Multidimensional analysis of New Zealand electricity prices

Matylda Jabłońska-Sabuka[a], Agnieszka Wyłomańska[b]

[a]*Lappeenranta University of Technology, Department of Mathematics and Physics, Skinnarilankatu 34, 53850 Lappeenranta, Finland, matylda.jablonska-sabuka[at]lut.fi*
[b]*Wrocław University of Technology, Institute of Mathematics and Computer Science, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, agnieszka.wylomanska[at]pwr.wroc.pl*

Abstract

Modeling electricity prices after market deregulation has become notoriously difficult and yet more important due to the increase of price volatility. The aim of this work is to propose two alternative multivariate autoregressive models with α-stable noise for modeling New Zealand electricity market prices. The models account for nodal price interrelations as well as price dependency on empirical factors. Moreover, a novel extension of classical approaches is provided by incorporating non-Gaussian noise structure to reproduce price spikes. The results are robust and show high accuracy in day-ahead forecasts and provide market participants with a sound basis for risk assessment.

Key words: electricity price, day-ahead forecast, multivariate autoregression, VAR, non-Gaussian noise

1. Introduction

Electricity prices have become some of the most challenging time series to be modeled mathematically. This is due to electricity market deregulation, which has by now been carried out in many countries across the world. Its aim was to allow fair competition in the market and thus open it to more participants. Indeed, the goal of reducing average price level through competition has been achieved, however, electricity prices have also become more volatile and less predictable than ever before.

An important factor which contributes to high volatility in most of the markets is the large variations in demand and supply of electricity, both of which are very uncertain in deregulated markets (Burger et al., 2004). For instance, temperature strongly affects the demand; in total, the demand varies between 50-100%. Thus, as some say, forecasting demand is almost equivalent to forecasting weather (Podraza, 2006). Next to any climatic factors, hydrological balance, demand and base load supply (Vehviläinen and Pyykkönen, 2005) can be considered with equal importance as the key spot price drivers.

Electricity prices can be studied both in their inter- and intraday format (hourly or half-hourly, depending on the market). A lot of research efforts have so far focused on investigating spot price interdependencies. For instance, the New Zealand spot prices can be divided into five intraday groups: overnight off-peak, morning peak, day-time off-peak, evening peak, and evening off-peak. Then it appears that prices within these groups are a lot more correlated than between these groups along different trading periods (Guthriea and Videbeck, 2007). Another work analyzed a group of models classified as Markov regime-switching (MRS) (Janczura and Weron, 2010). There, the focus was on the performance of different models in terms of statistical goodness-of-fit and the results showed that the best one was an independent spike 3-regime model with time-varying transition probabilities, heteroscedastic diffusion-type base regime dynamics and shifted spike regime distributions.

Thorough understanding of intraday price structure makes day-ahead forecasting these days possible up to a significant level of accuracy. However, also long-term predictions are equally important for risk management. It is known that electricity price trend and cyclic structure comes in big part from the influence of some well recognized factors. Among those one should list hydrological storage (for heavily hydro-dependent markets, like Nord Pool ad New Zealand) and thus rainfall, weather and demand (Vucetic et al., 2001; Ruibal and Mazumdar, 2008; Laitinen et al., 2000; Jabłońska et al., 2011). A stochastic model using a number of explanatory variables has been proposed, for instance, for the California electricity market (Kian and Keyhani, 2001), among others. For the case of New Zealand prices, there is lack of modeling approaches that would be able to thoroughly explain behavior of prices.

Only some works have just identified that the nature of prices has changed after market deregulation in various ways (Ying and Flynn, 2003). Some other researches have used discrete wavelet transforms to investigate the demand-price relationship (Qureshi et al., 2009) and market volatility in New Zealand with the use of GARCH models (Situ and Nair, 2007).

The main contribution of this study is twofold. Firstly, it exploits the significant interdependencies between the prices in different trading nodes, as well as external factors influencing them, like hydro storage, rainfall, air temperature and demand. This is done through multivariate autoregressive models which provide a lot of advantage over a classical regression analysis. Secondly, both proposed models consider non-Gaussian noise with α-stable structure, which allows to capture statistically the occurrence of price spikes. This is a novelty when compared to the classical ARMA-GARCH methodology. Results show that both models provide accurate day-ahead forecasts for the period of a year ahead. Therefore, we demonstrate that a novel combination of VAR models with non-Gaussian noise provides a powerful tool for analysis and day-ahead forecasting of electricity spot prices, including a new way of coping with simulation of price spikes.

This article is structured as follows. Section 2 describes the data set used in the study together with the main statistical features of the data. In Section 3, the multivariate autoregressive model with non-Gaussian structure is introduced. Section 4 applies the model to the real data and collects fit and forecast results. Section 5 concludes the results.

2. Main features of New Zealand electricity prices

2.1. New Zealand electricity market

New Zealand electricity market is a very interesting case study for spot prices analysis from many points of view. Its electricity sector is principally (70%) based on renewable energy sources such as hydropower, geothermal power and a steadily increasing wind energy, which makes New Zealand one of the most sustainable countries in the world when it comes to energy generation. On the other hand, its electricity demand has been significantly growing in the past by an average of 2.4% per year since 1974 and by 1.7% over 1997 – 2007.

New Zealand is characterized by a geographically unbalanced demand-supply relation. The highest electric power production takes place on the South Island whereas the highest demand comes from the more populated and industrialized North Island. Moreover, the electricity market in New Zealand is not pooled. The main participants are seven generators/retailers who trade at over 200 nodes across the transmission grid.

New Zealand market data is available free of charge from the Electricity Authority (former Electricity Commission) in the form of Centralized Dataset. The data consists of half-hourly information on nodal prices, bids and offers, meter data and binding constraints, as well as additional daily hydrology and network configuration data. The authors utilized price data from a number of nodes, as well as the centralized demand and hydrological storage values. Moreover, additional data on rainfall and temperatures in each of the considered nodes was used.

The analysis presented in this paper considers 11 trading nodes spread across both South and North Island, as presented in Fig. 1. Six of them are located near powerplants. In particular, Benmore, Tuai and Whakamaru that represent the hydro power generation, while Huntly, Otahuhu and Stratford are based on geothermal generation. The remaining 5 nodes are only splitting substations. The data covers a period from 1st of January, 1999 to 31st of July, 2009. Data are quoted daily.

The choice of the trading nodes was deliberate to properly represent the most important power generators as well as the key splitting substations. For instance, Haywards is the key node splitting power delivered through the High Voltage Direct Current (HVDC) connection between the South Island (originating in Benmore) and North Island. Numerous times, price spikes can occur in that node due to electricity transmission constraints. Also, fair balance between the South and North Island was needed as, for instance, prices in the South used to be on average lower than in the North until an additional fee was imposed on the generators for using the HVDC connector to transmit power.

Let us note that the 11 nodes create just a small representative subset of the total of over 200 nodes across the country. Theoretically, it is of course possible to work with all the grid nodes in one model. However, this would, in our opinion, make the system over-represented and, certainly, computationally heavy when it comes to parameter estimation (there would be many more parameters to be estimated that data points available).

Figure 1: Location of the 11 nodes in the New Zealand grid used in the analysis.

2.2. Relationship between respective price data corresponding to analyzed nodes

The first approach proposed in this study is based on relationships between price data for the 11 considered nodes. The analyzed data sets, which correspond to 11 vectors of electricity prices in chosen nodes in New Zealand during examined period, exhibit very similar behavior. In Fig. 2, we present the analyzed prices. Moreover the data sets have also similar statistical properties. The first property that we observe here is the non-stationarity of each vector. This non-stationarity can be easily proven by observing the autocorrelation functions (ACFs) of appropriate data sets. Moreover, in the ACFs we can also see the seasonal behavior for all of considered vectors of observations. The corresponding ACFs are presented in Fig. 3.

The second statistical property that is observable for all considered data sets is the weekly seasonality. This kind of seasonality can be seen in ACFs of differenced series. In Fig. 4 we present the autocorrelation functions for 11 electricity price vectors, where the weekly seasonality is easily observable for lags that are multiples of seven.

The last property that should be emphasized here is a strong relation between the analyzed data sets. The correlation coefficients are large and the test for its significance indicates at non-zero correlation between analyzed prices. That is, all the p-values (for testing the hypothesis of no correlation against the alternative that there is a non-zero correlation) are very close to zero. In particular, all correlation coefficients among the 11 price series themselves exceed 0.9. High correlations could raise a multicollinearity concern, however, this is problematic with classical regression models, whereas it has not been proven such in the case of autoregressive models. The idea of multivariate autoregressive models used in this study is that all the variables are equally important and we do not distinguish between dependent and independent variables. Therefore, there is no need to remove any of the highly correlated components.

2.3. Relationship between price data and deterministic factors

The second proposed approach is based on the relationship between price in one node and other variables that can have influence on the price. Along with the nodal spot price, the deterministic factors considered here are: rainfall in the analyzed node, country's hydro storage, electricity demand and temperature in the region.

The choice of the variables was deliberate with respect to their level of correlation with the prices, as well as the ability to represent both local and country-level situation. Therefore, some of them are considered as measurements

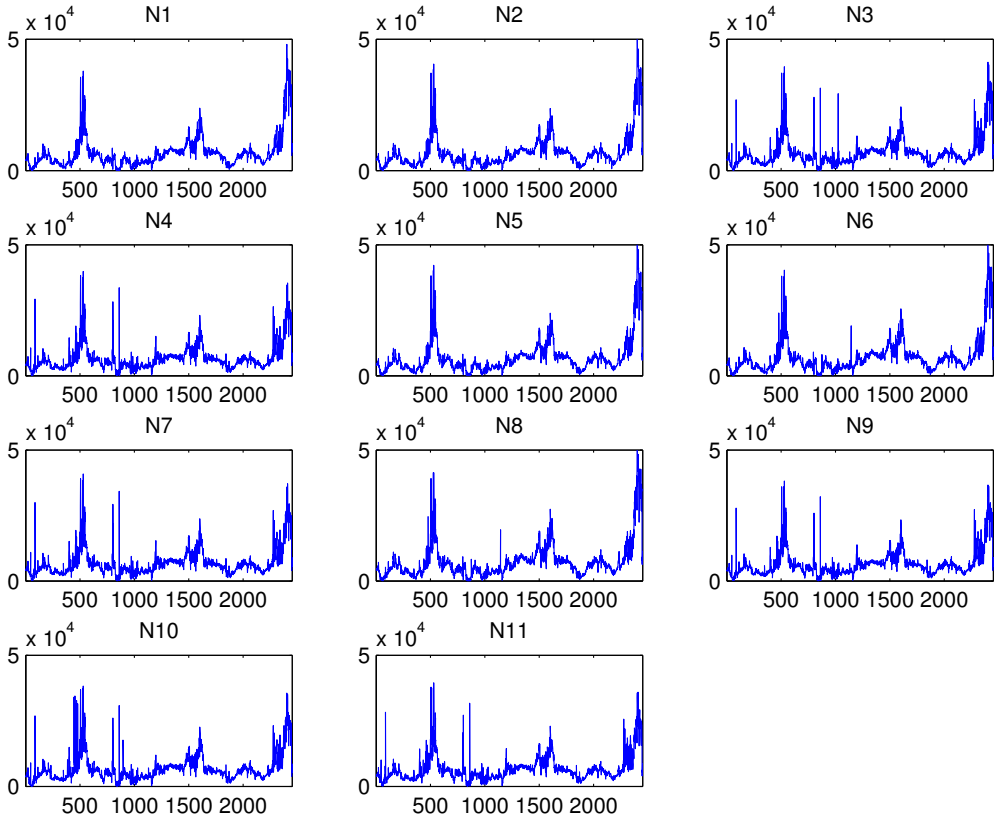

Figure 2: The examined 11 data sets corresponding to electricity prices in New Zealand in the examined period.

from the particular node location, and some are aggregated for the whole country. Firstly, as we mentioned before, the New Zealand electricity market is heavily hydro-dependent and the information on country's hydrological information is of key importance to price modeling (Tipping et al., 2004). Therefore, we include two variables which can account for this feature. The amount of rainfall is measured regionally in millimeters per square meter. This variable tells a lot about whether a given season/year is dry or not in the region. On the other hand, we know that prices in all nodes depend on the entire country's hydro reservoir levels, with correlations between the spot price and the hydro information ranging between -0.39 and -0.49 in various nodes. Namely, when the overall reserves drop, this triggers the prices to raise. Therefore, we add another variable which is the New Zealand's total hydro storage, since the study considers trading nodes which are representing not only hydro powerplants but also other plants and purely splitting substations. In the same fashion we include two demand-related variables, as it is known that when the consumption significantly increases, it may cause the prices to go the same way. One of them is regionally measured air temperature, and another is the total demand calculated for the entire country.

In Fig. 5 we present the five analyzed data sets. As we observe, there is a strong relation between the studied variables, which is especially visible in the seasonal behavior of all data series.

Similarly, as in the first approach, here we should also emphasize the non-stationarity of the data. This non-stationarity can be proved by using ACFs, where we can observe a strong relation between the analyzed processes, see Fig. 6. The only exception is rainfall, not showing non-stationarity.

It is worth mentioning that some of the analyzed vectors of observations exhibit also weekly seasonal behavior. We observe this especially for electricity price data and electricity demand. In Fig. 7, we present ACFs of the differenced series, where weekly seasonality is visible for price and demand data.

The last property is a strong relation between the analyzed data sets that can be proved not only by visualization

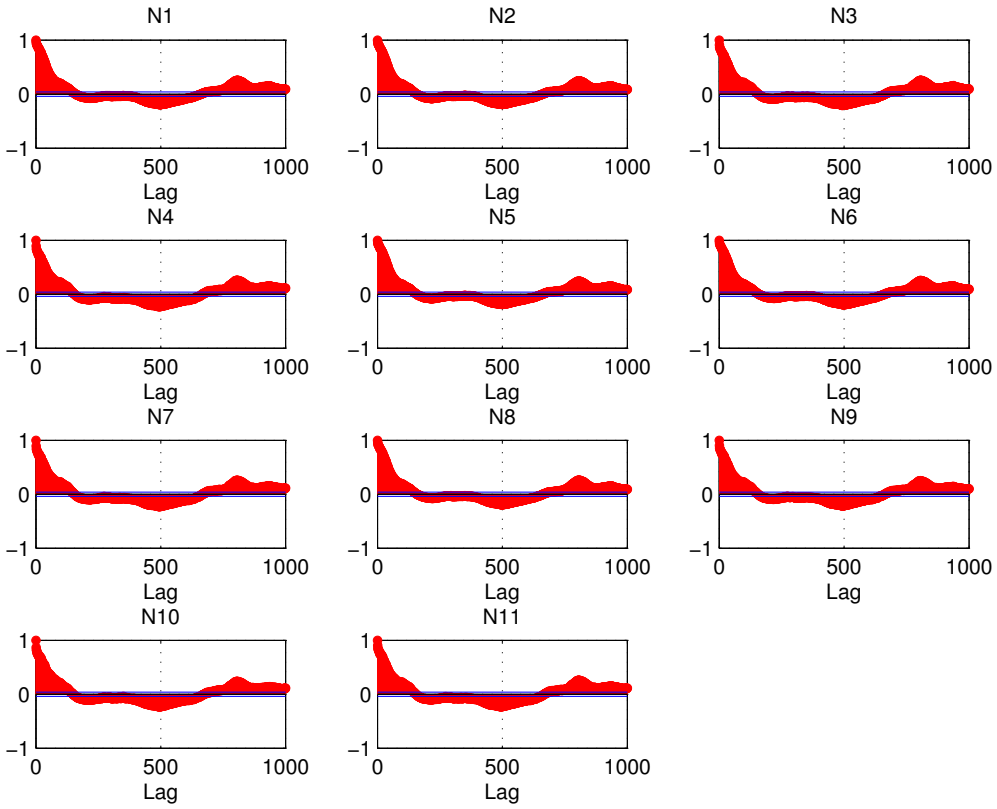

Figure 3: The ACFs of electricity prices in New Zealand in the examined period.

but also a strong statistical test. The correlation coefficients are large and the test for its significance indicates at non-zero correlation between analyzed variables. All p-values for testing the hypothesis of no correlation against the alternative that there is a non-zero correlation are very close to zero which indicates the correlations are significantly different from zero. In particular, the correlation coefficients between the price and other variables vary between -0.54 to 0.2. Only one exception is the correlation between rainfall and temperature. For those variables the p-value of the test for no-correlation is equal 0.6954 that means the correlation is at zero level.

3. Multivariate autoregressive model with non-Gaussian structure

An m-dimensional process $\{\mathbf{X}_t\} = \{X_t^1, ...X_t^m\}$ is an m-variate autoregressive moving average (ARMA) of order (p, q) if it is stationary and if for every t it satisfies the following equation (Brockwell and Davis, 1996):

$$\mathbf{X}_t - \Phi_1\mathbf{X}_{t-1} - ... - \Phi_p\mathbf{X}_{t-p} = \epsilon_t + \Theta_1\epsilon_{t-1} + ... + \Theta_q\epsilon_{t-q}, \tag{1}$$

where $\{\epsilon_t\}$ is an m-variate white noise of mean vector $\mathbf{0}$ and covariance matrix $\Gamma(t, t+h)$ that is independent of t and has the following form:

$$\Gamma(t, t+h) = \Gamma(h)\begin{cases} \Sigma & \text{for } h = 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

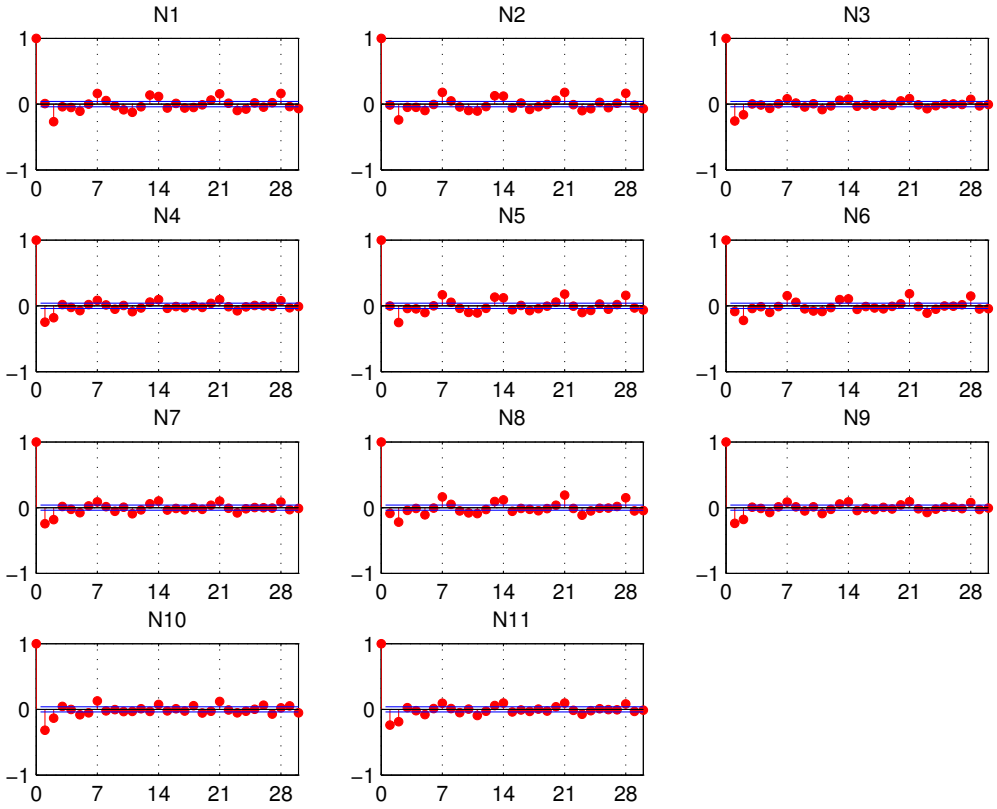

Figure 4: The ACFs of differenced series. The larger values of autocorrelations at lags being multiple of seven indicate at week seasonality.

Let us remind the covariance matrix for m-dimensional vector is the matrix $\Gamma(t, t+h)$ defined as follows:

$$\Gamma(t, t+h) = \begin{bmatrix} \gamma_{11}(t, t+h) & \dots & \gamma_{1m}(t, t+h) \\ \dots & \dots & \dots \\ \gamma_{m1}(t, t+h) & \dots & \gamma_{mm}(t, t+h) \end{bmatrix} \tag{3}$$

In the above definition $\gamma_{ij}(t, t+h) = Cov(X_t^i, X_{t+h}^j)$. In the further analysis we denote $\{\epsilon_t\} \sim WN(0, \Sigma)$ (white noise). The process $\{\mathbf{X}_t\}$ is called ARMA(p, q) with mean μ if $\{\mathbf{X}_t\} - \mu$ is an ARMA(p, q) system.

The multivariate ARMA processes are very useful in practice. They found many applications, especially, because of their simple form. They are extensions of known one-dimensional ARMA models extremely popular in different fields. Some of interesting applications of multivariate ARMA models can be found in Ranam and Sunilkumar (1995); Tiao and Tsay (1989); Stathopoulos and Karlaftis (2003); du Preeza and Witt (2003).

In the classical approach, the residual series $\{\epsilon_t\}$ is assumed to be an m-variate Gaussian distributed random variable, i.e. random variable of the following probability density function:

$$f(\mathbf{x}) = (2\pi)^{-m/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)\right\}, \tag{4}$$

where $\mathbf{x} = (x^1, ..., x^m)$ and $\mu = (\mu^1, ..., \mu^m)$ is the mean vector. Let us mention that if the random variable has m-dimensional Gaussian distribution, then each component has one-dimensional Gaussian distribution with appropriate parameters.

For simplicity, in the further analysis we concentrate only on multivariate autoregressive (AR) models that we denote as VAR (vector AR). In this case, the procedure of estimating the parameters is based on the method of

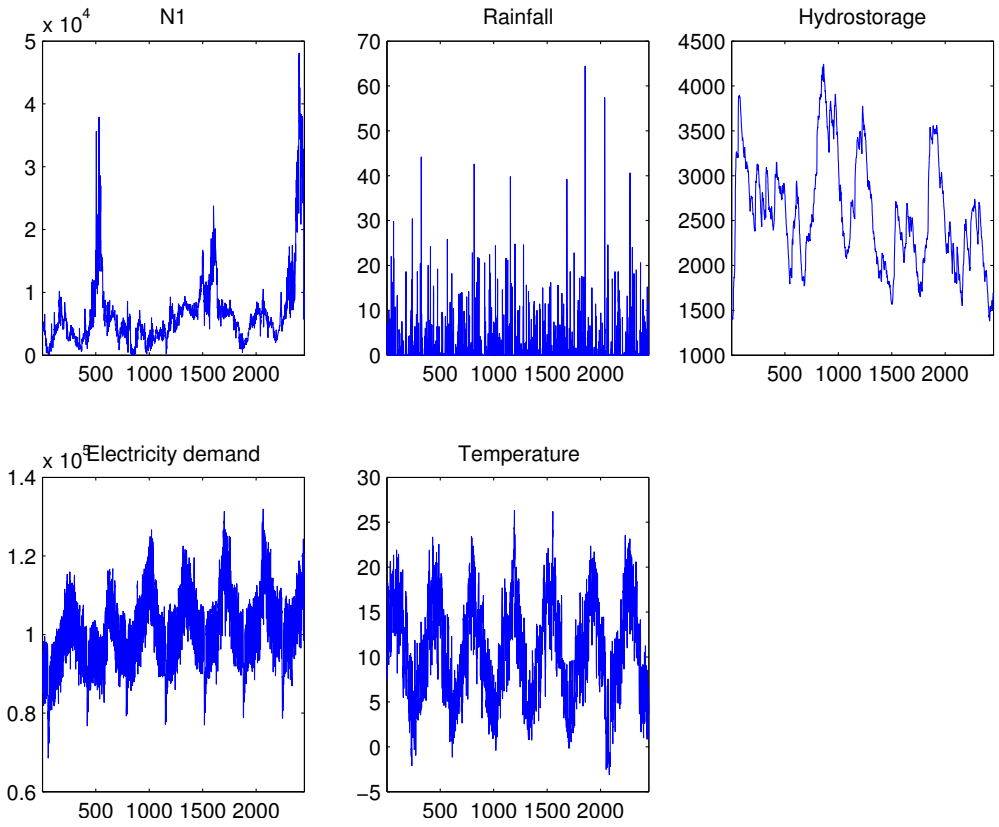

Figure 5: The five examined data sets corresponding to electricity price $N1$, rainfall, hydro storage, electricity demand and temperature. The rainfall and temperature are measured in the region corresponding to node 1.

moments and it is an extension of the popular Yule-Walker method applied to one-dimensional AR systems (Brockwell and Davis, 1996). The method applied to multidimensional AR systems is called the Whittle's algorithm (Whittle, 1963). If we multiply the causal VAR process:

$$\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + ...\Phi_p \mathbf{X}_{t-p} + \epsilon_t \tag{5}$$

by $\mathbf{X}'_{t-j}$ for $j = 0, 1, .., p$ and take expectations of both sides, then we obtain the following equations:

$$\Sigma = \Gamma(0) - \sum_{j=1}^{p} \Phi_j \Gamma(j), \;\; \Gamma(i) = \sum_{j=1}^{p} \Gamma(i-j), \;\; j = 1, 2, ...p$$

Now, taking the empirical equivalences of matrices $\Gamma(0), ..., \Gamma(p)$, we can estimate the matrices $\Phi_1, ..., \Phi_p$ and Σ. Let us remind that a natural estimator of covariance matrix is the sample covariance, that for the vector $\mathbf{X} = (X^1, ..., X^m)$ (each component of n elements) is given by:

$$\hat{\Gamma}(h) = \begin{cases} n^{-1} \sum_{j=1}^{n-h} (\mathbf{X}_{t+h} - \hat{\mu})(\mathbf{X}_t - \hat{\mu}) & \text{for } 0 \le h \le n-1 \\ \\ \hat{\Gamma}'(-h) & \text{for } -n+1 \le h < 0 \end{cases} \tag{6}$$

In the above formula, $\hat{\mu}$ denotes the sample mean vector.

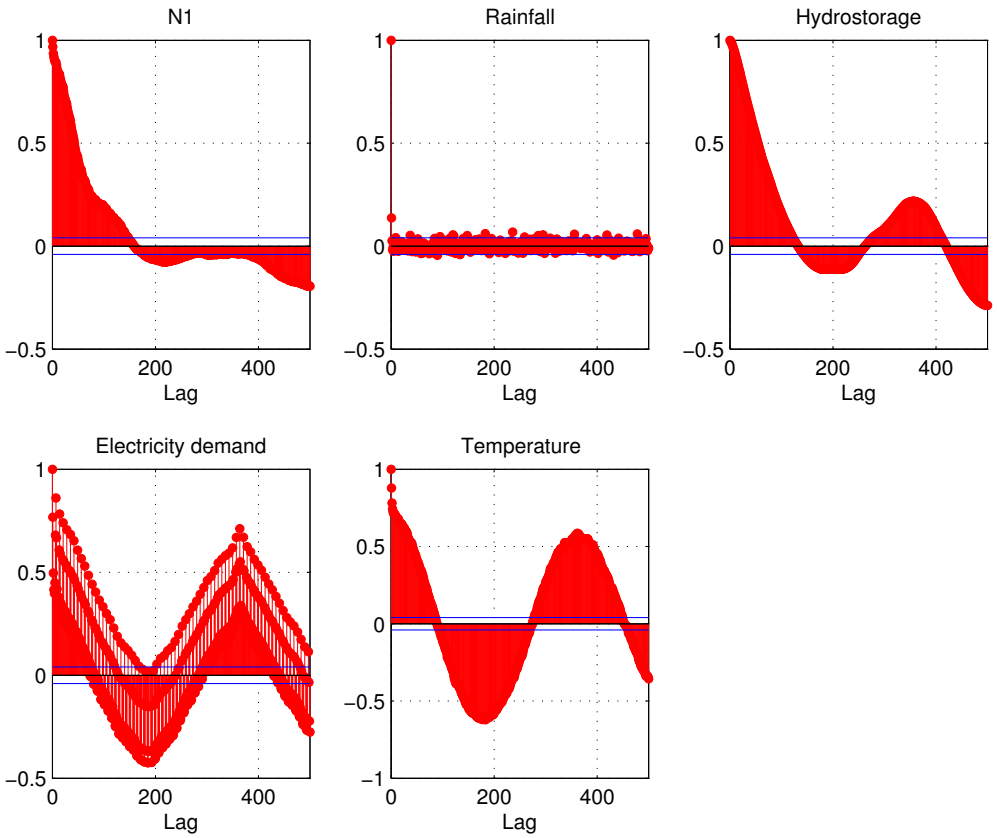

Figure 6: The ACFs of examined 5 data sets corresponding to electricity price $N1$, rainfall, hydro storage, electricity demand and temperature.

3.1. The VAR model with non-Gaussian structure

Because many real data series exhibit behavior not adequate for Gaussian systems, in this paper we extend the classical VAR model presented above by replacing the m-dimensional Gaussian distribution with more general class of distributions, namely, $\alpha-$stable (called also stable), which belong to the family of the so-called heavy-tailed distributions. The $\alpha-$stable distribution is flexible for data modeling and includes Gaussian distribution as a special case. The importance of this class of distributions is strongly supported by the limit theorems which indicate that the stable distribution is the only possible limiting distribution for the normed sum of independent and identically distributed random variables. The interesting applications of a class of $\alpha-$stable distributions can be found, for instance, in Mittnik and Ratchev (2000); Fichea et al. (2013); Stuck and Kleiner (1974). See also Nowicka-Zagrajek and Wyłomańska (2006, 2008).

We start by introducing a one-dimensional $\alpha-$stable random variable. A random variable S has one dimensional stable distribution if, for any numbers $A, B > 0$, there exist numbers $C > 0$ and $D \in R$ such that

$$AS_1 + BS_2 \stackrel{d}{=} CS + D, \tag{7}$$

where S_1 and S_2 are two independent copies of S (Samorodnitsky and Taqqu, 1994). For each stable random variable S there exists a number $\alpha \in (0, 2]$ such that the constant C in equation (7) satisfies the following relation:

$$C^\alpha = A^\alpha + B^\alpha.$$

The second equivalent definition is based on the characteristic function of random variable S.

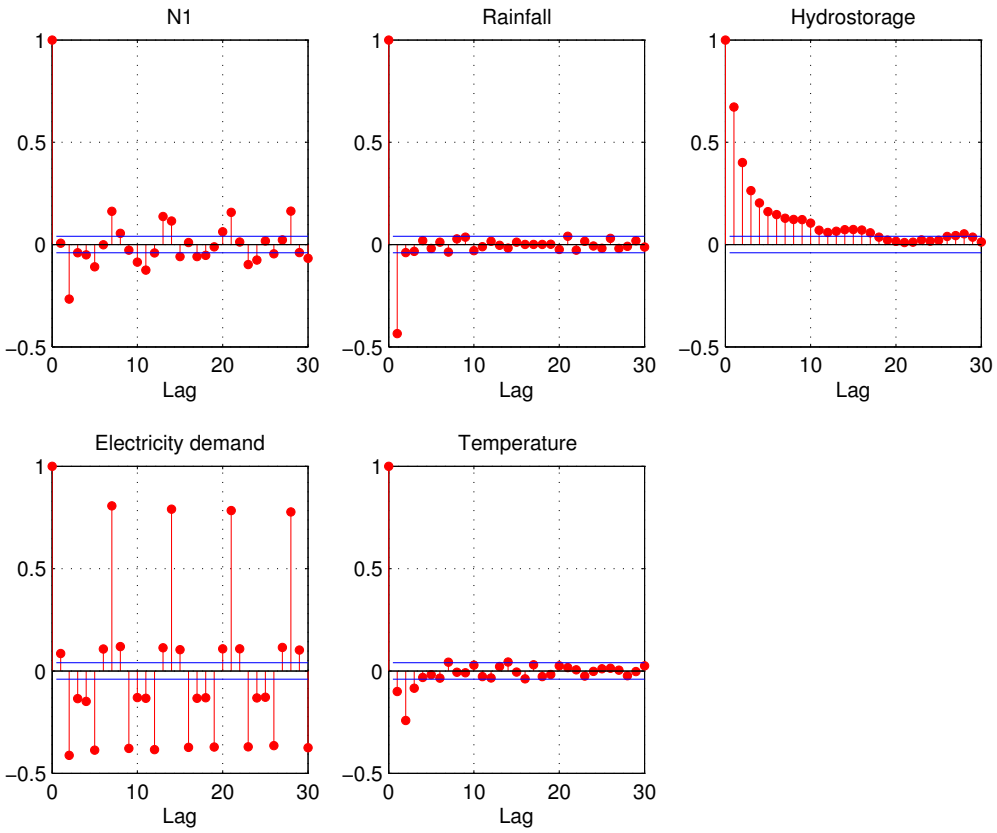

Figure 7: ACFs of the five differenced data sets corresponding to electricity price $N1$, rainfall, hydro storage, electricity demand and temperature.

The random variable S has $\alpha-$stable distribution if there exist parameters $\alpha \in (0,2]$, $\sigma > 0$, $-1 \leq \beta \leq 1$ and $\mu \in R$ such that the characteristic function of S takes form:

$$\mathbb{E}e^{ixS} = \begin{cases} \exp\{-\sigma^{\alpha}|x|^{\alpha}(1 - i\beta sign(x)tan(\pi\alpha/2)) + i\mu x\} & \text{for } \alpha \neq 1 \\ \exp\{-\sigma|x|(1 + i\beta(2/\pi)sign(x)\log(|x|)) + i\mu x\} & \text{for } \alpha = 1. \end{cases} \tag{8}$$

Stability index α, scale parameter σ, skewness parameter β and shift parameter μ in a unique way define the distribution of a random variable S. In the further analysis, we denote $S \sim S(\alpha, \sigma, \beta, \mu)$. The probability density function for most $\alpha-$stable distributions has no explicit form. However, there are three exceptions, namely, Gaussian (for $\alpha = 2$), Cauchy (for $\alpha = 1$ and $\beta = 1$) and Lévy (for $\alpha = 0.5$ and $\beta = 1$).

The extension of a univariate stable distribution is the multivariate one. A characteristic function of an m-dimensional random vector $\mathbf{S} = (S^1, ..., S^m)$ is defined as follows:

$$\Phi(\mathbf{x}) = Ee^{i\sum_{k=1}^{m} x_k S_k}.$$

The m-dimensional vector $\mathbf{S}$ is $\alpha-$stable in R^m if and only if there exists a finite-dimensional measure G on the unit sphere S_m and the vector μ such that:

$$\Phi(\mathbf{x}) = e^{-I(\mathbf{x})+i\sum_{k=1}^{m} x_k \mu^k},$$

where $I(\mathbf{x}) = \int_{S_m} \psi(\sum_{k=1}^{m} s_i x_i) G(ds_1, ..., ds_m)$ and

$$\psi(u) = \begin{cases} |u|^{\alpha} \left(1 - isign(u) tan(\pi\alpha/2)\right), & \text{for } \alpha \neq 1 \\ |u| \left(1 + i\frac{2}{\pi} sign(u) \log(|u|)\right), & \text{otherwise} \end{cases}$$

Moreover, (G, μ) is unique. It is worth mentioning that the stability index α, measure G (called spectral measure), and the shift vector μ in a unique way define the m-dimensional stable distribution denoted as $S^m(\alpha, G, \mu)$. If $\mathbf{S}$ has m-dimensional stable distribution, then each component has one-dimensional stable distribution with appropriate parameters.

The multidimensional AR model with $\alpha-$stable structure is the system defined in (1) for which the residual series $\{\epsilon_t\}$ comes from an m-dimensional $\alpha-$stable distribution.

4. Real data analysis

4.1. Model 1

In the first proposed model, we use the relationship between prices corresponding to 11 nodes. In the first step of our analysis we remove the seasonal behavior from the original data sets, which is related to the annual seasonality observable in electricity prices. In order to do this, we fit to all 11 vectors the sum of sinusoidal functions by using the least squares method. After removing the sinusoidal functions, we differentiate the data. For each vector a different function is fitted (we choose the best fitting) but results after those transformations are similar. This is especially observable in the autocorrelation functions where only weekly seasonality is visible. As an example, we present the vector $N1$ after mentioned transformation (see Fig.8 top panel) and its autocorrelation function (see Fig. 8 bottom panel).

Because the autoregressive models can be used only to stationary series, before we fit the VAR model, we have to remove weekly seasonality from the data. The simplest method is to remove the seasonal mean that is calculated by using the data corresponding to each season. More precisely, from the analyzed data set $X_1, ..., X_{nT}$ (T is the observed season) we remove function w which is calculated as follows:

$$w(t) = \frac{1}{n} \sum_{k=0}^{n-1} X_{kT+t}, \ t = 0, 1, 2, ... \tag{9}$$

After removing the seasonal behavior from the original data sets, we can fit the best VAR model. In order to find the best order p we use the Schwartz-Bayesian criterion (BIC) which, for the vector $X^1, ..., X^m$ (each of n observations), is defined as follows:

$$BIC(p) = \log(|\Sigma(p)|) + \frac{\log(m)}{m} p n^2, \tag{10}$$

where $\Sigma(p) = \frac{1}{m} \sum_{t=1}^{m} \epsilon_t \epsilon_t'$ is the residual covariance matrix from a VAR(p) model. The general approach is to fit a VAR(p) model with order $p = 0, ..., p_{max}$ and choose the value of p which minimizes the selection criterion. In Fig. 9, we present the BIC statistic for order p taking values between 1 and 15. The plot clearly indicates that the best model is VAR(4).

After selecting an appropriate order p, we can estimate the parameters. As it was mentioned earlier, here we use the Whittle's algorithm which is based on the method of moments (Brockwell and Davis, 1996). The model contains $11 * 11 * 4 = 5324$ parameters, therefore, we do not present the estimated values. In the next step of the analysis, we examine the residual series. Since we consider the multidimensional model, the series is also multidimensional. We test the distribution of the residual series for each vector separately. For all vectors the residuals exhibit non-Gaussian behavior. This can be proved by using the Jacque-Bera test for normality. The statistic of JB test for the vector $\epsilon_1, ..., \epsilon_n$ is defined as follows (Jarque and Bera, 1980; Burnecki et al., 2011):

$$JB = \frac{n}{6} \left(S^2 + \frac{K - 3)^2}{4} \right), \tag{11}$$

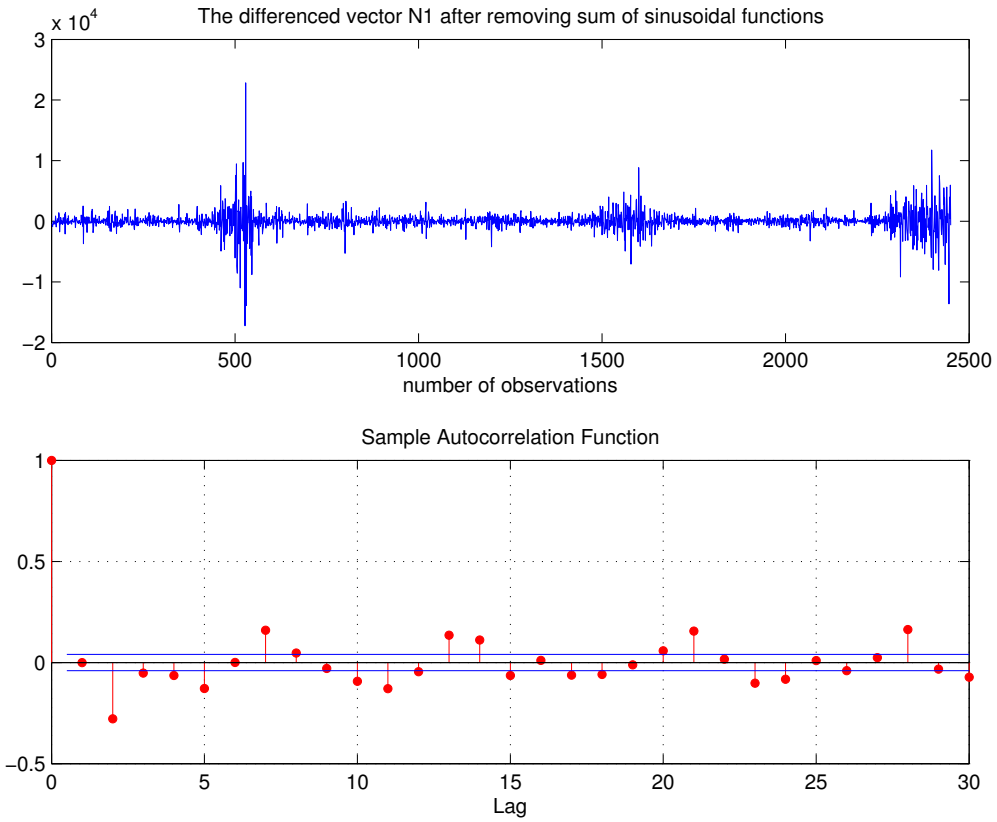

Figure 8: The exemplary differenced vector $N1$ after removing sum of sinusoidal functions (top panel) and its autocorrelation function (bottom panel). The ACF indicates the existence of weekly seasonality in the data set.

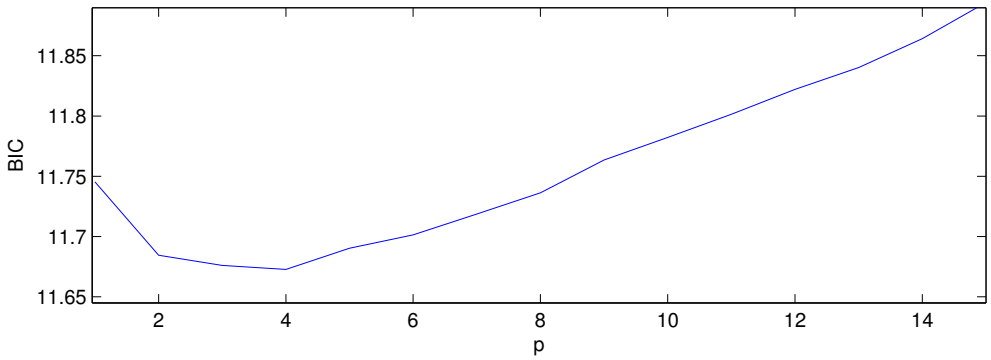

Figure 9: The BIC statistic for selection of the best model that takes under consideration electricity prices relation. For considered data the best model is VAR(4).

where S and K are the sample skewness and kurtosis, respectively, namely:

$$S = \frac{1/n \sum_{i=1}^{n} (\epsilon_i - \bar{\epsilon})^3}{\left(\sqrt{1/n \sum_{i=1}^{n} (\epsilon_i - \bar{\epsilon})}\right)^2},$$

$$K = \frac{1/n \sum_{i=1}^{n} (\epsilon_i - \bar{\epsilon})^4}{\left(\sqrt{1/n \sum_{i=1}^{n} (\epsilon_i - \bar{\epsilon})} \right)^2}.$$

The value of the JB statistic given by (11) forms a random variable which converges to zero if the underlying distribution has skewness zero and kurtosis 3 (e.g., Gaussian). Any deviation of skewness from zero and deviation of kurtosis from 3 increases the JB statistic. For distributions with infinite kurtosis (like $\alpha-$stable with $\alpha < 2$) it diverges to infinity. The test is quite standard and implemented in various numerical packages, like, for example, R or MATLAB. Similar as for all statistical tests, the calculated p-value indicates whether the zero hypothesis can be accepted at the given significance level. If p-value is small, the hypothesis (here of Gaussian distribution) should be rejected. For the analyzed residuals from VAR(4) model the obtained p-values are at the level of 0.001 that indicates non-Gaussian distribution. In order to show the analyzed residual series comes from an $\alpha-$stable distribution, we use two goodness of fit test statistics, namely, Kolmogorov-Smirnov and Anderson-Darling.

The most well-known supremum statistic is the Kolmogorov-Smirnov (KS) statistic. It is just the supremum of the set of distances:

$$KS = sup_x \left| ECDF(x) - F(x) \right|, \tag{12}$$

where F is the cumulative distribution function of the testing distribution and $ECDF(x)$ is the empirical cumulative distribution function, which for the series $\epsilon_1, ..., \epsilon_n$ is calculated as follows:

$$ECDF(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1} \{\epsilon_i \leq x\} \tag{13}$$

In the above definition $\mathbf{1}\{A\}$ denotes the indicator of a set A.

The Anderson-Darling statistic, belongs to the Cramer-von-Misses family of statistics which incorporate the idea of quadratic norm. The Cramer-von-Misses statistic for vector $\epsilon_1, ..., \epsilon_n$ is defined by (Burnecki et al., 2011)

$$Q = n \int_{-\infty}^{\infty} \left(ECDF\left(x\right) - F\left(x\right) \right)^2 \phi\left(x\right) \, \mathrm{d}x \tag{14}$$

where $\phi\left(x\right)$ is a suitable function which puts weights to the squared difference $\left(ECDF\left(x\right) - F\left(x\right) \right)^2$. When $\phi(x) = 1$, Q is called the Cramer-von-Misses statistic. If $\phi\left(x\right) = \left[F\left(x\right) \left(1 - F\left(x\right)\right) \right]^{-1}$, the above definition yields the Anderson-Darling (AD) statistic. Similar as for the JB test, there exists a statistical test that allows to test the proper distribution of examined data. More details can be found in Anderson and Darling (1952, 1954); Burnecki et al. (2012).

In Tab. 1 we present the values of the KS and AD statistics and corresponding p-values of tests for $\alpha-$stable distribution for the 11 analyzed data sets. For the $\alpha-$stable distribution the cumulative distribution function is not given explicitly, thus to obtain the corresponding p-values we use the Monte Carlo method with 1000 repetitions. As we observe, the obtained p-values significantly exceed the significance level 0.05 therefore we can not reject the hypothesis of $\alpha-$stable distribution of residual series.

Next we estimate the parameters of $\alpha-$stable distribution for all residual series. We use here the regression method (Koutrouvelis, 1980; Kogon and Williams, 1998). The method is based on the form of the characteristic function of an $\alpha-$stable random variable $S\left(\alpha, \sigma, \beta, \mu\right)$ given in (8). For a random sample $\epsilon_1, ..., \epsilon_n$ from the $\alpha-$stable distribution, we calculate the empirical characteristic function:

$$ECHF(x) = \frac{1}{n} \sum_{j=1}^{n} e^{ix\epsilon_j}. \tag{15}$$

In the regression method, the empirical characteristic function is compared to the theoretical one and by using least squares method one can estimate the parameters. For more details see Borak et al. (2005). In Tab. 2, we present the estimated parameters for all 11 vectors of residuals. As we observe the estimated α parameters of the residual series are on the level $1.3 - 1.4$, so tails of distributions of data related to the considered nodes are very similar. Moreover the σ parameter in all considered cases is approximately 500, so the scaling in all cases is the same. The β parameters indicate the data are symmetric (β close to zero). Only the shift parameter μ indicates at differences between the distribution of residuals.

Table 1: The values of KS and AD statistics and corresponding p-values of goodness of fit tests.

vector	KS	p-vale (KS)	AD	p-value (AD)
$N1$	0.55	0.60	0.29	0.68
$N2$	0.43	0.95	0.16	0.95
$N3$	0.54	0.60	0.61	0.48
$N4$	0.72	0.27	0.46	0.54
$N5$	0.52	0.65	0.23	0.86
$N6$	0.37	0.97	0.18	0.89
$N7$	0.53	0.67	0.35	0.75
$N8$	0.52	0.63	0.24	0.73
$N9$	0.74	0.18	0.61	0.40
$N10$	0.44	0.88	0.29	0.74
$N11$	0.72	0.23	0.54	0.48

Table 2: The estimated parameters of $\alpha-$stable distribution for the residual series from Model 1.

vector	α	σ	β	μ
$N1$	1.39	463.73	−0.02	−9.42
$N2$	1.42	483.86	−0.04	−9.23
$N3$	1.28	523.45	0.03	21.50
$N4$	1.28	536.31	0.04	17.81
$N5$	1.40	488.23	−0.03	−12.53
$N6$	1.38	513.43	−0.04	−21.46
$N7$	1.28	559.92	0.02	−3.09
$N8$	1.38	540.24	−0.03	−13.56
$N9$	1.27	513.06	0.05	37.71
$N10$	1.28	553.97	−0.03	−42.38
$N11$	1.28	530.19	0.06	36.20

By using the fitted model, we can simulate the VAR(4) system with $\alpha-$stable residuals. In Fig. 10 we present the exemplary vector $N1$ together with the quantile lines on the $10\%,\ldots,90\%$ levels from the simulated samples. The quantile lines are constructed by using Monte Carlo methods with 1000 repetitions. Note that the observed price does not exceed the bounds given by the 10% and 90% quantile lines.

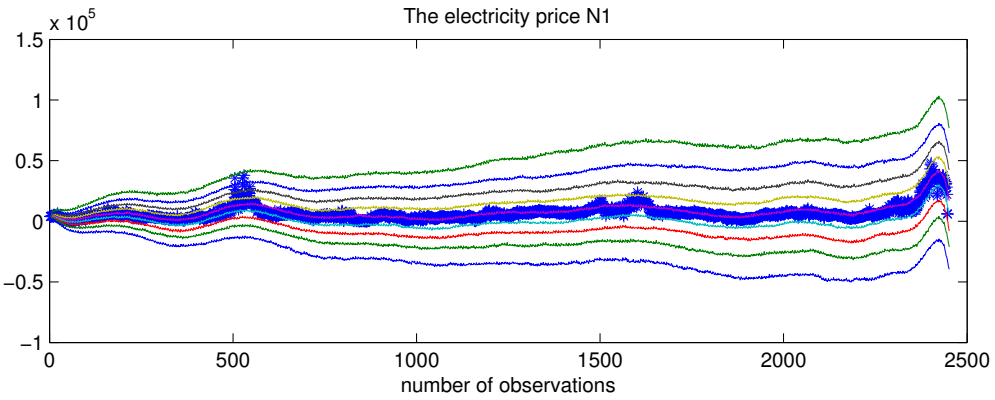

Figure 10: Quantile lines of level 10%, 20%, . . . , 90% and the measured price $N1$ (thick blue line)

4.2. Model 2

In the second approach we propose to use the relationship between the electricity price corresponding to node 1 and other variables such as rainfall, hydro storage, electricity demand and temperature. Let us remind that the rainfall and temperature were measured for the region related to node 1, whereas the storage and demand have been aggregated for the entire country (that is the original format in which the data has been provided by Electricity Commission).

Similarly as in the first approach, in the first step of the analysis we remove the non-stationarity of the data by fitting the sum of sinusoidal functions for all considered vectors of observations. Only one exception is the rainfall where the non-stationarity is not visible, see Fig. 6. Next, we differentiate the data after removing the fitted functions. In the second step, for the data with weekly seasonal behavior (price and electricity demand), we calculate the periodic mean and remove it. The formula for periodic mean is presented in (9).

Next, a proper order p should be calculated. Similar as in the previous approach, here we use the BIC criterion defined in (10). In Fig. 11 we present the plot of BIC statistic for order p varied between 1 and 15.

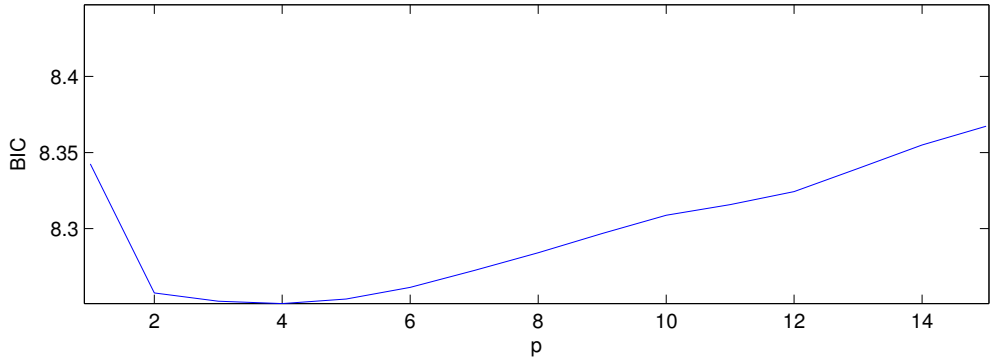

Figure 11: The BIC statistic for selection of the best model that takes under consideration the relationship between electricity price $N1$ and rainfall, hydro storage, electricity demand and temperature. For considered data the best model is VAR(4).

After the proper model has been fitted, we can analyze the residuals. First, we check if they constitute samples from Gaussian distribution. Similar as in Model 1, we use the JB statistic defined in (11) and test based on it. In Tab. 3 we present the p-values of the test for gaussianity. Let us remind, the small p-value (less than significance level 0.05) indicates on non-Gaussian behavior of underlying random sample. As we observe, only the residuals related to temperature can be treated as a sample from Gaussian distribution. The estimated parameters of Gaussian distribution for temperature residuals are $\mu = 0.0034$ and $\sigma = 2.45$.

Table 3: The p-values of the JB test for gaussianity.

price	rainfall	hydro storage	demand	temperature
0.001	0.001	0.001	0.001	0.1786

Next, we check if the the residuals related to price, rainfall, hydro storage and demand constitute samples from $\alpha-$stable distribution. In order to do this, we use the KS and AD goodness of fit tests presented above. In Tab. 4 we present the values of the statistics and corresponding p-values calculated on the basis of 1000 Monte Carlo simulations.

Table 4: The values of KS and AD statistics and corresponding p-values of goodness of fit tests.

vector	KS	p-vale (KS)	AD	p-value (AD)
price	0.64	0.39	0.31	0.72
rainfall	49.37	0	0.26	0.56
hydro storage	49.55	0	1.71	0.82
demand	1.02	0	2.5	0.13

As we observe, the KS test indicates that only the price comes from an $\alpha-$stable distribution. This can be seen from the prominent values of the statistic. But the AD test indicates the residuals of mentioned in Tab. 4 variables can be treated as from $\alpha-$stable family. Therefore, we estimate the parameters of this distribution. In Tab. 5, we present the estimated parameters by using the regression method. Here the situation is different than in Model 2. As we observe there is a difference in α parameters so the tail behavior is different for the considered variables. The other parameters also are different for different variables.

Table 5: The estimated parameters of $\alpha-$stable distribution for the residual series from Model 2.

vector	α	σ	β	μ
price	1.4	457.49	0.07	26.60
rainfall	1.10	1.11	−0.18	108.96
hydro storage	1.19	7.46	0.56	751.87
demand	1.58	1.390	0.19	237.75

By using the fitted model we can simulate the VAR(4) system with residuals which come from the fitted distributions. In Fig. 12 we present the price vector $N1$ together with the quantile lines on the level $10\%,\ldots,90\%$ from the simulated model. The quantile lines are constructed by using Monte Carlo methods with 1000 repetitions. Note that the observed price does not exceed the bounds given by the 10% and 90% quantile lines.

Finally, in order to illustrate how beneficial the obtained models might be in the problem of electricity price description, we calculate the price prediction for the next year. The obtained predicted values are then validated by comparing them to the actual values. To this end, we fit the model to the first 2091 observations and then based on the obtained estimates we calculate the prediction for next 360. The obtained values are plotted in Fig. 13. As can be observed in the figure, the predicted values visually resemble the actual values for both considered models.

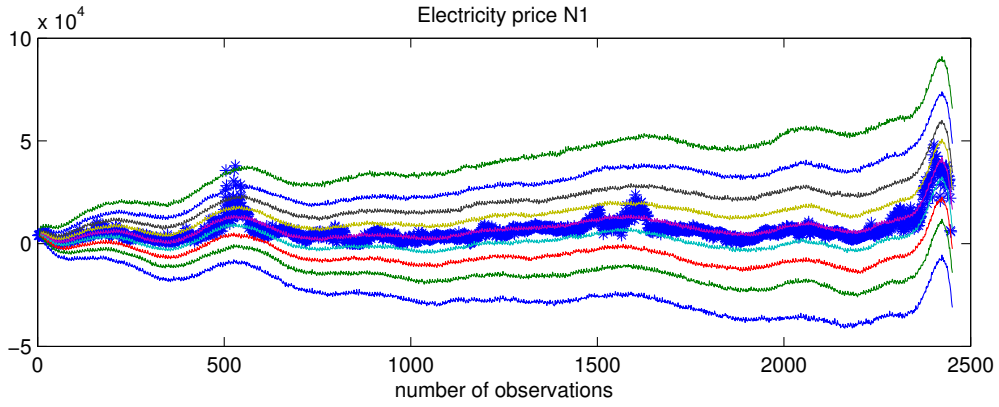

Figure 12: Quantile lines of level $10\%, 20\%, \ldots, 90\%$ and the electricity price $N1$ (blue thick line)

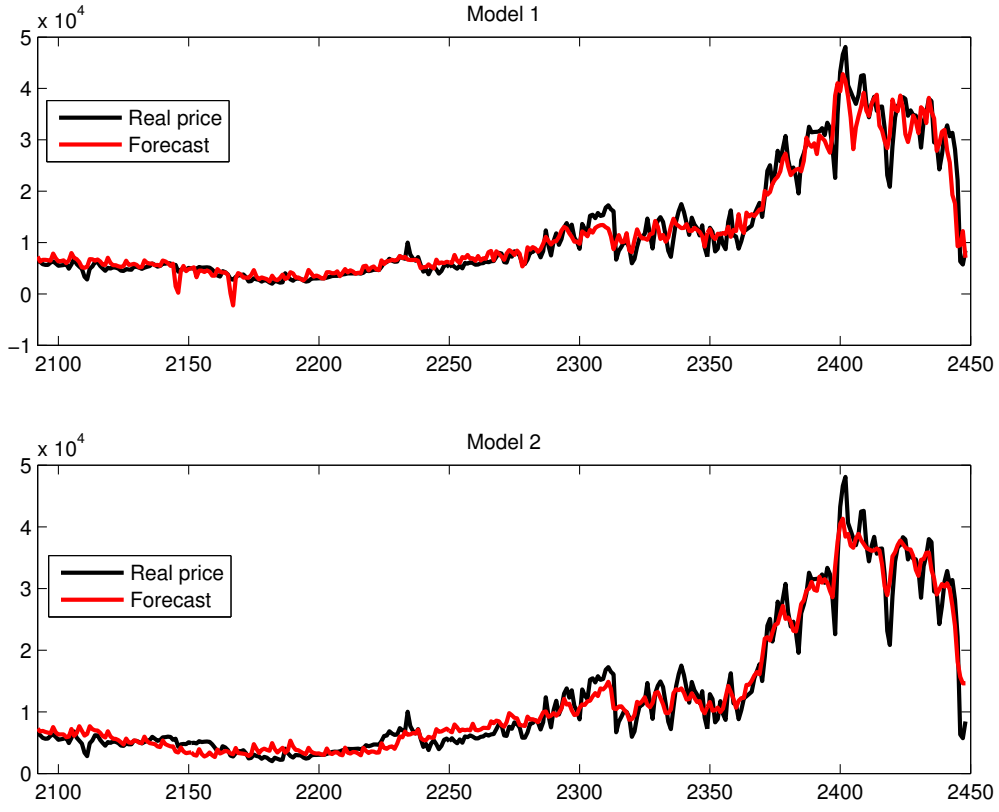

Figure 13: Values of electricity prices $N1$ together with the prediction for the next year. The predicted values were calculated by using the Model 1 (top panel) and Model 2 (bottom panel).

In order to check which model gives better prediction results, we calculate three measures: mean squared error (MSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) of the forecast for the next year. The

measures are defined as:

$$MSE = \frac{1}{k}\sum_{t=1}^{k}(X_t - \hat{X}_t)^2, \quad MAE = \frac{1}{k}\sum_{t=1}^{k}|X_t - \hat{X}_t|, \quad MAPE = \frac{1}{k}\sum_{t=1}^{k}\frac{|X_t - \hat{X}_t|}{X_t},$$

where k is the number of predicted values, X_t is the measured observation at point t and $\hat{X}_t$ its prediction. The obtained values of mentioned measures are given in Tab. 5.

Table 6: Mean square error (MSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) of the prediction for the next year

	MSE	MAE	MAPE
Model 1	$7.48 * 10^6$	$1.73 * 10^3$	17%
Model 2	$4.99 * 10^6$	$1.59 * 10^3$	19%

5. Conclusions

In this work two multivariate autoregressive models with non-Gaussian noise structure for forecasting day-ahead electricity prices in New Zealand have been proposed. The approaches have clear advantages over classical ARMA-GARCH type modeling. Firstly, they do not rely only on the historical values of the price itself, but account for other data series as well. In one of the models we have referred one of the nodal prices to ten other nodes. This is explained by the fact that prices throughout the country are interconnected and strongly correlated. The choice of nodes (11 out of over 200), has been carefully designed to pick representative nodes, but not to make the model too big.

The second approach modeled one of the nodal prices with the used of the regressed price itself together with some deterministic factors having significant influence on price dynamics. These were country's hydrological storage, centralized demand, node's rainfall and air temperature from the studies region. This model proved better in forecasting for two main reasons. Firstly, the forecasting skill was higher than for Model 1. Secondly, the model had less parameters than the first one, which made it more appropriate from modeling and parameter estimation point of view. Finally, second model better results in comparison to Model 1 show that incorporation of deterministic factors in price modeling is of key importance to high forecasting accuracy.

The robustness of our results was confirmed by the fact that both applied models were optimal with the same order, namely, VAR(4). Moreover, the models had non-Gaussian noise included which allowed us to capture statically the non-Gaussian distribution of prices themselves. One could argue that more trading nodes should be included in the analysis. However, we argue that this would unnecessarily complicate parameter estimation without much added value.

Our work provides a sound basis for risk analysis for electricity market traders in New Zealand. Further steps should include the use of Markov Chain Monte Carlo Methods in parameter estimation, to allow proper understanding of price noise structure.

References

Anderson, T., Darling, D., 1952. Asymptotic theory of certain "'goodness-of-fit"' criteria based on stochastic processes. Annals of Mathematical Statistics 23, 193–212.

Anderson, T., Darling, D., 1954. A test of goodness-of-fit. Journal of the American Statistical Association 49, 765–769.

Borak, S., Haerdle, W., Weron, R., 2005. Stable distributions in Statistical Tools for Finance and Insurance. Springer, Berlin.

Brockwell, P., Davis, R., 1996. Introduction to Time Series and Forecasting. Springer, New-York.

Burger, M., Klar, B., Müller, A., Schindlmayr, G., 2004. A spot market model for pricing derivatives in electricity markets. Quantitative finance 4, 109–122.

Burnecki, K., Janczura, J., Weron, R., 2011. Building loss models, in Statistical Tools for Finance and Insurance. Springer, Berlin.

Burnecki, K., Wyłomańska, A., Beletskii, A., Gonchar, V., Chechkin, A., 2012. Recognition of stable distribution with levy index alpha close to 2. Phys. Rev. E 85, 056711.

du Preeza, J., Witt, S., 2003. Univariate versus multivariate time series forecasting: an application to international tourism demand. International Journal of Forecasting 19 (3), 435–451.
Fichea, A., Cexusa, J.-C., Martin, A., Khenchaf, A., 2013. Features modeling with an α-stable distribution: Application to pattern recognition based on continuous belief functions. Information Fusion 14 (4), 504–520.
Guthriea, G., Videbeck, S., 2007. Electricity spot price dynamics: Beyond financial models. Energy Policy 35, 5614–5621.
Jabłońska, M., Nampala, H., Kauranne, T., Summer 2011. Multiple mean reversion jump diffusion model for nordic electricity spot prices. The Journal of Energy Markets 4 (2).
Janczura, J., Weron, A., 2010. An empirical comparison of alternate regime-switching models for electricity spot prices. Energy Economics 32, 1059–1073.
Jarque, C., Bera, A., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters 6 (3), 255–259.
Kian, A., Keyhani, A., 2001. Stochastic price modeling of electricity in deregulated energy markets. In: IEEE Proceedings of the 34th Hawaii International Conference on System Sciences.
Kogon, S., Williams, D., 1998. A Practical Guide to Heavy Tails. Birkhauser, Boston.
Koutrouvelis, I., 1980. Regression-type estimation of the parameters of stable laws. J. Amer. Statist. Assoc. 75, 918.
Laitinen, K., Hovila, J., Mannila, T., Korpinen, L., 2000. The influences of climatic factors on electricity prices in liberalized market in finland. In: International Conference on Electric Utility Deregulation and Restructuring and Power Technologies., 544–548.
Mittnik, S., Ratchev, S., 2000. Stable Paretian Models in Finance. Wiley, New York.
Nowicka-Zagrajek, J., Wyłomańska, A., 2006. The dependence structure for parma models with a-stable innovations. Acta Phys. Polon. B 37 (1), 3071–3081.
Nowicka-Zagrajek, J., Wyłomańska, A., 2008. Measures of dependence for stable ar(1) models with time-varying coefficients. Stochastic Models 24 (1), 58–70.
Podraza, E., Fall 2006. Challenges in forecasting electric load in deregulated markets. The Journal of Business Forecasting 25 (3), 31–35.
Qureshi, W., Nair, N.-K., Farid, M., 2009. Demand-price analysis of electricity market using discrete wavelet transform. Power Engineering Conference, AUPEC 2009. Australasian Universities.
Ranam, H., Sunilkumar, N., 1995. Multivariate modeling of water resources time series using artificial neural networks. Hydrological Sciences Journal 40 (2), 145–163.
Ruibal, C., Mazumdar, M., 2008. Forecasting the mean and the variance of electricity prices in deregulated markets. IEEE Trans. Power Systems 23 (1), 25–32.
Samorodnitsky, G., Taqqu, M., 1994. Stable Non-Gaussian Random Processes. Chapman and Hall, New York.
Situ, R., Nair, N.-K., 2007. Applying wavelet transform techniques for new zealand electricity market volatility analysis. Power Engineering Conference, AUPEC 2007. Australasian Universities.
Stathopoulos, A., Karlaftis, M., 2003. A multivariate state space approach for urban traffic flow modeling and prediction. Transportation Research Part C: Emerging Technologies 11 (2), 121–135.
Stuck, B., Kleiner, B., 1974. A statistical analysis of telephone noise. Bell Syst. Tech. J. 53, 1263–1320.
Tiao, G., Tsay, R., 1989. Model specification in multivariate time series. Journal of the Royal Statistical Society. Series B 51 (2), 157–213.
Tipping, J., Grant Read, E., Mcnickle, D., 2004. The incorporation of hydro storage into a spot price model for the new zealand electricity market.
Vehviläinen, I., Pyykkönen, T., 2005. Stochastic factor model for electricity spot price – the case of the nordic market. Energy Economics 27, 351–367.
Vucetic, S., Tomsovic, K., Obradovic, Z., 2001. Discovering price-load relationships in california's electricity market. IEEE Trans. Power Systems 16 (2), 280–286.
Whittle, P., 1963. On the fitting of multivariate autoregressions and the approximation of canonical factorization of a spectral density matrix. Biometrika 40, 129–134.
Ying, Y., Flynn, P., 2003. Power price in deregulated markets. IEEE Power Engineering Society General Meeting 2.