Electric Power Systems Research 142 (2017) 58–73

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

A novel multi-time-scale modeling for electric power demand forecasting: From short-term to medium-term horizon

Kianoosh G. Boroojeni[a], M. Hadi Amini[b,c,d,e,*], Shahab Bahrami[f], S.S. Iyengar[a], Arif I. Sarwat[g], Orkun Karabasoglu[b,c,d,e,*]

[a] School of Computing and Information Sciences, Florida International University, Miami, FL 33199, USA
[b] Sun Yat-Sen University – Carnegie Mellon University Joint Institute of Engineering, Guangzhou, Guangdong 510006, China
[c] Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[d] SYSU-CMU Shunde International Joint Research Institute, Shunde, China
[e] School of Electronics and Information Technology, SYSU, Guangzhou, China
[f] Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada
[g] Department of Electrical and Computer Engineering, Florida International University, Miami, FL 33174, USA

ARTICLE INFO

Article history:
Received 22 April 2016
Received in revised form 23 August 2016
Accepted 26 August 2016

Keywords:
Autoregressive model
Moving-average model
Time-series forecasting
Electric power demand forecast
Akaike information criterion
Bayesian information criterion

ABSTRACT

Short-term load forecasting is essential for reliable and economic operation of power systems. Short-term forecasting covers a range of predictions from a fraction of an hour-ahead to a day-ahead forecasting. An accurate load forecast results in establishing appropriate operational practices and bidding strategies, as well as scheduling adequate energy transactions. This paper presents a generalized technique for modeling historical load data in the form of time-series with different cycles of seasonality (e.g., daily, weekly, quarterly, annually) in a given power network. The proposed method separately models both non-seasonal and seasonal cycles of the load data using auto-regressive (AR) and moving-average (MA) components, which only rely on historical load data without requiring any additional inputs such as historical weather data (which might not be available in most cases). The accuracy of data modeling is examined using the Akaike/Bayesian information criteria (AIC/BIC) which are two effective quantification methods for evaluation of data forecasting. In order to validate the effectiveness and accuracy of the proposed forecaster, we use the hourly-metered load data of PJM network as a real-world input dataset.

1. Introduction

1.1. Motivation

Electricity demand forecasting plays a pivotal role in power systems management, especially for operation and maintenance purposes [1]. It is particularly more important for deregulated power systems, where the forecast inaccuracies have significant implications for market operators, transmission owners, and market participants. Load forecasting is categorized based on the time-scale into short-term, medium-term, and long-term forecasting. These three kinds are utilized for power systems scheduling and control, operation and planning, and generation/transmission expansion planning, respectively [1,2]. Short-term load forecasting (STLF) is required for generation scheduling, security assessment of system operation, and hourly economic dispatch information [3]. From the demand side's point of view, there is an exigent requirement to accurately estimate demand to achieve a more reliable operation of the power systems [4,5]. Future power systems, namely smart grids, are emerging with the concept of advanced metering infrastructure to ameliorate the reliability of the conventional power systems in demand side [6,7]. Furthermore, demand side management is widely used for residential [8,9] and industrial load control [10] and smart energy hub applications [11]. Although the utilization of demand response and other demand side resources improves the reliable operation of power systems, accurate demand forecasting is an inevitable obligation to maintain the load-generation balance. The accurate demand forecast will improve the real-time and long-term performance of power systems based on the available historical data.

* Corresponding authors at: Sun Yat-sen University–Carnegie Mellon University Joint Institute of Engineering (JIE), School of Electronics and Information Technology, SYSU, Guangzhou, Guangdong, China; Carnegie Mellon University, Pittsburgh, PA, USA; SYSU-CMU Shunde International Joint Research Institute, Guangdong, China. Tel.: +86 13226996631.
E-mail addresses: kghol002@fiu.edu (K.G. Boroojeni), amini@cmu.edu (M.H. Amini), bahramis@ece.ubc.ca (S. Bahrami), iyengar@cis.fiu.edu (S.S. Iyengar), asarwat@fiu.edu (A.I. Sarwat), karabasoglu@cmu.edu (O. Karabasoglu).

http://dx.doi.org/10.1016/j.epsr.2016.08.031
0378-7796/

Nomenclature

L^i	the lag operator on time-series x_t returning its ith previous value x_{t-i}
D	the differentiation operator on time-series x_t returning $x_t - x_{t-1}$
$\mathrm{corr}(\cdot,\cdot)$	the correlation function which gets two random variables as input and returns their correlations
AC_x	the auto-correlation function of a wide-sense stationary process x_t
PAC_x	the partial auto-correlation function of a wide-sense stationary process x_t
$P_{t,l}(x)$	the projection of time-series x onto the space spanned by $x_{t+1}, x_{t+2}, \ldots, x_{t+l-1}$
X_t	time-series representing the original daily load values
$\hat{X}_t$	time-series representing the forecast daily load values
τ_d	logarithmic Box–Cox transformation for homogenizing the variance of the daily training set
$D^{(d)}_{\mathrm{weekly}}$	weekly differentiating transformation on a time-series with daily values, where $D^{(d)}_{\mathrm{weekly}} = 1 - L^7$
$D^{(d)}_{\mathrm{annual}}$	annual differentiating transformation on a time-series with daily values, where $D^{(d)}_{\mathrm{weekly}} = 1 - L^{365}$
$D^{(h)}_{\mathrm{daily}}$	daily differentiating transformation on a time-series with hourly values, where $D^{(h)}_{\mathrm{daily}} = 1 - L^{24}$
$D^{(h)}_{\mathrm{weekly}}$	weekly differentiating transformation on a time-series with hourly values, where $D^{(h)}_{\mathrm{weekly}} = 1 - L^{168}$
η_t	transformed hourly load values as a stationary time-series
δ_t	transformed daily load values as a stationary time-series
$\mathcal{L}$	maximum likelihood of a forecasting model
$\mathcal{K}$	number of estimated parameters of a forecasting model
$\hat{\delta}_t$	forecast time-series corresponding to δ_t
$\phi^{(h)}_{\cdot}$	hourly auto-regressive coefficient of η_t
$\theta^{(d)}_{\cdot}$	daily moving-average coefficient of η_t
$\theta^{(w)}_{\cdot}$	weekly moving-average coefficient of η_t
$\varepsilon^{(h)}_t$	error values of η_t in ARMA model
$\varepsilon^{(d)}_t$	error values of δ_t in ARMA model
$\psi^{(d)}_{\cdot}$	daily auto-regressive coefficient of δ_t
$\gamma^{(w)}_{\cdot}$	weekly moving-average coefficient of δ_t
$\gamma^{(y)}_{\cdot}$	annual moving-average coefficient of δ_t
$\rho^{(d)}_t$	residual time-series of the day-ahead forecaster
Y_t	time-series representing the original hourly load values
$\hat{Y}_t$	time-series representing the forecast hourly load values
τ_h	logarithmic Box–Cox transformation for homogenizing the variance of the hourly training set
$\rho^{(h)}_t$	residual time-series of the hour-ahead forecaster

1.2. Literature review

A comprehensive survey on demand forecasting approaches has been provided in [12]. This paper classified the methods into four groups: very short-term load forecast (VSTLF) [13,14], STLF [15–17], medium-term load forecast (MTLF) [18], and long-term load forecast (LTLF) [19,20] that a literature survey represented 6%, 58%, 20% and 16% of the past research efforts focused on these load forecast horizons, respectively. The authors in [13] applied artificial neural networks to model load dynamics for VSTLF application. The proposed VSTLF approach was tested for online load forecasting in a power utility in the United States. Bagged neural network is deployed for STLF in [21]. In [14], 5-min moving time windows are used to determine the hourly ahead load. This paper adopted wavelet neural networks with data pre-filtering to forecast the load in very short time slots by minimizing the effect of noisy data. ISO New England data set has been employed for validation of the proposed approach in [14]. Piras et al. in [15] proposed heterogeneous neural network architecture composed of an unsupervised part to detect some features of the data and suggest regression variables. To obtain a smooth transition between submodels, a weighted fuzzy average was deployed to integrate the outputs of each submodel.

Amjady in [16] composed a forecast-aided state estimator (FASE) and the multi-layer perceptron (MLP) neural network to build a short term load forecaster. The method trains the MLP to determine the mapping function which is required for FASE (input features) and the output (real load). Dove et al. implemented different regularization procedures for training purposes in neural networks for medium term load forecasting. This approach referred to as feed forward neural network (FNN) model [18]. In [19], a knowledge-based expert system was developed for choosing the most appropriate annual prediction model. The selected model was then utilized medium/long term power system planning. In [22], kernel-based multi-task learning methods are employed to forecast the electric power demand at the distribution network. Authors in [20], proposed a forecasting method to predict the long-term peak demand. Different uncertainties that can affect the peak demand were considered during the LTLF, including population growth, economic conditions, and weather conditions. Furthermore, large-scale utilization of electric vehicles can increase the uncertainty due to different driving behaviors and charging patterns [23–26].

It also can be used for the power system expansion problems [27].

In addition to the time slot based classification of forecasting methods, we can classify the existing approaches according to the applied techniques. One method is to use historical load pattern as a time-series to forecast the demand using time-series analysis methods. The second method is based on the correlation of load pattern and weather variables. This approach constructs the relation between historical load and weather conditions to forecast demand.

Traditional load forecasting approaches, such as regression and interpolation, may not lead to accurate results. On the other hand, complex forecasting methods, which are computationally-burdened converge slower. Several number of studies have been focused on prediction techniques, including fuzzy logic approach and artificial neural network [16,28], linear regression [29], and data mining [30], transfer functions [31], Bayesian statistics [32], judgmental forecasting [33], and grey dynamic models [34]. In [20], a methodology to forecast electricity demand up to 10 years ahead was proposed. Artificial neural networks (ANN) has been widely used for electricity demand forecasting [35–37]. According to [35], the generalized Delta rule (GDR) was utilized for training neurons in ANN and the output vector is used as an input pattern to the network. In [36], ANN is used for short term demand forecasting. Lee et al. used ANN for weekdays and weekends separately to achieve more accurate results. In [38], artificial neural network is utilized to predict the interruptions in smart grids based on weather condition and historical interruptions in the analyzed network. A comprehensive literature survey of the neural network application in short term demand forecasting is provided in [37]. Amjady in

[39] proposed a short term forecaster that differentiates between weekdays, weekends, and public holidays to improve the accuracy. There are four major concerns related to the previous studies on the demand forecasting: (1) Long process of load forecasting (flow of load forecast process is too long) such as [40]; (2) Small data set is used for model validation [41]; (3) Large amount of data is required in the training phase [42]; and (4) Big error in forecasting [43].

Electricity load demand over a period of time is a seasonal non-stationary time-series. Many attempts have been made by statisticians to create forecasting models for this kind of time-series. Dudek in [44] proposed linear regression models for pattern-based short-term load forecasting in which multiple seasonal cycles of the forecasting time-series is filtered out and non-stationary in mean and variance is eliminated. A step-by-step procedure has been developed [45] for applied seasonal non-stationary time-series modeling following the Box-Jenkins methodology which is a known modeling methodology first created by Box and Jenkins in 1976 [46]. Additionally, Pappas et al. in [47] presented a new method for electricity demand load forecasting using the multi-model partitioning theory, first filters out the seasonality and non-stationary of the actual data using the modeling and forecasting electricity loads and prices (MFE) toolbox for Matlab and then applies a multi-variate auto-regressive moving average (ARMA) model to forecast the electricity demand of the Hellenic power system. Moreover, Desouky and Elkateb in [48] utilized a hybrid of ARMA and ANN methods in order to obtain a more promising forecaster compared with previous works using time-series method. In addition, Huang et al. proposed a short-term load forecaster based on ARMA model including non-Gaussian process considerations [49]. Furthermore, Contreras et al. in [50] analyzed different ARIMA models for predicting next-day electricity load.

1.3. Our contribution

This paper presents a generalized technique for modeling historical load data in the form of time-series with different cycles of seasonality (e.g., daily, weekly, quarterly, annually) in a given power network. The proposed method separately models both non-seasonal and seasonal cycles of the load data using auto-regressive (AR) and moving-average (MA) components, which only rely on historical load data without requiring any additional inputs such as historical weather data (which might not be available in most cases). The accuracy of data modeling is examined not only by calculating the conventional forecasting errors but utilizing the Akaike/Bayesian information criteria (AIC/BIC) which are two effective quantification methods that penalize the complexity of a model and reward its fitness and accuracy. If the forecaster's flexibility is our most important concern, i.e. our main objective is to develop a forecaster which can be broadly used for different data training sets without deteriorating the forecaster's performance, the BIC is a better alternative (compared with AIC) for quantifying the forecaster's utility since BIC penalizes the complexity of forecasting models more than AIC. However, AIC is a better criterion if we choose to compromise model flexibility by some degree to gain more fitness and improve the accuracy of forecaster by reducing the forecasting error. The main contributions of this paper are as follows:

- Enhancing the Box-Jenkins methodology for modeling of the historical electricity load data over a period of time in order to create an accurate electric power demand forecaster.
- Adding multiple seasonality cycles to create a multi-time-scale modeling for electricity power demand forecasting which is more flexible to the daily, weekly, and annual seasonal nature of the electricity data. Consequently, our model can be considered as an extended seasonal auto-regressive integrated moving average model (SARIMA) [51].
- Using Bartlett's periodogram-based test to prove that the residual time-series (forecasting error) is a white noise and is meaningless. That is, the proposed model has extracted any meaningful information from the input training set in order to obtain the most accurate forecast.
- Using Akaike/Bayesian information criteria (AIC/BIC) instead of the conventional error measurements for model evaluation and fine-tuning: AIC/BIC deals with the trade-off between appropriate fitness of the model and the model complexity. It considers some penalty value for complex models, which may only work precisely for the under-studied training set; however, the conventional error measurements are considerably dependent on the given training set and measure the fitness of the model merely with respect to the training set.
- Finally, we show that by changing the size or time of training set, the model remains robust and the forecasting error almost remains constant.

1.4. Organization of the paper

The rest of this paper is organized as follows. Section 2 introduces preliminaries of the time-series modeling. Section 3 presents a detailed explanation of the proposed multi-time-scale model. In Section 4, a practical case study is presented and the superior performance of our novel methodology is illustrated. Section 5 concludes the paper and outlooks the forecasting results. The detailed representation of utilized data-set is provided in the appendix.

2. Preliminaries

A time-series is a sequence of data points, typically consisting of successive measurements made over a time interval. Time-series are very frequently plotted via line charts and are used in any field which involves temporal measurements. Time-series based analysis comprises methods for analyzing data in order to extract meaningful statistics and other characteristics of the data. It also used for forecasting which is based on the utilization of a model to predict future values based on previously observed values. Time-series x_t is wide-sense stationary (WSS) if its mean does not vary over time ($\mathbb{E}[x_t] = \mu_x$) and its autocorrelation function, defined as:

$$R_x(t_1, t_2) = \mathbb{E}[(x_{t_1} - \mu_x)(x_{t_2} - \mu_x)]$$

depends only on the value of $(t_2 - t_1)$ for every $t_1, t_2 \in \mathbb{R}_{\geq 0}$. A seasonal time-series of seasonal cycle T is called stationary over its seasonal cycle if

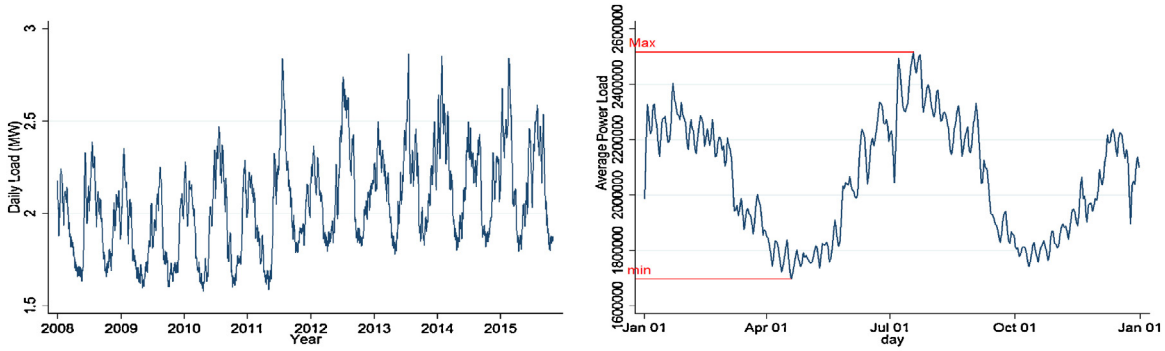
$$\mathbb{E}[x_t] = \mathbb{E}[x_{t+T}],$$

and $R_x(t, t+nT)$ is independent of t for every $t > 0$; e.g. if a time-series with daily value and weekly seasonality is *weekly-stationary* if it has constant autocorrelations on multiples of 7.

2.1. Autoregressive models

In statistics, an auto-regressive (AR) model is a representation of a type of random process showing (wide-sense) stationary behavior. In contrast with the regression analysis which models the output variable as a function of multiple independent variables, the autoregressive model considers the output variable as a linear function of its own previous values and on a stochastic term (a non-deterministic, imperfectly predictable term). Thus, the model is in the form of a stochastic difference equation:

$$x_t = \left(\sum_{i=1}^{p} \phi_i L^i \right) x_t + \varepsilon_t,$$



(a) Seasonality of daily load (big picture). (b) Average annual load data of PJM

Fig. 1. Statistical analysis of the load data of PJM network over a period of 8 years.

where x_t specifies the representing random process of a given time-series, ϕ_i is the coefficient of the ith AR term, L is the lag operator acting on x_t as $L : x_t \mapsto x_{t-1}$, and ε_t denotes the stochastic term called the *error value* of the model.

AR model is a special case of the more general ARMA model of time-series, which has a more complicated stochastic structure. Given a time-series of data x_t where t is an integer index and the x_t are real numbers, an ARMA(p, q) model is given by:

$$x_t = \underbrace{\left(\sum_{i=1}^{p} \phi_i L^i\right) x_t}_{\text{AR-Part}} + \underbrace{\left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t}_{\text{MA-Part}}, \tag{1}$$

where θ_i's are the parameters of the moving average part.

These models are fitted to time-series data to predict future points in the series (forecasting). An ARIMA model is an extension of an ARMA. They are applied in some cases where data are non-stationarity and an initial differentiating step (corresponding to the "integrated" part of the model) can be applied to reduce the non-stationarity. In fact, the ARIMA(p, d, q) model can be viewed as a "cascade" of two models. The first applies on a non-stationary time-series x_t:

$$x'_t = (1 - L)^d x_t$$

while the second part models the WSS time-series x'_t using ARMA(p, q) described in Eq. (1).

When two out of the three terms are zeros, the model may be referred to based on the non-zero parameter, dropping "AR", "I" or "MA" from the corresponding acronym. For example, ARIMA (1,0,0) is AR(1), ARIMA(0,1,0) is I(1), and ARIMA(0,0,1) is MA(1). Before explaining the proposed methodology, we introduce two functions that are assigned to a WSS time-series x_t: auto-correlation function[1] (ACF) and partial auto-correlation function[2] (PACF). The behavior of these two functions reveals that how a WSS time-series can be formulated as an ARMA model. More details regarding how the behavior of ACF and PACF is interpreted is presented in the next section.

[1] The ACF of a WSS time-series x_t is defined as function $\mathrm{AC}_x(l) = \mathrm{corr}(x_{t+l}, x_t)$, where $\mathrm{corr}(\cdot, \cdot)$ is the correlation function defined as $\mathrm{corr}(x_{t+l}, x_t) = \mathbb{E}\left[\frac{(x_t - \mu_x)(x_{t+l} - \mu_x)}{\sigma_x^2}\right]$, where μ_x and σ_x^2 are the mean and variance of stationary process x_t, respectively. Variance of the wide-sense stationary process x_t is defined as $\sigma_x^2 = \mathbb{E}[(x_t - \mu_x)^2]$, where μ_x is the mean of x_t.

[2] The PACF of a wide-sense stationary time-series x_t is represented by $\mathrm{PAC}_x(l)$ and is defined as $\mathrm{AC}_x(l)$ if $l = 1$, and $\mathrm{corr}(x_{t+l} - P_{t,l}(x_{t+l}), x_t - P_{t,l}(x_t))$ otherwise; where $P_{t,l}(x)$ denotes the projection of x onto the space spanned by $x_{t+1}, x_{t+2} \ldots, x_{t+l-1}$.

3. The proposed methodology

The chronologically-ordered power load values of a specific part of the power systems is a time-series that occurs over a period of time in a seasonal manner. To model such time-series as a function of its past values, we analyze the pattern with the assumption that the general pattern will persist in the future. This sections is devoted to identify the best model that matches the statistical behavior that the observed electricity demand data shows over a few years. The quality of modeling the power load time-series determines how accurate the power load data is estimated and forecast in the future. In order to evaluate the model identification, the section utilizes the Akaike/Bayesian information criteria which help us to figure out how close the estimated time-series fits the observed power load data. Analyzing the residual time-series (the actual difference between the observed and estimated load power) can also determine the quality of a model where if the residual time-series is non-deterministic and has no meaningful part, the model accurately represents the observed data. During this section, we utilize the actual daily load data of PJM network over the time period of 2008–2014 as our training-set in order to illustrate the steps of creating the model. Then, we use the load data of the same network in 2015 to evaluate the model and compare the observed and forecast values. Fig. 1 shows a statistical insight into the load data of PJM network.

3.1. Outline of the proposed methodology

First, we check whether the time-series representing the load data has homogeneous variance. If not, we apply an appropriate logarithmic transformation to the time-series to make it variance homogeneous. To obtain the effective transformation, we apply the Box–Cox method [52], which is explained later.

Second, we examine the ACF plot of the transformed time-series at the non-seasonal level and seasonal cycles (weekly, annually, etc.) to find any indication of being non-stationary. If the ACF of the time-series either falls off or declines in the first few lags, it is considered as stationary. If the ACF values either falls off after a considerable number of lags or declines quite slowly, it should be marked as non-stationary. Notice that the ACF of a time-series may show different behaviors in different seasonal cycles; e.g. the ACF of a time-series with weekly lags may show stationary sign; while, considering daily lags, the same ACF may die down very slowly (non-stationary behavior). In order to filter the non-stationary indications, we repeatedly apply differencing transformations on the time-series.

Finally, the transformed stationary time-series is modeled to a moving average and/or autoregressive model based on how ACF and PACF of the time-series behave. Finally, the detailed parameters

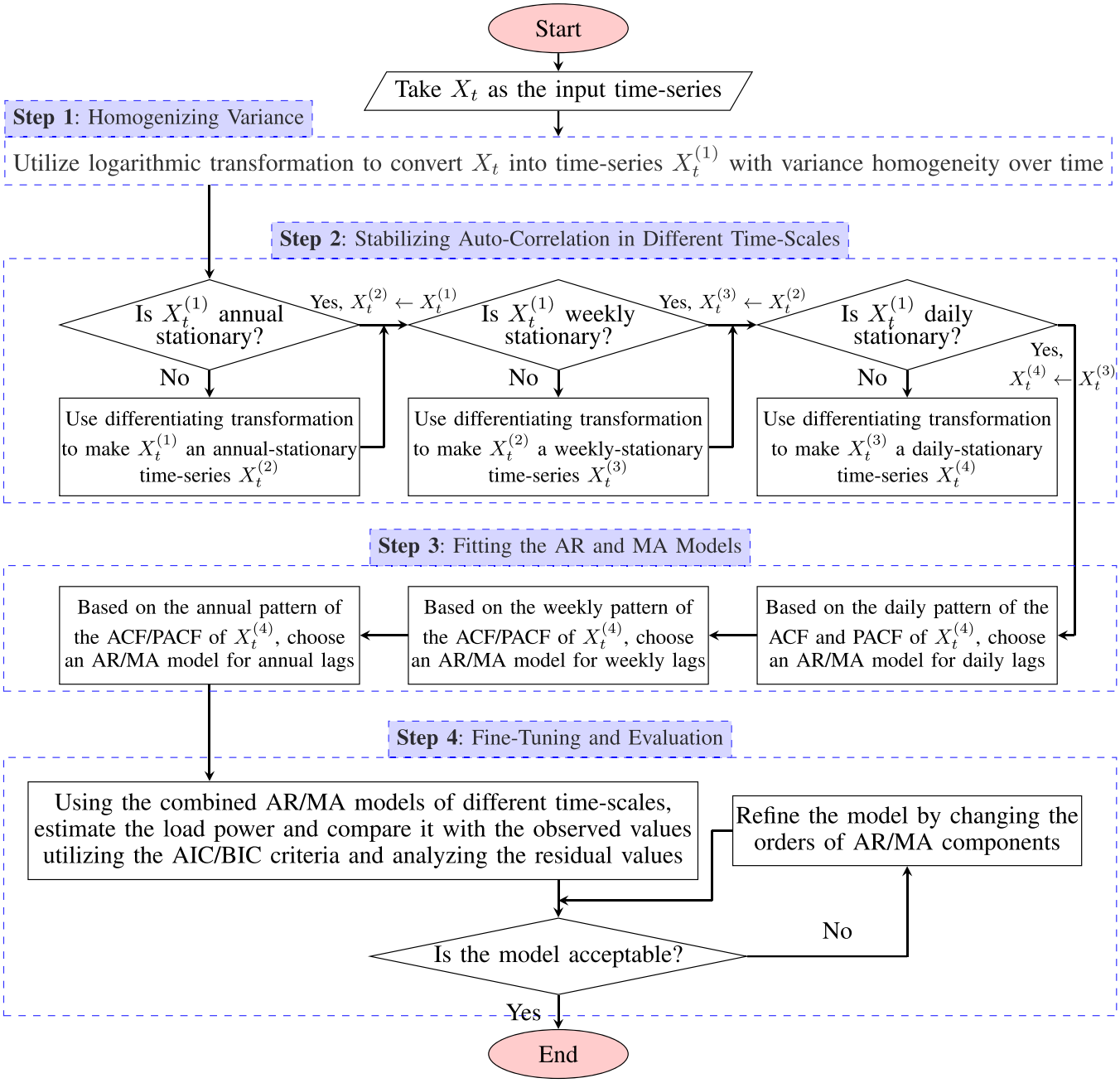


Fig. 2. The flowchart representation of the proposed methodology for creating a forecasting model for daily load values.

of the model is estimated by trying multiple values and validating the forecasts using the AIC. Fig. 2 shows the flowchart representation of the proposed methodology.

3.2. Homogenizing variance

When a time-series does not have variance homogeneity,[3] some transformations are applied on the time-series in order to homogenize (stabilize) its variance over time. Here, we use a special case of Box–Cox transformation [52] to obtain a homogeneous time-series. It is common that the standard deviation of a time-series (especially electric load data) proportionally changes with its mean over time, i.e. $Var[x_t] = (a\mu_t + b)^2$ for some a and b. Let τ_d denote the function that transforms x_t to a time-series with homogeneous variance: $Var[\tau_d(x_t)] =$ constant. Assuming that τ_d' represents the derivative function of τ_d, we approximate $\tau_d(x_t)$ by its first-order Taylor series around its mean μ_t:

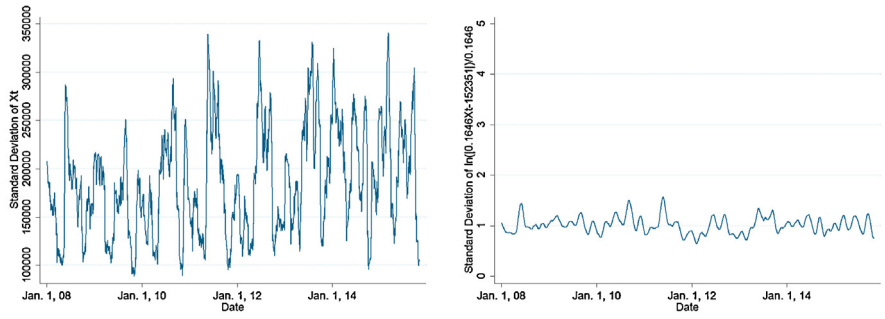
$$\tau_d(x_t) \approx \tau_d(\mu_t) + \tau_d'(\mu_t)(x_t - \mu_t).$$

Consequently, we obtain the following approximation for the variance of transformed time-series $\tau_d(x_t)$:

$$\begin{aligned} Var[\tau_d(x_t)] &\approx Var[\tau_d(\mu_t) + \tau_d'(\mu_t)(x_t - \mu_t)] \\ &= (\tau_d'(\mu_t))^2 Var[x_t] = (\tau_d'(\mu_t)(a\mu_t + b))^2. \end{aligned} \quad (2)$$

To obtain time-series $\tau_d(x_t)$ with variance-homogeneity, the following condition should hold: $\tau_d'(\mu_t)(a\mu_t + b) =$ constant. That is,

[3] A time-series has variance homogeneity if its variance does not change over time.



(a) The instantaneous standard deviation before applying logarithmic Box-Cox transformation.

(b) The stabilized standard deviation after applying logarithmic Box-Cox transformation.

Fig. 3. Homogenizing the variance of daily values of electric load in PJM power network over the period of 2008–2014.

τ_d is a logarithmic transformation: $\tau_d(x_t) = \ln(|ax_t + b|)/a$. Values of a and b are obtained by linear regression analysis of σ_t (standard deviation of x_t); where μ_t is the independent variable.

In order to illustrate the aforementioned variance-homogenizer method, consider the daily values of electric load in PJM power network over the period of 2008–2014. Let X_t denote the time-series representation of the daily load values. Fig. 3 shows the plot of standard-deviation σ_{X_t} of X_t ($\sqrt{\mathrm{Var}(X_t)}$) over time. It can be seen that the time-series is not stationary in variance. Using a linear regression, we obtain the relation of $\sigma_{X_t} = 0.1646\mu_{X_t} - 152351$ between instantaneous mean and standard-deviation of the time-series where the coefficient of determination for linear regression is $R^2 = 0.86$. As a result, the Box–Cox transformation for making the time-series stationary is in the following form: $\tau_d(X_t) = \ln(0.1646X_t - 152351)/0.1646$.

3.3. Stabilizing auto-correlation in different time scales

Any time-series that shows periodic similarities over time is referred to as seasonal. The electric power demand values follow multiple periodic patterns. For example, the load demand values on Sundays follow the same pattern; while they are different from Mondays'. Also, special holidays have specific annually-repetitive load patterns. The seasonal behaviors appear at the exact seasonal lags L, $2L$, $3L$, $4L$, and $5L$. For daily load data, the time-series shows repetitive pattern in the lags 7, 14, 21, etc. (weekly seasonality). Also, for hourly load time-series, data shows daily seasonality with $L = 24$, i.e. the seasonal lags are 24, 48, 72, and so on. In general, the transformed time-series values are considered stationary if the ACF satisfies the following conditions:

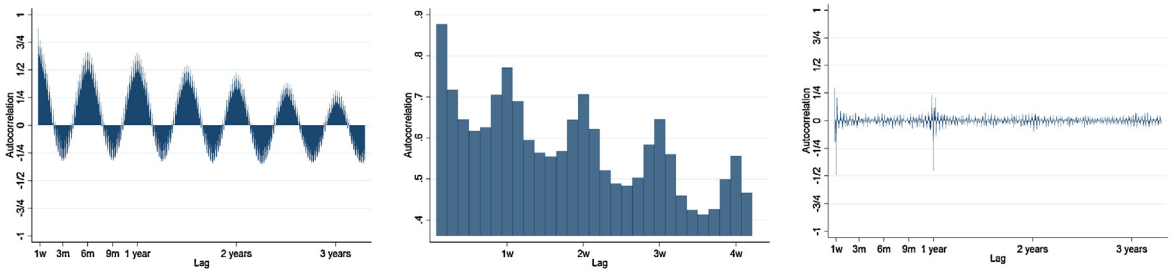
(1) Suddenly falling off or declining quickly at the non-seasonal level.

(2) Suddenly falling off or declining quickly at the seasonal cycles (exact seasonal lags or near seasonal lags).

Otherwise, these values are not considered as (wide-sense) stationary. For example, consider the daily load data of PJM network again. As mentioned before, $X_t^{(1)}$ is the logarithmic transformed load time-series with stabilized variance. Fig. 4(a) shows the ACF plot of $X_t^{(1)}$ for the first 1200 lags. It can be seen that the auto-correlation of data does not fall off or decline after the first few annual lags; instead, it decreases with a slow pace. Consequently, $X_t^{(1)}$ shows non-stationary behavior in annual seasonal cycle. Additionally, considering Fig. 4(b) which depicts the ACF plot of $X_t^{(1)}$ in the first few weekly lags, the auto-correlation $X_t^{(1)}$ in weekly and daily lags declines very slowly. As a result, $X_t^{(1)}$ is not a stationary time-series in both weekly lags (weekly seasonal cycle) and daily lags (non-seasonal level).

Henceforth, in order to make $X_t^{(1)}$ stationary, we need to transform it using multiple differentiating transformations in annual, weekly and daily levels. By transforming the time-series $X_t^{(1)}$ using the annual differentiation $D^{(d)}_{\text{annual}} = 1 - L^{364}$, the annual auto-correlation values fall off after the first annual lag (see the ACF plot of $X_t^{(2)} = D^{(d)}_{\text{annual}}(X_t^{(1)})$ depicted in Fig. 5(a)). Notice that we consider 364 as the number of days in a year as it is the closest multiple of 7 to the size of a year and it will make the later computations simpler.

Additionally, by transforming the time-series $X_t^{(2)}$ using the differentiation $D^{(d)}_{\text{weekly}} = 1 - L^7$, the weekly auto-correlation values fall off substantially after the first-week lag (see the ACF of $X_t^{(3)} = D^{(d)}_{\text{weekly}}(X_t^{(2)})$ in Fig. 5(b)). The daily differentiation $D = 1 - L$ transforms $X_t^{(3)}$ into a time-series with stationary behavior in all annual, weekly, and daily lags. As Fig. 5(c) depicts, the auto-correlation of

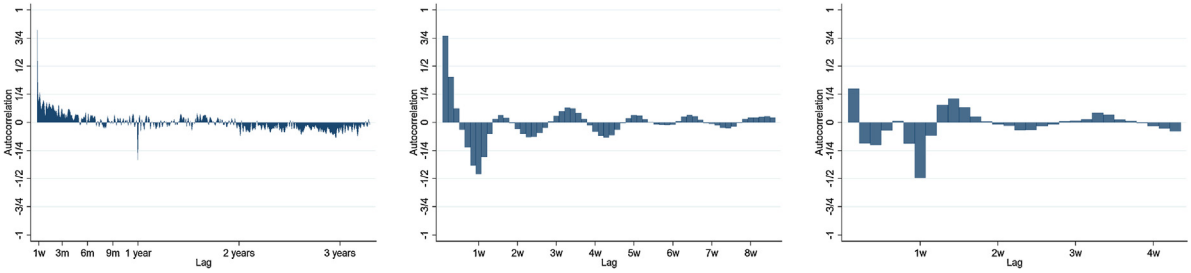


(a) ACF of $X_t^{(1)}$. The quarterly and annual seasonal patterns are illustrated.

(b) ACF of $X_t^{(1)}$ showing its non-stationary behavior in the first few weekly and daily lags.

(c) ACF of $X_t^{(4)}$ which is obtained by equation $X_t^{(4)} = D^{(d)}_{\text{annual}}(D^{(d)}_{\text{weekly}}(D^{(d)}_{\text{daily}}(X_t^{(1)})))$.

Fig. 4. ACF plots of daily load data in PJM network. Sub-figure (a) show the quarterly, and annual seasonality of variance stabilized load data. Sub-figure (b) shows the non-stationary behavior of data in both daily level and weekly cycle. Sub-figure (c) illustrates how multiple differentiating transformations have led us to obtain a stationary time-series in all daily, weekly, and annual levels.



(a) ACF of $X_t^{(2)}$ (b) ACF of $X_t^{(3)}$ (c) ACF of $X_t^{(4)}$

Fig. 5. Stabilizing the auto-correlation of daily load data of PJM network in daily, weekly, and annual levels utilizing differentiating transformations. Sub-figures (a), (b), and (c) show the ACF of $X_t^{(1)}$ transformed by annual, weekly and daily differentiation, respectively.

Table 1
Behavior of autocorrelation and partial-autocorrelation functions of the daily load data in non-seasonal and multiple seasonal levels.

Non-seasonal (daily)			Weekly seasonality			Annual seasonality		
Lag	ACF Sine-wave declining	PACF Falling off at *Lag* = 2	Lag	ACF Falling off at *Lag* = 7	PACF Exponentially declining	Lag	ACF Falling off at *Lag* = 364	PACF Exponentially declining
1	+0.300	+0.300	7	−0.494	−0.457	364	−0.450	+0.249
2	−0.187	−0.304	14	−0.017	−0.276	728	+0.030	+0.106
3	−0.199	−0.042	21	+0.015	−0.252	1092	+0.033	+0.031
4	−0.074	−0.048	28	−0.034	−0.190	–	–	–
5	+0.014	−0.010	35	+0.068	−0.118	–	–	–
6	−0.189	−0.275	42	−0.053	−0.151	–	–	–

$X_t^{(4)}$ exponentially decreases three or four days of every weekly lag. It shows that $X_t^{(4)}$ is a stationary time-series in non-seasonal level. Also, the plot in Fig. 5(c) verifies that the autocorrelation values of $X_t^{(4)}$ in weekly lags is negligible except in the seventh lag (proof of weekly stationary). Finally, the annual auto-correlation values of $X_t^{(4)}$ shown in Fig. 4(c) prove that $X_t^{(4)}$ shows stationary behavior in annual lags too (the auto-correlation in annual lags except the first one has very low values).

In summary, the series of multi-time-scale differentiating transformations $D^{(d)}_{\text{annual}} o D^{(d)}_{\text{weekly}} o D$ has converted the variance-stabilized time-series $X_t^{(1)}$ into a (wide-sense) stationary time-series $X_t^{(4)}$:

$$X_t^{(4)} = D^{(d)}_{\text{annual}} o D^{(d)}_{\text{weekly}} o D(X_t^{(1)}) = (1 - L^1 - L^7 + L^8 - L^{364} + L^{365} + L^{371} - L^{372}) X_t^{(1)}. \quad (3)$$

In the rest of the paper, we use δ_t to denote the stationary time-series $X_t^{(4)}$.

3.4. Fitting the AR and MA models

This step creates a model of the transformed stationary time-series δ_t (obtained in Eq. (3)) with a moving average and/or autoregressive components. The model is made based on the behavior of the ACF and PACF of δ_t. Table 1 shows the values of auto-correlation and partial auto-correlation of δ_t in multiple non-seasonal and seasonal levels.

The autocorrelation of the subsequent days (the non-seasonal part of the time-series) drops down suddenly after the second lag (see the values of the second column of Table 1). Additionally, the partial autocorrelations of δ_t in subsequent days are reduced exponentially (values of the third column of the same table supports the claim). The combinational behavior of the ACF and PACF functions at the non-seasonal level, which is depicted in Figs. 6c and f, implies that the transformed daily load values (δ_t) follows a moving-average of order two; i.e. the time-series consists of two moving-average terms:

$$\gamma_1^{(d)} \varepsilon_{t-1}^{(d)} + \gamma_2^{(d)} \varepsilon_{t-2}^{(d)},$$

where $\gamma_1^{(d)}$ and $\gamma_2^{(d)}$ denote the two coefficients of the auto-regressive terms that will be computed in the model fine-tuning step (next step). Moreover, $\epsilon_t^{(d)}$ shows the non-deterministic part of the whole signal (δ_t) and is traditionally called the "error value" of the ARMA model that will be built subsequently.

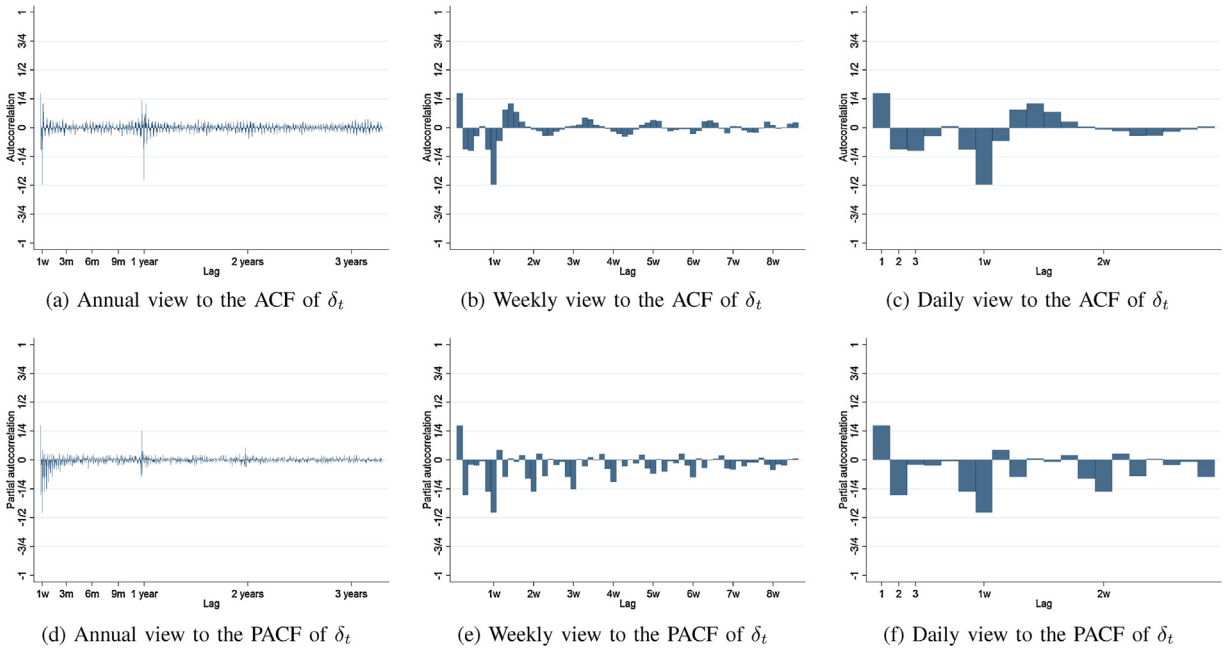
The values of the fifth and sixth columns of Table 1 illustrate that the autocorrelation values of the transformed daily load time-series in weekly lags ($L = 7, 14, 21, \ldots$) will substantially decrease after the first week; while its partial autocorrelation dies down exponentially with respect to the weekly lags. This behavior, which is depicted in Fig. 6b and e, implies that the time-series includes the moving-average of order one in the weekly seasonal cycle, i.e., the transformed daily load data contains the following moving-average signal which represents the non-deterministic identity of the time-series: $\gamma_1^{(w)} \varepsilon_{t-7}^{(d)}$; where $\gamma_1^{(w)}$ represents the coefficient of the moving-average term of the model that will be specified in the next step of the model creation process.

The results of Table 1 and Fig. 6 validate that the best real-time forecaster should have a moving-average element of lag two in the non-seasonal level and one moving-average term at the weekly cycle:

$$\delta_t = \underbrace{\psi_1 \delta_{t-1}^{(d)} + \psi_2^{(d)} \delta_{t-2}}_{\text{non-seasonal}} + \underbrace{\gamma_1^{(w)} \varepsilon_{t-7}^{(d)}}_{\text{weekly-seasonal}} + \underbrace{\gamma_1^{(y)} \varepsilon_{t-364}^{(d)}}_{\text{annual-seasonal}} + \varepsilon_t^{(d)}. \quad (4)$$

3.5. Fine-tuning and evaluation

In the final step of constructing the model, the detailed parameters of the AR/MA components are determined by trying multiple values and finding the forecasting model with minimum AIC/BIC values.



(a) Annual view to the ACF of δ_t (b) Weekly view to the ACF of δ_t (c) Daily view to the ACF of δ_t

(d) Annual view to the PACF of δ_t (e) Weekly view to the PACF of δ_t (f) Daily view to the PACF of δ_t

Fig. 6. ACF and PACF plots of stationary time-series δ_t obtained by transforming daily load values of PJM network.

3.5.1. AIC/BIC

The AIC is a measure of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Hence, AIC provides a means for model selection. AIC is founded on information theory: it offers a relative estimate of the information lost when a given model is used to represent the process that generates the data. To this end, it deals with the trade-off between the appropriate fitness of the model and the model complexity. AIC does not provide a test of a model in the sense of testing a null hypothesis, i.e. AIC can tell nothing about the quality of the model in an absolute sense. If all the candidate models fit poorly, AIC will not give any warning of that. Suppose that we have a statistical model of some data. Let $\mathcal{L}$ be the maximum value of the likelihood function for the model; let $\mathcal{K}$ be the number of estimated parameters in the model. Then, the AIC value of the model is the following:

$$\mathrm{AIC} = 2\mathcal{K} - 2\ln(\mathcal{L}).$$

Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value. Hence, AIC rewards goodness of fit (as assessed by the likelihood function), but it also includes a penalty that is an increasing function of the number of estimated parameters. The penalty discourages overfitting (increasing the number of parameters in the model almost always improves the goodness of the fit).

BIC is a criterion for model selection among a finite set of models; the model with the lowest BIC is preferred. It is based, in part, on the likelihood function and it is closely related to AIC. When fitting models, it is possible to increase the likelihood by adding parameters, but doing so may result in over-fitting. Both BIC and AIC resolve this problem by introducing a penalty term for the number of parameters in the model; the penalty term is larger in BIC than in AIC. The BIC is formally defined as

$$\mathrm{BIC} = -2 \cdot \ln \mathcal{L} + \mathcal{K} \cdot \ln(n),$$

where n is the size of historical data which the model is based on.

3.5.2. Fine-tuning the model

As mentioned before in Eq. (4), the behavior of ACF and PACF of time-series δ_t suggests that there should be two daily-AR terms, one weekly-MA term, and one annual-MA term in the model. In this step, we construct different forecasting models with multiple number of AR and MA terms and compare their corresponding AIC/BIC values to find out which model has the least AIC/BIC and therefore forecasts in the most accurate way. In other words, a sensitivity

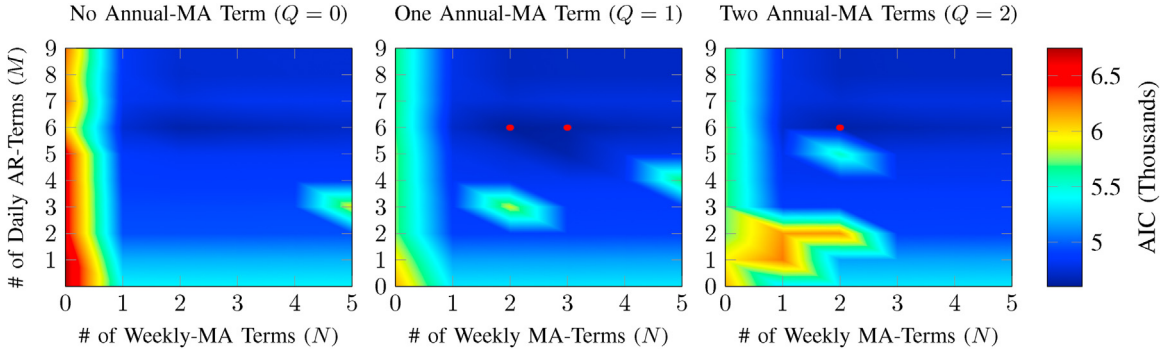


Fig. 7. AIC values for different settings of the proposed AR/MA model. Red markers show three smallest AIC values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

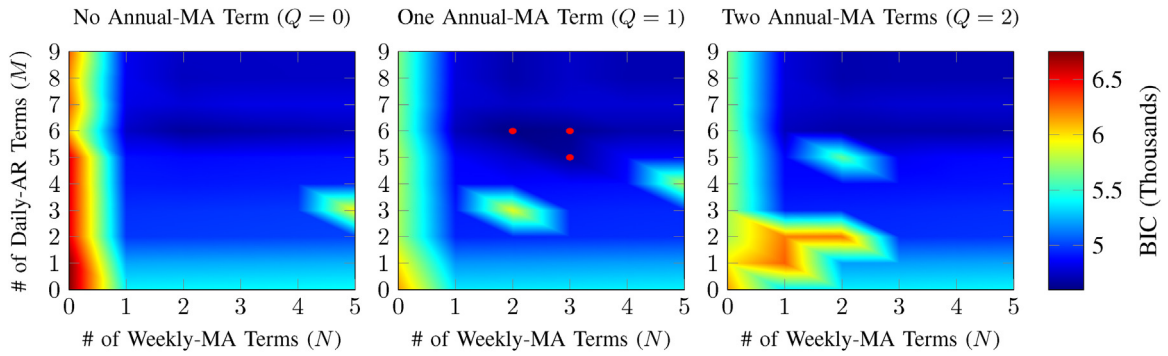


Fig. 8. BIC values for different settings of the proposed AR/MA model. Red markers show three smallest BIC values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

analysis is necessary in order to find the best combination of M (number of daily-AR terms), N (number of weekly-MA terms) and Q (number of annual-MA terms) where the corresponding model has the following form:

$$\delta_t = \underbrace{\sum_{i=0}^{M} \psi_i^{(d)} \delta_{t-i}}_{\text{daily-AR}} + \underbrace{\sum_{j=0}^{N} \gamma_j^{(w)} \varepsilon_{t-7j}^{(d)}}_{\text{weekly-MA}} + \underbrace{\sum_{k=0}^{Q} \gamma_Q^{(y)} \varepsilon_{t-364k}^{(d)}}_{\text{annual-MA}} + \varepsilon_t^{(d)}, \tag{5}$$

and $\psi_0^{(d)} = \gamma_0^{(w)} = \gamma_0^{(y)} = 0$. By changing any of these three parameters, we obtain a different forecasting model with different evaluation criteria (AIC, BIC, RMSE, MAE, and MAPE). Figs. 7–10 visualize the results of sensitivity analysis. In Figs. 7 and 8, the AIC and BIC values of $10 \times 6 \times 3 = 180$ different forecasters are depicted in terms of three thermal graphs since the triple of parameters (M, N, Q) belongs to the set $\{0, 1, \ldots, 9\} \times \{0, 1, \ldots, 5\} \times \{0, 1, 2\}$. In these two figures, the changing pattern of AIC and BIC is represented over multiple settings of the forecaster parameters. The darkest blue

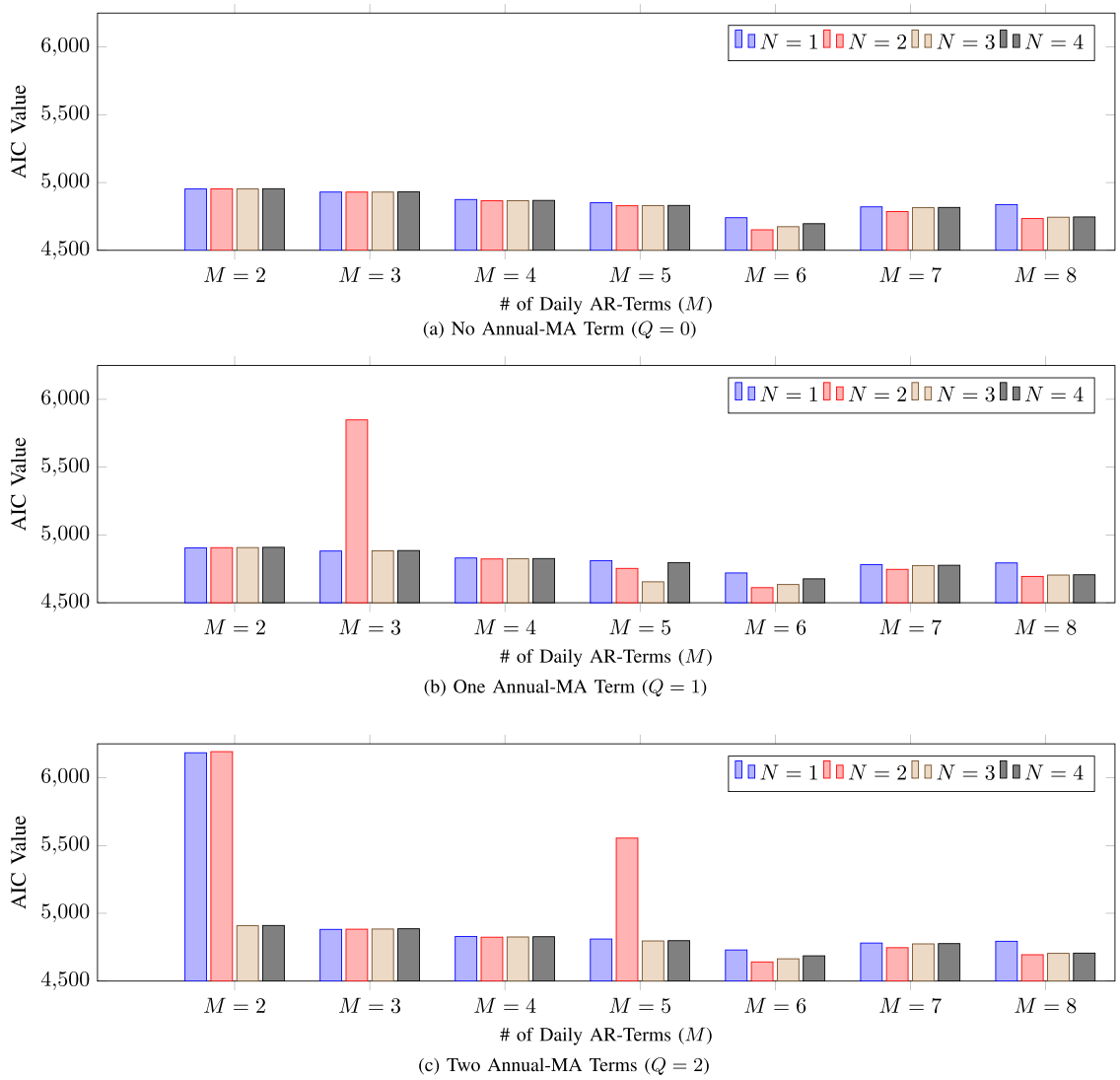


Fig. 9. Bar chart of the AIC values corresponding to different settings of the proposed AR/MA model.

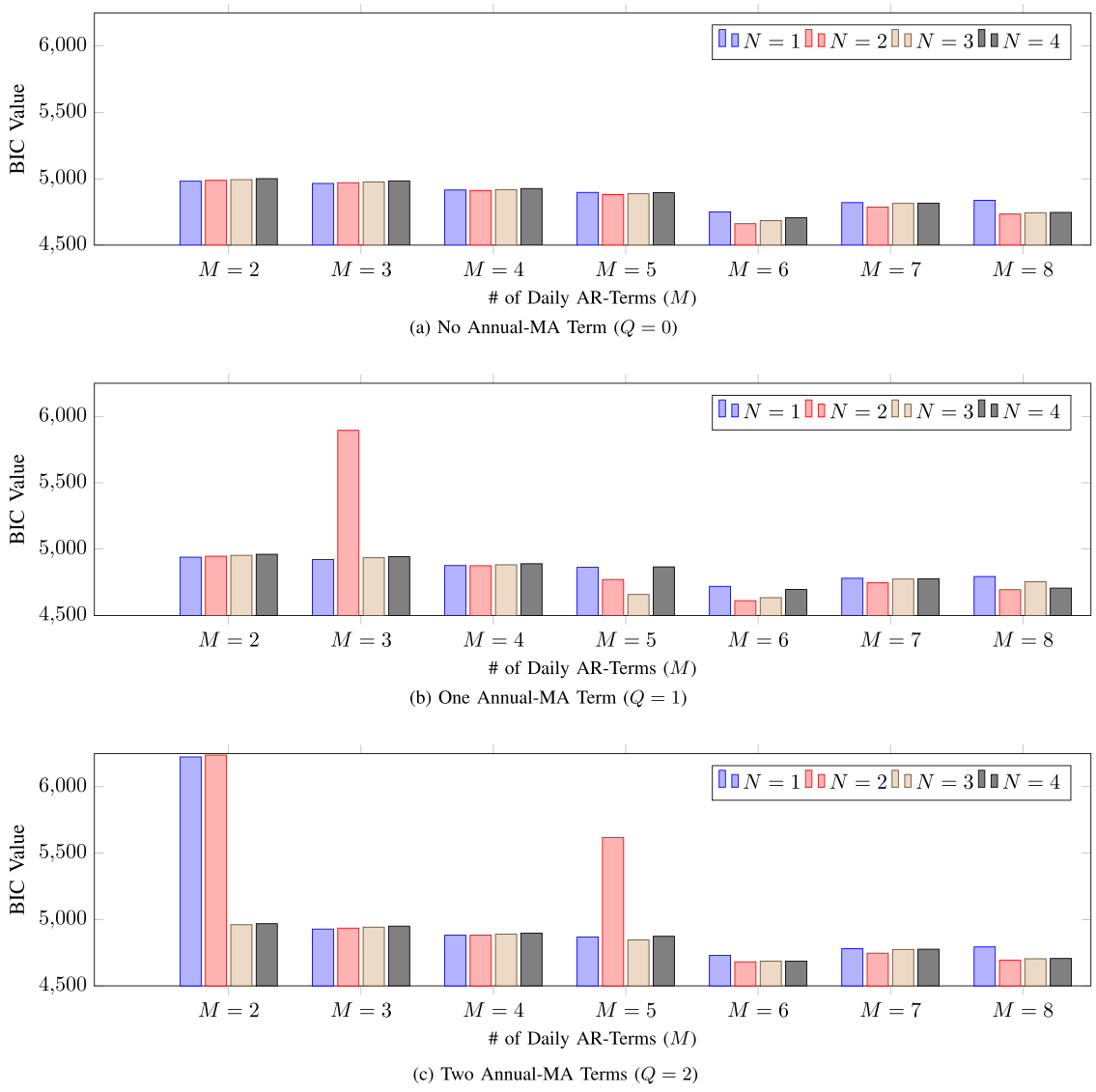


Fig. 10. Bar chart of the BIC values corresponding to different settings of the proposed AR/MA model.

regions in these figures, marked by red spots, specify the parameter settings of the forecaster which lead to the lowest AIC/BIC values. Fig. 9 shows the changes of AIC over three different values of *Q*. Each of sub-figures also represents a sensitivity analysis over *M* and *N* for a specific *Q*. Similarly, in Fig. 10, the changing pattern of BIC values is shown over three different values of *Q*. Each of sub-figures also represents a sensitivity analysis over *M* and *N* for a specific *Q*.

Fig. 11 compares the values of three different error types for multiple settings of the proposed model. Fig. 11a–c specify the values of Root-Mean-Square Error (RMSE), Mean-Absolute-Percent-Error (MAPE), and Mean-Absolute-Error (MAE)[4] respectively for different values of *M*, *N* and *Q*. The minimum error values of our proposed model happens in the minimum points of the plots depicted in Figs. 7, 8 and 11. Table 2 summarizes the best choices for triple (*M*, *N*, *Q*) which minimize the values of AIC/BIC and three different error types and AIC/BIC.

Table 2 suggests that the model with setting (*M*, *N*, *Q*) = (2, 6, 1) has the least value of AIC and BIC and second-least error value among all of the considered settings. Additionally, as mentioned before, AIC and BIC not only reward the model goodness of fit, but they also include penalty function which are increasing functions of the number of estimated parameters and discourage overfitting. Consequently, they are better criteria for evaluating the forecasting models than error values which merely quantify how fit the model is for the training set. As a result, we consider (*M*, *N*, *Q*) = (6, 2, 1) as the best setting of our proposed model for forecasting the daily electrical load. After creating the AR/MA-based model for δ_t and computing the forecast values of $\hat{\delta}_t$, the forecast load values $\hat{X}_t$ are obtained by applying the reverse transformations of steps one and two on $\hat{\delta}_t$, i.e. regarding Eq. (3), time-series $\hat{X}_t$ is computed by the following equation:

$$\hat{X}_t = \tau_d^{-1}(\hat{\delta}_t + [(L^1 + L^7 - L^8 + L^{364} - L^{365} - L^{371} + L^{372}) \circ \tau_d]X_t), \tag{6}$$

where τ_d is the logarithmic Box–Cox transformation introduced earlier and

$$\hat{\delta}_t = \sum_{i=1}^{6} \psi_i^{(d)} \delta_{t-i} + \sum_{j=1}^{2} \gamma_j^{(w)} \varepsilon_{t-7j}^{(d)} + \gamma_1^{(y)} \varepsilon_{t-364}^{(d)} + \varepsilon_t^{(d)}. \tag{7}$$

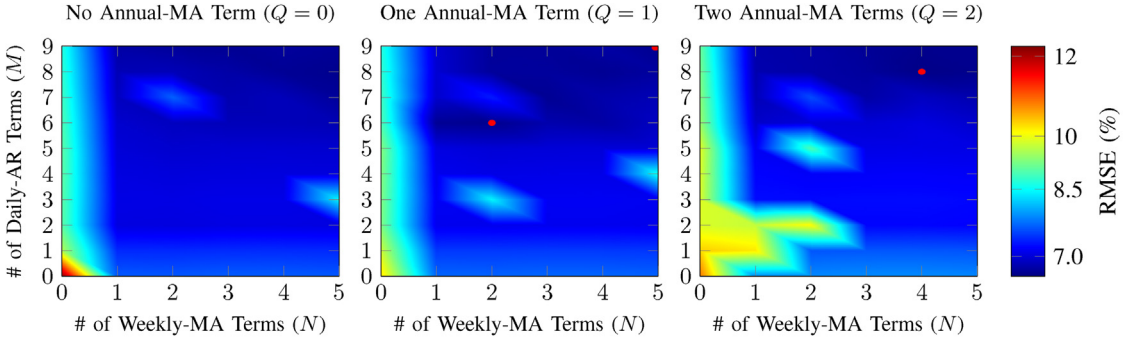
[4] Assuming o_i and f_i as the observed and forecast values ($i = 1, 2, \ldots, n$ and $o_i, f_i > 0$), the RMSE, MAPE, and MAE are respectively defined as $\frac{100}{n\bar{o}}\sum_{i=1}^{n}(f_i - o_i)^2$, $\frac{100}{n}\sum_{i=1}^{n}\frac{|f_i - o_i|}{o_i}$, and $\frac{100}{n\bar{o}}\sum_{i=1}^{n}|f_i - o_i|$.

Table 2
Choosing the best setting of the proposed model for forecasting the daily load values of PJM network in 2015.

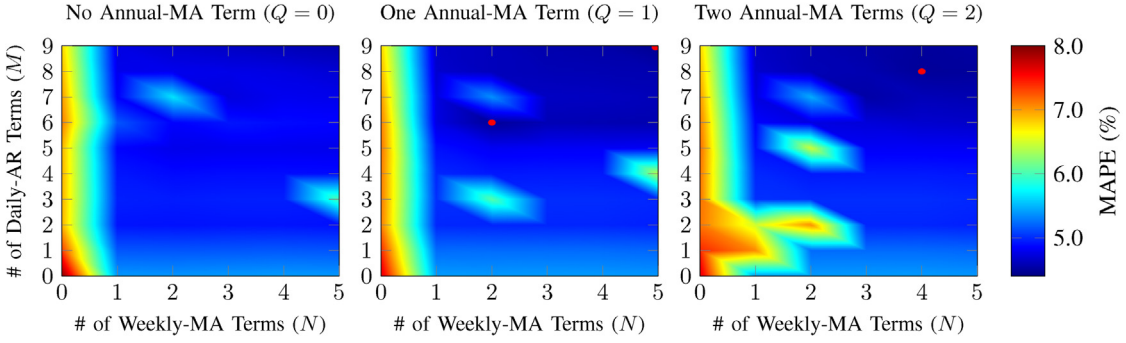
Selection criteria	# of Daily-AR terms (M), # of weekly-MA terms (N) and # of annual-MA terms (Q)	AIC (thousands)	BIC (thousands)	RMS error (%)	MAP error (%)	MA error (%)
Minimizing AIC	$M=2, N=6, Q=1$	4.61	4.76	6.60	4.49	4.68
	$M=3, N=6, Q=1$	4.63	4.79	6.80	4.59	4.73
	$M=2, N=6, Q=2$	4.64	4.80	7.02	4.75	5.42
Minimizing BIC	$M=2, N=6, Q=1$	4.61	4.76	6.60	4.49	4.68
	$M=3, N=6, Q=1$	4.63	4.79	6.80	4.59	4.73
	$M=3, N=5, Q=1$	4.65	4.81	7.09	4.82	4.98
Minimizing RMS/MAP/MA errors	$M=5, N=9$, and $Q=1$	4.71	4.85	6.60	4.46	4.67
	$M=2, N=6$, and $Q=1$	4.61	4.76	6.60	4.49	4.68
	$M=4, N=8$, and $Q=2$	4.71	4.87	6.62	4.49	4.70

3.5.3. *Comparing the residual time-series with white-noise*

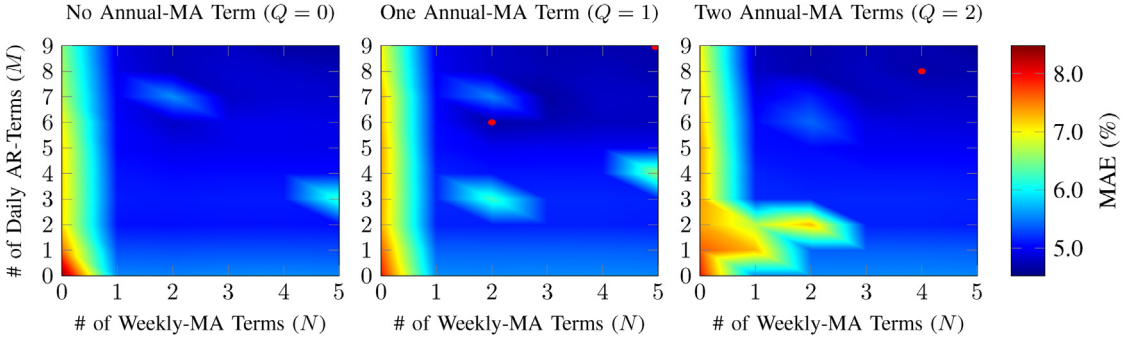
As mentioned before, one way of evaluating the goodness of a forecaster is to compare its residual time-series $\rho_t^{(d)} = \hat{X}_t - X_t$ (error values over time) with white noise. The ideal forecasting model extracts any meaningful information from the input training set to obtain the most accurate forecast. Henceforth, no meaningful signal will remain in the residual time-series, i.e. the residual time-series will be a *white noise*. Here, we use the cumulative periodogram of $\rho_t^{(d)}$ to show how close the residual time-series is to a *white noise*.



(a) RMS Errors in forecasting the daily load of year 2015 when the proposed AR/MA model is applied with different settings.



(b) MAP Errors in forecasting the daily load of year 2015 when the proposed AR/MA model is applied with different settings.



(c) MAE values in forecasting the daily load of year 2015 when the proposed AR/MA model is applied with different settings.

Fig. 11. Comparing the values of three different error types for different settings of the proposed AR/MA model. The points with three smallest error values have been specified with red markers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

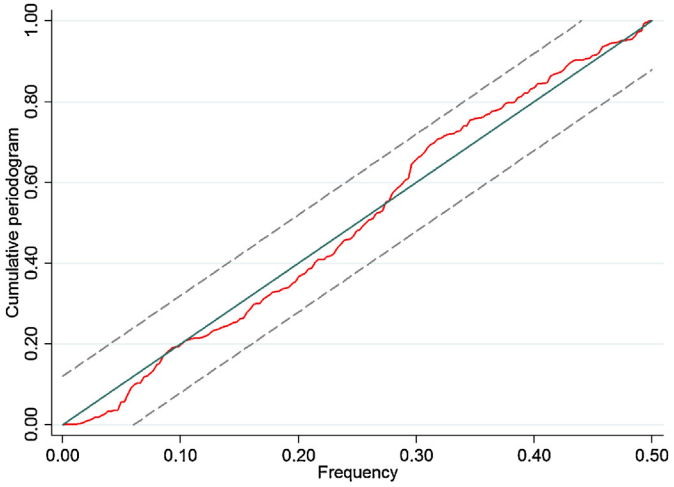


Fig. 12. The cumulative periodogram of residual time-series $\rho_t^{(d)}$. The dashed lines specify the 99%-confidence band for Bartlett's test to prove that the time-series is a white noise of constant mean and variance.

Let $\hat{f}(\omega)$ denote the sample spectral density of the residual time-series $\rho_t^{(d)}$:

$$\hat{f}(\omega) = \frac{1}{n}\left|\sum_{t=1}^{n}\rho_t^{(d)} e^{2\pi i(t-1)\omega}\right|^2,$$

where $\omega \in [0, 0.5]$. The cumulative periodogram of $\rho_t^{(d)}$ is defined as the plot of $\sum_{j=1}^{k}\hat{f}(\omega_j)/\sum_{j=1}^{q}\hat{f}(\omega_j)$ versus $\omega_k = (k-1)/n$ for $k = 1, 2, \ldots, q = \lfloor n/2 \rfloor + 1$. Fig. 12 depicts the cumulative periodogram of $\rho_t^{(d)}$ with red color. This plot is used in Bartlett's test which is a test of the null hypothesis that the residual time-series comes from a white-noise process of uncorrelated random variables having a constant mean and variance. In this figure, two dashed lines specify the 99%-confidence band for Bartlett's test; i.e. since the periodogram resides in the band, the residual time-series is a white-noise with confidence level of at-least 99%.

4. Case study

In this section, our approach is implemented on the PJM hourly-metered load data available in [53] and the superior performance of our proposed methodology is illustrated compared with the currently existing models. Throughout this section, we build an hour-ahead forecaster for the PJM load data. Similar to the model built in previous section, the forecaster is trained using the historical hourly load data of PJM network in years 2013–2014; while it will be evaluated and fine-tuned using the same network load data in 2015.

4.1. Homogenizing the variance

As mentioned in the previous section, the first step for modeling the load data is to transform the original load data Y_t to time-series $Y_t^{(1)} = \tau_h(Y_t)$ with variance homogeneity. To this end, we use the same technique utilized in previous section to create the following logarithmic Box–Cox transformation: $\tau_h(\cdot) = 0.1232\ln(\cdot) - 3880.71$.

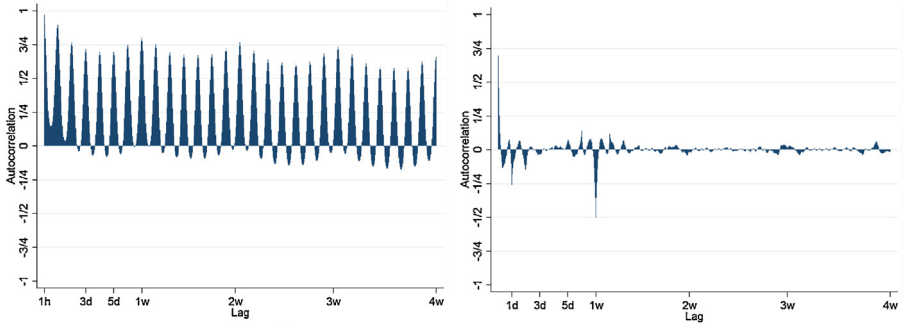
4.2. Stabilizing auto-correlation in different time-scales

In the next step, the time series $Y_t^{(1)}$ with variance-homogeneity is transformed to a WSS time-series η_t by three chained differentiating transformations $\eta_t = D^{(h)}_{\text{weekly}} \circ D^{(h)}_{\text{daily}} \circ D(Y_t^{(1)})$. Fig. 13, respectively depicts the ACF plots of $Y_t^{(1)}$ and η_t for the first four weeks lags. As Fig. 13a illustrates, $Y_t^{(1)}$ is a daily and weekly seasonal time-series and does not show WSS behavior as its ACF values decrease extremely slowly in daily and weekly cycles; however, the ACF plot of η_t in Fig. 13b proves that η_t is a weekly-stationary, daily-stationary, and hourly-stationary time-series as its autocorrelation suddenly falls off after the first few hourly, daily, and weekly lags.

4.3. Fitting the AR and MA models

In this step, the WSS time-series η_t obtained in previous step is modeled to an AR/MA model based on the behavior of its ACF and PACF. Table 3 shows the values of ACF and PACF of η_t in the first few hourly, daily, and weekly lags. According to this table, in the first few hourly lags, the ACF substantially decreases in a Sine-wave manner with the respect to the lag parameter (see the values of the second column of Table 3). Additionally, the PACF falls off suddenly at lag $L = 1h$ (values of the third column of the same table supports the claim). The combinational behavior of the ACF and PACF functions at the non-seasonal level implies that the WSS time-series η_t follows an auto-regressive model of order one, i.e. the time-series consists of the auto-regressive term $\phi_1^{(h)}\eta_{t-1}$, where variable $\phi_1^{(h)}$ denotes the coefficient of the auto-regressive term.

The values of the fifth and sixth columns of Table 3 illustrate that the autocorrelation of η_t in daily lags ($L = 24, 48, 72, \ldots$) will substantially decrease after the first two days while its partial autocorrelation dies down exponentially with respect to the daily lags. This behavior induces that the time-series η_t should include the moving-average of the second order in the daily seasonal cycles; i.e. η_t includes the summation of two moving-average signals: $\theta_1^{(d)}\varepsilon_{t-24}^{(h)} + \theta_2^{(d)}\varepsilon_{t-48}^{(h)}$, where $\theta_1^{(d)}$ and $\theta_2^{(d)}$ represent the coefficients of the moving average terms of the model. Moreover, $\varepsilon_t^{(h)}$ shows the non-deterministic part of η_t.



(a) ACF plot of $Y_t^{(1)}$. (b) ACF plot of η_t.

Fig. 13. ACF plots of $Y_t^{(1)}$ and its transformed time-series η_t.

Table 3
Behavior of autocorrelation and partial-autocorrelation functions of the hourly load data in non-seasonal and multiple seasonal levels.

Non-seasonal (hourly)			Daily seasonality			Weekly seasonality		
Lag	ACF Sine-wave declining	PACF Falling off at *Lag* = 1	Lag	ACF Falling off at *Lag* = 48	PACF Exponentially declining	Lag	ACF Falling off at *Lag* = 168	PACF Exponentially declining
1	+0.697	+0.697	24	−0.281	−0.229	168	−0.490	−0.156
2	+0.428	−0.111	48	−0.135	−0.130	336	−0.009	−0.089
3	+0.218	−0.071	72	−0.033	−0.096	504	+0.013	−0.061
4	+0.084	−0.020	96	+0.008	−0.064	672	–	–
5	−0.001	−0.030	120	+0.050	−0.019	840	–	–
6	−0.060	−0.045	144	+0.149	+0.090	1008	–	–

The values of the last two columns of Table 3 determine that the ACF and PACF of η_t in weekly lags ($L = 168, 336, \ldots$) show similar behavior as in daily lags. In fact, the values of ACF will suddenly falls off after the first weekly lag; while PACF declines exponentially with respect to the weekly lags. This combination of behaviors lead us to the conclusion that the time-series η_t should include the moving-average of the first order in the weekly seasonal cycle, i.e. η_t has a non-deterministic moving-average part in the form of $\theta_1^{(w)}\varepsilon_{t-168}^{(h)}$ where $\theta_1^{(w)}$ represents the coefficient of the first moving-average term in weekly lags.

Table 3 suggests the following AR/MA-based model for WSS time-series η_t

$$\eta_t = \underbrace{\phi_1^{(h)}\eta_{t-1}}_{\text{non-seasonal}} + \underbrace{\theta_1^{(d)}\varepsilon_{t-24}^{(h)} + \theta_2^{(d)}\varepsilon_{t-48}^{(h)}}_{\text{daily-seasonal}} + \underbrace{\theta_1^{(w)}\varepsilon_{t-168}^{(h)}}_{\text{weekly-seasonal}} + \varepsilon_t^{(h)}. \quad (8)$$

4.4. Fine-tuning and evaluation

As mentioned before, in the final step of constructing the model, the detailed parameters of the AR/MA components are determined by trying multiple values and finding the forecasting model with minimum AIC/BIC values. In the case of hourly forecaster, we figured out that the following model settings will give the minimum AIC/BIC when forecasting the hourly electrical load of PJM network in 2015:

$$\eta_t = \sum_{i=1}^{12}\phi_i^{(h)}\eta_{t-i} + \sum_{j=1}^{3}\theta_j^{(d)}\varepsilon_{t-24j}^{(h)} + \theta_1^{(w)}\varepsilon_{t-168}^{(h)} + \varepsilon_t^{(h)} \quad (9)$$

Table 4 summarizes the performance of our model comparing the other forecasters of current literature.

5. Summary and outlook

This work proposes a novel step-by-step technique for creating an AR/MA-based forecasting model. This model is used for electric power demand forecasting from short-term to medium-term horizon. In the first step, the time-series corresponding to the historical load data is evaluated to make sure that it has homogeneous variance over time. In the case that the time-series is not variance homogeneous, we apply an appropriate logarithmic Box–Cox transformation in order to convert it into a time-series with variance homogeneity. In the next step, the ACF values of the transformed time-series are examined to see if there exists any indication of being non-stationary at the non-seasonal level and seasonal cycles (daily, weekly, and annually). If the ACF values either fall off or decline in the first few lags, it should be considered stationary. On the other hand, if the ACF values either fall off after a considerable number of lags or declines quite slowly, it should be considered as non-stationary. To eliminate the non-stationary indications, we iteratively apply differentiating transformations on the time-series. In the next step, the stationary time-series, resulted from the second step, is modeled to a moving average and/or autoregressive model based on the behavior of its corresponding ACF and PACF in different time resolutions (hourly, daily, weekly, and annually). For any given time resolution, if the ACF values of the time-series cut off in the first few lags, while the PACF values are reduced exponentially or in a Sine-wave manner, the time-series in that specific time resolution is best fit to a moving-average model. On the other hand, if the PACF values of the time-series substantially fall off in the first few lags, while the ACF values are reduced exponentially or in a Sine-wave way, the best model for that specific time resolution is autoregressive.

In order to fine-tune the model, we utilize AIC/BIC which deals with the trade-off between appropriate fitness of the model and the model complexity. To this end, we consider a penalty value for

Table 4
Comparison the accuracy of forecasters presented in Pappas et al. [47], Shaker et al. [54], and Dudek [44]

Work	Year	Location	Method	WN Test[i]	MAPE (%)	RMSE (%)	MAE (%)	Type	AIC (1000's)	BIC (1000's)
Pappas et al.	2010	Greece	MMPF[a]	Tested	1.87	Not tested	Not tested	Short-term	Not tested	Not tested
Pappas et al.	2010	Greece	AICC[b]	Not tested	1.98	Not tested	Not tested	Short-term	Not tested	Not tested
Shaker et al.	2014	Alberta Canada	WNN[c]	Not tested	Not tested	1.328	0.983	Short-term	Not tested	Not tested
Shaker et al.	2014	Alberta Canada	MLPNN[d]	Not tested	Not tested	3.724	2.556	Short-term	Not tested	Not tested
Shaker et al.	2014	Alberta Canada	RBFNN[e]	Not tested	Not tested	2.287	1.712	Short-term	Not tested	Not tested
Dudek	2016	Polish Network	PCR[f]	Not tested	1.15	Not tested	Not tested	Short-term	Not tested	Not tested
Dudek	2016	Polish Network	PLSR[g]	Not tested	1.09	Not tested	Not tested	Short-term	Not tested	Not tested
This work	2016	PJM Network	MTS ARMA[h]	Tested	0.86	1.24	0.92	Short-term	Tested	Tested

[a] Multi-model partitioning filter.
[b] Corrected Aikaike information criterion.
[c] Wavelet neural network.
[d] Multi-layer perceptron neural network.
[e] Radial basis function neural network.
[f] Principle component regression.
[g] Partial least-square regression.
[h] Multi time-scale ARMA.
[i] WN test: white noise test.

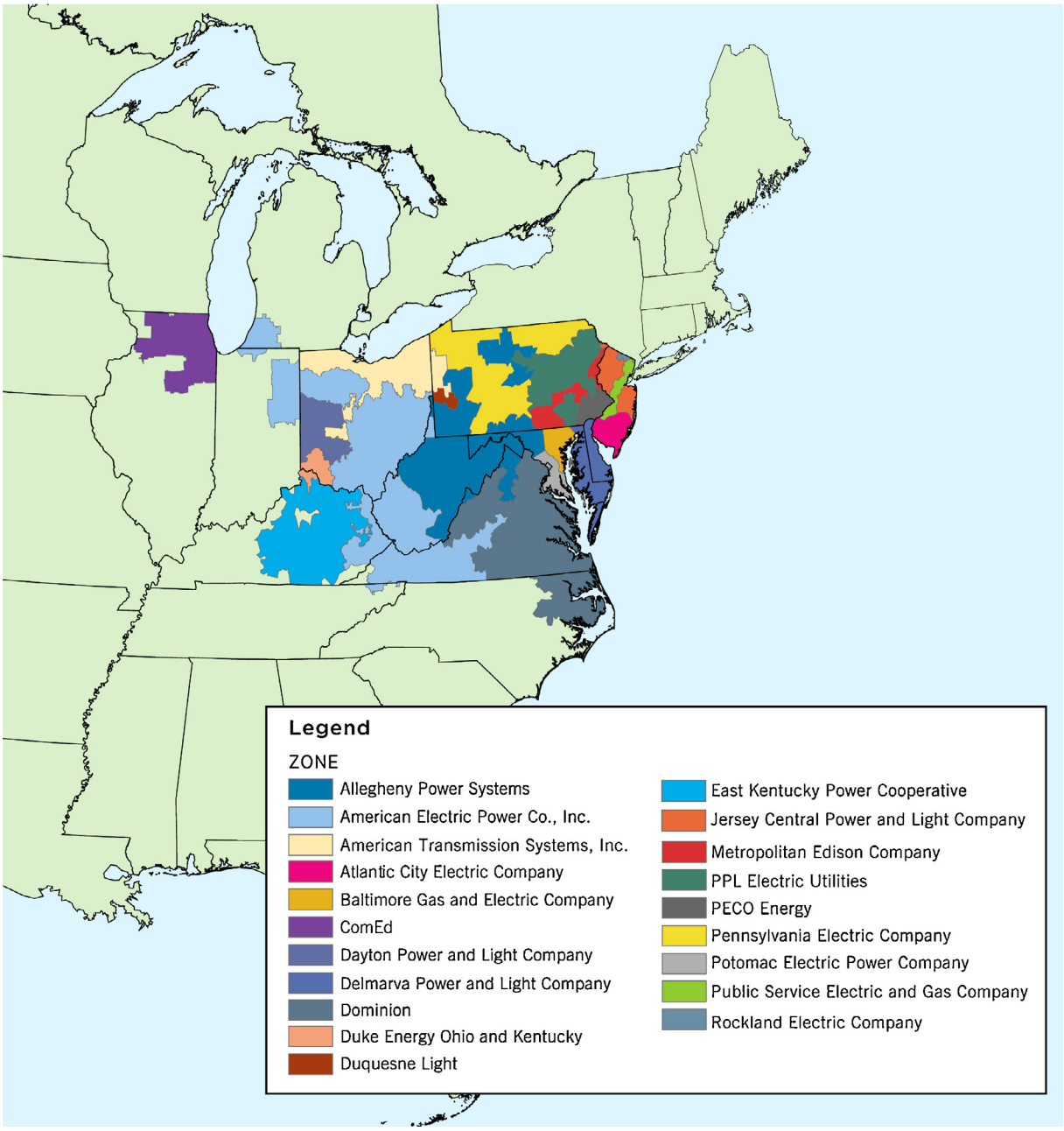


Fig. 14. Geographical regions covered by PJM interconnection [56].

complex models which may only work precisely for the analyzed training set; however, the conventional error measurements are considerably dependent on the given training set and measure the fitness of the model merely with respect to the training set. If the forecaster's flexibility is our most important concern, i.e. our main objective is to develop a forecaster which can be broadly used for different data training sets without deteriorating the forecaster's performance, the BIC is a better alternative (compared with AIC) for quantifying the forecaster's utility since BIC penalizes the complexity of forecasting models more than AIC. However, AIC is a better criterion if we choose to compromise model flexibility by some degree in order to gain more fitness and improve the accuracy of forecaster by reducing the forecasting error.

In order to validate the effectiveness of the proposed forecaster, we implement it for forecasting the electric power demand of a real-world example from PJM interconnection. The accuracy of the proposed forecaster is evaluated by Bartlett's periodogram-based test and quantified by multiple conventional error types. The our performance of the proposed forecaster, compared with the previous studies, is validates using three error metrics (MAPE, RMSE, and MAE). For instance, the MAPE value of our forecaster for PJM hourly-metered load data is 0.86%, which shows 21% reduction compared with Dudek's forecaster which is developed in 2016 [44].

Appendix A. Detailed information about the hourly-metered load data of PJM

PJM interconnection is a regional transmission organization (RTO) that manages the electricity in the following regions in the United States: Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia and the District of Columbia [55]. In

Table 5
The daily load in the first six months of 2015 in PJM. The numbers are shown in kWh.

Day/month	January	February	March	April	May	June
1	2,224,247.0	2,287,549.2	2,442,177.8	2,024,486.3	1,847,056.1	2,101,982.5
2	2,214,133.4	2,467,346.6	2,437,075.9	1,974,545.2	1,681,862.1	1,943,814.3
3	2,146,834.0	2,619,797.2	2,529,970.7	1,807,215.0	1,657,509.5	1,922,430.1
4	1,974,231.4	2,445,748.3	2,351,650.9	1,730,857.6	1,944,146.7	2,008,393.8
5	2,445,435.6	2,571,644.5	2,564,235.5	1,692,526.7	2,016,051.7	2,033,350.8
6	2,652,063.8	2,681,963.0	2,705,979.7	1,898,950.8	2,030,419.7	1,871,852.8
7	2,843,836.9	2,243,782.4	2,369,589.7	1,922,330.2	2,067,433.1	1,868,774.4
8	3,023,713.9	2,025,259.8	1,954,528.4	1,970,665.4	2,131,769.8	2,238,112.9
9	2,736,306.6	2,347,400.7	2,139,678.0	2,028,948.3	1,927,307.7	2,270,840.2
10	2,679,408.0	2,504,609.6	2,165,532.0	1,943,293.5	1,931,900.8	2,382,590.1
11	2,463,715.8	2,459,178.1	2,055,011.6	1,718,914.8	2,239,332.3	2,542,460.0
12	2,461,286.0	2,547,267.4	2,090,851.4	1,688,730.1	2,169,028.7	2,578,623.9
13	2,561,561.2	2,755,521.6	2,127,633.3	1,886,606.5	1,922,147.6	2,352,410.3
14	2,690,702.6	2,543,273.5	1,932,114.2	1,879,084.5	1,890,764.5	2,353,527.6
15	2,547,062.4	2,809,745.1	1,884,176.9	1,882,273.2	1,928,452.4	2,621,394.3
16	2,453,213.0	3,010,474.6	2,060,271.8	1,884,624.7	1,904,483.3	2,590,952.4
17	2,304,689.7	2,802,988.6	2,017,453.2	1,876,043.3	2,001,731.2	2,332,716.6
18	2,159,352.6	2,803,212.0	2,184,042.3	1,705,664.2	2,281,909.0	2,393,511.2
19	2,297,892.3	3,052,242.2	2,190,832.1	1,629,643.8	2,160,130.8	2,338,444.6
20	2,345,274.7	3,104,390.6	2,226,302.0	1,901,434.6	1,952,688.8	2,161,401.8
21	2,412,450.4	2,720,441.3	1,958,744.1	1,879,815.0	1,881,833.7	2,312,095.6
22	2,416,012.7	2,306,819.3	1,896,000.7	1,926,903.3	1,847,969.6	2,638,997.9
23	2,399,030.1	2,670,971.2	2,203,770.5	1,974,713.6	1,665,668.7	2,694,198.8
24	2,247,946.0	2,868,672.0	2,220,515.3	1,947,838.4	1,701,091.1	2,399,544.4
25	2,152,295.5	2,609,601.0	2,179,801.4	1,804,096.3	1,910,615.1	2,307,415.4
26	2,492,879.1	2,609,484.5	2,066,330.6	1,707,445.2	2,285,864.7	2,234,547.0
27	2,571,906.4	2,640,380.4	2,137,934.4	1,919,708.2	2,350,697.4	1,929,125.4
28	2,607,071.1	2,539,066.3	2,161,144.5	1,902,106.7	2,359,577.4	1,851,272.4
29	2,562,531.5	–	2,101,593.8	1,882,956.0	2,336,808.4	2,087,940.5
30	2,481,083.7	–	2,093,609.3	1,867,538.4	2,178,553.2	2,237,095.2
31	2,484,570.9	–	2,048,286.0	–	2,056,848.5	–

some of the mentioned regions, PJM is coordinating the wholesale electricity for the whole region. However, for other areas, it only serves a part of the region. Its main task is to provide reliable electricity for more than 61 million customers. Fig. 14 represents the regions covered by PJM interconnection. Tables 5 and 6 represent the historical daily load for the first and second six months of 2015 in the regions covered by PJM interconnection, respectively.

Table 6
The daily load in the last six months of 2015 in PJM. The numbers are shown in kWh.

Day/month	July	August	September	October	November	December
1	2,235,981.1	2,326,506.7	2,675,402.2	1,871,454.0	1,625,687.1	2,089,866.7
2	2,098,594.8	2,326,394.3	2,686,777.1	1,848,997.8	1,900,526.7	2,076,625.0
3	1,947,908.6	2,560,275.1	2,719,161.1	1,756,196.2	1,932,687.6	2,161,429.0
4	1,897,211.1	2,574,999.1	2,599,482.7	1,674,459.5	1,925,894.8	2,175,708.5
5	2,000,731.8	2,498,275.3	2,280,105.6	1,869,228.5	1,949,722.0	2,055,484.6
6	2,428,637.2	2,325,697.5	2,215,356.6	1,897,800.1	1,931,240.4	2,013,747.7
7	2,523,706.6	2,284,944.6	2,375,704.9	1,914,453.2	1,739,918.6	2,173,588.0
8	2,399,394.5	2,172,007.5	2,679,940.3	1,930,268.7	1,746,607.9	2,164,163.3
9	2,428,976.7	2,169,595.7	2,630,928.8	1,918,961.5	2,026,658.3	2,168,256.1
10	2,366,755.8	2,391,373.9	2,366,200.6	1,667,591.6	1,969,193.8	2,065,124.2
11	2,169,871.9	2,437,264.5	2,193,611.3	1,641,818.7	1,945,546.9	1,995,847.2
12	2,157,110.2	2,334,476.4	1,888,958.6	1,874,666.6	1,970,728.4	1,797,150.0
13	2,442,000.9	2,317,973.4	1,750,362.8	1,892,341.3	1,951,856.9	1,746,418.1
14	2,495,718.8	2,400,619.0	1,954,813.6	1,866,502.2	1,858,388.4	1,963,965.5
15	2,338,597.1	2,336,837.2	2,081,851.1	1,873,318.1	1,800,851.9	2,004,113.1
16	2,221,536.9	2,409,186.7	2,178,468.4	1,844,691.9	1,964,457.0	2,048,928.6
17	2,404,239.5	2,698,831.6	2,240,706.4	1,750,889.6	1,961,147.6	2,087,605.0
18	2,478,797.6	2,636,636.9	2,230,933.6	1,791,688.1	1,941,140.4	2,187,556.3
19	2,589,602.2	2,634,215.7	2,039,646.9	2,015,386.1	1,928,128.6	2,176,180.8
20	2,824,010.1	2,497,295.3	1,847,147.5	1,952,358.9	1,975,289.0	2,107,844.3
21	2,675,592.0	2,347,128.5	1,964,309.2	1,917,727.2	1,927,928.7	2,155,731.4
22	2,412,955.5	2,094,643.9	2,006,255.5	1,903,754.3	1,952,532.7	2,011,890.0
23	2,379,320.3	2,079,893.8	2,036,401.2	1,849,517.5	2,237,398.5	1,958,759.2
24	2,408,843.6	2,363,988.9	2,043,572.7	1,718,982.9	2,181,561.0	1,767,741.6
25	2,352,382.1	2,274,723.0	1,980,467.2	1,657,375.6	2,097,915.4	1,669,885.1
26	2,361,044.4	2,128,817.6	1,767,919.0	1,895,559.4	1,745,687.3	1,726,500.0
27	2,638,585.8	2,121,712.1	1,752,623.6	1,963,221.4	1,745,832.1	1,743,846.6
28	2,770,611.5	2,141,847.0	2,119,168.7	1,953,809.3	1,758,048.5	2,052,214.7
29	2,813,593.6	2,086,999.4	2,172,743.8	1,913,205.3	1,820,186.8	2,028,147.0
30	2,725,597.2	2,190,522.9	2,073,200.0	1,902,573.2	2,115,151.5	2,011,894.9
31	2,562,081.4	2,561,310.6	–	1,781,476.8	–	1,968,810.2

References

[1] E. Gonzalez-Romera, M.A. Jaramillo-Moran, D. Carmona-Fernandez, Monthly electric energy demand forecasting based on trend extraction, IEEE Trans. Power Syst. 21 (4) (2006) 1946–1953.
[2] J.W. Taylor, P.E. McSharry, Short-term load forecasting methods: an evaluation based on European data, IEEE Trans. Power Syst. 22 (4) (2007) 2213–2219.
[3] M. Shahidehpour, H. Yamin, Z. Li, Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management, Wiley Online Library, 2002.
[4] K. Moslehi, R. Kumar, A reliability perspective of the smart grid, IEEE Trans. Smart Grid 1 (1) (2010) 57–64.
[5] M.H. Amini, O. Karabasoglu, M.D. Ilić, K.G. Boroojeni, S.S. Iyengar, ARIMA-based demand forecasting method considering probabilistic model of electric vehicles' parking lots, in: IEEE PES General Meeting 2015, Denver, CO, USA, July 26–30, 2015.
[6] F. Rahimi, A. Ipakchi, Demand response as a market resource under the smart grid paradigm, IEEE Trans. Smart Grid 1 (1) (2010) 82–88.
[7] S. Bahrami, V.W.S. Wong, An autonomous demand response program in smart grid with foresighted users, in: Proc. of IEEE Smart Grid Communications (SmartGridComm), Miami, FL, 2015, pp. 205–210.
[8] F. Kamyab, M.H. Amini, S. Sheykhha, M. Hasanpour, M.M. Jalali, Demand response program in smart grid using supply function bidding mechanism, IEEE Trans. Smart Grid 7 (2) (2016) 1277–1284.
[9] N. Mahmoudi, New Demand Response Framework and Its Applications for Electricity Markets (PhD Dissertation), The University of Queensland, 2015.
[10] X. Zhang, G. Hug, J. Zico Kolter, I. Harjunkoski, Model predictive control of industrial loads and energy storage for demand response, in: IEEE PES General Meeting, 2016.
[11] S. Bahrami, A. Sheikhi, From demand response in smart grid toward integrated demand response in smart energy hub, IEEE Trans. Smart Grid 7 (2) (2016) 650–658.
[12] L. Hernandez, et al., A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings, IEEE Commun. Surv. Tutor. (2014) 1–36.
[13] W. Charytoniuk, M.-S. Chen, Very short-term load forecasting using artificial neural networks, IEEE Trans. Power Syst. 15 (1) (2000) 263–268.
[14] C. Guan, P.B. Luh, L.D. Michel, Y. Wang, P.B. Friedland, Very short-term load forecasting: wavelet neural networks with data pre-filtering, IEEE Trans. Power Syst. 28 (1) (2013) 30–41.
[15] A. Piras, A. Germond, B. Buchenel, K. Imhof, Y. Jaccard, Heterogeneous artificial neural network for short term electrical load forecasting, IEEE Trans. Power Syst. 11 (1) (1996) 397–402.
[16] N. Amjady, Short-term bus load forecasting of power systems by a new hybrid method, IEEE Trans. Power Syst. 22 (1) (2007) 333–341.
[17] L. Zjavka, V. Snášel, Short-term power load forecasting with ordinary differential equation substitutions of polynomial networks, Electr. Power Syst. Res. 137 (2016) 113–123.
[18] E. Doveh, P. Feigin, D. Greig, L. Hyams, Experience with FNN models for medium term power demand predictions, IEEE Trans. Power Syst. 14 (2) (1999) 538–546.
[19] M.S. Kandil, S.M. El-Debeiky, N.E. Hasanien, Long-term load forecasting for fast developing utility using a knowledge-based expert system, IEEE Trans. Power Syst. 17 (2) (2010) 491–496.
[20] R.J. Hyndman, S. Fan, Density forecasting for long-term peak electricity demand, IEEE Trans. Power Syst. 25 (2) (2010) 1142–1153.
[21] A.S. Khwaja, M. Naeem, A. Anpalagan, A. Venetsanopoulos, B. Venkatesh, Improved short-term load forecasting using bagged neural networks, Electr. Power Syst. Res. 125 (2015) 109–115.
[22] J. Fiot, F. Dinuzzo, Electricity demand forecasting by multi-task learning, IEEE Trans. Smart Grid 99 (2016).
[23] M.H. Amini, A. Islam, Allocation of electric vehicles' parking lots in distribution network, in: IEEE Innovative Smart Grid Technologies Conference (ISGT), 2014.
[24] I. Rahman, et al., Review of recent trends in optimization techniques for plug-in hybrid, and electric vehicle charging infrastructures, Renew. Sustain. Energy Rev. 58 (2016) 1039–1047.
[25] S. Bahrami, M. Parniani, Game theoretic based charging strategy for plug-in hybrid electric vehicles, IEEE Trans. Smart Grid 5 (5) (2014) 2368–2375.
[26] S. Bahrami, V.W.S. Wong, A potential game framework for charging PHEVs in smart grid, in: IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM), Victoria, Canada, 2015, pp. 28–33.
[27] H. Arasteh, et al., SoS-based multiobjective distribution system expansion planning, Electr. Power Syst. Res. 141 (2016) 392–406.
[28] K.B. Song, Y.S. Baek, D.H. Hong, G. Jang, Short-term load forecasting for the holidays using fuzzy linear regression method, IEEE Trans. Power Syst. 20 (1) (2005) 96–101.
[29] P.E. McSharry, S. Bouwman, G. Bloemhof, Probabilistic forecast of the magnitude and timing of peak electricity demand, IEEE Trans. Power Syst. 20 (2) (2005) 1166–1172.
[30] B.J. Chen, M.W. Chang, C.-J. Lin, Load forecasting using support vector machines: a study on EUNITE competition 2001, IEEE Trans. Power Syst. 19 (4) (2004) 1821–1830.
[31] A. Pardo, V. Meneu, E. Valor, Temperature and seasonality influences on Spanish electricity load, IEEE Energy Econ. 24 (2002) 55–1830.
[32] A.P. Douglas, A.M. Breipohl, F.N. Lee, R. Adapa, Load forecasting using support vector machines: a study on EUNITE competition 2001, IEEE Trans. Power Syst. 13 (4) (1998) 1507–1513.
[33] M.S. Kandil, S.M. El-Debeiky, N.E. Hasanien, Overview and comparison of long-term forecasting techniques for a fast developing utility: Part I, Electr. Power Syst. Res. 58 (1) (2001) 11–17.
[34] H. Morita, T. Kase, Y. Tamura, S. Iwamoto, Interval prediction of annual maximum demand using grey dynamic model, Int. J. Electr. Power Energy Syst. 18 (7) (1996) 409–413.
[35] D.C. Park, et al., Electric load forecasting using an artificial neural network, IEEE Trans. Power Syst. 6 (2) (1991) 442–449.
[36] K.Y. Lee, Y.T. Cha, J.H. Park, Short-term load forecasting using an artificial neural network, IEEE Trans. Power Syst. 16 (1) (2001) 124–132.
[37] H.S. Hippert, C.E. Pedreira, R.C. Souza, Neural networks for short-term load forecasting: a review and evaluation, IEEE Trans. Power Syst. 16 (1) (2001) 44–55.
[38] A.I. Sarwat, M.H. Amini, A. Domijan Jr., A. Damnjanovic, F. Kaleem, Weather-based interruption prediction in the smart grid utilizing chronological data, J. Mod. Power Syst. Clean Energy 4 (2) (2016) 308–315.
[39] N. Amjady, Short-term hourly load forecasting using time-series modeling with peak load estimation capability, IEEE Trans. Power Syst. 16 (3) (2001) 498–505.
[40] W. Charytoniuk, M.-S. Chen, Very short-term load forecasting using artificial neural networks, IEEE Trans. Power Syst. 15 (1) (2002) 263–268.
[41] Y. Wang, Q. Xia, C. Kang, Secondary forecasting based on deviation analysis for short-term load forecasting, IEEE Trans. Power Syst. 26 (2) (2011) 500–507.
[42] C. Garca-Ascanio, C. Mate, Electric power demand forecasting using interval time series: a comparison between VAR and iMLP, Energy Policy 38 (2) (2010) 715–725.
[43] D.C. Park, M.A. El-Sharkawi, R.J. Marks II, L.E. Atlas, M.J. Damborg, Electric load forecasting using an artificial neural network, IEEE Trans. Power Syst. 6 (2) (1991) 442–449.
[44] G. Dudek, Pattern-based local linear regression models for short-term load forecasting, Electr. Power Syst. Res. 130 (2016) 139–147.
[45] T.H.D. Ngo, The Box–Jenkins methodology for time series models, SAS Glob. Forum 54 (2013).
[46] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, G.M. Ljung, Time-Series Analysis: Forecasting and Control, Holden-Day, CA, 1976.
[47] S.S. Pappas, L. Ekonomou, P. Karampelas, D.C. Karamousantas, S.K. Katsikas, G.E. Chatzarakis, P.D. Skafidas, Electricity demand load forecasting of the Hellenic power system using an ARMA model, Electr. Power Syst. Res. 80 (3) (2010) 256–264.
[48] A.A. El Desouky, M.M. Elkateb, Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA, IEE Proc. Gener. Transm. Distrib. 147 (4) (2000) 213–217.
[49] S.J. Huang, K.R. Shih, Short-term load forecasting via ARMA model identification including non-Gaussian process considerations, IEEE Trans. Power Syst. 18 (2) (2003) 673–679.
[50] J. Contreras, R. Espinola, F.J. Nogales, A.J. Conejo, ARIMA models to predict next-day electricity prices, IEEE Trans. Power Syst. 18 (3) (2003) 1014–1020.
[51] J.W. Taylor, Triple seasonal methods for short-term load forecasting, Eur. J. Oper. Res. 204 (1) (2010) 139–152.
[52] G.E.P. Box, D.R. Cox, An analysis of transformations, J. R. Stat. Soc. Ser. B 26 (1964) 211–252.
[53] http://www.pjm.com/markets-and-operations/ops-analysis/historical-load-data.aspx.
[54] H. Shaker, H. Chitsaz, H. Zareipour, D. Wood, On comparison of two strategies in net demand forecasting using wavelet neural network, in: North American Power Symposium (NAPS), 2014, pp. 1–6.
[55] Available Online: http://pjm.com/about-pjm/who-we-are.aspx.
[56] Available Online: http://pjm.com//media/about-pjm/pjm-zones.ashx.