Pointwise forecast and prediction intervals in electricity demand and price

Paula Raña, Juan Vilar y Germán Aneiros

Departamento de Matemáticas, Universidade da Coruña

Machine Learning Workhop Galicia 2016
27 October 2016

UNIVERSIDADE DA CORUÑA

gmodes

Spanish Electricity Market

"Mercado Ibérico de la Electricidad" (MIBEL)

- Market operator: "Operador del Mercado Ibérico Español" (OMIE)
- System operator: "Red Eléctrica de España" (REE)

Daily market

24 hourly offers for the quantity of power (measured in Mega Watts per hour, MWh) and its price (measured in €/MWh)

Functional time series

Daily curves of electricity demand or price along 2012: $\{\chi_i\}_{i=1}^{365}$
Discretized curves: $\chi_i(t_j), j = 1, \ldots, 24$.
Functional time series: $\{\chi_i\}_{i=1}^{n}$

Electricity demand

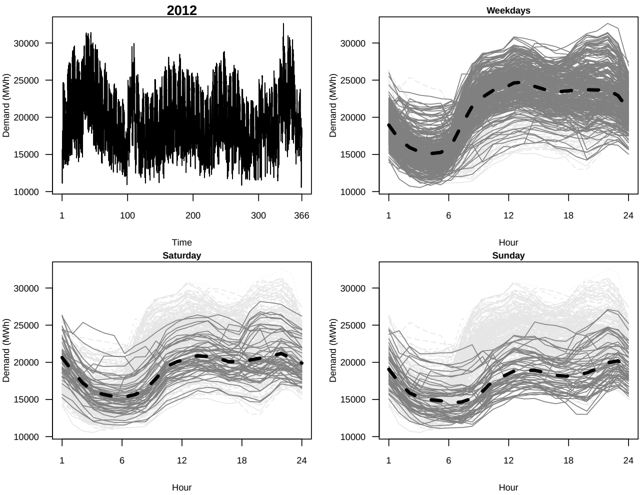


Figure : Electricity demand in 2012.

Electricity price

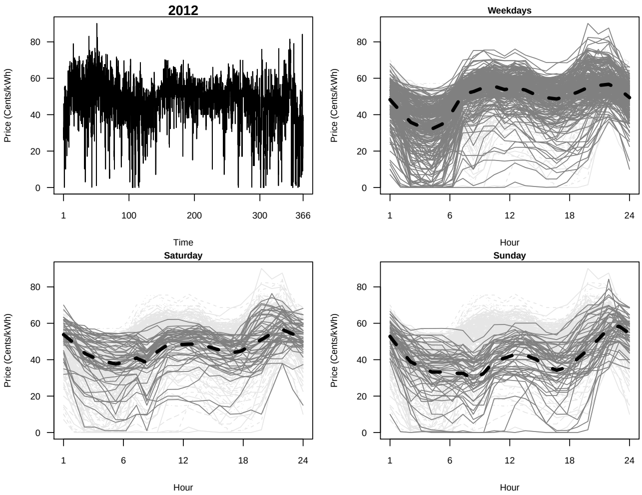
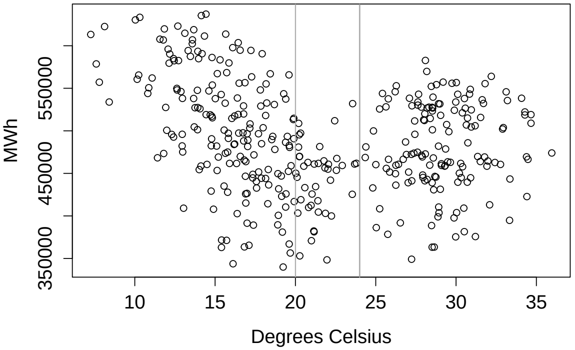


Figure : Electricity price in 2012.

Temperature

Daily maximum temperature in Spain
Non-linear relationship $\longrightarrow$ U-shaped
Cumulative daily demand $\sim$ Maximum daily temperature
Comfort zone (20,24) with no effect.



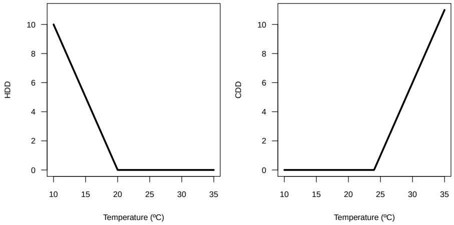
Temperature–derived functions

Heating/Cooling Degree Days

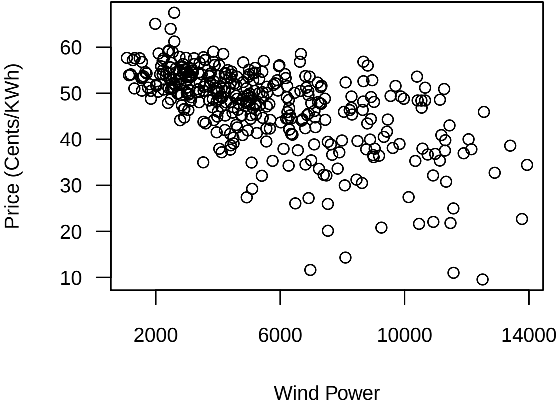
HDD/CDD: Measurement of the amount of energy needed to heat/cool a building.

$$HDD(t) = \max\{20 - T(t), 0\}$$

$$CDD(t) = \max\{T(t) - 24, 0\}$$



Wind Power Production



Introduction

Comparative study of prediction methods applied to electricity data: demand and price.

Extension of:

- Vilar JM, Cao R, Aneiros G. (2012), Forecasting next-day electricity demand and price using nonparametric functional methods, *International Journal of Electrical Power and Energy Systems*, 39: 48–55.

Functional regression methods: functional response and external covariates. Combined predictions.

- Aneiros, G., Vilar, J., and Raña, P. (2016)
  Short-term forecast of daily curves of electricity demand and price. *International Journal of Electrical Power and Energy Systems*, 80, 96–108.

Comparison methods

Naive

Demand/price in one day = demand/price in the previous day of the same class.

RFPCA

Hyndman and Ullah (2007), available at ftsa package in R.

ARIMA

24 separate univariate time series, one for each hour of the day.

Functional Nonparametric model with functional response

- Functional response variable: χ_{i+1}.
- Functional explanatory variable: χ_i.

$\mathcal{I}_0 = \{N-364, N-363, \ldots, N-1, N\}$,
$\mathcal{I}_{Sat} = \left\{ j \in \mathcal{I}_0, \text{ such that } \chi_j \text{ is a Saturday } \right\}$

Model

$$\chi_{i+1} = m\left(\chi_i\right) + \varepsilon_{i+1}, \;\; i+1 \in \mathcal{I}_{Sat}$$

$$\widehat{\chi}_{N+1} = \widehat{m}_h^{FNP}\left(\chi_N\right)$$

3 models: Weekdays, Saturday, Sunday.

FNP model

m will be estimated using a Nadaraya-Watson type estimator:

$$\widehat{m}_h^{FNP}(\chi_N) = \sum_{i \ /i+1 \in \mathcal{I}_{Sat}} w_h(\chi_N, \chi_i)\chi_{i+1}$$

$$w_h(\chi_N, \chi_i) = \frac{K\left(d(\chi_N, \chi_i)/h\right)}{\sum_{j \ /j+1 \in \mathcal{I}_{Sat}} K\left(d(\chi_N, \chi_j)/h\right)},$$

K is a kernel function, h is a smoothing parameter and d is a semimetric.

FNP model with scalar response

- Scalar response variable: $\chi_{i+1}(j)$, $j = 1, \ldots, 24$.
- Functional explanatory variable: χ_i.

24 models

$$\chi_{i+1}(j) = m\left(\chi_i\right) + \varepsilon_{i+1}, \;\; i+1 \in \mathcal{I}_{Sat}, j = 1, \ldots, 24$$

$$\widehat{\chi}_{N+1}(j) = \widehat{m}_h^{FNP}\left(\chi_N\right), j = 1, \ldots, 24$$

72 models: 24 $\times$ Weekdays, Saturday, Sunday.

Semi–functional partial linear model

- Functional response variable: χ_{i+1}.
- Functional explanatory variable: χ_i.
- Exogenous scalar variables: $\boldsymbol{X}_{i+1}^T = (x_{i+1\ 1}, \ldots, x_{i+1\ p})$.

Model

$$\chi_{i+1} = \boldsymbol{X}_{i+1}^T \beta + m(\chi_i) + \varepsilon_{i+1}, \ \ i+1 \in \mathcal{I}_{Sat}$$

$$\widehat{\chi}_{N+1} = \boldsymbol{X}_{N+1}^T \widehat{\beta} + \widehat{m}(\chi_N)$$

SFPL model

$$\widetilde{\mathbf{X}}_h = (\mathbf{I} - \mathbf{W}_h)\mathbf{X} \text{ and } \widetilde{\chi}_h = (\mathbf{I} - \mathbf{W}_h)\chi$$

where $\mathbf{W}_h = (w_h(\chi_i, \chi_j))_{i+1, j+1 \in \mathcal{I}_{Sat}}$, $\mathbf{X} = (x_{i+1j})_{\substack{i+1 \in \mathcal{I}_{Sat} \\ 1 \leq j \leq p}}$ and $\chi = (\chi_{i+1})_{i+1 \in \mathcal{I}_{Sat}}$

The estimator for β is defined by:

$$\widehat{\beta}_h = (\widetilde{\mathbf{X}}_h^T \widetilde{\mathbf{X}}_h)^{-1} \widetilde{\mathbf{X}}_h^T \widetilde{\chi}_h$$

and finally

$$\widehat{m}_h^{SFPL}(\chi) = \sum_{i+1 \in \mathcal{I}_{Sat}} w_h(\chi, \chi_i) \left(\chi_{i+1} - \boldsymbol{X}_{i+1}^T \widehat{\beta}_h \right)$$

SFPL model

SFPL1: 3 models for Weekdays, Saturday and Sunday

Demand

$$\boldsymbol{X} = (x_1, x_2)^T = (HDD, CDD)^T$$

Price

$$\boldsymbol{X} = (x_1, x_2)^T = (D, WPP)^T$$

SFPL2: 1 model

$x_3 = I_{Saturday}$, $x_4 = I_{Sunday}$ and $x_5 = I_{Monday}$ $\boldsymbol{X} = (x_1, x_2, x_3, x_4, x_5)^T$

SFPL model with scalar response

- Scalar response variable: $\chi_{i+1}(j)$, $j = 1, \ldots, 24$.
- Functional explanatory variable: χ_i.
- Exogenous scalar variables: $\boldsymbol{X}_{i+1}^{T} = (x_{i+1\ 1}, \ldots, x_{i+1\ p})$.

24 models

$$\chi_{i+1}(j) = \boldsymbol{X}_{i+1}^{T}\beta + m\left(\chi_i\right) + \varepsilon_{i+1},\ \ i+1 \in \mathcal{I}_{Sat}, j = 1, \ldots, 24.$$

$$\widehat{\chi}_{N+1}(j) = \boldsymbol{X}_{N+1}^{T}\widehat{\beta} + \widehat{m}\left(\chi_N\right), j = 1, \ldots, 24.$$

Combined methods

Combinations of the methods to forecast curves.

Combined forecasting 1

Average all the individual methods.

Combined forecasting 2

For each group of days (weekdays, Saturday and Sunday), average the two best predictors.

Error measurement

Demand errors: Integrated absolute percentage error (IAPE)

$$IAPE_{N+1} = \frac{1}{24}\int_0^{24} APE_{N+1}(t)dt \approx \frac{1}{24}\sum_{j=1}^{24} APE_{N+1}(j),$$

where

$$APE_{N+1}(t) = 100 \times \left| \frac{\widehat{\chi}_{N+1}(t) - \chi_{N+1}(t)}{\chi_{N+1}(t)} \right|.$$

Price errors: Integrated absolute error (IAE)

$$IAE_{N+1} = \frac{1}{24}\int_0^{24} AE_{N+1}(t)dt \approx \frac{1}{24}\sum_{j=1}^{24} AE_{N+1}(j),$$

where

$$AE_{N+1}(t) = \left|\widehat{\chi}_{N+1}(t) - \chi_{N+1}(t)\right|.$$

Table : Mean of the IAPE for the electricity demand curves.

	Method	2012
Functional	Naïve	6.39
	RFPCA	6.00
	FNP	6.05
	SFPL1	**5.78**
	SFPL2	6.03
Combined	CF1	5.46
	CF2	**5.40**
Scalar	ARIMA	**5.90**
	FNP sc	6.20
	SFPL1 sc	5.97

Table : Mean of the IAE for the electricity price curves.

	Method	2012
Functional	Naïve	6.83
	RFPCA	5.94
	FNP	6.36
	SFPL1	5.15
	SFPL2	**5.00**
Combined	CF1	5.11
	CF2	**4.94**
Scalar	ARIMA	6.27
	FNP sc	6.41
	SFPL1 sc	**5.22**

Introduction

Scalar response

$$Y = m(\boldsymbol{\chi}) + \varepsilon$$

Confidence Interval

$$\mathbb{E}(Y|\boldsymbol{\chi} = \chi) = m(\chi)$$

Prediction Interval

$$Y/\chi = m(\chi) + \varepsilon$$

- Asymptotic theory
- Bootstrap

Prediction Interval in FNP model

$$Y/\chi = m(\chi) + \varepsilon = \widehat{m}_h(\chi) + m(\chi) - \widehat{m}_h(\chi) + \varepsilon.$$

where we can approximate

- $m(\chi) - \widehat{m}_h(\chi)$ by $\widehat{m}_b(\chi) - \widehat{m}^*_{hb}(\chi)$
- ε by $\widetilde{\varepsilon}$

Bootstrap $(1-\alpha)$-prediction intervals for Y/χ

$$I^*_{\chi,1-\alpha} = (\widehat{m}_h(\chi) + q^*_{\alpha/2}(\chi), \widehat{m}_h(\chi) + q^*_{1-\alpha/2}(\chi))$$

Prediction Interval in SFPL model

$$
\begin{aligned}
Y/\{\boldsymbol{X},\chi\} = \quad & \boldsymbol{X}^T\boldsymbol{\beta} + m(\chi) + \varepsilon = \\
& \boldsymbol{X}^T\widehat{\boldsymbol{\beta}}_h + \boldsymbol{X}^T(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_h) + \widehat{m}_h(\chi) + m(\chi) - \widehat{m}_h(\chi) + \varepsilon,
\end{aligned}
$$

where we can approximate

- $\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_h$ by $\widehat{\boldsymbol{\beta}}_b - \widehat{\boldsymbol{\beta}}_b^*$
- $m(\chi) - \widehat{m}_h(\chi)$ by $\widehat{m}_b(\chi) - \widehat{m}_{hb}^*(\chi)$
- ε by $\widetilde{\varepsilon}$

Bootstrap $(1-\alpha)$-prediction intervals for Y/χ

$$
I^*_{\boldsymbol{X},\chi,1-\alpha} = (\boldsymbol{X}^T\widehat{\boldsymbol{\beta}}_h + \widehat{m}_h(\chi) + q^*_{\alpha/2}(\boldsymbol{X},\chi), \boldsymbol{X}^T\widehat{\boldsymbol{\beta}}_h + \widehat{m}_h(\chi) + q^*_{1-\alpha/2}(\boldsymbol{X},\chi))
$$

Electricity demand

Dataset: workdays of the second quarter of the year 2012.
Predict one day (24 hours).

FNP model

$$\chi_{i+1}(t) = m_t(\chi_i) + \varepsilon_{i,t} \ (t = 1, \ldots, 24, \ i = 1, \ldots, n);$$

SFPL model

$$\chi_{i+1}(t) = \boldsymbol{X}_i^T \beta + m_t(\chi_i) + \varepsilon_{i,t} \ (t = 1, \ldots, 24, \ i = 1, \ldots, n);$$

Temperature covariates: $\boldsymbol{X}_i = (X_{i1}, X_{i2})^T = (HDD_i, CDD_i)^T$

Prediction intervals for electricity demand

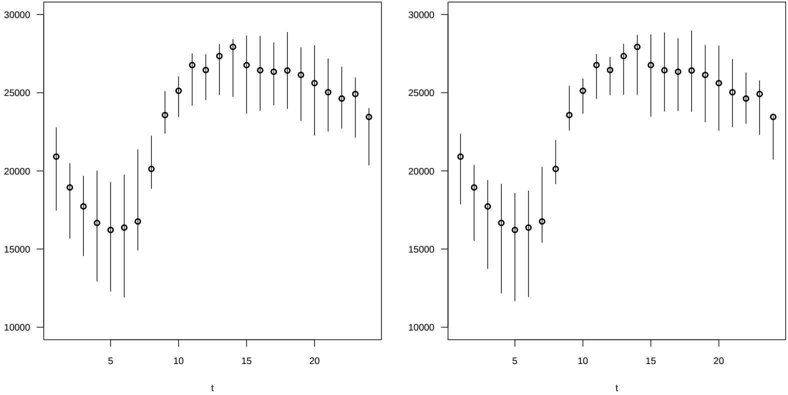


Figure : Bootstrap prediction intervals for the electricity demand. Left: FNP, right: SFPL.

Prediction density

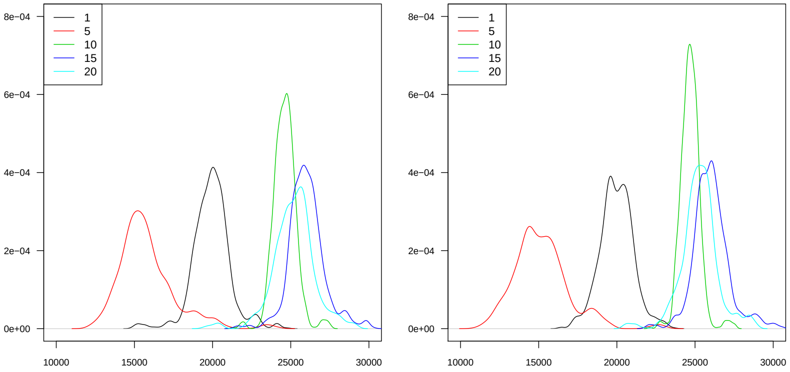


Figure : Prediction density for the electricity demand using the FNP (left) or SFPL (right), along June 29, 2012.

Software

R

Packages:

- fda.usc
- npfda routines
- ftsa
- TSA

Execution time

- FNP: fast
- SFPL: slow
  - predict one day (functional response): 30'
  - predict one hour (scalar response):6' $\Rightarrow$ 24 hours: 2.5h.

References

- Aneiros, G., Vilar, J., and Raña, P. (2016), Short-term forecast of daily curves of electricity demand and price, *International Journal of Electrical Power and Energy Systems*, 80, 96–108.
- Vilar JM, Cao R, Aneiros G. (2012), Forecasting next-day electricity demand and price using nonparametric functional methods, *Electr Power Energy Syst*, 39: 48–55.
- Raña, P., Aneiros, G., Vilar, J. and Vieu, P. (2016) Bootstrap confidence intervals in functional nonparametric regression under dependence. *Electronic Journal of Statistics*, 10(2), 1973–1999.

Pointwise forecast and prediction intervals in electricity demand and price

Paula Raña, Juan Vilar y Germán Aneiros

Departamento de Matemáticas, Universidade da Coruña

Machine Learning Workhop Galicia 2016
27 October 2016