Short-Term Residential Electric Load Forecasting: A Compressive Spatio-Temporal Approach

Akin Tascikaraoglu[a,*], Borhan M. Sanandaji[b]

[a] *Department of Electrical Engineering, Yildiz Technical University, Istanbul, Turkey.*

[b] *Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA*

Abstract

Load forecasting is an essential step in power systems operations with important technical and economical impacts. Forecasting can be done both at aggregated and stand-alone levels. While forecasting at the aggregated level is a relatively easier task due to smoother load profiles, residential forecasting (stand-alone level) is a more challenging task due to existing diurnal, weekly, and annual cycles effects in the corresponding time series data and fluctuations caused by the random usage of appliances by end-users. Exploring the available historical load data, it has been discovered that there usually exists an interesting trend between the data from a target house and the data from its surrounding houses. This trend can be exploited for improving the forecast accuracy. One can define several different features for each house, including house size, occupancy level, and usage behavior of appliances. While the number of such features can be large, the main challenge is how to determine the best candidates (features) for an input set without increasing the forecasting computational costs. With this objective in mind, we present a forecasting approach which combines ideas from Compressive Sensing (CS) and data decomposition. The idea is to provide a framework which facilitates exploiting the existing low-dimensional structures governing the interactions among residential houses. The effectiveness of the proposed algorithm is evaluated using real data collected from residential houses in Texas, USA. The comparisons against benchmark methods show that the

*Corresponding author. Tel.: +90 212 3835866; fax: +90 212 3835858.
E-mail address: atasci@yildiz.edu.tr, akintasci@gmail.com (A. Tascikaraoglu)

Preprint submitted to Energy and Buildings *November 26, 2015*

proposed approach significantly improves the short-term forecasts.

Keywords: Electric load forecasting; Compressive Sensing; Spatial correlation; Residential buildings; Data decomposition.

1. Introduction

1.1. Demand Forecasting Methods

Electricity generation in conventional power systems is highly concentrated on large production units. This centralized power generation and the resulting unidirectional power flow hinder the inclusion of advanced technologies and services such as renewable energy integration and demand response [1, 2]. In order to overcome these challenges with a reliable and cost-effective method, smart grid, a modernized network based on information and communication technologies as well as automated control, has been introduced in the last decades.

Among the various components required to realize a smart grid, demand forecasting is of great importance since these forecasts enable the adjustment of generation and consumption while minimizing the operating cost, decreasing reserve capacity while maintaining the system security and removing the requirement of expensive energy storage systems. Demand forecasting is defined as the expected load information in a look-ahead time period predefined depending on the requirement in power system operation. These approaches can be classified in various ways: (i) very short-term (minutes ahead), short-term (hours or days ahead), middle-term (months ahead) and long-term (years ahead) approaches in terms of prediction horizon length, (ii) electrical appliance-level, house-level, distribution transformer-level, region-level and country-level approaches in terms of the range of the area at which forecasts are carried out and aggregated, (iii) point and probabilistic approaches in terms of output type, (iv) parametric, nonparametric and Artificial Intelligence (AI)-based approaches in terms of model structure and, (v) temporal, spatial and spatio-temporal approaches in terms of input data type.

In the literature, there has been a large number of studies which investigate the performance of many approaches on the demand forecasting of large areas, such as time series models [3–5], Neural Network (NN) [6, 7] and Extreme Learning Machine (ELM) [8]. Furthermore, some papers consider the forecast of large non-residential buildings [9–11]. These studies generally achieve reasonable results thanks to the relatively regular load profiles, resulted by the aggregation of a large number of loads. However, a limited number of studies regarding demand forecasting in residential units has so far been presented due especially two reasons: (i) more stochastic and more dynamic load curves caused by the random behaviors of consumers, seasonal factors, climatic conditions, etc., which make the forecasts more difficult, and (ii) lack of data availability for these units.

Jetcheva et al. [12] introduce an NN-based model for day-ahead demand forecasting of residential houses. Another building-level method based on NN is presented for load prediction of micro-grids [13]. Based on context information and its daily schedule patterns, an approach is proposed for modeling a household load [14]. Edwards et al. [15] investigate the performances of various machine learning techniques for the electrical consumption forecast in a residential level. The main objective of such studies is to increase the prediction accuracy while generally using a great deal of data and employing various approaches together. Any increase in the demand prediction accuracy could be considered to be very valuable for the power companies since over-predictions cause increments in the start-up and fixed costs of power generating units and under-predictions lead to purchases of expensive peaking power.

1.2. Motivation Behind Spatio-Temporal Forecasting Approaches

This paper aims at focusing on residential end-users and developing a load forecasting approach which exploits a large data set to improve the forecasting performance. There exist several factors, such as outdoor temperature values, humidity and the social activities specific to some time periods (e.g., holiday activities, public holiday activities, etc.) that affect the energy consumption of a target house. The proposed forecasting method incorporates all such data recorded from all houses in the vicinity of the target house. In fact, these houses face similar conditions with the target house; however, many of them exhibit a different energy consumption profile. In addition to the data taken from the houses with the similar consumption pattern, which significantly contributes to the forecasting performance, the data from slightly correlated houses also provides useful information to the forecasting model. For example, if one conducts an energy forecast using the data from the last month (i.e., the training data of one month), it is likely to obtain the forecasts between the minimum and the maximum values of the training set. However, the weather might be colder or warmer in the given prediction horizon and hence the demand can be significantly different from the values seen in the training set because of the working condition of the corresponding Air Conditioner (AC). In case of using spatial data in forecasting (i.e., data from surrounding houses), more accurate forecasts can be realized since in this case the training data set also contains values beyond the extremum values of the target house, resulted by different AC energy usage patterns in different houses.

In fact, using spatial information for demand forecasting has recently gained so some attention. Melo et al. [16] present an approach which takes the distribution of the load growth in a city and the relationships among different areas into account. For a similar forecasting structure, Carreno et al. [17] propose Cellular Automata (CA) for spatio-temporal allocation of new loads. In order to investigate the spatial correlation of demand changes and to exploit this information in forecasting, Wu and Lu [18] employ a data mining technique. Li et al. [19] study on a Radial Basis Function (RBF) NN forecast method, also using a spatial correlation analysis for inputs.

1.3. Contributions

It is evident that incorporating spatial information can help improve the forecast accuracy and thus the trend has been towards exploiting all of the available data in forecasting. The question here is how to determine the best candidate houses that would contribute to the forecasting performance without increasing computational burden. With this objective, a preliminary analysis can be carried out with some techniques presented to filter the redundant data; however, these methods generally provide good results for certain conditions. In other words, new analyses are required for different data sets and prediction horizons. To this end, we propose a method that is based on a claim that there exists a sparse relational pattern among the correlations between time series of surrounding houses and exploiting such low-dimensional structures considerably improves the forecasting performance. Consequently, the forecasting problem in this study is casted as recovery of a block-sparse coefficient vector from a set of linear equations.

Data decomposition has also gained some recent attention in demand forecasting [7, 8, 13] where the original load data is decomposed into various subsets in order to perform higher-precision forecasts with more regular subseries. The forecasting literature [20] shows that any decomposition method would improve the forecasting performance; however, the methods which are widely used for this purpose might create some challenges. The first problem is that, contrary to the wind and solar time series, every increase and decrease in the load time series has an obvious reason and data decomposition carried out with well-known methods such as Wavelet Transform (WT) and Fast Fourier Transform (FFT) would cause to obtain various subsets in which the load changes generally do not make any sense. Therefore, it would be hard to make a comment about the forecasts of the subsets and about how to improve the performance. In other words, without knowing the reason of cycles and high ramps seen in the load profiles, the contribution of the data decomposition, particularly for the spatial data, would be limited. The second reason is that a large number of subsets are required to derive the daily and weekly cycles embedded in load profile, which results in a great deal of computational costs. The difficulties in choosing the suitable structures and parameters for the data decomposition methods, such as mother wavelet and decomposition level numbers in WT, can be pointed out as another disadvantage of these methods. Considering the challenges mentioned, instead of using a conventional decomposition method, we first divide the energy consumption of the houses into three subsets before forecasting process: (i) AC load, which constitutes the largest fraction of the total load [21], (ii) Electric Vehicle (EV) load, which generally causes peaks in demand and, (iii) the remaining load consisting of the consumption of other appliances. The decomposed subseries are then forecasted individually using different model configurations, followed by an aggregation process. The performance of the proposed *spatio-temporal* algorithm is evaluated using real energy consumption data collected by the Pecan Street Research Consortium from 173 houses in Texas for the year 2014, where the data are provided for the appliances using in-home measurement equipment.

1.4. Paper Organization

The paper is organized as follows: In Section 2, the load forecasting problem is formulated. The proposed approach is presented in Section 3 and performance comparison of the proposed approach with the benchmark methods is provided in Section 4. Section 5 presents some additional features of the proposed approach. Section 6 summarizes our concluding remarks.

2. Multivariate Autoregressive (M-AR) Model

Autoregressive (AR) approaches have been used effectively in the forecasting literature, assuming that the weighted linear combination of past values of a system can be used for estimation of its output variable. In order to apply this method to multivariate time series, a generalized model, namely Multivariate Autoregressive (M-AR) model, is constructed as follows:

$$\begin{aligned}\boldsymbol{y}(t) &= X_1\boldsymbol{y}(t-1) + \cdots + X_n\boldsymbol{y}(t-n) + \boldsymbol{e}(t) \\ &= \sum_{j=1}^{n} X_j\boldsymbol{y}(t-j) + \boldsymbol{e}(t),\end{aligned} \tag{1}$$

where $\boldsymbol{y}(t) \in \mathbb{R}^P$ is a P-dimensional output at time t, $X_j \in \mathbb{R}^{P\times P}$ is a coefficient matrix for the j-th time lag, n is the model order and $\boldsymbol{e}(t)$ is the noise component. For load data measured at P houses ($p = 1, 2, \ldots, P$), define y_t^p as the measured load value at the p-th house at time t ($t = 1, 2, \ldots, M + n$) and let y^{p^*} be the target variable. Let $N := nP$. The M-AR model given in (1) can be modified in a format as in (2-4). We have

$$\begin{aligned}\boldsymbol{b} &= T_{Mn}^1\boldsymbol{x}^1 + T_{Mn}^2\boldsymbol{x}^2 + \cdots + T_{Mn}^P\boldsymbol{x}^P \\ &= \sum_{i=1}^{P} T_{Mn}^i\boldsymbol{x}^i,\end{aligned} \tag{2}$$

where

$$\boldsymbol{b} = \begin{bmatrix} y_{n+1}^{p^\star} \\ y_{n+2}^{p^\star} \\ \vdots \\ y_{n+M}^{p^\star} \end{bmatrix} \tag{3}$$

and T_{Mn}^p is an $M \times n$ Toeplitz matrix given by

$$T_{Mn}^p = \begin{bmatrix} y_n^p & \cdots & y_1^p \\ y_{n+1}^p & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ y_{n+M-1}^p & \cdots & y_M^p \end{bmatrix}. \tag{4}$$

The summation given in (2) can be compactly written as $\boldsymbol{b} = A\boldsymbol{x}$ where

$$\boldsymbol{x} = \begin{bmatrix} x_1^1 \\ \vdots \\ x_n^1 \\ \hline \vdots \\ \vdots \\ \hline x_1^P \\ \vdots \\ x_n^P \end{bmatrix} \tag{5}$$

and

$$A = \left[\; T_{Mn}^1 \;\middle|\; \dots \;\middle|\; T_{Mn}^P \;\right] \in \mathbb{R}^{M \times N}. \tag{6}$$

The objective in the training phase is to determine a vector $\boldsymbol{x} \in \mathbb{R}^N$ that best explains the observations $\boldsymbol{b} \in \mathbb{R}^M$ and $A \in \mathbb{R}^{M \times N}$. As can be seen from (5), $\boldsymbol{x}$ has a block structure and the coefficients of each unique variable appear in one vector-block.

3. Compressive Spatio-Temporal Load Forecasting (CST-LF)

In this study, it is assumed that only a few houses have a high correlation with the target house in a large collection of houses, and thereby only these houses contribute to the output of the target house. There will be a distinct structure to the solution $\boldsymbol{x}$ under the assumption of *sparsity* of the interconnections. According to model assumptions adopted, a coefficient vector $\boldsymbol{x}$, called *block-sparse*, will have very few non-zero entries clustered in few locations. The block number corresponds to the number of links that contribute to the output of the target house. For a given target house, we solve the following minimization problem:

$$\min_{\boldsymbol{x}} \|\boldsymbol{b} - A\boldsymbol{x}\|_2 \quad \text{subject to} \quad (\boldsymbol{x} \text{ is block-sparse}). \tag{7}$$

3.1. Uniform Compressive Spatio-Temporal Load Forecasting (CST-LF)

We use tools from Compressive Sensing (CS) for recovery of a block-sparse $\boldsymbol{x}$ (Please see [22] and [23] for further information about CS recovery problem).

Algorithm 1 The BOMP algorithm for block-sparse recovery

Require: A, $\boldsymbol{b}$, $\{n_i\}_{i=1}^{P}$, stopping criteria
Ensure: $\boldsymbol{r}^0 = \boldsymbol{b}$, $\boldsymbol{x}^0 = \boldsymbol{0}$, $\Lambda^0 = \emptyset$, $l = 0$
repeat
  1. match: $\boldsymbol{h}_i = A_i^{\text{tr}} \boldsymbol{r}^l, \quad i = 1, 2, \cdots, P$
  2. identify support: $\lambda = \arg\max_i \|\boldsymbol{h}_i\|_2$
  3. update the support: $\Lambda^{l+1} = \Lambda^l \cup \lambda$
  4. update signal estimate:
   $\boldsymbol{x}^{l+1} = \arg\min_{\boldsymbol{s}:\text{supp}(\boldsymbol{s}) \subseteq \Lambda^{l+1}} \|\boldsymbol{b} - A\boldsymbol{s}\|_2$,
   where supp($\boldsymbol{s}$) indicates the blocks
   on which $\boldsymbol{s}$ may be non-zero
  5. update residual estimate: $\boldsymbol{r}^{l+1} = \boldsymbol{b} - A\boldsymbol{x}^{l+1}$
  6. increase index l **by** 1
until stopping criteria true
output: $\widehat{\boldsymbol{x}} = \boldsymbol{x}^l = \arg\min_{\boldsymbol{s}:\text{supp}(\boldsymbol{s}) \subseteq \Lambda^{l}} \|\boldsymbol{b} - A\boldsymbol{s}\|_2$

Definition 1 (Block K-Sparse Signal). *Let $\boldsymbol{x} \in \mathbb{R}^N$ be a concatenation of P vector-blocks $\boldsymbol{x}_i \in \mathbb{R}^n$, i.e.,*

$$\boldsymbol{x} = [\boldsymbol{x}_1^{tr} \cdots \boldsymbol{x}_i^{tr} \cdots \boldsymbol{x}_P^{tr}]^{tr}, \tag{8}$$

where $N = nP$. A signal $\boldsymbol{x} \in \mathbb{R}^N$ is called block K-sparse *if it has $K < P$ non-zero blocks.* □

The Block Orthogonal Matching Pursuit (BOMP) algorithm, among many other extensions of the CS recovery algorithms for block sparse recovery, is used in this paper [24]. To find a block-sparse solution to the equation $\boldsymbol{b} = A\boldsymbol{x}$, the steps of BOMP are given in Algorithm 1, where $A \in \mathbb{R}^{M \times N}$ is a concatenation of P matrix-blocks $A_i \in \mathbb{R}^{M \times n}$ $(\forall n_i = n)$ as

$$A = [A_1 \cdots A_i \cdots A_P]. \tag{9}$$

3.2. Nonuniform CST-LF

In a uniform CST-LF, we assume that the interactions between houses are governed by a uniform M-AR model (AR models with the same order values for all interconnections) as given in (2). In this section, we consider a more generalized CST-LF algorithm where the houses are related by AR models of different orders. This model structure distinguishes between the houses based on their cross-correlations with the target house. Let n_i be the order associated with the i-th house for $i = 1, 2, \dots, P$. A Nonuniform Multivariate Autoregressive (NM-AR) version of (2) can be defined as shown in (10-13), where $n_{\max} \geq \max_i n_i$ and $N := \sum_{i=1}^{P} n_i$. We have

$$\begin{aligned} \boldsymbol{b} &= T^1_{Mn_1}\boldsymbol{x}^1 + T^2_{Mn_2}\boldsymbol{x}^2 + \cdots + T^P_{Mn_P}\boldsymbol{x}^P \\ &= \sum_{i=1}^{P} T^i_{Mn_i}\boldsymbol{x}^i, \end{aligned} \tag{10}$$

where

$$\boldsymbol{b} = \begin{bmatrix} y^{p^\star}_{n_{\max}+1} \\ y^{p^\star}_{n_{\max}+2} \\ \vdots \\ y^{p^\star}_{n_{\max}+M} \end{bmatrix} \quad \text{and} \quad T^p_{Mn_p} = \begin{bmatrix} y^p_{n_{\max}} & \cdots & y^p_{n_{\max}-n_p+1} \\ y^p_{n_{\max}+1} & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ y^p_{n_{\max}+M-1} & \cdots & y^p_{n_{\max}-n_p+M} \end{bmatrix}. \tag{11}$$

The summation given in (10) can be compactly written as $\boldsymbol{b} = A\boldsymbol{x}$ where

$$A = \left[\ T^1_{Mn_1} \,\middle|\, \ldots \,\middle|\, T^P_{Mn_P}\ \right] \tag{12}$$

and

$$\boldsymbol{x} = \left[\begin{array}{c} x^1_1 \\ \vdots \\ \vdots \\ \vdots \\ x^1_{n_1} \\ \hline \vdots \\ \hline x^P_1 \\ \vdots \\ x^P_{n_P} \end{array}\right] \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \\ \\ \end{array}\right\}\text{Block 1} \\ \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\}\text{Block } P \end{array} \tag{13}$$

This model structure results in a nonuniform block-sparse coefficient vector $\boldsymbol{x}$ with different block lengths.

Definition 2 (Nonuniform Block K-Sparse Signal). *Let $\boldsymbol{x} \in \mathbb{R}^N$ be a concatenation of P vector-blocks $\boldsymbol{x}_i \in \mathbb{R}^{n_i}$ where $N = \sum_{i=1}^{P} n_i$. A signal $\boldsymbol{x} \in \mathbb{R}^N$ is called* nonuniform block K-sparse *if it has $K < P$ non-zero blocks.* □

Given $\{n_i\}_{i=1}^{P}$, the BOMP algorithm can be used for recovery of $\boldsymbol{x}$ with $A_i \in \mathbb{R}^{M \times n_i}$. In order to find the set of order values $\{n_i\}_{i=1}^{P}$, we use a correlation analysis and adjust the orders to accomplish the best forecasting performance.

4. Case Study of 173 houses in Austin, Texas

In the simulation studies, we use a high-quality hourly real load data provided by a global collaboration, Pecan Street Research Consortium [25], for 173 houses in Austin, Texas.

4.1. Data Description

Pecan Street data provides high-resolution individual measurements collected from various electrical appliances. Among the available data, hourly electricity meter measurements from year 2014 including the total usage and consumption of AC and EV are incorporated in the case studies. In order to study on a relatively hard forecasting task and to highlight the contribution of the proposed models, first, some plausible candidate houses, which have: (i) at least one AC, (ii) at least one EV, (iii) a reasonable consumption data during all days of a week (i.e., not occupied in certain days such as weekdays only) and (iv) a profile with drastic changes, are determined. Subsequently, two houses, which have the highest mean hourly power consumption (3.6020 kW and 2.1640 kW for the first and second houses, respectively) over the test period are chosen as the target houses among the candidate house set. Other main descriptive statistics, such as median and quartiles, during the considered test period are illustrated in Fig. 1, where the middle horizontal line in the box shows the median and the box spans the first (below the median) and third (above the median) quartiles. Moreover, the smallest and highest measurements are marked with the bottom and top horizontal lines extended out from the box. As can be seen from Fig. 1, the maximum power values of the selected houses are considerably high and also farther above the median, both of which imply the existence of high ramps in the load profiles.

In the case studies, a time period from June 1, 2014 to August 31, 2014 (a period of 92 days) is considered as a test set for all houses due to: (i) the lack of data availability for a large collection of houses for the whole year and (ii) the fact that this period contains the highest and the most variable power consumption data because of intensive AC usage in summer months in the considered area. Also, a period of 14 days before the first day of test set is used as an initial training set. It is expected that high correlations between the consumption values, particularly between the AC data, would help improve the prediction accuracy. Using BOMP, houses with a high correlation with the target houses are chosen and the ones with a very different load pattern are omitted in the forecasting process by assigning them a high or low coefficient proportional to their correlations. As an example, Fig. 2 shows the electricity consumption of one of the target houses as well as of two different houses in the first week of the test data set. Regarding the similarity of the load consumption profiles between the given houses, those shown in Fig. 2(a) can be considered as houses that would probably have a high coefficient in the considered period while the houses shown in Fig. 2(b) are not likely to be assigned a considerable value.

The effectiveness of the proposed algorithm is investigated through different case studies (two different houses and two different prediction horizons (6 and 24 hours)). The simulations are carried out separately for each case study for the given test period; however, the comparison results are presented for a time period of one week only, which includes the highest weekly mean power consumption measured in year 2014, in order to observe the performance of forecasting methods clearly.

4.2. Comparison with Other Benchmark Algorithms for 6-h-ahead forecasts

In the first set of results, 6-h-ahead forecasts of CST-LF methods as well as various temporal and spatio-temporal benchmark methods are compared. For a temporal forecasting method, we first consider the well-known *persistence* method, which is reasonable for particularly short terms due to the relatively regular behavior of load over such short terms [26]. An AR model is also used as a temporal benchmark model, in which the forecasts are performed in a recursive manner with 6-hour updates. Moreover, a combined model, which consists of WT and Artificial Neural Network (ANN) methods, presented by the authors in a previous study [27], is included to the comparisons due to its high capability of modeling nonlinearities in the time series. The forecasting results of considered temporal methods are presented in Figs. 3 and 4 for the first and second target houses, respectively. As can be seen, the WT-ANN outperforms the other methods for both houses, by especially predicting the high values more accurately.

Also, two spatio-temporal approaches, an M-AR approach using Least Squares (LS) to calculate the model coefficients [28] and an ANN-based spatio-temporal approach [29], are employed in order to examine the contribution of spatial information on the forecast accuracy. Figs. 5 and 6 show the improvement of forecasts by incorporating the available spatial data.

After evaluating the performance of the various approaches on the two houses chosen for the case studies, we now apply the proposed CST-LF algorithms with the objective of improving the prediction accuracy without increasing the computational burden. With this goal in mind, first, a nonuniform M-AR model exploiting all the available data is presented, in which a new coefficient vector $\boldsymbol{x}$ is calculated every prediction time horizon and used in load forecasts. It is noted that the load forecasts at time $n + M + 1$ for the 173 houses $(\hat{y}^i_{n+M+1}, \forall i)$ are inserted in the A matrix for forecasting the hourly load at time $n + M + 2$ $(\hat{y}^i_{n+M+2}, \forall i)$ and so on. After the forecasting of last step, the elements of A matrix are replaced with the measurements collected during the prediction horizon, which helps the model exhibit the latest trend of load profile. This adaptive structure allows to deal with the drifting problem that is a result of input and output data changes over time. Then, the recursive process begins for another 6 time steps. Another important point here worth mentioning is that a normalization process is employed to the time series from all houses to limit the dominancy of larger load values over the smaller ones in calculating the coefficients. Figs. 7(a) and 7(b) show the results of the nonuniform M-AR model for the first and second target houses, respectively. As can be seen, the results of BOMP-based models are superior to the all benchmark approaches presented for both houses. Thus, it is also possible to state that BOMP algorithm is more effective compared to the LS method in exploiting such a large data.

Moreover, the nonuniform M-AR is combined with WT in order to further improve the accuracy by decomposing the load data into relatively more regular subsets and thus forecasting each subset separately with a higher accuracy. The results are depicted in Figs. 8(a) and 8(b). It is evident that the WT-based method achieves more-accurate forecasts compared to the nonuniform M-AR method. Finally, the nonuniform M-AR based on appliance-level forecasting is employed and the results of single and total use predictions are given in Figs. 9 and 10 for target houses 1 and 2, respectively. In particular, the forecasts of AC usage are generally aligned with the measured values, resulting a considerable improvement in total usage forecasts. With respect to the remaining load forecasts, it can be indicated that removing the consumption of dominant appliances increases the accuracy by facilitating the prediction task. However, contrary to the other two components, the forecasts of EV are mostly not good because of its highly stochastic and complex profile. The main reasons of these ineffective EV power forecasts can be listed as follows: (i) Some of the houses have no EV while some have more than one EV, which complicates the load profiles, (ii) the types of the EVs are mostly different for each house and each type has different specifications such as different charging and discharging times, power ratings and maximum allowable mileages, (iii) the consumers do not use their EVs everyday or on certain days in a week, and (iv) they mostly do not wait until the end of charging process, which leads to a large number of different power consumption values drawn from the electricity system. Nevertheless, the accuracy of the EV forecasts does not affect the total performance of the proposed model due to the relatively short consumption periods of EVs in one week.

In order to measure the effectiveness of the proposed methods, various performance metrics for all the methods given in this paper are compiled in Tables 1 and 2. All these negatively oriented error metrics can provide an idea about the accuracy of the models. For instance, a higher Mean Absolute Error (MAE) indicates a permanent deviation between the measured and predicted data, and a higher Root Mean Squared Error (RMSE) implies mostly large error values. However, considering Normalized Root Mean Squared Error (NRMSE) for comparison between the results from two different houses is more appropriate due to its independency on the range of measured data. Regarding the NRMSE shown in Tables 1 and 2, the proposed nonuniform CST-LF approaches evidently outperform both the temporal and spatio-temporal models and the least error metric is achieved by the nonuniform CST-LF based on appliance-level forecasting. As an example, the final model provides an average improvement of 10.39% and 8.85 % for two houses compared to the conventional LS method and spatio-temporal ANN method, respectively. It is to be noted that the MAE and RMSE metrics of target house 1 are worse than those of target house 2 due to the higher maximum value and variance of target house 1; however, the wider range of measured data of house 1 leads to obtain lower NRMSE values compared to those of house 2. Also, the results are consistent with the fact that spatial information from the surrounding houses significantly contribute to the accuracy of time series models.

4.3. Comparison with Other Benchmark Algorithms for 24-h-ahead forecasts

In order to investigate the performance of the proposed approach on longer horizons which are generally of great importance for power system operations such as unit commitment, scheduling and load dispatch, we also perform the simulations for daily predictions. Considering the inefficiency of the temporal benchmark methods discussed in this study (i.e., persistence, AR and WT-ANN models) over such a long horizon, only the results belonging to the spatio-temporal benchmark methods are depicted in Figs. 11 and 12 for 24-h-ahead forecasts. Then, similar to the case in the 6-h-ahead forecasts, the forecasts are realized with the nonuniform M-AR and nonuniform M-AR based on WT (Figs. 13 and 14). The improved results accomplished with exploiting spatial information with BOMP algorithm and with including a data decomposition method confirm the effectiveness of the employed methods also for longer prediction horizons. Lastly, Figs. 15 and 16 demonstrate the forecasted time series with nonuniform M-AR based on appliance-level forecasting technique. Considering the NRMSE measures again, it can be concluded that the proposed forecasting structure is the most effective approach in longer prediction horizons, which is mainly stemmed from the high-accuracy forecasts of AC. The consumption of AC load is mainly based on ambient temperature and occupancy level of house, and hence it is generally more stable and more predictable compared to the other electrical appliances. Therefore, there is mostly a small difference in accuracy between forecasting a shorter horizon and a longer horizon for ACs, which makes this approach more promising for such longer periods in residential houses.

The performance metrics are also calculated for the daily forecasts and provided in Tables 3 and 4. As expected, the resulted error metrics are consistent with those in 6-h-ahead forecasts and the nonuniform M-AR based on appliance-level forecasting is still superior to the all other considered approaches. With respect to comparing the performance of proposed approach on different house profiles for longer horizons, it can be concluded that the approach is more effective in the load time series having more variability and higher power values.

Additionally, in terms of analyzing the sparsity in the coefficient vectors, Fig. 17 illustrates the corresponding vector calculated at the end of the training stage for the nonuniform CST-LF model developed for daily forecasts of target house 2. In consistent with the assumption made at the model construction, only a few of the coefficients or blocks (17 out of 173 houses) in the model are non-zero, which results in a sparse $\boldsymbol{x}$. As we shift the forecast horizon, a slowly-varying sparse structure appears in other consecutive vectors and that validates our idea behind using the low-dimensional structures in spatio-temporal demand forecasting.

5. Discussion

In the literature of load forecasting, the historical data is generally classified based on the day of the week into two different groups; (i) weekdays (Monday to Friday) and (ii) weekend (Saturday and Sunday). The main reason behind this data separation is to benefit from the similarity of the days, particularly weekdays in which load profiles are more regular and almost the same for each day, in the forecasting training phase. Contrary to these widely-preferred methods which consider daily as well as hourly cycles in load profiles [12, 14, 30, 31], no data classification has been considered in our study due to the following reasons: (i) There is no significant difference between the consumers' power consumption behavior during weekdays and weekends in the considered test data, which is mainly caused by the permanent use of ACs regardless of the day of week, (ii) it will discard the time-consuming data preprocessing step, and (iii) the proposed algorithm has a high performance even in these cases.

It is to be noted that the proposed spatio-temporal model is particularly designed to benefit from the relatively easy predictability of AC loads due to their strong dependency on the weather parameters and the high inter-house correlations. Therefore, it can be indicated that one of the main reasons for the high-accuracy forecasts achieved in the case studies is the efficiency of the proposed method on the forecasts of AC power consumption, which is nearly half of the total consumption during most of the summer days. The proposed model can still provide reasonable results in case of demand profiles with less AC usage ratio; however, a better way to improve the forecasts is to do a detailed evaluation of the components of the total load and select a subset of those components to be forecasted individually using our proposed method. Any load component satisfying two criteria; (i) continuous and deterministic usage during a given period, and (ii) correlated profile with the surrounding houses, could be a strong candidate to be applied to the decompose-and-forecast process in order to improve forecast performance.

It is worth noting that another advantage of the proposed algorithms is the high performance of the models in case of small training data sets. As mentioned above, a period of 14 days is used in the training stage; however, for instance, a larger historical data set (a period of 30 days) should be applied to train the ANN models employed in the case studies due to their high data requirement for learning the dependencies between the output power values and the input data.

Besides, other features (apart from electricity) are usually included in various load forecasting approaches presented in the literature for the purpose of increasing accuracy. Our proposed algorithm can remove the requirement of such a feature selection process in forecasting. This is because the algorithm already assigns a higher value to the most informative variables in a continuously updated coefficient determination process. Therefore, all the available features, if there exist, could be applied to the proposed forecasting approach without concerning about a time-consuming data selection treatment.

6. Conclusion

Under the assumption that there usually exists an intrinsic low-dimensional structure governing the data recorded from a collection of residential houses and that using this structure in load forecasting can help improve the forecasting performance, a compressive load forecasting approach incorporating both temporal and spatial information is presented in the paper. The proposed method is called nonuniform CST-LF as it is inspired by CS and structured-sparse recovery algorithms and is testified against various benchmark models using real and high-quality data, showing that proposed approach considerably improves the forecasts for short terms. Moreover, we illustrated that by decomposing the load data into several components and then forecasting each of these components (usually belonging to the high-power appliances) individually, would help further increase the accuracy, particularly for extended prediction horizons.

Including some additional features such as outdoor temperature and humidity is considered as a future study. Incorporating extra features in forecasting is found to be effective in the literature for prediction horizons more than several hours. Moreover, incorporating temperature forecasts in the model is another research path due to the close relationship between ambient temperature and AC consumption, particularly in summer months.

As another future direction, our target is to demonstrate the effectiveness of the algorithm for longer prediction horizons using a much larger collection of houses located in a wider area and using data from previous years to take advantage of the high similarity of the energy consumption behaviors in the same periods over consecutive years. Yet another path for us is to consider more detailed information about the houses (e.g., data gathered from surveys), such as their exact locations and the number of floors, and more detailed information about the household, such as occupancy level, for refining the input set and enabling more detailed interpretations about the results.

7. Acknowledgments

This work is supported by The Scientific and Technological Research Council of Turkey under grant TUBITAK-2219.

References

[1] E. L. Cano, M. Groissböck, J. M. Moguerza, M. Stadler, A strategic optimization model for energy systems planning, Energy and Buildings 81 (2014) 416–423.

[2] F. Farzan, S. A. Vaghefi, K. Mahani, M. A. Jafari, J. Gong, Operational planning for multi-building portfolio in an uncertain energy market, Energy and Buildings.

[3] N. Ding, Y. Bésanger, F. Wurtz, Next-day mv/lv substation load forecaster using time series method, Electric Power Systems Research 119 (2015) 345–354.

[4] Y. Goude, R. Nedellec, N. Kong, Local short and middle term electricity load forecasting with semi-parametric additive models, Smart Grid, IEEE Transactions on 5 (1) (2014) 440–446.

[5] W.-J. Lee, J. Hong, A hybrid dynamic and fuzzy time series model for mid-term power load forecasting, International Journal of Electrical Power & Energy Systems 64 (2015) 1057–1062.

[6] A. Khwaja, M. Naeem, A. Anpalagan, A. Venetsanopoulos, B. Venkatesh, Improved short-term load forecasting using bagged neural networks, Electric Power Systems Research 125 (2015) 109–115.

[7] M. Ghayekhloo, M. Menhaj, M. Ghofrani, A hybrid short-term load forecasting with a new data preprocessing framework, Electric Power Systems Research 119 (2015) 138–148.

[8] S. Li, P. Wang, L. Goel, Short-term load forecasting by wavelet transform and evolutionary extreme learning machine, Electric Power Systems Research 122 (2015) 96–103.

[9] A. H. Neto, F. A. S. Fiorelli, Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption, Energy and Buildings 40 (12) (2008) 2169–2176.

[10] Q. Dai, T. Cai, S. Duan, F. Zhao, Stochastic modeling and forecasting of load demand for electric bus battery-swap station, Power Delivery, IEEE Transactions on 29 (4) (2014) 1909–1917.

[11] J. Massana, C. Pous, L. Burgas, J. Melendez, J. Colomer, Short-term load forecasting in a non-residential building contrasting models and attributes, Energy and Buildings 92 (0) (2015) 322 – 330.

[12] J. G. Jetcheva, M. Majidpour, W.-P. Chen, Neural network model ensembles for building-level electricity load forecasts, Energy and Buildings 84 (2014) 214–223.

[13] H. Chitsaz, H. Shaker, H. Zareipour, D. Wood, N. Amjady, Short-term electricity load forecasting of buildings in microgrids, Energy and Buildings 99 (2015) 50–60.

[14] Y.-H. Hsiao, Household electricity demand forecast based on context information and user daily schedule analysis from meter data, Industrial Informatics, IEEE Transactions on 11 (1) (2015) 33–43.

[15] R. E. Edwards, J. New, L. E. Parker, Predicting future hourly residential electrical consumption: A machine learning case study, Energy and Buildings 49 (2012) 591–603.

[16] J. D. Melo, E. M. Carreno, A. Padilha-Feltrin, Multi-agent simulation of urban social dynamics for spatial load forecasting, Power Systems, IEEE Transactions on 27 (4) (2012) 1870–1878.

[17] E. M. Carreno, R. M. Rocha, A. Padilha-Feltrin, A cellular automaton approach to spatial electric load forecasting, Power Systems, IEEE Transactions on 26 (2) (2011) 532–540.

[18] H.-C. Wu, C.-N. Lu, A data mining approach for spatial modeling in small area load forecast, Power Systems, IEEE Transactions on 17 (2) (2002) 516–521.

[19] W. Li, Z. Wei, G. Sun, Z. Wan, W. Miao, Multi-interval wind speed forecast model based on improved spatial correlation and rbf neural network, Electric Power Automation Equipment 6 (2009) 021.

[20] A. Tascikaraoglu, A. Boynuegri, M. Uzunoglu, A demand side management strategy based on forecasting of residential renewable sources: A smart home system in Turkey, Energy and Buildings 80 (2014) 309–320.

[21] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy and buildings 40 (3) (2008) 394–398.

[22] B. M. Sanandaji, A. Tascikaraoglu, K. Poolla, P. Varaiya, Low-dimensional models in spaito-temporal wind speed forecasting, in Proceedings of the 2015 American Control Conference – ACC accepted.

[23] E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Transactions on Information Theory 52 (2) (2006) 489–509.

[24] B. M. Sanandaji, T. L. Vincent, M. B. Wakin, A review on sufficient conditions for structure identification of interconnected systems, Proceedings of the 16th IFAC Symposium on System Identification (2012) 1623–1628.

[25] Pecan Street Research Consortium, Dataport.
URL `http://dataport.pecanstreet.org`

[26] A. Tascikaraoglu, M. Uzunoglu, A review of combined approaches for prediction of short-term wind speed and power, Renewable and Sustainable Energy Reviews 34 (2014) 243–254.

[27] A. Tascikaraoglu, M. Uzunoglu, B. Vural, The assessment of the contribution of short-term wind power predictions to the efficiency of stand-alone hybrid systems, Applied Energy 94 (2012) 156–165.

[28] J. Tastu, P. Pinson, E. Kotwa, H. Madsen, H. A. Nielsen, Spatio-temporal analysis and modeling of short-term wind power forecast errors, Wind Energy 14 (1) (2011) 43–60.

[29] M. Bilgili, B. Sahin, A. Yasar, Application of artificial neural networks for the wind speed prediction of target station using reference stations data, Renewable Energy 32 (14) (2007) 2350–2360.

[30] S. Karatasou, M. Santamouris, V. Geros, Modeling and predicting building's energy use with artificial neural networks: Methods and results, Energy and Buildings 38 (8) (2006) 949–958.

[31] G. Escrivá-Escrivá, C. Álvarez-Bel, C. Roldán-Blay, M. Alcázar-Ortega, New artificial neural network prediction method for electrical consumption forecasting based on building end-uses, Energy and Buildings 43 (11) (2011) 3112–3119.

Table 1: Statistical Error Measure Comparison for 6-h-ahead forecasts - Target House 1

Forecasting approach	MAE (m/s)	RMSE (m/s)	NRMSE (%)
Persistence	1.4509	1.9739	18.83
AR	1.2308	1.6243	15.50
WT-ANN	1.0203	1.3855	13.22
LS-based ST	1.0199	1.3156	12.55
ANN-based ST	1.0064	1.2904	12.31
NM-AR	0.9892	1.2614	12.04
WT-NM-AR	0.9860	1.2487	11.91
Proposed Approach	0.9856	1.2453	11.88

Table 2: Statistical Error Measure Comparison for 6-h-ahead forecasts - Target House 2

Forecasting approach	MAE (m/s)	RMSE (m/s)	NRMSE (%)
Persistence	1.1203	1.6872	19.21
AR	1.0428	1.6374	18.64
WT-ANN	1.0331	1.4455	16.46
LS-based ST	1.0016	1.4106	16.06
ANN-based ST	0.9944	1.3910	15.83
NM-AR	0.9155	1.3166	14.99
WT-NM-AR	0.8445	1.2234	13.93
Proposed Approach	0.8287	1.1929	13.58

Table 3: Statistical Error Measure Comparison for 24-h-ahead forecasts - Target House 1

Forecasting approach	MAE (m/s)	RMSE (m/s)	NRMSE (%)
Persistence	1.8688	2.3953	22.85
AR	1.5944	1.8880	18.01
WT-ANN	1.3882	1.7199	16.41
LS-based ST	1.3067	1.6255	15.51
ANN-based ST	1.2223	1.5842	15.11
NM-AR	1.1675	1.4387	13.72
WT-NM-AR	1.1217	1.3742	13.11
Proposed Approach	1.1196	1.3549	12.92

Table 4: Statistical Error Measure Comparison for 24-h-ahead forecasts - Target House 2

Forecasting approach	MAE (m/s)	RMSE (m/s)	NRMSE (%)
Persistence	1.4300	1.9263	21.93
AR	1.1765	1.7049	19.42
WT-ANN	1.0976	1.5499	17.65
LS-based ST	1.0465	1.5109	17.21
ANN-based ST	1.0444	1.5022	17.10
NM-AR	0.9603	1.4199	16.17
WT-NM-AR	0.8553	1.2651	14.41
Proposed Approach	0.8462	1.2349	14.06

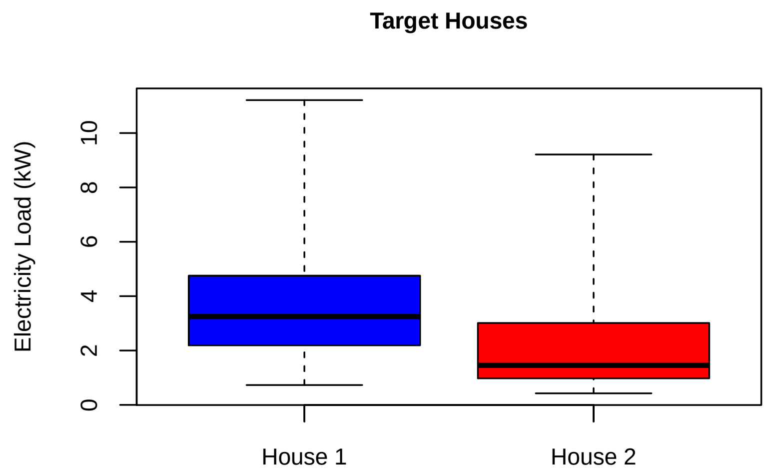
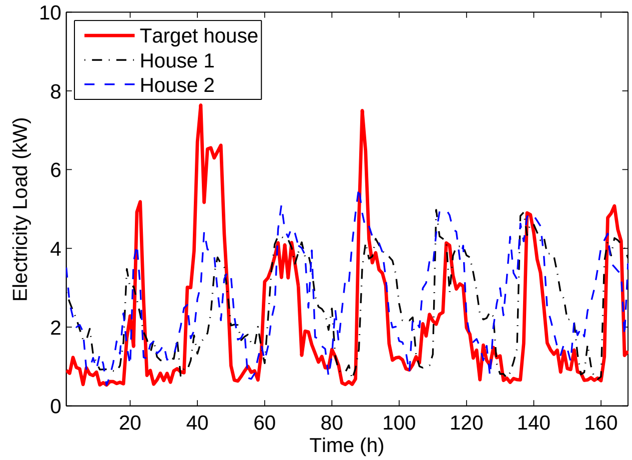
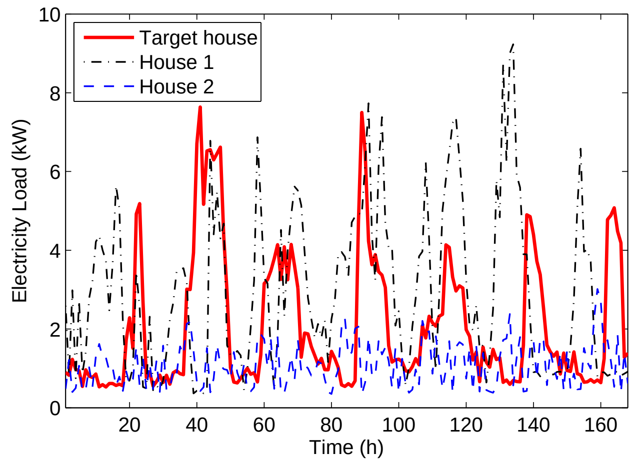


Figure 1: Descriptive statistics of the selected houses for the test period.

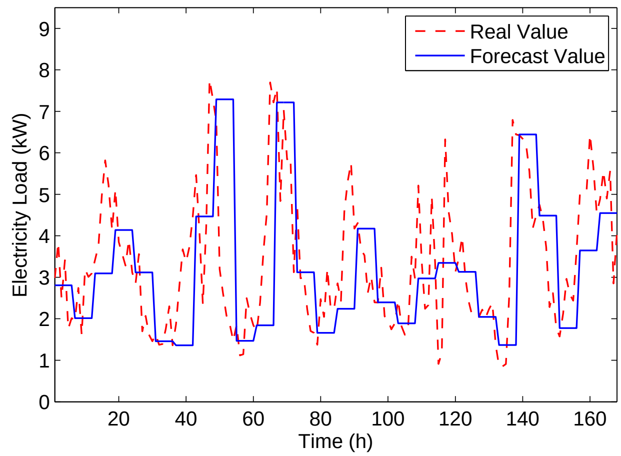


(a) Correlated houses

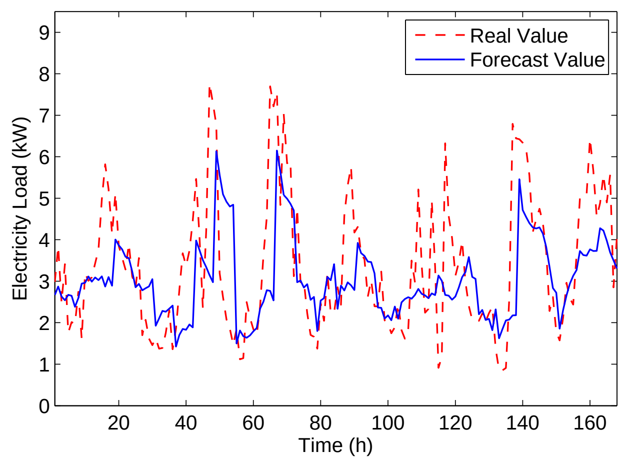


(b) Noncorrelated houses

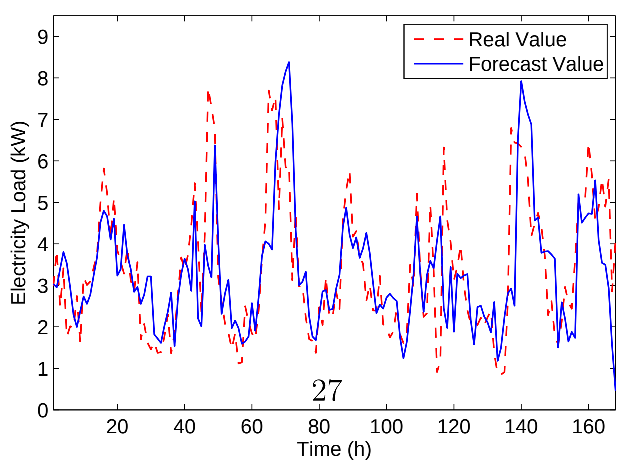
Figure 2: Consumption data comparison of different houses with target house over one week period.



(a) Persistence

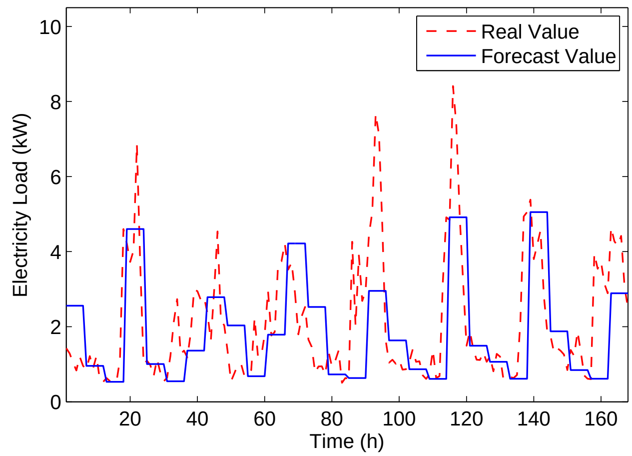


(b) AR

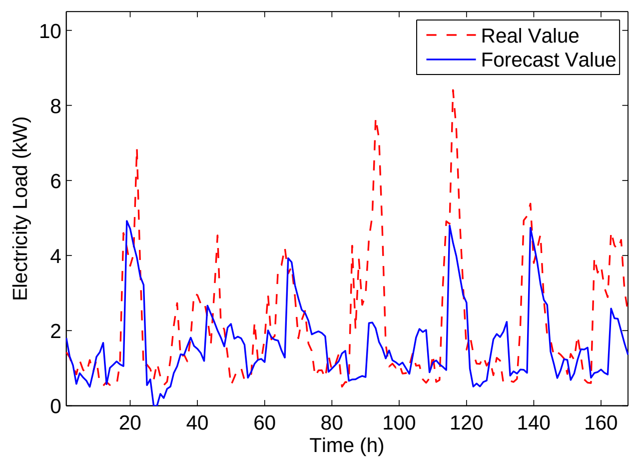


(c) WT-ANN

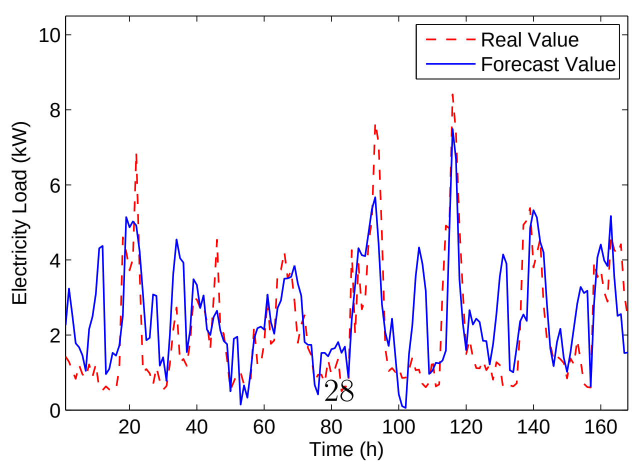
Figure 3: Comparison of temporal benchmark algorithms for 6-h-ahead forecasting - Target house 1.



(a) Persistence

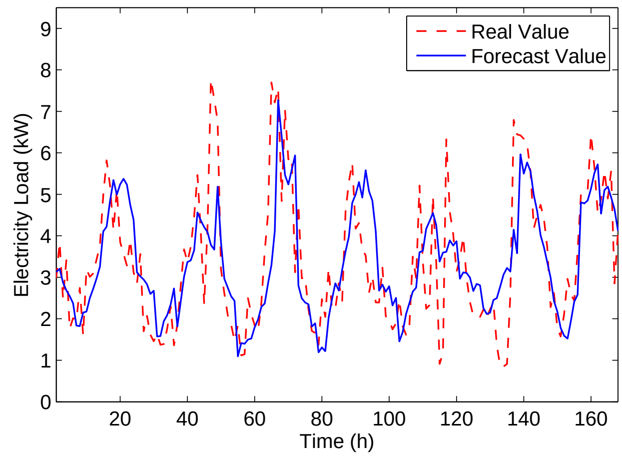


(b) AR

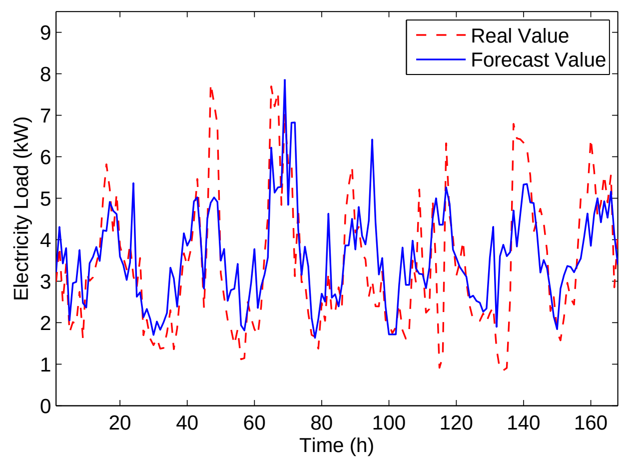


(c) WT-ANN

Figure 4: Comparison of temporal benchmark algorithms for 6-h-ahead forecasting - Target house 2.

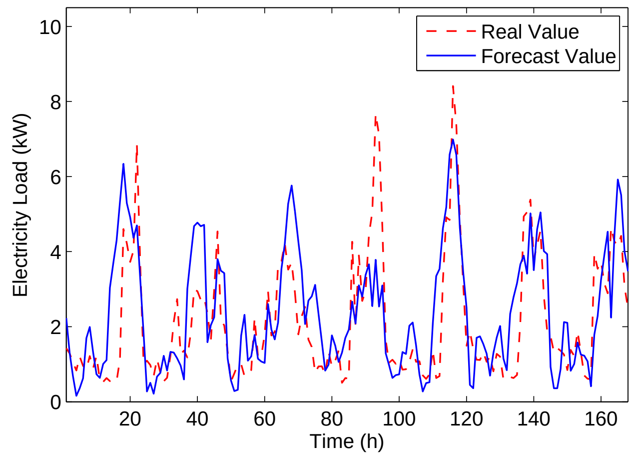


(a) LS-based ST

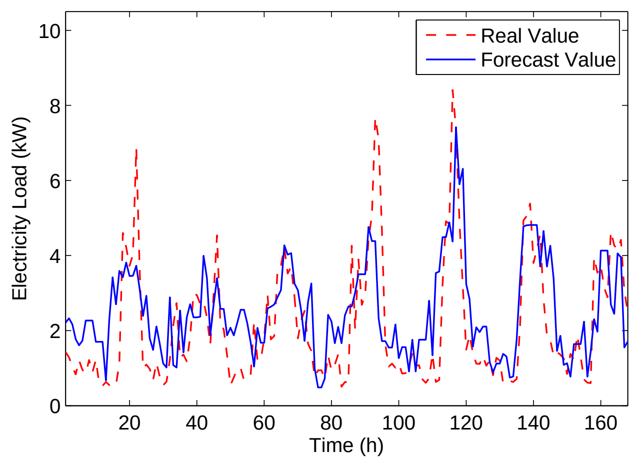


(b) ANN-based ST

Figure 5: Comparison of spatio-temporal benchmark algorithms for 6-h-ahead forecasting - Target house 1.

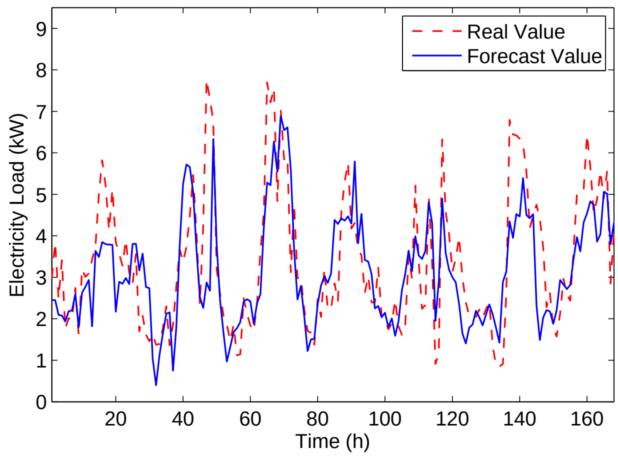


(a) LS-based ST

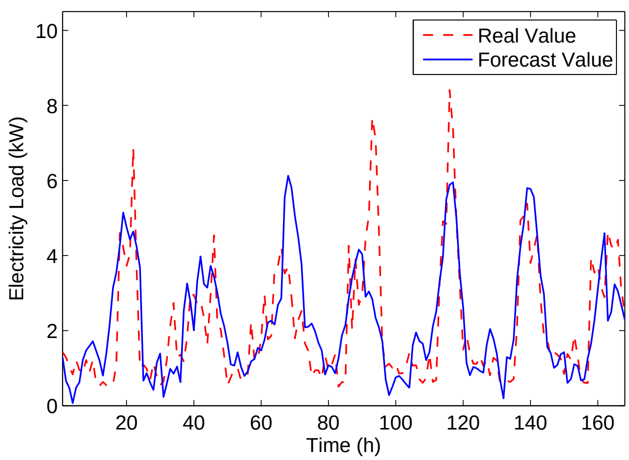


(b) ANN-based ST

Figure 6: Comparison of spatio-temporal benchmark algorithms for 6-h-ahead forecasting - Target house 2.

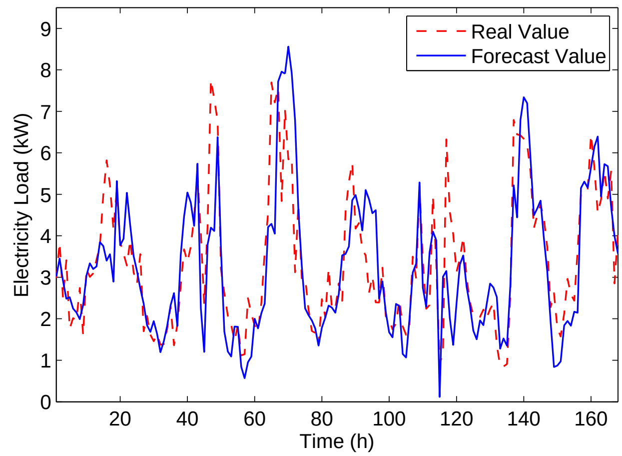


(a) Target House 1

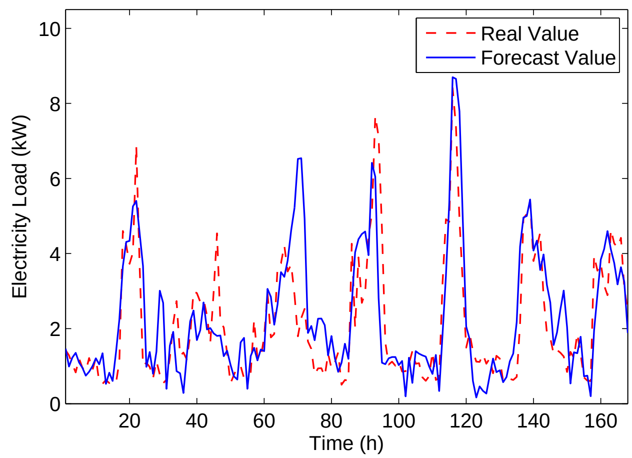


(b) Target House 2

Figure 7: Nonuniform M-AR algorithm results for 6-h-ahead forecasting. a) Target house 1 b) Target house 2.

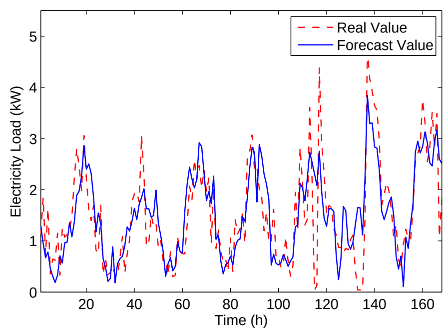


(a) Target House 1

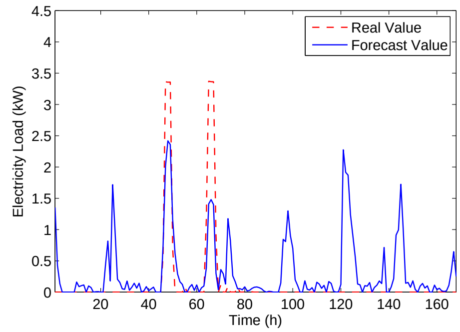


(b) Target House 2

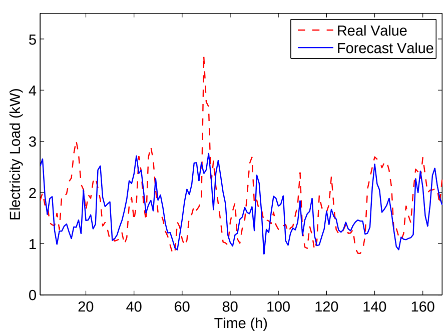
Figure 8: Nonuniform M-AR using WT algorithm results for 6-h-ahead forecasting. a) Target house 1 b) Target house 2.



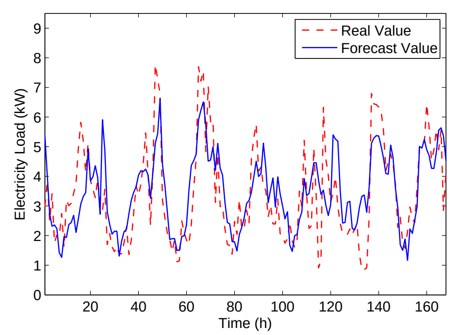
(a) AC load



(b) EV load

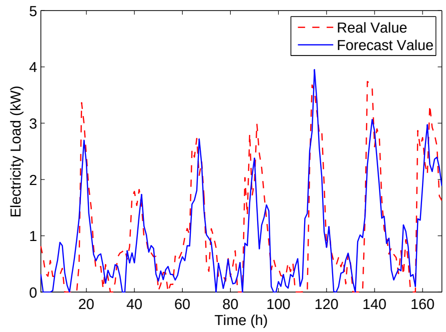


(c) Remaining load

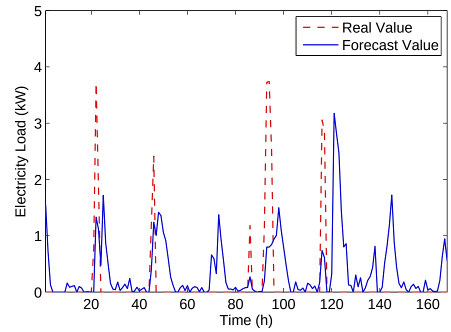


(d) Nonuniform M-AR based on individual forecasting

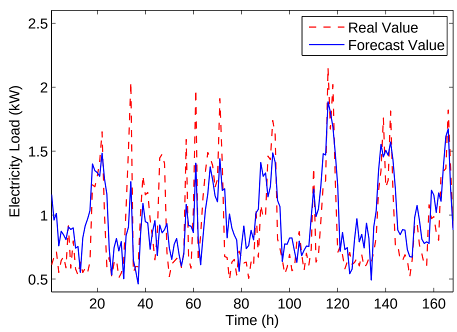
Figure 9: The proposed nonuniform CST-LF algorithm based on aggregation of single load forecasts for 6-h-ahead forecasting - Target house 1.



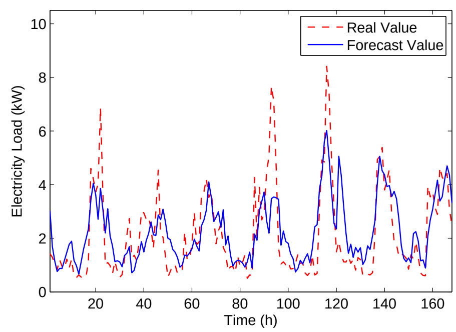
(a) AC load



(b) EV load

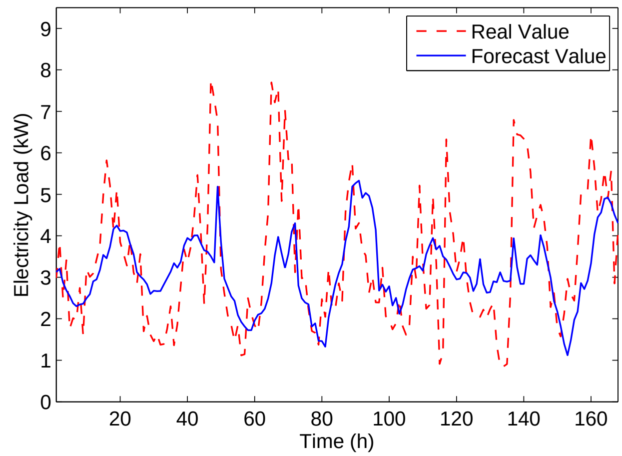


(c) Remaining load

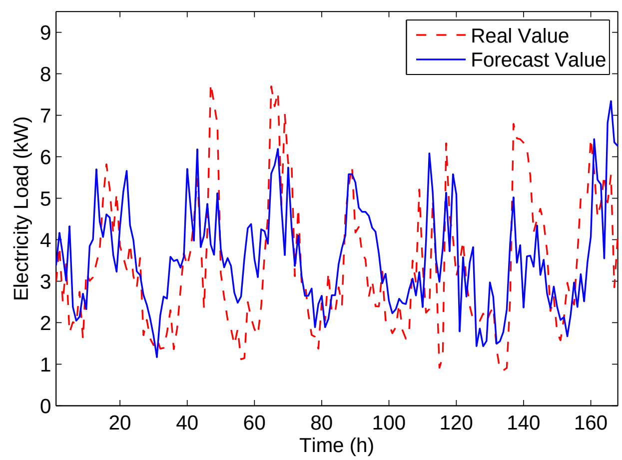


(d) Nonuniform M-AR based on individual forecasting

Figure 10: The proposed nonuniform CST-LF algorithm based on aggregation of single load forecasts for 6-h-ahead forecasting - Target house 2.

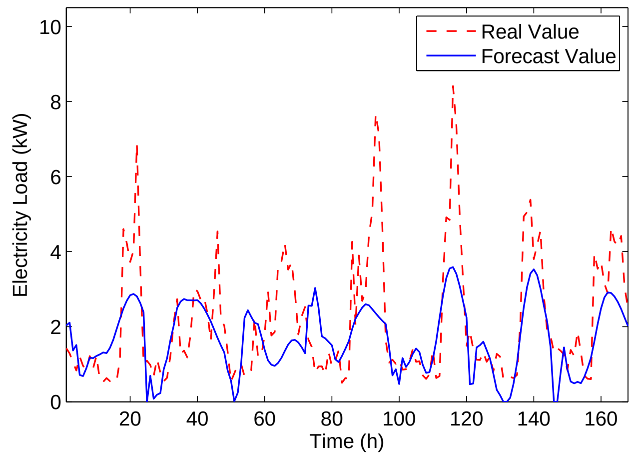


(a) LS-based ST

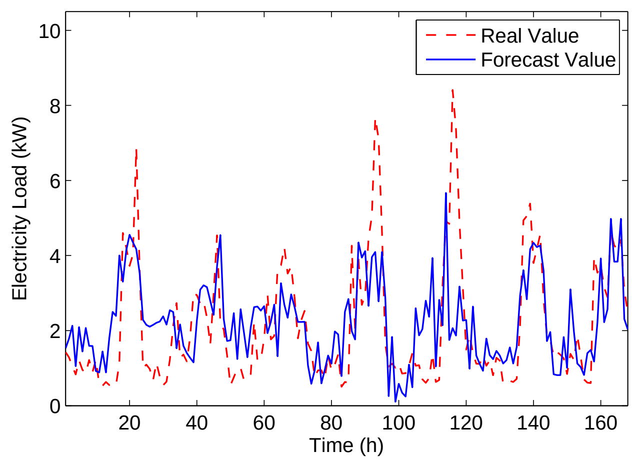


(b) ANN-based ST

Figure 11: Comparison of spatio-temporal benchmark algorithms for 24-h-ahead forecasting - Target house 1.

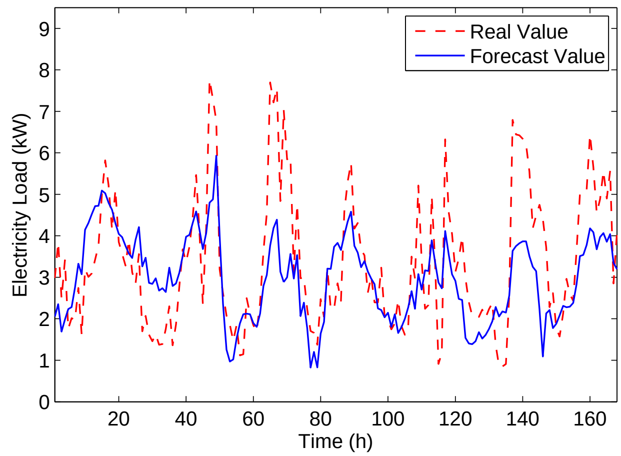


(a) LS-based ST

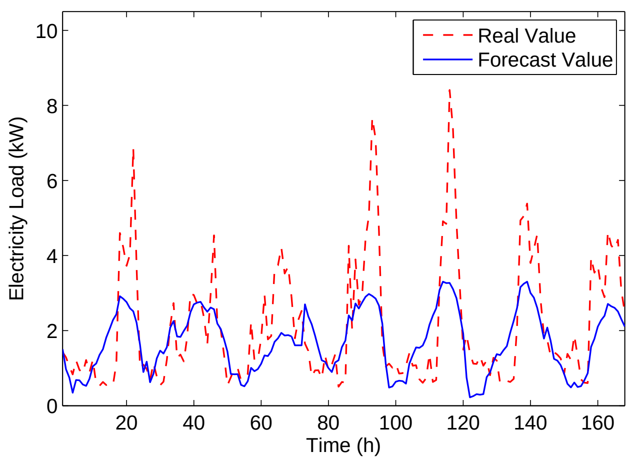


(b) ANN-based ST

Figure 12: Comparison of spatio-temporal benchmark algorithms for 24-h-ahead forecasting - Target house 2.

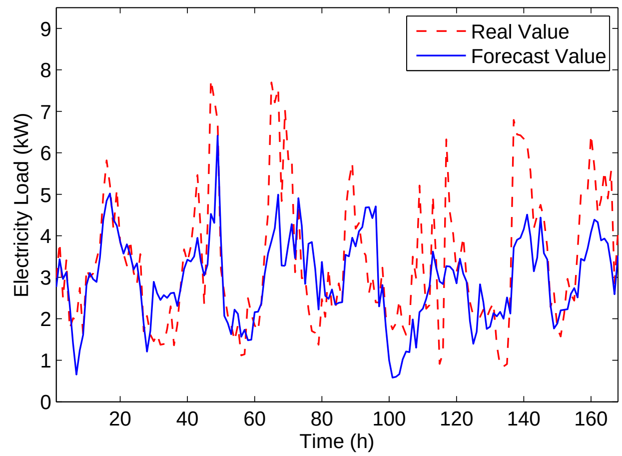


(a) Target House 1

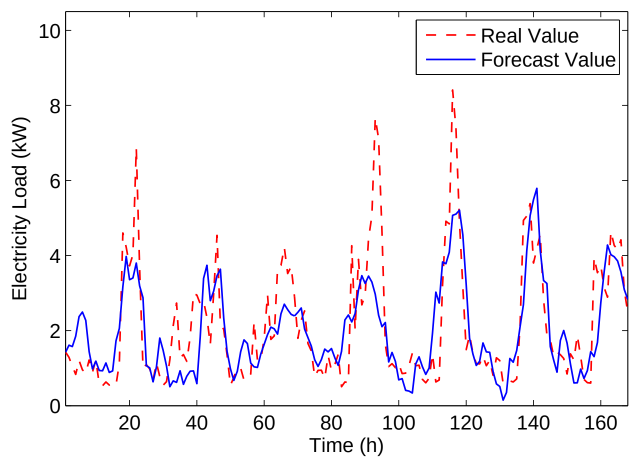


(b) Target House 2

Figure 13: Nonuniform M-AR algorithm results for 24-h-ahead forecasting. a) Target house 1 b) Target house 2.

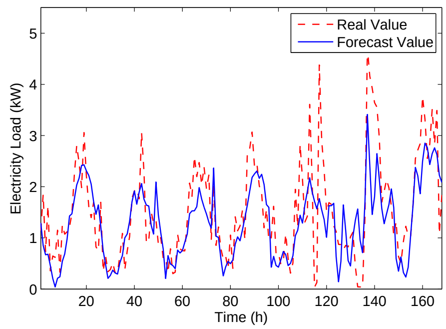


(a) Target House 1

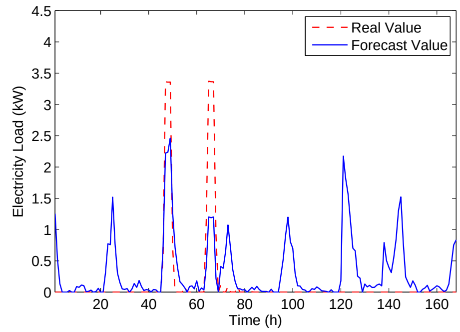


(b) Target House 2

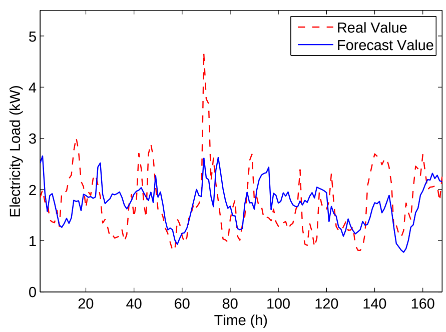
Figure 14: Nonuniform M-AR using WT algorithm results for 24-h-ahead forecasting. a) Target house 1 b) Target house 2.



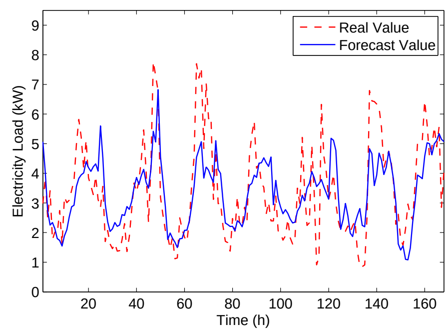
(a) AC load



(b) EV load

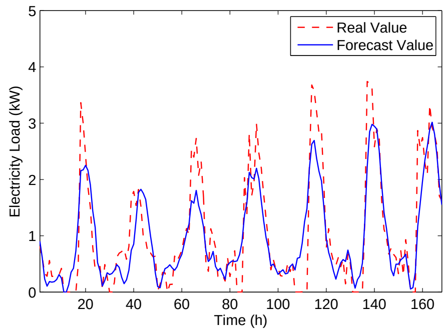


(c) Remaining load

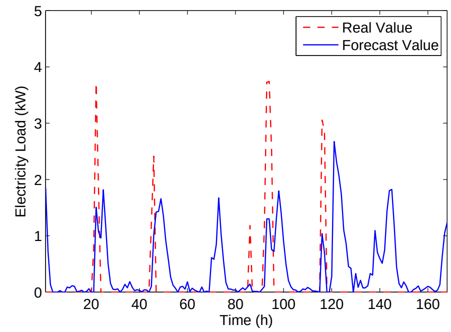


(d) Nonuniform M-AR based on individual forecasting

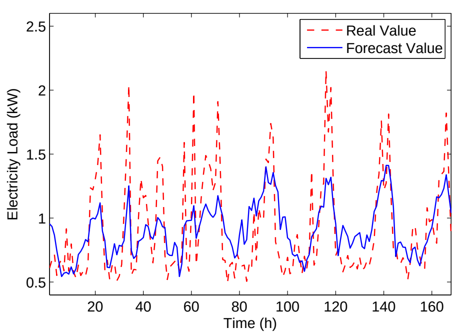
Figure 15: The proposed nonuniform CST-LF algorithm based on aggregation of single load forecasts for 24-h-ahead forecasting - Target house 1.



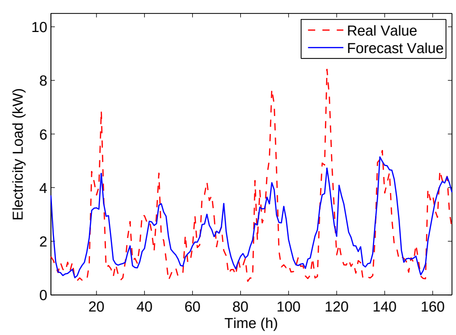
(a) AC load



(b) EV load



(c) Remaining load



(d) Nonuniform M-AR based on individual forecasting

Figure 16: The proposed nonuniform CST-LF algorithm based on aggregation of single load forecasts for 24-h-ahead forecasting - Target house 2.

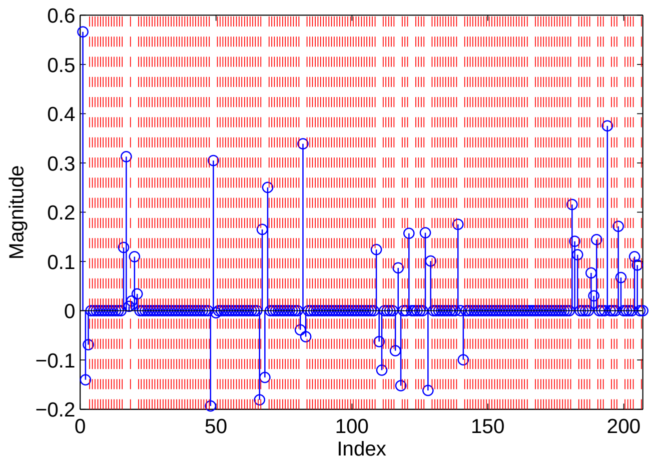


Figure 17: Sparse coefficient vector of Nonuniform CST-LF with different orders (vector-block lengths). The red dashed lines specify 173 vector-blocks of the coefficient vector.