

Article

Long-Term Electricity Load Forecasting Considering Volatility Using Multiplicative Error Model

Swasti R. Khuntia ^{1,*} , Jose L. Rueda ¹ and Mart A.M.M. van der Meijden ^{1,2}

¹ Department of Electrical Sustainable Energy, Delft University of Technology, 2600 GA Delft, The Netherlands; j.l.ruedatorres@tudelft.nl (J.L.R.); m.a.m.m.vandermeijden@tudelft.nl (M.A.M.M.v.d.M.)

² TenneT TSO B.V., 6812AR Arnhem, The Netherlands

* Correspondence: s.r.khuntia@tudelft.nl; Tel.: +31-15-27-84051

Received: 13 September 2018; Accepted: 27 November 2018; Published: 27 November 2018



Abstract: Long-term electricity load forecasting plays a vital role for utilities and planners in terms of grid development and expansion planning. An overestimate of long-term electricity load will result in substantial wasted investment on the construction of excess power facilities, while an underestimate of the future load will result in insufficient generation and inadequate demand. As a first of its kind, this research proposes the use of a multiplicative error model (MEM) in forecasting electricity load for the long-term horizon. MEM originates from the structure of autoregressive conditional heteroscedasticity (ARCH) model where conditional variance is dynamically parameterized and it multiplicatively interacts with an innovation term of time-series. Historical load data, as accessed from a United States (U.S.) regional transmission operator, and recession data, accessed from the National Bureau of Economic Research, are used in this study. The superiority of considering volatility is proven by out-of-sample forecast results as well as directional accuracy during the great economic recession of 2008. Historical volatility is used to account for implied volatility. To incorporate future volatility, backtesting of MEM is performed. Two performance indicators used to assess the proposed model are: (i) loss functions in terms of mean absolute percentage error and mean squared error (for both in-sample model fit and out-of-sample forecasts) and (ii) directional accuracy.

Keywords: load forecast; long-term horizon; multiplicative error model; time-series forecasting; volatility

1. Introduction

Load forecasting in the long-term horizon is important for electric utilities and planners in terms of grid expansion planning, future investments, and revenue analysis for long-term decision-making process. Moreover, it plays a crucial role in the economic and social development of a country (or specific region in case of some utilities). A more realistic range of future generation scenarios can be modeled when the electricity consumption is increasing at a faster rate in this globalizing world. For instance, annual load forecasting is favored among utilities and is one of the common long-term load forecasting problems. It can alleviate the disparity between demand and generation, thereby maintaining the required level of security of supply. Choosing the right forecast horizon for long-term varies from one utility to another based on their policies. Usually a monthly or yearly time-step for one to ten years ahead in long-term load forecasting is helpful in inter-tie tariff setting and long-term grid investment return problems.

It is often difficult to accurately forecast load over such a long planning horizon and it is due to the stochastic nature of demand growth and the influential parameters. Most of these parameters are, by nature, unpredictable and uncontrollable. Some examples are socio-economic development, the occurrence of special events, and/or climatic conditions, and regulatory requirements.

Any considerable deviation in forecast results in over expenditure on generation/transmission infrastructure or energy resource waste. Hence, in order to improve the forecast accuracy in the long-term horizon, attention is needed either in terms of improvement of existing employed techniques or development of a new technique to consider all the aforementioned parameters. Forecast accuracy influences the decision making of generation and transmission companies on their plans to address the future load growth and market volatilities. Based on the forecast, electric utilities coordinate their resources to meet the actual demand using a cost-effective plan. As we look into the future, in the energy consumption scenario, it is expected that electricity will play a major role as we move towards the declined use of de-carbonized heat pumps in the residential sector and the addition of electric vehicles (EVs) and other hybrid vehicles in the transportation sector. Such a transition will have a significant impact on overall load profile. To visualize, Figure 1 depicts the future complexity in the context of load forecast, various players in action and the inter-dependency that needs attention as well. Stochasticity in future scenarios, such as economy (GDP: Gross Domestic Product), demographics (change in energy usage of end-users), energy generation mix (renewable energy generation depends on forecasting of a resource like wind, solar irradiation), and technology advancement, influences the forecast accuracy to a large extent. For utilities, the evolution of prosumer complicates the forecasting methodology. A prosumer (combination of words producer and consumer) is a consumer that has its own production facility (e.g., rooftop solar panel). Because of decentralized power generation, prosumers have evolved recently and played an active role in electricity generation and the provision of grid services. Not to be forgotten is the spatial complexity as the area expands from the distribution system operator (DSO) level to multi-transmission system operator (TSO) level. Decisions about grid development are mostly based on the accuracy of predictions of both the scale and occurring geographical locations of energy consumption in the long-term horizon. Any change in the modification of electric consumption patterns is compensated by financial incentives and/or electricity tariffs. In such cases, accurate forecasting will certainly provide support for the utility (TSO/DSO) to estimate the amount of investment needed.

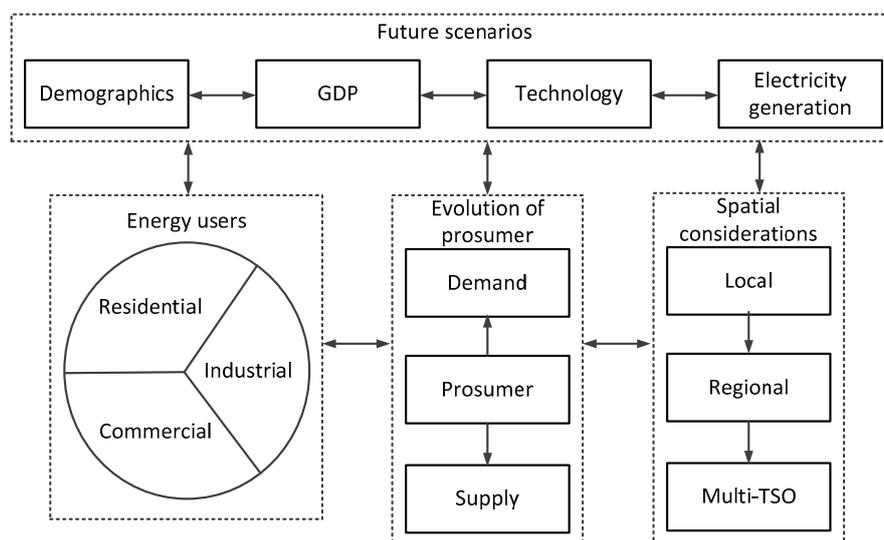


Figure 1. Complexity in long-term electricity load forecast (GDP: Gross Domestic Product, DSO: Distribution System Operator, TSO: Transmission System Operator).

This research sums up the need for accurate load forecasting in the long-term horizon as,

- Firstly, moving towards a greener future is accredited with development in new technology and the integration of renewable energy into primary grid while discarding fossil fuels is becoming important. In the Paris Agreement 2016 [1], it was agreed upon to move towards renewable energy from the more conventional energy. Such a move is realized with an accurate and reliable

forecast of the electrical energy demand. Despite advancements in battery technology, storing energy for the long-term purpose is not the viable option. Thus, accurate and reliable forecasting is required for planning the right tools.

- Secondly, considerable changes in weather factors, like temperature, rainfall, and hot/cold days. Any change in climatic variables will have a direct impact on the demand pattern. Erratic weather events posed due to climate change pose some serious burden on forecasters to accurately model load growth when considering the long-term horizon.
- Thirdly, maintaining the security of supply during the energy transition. In today's date, existing grids are performing under stress to deliver the growing demand in the presence of variable stochastic renewable energy sources.
- Lastly, black swan events, like the great economic recession of 2008, jolted the economic backbone of many countries. Its effect was widespread and energy investment worldwide plunged into tougher financing environment and weakening final demand for energy [2]. This reminds the importance to study the financial aspects of long-term forecasting by energy forecasters in electric utilities.

Based on the needs, the key contributions of this research can be listed as:

- Reviews recent advancement in long-term electricity load forecasting in terms of techniques and models developed.
- Provides a comprehensive and critical evaluation of long-term electricity load forecasting while considering historical volatility.
- Proposes the use of multiplicative error model (MEM) to model conditional mean and to forecast aggregated zonal monthly load. In this research, we consider a forecast horizon of four years as a solution for electric utilities and planners based on the fact that construction of offshore wind farms takes approximately 3–4 years depending on the capacity [3].

The rest of the paper is organized, as follows: Section 2 gives a background on long-term electricity load forecast and the need for accounting volatility, Section 3 introduces multiplicative error model, and is followed by forecast methodology in Section 4. Section 5 analyses the forecast results based on real data. Finally, Section 6 concludes the paper.

2. Background on Long-Term Electricity Load Forecast

Electricity load forecasting in the long-term horizon is an important part of the transformation of electric power systems and it has appealed more and more attention from both academics and industry. By principle, a load forecasting model aims at a mathematical representation of the relationship between load and influential parameters. Such a model is identified with coefficients that are used to forecast future values by extrapolating the relationship to the desired lead time. Eventually, the accuracy of the model depends on both the selected model as well as the accuracy of the estimated parameters. Literature study reveals that long-term load forecasting received less attention when compared to short-term load forecasting. This is because of the complexity involved in achieving an accurate forecast. Long-term load forecasting is based on the integration of concepts from theoretical foundations of economic theory with knowledge on financial, statistical, probability, and applied mathematics to make inferences about the load growth/fall and technology evolution. Reference [4] illustrates rationally the concept of long-term load forecasting and also presents recent development within the electric power industry. Reference [5] performed a study on past, current, and future trends in energy forecasting while stating the trend in spatial, short-term, and long-term load forecasting, and energy price forecasting in a lucid manner. Reference [6] proposed three methods that are suitable for long-term forecasting as time-series approach, econometric approach, and end-use approach. For long-term forecasting, all approaches require historical data and they are broadly categorized into traditional (or statistical) and non-traditional (or artificial intelligence AI) based methodologies. In recent times, the use of probabilistic methods in long-term forecasting is also reported [7,8].

Traditional methods include regression-based model and time-series methods. Reference [9] proposed univariate autoregressive models to forecast load with monthly time-step in Lebanon. Multiple linear regression models were proposed in [10]. Reference [11] implemented a knowledge-based expert system to support the choice of the most suitable load forecasting model with practical application. However, traditional methods are criticized for their weakness of non-linear fitting capability. In AI-based techniques, artificial neural network (ANN) is one of the most popular models. Its application in forecasting Greek long-term energy consumption for the years ahead is reported in [12]. Reference [13] used ANN on the Egyptian electrical network for long-term peak load forecasting. Reference [14] reported the superiority of ANN for medium and long-term load forecasting in terms of accuracy and robustness. Hybrids of fuzzy and ANN are reported in [15] for forecasting Taiwan's annual electricity load and in [16] for long-term electrical energy consumption in India. Use of Bayesian statistics for load modeling is reported in [17] and for long-term forecasting is reported in [18]. Other AI techniques include support vector regression models (SVR) [19,20] and SVR with simulated annealing algorithms [21].

Use of metaheuristic methods, such as genetic programming [22], fruit-fly algorithm [23], gravitational search algorithm [24], and particle swarm optimization (PSO) [25,26] are also reported. Other methods include long-term forecasting based on partial least squares method [27] and complete decomposition method [28]. A recent study includes forecasting for country-specific, such as Spain [29], Greece [12,30], Lebanon [9], and Turkey [31–33]. More recently, reference [34] used gene expression programming for long-term prediction of electrical energy consumption in ASEAN (Association of Southeast Asian Nations, a regional intergovernmental organization comprising ten Southeast Asian countries)-5 countries and projected up to 2030 according to rolling-based forecasting procedure. The results are compared with those that were obtained from ANN, SVR, adaptive neuro-fuzzy inference system (ANFIS), rule-based data mining algorithm, gene expression programming (GEP), linear and quadratic models optimized by PSO, cuckoo search algorithm (CSA), and backtracking search algorithm (BSA).

It was learned that the developed models aim at predicting accurate peak load or electrical energy consumption while comparing with any classical model. However, one aspect that has received less attention in long-term electricity load forecasting when the whole energy scenario is growing in terms of complexity and dynamics is volatility. The concept of volatility is prevalent in financial markets and it refers to the degree of erratic variations of a process over time. It is used as a criterion to study the risk associated with a financial asset. Reference [35] showed that power markets have greater volatility levels than other financial markets, like crude oil, natural gas, or stock prices. Literature study reveals volatility studies on various electricity markets: Spanish, Californian, UK, and PJM (Interconnection, a regional transmission organization (RTO) in the eastern United States that operates one of the world's largest competitive wholesale electricity markets) electricity markets [36], Ontario and some of its neighboring markets [35], German market [37], Australian electricity market [38], to name a few. Reference [39] examined and compared the volatility of 14 deregulated markets through the "price velocity" metric. The Nordic pool was studied in [40] while considering volatility clustering, log-normal distribution, and long-range correlations. In time-series forecasting of electrical load, volatility is defined as a deviation from the mean, which corresponds to risk. An advantage of such an approach is that once the time-series model is understood, it is possible to simulate the data generation for any lead time in the future. Reference [41] explained the importance of volatility in long-term electricity load forecast, which no work reported earlier. Extending the concept of volatility forecast to electricity load forecast in the long-term horizon is adopted in this research.

Volatility is a fundamental issue in financial and econometrics domain, and virtually present in all financial decision making. The concept of volatility in financial markets refers to the degree of unpredictable fluctuations of a process over time. Volatility can be broadly classified into five major types: price volatility, stock volatility, historical volatility, implied volatility, and market volatility [42]. In this study, historical volatility is used to account for implied volatility. In the financial market,

historical volatility is understood as how much volatility that a stock has had over the past time-frame (say, 12 months). If the stock price varied widely in the past 12 months, it is considered to be more volatile and riskier. Implied volatility is understood as how much volatility the stock will have in the future. In fact, volatility is forecastable because of a number of persistent properties: (i) it appears in clusters, (ii) it changes over time and has unusual jumps, (iii) it does not grow to infinity and is persistent in specific time-span, and (iv) it reacts different for an increase or decrease of the considered entity. For instance, load forecast in the long-term horizon takes into account socio-economic factors, like population growth and gross domestic product (GDP), along with explicit factors, like historical load and weather data. Presence of economic factors induces volatility, or what is called as implied volatility. In fact, implied volatility is generally treated to be the best available forecast as it has certain characteristics that can increase the accuracy of forecast values. Likewise electricity load, future volatility prediction is an extremely difficult task as well because the actual realization of the future process volatility will be influenced by events that happen in the future. Thus, it is important to develop a model that can fit the sequence of calm and turbulent periods. Studies reveal that linear regression techniques, for example autoregressive integrated moving average (ARIMA) model, one of the most well-known forecasting models, is inadequate in long-term forecasting task because it suffers from the mean convergence problem. It means that ARIMA forecast converges to the mean of the observations as the forecast horizon grows. To address the short-coming and treating volatility as an influential parameter, the next section introduces the concept of MEM and its application to load forecast in the long-term horizon.

3. Multiplicative Error Model for Long-Term Electricity Load Forecast

Multiplicative error model (MEM) was introduced by Engle in 2002 [43] as an adaptation of the autoregressive conditional duration model [44] to be used for time-series that always receive positive values. Literature study on MEM reveals its application in financial risk and volatility forecasting [45–47]. A search about the application of MEM in electricity load forecasting reveals no information, not even for short-term forecasting, which is common among forecasters. Hence, the proposed model is first-of-its-kind to introduce MEM for electricity load forecasting. As the electric load is always represented as a non-negative time-series, MEM is presumed to be a good fit to forecast. The MEM for a non-negative time-series (y_t) on $[0, +\infty)$ and considering \mathcal{F}_{t-1} as information available for forecasting y_t is written as [43]:

$$y_t = \mu_t \varepsilon_t \quad (1)$$

where, the range of the disturbance ε_t is non-negative on $[0, +\infty)$, unit mean, and unknown constant variance given as $\varepsilon_t | \mathcal{F}_{t-1} \sim D(1, \psi)$ for positive distribution D . μ_t is conditional on \mathcal{F}_{t-1} and positive, described on a parameter vector θ as:

$$\mu_t = \mu(\theta, \mathcal{F}_{t-1}) \quad (2)$$

When \mathcal{F}_{t-1} includes only historical values of the series, μ_t can be generalized as:

$$\mu_t = \delta + \sum_{i=1}^p \alpha_i \mu_{t-i} + \sum_{j=1}^q \beta_j y_{t-j} \quad (3)$$

where, δ is constant, term $\sum_{i=1}^p \alpha_i \mu_{t-i}$ represent an inertial component, and term $\sum_{j=1}^q \beta_j y_{t-j}$ represent more recent observation. Equation (3) is referred to as referenced MEM of order (p, q) . Model specifications can be modified to adapt to the needs of the load forecast. For instance, residuals at t -th observation denoted as $\vartheta_t = y_t - \mu_t$ and $\alpha_1^* = \alpha_1 + \beta_1$, Equation (3) can be written as:

$$y_t = \delta + \alpha_1^* y_{t-1} + \vartheta_t - \alpha_1 \vartheta_{t-1} \quad (4)$$

Equation (4) represents an ARMA model with heteroskedastic errors and is the cornerstone of this modeling approach. The procedure of finding and validating a suitable MEM for a given dataset is discussed in the next section.

4. Forecast Methodology Considering Real Data

In order to realize a suitable long-term forecasting model, one must start with a rich historical database, construct the model, identify the appropriate model, and finally evaluate the forecast results. Figure 2 shows the steps to forecast load using MEM. Since MEM falls under time-series models, we follow the Box-Jenkins methodology of building a model with certain adaptations [48]. Starting from data preparation, the first part involves stationarity checking, data fitting, and model identification while checking various statistical properties of the time-series. Identifying the right model, estimating parameters, and checking the model adequacy falls under this part. In the second part, MEM is validated for forecasting both as the in-sample model fit and out-of-sample prediction. Modeling of MEM starts with the identification of autoregressive and moving average parameters of non-negative time-series that has predictive power regarding the directional change, and later added by persistent error specifications that eventually improves forecast. MEM differs from linear regression models in the sense that the mean equation, which is a scalar factor, is multiplied with the independent and identically distributed (*i.i.d.*) error term. The scalar factor evolves in a conditionally autoregressive manner, and hence is favorable for forecasting. The assumption of *i.i.d.* means that the error terms behave randomly with constant mean and variance over a considered time-horizon. However, in reality, both the original time-series as well as error time-series are highly correlated and do not behave as an *i.i.d.* process.

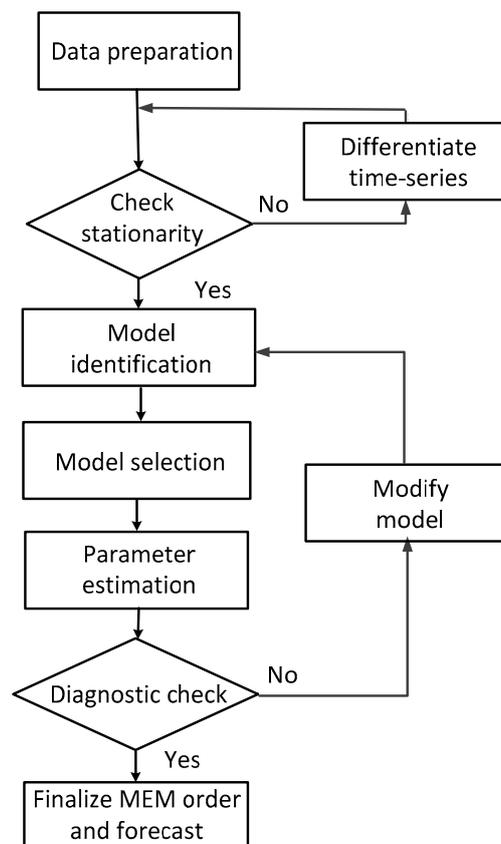


Figure 2. Flowchart for model identification, selection, and forecast.

4.1. Database Generation

Forecast accuracy strongly depends on the quality of available historical data. A poor history, composed only by anomalous or average events, may polarize the analysis and affect the quality of the forecast values. For this study, historical data of a specific load zone region under a United States (U.S.) regional transmission operator [49] and data describing the economic recession as extracted from National Bureau of Economic Research (NBER) [50] is considered. Hourly load data for years 1993–2016 is extracted and sampled to monthly aggregated load, as shown in Figure 3. The use of monthly time-stamp enables in understanding the monthly energy consumption. Recession data for years 1993–2012 is used to build the predictor. From Figure 3, it is evident that the load growth is on a fairly positive side apart from a few incidents where a downturn in demand is observed. Such an incident is the year span of 2006–2009, where the year 2007–2008 is identified by large variability in demand value because of spikes and negative demand growth coincide with the great recession of 2008. The de-seasonalized data, as shown in blue color in Figure 3, help in obtaining a goodness-of-fit measure. It is achieved by dividing the original data by the seasonal factors (12 months as it is monthly aggregated data) to get something that looks more like a straight line.

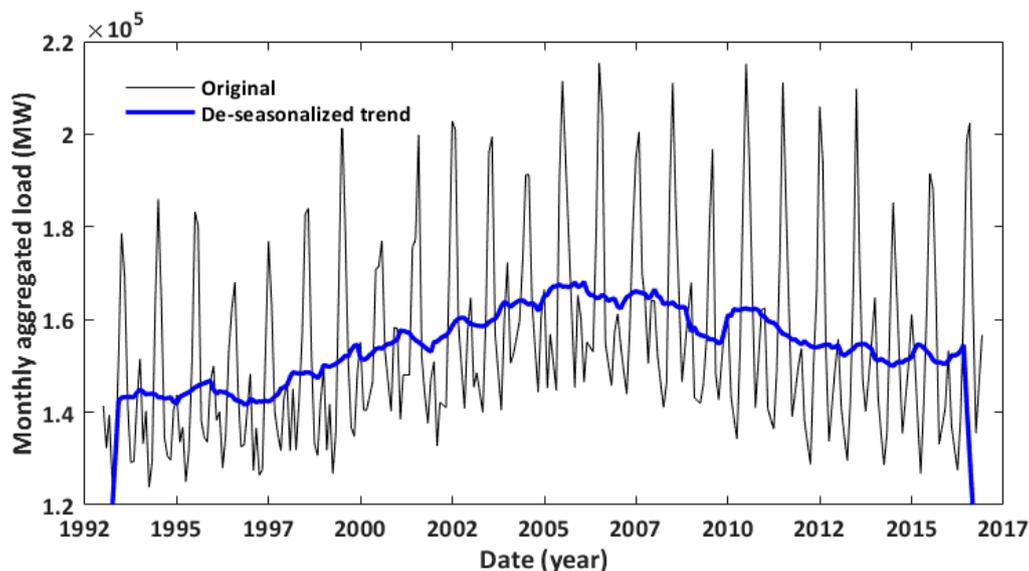


Figure 3. Monthly aggregated load data and de-seasonalized trend for years 1993–2016.

4.2. Stationarity and Autocorrelation Test

A visual inspection of Figure 3 suggests non-stationary time-series with a linear trend and seasonal periodicity. Tests reveal that non-stationarity is apparent as both mean and variance increase with time. The class of MEM requires time-series to be stationary so that its statistical (up to the second order moment) properties do not depend on time. This is coherent with any time-series forecasting because non-stationary time-series are erratic and unpredictable. Phillips-Perron (PP) test is used for stationarity check [51]. For any time-series $x_t = ax_{t-1} + e_t$ where e_t is the residual, PP test checks for the null hypothesis ($H_0 : a = 0$ vs. $H_1 : a \neq 0$). The use of PP test is preferred over the widely used Augmented Dickey-Fuller (ADF), because of its non-parametric nature. In addition to the steps from ADF test, PP test corrects the statistics to account for autocorrelations and heteroscedasticity. The time-series is checked for 0 lags and both the tests reject the null hypothesis with p -value of 0.001. In some cases, unit root tests cannot distinguish highly persistent stationary processes from non-stationary processes very well and thus a modified unit root test is preferred. A more powerful and reliable test for small data by Ng and Perron (2001) is adopted [52] and for further reading [53] is recommended. When compared to the PP test, the Ng and Perron test is argued to exhibit excellent power and size performance and the accompanying code is available here [54]. The test rejects the null

hypothesis and, thus, the time-series is differenced to obtain a stationary time-series. Next step is to determine the presence of autoregressive or moving average terms to correct any autocorrelation that exists in the differenced time-series.

The two tests used to check the null hypothesis (H_0 : no autocorrelation vs. H_1 : correlation) are Ljung-Box Q-test (Q) and Durbin-Watson (DW) test [55]. The Q-test statistic for R residuals, L lags is written as,

$$Q = R(R + 2) \sum_{l=1}^L \left(\frac{\rho(l)^2}{(R - l)} \right) \quad (5)$$

where, $\rho(l)$ is the autocorrelation coefficient at lag l . The Durbin-Watson statistic (DW) is conditioned on the order of the observations (rows) or the number of months in our study. The DW -test statistic for n -observations is written as:

$$DW = \frac{\sum_{i=2}^n (R_i - R_{i-1})^2}{\sum_{i=1}^n R_i^2} \quad (6)$$

Presence of autocorrelation in a time-series indicates that the values of adjacent observations are correlated. Figure 4 shows the autocorrelation function (ACF) and partial autocorrelation function (PACF) plot giving evidence of the presence of autoregressive and moving average parameters. The ACF plot reveals the presence of significantly large autocorrelations, particularly at every 12th lag. The presence of autocorrelation suggests that the data is dependent and correlated and needs modification. Table 1 displays the detailed statistics of original and differenced time-series. Taking a look at the third and fourth moments of distribution (skewness and kurtosis), it is realized that the differenced time-series is close to normal. It is slightly left-skewed, which means that the left tail is longer. Skewness involves the third moment of the distribution and kurtosis involves the fourth moment. The outliers in the distribution, therefore, have even more effect on the kurtosis than they do on the skewness. In a symmetric distribution, both tails increase the kurtosis, unlike skewness, where they offset each other. The mean and standard deviation have the same units as the original data, and the variance has the square of those units. However, the kurtosis, like skewness, has no units: it is a pure number. The standard value of kurtosis of a normal distribution is 3.

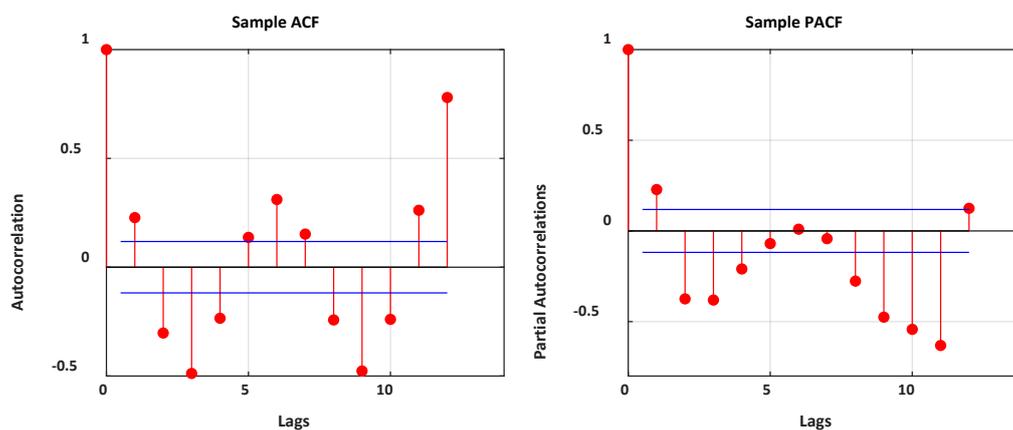


Figure 4. Sample autocorrelation function (ACF) and partial autocorrelation function (PACF) plot of differenced time-series.

Table 1. Detailed statistics of load time-series.

	Original	Differenced
Mean	154,581.1	53.40888
Max	215,379.9	45,842.8
Min	123,786.5	−49,286.9
Median	149,928.9	433.1667
Standard Deviation	20,514.4	17,879.6
Skewness	1.015012	−0.10097
Extra Kurtosis	3.449577	3.005507

4.3. Volatility Check and Multiplicative Error Modeling

The next step in modeling is to check if the differenced time-series shows any cluster of volatility and satisfy the homoscedastic assumption of constant variance or heteroskedastic behavior. It may happen that squared values of the differenced time-series exhibit significant serial correlation. It means that values are again dependent but serially uncorrelated. So, the sample autocorrelation and partial autocorrelation test is repeated for squared residual followed by Q-test and DW-test. The tests re-confirm our model selection [56], and the corresponding plot of autocorrelation and partial autocorrelation function is shown in Figure 5.

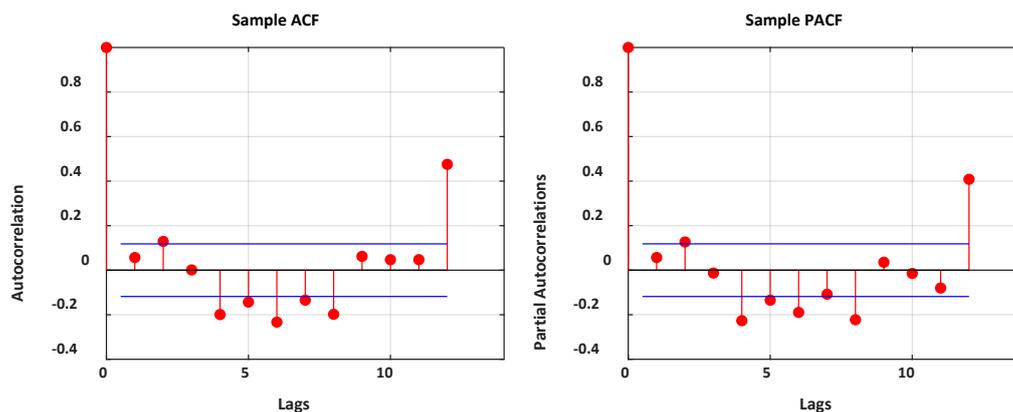


Figure 5. Sample autocorrelation function (ACF) and partial autocorrelation function (PACF) plot of squared residuals.

The ACF and PACF plots in Figure 5 verify the presence of conditional heteroscedasticity and also facilitates in identifying the appropriate order for de-seasonalized differenced time-series. After checking for heteroscedasticity and before proceeding to forecast, we checked the presence non-linearity in our residual by using BDS test [57]. The BDS test (after the initials of W. A. Brock, W. Dechert, and J. Scheinkman) detects non-linear serial dependence in time-series and checks for the null hypothesis that the residuals are *i.i.d.* It is proven to be powerful against a wide range of non-linear alternatives and it does not require any distributional assumptions on the tested data. For further reading, reference [58] is recommended. The BDS code for MATLAB can be accessed at [59]. Upon testing the residuals, the BDS test rejected the null hypothesis, hinting the presence of non-linearity or non-stationarity or other types of structure missed out by de-trending or model fitting in residuals. This is followed by conditional heteroscedasticity modeling.

As stated in [43], generalized autoregressive conditional heteroscedasticity (GARCH) models are a form of MEM and form the basis of the proposed model. With reference to Equation (1), if μ_t is the conditional expectation of y_t w.r.t available information (or historical values), its parameters can be estimated by specifying a GARCH for the conditional second moment of $\sqrt{y_t}$ while imposing its conditional mean to be zero. Reference [60] augmented the regression model with GARCH error

modeling, and the same concept is adapted for this study. The standard model that is common to both the processes and its square while rewriting Equation (1) is:

$$\begin{aligned} Z_t &= \sqrt{h_t} e_t \\ Z_t^2 &= h_t e_t^2 \end{aligned} \quad (7)$$

In the squared equation, the dependent variable (Z_t) is non-negative with conditional mean h and a non-negative multiplicative error term $e_t \sim i.i.d. (0, 1)$ with unit mean. This can be estimated by taking the load residual as the dependent variable of a GARCH model. The GARCH model is an extension of the ARCH model, in the way that it allows current volatility to be dependent on its lagged values directly. For more information on ARCH and GARCH, reference [61] is recommended. The model can be estimated by taking Z_t as the dependent variable, with specifications of zero mean and an error process. In such case, the conditional variance is then the conditional mean of Z_t^2 [62]. Rewriting Equation (1), the GARCH model with order $p \geq 0$ and $q \geq 0$ is defined as [61]:

$$Z_t = \sqrt{h_t} e_t \quad (8)$$

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i Z_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2 \quad (9)$$

for the square root of duration, and where $\alpha_0 > 1$, $\alpha_i \geq 0$, and $\beta_j \geq 0$ are constants with

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1 \quad (10)$$

and $e(t)$ is independent of Z_{t-k} , $k \geq 1$.

Selecting the right order (p, q) is achieved by following one of the many order selection tests. Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) tests are chosen in this study. The reasons for choosing the two criteria are that both of the tests assess the fit between model predicted and original values and penalize models with a larger number of parameters. Tests confirmed the use of order (1, 1) multiplicative error model. Figure 6 shows the innovation plot for a sample size of 101 (0–80 range shown in Figure 6), and it can be concluded that clusters of volatility appear in a periodic manner. The innovation is the difference between the observed value of a variable at time t and the optimal forecast of that value based on information available prior to time t . Thus, the movement of non-linearity is not only dependent on the previous values but for the whole time-series, it is uncorrelated. Innovations display non-seasonal periods of high and low variances. Volatility tends to cluster into periods with higher and lower volatility. This effect proves that volatility at some time must be dependent on its historical values, say with some degree of dependence.

After the model order is identified, maximum likelihood estimator is used to estimating the parameters. Regardless of the low standard errors, parameter estimation is still feasible. As the sample size runs from $N \rightarrow \infty$, the probability that the value of the estimators shows a large divergence from the true (which is unknown) parameter values goes to 0, making it a consistent estimator. Estimation is achieved with conditional variance $h_t \sim i.i.d.(0, 1)$ and with an assumption that error distribution follows student t-distribution, a version of the generalized error distribution, whose density is given as,

$$f(x) = \frac{v e^{(-\frac{1}{2} | \frac{x}{\lambda \mu} |^v) \frac{1}{\mu}}}{\lambda 2^{(1+\frac{1}{v})} \Gamma(\frac{1}{v})} \quad (11)$$

where v is a positive measure of fat tail, $\lambda = \sqrt{2^{-(\frac{2}{v})} \frac{\Gamma(\frac{1}{v})}{\Gamma(\frac{3}{v})}}$, and $\Gamma(\cdot)$ is the gamma function defined as $\Gamma(x) = \int_0^\infty y^{x-1} e^{-y} dy$. This assumption helps in the better modeling of excess kurtosis (in Table 1).

It also approximates the normal distribution as the degrees of freedom grow to infinity. Presence of fat tail is evident from Q-Q plot in Figure 7.

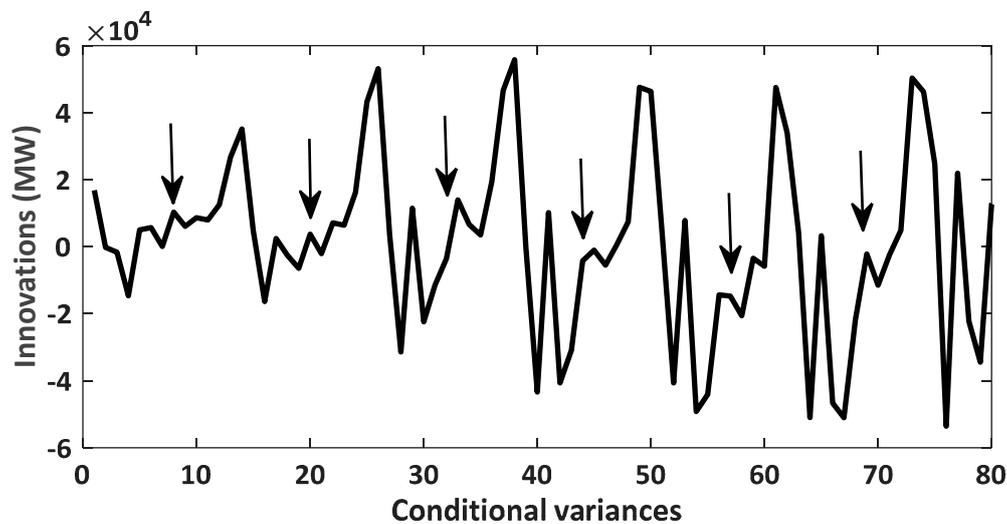


Figure 6. Innovation plot showing clusters of volatility (pointers show the clusters).

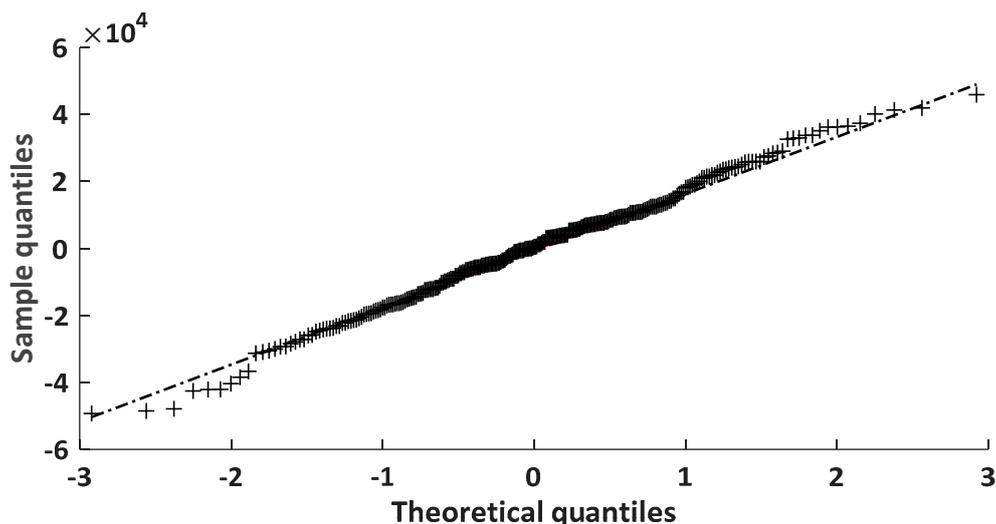


Figure 7. Q-Q plot of residuals.

5. Results and Analysis

Result analysis consists of two parts: first part consist of in-sample model fit using load and economic data followed by out-of-sample forecast, and the second part is checking directional accuracy by forecasting for the year 2008 during the great economic recession.

5.1. In-Sample Model Fit and Out-of-Sample Forecast

A forecast horizon of 4 years is chosen in this study for both in-sample model fit as well as out-of-sample forecast. It is based on the assumption that off-shore wind-farm plant needs at least 3–4 years for completion, which is itself a long-term grid development action [3]. For the in-sample model fit, the study embodies fitting the MEM using load data and recession data of years 1993–2008 and then evaluating its performance on the dataset for years 2009–2012. When assessing point forecasts with mean square errors, it appears to be useful to use a longer in-sample period for model estimation, as followed in this study. Figure 8 shows the in-sample model fit for the years 2009–2012. It is clear from Figure 8 that the in-sample model fit is below the original values for the last three years (2010–2012) but

the year 2009. For the year 2009, the original values show a lower peak aggregated load as compared to model fit prediction. Hereby, it is essential to note that the economy was just reviving after the great economic recession of 2008. Thus, the year 2009 shows a lower peak while the in-sample model fit does not recognize and is higher. However, the model learns and consequently the peak is lower. To check the model accuracy, apart from visual inspection, an expected loss function is required to assess the model performance and check the model accuracy. Use of appropriate loss function also aims at summarizing the accuracy of the point estimate and future distribution. The two loss functions used in this research are Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE), both being unit-free measures. While the optimal point forecast under mean absolute error (without percentage) is the median, MSE represents the (conditional) mean.

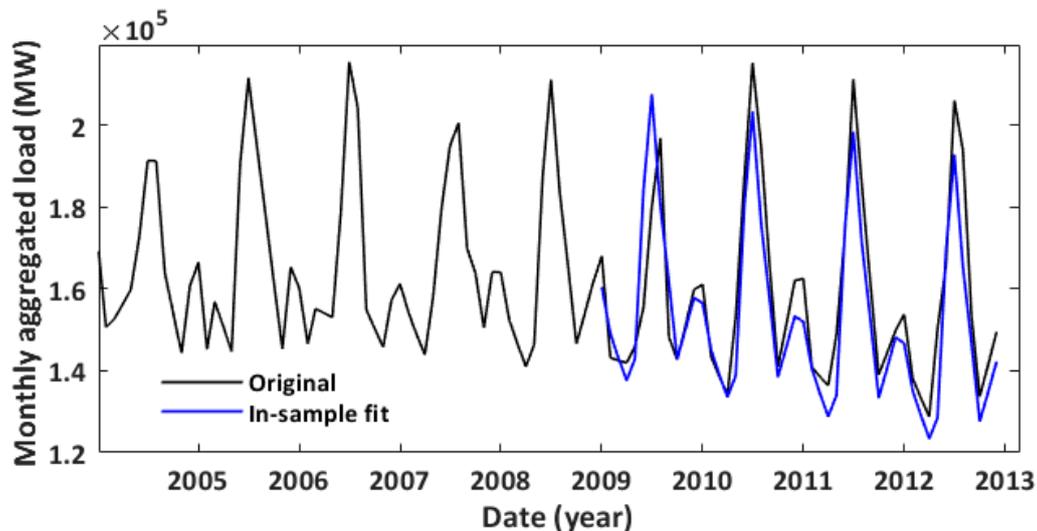


Figure 8. In-sample model fit for years 2009–2012 with a Mean Absolute Percentage Error (MAPE) of 4.98%.

For two sets of n -observations $(x_{i,\dots,n}, y_{i,\dots,n})$, MAPE is defined as

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - x_i}{y_i} \right| \quad (12)$$

where y_i is the original monthly aggregated load and x_i is the predicted monthly aggregated load. In-sample model fit accuracy is achieved with MAPE of 4.98%. MSE is defined as

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2 \quad (13)$$

The MSE, a quadratic and symmetric function, is a measure of how close a fitted line is to data points. The smaller the MSE, the closer the fit is to the data. However, studies reveal that median error measures are not sensitive [63]. Thus, we employ Root Mean Squared Error (RMSE), which is just the square root of the MSE and it measures the deviation in terms of MW. The RMSE describes the magnitude of the error in terms that would be relatively more useful to decision makers. It can be argued that both MAPE and MSE are less appealing, because percentages do not have obvious implications for decision making. While MAPE is scale independent, MSE is more sensitive to a few large errors than to many small errors. In addition, squared error terms may be more difficult for decision makers to understand. The RMSE is calculated as the distance, on average, of a data point from the fitted line, measured along a vertical line. For in-sample model fit, calculated RMSE is 7.7×10^3 MW.

No forecasting analysis is complete without performing out-of-sample forecasts. For better out-of-sample forecasts, the most crucial choice is splitting the series between training and test periods. Unfortunately, no study exists so far that discusses how to choose the decision point [64]. In this study, the training dataset of 1993–2012 is chosen to forecast the next four years (2013–2016). The accuracy of the MEM model is improved with backtesting technique, where the aim is to achieve a dynamic model that can address the future volatilities. With 48 months as forecast horizon and monthly timestamp, the MEM is built every month and forecasts ahead 48 months. The forecast result compares with original values and averages the error. In such a manner, the out-of-sample result improves as the model learns and adapts from past results. Figure 9 shows the out-of-sample forecast results with MAPE of 7.09% and RMSE of 1.09×10^4 MW. A high error percentage and MW error as compared to the in-sample model fit are understood from the long out-of-sample forecast horizon.

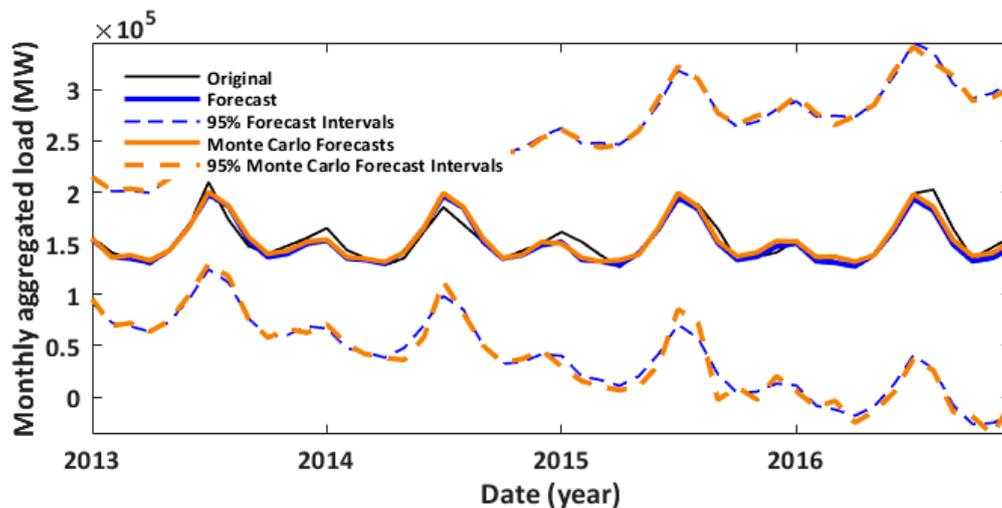


Figure 9. Out-of-sample forecast for years 2013–2016 with 95% confidence interval both upper and lower and a MAPE of 7.09%.

To better evaluate the model accuracy, Monte Carlo simulation is run for 500 sample paths by choosing a confidence interval of 95%. The motivation behind calculating range forecasts this way is to evaluate the likelihood that a particular forecast will be accurate within a specified confidence bound. In this way, the values within the confidence interval of the conditional mean describe the considerable range of values of the point on the line. Thus, the conditional mean for all the values of time-series indicates how much the entire MEM prediction can considerably move from sample to sample. It eases in predicting the range of likelihood values that an observation in the next time step may take. The confidence interval of the out-of-sample forecast presents a range for the mean rather than the distribution of individual data points. Figure 9 shows a comparative analysis of out-of-sample forecast and the Monte Carlo simulation results. Both the forecast as well as confidence intervals from the two outputs are virtually indistinguishable. To understand the intervals, a value of 0.05 corresponds to predicted upper and lower intervals where there is a 5% chance that original values will not be in that range.

It is implicit that forecast error measure increases with forecast horizon. Thus, to check the forecasts, both in-sample model fit as well as out-of-sample forecasts are evaluated by means of their MSE values. A comparison of MSE for the 48 months horizon for both in-sample model fit and the out-of-sample forecast is shown in Figure 10. The error difference is quite large at the beginning of the horizon and increases in the middle, evident in 13–36 months. An exceptional performance of out-of-sample forecast is observed when it outperforms the in-sample model fit results. The error comparison graph reveals that out-of-sample forecasts better reflect the information that is available to the forecaster in real time.

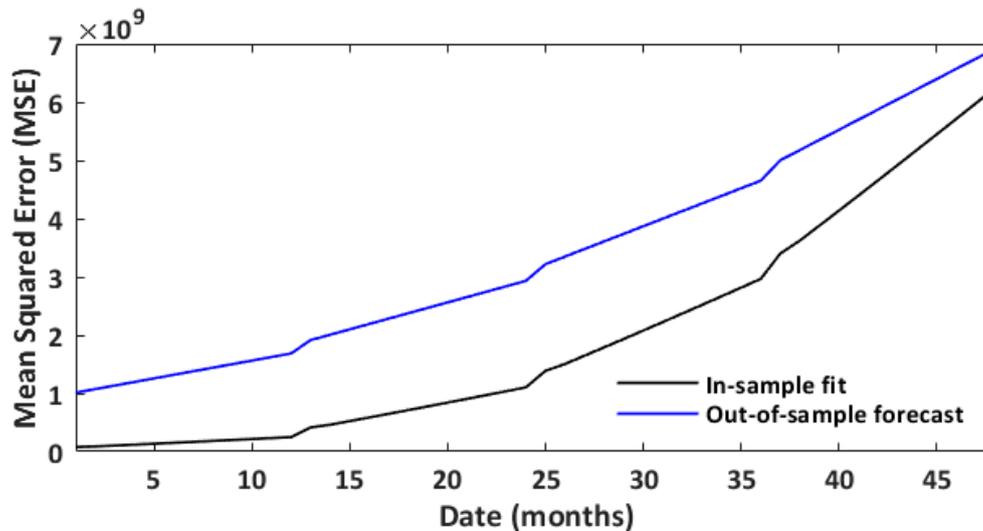


Figure 10. Comparison of Mean Squared Error (MSE) values for in-sample model fit and out-of-sample forecast.

While evaluating the forecast results, we take a glimpse back at Figure 7 and observe that the data is skewed to the right. The Q-Q plot also displays sizeable excess kurtosis or fat tails. Also referring to Table 1, the high skewness and kurtosis value is an indicator of non-normal time-series. To verify the claim, the Jarque–Bera (JB) test is considered in our study. It is usually used for large datasets because other normality tests are not reliable for large datasets. The JB-test verifies the null hypothesis (H_0 : normal vs. H_1 : non – normal distribution). The JB-test statistic is written as [65]:

$$JB = N \left(\frac{s^2}{6} + \frac{(k-3)^2}{24} \right) \quad (14)$$

where, N is the sample size, s is the skewness coefficient, and k is the kurtosis coefficient. A value of 1 from JB-test indicates the data is non-normally distributed. The residual distribution is fitted with Student's t -distribution, which has a thicker tail and is thus more tolerant of extremes. The study is repeated by including both fat-tails and volatility to verify if the forecast improves and the result is significant. Figure 11 shows the forecast for years 2013–2016 with and without accounting for fat-tails. The inclusion of fat-tail is significant, because it represents a greater likelihood of extreme events occurring similar to the financial crisis, also called the black swan event [66]. Some notable features of volatility that should be clearly mentioned are: volatility appears in clusters apparent from Figure 7, volatility changes over time and that jumps in the volatility are unusual, volatility does not grow to infinity; it rather stays within some spans, and the fourth characteristic is that the volatility reacts differently on a drop in the demand than it does for an increase in the demand. The estimated MEM parameters are shown in Table 2. To support the range for in-sample model fit, one of the assumptions in the study is that a t -statistic > 2 in magnitude correspond to approximately a 95% confidence level. The t -statistic column is the parameter value divided by the standard error and is normally distributed for large samples. It measures the number of standard deviations that the parameter estimate is away from zero.

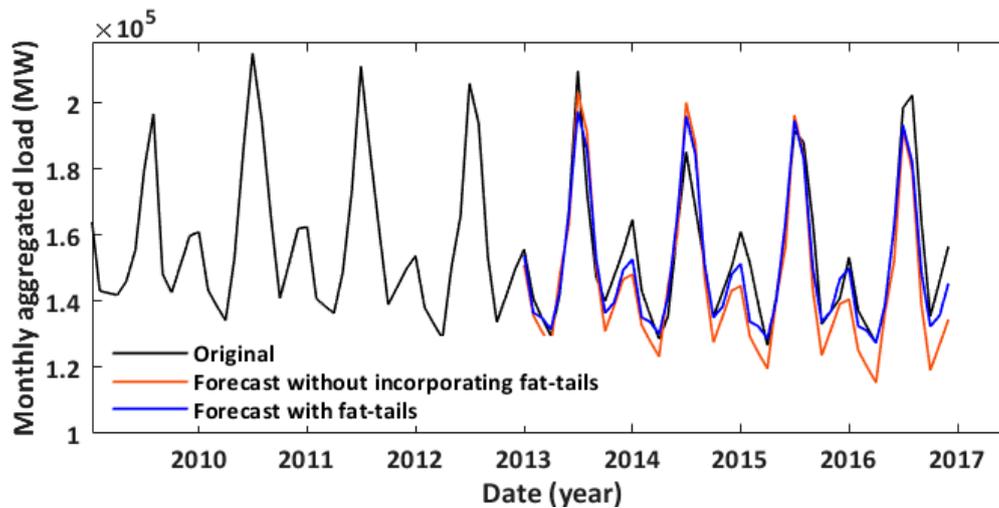


Figure 11. Forecast for years 2013–2016 with and without incorporating fat-tails.

Table 2. Multiplicative Error Model Parameters.

Parameter	MEM Values	Standard Errors	<i>t</i> -Statistic
α	1.368271×10^7	0.000153268	7.22314×10^{10}
α_1	0.703561	0.012491	40.1238
β_1	0.0241376	0.0201932	1.5713

5.2. Directional Accuracy of Forecast Methodology

The second part of result analysis is checking directional accuracy during the year 2008 when the great economic recession hit the whole world and the U.S. was largely affected. Since the data is from U.S. utility, it was decided to check the robustness of the model during that period. The loss functions were used to evaluate the accuracy of the proposed methodology. However, their usage does not distinguish the direction of errors. In other words, the positive forecast errors (i.e., when historical values more than forecasted values) and negative forecast errors (i.e., when historical values less than forecasted values) are counted equally in the error metrics. It can be agreed that ignoring the direction of errors simplifies the efforts of evaluating forecasting accuracy. Nevertheless, it should be noted that the directions of errors often have an economic impact in long-term forecast applications. For example, in power system planning, positive forecast errors (i.e., historical values more than forecast values) can result in planning inadequate capacity, and in turn, loss of load service. On the other hand, negative forecast errors (historical values less than forecast values) can result in wasting of resources by deploying more capacity than necessary. Note that economic loss corresponding to losing load (due to positive errors) is often different from that corresponding to resource wasting (due to negative errors).

An out-of-sample forecast is performed for the year 2008 with a training dataset of years 1993–2007. When economic factors play a pivotal role, the need to study market movements is important. Not many forecast studies include the significance of directional forecasting and how its accuracy supports the statistical parameters. Figure 12 shows the two overlapped time-series. A long period of uniform load growth was interrupted in the early 2000s till mid-2000s. In fact, the 2000s show two distinct jumps in historical load data (seen in Figure 3): one was triggered by energy crisis because of fluctuating oil prices, and one was prompted by the great recession of 2008. Since then, load growth has regularly displayed volatility relative to the pre-2000s. As the real load growth has not changed much over time, still large fluctuations tend to be concentrated over somewhat short periods, thus embodying directional accuracy along with an improved and accurate forecast result is preferred.

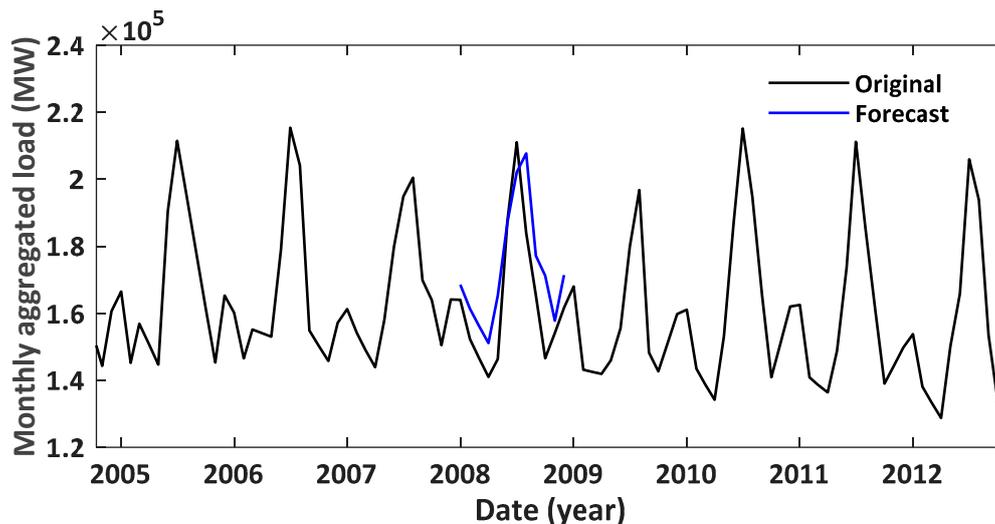


Figure 12. Directional accuracy of multiplicative error model (MEM) during great recession of 2008.

6. Conclusions

This paper proposed a novel implementation of MEM to forecast electricity load in the long-term horizon and used historical volatility to account for implied volatility. The proposed model aimed at accurate forecasting of the monthly aggregated load for a horizon of four years (or 48 months) and the error metrics used to validate the model are mean squared error and mean average percentage error. To account for future volatility, the proposed MEM is built upon recession data accessed from NBER. MEM is able to model the volatility in time-series by accounting for conditional variance. The term conditional variance in MEM denotes the dependency on past sequence of events and is quite contrasting to unconditional, which implies long-term behavior assuming null knowledge on past events. Moreover, the discovered non-linearity could be well handled using MEM. The two performance indicators for this research are: point forecast with a low error percentage as proved by mean error metrics for both in-sample and out-of-sample forecasts, and the directional accuracy during the recession. The inclusion of heteroscedastic errors improved forecast performance and also showed that it is possible to predict the direction of change of residuals in the presence of conditional heteroscedasticity, even if the residuals themselves cannot be predicted.

Author Contributions: S.R.K. contributed to data preparation, methodology formulation, validation and writing. J.L.R. and M.A.M.M.v.d.M. contributed to review and editing.

Funding: The research leading to this result has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 608540 GARPUR project <http://www.garpur-project.eu>. The scientific responsibility rests with the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Available online: http://unfccc.int/paris_agreement/items/9485.php (accessed on 20 July 2018).
2. Available online: <https://www.iea.org/publications/freepublications/publication/impact.pdf> (accessed on 20 July 2018).
3. Khuntia, S.R.; Tuinema, B.W.; Rueda, J.L.; van der Meijden, M.A.M.M. Time-horizons in the planning and operation of transmission networks: An overview. *IET Gen. Trans. Distr.* **2016**, *10*, 841–848. [[CrossRef](#)]
4. Khuntia, S.R.; Rueda, J.L.; van der Meijden, M.A.M.M. Forecasting the load of electrical power systems in mid-and long-term horizons: A review. *IET Gen. Trans. Distr.* **2016**, *10*, 3971–3977. [[CrossRef](#)]
5. Hong, T. Energy forecasting: Past, present, and future. *Foresight* **2014**, *32*, 43–48.
6. Feinberg, E.A.; Genethliou, D. *Load Forecasting*; Springer: New York, NY, USA, 2005.

7. Hong, T.; Fan, S. Probabilistic electric load forecasting: A tutorial review. *Int. J. Forecast.* **2016**, *32*, 914–938. [[CrossRef](#)]
8. Hong, T.; Wilson, J.; Xie, J. Long term probabilistic load forecasting and normalization with hourly information. *IEEE Trans. Smart Grid* **2014**, *5*, 456–462. [[CrossRef](#)]
9. Saab, S.; Badr, E.; Nasr, G. Univariate modeling and forecasting of energy consumption: The case of electricity in Lebanon. *Energy* **2001**, *26*, 1–14. [[CrossRef](#)]
10. Mohamed, Z.; Bodger, P. Forecasting electricity consumption in New Zealand using economic and demographic variables. *Energy* **2005**, *30*, 1833–1843. [[CrossRef](#)]
11. Kandil, M.S.; El-Debeiky, S.M.; Hasanien, N.E. The implementation of long-term forecasting strategies using a knowledge-based expert system: Part-II. *Elect. Power Syst. Res.* **2001**, *58*, 19–25. [[CrossRef](#)]
12. Ekonomou, L. Greek long-term energy consumption prediction using artificial neural networks. *Energy* **2010**, *35*, 512–517. [[CrossRef](#)]
13. Abou El-Ela, A.A.; El-Zeftawy, A.A.; Allam, S.M.; Atta, G.M. Long-term load forecasting and economical operation of wind farms for Egyptian electrical network. *Elect. Power Syst. Res.* **2009**, *79*, 1032–1037. [[CrossRef](#)]
14. Xia, C.; Wang, J.; McMenemy, K. Short, medium and long term load forecasting model and virtual load forecaster based on radial basis function neural networks. *Int. J. Electr. Power Energy Syst.* **2010**, *32*, 743–750. [[CrossRef](#)]
15. Chen, T. A collaborative fuzzy-neural approach for long-term load forecasting in Taiwan. *Comput. Ind. Eng.* **2012**, *63*, 663–670. [[CrossRef](#)]
16. Padmakumari, K.; Mohandas, K.P.; Thiruvengadam, S. Long term distribution demand forecasting using neuro fuzzy computations. *Int. J. Electr. Power Energy Syst.* **1999**, *21*, 315–322. [[CrossRef](#)]
17. Chiodo, E.; Lauria, D. Probabilistic description and prediction of electric peak power demand. In Proceedings of the 2012 Electrical Systems for Aircraft, Railway and Ship Propulsion, Bologna, Italy, 16–18 October 2012.
18. Rivero, C.R.; Sauchelli, V.; Daniel Patiño, H.; Pucheta, J.A.; Laboret, S. Long-term power consumption demand prediction: A comparison of energy associated and Bayesian modeling approach. In Proceedings of the 2015 Latin America Congress on Computational Intelligence (LA-CCI), Curitiba, Brazil, 13–16 October 2015.
19. Hong, W.C. Electric load forecasting by support vector model. *Appl. Math. Model.* **2009**, *33*, 2444–2454. [[CrossRef](#)]
20. Jianjun, W.; Li, L.; Ding, L. Application of SVR with backtracking search algorithm for long-term load forecasting. *J. Intell. Fuzzy Syst.* **2016**, *31*, 2341–2347. [[CrossRef](#)]
21. Pai, P.F.; Hong, W.C. Support vector machines with simulated annealing algorithms in electricity load forecasting. *Energy Convers. Manag.* **2005**, *46*, 2669–2688. [[CrossRef](#)]
22. Karabulut, K.; Alkan, A.; Yilmaz, A.S. Long term energy consumption forecasting using genetic programming. *Math. Comput. Appl.* **2008**, *13*, 71–80. [[CrossRef](#)]
23. Li, H.Z.; Guo, S.; Li, C.J.; Sun, J.Q. A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. *Knowl-Based Syst.* **2013**, *37*, 378–387. [[CrossRef](#)]
24. Abdi, H.; Beigvand, S.D. Long-term load forecasting based on gravitational search algorithm. *J. Intell. Fuzzy Syst.* **2016**, *30*, 3633–3643. [[CrossRef](#)]
25. Unler, A. Improvement of energy demand forecasts using swarm intelligence: The case of Turkey with projections to 2025. *Energy Policy* **2008**, *36*, 1937–1944. [[CrossRef](#)]
26. AlRashidi, M.R.; El-Naggar, K.M. Long term electric load forecasting based on particle swarm optimization. *Appl. Energy* **2010**, *87*, 320–326. [[CrossRef](#)]
27. Meng, M.; Niu, D. Annual electricity consumption analysis and forecasting of China based on few observations methods. *Energy Convers. Manag.* **2011**, *52*, 953–957. [[CrossRef](#)]
28. Sun, J.W. Energy demand in the fifteen European Union countries by 2010—A forecasting model based on the decomposition approach. *Energy* **2001**, *26*, 549–560. [[CrossRef](#)]
29. Moral-Carcedo, J.; Pérez-García, J. Integrating long-term economic scenarios into peak load forecasting: An application to Spain. *Energy* **2017**, *140*, 682–695. [[CrossRef](#)]
30. Angelopoulos, D.; Psarras, J.; Siskos, Y. Long-term electricity demand forecasting via ordinal regression analysis: The case of Greece. In Proceedings of the IEEE PowerTech, Manchester, UK, 18–22 June 2017.

31. Yumurtaci, Z.; Asmaz, E. Electric energy demand of Turkey for the year 2050. *Energy Source* **2004**, *26*, 1157–1164. [[CrossRef](#)]
32. Dilaver, Z.; Hunt, L.C. Turkish aggregate electricity demand: An outlook to 2020. *Energy* **2011**, *36*, 6686–6696. [[CrossRef](#)]
33. Hamzacebi, C.; Es, H.A. Forecasting the annual electricity consumption of Turkey using an optimized grey model. *Energy* **2014**, *70*, 165–171. [[CrossRef](#)]
34. Kaboli, S.H.A.; Fallahpour, A.; Selvaraj, J.; Rahim, N.A. Long-term electrical energy consumption formulating and forecasting via optimized gene expression programming. *Energy* **2017**, *126*, 144–164. [[CrossRef](#)]
35. Zareipour, H.; Bhattacharya, K.; Canizares, C. Electricity market price volatility: The case of Ontario. *Energy Policy* **2007**, *35*, 4739–4748. [[CrossRef](#)]
36. Benini, M.; Marracci, M.; Pelacchi, P.; Venturini, A. Day-ahead market price volatility analysis in deregulated electricity markets. In Proceedings of the 2002 IEEE Power Engineering Society Summer Meeting, Chicago, IL, USA, 25 July 2002.
37. Auer, B.R. How does Germany's green energy policy affect electricity market volatility? An application of conditional autoregressive range models. *Energy Policy* **2016**, *98*, 621–628. [[CrossRef](#)]
38. Boland, J.; Filar, J.A.; Mohammadian, G.; Nazari, A. Australian electricity market and price volatility. *Ann. Oper. Res.* **2016**, *241*, 357–372. [[CrossRef](#)]
39. Li, Y.; Flynn, P.C. Deregulated power prices: Comparison of volatility. *Energy Policy* **2004**, *32*, 1591–1601. [[CrossRef](#)]
40. Simonsen, I. Volatility of power markets. *Phys. Stat. Mech. Appl.* **2005**, *335*, 10–20. [[CrossRef](#)]
41. Khuntia, S.R.; Rueda, J.L.; van der Meijden, M.A.M.M. Volatility in electrical load forecasting for long-term horizon—An ARIMA-GARCH approach. In Proceedings of the IEEE PMAPS, Beijing, China, 16–20 October 2016.
42. Available online: <https://www.thebalance.com/volatility-definition-and-types-3305968> (accessed on 17 October 2018).
43. Engle, R.F. New frontiers for ARCH models. *J. Appl. Econom.* **2002**, *17*, 425–446. [[CrossRef](#)]
44. Engle, R.F.; Russell, J.R. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* **1998**, *66*, 1127–1162. [[CrossRef](#)]
45. Lanne, M. A mixture multiplicative error model for realized volatility. *J. Financ. Econom.* **2006**, *4*, 594–616. [[CrossRef](#)]
46. Han, H.; Park, M.D.; Zhang, S. A multiplicative error model with heterogeneous components for forecasting realized volatility. *J. Forecast.* **2015**, *34*, 209–219. [[CrossRef](#)]
47. Caporin, M.; Rossi, E.; de Magistris, P.S. Chasing volatility: A persistent multiplicative error model with jumps. *J. Econom.* **2017**, *198*, 122–145. [[CrossRef](#)]
48. Box, G.E.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
49. Available online: <http://www.pjm.com/markets-and-operations/ops-analysis/> (accessed on 20 July 2018).
50. National Bureau of Economic Research (NBER). Business Cycle Expansions and Contractions. Available online: <http://www.nber.org/cycles/cyclesmain.html> (accessed on 20 July 2018).
51. Elder, J.; Kennedy, P.E. Testing for unit roots: What should students be taught? *J. Econ. Educ.* **2001**, *32*, 137–146. [[CrossRef](#)]
52. Ng, S.; Perron, P. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* **2001**, *69*, 1519–1554. [[CrossRef](#)]
53. Haldrup, N.; Jansson, M. *Improving Size and Power in Unit Root Testing*; Aarhus University Economics Paper No. 2005-02; Aarhus University: Aarhus, Denmark, 2005.
54. Available online: <http://people.bu.edu/perron/code.html> (accessed on 6 October 2018).
55. Johnson, A.C., Jr.; Johnson, M.B.; Buse, R.C. *Econometrics: Basic and Applied*; Academic Press: New York, NY, USA, 1987; p. 90.
56. Ljung, G.M.; Box, G.E. On a measure of lack of fit in time series models. *Biometrika* **1978**, *65*, 297–303. [[CrossRef](#)]
57. Brock, W.; Scheinkman, J.A.; Dechert, W.D.; LeBaron, B. A test for independence based on the correlation dimension. *Econom. Rev.* **1996**, *15*, 197–235. [[CrossRef](#)]
58. Zivot, E.; Wang, J. *Modeling Financial Time Series with S-PLUS*; Springer: New York, NY, USA, 2007.

59. Kanzler, L. *BDS: MATLAB Module to Calculate Brock, Dechert & Scheinkman Test for Independence*; Statistical Software Components T871803; Boston College Department of Economics: Chestnut Hill, MA, USA, 1998.
60. Ahoniemi, K. *Modeling and Forecasting Implied Volatility—An Econometric Analysis of the VIX Index*; Working Paper; Helsinki Center of Economic Research: Helsinki, Finland, 2006.
61. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [[CrossRef](#)]
62. Brooks, C.; Oozeer, M.C. Modelling the implied volatility of options on long gilt futures. *J. Bus. Financ. Acc.* **2002**, *29*, 111–137. [[CrossRef](#)]
63. Armstrong, J.S.; Collopy, F. Error measures for generalizing about forecasting methods: Empirical comparisons. *Inter. J. Forecast.* **1992**, *8*, 69–80. [[CrossRef](#)]
64. Hansen, P.R.; Timmermann, A. *Choice of Sample Split in Out-of-Sample Forecast Evaluation*; Working Paper; Stanford University: Stanford, CA, USA, 2012.
65. Jarque, C.M.; Bera, A.K. A test for normality of observations and regression residuals. *Int. Stat. Rev.* **1987**, *55*, 163–172. [[CrossRef](#)]
66. Taleb, N.N. Black swans and the domains of statistics. *Am. Stat.* **2007**, *61*, 198–200. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).