Prediction models of electricity demand with high-frequency data for Uruguay

Bibiana Lanzilotta - Silvia Rodríguez-Collazo

Facultad de Ciencias Económicas, FCEA, UdelaR - Centro de Investigaciones Económicas

1 de septiembre de 2021

Overview

Objectives

- Characterize and model daily demand of electric energy in Uruguay
- Identifying the incidence of specific events, individually and integrated in the short term dynamic
- Exploring the non-linear association between energy consumption and climatic variables.

Data sources

- UTE
- INIA, Banco datos agroclimático
- Instituto de Relaciones laborales, UCU. Índice de Conflictividad Laboral (general strikes)

Main features of the electric energy demand in Uruguay

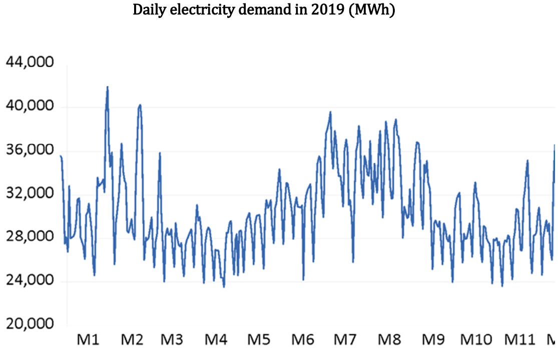
Significant growth in energy demand over the last 30 years

- 2010 - 2019: 25,754 MWh - 30,399 MWh (daily average)
- Average annual growth rate 2010-2019: 1.6 %

Main features

- Evolution of average daily demand (trend)
- Multiple seasonalities
- Changes in seasonality

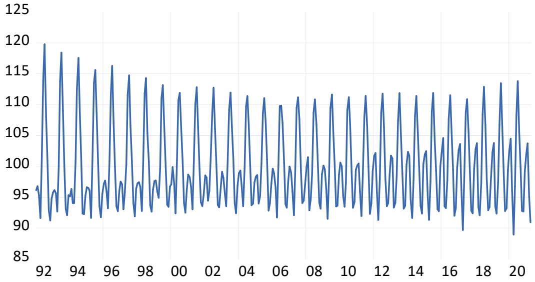
Seasonal patterns of electricity demand in 2019 (MWh)



- The highest daily peak take place in summer:January 29, with a 41,905 MWh demand.
- The secondly highest peak took place in winter: July with 34,762 MWh daily consumption.
- In the fall and spring the demand is lower due to the moderation of the climatic variables in these seasons.

Changes in seasonal patterns 1992-2019

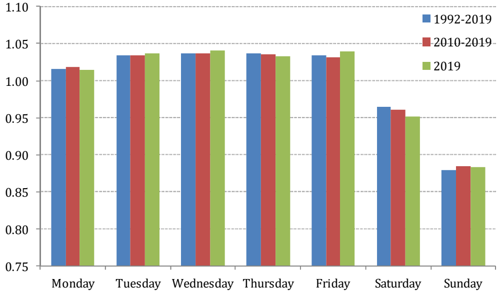
Graph 2. Seasonal factor of electricity demand (1992-2019)



- Seasonality (associated with the seasons) has changed in time, diminishing the gap between winter and summer
- Between 1992 and 2000 the energy consumption during summer represented 88 % of that of the winter
- Between 2010 and 2019, the energy consumption during summer increased, and represented 93 %

Weekly seasonality

Graph 3. Average daily consumption by day of the week, (% of the weekly average consumption)



- There is no noticeable change in this pattern throughout the period analyzed.
- Saturdays and Sundays average consumption represents nearly 95 % and 85 % of the average consumption of the trading days (Monday, Tuesday, Wednesday, Thursday, Friday), respectively.

Methodological approach

- Based on Cancelo and Espasa (1996):
  Modelling the non linear link between energetic consumption and climate variables, incorporating a detailed analysis of intervention on special days (holidays, vacations, strikes).

$$lnD_t = FS_t + ES_t + CDE_t + CVC_t + \epsilon_t \tag{1}$$

lnD_t the natural logarithm of the daily energy demand in Uruguay

FS_t a trend associated to socioeconomic factors that influence the energetic demand

ES_t the weekly seasonal pattern

CDE_t the special days contribution to the energetic demand

CVC_t the climate variables contribution

ϵ_t random shocks not taken into account in any previous variables

Methodological background II

$$lnDB_t = lnD_t - CDE_t - CVC_t = FS_t + ES_t + \epsilon_t \quad (2)$$

The basic demand DB_t is corrected for the effects of special days and the effect of climatic variables.

DB_t contains the more stable components of demand that could be expressed as an ARIMA model.

$$\Delta\Delta_7 lnDB_t = \eta_t \quad (3)$$

Being η_t a stationary ARMA (p,q) process.

$$\Delta\Delta_7 lnD_t = \Delta\Delta_7 CDE_t + \Delta\Delta_7 CVC_t + \eta_t \quad (4)$$

Methodological background III

$$CDE_t = f_1(L)DE_t = \sum_{i=1}^{i=m} f_{1,i}(L)DE_{i,t} \qquad (5)$$

$$CVC_t = f_2(L)VC_t = \sum_{j=1}^{j=n} f_{2,j}(L)VC_{j,t} \qquad (6)$$

DE_t a matrix conformed by m dummy variables used to model special days
f_1 a vector of polynomials associated to the L operator.

VC_t corresponds to the matrix composed by n variables used to model the effects of climate variables
f_2 refers to the polynomials vector associated to L.

f_1 and f_2 will represent the *specific dynamics* of each of these variables respectively.

$$\Delta\Delta_7 lnD_t = \Delta\Delta_7 f_1(L)DE_t + \Delta\Delta_7 f_2(L)VC_t + \eta_t \qquad (7)$$

The treatment of the effects of the special days II

$$
\begin{aligned}
\Delta\Delta_7 lnD_t = \\
& \sum_{j=1}^{j=4} f_1(L)\Delta\Delta_7 G_{j,t} * Sun_t + \sum_{j=1}^{j=4} f_2(L)\Delta\Delta_7 G_{j,t} * Sat_t \\
& + \sum_{j=1}^{j=4} f_3(L)\Delta\Delta_7 G_{j,t} * Fri_t + \sum_{j=1}^{j=4} f_4(L)\Delta\Delta_7 G_{j,t} * Thu_t \\
& + \ldots\ldots + \Delta\Delta_7 f_{10}(L) * Easter_t + \Delta\Delta_7 f_{11}(L) * Carninval_t \\
& + \Delta\Delta_7 f_{12}(L) * Strike_t + \varphi_t
\end{aligned} \tag{8}
$$

- The interaction of holiday variables (grouped) and day of the week variables allows us to measure the impact of the holidays depending on the week day they fall
- Easter indicates Easter Sunday
- Carnival designates Carnival Tuesday
- Strike assembles all general activity strikes
- $f_i(L)$ capture the dynamic of the calendar effects

The non-linear model applied

One of the unarguable features of the link between climate variables and energetic demand is its **non linearity**.

- Our contribution consists on the usage of Markov´s Switching Models to estimate breaking points on the demand function
- Include different climatic variables: wind, relative humidity and the solar light (heliophany) that affect apparent temperature and therefore individual‘s thermal comfort needs

Two step procedure for estimating breaking points

- Step 1: Estimation a linear function to determine relevant climatic variables to model energetic demand, as well as its structure and main outliers.
- Step 2: Estimation of breaking points. Following Markov´s Switching Model methodology

$$
\begin{aligned}
\Delta\Delta_7 lnDAD_t = \\
& \Delta\Delta_7 f_{21}(L) * Warm_t + \\
& \Delta\Delta_7 (Temp_t - W_t^i) * f_{22}(L) * Warm_t + \\
& \Delta\Delta_7 f_{23}(L) * Cold_t + \\
& \Delta\Delta_7 (Temp_t - C_t^i) * f_{24}(L) * Cold_t + \\
& \Delta\Delta_7 Heliophany * f_{25}(L) * Warm_t + \\
& \Delta\Delta_7 Heliophany * f_{26}(L) * Cold_t + \\
& \Delta\Delta_7 RH * f_{27}(L) * Warm_t + \\
& \Delta\Delta_7 RH * f_{28}(L) * Cold_t + \\
& \Delta\Delta_7 Wind * f_{29}(L) * Warm_t + \\
& \Delta\Delta_7 Wind * f_{30}(L) * Cold_t + \\
& \Delta\Delta_7 Hour_t + \Delta\Delta_7 Veranillo_t + \Delta\Delta_7 Save_t + \\
& \sum_{j=1}^{j=11} \Delta\Delta_7 Month_{i,t} + \sum_{j=1}^{j=s} \Delta\Delta_7 Outlier_{i,t} + \frac{\theta(L)}{\phi(L)} a_t
\end{aligned} \tag{9}
$$

Search for the breaking point

- W_t^i and C_t^i are threshold variables
- $W_t^i = max\,[0, k_i - Temp]$ y $C_t^i = max\,[0, k_i - Temp]$
- k_i the *ith* break found in the link between energetic consumption and temperature, in warm and cold season respectively

- Find k_i, following Markov´s Switching Models methodology, and choosing as break point candidate the temperature that minimized the squared sum of errors of the regression.
- Second stage:Tested the existence of significant differences among parameters associated with the *Temp* variable for values above and below k_i.
- This process is repeated until we no k_i for which estimated parameters are not significantly different is found
- We do not take into account for the estimation of breaking points, 5 % of the lowest and highest observed temperatures on each season, (warm and cold)

The final equation

$$
\begin{aligned}
\Delta\Delta_7 lnDAD_t = & \\
& \sum\nolimits_{j=1}^{j=v} \Delta\Delta_7 W_t^i f_{31}(L) * Warm_t + \\
& \sum\nolimits_{j=1}^{j=v} \Delta\Delta_7 (Temp_t - W_t^i) * f_{32}(L) * Warm_t + \\
& \sum\nolimits_{j=1}^{j=v} \Delta\Delta_7 C_t^i f_{33}(L) * Cold_t + \\
& \Delta\Delta_7 (Temp_t - C_t^i) * f_{34}(L) * Cold_t + \\
& \Delta\Delta_7 Heliophany * f_{35}(L) * Warm_t + \\
& \Delta\Delta_7 Heliophany * f_{36}(L) * Cold_t + \\
& \Delta\Delta_7 RH * f_{37}(L) * Warm_t + \Delta\Delta_7 RH * f_{38}(L) * Cold_t + \\
& \Delta\Delta_7 Wind * f_{39}(L) * Warm_t + \Delta\Delta_7 Wind * f_{40}(L) * Cold_t + \\
& \Delta\Delta_7 Hour_t + \Delta\Delta_7 Save_t + \Delta\Delta_7 Veranillo_t + \\
& \sum\nolimits_{j=1}^{j=11} \Delta\Delta_7 Month_{i,t} + \sum\nolimits_{j=1}^{j=s} \Delta\Delta_7 Outlier_{i,t} + \frac{\theta(L)}{\phi(L)} a_t
\end{aligned}
\tag{10}
$$

Climatic inertia

- Residences keep ambient temperature, at least for a few days.
- A given temperature does not have the same effect in summer than in winter, and different breaking points have to be found for different seasons.
- Two qualitative variables are included: **warm and cold**; that reflect the months when average temperature is higher than the year average temperature (the first one) and those when it is not (the second).
- Two variables interact with the observed temperature
  - cold dummy: months May, June, July, August, September and October.
  - Warm dummy is defined by difference.

The treatment of the effects of the special days

- National holidays were included, both workable and not: Easter, Carnival, holidays and strikes.
- Their impact is different according to which week day, and they also have lagged and forward effects on the demand.
- Four groups in which we grouped the different holidays according to its incidence in energetic consumption.

Table 1. Holiday Groups and effects (average dynamic effects)

		Sample 2010 -2019
Group 1	January 1, December 25	-0.097
Group 2	May 1, August 25, March 1	-0.061
Group 3	January 6, July 18, November 2	-0.036
Group 4	April 19,May 18, June 19, October 12	-0.015
Easter		-0.051
Carnival		-0.037

Note: Impacts are on $\Delta\Delta_7 lnD_t$.
Source: Authors estimations.

Temperature thresholds and nonlinear modeling

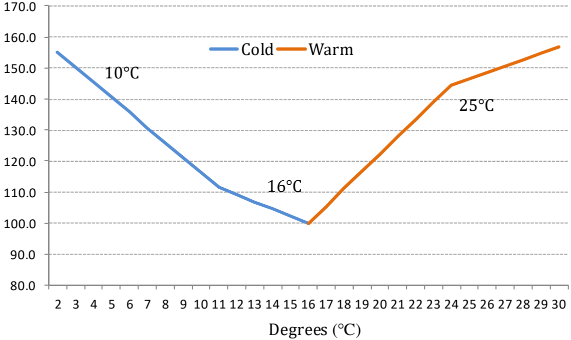
Table 2.Breaking point estimation on the electric demand function

	Function section	lag	Coeff	$\Sigma Coeff$
Warm	Between 16 C° – 25$C°$	0	0.742 %	
		1	0.149 %	0.891 %
	More than 25 C°	0	0.294 %	
		1	0.029 %	0.323 %
Cold	Less than 10 C°	0	0.522 %	
		1	0.254 %	0.776 %
	Between 10 C° – 16$C°$	0	0.233 %	
		1	0.135 %	0.368 %

Source: Authors estimations.

Temperature thresholds and nonlinear modeling

Graph 5. Graphic representation of temperature impact on the daily weekly growth rate energetic demand, Base Index 100=0% growth



Climatic Variables Effects on Electricity Demand

Table 4.Climatic Variables Effects on Electricity Demand

	Climatic Variable	lag	Coeff	Coeff
Warm	Heliophany	0	0.007 %	
		1	0.036 %	0.043 %
	Relative humidity	0	0.026 %	0.026 %
Cold	Heliophany	1	0.026 %	0.026 %
	Wind	1	0.0013 %	0.0013 %

Source: Authors estimations.

Predictive evaluation

- In order to assess the model´s predictive capability we left out the last year of the sample (from 01/01/2019 to 12/31/3019) and calculate the forecast error for each of the following months.

Table 5.Mean absolute relative errors for 7,and 14 prediction horizons

Steps	January	February	March	April	May	June	July
h=7	3.4	3.3	2.0	1.6	2.0	2.8	2.8
h=14	3.7	4.0	2.3	1.9	2.3	3.5	3.3
Steps	August	Sep.	October	Nov	Dec		Annual Average.
h=7	3.0	2.1	3.9	2.1	2.9		2.7
h=14	3.7	2.6	4.7	2.4	3.3		3.1

Note: Predictions during 4 weeks for each month
Source: Authors estimations

- Both in the 7-step and in in the 14-step prediction, October is the month with the highest error.
- These poor results during the month of October 2019 are due to the fact that substantially higher errors were recorded during the first two weeks of the month than in the last two weeks.

Predictive evaluation II

If another, less demanding indicator is used, such as the relative error at 7 and 14 steps.

Table 6.Mean relative errors for 7,and 14 prediction horizons

Steps	January	February	March	April	May	June	July
h=7	-0.64	0.6	-0.03	-0.1	-0.31	0.11	0.1
h=14	0.02	0.2	0.14	-0.1	-0.04	-0.04	0.4
Steps	August	Sep.	October	Nov	Dec		Annual Av.
h=7	0.63	0.2	-1.6	-1.0	-0.30		-0.19
h=14	0.03	-0.3	-0.4	-0.4	-0.03		-0.05

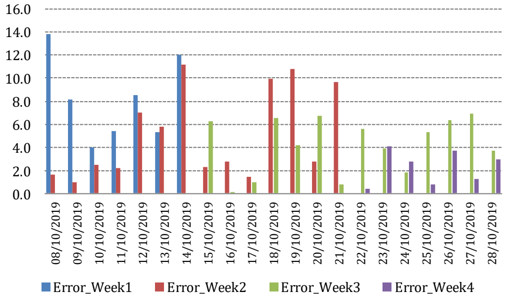
Note: Predictions during 4 weeks for each month
Source: Authors estimations

- As expected, the largest errors are concentrated in the month of October.
- As we saw in the previous Table, these poor results for the first two weeks of October 2019 negatively affect the relative errors for the month.
- The relative error during the first two weeks of the month was -3.25 %, while in the following two weeks it was 0.04 %.

Update of predictions

- Note that errors were estimated from predictions with exogenous variables already observed.
- Prediction update results: each bar indicates the absolute relative error level, and the overlap of bars indicates the reduction of the error when new information is incorporated.
- Considering the negative results obtained for the month of October 2019, we analyze how the forecasting update process operates in that month.

Graph 6: Prediction adjustment process. October 2019



Main conclusions

- This paper revise and update a preliminary work that with estimations to 2012.
- Our results show the incidence of the special days (calendar effects, holidays) and saving energy measures.The relevance of capturing these effects where the heterogeneity of the joint impact of the holidays can be captured according to the day of the week on which they fall and their temporal dynamics.
- Modeling the association between energy consumption and climatic variables (temperature, humidity, winds and heliophany) with a non-linear model with estimated breaks (estimated by applying Markov´s Switching Models).
- Breaks were identified by considering the sample split in warms and cold months at $16C^{\circ}$, $25C^{\circ}$ (in warm months) and at $10C^{\circ}$ in cold months.

Main conclusions

- The estimated coefficients show that the electricity demand function changed compared with the preliminary paper.
- At high temperatures, the demand function increase at a major rate, and therefore, the curve is sharper. A saturation temperature is reached.
- The section of the function corresponding to colder temperatures remains relatively similar.
- These changes are probably associated with the increased use and availability of refrigeration equipment by households.
- The predictive assessment conducted during 2019, indicates that the model is adequate to predict the next 7 days.
- Further research must focus on the analysis of the month of October where climatic variability is significantly higher.

References

- Bogard, C., George, G., Jenkins, G.M. y McLeod, G. (1982). "Analyzing a large number of energy time series for utility company ", in *Case Studies in Time Series Analysis*, Jenkins, G.M. and McLeod (eds.), Gwilym Jenkins Partners Ltd., Lancaster, chapter 5.
- Bessec, M. J. Fouquau (2008). "The non-linear link between electricity consumption and temperature in Europe: A threshold panel approach". *Energy Economics* 30, 2705–2721.
- Bunn, D. W. Farmes, E. D. (1985). "Economic and operational context of electric load prediction", in *Comparative Models for Electrical Load Forecasting*, Bunn, D. W. Farmer, E.D. (eds.), John Wiley Sons, New York, Chapter 1.
- Cancelo, J.R. Espasa, A. (1995). "Algunas consideraciones sobre la modelización de series diarias de actividad económica", *Actas de las X Jornadas de Economía Industrial*, 195-201.
- Cancelo, J.R., y Espasa, A. (1996). "Modeling and forecasting daily series of electricity demand." *Investigaciones Económicas*. Vol XX(3), 359-376.
- Cancelo, J.R., Espasa, A. (2001). "Using high-frequency data and time series models to improve yield management". *International Journal of Services Technology and Management*, 2, 59-70.

References

- Cancelo, J.R., Espasa, A. Grafe R. (2008), "Forecasting the electricity load from one day to one week ahead for the Spanish system operator". *International Journal of Forecasting* 24, 588-602.
- Darbellay, G.A., Slama, M. (2000). "Forecasting the short-term demand for electricity. Do neural networks stand a better chance?" *International Journal of Forecasting*, 16, 71-83.
- Engel, R.F., Granger, C.W.J., Rice, J., Weiss, A. (1986). "Semi-parametric estimates of the relation between weather and electricity sales", *Journal of the American Statistical Association*, 81, 310-320.
- Espasa, A. (1993). "Modeling daily series of economic activity" *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 313-318.
- Hippert, H.S., Bunn, D.W., Souza, R. C. (2005). "Large neural networks for electricity load forecasting: are they overidentified?" *International Journal of Forecasting*, 21, 425-434.

References

- Lanzilotta, B., Carlomagno, G: Rosá, T. (2012) "Un sistema de predicción y simulación para la demanda de energía eléctrica en Uruguay". Informe de Proyecto ANII-FMV 2009 Lanzilotta, B. (resp.), Carlomagno, G: Rosá, T.Piggot, J. L. (1985). "Short-term forecasting at British Gas", in Comparative Models for Electrical Load Forecasting, Bunn, D. W. and Farmer, E.D.(eds.), John Wiley Sons, New York, ch.1.
- Lanzilotta, B., Rodríguez-Collazo, S.(2016). "Modelos de predicción de demanda de energía eléctrica con datos horarios para Uruguay". Cuadernos del CIMBAGE, (18).
- Ramanathan, R., Engle, R., Granger, C.W.J., Vahid-Araghi, F., Brace, C. (1997). "Short-run forescasts of electricity loads and peaks". International Journal of Forecasting, 13, 161-174.
- Smith, M. (2000). "Modeling and short-term forecasting of New South Wales electricity system load". *Journal of Business and Economic Statistics*, 18, 465-478.
- Soares L. J, Souza, L.R. (2008). "Forecasting electricity demand using generalized long memory", *International Journal of Forecasting* 22, 17– 28.

thanks