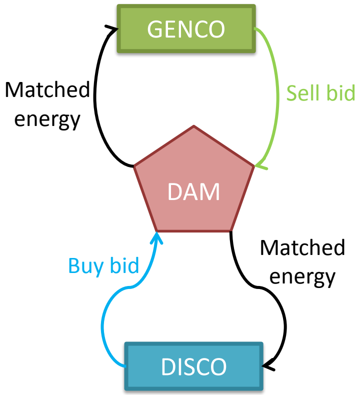
Improving electricity market price scenarios by means of forecasting factor models

M.P. Muñoz, C. Corchero and F.J. Heredia
Department of Statistics and Operational Research
Universitat Politècnica de Catalunya
pilar.munyoz@upc.edu

August 19, 2009

Day-Ahead Market Structure



GENCO:

- Generation Company
- Daily participation in the pool
- Bidding strategy to sell the energy production

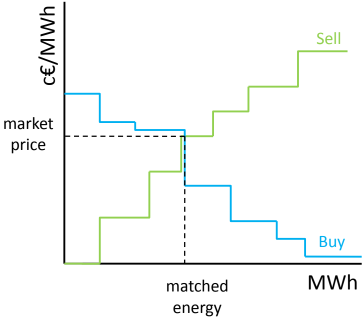
DAM:

- Day-Ahead Market
- Hourly auction
- Matching procedure 24h before the delivery period

DISCO:

- Distribution Company

Matching procedure

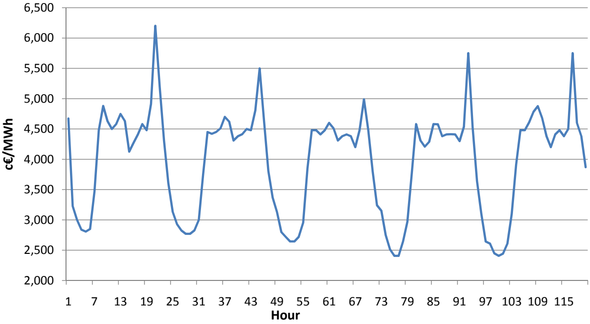


The market price results from the matching procedure:

- An offer consists in a pair of energy and price.
- The offers for selling and buying energy are sorted and the matched energy and price is fixed.
- This process takes place the day before the energy is produced.

The GenCo needs to know the price at which the energy will be paid in order to decide how to bid and how to schedule their resources for maximizing their profit.

Price characteristics



Electricity spot prices exhibit:

- Non-constant mean and variance
- Daily and weekly seasonality
- Calendar effects
- High volatility and presence of outliers

Objectives

Forecasting objectives

To apply the methodology of factor models to forecast electricity market prices in a short-term horizon

Application objectives

To include short-term forecasting of the electricity market spot price in a stochastic optimization model for the management of a GenCo in order to obtain realistic market price scenarios in which the GenCo should decide how to optimally operate.

Objectives

Forecasting objectives

To apply the methodology of factor models to forecast electricity market prices in a short-term horizon

Application objectives

To include short-term forecasting of the electricity market spot price in a stochastic optimization model for the management of a GenCo in order to obtain realistic market price scenarios in which the GenCo should decide how to optimally operate.

Previous approaches

Parametric/non-parametric and conditional homocedastic/heterocedastic approaches has been proposed. For example:

- Non-parametric statistic methods - such as clustering or bootstrapping - applied to historical data.
- ARIMA models
- Neural networks models
- Dynamic regressions

However, the residuals exhibit non stationary conditional variance in most of the analyzed models

New approach

To apply the methodology of factor models in the next way:

- The spot price is interpreted not as a single time series but a set of 24 time series, one for each hour.
- The factor model allows to identify common unobserved factors which represent the relationship between the hours of a day.
- The forecasting model provide suitable scenarios for the optimization model.

Schema

Hourly electricity prices

↓

Factor analysis

⇓

One-step-ahead forecast price and confidence interval

⇓

Scenario generation

⇓

Optimization model

Time Series Factor Analysis

Time Series Factor Analysis

Time Series Factor Analysis [a] (*TSFA*) estimates measurement model for time series data with as few assumptions as possible about the dynamic process governing the factors. It estimates parameters and predicts factor scores.

[a] Gilbert P.D., Meijer E. (2005). Time Series Factor Analysis with an Application to Measuring Money

Time Series Factor Analysis vs. other methods

TSFA vs. standard Factor Analysis

- the factor model has a nontrivial mean structure
- the observations are allowed to be dependent over time
- the data does not need to be covariance stationary as long as differenced data satisfies a weak boundedness condition.

TSFA vs. Dynamic Factor Analysis

- TSFA estimates parameters and predicts factor scores with few assumptions about factor dynamics
- DFA assumes a predetermined relationship between the factors at time t and at time t-1, if this relationship is misspecified, the factors estimated by DFA can be biased.

Factor Model Estimation (I/II)

Let y_t be a M-vector of observed time series of length T and k unobserved factors ($k << M$) collected in the K-vector ξ. The relationship between the observed time series and the factors is assumed to be linear and described by equation:

$$y_t = \alpha_t + B\xi_t + \epsilon_t$$

where α_t is an M-vector of intercept parameters, B is an Mxk matrix parameter of loadings, assumed time-invariant, and ϵ is a random M-vector of measurement errors.

Factor Model Estimation (II/II)

Defining D as the difference operator $y_t = \alpha_t + B\xi_t + \epsilon_t$ becomes: $Dy_t = \tau_t + BD\xi_t + D\epsilon_t$ and the following conditions are assumed:

$$\sum_{t=1}^{T} \frac{D\xi_t}{T} \xrightarrow{d} \kappa$$

$$\sum_{t=1}^{T} \frac{(D\xi_t-\kappa)(D\xi_t-\kappa)'}{T} \xrightarrow{d} \Phi$$

$$\sum_{t=1}^{T} \frac{D\epsilon_t D\epsilon_t'}{T} \xrightarrow{d} \Omega$$

The sample covariance of the differenced series Dy_t is denoted by S_{D_y} and from the previous assumptions, it follows that $S_{D_y} \xrightarrow{d} \Sigma \equiv B\Phi B' + \Omega$

Parameters are estimated by maximum likelihood, minimizing the function:

$$L \equiv \lg \det \Sigma + \mathrm{tr}(\Sigma^{-1} S_{D_y})$$

Forecasting model

The factors obtained have to be implemented into a forecasting model in order to obtain the price forecasts.

A one-step-ahead forecasting model is specified and estimated as a linear multiple regression model with the factors as predictors[1]:

$$y_{t+1} = \beta\hat{\xi}_t + \alpha(L)y_t + \varepsilon_{t+1}$$

The out of the sample forecast for $y_{T+1|T}$ is given by the conditional expectation

$$y_{T+1|T} = \hat{\beta}\hat{\xi}_T + \hat{\alpha}(L)y_T$$

[1] Stock J., Watson M.W. (2002). Forecasting Using Principal Components From a Large Number of Predictors

Stochastic optimization model

Stochastic optimization

Optimization models which incorporate probabilistic elements, either in the problem data. In this case, the market incomes depend on a random variable, the market price.

Generation Company Management

The Electricity Generation Companies must bid daily in the Day-Ahead Market in order to sell its energy. At the same time, they must optimize its production taking into account the operational limits of the units.

Formulation (I/II)

The stochastic approach implies the discretization of the random variable in a set of scenarios with its corresponding probabilities.

It is maximized the expected benefits expressed as the difference between the production costs and the market incomes.

In this work, the set of scenarios is build from the discretization of the confidence interval of the forecast.

Formulation (II/II)

maximize E_y(market incomes - production costs)

subject to Futures contracts covering

Zero-price bid

Maximum and minimum production level

Other operational constraints

Data analysis

- Random variable: Iberian Day-Ahead Market electricity prices
- Data set: work days from January 1^{rts}, 2007 to March 30^{th}, 2008.
- 3 significant factors, based on eigenvalues of the sample correlation matrix.
- The data has been analyzed using R (version 2.7.0) with the library TSFA available at CRAN (*www.cran.r-project.org*).

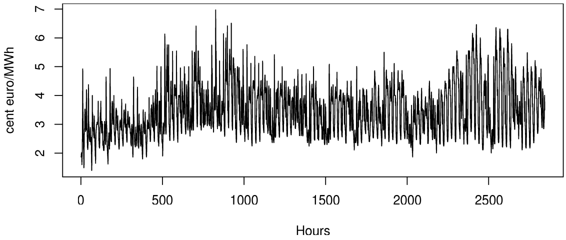


Figure 1: Iberian Day-Ahead Electricity Market price (January 1rst, 2007 - March 30th, 2008)

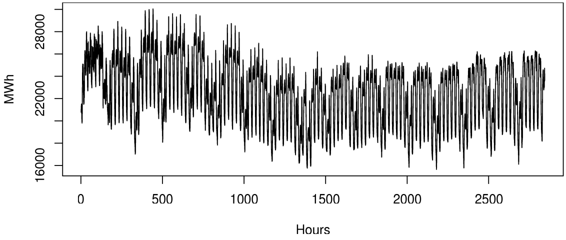


Figure 2: Iberian Day-Ahead Electricity Market demand (January 1rst, 2007 - March 30th, 2008)

24 Time Series

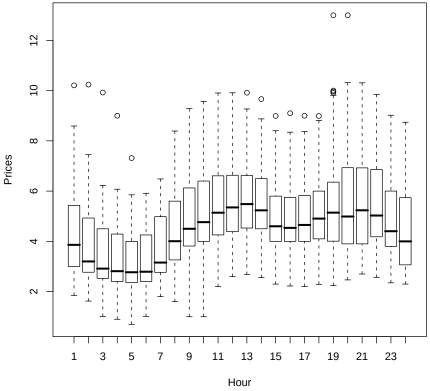


Figure 3: Iberian Day-Ahead Electricity Market price for each hour(January 1rst, 2007 - March 30th, 2008)

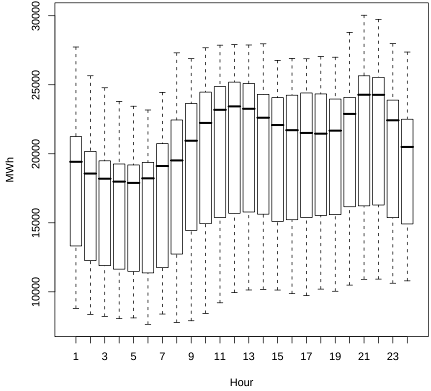


Figure 4: Iberian Day-Ahead Electricity Market demand for each hour (January 1rst, 2007 - March 30th, 2008)

Price vs. Demand

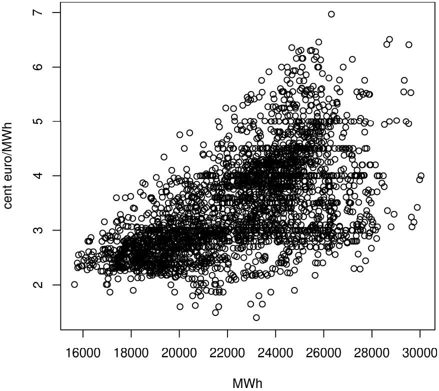


Figure 5: Iberian Day-Ahead Electricity Market price vs. demand (January 1rst, 2007 - March 30th, 2008)

Factor model results

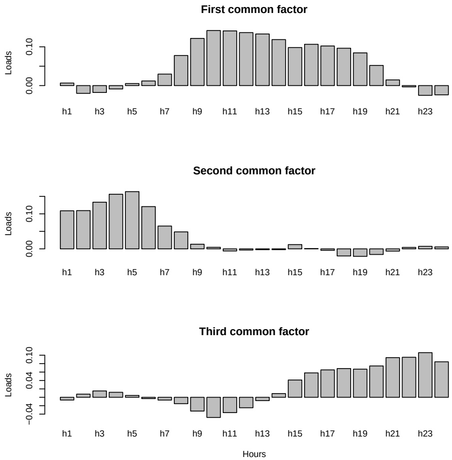


Figure 6: Loads of the common factors

Out of sample forecasting results (I/II)

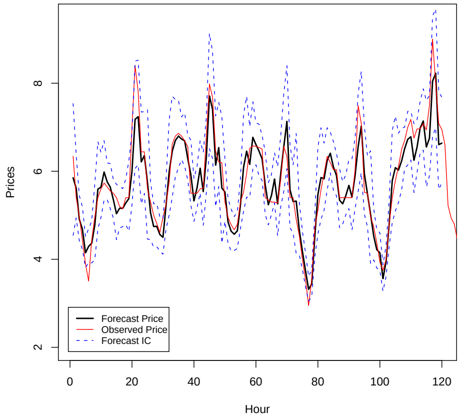


Figure 7: One-step-ahead forecast prices

Out of sample forecasting results (II/II)

Hour	1	2	3	4	5	6
R^2	99.1	95.3	97.1	99.8	99.8	97.6
MSE	0.017	0.004	0.003	0.003	0.002	0.002
Hour	7	8	9	10	11	12
R^2	96.0	99.6	99.7	99.8	96.3	98.3
MSE	0.003	0.008	0.008	0.004	0.003	0.001
Hour	13	14	15	16	17	18
R^2	99.9	97.7	99.8	99.9	99.9	97.1
MSE	0.002	0.002	0.004	0.002	0.002	0.002
Hour	19	20	21	22	23	24
R^2	99.7	96.6	94.2	99.7	99.7	95.1
MSE	0.006	0.005	0.007	0.007	0.007	0.005

Table 1: Summary of the forecast models for each hour

Optimization model results

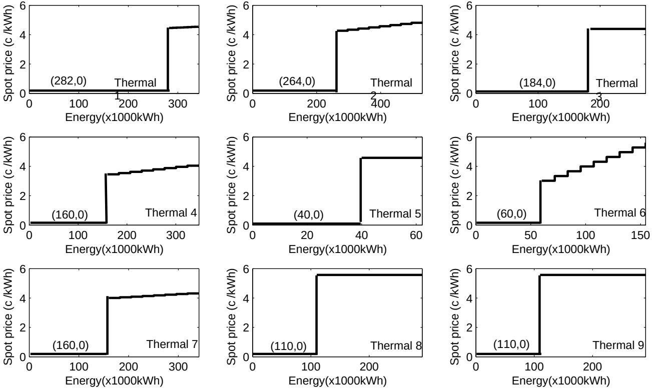


Figure 8: Bidding curve for each unit at hour 20

Conclusions

- The forecast procedure based on factor models gives suitable results.
- These results are equivalent to the ones obtained through an ARIMA model.
- The advantage of the procedure presented lies in its simplicity, easy to implement and to present.
- The improved forecasts have been used to successfully generate a set of scenarios to feed the stochastic optimization model.
- This set of scenarios is also built in an automatic way from the forecast confidence interval.

Thanks!