



**HAL**  
open science

## Modèles de prévision de demande d'électricité. Application au système uruguayenne.

Andrés Castrillejo, Jairo Cugliari, Fernando Massa, Ignacio Ramirez

► **To cite this version:**

Andrés Castrillejo, Jairo Cugliari, Fernando Massa, Ignacio Ramirez. Modèles de prévision de demande d'électricité. Application au système uruguayenne.. Journées de Statistique, May 2017, Avignon, France. hal-01539443

**HAL Id: hal-01539443**

**<https://hal.science/hal-01539443>**

Submitted on 18 Sep 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MODÈLES DE PRÉVISION DE DEMANDE D'ÉLECTRICITÉ. APPLICATION AU SYSTÈME URUGUAYENNE.

Andrés Castrillejo <sup>1</sup> & Jairo Cugliari <sup>2</sup> & Fernando Massa <sup>3</sup> & Ignacio Ramirez <sup>4</sup>

<sup>1</sup> *Iesta, Fcea, UdelaR, Uruguay, andres@iesta.edu.uy*

<sup>2</sup> *Lab ERIC, Univ Lyon, France, Jairo.Cugliari@univ-lyon2.fr*

<sup>3</sup> *Iesta, Fcea, UdelaR, Uruguay, fmassa@iesta.edu.uy*

<sup>4</sup> *IIE, Fing, UdelaR, nacho@iie.edu.uy*

**Résumé.** Le développement de nouvelles technologies de génération d'électricité donnent aux économies émergentes des possibilités nouvelles. Ces économies sont souvent dépendantes des énergies fossiles et ainsi du cours du pétrole. L'Uruguay a terminé sa transformation énergétique, ayant aujourd'hui une très large participation des sources renouvelables dans son mix de production.

Le changement rapide a demandé des nouvelles méthodes mathématiques et informatiques pour la gestion et pilotage de la charge du système. Dans ce travail nous présentons **enercast**, un package R qui contient des modèles de prévision de demande qui peuvent être utilisés par l'opérateur du réseau.

**Mots-clés.** Prévision, Demande d'électricité

**Abstract.** The development of new electricity generation technologies has given new opportunities to developing economies. These economies are often highly dependent on fossil sources and so on the price of petrol. Uruguay has finished the transformation of its energetic matrix, presenting today a very large participation of renewable sources among its production mix.

This rapid change has demanded new mathematical and computing methods for the administration and monitoring of the system load. In this work we present **enercast**, a R package that contains prediction models that can be used by the network operator.

**Keywords.** Forecasting, Electricity demand

## 1 Introduction

Le développement de nouvelles technologies de génération d'électricité donnent aux économies émergentes des possibilités nouvelles. Ces économies sont souvent dépendantes des énergies fossiles et ainsi du cours du pétrole. L'Uruguay a terminé sa transformation énergétique, ayant aujourd'hui une très large participation des sources renouvelables dans son mix de production.

Le changement rapide a demandé des nouvelles méthodes mathématiques et informatiques pour la gestion et pilotage de la charge du système.

Nous proposons dans ce travail un package R qu'on appelle **enercast** qui implémente des méthodes et techniques de prévision de la demande électrique. Les méthodes disponibles sont en lien avec le besoin d'un système électrique de taille réduite avec une rare disponibilité des données météorologiques. Ceci est le cas notamment dans des zones régionales des systèmes plus larges ou dans des petites économies émergentes comme l'Uruguay.

## 2 Modèles de prévision

Nous décrivons dans cette section quelques modèles que nous avons utilisé dans la prévision de la charge du système électrique uruguayen.

**Modèle de régression de Hong** (Hong Vanilla Benchmark, [?]) Proposé comme référence pour la compétition Gefcom 2012, ce modèle prévision très simple utilise une fonction de régression linéaire multiple pour décrire la dépendance de la charge à la température, des variables temporelles (heure, jour, mois) ainsi que quelques interactions. Il incorpore aussi un terme de tendance linéaire .

Concrètement le modèle est spécifié comme suit

$$\begin{aligned} \mathbb{E}(\text{load}_t) = & \beta_0 + \beta_1 t + \beta_2 \text{Day}_t \times \text{Hour}_t + \beta_3 \text{Month}_t + \beta_4 \text{Month}_t \times \text{Temp}_t \\ & + \beta_5 \text{Month}_t \times \text{Temp}_t^2 + \beta_6 \text{Month}_t \times \text{Temp}_t^3 + \beta_7 \text{Hour}_t \times \text{Temp}_t \\ & + \beta_8 \text{Hour}_t \times \text{Temp}_t^2 + \beta_9 \text{Hour}_t \times \text{Temp}_t^3 \end{aligned}$$

où **Hour**, **Day** y **Month** sont les variables liées au temps; **Temp** est la température horaire. Le modèle incorpore un certain nombre d'interactions.

**Modèles de Séries Temporelles de Référence (Benchmark)** Très populaires dans la littérature économétrique, ces modèles étendent les modèles linéaire comme celui du paragraphe précédant en incorporant des versions retardées de la courbe de charge (variable **load**) comme prédicteur.

Nous considérons cette approche comme un *benchmark* à partir d'un ajustement automatique à l'aide du package **forecast**. Le package cherche le meilleur élément de la famille de modèles décrits par

$$\Phi_{P_2}(L^{s_2})\Phi_{P_1}(L^{s_1})\phi_p(L)\nabla_{s_2}^{D_2}\nabla_{s_1}^{D_1}\nabla^d(\text{load}_t - X_t\beta) = \Theta_{Q_2}(L^{s_2})\Theta_{Q_1}(L^{s_1})\theta_q(L)\epsilon_t$$

où  $X_t$  c'est la matrice des variables exogènes au moment  $t$ ,  $\beta$  le vecteur de coefficients de régression et  $\epsilon_t$  est une suite de type bruit blanc.

Les polynômes autorégressifs  $\Phi_{P_1}(L^{s_1})$  et  $\Phi_{P_2}(L^{s_2})$  sont associés aux comportements saisonniers, dans notre cas  $s_2$  représente les semaines et  $s_1$  les jours. Finalement  $\phi_p(L)$

décrit le comportement à très court terme (inférieur à la journée). De manière analogue l'on définit les polynômes des moyennes mobiles  $\Theta_{Q_2}(L^{s_2})$ ,  $\Theta_{Q_1}(L^{s_1})$  et  $\theta_q(L)$ .

L'ajustement par **forecast** trouve les index  $p, d, q, P_1, D_1, Q_1, P_2, D_2$  et  $Q_2$  qui minimisent un certain critère d'erreur. Typiquement les valeurs des opérateurs de différenciation  $d, D_1$  et  $D_2$  sont soit 0 soit 1, tandis que les valeurs des indices autorégressifs et des moyennes mobiles peuvent être plus larges.

Nous remarquons que cette variante est une alternative à un seul modèle. Voir Cancelo et al (2008) pour une approche où 24 modèles horaires sont traités conjointement pour décrire la courbe de charge journalière.

**Modèle d'espace à états** (MEE) Cette famille de modèles est une extension des modèles classiques de séries temporelles et des modèles de régression. Plusieurs MEE peuvent être obtenus à partir des 2 équations qui suivent

$$\begin{aligned}\alpha_t &= c_t + T_t \alpha_{t-1} + v_t \\ Y_t &= d_t + Z_t \alpha_t + w_t.\end{aligned}\tag{1}$$

La première équation, appelée équation d'états, décrit l'évolution temporelle du vecteur (non observable)  $\alpha_t$  avec un vecteur de "contrôle"  $c_t$  qui modifie sa moyenne conditionnelle, une matrice  $T_t$  qui modifie sa moyenne conditionnelle et un vecteur de bruit  $v_t$  avec  $H_t$  sa matrice de covariance.

La deuxième équation, appelée équation d'observations, modélise l'évolution des données observées  $Y_t$ . La composante  $d_t$  contient des éléments qui modifient la moyenne des observations,  $Z_t$  fait le lien entre le vecteur d'états et le vecteur d'observations et  $w_t$  est un vecteur aléatoire avec une matrice de covariances  $G_t$ .

Nous suivons Dordonnat et al. (2008) où les auteurs proposent la construction de 24 modèles horaires. Le modèle contient une régression linéaire avec des paramètres qui peuvent changer dans le temps en suivant une marche aléatoire. Le modèle s'écrit:

$$\begin{aligned}\alpha_t &= \alpha_{t-1} + v_t \\ Y_t &= X_t^f \beta_t + X_t^a \alpha_t + w_t\end{aligned}\tag{2}$$

Les variables dont les coefficients sont fixes (constants dans le temps) sont dans  $X_t^f$ , et les variables dont les coefficients peuvent évoluer sont dans  $X_t^a$ .

Dans notre application nous incorporons dans  $X_t^f$  les variables qui indiquent la présence des jours spéciaux (e.g. fériés), et dans  $X_t^a$  nous incorporons une tendance, des fonctions sinusoïdales pour décrire les cycles (journalières, hebdomadaires, annuels) et des termes d'un spline cubique pour capturer le lien entre température et demande.

### 3 Expériences numériques

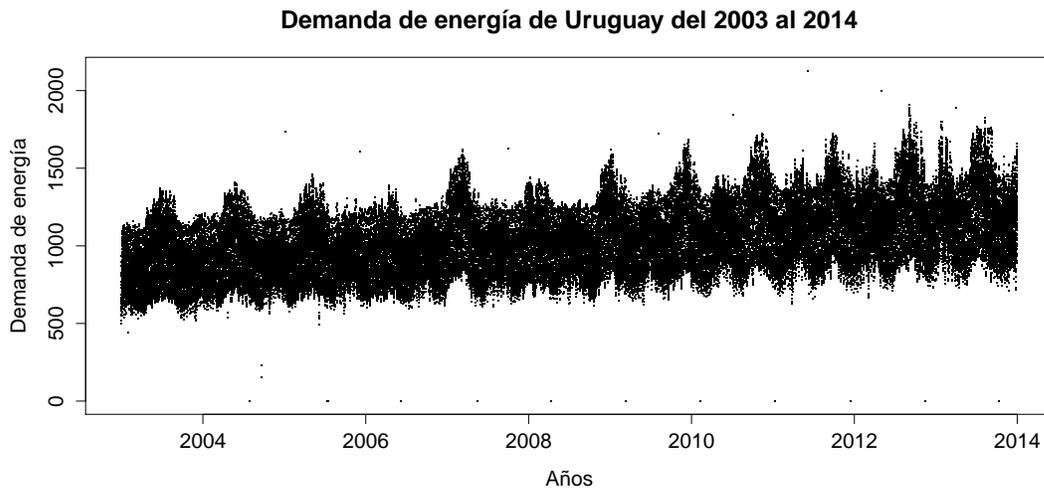
#### 3.1 Données.

Les données horaires de demande ont été fourni par le *Despacho Nacional de Carga* de UTE<sup>1</sup>. Les données disponibles couvrent la période allant du 1 janvier 2003 jusqu'au 31 décembre 2014. Les données de charge sont mesurées à pas horaires. Voici un aperçu des données

	année	mois	jour	heur	demande (kWh)	jour_semaine
1	2003	1	1	1	828,220	mercredi
2	2003	1	1	2	766,573	mercredi
3	2003	1	1	3	686,091	mercredi
4	2003	1	1	4	624,981	mercredi
5	2003	1	1	5	591,963	mercredi
6	2003	1	1	6	552,376	mercredi

Table 1: Aperçu du jeu de données.

Le graphique 3.1 met en évidence la présence d'une tendance croissante ainsi que de la saisonnalité annuelle de la série.



---

<sup>1</sup><http://portal.ute.com.uy/>

### 3.2 Procédure de prévision.

Les méthodes introduites dans la Section 2 ont été programmées et mises a dispositions dans le package R `enercast`<sup>2</sup>.

Nous suivons une procédure habituelle en séparant le jeu de données en deux

- Calibration: 00:00 du 01/01/2007 jusqu'à 23:00 du 31/12/2013.
- Validation: 00:00 du 01/01/2014 jusqu'à 23:00 du 31/12/2014.

Nous utilisons la MAPE (*Mean Absolute Proportional Error*) comme critère d'erreur qui est définie comme suit,

$$\text{MAPE}(y, \hat{y}) = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

où  $y = (y_1, \dots, y_n)$  est la demande effective,  $\hat{y} = (\hat{y}_1, \dots, \hat{y}_n)$  est la prévision et  $n$  c'est le nombre d'observations (ici  $n = 24$  ce qui correspond à des données horaires).

### 3.3 Quelques résultats

**Modèle de régression - HVB** Le temps d'exécution est d'environ 5 seconds .

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2.12	4.26	5.49	5.972	6.87	30.85

Table 2: Resumen de MAPE para el año 2014

	regular	feriado
Min.	2.120	3.19
1st Qu.	4.238	5.56
Median	5.470	9.08
Mean	5.730	13.76
3rd Qu.	6.828	22.08
Max.	20.960	30.85

Table 3: MAPE por tipo de día para el año 2014

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Min.	2.290	2.770	2.500	2.180	2.820	2.120	3.190
1st Qu.	4.128	4.198	4.270	3.982	4.258	4.625	5.002
Median	5.170	5.210	5.100	5.085	5.505	5.905	5.915
Mean	5.901	5.485	5.653	6.261	6.176	5.955	6.378
3rd Qu.	6.358	6.910	6.050	6.695	7.185	7.200	6.950
Max.	26.990	11.010	17.180	30.850	20.960	13.720	12.060

Table 4: MAPE por día para el año 2014

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.65	2.23	3.14	4.127	5.11	22.03

Table 5: Resumen de MAPE para el año 2014

**Modèles de séries temporelles** Les tables ci-dessus montrent les résultats de prévision pour les modèles de séries temporelles de référence.

## Remerciements

Ce travail a bénéficié du support de la *Agencia Nacional de Investigacion e Innovacion* (ANII) de l'Uruguay. Il a pris place dans le cadre du projet FSE.1.2013.1.10886 " *Modelos de prevision de demanda de corto plazo*".

<sup>2</sup><https://github.com/cugliari/enercast>