Journal of Risk and Financial Management

MDPI

Article

Coupling the Empirical Wavelet and the Neural Network Methods in Order to Forecast Electricity Price

Heni Boubaker [1,2,*] and Nawres Bannour [3,4,5]

[1] Institute of High Commercial Studies (IHEC) of Sousse, LaREMFiQ, B.P. 40, Sousse 4054, Tunisia
[2] IPAG LAB, IPAG Business School, 184 Boulevard Saint-Germain, 75006 Paris, France
[3] Faculté des Sciences Economique et de Gestion de Sousse, Cité Riadh, Sousse 4023, Tunisia
[4] Université Vytautas-Magnus (VDU Rektoratas), K. Donelaičio g. 58, 44248 Kaunas, Lithuania
[5] Hungarian University of Agriculture and Life Sciences (MATE), Páter Károly U 1, 2100 Gödöllő, Hungary
* Correspondence: heniboubaker@gmail.com

Abstract: This paper aims to evaluate the forecast capability of electricity markets, categorized by numerous major characteristics such as non-stationarity, nonlinearity, highest volatility, high frequency, mean reversion and multiple seasonality, which give multifarious forecasts. To improve it, this investigation proposes a new hybrid approach that links a dual long-memory process (Gegenbauer autoregressive moving average (GARMA) and generalized long-memory GARCH (G-GARCH)) and the empirical wavelet transform (EWT) and local linear wavelet neural network (LLWNN) approaches, forming the k-factor GARMA-EWLLWNN model. The future hybrid model accomplished is assessed via data from the Polish electricity markets, and it is matched with the generalized long-memory k-factor GARMA-G-GARCH process and the hybrid EWLLWNN, to demonstrate the robustness of our approach. The obtained outcomes show that the suggested model presents important results to define the relevance of the modeling approach that offers a remarkable framework to reproduce the inherent characteristics of the electricity prices. Finally, it is presented that the adopted methodology is the most appropriate one for prediction as it realizes a better prediction performance and may be an answer for forecasting electricity prices.

Keywords: electricity cost; empirical wavelet; dual long memory; hybrid estimation; prediction performance

JEL Classification: C13; C14; C22; C45; C53

Citation: Boubaker, Heni, and Nawres Bannour. 2023. Coupling the Empirical Wavelet and the Neural Network Methods in order to Forecast Electricity Price. *Journal of Risk and Financial Management* 16: 246. https://doi.org/10.3390/jrfm16040246

Academic Editor:Shigeyuki Hamori

Received: 28 January 2023
Revised: 6 April 2023
Accepted: 9 April 2023
Published: 18 April 2023

1. Introduction

In electricity markets, an analysis of the price has become a significant subject for all its contributors. The background information about the electricity price is for decisive risk management. Moreover, it signifies a gain for a market actor leading against their competitors. As part of this framework, both suppliers and customers depend on the forecasting of price data to move forward with their corresponding bidding strategies. If a supplier has a precise price expected, they are able to improve their bidding strategy to take full advantage of the profit. Contrariwise, if an exact expected price is accessible, a customer can plot to decrease their own electricity fees. Therefore, the player's advantage is significantly impacted by the exactitude of the price expected. Nevertheless, the electricity price performance is different from other commodity markets and financial markets, whereby it has particular features which are connected to its characteristics and can noticeably disturb the prices.

Indeed, in antagonizing commodities markets, the unfeasibility of storing electrical energy to transport in future periods, related with the requirement of guaranteeing a continuous balance between supply and demand, makes the way electricity markets should

be activated ultra-unusual. Moreover, delivering electrical energy necessities to please the corporal and technical boundaries of the system makes the market more complex. These worries, as well as the impact of weather conditions and production materials prices, significantly affect the organization of the electricity markets and lead to special features of the electricity prices and to the spot price[1]. More precisely, the electricity spot prices are considered by their extreme weight fluctuations, which make large and uncommon leaps (Weron et al. 2004). Moreover, electricity prices display some characteristics, such as non-stationarity, high frequency, multiple seasonality (on annual, weekly and daily levels) (Escribano et al. 2011; Koopman et al. 2007; Knittel and Roberts 2005), hard nonlinearity, high volatility, long memory, a high percentage of unusual prices, the calendar effect, price spikes and mean reversion. Subsequently, these behaviors will possibly affect the spot prices dramatically.

To tackle these challenges, this paper aims at theoretically and empirically investigating the quality of a hybrid system in resolving the difficulty of the electricity spot price. The objective of this paper is to illustrate a robust approach to modeling electricity prices. To accomplish the aforementioned goal, the log-return of the electricity price for the Polish market is used in this study to demonstrate the relevance and usefulness of the model for time-series forecasting. The remainder of this paper is organized as follows: The following section provides a short review of the literature. Section 3 illustrates the econometric methodology which contains the theoretical notions of the k -factor GARMA model and the wavelet local linear neural network model and explains the hybrid k -factor GARMA-WLLWNN method and the k -factor GARMA-G-GARCH process, using a wavelet estimation approach. Section 4 outlines the empirical framework, where the suggested hybrid model is executed to demonstrate log-return electricity spot price forecasting, and its performance is compared with the individual WLLWNN model and the generalized long-memory k -factor GARMA-G-GARCH model. Section 5 concludes this paper.

2. Literature Review

It is vital to remember that a convenient forecast model for electricity prices should reflect the characteristics outlined above. In the area of electricity price forecasting, two methods, statistical or econometric time-series models, are largely considered as a parametric tool, and soft computing models are considered a nonparametric tool. These two methods are established to have been useful. In statistical models, the autoregressive integrated moving average (ARIMA) (Contreras et al. 2003; Liu and Shi 2013) and generalized autoregressive conditional heteroscedasticity (GARCH) (Garcia et al. 2005; Ghosh and Kanjilal 2014; Girish 2016) models are applied widely. On the other hand, these models do not allow to take into consideration the long-memory behavior that indicates the electricity prices. To address this restriction, Granger and Joyeux (1980) and Hosking (1981) established the fractional autoregressive moving average (ARFIMA) model. Baillie et al. (1996) and Bollerslev and Mikkelsen (1996) confirmed the use of the fractionally integrated generalized autoregressive conditional heteroscedasticity (FIGARCH) process to present finite persistence in the conditional variance. Actual research has executed these methodologies for the electricity prices (Koopman et al. 2007). In the spectral area, these procedures have the highest value for ultra-low rates near to zero frequency. It is noteworthy that the ARFIMA model is not capable of presenting the determined periodic or cyclical behavior in the time series.

To address this limitation, Gray et al. (1989) submitted another group of long-memory models, including the generalized (seasonal) long-memory or Gegenbauer autoregressive moving average (GARMA) process, which has been recognized for evaluating equally the seasonality and the persistence in the data. Alternatively, in the frequency domain, the spectral density is not essentially unlimited at the origin, as with the ARFIMA model, but for any frequency λ along the interval $[0, \pi]$. In that way, the GARMA process provides long-memory cyclical behavior at a single frequency λ, ; therefore, it is able

to judge a single persistent periodic component. To address this pitfall, Woodward et al. (1998) generalized the single frequency GARMA process to the so-called k -factor GARMA process that permits the spectral density function to be not just positioned at a single frequency but nonetheless defined at a fixed number k of frequencies in $[0,\pi]$, recognized as the Gegenbauer frequencies or G-frequencies. The key distinguished feature of this k -factor GARMA process is that it permits additional variety in the structure of the covariance of a variable founded equally through the spectral density function and the autocorrelation function that offers k singularities. The k -factor GARMA model has been used by numerous researchers to duplicate the seasonal patterns in addition to the persistent special effects in the stock markets (Caporale and Gil-Alana 2014; Boubaker and Sghaier 2015). Notwithstanding the compliance of this modelization with the features of electricity prices, some uses are concerned with the electricity market (Diongue et al. 2009; Soares and Souza 2006). Regarding the approximation of the parameter's k -frequency GARMA process, Gray et al. (1989), Beran (1994) and Woodward et al. (1998) measured the time-domain maximum likelihood method. Whitcher (2004) added an assessment method in the wavelet domain, established on the maximal overlap discrete wavelet packet transform (MODWPT). In comparison with the Fourier analysis, the efficiency of the wavelet analysis connects its capability to identify a process simultaneously in the time and frequency domains equally (Mallat and Zhang 1993; Mallat 1999; Boubaker and Boutahar 2011; Boubaker 2015, 2016).

To duplicate these models, two methods have been approved in this study: the nonparametric methods, for instance, the neuronal networks, and the parametric models, termed the generalized GARCH (G-GARCH) process. In the fundamental approach, artificial neural networks (ANN) have been frequently used in modeling and forecasting electricity spot prices; Wang and Ramsay (1998), Szkuta et al. (1999), Anbazhagan and Kumarappan (2014) and Panapakidis and Dagoumas 2016) implemented neural networks to present and forecast the dynamics of intra-day prices. Zhang and Benveniste (1992) recommended wavelet neural networks as a solution to traditional NNs (such as feedforward NNs) to decrease the faults linked to each methodology. WNs are unknown depth networks, which allow for a wavelet as an activation function. The WNs have been well applied in time-series forecasting (Cao et al. 1995; Cristea et al. 2000; Jiang et al. 2020; Jabeur et al. 2021) and in short-term electricity prices forecasting (Bashir and El-Hawary 2000; Yao et al. 2000; Gao and Tsoukalas 2001; Benaouda et al. 2006; Rana and Koprinska 2016). Yet the main weakness of the WNN is that it requires various hidden layer units for advanced dimensional problems. To maintain the improvement correlated to the local ability of the wavelet basis purposes while not using many hidden layers, Chen et al. (2004) established a novel kind of wavelet neural network, named the local linear wavelet neural network (LLWNN). In this model, a local linear model replaces the connection weights between the hidden layer units and the output units. For that reason, this network needs less significant wavelets for a certain hardness, assimilating to the wavelet neural networks. Quite a number of scientists have widely used the LLWNN model for electricity price forecasting (Pany 2011; Chakravarty et al. 2012; Athanassios et al. 2015).

Truthfully, both the k -factor GARMA model, as a dominant statistical method, and the LLWNN model, as an advanced artificial intelligence methodology, have reached achievements in their specific nonlinear parametric and nonparametric domains individually. Nevertheless, neither of these are a global model that is appropriate for all situations. That is, a time series is habitually complicated in nature and an individual model will possibly not allow to notice diverse forms in a similar way; at this time, no methodology is the best for all circumstances. For that reason, accepting a hybrid method or joining numerous modulizations (Granger 1989) has converted a common practice, jumping the limits of the utilization of a single model to improve the forecasting precision. In the literature, various mixed methodologies have been recommended to prevent the limitations linked to unique models (Yu et al. 2005; Armano et al. 2005; Tseng et al. 2002; Zhang

2003; TaskaValenzuela et al. 2008; Khashei and Bijari 2010; Tan et al. 2010; Sharkey 2002; Shafie-khah et al. 2011; Jiang et al. 2017; Zhang et al. 2018; Grossi and Nan 2019).

In the next approach, to overcome the restriction of the k -factor GARMA model, Boubaker (2015) utilized the GARCH model, proposed by Engle (1982) and Bollerslev (1986). In one more study, Boubaker and Boutahar (2011) submitted the k -factor GARMA-FIGARCH to assess the long-memory behavior in the conditional variance of the exchange rate. However, these modulizations are not completely suitable in modeling the volatility of intra-daily financial time series. The core characteristic of such data is the robust indication of cyclical patterns in the volatility. The empirical result highlights the significance of modeling the periodic dynamics of the volatility. For this target, Bordignon et al. (2007, 2008) advised a first-hand type of GARCH model categorized by periodic long-memory behavior. This sort of modulization familiarizes Gegenbauer polynomials into the equation of the standard GARCH model, measured as generalized periodic long-memory filters to evaluate the time-varying volatility. These processes are called the periodic long-memory GARCH (PLM-GARCH) and generalized long-memory GARCH (G-GARCH). In the literature, the generalized long-memory GARCH model (or G-GARCH) is applied to weigh the financial time series, such as the exchange rate, using Monte Carlo simulations (Bordignon et al. 2007; and Caporin and Lisi 2010). Rarely has research carried out this process for modeling the electricity spot price (Diongue et al. 2009).

It is worth taking into consideration that in the works of generalized long-memory models, the researchers approve either the k -factor GARMA model or the G-GARCH model to evaluate the conditional mean and the conditional variance of the time series, correspondingly. None of them take into account the presence of long-memory and cyclical behavior in the conditional mean and conditional variance together.

The objective of the current work is to address the challenges of modelizing and expecting numerous components of electricity prices, mainly the presence of seasonal long-memory behavior in the conditional mean and conditional variance. With this in mind, this paper provides three contributions. The main goal is to increase the LLWNN model's predicting accuracy. Instead of adding the historical price straight to the network (at the input layer), this is achieved by implementing wavelet theories to deconstruct it and evaluate the impact of varying heights of decomposition on forecasting accuracy. This strategy, designated as EWLLWNN, enables the learner's network to recognize the presence of seasonal long-memory behavior and therefore further evaluate the data. In reality, the prior research integrated the wavelet decomposition with the ANN (Aggarwal et al. 2008), demonstrating its effectiveness, but this technique is not used with the LLWNN. A few studies (Pany 2011; Chakravarty et al. 2012) hypothesized that this model might offer an accurate forecast because the features related to electricity prices can be recognized in the unseen layer and use the wavelet activation function, deprived of the use of an external decomposer/composer. Lastly, so as to define the best architecture for the proposed EWLLWNN, we will compare two distinct learning methods, namely the back propagation (BP) and particle swarm optimization (PSO) algorithms, and then choose the one that minimizes the errors.

Second, a novel hybrid model is suggested so as to take advantage of the advantages of both semi-parametric and nonparametric approaches, which actually results in the k -factor GARMA-EWLLWNN process, which makes it possible for long-memory behavior, related with the frequency, and involves an EWLLWNN-type model to explain time-varying volatility. We employ a wavelet estimation method based on Whitcher's (2004) suggested maximal overlap discrete wavelet packet transform (MODWPT) to estimate the parameters k -factor GARMA process (for additional details, see Boubaker 2015).

Finally, we devised a new methodology based on a dual generalized long-memory method that involves both the k -factor GARMA model and the G-GARCH model to accommodate for various stylized facts including the stochastic volatility, long-range dependence, and seasonality characteristics found in electricity spot prices. As a consequence, the forecast accuracy of the k -factor GARMA-G-GARCH model is compared to

the suggested k -factor GARMA-EWLLWNN model. In fact, this stage entails comparing the new EWLLWNN's efficiency with a parametric model (G-GARCH) in forecasting and predicting the conditional variance's periodic long-memory behavior.

3. Econometric Methodology

3.1. The k-factor GARMA model

Gray et al. (1989) suggested a k -frequency GARMA model to generalize the ARFIMA model, permitting periodic or quasi-periodic movement in the signal, which is presented as

$$\Phi(B)\prod_{i=1}^{k}\left(I-2\nu_{m,i}B+B^{2}\right)^{d_{m,i}}\left(y_t-\mu\right)=\Theta(B)\varepsilon_t \tag{1}$$

where $\Phi(B)$ and $\Theta(B)$ are the polynomials of the backshift operator B such that all the roots of $\Phi(z)$ and $\Theta(z)$ lie outside the unit circle. The parameters $\nu_{m,i}$ offer information about periodic movement in the conditional mean, ε_t is a white noise disturbance sequence with variance σ_{ε}^{2}, k is a finite integer, $\left|\nu_{m,i}\right|<1$, $i=1,2,\ldots,k$, $d_{m,i}$ are long-memory parameters of the conditional mean indicating how slowly the autocorrelations are damped, μ is the mean of the process, $\lambda_{m,i}=\cos^{-1}(\nu_{m,i})$, $i=1,2,\ldots,k$, denote the Gegenbauer frequencies (G-frequencies). The GARMA model with k -frequency is stationary when $\left|\nu_{m,i}\right|<1$ and $d_{m,i}<1/2$ or when $\left|\nu_{m,i}\right|=1$ and $d_{m,i}<1/4$; the model exhibits a long memory when $d_{m,i}>0$. However, when it comes to a single frequency $\nu=1$, the process is an ARFIMA (p,d,q) model, and when $\nu=1$ and $d=1/2$, the process is an ARIMA model. Finally, when $d=0$, we acquire a stationary ARMA model. Cheung (1993) defines the spectral density function and displays that for $d>0$ the spectral density function has a pole at $\lambda=\cos^{-1}(\nu)$, which differs in the interval $[0,\pi]$. It is important to note that when $|\nu|<1$, the spectral density function is bounded at the origin for GARMA processes and thus does not suffer from many problems related with ARFIMA models (for more details, see Boubaker and Boutahar 2011; Boubaker 2015).

3.2. The Wavelet Local Linear Wavelet Neural Network

The wavelet local linear wavelet neural network (WLLWNN) involves the following. Initially, the historical data have been decomposed into wavelet-domain constitutive subseries using wavelet area and then using the local linear wavelet neural network (LLWNN) to form the WLLWNN forecasting model. The reader should consult Ben Amor et al. (2018) and Boubaker et al. (2020) for further details.

3.2.1. Theoretical Concepts of Wavelet

Wavelet decompositions are a useful tool in producing better local representation of the series in both frequency and time domain (Nicolaisen et al. 2000; Boubaker 2015, 2016). Wavelets are orthonormal bases attained through dyadically dilating and translating a pair of specially constructed functions denoted by ϕ and ψ, which are father wavelet (detected the smooth and low-frequency part of the time series) and mother wavelet (defined the detail and the high-frequency components), respectively, given by

$$\int\phi(t)dt=1 \text{ and } \int\psi(t)dt=0. \tag{2}$$

The obtained wavelet basis is $\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)$ and $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$, where $j = 1, \ldots, J$ indexes the scale and $k = 1, \ldots, 2^j$ indexes the translation. The parameter j is adopted as the dilation parameter of the wave's functions. This parameter j adjusts the support of $\psi_{j,k}(t)$ to locally detect the features of high or low frequencies. The parameter k is used to relocate the wavelets in the temporal scale. The number of observations limits the maximum number of scales that can be employed in the analysis $(T \geq 2^J)$. The localization property is a special property of the wavelet expansion, where the coefficient of $\psi_{j,k}(t)$ reveals information content of the function at approximate location $k2^{-j}$ and frequency 2^{-j}. According to wavelets, all functions in $L^2(\bullet)$ can be extended over the wavelet basis, exceptionally, as a linear combination at arbitrary level $J_0 \in \mathbb{N}$ through different scales of the type

$$f(t) = \sum_k s_{J_0,k}\phi_{J_0,k}(t) + \sum_{j \geq J}\sum_k d_{j,k}\psi_{j,k}(t) \tag{3}$$

where $\phi_{J_0,k}$ is a scaling function with the corresponding coarse scale coefficients $s_{J_0,k}$ and $d_{j,k}$ are the detail coefficients, which measures the contribution of the corresponding wavelet to the function and constitutes the wavelet multiresolution analysis (MRA), given, respectively, by

$$s_{J_0,k} = \int f(t)\phi_{J_0,k}(t)dt \text{ and } d_{j,k} = \int f(t)\psi_{j,k}(t)dt \tag{4}$$

In the following, the coefficients of the discrete wavelet transform can be deemed from the recursive MRA scheme, which is implemented by a two-channel filter bank illustration of the wavelet transform (i.e., a high-pass wavelet filter $\{h_l, \quad l = 0, \ldots, L-1\}$ and its associated low-pass scaling filter $\{g_l, \quad l = 0, \ldots, L-1\}$, satisfying the quadrature mirror relationship given by $g_l = (-1)^{l+1} h_{L-1-l}$ for $l = 0, \ldots, L-1$, where $L \in \mathbb{N}$ is the length of the filter). Moreover, Daubechies (1992) has constructed a class of wavelet functions which forms an orthonormal basis that possesses the smallest support of a given number of vanishing moments (the extremal phase filters $D(L)$ and the least asymmetric filters $La(L)$).

3.2.2. Empirical Wavelet Transforms

Conventional data pre-processing methods such as wavelet alter and Fourier convert approaches have some shortcomings , for example, the difficulty in choosing the mother wavelet. Similarly, empirical mode decomposition (EMD) which began with Hilbert–Huang transform (HHT) and was advanced by Huang et al. (1998), variational mode decomposition (VMD) by Dominique and Dragomiretskiy (2014) and ensemble empirical mode decomposition (EEMD) developed by Wu and Huang (2009) also have some disadvantages. In contrast, empirical wavelet transforms (EWT) were suggested by Gilles (2013) as a new adaptive time-frequency analysis method, taking into account the advantage of the empirical mode decomposition method and traditional wavelet analysis in signal processing. In the frequency-domain analysis of the signal, the Fourier spectrum is adaptively segmented by capturing the frequency-domain maximum point, and finally, the separation of different modes in the signal frequency component is realized. Indeed, the EWT is used as a processing tool to decompose the series into specific modal components according to the features of the signal. The main idea of EWT is based on the divided Fourier spectrum.

Note that the Fourier support interval $[0,\pi]$ is split into N consecutive parts and ω_n represents the boundary between each segment with $\lambda_0 = 0$ and $\lambda_n = \pi$, and every pass filter can be defined as $\Lambda_n = [\lambda_{n-1}, \lambda_n]$ and $\bigcup_{n=1}^{N} \Lambda_n = [0,\pi]$, $\forall n > 0$. Following Cheng et al. (2019), the empirical wavelet's function $\hat{\psi}_n(\lambda)$ and the empirical scaling function $\hat{\phi}_n(\lambda)$ are given by the following formula:

$$\hat{\psi}_n(\lambda) = \begin{cases} 1 \quad if\ \lambda_n + \tau_n \le |\lambda| < \lambda_{n+1} - \tau_{n+1} \\ cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}\left(|\lambda| - \lambda_{n+1} + \tau_{n+1}\right)\right)\right] \quad if\ \lambda_{n+1} - \tau_{n+1} \le |\lambda| < \lambda_{n+1} + \\ sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}\left(|\lambda| - \lambda_n + \tau_n\right)\right)\right] \quad if\ \lambda_n - \tau_n \le |\lambda| < \lambda_n + \tau_n \\ 0 \quad \text{otherwise} \end{cases} \tag{5}$$

and

$$\hat{\phi}_n(\lambda) = \begin{cases} 1 \quad if\ |\lambda| < \lambda_n - \tau_n \\ cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}\left(|\lambda| - \lambda_n + \tau_n\right)\right)\right] \quad if\ \lambda_n - \tau_n < |\lambda| < \lambda_n + \tau_n, \\ 0 \quad \text{otherwise} \end{cases}$$

where the proportional relationship between τ_n and λ_n is $\tau_n = \gamma\lambda_n$, $0 < \gamma < 1$ and $\beta(x)$ satisfy $\beta(x) = \begin{cases} 0, \forall x \le 0 \\ 1, \forall x > 1 \end{cases}$ along with $\beta(x) + \beta(1-x) = 1, \forall x \in [0,1]$. The set $\left\{\phi_1(t), \{\psi_n(t)\}_{n=1}^{N}\right\}$ is a tight frame of $L^2(\cdot)$. Thus, the EWT can be implemented in the same way as the classic wavelet transform, and the corresponding mathematical expression is as follows:

$$W_f^{\varepsilon}(n,t) = \langle f(t), \psi(n,t) \rangle = \int f(\tau)\overline{\psi_n(t-\tau)}\, d\tau = F^{-1}\left(\hat{f}(\omega), \overline{\hat{\psi}_n(\omega)}\right). \tag{6}$$

where $W_f^{\varepsilon}(n,t)$ is different frequency component of the empirical wavelet, and $F^{-1}(\bullet)$ is the inverse Fourier transformation.

Ultimately, we can reconstruct the original signal as

$$\begin{aligned} f(t) &= W_f^{\varepsilon}(0,t) * \phi_1(t) + \sum_{n=1}^{N} W_f^{\varepsilon}(n,t) * \psi_n(t) \\ &= F^{-1}\left(\hat{W}_f^{\varepsilon}(0,\omega) * \hat{\phi}_1(\omega) + \sum_{n=1}^{N} \hat{W}_f^{\varepsilon}(n,\omega) * \hat{\psi}_n(\omega)\right) \end{aligned} \tag{7}$$

where $\hat{W}_f^{\varepsilon}(0,\omega)$ and $\hat{W}_f^{\varepsilon}(n,\omega)$ are, respectively, the Fourier transformations of $W_f^{\varepsilon}(0,t)$ and $W_f^{\varepsilon}(n,t)$; $*$ is convolution symbol. Therefore, the empirical wavelet

decomposition is expressed by $f_0(t) = W_f^{\varepsilon}(0,t) * \phi_1(t)$ and $f_k(t) = W_f^{\varepsilon}(k,t) * \phi_k(t)$. So, any series $f(t)$ can be defined as intrinsic mode function as $f(t) = \sum_{k=0}^{N} f_k(t)$.

3.2.3. The Local Linear Wavelet Neural Network

The local linear wavelet neural network (LLWNN) of time-series forecasting was proposed by Chen et al. (2004) who proved that this modelization has more accuracy than the classic WNN. The LLWNN includes an input layer, hidden layer and linear output layer. In this case, the input series in the input layer of the network are diffused straight into the wavelet layer. The hidden layer neurons make use of wavelets as activation functions; these neurons are habitually called 'wavelons' (see Ben Amor et al. 2018 for more details about the method). As an alternative to exploiting multilayered neural networks and several variants, a WLLWNN is used for forecasting data. Regarding wavelet transformation theory, wavelets in the following form are a family of functions, produced from one single function $\psi(x)$ by the operation of dilation and translation.

$$\psi(x) = \left\{ \psi_i = |a_i|^{-1/2} \psi\left(\frac{x - b_i}{a_i}\right); \quad a_i, b_i \in \mathbb{R}^n, i \in \mathbb{Z} \right\} \tag{8}$$

$x = (x_1, x_2, \ldots, x_n)$,

$a_i = (a_{i1}, a_{i2}, \ldots, a_{in})$,

$b_i = (b_{i1}, b_{i2}, \ldots, b_{in})$.

$\psi(x)$ is localized in both time space and the frequency space, called a mother wavelet, and the parameters a_i and b_i are the scale and translation parameters, respectively. Instead of the straightforward weight w_i (piecewise constant model), a linear model $v_i = w_{i0} + w_{i1}x_1 + \ldots + w_{in}x_n$ is introduced.

The activities of the linear models v_i $(i = 1, \ldots, n)$ are resolved by the associated locally active wavelet functions $\psi_i(x)$ $(i = 1, \ldots, n)$; thus, v_i is the single significant variable. Nonlinear wavelet basis functions (named wavelets) are contained equally in time space and frequency space. Here, $m = n$ and output (Y) of the suggested model is considered as

$$Y = \sum_{i=1}^{M} (w_{i0} + w_{i1}x_1 + \ldots + w_{in}x_n)\psi_i(x) \tag{9}$$

The mother wavelet is

$$\psi(x) = \frac{-x^2}{2} e^{\frac{-x^2}{\sigma^2}} \tag{10}$$

$$\psi(x) = e^{-\left(\frac{x-c}{\sigma}\right)^2} \tag{11}$$

where

$$x = \sqrt{d_1^2 + d_2^2 + \ldots + d_n^2} \tag{12}$$

To optimize the neural networks, the widely used learning algorithm is the back propagation (BP). Firstly, the parameters are randomly prepared, and after the algorithm

measures the error between the output value and the real value, it lastly adjusts the weights in the direction of descendent gradient. The equations of the BP algorithm are offered in Burton and Harley (1994).

Moreover, Kennedy and Eberhart (1995) established the PSO as an optimization technique. In accord with other learning algorithms, the PSO clearly showed its efficacy. PSO algorithm is made through simulation of bird flocking in two-dimensional space. The position of each agent is denoted by XY -axis position, and the velocity is represented by vx and vy. The agent position's adjustment is documented by the location and the velocity information. The bird flocking optimizes the objective function. Each agent identifies its best value so far $(pbest)$ and its XY place. In addition, each agent identifies the top value so far in the group $(gbest)$ between $(pbest)$. Mostly, each agent attempts to regulate its position using the following information:

(a) The distance between current position and $pbest$.

(b) The distance between the current position and $gbest$.

Velocity of each agent can be updated by the following equation:

$$v_i^{p+1} = wv_i^p + c_1 rand_1 \times (pbest_1 - s_i^p) + c_2 rand_2 (gbest - s_i^p) \tag{13}$$

where v_i^p is the velocity of agent i at iteration p, w is the weight function, c_j is weighting factor, s_i^p is the current position of agent i at iteration p, $pbest_i$ is the $pbest$ of agent i and $gbest$ is the $gbest$ of the group.

The velocity, which progressively gets closer to $pbest$ and $gbest$ can be computed using the above equation. The actual position, which characterizes the searching point in the solution space, can be updated using the following equation:

$$s_i^{p+1} = s^p + v_i^{p+1} \tag{14}$$

The PSO algorithm escapes from convergence toward a local minimum, because it is not based on gradient information contrary to the BP case (Abbass et al. 2001; Boubaker et al. 2020). The objective of the PSO is to produce the best set of weights (particle position) where numerous particles are moving to obtain the best solution, where the total number of weights characterizes the dimension of the search space. The optimization is finished when the personal best solution of each particle and the global best amount of the entire swarm are attained.

3.3. The Hybrid k-Factor GARMA-WLLWNN Model

Our hybrid approach combines a semi-parametric k -factor GARMA model and the proposed WLLWNN model. The k -factor GARMA model offers better flexibility in modeling simultaneous short- and long-term behavior of a seasonal time series. In addition, the selection of WLLWNN in our hybrid model is encouraged by the wavelet decomposition and its local linear modeling ability. It may be useful to think of time series as having two components: The first is a parametric form with unknown parameters, for which a parametric technique appears suitable. The residuals are the second component; this section normally has no special procedure. As a result, determining the right model to deal with this component of the time series is problematic. As a result, a nonparametric model appears to be the best choice for modeling the residuals. This choice is based on the notion that nonparametric models can decrease modeling bias by enforcing no specific model structure rather than a smoothness assumption and are thus particularly useful when we have little data or wish to be flexible with the underpinning model.

We consider a two-step approach; in the first stage, the purpose is modeling the conditional mean using a semi-parametric k -factor GARMA model. Conversely, residuals are important in forecasting time series; they may contain some information that is able to develop forecasting performance. Thus, in the second phase, the residuals resulting

from the first step will be treated according to a novel wavelet local linear wavelet neural network (WLLWNN) model.

Hence, a time series can be written as

$$y_t = \mu_t + \varepsilon_t \tag{15}$$

where μ_t denotes the conditional mean of the time series, and ε_t is the residuals. In the head stage, the key purpose is the parametric modeling; therefore, the k-factor GARMA model is used to replicate the conditional mean (equation 1).

In the next stage, the residuals from the parametric model are used as a proxy for the corresponding volatility and modeled using the WLLWNN model. Let ε_t denote the residuals at time t from the k-factor GARMA model, and then

$$\varepsilon_t = y_t - \hat{\mu}_t \tag{16}$$

The first stage's results are the forecast values and residuals of the semi-parametric modeling. In the following stage, the target is the modeling of the residuals using the WLLWNN with n input nodes, and the WLLWNN for the residuals is

$$\varepsilon_t = f\left(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots \varepsilon_{t-n}\right) \tag{17}$$

where each ε_{t-i} is decomposed using the wavelet transform (Equation (3)), f is a nonlinear, nonparametric function determined by the neural network with the reference to the current state of the data, during the training of the neural network. The output layer of the network (Equation (9)) gives the forecasting results.

$$\hat{y}_t = \hat{\mu}_t + \hat{\varepsilon}_t \tag{18}$$

As a consequence, this global prediction is the outcome of anticipating the time series' conditional mean and conditional variance.

3.4. The k-Factor GARMA-G-GARCH Model

The conditional variance is assumed to be constant throughout time in the k-frequency GARMA model. Many time series are known to display volatility clustering, where time series have both high and low periods of volatility, according to empirical investigations. To duplicate these patterns, we added a fractional filter to the conditional variance equation of the k-factor GARMA model described above. As a result, we present the dual generalized k-factor GARMA-G-GARCH model, which can incorporate seasonality and long-memory dependency in the conditional mean and variance. The basic concept behind this model is to incorporate the generalized long-memory process into the GARCH equation that describes the evolution of conditional variance. This is why this new model category is known as Gegenbauer-GARCH (G-GARCH). To account for the presence of a time-varying conditional variance, we investigate the following k-factor GARMA process with G-GARCH-type innovations.

$$y_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t z_t \tag{19}$$

where μ_t is the conditional mean of y_t modeling using the following k-factor GARMA process

$$\Phi(B)\prod_{i=1}^{k}\left(I - 2v_{m,i}B + B^2\right)^{d_{m,i}}\left(y_t - \mu\right) = \Theta(B)\varepsilon_t \tag{20}$$

$$\varepsilon_t / I_{t-1} \sim \mathcal{D}\left(0, \sigma_t^2\right) \tag{21}$$

where σ_t^2 is the conditional variance, I_{t-1} is the information up to time $t-1$, z_t is a $i.i.d$ random variable with zero mean and unitary variance and $\mathcal{D}(\cdot)$ is a probability

density function. To define the conditional variance's dynamics, the starting point is the dynamics of ε_t^2. We assume that ε_t^2 follow a k-factor GARMA model, which describes a cyclical pattern of length S.

$$\left[(I-B)^{d_{v0}}(I+B)^{d_{v,k}I(E)}\prod_{i=1}^{k-1}\left(I-2v_{v,i}B+B^2\right)^{d_{v,i}}\right]\alpha(B)\varepsilon_t^2=\gamma+\left[I-\beta(B)\right]\vartheta_t,$$
$$P_v(B)\alpha(B)\varepsilon_t^2=\gamma+\left[I-\beta(B)\right]\vartheta_t. \tag{22}$$

where $\alpha(B)=1-\sum_{i=1}^{q}\alpha_i B^i$ and $\beta(B)=1-\sum_{i=1}^{p}\beta_i B^i$ are suitable polynomials in the lag operator B and $\vartheta_t=\varepsilon_t^2-\sigma_t^2$ is a martingale difference, $d_{v,0}=d_v/2$, $I(E)=1$ if S is even and zero otherwise.

With this assumption, the corresponding GARCH-type dynamics for conditional variance are given by

$$\sigma_t^2=\gamma+\beta(B)\sigma_t^2+\left[I-\beta(B)-P_v(B)\alpha(B)\right]\varepsilon_t^2 \tag{23}$$

This indicates that in the G-GARCH framework, each frequency has been modeled by means of a specific long-memory parameter $d_{v,i}$ (differencing parameter of the conditional variance). When $d_{v,0}=d_{v,1}=\ldots=d_{v,k}$, all the involved frequencies have equal degree of memory. Model (23) can deliver, in this case, most of the existing GARCH models. For instance, standard GARCH models (including seasonal GARCH (Bollerslev and Hodrick (1992)) can be obtained by putting $d_{v,i}=0,\ \ i=0,\ldots,k$. Similarly, the FIGARCH model is equivalent to $S=1$ and $0<d_{v,0}<1$. It is worth noting that generalized long-memory filters can theoretically be applied to any GARCH structure category. However, because G-GARCH models are not always practical due to the limitations required for conditional variance positive, Bordignon et al. (2007) advocated modeling the conditional variance logarithm. As a result, applying the filter to a generalized log-GARCH model provides a viable computational solution. This entails starting with the phrase

$$P_v(B)\alpha(B)\left[ln(\varepsilon_t^2)-\tau\right]=\gamma+\left[I-\beta(B)\right]\vartheta_t \tag{24}$$

where $P_v(B)$ is the generalized long-memory filter introduced into a GARCH structure, $\vartheta_t=ln\left(\varepsilon_t^2\right)-\tau-ln\left(\sigma_t^2\right)$ is a martingale difference and $\tau=E\left[ln\left(z_t^2\right)\right]$. The expected τ value depends on the distribution of the idiosyncratic shock and ensures that ϑ_t is a martingale difference, given that $ln\left(\varepsilon_t^2\right)=ln\left(\sigma_t^2\right)+ln(\ \ )z_t^2$. Under the Gaussian assumption, $\tau=-1.27$. The expression for conditional variance implied by (23) is

$$ln\left(\sigma_t^2\right)=\gamma+\beta(B)ln\left(\sigma_t^2\right)+\left[I-\beta(B)-P_v(B)\alpha(B)\right]\left[ln\left(\varepsilon_t^2\right)-\tau\right] \tag{25}$$

Because we are modeling $\ln(\sigma_t^2)$ instead of σ_t^2, there are no requirements for variance positivity limitations. Adopting EGARCH versions of our model is another way to get around the problem of parameter limitations. To sum up, the purpose of this research consists of modeling the different patterns in the electricity time series to provide the best forecasting methods. For this determination, we exploit a hybrid methodology based on combining the semi-parametric k-factor GARMA model with a novel neural network named the EWLLWNN model. The performances of the proposed hybrid k-factor GARMA-EWLLWNN model are evaluated using data from the Polish electricity markets and matched with the dual comprehensive long-memory k-factor GARMA-G-GARCH

model and the individual (E)WLLWNN, in order to demonstrate the resilience of the hybrid model we have proposed. The flow-chart construction of the hybrid GARMA-EWLL-WNN model appears in Figure 1.

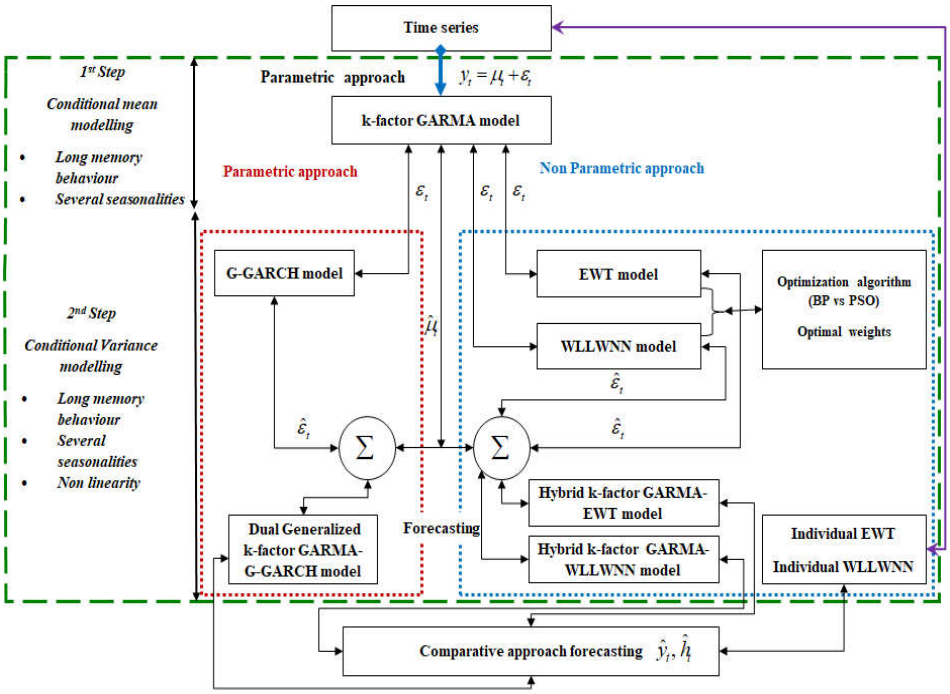

Figure 1. The implemented hybrid procedure.

4. Empirical Methodology

4.1. Data Description and Preliminary Study

After the liberalization of the electricity sector in 1999, the Polish power market was established, with electrical power being handled like any other commodity in a competitive market (instead the monopolistic market). The technique provided in this study is validated using hourly spot prices on the Polish power market between 10 January 2021 and 10 August 2022. Overall $T = 5390$ hourly observations are shown in Figure 2. This information was taken from the Polish Power Exchange's official website. We look at the logarithm of these series in this study because using the difference logarithm can make the series stationary and allow modeling of returns series ($R_t = \Delta LogP_t$). As a result, we investigate the statistical and econometric characteristics of log-return electricity spot price series.

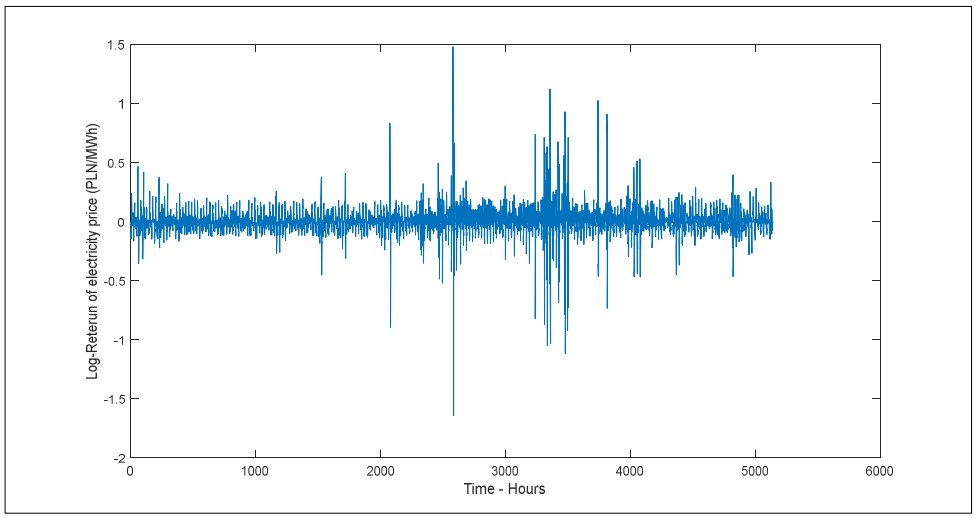

Figure 2. Hourly log-return of spot price for Polish electricity market.

The log-return electricity price (abbreviated L-REP) shown in Figure 2 seem to fluctuate randomly around zero, while the variance varies over time with the alternation of volatile and tranquil periods. Moreover, L-REP series appears to have remained stable throughout time. Furthermore, the application of standard unit root tests and unit root tests show evidence of stationarity[2]. The return series illustrated in Figure 1 seem to fluctuate randomly around zero while the variance varies over time with alternation of volatile and tranquil periods.

Table 1 shows the L-descriptive REP's data in summary form. The standard deviation is low, indicating a non-symmetric distribution, whereas the series exhibit negative skewness and show excess kurtosis. Furthermore, the high kurtosis statistic indicates that the underlying data are leptokurtic. The observed asymmetry may indicate the presence of nonlinearities in the evolution process of returns. The departure from normality is confirmed by the Jarque–Bera test. Indeed, this test strongly rejects the null hypothesis of normality for the series, which means that minimum and maximum values deviate in higher number from the mean calculated. These findings clearly show that the probability of observing extremely negative and positive realizations for our return series is higher than that of a normal distribution.

Table 1. Descriptive statistics of the spot prices time series.

The Log-Returns Electricity Price	
Mean	-1.5478×10^{-5}
Standard Deviation	0.1079
Skewness	−0.2634
Excess Kurtosis	37.2358
Jarque–Bera	2.9826×10^{5} ***

Note: *** denotes significance at 1% level.

As shown in Figure 3, for the log-return electricity price series (L-REP), the spectral density, traced by the periodogram, displays numerous peaks at equidistant frequencies, which demonstrates the presence of many seasonalities.

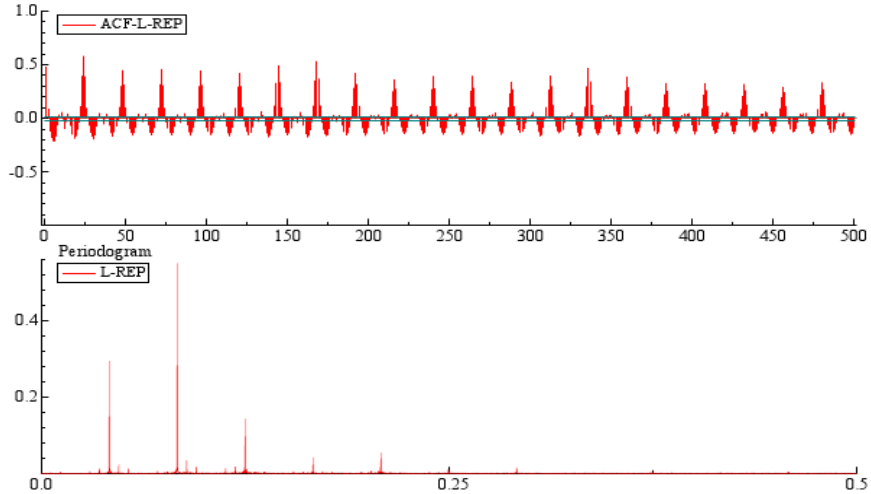

Figure 3. Polish L-REP ACF and Periodogram.

In addition, we use the GPH (Geweke and Porter-Hudak 1983) and local Whittle (LW) (Robinson 1995) semi-parametric techniques approach. For the GPH and LW tests, we need to choose a bandwidth, balancing a high variance caused by staying too close to the origin and using too little information and a bias induced by the contamination of the estimation through the short-memory component of the process. We apply three different bandwidths $T^{0.6}$, $T^{0.7}$ and $T^{0.8}$, where T is the sample size. The results of the GPH and LW tests do not prove to be sensitive to the choice of the bandwidth. Table 2 indicates that the log-return electricity price is stationary and mean reverting.

Table 2. Results of GPH and LW tests in the conditional mean.

	Bandwidth	GPH			LW		
		$\hat{d}_m$	Standard error	p-value	$\hat{d}_m$	Standard error	p-value
L-REP	$T^{0.6}=169$	−0.51683	0.0558	0.0000	−0.5588	0.0403	0.0000
	$T^{0.7}=397$	−0.61336	0.0372	0.0000	−0.5775	0.0269	0.0000
	$T^{0.8}=934$	−0.4786	0.0253	0.0000	−0.5892	0.0178	0.0000

4.2. Estimation Results

4.2.1. The k-Factor GARMA Estimation Results

The seasonality can be detected without difficulty in the frequency domain $\lambda_i = 1/T$, where λ is the frequency of the seasonality and T is the period of seasonality. As revealed, the spectral densities, denoted in the periodogram (see Figure 3), are unrestrained at equidistant frequencies, which demonstrates the existence of numerous seasonalities. The estimation results of the k-factor GARMA model, using the wavelets method, are reported in Table 3. This result illustrates unusual peaks at frequencies $\hat{\lambda}_{m;1} = 0.0417$ ($T = 23.9808 \approx 1$ day), $\hat{\lambda}_{m,2} = 0.0834$ ($T = 11.9904 \approx 1/2$ day) and $\hat{\lambda}_{m,3} = 0.1262$ ($T = 7.9239 \approx 1/3$ day), corresponding to cycles with daily, semi-daily and third-daily periods, respectively. The results reveal that the k-factor GARMA adaptation appears to

be the most satisfactory representation to describe the seasonal long-memory behavior of the L-REP series.

Table 3. Estimation of the *k*-factor GARMA model: a wavelet-based approach.

Parameters	k *-Factor GARMA Model Estimation*
$\hat{\Phi}$	0.5132 ***
$\hat{\Theta}$	-
μ	-
$\hat{d}_{m,1}$	0.1687 ***
$\hat{d}_{m,2}$	0.2349 ***
$\hat{d}_{m,3}$	0.4137 ***
$\hat{\lambda}_{m;1}$	0.0417 ***
$\hat{\lambda}_{m,2}$	0.0834 ***
$\hat{\lambda}_{m,3}$	0.1262 ***

Notes: *** denotes significance at 1% level.

The next step consists of modeling the conditional variance, so the residuals of the k -factor GARMA estimation are formed using a new EWLLWNN as a first approach, and then addressed using the generalized GARCH model, known as G-GARCH, as a second approach, to determine the appropriate strategy.

4.2.2. The EWLLWNN Estimation Results

The residuals from the k -factor GARMA modeling are used as input for the innovative EWLLWNN and WLLWNN models to assess the conditional variance. All the inputs are normalized within a range of [0, 1] using the following formula before being applied to the network to eliminate the risk of coupling among distinct inputs and to accelerate convergence, which is regarded as the most generally used data smoothing approach: $y_{norm} = \frac{y_{org} - y_{\min}}{y_{\max} - y_{\min}}$, where y_{norm} is the normalized value, y_{org} is the original value and $y_{\min}$ and $y_{\max}$ are the minimum and maximum values of the corresponding residuals data, respectively. These normalized data are then decayed by means of the MODWT[3] with Daubechies least asymmetric (La) wavelet filter of length $L = 8$ ($La(8)$). This wavelet filter has been recurrently assumed in the financial literature and it has been shown that $La(8)$ offers the greatest performance for the wavelet time-series decomposition. Our MODWT decomposition goes up to level $J = 12$ which is specified by $J \le \log_2\left[\frac{T}{L-1} + 1\right]$, i.e., where T denotes the distance of the given time series and L denotes the length of the filter (see Percival and Walden 2000; and Gençay et al. 2002). Indeed, our time series is decomposed into 12 components.

4.2.3. The LLWNN Modeling

The datasets are split into three parts: (a) a sample of 500 observations to start the network training, (b) a training set (4815 observations) and (c) a test set (72 observations). The forecasting experiment is carried out using an iterative forecasting method across the test set, with the model forecasting for 6, 12, 24, 48 and 72 h ahead. The parameters are randomly initialized at first to discover the optimum neural network design. The parameters are then modified using two separate algorithms: the back-propagation algorithm (BP) and the particle swarm optimization algorithm (PSO) to reduce the error between the output values and the real values throughout the network's training.

4.2.4. The k -Factor GARMA-G-GARCH Estimation Results

The spectral distribution, Figure 4, traced by the periodogram, for the squared residuals of the k -factor GARMA model (the squared log-returns are used as a proxy for the associated volatility) exhibits numerous peaks at equidistant frequencies, proving the existence of various seasonalities. The findings of the GPH and LW demonstrate the existence of long memory in the conditional variance, as shown in Table 4.

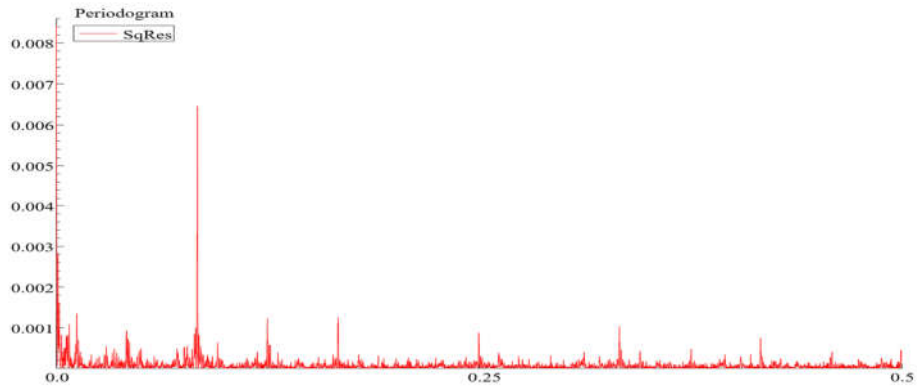

Figure 4. Periodogram of the squared residuals of the k -factor GARMA model.

The G-GARCH model is used to evaluate the seasonal long-memory behavior in the conditional variance using the residuals from the k-factor GARMA. At equidistant frequencies, the spectral densities are unbounded, indicating the presence of multiple seasonality's. The estimation results reported in Table 5, indicating the long-range dependence in squared residuals of the GARMA process and strongly supporting the estimation of dynamic returns that allow for time-varying correlations. Moreover, the three different estimates of the long memory of the conditional variance are close to one another and less than 0.5, implying that the squared returns series is stationary and mean reverting, and long-memory processes require the use of some fractionally integrated method to estimate such processes.

Table 4. Results of GPH and LW tests in the conditional variance.

	Bandwidth	GPH			LW		
		$\hat{d}_v$	Standard error	p-value	$\hat{d}_v$	Standard error	p-value
L-REP	$T^{0.6} = 169$	0.5275	0.0523	0.0000	0.4982	0.0381	0.0000
	$T^{0.7} = 397$	0.3678	0.0329	0.0000	0.3874	0.0246	0.0000
	$T^{0.8} = 934$	0.2379	0.0218	0.0000	0.2453	0.0249	0.0000

The results of estimation of the *k*-factor GARMA-G-GARCH model are described in Table 5. They show special peaks at frequencies $\hat{\lambda}_{v;1} = 0.0411$ (≈1 day), $\hat{\lambda}_{v,2} = 0.0854$ (≈1/2 day) and $\hat{\lambda}_{v,3} = 0.1197$ (≈1/3 day), which correspond to cycles with daily, semi-daily and third-daily periods, respectively.

Table 5. The k -factor GARMA-G-GARCH estimation results.

k **-Factor GARMA Model Estimation**		**The G-GARCH Model Estimation**	
$\hat{\Phi}$	0.6327 ***	$\hat{\psi}$	0.7156 ***
$\hat{\Theta}$	-	$\hat{\beta}$	0.5139 ***
μ	-	$\hat{\gamma}$	-
$\hat{d}_{m,1}$	0.1698 ***	$\hat{d}_{v,1}$	0.1497 ***
$\hat{d}_{m,2}$	0.2568 ***	$\hat{d}_{v,2}$	0.2678 ***
$\hat{d}_{m,3}$	0.3978 ***	$\hat{d}_{v,3}$	0.4319 ***
$\hat{\lambda}_{m;1}$	0.0405 ***	$\hat{\lambda}_{v;1}$	0.0411 ***
$\hat{\lambda}_{m,2}$	0.0838 ***	$\hat{\lambda}_{v,2}$	0.0854 ***
$\hat{\lambda}_{m,3}$	0.1289 ***	$\hat{\lambda}_{v,3}$	0.1197 ***

Notes: *** denotes significance at 1% level.

4.3. Forecasting Results: A comparative Approach

In a multi-step-ahead forecasting task, this section is devoted to evaluating the estimated models. We prefer to use out-of-sample criteria because predicting is essentially an out-of-sample problem. To ensure the accuracy and robustness of the modeling and forecasting findings, 5 distinct periods (6 h, 12 h, 1 day, 2 days and 3 days) were chosen. To assess the accuracy of the forecasting, we adopt two evaluation criteria, the mean absolute error (MAE) and the mean squared error (MSE), given, respectively, by $MAE = \frac{1}{N-t_1}\sum_{t=t_1}^{N}\left|\left(y_{t+h} - \hat{y}_{t,t+h}\right)\right|$ and $MSE = \frac{1}{N-t_1}\sum_{t=t_1}^{N}\left(y_{t+h} - \hat{y}_{t,t+h}\right)^2$, where N is the number of observations, $N - t_1$ is the number of observations for predictive performance, y_{t+h} is the log-return series through period $t+h$, $\hat{y}_{t,t+h}$ is the predictive log-return series of the predictive horizon h at time t.

To evaluate the suggested hybrid methodology's prediction performance, the k -factor GARMA-WLLWNN was evaluated by two different models: the individual WLLWNN model and the k -factor GARMA-G-GARCH model. In terms of the network training, we used two alternative learning methods (BP and PSO). Furthermore, we used 5 time horizons for forecasting, 6 h, 12 h, 1 day, 2 days and 3 days ahead forecasting, using the MAE and the MSE out-of-sample criteria. Furthermore, we use Diebold and Mariano's (1995) statistical test to conclude that the projections are equally accurate. The test statistics noted (DM) is asymptotically $N(0,1)$ distributed under the null hypothesis of no

difference. However, the DM test requires the loss differential to be covariance stationary. Nevertheless, it may not be strictly necessary in some cases (Diebold 2015). In addition, the DM statistic can be obtained by regressing the loss differential on a constant, using Newey–West standard errors. To estimate the performance of the hybrid methodology forecast, we consider the k -factor *GARMA-G-GARCH* model for the purpose of comparison between the forecast results of all the other models. The significance level at 5% and 1% is displayed by ** and ***, respectively. Moreover, to better approve the predictive performances of our models, we use the model confidence set (MCS), introduced by Hansen et al. (2003), for the comparison of multiple forecast models at once. This model selection method is an innovative way to deal with the issue of selecting the best forecast model(s) using an out-of-sample evaluation under a specified loss function. In this work, we use the block bootstrap procedure and significance level $\alpha = 5\%$ to determine the MCS *p*-values. The forecast models that are included in the 95% MCS are identified by an asterisk on the MSE. The forecast evaluation results are described in Table 6.

According to this table, it can be seen that the individual EWLLWNN-based PSO method outperforms the individual EWLLWNN-based BP algorithm, demonstrating the superiority of the PSO algorithm for neural network model training. This finding can be clarified by the fact that weights are updated in the direction of the negative gradient in BP algorithms. As a result, network training using BP methods has several disadvantages, such as sluggish convergence to a local minimum. Weights are branded by the particle position in the case of training with the PSO algorithm. The velocity and location of these particles are adjusted so as to find personal and global optimal values. This prevents the weights from converging to a local minimum.

Furthermore, the hybrid k-factor GARMA-EWLLWNN model outperforms all other computing approaches, as shown in Table 6. However, from this table, the ARFIMA-HY k-factor GARMA-EWLLWNN seems to be the only model included in the 95% MCS for all horizons under the MSE. In reality, this model combines the advantages of three techniques: first, the semi-parametric k-factor GARMA model, which detects and estimates both long memory and seasonality in the conditional mean; second, the empirical wavelet decomposition can produce a good local representation of the signal in both the time and frequency domains, making it a useful tool for revealing hidden patterns in electricity prices, such as high volatility, corrupted by occasional spikes, and followed by multiple seasonalities; and finally, the LLWNN model's capability as a nonlinear, nonparametric model, as well as its uniqueness in having a wavelet activation function and local linearity, allows it to capture more nuanced characteristics of the data. As a consequence, the suggested hybrid k-factor GARMA-EWLLWNN is a robust tool that can handle the characteristics of power pricing while providing the best forecasting results. In sum, for all the evaluation criteria and forecast time horizons, the k-factor GARMA-EWLLWNN model prediction errors are the smallest.

Table 6. Out of sample forecasts results.

Models	*Criterion*	$h=6$	$h=12$	$h=24$
WLLWNN-based BP algorithm	MAE	6.3134×10^{-5}	2.3486×10^{-5}	7.5648×10^{-5}
	MSE	4.2321×10^{-9}	2.2457×10^{-10}	5.7365×10^{-9}
	DM	1.1545	1.2874	1.3786
EWLLWNN-based BP algorithm	MAE	5.5623×10^{-5}	1.5893×10^{-5}	6.9873×10^{-5}
	MSE	3.4285×10^{-9}	2.0445×10^{-10}	4.8745×10^{-9}
	DM	1.2334	1.2982	1.4684
WLLWNN-based PSO algorithm	MAE	2.3458×10^{-7}	3.9523×10^{-7}	10.2415×10^{-7}
	MSE	6.9982×10^{-14}	2.2341×10^{-13}	1.2134×10^{-12}
	DM	1.3566	1.5562	1.6414

EWLLWNN-based PSO algorithm	MAE	2.2412×10^{-7}	3.5932×10^{-7}	9.9865×10^{-7}
	MSE	6.4532×10^{-14}	2.0126×10^{-13}	1.1625×10^{-12}
	DM	1.3751	1.5875	1.7486
The hybrid k-factor GARMA-WLLWNN-based BP algorithm	MAE	4.4627×10^{-5}	1.4341×10^{-5}	6.6247×10^{-5}
	MSE	1.8653×10^{-9}	1.8968×10^{-10}	4.8791×10^{-9}
	DM	1.9874 **	2.3536 **	2.7861 ***
The hybrid k-factor GARMA-EWLLWNN-based BP algorithm	MAE	4.2358×10^{-5}	1.3652×10^{-5}	6.1762×10^{-5}
	MSE	1.8342×10^{-9}	1.7981×10^{-10}	4.3871×10^{-9}
	DM	2.0143 **	2.5462 **	2.8871 ***
The hybrid k-factor GARMA-WLLWNN-based PSO algorithm	MAE	1.4760×10^{-9}	1.9674×10^{-8}	3.4597×10^{-9}
	MSE	2.2175×10^{-18}	5.3560×10^{-16}	1.5261×10^{-17}
	DM	2.6716 ***	2.9873 ***	3.5728 ***
The hybrid k-factor GARMA-EWLLWNN-based PSO algorithm	MAE	1.4372×10^{-9}	1.9247×10^{-8}	3.2655×10^{-9}
	MSE	2.0716×10^{-18} *	5.1038×10^{-16} *	1.4234×10^{-17} *
	DM	2.7861 ***	3.0837 ***	3.7981 ***
The k-factor GARMA-G-GARCH model	MAE	7.4882×10^{-5}	7.8547×10^{-5}	3.7866×10^{-7}
	MSE	4.5763×10^{-14}	9.5376×10^{-14}	1.8527×10^{-15}
Models	*Criterion*	$h=48$		$h=72$
WLLWNN-based BP algorithm	MAE	1.2834×10^{-4}		3.4326×10^{-4}
	MSE	1.3543×10^{-8}		7.3536×10^{-8}
	DM	1.4932		1.6133
EWLLWNN-based BP algorithm	MAE	1.1382×10^{-4}		2.9546×10^{-4}
	MSE	1.3312×10^{-8}		6.8745×10^{-8}
	DM	1.5623		1.7593
WLLWNN-based PSO algorithm	MAE	2.7863×10^{-7}		3.9472×10^{-7}
	MSE	1.1568×10^{-13}		2.2486×10^{-13}
	DM	1.9645 **		2.2514 **
EWLLWNN-based PSO algorithm	MAE	2.5643×10^{-7}		3.6765×10^{-7}
	MSE	1.1358×10^{-13}		2.1378×10^{-13}
	DM	20342 **		2.3436 **
The hybrid k-factor GARMA-WLLWNN-based BP algorithm	MAE	7.7763×10^{-5}		8.9761×10^{-5}
	MSE	7.1321×10^{-9}		9.1497×10^{-9}
	DM	2.9784 ***		3.4836 ***
The hybrid k-factor GARMA-EWLLWNN-based BP algorithm	MAE	7.6408×10^{-5}		8.4873×10^{-5}
	MSE	7.0265×10^{-9}		8.9870×10^{-9}
	DM	3.1426 ***		3.7625 ***
The hybrid k-factor GARMA-WLLWNN-based PSO algorithm	MAE	4.4379×10^{-9}		2.4687×10^{-8}
	MSE	2.7859×10^{-17} *		7.5843×10^{-16} *
	DM	3.7633 ***		3.8921 ***
	MAE	4.2465×10^{-9}		2.2274×10^{-8}

The hybrid k-factor GARMA-EWLLWNN-based PSO algorithm	MSE	2.7152×10^{-17} *	7.0872×10^{-16} *
	DM	3.9562 ***	4.1201 ***
The k-factor GARMA-G-GARCH model	MAE	4.6851×10^{-8}	5.1257×10^{-8}
	MSE	2.3420×10^{-15}	3.7871×10^{-15}

5. Conclusions

In this research, we combine the parametric k-factor GARMA and the novel EWLLWNN models to produce a combined approach for electricity price forecasting. The k-factor GARMA model is used to evaluate the conditional mean of the time series because it can estimate the periodic long-memory behavior in the data. Second, the residuals from the k-factor GARMA model are utilized as a proxy for the associated volatility, and the empirical wavelet local linear neural network (EWLLWNN) model is used to estimate it. The data were divided into a wavelet-domain constitutional subseries and then injected into the network to construct the set of input variables in this network using empirical wavelet transform and form the proposed EWLLWNN forecasting model. On the other hand, while working with neural networks, it is critical to select the right training method; hence, this research compares two learning algorithms: the BP and PSO algorithms.

The projected hybrid model's performance is assessed by means of data from the Polish electricity markets. To demonstrate the robustness of our suggested hybrid model, it is collated with the dual generalized long-memory k-factor GARMA-G-GARCH model and the individual (E) WLLWNN. The proposed k-factor GARMA-EWLLWNN approach is the best acceptable price forecasting strategy according to the empirical data. Because it produces fewer predicting errors than the other computing techniques, it is preferred. It could be considered a powerful forecasting method, especially when higher forecasting accuracy is required.

Related to other approaches, the innovation in our methodology lies in the combination of the seasonal long-memory estimation process, empirical wavelet decomposition and learning algorithms. Indeed, compared to other time-frequency analyses, the empirical wavelet method is used to improve the prediction accuracy of the artificial structure of the neural network. Moreover, we argue that this approach is a promising tool for modeling and forecasting the price of electricity.

In conclusion, the empirical example proved that the suggested hybrid algorithm can produce a superior predicting performance and that it is the best approach for forecasting the power market. Forecasts derived from benchmark datasets further demonstrate the usefulness of our technique in terms of reliability evaluation and prediction optimization. Finally, our results are of major interest to researchers, regulators and market participants. In fact, our findings may be valuable for energy market players, international asset allocation and risk management as a strategy for predicting power prices.

As a proposal for future work, we suggest comparing our hybrid approach to other artificial neural network algorithms as well as other recently used decomposition procedures in order to validate the stability of the prediction results obtained for different forecast horizons.

Author Contributions: H.B.: Initiated the subject, analysed the data in Matlab, contributed to the methodologies and interpretation, discussion of results, clarification of the analysis and editing of the manuscripts. N.B.: co-Initiated the subject, review of literature, discussion of the results and contributed to the development of the interpretations. The authors read and approved the final manuscript.

Funding: This research received no external funding.

Acknowledgements: The authors express their sincere thanks to the editor and anonymous referee for their helpful comments and suggestions which have led to substantial improvement in the presentation of this paper.

Conflicts of Interest: The authors declare that they have no competing interests.

Notes

1 We refer to Bunn and Karakatsani (2003) and Huisman et al. (2007), among others, for an overview on (hourly specific) day-ahead price characteristics.

2 To check the stationarity, we apply the unit root tests without and with structural breaks. We find evidence of stationarity. These results are not reported here but are available upon request.

3 For more details, see Daubechies (1992) and Gençay et al. (2002).

References

Abbass, Hussein A., Ruhul Sarker, and Charles Newton. 2001. A Pareto-frontier Differential Evolution Approach for Multi-objective Optimization Problems. *IEEE Congress on Evolutionary Computation* 2: 971–978.

Aggarwal, S. K., L. M. Saini, and Ashwani Kumar. 2008. Price Forecasting Using Wavelet Transform and LSE Based Mixed Model in Australian Electricity Market. *International Journal of Energy Sector Management* 2: 521–46. https://doi.org/10.1108/17506220810919054.

Anbazhagan, S., and N. Kumarappan. 2014. Day-Ahead Deregulated Electricity Market Price Forecasting Using Neural Network Input Featured by DCT. *Energy Conversion and Management* 78: 711–19. https://doi.org/10.1016/j.enconman.2013.11.031.

Armano, Giuliano, Michele Marchesi, and Andrea Murru. 2005. A Hybrid Genetic-Neural Architecture for Stock Indexes Forecasting. *Information Sciences* 170: 3–33.

Athanassios, Kintsakis, Antonios Chrysopoulos, and Pericles A. Mitkas. Electrical and Computer Engineering Dept., Aristotle University of Thessaloniki, Thessaloniki, Greece. 2015. Agent-based Short-Term Load and Price Forecasting Using a Parallel Implementation of an Adaptive PSO-Trained Local Linear Wavelet Neural Network. Paper presented at International Conference on the European Energy Market, Lisbon, Portugal, May 19–22.

Baillie, Richard T., Bollerslev Tim Mikkelsen and Hans Ole Mikkelsen. 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74: 3–30

Bashir, Zidan and M. E. El-Hawary. 2000. Short Term Load Forecasting by Using Wavelet Neural Networks. Paper presented at 2000 Canadian Conference on Electrical and Computer Engineering. Conference, Halifax, NS, Canada, May 7–10, vol. 1, pp. 163–66. https://doi.org/10.1109/ccece.2000.849691

Ben Amor, Souhir, Heni Boubaker, and Lotfi Belkacem. 2018. Forecasting Electricity Spot Price for Nord Pool Market with a Hybridk-Factor Garma-LLWNN Model. *Journal of Forecasting* 37: 832–51. https://doi.org/10.1002/for.2544.

Benaouda, D., F. Murtagh, J.-L. Starck, and O. Renaud. 2006. Wavelet-Based Nonlinear Multiscale Decomposition Model for Electricity Load Forecasting. *Neurocomputing* 70: 139–54. https://doi.org/10.1016/j.neucom.2006.04.005.

Beran, Jan. 1994. *Statistics for Long Memory Processes*. New York: Chapman & Hall.

Bollerslev, Tim. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31: 307–27.

Bollerslev, Tim, and Hans Ole Mikkelsen. 1996. Modeling and Pricing Long Memory in Stock Market Volatility. *Journal of Econometrics* 73: 151–84. https://doi.org/10.1016/0304-4076(95)01736-4.

Bollerslev, Tim, and Robert Hodrick. 1992. *Financial Market Efficiency Tests*. Cambridge: National Bureau of Economic Research. https://doi.org/10.3386/w4108.

Bordignon, Silvano, Massimiliano Caporin, and Francesco Lisi. 2007. Generalised Long-Memory GARCH Models for Intra-Daily Volatility. *Computational Statistics & Data Analysis* 51: 5900–12. https://doi.org/10.1016/j.csda.2006.11.004.

Bordignon, Silvano, Massimiliano Caporin, and Francesco Lisi. 2008. Periodic Long-Memory GARCH Models. *Econometric Reviews* 28: 60–82. https://doi.org/10.1080/07474930802387860.

Boubaker, Heni, and Mohamed Boutahar. 2011. A Wavelet-Based Approach for Modelling Exchange Rates. *Statistical Methods & Applications* 20: 201–20. https://doi.org/10.1007/s10260-010-0152-x.

Boubaker, Heni, and Nadia Sghaier. 2015. Semiparametric Generalized Long-Memory Modeling of Some MENA Stock Market Returns: A Wavelet Approach. *Economic Modelling* 50: 254–65. https://doi.org/10.1016/j.econmod.2015.06.027.

Boubaker, Heni, Souhir, Ben Amor, and Hichem Rezgui. 2020. A New Hybrid Wavelet-Neural Network Approach for Forecasting Electricity. *Energy Studies Review* 24. https://doi.org/10.15173/esr.v24i1.4135.

Boubaker, Heni. 2015. Wavelet Estimation of Gegenbauer Processes: Simulation and Empirical Application. *Computational Economics* 46: 551–74. https://doi.org/10.1007/s10614-014-9471-6.

Boubaker, Heni. 2016. A Comparative Study of the Performance of Estimating Long-Memory Parameter Using Wavelet-Based Entropies. *Computational Economics* 48: 693–731. https://doi.org/10.1007/s10614-015-9541-4.

Bunn, Derek W., and Nektaria Karakatsani. 2003. *Modeling and Forecasting Electricity Prices*. London: London Business School.

Burton, Bruce, and Ronald G. Harley. 1994. Reducing the computational demands of continually online trained artificial neural networks for system identification and control of fast processes. Paper presented at 1994 IEEE Industry Applications Society Annual Meeting, Denver, Colorado, USA, October 2–5, vol. 2, pp. 1836–43.

Cao, Liangyue,Yiguang Hong, Haiping Fang and Guowei He. 1995. Predicting Chaotic Time Series with Wavelet Networks. *Physica D: Nonlinear Phenomena* 85: 225–38. https://doi.org/10.1016/0167-2789(95)00119-o.

Caporale, Guglielmo Maria, and Luis Gil-Alana. 2014. Long-Run and Cyclical Dynamics in the US Stock Market. *Journal of Forecasting* 33: 147–61. https://doi.org/10.1002/for.2282.

Caporin, Massimiliano, and Francesco Lisi. 2010. MISSPECIFICATION Tests for Periodic Long Memory GARCH Models. *Statistical Methods and Applications* 19: 47–62. https://doi.org/10.1007/s10260-009-0118-z.

Chakravarty, S., Maya Nayak, and R. Bisoi. 2012. Particle Swarm Optimization Based Local Linear Wavelet Neural Network for Forecasting Electricity Prices. Presented at 2012 International Conference on Energy, Automation and Signal, Bhubaneswar, India, December 28–30. https://doi.org/10.1109/iceas.2011.6147149.

Chen, Yuehui, Jiwei Dong, Bo Yang, and Yong Zhang. 2004. A local linear wavelet neural network. Paper presented at Fifth World Congress on Intelligent Control and Automation (IEEE Cat. No. 04EX788), Hangzhou, China, June 15–19, vol. 3, pp. 1954–57.

Cheng, Hangyang, Xiangwu Ding, Wuneng Zhou, and Renqiang Ding. 2019. A Hybrid Electricity Price Forecasting Model with Bayesian Optimization for German Energy Exchange. *International Journal of Electrical Power & Energy Systems* 110: 653–66. https://doi.org/10.1016/j.ijepes.2019.03.056.

Cheung, Yin-Wong. 1993. Long Memory in Foreign-Exchange Rates. *Journal of Business & Economic Statistics* 11: 93–101. https://doi.org/10.1080/07350015.1993.10509935.

Contreras, Javier, Rosario Espinola, Fransicico J. Nogales, and Antonio J. Conejo. 2003. Arima Models to Predict next-Day Electricity Prices. *IEEE Transactions on Power Systems* 18: 1014–20. https://doi.org/10.1109/tpwrs.2002.804943.

Cristea, Paul, Rodica Tuduce, and Alexandra Cristea. 2000. Time Series Prediction with Wavelet Neural Networks. Paper presented at 5th Seminar on Neural Network Applications in Electrical Engineering, NEUREL 2000 (IEEE Cat. No.00EX287), Belgrade, Yugoslavia, September 27. https://doi.org/10.1109/neurel.2000.902374.

Daubechies, Ingrid. 1992. *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611970104.

Diebold, Francis X. 2015. Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests. *Journal of Business and Economic Statistics* 33: 1–9.

Diebold, Francis X., and Roberto S. Mariano. 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* 13: 253–63. https://doi.org/10.1080/07350015.1995.10524599.

Diongue, Abdou Kâ, Dominique Guégan, and Bertrand Vignal. 2009. Forecasting Electricity Spot Market Prices with a K-Factor GIGARCH Process. *Applied Energy* 86: 505–10. https://doi.org/10.1016/j.apenergy.2008.07.005.

Dominique, Zosso, and Dragomiretskiy Konstantin. 2014. Variational Mode Decomposition. *IEEE Transactions on Signal Processing* 62: 531–44. https://doi.org/10.1109/tsp.2013.2288675.

Engle, Robert F. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50: 987. https://doi.org/10.2307/1912773.

Escribano, Alvaro, J. Ignacio Peña, and Pablo Villaplana. 2011. Modelling Electricity Prices: International Evidence*. *Oxford Bulletin of Economics and Statistics* 73: 622–50. https://doi.org/10.1111/j.1468-0084.2011.00632.x.

Gao, Rong, and Lefteri H. Tsoukalas. 2001. Neural-Wavelet Methodology for Load Forecasting. *Journal of Intelligent and Robotic Systems* 31: 149–57. https://doi.org/10.1023/a:1012205313137.

Garcia, Reinaldo C., Javier Contreras, Macro vanAkkeren, and João Batista C. Garcia. 2005. A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. *IEEE Transactions on Power Systems* 20: 867–74. https://doi.org/10.1109/tpwrs.2005.846044.

Gençay, Ramazan, Faruk Selçuk, and Brandon Whitcher. 2002. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics. Waltham*. Cambridge: Academic Press, pp. 1–14. https://doi.org/10.1016/b978-012279670-8.50004-5.

Geweke, John, and Susan Porter-Hudak. 1983. The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis* 4: 221–38. https://doi.org/10.1111/j.1467-9892.1983.tb00371.x.

Ghosh, Sajal, and Kakali Kanjilal. 2014. Modelling and Forecasting of Day-Ahead Electricity Price in Indian Energy Exchange—Evidence from MSARIMA-EGARCH Model. *International Journal of Indian Culture and Business Management* 8: 413. https://doi.org/10.1504/ijicbm.2014.060367.

Gilles, Jerome. 2013. Empirical Wavelet Transform. *IEEE Transactions on Signal Processing* 61: 3999–4010. https://doi.org/10.1109/tsp.2013.2265222.

Girish, Godekere P. 2016. Spot Electricity Price Forecasting in Indian Electricity Market Using Autoregressive-GARCH Models. *Energy Strategy Reviews* 11–12: 52–57. https://doi.org/10.1016/j.esr.2016.06.005.

Granger, Clive W. 1989. Invited Review Combining Forecasts—Twenty Years Later. *Journal of Forecasting* 8: 167–73. https://doi.org/10.1002/for.3980080303.

Granger, Clive W. and Roselyne Joyeux. 1980. An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis* 1: 15–29. https://doi.org/10.1111/j.1467-9892.1980.tb00297.x.

Gray, Henry L., Nien-Fan Zhang, and Wayne A. Woodward. 1989. On Generalized Fractional Processes. *Journal of Time Series Analysis* 10: 233–57. https://doi.org/10.1111/j.1467-9892.1989.tb00026.x.

Grossi, Luigi, and Fany Nan. 2019. Robust Forecasting of Electricity Prices: Simulations, Models and the Impact of Renewable Sources. *Technological Forecasting and Social Change* 141: 305–18. https://doi.org/10.1016/j.techfore.2019.01.006.

Hansen, Peter Reinhard, Asger Lunde, and James M. Nason. 2003. Choosing the best volatility models: the model confidence set approach. *Oxford Bulletin of Economics and Statistics* 65: 839–61.

Hosking, J. R. 1981. Fractional Differencing. *Biometrika* 68: 165–76. https://doi.org/10.1093/biomet/68.1.165.

Huang, Norden E., Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H. Liu. 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: 903–95. https://doi.org/10.1098/rspa.1998.0193.

Huisman, Ronald, Christian Huurman, and Ronald Mahieu. 2007. Hourly Electricity Prices in Day-Ahead Markets. *Energy Economics* 29: 240–48. https://doi.org/10.1016/j.eneco.2006.08.005.

Jabeur, Sami Ben, Salma Mefteh-Wali, and Jean-Laurent Viviani. 2021. Forecasting Gold Price with the XGBOOST Algorithm and Shap Interaction Values. *Annals of Operations Research*. https://doi.org/10.1007/s10479-021-04187-w.

Jiang, Manrui, Lifen Jia, Zhensong Chen, and Wei Chen. 2020. The Two-Stage Machine Learning Ensemble Models for Stock Price Prediction by Combining Mode Decomposition, Extreme Learning Machine and Improved Harmony Search Algorithm. *Annals of Operations Research* 309: 553–85. https://doi.org/10.1007/s10479-020-03690-w.

Jiang, Ping, Feng Liu, and Yiliao Song. 2017. A Hybrid Forecasting Model Based on Date-Framework Strategy and Improved Feature Selection Technology for Short-Term Load Forecasting. *Energy* 119: 694–709. https://doi.org/10.1016/j.energy.2016.11.034.

Kennedy, James, and Russell Eberhart. 1995. Particle Swarm Optimization. Paper presented at Proceedings of ICNN'95—International Conference on Neural Networks, Perth, WA, Australia, November 27–December 1. https://doi.org/10.1109/icnn.1995.488968.

Khashei, Mehdi, and Mehdi Bijari. 2010. An Artificial Neural Network (p,D,Q) Model for Timeseries Forecasting. *Expert Systems with Applications* 37: 479–89. https://doi.org/10.1016/j.eswa.2009.05.044.

Knittel, Christopher R., and Michael R. Roberts. 2005. An Empirical Examination of Restructured Electricity Prices. *Energy Economics* 27: 791–817. https://doi.org/10.1016/j.eneco.2004.11.005.

Koopman, Siem Jan, Marius Ooms and M. Angeles Carnero. 2007. Periodic Seasonal Reg-Arfima–GARCH Models for Daily Electricity Spot Prices. *Journal of the American Statistical Association* 102: 16–27. https://doi.org/10.1198/016214506000001022.

Liu, Heping, and Jing Shi. 2013. Applying Arma–GARCH Approaches to Forecasting Short-Term Electricity Prices. *Energy Economics* 37: 152–66. https://doi.org/10.1016/j.eneco.2013.02.006.

Mallat, Stéphane. 1999. A Wavelet Tour of Signal Processing. New York: Academic Press.

Mallat, Stéphane G., and Zhifeng Zhang. 1993. Matching Pursuits with Time-Frequency Dictionaries. *IEEE Transactions on Signal Processing* 41: 3397–3415. https://doi.org/10.1109/78.258082.

Nicolaisen, James D., C. W. Richter, and Gerald B. Sheble. 2000. Price Signal Analysis for Competitive Electric Generation Companies. Paper presented at DRPT2000. International Conference on Electric Utility Deregulation and Restructuring and Power Technologies. Proceedings (Cat. No.00EX382), London, UK, April 4–7. https://doi.org/10.1109/drpt.2000.855640.

Panapakidis, Ioannis P., and Athanasios S. Dagoumas. (2016). Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy* 172: 132–51.

Pany, Prasanta Kumar. 2011. Short-Term Load Forecasting Using PSO Based Local Linear Wavelet Neural Network. *International Journal of Instrumentation Control and Automation* 1: 163–68. https://doi.org/10.47893/ijica.2011.1030.

Percival, Donald B., and Andrew T. Walden. 2000. *Wavelet Methods for Time Seriesanalysis*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511841040.

Rana, Mashud, and Irena Koprinska. 2016. Forecasting Electricity Load with Advanced Wavelet Neural Networks. *Neurocomputing* 182: 118–32. https://doi.org/10.1016/j.neucom.2015.12.004.

Robinson, P. M. 1995. Log-Periodogram Regression of Time Series with Long Range Dependence. *The Annals of Statistics* 23. https://doi.org/10.1214/aos/1176324636.

Shafie-khah, M., M. Parsa Moghaddam, and M.K. Sheikh-El-Eslami. 2011. Price Forecasting of Day-Ahead Electricity Markets Using a Hybrid Forecast Method. *Energy Conversion and Management* 52: 2165–69. https://doi.org/10.1016/j.enconman.2010.10.047.

Sharkey, Amanda J. 2002. Types of Multinet System. In *Multiple Classifier System.* Berlin and Heidelberg: Springer, pp. 108–17. https://doi.org/10.1007/3-540-45428-4_11.

Soares, Lacir Jorge, and Leonardo Rocha Souza. 2006. Forecasting Electricity Demand Using Generalized Long Memory. *International Journal of Forecasting* 22: 17–28. https://doi.org/10.1016/j.ijforecast.2005.09.004.

Szkuta, Boguslaw R., L. Augusto Sanabria, and Tharam S. Dillon. 1999. Electricity Price Short-Term Forecasting Using Artificial Neural Networks. *IEEE Transactions on Power Systems* 14: 851–57. https://doi.org/10.1109/59.780895.

Tan, Zhongfu, Jinliang Zhang, Jianhui Wang, and Jun Xu. 2010. Day-Ahead Electricity Price Forecasting Using Wavelet Transform Combined with Arima and GARCH Models. *Applied Energy* 87: 3606–10. https://doi.org/10.1016/j.apenergy.2010.05.012.

Tseng, Fang-Mei, Hsiao-Cheng Yu, and Gwo-Hsiung Tzeng. 2002. Combining Neural Network Model with Seasonal Time Series Arima Model. *Technological Forecasting and Social Change* 69: 71–87. https://doi.org/10.1016/s0040-1625(00)00113-x.

Valenzuela, Olga, I. Rojas, F. Rojas, H. Pomares, L. J. Herrera, A. Guillen, L. Marquez, and M. Pasadas. 2008. Hybridization of Intelligent Techniques and ARIMA Models for Time Series Prediction. *Fuzzy Sets and Systems* 159: 821–45. https://doi.org/10.1016/j.fss.2007.11.003.

Wang, A. J., and B. Ramsay. 1998. A Neural Network Based Estimator for Electricity Spot-Pricing with Particular Reference to Weekend and Public Holidays. *Neurocomputing* 23: 47–57. https://doi.org/10.1016/s0925-2312(98)00079-4.

Weron, Rafał, Ingve Simonsen, and Piotr Wilman. 2004. Modeling Highly Volatile and Seasonal Markets: Evidence from the Nord Pool Electricity Market. In *The Application of Econophysics*. Tokyo: Springer, pp. 182–91. https://doi.org/10.1007/978-4-431-53947-6_25.

Whitcher, Brandon. 2004. Wavelet-Based Estimation for Seasonal Long-Memory Processes. *Technometrics* 46: 225–38. https://doi.org/10.1198/004017004000000275.

Woodward, Wayne A., Qin C. Cheng, and Henry L. Gray. 1998. A K-Factor Garma Long-Memory Model. *Journal of Time Series Analysis* 19: 485–504. https://doi.org/10.1111/j.1467-9892.1998.00105.x.

Wu, Zhaohua, and Norden E. Huang. 2009. Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Advances in Adaptive Data Analysis* 1: 1–41. https://doi.org/10.1142/s1793536909000047.

Yao, S. J., Y. H. Song, L. Z. Zhang, and X. Y. Cheng. Wavelet Transform and Neural Networks for Short-Term Electrical Load Forecasting. *Energy Conversion and Management* 41: 1975–88. https://doi.org/10.1016/s0196-8904(00)00035-2.

Yu, Lean, Shouyang Wang, and Kin Keung Lai. 2005. A Novel Nonlinear Ensemble Forecasting Model Incorporating GLAR and Ann for Foreign Exchange Rates. *Computers & Operations Research* 32: 2523–41. https://doi.org/10.1016/j.cor.2004.06.024.

Zhang, G. Peter. 2003. Time Series Forecasting Using a Hybrid Arima and Neural Network Model. *Neurocomputing* 50: 159–75. https://doi.org/10.1016/s0925-2312(01)00702-0.

Zhang, Jinliang, Yi-Ming Wei, Dezhi Li, Zhongfu Tan, and Jianhua Zhou. 2018. Short Term Electricity Load Forecasting Using a Hybrid Model. *Energy* 158: 774–81. https://doi.org/10.1016/j.energy.2018.06.012.

Zhang, Qinghua, and Albert Benveniste. 1992. Wavelet Networks. *IEEE Transactions on Neural Networks* 3: 889–98. https://doi.org/10.1109/72.165591.