

# Modeling Long-term Seasonality and Spikes of the Spot Electricity Prices in Turkey

Erdinç Akyıldırım\*

Akdeniz University

Ahmet Göncü\*\*

Xi'an Jiaotong Liverpool University

Albert Altarovici\*\*\*

ETH Zürich

## Abstract

Stochastic models of electricity spot prices depend on price spikes and long-term seasonality. Therefore it is crucial to determine suitable methods for the identification of price spikes and the modeling of long-term seasonal components (LTSC). Following recent studies (Janczura and Weron, 2010; Janczura et al., 2013), we compare the proportion of observations identified as outliers for five different outlier detection methods and three approaches to long-term seasonality modeling. After removing the effects of outliers, we compare the out-of-sample forecasting performance for three categories of long-term seasonality models: dummies, Fourier series, and wavelet-based methods. We consider various combinations of each approach and perform a comprehensive backtesting comparison at different forecasting horizons for the recently liberalized Turkish electricity market.

*Keywords: electricity price spikes, long-term seasonality modeling, Turkish electricity prices, wavelets.*

## Türkiye Elektrik Piyasasında Elektrik Fiyatlarındaki Uzun Dönemli Mevsimselliklerin ve Ani ve Ciddi Hareketlerin Modellenmesi

### Özet

Spot elektrik fiyatlarının stokastik modellenmesi fiyatlardaki ani ve ciddi yükselişlere (ve düşüşlere) ve uzun dönemli mevsimselliklere bağlıdır. Bu nedenle ani ve ciddi fiyat yükselişlerini (ve düşüşlerini) tespit edecek metotların belirlenmesi ve uzun dönemli mevsimselliklerin bileşenlerinin modellenmesi büyük önem arz etmektedir. Bu çalışmada, Janczura ve Weron (2010) ve Janczura ve diğ. (2013) makalelerindeki yöntem takip edilerek beş değişik aykırı gözlem belirleme metodunun sonuçları değerlendirilmiştir. Ayrıca aykırı gözlemlerin etkileri giderildikten sonra, üç değişik mevsimsellik yöntemine (kukla değişkenler, Fourier serileri, dalgacık tabanlı yöntemler) göre örneklem dışı gözlemlerin performansı karşılaştırılmıştır. 2009-2013 yıllarında arasında Türkiye'deki spot elektrik fiyatlarını kullanan ve bu alanda ilk olan çalışmamızda, elektrik fiyatlarının doğru olarak modellenerek ekonomiye katkı sağlanması hedeflenmiştir.

*Anahtar kelimeler: elektrik fiyatlarındaki ani ve ciddi yükseliş ve düşüşler, uzun dönemli mevsimsellik modellenmesi, Türkiye elektrik fiyatları, dalgacık methodu.*

\* Erdinç Akyıldırım is an Assistant Professor in the Department of Banking and Finance at Akdeniz University, Dumlupınar Bulvarı, 07058, Turkey. E-mail: erdincakyildirim@akdeniz.edu.tr

\*\* Ahmet Göncü is an Assistant Professor in the Department of Mathematical Sciences at Xi'an Jiaotong Liverpool University, Suzhou, China, E-mail: ahmet.goncu@xjtlu.edu.cn

\*\*\* Albert Altarovici is a PhD student in the Department of Mathematics at ETH Zurich, Raemistrasse 101, 8092, Zurich, Switzerland. E-mail: albert.altarovici@math.ethz.ch

Electricity is a non-storable commodity with consumption dependent on temperature, weather conditions, and various economic factors that might be cyclical. Electricity supply is also affected by power plant outages or problems related to transmission grids or lines. In liberalized electricity markets, such as in the USA and in Europe, spot prices are determined by supply and demand. In this study we focus on the Turkish electricity market which was liberalized in 2009. Active trading commenced in December 2009.

According to TEIAS (Turkish Electricity Transmission Company), Turkey's electricity generation capacity rose to 56GW at the end of 2012, up from 42GW in 2008, and electricity generation capacity is expected to reach 66GW by 2015. Per capita electricity consumption in Turkey is approximately 3000kWh (see Hamzacebi and Es, 2014). This is less than the average consumption in OECD countries which is near 8000kWh. Between 2004 and 2013, the Turkish economy grew on average by 4.9%, while electricity consumption grew at a faster rate of 5.7% (Source: Turkish Statistical Institute). Over the last ten years, Turkey has seen the highest increase in electricity consumption of all European countries, and places second in the world after China in this category. This shows that the Turkish economy has an important potential for further growth in its electricity generation capacity in order to meet the increasing energy demands of its expanding economy. In an effort to optimize the use of its existing electricity generation capacity, the Turkish government introduced in 2009 the electricity spot trading system which has led to the determination of spot electricity prices one day ahead in the market.

To effectively manage existing resources, it is crucial to have reliable forecasts of electricity prices (Weron, 2006). The regulatory issues regarding the energy market in Turkey are discussed by Çetin and Oğuz (2007), Bağdadioğlu and Odyakmaz (2009), and Erdoğan (2010). Although there are many studies that model and explain the electricity consumption and demand in Turkey (e.g. Altınay and Karagöl, 2005; Hamzacebi and Es, 2014; Akay and Atak, 2007; Hamzacebi, 2007; Ediger and Tatlıdil, 2002; Bölük and Koç, 2010), there are very few studies which model electricity spot prices after the establishment of the spot market in 2009. Gökgöz and Atmaca (2012) use spot electricity prices to determine optimal electricity generation asset allocation. To the best of our knowledge, the only studies that model spot prices are provided by Azize and Talaslı (2014), Stevenson et al. (2006), and Nowotarski et al. (2013).

Electricity prices exhibit strong seasonality, mean reversion, and spikes (as is documented in Azize and Talaslı, 2014; Weron, 2008; Küçükali and Barış, 2010), and stochastic models with mean reversion, spikes, and seasonality are commonly utilized. Another approach is the use of regime switching models as in Weron (2008). Furthermore, stochastic models allow for the consistent pricing of derivatives instruments written on spot electricity prices. The studies provided by Küçükali and Barış (2010), Azize and Talaslı (2014), and Benth et al. (2012) consider mean reverting diffusion processes with seasonality and spikes. Multi-factor stochastic modeling of Turkish electricity spot prices can be found in Azize and Talaslı (2014). The main methodologies used in electricity price forecasting have been reviewed by Aggarwal et al. (2009). They also present the

application of various models on different electricity markets.

The choices of method for detecting price spikes and modeling long-term seasonal components have strong consequences for the parameter estimates and the fit of the stochastic models. There are different methodologies in the literature to model the trend of the seasonal component of electricity spot prices. One commonly used simple method is the application of piecewise constant functions (or dummies) for the months (see Fanone et al., 2013; Fleten et al., 2011; Haldrup et al. 2010; Higgs and Worthington, 2008). Another method is the use of sums of sinusoidal functions of different frequencies (see Keleş et al., 2012; Erlwein et al., 2010; Koopman et al., 2007; Cartea and Figueroa, 2005). Another way to model the long term seasonal component is the application of wavelet decomposition and smoothing techniques (see Conejo et al., 2005; Janczura and Weron, 2010, 2012; Schlueter, 2010; Stevenson, 2001; Stevenson et al., 2006; Weron, 2006, 2009; Weron et al., 2004a,b).

In this study we consider a wide range of outlier detection methods and long-term seasonality models which have not yet been considered for Turkish spot electricity prices. Based on our analysis, we then provide recommendations for the modeling of spot electricity prices in Turkey.<sup>[1]</sup> Our findings can also provide guidance for other energy markets that were recently liberalized or which are in the process of liberalization.

## Data

Our dataset consists of 1509 observations of daily electricity spot prices in Turkey from December 1, 2009 to February 13, 2014. There are two ways to represent electricity prices for stochastic models. We can either model the spot price  $p_t$  directly or the log price time series, i.e.  $\ln(p_t)$ . In the first approach, we have  $p_t = T_t + s_t + X_t$ , where  $T_t$  is the long-term trend component,  $s_t$  is the short term seasonality component, and  $X_t$  is the stochastic component. For the second approach, we write  $p_t = \exp(T_t + s_t + X_t)$ , which is in multiplicative form and ensures non-negativity of spot prices. In some datasets, the spot price of electricity can be negative since it is usually costly for power companies to shut down and restart production. However, though uncommon, we have no negative prices in our dataset. Notwithstanding, we still prefer to use the multiplicative form which guarantees positive prices. The dataset of spot electricity prices, displayed in Figure 1, exhibits spikes and seasonality.

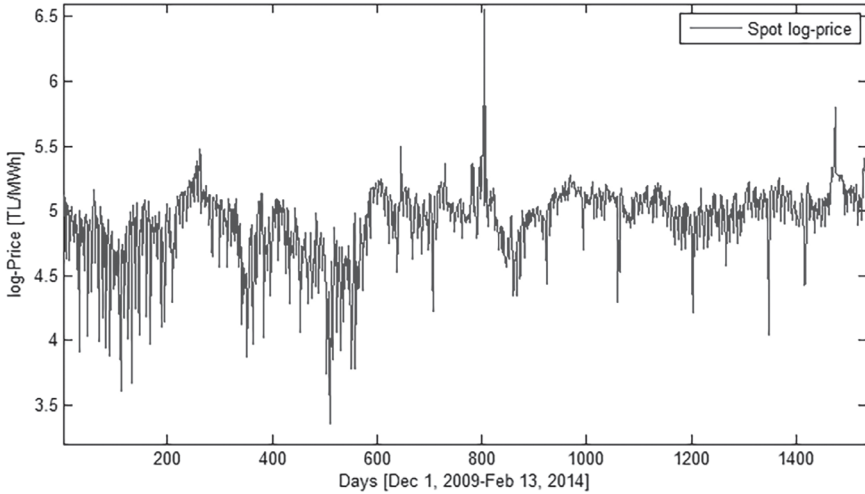
## Detecting Price Spikes

The proper identification of spikes is a crucial part of models for electricity spot prices since spikes can significantly influence the deterministic and stochastic components of a model. In the literature, the following spike identification methods have been applied: (1) fixed price thresholds where all prices exceeding some suitably chosen price level are classified as spikes (e.g. see Boogert and Dupont, 2008); (2) Variable price thresholds

---

<sup>[1]</sup> A large set of MATLAB functions for different outlier and long-term seasonality models are provided by R. Weron on the website: <http://ideas.repec.org/e/pwe42.html>, which we used in our study.

Figure 1  
 Mean Daily Electricity Spot Prices from December 1, 2009 to February 13, 2014



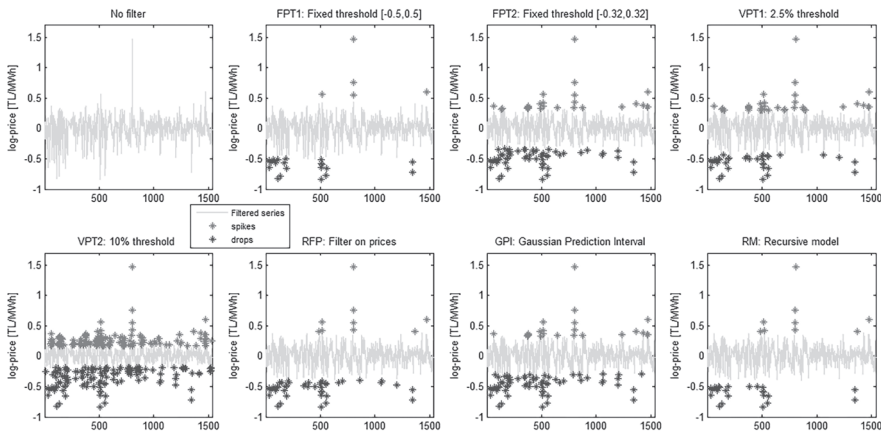
where a certain percentage of the highest (and/or lowest) prices, e.g., the upper 1% of prices are classified as outliers (see Mayer et al., 2012); (3) Fixed price change thresholds where price increments or price returns exceeding some threshold are classified as outliers (Bierbrauer et al., 2004); (4) Variable price change thresholds, more commonly known as the “recursive filter” technique, where prices corresponding to the price increments (or returns) exceeding three standard deviations of all returns are removed one by one in an iterative procedure (c.f. Weron, 2006); (5) Wavelet filtering where the signal (the price series) is first decomposed using the wavelet transform, then reconstructed up to a certain level of detail (c.f. Trück et al., 2007); (6) Thresholds implied by the Gaussian 90% prediction intervals (see Borovkova and Permana, 2006); (7) Thresholds yielding the best model in terms of matching kurtosis (see Geman and Roncoroni, 2006); (8) Markov regime switching model (see Weron, 2008); and (9) Recursive seasonal models where at each step the number of outliers is reduced based on the marginal reduction of the mean squared error (MSE) obtained from fitting a seasonal pattern to the new series (Janczura and Weron, 2010).

In this study we apply the methods (1), (2), (4), (6), (9). We did not apply (3) since we are not working with log-returns or price returns for our time series. We also did not employ (5) since we eliminate long term seasonality with wavelet filtering. The other methods applied in this paper are the most common methods used in the literature. In Janczura and Weron (2010), which is the inspiration for this study, they employ fixed price thresholds, variable thresholds, recursive filters, recursive seasonal and Markov regime switching methods for the identification of spikes. Since all of the above methods are explained in detail and our procedure is exactly the same as in Janczura and Weron (2010) we only present our results.

The identification of spikes has a direct impact on the long-term and the stochastic components. If the outlier detection method produces too few spikes, then the forecasting performance of the long-term mean component will be poor and will fluctuate more often, thus overshooting or undershooting the correct level. However, if too many spikes are observed, the long-term component will be too smooth and the stochastic component will underestimate the variation in prices.

Figure 2

**Detecting price spikes/outliers using the following methods: Fixed Price Threshold; Variable price threshold; Recursive filter on prices; Gaussian prediction interval; Recursive seasonal model. The deseasonalization approach used here is a wavelet approximation for the LTSC.**



As can be seen in Table 1 and Figure 2, different outlier detection methods yield significantly different results in terms of both the number of spikes and drops and also the average magnitude of spikes and drops. Figure 2 illustrates the results for a wavelet long-term seasonality model. However as Table 1 indicates, the results are generally the same across the sin-EWMA long-term seasonality models. According to the wavelet LTSC in Table 1, FPT1 (fixed price threshold with deseasonalized log-prices exceeding the range of  $(-0.5, 0.5)$ ) produces the smallest number of price spikes (0.32%); and VPT2 (variable price threshold with 10% highest and 10% lowest deseasonalized log-prices treated as outliers) yields the highest number of price spikes (10.03%). Following VPT2, VPT1 (variable price threshold with 2.5% highest and 2.5% lowest deseasonalized log-prices treated as outliers) classifies 2.47% of the observations as spikes. FPT1 (thresholds determined by the mean-excess function as  $(-0.32, 0.32)$ ) and GPI produce similar results and RFP and RM identify the same number of price spikes. We expect that the model that identified the highest number of the price spikes produced the lowest average value of a spike. For example, VPT2 has the lowest average value of 5.03 while FPT1 has the highest value with 5.72. For the other methods, the average log-price spike lies between TL 5.12 and TL 5.47. The results for price drops are similar to the price spikes.

According to the wavelet based modeling, FPT1 yields the smallest number of price drops while VPT2 produces the highest number of drops. The lowest average log-price drop (3.94) is captured by FPT1 and the highest (4.45) comes from VPT2 as predicted.

Table 1  
**Frequency and Average Magnitude of Price Spikes and Drops  
 under Different Outlier Detection Methods Considered**

	Frequency of price spikes		Average magnitude of log-price spike	
Filter	Wavelet	Sin-EWMA	Wavelet	Sin-EWMA
FPT1	0.32%	0.26 %	5.72	5.96
FPT2	1.76%	3.06 %	5.22	5.21
VPT1	2.47 %	2.47 %	5.12	5.24
VPT2	10.03%	10.03%	5.03	5.11
RFP	0.59%	0.59 %	5.47	5.55
GPI	1.56 %	1.82 %	5.25	5.28
RM	0.59%	18.75%	5.47	5.07
	Frequency of price drops		Average magnitude of log-price drop	
Filter	Wavelet	Sin-EWMA	Wavelet	Sin-EWMA
FPT1	1.43%	2.02 %	3.94	3.98
FPT2	4.17%	4.88 %	4.17	4.21
VPT1	2.47 %	2.47 %	4.01	4.04
VPT2	10.03%	10.03%	4.45	4.42
RFP	2.99 %	2.73 %	4.06	4.06
GPI	4.75%	4.69 %	4.21	4.20
RM	1.63%	17.77%	3.96	4.59

## Modeling Long-Term Seasonality

We decompose the log-spot prices into three components: the stochastic part  $X_t$ , the short term component  $s_t$ , and the long-term seasonality component  $T_t$ . We investigate three approaches to modeling the log-term seasonality  $T_t$ : 1) dummies or piecewise constant functions, 2) sine/cosine functions, and 3) wavelets. As in Janczura et al. (2013), we apply the same set of models to test the forecasting performance of different long-term seasonality models in each of the three approaches.

In the first approach, each month is fitted to the data. Forecasting using this approach is straightforward, however it does not yield a smooth seasonal component and therefore smoothing techniques may be necessary. In the second approach, log-prices are modeled as linear combinations of sine and cosine functions. Sine and cosine functions yield periodic results, however, electricity prices often do not show periodic behavior. Therefore, multiple sine and cosine terms at different frequencies are utilized.



The wavelet decomposition is more robust for outliers and is a less periodic alternative to the Fourier series approach. As in Janczura et al. (2013), we consider a variety of wavelet models that differ from each other in terms of the way the signal is extrapolated before applying the Discrete Wavelet Transform (DWT) and in the choice of the input signal.

The details of these methods are given in Janczura et al. (2013). A total of 152 models are considered using the three approaches described above. We follow exactly the same model codes as given in Table 1 of [25].

In Tables 2 and 3, we present a classification of the models used in this study.

Table 2

**The six digit codes of the 152 models tested in this study, part I. A star (\*) indicates that a certain digit can take one of a few values and is used to represent subgroups of models. A square cup (‘L’) identifies the digit of interest.**

Digit	Value	Meaning
		Simple models (1000*) → 8 models in total
1000L	1	Mean price in the calibration window
	2	Extrapolated linear regression of prices in the calibration window
	3	Median price in the calibration window
	4	Exponential decay to the median with the decay parameter $\lambda = \frac{1}{30}$
	5	Exponential decay to the median with the decay parameter $\lambda = \frac{1}{180}$
	6	Linear decay to the median
	7	Mean based monthly dummies
	8	Median based monthly dummies
		Sines fitted to raw prices (2**00) → 12 models
2L*00	1,2,3,4	Number of sines used to represent the LTSC
2*L00	1	Periods of the given number of sines estimated
	2	Periods set to 1, $\frac{1}{2}$ , $\frac{1}{3}$ , $\frac{1}{4}$ of a year, respectively
	3	Period of the 1st sine estimated, remaining periods set to 1, $\frac{1}{2}$ , $\frac{1}{3}$ of a year
		Sines fitted to spike-filtered prices (3***0) → 24 models
3L*00	1,2,3,4	Number of sines used to represent the LTSC
3*L00	1	Periods of the given number of sines estimated
	2	Periods set to 1, $\frac{1}{2}$ , $\frac{1}{3}$ , $\frac{1}{4}$ of a year, respectively
	3	Period of the 1st sine estimated, remaining periods set to 1, $\frac{1}{2}$ , $\frac{1}{3}$ of a year.
3**L0	1	Spikes replaced by the mean of the deseasonalized prices
	2	Spikes replaced by the upper/lower 2.5% quantiles of the deseasonalized prices
		Wavelets with an exponential decay to the median fitted to raw prices (4**0*) 24 models
4L*0*	1	Daubechies wavelet family of order 12 (db12)
	2	Daubechies wavelet family of order 24 (db24)
	3	Coiets wavelet family of order 2 (coif2)
	4	Coiets wavelet family of order 4 (coif4)
4*L0*	1	$S_6$ approximation level
	2	$S_7$ approximation level
	3	$S_8$ approximation level
4**0L	1	Exponential decay to the median with the decay parameter $\lambda = \frac{1}{30}$
	2	Exponential decay to the median with the decay parameter $\lambda = \frac{1}{180}$

Table 3

The six digit codes of the 152 models tested in this study, part II. A star (\*) indicates that a certain digit can take one of a few values and is used to represent subgroups of models. A square cup ('∩') identifies the digit of interest.

Digit	Value	Meaning
		Wavelets with a linear decay to the median fitted to raw prices (5**00) 12 models
5 ∩ *00	1	Daubechies wavelet family of order 12 (db12)
	2	Daubechies wavelet family of order 24 (db24)
	3	Coiets wavelet family of order 2 (coif2)
	4	Coiets wavelet family of order 4 (coif4)
5 * ∩00	1	$S_6$ approximation level
	2	$S_7$ approximation level
	3	$S_8$ approximation level
		Wavelets with an exponential decay to the median fitted to spike-filtered prices (6****) 48 models
6 ∩ * * *	1	Daubechies wavelet family of order 12 (db12)
	2	Daubechies wavelet family of order 24 (db24)
	3	Coiets wavelet family of order 2 (coif2)
	4	Coiets wavelet family of order 4 (coif4)
6 * ∩ * *	1	$S_6$ approximation level
	2	$S_7$ approximation level
	3	$S_8$ approximation level
6 * * ∩ *	1	Spikes replaced by the mean of the deseasonalized prices
	2	Spikes replaced by the upper/lower 2.5% quantiles of the deseasonalized prices
6 * * * ∩	1	Exponential decay to the median with the decay parameter $\lambda = \frac{1}{30}$
	2	Exponential decay to the median with the decay parameter $\lambda = \frac{1}{180}$
		Wavelets with a linear decay to the median fitted to spike-filtered prices (7***0) 24 model
7 ∩ * * 0	1	Daubechies wavelet family of order 12 (db12)
	2	Daubechies wavelet family of order 24 (db24)
	3	Coiets wavelet family of order 2 (coif2)
	4	Coiets wavelet family of order 4 (coif4)
7 * ∩ * 0	1	$S_6$ approximation level
	2	$S_7$ approximation level
	3	$S_8$ approximation level
7 * * ∩ 0	1	Spikes replaced by the mean of the deseasonalized prices
	2	Spikes replaced by the upper/lower 2.5% quantiles of the deseasonalized prices

## Empirical Results

We compute three measures of forecasting accuracy for each model and for each forecast horizon, namely the mean absolute error (MAE), the mean squared error (MSE), and the mean absolute percentage error (MAPE). Then, we rank the models from 1 to 152 based on the values of MAE, MSE, and MAPE for the various forecasting horizons.



In Table 4, the top twenty models together with the best models from each family are listed according to the three measures of forecast accuracy. We use the indices a, b, c, d, e for MAE; 1, 2, 3, 4, 5 for MSE; and A, B, C, D, E for MAPE to indicate the best five models in terms of each measure.

Table 4

**Top 20 Models According to Each of the Three Forecast Error Measures: MAE in columns 23, MSE in columns 45 and MAPE in columns 67. The best five models in terms of each measure are shown in bold with the index indicating their rank.**

No.	MAE	Model	MSE	Model	MAPE	Model
1	0.1401	<b>10004<sup>a,A</sup></b>	0.0327	<b>64121<sup>c,1,C</sup></b>	2.806%	<b>10004<sup>a,A</sup></b>
2	0.1402	<b>61121<sup>b,2,B</sup></b>	0.0327	<b>61121<sup>b,2,B</sup></b>	2.807%	<b>61121<sup>b,2,B</sup></b>
3	0.1402	<b>64121<sup>c,1,C</sup></b>	0.0328	<b>62121<sup>d,3,D</sup></b>	2.808%	<b>64121<sup>c,1,C</sup></b>
4	0.1402	<b>62121<sup>d,3,D</sup></b>	0.0328	<b>63121<sup>4</sup></b>	2.808%	<b>62121<sup>d,3,D</sup></b>
5	0.1403	<b>62221<sup>e,E</sup></b>	0.0329	<b>44101<sup>5</sup></b>	2.809%	<b>62221<sup>e,E</sup></b>
6	0.1403	41101	0.0329	41101	2.810%	41101
7	0.1403	<b>44101<sup>5</sup></b>	0.0329	<b>62221<sup>e,E</sup></b>	2.811%	<b>44101<sup>5</sup></b>
8	0.1403	42101	0.0329	61221	2.811%	42101
9	0.1404	64221	0.0329	42101	2.812%	64221
10	0.1404	61221	0.0329	<b>10004<sup>a,A</sup></b>	2.813%	61221
11	0.1405	42201	0.0330	43101	2.814%	<b>63121<sup>4</sup></b>
12	0.1405	43101	0.0331	41201	2.815%	42201
13	0.1405	<b>63121<sup>4</sup></b>	0.0331	42201	2.815%	43101
14	0.1406	41201	0.0331	64221	2.817%	41201
15	0.1407	44201	0.0332	64111	2.819%	44201
16	0.1413	63221	0.0333	62111	2.830%	63221
17	0.1415	43201	0.0333	61111	2.835%	62111
18	0.1416	62111	0.0333	44201	2.835%	43201
19	0.1417	61111	0.0334	63111	2.837%	61111
20	0.1417	64111	0.0334	10003	2.837%	64111
71	0.1470	72110	.	.	2.957%	72110
76	.	.	0.0382	72110	.	.
94	0.1494	52200	.	.	3.008%	52200
100	.	.	0.0404	53300	.	.
103	.	.	0.0410	10008	.	.
113	0.1524	10008	.	.	3.054%	10008
114	.	.	0.0471	31210	.	.
115	0.1740	32210	.	.	3.476%	32210
117	.	.	0.0481	21200	.	.
121	0.1778	22200	.	.	3.554%	22200
131	0.1845	10007	0.0559	10007	3.686%	10007

According to MAPE, four out of the top five models are from the 6\*\*\*\* wavelets group with an exponential decay to the median fitted to spike-filtered prices. In particular, these four models differ only in the choice of the wavelet family (61121<sup>b,2,B</sup>, 62121<sup>d,3,D</sup>, 64121<sup>c,1,C</sup>) and in the choice of the approximation level (62221<sup>e,E</sup>). Surprisingly, the best model in terms of MAPE comes from the simple models, i.e. 10004<sup>a,A</sup>. However, it is also consistent with the other wavelet based models in the sense that both models support exponential decay to the median with the decay parameter  $\lambda = 1/30$ .

The first three models ranked according to MSE also appear in the list of MAPE and the fourth best model (63121<sup>4</sup>) is also from the same wavelet family. The fifth model is again from the wavelet family, but fitted to raw prices (without filtering spikes).

The models in the top twenty list are dominated by two sets of models. Eleven of them are from the 6\*\*\*\* family and another eight of them belong to 4\*\*\*\* family according to all three measures, MSE, MAE, and MAPE. The best models from the other families lag far behind the top models, ranking near the middle of the list. The best models from the 7\*\*\*\* and 5\*\*\*\* families are ranked at 71 and 94 for MAE (and MAPE), and 76 and 100 for MSE, respectively. Following them, the median based monthly dummy models rank at 113 for MAE (and MAPE) and 103 for MSE. Sine based models fitted to raw and spike filtered prices (2\*\*\*\* and 3\*\*\*\*) are listed with ranks at 121 and 115 for MAE (and MAPE) and at 117 and 114 for MSE, respectively.

Among the top 20 models, models 6\*\*\*\* with wavelets with an exponential decay to the median provide a better fit to the data if the spikes are replaced by the upper/lower 2.5% quantiles of the deseasonalized prices (models with '2' as the fourth digit). We also observe that for the models 4\*\*\*\* generally  $S_6$  approximation performs better than the  $S_7$  approximation level. It is also clear from Table 4 that for all the models in the top twenty, the models with exponential decay to the median with decay parameter  $\lambda = 1/30$  are far superior to models with decay parameter  $\lambda = 1/180$ .

In Figure 6, we use histograms to plot the number of times models from a given family are ranked in the top five (which roughly corresponds to the top 3.2%), top ten (or 6.5%) and top fifty (or 32.8%) of all 152 models according to  $MAE_h$ ,  $MSE_h$ , and  $MAPE_h$  for each of the six forecast horizons  $h = 1, 2, \dots, 6$ . The expected number of models from a given family is represented by the white bars in Figure 6 with the assumption that all model families are proportionately represented in the "top ..." models. For example, the expected number of the 6\*\*\*\* models in the "top 5" is equal to  $(48/152) \cdot 5 \cdot 6 = 9.47$  where '6' represents the number of forecast horizon  $h$ .

Figure 6 provides results consistent with our observations from Table 4. Models from the 6\*\*\*\* family perform much better than the ones in the other families and significantly better than the number of expected values. Models from the families 4\*\*\*\*, 7\*\*\*\*, 5\*\*\*\*, and 1\*\*\*\* come in order after the 6\*\*\*\* models. There is clear evidence in Figure 6 that sin-based models do not perform well for modeling long-term seasonality.

Figure 3  
**Sample Fits of the Wavelet Based and Simple LTSC Models**

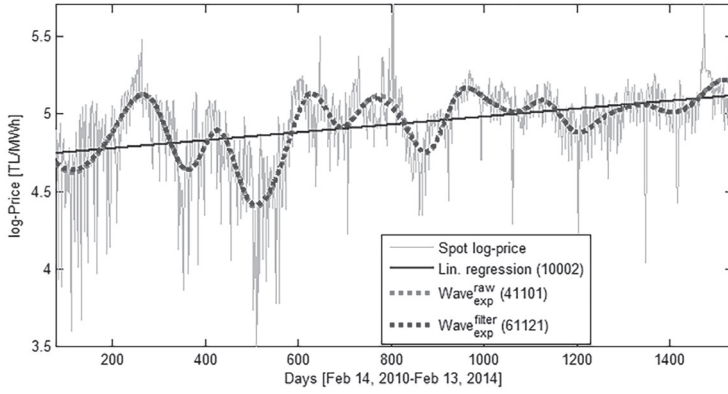


Figure 4  
**Sample Fits of the Wavelet Based and Simple LTSC Models**

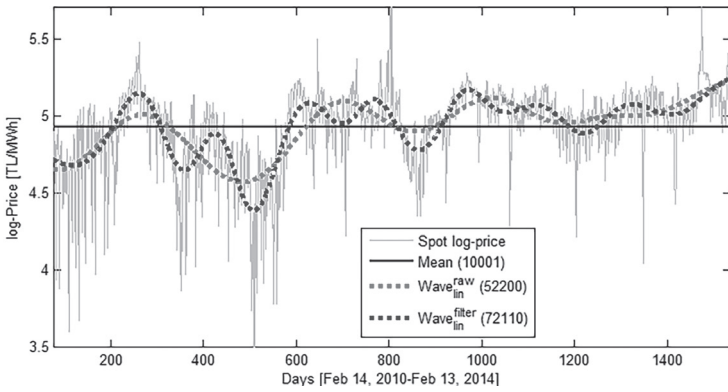
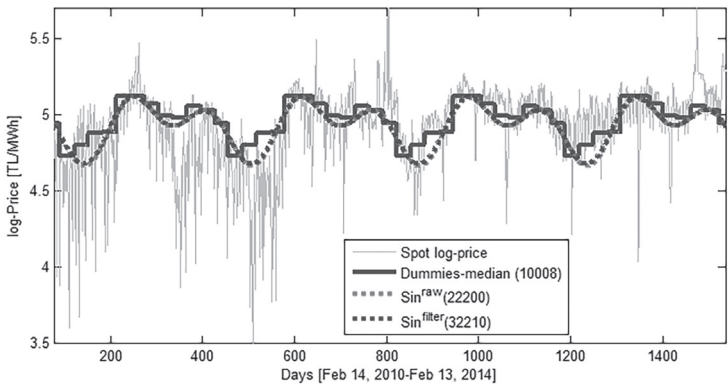


Figure 5  
**Sample Fits of the Fourier Decomposition Based and Simple LTSC Models**



## Performance across Different Forecasting Horizons

Table 5 and Table 6 contain lists of the top five models according to the three error measures  $MAE_h$ ,  $MSE_h$ , and  $MAPE_h$  for  $h = 1, \dots, 6$ . We observe that two models among the overall best performing models (emphasized in bold in Table 4) are listed in the top five for the 1-7 day forecasting horizon. Similarly three of the best performing models are ranked in the list of the top five for the 8-30 day, 31-90 day, and 91-182 day forecasting horizons.

None of the overall best performing models appear in the top five list for the third and fourth quarters of the one-year forecasting window. As expected, the forecasting power of the overall best performing models decreases as the time horizon increases. The results show that the third and fourth quarters of the one-year forward curve require different models.

### 1-7 Day Forecasting Horizon

For the 1-7 day forecasting horizon, 6\*\*\*\* type models (wavelets with an exponential decay to the median fitted to spike-filtered prices) dominate the other models. It is important to note that all 6\*\*\*\* models come from an  $S_6$  approximation level which is a more sensitive approximation. This result is consistent with the belief that a more sensitive approximation level provides a better local calibration and in turn a better short term estimation. Among the 6\*\*\*\* models, the ones from the Daubechies wavelet family of order 12 perform better than other Daubechies and Coiflets models.

### 8-30 Day and 31-90 Day Forecasting Horizons

For the 8-30 day forecasting horizon, there is no model other than 6\*\*\*\* in the top five list. This result underlines the superior forecasting power of the wavelets family with an exponential decay to the median fitted to spike-filtered prices to the one month period. We also observe that for both the 8-30 day and the 31-90 day periods, all of the 6\*\*\*\* models also belong to the  $S_6$  approximation level. It is also interesting to note that all of the models in the 8-30 day, 31-90 day, and 91-182 day periods exhibit relatively fast exponential decay to the median with  $\lambda = 1/30$ . Furthermore, consistent with our previous results, the only simple model which appears in the top five lists of different forecasting horizons is 10004, i.e. exponential decay to the median with the decay parameter  $\lambda = 1/30$ .

### 91-182 Day and 183-274 Day Forecasting Horizons

For the 2nd quarter (91-182 day), although there are two 6\*\*\*\* models from the overall best performing models, 4\*\*\*\* types of family (wavelets with an exponential decay to the median fitted to raw prices) dominate the top five lists according to the three different error measures. Surprisingly, although the best forecasting model (42101) comes from the family which is fitted to raw prices, the second best model (62121) comes from the family which is fitted to spike filtered prices. Good performance of the 4\*\*\*\* models continues also into the 3rd quarter (i.e. 183-274 day forecasting horizon) as they are ranked among the best three models. For this period, all of the 4\*\*\*\* and 6\*\*\*\* models show slow exponential decay to the median with  $\lambda = 1/180$ .

Table 5

**Top Five Models According to the Three Error Measures. The models are ranked with respect to MAPE, independently for each of the three shorter forecasting horizons: 1–7 days, 8–30 days and 31–90 days.**

Forecasting Horizon 1-7 days						
Model	$MAE_1$	Rank	$MSE_1$	Rank	$MAPE_1$	Rank
<b>61121</b> <sup>b,2,B</sup>	0.1060	1	0.0263	1	2.120%	1
61111	0.1071	4	0.0278	25	2.136%	2
61112	0.1070	3	0.0283	42	2.139%	3
61122	0.1069	2	0.0268	2	2.142%	4
62112	0.1073	6	0.0285	48	2.145%	5
62122	0.1072	5	0.0271	10	2.147%	7
<b>64121</b> <sup>c,1,C</sup>	0.1087	19	0.0269	3	2.179%	20
73220	0.1094	24	0.0269	4	2.192%	23
41101	0.1082	15	0.0270	5	2.165%	15
Forecasting Horizon 8-30 days						
Model	$MAE_2$	Rank	$MSE_2$	Rank	$MAPE_2$	Rank
62111	0.1292	2	0.0351	6	2.582%	1
<b>62121</b> <sup>d,3,D</sup>	0.1292	3	0.0346	2	2.586%	2
<b>61121</b> <sup>b,2,B</sup>	0.1292	1	0.0343	1	2.587%	3
61111	0.1297	4	0.0349	5	2.594%	4
64111	0.1302	5	0.0348	4	2.605%	5
<b>64121</b> <sup>c,1,C</sup>	0.1306	7	0.0347	3	2.615%	7
Forecasting Horizon 31-90 days						
Model	$MAE_3$	Rank	$MSE_3$	Rank	$MAPE_3$	Rank
62111	0.1472	1	0.0397	1	2.952%	1
<b>64121</b> <sup>c,1,C</sup>	0.1476	2	0.0400	3	2.961%	2
<b>10004</b> <sup>a,A</sup>	0.1476	3	0.0404	10	2.962%	3
64111	0.1477	5	0.0400	2	2.962%	4
<b>62121</b> <sup>d,3,D</sup>	0.1477	4	0.0401	4	2.965%	5
61111	0.1480	7	0.0402	5	2.968%	8

### 275-365 Day Forecasting Horizon

It is clear from the 3rd and 4th quarter panels that, towards the very end of the forward curve, electricity prices require a different kind of modeling. For instance, although 7\*\*\*\* models perform poorly for the forecasting period of a whole year, they are ranked in the top five lists for the 3rd and 4th quarter forecasting horizons. Surprisingly, no 4\*\*\*\* or 6\*\*\*\* type models exist in the 4th quarter. Apparently, linear decay to the median is superior to the exponential decay to the median for the long end of the forward curve.

If one searches for the best performing models, then few candidates exist. Three models, namely 61121, 62121, and 64121, which appear in the top five models (emphasized in bold) in Table 4 also enter into the list of the best five models of 1-7 day, 8-30 day, 31-90 day, and 91-182 day forecasting horizons. In other words, appropriate models

for up to a 182 day forecasting horizon are the wavelet-based,  $S_6$  approximation level with spikes replaced by the upper/lower 2.5% quantiles of the deseasonalized prices and with exponential decay to the median with decay parameter  $\lambda = 1/30$  models. These best performing models differ only in the choice of the wavelet family.

Table 6

**Top Five Models According to the Three Error Measures and over All Six Datasets. The models are ranked with respect to MAPE h., independently for each of the three longer forecasting horizons: 91–182 days, 182–**

Forecasting Horizon 91-182 days (2nd quarter)						
Model	$MAE_4$	Rank	$MSE_4$	Rank	$MAPE_4$	Rank
42101	0.1351	2	0.0283	1	2.711%	1
<b>62121<sup>d,3,D</sup></b>	0.1355	3	0.0285	2	2.719%	2
43301	0.1350	1	0.0298	26	2.723%	3
<b>62221<sup>e,E</sup></b>	0.1359	7	0.0287	5	2.725%	4
41101	0.1358	4	0.0286	3	2.725%	5
<b>44101<sup>5</sup></b>	0.1358	5	0.0287	6	2.727%	6
43101	0.1360	8	0.0286	4	2.729%	8

Forecasting Horizon 183-274 days (3rd quarter)						
Model	$MAE_5$	Rank	$MSE_5$	Rank	$MAPE_5$	Rank
41202	0.1373	1	0.0303	1	2.754%	1
42202	0.1374	2	0.0303	2	2.755%	2
43302	0.1376	4	0.0307	7	2.763%	3
71220	0.1375	3	0.0315	20	2.765%	4
63322	0.1378	9	0.0307	8	2.767%	5
72220	0.1377	5	0.0316	25	2.769%	10
61222	0.1381	12	0.0305	3	2.769%	9
62222	0.1381	14	0.0305	4	2.770%	11
44202	0.1381	15	0.0306	5	2.770%	12

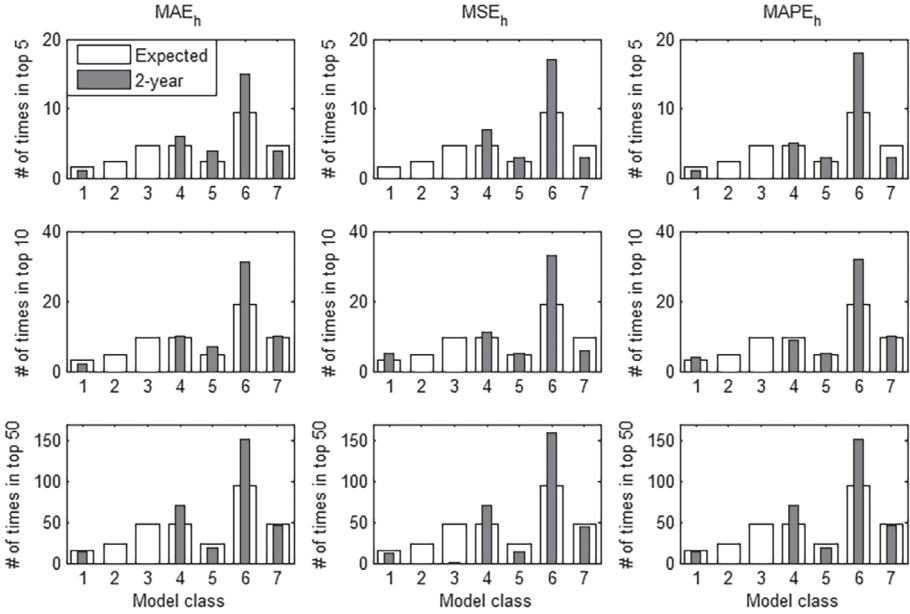
  

Forecasting Horizon 275-365 days (4th quarter)						
Model	$MAE_6$	Rank	$MSE_6$	Rank	$MAPE_6$	Rank
53200	0.0303	1	0.1345	1	2.691%	1
73220	0.0305	2	0.1349	2	2.700%	2
54200	0.0305	3	0.1350	3	2.701%	3
51100	0.0306	5	0.1352	4	2.706%	4
74220	0.0306	4	0.1352	6	2.706%	5
54100	0.0307	6	0.1352	5	2.707%	6



Figure 6

Histograms show how many times models from a given family (1\*\*\*\*, 2\*\*\*\*, 3\*\*\*\*, 4\*\*\*\*, 5\*\*\*\*, 6\*\*\*\*, 7\*\*\*\*) are ranked in the top 5 (top row), top 10 (center row) and top 50 (bottom row) of all 152 models according to MAE, MSE, and MAPE, (in columns, from left to right) for each of the six forecast horizons  $h = 1, \dots, 6$



### Conclusion

Spot electricity prices are often characterized by price spikes and seasonality. The prices observed in the Turkish electricity market are no exception. Therefore, it is important to find an appropriate outlier detection method and then model long-term seasonality in prices. We apply different outlier detection methods following Janczura and Weron (2010). For modeling the long-term seasonality component, we consider three approaches following Janczura et al. (2013), namely, (1) Fourier decomposition, (2) piecewise constant functions or dummies, and (3) wavelet decomposition. After removing the long-term seasonality in the price time series under three different approaches, we remove price spikes and compare different outlier detection methods. We observe that FPT1 (fixed price threshold with deseasonalized log-prices exceeding the range of (-0.5,0.5) ) produces the smallest number of price spikes (0.32%), and VPT2 (variable price threshold with 10% highest and 10% lowest deseasonalized log-prices treated as outliers) yields the highest number of price spikes (10.03%).

Forecasting performance of long-term seasonality models are compared at six different forecast horizons with a maximum forecasting horizon of one year. We replace

the spikes by the mean of the deseasonalized prices, or, alternatively, we replace spikes by the upper/lower 2.5% quantiles of deseasonalized prices. In these three approaches to modeling long-term seasonality, we utilize a total of 304 models and find that the models based on wavelet decomposition perform best in forecasting spot electricity prices in Turkey. On the other hand, Fourier decomposition models tend to overshoot or undershoot. To the best of our knowledge, this is the first study that offers a comprehensive comparison of different long-term seasonality and outlier detection models for spot electricity prices in Turkey.

## References

- Aggarwal, S.K., Saini, L.M., and Kumar, A. (2009). "Electricity Price Forecasting in Deregulated Markets: A Review and Evaluation," *International Journal of Electrical Power and Energy Systems*, 31(1): 13-22.
- Akay, D. and Atak, M. (2007). "Grey Prediction with Rolling Mechanism for Electricity Demand Forecasting of Turkey," *Energy*, 32: 1670-1675.
- Altınay, G. and Karagöl, E. (2005) "Electricity Consumption and Economic Growth: Evidence from Turkey," *Energy Economics*, 27: 849-856.
- Azize, H. and Talaslı, I. (2014). "Stochastic Multifactor Modeling of Spot-Electricity Prices," *Journal of Computational and Applied Mathematics*, 259: 434-442.
- Bağdadıoğlu, N. and Odyakmaz, N. (2009). "Turkish Electricity Reform," *Utilities Policy*, 17: 144-152.
- Benth, F.E., Kiesel, R., and Nazarova, A. (2012). "A Critical Empirical Study of Three Electricity Spot Price Models," *Energy Economics*, 34: 1589-1616.
- Bierbrauer, M., Trück, S., and Weron, R. (2004). "Modeling Electricity Prices with Regime Switching Models," *Lecture Notes Computer Science* 3039, 859867.
- Boogert, A. and Dupont, D. (2008). "When Supply Meets Demand: The Case of Hourly Spot Electricity Prices," *IEEE Transactions Power Systems*, 23(2): 389-398.
- Borovkova, S. and Permana, F.J. (2006). "Modelling Electricity Prices by the Potential Jump-diffusion," in Shiryaev, A.N., et al. (eds.), *Stochastic Finance Proceedings of StochFin2004*. Springer: 239264.
- Bölük, G. and Koç, A.A. (2010). "Electricity Demand of Manufacturing Sector in Turkey: A Translog Cost Approach," *Energy Economics*, 32: 609-615.
- Cartea, A. and Figueroa, M.G. (2005). "Pricing in Electricity Markets: a Mean Reverting Jump Diffusion Model with Seasonality" *Applied Mathematical Finance*, 12(4): 313-335.
- Çetin, T. and Oğuz, F. (2007). "The Politics of Regulation in the Turkish Electricity Market," *Energy Policy*, 35: 1761-1770.
- Conejo, A J., Plazas, M.A., Espinola, R., and Molina, A.B. (2005). "Day-ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models," *IEEE Transactions on Power Systems*, 20(2): 1035-1042.
- Ediger, V.S. and Tatlıdil, H. (2002). "Forecasting the Primary Energy Demand in Turkey and Analysis of Cyclic Patterns," *Energy Conversion and Management*, 43: 473-487.
- Erdoğan, E. (2010). "A Paper on the Unsettled Question of Turkish Electricity Market: Balancing and Settlement System (Part I)," *Applied Energy*, 87: 251-258.

- Erlwein, C., Benth, F.E., and Mamon, R. (2010). "HMM Filtering and Parameter Estimation of an Electricity Spot Price Model," *Energy Economics*, 32(5): 1034-1043.
- Fanone, E., Gamba, A., and Prokopczuk, M. (2013). "The Case of Negative Day-ahead Electricity Prices," *Energy Economics*, 35: 22-34.
- Fleten, S.E., Heggedal, A.M., and Siddiqui, A. (2011). "Transmission Capacity between Norway and Germany: A Real Options Analysis," *Journal of Energy Markets*, 4(1): 121-147.
- Geman, H. and Roncoroni, A. (2006). "Understanding the Fine Structure of Electricity Prices," *Journal of Business* 79: 1225-1261.
- Gökgöz, F. and Atmaca, M.E. (2012). "Financial Optimization in the Turkish Electricity Market: Markowitz's Mean-variance Approach," *Renewable and Sustainable Energy Reviews*, 16: 357-368.
- Haldrup, N., Nielsen, F.S., and Nielsen, M.O. (2010). "A Vector Autoregressive Model for Electricity Prices Subject to Long Memory and Regime Switching," *Energy Economics*, 32(5): 1044-1058.
- Hamzacebi, C. (2007). "Forecasting of Turkey's Net Electricity Energy Consumption on Sectoral Bases," *Energy Policy*, 35: 2009-2016.
- Hamzacebi, C. and Es, H.A. (2014). "Forecasting the Annual Electricity Consumption of Turkey Using an Optimized Grey Model," *Energy*, 70: 165-171.
- Higgs, H. and Worthington, A. (2008). "Stochastic Price Modeling of High Volatility, Mean-reverting, Spike-prone Commodities: The Australian Wholesale Spot Electricity Market," *Energy Economics*, 30(6): 3172-3185.
- Janczura, J. and Weron, R. (2010). "An Empirical Comparison of Alternate Regime-switching Models for Electricity Spot Prices," *Energy Economics*, 32(5): 1059-1073.
- Janczura, J. and Weron, R. (2012). "Efficient Estimation of Markov Regime-switching Models: An Application to Electricity Spot Prices," *AStA Advances in Statistical Analysis*, 96(3): 385-407.
- Janczura, J., Trück, S., and Weron, R. (2013). "Identifying Spikes and Seasonal Components in Electricity Spot Price Data: A Guide to Robust Modeling," *Energy Economics*, 38: 96-110.
- Keleş, D., Hartel, R., Möst, D., and Fichtner, W. (2012). "Compressed-air Energy Storage Power Plant Investments under Uncertain Electricity Prices: An Evaluation of Compressed-air Energy Storage Plants in Liberalized Energy Markets," *Journal of Energy Markets*, 5(1): 53-84.
- Koopman, S.J., Ooms, M., and Carnero, M.A. (2007). "Periodic Seasonal Reg-arma-garch Models for Daily Electricity Spot Prices," *Journal of the American Statistical Association*, 102(477): 16-27.
- Küçükali, S. and Barış, K. (2010). "Turkey's Short-term Gross Annual Electricity Demand Forecast by Fuzzy Logic Approach," *Energy Policy*, 38: 2438-2445.
- Mayer, K., Schmid, T., and Weber, F. (2012). "Modeling Electricity Spot Prices: Combining Mean Reversion, Spikes and Stochastic Volatility," *European Journal of Finance*, 1-24. Available [online] at: <http://dx.doi.org/10.1080/1351847X.2012.716775>
- Nowotarski, J., Tomczyk, J. and Weron, R. (2013). "Robust Estimation and Forecasting of the Long-term Seasonal Component of Electricity Spot Prices," *Energy Economics*, 39: 13-27.
- Trück, S., Weron, R., and Wolff, R. (2007). "Outlier Treatment and Robust Approaches for Modeling Electricity Spot Prices. Proceedings of the 56th Session of the ISI. Available [online] at: MPRA: <http://mpra.ub.unimuenchen.de/4711/>
- Schlueter, S. (2010). "A Long-term/short-term Model for Daily Electricity Prices with Dynamic Volatility," *Energy Economics*, 32(5):1074-1081.

Stevenson, M. (2002). "Filtering and Forecasting Spot Electricity Prices in the Australian Electricity Market," working paper, University of Technology, Sydney.

Stevenson, M.J., Amaral, J.F.M., and Peat, M. (2006). "Risk Management and the Role of Spot Price Predictions in the Australian Retail Electricity Market," *Studies of Nonlinear Dynamic Equations*, 10(3): (Article 4).

Talashlı, I. (2012). Dissertation, The Graduate School of Applied Mathematics, Middle East Technical University.

Weron, R. (2006). *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. John Wiley and Sons.

-----, (2009). "Heavy-tails and regime-switching in electricity Prices," *Mathematical Methods of Operations Research*, 69(3): 457-473.

-----, (2008). "Market Price of Risk Implied by Asian-style Electricity Options and Futures," *Energy Economics*, 30: 1098-1115.

Weron, R., Bierbrauer, M., and Trück, S. (2004a). "Modeling Electricity Prices: Jump Diffusion and Regime Switching," *Physica A: Statistical Mechanics and its Applications*, 336(1): 39-48. ce from the Nord Pool electricity market. In *The application of econophysics* (pp. 182-191). Springer Japan.

Yıldırım, M.H., Özmen, A., Bayrak, O.T.B. and Weber, G.W. (2012). "Electricity Price Modelling for Turkey," *Operations Research Proceedings*, Springer: 34-44.