Hybrid Models for Short-term Load Forecasting Using Clustering and Time Series

Wael Alkhatib, Alaa Alhamoud, Doreen Böhnstedt, and Ralf Steinmetz

Technische Universität Darmstadt,Fachgebiet Multimedia Kommunikation, S3/20, Rundeturmstr. 10, 64283 Darmstadt,Germany
{wael.alkhatib,alaa.alhamoud,doreen.boehnstedt,ralf.steinmetz}@kom.tu-darmstadt.de

Abstract. Short-term forecasting models on the micro-grid level help guaranteeing the cost-effective dispatch of available resources and maintaining shortfalls and surpluses to a minimum in the spot market. In this paper, we introduce two time series models for forecasting the day-ahead total power consumption and the fine-granular 24-hour consumption pattern of individual buildings. The proposed model for predicting the consumption pattern outperforms the state-of-the-art algorithm of Pattern Sequence-based Forecasting (PSF). Our analysis reveals that the clustering of individual buildings based on their seasonal, weekly, and daily patterns of power consumption improves the prediction accuracy and increases the time efficiency by reducing the search space.

Keywords: Smart grid; Sequence-based Forecasting; Time series models; K-means; Hierarchical clustering.

1 Introduction

Until 1998, the electricity market was divided into a set of power supply areas where the grid of each area is owned and supplied by one utility. However, this monopoly ended after the electricity market liberalization in Europe which increased the competition and led to more efficient production and supply of electricity. Power utilities are primarily involved in the trading with electricity suppliers on three different levels, namely: forward, day-ahead and intraday markets [5]. In forward market, utilities and suppliers agree on the deliveries of each year and up to six years with particularly liquid trading for the next three years. In the day-ahead trading, companies agree on the power deliveries for the next day and the deliveries are auctioned during the 12 midday before. However, oscillation of the spot market power consumption can happen due to unexpected events i.e. temperature changes, customers consumption pattern etc. To cope with these issues and to ensure the cost-effective dispatch of the available power generation facilities, the companies can after the day-ahead auction closure, trade on the intraday market level and agree on power deliveries on a very short-term basis from quarter hour to hour blocks which trigger the vital need for forecasting the day-ahead total power consumption and the fine-granular

24-hour consumption pattern of individual buildings. Beside the importance of accurate forecasting of power consumption on a very short timescale for utilities, the grid scale hourly power consumption prediction can assist the planning of duty cycles of A/C and ventilators in building management systems to flat the daily power consumption and/or to leverage low price periods based on signals from the power utilities monitoring systems.

The majority of previous research works have concentrated on aggregated energy consumption showing that accurate short-term consumption forecasting at the portfolio of buildings level can be achieved [16, 14, 12]. However, there is a lack of results related to estimating short-term power consumption of individual customers. In this work, we introduce two time series models, namely the Total Consumption Pattern Matching (TCPM) forecasting model which is used to predict the day-ahead total power consumption, and the Hourly Consumption Pattern Matching (HCPM) model which is used to predict the 24-hours consumption pattern of individual buildings i.e. the detailed consumption values at each hour of the day.

This paper is organized as follows: Section 2 gives an overview of related work in the domain of power consumption forecasting. In Sect. 3, we introduce our concept for the long-term and short-term forecasting of power consumption. Section 3 presents in details the short-term prediction models and their comparative analysis. Finally, Sect. 5 summarizes the paper and discusses future work.

2 Related Work

Short-term forecasting of power consumption is considered by the power utilities for economic scheduling [1] and real-time control where accurate and robust short-term forecasting guarantees the cost-effective dispatch of available resources while keeping shortfalls and surpluses to a minimum in the spot market. S. Aman et al. presented in [2] an empirical comparison between several averaging models for short-term forecasting of power consumption. Averaging models are characterized by their simplicity as they make their prediction based on averaging the power consumption of similar points of time horizon such as day, month, and year [9, 8].

Time series (TS) models predict future values based on previous historical observations. The commonly used approaches include Moving Average (MA), Auto-Regressive Integrated Moving Average ($ARIMA$) and the Pattern Sequence-based Forecasting (PSF) [12, 14]. Also regression models have been widely adapted to model and forecast power consumption [7, 6]. A Semi-Parametric Additive model for short-term (half-hourly) load forecasting model is proposed by Fan et al. [4]. In this work, the authors aim to tackle the non-linearity, volatile consumption pattern and interpret the effect of demand external drivers on power consumption prediction. The study integrates the non-linear and nonparametric driver factors within the regression framework. This is done by proposing semi-parametric additive model to estimate the relation between power consumption and affecting variables. The load demand drivers include the seasonality factors (weekdays, holidays and day of the year). In addition to the previous factors, the historical consumption and temperature data from the previous three hours as well as the same period from the previous six days are used.

Artificial intelligence techniques such as neural networks, support vector machines, and pattern matching techniques are widely applied to predict short-term power consumption[13, 11]. An overview of different AI techniques is provided in [10]. Among all AI-based methods, the technique of artificial neural networks (ANNs) has received substantial attention in forecasting power consumption due to its flexibility in learning load series and modeling the non-linearity between power consumption and the exogenous variables influencing it as well as providing fairly acceptable results. S. Wan et al. developed an artificial neural network model for modeling the electricity load of campus buildings in [17]. The input data of the network includes the consumption history of buildings and the time-dependent climate variables such as dew point, rainfall rate, pressure, wind speed, humidity and temperature.

In this work, we improve on previous works by dividing the buildings into homogeneous groups based on extracting features characterizing their consumption pattern over different tie horizons i.e. daily, weekly and monthly. Then we introduce two time series models namely, TCPM which is used to predict the day-ahead total power consumption, and HCPM model which is used to predict the 24-hours consumption pattern of individual buildings i.e. the detailed consumption values at each hour of the day.

3 Hybrid Short-term Power Consumption Forecasting Model

In this section, we present our design of two time-series models for short-term forecasting of power consumption inspired from the PSF algorithm, TCPM for predicting individual customers day-ahead power consumption and HCPM for day-ahead 24-hours consumption pattern forecasting. Both models consist of two phases. As the first phase of grouping the buildings based on time-derived discriminative features is common for both models, we present it separately in the next section.

3.1 First Phase: Clustering of Buildings Using Time-Derived Discriminative Features

The goal of this phase is to divide the buildings into several homogeneous groups where each group contains only buildings with similar power consumption patterns. As a first step, we extract discriminative features that characterize the power usage pattern of individual buildings by leveraging the seasonal, weekly, and daily patterns in their historical consumption data. The features are extracted with regard to different time horizons, namely day segments, daily, monthly, and yearly. The extracted features are as follows:

- Average power consumption for each week day and each month.
- Percentage of power consumption over six day segments: early morning, morning, early afternoon, afternoon, early night, and late night similar to previous work [15].
- Total power consumption over the trial in kWh.
- Percentage of power consumption over business days, Sunday, and Saturday.

The vital point of using time discriminative features is to identify different classes of buildings without previous knowledge about them and only by relying on the shape of their power consumption during the trial. After the features extraction, the proposed clustering methodology consists of the following steps:

1. Normalizing the extracted features in the range [0,1] and assign different importance levels to them.
2. Clustering buildings using k-means and hierarchical clustering algorithms.
3. Selecting the optimal number of clusters based on validity indexes of the clustering process namely Dunn, Silhouette, and Davies-Bouldin.
4. Selecting the optimal clustering algorithm using the same validity indexes.

Two clustering algorithms, namely k-means and hierarchical clustering are used to identify the different groups of buildings in the designed feature space. In section 4.1, we evaluate our clustering methodology in order to determine the optimal clustering algorithm as well as the optimal number of clusters to which the buildings should be divided.

3.2 Hourly Consumption Pattern Matching (HCPM) Model

As an output of the first phase, we get a set of clusters where each cluster represents buildings with similar consumption pattern. The second phase in HCPM model consists of two independent steps. In the first step, each group from the initial clustering process in the first phase is clustered using hierarchical clustering. Differently from first phase, the clustering in each group is done based on their detailed 24-hour consumption pattern in order to produce a set of different patterns which characterize the different daily consumption patterns of each building. Then a label will be assigned to each resulting cluster. The assigned labels will be used in the next stage for pattern sequence matching. The number of clusters is selected based on majority vote of the previously used clustering indexes. As a result of the labeling process, we will have our 24-hour daily patterns in each group represented as a time series of subsequent labels where each label corresponds to one day. In the second step, HCPM utilizes the labels produced by the clustering phase in the pattern matching process in order to forecast the next day 24-hours consumption pattern. The forecasting process starts by searching for days with same historical power consumption pattern with the window size confirming to the length of the labels sequence.

The algorithm tries to find a matched pattern in the historical consumption data of same building . In case of multiple matches, it takes the closest match in time horizon because the consumption behavior should be more similar in recent past. If the pattern is not found in the historical data of same building, the algorithm searches in the repository of historical consumption data of all buildings. If multiple matches are found, the majority vote over all matched patterns is considered as the predicted pattern. Finally, if no sequence is found in the repository corresponding to window size W, the algorithm searches for the sequences of labels equals to W-1 and thus successively. The searching procedure for a match uses only the historical data of the same day of the week. The full procedure of searching for an equal pattern in HCPM is illustrated in Fig. 1.

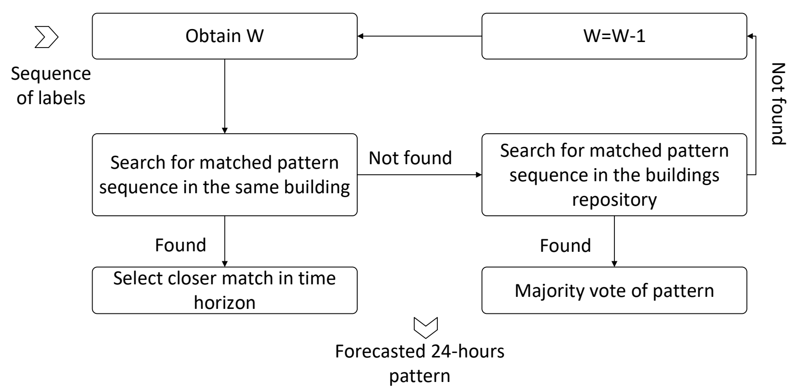

Fig. 1: Workflow of the hourly consumption pattern matching algorithm (HCPM).

3.3 Total Consumption Pattern Matching (TCPM) Model

TCPM forecasting model depends in its second phase on the total power consumption in each day rather than the used labels of each day's 24-hour consumption pattern in HCPM. The percentage error defined in (1) is used as a metric to decide whether two days are considered as a match or not, where $C1$ and $C2$ are the total power consumption in the days to be compared.

$$PR(C1, C2) = |\frac{C1 - C2}{C1}| \tag{1}$$

TCPM follows the same steps as in HCPM. The difference is that the numerical values of the power consumption are the input for the algorithm instead of the sequence of labels in HCPM. Also if multiple matches are found in the repository of all buildings, the average consumption is considered as the final prediction while the majority vote is used in case of HCPM.

4 Evaluation

The used dataset in this work was provided by the Commission of Energy Regulation (CER) in Ireland [3]. Raw data representing the 30-minute power consumption readings in kWh of individual buildings was collected over a period of eighteen months during 2009 and 2010. More than 5000 smart meters were installed in Irish homes and businesses in eight urban areas and three villages. Pre-trial and post-trial surveys were conducted for both residential and business participants. The questions were related to their demographic features, building characteristics, life style and usage patterns of different appliances. For the evaluation of our proposed short-term power consumption forecasting models, we used the data related to small and medium enterprise (SME) buildings as the oscillation in their power consumption have a more regular pattern compared to the residential buildings.

4.1 Buildings Clustering Using Time Derived Discriminative Features

The average daily, monthly, and total power consumption during the trial were given a higher importance level. The different importance levels can be given after normalizing all features values in the range [0,1] by multiplying the range with a constant factor. Based on extensive analysis of different features combinations, the final set of used features consisted of the total consumption as well as the consumption percentages on Sunday, Saturday, and different day segments. We have given the percentage of consumption on Sunday, Saturday and different day segments the same importance level whereas the total consumption was given the highest importance level.

The optimal number of clusters for k-means and hierarchical clustering was selected based on the majority vote of the three validity indexes. Lower value of Davies Bouldin index indicates better clusters quality while higher values for Silhouette and Dunn indexes prove better clustering quality. Figure 2 illustrates the clustering quality for hierarchical and k-means algorithms. Apparently the indices reached their optimal values when the number of clusters K=8. We proceeded using hierarchical clustering approach based on the relative comparison of the indexes' scores for both algorithms.

Figure 3 illustrates the average daily power consumption in kWh of each cluster generated using hierarchical clustering with an optimal number of clusters K=8. There is a clear separation between groups 8, 6, 4, and 2. This indicates that the total power consumption is the main factor to distinguish between these groups. However, groups 1, 3, and 7 have close daily average power consumption and the total power consumption is not the main factor to distinguish between these three classes but rather the percentage of power consumed during the weekend in comparison to the whole week as well as the percentage of power consumed over different day segments.

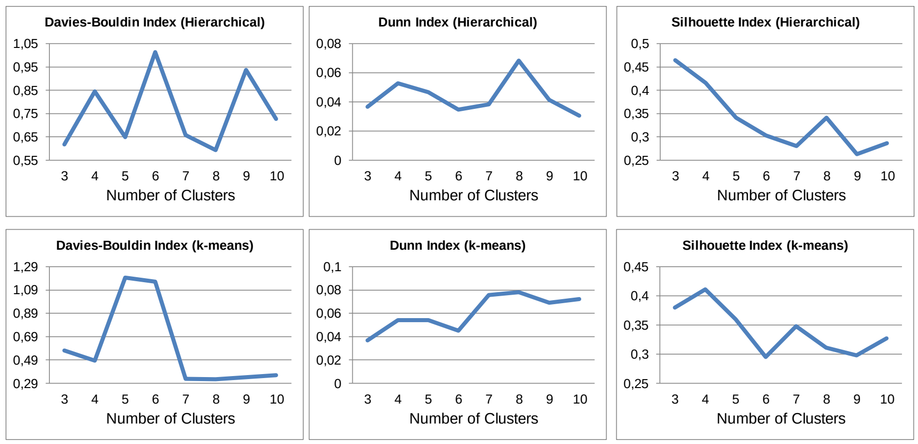

Fig. 2: Derived features clustering indexes.

4.2 HCPM Model

For the evaluation of HCPM model, one group named group 3 was selected with 59 SME buildings. A subset of 5 months of the trial was used as a repository of power consumption historical data that can be used to predict the day-ahead power consumption pattern. Also one month was excluded from the clustering step for out-of-sample evaluation. The input for the hierarchical clustering algorithm is the dataset of detailed 24-hours power consumption for all buildings in the repository of power consumption historical data. Each data-point is a sequence of 24 points representing the consumption during 24 hours in kWh. The optimal number of clusters was selected based on the majority vote of the same validity indexes. Figure 4 demonstrates the quality of clusters generated using hierarchical clustering algorithms.

Based on the majority vote, K=10 and K=9 were selected as the optimal number of clusters. Mean Error Relative (MER) is used for analyzing the effect of the obtained window size on the prediction error. Figure 8 demonstrates the relation between the window size and the prediction MER. Mean Error Relative (MER) starts with 24% and falls down to a minimum of 17.1 % for a window size equals to 9.

Figure 5 demonstrates the relation between the window size and the percentage of correctly predicted labels. We define the precision as a metric reflecting the percentage of correctly predicted labels divided by the total number of tested samples. It shows that the precision increases significantly from 62.5% for W=1 to 82.1% for W=10. Also the number of found matches will decrease with an increasing window size. However, more matches can be found in buildings which are following a consumption pattern close to the targeted building. The results in Fig. 6 are consistent with the MER statistics in Fig. 8.

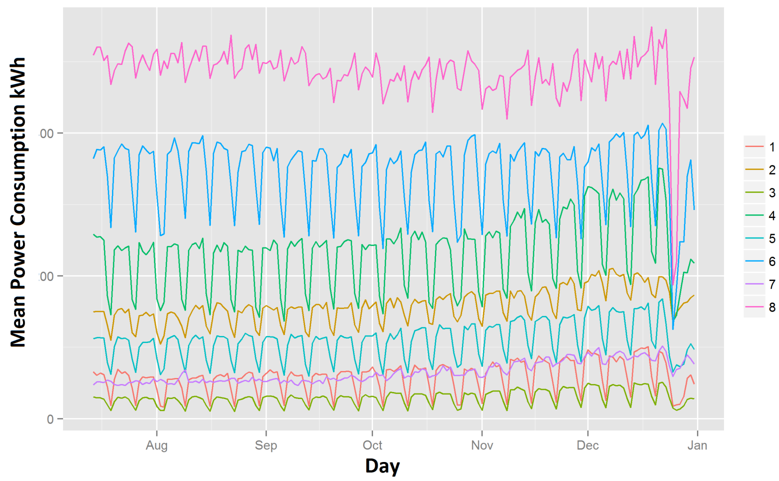

Fig. 3: Daily mean power consumption of each cluster generated using hierarchical clustering algorithm with an optimal number of clusters K=8.

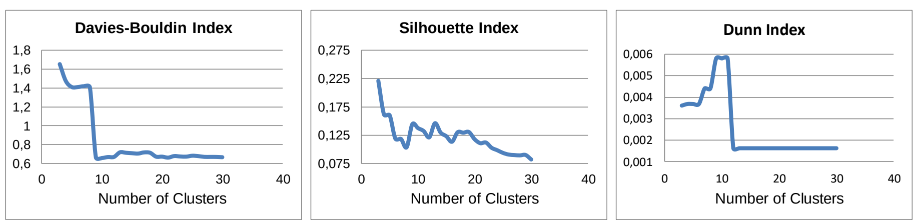

Fig. 4: Selecting the optimal number of clusters for group 3.

4.3 TCPM Model

For the evaluation of the TCPM model, the same group of 59 SME buildings was used. We considered a percentage error of less than 10% to count for a match between a day in the repository of historical data of the buildings and a day in the consumption sequence of previous days of the day in question. Figure 8 depicts the relation between the window size and the prediction MER. The MER is falling gradually from 22.6% and is reaching a minimum value of 9.1% for a window size equal to 8. Then, it slightly increases for a window size W=9 and again decreases to 9.2% for W=10. As a conclusion, more knowledge about the power consumption of previous days can help improving the prediction accuracy. However, a larger window might reduce the accuracy as we notice for window size of 9. Moreover, a larger window size will increase the possibility of not finding a match. Figure 7 demonstrates the drop in the number of found patterns based on

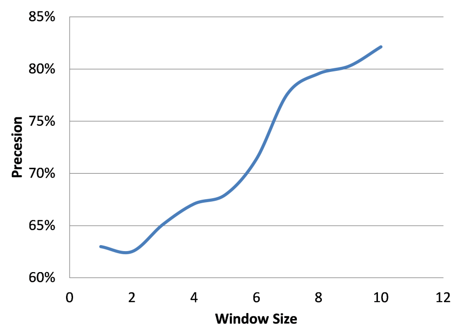

Fig. 5: HCPM precision for different values of window size.

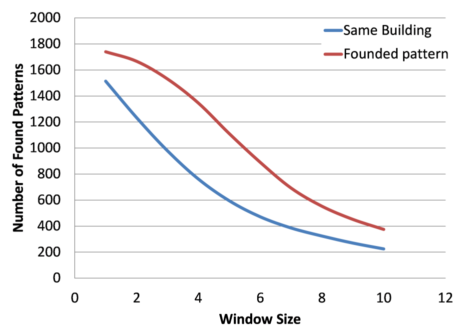

Fig. 6: HCPM number of matches for different values of window size.

the increased window size also the relation between the number of found patterns and whether they belong to the consumption history of the same building or not. With a larger window size, the possibility to find a match in the historical data of the same building is higher than finding this pattern in other buildings.

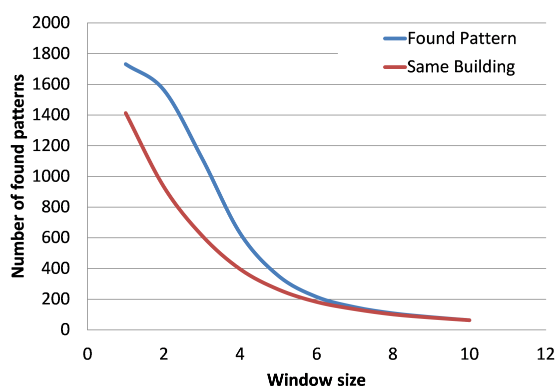

Fig. 7: The effect of window size on the number of found matches for TCPM model.

4.4 Comparative Evaluation of HCPM against TCPM

Figure 8 shows that pattern matching using the total power consumption TCPM significantly outperforms hourly consumption pattern matching (HCPM) in terms of mean error relative (MER). The reason is that the used Euclidean distance to measure the similarity level between two days based on the 24-hours consumption pattern might assign the same label to two days with a large difference in total power consumption just because both have small distances to the centroid of the cluster they belong to. To sum up, the total power consumption is more appropriate for predicting the day-ahead power consumption.

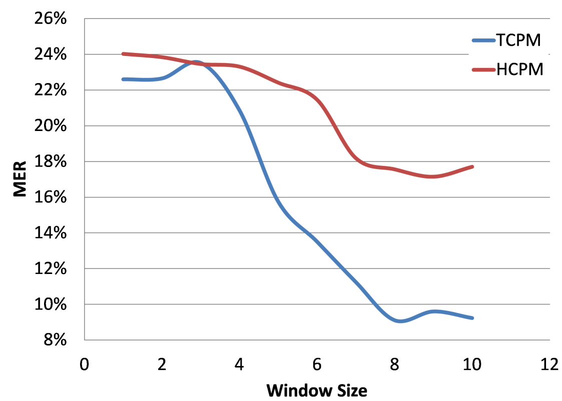

Fig. 8: Relative comparison of MER values.

4.5 Comparative Evaluation of HCPM against PSF

PSF utilizes the k-means algorithm to realize the clustering of the dataset. The first step was to determine the optimal number of clusters. For this, same indexes were used to validate the clustering quality and select the number of clusters based on the majority vote. The optimal number of clusters has produced clusters with a big gap in the average hourly power consumption as shown in Table 1. Group 4 contains days with an average hourly power consumption around 35.66 kWh. Therefore, we can roughly conclude that it represents retail buildings with high power consumption and not days with different consumption pattern based on day segments.

For the comparative analysis of prediction accuracy, we evaluated the PSF using four different window size values $W = 7, W = 8, W = 9, W = 10$. With $W \geqslant 7$ PSF should be capable of capturing the similarity over a week of consumption data. Table 2 shows the comparative analysis between HCPM and PSF. The proposed model outperformed PSF in terms of time efficiency and prediction accuracy by introducing three enhancement on the original PSF approach. Firstly, the proposed model clusters the buildings based on derived features using the seasonal, weekly and daily pattern of the historical power consumption data. This step makes substantial contribution to the overall performance by reducing the search space to include data for buildings belonging to the same cluster only. Moreover, by giving high importance to the total power consumption as one dimension for clustering, we separate buildings belonging to different power

Table 1: Average hourly power consumption in kWh of the k-means clusters.

Cluster Number	**Average Hourly Power Consumption (kWh)**
1	6.62
2	10.03
3	4.46
4	35.66
5	1.32
6	18.73
7	15.14

Table 2: Comparative analysis between HCPM and PSF.

Window Size	**MER**	
W	PSF	HCPM
7	46.73%	24.02%
8	45.11%	23.84%
9	43.15%	17.14%
10	39.91%	17.70%

consumption classes. In addition, the algorithm tries to find a matched pattern in the same building historical data as a first option. This heuristic reduces the search space and increases the prediction accuracy in case a match is found in same building historical data. With more historical data of each building, the chance for this heuristic to successfully find a match is higher. Finally, searching for a match using only the same day of the week also reduces the search space more. This heuristic will reduce the probability to find a match in the historical data, however, this issue can be overcome by providing more historical consumption data. The experiments prove that the used traditional time series model of PSF cannot forecast the day-ahead power consumption of individual customers in case of inhomogeneous groups. However, additional heuristics can improve the traditional methods ability to predict the individuals short-term consumption.

5 Discussion and Future Work

In this work, two time series models TCPM and HCPM were designed to predict individual customers day-ahead power consumption and the 24-hours consumption pattern respectively. The comparison against PSF showed that the proposed model of HCPM significantly outperformed PSF, also our analysis revealed that, clustering buildings based on their seasonal, weekly and daily patterns can improve the prediction accuracy and increase time efficiency by reducing the search space. In future work, HCPM model will be tested on the post-trial dataset in which the participants were allocated different tariffs, in order to check whether the proposed model is able to handle the changes in customers demand pattern.

References

1. Alfares, H.K., Nazeeruddin, M.: Electric Load Forecasting: Literature Survey and Classification of Methods. International Journal of Systems Science 33(1), 23–34 (2002)
2. Aman, S., Frincu, M., Chelmis, C., Noor, M.U.: Empirical Comparison of Prediction Methods for Electricity Consumption Forecasting. Department of Computer Science, University of Southern California, Los Angeles, CA, 90089 (2012)
3. Commission for Energy Regulation (CER): Electricity Smart Metering Technology Trials Findings Report. ESB Networks, Belgard Square North, Tallaght, Dublin 24. (2011)
4. Fan, S., Hyndman, R.: Short-term Load Forecasting Based on a Semi-parametric Additive Model. Power Systems, IEEE Transactions on 27(1), 134–141 (Feb 2012)
5. Gabriel, S.: An Electricity Market for Germany's Energy Transition. Federal Ministry for Economic Affairs and Energy (BMWi),PRpetuum GmbH, Munchen (2014)
6. Hong, T., Gui, M., Baran, M., Willis, H.: Modeling and Forecasting Hourly Electric Load by Multiple Linear Regression with Interactions. In: Power and Energy Society General Meeting, 2010 IEEE. pp. 1–8 (July 2010)
7. Jiang, H., Lee, Y., Liu, F.: Anomaly Detection, Forecasting and Root Cause Analysis of Energy Consumption for a Portfolio of Buildings Using Multi-step statistical Modeling (2012), uS Patent App. 13/098,044
8. K. Coughlin, M. A. Piette, C.G., Kiliccote, S.: Emergency Demand Response Program Manual, Sec 5.2: Calculation of Customer Baseline Load (CBL). New York Independent System Operator, Southern California Edison, Tech. Rep. (2010)
9. K. Coughlin, M. A. Piette, C.G., Kiliccote, S.: 10-Day Average Baseline and Day-Of Adjustment. Southern California Edison, Tech. Rep. (2011)
10. K. Metaxiotis, A. Kagiannas, D.A., Psarras, J.: Artificial Intelligence in Short-term Electric Load Forecasting: a State-of-the-art Survey for the Researcher, Energy Conversion and Management (2003)
11. Khotanzad, A., Afkhami-Rohani, R., Lu, T.L., Abaye, A., Davis, M., Maratukulam, D.: Annstlf-a Neural-network-based Electric Load Forecasting System. Neural Networks, IEEE Transactions on 8(4), 835–846 (Jul 1997)
12. Martinez Alvarez, F., Troncoso, A., Riquelme, J., Aguilar Ruiz, J.: Energy Time Series Forecasting Based on Pattern Sequence Similarity. Knowledge and Data Engineering, IEEE Transactions on 23(8), 1230–1243 (Aug 2011)
13. Rui, Y., El-Keib, A.: A Review of Ann-based Short-term Load Forecasting Models. In: System Theory, 1995., Proceedings of the Twenty-Seventh Southeastern Symposium on. pp. 78–82 (Mar 1995)
14. Shen, W., Babushkin, V., Aung, Z., Woon, W.: An Ensemble Model for Day-ahead Electricity Demand Time Series Forecasting. In: Proceedings of the Fourth International Conference on Future Energy Systems. pp. 51–62. ACM, New York, NY, USA (2013)
15. Silipo, R., Winters, P.: Big Data, Smart Energy, and Predictive Analytics Time Series Prediction of Smart Energy Data. KNIME.com (2013)
16. Simmhan, Y., Noor, M.: Scalable Prediction of Energy Consumption Using Incremental Time Series Slustering. In: Big Data, 2013 IEEE International Conference on. pp. 29–36 (Oct 2013)
17. Wan, S., Yu, X.H.: Facility Power Usage Modeling and Short-term Prediction with Artificial Neural Networks. In: Advances in Neural Networks - ISNN 2010. pp. 548–555 (Mar 2010)