The long-term trends on Russian electricity market: comparison of empirical mode and wavelet decompositions

Dmitriy Afanasyev and Elena Fedorova

Financial University under the Government of Russian Federation, National Research University Higher School of Economics

24. February 2015

The long-term trends on Russian electricity market: comparison of empirical mode and wavelet decompositions ☆

Dmitriy O. Afanasyev[a,*], Elena A. Fedorova[b]

[a] *Financial University under the Government of Russian Federation, Moscow, Russia*
[b] *National Research University Higher School of Economics, Moscow, Russia*

Abstract

The problem of trend-cyclic component filtering from price time-series arises in many commodity market studies, including those of wholesale electricity market. The long-term component filtering is an important part of price analysis since incorrect determination of this component may result in substantial risk underestimation, distorted expectations of both consumers and power generating companies, as well as financial losses. A great strand of literature on this topic proposes quite a lot of approaches and procedures for solving this problem, but all of them suffer from two principal flaws: (1) inability to deal with non-stationary and nonlinear processes; (2) assumption of an "a priori", knowledge of the phenomenon being studied. The complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) allows to effectively overcome these flaws and is expected to produce more adequate results as compared to other methods. In order to check this, we compare the performance of CEEMDAN with the ordinary EMD and yet another well-known approach - the wavelet-decomposition, with an example of the Russian day-ahead electricity market (price zones Europe-Ural and Siberia). Our results shows that the CEEMDAN is much more effective than the standard EMD and is comparable with the wavelet-decomposition (in terms of trend estimation error). At the same time, we found that there are some real data problems with the criterion of the number of low-frequency modes that are included into trend.

Keywords: electricity market, trend-cyclic component, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), wavelet-decomposition

☆Matlab codes used in this article are available for download here: `http://dmafanasyev.ru/en/ceemdan-vs-wave-en/`. CEEMDAN implementation is available here: `http://www.bioingenieria.edu.ar/grupos/ldnlys/`. If you use the codes from this paper in yours own researches, please do not forget to cite this paper as well as Colominas et al. (2012).

*Corresponding author. Financial University under the Government of Russian Federation, Financial management department. Post address: 125993, Leningradskiy av. 49, Moscow, Russia. Phone: +7 926 6320115

Email addresses: `dmafanasyev@gmail.com` (Dmitriy O. Afanasyev), `ecolena@mail.ru` (Elena A. Fedorova)
URL: `http://dmafanasyev.ru` (Dmitriy O. Afanasyev)

1. Introduction

In many applied econometric studies concerned with financial time-series analysis, authors face the problem of separation of long-term dynamics and short-term fluctuations in the variables studied. Mathematically, this problem means the necessity of extraction of trend T_t and cyclical C_t components out of a time-series[1]. The residual of such decomposition can be considered as stochastic part S_t. Traditionally, their combination in an additive ($P_t = T_t + C_t + S_t = TC_t + S_t$) or a multiplicative ($P_t = T_t \times C_t \times S_t = TC_t \times S_t$) form is used for the original time-series reconstruction. It is worth noting that the second form for positive-valued time-series can be brought to the first one by the standard logarithmic transformation.

Despite apparent conceptual simplicity of this approach, there is a whole number of problems with its practical implementation. First of all, the absence of an unambiguous definition of "trend-cyclical component" term. It is intuitively clear that this term reflects the low-frequency oscillations in the time-series analyzed. Though, the exact quantitative criteria for its identification do not follow from this intuition. This results in the second problem: the absence of a generally accepted method for trend filtering. In a great strand of literature on this topic, there exist a lot of methods and procedures to solve the latter problem. But, even provided this diversity of methods, there is a third difficulty which is that most of these methods are not able to deal with either non-stationary or non-linear time-series, while in practice we usually meet such complicated financial processes.

The specified problems are also directly related to the wholesale electricity market where the commodity price[2] shows intraday, weekly, and annual seasonality. In addition to specified above, electricity market has other peculiarities: impossibility to store the commodity on the market for long time; coincidence of production and consumption time-moments; presence of producers (generation technologies) that can not cancel product delivery due to price decrease; low short-term price elasticity of demand; occurrence of substantial price outliers (positive - "spikes" and negative - "drops"); mean long-term price reversion. At the same time, the long-term component is an important part of price modeling since incorrect determination of this component may result in substantial risk underestimation, distorted expectations of both consumers and power generating companies, as well as financial losses. This is why in the current study we focus on trend component filtering, leaving the questions of short-term price fluctuations in the background.

Let us recall some most common approaches to trend component filtering that are typically used in electricity market studies:

- Polynomial regressions with different powers (though, linear regression is most frequently used) (De Jong, 2006; Weron et al., 2004).
- Linear regression on time-variable t given a moving window or a so-called "loess-regression" (Bordignon et al., 2013; Veraart and Veraart, 2012).
- Moving average (including exponentially weighted moving average) or moving median given a moving window (De Jong, 2006; Trück et al., 2007; Nowotarski et al., 2013).

[1] Though, in our study, traditionally for electricity market studies, we do not separate these components and consider an aggregate trend-cyclical component as trend $TC_t = T_t + C_t$.

[2] Hereafter, when speaking about electricity price, we assume the logarithm of that price, which is commonly accepted in econometric studies.

- Dummy variables regression (piecewise continuous functions) for each month in the year (Lucia and Schwartz, 2002; Haldrup et al., 2010; Fanone et al., 2013).

- Fourier transform, i.e. decomposition into a sum of sine functions of different frequencies (De Jong, 2006; Janczura et al., 2013).

- Wavelet-decomposition where different families of wavelets (localized in time, auto-modal (self-similar) functions with zero mean) are used as the basis of decomposition (Trück et al., 2007; Janczura et al., 2013).

- Hodrick-Prescott (HP) filter (Hodrick and Prescott, 1997), which is based on the linear minimization problem being solved for a specified smoothing parameter (Weron and Zator, 2014).

These approaches can be used both independently and as a combination (for instance, linear trend with a sum of sine functions; Fourier transform with exponentially weighted moving average; etc.). The main problems occurring while applying these methods are: (1) the necessity to "a priori" set the value for some unknown parameters; (2) the inability of these methods to deal both with non-stationary and non-linear processes. This allows to conclude that it is necessary to use a new approach to electricity market studying, which will effectively overcome these problems. We propose to consider the empirical mode decomposition (EMD) (Huang et al., 1998) as such an approach.

The central idea of EMD is a local and highly data-adaptive decomposition of a time-series into intrinsic mode functions (IMF) with different average periods: from low-frequency to high-frequency components. The main advantage of EMD is its intrinsic ability to deal with non-stationary and non-linear processes since there are no "a priori" assumptions on these properties. Also, this approach does not require an a priori specification of any parameters (unlike, for example, in wavelet-decomposition or HP filter application).

Nevertheless, empirical mode decomposition (Huang et al., 1998) in its classical form has several flaws that we consider further. In order to get rid of them, Wu and Huang (2009) proposed to use ensemble empirical mode decomposition (EEMD). But, at the same time, EEMD both introduces additional noise into the results of decomposition and does not produce a stable number of IMFs after applying to the same time-series. Complete ensemble empirical mode decomposition with adaptive noise, proposed in Colominas et al. (2012), solves these problems, being at the same time quite parsimonious to computing resources.

There are not so many electricity market studies that employ EMD (Kurbatsky and Tomin, 2010; An et al., 2013; Ismail, 2013; Ghelardoni et al., 2013). Its application is limited to filtering noise components out of time-series. There are even fewer studies that address the issues of comparing EMD with other popular methods of trend filtering. Mhamdi et al. (2010) compared EMD and HP filter on the data on peak electricity loads and showed that EMD provides quite adequate results while not requiring selection of an optimal value of a smoothing parameter which is required for HP filtering.

In all the above-mentioned studies the authors considered the residual of decomposition without oscillatory component as trend component. Moghtader et al. (2011) proposed a more advanced approach which, based on certain criteria, allows to include low-frequency IMFs into trend, thus resulting in its more exact filtering since in that case possible changes in trend direction are taken into account. In its essence, thus calculated estimate is the trend-cyclical component containing the long-term fluctuations. Moghtader et al. (2011) compared their approach to HP filtering on simulated data and concluded that their approach is quite effective (the same conclusions were obtained in Mhamdi et al. (2010)).

As of now, as far as we know, there are no studies that thoroughly compare the approach of Moghtader et al. (2011) with another popular approach - the wavelet-decomposition. This may be especially topical since relatively recently in Nowotarski et al. (2013) it was shown that using wavelet-decomposition for trend filtering (in order to forecast electricity prices) shows much better performance than the Fourier transform and dummy variables regression. Moreover, as of our knowledge, using CEEMDAN in combination with Moghtader et al. (2011) approach for trend filtering was not proposed in previous studies.

In order to fill this gap, in our study we empirically compare the wavelet-decomposition, EMD and CEEMDAN in the context of price trend filtering on Russian wholesale electricity market. We also propose a new "low-frequency" criterion for determining the number of IMFs included into trend, taking into account their statistical significance. To model the stochastic part of decomposition, we apply the approach of Janczura et al. (2013), but in contrast to this study, we consider a skewed generalized normal distribution (GND) to investigate the price peak regimes.

The rationale is as follows. First of all, the specified distribution has not been used in solving such problems and thus has some methodological potential. Second, it allows us not to specify a strictly low (in "spike" regimes) and a strictly high (in "drop" regimes) values of electricity price, but at the same time allows to take into account theoretically underpinned skewness of density function of the price distribution.

The rest of the paper is organized as follows. In section 2 we consider the methodology of the research: CEEMDAN, criteria for trend-cyclical component filtering, and simulation experiment design. In section 3 the dataset used for calculations is described. Section 4 contains the discussion of the results obtained. Section 5 concludes.

2. Research methodology

2.1. Empirical mode decomposition

Empirical mode decomposition (the Huang transform) was first proposed in Huang et al. (1998) and is itself a highly data-adaptive method allowing to obtain the original signal (time-series) decomposition into oscillatory components. Using EMD, the original signal is decomposed into a sum of intrinsic mode functions (IMFs) that meet two conditions (see (Huang et al., 1998)): (1) the number of extrema and the number of transitions through zero (intersection of the time axis) is either equal or differ by no more than one; (2) at any time, the average value of the envelope constructed on the local maxima and the envelope constructed on the local minima is equal to zero. In order to find the IMFs, an iterative sieving algorithm proposed in Huang et al. (1998) is used. The original signal $x[t]_{t\in(1,T)}$ may be reconstructed as a sum I of the obtained IMFs and the residual $r[t]$ (the empirical basis of decomposition):

$$x[t] = \sum_{i=1}^{I} IMF_i[t] + r[t]. \quad (1)$$

But the classical EMD has several flaws, specifically (1) a substantial influence of boundary effects on the decomposition components obtained; (2) the problem of mode mixing (modes with different average frequencies are mixed within one IMF or a mode corresponding to one frequency is included into several IMFs). In order to overcome these problems, Wu and Huang (2009) proposed to use an ensemble empirical mode decomposition (EEMD). The central idea of EEMD is that the EMD procedure is run for quite a substantial number of times, and at each iteration the original signal is augmented by different realizations of white noise $n_k[t] \sim N(0,1)$ with limited amplitude α: $x_k[t] = x[t] + \alpha n_k[t]$. As a result, a set of $IMF_i^k[t]$ is formed, where $k = (1, \dots, K)$ is the iteration number. A more exact estimate $\overline{IMF_i}$ is calculated by (simple) averaging of the IMFs obtained.

EEMD allows to effectively solve the problems of mode mixing and boundary effects, but introduces two additional difficulties: (1) the reconstructed signal $x[t]$ contains residual noise; (2) the number of IMFs may differ for the same decomposition. To solve these difficulties, Colominas et al. (2012) proposed to use a complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN). The main difference of CEEMDAN from EEMD is in the way of the white-noise component addition. In the case of EEMD, each realization of the signal with noise is decomposed into modes independently, and the residuals obtained for each realization are also independent, providing their individual contribution to the resulting residual of the decomposition. CEEMDAN adds noise not to the original noise but to the resulting residual from the previous iteration. Also, it is not the noise itself that is used, but the noise's mode corresponding to the iteration and obtained with EMD. Thus, noise in CEEMDAN is adaptive and does not create additional input to the original signal since its influence is averaged at each iteration. Following Colominas et al. (2012), let $E_i(\cdot)$ be the operator of the i-th mode extraction by using EMD (with $E_0(x[t]) = x[t]$), and let $\widetilde{IMF_i}$ be the mode which is extracted by using CEEMDAN. Setting $r_0[t] = x[t]$, $i = 1$, the algorithm of this method can be described as follows (see (Colominas et al., 2012)):

1. Extract the first mode for K realizations[3] of white noise $r_{i-1}[t] + \alpha_{i-1}E_{i-1}(n_k[t])$ and find the i-th mode of the original signal by averaging the result obtained:

$$\widetilde{IMF_i}[t] = \frac{1}{K}\sum_{k=1}^{K} E_1(r_{i-1}[t] + \alpha_{i-1}E_{i-1}(n_k[t])). \tag{2}$$

2. Calculate the i-th residual with $r_i[t] = r_{i-1}[t] - \widetilde{IMF_i}[t]$.

3. If $r_i[t]$ has at least two extrema, repeat the procedure for the following i.

The original signal $x[t]$ may be reconstructed with the obtained modes $\widetilde{IMF_i}[t]$ using the formula analogous to (1), which makes the decomposition complete. It was shown in Colominas et al. (2012) that CEEMDAN (unlike EEMD) is robust to changes in added noise amplitude. The precision of the original signal reconstruction does not depend significantly on the signal-noise ratio (SNR), while the global minimum of signal reconstruction error is observed at the amplitude equal to 0.2. In addition to the advantages described above, CEEMDAN is also more parsimonious with respect to computational resources. Motivated by these facts, in our study we use specifically CEEMDAN to decompose the time-series of electricity prices into IMFs.

[3] In this study, we use $K = 300$ realizations of a white noise.

2.2. Trend filtering with CEEMDAN

It is intuitively clear that, when decomposing the original signal into several IMFs, the trend-cyclical component is represented by a sum of several low-frequency modes and the decomposition residual. Of special interest is the number i^* starting with which to compose the modes into trend. Flandrin et al. (2004) were the first to show that EMD can be effectively used for trend filtering based on the statistical properties of IMFs. According to these authors, i^* corresponds to the number of IMF starting with which the standardized mean of partially reconstructed signal $\widehat{x}[t] = \sum_{i=1}^{i^*} IMF_i[t]$ is significantly different from zero. But this rule is not exact, and in order to increase its precision the authors considered the statistical properties of IMFs of a fractional Gaussian noise (fGn) for different values of the Hurst exponent H (Hurst, 1951). The following formula was proposed for calculation of confidence interval C_i^H for the i-th mode's energy G_i:

$$\log_2(\log_2(C_i^H/\widehat{W}_i^H)) = a_H i + b_H, \tag{3}$$

where $\widehat{W}_i^H = (\widehat{W}_1^H/\beta_H)\rho_H^{-2(1-H)i}$ is the estimated value of the i-th mode's energy ($i \geqslant 2, \rho_H \approx 2$) of a fractional Gaussian noise (the first mode of the initial signal is considered equivalent to noise, i.e. $\widehat{W}_1^H = G_1$). If $G_i > C_i^H$ then the i-th IMF of the initial signal is recognized statistically significant and, according to Flandrin et al. (2004), is included into the trend component. To apply the statistical criterion, it is necessary to determine the value of H. To do this, we were guided by the following qualitative considerations. It is well known (Peters, 1991) that a process with $0 \leqslant H < 0.5$ is anti-persistent, i.e. it has the mean-reversion property. Taking this into account, we choose $H = 0.2$ for electricity prices. The choice of this exact value out of a whole range of values is dictated by the availability of empirical coefficients β_H, a_H, b_H for different levels of confidence γ (Flandrin et al. (2004)).

In Moghtader et al. (2011), two approaches to finding i^* were proposed: energy and zero-crossing ratio, and the efficiency of their simultaneous usage was shown. According to the energy approach, i^* is the smallest of $i \geqslant 2$ for which $G_i > G_{i-1}$, i.e. the energy of the current mode is greater than the energy of the previous mode. This follows from the fact that, as shown in Flandrin et al. (2004), if the studied process is a generalized broadband signal, then the energy of its IMFs decreases as index i increases. Thus, if energy rises around some index i, then this index is a candidate for i^*.

The second approach proposed in Moghtader et al. (2011) implies the analysis of a ratio of zero-crossing number (RZCN) value $R_i = Z_{i-1}/Z_i$, where Z_i is the number of points where the i-th IMF crosses zero. The authors showed that if the studied process is again a generalized broadband signal then $R_i \approx 2$. Thus, i^* is the smallest of indexes i such that R_i is "significantly different from 2". In order to specify the value of this difference, the authors ran a number of simulations and constructed the empirical distribution of R_i. In their paper, for a number of most common significance levels α, the left and right boundaries were found, values beyond which are considered as significant deviations from 2.

Since the two approaches described above are independent, their combination allows to increase the accuracy of i^* determination, as it was shown in Moghtader et al. (2011). Thus, the smallest of i ($2 \leqslant i \leqslant I$) such that $G_i > G_{i-1}$ and R_i is significantly different from 2, can be considered i^*, while the trend can be obtained as a sum of the residual and all the IMFs with $i \geqslant i^*$.

In our research, we have modified this approach. First, we propose to use CEEMDAN instead of EMD as the decomposition procedure. Second, as potential components to include in trend, we consider only the IMFs with indexes i such that $I/2 + 1 \leqslant i \leqslant I$ (we call it a "low-frequency" criterion). This is motivated by the fact that, in our opinion, only the IMFs with rather high indexes (the low-frequency modes) should be taken into consideration. As we showed earlier in Afanasyev et al. (2014), the application of this criterion allows to cut off the high-frequency IMFs of electricity prices, which, on the one hand, satisfy the two other criteria, but on the other hand, logically should not be included into the trend component. Third, in addition to the above-described criteria for i^* determination, we propose to take into account only the IMFs that are statistically different from the white noise with a given Hurst exponent H (we call this a "statistical" criterion). For these IMFs, the condition $G_i > C_i^H$ should hold, where C_i^H is a confidence interval which can be determined from (3).

To compare CEEMDAN with EMD, we also consider this latter method in our research. Taking into account the fact that EEMD has substantial disadvantages, in particular, this method is influenced by the residual noise, which results in obtaining trends with high-frequency fluctuations with small amplitudes, we do not consider EEMD in this research. In our opinion, EEMD can be thought of as a transitory step between EMD and CEEMDAN.

2.3. Trend filtering with wavelet-decomposition

The usage of wavelet-decomposition for trend filtering from electricity prices was thoroughly studied in many papers (see, for example, Janczura et al., 2013; Nowotarski et al., 2013). We do not dwell here on the wavelet methodology, referring the interested reader to the references on the topic. It should only be noted that, using the results from Nowotarski et al. (2013), we base our study on Daubechies' wavelet of order 24, as well as consider 3 scale parameters $m = 6, 7, 8$ (hereinafter denoted by W_6, W_7, W_8) that correspond to smoothing at time-scales $2^6 \approx 1$ month, $2^7 \approx 2$ months, and $2^8 \approx 4$ months.

2.4. Modeling the stochastic component

For modeling the stochastic component of decomposition, we use a modification of an approach proposed earlier in Janczura and Weron (2010), which is based on Markov regime-switching model.

The main property of regime-switching models is their ability to capture substantial structural breaks in studied processes. This follows from the fact that the parameters of these models discretely vary in time, i.e. there are several regimes or "states". This class of models was considerably developed in Hamilton (1989); Hamilton and Lin (1996) where an unobserved first order Markov chain is considered as the control process for regime switching. The transition probability from regime j to regime i is determined by a square matrix $\mathbf{P} = (p_{ij})$ with dimensions $r \times r$, where r is the number of regimes.

Following Janczura and Weron (2010), we assume that in the basic operating market regime the stochastic component of electricity price follows the Ornstein-Uhlenbeck (OU) process. This specification allows to take into account the well-known property of electricity price – mean-reversion (Janczura et al., 2013). In discrete time, the OU-process is transformed into an auto-regressive process of order 1 (AR(1)).

In this study we do not perform pre-removal of the component corresponding to weekly seasonality in the data from the stochastic part of the process. Instead of this, we include the 7-th lag into the basic regime, which reflects the influence of a short-term calendar effect. As a result, our specification of electricity price $P_{t,b}$[4] in the basic regime is as follows:

[4]Unlike Janczura and Weron (2010) we use $P_{t,b}$ to denote the logarithm of electricity price.

$$P_{t,b} = \mu_b + \psi_1 P_{t-1,b} + \psi_7 P_{t-7,b} + \sigma_b \epsilon_t, \tag{4}$$

where b denotes the basic regime, ψ_i is the coefficient of the i-th lag of price, μ_b is the constant term (for AR(1) process, $\frac{\mu_b}{1-\psi_1}$ is the long-term mean price level), σ_b is the volatility of residuals, $\epsilon_t \sim i.i.d.\, N(0,1)$. Given that the stochastic part does not contain a trend component, we expect that $\mu_b \approx 0$.

A characteristic property of electricity prices is a non-periodic occurrence of sudden "outliers" (Weron et al., 2004; De Jong, 2006; Janczura et al., 2013). They can be caused by such events as climatic anomalies, failure of the power generating equipment of a major power supplier, or problems in electricity transmission networks. These outliers can be taken into account with two additional model regimes: the "spike" regime s and the "drop" regime d. Janczura and Weron (2010) propose to use a truncated log-normal distribution for s-regime modeling, and an inverted truncated log-normal distribution for d-regime modeling. Janczura et al. (2013) propose to use the 1st and the 3rd quantiles for truncation.

In this study, we consider the logarithms of prices in these outlier regimes as a mean augmented with a stochastic disturbance following a skewed generalized normal distribution (GND) of the following kind (with zero mean):

$$f_{GND}(0,\sigma,k) = \frac{\phi(y)}{\sigma - kP_t}, \tag{5}$$

$$y = \begin{cases} -\frac{1}{k}\log\left[1 - \frac{kP_t}{\sigma}\right], & k \neq 0 \\ \frac{P_t}{\sigma}, & k = 0 \end{cases},$$

where σ is the scale parameter (the variance of the distribution), k is the skewness parameter, ϕ is the density function of a standard normal distribution. If $-1 \leqslant k < 0$, then the distribution is right-skewed and left-truncated (we expect this to occur in s-regimes); if $0 < k \leqslant 1$ then the distribution is left-skewed and right-truncated (which should occur in d-regimes); if $k = 0$ then the distribution is the standard normal distribution. The specification of our model in s- and d-regimes is as follows:

$$P_{t,r} = \mu_r + \epsilon_{t,r}, \tag{6}$$

$$\epsilon_{t,r} \sim GND(0, \sigma_r, k_r),$$

where $r = \{s, d\}$ is the regime of price process, μ_r is the price mean in regime r. For the mean values we use the 5%-th percentile of the sample (for downward outliers) and the 95%-th percentile of the sample (for upward outliers). Thus, unlike Janczura and Weron (2010), we impose fewer restrictions on the parameters of the model, allowing the price in an outlier regime to be distributed around a given deterministic mean, and the price distribution to be skewed. To estimate the parameters of a Markov regime-switching model, we use the maximum likelihood routine implemented in a MATLAB library MS_Regress (Perlin, 2012) and modified by us to include a skewed GND.

2.5. Simulation experiment design

The main problem occuring when comparing different methods of trend-cyclical component filtering from actual data[5] is that the "true" trend is unobserved, and thus, there is no possibility to test the methods against that "true" trend.

In order to overcome this problem we use a simulation approach analogous to the one proposed earlier in Janczura et al. (2013) and Weron and Zator (2014). The design of our simulation experiment is as follows.

1. Estimate the trend-cyclical and stochastic components by using each of the methods under comparison (W_6, W_7, W_8, EMD, CEEMDAN; we call them the "input" methods).

2. Model the stochastic component using the approach described in 2.4, and then estimate the parameters of the model.

3. Generate $N = 500$ trajectories of the stochastic part with the parameters estimated at the previous step, and add them to the previously obtained trend-cyclical components in order to get 500 "artificial" trajectories.

4. Extract the trend-cyclical components from these "artificial" trajectories using each of the methods under comparison (W_6, W_7, W_8, EMD, CEEMDAN; we call them the "output" methods).

5. Calculate the error of the "output" methods by obtaining the bootstrapped medians of the distributions of RMSE (root mean square error), calculated for the original and estimated trends. We use 5000 re-samplings for the bootstrap procedure.

It can be seen that here we have modified the approach of Janczura et al. (2013), using the bootstrapped median of a root mean square error (RMSE) distribution instead of the mean of a mean squared error (MSE) distribution. This bootstrapped median, in our opinion, is more natural and intuitive than the mean of MSE, as well as more resistant to outliers in the sample.

3. Data

In this study we consider the results of functioning of Russian wholesale electricity market during the period from February, 10, 2011, to December, 31, 2013 in two major price zones: Europe-Ural (the first zone) and Siberia (the second zone). The functioning of a competitive electricity day-ahead market (DAM) was established in these zones in 2006.

The electricity DAM in Russia is a mechanism of competitive selection (auction) of price claims of electricity suppliers and buyers one-day-ahead of electricity delivery with determination of the prices and the supply volumes for each hour of the day. The selection is organized by a commercial operator (an open joint stock company "ATS").

[5]By "actual data" we mean the data on actual financial indicators such as electricity prices, commodity demand, security prices, etc., but not artificially generated time-series where the trend is a priori given. It is worth noting that in previous studies the comparison of EMD with other methods was performed on simulated data only, which in general is simpler than using actual data since it does not require choosing specific approaches to take the peculiarities of actual data into account.

Table 1: Descriptive statistics of electricity prices for the period 10.02.2011–31.12.2013 in Europe-Ural and Siberia zones.

Statistic	Europe-Ural zone	Siberia zone
Average	6.903 (1001.6)	6.466 (650.4)
Median	6.896 (988.1)	6.494 (661.4)
Standard deviation	0.113 (115.4)	0.153 (97.1)
Coefficient of variation	1.6%	2.4%
Skewness	0.240	-0.377
Excess kurtosis	-0.268	-0.529

Note: For some of the statistics, in addition to the values of logarithmic prices, the parentheses contain the values of the initial levels of prices.

It should be noticed that there exists a margin pricing on Russian electricity DAM. This means that electricity price is determined via equalizing electricity demand and supply, which is fair for each market participant. As an auction result, there appears a single equilibrium electricity price which is the highest of the prices at which producers are willing to meet the demand. The price indexes and the volumes traded at the DAM are published daily on the website of OJSC "ATS" (`http://www.atsenergo.ru`).

The descriptive statistics of the electricity prices sample for both zones are given in Table 1, while the dynamics of these prices are shown in Fig. 1. It can be seen that the average logarithmic price in Europe-Ural zone is higher than the one in Siberia zone by approximately 7% (by 54% for the initial levels of prices). This reflects the fact in the technological structure of Siberian energy industry there dominates a comparatively cheaper hydroelectric power generation (which accounts for about 40% of the total energy generation in this zone) while in Europe-Ural zone most of the energy is generated by a more expensive thermal power generation (about 69% of the total zone generation).

This also causes the differences in the structure of electricity consumption in the two zones. In Siberia zone, a significant part of electricity consumption is attributed to aluminum production that gives about 90% of the total aluminum produced in Russia. This concentration of aluminum production in the region is primarily caused by cheaper electricity which, depending on the technological process, makes up 25%–40% of expenditures in the cost of aluminum.

Both the standard deviation and the coefficient of variation of electricity price in Siberia zone are higher than in Europe-Ural zone indicating that the prices are more volatile in the former zone. Though, the coefficient of variation in both zones is not that large: the relative variation of prices in Europe-Ural zone is 1.6%, while in Siberia zone it is 2.4%. Thus, the deviation of electricity price from its average is not that substantial in the studied period. Still, despite this fact, even a visual analysis of Fig. 1 shows that there are significant outliers in the electricity price time-series, which should be taken into account in subsequent modeling of price behavior.

The skewness of electricity price distribution for Europe-Ural zone is positive and is equal to 0.240. This suggests that the electricity prices here typically deviate to values lower than the average. Therefore, when modeling the stochastic part of the price time-series, we can expect a tendency for "drops" to be discovered in this zone. At the same time, in Siberia zone the situation is the opposite: the skewness is negative and is equal to -0.377, indicating the prevalence of (positive) price "peaks". The excess kurtosis is negative in both zones, which also indirectly confirms the presence of outliers in the sample.

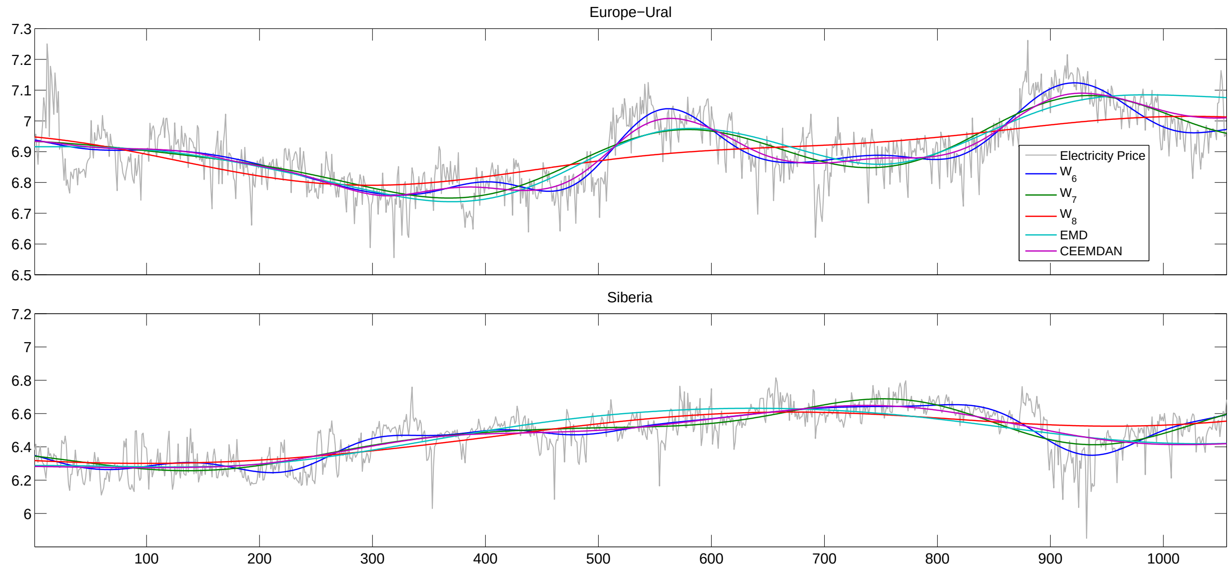

Figure 1: Europe-Ural and Siberia electricity prices and their trend-cyclical components obtained by the five methods under consideration: W_6, W_7, W_8, EMD, CEEMDAN.

Given the significant difference in consumption and generation technologies, as well as minor amounts of electricity flows between the two zones, these zones can be considered as two independent markets. This allows us to perform our numerical experiment more objectively, since the pricing processes in these two markets are sufficiently different, while the impact of their specific features can be identified as a result of comparison of trend-cyclical component filtering methods.

4. Results

4.1. Trend-cyclical and stochastic components filtering

In this section we analyze the obtained results. Fig. 1 shows the trend-cyclical components of the electricity prices obtained by the five considered methods: W_6, W_7, W_8, EMD, CEEMDAN. It can be clearly seen that for Europe-Ural zone the application of W_8 results in a considerably lower frequency component as compared to the results of the other methods. At the same time, W_6 and CEEMDAN, as well as W_7 and EMD, in a pair-wise manner demonstrate quite close estimates of the trend component. Taking this into account, we may expect their mean squared errors to be minimal. For the electricity price in Siberia zone, most methods (except W_6) give similar results over almost all the time period under study. This is why we cannot assume anything about the errors of trend filtering in this case.

Now we briefly consider the results of electricity price stochastic part modeling using CEEMDAN with a three-state Markov regime-switching model. Table 2 provides the estimates of the parameters of this model for the two zones, as well as the probabilities of staying in each of the regimes r between time moments t and $t+1$. All the parameter estimates are statistically significant at least at the 10% level of significance, except coefficient K in Siberia zone. The results show that in Europe-Ural zone there is a significant price self-dependence on both the one-day and seven-day lags in the base regime. The extent of these lags' influence is approximately equal (the coefficients ψ_1 and ψ_7 are approximately the same, taking into account their standard errors). In

Table 2: Parameter estimates of the Markov regime-switching model for electricity price's stochastic component obtained after trend-cyclical component filtering using CEEMDAN.

Parameter	*Base regime*	*Spike regime*	*Drop regime*
Europe-Ural zone			
μ	0.003^{*} (0.002)	0.074	-0.082
ψ_1	0.349^{***} (0.028)	0	0
ψ_7	0.402^{***} (0.024)	0	0
σ	0.041^{***} (0.010)	0.052^{*} (0.029)	0.052^{**} (0.024)
K	0	-0.603^{***} (0.227)	0.237^{*} (0.143)
p_{rr}	0.97^{***} (0.03)	0.78^{***} (0.15)	0.77^{***} (0.10)
Siberia zone			
μ	0.004^{*} (0.002)	0.185	-0.200
ψ_1	0.650^{***} (0.024)	0	0
ψ_7	0.099^{***} (0.021)	0	0
σ	0.053^{***} (0.012)	0.124^{**} (0.057)	0.134^{**} (0.067)
K	0	0.399 (0.286)	0.158 (0.125)
p_{rr}	0.97^{***} (0.03)	0.65^{***} (0.14)	0.71^{***} (0.11)

Note: Standard errors of coefficients are given in parentheses. The coefficients without standard errors are fixed in the model. The levels of significance notation: *** – 1%, ** – 5%, * – 10%. p_{rr} is the probability of regime r between time moments t and $t+1$.

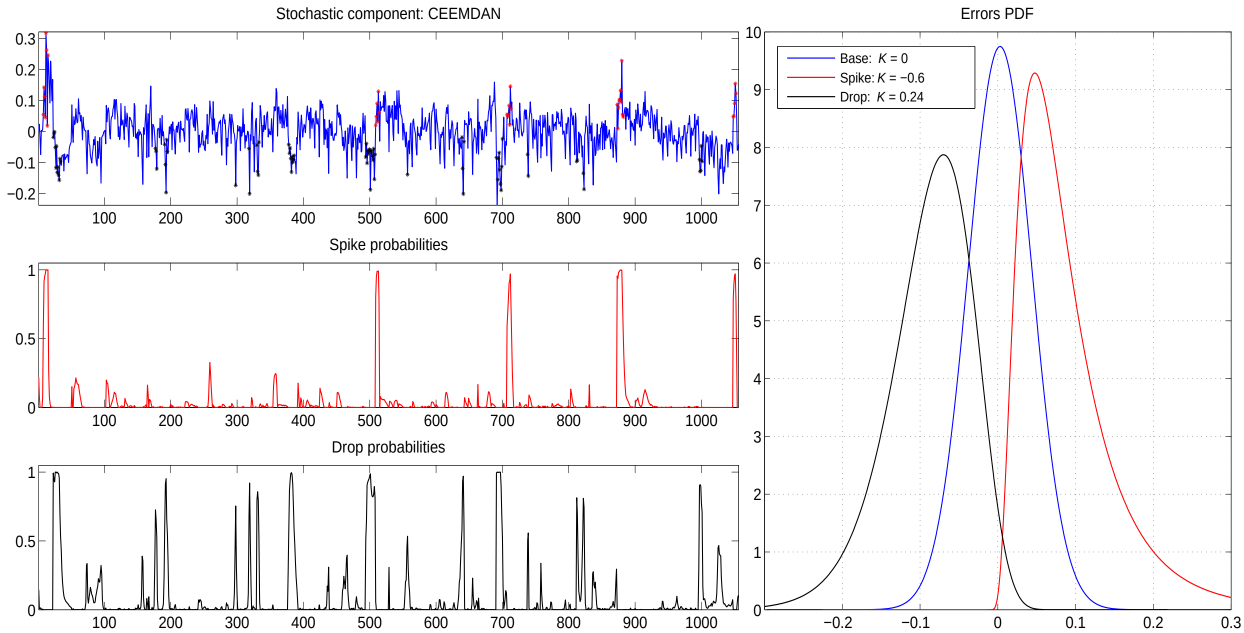

Figure 2: The results of modeling of Europe-Ural electricity price's stochastic component obtained after removing its trend-cyclical component estimated with CEEMDAN. The left panel (top to bottom) shows the stochastic component and the specially marked outliers (red asterisks are "spikes", black asterisks are "drops"), as well as the probabilities of positive and negative outliers. The right panel shows the density graphs of model errors in different regimes.

Table 3: Comparison of errors of trend-cyclical component filtering by the methods under study. For EMD and CEEMDAN, the number of IMFs included in the trend component is determined based on the four criteria: energy, zero-crossing ratio, statistical, and low-frequency.

Output	Input				
	W_6	W_7	W_8	EMD	$CEEMDAN$
Europe-Ural					
W_6	0.0	*29.7*	50.4	*25.8*	0.0
W_7	66.9	0.0	*22.0*	0.0	*17.0*
W_8	179.5	135.7	0.0	65.4	115.4
EMD	105.3	42.9	32.7	*28.2*	38.4
$CEEMDAN$	113.6	*28.4*	*8.2*	*20.3*	37.0
Siberia					
W_6	0.0	35.5	67.7	74.1	41.3
W_7	61.8	0.0	*24.6*	*28.1*	*5.8*
W_8	157.4	153.2	0.0	0.0	33.5
EMD	138.0	121.6	*29.6*	*24.5*	*29.4*
$CEEMDAN$	99.6	70.2	*26.2*	*11.7*	0.0

Note: The columns correspond to the input methods of trend filtering; the rows correspond to the output methods. The measure of deviation is a bootstrapped median of distribution of RMSE between the initial trend and its estimate calculated for each artificial trajectory. For illustrative purposes, the table contains percentage deviation of absolute values of errors from their minimum for each of the input methods (the minimum value in each column). The deviations of less than 30% are given in italics.

Siberia zone, the dependence of the current electricity price on its one-day lag is much stronger than on its seven-day lag (ψ_1 is 6.5 times greater than ψ_7), which speaks in favor of a weak weekly seasonality in the electricity prices in that zone.

The parameters of outlier regimes in Europe-Ural zone are very close to each other except for coefficient K which reflects the skewness of price distributions in these regimes. Specifically, $K_s = -0.603$ which speaks in favor of a quite strongly right-skewed distribution. It is clearly seen in Fig. 2 on its right panel giving the graphs of the obtained distributions. In "drop" regime $K_d = 0.237$, and there is a left-skewed distribution, though the degree of skewness is less than for the "spike" regimes s. This is why, having the same variance as for the "spike" regimes, the distribution in "drop" regime d is more wide-spread, and as a result there are more observations that are classified as "drops". This also coincides with our assumptions made earlier when carrying out the analysis of skewness of the price distribution in Europe-Ural zone (see Table 1). In Siberia zone the skewness parameters for both regimes are not significant, thus, the price distributions in these regimes are better captured by a standard normal distribution rather than by a skewed generalized normal distribution (GND). We do not dwell here on the results obtained for the corresponding stochastic parts since this is not the primary objective of our research. We now switch to comparison of trend-cyclical component filtering methods.

4.2. Performance comparison of trend-cyclical component filtering methods

Table 3 shows percentage deviations of absolute values of errors (the bootstrapped median of distribution of RMSE) from their minimum values (equal to 0) for each of the input methods (given in the columns of the table; the rows correspond to the output methods). The deviations of less than 30% are given in italics. Though this latter value may seem quite large, still, in our opinion, it is quite adequate as a threshold, taking into account the spread of errors obtained for different methods of trend filtering.

It can be seen from Table 3 that in the case of wavelet-decomposition, for each considered scale parameter m, the coincidence of the output and the input methods results in the best performance, which is to some extent expected. At the same time, this is not always true for EMD and CEEMDAN. Specifically, for input EMD for the Europe-Ural zone prices method W_7 is the best in the sense of trend-estimation error W_7, while for the Siberia zone prices W_8 is the best method. Moreover, for the former zone EMD shows non-satisfactory results for almost all input methods (except for EMD itself), while for the latter zone the results of EMD are non-satisfactory only for two of the input methods. Formally speaking, in 6 out of 10 cases EMD does not give satisfactory results, thus, taking into account such imbalance, we cannot make an unambiguous conclusion about the quality of trend filtering with this method as compared to the wavelet-decomposition. Still, we are inclined to say that the performance of EMD is not well.

CEEMDAN both for the Europe-Ural and the Siberia zone prices in 3 out of 5 cases shows the deviations of less than 30%, which is quite a satisfactory result. Moreover, the deviations are less than 12% for the input methods W_8 (for the former zone) and EMD (for the latter zone), being quite small amongst all the obtained deviations. Even in the cases when CEEMDAN trend-estimation error is non-satisfactory, still, it is within the first two methods following the best output method for a specified input method. In Siberia zone, CEEMDAN has the minimal error when input and output methods coincide.

Comparing the obtained results, it can be concluded that CEEMDAN has the best performance in the sense of trend-filtering quality as compared to EMD. Still, comparing the performance of CEEMDAN with the performance of wavelet-decomposition, we cannot conclude the same. On the one hand, the former method has quite a satisfactory performance in most cases. On the other hand, there is only one time when this method has the best performance among the other methods. In addition, comparing the errors of trend component filtering for the Europe-Ural zone prices (see Fig. 1, the upper panel), which visually look almost identical (W_6 and CEEMDAN, W_7 and EMD), it can be noted that the corresponding wavelet-decompositions (as output methods for the input methods EMD and CEEMDAN) result in minimum errors, while the converse is not true.

In order to understand the rationale for the above-mentioned results for CEEMDAN, we consider 2 additional aspects of our numerical experiment: RMSE distributions and possible dependence between error and the size of samples generated for calculations.

4.3. Analysis of the trend filtering criteria issues

Fig. 3 shows the obtained empirical distributions of RMSE for Europe-Ural zone (as before, the columns correspond to the input methods, the rows correspond to the output methods). It can be seen that for CEEMDAN (in the case of W_6 and CEEMDAN input methods) the distributions have 2 distinct peaks. Visual analysis of trends that are filtered out by these methods from the artificially generated trajectories also shows that there are 2 types of output trends: the first one has a higher frequency and resembles the initial trend (see the top panel in Fig. 1, CEEMDAN method); while the second one has a lower frequency and resembles the input trend obtained by W_8 method.

In our opinion, the rationale for this is that the ratio of zero-crossing number (RZCN) criterion, which we also used to determine the minimal index i^* (starting with which the IMFs are included in trend), is an empirical one and in some cases is not robust. The critical bounds for the "significant deviation from 2" (obtained earlier in Moghtader et al. (2011)) are found by simulation

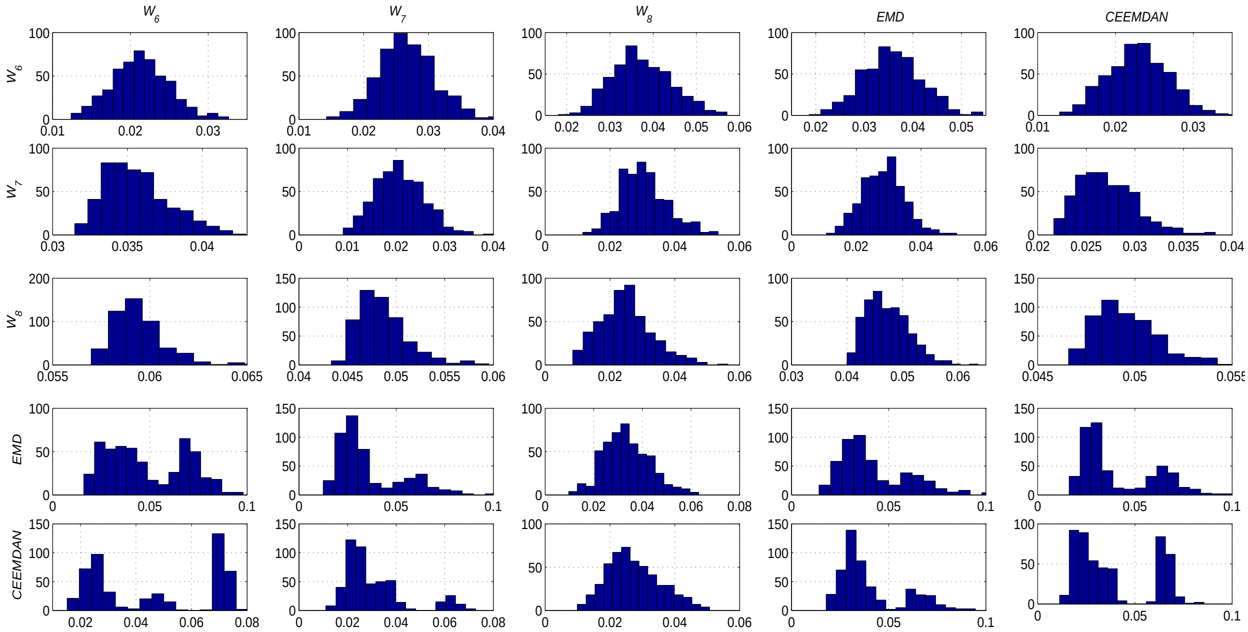

Figure 3: The RMSE distributions of different trend filtering methods for the Europe-Ural zone prices. The columns correspond to the input methods, the rows correspond to the output methods. It can be clearly seen that the distributions for W_6 and CEEMDAN input methods have two peaks.

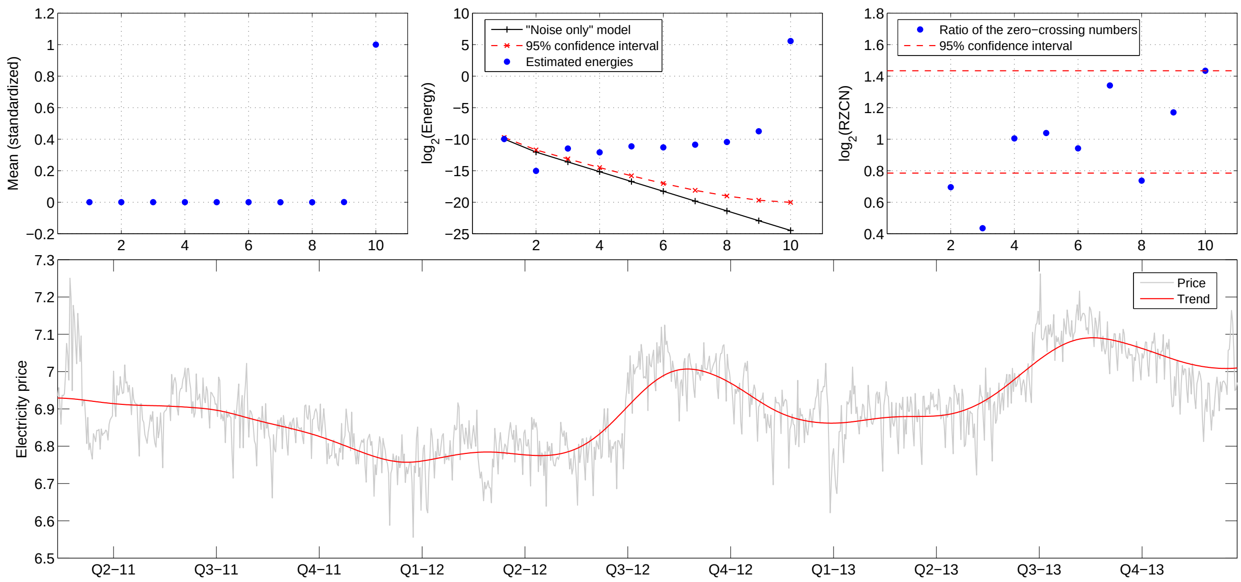

Figure 4: Electricity price trend for Europe-Ural zone. Top panel: the left graph shows the evolution of standardized average of the number of IMFs summed; the middle graph shows energy $\widehat{W}_i^H$ of the IMFs of fractional Gaussian noise with $H = 0.2$ (solid line with markers +), the 95% confidence interval C_i^H (dashed line with markers *) and energy G_i of price IMF (dots); the right graph shows ratio of zero-crossing number R_i (dots) and the 95% confidence interval (dashed lines). Bottom panel: the graph shows the electricity price and the trend for $i^* = 8$.

Table 4: Comparison of errors of trend-component filtering by the methods considered in the study. For EMD and CEEMDAN, the number of IMFs included in the trend component is constant and equal to the number of IMFs constituting the trend for the initial data.

	W_6	W_7	W_8	EMD	$CEEMDAN$
Europe-Ural zone					
W_6	0.0	*29.6*	50.3	*25.8*	*8.3*
W_7	66.9	0.0	*22.0*	0.0	*26.8*
W_8	179.5	135.6	0.0	65.4	133.4
$EMD_{(8/2)}$	121.5	49.0	*24.3*	*24.5*	86.8
$CEEMDAN_{(10/3)}$	*15.6*	*17.2*	31.9	*12.2*	0.0
Siberia zone					
W_6	0.0	35.4	67.7	74.1	33.6
W_7	61.8	0.0	*24.6*	*28.1*	0.0
W_8	157.4	153.1	0.0	0.0	*26.1*
$EMD_{(9/2)}$	146.1	129.0	*30.4*	32.2	*23.0*
$CEEMDAN_{(11/3)}$	101.6	69.2	*27.1*	*19.7*	*0.3*

Note: see Table 3. For the output methods EMD and CEEMDAN the subscript contains: (1) the number of IMFs obtained by decomposition of the initial data, (2) the number of low-frequency IMFs forcibly included in the trend at the step of artificial trajectories processing.

and averaging over 21 types of broadband processes, so they may contain some bias. This results in the fact that in some cases we get $i^* = 8$, and only 3 last IMFs[6] and the residual are included in trend, which causes a higher frequency output trend and a peak in RMSE distribution with a smaller error. In other cases $i^* = 10$, and only the last IMF and the residual are included in trend, which causes a lower frequency output trend and a peak in RMSE distribution with a larger error.

This is illustrated in Fig. 4 which presents the applied criteria for i^* determination for the initial trend in the Europe-Ural zone prices. It can be seen from the top-right graph that RZCN for $i = i^* = 8$ is not that far out from the lower bound, while for $i = 9$ it is within the critical bounds. For $i = 10$, RZCN is infinitely large since the last IMF does not cross zero point, which is why we simply depict this point on the upper bound. Thus, even smallest changes in the initial time-series of electricity prices (that are introduced by us in artificial trajectories) are able to result in $RZCN_8$ being higher than the lower critical bound and the trend comprising only IMF_{10} and the residual. It is intuitively clear that, since the deviation of $RZCN_8$ from the lower bound is very small, for the generated trajectories we get approximately the same number of trends of the two types, because RZCN oscillates around the critical value. It is clearly seen in the RMSE distributions (see Fig. 3), where the heights of 2 peaks are almost equal.

It is worth noting that the trend filtered out from the Siberia zone prices contains only 3 out of 11 IMFs. Moreover, for each of the low frequency IMFs, indicator RZCN appeared to be infinitely large, while the criterion itself is robust for different stochastic components generated when simulating the artificial trajectories. As a result, the distributions of RMSE obtained for Siberia zone do not have distinct double peaks. In our opinion, the unsatisfactory CEEMDAN error values for the input methods W_6 and W_7 are to a large extent caused by the fact that the

[6]When decomposing the initial Europe-Ural zone prices, just as the vast majority of artificially generated trajectories, CEEMDAN results in 10 IMFs.

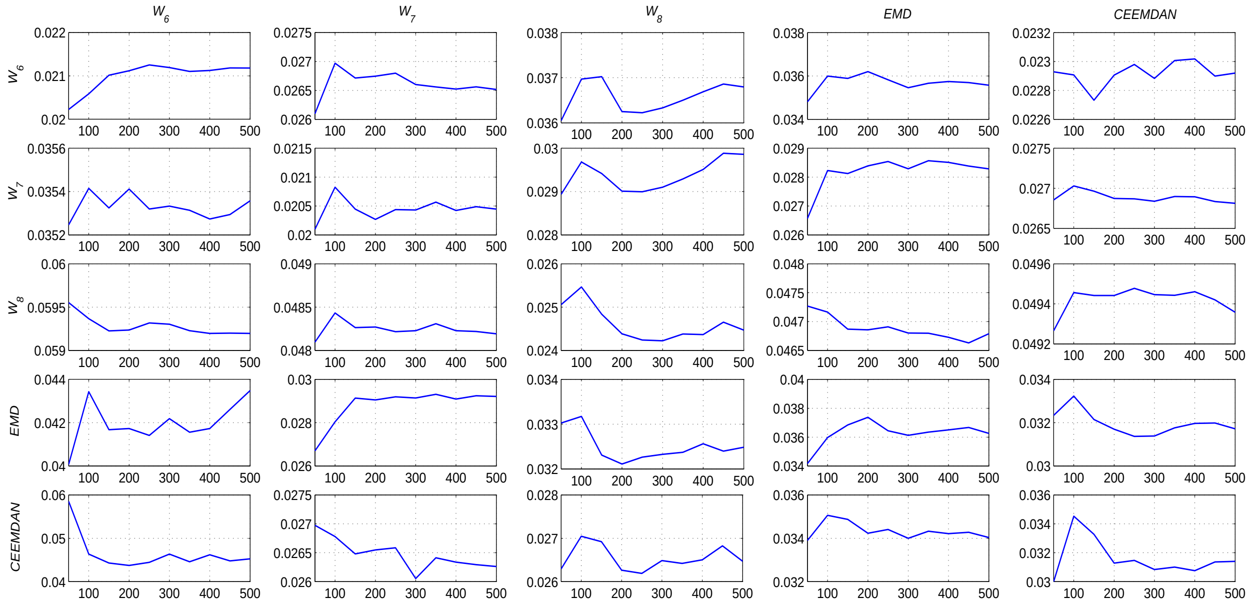

Figure 5: RMSE median changes depending on the artificial trajectories sample size for different methods of trend filtering in Europe-Ural zone. The columns correspond to the input methods, the rows correspond to the output methods.

time-scale of smoothing for these methods is too small for sufficiently smooth long-term fluctuations in the Siberia zone prices (see Fig. 1). This is why the initial trend in such cases is likely to be incorrect and significantly diverges from the one estimated by CEEMDAN.

In order to level RZCN oscillations around the critical value and to understand how this influence the results of methods performance comparison (especially, for Europe-Ural zone), we have modified the algorithm of the experiment. To this end, we determined the number of low-frequency IMFs that were in the trend at the step of initial data processing. Using this number, we repeated the procedure of trend filtering for the artificial trajectories, each time including into the trend this specific quantity of IMFs and the residual of decomposition. Thus, we do not determine index i^* at each iteration, but we always use the initial value corresponding to the initial data. Taking into account the robustness of RZCN criterion for the Siberia zone prices, we do not expect such an approach for this zone to significantly influence the results of methods performance comparison, unlike for Europe-Ural zone.

Table 4 shows the results of this modified algorithm performance. It can be seen that for EMD the situation has not changed significantly, but CEEMDAN shows significantly improved results for the Europe-Ural zone prices. Specifically, for input method W_6, the error is not only satisfactory (15.6%) but also is one of the smallest, while for input method CEEMDAN the error is the smallest. The same is true for the rest of the methods that demonstrate much smaller deviations than the established threshold of 30%. The only exclusion here is method W_8 which at the input gives quite a low-frequency trend which is significantly different from the ones obtained by other methods (see Fig. 1). In Siberia zone, just as expected, we were not able to determine significant changes in methods performance results which confirms robustness of RZCN criterion in this case.

The second aspect that we have taken into account when analyzing the results of our numerical experiment is a possible error's dependence on the size of the sample that was generated for the experiment. Since the number of artificial trajectories used may potentially influence the value of RMSE median, it can be assumed that the above-mentioned problems with CEEMDAN are associated not only with i^* selection criteria, but also with the parameters of the experiment itself. In order to check this, for all the input and output methods we have monitored the dynamics of errors in the range from 50 to 500 trajectories with the increment of 50 trajectories. The obtained results are shown in Fig. 5. It can be seen that the used experiment design is quite robust since the error is practically independent of the sample size, and it is 100 trajectories that are quite enough to reach the established level of error.

For a more objective analysis, we further compared the methods, using for each of them the minimal errors obtained in the above-mentioned range of trajectories (i.e. for each method, as an error we have adopted the minimum point on the graphs shown in Fig. 5). The results show that it has little effect on the general situation with methods performance comparison, only slightly changing the deviations of RMSE medians from the minimum values. Thus, we conclude that the parameters of our experiment are not crucial for the above-detected problem of CEEMDAN usage along with the considered criteria for IMFs selection when performing trend filtering.

5. Conclusion

In this research we study the problems of performance comparison of different trend-cyclical component filtering methods applied to electricity prices. The methods comprise wavelet - decomposition, EMD and CEEMDAN. In order to determine the number of low-frequency IMFs that are to be included in trend, applying of the latter two methods, we used both 2 criteria (energy and ratio of zero-crossing number) proposed earlier in Moghtader et al. (2011), and introduced by us the "low-frequency" criterion, as well as took into account the requirement of statistical significance of IMFs (Afanasyev et al., 2014). As the empirical basis for testing these methods, we used data from the day-ahead electricity market in 2 Russian major electricity price zones: Europe-Ural zone and Siberia zone.

Our results allow to conclude that the ordinary EMD does not provide satisfactory quality of trend estimation as compared to the wavelet-decomposition. In 6 out of 10 cases (4 for Europe-Ural zone and 2 for Siberia zone), its application results in error deviating from the minimum error by more than 30% for a specified input method of trend filtering. Application of complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) shows a much better result than EMD: in 6 out of 10 cases (3 in each zone) the deviation of error from the minimal error is less than 30%, while in the rest 4 cases it is among the top two smallest deviations. Still, when comparing CEEMDAN with wavelet-decomposition, the results are not that unambiguous. Unlike wavelet-decomposition which consistently produces best results when input and output methods coincide, CEEMDAN provides the same result only in Siberia zone, while in Europe-Ural zone it loses to the other two methods.

The analysis of these results shows that the main difficulty here is attributed to the ratio of zero-crossing number criterion. Its empirical nature and specific critical bounds, calculated in Moghtader et al. (2011) by averaging a limited number of certain types of broadband processes, result in the fact that even with a "sensible noising"[7] the value of RZCN may oscillate around the

[7]An artificial stochastic component was generated on the basis of parameters obtained within its preliminary modeling by a 3-state Markov regime-switching model.

critical value. This, in turn, makes the number of IMFs included in the trend component differ for different artificial trajectories, and, as we showed above, the distribution of RMSE acquires two distinct peaks.

In order to level this effect, we fixed the number of IMFs included in trend and made it equal to the number of IMFs found when processing the initial data, after which we re-ran the experiment. The results showed that the performance of EMD did not principally changed, while the performance of CEEMDAN improved significantly: in 7 out of 10 cases the deviation of error from its minimum value appeared to be significantly less than 30%.

Comparing the results for the two zones, we can assume that CEEMDAN has better performance for the time-series that visually demonstrate the presence of a potentially lower frequency trend-cyclical component (Siberia zone, see the bottom panel of Fig. 1), than the time-series with distinct oscillation in their trends (Europe-Ural zone, see the top panel of Fig. 1). Still, a thorough check of this assumption requires a separate research.

In general, the obtained results allow us to conclude that CEEMDAN, along with the considered criteria, performs much better than the ordinary empirical mode decomposition (EMD) and is comparable to wavelet-decomposition within the trend filtering framework. Moreover, CEEMDAN has important advantages: it demonstrates a high degree of adaptation to data, while not requiring specific a priori assumptions, as well as is able to deal with non-stationary and nonlinear processes. All this allows us to expect that the thus filtered trend-cyclical component is likely to better correspond to the "true" one, rather than the trend-cyclical components filtered by the methods without such properties. At the same time, the problems detected for the ratio of zero-crossing number criterion leave us a number of open question when applying it to actual data, and thus require further research.

6. Acknowledgements

The authors are grateful to Evgenii V. Gilenko, PhD, St. Petersburg State University, Russia, for his fruitful comments on the text of the paper. Responsibility for all errors and inaccuracies made by the authors rests solely on the part of the authors.

References

Afanasyev, D., Fedorova, E., Popov, V., 2014. Fine structure of the price-demand relationship in the electricity market: multi-scale correlation analysis.
URL http://mpra.ub.uni-muenchen.de/58827/

An, N., Zhao, W., Wang, J., Shang, D., Zhao, E., 2013. Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting. Energy 49, 279–288.

Bordignon, S., Bunn, D.W., Lisi, F., Nan, F., 2013. Combining day-ahead forecasts for british electricity prices. Energy Economics 35, 88–103.

Colominas, M., Schlotthauer, G., Torres, M., Flandrin, P., 2012. Noise-assisted EMD methods in action. Advances in Adaptive Data Analysis 4 (4).

De Jong, C., 2006. The nature of power spikes: a regime-switch approach. Studies in Nonlinear Dynamics and Econometrics 10 (3).

Fanone, E., Gamba, A., Prokopczuk, M., 2013. The case of negative day-ahead electricity prices. Energy Economics 35, 22–34.

Flandrin, P., Goncalves, P., Rilling, G., 2004. Detrending and denoising with empirical mode decomposition. EUSIPCO.

Ghelardoni, L., Ghio, A., Anguita, D., 2013. Energy load forecasting using empirical mode decomposition and support vector regression. IEEE Transactions on Smart Grid 4 (1), 549–556.

Haldrup, N., Nielsen, F., Nielsen, M., 2010. A vector autoregressive model for electricity prices subject to long memory and regime switching. Energy Economics 32, 1044–1058.

Hamilton, J., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357–384.

Hamilton, J., Lin, G., 1996. Stock market volatility and the business cycle. Journal of Applied Econometrics 11, 573–593.

Hodrick, R., Prescott, E., 1997. Postwar U.S. business cycles: An empirical investigation. Journal of Money, Credit and Banking 29 (1), 1–16.

Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., Yen, N., Tung, C., Liu, H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454 (1971), 903–995.

Hurst, H., 1951. The long-term storage capacity of reservoirs. Transactions of the American society of civil engineers 116, 770–799.

Ismail, Z., 2013. A new approach to peak load forecasting based on EMD and ANFIS. Indian Journal of Science and Technology 6 (12), 5600–5606.

Janczura, J., Truck, S., Weron, R., Wolff, R., 2013. Identifying spikes and seasonal components in electricity spot price data: A guide to robust modelings. Energy Economics 38, 96–100.

Janczura, J., Weron, R., 2010. An empirical comparison of alternate regime-switching models for electricity spot prices. Energy Economics 32, 1059–1073.

Kurbatsky, V., Tomin, N., 2010. Forecasting prices in the liberalized electricity market using the hybrid models. In: Energy Conference and Exhibition (EnergyCon), 2010 IEEE International. pp. 363–368.

Lucia, J., Schwartz, E., 2002. Electricity prices and power derivatives: evidence from the Nordic Power Exchange. Review of Derivatives Research 5 (1), 5–50.

Mhamdi, F., Jadane-Sadane, M., Poggi, J.-M., 2010. Empirical mode decomposition for trend extraction: application to electrical data. 19th International Conference on Computational Statistics.

Moghtader, A., Borgnat, P., Flandrin, P., 2011. Trend filtering: empirical mode decomposition versus l_1 and Hodrick-Prescott. Advances in Adaptive Data Analysis 3 (1 and 2), 41–61.

Nowotarski, J., Tomczyk, J., Weron, R., 2013. Robust estimation and forecasting of the long-term seasonal component of electricity spot prices. Energy Economics 39, 13–27.

Perlin, M., 2012. MS Regress - The Matlab package for Markov regime switching models. Available at SSRN.
URL `http://ssrn.com/abstract=1714016`

Peters, E., 1991. Chaos and order in the capital markets: a new view of cycles, prices, and market volatility. New York: Wiley.

Trück, S., Weron, R., Wolff, R., 2007. Outlier treatment and robust approaches for modeling electricity spot prices. Proceedings of the 56th Session of the ISI.
URL `http://mpra.ub.uni-muenchen.de/4711/1/MPRA_paper_4711.pdf`

Veraart, A., Veraart, L., 2012. Modelling electricity day-ahead prices by multivariate Lévy semi-stationary processes. CREATES Research Paper 2012-13. Aarhus University.

Weron, R., Bierbrauer, M., Trück, S., 2004. Modeling electricity prices: jump diffusion and regime switching. Physica A: Statistical Mechanics and its Applications 336, 39–48.

Weron, R., Zator, M., 2014. A note on using the Hodrick-Prescott filter in electricity markets. HSC Research Reports HSC/14/04, Hugo Steinhaus Center, Wroclaw University of Technology.
URL `http://prac.im.pwr.wroc.pl/~hugo/RePEc/wuu/wpaper/HSC_14_04.pdf`

Wu, Z., Huang, N., 2009. Ensemble empirical mode decomposition: A noise-assisted data analysis method. Advances in Adaptive Data Analysis 1 (1), 1–41.