Mathematics-in-Industry Case Studies Journal, Volume 2, pp. 86-110 (2010)

Stochastic simulation of the Uplift process for the Irish Electricity Market

Matylda Jabłońska * Arno Mayrhofer † James P. Gleeson ‡

Abstract. In the Irish electricity market participants declare their true marginal costs and therefore the Shadow Price alone does not guarantee that generators will recover their fixed running costs. The so-called uplift complements the price and ensures that the generators recover their total costs. The aim of this paper is to review purely stochastic features of the uplift and make an attempt to simulate a new process reconstructing the original data characteristics. We propose two alternative algorithms basing on the uplift wait-jump structure as well as daily and annual seasonality. Presented results show that this kind of reconstruction is possible up to a quantitatively comparable degree.

Keywords. Electricity Price, Irish All Island Market for Electricity, Uplift, Stochastic Simulation, Financial Time Series

1 Introduction

Electricity prices are as popular in research studies as any other financial time series. However, due to the main feature that differentiates electricity from other commodities, i.e. its non-storability, the electricity spot prices are one of the most challenging types of time series in terms of simulation and forecasting. Moreover, it has already been shown by many authors, e.g. [8], that their behaviour cannot be fully captured by classical time series models.

Methodologies for electricity price calculation may vary among different markets. For instance, in the Irish All Island Market for Electricity the System Marginal Price (SMP) calculated on a half-hourly basis with use of Market Scheduling and Pricing (MSP) Software consists of two components. The first one, *Shadow Price*, represents the marginal cost per 1 MW of power necessary to meet

*Corresponding author. Lappeenranta University of Technology, Department of Mathematics and Physics, Skinnarilankatu 34, 53850 Lappeenranta, Finland, tel. (+358) 40 53 13041, fax (+358) 5 621 2898, matylda.jablonska@lut.fi

†University of Vienna, Department of Mathematics, Dr.-Karl-Lueger-Ring 1, 1090 Vienna, Austria, arnom@amconception.de

‡Department of Mathematics and Statistics, University of Limerick, Ireland, james.gleeson@ul.ie

demand in a particular half-hour trading period. It is considered as within an unconstrained schedule, i.e. with no power transmission congestions. What complements the half-hourly SMP values is the so-called *Uplift*, which, added on top of the Shadow Price, makes sure that all the generators recover their total costs, including any expenses associated with start up and no-load costs.

The problem of models for uplift calculation has already been addressed in various studies, with example of [6], [5] and [2]. Thus the process values are obtained from software tools, which solve complicated optimization programmes with constraints based on knowledge of generation and demand. However, an interesting question emerges – whether the uplift process can be described and simulated as an individual stochastic process, with no background or constricting variables. This issue has been posed by Bord Gáis company at the 70^{th} European Study Group with Industry (hosted by Mathematics Applications Consortium for Science and Industry) and provided the basis for this work (see [4]). Even though there is no open market in uplift itself, having a stochastic model of the process is useful both to Bord Gáis and to other industry participants. Such a model can be used for VaR style analysis of the risk inherent in a book of power contracts, as well as being used to set prices for wholesale customers.

In this article we analyze the Irish uplift time series as a pure stochastic process, making note of its statistical features and proposing reasonable simulation approaches. The data set covers 451 days of half-hourly observations, which gives a decent background for a reliable statistical study. The aim of the simulations is to synthetically reconstruct a series which visually behaves similarly to the original uplift series and shows comparable statistical parameters (mean, standard deviation, skewness, kurtosis) and autocorrelation structure. The first simulation attempts are built on uplift features like jump waiting, jumps and zero-reversion. The other approach depends on probabilities for uplift price levels and constant plateaus for specific trading periods as well as seasonal components. All the proposed algorithms show that it is possible to reconstruct a non-negative process consisting of plateaus and jumps, but we also verify that the last proposed approach gives the best reconstruction of the uplift intra-day structure.

The paper is organized as follows: Section 3 presents a simulation approach and its results based on uplift jump waiting, jumps and zero-reversion features. Section 4 contains simulation algorithms based on uplift behaviour conditional on time of the day and presents respective results. Finally, Section 5 concludes and gives suggestions for future work.

2 Introduction to uplift calculation and study motivation

In the Irish Electricity Market, the System Marginal Price for each half-hourly trading period (SMP_h) consists of two components. The first one, Shadow Price (SP_h) representing the marginal price of electricity per MWh in each half-hourly trading period based on the information provided by the generators and the uplift (UP_h) represents the correction applied retrospectively to the shadow prices to ensure the fixed running costs recovery for all generators.

Every day the uplift process values are determined (see [9] for more details) by the Single Electricity Market Operator (SEMO) by solving a quadratic program that minimizes both uplift revenues (the Cost objective) and the Shadow Price distortion (the Profile objective).

$$\min_{UP_h,\, h=1,\dots,48} F(UP_h) \equiv \underbrace{\alpha \sum_h \left[(SP_h + UP_h) \sum_g Q_{gh} \right]}_{\text{Cost objective}} + \underbrace{\beta \sum_h UP_h^2}_{\text{Profile objective}}$$

subject to

$$\begin{aligned} \textstyle\sum_h [(SP_h + UP_h) Q_{gh}] \geq CR_g \quad & \text{for all } \; g = 1, 2, \dots, G \\ UP_h \geq 0 \quad & \text{for all } \; h = 1, 2, \dots, 48 \end{aligned}$$

where $F(UP_h)$ represents the uplift function, and Q_{gh} means quantity of electricity produced by generator g in half-hourly trading period h. Parameters α and β stand for importance of the uplift Cost objective and importance of the uplift Profile objective respectively. CR_g is the total cost of running for generator g, given by

$$CR_g = \sum_h [Q_{gh} C_u + NLC_g \mathbb{I}_{Q_{gh} > 0}] + ST_g$$

where C_u is the variable fuel cost per unit, NLC_g is the no-load cost of generator g representing the generator's expenditure when operating in stand-by mode and not producing electricity, and ST_g is the start-up cost of turning on a generator g that stays switched off as long as no production takes place. These costs will be considered constant for all half-hourly trading periods h for all days t. The first listed constraint ensures that each generator g recovers its costs CR_g and the second one certifies that all uplift values stay positive.

Also, there have been some alternatives of objective functions and constraints studied as can be found from [4]. Thus we can see that methodology of uplift calculation is well established. However, as typical for highly volatile electricity price markets, there appears a need for a more statistical analysis of the uplift process. The prices are determined on a day-to-day basis, whereas risk models require long-term view on the price behaviour. Companies tend to analyze risks and plan preferably for the whole year. But exact generation and consumption quantities can not be predicted for such a long time horizon. And this is where stochastic analysis comes to play a significant role.

As soon as one is able to investigate statistical features driving a given process it is possible, by using Monte Carlo simulations, to get more information on the distribution of the uplift prices. We assume that this knowledge would leave that general patterns of the process unchanged, and would rather give a better view on process behaviour and would support electricity companies in risk analysis (including probability of outstanding values/spikes occurrence) and contracting prices for the customers.

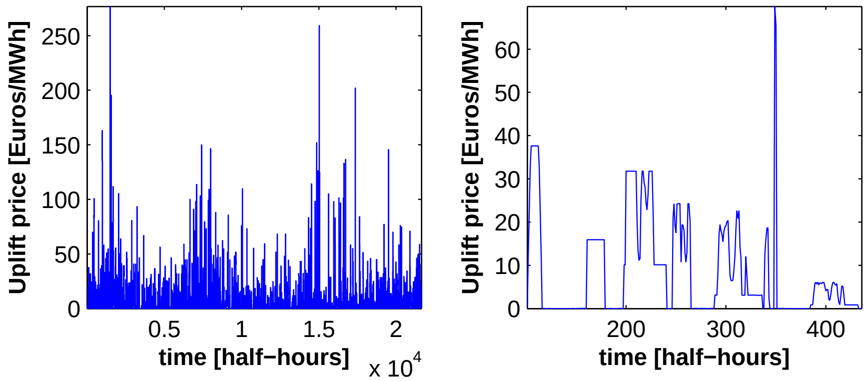

Figure 1: *Uplift half-hourly over 451 days (left panel) and 7 days (right panel).*

3 Uplift as a jump-waiting and jumping / zero-reverting process

3.1 Properties

The data set covers 451 days of half-hourly observations. Figure 1 (left panel) presents the original half-hourly uplift data series. The definition of uplift as a complement of shadow price states its first important feature, i.e. non-negativity. Moreover, from Figure 1 (right panel) we can see that the process has a clear step structure, i.e. there are plateaus and jumps. The presented analyzes will be performed on horizon equal precisely to one year.

Usually, when dealing with 'easy' and predictable financial or economical time series, we first think of a classical time series approach, i.e. using Autoregressive Moving Average (ARMA) models (see [3]), optionally extended by a Generalized Autoregressive Conditionally Heteroscedastic (GARCH) approach (see [1]). In case of uplift a first visual investigation tells us that ARMA models are not applicable here. However, we do use a piece of the classical theory. Since we expect the price series to be strongly seasonal (prices depend on demand which is seasonal), we use the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) to investigate the uplift periodicity. Figure 2 presents the respective results. As expected, we clearly see the humps in the ACF repeating with 48-lag regularity. Moreover, they are slightly locally maximal for every 336-th lag (48 half-hours times 7 week days), showing the weekly periodicity as well. This weekly periodicity is not as significant as the daily periodicity. To reproduce the 48 half-hour periodicity in the simulation we will base the current observation on the one that occurred 48 periods earlier, with a regression-estimated coefficient.

As we have already seen from the visual representation of the uplift data, we know that the process has a step (up and down) structure. Therefore, consistent with typical approaches (see [7]), we decide to use the Poisson distribution to model the plateaus sizes in the uplift series. Moreover, as can be seen from Figure 1, the jump waiting times are not distributed uniformly, but

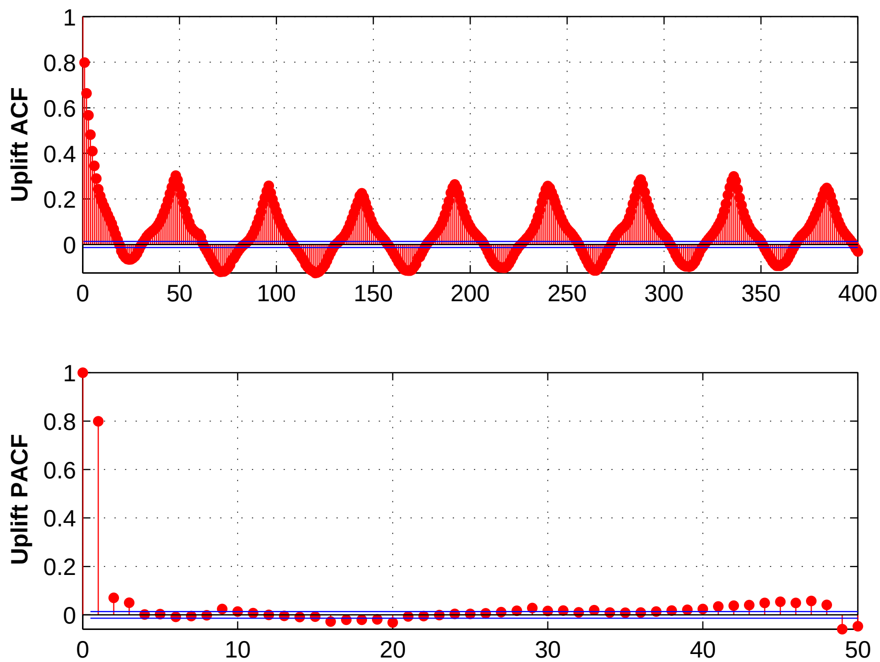

Figure 2: *ACF and PACF of uplift series.*

rather depend on the current price level. In particular, the zero level constant parts of the process are considerably longer, whereas when the process reaches relatively high values, it jumps down almost immediately. Therefore, we predetermined 4 different heuristic price levels for which we used different values of the Poisson parameter λ, estimated as an average length of plateau within the specified uplift level. They are as follows:

- $\lambda = 20$ for $U_h = 0$
- $\lambda = 8$ for $0 < U_h \leq 40$
- $\lambda = 1$ for $40 < U_h \leq 100$
- $\lambda = 0.05$ for $U_h > 100$

Also, we state that the probability whether the process jumps up or down at a given time point depends on the current price level, i.e. the higher the current uplift is, the more likely it is to jump down. Moreover, if the current price crosses a considerably high price level (also defined heuristically), the process will continue jumping down until it reaches zero. This tends to occur after a small number of half hours. Figure 3 presents normalized histograms for jumps up and down with respect to different price levels.

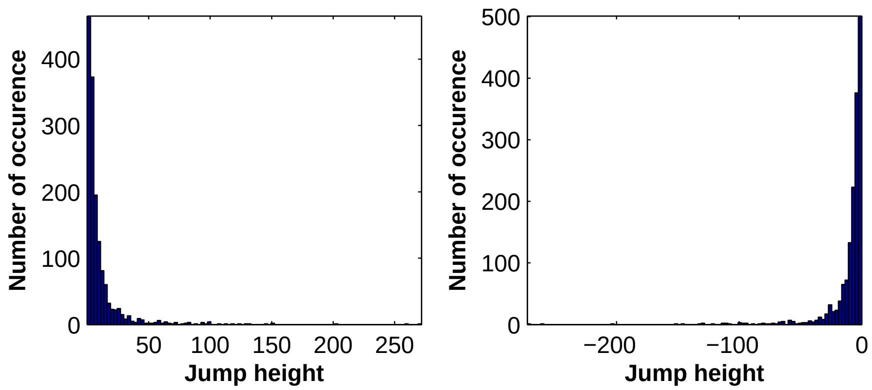

Figure 3: *Histograms for occurrences of jumps up (left panel) and down (right panel) depending on current uplift level.*

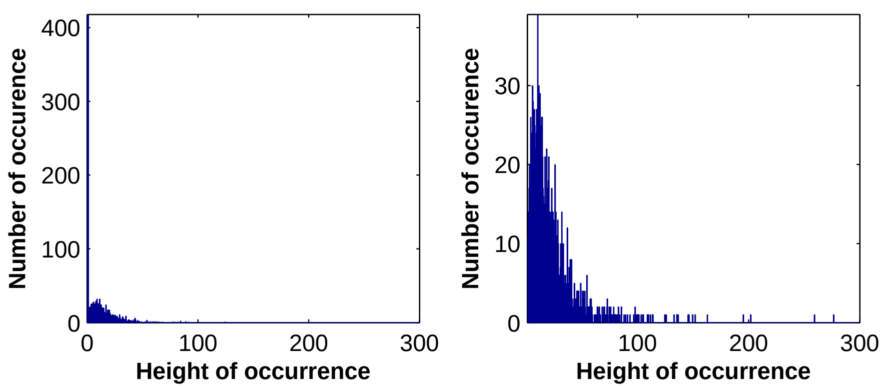

Figure 4: *Empirical histograms for uplift jump heights up (left panel) and down (right panel).*

Along with the up/down jump probabilities there comes a need for the study of jump heights. We identify the empirical unconditional distributions of jump magnitudes both up and down (see Figure 4) and use those later for jump height simulations.

As we can see from the plots, the histograms seem to be of exponential shape. For the purpose of sampling, we build empirical cumulative distribution functions and use those for random number generation, rather than estimate possible exponential distribution parameters.

We noticed that in the real data it is very likely that the prices are in most cases constant (usually zero) in the night hours. This characteristic is accounted for in the final process formulation as follows: after having the base process simulated, we set the night hours uplift to zero, with uniformly distributed probability of zero level being from 6 to 10 hours long.

3.2 Model and simulation algorithm

The review and discussion of Section 3.1 supplies insights about the structure and specific statistical features of the uplift price process. Using this information, we build a simulation algorithm implementing particular types of series behaviour as follows:

- non-negativity and zero-reversion
- strong 48 half-hourly periodicity
- constant price steps with length depending on price level
- jumps up and down with direction depending on price level
- probability of uplift staying constant or in particular zero being higher for night hours.

Based on the features we can formulate the uplift model as follows:

$$U_h = \begin{cases} J_u, & U_{h-1} = 0 \\ J_d, & U_{h-1} > M \\ \beta U_{h-48} + J_u \cdot v_u + J_d \cdot v_d, & \text{otherwise} \end{cases}$$

$$U_{h+1:h+w|_{U=u}} = U_h$$

where

- U_h and U_{h-1} is the uplift for trading period h and $h-1$ respectively
- J_u is a jump up with empirical distribution of upward jumps
- J_d is a jump down with empirical distribution of downward jumps
- M is a heuristic *large* uplift threshold
- β is the estimated regression coefficient for the 48-half-hourly periodicity
- U_{h-48} is the uplift for trading period $h-48$, i.e. the respective half-hourly trading period on the previous day
- v_u is a binary-distributed variable for jumps up
- v_d is a binary-distributed variable for jumps down, and $v_d = 1 - v_u$
- $U_{h+1:h+w}$ is the uplift for w consecutive trading periods
- $w|_{U=u}$ is the Poisson-distributed process waiting time conditional on the uplift level U.

This model strongly underlines the daily periodicity of uplift behaviour, as well as the fact that process waiting times are not the same for different price levels. Also the fact whether the next process move goes upwards or downwards is related to the current uplift status. This behaviour is expected to be related to particular daily electricity consumption patterns – highest in the peak morning and afternoon hours, and lowest at night.

Having identified the main features and defined the model we can build an algorithm for process simulation. The aim of this simulation is not to precisely reconstruct the real uplift series, but rather to synthetically produce a process that quantitatively behaves similarly to the original data in the long and short term horizon. We write down the simulation algorithm in a form of pseudo-code as follows:

1. compute regression coefficient for U_h and U_{h+48} dependence, where U_h is the uplift value in moment h
2. set the first 48 simulation values as the first 48 observations from the real uplift
3. initiate the Poisson parameters for 4 price thresholds
4. generate jump waiting time based on the current price level
5. set current time point as sum of last time point and the new generated waiting time
6. set the uplift values within the waiting time equal to the previous uplift value
7. then
    a. if the last price value after the last jump is higher than a predetermined threshold, force only jumps down until the uplift reaches level zero; if jumps down make uplift go below zero, set the last price to zero
    b. else, if the last price value after the last jump equals zero, force only jump up by a magnitude generated from empirical distribution
    c. else, based on the current price level sample whether the process jumps up or down and then sample the jump magnitude
    d. add the sampled jump magnitude to the price value that occurred 48 periods ago, multiplied by the earlier estimated regression coefficient
    e. move the current time point by one step ahead
8. return to point 4 while the current time point does not exceed the target process length to be simulated
9. set the night hours equal to zero with uniformly distributed probability of zero level being from 6 to 10 hours long.

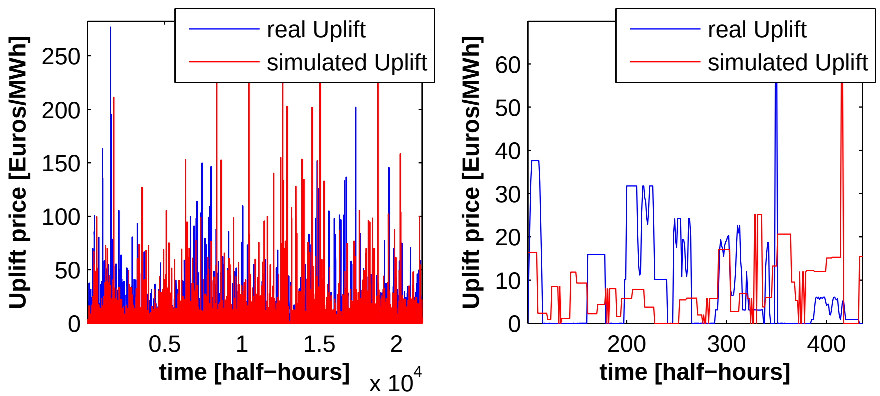

Figure 5: *Uplift half-hourly over 451 days (left panel) and 7 days (right panel).*

3.3 Results

The simulation was run using the algorithm presented in Section 3.2 for a sample as long as the real data set to get an insight of whether not only the short but also long term reconstruction gives any reasonable results. Figure 5 (left panel) presents the whole simulated realization which is quantitatively comparable with the original uplift series. Figure 5 (right panel) presents a slice of 7 days from the simulation, confirming that the general behaviour of the synthetic process is comparable with the original data.

We can see that our simulation is also able to produce values significantly standing out from the process mean, as it is for the real data. An additional aim of the simulation was to restore ACF and PACF structures similar to those of the real data, which is displayed in Figure 6. The autocorrelations of the original series was showing clear half-hourly periodicity (significant humps at every 48th lag) – we managed to reconstruct that feature up to a certain degree. The humps for the simulated series ACF are skewed. Also, there are no significant humps on the negative side, whereas this was the case for the real data. The PACFs of both original and synthetic data look comparable.

Moreover, we compare the real and simulated uplift distribution parameters to verify differences between (features like) mean value, standard deviation, skewness and kurtosis. Table 1 collects the mentioned figures for 5 different simulation runs against the original uplift parameters. We can see that the mean values of the generated series are comparable with the real data. So is skewness of ran simulations. Kurtosis results seem to present some controversy – even though simulation does not seem robust with respect to this parameter, the original kurtosis falls within the standard deviation (STD Sim) neighbourhood around the kurtosis' mean value of the five simulations (Mean Sim). Finally, standard deviation of the produced series remains regularly too low with respect to true uplift.

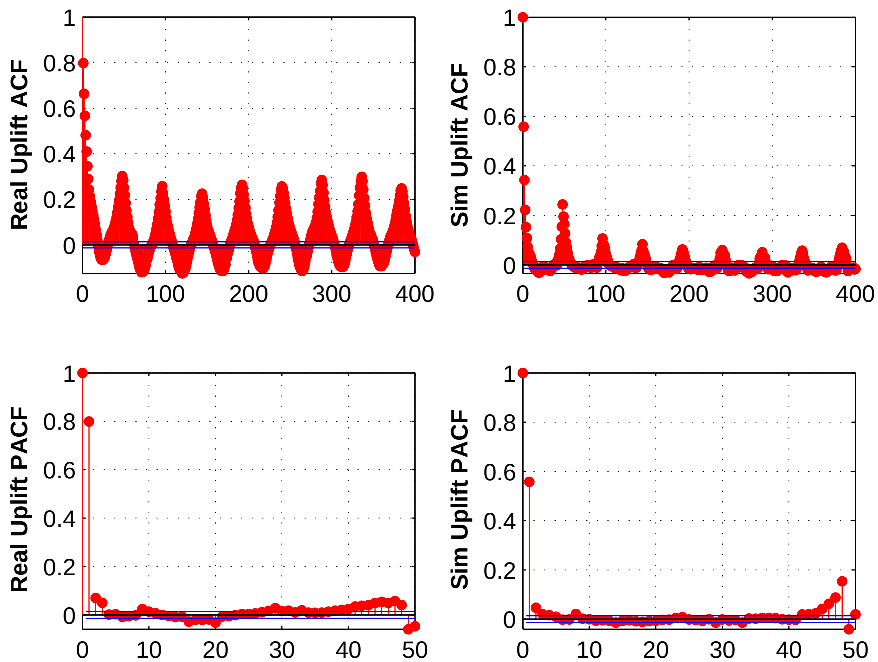

Figure 6: *ACF and PACF of the real and simulated uplift process.*

Table 1: *Real and simulated uplift distribution parameters.*

	mean	st. dev.	skewness	kurtosis
Real uplift	7.28	13.08	4.72	51.16
Simulation 1	7.36	10.35	4.60	56.17
Simulation 2	7.13	9.72	3.84	46.93
Simulation 3	7.01	10.01	4.65	61.33
Simulation 4	7.39	10.56	5.86	98.42
Simulation 5	6.91	10.05	4.45	49.05
Mean Sim	7.16	10.14	4.68	60.38
STD Sim	0.21	0.32	0.73	22.13

Finally, we use one more technique to verify statistical properties of the real and simulated series, that is we compare probabilities of exceedance for different level prices. In particular, we split the prices in slices by every 20 Euro for the range from 0 to 140 and verify what is the frequency of prices crossing the given thresholds. This gives a view on chances of uplift reaching particular elevations, including the highest spiky observations.

Figure 7 presents results for eight different price levels, from 0 to 140 Euro split by every 20 Euro. The probabilities are computed for 100 independent uplift simulations. We can see that for all levels except $thr = 0$ the simulated probabilities are about two to three times smaller than the observed ones. However, we do not notice that in the overall process mean estimates compared with the original parameter, as the simulation is on the other side more likely to give the price non-zero values.

We can see that the proposed algorithm reproduces quite a lot of the original data behaviour. Most important parameters fall into reasonable neighbourhood of the ones for real uplift. Also, the general quantitative look of the process for both long and short term horizon seems to resemble the real uplift structure up to a significant level. The simulated time series reproduces some of the 48-half-hourly periodicity as well.

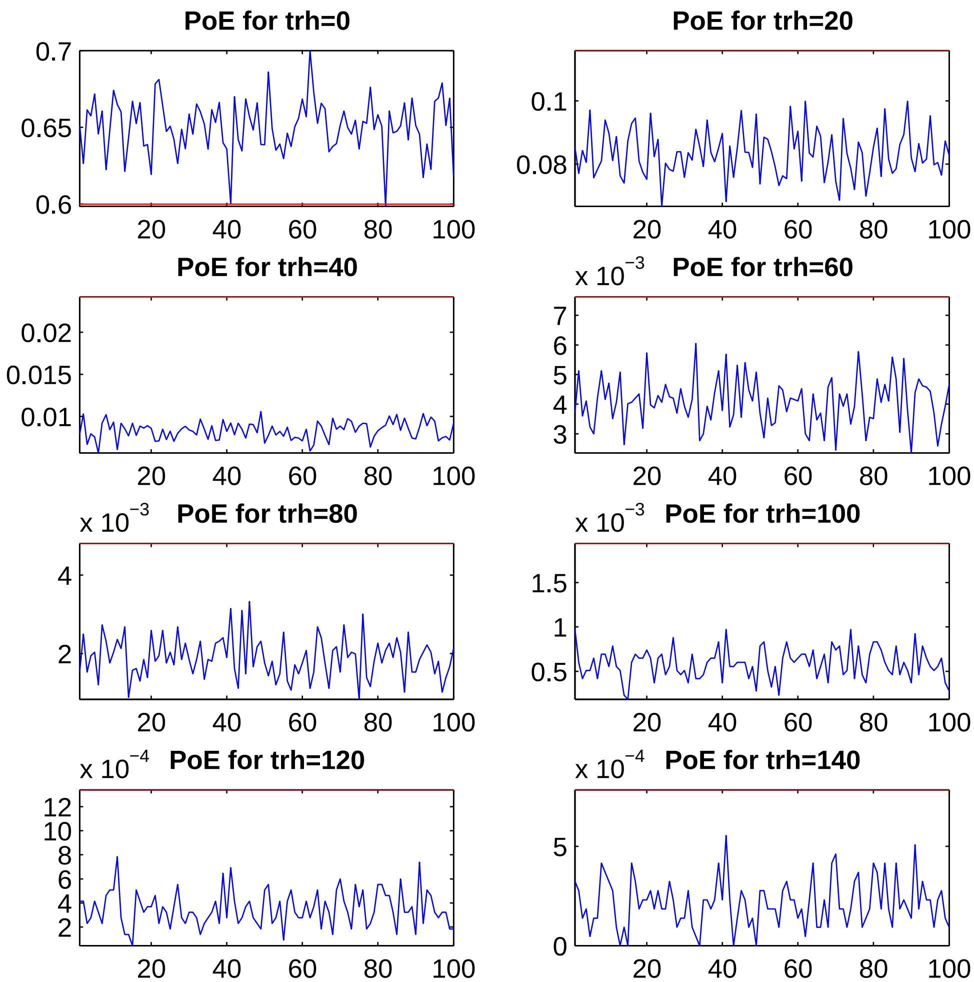

Figure 7: *Probabilities of uplift exceeding certain price levels.*

4 Uplift as a seasonal process depending on time of the day

The simulation discussed in the previous section has a few disadvantages like heuristic parameters for the plateau length distribution. Furthermore, as an examination of the plateau lengths shows they are not actually Poisson distributed but rather follow a more difficult distribution. Considering uplift only on a specific time of day was the initial idea for a new simulation with the characteristics described below.

4.1 Properties

For a suitable reproduction of the data it proved vital to split the problem into two parts:

1. uplift is constant
2. uplift price depends on time of day and month.

In the following sections these two properties shall be examined closely and described in a way so that they can be used for the simulation.

4.1.1 Constant uplift

Two observations can be made by looking at the data set when it comes to constant uplift. The first observation is the dependence of the constant parts on the current time of day. It is much more likely that uplift is constant during the night and the early morning (00:00-07:00). The second observation is, that uplift being constant is related to the current uplift, i.e. low uplift is more likely to be constant. These two observations will now be verified and will prove sufficient in describing the constant plateaus.

The probability P_{tod}, i.e. the probability of uplift being constant depending on the current time of day, can be seen in Figure 8. The uplift is reset each day at 06:00 and is given in half-hourly intervals. Thus the following figures with time of day dependence will always be from 06:00 to 05:30. A clear periodicity can be seen and during times where electricity demand changes most, less constant parts are observed. To include the reset at 06:00 a new uplift price will be calculated every time according to the properties discussed in the next section.

To verify the other observation, the probability P_{U}, i.e. the probability of constant uplift depending on the uplift, is displayed in Figure 9. Due to the small amount of data available only $U \in [0, 10]$ can be considered accurate and an exponential is used for fitting.

It can be argued that this might not be the best approach since P_{tod} and P_{U} are not actually independent. Indeed as will be shown later, this might be a problem that needs to be taken into consideration. However due to the small amount of data available no other approach could be found.

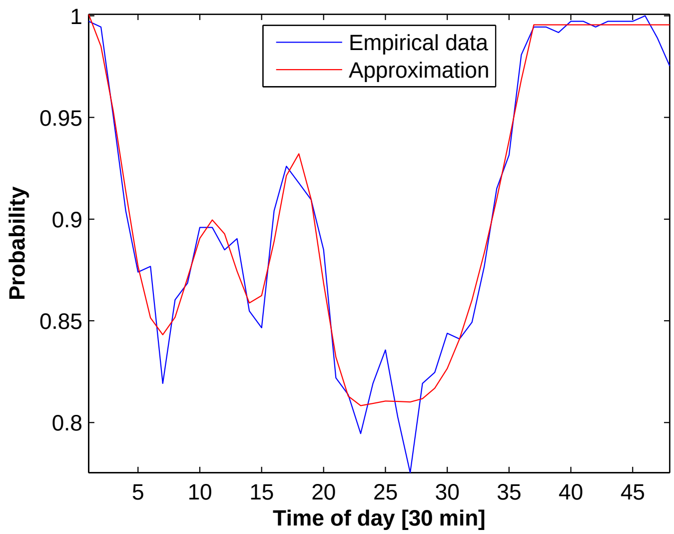

Figure 8: *Probability of constant uplift depending on time of day.*

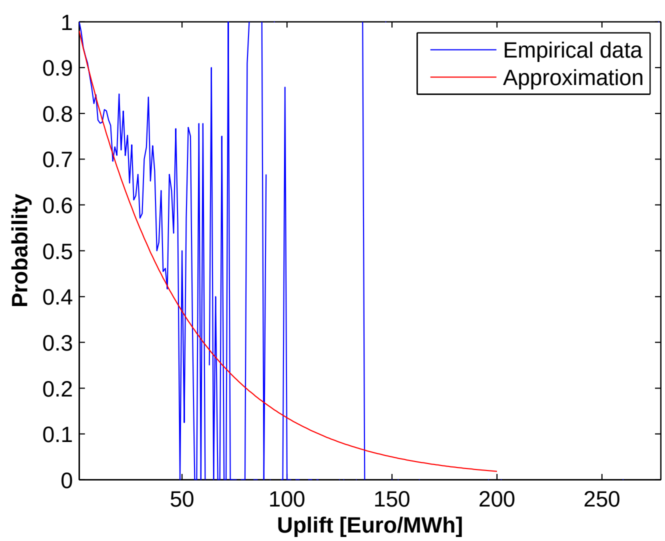

Figure 9: *Probability of constant uplift depending on uplift.*

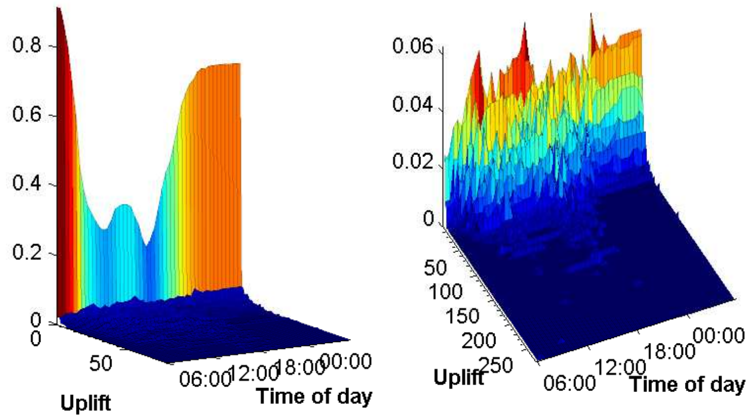

Figure 10: *Uplift histogram with entire data.*

In the simulation we will then use these probabilities to determine the probability of constant uplift as

$$P(U_i = U_{i-1}) = P_{\text{tod}}(i\%48) * P_{\text{U}}(U_{i-1}) \tag{1}$$

where % is the modulo operator.

4.1.2 Periodicities in uplift prices

At first an uplift histogram was considered separately for each time of day. The resulting graph can be seen in Figure 10. The plot on the left hand side shows the full histogram, emphasizing the probability of uplift being zero. The right panel neglects zero uplift to demonstrate daily periodicities of uplift greater than zero.

A first version of the simulation used this data, but the results showed a low mean (by a factor of about two). When considering a smoothed picture of the uplift, cf. Figure 11, annual seasonality shows up. Although only 1 year and 3 months of data was available it appeared necessary to include those changes. One large difference between summer and winter is the uplift around 12:00. This inhibited large uplifts around that time and thus reduced the mean of the simulation.

Due to that the same histogram as above was considered, this time for each month separately. It can be seen from Figure 12 (and in every other month) that there is a clear discontinuity between uplift being zero and non-zero, i.e. the most probable uplift value is zero, with much lower probabilities for non-zero values. Leading to the necessity of treating those two cases separately.

For $U > 0$ this resulted in approximately 30 data points for each histogram at a specific time of day. Figure 13 presents the one for May at 12:00. The bin size used for the histogram is 1. Even though the distribution can be suspected to be of Poisson type, due to a small number of data points it is not possible to have complete certainty and, therefore, the empirical distributions are

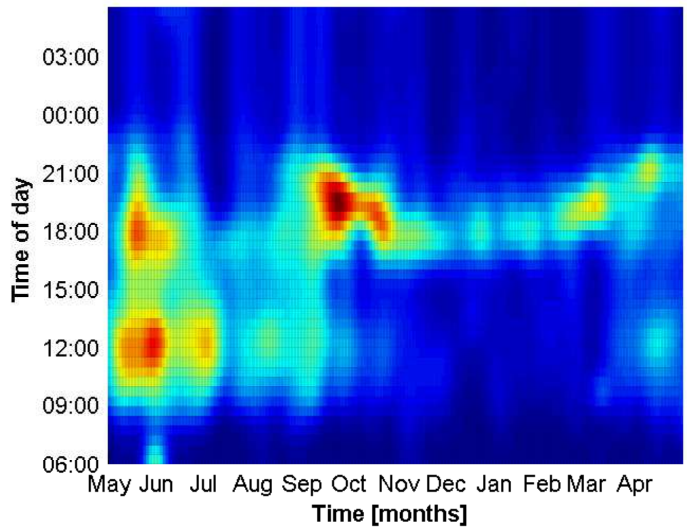

Figure 11: *Uplift (smoothed in horizontal direction).*

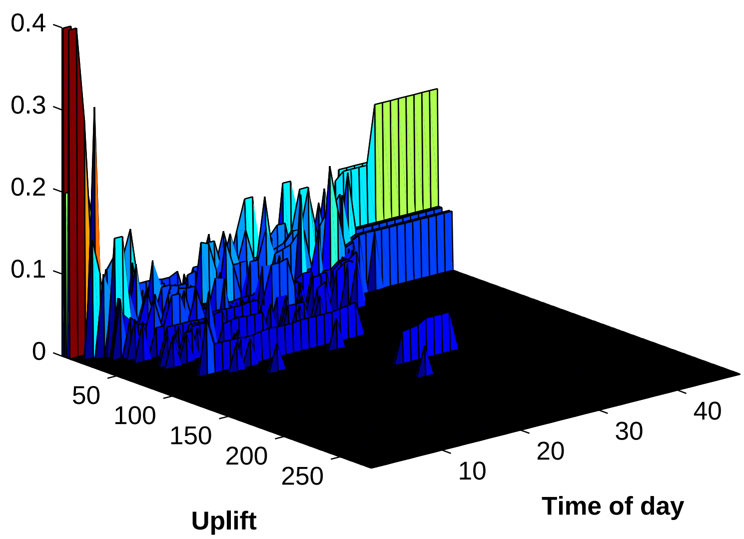

Figure 12: *Uplift histogram for May*

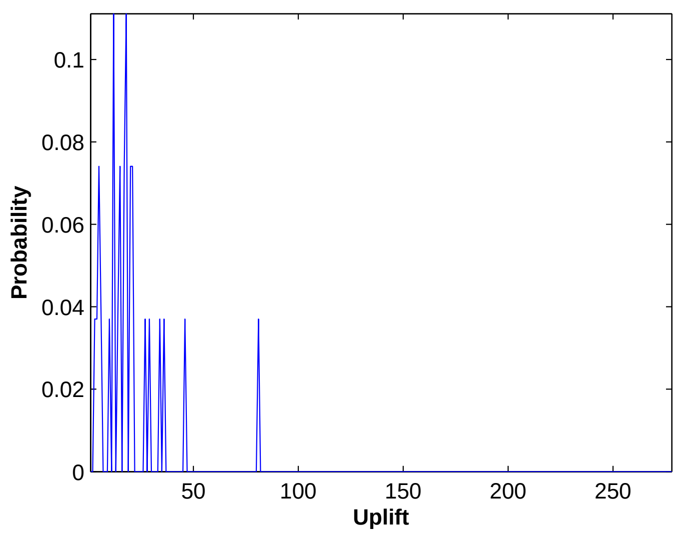

Figure 13: *Uplift histogram for May at 12:00.*

advised to be used for the simulation. Again a clear daily periodicity can be observed, as well as differences between distinct months.

The probability $P(U_i = 0)$ shows inverse proportionality to the mean of the $U > 0$ histogram and maybe it is possible to identify a correlation between those two parameters in the future. As noted in the previous section the discontinuity between 05:30 and 06:00 can be also seen from $P(U_i = 0)$.

4.2 Model and simulation algorithm

With the properties mentioned and characterized above it is now possible to formulate a new probability-based model for the uplift process

$$U_h = U_{h-1} \cdot v_1 + 0 \cdot v_2 + (\widetilde{U} + \beta U_{h-48}) \cdot (1 - v_1) \cdot (1 - v_2)$$

where

- U_h is the uplift for trading period h
- v_1 is a binary-distributed variable with success probability $P_{\text{tod}}(tod) * P_{\text{U}}(U_{i-1})$, where $P_{\text{tod}}(tod)$ is the probability of uplift being constant for given time of the day tod and $P_{\text{U}}(U_{i-1})$ is the probability of uplift being constant provided that the uplift on the previous trading period is equal U_{i-1}
- v_2 is a binary-distributed variable with success probability $P(U_i = 0)(m, tod)$ which stands for likelihood of uplift being zero provided that the trading period h falls into month m and time of the day tod

- $\widetilde{U}(m, tod)$ comes from empirical distribution of uplift values strictly greater than zero $U_h > 0$ for specific month m and time of the day tod

- β is the estimated regression coefficient for the 48-half-hourly periodicity.

Having the model formulated a new approach to the simulation can be constructed. The algorithm, implemented in Matlab, can be described as follows:

1. create regression coefficient reg for U_h and U_{h+48} dependence

2. use the first 48 half-hours of the real data for initiation

3. start the loop to create as much data as there is available from the real data

4. calculate current time tod and month m

5. if random number r_1 is smaller than $P_{\text{tod}}(tod) * P_{\text{U}}(U_{i-1})$ and $tod \neq 06:00$ set $U_i = U_{i-1}$

6. else if random number r_2 is smaller than $P(U_i = 0)(m, tod)$ set $U_i = 0$

7. else the $U > 0$ histogram is used for the specific tod and m to find a random variable $\widetilde{U}(m, tod)$ distributed accordingly (via the inverse cumulative distribution function), then $U_i = \widetilde{U} + \beta U_{i-48}$.

The model and respective algorithm give credit not only to the above mentioned daily (48-half-hourly) periodicity, but also to monthly seasonal patterns. It emphasizes the fact that electricity price behaviour in weather-dependent countries tends to have different behaviour in different months of the year.

4.3 Results

The first simulations were conducted without considering the monthly changes. As a result the values for mean, standard deviation and kurtosis were too low. However, it showed that the constant parts correlated very well with the original data, since it was possible to see the typical structures which showed up at low uplift. After improving the model by including the monthly dependence the performance improved drastically. In Figure 14 (left panel) simulated uplift for one year and three months can be seen. The seasonality and high spikes show up similar to those of the real data. The appearing structures can be further examined in Figure 14 (right panel) where data for a single week is extracted.

As mentioned in Section 3.3 another important feature of the uplift data is the characteristic ACF and PACF. For the new simulation these two graphs can be seen in Figure 15. Compared to the first model the negative parts of the ACF can now also be captured while the positive parts show similar good accordance.

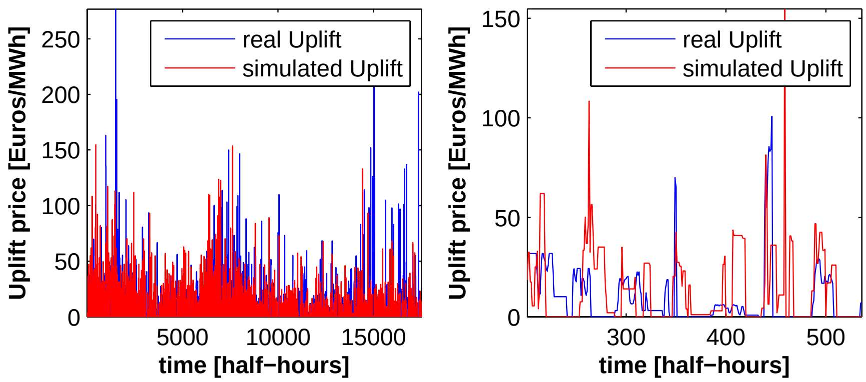

Figure 14: *Comparison of real and simulated uplift over (left panel) 451 days and (right panel) 7 days.*

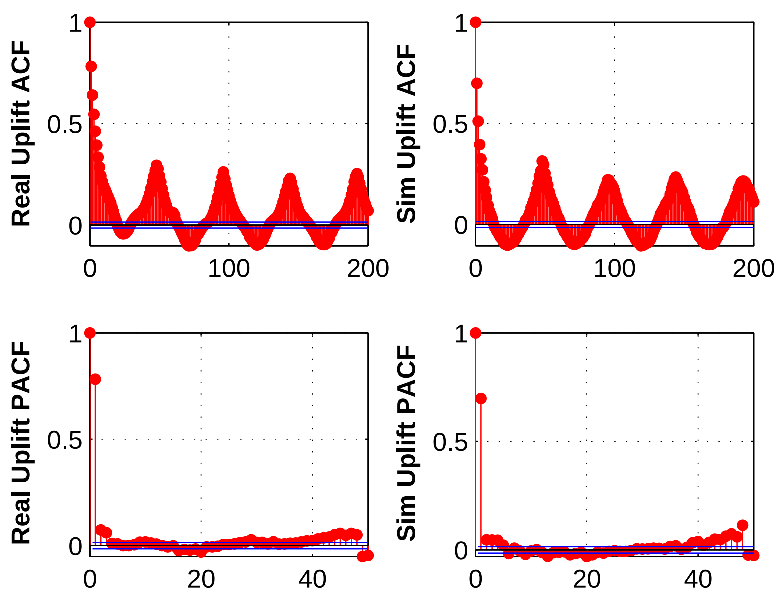

Figure 15: *Comparison of ACF and PACF for real and simulated uplift.*

Table 2: *Real and simulated uplift distribution parameters.*

	mean	st. dev.	skewness	kurtosis
Real uplift	7.28	13.08	4.73	51.16
Simulation 1	5.76	11.49	3.56	25.43
Simulation 2	5.61	11.74	3.94	35.05
Simulation 3	5.39	11.47	4.36	41.70
Simulation 4	5.50	11.37	3.72	27.69
Simulation 5	5.54	11.78	4.74	53.44
Mean Sim	5.36	11.57	4.06	33.66
STD Sim	0.34	0.18	0.48	11.36

Comparing the four characteristic values, mean, standard deviation, skewness and kurtosis (see Table 2), for the real data with five simulations it is possible to see that most values are close together. However, all the parameters are generally too low with respect to the true estimates. It can be expected that this issue becomes improved as soon as more data is available. As mentioned before, due to a short data horizon available, there were only 30 or 60 prices available for determining the distributions for each trading period within different months separately. As long as there is from 3 to five years of data available, we expect significant improvement for the second algorithm's results. Another possible explanation for the slightly lower mean is the constant uplift model. As noted earlier two probabilities P_{tod} and P_{U} are multiplied although they are not independent. This results in comparably shorter plateaus at middle and especially high uplift and since constant parts are followed by a jump, usually to a lower price, this lowers the mean as well as the standard deviation.

Table 2 also shows the strong fluctuations in kurtosis, due to the coarse distributions for uplift prices ($U > 0$) this is not surprising. Note that these fluctuations were also visible in the previous simulation (cf. Section 3.3).

Analog to the first proposed methodology, for these results we again verify the uplift probabilities of exceedance for specific price levels split by each 20 Euro, from zero up to 140 Euro, again collecting the outcome for 100 independent stochastic simulations. As Figure 16 shows, the second approach reduces probability differences with respect to the original ones for some price levels. On the other hand, the likelihood for level zero is now also lower than the true value and this fact is reflected in the significant difference between mean values of the original and simulated prices. Despite the differences between the true and simulated probabilities, we can see that all independent runs result in similar levels of likelihood, proving that the simulation algorithm brings robust outputs.

As mentioned in Section 3.1, the studies, i.e. all distributions and parameter analyses were run on exactly one year of data. Therefore, we use a part of the remaining observations to verify

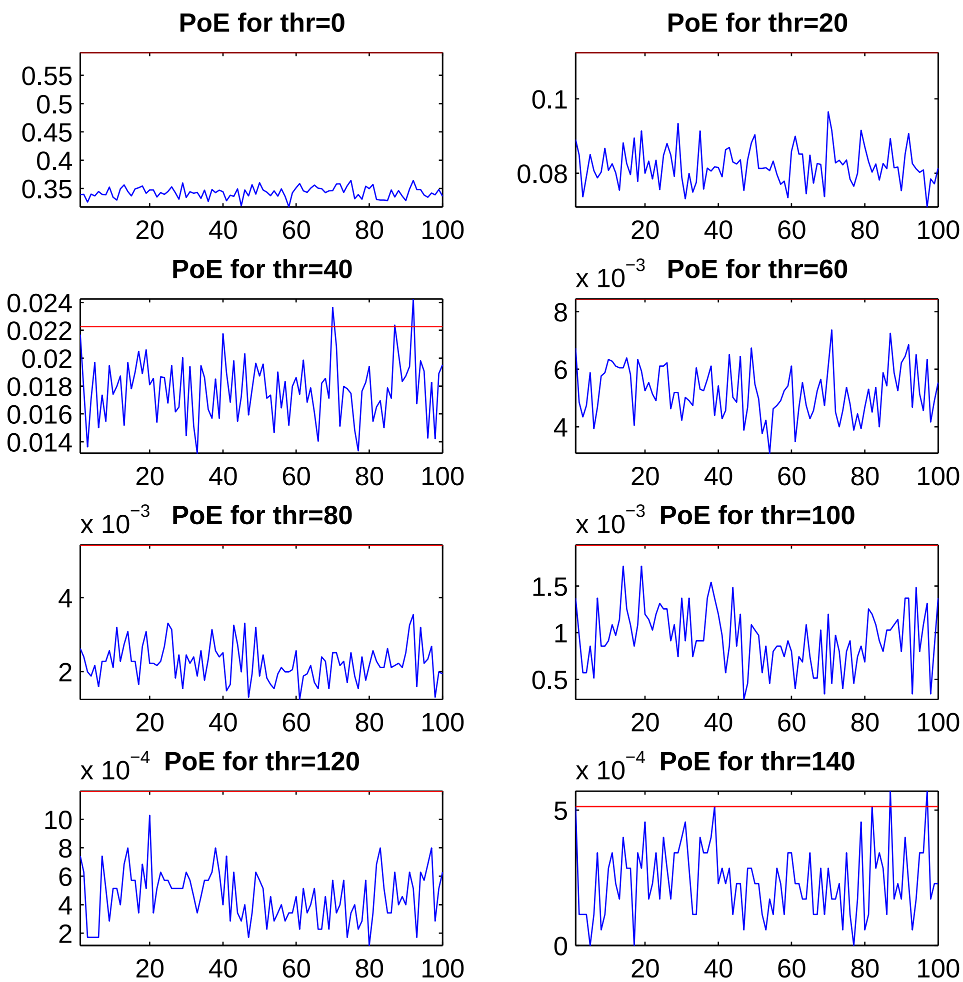

Figure 16: *Probabilities of uplift exceeding certain price levels.*

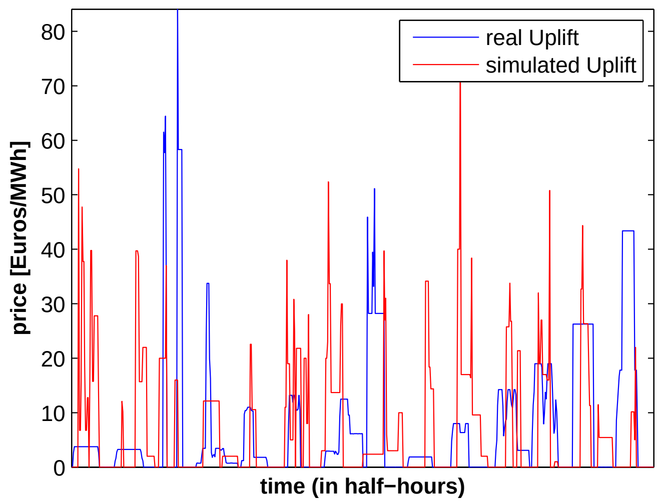

Figure 17: *Out-of-sample simulation for algorithm 2.*

the simulation performance on out-of-sample data. As the second algorithm was more convenient for reconstruction of intra-day specific uplift behaviour, we pick this one for the final comparison. Figure 17 presents simulation of the two weeks following the one-year part of data used for estimation.

Clearly, the general process intra-day behaviour is similar to the true path. Moreover, the simulation is able to reproduce both high and close to zero values, analogical to the real uplift characteristics. Table 3 collects basic statistics for the real and five times independently simulated out-of-sample data. The differences are acceptably low. Note that for all parameters the original values fall into the Mean Sim ± STD Sim intervals.

Table 3: *Real and simulated uplift distribution parameters on out-of-sample 2-week horizon.*

	mean	st. dev.	skewness	kurtosis
Real uplift	6.88	11.80	2.59	10.74
Simulation 1	6.72	10.93	2.04	8.01
Simulation 2	7.63	13.18	1.98	7.14
Simulation 3	4.70	11.29	3.80	23.62
Simulation 4	8.35	16.76	3.34	19.68
Simulation 5	8.02	16.02	1.31	4.53
Mean Sim	7.08	13.64	2.49	12.60
STD Sim	1.47	2.67	1.04	8.48

5 Conclusion and suggestions for future work

The aim of this study was to reconstruct behaviour of the uplift process coming from the Irish All Island Market for Electricity based only on the process itself. For this purpose we reviewed different statistical features of the process and proposed two alternative simulation algorithms. The suggested methodologies were able to reconstruct the real uplift behaviour up to different degree levels.

The first approach was based on finding respective distributions for process jump waiting times as well as jump direction and sizes, dependent on the given price level. The method did manage to produce a plateau-step structure similar to the original uplift path. It also returned main distribution parameters quantitatively comparable with the real ones, except for kurtosis. Also standard deviation seem to regularly differ from the original one by approximately 30%. Moreover, this approach failed to sufficiently reconstruct the autocorrelation and partial autocorrelation seasonal structure.

The aim of the second simulation was to eliminate certain heuristic parameters but eventually a different approach could be found by utilising other uplift characteristics. It proved of importance to include annual seasonality as well as the dependence on the current time of day. Although empirical distributions are used for most of the simulation it can be suspected that it might be possible to find analytical expressions as soon as more data is available. The simulation provided promising results which were approximately 25% too low. Two possible explanations for this phenomena have been given. Amongst them is the mathematically incorrect treatment of dependent probabilities. With two or three years more data it might be possible to also find annual seasonality in P_{tod} which can be exploited to eliminate the P_{U} term.

Finally, we verified the results of the proposed algorithms by comparing the uplift probabilities of exceeding certain, evenly distributed price levels. Even though we could notice differences between the desired and simulated likelihood outcomes, the results were robust over a number of independent simulations. Moreover, as each simulation brings independent uplift paths, the methodology could be employed within Monte Carlo framework, where a combined simulation would result in a broader view on the probability distributions describing the price behaviour. Doing this is recommended as soon as more data (at least 2 years) is available, and the seasonal probability components can be estimated more reliably.

This article studied an electricity price series different from those most commonly known (like in Scandinavian or New Zealand markets). It made it more challenging, since popular ARMA-GARCH or mean-reverting jump diffusion models could not be used. Nevertheless, we did obtain very good results with still much room for future improvement.

Acknowledgements

The authors wish to acknowledge the support of Science Foundation Ireland though awards 06/MI/005 (MACSI) and 06/IN.1/I366. The support of Gavin Hurley and Sile Bourke is acknowledged, for their specialist insight to the study. The authors would also like to express their gratitude for the two anonymous reviewers and MICS Journal editor, professor Matt Davison, whose invaluable comments helped understand and eliminate weaknesses of the initial submission.

References

[1] T. Bollerslev, *Generalized autoregressive conditional heteroscedasticity*, Journal of Econometrics **31** (1986), 307–327. 89

[2] F. Bouffard and F.D. Galiana, *Generalized Uplifts in Pool-Based Electricity Markets*, Analysis, Control and Optimization of Complex Dynamic Systems, Springer US, December 5, 2005. 87

[3] G. Box, G. Jenkins, and G. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed., Prentice Hall, Engewood Cliffs, NJ, 1994. 89

[4] M. Carey, C. Houghton, M. Jabłońska, and J. Kinsella, *Bord Gáis. Uplift Problem.*, 70th European Study Group with Industry, September 3, 2009, Limerick, Ireland. 87, 88

[5] P.R. Gribik, W.W. Hogan, and S.L. Popeii, *Market-Clearing Electricity Prices and Energy Uplift*, December 31, 2007. 87

[6] W.W. Hogan and B.R. Ring, *On Minimum-Uplift Pricing for Electricity Markets*, March 19, 2003. 87

[7] B.J.T. Morgan, *Applied Stochastic Modelling*, 2nd ed., CRC Press, Tylor & Francis Group, 2009. 89

[8] P. Ptak, M. Jabłońska, D. Habimana, and 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. 86

[9] SEMO, *SEM Trading and Settlement Code: Appendices, v.6.1,* A*ppendix* N*: Operation of the* MSP *software.*, January 29, 2010. 88