Energy Conversion and Management 80 (2014) 582–590

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Ensemble re-forecasting methods for enhanced power load prediction

Amanpreet Kaur, Hugo T.C. Pedro, Carlos F.M. Coimbra *

Department of Mechanical and Aerospace Engineering, Jacobs School of Engineering, Center for Energy Research and Center of Excellence in Renewable Resource Integration, University of California San Diego, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:
Received 14 October 2013
Accepted 3 February 2014
Available online 25 February 2014

Keywords:
Load forecast
Re-forecast
Hour-ahead market
Day-ahead market
Ensemble method

ABSTRACT

Electric load forecasting is a key element for management and operation of the electric grid. In this study we introduce ensemble re-forecast methods that take an initial forecast and produce a better prediction by extracting information from the structured errors. The models in the ensemble rely upon the real-time information obtained from load measurements and estimates over a state-wide domain. The weights in the ensemble are optimized in three different ways based on global, hourly, and weekly performance of the models. The proposed methodology is applied to predict hour-ahead market (HAM) and day-ahead market (DAM) load for California Independent System Operator (CAISO) and Electric Reliability Council of Texas (ERCOT) respectively. Proposed models showed consistent performance enhancements for all the cases. HAM predictions show an improvement of 47% and 36% in terms of Mean Absolute Percentage Error over the forecasts provided by CAISO and ERCOT. For DAM, the improvements are 34% for CAISO and 47% for ERCOT. Temporal analysis comparing the internal forecast produced by the ISOs and re-forecasts shows significant improvement during off-peak hours and small improvement for on-peak hours. Results validate the potential of the proposed methodology to enhance the forecast accuracy, independent of load profile or forecast horizon.

1. Introduction

Short term load forecast (STLF) plays a key role in operation, control, and management of the grid. Each forecast depending on its forecast horizon has a specific application for grid balancing and scheduling. Day-ahead market (DAM) load forecasts are required by Independent System Operators (ISOs), utilities, and electricity market participants for operation planning and unit commitment of generating plants. Hour-ahead market (HAM) load forecasts are needed for real-time control and load following. To compensate for the uncertainty in load forecasts ISO maintains an operating reserve [1], which increases cost for ISOs, utilities, and customers [2].

Continuous research is being conducted to lower the uncertainty and increase the accuracy of load forecasts. Various methodologies based on time-series, regression, fuzzy logic, Artificial Neural Networks (ANNs), expert systems, hybrid models, etc., have been proposed [3]. Most of these methods rely on external inputs such as meteorological forecasts, temperature, dew point, etc., [4,5]. Since no forecast is completely accurate, the errors in meteorological forecasts are propagated to the load forecasts [4]. Thus, the study of the forecast error distribution is essential to reveal the error structure which, if systematic and non-random, could be refined for better prediction.

In this work, we propose refining techniques that extracts the information from the structured forecast error needed to enhance forecast accuracy. The predictions are ensembled using optimized weights for the models, depending on their global, hourly, and weekly performance. The proposed methodology is applied to forecast load. We use DAM load estimates provided by the ISOs as an exogenous input to re-forecast HAM and DAM loads.

The major contribution of this paper is the demonstration that with the proper re-forecasting and ensemble techniques it is possible to substantially improve publicly available utility and ISOs load forecasts. Moreover, we demonstrate that it can be achieved without requiring any additional information (such as weather forecast). Another contribution is that our algorithms can be used to produce short-term forecasts based on the utility predictions, without requiring exogenous information. The proposed ensemble re-forecast methods for load prediction have several advantages: (1) methods are independent of seasonal cycles; rather they depend on the structure in the forecast residues. However, seasonal effects in the base forecast may influence the re-forecast produced by the proposed algorithm, (2) the weights of the time-variant model are updated as the new data becomes available, so, the

* Corresponding author. Tel.: +1 (858) 534 4285.
E-mail address: ccoimbra@ucsd.edu (C.F.M. Coimbra).

http://dx.doi.org/10.1016/j.enconman.2014.02.004
0196-8904/

Nomenclature

δ	detrend polynomial
ϵ	error in individual forecast model
$\hat{\cdot}$	forecast of $\cdot$
$\hat{\phi}$	ensemble re-forecast
S	simulation set
T	training set
V	validation set
$\mathcal{T}$	vector for the timestamps of the data $[t_1, t_2, \ldots, t_k]^T$
μ	detrended input load
ψ	detrended actual load
$A(q)$	polynomial with an order n_a
$B(q)$	polynomial with an order n_b
$C(q)$	polynomial with an order n_c
$D(q)$	polynomial with an order n_d
E	vector representing the error $[e_1, e_2, \ldots e_k]^T$
$F(q)$	polynomial with an order n_f
h	hour of the day
n_k	input–output delay parameter
q	shift operator
t	time
u	input for the model
W	vector representing the weights $[\omega_1, \omega_2, \ldots \omega_k]^T$
w	day of the week
Y	vector representing the actual load $[y_1, y_2, \ldots y_k]^T$
y	actual load
ARMAX	Autoregressive moving average model with exogenous input
ARX	Autoregressive model with exogenous input
BJ	Box–Jenkins model
CAISO	California Independent System Operator
DAM	day-ahead market
ERCOT	Electric Reliability Council for Texas
GM	generalized model
HAM	hour-ahead market
ISO	Independent System Operator
LS	least squares
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NARX	non-linear ARX model
RLS	recursive least squares
STLF	short term load forecast

model is robust for the frequent changes in the time-series, and (3) the models can make predictions for any forecast horizon given that it is equal to or smaller than the forecast horizon of the input estimate.

The load data used in this study is from California Independent System Operator (CAISO) and Electric Reliability Council of Texas (ERCOT). Real-time load measurements, DAM and HAM load forecasts are publicly available on their websites [6,7].

The organization of the paper is as follows: literature review about the load forecast, ensemble methods, and re-forecast is presented in Section 2; proposed methodology is described in Section 3; dataset used for validation is explained in Section 4; results are discussed in Section 5 and the conclusions are drawn in Section 6.

2. Literature review

Several comprehensive reviews on load forecasting techniques for STLF have been published as research has progressed in this field. Comprehensive reviews on time-series modeling and forecasting can be found in [4,8,9]. The authors in [9] concluded that Box–Jenkins time-series models are well suited for STLF, but they do not capture the non-linearities in the load time-series. To circumvent this problem, they proposed a polynomial regression to linearize the relationship between load and temperature before using the time-series models.

A comparative evaluation of linear regression, stochastic time-series, general exponential smoothing, states-space methods, and knowledge-based approaches was presented in [10] by comparing their performance for the same load time-series. They showed that the regression and state space methods do not respond to sudden changes in load. Hence, the authors recommended to update the model parameters automatically and develop models specific to weekends and holidays.

In the 1990s, Artificial Neural Networks (ANNs), fuzzy logic, and Computational Intelligence (CI) techniques became popular and several studies reviewing their application for load forecasting were published [5,11–13]. The studies [5,14] emphasized the need for more rigorous standard tests to compare the ANN model performance and check for over-fitting and over-parameterization. Despite these concerns, [11] stated that ANNs capture the non-linearities in load data better than time-series based regression methods, and the potential of ANNs should not be overlooked.

Results from the Artificial Neural Networks and Computational Intelligence Forecasting competition (NN3 competition) were presented in [15] with a detailed discussion on merits, demerits, and performance of CI and statistical methods. It was observed that ensemble models outperformed the individual models. More recent forecast and optimization methods like support vector regression (SVR), simulated annealing, ant colony optimization, and other hybrid methods for energy forecasting were reviewed in [16,3,17]. All these reviews highlight the promising application of optimized hybrid and ensemble methods in load and other areas of energy forecasting [18–22].

Following the conclusions that combining several methods outperform single methods, we focus on ensemble optimization. Ensemble methods have been rigorously applied and tested for load forecasting (see Table 1). For instance in [23], local predictors were trained on selected inputs using k Nearest Neighbors (kNN). This was beneficial because local predictors were able to retain the forecast accuracy for longer lead times than global predictors. The studies [24,25] proposed using forecasted weather ensembles for load prediction. In [26], different training sets were used to train a network committee. The study [27] could be seen as a hybrid of [25,26]. In this work various weather forecasts from different forecast services were combined using adaptive weights and an ensemble comprised of five ANNs.

The study [28] proposed an ensemble of kernels-based gaussian processes (GPs) where the linear model parameters were selected using genetic algorithm (GA), and showed that the ensemble outperforms SVR and ARMA. Various ensemble methods like simple averaging, singular value decomposition (SVD) based weighted averaging (for each hour separately), principal component analysis (PCA), and blind source separation were investigated in [29]. The best results were obtained by blind source separation method, where the data was decomposed into statistically independent components and the time-series was reconstructed by filtering the noise.

The studies [30,31] proposed using selective ensemble techniques, i.e., a combination of models was optimally selected to

Table 1
Ensemble models applied for short term load forecasting.

Ref.	Exogenous inputs	Training set selection	No. of models	Forecast models	Ensemble methods
[23]	Average temperature, maximum and minimum temperature	k Nearest Neighbor (kNN)	2	ANNs	Mean
[24]	51 Forecasted weather ensembles for temperature, wind speed, and cloud cover		51	Regression model	Mean
[25]	51 Forecasted weather ensembles for temperature, wind speed, and cloud cover		51	ANNs	Mean
[26]	Temperature and day-type variables		3	ANN committee, abductive network committees	Mean
[33]	Temperature, rainfall index, and holiday index			Fuzzy hyperrectangular composite neural network	Chaos-search genetic algorithm (CGA) and simulated annealing (SA)
[27]	Forecasted weather from different services combined using adaptive weights	Bootstrap sampling	5	ANNs	Median
[30]	Temperature		20	ANNs	Mean and entropy weighted method for models selected using discrete differential algorithm
[29]	Type of the day, season of the year		3	Multilayer perceptron layer (MLP), self organizing map (SOM), and fuzzy SOM	Mean, singular value decomposition (SVD) based weighted averaging (for each hour separately), principal component analysis (PCA) decomposition, and blind source separation (BSS)
[28]			3	Kernel-based gaussian processes (GPs) regression	Evolutionary optimization
[31]	Temperature (if available)		7	Random walk (RW) algorithm, autoregressive moving average (ARMA), similar days algorithm, layer recurrent neural network (LRNN), MLP, ν-support vector regression (ν-SVR), and robust LS-SVM (RobLSSVM)	Meta learning

form an ensemble. A discrete differential algorithm was used in [30] to select a combination of ANN models whereas a meta-learning technique was applied in [31] to select models based on meta-features like highest ACF, granularity, fickleness, etc. Results for various load profiles with different forecast horizons were presented to show that the meta-learning based ensemble is independent of load profile, forecast horizon, and outperforms various well established load forecasting models.

Motivated by the results from ensemble methods we propose novel re-forecast ensemble methods that filters the structured noise from a given estimate to produce a better prediction. We address the issue of extracting the information from non-random errors to further improve the forecasts. Previously, the re-forecast concept was investigated by [32] for DAM load forecast, where the forecast was updated to account for sudden changes in the weather. In this study no sudden update is made once the forecast has been issued.

3. Proposed methodology

The schematic for the proposed methodology to improve the accuracy of publicly available utility and ISOs load forecasts and generate forecasts for shorter forecast horizon with no exogenous input is shown in Fig. 1. A base forecast with given forecast horizon k' is required to apply this methodology to produce k steps ahead forecast s.t. $k \leqslant k'$. For $k = k'$, if the forecast residuals are white then it is considered as the final forecast otherwise re-forecast and ensemble techniques are applied unless the residuals are white. Similarly, for $k < k'$, re-forecast and ensemble technique are applied directly to produce a short-term forecast and then again the residues are checked for whiteness. In this study, the proposed methodology is applied for $k = 1$ and $k \in \{1, 2, \ldots, 24\}$ forecast horizon and the details are explained below.

3.1. HAM re-forecast

Using real-time system measurements, a given input $u(t)$ is processed every hour by ensemble re-forecast models (m) to produce HAM point forecast $\hat{y}(t+k)$,

$$\hat{y}(t+k) = m(y(t), y(t-1), \ldots, u(t), u(t-1), u(t-2), \ldots), \quad (1)$$

where t is time and $k = 1$ is the forecast horizon. We use 24-hourly DAM load estimates provided by the ISO as an exogenous input, filter it every hour and produce HAM load forecasts for the system. For example, ISO DAM forecast produced on June 11, 2013 for the following 24 h is corrected hourly by our re-forecast in order to produce HAM point forecast.

3.2. DAM re-forecast

The above idea is extended for DAM load prediction with forecast horizons $k \in \{1, 2, \ldots, 24\}$. In this case, hourly forecasts are made at midnight for each hour $h \in \{1, 2, \ldots, 24\}$, i.e., $\hat{Y} = [\hat{y}(t+1), \hat{y}(t+2), \ldots, \hat{y}(t+24)]^T$. The time-series models $m(h)$ are built specific for each hour of the day with an exogenous input consisting of estimates for that hour only $u(t-24n)|_{n=0,1,\ldots}$ and corresponding actual measurements $y(t-24n)|_{n=1,2,\ldots}$. In summary, we use the twenty-four hourly DAM estimated load provided by the ISO as an input, refine it by applying the models developed for each hour, combine the forecast from all the models and produce a vector of 24 load forecasts corresponding to all the hours of the next day.

3.3. Preprocessing

The time-series is preprocessed to remove the outliers and days with missing data. After preprocessing, the hourly time-series is

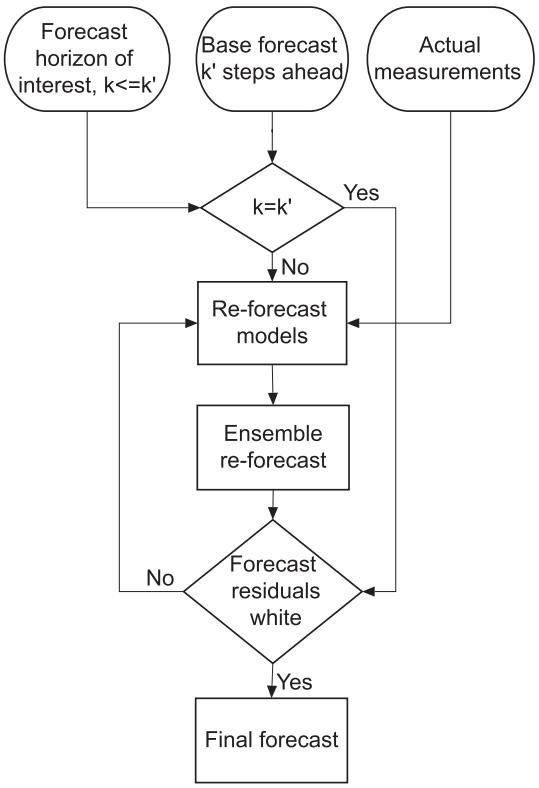

Fig. 1. Schematic of the proposed methodology for ensemble re-forecasting. A base forecast up to k' steps ahead forecast horizon is required to produce k steps ahead forecast s.t. $k \leqslant k'$. If the forecast residuals are white then it is consider as the final forecast else re-forecast and ensemble techniques are applied unless the residuals are white.

generated and separated into three disjoint datasets: training set (**T**), simulation set (**S**) and validation set (**V**). The data points in **T** were used to derive a polynomial fit to detrend the time-series and develop the models which were then tested using the **S** dataset. The forecasts from **T** and **S** were combined to optimize the weights for the ensemble.

3.4. Detrending

In this work we applied a simple polynomial fit that captures the daily trend for the electric load. Using a simple least squares method (from the Curve Fitting toolbox available in Matlab) we determine that daily trends are best fitted by a 6th order polynomial as a function of time of the day t, specific for each weekday $w \in \{1, 2, \ldots, 7\}$ where Sunday is represented as 1, Monday as 2, and so on. The actual load data can be expressed as a sum of the polynomial fit and the detrended load, i.e.,

$$y(t, w) = \psi(t, w) + \delta(t, w), \tag{2}$$

where y represents the actual load and ψ is the corresponding detrended load using a polynomial fit for each day of the week $\delta(t, w)$. Similarly, the input for the models, i.e., a given estimate u is detrended,

$$u(t, w) = \mu(t, w) + \delta(t, w), \tag{3}$$

where μ is the detrended input using the polynomial fit $\delta(t, w)$.

3.5. Re-forecast models

This study focuses on extracting the information from non-random noise to produce better forecasts using time-series models with the assumption that the model's residue is a white Gaussian noise. The filtered output produced by any of the models is called a re-forecast.

Various time-series models are applied at re-forecast stage. The linear and non-linear time-series models used in this work are well defined in the literature [4,8,9,34]. Only a brief description is provided here.

3.5.1. Linear models

A generalized model (GM) that linearly combines current and past values of the input $\mu(t)$ and past values of the output $\psi(t)$ to model current output can be defined as [34],

$$A(q)\psi(t) = \frac{B(q)}{F(q)}(q)\mu(t - n_k) + \frac{C(q)}{D(q)}(q)\epsilon(t), \tag{4}$$

where t represents time, n_k is the input–output delay parameter, ϵ is assumed to be white noise, q is a shift operator, and $A(q), B(q), C(q), D(q)$, and $F(q)$ are the polynomials of order $n_a, n_b - 1, n_c, n_d$, and n_k, i.e.,

$$q^{\pm N}\psi(t) = \psi(t \pm N), \tag{5}$$

$$A(q) = 1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a}, \tag{6}$$

$$B(q) = b_1 + \ldots + b_{n_b} q^{-n_b+1}, \tag{7}$$

$$C(q) = 1 + c_1 q^{-1} + \ldots + c_{n_c} q^{-n_c}, \tag{8}$$

$$F(q) = 1 + f_1 q^{-1} + \ldots + f_{n_f} q^{-n_f}, \tag{9}$$

$$D(q) = 1 + d_1 q^{-1} + \ldots + d_{n_d} q^{-n_d}. \tag{10}$$

Based on the application of GM and information about the physical system one or many polynomials can be fixed to unity. In this study we also apply the other forms of GM such as Autoregressive model with exogenous input (ARX), Autoregressive moving average model with exogenous input (ARMAX), and Box–Jenkins model (BJ). Table 2 lists the polynomials used in each one of these models.

3.5.2. Non-linear model

To model the non-linearities in the load time-series we use the Non-linear Autoregressive model (NARX). In this model the current output is expressed as a non-linear combination of the input and the past values of the output,

$$\psi(t) = f(\psi(t-1), \psi(t-2), \ldots, \mu(t), \mu(t-1), \mu(t-2), \ldots), \tag{11}$$

where $f(\cdot)$ can be computed using any non-linear estimator. In this study, we use a wavelet network non-linear estimator. The prediction error method (PEM) is applied to derive the model parameters that minimize the weighted norm of the prediction error for the dataset **T**.

3.6. Ensemble re-forecast

Ensemble re-forecast is produced by combining the above re-forecasts. The trend $\delta(t, w)$ is added to the re-forecast $\hat{\psi}(t)$ to produce a final re-forecast $\hat{y}(t)$,

$$\hat{y}(t, w) = \hat{\psi}(t, w) + \delta(t, w). \tag{12}$$

We investigate three ways of forming an ensemble by optimizing the weights based on global, hourly, and weekly performance of the models. The mathematical details about the proposed ensembles are provided in the next subsections.

Table 2
Polynomials used for various model structures.

Model parameters	$A(q)$	$B(q)$	$C(q)$	$D(q)$	$F(q)$
ARX	✓	✓			
ARMAX	✓	✓	✓		
BJ		✓	✓	✓	✓
GM	✓	✓	✓	✓	✓

3.6.1. Least squares model ensemble (LS-ME)

An ensemble forecast $\hat{\phi}$ for any given time t is produced by linearly combining forecasts from n forecasting models for k forecast horizons,

$$\hat{Y} = \begin{pmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} & \dots & \hat{y}_{1,n} \\ \hat{y}_{2,1} & \hat{y}_{2,2} & \dots & \hat{y}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{k,1} & \hat{y}_{k,2} & \dots & \hat{y}_{k,n} \end{pmatrix}, \tag{13}$$

using the weights $W = [\omega_1, \omega_2, \dots, \omega_n]^T$, i.e.,

$$\hat{\phi} = \hat{Y}W \tag{14}$$

For HAM load prediction $k = 1$ and for DAM $k = \{1, 2, \dots, 24\}$. The weights ω_i are optimized based on the global performance of the models applied to the T and S datasets, which consist of p actual load measurements $Y = [y_1, y_2, \dots, y_p]^T$. Each measurement can be uniquely identified by a timestamp $\mathcal{T} = [t_1, t_2, \dots, t_p]^T$. Correspondingly, there are p hourly forecasted data points from n forecasting models $\hat{y}_{i,j}$ for $i = 1, 2, \dots, p$ and $j = 1, 2, \dots, n$,

$$\hat{Y} = \begin{pmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} & \dots & \hat{y}_{1,n} \\ \hat{y}_{2,1} & \hat{y}_{2,2} & \dots & \hat{y}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{p,1} & \hat{y}_{p,2} & \dots & \hat{y}_{p,n} \end{pmatrix}. \tag{15}$$

The forecasted data points $\hat{Y}$ are combined using $W = [\omega_1, \omega_2, \dots, \omega_n]^T$ to produce a final forecast $\hat{\phi} = [\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p]^T$ with an error $E = [e_1, e_2, \dots, e_p]^T$,

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} & \dots & \hat{y}_{1,n} \\ \hat{y}_{2,1} & \hat{y}_{2,2} & \dots & \hat{y}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{p,1} & \hat{y}_{p,2} & \dots & \hat{y}_{p,n} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{pmatrix} = \begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{pmatrix}, \tag{16}$$

such that $E = Y - \hat{Y}W$, where W is approximated using linear least squares with an objective function of minimizing the sum of squares of errors, i.e., $G(T, W) = \frac{1}{2}E^T E$, whose global solution is $W = (\hat{Y}^T\hat{Y})^{-1}\hat{Y}^T Y \in \mathbb{R}^n$.

3.6.2. Least squares hourly ensemble (LS-HE)

In this model an ensemble forecast $\hat{\phi}(h)$ depends on the hour of the day h, which can be easily retrieved from the timestamp t, i.e., $h \subset t$. $\hat{\phi}(h)$ is produced by linearly combining forecasts for the hth hour from n forecasting models with k forecast horizons $\hat{y}_{ij}(h), i \in \{1, 2, \dots, k\}$ and $j \in \{1, 2 \dots, n\}$ using the weights $\omega_j(h)$,

$$\hat{\phi}(h) = \sum_{i=1}^{k}\sum_{j=1}^{n}\hat{y}_{i,j}(h)\omega_j(h). \tag{17}$$

Using the timestamps $\mathcal{T}$, only q data points with $q \leqslant p$ corresponding to hth hour are considered to compute the weights $w_i(h)$ and the same procedure as defined above is followed, i.e.,

$$\begin{pmatrix} y_1(h) \\ y_2(h) \\ \vdots \\ \hat{y}_q(h) \end{pmatrix} = \begin{pmatrix} \hat{y}_{1,1}(h) & \hat{y}_{1,2}(h) & \dots & \hat{y}_{1,n}(h) \\ \hat{y}_{2,1}(h) & \hat{y}_{2,2}(h) & \dots & \hat{y}_{2,n}(h) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{q,1}(h) & \hat{y}_{q,2}(h) & \dots & \hat{y}_{q,n}(h) \end{pmatrix} \begin{pmatrix} \omega_1(h) \\ \omega_2(h) \\ \vdots \\ \omega_n(h) \end{pmatrix} + \begin{pmatrix} e_1(h) \\ e_2(h) \\ \vdots \\ e_q(h) \end{pmatrix} = \begin{pmatrix} \hat{\phi}_1(h) \\ \hat{\phi}_2(h) \\ \vdots \\ \hat{\phi}_q(h) \end{pmatrix} + \begin{pmatrix} e_1(h) \\ e_2(h) \\ \vdots \\ e_q(h) \end{pmatrix}. \tag{18}$$

The above algorithm is repeated twenty-four times $\forall h \in \{1, 2, \dots, 24\}$ to compute individual weights for each hour of the day for all n models. In this case, $W \in \mathbb{R}^{n,h}$.

3.6.3. Least squares weekday ensemble (LS-WE)

This model produces ensemble forecast $\hat{\phi}(d)$ depending on the day of the week $d \in \{1, 2, \dots, 7\}$, which can be computed from the timestamp t, i.e., $d \subset t$. $\hat{\phi}(d)$ is produced by linearly combining forecasts from n forecasting models with k forecast horizons $\hat{y}_{ij}(h), i \in \{1, 2, \dots, k\}$ and $j \in \{1, 2 \dots, n\}$ using the weights $\omega_i(d)$ specific for each weekday,

$$\hat{\phi}(d) = \sum_{i=1}^{k}\sum_{j=1}^{n}\hat{y}_{i,j}(d)\omega_j(d). \tag{19}$$

In this case, from the dataset using $\mathcal{T}$, only r data points with $r \leqslant p$ corresponding to d are considered to compute the weights $\omega_i(d)$,

$$\begin{pmatrix} y_1(d) \\ y_2(d) \\ \vdots \\ \hat{y}_r(d) \end{pmatrix} = \begin{pmatrix} \hat{y}_{1,1}(d) & \hat{y}_{1,2}(d) & \dots & \hat{y}_{1,n}(d) \\ \hat{y}_{2,1}(d) & \hat{y}_{2,2}(d) & \dots & \hat{y}_{2,n}(d) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{r,1}(d) & \hat{y}_{r,2}(d) & \dots & \hat{y}_{r,n}(d) \end{pmatrix} \begin{pmatrix} \omega_1(d) \\ \omega_2(d) \\ \vdots \\ \omega_n(d) \end{pmatrix} + \begin{pmatrix} e_1(d) \\ e_2(d) \\ \vdots \\ e_r(d) \end{pmatrix} = \begin{pmatrix} \hat{\phi}_1(d) \\ \hat{\phi}_2(d) \\ \vdots \\ \hat{\phi}_r(d) \end{pmatrix} + \begin{pmatrix} e_1(d) \\ e_2(d) \\ \vdots \\ e_r(d) \end{pmatrix}. \tag{20}$$

The above algorithm is repeated seven times $\forall d \in \{1, 2, \dots, 7\}$ to compute individual weights for each weekday and forecasting model. For this ensemble $W \in \mathbb{R}^{n,d}$.

3.7. Recursive least squares model ensemble (RLS-ME)

This method only applies to point forecast with one forecast horizon. For real-time application, this model produces an hourly ensemble forecast $\hat{\phi}(t)$ depending on the time of the forecast t. $\hat{\phi}(t)$ is produced by linearly combining forecasts from n forecasting models, $\hat{y}_i(t), i \in \{1, 2, \dots, n\}$ using the weights $w_i(t)$ updated based on the last measurement,

$$\hat{\phi}(t) = \sum_{i=1}^{n}\hat{y}_i(t)\omega_i(t). \tag{21}$$

The model is initialized using the weights computed by the LS-ME method as they represent the best estimate of weights for the individual model. The weights $w_i(t)$ are updated using the following RLS algorithm [35],

$$\omega(t) = \omega(t-1) + K(t)(y(t) - \hat{y}(t)')\omega(t-1), \tag{22}$$

$$K(t) = \frac{P(t-1)\hat{y}(t)}{1 + \hat{y}(t)'P(t-1)\hat{y}(t)}, \tag{23}$$

and

$$P(t) = (I - K(t)\hat{y}(t)')P(t-1). \tag{24}$$

4. Data

As explained in Section 3.3, the collected data was preprocessed and separated into three disjoint datasets: training set (**T**), simulation set (**S**), and validation set (**V**). For details see Table 3 and Fig. 2. The results presented in this study are for the **V** dataset. As mentioned above, we implemented load re-forecast models for two ISOs: CAISO and ERCOT.

Table 3
Details for the dataset used for HAM and DAM forecast.

Data source	Dataset label	Time-period	Data points
CAISO	**T**	11/2/2012 to 2/28/2013	2256
	S	3/1/2013 to 3/17/2013	360
	V	3/18/2013 to 6/11/2013	2016
ERCOT	**T**	12/01/2012 to 04/14/2013	1848
	S	04/15/2013 to 04/30/2013	336
	V	05/01/2013 to 06/20/2013	528

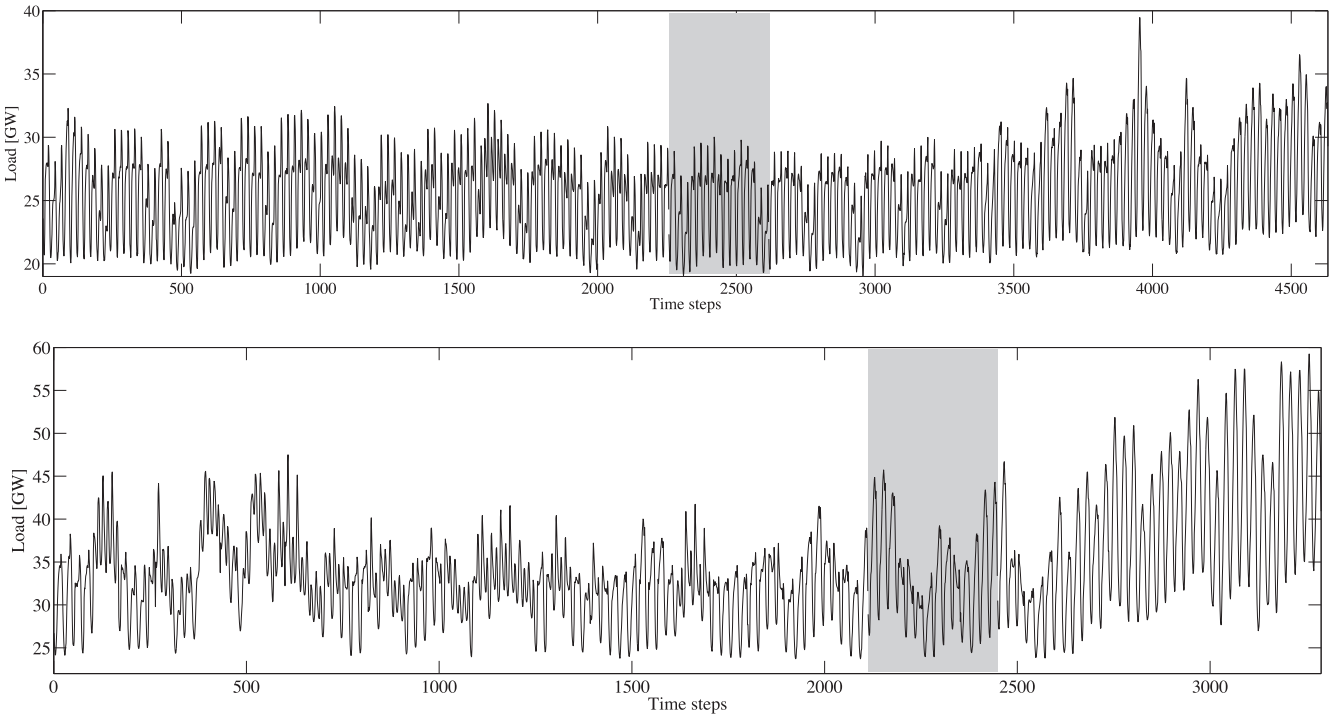

Fig. 2. The CAISO time-series (top) shows load measurements from November 2, 2012 to June 11, 2013. The ERCOT time-series (bottom) shows load measurements for Texas, ranging from December 1, 2012 to June 20, 2013. The measurements on the left of the shaded region represent training set **T**, the shaded portion highlights simulation set **S**, and the time-series on the right of the shading region is validation set **V**. The influence of heat and summer weather can be seen in **V** with sudden spikes and increase in load demand. It can be observed that the ERCOT load profile is highly variable and even daily trends are change as the seasons change from winter to summer season.

4.1. California Independent System Operator (CAISO)

CAISO operates the power grid and electricity market for California (CA). It covers almost 80% of the California-Mexico (CAMX) power area [36,37]. The remaining 20% of CA's state load is covered by the Los Angeles Department of Water and Power, the Sacramento Municipal Utility District, and the Imperial Irrigation District. Their load forecasts are based on the weather data provided by the NOAA National Weather Service for various stations in CA.

4.2. Electric Reliability Council of Texas (ERCOT)

ERCOT covers almost 85% of Texas' load [38]. With centralized control ERCOT is responsible for transmission reliability, wholesale open access, and management of wholesale market for balancing energy and ancillary services for Texas.

5. Results and discussion

The above methods were implemented in Matlab using the system identification toolbox. The results and discussion presented here are for the re-forecasts and ensemble re-forecasts produced for the validation set, i.e., $\hat{y}, \hat{\phi} \in \mathbf{V}$. For the RLS algorithm, the HAM re-forecast started two days before the validation set to discard the initial effect of sudden unstable changes in weights. However, only the results from **V** were used in the error analysis.

For HAM, the parameters of the CAISO time-series model were selected using minimum description length (MDL) criterion. Using those parameters as reference, similar parameters were selected for ERCOT with small changes (Table 4). For DAM, global search was performed using data points in **T** and **S**. Before applying least squares ensemble methods for DAM load, the forecast vector was produced by combining the re-forecast from various models. They are denoted as ARX-E, ARMAX-E, BJ-E, GM-E, and NARM-E to represent their ensemble for DAM load forecasts.

5.1. Error statistics

Standard error metrics commonly used to compare the performance of the forecast models are: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Mean Bias Error (MBE), and Root Mean Square Error (RMSE). MAPE measures the accuracy of a method in terms of percentage error. A MAPE of zero implies a perfect fit, but there is no upper bound on its value. The mean of

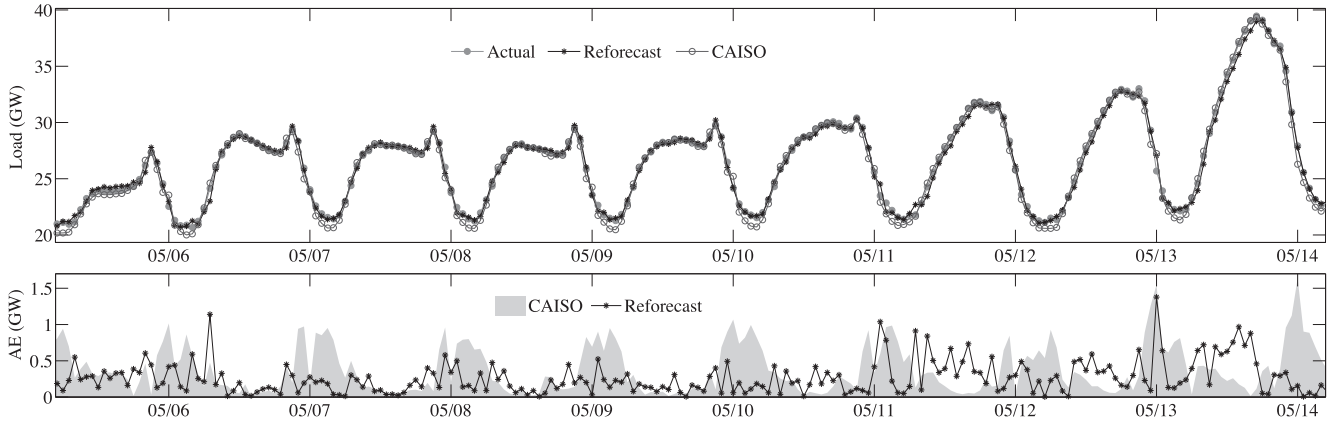

Fig. 3. Time-series of actual CAISO load and HAM load prediction for several days in 2013 by CAISO and NARX re-forecast model (top) and the instantaneous absolute error (bottom). The time-series shows change in load profiles for weekdays and weekends. In addition, it can be seen that a day like 05-13-2013 has a different load profile than the rest of the days, which can be attributed to the weather conditions. The CAISO model does not perform well during off-peak hours (22:00–8:00 PDT) and the forecast for this period is corrected by the re-forecast model.

Table 4
Number of model parameters used for CAISO and ERCOT re-forecast time-series models.

Model parameters	n_a	n_b	n_c	n_d	n_f	n_k
CAISO-HAM	2	3	5	5	3	24
ERCOT-HAM	2	3	5	5	3	23
CAISO-DAM	2	2	2	2	2	0
ERCOT-DAM	2	2	2	2	2	0

absolute errors is represented by MAE and bias in the forecast is represented by MBE. RMSE measures the average squared distance between the measured and forecasted load. The results after applying these metrics for the proposed models are presented in Tables 5 and 6. The individual models are well-established base models for load forecasting. They were evaluated before combining to compare the re-forecasts using a single model and ensemble model. This gives an insight to the reader how much improvement is achieved by re-forecast using a single model and then further by combining them.

5.1.1. HAM re-forecast

For CAISO, NARX performed the best with a MAPE of 0.85% which presents a 47% improvement over the internal HAM forecast produced by CAISO (see Fig. 3). The GM forecast had the lowest bias in the forecast. NARX and LS-ME performance is almost the same in terms of RMSE. Although BJ showed improvement in forecast when compared to CAISO, it performed the worst out of the other re-forecast models.

For ERCOT, the re-forecast showed an improvement of 36% over the existing HAM forecast produced by the ISO. Compared to all the models, LS-HE performed the best with the lowest MAPE of 1.39% and RMSE of 707.58 MW. As in the previous case, all the models showed improvement over the internal forecast provided by the ISO. Amongst all the forecast methods implemented in this study, ARX performed the worst. In terms of MBE, no model was able to perform better than the ERCOT forecast. In both cases, the RLS-ME forecast model performed similar to any other LS-based ensemble model.

5.1.2. DAM re-forecast

The re-forecasts combined from various GM models (GM-E) showed the best results with a MAPE of 1.47% and 2.00% in comparison to all other methods for both CAISO and ERCOT datasets. In terms of MBE, LS-based ensemble methods performed the best for both cases.

Comparing the results from both HAM and DAM re-forecasts, it can be inferred that ensemble methods are robust for any forecast horizon or error statistics. This highlights the promising application of ensemble re-forecasting for better prediction. Moreover, all the time-series models used are based on the assumption that ϵ is white-noise which is validated using a correlation analysis presented in the next section.

5.2. Residual analysis

5.2.1. Correlation analysis

The assumption in time-series forecasting is that the residue of the forecast should be a white noise. If the residues are not white, then the information in the residues can be used to enhance model

Table 5
Statistical error metrics for HAM load forecast using various models for CAISO (3-18-2013 to 06-11-2013) and ERCOT (05-01-2013 to 06-20-2013).

Model	CAISO				ERCOT			
	MAPE (%)	MBE (MW)	MAE (MW)	RMSE (MW)	MAPE (%)	MBE (MW)	MAE (MW)	RMSE (MW)
ISOs	1.57	255.64	379.99	519.05	2.20	26.32	798.75	1009.12
ARX	0.88	45.99	224.13	303.66	1.91	286.80	695.64	897.12
ARMAX	0.88	39.06	224.37	303.48	1.74	225.46	630.53	818.15
BJ	0.90	34.63	228.67	306.75	1.77	211.65	644.91	836.70
GM	0.90	−0.03	229.52	308.60	1.68	203.95	614.67	792.07
NARX	0.84	34.60	215.42	292.90	1.79	313.26	650.85	853.72
LS-ME	0.85	31.70	218.09	294.37	1.63	217.28	593.16	769.44
LS-DE	0.89	28.76	225.52	304.72	1.64	214.80	598.57	771.69
LS-HE	0.88	52.43	227.42	300.11	1.39	201.37	518.64	707.58
RLS-ME	0.88	58.65	224.77	302.63	1.70	231.16	650.03	815.37

Table 6
Statistical error metrics for the DAM load forecast using various models for CAISO (3-18-2013 to 06-11-2013) and ERCOT (05-01-2013 to 06-20-2013).

Model	CAISO				ERCOT			
	MAPE (%)	MBE (MW)	MAE (MW)	RMSE (MW)	MAPE (%)	MBE (MW)	MAE (MW)	RMSE (MW)
ISOs	2.26	370.37	570.55	718.86	3.76	−108.48	1359.71	1663.55
ARX-E	1.64	155.61	437.30	604.51	2.81	430.22	1098.04	1669.98
ARMAX-E	1.99	−52.29	518.97	683.69	2.98	522.77	1173.04	1667.83
BJ-E	1.48	55.71	393.91	558.10	2.76	129.95	1026.14	1357.29
GM-E	1.47	55.25	391.04	545.81	2.00	64.09	776.37	1140.42
NARX-E	1.61	72.91	429.30	597.77	3.00	703.77	1183.02	1778.51
LS-ME	1.47	8.87	391.11	544.35	2.15	49.94	819.27	1142.21
LS-DE	1.54	13.33	407.95	562.36	2.31	143.02	889.64	1283.66
LS-HE	1.55	2.27	411.30	563.74	2.38	293.66	920.19	1337.18

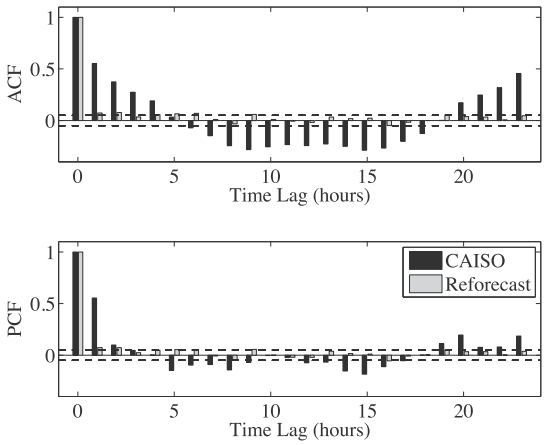

Fig. 4. Autocorrelation and partial autocorrelation for the residuals from CAISO and NARX re-forecast models for HAM load prediction. The dashed line represents the 95% confidence interval. It can be noticed that the residuals from re-forecast model are white whereas the ACF for CAISO residuals is greater than 0.05 which implies the presence of structured systematic errors. Similar pattern was observed for the ERCOT dataset.

performance. In our experience the residuals from CAISO and ERCOT forecasts showed a periodic correlation in ACF that served as a motivation to apply re-forecast. The autocorrelation function (ACF) and partial correlation function (PCF) of the forecast residuals produced by the ISO and the re-forecast model for HAM are shown in Fig. 4. The residuals from the re-forecast are white and performance of the models shows improvement, which validates the assumption. One drawback of this analysis is that it does not provide any information about the time of the errors which is very important in load forecasts as the errors during peak times have higher implications on the stability of the grid than off-peak times. Thus, a temporal analysis of errors is discussed below.

5.2.2. Temporal analysis

For CAISO load demand, midnight to 5:00 am PDT in the morning is considered super off-peak time, 6:00 am to 12:00 noon and 8:00 pm to midnight PDT is considered off-peak time, and noon to 8:00 pm PDT is considered an on-peak time. The grid is very sensitive during on-peak times because of high demand. Fig. 5 shows the mean and standard deviation in HAM load forecast errors from CAISO and re-forecast model (NARX). This analysis shows that the CAISO forecast is the worst during super off-peak times. As the time-period changes to off-peak and on-peak times, the forecast improves significantly. Hence, the forecast error is smallest during on-peak times with the lowest standard deviation.

The performance of the re-forecast model is consistent throughout the day. The re-forecast model significantly improves the forecast in super-off peak and off-peak times. During on-peak times, the re-forecast shows improvement only for certain hours (6:00 pm to 8:00 pm PDT). The variance in the CAISO forecast is smaller than the re-forecast with almost the same mean value. It can be inferred that for on-peak times CAISO forecast is the best (12:00 noon to 6:00 pm PDT) compared to any re-forecasts whereas for all other times the re-forecast has potential to improve the forecast produced by the ISO. Similar results were found for ERCOT (see Fig. 5) and DAM load forecasts.

6. Conclusions

In this study we proposed a re-forecast methodology to enhance performance of STLF load prediction by extracting the

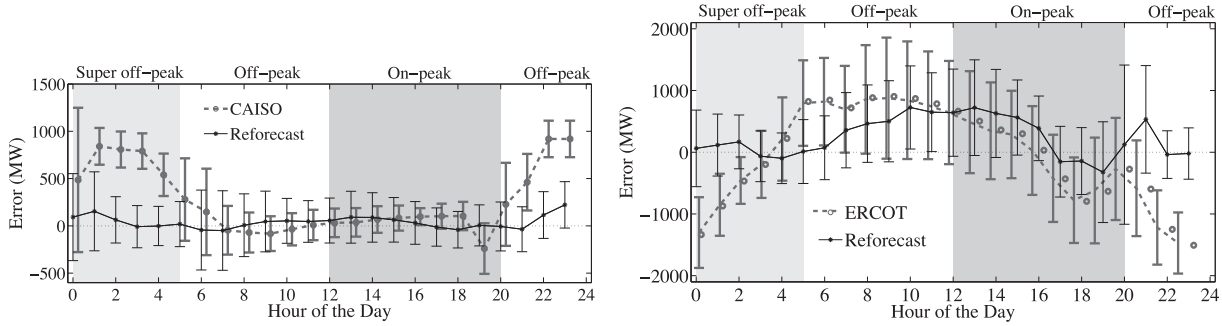

Fig. 5. Temporal distribution of the mean error (actual load-forecast HAM load) over the day for CAISO (top) and ERCOT (bottom). The error-bar shows the standard deviation in error for that hour of the day where 0 represents a perfect forecast with 0 MW error. It can be noticed that the accuracy of the CAISO model significantly improves during on-peak times than super off-peak and off-peak times whereas the performance of the re-forecast model (NARX) is consistent over the day. A similar pattern can be observed for ERCOT where re-forecast (LS-HE) shows improvement only for partial hours during on-peak times (5:00–9:00 pm) and significant improvement during super off-peak and off-peak times.

information from the structured non-random errors in the given estimate. The re-forecast ensembles consist of various time-series models combined using least squares optimization. Different ensemble combinations, with the weights optimized based on global, hourly, and weekday performance of the models, are proposed and investigated in this study. For real-time load prediction, the weights are updated at every time-step new data becomes available using recursive least squares. The results are presented for two types of load forecasts: HAM and DAM load for both CAISO and ERCOT. For HAM, point load forecast is produced at every hour. For DAM point forecasts are produced at midnight for all twenty-four hours of the next day. The re-forecast results showed an improvement over the internal forecast provided by the ISOs.

For HAM, NARX and LS-ME performed the best for CAISO. For ERCOT, LS-HE performed the best with a MAPE of 1.49% as compared to ERCOT forecast MAPE of 2.24%. RLS-ME improved forecast accuracy with a MAPE of 0.88% for CAISO and 1.70% for ERCOT. Similarly, for DAM load prediction, GM-E and LS-ME performed the best for both CAISO and ERCOT dataset. Therefore, we can conclude that the methods proposed in this paper can significantly enhance the accuracy of load prediction. This facilitates prospective improvements in all aspects of grid management and operations.

Correlation analysis of the forecast residuals from ISOs revealed structured non-random errors in their estimates. Application of re-forecast refined the forecast with white noise errors. Temporal analysis of the residuals showed that the internal forecasts by the ISOs have the highest error during off-peak times and their forecast accuracy significantly improves for on-peak times. The re-forecast refines the forecast with significant improvements during off-peak times and small improvements during on-peak times. Similar results from two ISOs with different load profiles and forecast horizons validates the robustness of our models. Hence, the proposed methodology is applicable to ISOs and power utilities, and has the ability to substantially enhance the accuracy of current load forecasts.

Acknowledgements

The authors gratefully acknowledge funding from the California Solar Initiative (CSI) Research, Development, Demonstration, and Deployment (RD&D) Program Grant III; and from the National Science Foundation (NSF) EECS-EPAS Award No. 1201986, which is managed by Dr. Paul Werbos. Partial support from the DOE SUNRISE project DE-EE0006330 is also gratefully acknowledged.

References

[1] WECC Standard BAL-STD-002-0 Operating Reserves; 2010.
[2] Ortega-Vazquez MA, Kirschen DS. Economic impact assessment of load forecast errors considering the cost of interruptions. In: Power engineering society general meeting 2006. IEEE; 2006.
[3] Suganthi L, Samuel AA. Energy models for demand forecasting a review. Renew Sust Energy Rev 2012;16(2):1223–40.
[4] Matthewman PD, Nicholosn H. Techniques for load prediction in the electricity-supply industry. Proc IEEE 1968;115(10):1451–7.
[5] Hippert HS, Pedreira CE, Souza RC. Neural networks for short-term load forecasting: a review and evaluation. IEEE Trans Power Syst 2001;16(1):44–55.
[6] CAISO - Today's Outlook; 2013. [cited July 2013].
[7] ERCOT forecast; 2013. [cited July 2013].
[8] Willis H, Northcotegreen J. Spatial electric-load forecasting-a tutorial reveiw. Proc IEEE 1983;71:232–53.
[9] Hagan MT, Behr SM. The time series approach to short term load forecasting. IEEE Trans Power Syst 1987;2(3):785–91.
[10] Moghram I, Rahman S. Analysis and evaluation of five short-term load forecasting techniques. IEEE Trans Power Syst 1989;4(4):1484–91.
[11] Metaxiotis K, Kagiannas A, Askounis D, Psarras J. Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. Energy Convers Manage 2003;44(9):1525–34.
[12] Tzafestas S, Tzafestas E. Computational intelligence techniques for short-term electric load forecasting. J Intell Robot Syst 2001;31(1/3):7–68.
[13] Alfares HK, Nazeeruddin M. Electric load forecasting: literature survey and classification of methods. Int J Syst Sci 2002;33(1):23–34.
[14] Adya M, Collopy F. How effective are neural networks at forecasting and prediction? a review and evaluation. J Forecast 1998;17:481–95.
[15] Crone SF, Hibon M, Nikolopoulos K. Advances in forecasting with neural networks? empirical evidence from the NN3 competition on time series prediction. Int J Forecast 2011;27(3):635–60.
[16] Hahn H, Meyer-Nieberg S, Pickl S. Electric load forecasting methods: tools for decision making. Euro J Operat Res 2009;199(3):902–7.
[17] Inman RH, Pedro HTC, Coimbra CFM. Solar forecasting methods for renewable energy integration. Prog Energy Combust Sci 2013;39(6):535–76.
[18] Kaur A, Pedro HT, Coimbra CFM. Impact of onsite solar generation on system load demand forecast. Energy Convers Manage 2013;75:701–9.
[19] Marquez R, Pedro HTC, Coimbra CFM. Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs. Solar Energy 2013;92(0):176–88.
[20] Marquez R, Coimbra CFM. Intra-hour DNI forecasting based on cloud tracking image analysis. Solar Energy 2013;91:327–36.
[21] Pedro HTC, Coimbra CFM. Assessment of forecasting techniques for solar power production with no exogenous inputs. Solar Energy 2012;86(7):2017–28.
[22] Marquez R, Gueorguiev VG, Coimbra CFM. Forecasting of global horizontal irradiance using sky cover indices. J Solar Energy Eng 2013;135(1):0110171–75.
[23] Drezga I, Rahman S. Short-term load forecasting with local ANN predictors. IEEE Trans Power Syst 1999;14(3):844–50.
[24] Taylor JW, Buizza R. Neural network load forecasting with weather ensemble predictions. IEEE Trans Power Syst 2002;17(3):626–32.
[25] Taylor JW, Buizza R. Using weather ensemble predictions in electricity demand forecasting. Int J Forecast 2003;19(1):57–70.
[26] Abdel-Aal RE. Improving electric load forecasts using network committees. Electric Power Syst Res 2005;74(1):83–94.
[27] Fan S, Chen L, Lee W-J. Short-term load forecasting using comprehensive combination based on multimeteorological information. IEEE Trans Ind Appl 2009;45(4):1460–6.
[28] Alamaniotis M, Ikonomopoulos A, Tsoukalas LH. Evolutionary multiobjective optimization of kernel-based very-short-term load forecasting. IEEE Trans Power Syst 2012;27(3):1477–84.
[29] Siwek K, Osowski S, Szupiluk R. Ensemble neural network approach for accurate load forecasting in a power system. Int J Appl Math Comput Sci 2009;19(2):303–15.
[30] Li Y, Wang D-F, Han P. Selective ensemble using discrete differential evolution algorithm for short-term load forecasting. In: 2009 International conference on machine learning and cybernetics, vol. 1; 2009. p. 80 –5.
[31] Matijas M, Suykens JAK, Krajcar S. Load forecasting using a multivariate meta-learning system. Expert Syst Appl 2013;40(11):4427–37.
[32] Ansarimehr P, Barghinia S, Habibi H, Vafadar N. Short term load forecasting for Iran national power system using artificial neural network and fuzzy expert system. In: Proceedings. PowerCon 2002. International conference on power system technology, 2002, vol. 2; 2002. p. 1082–5.
[33] Liao G-C, Tsao T-P. Application of a fuzzy neural network combined with a chaos genetic algorithm and simulated annealing to short-term load forecasting. IEEE Trans Evolut Comput 2006;10(3):330–40.
[34] Ljung L. System identification: theory for the user. New Jersey (USA): Prentice Hall PTR; 1999. No. ISBN 0-13-881640-9 025.
[35] Passino KM. Biomimicry for optimization, control, and automation. London (UK): Springer-Verlag; 2005.
[36] CAISO; 2013. <http://www.caiso.com> [cited July 2013].
[37] FERC; 2013. <http://www.ferc.gov/market-oversight/mkt-electric/california.asp> [cited July 2013].
[38] ERCOT; 2013. <http://www.ercot.com> [cited July 2013].