Lappeenranta University of Technology
Faculty of Technology
Department of Mathematics and Physics

# Is it possible to predict high volatility and price spikes in financial markets?

## An extensive statistical study of the wholesale electricity markets.

(research plan)

Lappeenranta, September 15, 2010

author: Matylda Jabłońska
supervisor: Ph.D. Tuomo Kauranne

**Tiivistelmä**

*Onko taloudellisen aikasarjan hintapiikkien tai korkean volatiliteetin ennustaminen mahdollista pörssin hintakehityksen matemaattisen mallinnuksen avulla?*
*Sähkömarkkinoiden hintakehityksen tilastollinen analyysi matemaattisen simuloinnin keinoin*

Sähkön spot-hinnat ovat eräs kaikkein volatiileimmista taloudellisista aikasarjoista. Niissä esiintyy voimakkaita hintapiikkejä. Tämän tutkimuksen tavoitteena on rakentaa yksi tai useampia matemaattisia malleja joiden avulla pyritään rekonstruoimaan spot-hinnan tilastollisia parametreja ja muita ominaisuuksia, joiden avulla voitaisiin ennustaa hinnan poikkeuksellisen suurta vaihtelua. Tässä tutkimussuunnitelmassa on tiivistetysti kuvattu tämän tutkimuksen menetelmät. Siinä tutkitaan myös kahta uutta menetelmää: Burgersin yhtälöä ja epälineaarisia stokastisia populaatiomalleja joiden avulla pyritään kuvaamaan markkinatoimijoiden mielentilan, Keynesin ns. "eläimellisten vaistojen", vaikutusta hintakehitykseen. Mikäli tällainen malli saadaan tilastollisesti käyttäytymään todellisen markkinan tavoin, voidaan sen avulla sekä tutkia vaihtoehtoisia tapoja säännöstellä markkinaa liian suuren volatiliteetin vähentämiseksi, kuten myös innovatiivisten markkinastrategioiden soveltuvuutta spot-markkinaan jossa esiintyy voimakkaita hintapiikkejä ja korkeaa volatiliteettia.

Matylda Jabłońska
Ruskonlahdenkatu 9 B 15
53850 Lappeenranta
matylda.jablonska [at] lut.fi

# Contents

# 1 Introduction

Electricity spot prices are the most volatile and most difficult time series, especially in terms of forecasting. Even though their long-term behaviour seems to follow some factor-dependent trends, various models still fail to predict the price microstructure. Here, I refer to very high and non-constant volatility, as well as resulting from it phenomenon – the so-called price spikes. They are the prices which within few hours can jump as much as tens times higher than the current spot price. These events are so far unpredictable. Even though in hindsight after the spike the power market specialists are almost always able to give specific reasons for these high price changes, the suggested factors are later never a hundred percent certain cause for next spikes.

What creates the base of my research is a hypothesis that we should stop looking at the electricity spot prices as classical time series, but start considering also human psychology driving traders' behaviour in financial markets. Those forces, known also as 'animal spirits', combined with the fact that electricity cannot be stored, may be the real reason standing behind the extreme values occurring in the electricity spot prices.

This research plan is an updated version of the original plan submitted with the postgraduate studies application, and has two main roles. The first one is to briefly present my whole work related to the mentioned topic, carried out or co-supervised within my postgraduate studies so far. From each research step I draw conclusions that influence the further approaches. The second goal is to present the implications of the results obtained so far, and suggest the main direction for continuation.

# 2 Background

## 2.1 Literature review

If one wants to investigate electricity price behaviour, it is helpful to work not only with price series themselves. Numerous studies, for example Ptak *et al.* [1] have shown that most popular heteroscedastic models are not sufficient to describe electricity spot prices. Traditional Box-Jenkins time series analysis predicts only some more general features, see *e.g.* Hadsell and Shawky [2]. Thus notice that in a regular trading market some non-zero values for certain lags in price return autocorrelations could indicate arbitrage opportunities for market participants. However, in case of electricity markets, the non-storability of electric power makes it very unlikely to benefit from such relationships, even with use of futures contracts.

Moreover, in equity markets one can notice that downward movements of prices are more often followed by higher volatilities than in case of upward movements of the same magnitude (see Campbell and Hentschel [3] and Nam *et al.* [4]). Thus is such asymmetry related to the electricity spot market or is it one more feature that differentiates regular stock trading from it?

Researchers have proposed innovative models based on supply and demand matching, which is what really drives electricity prices. As Hinz [5] states, predictability of demand plays a very significant role in understanding price behaviour. Thus time series models based on demand fluctuations may be helpful in defining day-ahead spot electricity prices.

On the other hand, demand still needs to be complemented by supply data to understand when the market faces constraints and perhaps produces higher than normal prices. Hadsell and Shawky [2] claim that transmission congestions are crucial for price behaviour. Hence, study of their importance against spike occurrence and significant volatility changes is important. We define a congestion as a situation when the flow of electricity on a transmission path equals the physical limit of that path, due to electrical resistance in transmission lines. It should be verified how the congestions appear when compared to differences between power bids and offers or generations within certain regions.

In recent years there appeared studies, *e.g.* Capasso and Morale [6, 7] and Morale *et al.* [8], of 2-level partial differential equations and stochastic differential equations systems for analysis of phenomena including particle interactions. These works are related to biological applications.

However, it may be interesting to use the proposed models to study possible *price herding*, meaning price behaviour leading to fat tailed distributions in financial markets. We propose to perform an empirical analysis of whether constraints or price herding is the principal cause of spikes. In case if constraints appear to be the explanation, regime specific models should be developed and calibrated, for example, using reversible jump MCMC algorithms. On the other hand, if price herding happens to be the reason for spikes, market dynamics could be described by Capasso and Morale type theoretical models, whose free parameters can be estimated with the use of data assimilation from empirical data that we have collected.

By basic decomposition of prices as performed in [9] we can see that magnitudes of spikes may differ a lot within different periods. There have also been studies carried out which show that magnitudes of spikes are not arbitrary at all. Moreover, Karakatsani and Bunn [10] and Kanamura and Ohashi [11] state that electricity prices work in two separate regimes: with and without generation/transmission constraints. Therefore, there arises a question whether it is possible to indicate the exact levels of factors which cause significant raise in volatility or price jumps.

There has also been an interesting proposal by Kanamura and Ohashi [12] of a structural model which is based on supply and demand curve characterization. An advantage of this proposal is that this model can describe and generate spikes in the series and it was proven to perform in this matter better that 3 other alternative models: Jump Diffusion model, Box-Cox model and extended Box-Cox model with seasonality.

Conejo [13] proposes wavelet transform decomposition of electricity prices. This may work out a combined wavelet and classical time series approach.

## 2.2 LUT past research in the field

In this section, I allow myself to skip wider reference to other researchers' works, because I cite them in my listed contributions. Here, I focus on the main steps and conclusions in the studies I performed myself, co-performed or co-supervised.

The first ideas of the aforementioned research dealt with combining classical time series theories with modern Markov Chain Monte Carlo approach. The classical information criteria used for ARMA-GARCH model choice were extended and then the parameters of a model picked by the methodology as optimal, were varied around the original values and 5000 simulations were run to create forecast predictive distribution. The results showed that the simulation of next steps was unable to capture the true price paths. For more details see [1].

As an alternative for classical ARMA-GARCH models we took a stochastic process of a mean-reverting character, i.e. Ornstein-Uhlenbeck process. Its main aim was to capture the fact that electricity spot prices, however far they deviate from current level, sooner or later come back to an overall mean level. The work was done by [14]. Later it was followed by [15], who compared performance of ARMA-GARCH models with Ornstein-Uhlenbeck-performed simulations. Conclusions from all works were similar. Firstly, residuals obtained from classical models were not normally distributed. Secondly, the stochastic process, even though well controlling the mean level, was spiking far too often and too low when compared to the original spot price returns.

Separately from building models describing the electricity spot prices I have worked on statistical analysis of price spikes (see [9]). The aim was to verify whether sudden price changes can be explained based on the year season, week day or price behaviour before the actual spike. The study showed that in New England, New Zealand, and in Nord Pool market, even though spikes were more common in particular year seasons (e.g. winter in Scandinavia), they were also frequently occurring in other times of the year. Also, there were no particular week days typical for spikes; for instance, in Norway there happened to be spikes on Sunday morning hours. The study of price volatility was of no help for spike prediction. Though we would expect to see some signals of higher price changes before the actual spikes, it did not appear to be a rule. However, the following were our observations:

- If a high spike occurs, the variance will jump and then gradually come back to its previous level.

- If a spike occurs after a long stable period, there will appear a similar (in magnitude) spike within 3-6 days after the first one.
- If variance rises twice or more, there will most probably either be a long period with persistently high variance or a spike will appear.

Under my supervision, a group of students worked on this problem (on a Student Mathematical Modelling Week) with a much more extensive data set (see [16]). The work aimed at explaining both spike occurrence and local price trends by the available background data for two electricity markets: New Zealand and Nord Pool. The built models were not ideal in terms of fit, but it was found to me mostly due to the long data horizon. Since there was potential for further work, it was continued by [17] who built an improved detrending and deseasonalising moving regression model, which was explaining significantly more of the price trends, though not spikes and high volatility. Moreover, for the New Zealand case [16], spike time instances were compared with data informing about transmission to the given price node reaching its physical limit. We expected to find that spikes occur particularly when the transmission limit is reached and there is a risk of congestion. This was not the case for New Zealand data.

A slightly different type of study did I do in Ireland, on a Mathematics with Industry Study Group. It also regarded electricity price stochastic modelling, but of a very distinct price series, so-called uplift (for more details see [18], [19]). Even though the data character was completely different, one of the approaches (jump-waiting) used for modelling them inspired me to apply a similar methodology on the regular spot prices. The main findings that helped to formulate a relevant model were [20], [21]:

- probability of spike occurrence depends on current price level
- spikes have a lot higher mean reversion rate than the around-mean process
- as studying spot hourly prices we took into account the trading characteristics – even though the prices are hourly, the trading is day ahead, and particular day hours are more likely to have higher prices than others, thus mean reversion is within 24h time intervals rather than hour to hour; whenever there is a spike in a particular hour, it is highly probable that there will be its influence visible in the respective hour on the next day.

We have considered Eq. (1) as an extension of Ornstein-Uhlenbeck and the other commonly used mean reverting jump diffusion models, to include a second drift of the spike regime.

$$dM_t = \gamma(t, M_t)dt + \sigma_t dW_t + J(t, M_t)dN_t \tag{1}$$

where, given a price threshold $M^*$ beyond which the prices are regarded spiky

$$\gamma(X) = \begin{cases} \alpha(X^* - X_t), & \forall_{X \le M^*} \\ \beta(Y^* - X_t), & \forall_{X > M^*} \end{cases} \tag{2}$$

The simulation results by general graphical as well as more detailed statistical comparison proved to resemble well the true data behavior. Both price and spike parameters did not differ significantly from these from counterparts. The small differences observed could be possibly still decreased by putting more emphasis on within-day price behavior, i.e. enriching the simulation with the probability structure of particular times of day that may be more spiky than others. Also the regular price path has a strong 24-hour periodic structure, which was not fully captured here.

Finally, we still kept verifying the specialists' hypothesis that the main reason of all spikes are congestions in transmission grid [22]. We studied three different ways to identify two regimes. One of the regimes was always the regular regime. The other one, generally called the non-regular regime, had three different names referring each to a different definition of non-regularity: the *spiky regime*, the *capacity-limited regime* and the *split regime*. Finland and Sweden were of interest for our study because a considerable amount of electricity which is used in these two countries comes from Norway. Then a part is transmitted within Sweden and part is forwarded to Finland. Thus in case of any congestions in the power grid between Sweden and Finland we expect the Finnish area price to rise. However, the correlations and co-occurrences of differently defined regimes remained statistically insignificant, even when properly normalized.

### 2.3 Conclusions influencing further objectives – current state of research

Based on the studies carried out so far, we drew the following conclusions:

- Background variables do not explain most of the volatility in prices, even less the spikes.
- ARMA-GARCH models do not capture markets (most likely because there is no coupling from price to volatility).
- Regular mean-reversion does not spike, or spikes far too often with lower magnitudes (a single mean reversion rate is not enough).
- Price distribution around mean level is asymmetric, but to a different direction du ring buyer's market and seller's market – coloured noise necessary.
- Volatility increases briefly after the spike, but otherwise spikes and high volatility are not significantly correlated.
- Spiking, two-regime pricing and transmission capacity saturation correlated only very weakly.
- Statistics of simulated and real prices can be very close but still visually a difference between the series is evident.
- The missing link – Human psychology; 'animal spirits' or a bidirectional nonlinear connection between price and volatility.
- **Better models needed!**

## 3 Objectives – research goal

The idea is to analyze the spot price behaviour based on daily data but then perhaps also move to a micro analysis of certain days by the half-hourly data. *The purpose of this study is to verify how much of spike characteristics (times of occurrence and magnitudes) and high price volatility can be explained by transmission grid binding constraints and which part of those comes rather from market dynamics.* My main hypothesis says that spikes can be only predicted in terms of occurrence probabilities, not as actual certain events to happen.

## 4 Research material and methods

### 4.1 Data

The research material consists of differently-scoped data from four different electricity markets: Nord Pool, New Zealand, Ireland and New England. The main focus is on the Nordic data set because it is most extensive. Also, Nord Pool was set up as the first international Power Exchange, and the way it functions is taken as example for new deregulated markets.

The available Nordic data set includes hourly data (covering period from 1 Jan 1999 to 31 Aug 2010) of the following variables:

- system and area spot prices
- area electricity production and consumption

- inter-area electricity transmission capacities and actual flows

Additional available background variables are temperatures in Scandinavian countries, hydrological power storage level and rainfall measurements.

The data from remaining three markets (New Zealand, Ireland and New England) were so far used for some of the specific mentioned past approaches. When the final methodology is built for Nord Pool, it can be implemented for other markets as well, respectively to their data availability.

## 4.2 Next research steps

The main focus of further research is to test the hypothesis that the true dynamics driving occurrence of spikes in the electricity spot prices are actually unpredictable on a day-to-day basis and their character is related to one of the following phenomena:
– turbulence in fluid dynamics
– animal spirits in human psychology.

To briefly motivate the approaches above, my recent simulations showed that a regular mean reverting jump diffusion model enriched with a power factor over the mean reversion part makes it possible to have the spikes reverted in a realistically short time. And the power which brings the simulation distribution closest to the real data appears to be $\alpha = 5/3$, same as the Kolmogorov constant known from turbulence theory. Also, one of the recent works on Burgers equation showed interesting simulation of fluid pressure measurements (in a single point) very closely resembling spot price realizations. Therefore, I intend to implement the Burgers equation with parameters estimated from the electricity spot price series. If we look at the philosophy behind, in the prices (and market, in general) we really can find correspondence to actual physical phenomena.

In particular, for the Burgers equation

$$u_t + \alpha u u_x + v u_{xx} = f(x,t)$$

we would have the following analogies:

- $u$ stands for the price,
- $f(x,t)$ describes the fundamentals (of a periodic character),
- $vu_{xx}$ is the diffusion term related to the fact that the spot market tends to reach the equilibrium price,
- $uu_x$ is the momentum term expressing traders' movement towards higher price,
- $u_x$ alone is the spread of bids for a given hour or day.

On the other hand, we know that so far there were no models built to account for human behaviour and psychology as one of the main factors influencing the spot price evolution. Whereas we know, that people, as all other species of animals, have 'animal spirits'. These, in financial market mean mostly fear and greed, influenced then by our common trading biases: herding, overconfidence and short-term thinking. Therefore, the second type of model I intend to implement and use for spot price simulation is Capasso-Morale system of stochastic differential equations (see [8]), used so far for modelling animal population dynamics. In my case, the population would be a group of traders in the spot market, and the measure of their distance would be the price. Traders do observe one another and thus create the general price path, which could be also understood as the global (in *macroscale*) population formation. However, there is a limit for overcrowding (in *microscale*) which in power trading could be interpreted as physical impossibility of two market participants to buy the same dose of electricity. Therefore, our suggestion for future work is to employ models proposed by [8] in mathematical biology. There, each individual price path simulated from the model proposed in this paper would represent a single trader, and the multiple simulation would provide coupling between the participants (in *mesoscale*). The movement of each particle might be driven by an external information coming from the environment, expressed via suitable potentials.

$$dX_N^k(t) = [\gamma_1 \nabla U(X_N^k(t)) + \gamma_2(\nabla(G - V_N) * X_N)(X_N^k(t))]dt$$
$$+ \sigma dW^k(t), \qquad k = 1, \ldots, N$$

# 5 Implementation, timetable and publications

My work timetable for the coming over a year time looks briefly as follows:

- Sep - Dec 2010 - close analysis of latest prominent spikes in Nordic prices (from winter 2010), i.e. extension for the currently available data laboratory,
- Jan - Mar 2011 - implementation and simulations of Burgers equation with parameters estimated from the electricity spot price series – publication,
- Mar - Jun 2011 - implementation and simulations of the Capasso-Morale stochastic differential equation system – publication,
- Jul - Aug 2011 - summary of the whole work, dissertation writing,
- Sep 2011 onwards - procedures towards final revision and defence.

# References


[1] P. Ptak, M. Jabłońska, D. Habimana, T. Kauranne, Reliability of ARMA and GARCH Models of Electricity Spot Market Prices., In: European Symposium on Time Series Prediction, Porvoo, Finland, September 17-19,, 2008.

[2] L. Hadsell, H. A. Shawky, Electricity Price Volatility and the Marginal Cost of Congestion: An Empirical Study of Peak Hours on the NYISO Market, 2001-2004., The Energy Journal, 2006.

[3] J. Y. Campbell, L. Hentschel, No News Is Good News: An Asymmetric Model of Changing Volatility in Stock Returns., Journal of Financial Economics, 1992.

[4] K. Nam, P. Chong Soo, K. Sei-Wan, Is Asymmetric Mean-reverting Pattern in Stock Returns Systematic? Evidence from Pacific-basin Markets in the Short-horizon., International Financial Markets, Institutions, and Money, 2003.

[5] J. Hinz, Modelling day-ahead electricity prices., Applied Mathematical Finance, 2003.

[6] V. Capasso, D. Morale, Stochastic modelling of tumour-induced angiogenesis., Journal of Mathematical Biology, 2009.

[7] V. Capasso, D. Morale, Asymptotic behavior of a system of stochastic particles subject to nonlocal interactions., University of Milan, Italy, 2004.

[8] D. Morale, V. Capasso, K. Oelschlager, An interacting particle system modelling aggregation behavior: from individuals to populations., Journal of Mathematical Biology, 2004.

[9] M. Jabłońska, Analysis of outliers in electricity spot prices with example of New England and New Zealand markets., Master's Thesis, Lappeenranta University of Technology, 2008.

[10] N. V. Karakatsani, D. W. Bunn, Intra-day and regime-switching dynamics in electricity price formation., Energy Economics, 2008.

[11] T. Kanamura, K. Ohashi, A structural model for electricity prices with spikes: Measurement of spike risk and optimal policies for hydropower plant operation., Energy Economics, 2007.

[12] T. Kanamura, K. Ohashi, On transition probabilities of regime switching in electricity prices., Energy Economics, 2008.

[13] A. J. Conejo, J. Contreras, R. Espínola, M. A. Plazas, Forecasting electricity prices for a day-ahead pool-based electric energy market., International Journal of Forecasting, 2005.

[14] E. Evarest, Princing of energy by mean of stochastic model., Master's Thesis, University of Dar es Salaam, 2008.

[15] M. Naeem, A comparison of electricity spot prices simulation using ARMA-GARCH and mean-reverting models., Master's Thesis, Lappeenranta University of Technology, 2010.

[16] H. Baya, C. Buchasia, M. Rykfors, P. d. Saint-Aubain, I. Vecchio, I. Wangwe, E. Yedeg, Influence of physical factors on electricity spot market price., Project work report, 23rd ECMI Modelling Week, Wrocław, Poland, 2009.

[17] B. N. Kirabo, Analysis of patterns in electricity spot market time series., Master's Thesis, Lappeenranta University of Technology, 2010.

[18] M. Carey, C. Houghton, M. Jabłońska, J. Kinsella, Bord Gáis. Uplift Problem., in: $70^{th}$ European Study Group with Industry, 2009, Limerick, Ireland.

[19] M. Jabłońska, A. Mayrhofer, J. Gleeson, Stochastic simulation of the Uplift process for the Irish Electricity Market., accepted for publication in Mathematics in Industry Case Studies Journal, 22 Jul 2010.

[20] H. Nampala, A stochastic mean-reverting jump-diffusion model with multiple mean reversion rates., Master's Thesis, University of Dar es Salaam, 2010.

[21] M. Jabłońska, H. Nampala, T. Kauranne, Multiple mean reversion jump diffusion model for Nordic electricity spot prices., under review for publication in Journal of Energy Markets, 2010.

[22] J.-P. Murara, Regime switching models for electricity time series that capture fat tailed distributions., Master's Thesis, Lappeenranta University of Technology, 2010.

[23] A. Mtunya, Modelling electricity spot price time series using coloured noise forces., Master's Thesis, University of Dar es Salaam, 2010.