. In this way, the more recent observations will be strongly considered in the updating of the prior distribution and the system can be more adaptive to possible changes.

4.3.1 Anticipatory Intervention Analysis

An anticipatory intervention can be described as including an extra component in the evolution noise or directly changing the prior moments. Denote an intervention by $I_t = \{a_t^i, R_t^i\}$, where a_t^i, R_t^i are respectively the elicited prior mean and variance. It is worth pointing out that the interventions in the state equation evolution have a permanent effect.

In the first approach, let $\eta \sim t_{n_t}(h_t, H_t)$ be the extra term added to the evolution error term, where t_{n_t} denotes a Student-t distribution with n_t degrees of freedom. So, formally we have: $\theta_t|D_{t-1} \sim t_{n_t}(a_t^i, R_t^i)$, where $a_t^i = a_t + h_t$ e $R_t^i = R_t + H_t$. If some components of the elements of h_t and $diag[H]_t$ are nil, those component of the state vector do not change.

The alternative approach is more interesting and closely related with the recent developments in elicitation of probability distributions (Garthwaite *et al.* (2005)) and allows an expert to determine in a coherent way the quantities a_t^i and R_t^i, anticipating the uncertainty effects that could occur. Our concern in this paper is to anticipate, in a subjective way, the effect of a known future event that will substantially change the form of the electricity load curve, as for example daylight saving time.

Let us assume that an expert is able to know in advance how some relevant points in the load curve will change. If our DLM model for the load curve has p parameters, then we ask the expert to predict the load curve at time t for p different values of the hourly index, τ. Suppose he is able to produce the forecast mean values: $\mathbf{f}^i = f(\tau_1^i), \cdots, f(\tau_p^i)'$. Then following Bedrick *et al.* (1996), we obtain: $a_t^i = X_t^{-1}\mathbf{f}^i$, where X_t is the $p \times p$ design matrix, which, without loss of generality, is assumed to be full rank p. Let us assume also that the uncertainty associated with this evaluation will be described by a discount factor δ^i. Then $I_t = \{a_t^i, R_t^i\}$ where $R_t^i = R_t/\delta^i$.

In order to show that this form of intervention corresponds to a second evolution over the original state parameters, let us define $\theta_t^i = \frac{\theta_t}{(\delta_t^i)^{1/2}} + (a_t^i - \frac{a_t}{(\delta_t^i)^{1/2}})$. From lemma 11.1 in West and Harrison (1997) it follows that $E[\theta_t^i] = a_t^i$ and $V[\theta_t^i] = R_t^i$. It is worth noting that θ_t^i is like a second evolution. Therefore, in terms of the original parameters we have the evolution $\theta_t = G_t^i \theta_{t-1} + \omega_t^i$ where $G_t^i = \frac{G_t}{(\delta_t^i)^{1/2}}$, $\omega_t^i = \frac{\omega_t}{(\delta_t^i)^{1/2}} + h_t$, with $h_t = a_t^i - \frac{a_t}{(\delta_t^i)^{1/2}}$ and $W_t^i = \frac{W_t}{\delta_t^i}$.

4.3.2 Bayesian Model Average: forecasting combination

Let m alternative models be described by the probability distributions $p(y|M_i, D)$, where M_i denotes the i^{th} model, D the observables and y a *future* observation, not yet observed. Denote the prior over the alternative models by $p(M_i)$. Applying Bayes theorem we easily obtain $p(M_i|D)$ as

$$p(M_i|y, D) \propto p(M_i|D)p(y|M_i, D)$$

Those posterior distributions are used in at least two alternative ways. The first corresponds to choosing the most probable model and the other to combining the alternative models before making predictions. Then

1. The best model is the one with maximum posterior probability, that is: $M^* = argmax_{M_i}\ p(M_i|y, D)$. Note that in this case all the predictions will be based on the conditional distribution: $p(y|M^*, D)$. Conditioning on a single selected model ignores model uncertainty, and thus leads to underestimation of uncertainty when making inferences about quantities of interest.

2. Using the model average alternative leads to the predictive distribution:

$$p(y|D) = \Sigma p(M_i|D)p(y|M_i, D)dM_i$$

This alternative seems to be more attractive, since often we do not know what the true model is. Then the best we can do is combine them using the posterior probabilities as weights (Clemen (1989)).

5 The Main Findings

The data set analyzed in this paper correspond to the hourly electricity consumption from 01 Jan 2002 to 20 Jan 2011. The graphical illustrations of the one-step-ahead prediction of the load curve refer only to week 15 of 2010. Logarithm transformation is used to facilitate the models' fit, although we are aware of all the inconvenience this can cause to decision makers. In the next two subsections we present some results on sensitivity analysis, model selection and also the predictive ability of our models. The models specified in the previous sections were applied and the main findings are graphically illustrated.

This section includes the main results about model selection. Altogether we compare seven univariate models, the simplest one being a linear growth plus an AR component and the effect of temperature. The other univariate models differ from this basic one by the inclusion of seasonal components (daily and weekly) and dummies characterizing the weekday type. From the multivariate standpoint, we compare two alternatives. The simplest one only includes the terms describing the extension of the Nelson-Siegel polynomials. This one gives origin to another one by the inclusion of the dummies. In this application, the multivariate model was further simplified. We assume that the covariance matrix is $\phi_t\Sigma = \phi I_m$. Therefore, the prior distribution for ϕ is assumed to be an $IG(5/2, 0.2/2)$ and the initial information is stated as $m_0 = 0$ and $C_0 \sim N(0, 10^5 I_{mp})$. Although we performed many simplifications in order to make feasible the inference in closed form, the computational time is very large due to the huge dimension of the state vector. All the calculation were developed in an Intel Core 2, 2.53 GHz and 4.00 GB of RAM, taking more than one hour and twenty minutes to process the multivariate model.

The two criteria used for this model selection are the root mean square error (RMSE) and log of predictive likelihood (LPL). Of course one model is better than the other if it has smaller RMSE or if it has greater LPL .

On the other hand, our main interest is to evaluate the predictive ability of the models to forecast the electricity load for one day ahead. This is accomplished by examination of the graphs below. Although our data set covers the period starting in 2001 and ending in 2011, only the years 2008 up to 2010 were selected to evaluate the models' performance numerically. For the univariate model we concentrate only on the time series of the loads at 9 a.m. and 7 p.m. and we elect the 15th week of 2010 to present the graphic illustrations of the load curves forecast by the univariate and multivariate models.

5.1 Univariate model selection

In this section we present an analysis of the impact of different discount factors when applied to various univariate models. The model comparison is based on the predictive likelihood, LPL, and on the root mean square errors, RMSE , for the one-step-ahead forecasting.

The main conclusion from the sensitivity analysis is that it is bestworth to use a discount factor of 0.95 in all cases. We choose the time series of the consumption of electricity at 9 a.m. (less variability) and 7 p.m. (more volatile) to illustrate our findings. The data analyzed reflect the effects of the international economic crisis of 2008. In Table 1, at 7 p.m., it is clear the importance of using a discount factor for the simplest model. The same conclusions apply to all the other models. Therefore, we decided to use the same discount factor for all the models

Table 1: Sensitivity analysis for the simplest model (M1) and full model (M7) at 9 a.m. and 7 p.m. based on the RMSE and on the LPL using the data from 2008 to 2010.

M1: Trend + Temp + AR					
		$\delta = 0.85$	$\delta = 0.90$	$\delta = 0.95$	$\delta = 0.99$
9 a.m.	RMSE	0.156	0.141	0.129	**0.123**
	LPL	526.602	608.080	689.606	**744.416**
7 p.m.	RMSE	0.069	0.063	**0.059**	0.062
	LPL	1404.044	1475.057	**1535.813**	1496.854
M7: M1 + Harm(d) + 2 Harm(w) + Dummies					
		$\delta = 0.85$	$\delta = 0.90$	$\delta = 0.95$	$\delta = 0.99$
9 a.m.	RMSE	0.100	0.079	0.064	**0.054**
	LPL	1282.601	1422.524	1574.717	**1655.778**
7 p.m.	RMSE	0.049	0.043	0.038	**0.035**
	LPL	1858.421	1986.359	2108.915	**2145.757**

In Table 2 we can note that the inclusion of dummy variables significantly improves the models´ performance. The RMSE of the models without dummies is bigger than the ones for models including the dummy variables (see, for instance, models $M4$, $M5$ and $M6$). The same happens with the LPL criterion. The conclusions are similar for both time series and model 7 is slightly better, consistently with both criteria.

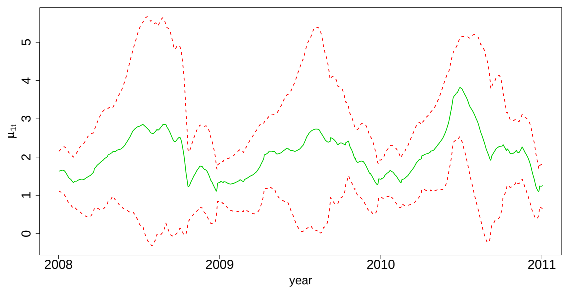
The smoothed mean coefficients for the main factors (level, temperature, daily seasonality and dummy coefficients) in model 7 can be seen in Figure 7, covering the period from 01 Jan 2008 to 01 Jan 2011.

Figure 7 shows the smoothed mean obtained with model 7, for all the model coefficients: level, temperature, the $AR(1)$ component, the daily seasonalities, and the dummies. The graphs cover the period from Jan 2008 to Jan 2011 (three years), allowing rough observation of some annual movements present in those components. The level coefficient, μ_{1t}, steadily decreases along the year and the uncertainty is bigger in the winter than in the summer. The same behavior is noted in the AR component. Although it is not easy to see, there are some changes of seasonality along the years. The dummy coefficients have a complementary time evolution. While the weekend dummy, respectively, decreases/increseas at the end/beginning of each year, the weekday dummy, excluding

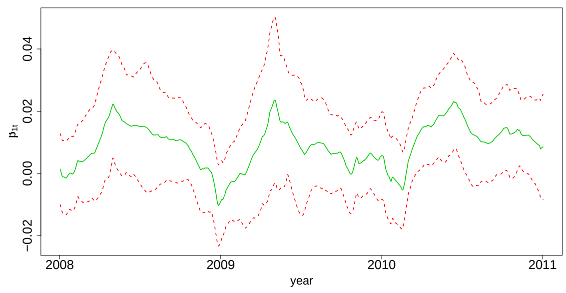
Table 2: Model selection for the 9 a.m. and 7 p.m. models based on the RMSE and on the LPL using a discount factor equal to 0.95.

	Model	*RMSE*	*LPL*
M1:	Trend + Temp + AR (9 a.m.)	0.129	689.606
M2:	M1 + Harm(d) (9 a.m.)	0.136	640.718
M3:	M2 + Harm(w) (9 a.m.)	0.101	953.340
M4:	M1 + Dummies (9 a.m.)	0.070	1465.784
M5:	M2 + Dummies (9 a.m.)	0.072	1415.683
M6:	M3 + Dummies (9 a.m.)	0.073	1413.417
M7:	M2 + 2 Harm(w) + Dummies (9 a.m.)	**0.064**	**1574.717**
M1:	Trend + Temp + AR (7 p.m.)	0.059	1535.813
M2:	M1 + Harm(d) (7 p.m.)	0.62	1500.305
M3:	M2+ Harm(w) (7 p.m.)	0.051	1727.553
M4:	M1+Dummies (7 p.m.)	0.042	2003.733
M5:	M2 + Dummies (7 p.m.)	0.043	1961.336
M6:	M3 + Dummies (7 p.m.)	0.042	1981.940
M7:	M2 + 2 Harm(w) + Dummies (7 p.m.)	**0.038**	**2108.915**

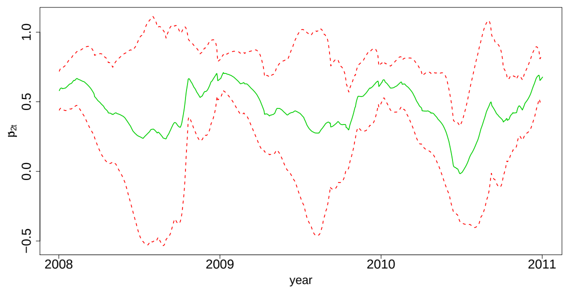
Monday, has the opposite behavior. Note that the dummies´ effects, in general, are negative.



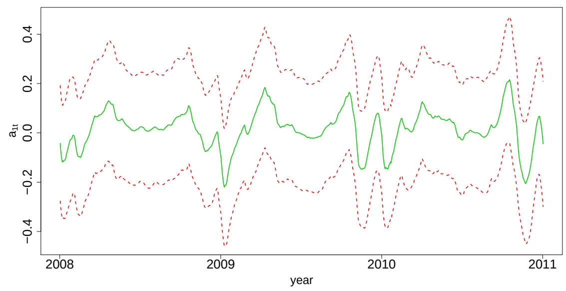
μ_{1t} - level



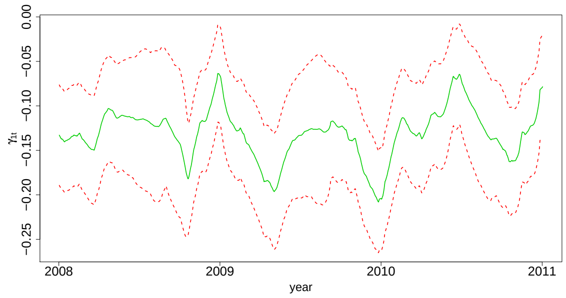
β_{1t} - tempetature



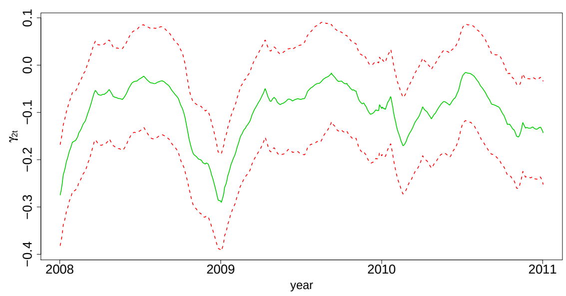
β_{2t} - autoregressive



a_{1t} - daily seasonal



γ_{1t} - dummy for weekend



γ_{2t} - dummy for weekday

Figure 7: *Smoothed mean value with 95% credibility intervals for model* M7, *at 7 p.m.*

5.2 Univariate Models - Daily Assessment

After obtaining the one-step-ahead forecast for each separate dynamic regression model, we join them to produce the load curve forecast.

We arbitrarily choose the 15th week of 2010 to graphically illustrate the electricity load curve predictive behavior using our preferred univariate model, for instance, model 7. For the numerical results, we use the years 2008 to 2010. Table 3 shows the RMSE and LPL , based on this period and on the relevant univariate models, M3, M4, M6 and M7. The two criterion functions used are defined as:

$$RMSE(k) = \sqrt{(mK)^{-1} \sum_{t=1}^{K} \sum_{\tau=1}^{m} (y_{k,\tau} - f_{k,\tau})^2}$$

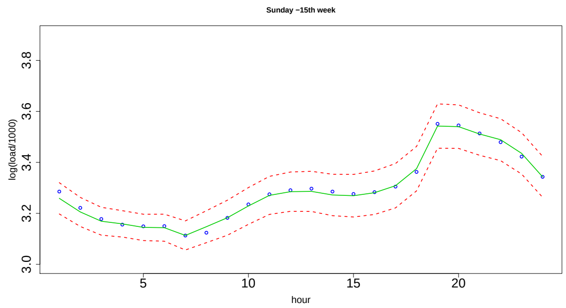
$$LPL(k) = -\left\{ \sum_{t=1}^{K} \sum_{\tau=1}^{m} \left[\frac{1}{2} log(2\pi q_{k,\tau}) + \frac{(y_{k,\tau} - f_{k,\tau})^2}{2q_{k,\tau}} \right] \right\}$$

where: $f_{k,\tau}$ and $q_{k,\tau}$ are the mean and variance of the predictive distribution, at hour τ on day $k = 1, \cdots, 7$, $m = 24$ and $K = T/7$. The main results obtained can be seen in the next table. Model 7 is the best following the LPL and the RMSE criteria. So we will use model 7 in the graphical illustrations.

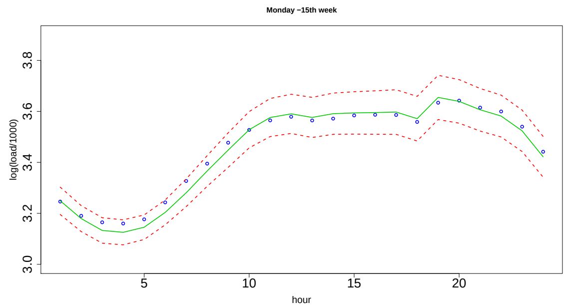
Table 3: Univariate model selection, based on the $RMSE$ and LPL, normalized by the numbers of days, using a discount factor equal to 0.95.

		Monday	Weekday	Weekend
M3	RMSE	0.057	0.061	0.101
	LPL	5198.681	5281.029	2416.219
M4	RMSE	0.047	0.052	0.050
	LPL	6623.291	6078.908	5863.215
M6	RMSE	0.053	0.054	0.050
	LPL	5986.958	5984.507	5851.575
M7	RMSE	**0.042**	**0.051**	**0.043**
	LPL	**6903.421**	**6169.371**	**6451.267**

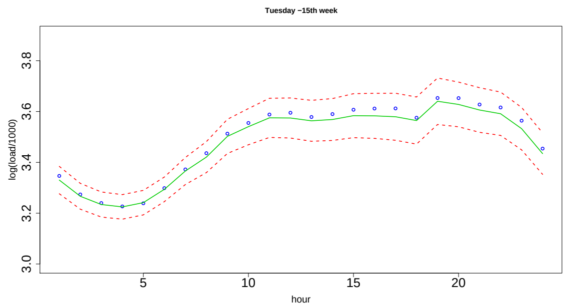
Figure 8 shows the mean of the one-step-ahead forecast distribution, based on model 7, for the 15th week of 2010, based on a discount factor $\delta = 0.95$. Note that the 95% credibility intervals (dashed lines) include almost all the observed values (dots).



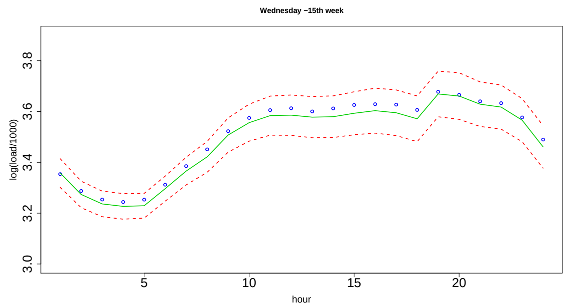
Apr 11, 2010



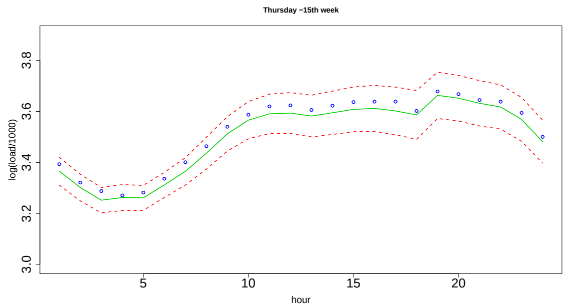
Apr 12, 2010



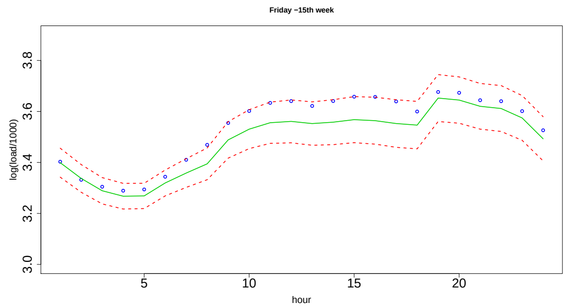
Apr 13, 2010



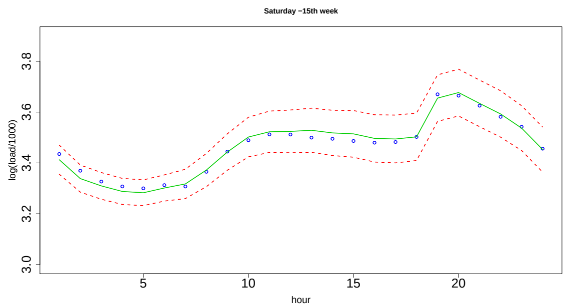
Apr 14, 2010



Apr 15, 2010



Apr 16, 2010



Apr 17, 2010

Figure 8: *One-step-ahead load curve forecasting (solid line) with* 95% *credibility interval (dashed line), Model 7, week 15, 2010* ($\delta = 0.95$)*. The observations are in dots.*

5.3 Multivariate Models Selection

As previously mentioned, multivariate models are time demanding. Therefore, the sensitivity analysis was only partially developed. First we chose the λ´s, evaluating the RMSE and LPL criteria for each weekday and some hours of the day. This was done using the best discount factor obtained for the univariate models and applying only the simplest multivariate model, that is, without dummies. We ended up with two solutions with almost the same value for both criteria. Therefore, we decided to vary the discount factor from 0.9 to 0.99, for each choice of λ´s, to see what their effects are. We concluded that the previously used discount factor is really the best one.

Table 4 presents the results obtained to compare the models with time varying level, with and without the dummies. We inspect in depth the cases: $(\lambda_1, \lambda_2) = (0.8, 0.5)$ e $(0.7, 0.3)$. All the results are based on a discount factor equal to 0.95 and the criteria functions evaluated as defined before.

Table 4 presents the results obtained to compare the models with time varying level, with and without the dummies. All the results are based on the discount factor equal 0.95 and the criterion functions evaluated as defined before.

Table 4: Multivariate model selection, based on the $RMSE$ and LPL (normalized by the number of days), using a discount factor equal to 0.95 and based on the data from 2008 to 2010.

RMSE					
		(λ_1, λ_2)	Monday	Weekday	Weekend
Model 1	no dummy	$(0.8, 0.5)$	0.053	0.058	0.068
		$(0.7, 0.3)$	0.053	**0.058**	0.068
Model 2	with dummy	$(0.8, 0.5)$	0.052	0.058	0.066
		$(0.7, 0.3)$	**0.052**	0.060	**0.066**
LPL					
		(λ_1, λ_2)	Monday	Weekday	Weekend
Model 1	no dummy	$(0.8, 0.5)$	5828.483	5819.685	5123.048
		$(0.7, 0.3)$	5833.942	**5820.913**	5125.742
Model 2	with dummy	$(0.8, 0.5)$	5872.553	5796.993	5169.429
		$(0.7, 0.3)$	**5922.842**	5779.842	**5209.486**

It is worth recalling that the alternative model we are comparing differs by the inclusion or not of dummy variables. We conclude that the dummy variables are significant and that the decay factors $\lambda_1 = 0.7$ and $\lambda_2 = 0.3$ are always the better.

It is worth recalling that the alternative model we are comparing differs by the inclusion or not of dummy variables. Since both criteria do not consistently favor any of the alternative models, particularly,

We used the complete model with $(\lambda_1; \lambda_2) = (0.7; 0.3)$ to make the graphical illustrations. Figure 9 shows the one-step-ahead forecasting mean value (solid line), the 95% credibility interval (dashed lines) and the observations (dots), obtained with the multivariate model 2, based on the data for the 15th week of 2010.

Finally, we can compare the performance of the univariate and multivariate dynamic regression models, by using the numbers in Tables 3 and 4. The multivariate model with dummy variables is not better than the univariate model including the seasonal effects only, for instance Model 7. Although this finding is not so intuitive, it could be derived from the substantial simplifications we imposed on the intraday dependence structure.

6 Concluding remarks and extensions

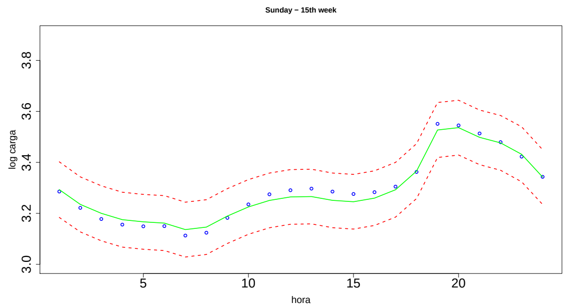
In this paper we discussed the implementation of two classes of alternative models. The first one is based on the extensions of Nelson-Siegel factor models with time varying parameters and the second one extends previous works in the recent literature and is based on separate regressions by hour. Our main contribution to this class of problem include the use of time varying Nelson-Siegel factor models and the use of dummy variables to modify the factors and allow the observational variance estimation.

The data analysis developed allows us to conclude that it is enough to include the temperature as linear regression and also that all models produce very similar forecasts with almost the same precision. In order to keep the computational effort under control, some simplifications were introduced. The decay factors in the Nelson-Siegel were kept fixed and time invariant, the observational variances were supposed known and the discount factors were used to describe the state evolution.

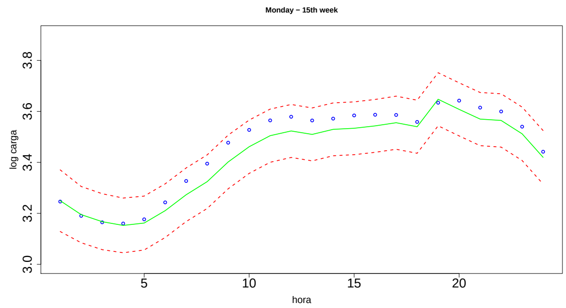
We plan to implement the full model using MCMC and also to explore results for the dynamic factor model. Modeling the temporal dependence structure in a sequence of variance matrices is of interest in multi and matrix variate time series with application to many different fields, such as econometrics, neuroscience etc. Three alternatives to approach this problem include: graphical models (Carvalho and West (2007) and Wang and West (2003)), sparse regularization and prediction (Friedman *et al.* (2008)) and autoregressive models for variance matrices (Fox and West (2011)).

Acknowledgements

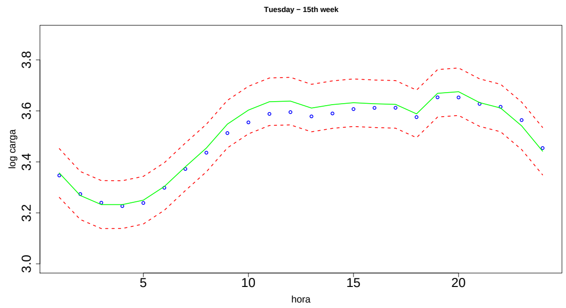
The first author acknowledges the financial support of research grants from CNPq and the second one the PhD scholarship sponsored by Capes. We both thank ONS - Operador Nacional do Sistema for supporting on all stages of this project.



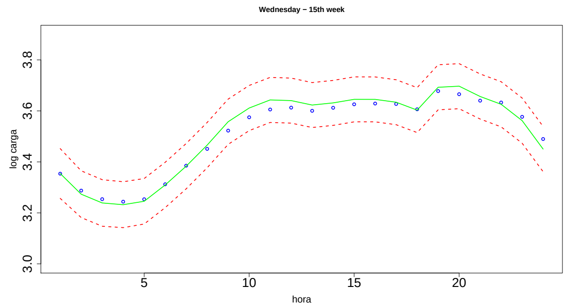
Apr 11, 2010



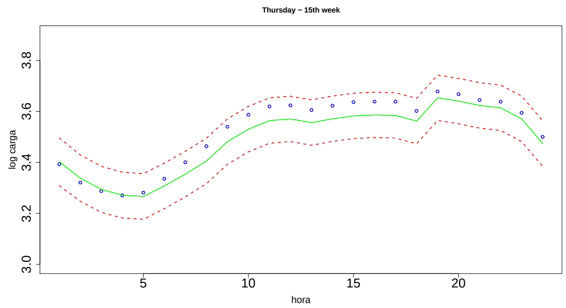
Apr 12, 2010



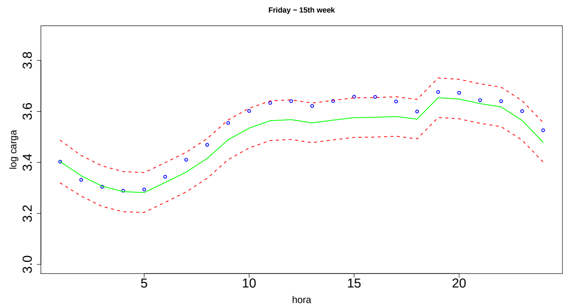
Apr 13, 2010



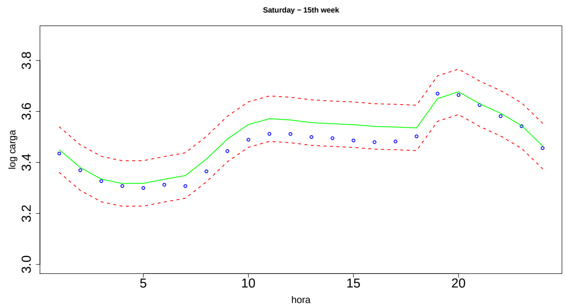
Apr 14, 2010



Apr 15, 2010



Apr 16, 2010



Apr 17, 2010

Figure 9: *One-step-ahead load curve forecasting (solid line) with* 95% *credibility interval (dashed line), Multivariate Model 2, week 15, 2010* ($\delta = 0.95$ *and* $(\lambda_1, \lambda_2) = (0.7, 0.3)$*). The observations are in dots.*

References

Alfares, H. K. and Nazeeruddin, M. Alfares and Nazeeruddin (2002) Electric load forecasting: literature survey and classification of methods. *International Journal of Systems Science*, **33,** 23–24.

Badran, I., El-Zayyat, H. and Halasa, G. Badran *et al.* (2008) Short-term and medium-term load forecasting for jordan's power system. *American Journal of Applied Sciences*, **(5) 7,** 763–768.

Bedrick, E. J., Christensen, R. and Johnson, W. Bedrick *et al.* (1996) A new perspective on priors for generalized linear models. *Journal of Operational Research Society*, **91,** 1450–1460.

Cancelo, J. R., Espasa, A. and Grafe, R. Cancelo *et al.* (2008) Forecasting the electricity load from one day to one week ahead for the spanish system operator. *International Journal of Forecasting*, **24,** 59–83.

Carvalho, C. M. and West, M. Carvalho and West (2007) Dynamic matrix-variate graphical models. *Bayesian Analysis*, **2,** 69–98.

Clemen, R.T. Clemen (1989) Combining forecasts: a review and annotated bibliography. *International Journal of Forecasting*, **5,** 559–583.

Cottet, R. and Smith, M. Cottet and Smith (2003) Bayesian modeling and forecasting of intraday electricity load. *Journal of the American Statistical Association*, **98,** 839–849.

De Pooter, M. De Pooter (2007) Examining the nelson-siegel class of term structure models.

Dias, R., Garcia, N. L. and Schmidt, A. M. A hierarchical model for aggregated functional data. Technical Report.

Diebold, F. X. and Li, C. Diebold and Li (2006) Forecasting the term structure of governments bond yields. *Journal of Econometrics*, **130,** 337–364.

Dordonnat, V., Koopman, S. J. and Ooms, M. Dynamic factors in periodic time-varying regressions with an application to hourly electricity load modelling. Technical Report.

Dordonnat, V., Koopman, S. J., Ooms, M., Dessertaine, A. and Cottet, J. Dordonnat *et al.* (2008) An hourly periodic state space model for modelling french national electricity load. *International Journal of Forecasting*, **24,** 566–587.

Fernandes, M., Schmidit, A. M. and Migon, H. S. Fernandes *et al.* (2009) Modelling zero inflated spatio-temporal processes. *Statistical Modelling*, **9,** 3–25.

Fox, E. B. and West, M. Fox and West (2011) Autoregressive models for variance matrices: Stationary inverse wishart processes. *Cornell University Library*, arXiv:1107.5239v1 [stat.ME].

Friedman, J., Hastie, T. and Tibshirani, R. Friedman *et al.* (2008) Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, **9,** 432–441.

Garthwaite, P. H., Kadane, J. B. and A., O'Hagan Garthwaite *et al.* (2005) Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, **100,** 680–700.

Kyung, M., Gill, J., Ghosh, M. and Casella, G. Kyung *et al.* (2010) Penalized regression, standard errors, and bayesian lassos. *Bayesian Analysis*, **5,** 369–412.

Litterman, R. B. Litterman (1986) Forecasting with bayesian vector autoregressions: Five years of experience. *Journal of Business & Economic Statistics*, **4,** 25–38.

Lourenço, J. and Santos, P. Short term load forecasting using gaussian process models. Technical Report.

Migon, H. S., Gamerman, D., Lopes, H. F. and Ferreira, M. A. R. *Dynamic models.* Bayesian Thinking, Modeling and Computation., pp. 557–592. Handbook of Statistics.

Nelson, C. R. and Siegel, A. F. Nelson and Siegel (1987) Parsimonious modeling of yield curve. *Journal of Business*, **60,** 473–489.

Taylor, J. M., de Menezes, L. M. and McSharry, P. E. A. Taylor *et al.* (2006) Comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, **22,** 1–16.

Taylor, J. W. Taylor (2003) Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of Operational Research Society*, **54,** 799–805.

Wang, H. and West, M. Wang and West (2003) Bayesian analysis of matrix normal graphical models. *Biometrika*, **96,** 821–834.

West, M. and Harrison, J. West and Harrison (1997) *Bayesian Forecasting and Dynamic Models.* Springer.