A New Model for Electricity Price Series Modelling and Forward and Volatility Curves Computation

Julián Barquín, *Member, IEEE,* Ángel Garro, Eugenio Fco. Sánchez-Úbeda, Santiago Tejero

***Abstract*—As a result of the deregulation processes, liberalized markets, where electricity futures and derivatives are traded, have arisen all over the world. Utilities, consumers, traders and, generally, market agents must do quantitative assessments of their positions. Basic analytical data are the forward and volatility curves of the traded products. Ideally, these curves should be derived from publicly available traded prices. However, as electricity derivative markets are usually rather thin, alternative procedures based on spot prices and theoretical computations must be used instead. In this paper, a new spot price evolution model is proposed and the resulting theoretical forward and volatility curves derived.**

***Index Terms*— Electricity price, spot price, futures, derivatives, volatility.**

I. INTRODUCTION

During the 90's, a great number of countries undertook restructurating plans in electricity industry and still today reforms are carried out. One consequence is the creation of a pool market, where electricity prices are determined by the competitive action of buyers and sellers. Uncertainty in prices and other many issues in the new framework as regulation changes or the entrance of new competitors need new approaches where risk performs a critical role. This paper is focused on the study of prices behaviour and futures valuation, covering a key part in risk analysis.

A worthy measure of uncertainty is the volatility based on market prices which is generally higher in electricity markets than in financial ones. Electricity prices have a wide range of features as we detail next [1,2,3].

- Market prices are strongly seasonal because of storing difficulties and strong seasonal consumption patterns.
- Generally, spot prices are subject to frequent jumps and spikes caused by either transmission and capacity constraints or oligopolistic behaviour where different agents with market power drive prices upward. Depending on facilities and agents in every market, frequency and intensity of jumps can vary. When the spikes are not very frequent, the study of jumps and spikes is more important in stress testing than valuing forwards and futures, where extreme values are not expected by the market, so in order to ensure quantitative methods work reasonably, a sensible approach should be used to eliminate them.
- Owing to the non storability of electricity before mentioned, prices in temporally extreme values revert within hours or days to a more stable level. This effect is called mean reversion and it is characteristic of flow commodities.

Therefore, electricity price models are very different from traditional financial price models, which do not share these characteristics.

Futures and other electricity derivatives are an increasingly used tool for managing market risk. In this regard, basic analytical data are the forward and volatility curves of the traded products. Ideally, these curves should be derived from publicly available traded prices. However, as electricity derivative markets are usually rather thin, alternative procedures based on spot prices and theoretical computations must be used instead.

This paper proposes a new model able to cope with the specific characteristics of electricity prices. A model estimation procedure as well as the computation of the forward and volatilty curves are also proposed.

The sequel of this paper is organized as follows. In section II the model and the estimation procedure are stated. Next section deals with the forward and volatility curves computations. Next, section IV applies our procedures to real data. Finally, we conclude.

II. THE SPOT PRICE MODEL

The model describes the evolution of spot prices in terms of two components. The first one is a deterministic component, which explains the strong level of seasonality and regular behaviour that this kind of series displays, such as a trend or temporary patterns. The second one is the stochastic component of the model and refers to short-term deviation in prices. Moreover, additional components can be incorporated into the model with the intention to deal with jumps or long-term evolution of prices. Nevertheless, in this paper, only the

Julián Barquín and Eugenio Fco. Sánchez-Úbeda are with IIT, Pontificia Comillas University, C/ Alberto Aguilera 23, 28015 Madrid, Spain.

Angel Garro was with IIT as well at the moment of doing the work herein reported. He is presently with Iberdrola Gas Trading, C/ Tomás Redondo 1, 28033 Madrid, Spain.

Santiago Tejero is with Endesa, C/ Ribera del Loira 60, 28042 Madrid, Spain.

Corresponding author is Dr. Barquín (e-mail: julian.barquin@iit.upco.es).

two former components are taken into account, and the price evolution is defined as the sum of them.

Let P_t be the spot price of a commodity at each instant of time t, μ_t and y_t the deterministic and variable components respectively.

$$\ln\left(P_t\right) = \mu_t + y_t \tag{1}$$

In this paper, daily electricity prices are been studied, as many derivatives involve these prices. Therefore, subscript *t* denotes days. Traded derivatives also involve specific hourly averages, as "peak", "off-peak", "morning" and "business" prices. Proper modelling of these products requires either hourly or multivariate time series models. Although this is an important issue, it will not further pursued in the sequel.

A. The deterministic component

Computation of the deterministic component is done on the basis of an assumed time series seasonality. For instance, a monthly an a weekday seasonality could be, as in the example at the end of the paper, assumed. In this case, the deterministic component is written as:

$$\begin{aligned} \mu_t \quad = \quad & \sum_{month} \alpha_{month} c_{month}\left(t\right) + \\ & \sum_{week-day} \alpha_{week-day} c_{week-day}\left(t\right) \end{aligned} \tag{2}$$

where $c_{month}\left(t\right)$ and $c_{week-day}\left(t\right)$ are indicator functions, that is, $c_{January}\left(t\right) = 1$ if sample t belongs to January and 0 otherwise, and so on.

Estimation of coefficients $\alpha_{week-day}$ and α_{month} is done by recursive averaging. That is, α_{month} is firstly estimated by computing the time series average for each different month. For instance:

$$\alpha_{January} = \frac{\sum_{t \in January} \ln\left(P_t\right)}{\text{number of days in Januarys}} \tag{3}$$

and so on. Now, the weekday coefficients are computing by averaging the residual series. For instance:

$$\alpha_{Monday} = \frac{\sum_{t \in Monday} \left[\ln\left(P_t\right) - \sum_{month} \alpha_{month} c_{month}\left(t\right)\right]}{\text{number of days which are mondays}} \tag{4}$$

More complex seasonality structures (as those taking into account public holydays) can be handled in a similar way.

B. The stochastic component

Once that the deterministic component has been computed, historical stochastic time series y_t can be found by subtracting the deterministic component from the original time series. Now, a vector autoregressive model composed by several autoregressive lag polynomials is fitted to the component y_t. Lag selection criterion is based on the periodic patterns displayed by spot prices due to the deterministic component is not able to capture all the seasonality of the market.

Given a certain polynomial structure, the computation of the coefficient values is done by a least squares method. The identification of the polynomials order is based on the computation of the Aikine index (See [4]).

In detail, the stochastic component y_t is modelled according to:

$$H\left(q^{-1}\right) y_t = u_t \tag{5}$$

Being q^{-1} the lag operator and u_t a residuals series. Furthermore, $H\left(q^{-1}\right)$ is the product of possibly several polynomials.

$$H\left(q^{-1}\right) = \prod_i H_i\left(q^{-1}\right) \tag{6}$$

Each polynomial has a characteristic base lag. For instance, in daily electricity prices, we expect characteristic lags to be daily and weekly ones. Therefore, an admissible model is:

$$\begin{aligned} & H\left(q^{-1}\right) = H_1\left(q^{-1}\right) H_2\left(q^{-1}\right) \\ & H_1\left(q^{-1}\right) = 1 - a_1 q^{-1} - a_2 q^{-2} - a_3 q^{-3} \\ & H_2\left(q^{-1}\right) = 1 - b_1 q^{-7} - b_2 q^{-14} \end{aligned} \tag{7}$$

Computation of coefficients a_1, a_2, a_3, b_1 and b_2 is done by least squares, minimizing the sum of the squares of the residuals. Order computation (how many a's and b's) is done by applying the Aikine criterion, that is, minimizing the addition of the sum of the squares of the residuals plus a term related to the number of estimated parameters [4].

C. The residuals

As a subproduct of the stochastic component model estimation, a time series of historical residuals u_t is obtained. Usually, squared residuals exhibit strong autocorrelation. This means that they are not independent and also indicates the existence of clusters in volatility. Tracking commodities prices series, it is seen that their behaviour is more normal during quiet periods and more anomalous during crisis periods, so these series exhibit time varying volatility, and Generalized Autoregressive Conditional Heteroskedastic (GARCH) models are reasonable models to identify these temporary volatility changes in the residuals [5]:

$$\begin{aligned} & u_t = k_1 + \varepsilon_t \\ & \sigma_{\varepsilon,t}^2 = k_2 + \sum_{i=1}^{P} a_i \sigma_{\varepsilon,t-i}^2 + \sum_{i=1}^{Q} b_i \varepsilon_{t-i}^2 \\ & \varepsilon_t \sim \mathrm{N}\left(0, \sigma_\varepsilon\right) \end{aligned} \tag{8}$$

Computation, from the historical residuals series, of the number and values of the GARCH model parameters is done by using the Matlab Financial Toolbox.

III. Valuing Futures Contracts: Forward and Volatility Curves

The basic quantitative tools in derivatives trading are the forward and the volatility curves.

A. *The forward curve*

The forward curve is the discounted risk-neutral expectation of the electricity prices for several delivery dates. Several remarks are in order:

1. Risk-neutral expectation is generally different from natural expectation, because of agents risk aversion. However, it will be assumed in the sequel that both operators are identical. This can be justified along the lines in [1]. Alternatively, a risk aversion parameter could be added (as in, for instance, [2]) and calibrated from market forward curves.
2. The rate of interest is variable over time. However currently interest rates are very low, so variations in risk-free rate have also low impact. Therefore, a constant and known interest rate will be assumed in the sequel.
3. Most electricity derivative contracts are based on the average price of electricity in a specified period (for instance, average electricity March price). Even if the daily prices were described by a log-normal distribution (which they are not, because of GARCH effects), the average of log-normal distributions is not log-normal. So, analytical forms of the forward (and volatility) curves are very difficult or impossible to obtain. Most of the results in the sequel are useful approximations to the "exact" theoretical curves.

Specifically, let $F_{t,T}$ be the market price at time t for a futures contract with time maturity time T, and r the rate of interest.

$$F_{t,T} = \mathrm{E}_t\left[P_T\right]e^{-r(T-t)} = \mathrm{E}_t\left[e^{\tilde{y}_{t,T}+\mu_T}\right]e^{-r(T-t)} \tag{9}$$

$\mathrm{E}_t\left[\bullet\right]$ is the expectation operator at time t. A tilde is added when the stochastic character of a variable is to be emphasized.

Stationary autoregressive models are equivalent to infinite moving average ones. Specifically, let us define the polynomial $G\left(q^{-1}\right)$:

$$G\left(q^{-1}\right)H\left(q^{-1}\right)=1 \tag{10}$$

Even if the order of $G\left(q^{-1}\right)$ is infinite, its coefficients are rapidly decreasing and can be computed efficiently in great number. So, in a moving average representation:

$$\tilde{y}_{t,T} = \sum_{l=1}^{R} M_l^T y_l^0 + \sum_{n=1}^{T-t} G_n \tilde{u}_{T-n+1} \tag{11}$$

Where G_n is the *n*-th coefficient of polynomial $G\left(q^{-1}\right)$, y_l^0 the R initial conditions, and M_l^T (that have a decreasing influence on initial conditions) is defined as:

$$M_l^T = \sum_{n=1}^{l} G_{R+n-l} G_{T-n+1} \tag{12}$$

Fig. 1 shows the structure of the previous equation. Valuation (*t*) and maturity *(T)* times have been greyed.

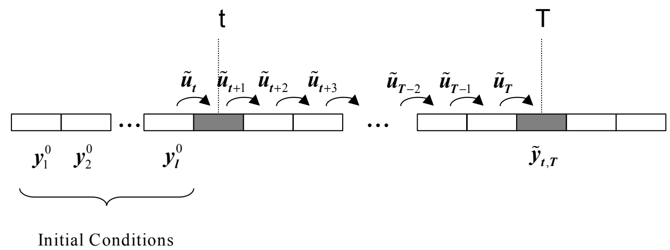

Fig. 1. Structure of equation (12).

Now, the value of the future on average prices in the period $\left[T_1, T_2\right]$ at current date $t=0$ can be computed. These computations take two steps: firstly to calculate the expectation of prices at the beginning of delivery period, considering the initial conditions as known. And secondly to develop the expectancy of the first step term at current time, where initial conditions are not known and become the variability source. Figure 2 represents the approach. The greyed boxes represent the period which defines the product, and $\overline{S}_{T_1\to T_2}$ the discounted to T_1 average spot price.

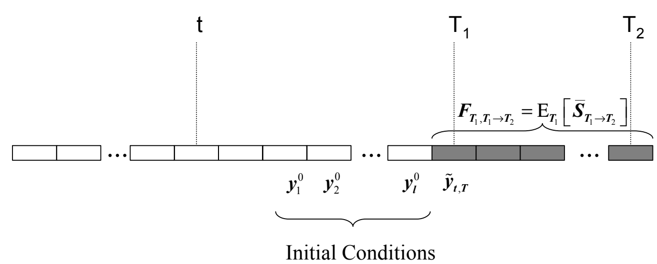

Fig. 2. Computing the value of the future

Then, let $F_{T_1,T_1\to T_2}$ the future contract value at time T_1 with delivery period from T_1 to T_2. This value is the expectation at time T_1 of average discounted spot price $\overline{S}_{T_1\to T_2}$:

$$\begin{aligned} F_{T_1,T_1\to T_2} &= \mathrm{E}_{T_1}\left[\overline{S}_{T_1\to T_2}\right] \\ &= \mathrm{E}_{T_1}\left[\frac{1}{T_2-T_1+1}\sum_{t=T_1}^{T_2} e^{\tilde{y}_{T_1,t}+\mu_t-r(t-T_1)}\right] \\ &\approx \frac{1}{T_2-T_1+1}\sum_{t=T_1}^{T_2} e^{\sum_{l=1}^{R} y_l^0 M_l^{t-T_1}+\mu_t-r(t-T_1)+\frac{1}{2}\mathrm{var}\left[\sum_{n=0}^{t-T_1} G_n\tilde{u}_{t-n}\right]} \end{aligned} \tag{13}$$

Where

$$\text{var}\left[\sum_{n=0}^{t-T_1} G_n \tilde{u}_{t-n}\right] = \sum_{n=0}^{t-T_1} G_n^2 \sigma_{t-n}^2 \tag{14}$$

This formula is approximate because prices distribution are not exactly log-normal ($\approx$ would be $=$ were $\tilde{y}_{T_1,t}$ normal [6]). However, as they are almost log-normal (but in the far-off tails), this is a useful and accurate approximation, as checked by Montecarlo simulation.

The variances of autoregressive model residuals σ_t^2 are forecasted using GARCH process parameters. At the beginning of the forecast, volatility depends strongly on the current variance of the valuation date, although the variance tends to an average value similar to the historical residual average variance.

In the second step the R initial conditions become stochastic variables:

$$y_l^0 = \tilde{y}_{0,T_1-R+l} \tag{15}$$

To work easily $\tilde{\varphi}_{t,T_1}$ is defined as the almost lognormally distributed variable:

$$\tilde{\varphi}_{t,T_1} = e^{\sum_{l=1}^{R} M_l^{t-T_1} \tilde{y}_{0,T_1-R+l} + f_{t,T_1}} \tag{16}$$

Also, other ancilliary variables are necessary

$$f_{t,T_1} = \mu_t - r\left(t-T_1\right) + \frac{1}{2}\text{var}\left[\sum_{n=0}^{t-T_1} G_n \tilde{u}_{t-n}\right] \tag{17}$$

$$f'_{t,T_1} = f_{t,T_1} + \sum_{l=1}^{R}\sum_{k=1}^{R} M_l^{t-T_1} y_k^0 M_k^{t-R+l} \tag{18}$$

Finally the future contract value at current time 0 is:

$$F_{0,T_1\to T_2} = \frac{e^{-rT_1}}{T_2-T_1+1}\sum_{t=T_1}^{T_2} \text{E}_0\left[\tilde{\varphi}_{t,T_1}\right] = \frac{e^{-rT_1}}{T_2-T_1+1}\sum_{t=T_1}^{T_2} e^{f'_{t,T_1} + \frac{1}{2}\text{var}\left[\sum_{l=1}^{R}\sum_{n=1}^{t-R+l} M_l^{t-R+l} G_n \tilde{u}_{t-R+l-n}\right]} \tag{19}$$

where

$$\text{var}\left[\sum_{l=1}^{R}\sum_{n=1}^{t-R+l} M_l^{t-R+l} G_n \tilde{u}_{t-R+l-n}\right] = \sum_{l=1}^{R}\sum_{n=0}^{t-R+l}\left(M_l^{t-R+l} G_n \sigma_{t-R+l-n}\right)^2 + 2\sum_{m=1}^{R}\sum_{\substack{p=1\\m<p}}^{R}\sum_{n=0}^{t-R+m} M_m^{t-R+m} M_p^{t-R+p} G_n G_{n+p-m} \sigma_{t-R+m-n}^2 \tag{20}$$

B. The volatility curve

Volatility is defined as the standard deviation of the distribution logarithm. As the assets are only approximately log-normal, the volatility so defined can not be used for valuing very out-of-the-money options or other similar derivatives.

Taking into account these *caveats*, let us compute the covariance matrix:

$$\text{cov}_0\left[\ln \tilde{F}_{0,T_1\to T_2}\ \ln \tilde{F}_{0,T_3\to T_4}\right] = \ln\left(\frac{\text{E}_0\left[\tilde{F}_{0,T_1\to T_2}\tilde{F}_{0,T_3\to T_4}\right]}{\text{E}_0\left[\tilde{F}_{0,T_1\to T_2}\right]\text{E}_0\left[\tilde{F}_{0,T_3\to T_4}\right]}\right) = \ln\left(\frac{\sum_{i=T_1}^{T_2}\sum_{j=T_3}^{T_4}\text{E}_0\left[\tilde{\varphi}_{i,T_1}\tilde{\varphi}_{j,T_3}\right]}{\sum_{i=T_1}^{T_2}\sum_{j=T_3}^{T_4}\text{E}_0\left[\tilde{\varphi}_{i,T_1}\right]\text{E}_0\left[\tilde{\varphi}_{j,T_3}\right]}\right) \approx \frac{\sum_{i=T_1}^{T_2}\sum_{j=T_3}^{T_4}\text{cov}\left[\ln\tilde{\varphi}_{i,T_1}\ \ln\tilde{\varphi}_{j,T_3}\right]}{\left(T_2-T_1+1\right)\left(T_4-T_3+1\right)} \tag{21}$$

Note that index *i* runs over the time periods belonging to the first product and *j* over those belonging to the second one. Besides,

$$\text{cov}\left[\ln\tilde{\varphi}_{i,T_1}\ \ln\tilde{\varphi}_{j,T_1}\right] = \tag{22}$$

$$\text{cov}\left[\sum_{l=1}^{R}\sum_{n=1}^{i-R+l} M_l^{i-R+l} G_n \tilde{u}_{i-R+l-n}, \sum_{p=1}^{R}\sum_{m=1}^{j-R+p} M_l^{j-R+p} G_m \tilde{u}_{j-R+p-m}\right] = \sum_{l=1}^{R}\sum_{\substack{p=1\\p\geq l}}^{R}\sum_{n=0}^{i-R+l} M_l^{i-R+l} M_p^{j-R+p} G_n G_{n+p-l}\sigma_{i-R+l-n}^2 + \sum_{l=1}^{R}\sum_{\substack{p=1\\p<l}}^{R}\sum_{n=0}^{j-R+p} M_l^{i-R+l} M_p^{j-R+p} G_n G_{n+l-p}\sigma_{j-R+p-n}^2$$

Again, some results collected in [6] have been applied. The volatility is an specific case from the previous expressions (being $T_2 = T_4$ and $T_1 = T_3$):

$$\sigma_{F_{0,T_1\to T_2}} = \sqrt{\text{var}_0\left[\left(\ln\tilde{F}_{0,T_1\to T_2}\right)^2\right]} \approx \sqrt{\frac{\sum_{i=T_1}^{T_2}\sum_{j=T_1}^{T_2}\text{cov}\left[\ln\tilde{\varphi}_{i,T_1}\ \ln\tilde{\varphi}_{j,T_1}\right]}{\left(T_2-T_1+1\right)^2}} \tag{23}$$

The annualized option volatility is often quoted:

$$\sigma_{Option} = \frac{\sigma_{F_{0,T_1\to T_2}}}{\sqrt{T_1}} \tag{24}$$

IV. Numerical Results

In this section we discuss the spot price data collected from European Energy Exchange (EEX) where each day during the year, hourly power contracts are traded on the Day-Ahead Market. We use daily average spot prices from 16/06/2000 to

20/09/2002. These data are publicly available from the web page: www.eex.de .

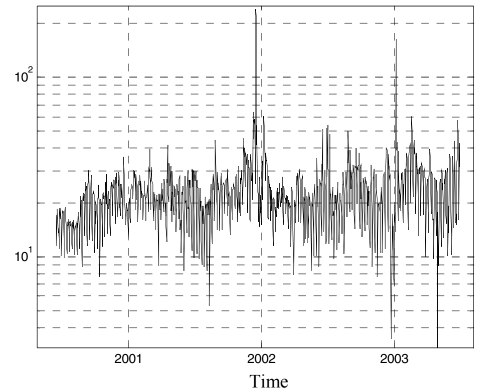

Fig. 1. Germany Average Daily Prices (€/MWh)

Concerning to jumps, there is a significant upward spike in mid-December 2001, when allegedly several agents exercised their market power and prices went up drastically [7]. In this work we have not taken into account spikes and we have eliminated them using a jumps filter, as described in [3].

Such as other electricity markets, the German Market shows strong seasonality.

Fig. 2 shows the deterministic component which reflects this seasonal effects. The deterministic component includes month and weekday terms.

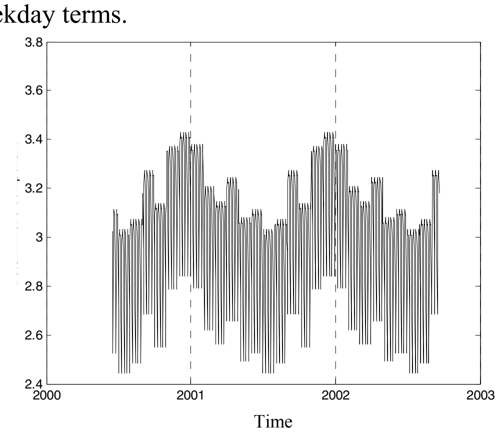

Fig. 2. Deterministic Component (€/MWh)

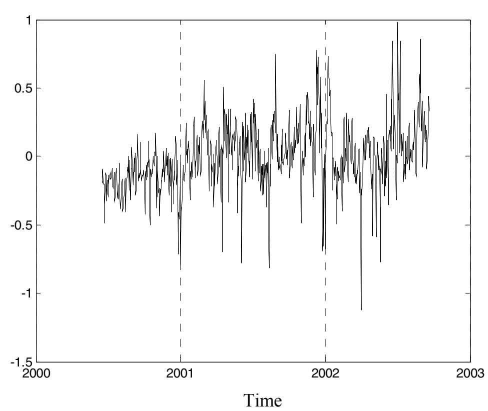

Fig. 3. Stochastic Component (€/MWh)

The stochastic component is modelled by fitting an AR structure incorporating, because of physical reasons, daily and weekly delays. Optimal order and parameters values, as given by the Aikine index and least squares fitting, are:

$$H_1 = 1 - 0.520q^{-1} - 0.060q^{-2} - 0.125q^{-3}$$
$$H_2 = 1 - 0.129q^{-7} \tag{25}$$

To check the fitting goodness, the autocorrelation graph (Fig. 4) leads to the conclusion that the spot prices and theoretic autocorrelation of the model exhibit similar patterns.

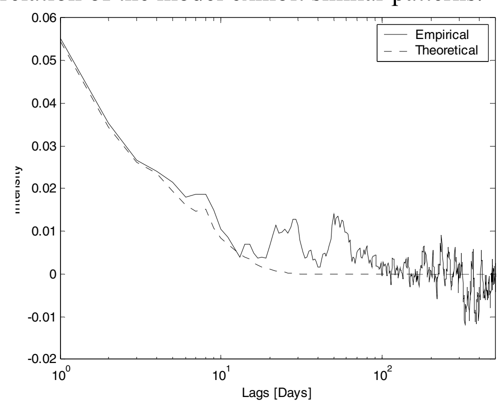

Fig. 4. Autocorrelation Graph

Residuals distribution (Fig. 5) displays fat tails and non normal behaviour, so further analysis should be investigated. Ljung-Box-Pierce Q-Test and Engle's ARCH Test, besides leptokurtic effect previously mentioned, show significant evidence in support of GARCH effects.

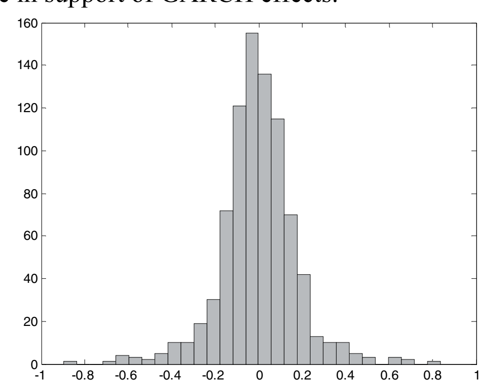

Fig. 5. Residuals Histogram

The application of the Aikake information criterium indicates that there is evidence in support of the GARCH(2,1) model, being the parameters values (as computed with the Matlab Financial Toolbox):

$$u_t = -0.009 + \varepsilon_t$$
$$\sigma_{\varepsilon,t}^2 = 0.004 + 0.386\sigma_{\varepsilon,t-1}^2 + 0.342\sigma_{\varepsilon,t-2}^2 + 0.157\varepsilon_{t-1}^2 \tag{26}$$

Fig. 6 shows the historical GARCH volatility evolution, as computed from the previous model.

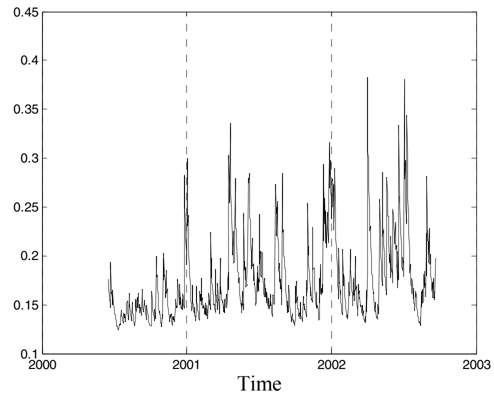

Fig. 6. GARCH historical volatility

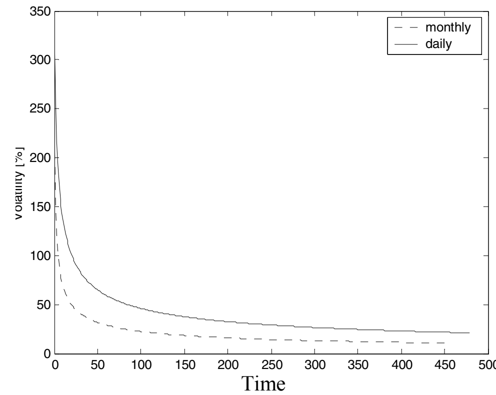

Fig. 9. Option Volatility for Daily and Monthly Futures (per cent annualised)

In order to forecast futures, an interest rate is required. This is directly collected from economical data. By applying previous formulae, the following forward curves are obtained:

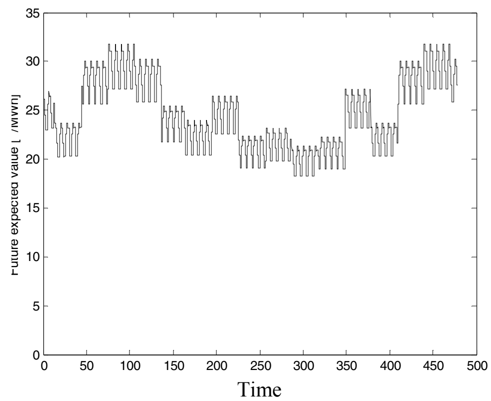

Fig. 7. Daily price forward curve (€/MWh)

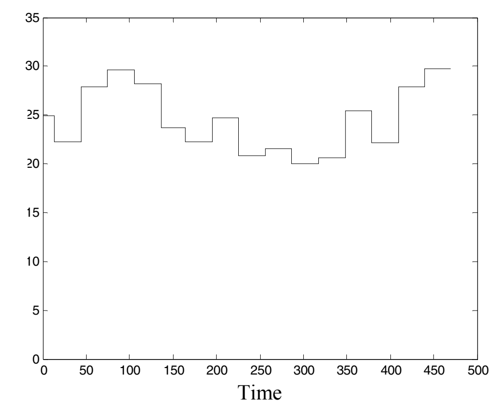

Fig. 8. Average monthly price forward curve (€/MWh)

Futures contracts with maturity dates on first days after valuation trading date, have high volatilities which are strongly linked to average volatility during the most recent period. Conversely, for longer maturities, volatility behaviour becomes more stable reducing uncertainty as contracts expire.

Useful contracts to reduce volatility are Monthly Futures whose fixed price is a monthly average price. The averaging allows variance to be lower on monthly contracts than on daily ones (Fig. 9).

V. Conclusions

A new model able to deal with electricity spot price series has been proposed. This model considers specific characteristics of electricity, such as seasonality, variable volatility and jumps. Way of computing forward and volatility curves (the basic analytic tools in market risk analysis and trading) from the model have been also stated.

The model and the curves computation procedures have been applied to real data: the European Energy Exchange daily spot prices. Volatility values from short and medium term maturities are reasonable, displaying reduced volatility for longer futures contracts (compare daily anf monthly futures curves in fig. 9), although this reduction is not so great as in the Black-Scholes model because of correlation among successive daily prices.

More arguable is the shape of the volatility curves for long maturities (longer than one year). Asymptotic volatility value is zero, as the price model is mean-reverting. Additional long-term factors, as in [8], could change the volatility curves in this region. However, their statistical estimation is not easy, because of relatively short and noisy time series. More research is needed.

Acknowledgment

The authors wish to thank Dr. Jaime Román for his help and kind remarks. They also wish to thank reviewer 2 his careful and critical reading and their helpful observations.

References

[1] Shijie Deng, "Financial Methods in Deregulated Electricity Markets", Ph.D. Thesis, University of California at Berkeley, 1999.

[2] D. Duffie, S. Gray, P. Hoang, "Volatility in energy prices", Managing energy price risk, Risk books, pp. 273-288, 1999.

[3] L. Clewlow and Ch. Strickland, "Energy derivatives. Pricing and risk management", Lacima publications, 2000.

[4] H. Lutkepohl, "Introduction to multiple time series analysis", Springer-Verlag, 1993.

[5] Carol Alexander, "Market Models", John Wiley and Sons, 2001.

[6] Marvin K. Simon, "Probability Distributions Involving Gaussian Random Variables", Kluwer Academic Publishers, 2002.

[7] Tobias Federico, "Out on its own", *Energy and Power Risk Management,* nr. 2, pp. 24-28.

[8] Eduardo Schwartz and James E. Smith, "Short-Term Variations and Long-Term Dynamics in Commodity Prices", *Management Science,* vol. 46, no. 7, pp. 893-911.