

# Machine Learning Approaches to Electricity Consumption Forecasting in Automated Metering Infrastructure (AMI) Systems: An Empirical Study

A. Jayanth Balaji<sup>(✉)</sup>, D.S. Harish Ram, and Binoy B. Nair

Department of Electronics and Communication Engineering,  
Amrita School of Engineering, Coimbatore Amrita Vishwa Vidyapeetham,  
Amrita University, Coimbatore 641112, India  
{a\_jayanthbalaji, ds\_harishram, b\_binoy}@cb.amrita.edu

**Abstract.** In a Smart grid, implementation of value-added services such as distribution automation (DA) and Demand Response (DR) [1] rely heavily on the availability of accurate electricity consumption forecasts. Machine learning based forecasting systems, due to their ability to handle nonlinear patterns, appear promising for the purpose. An empirical evaluation of eight machine learning based systems for electricity consumption forecasting, based on Extreme Learning machines (ELM), Ensemble Regression Trees (ERT), Artificial Neural Network (ANNs) and regression is presented in this study. Forecasting systems thus designed, are validated on consumption data collected from 5275 users. Result indicate that ELM based electricity consumption forecasting systems are not only more accurate than other systems considered, they are considerably faster as well.

**Keywords:** Machine learning · Extreme Learning machines (ELM) · Ensemble Regression Trees (ERT) · Artificial Neural Network (ANNs) · Hodrick-Prescott (HP)

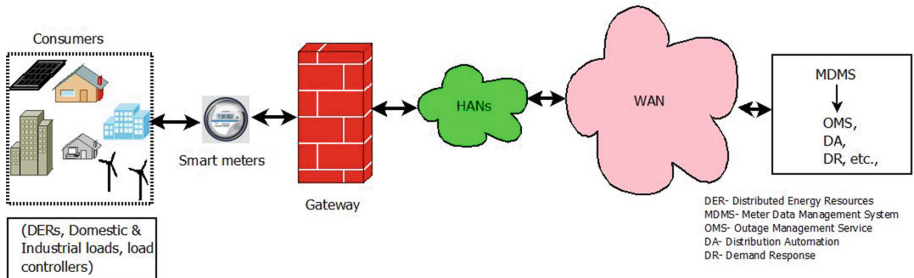
## 1 Introduction

Automated Metering Infrastructure (AMI) is an integral part of a smart grid. AMI not only allows the consumers facilities such as real time consumption monitoring and automated billing etc., it also allows the power utilities to implement value-added services such as DA and DR. A typical AMI system is represented in Fig. 1 [2]. One of the major aspects of AMI is its ability to learn customer electricity consumption behavior and accordingly optimize the quality of smart grid services delivered to the user. A detailed analysis of the issues related to load forecasting and DR has been presented in [3, 4]. AMI applications to scheduling

---

The CER Electricity Dataset is sourced from Irish Social Science Data Archive (ISSDA).

of residential microgrids [5], weather based load normalization and increasing the forecast accuracy for hourly forecast consumption is done in [6] and the user segmentation based on users power consumption profile forecasting for hourly power consumption data and thereby maximizing the benefits of smart grid services especially on DR is done in [7].



**Fig. 1.** AMI block diagram [2]

However, it is observed that there are few studies on identification of suitable techniques for electricity consumption forecasting. It is also observed that the consumption patterns of users change according to the time of the day and according to the seasonal variations. This aspect has also not received much attention. Consumption behavior is also seen to be very different for domestic consumers when compared to industrial users. The earlier work by the authors on the topic [2] attempted to empirically address these aspects of consumption forecasting by analyzing forecast performance of machine learning based systems for 485 Small and Medium Enterprise consumers. This paper presents a much more comprehensive study with a total of 5275 users and four different forecasting techniques.

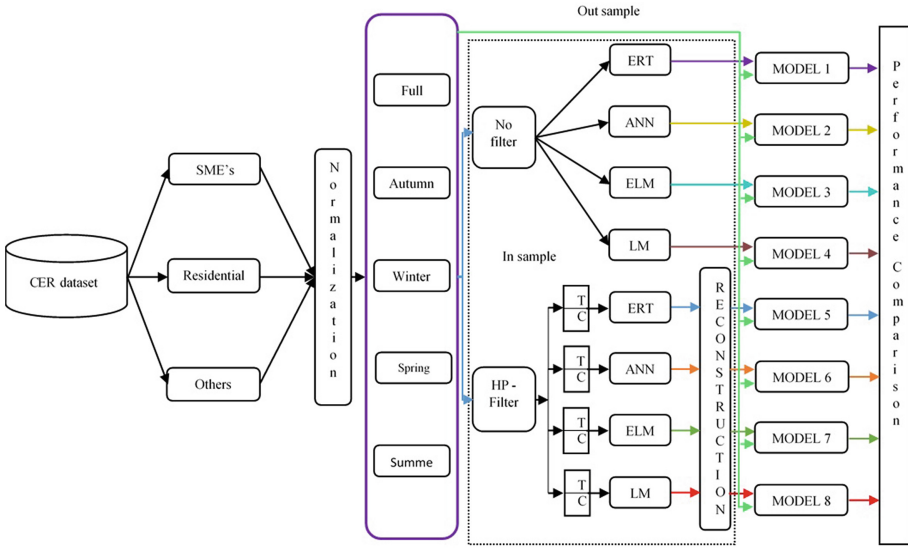
Rest of the paper is organised as follows: Sect. 2 presents the description of the proposed system and the methodology, Sect. 3 details the experimental results and conclusions are presented in Sect. 4.

## 2 System Description and Methodology

The block diagram of the proposed system is presented in Fig. 2.

### 2.1 Dataset and Preprocessing

All the models were developed for CER electricity dataset acquired from Irish Social Science Data Archive (ISSDA) [8] comprising half hourly interval electricity consumption data for 6445 users, subcategorized as (a) Small and Medium Enterprise (SME) consisting of 485 users (b) Residential (RES) consisting of



**Fig. 2.** Proposed system block diagram

4225 users and (c) ‘Others’ (OTH) consisting of 1735 users. The data has been recorded over the period 00:00 h, 14/07/2009 to 23:59 h, 29/12/2010. It was observed on inspection that in the OTH category, the consumption data for 1170 users is not available for the entire duration listed above and was not considered for further analysis. The total number of users considered, then, is 5275.

As the first preprocessing step, the consumption data for each user is split into four subsets: summer (May to July), winter (November to January), autumn (August to October) and spring (February to April). Different forecasting models are trained for each season separately. For comparison purposes, the models were also trained on the data without seasonal splits.

As the next step, after normalization of the data, the trend and cyclic components of the data are separated using Hodrick-Prescott (HP) filter [9]. HP filter has been successfully used in [10–13] as a preprocessing technique for financial forecasting systems. It has also been reported to offer an enhancement in performance for electricity consumption forecasting systems in [2]. Machine learning based forecasting systems described later in the paper are separately trained to forecast trend and cyclic components. Finally both the forecasts are added to yield the estimated consumption for that instant. Systems that forecast the consumption based directly on the normalized data alone (i.e. not employing the HP filter) are also evaluated.

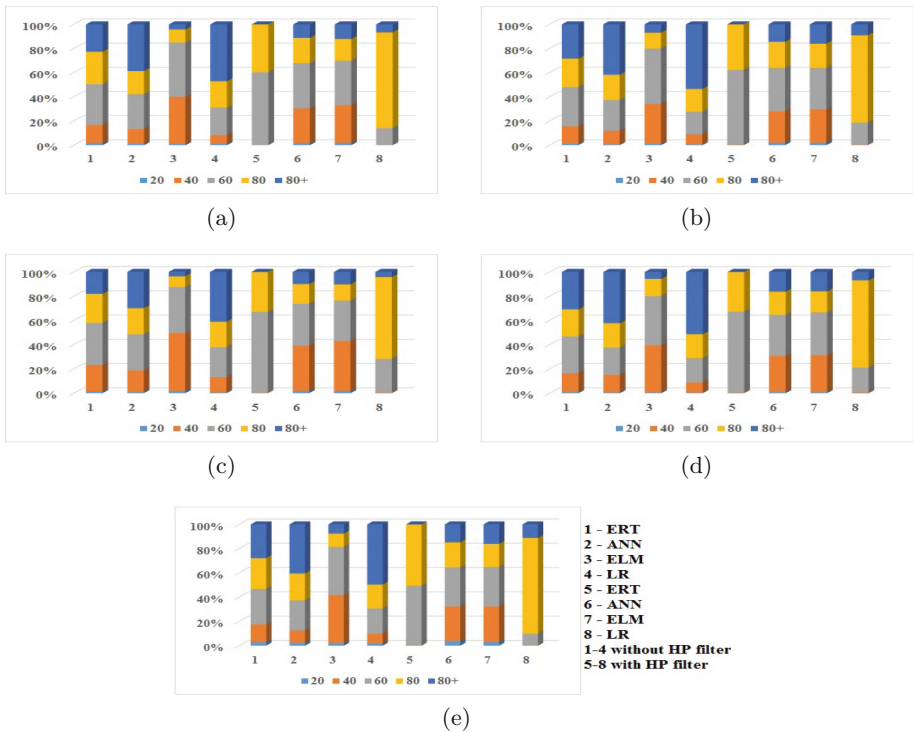
Next step in the process is to train a machine learning system to estimate the electricity consumption. Three different machine learning techniques as well as linear regression (LR) models are investigated for the purpose. Each of these techniques is described in the next sub-section.

## 2.2 Machine Learning Techniques Considered

A total of four techniques, namely, ELM, ANN, ERT and LR were evaluated in the present study. Hence a total of eight different forecasting models were evaluated (see Fig. 2). Each of the techniques used and the system parameters are described below.

### 2.2.1 Extreme Learning Machines (ELMs)

ELMs are basically single hidden layer feedforward neural network (SLFNN), however in ELMs the weights connecting input and hidden layer are randomly fixed and not updated. The hidden to output layer weights are learnt in a single step with the help of Pseudoinverse technique which makes the architecture much faster and generalized when compared to any other neural network, with minimal training errors and smaller norm of weights [14, 15]. Kernel based ELMs have also been proposed [16]. The results of employing kernel ELMs have been encouraging, as seen in [16, 17]. A detailed treatment on kernel ELMs can be found in [15]. In the proposed work, the ELM Models are designed using linear kernel with regularization coefficient as ‘10’. The kernel function and the regularization coefficient value were arrived at, using trial-and-error. The estimated



**Fig. 3.** MdAPE plots for all models with RES category users (a) all seasons together (b) Autumn (c) Winter (d) Spring and (e) Summer.

output from the ELM,  $\hat{Y}$  can be expressed as follows:

$$\hat{Y} = W_{H-O}\sigma_A(W_{I-H}X) \tag{1}$$

where:  $W_{I-H}$  = input - hidden layer weight vector,  $W_{H-O}$  = hidden - output layer weight vector and  $\sigma_A$  = activation function.

Algorithm:

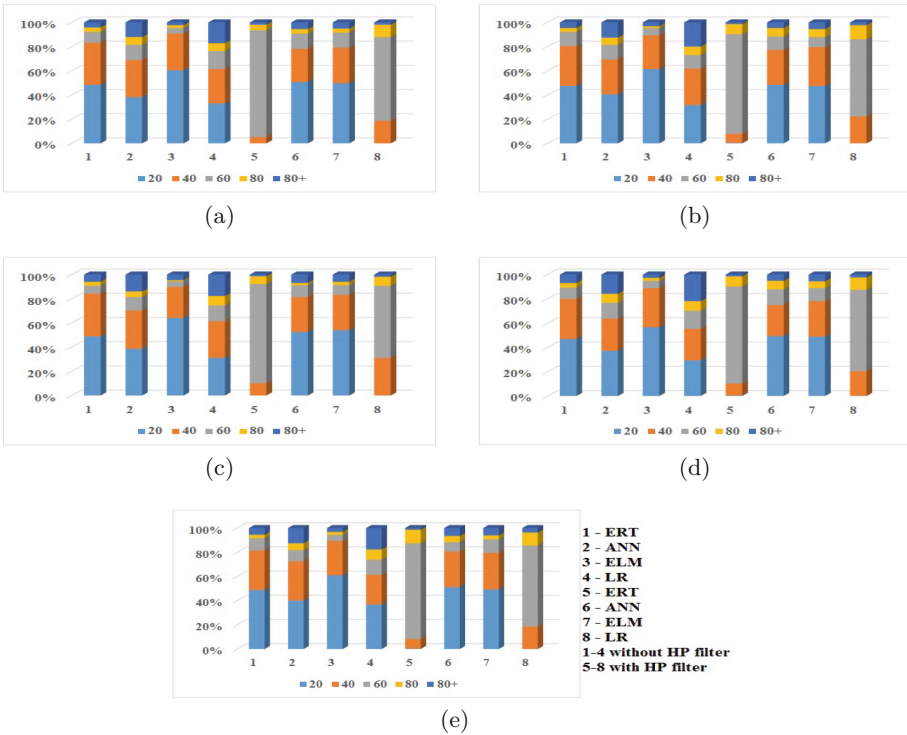
Step 1: Initialize  $W_{I-H}$  with random weights.

Step 2: Compute  $W_{H-O}$  using pseudoinverse:

$$W_{H-O} = \sigma_A(W_{I-H}X)^+Y \tag{2}$$

### 2.2.2 Artificial Neural Networks (ANNs)

ANNs have been very widely used for forecasting purposes [18]. In the present study single hidden layer feedforward ANNs trained using Levenberg-Marquardt (LM) learning algorithm and 5 hidden neurons (number of hidden neurons is arrived at, using trial and error) are employed.



**Fig. 4.** MdAPE plots for all models with SME users (a) all seasons together (b) Autumn (c) Winter (d) Spring and (e) Summer.

### 2.2.3 Ensemble Regressing Trees (ERTs)

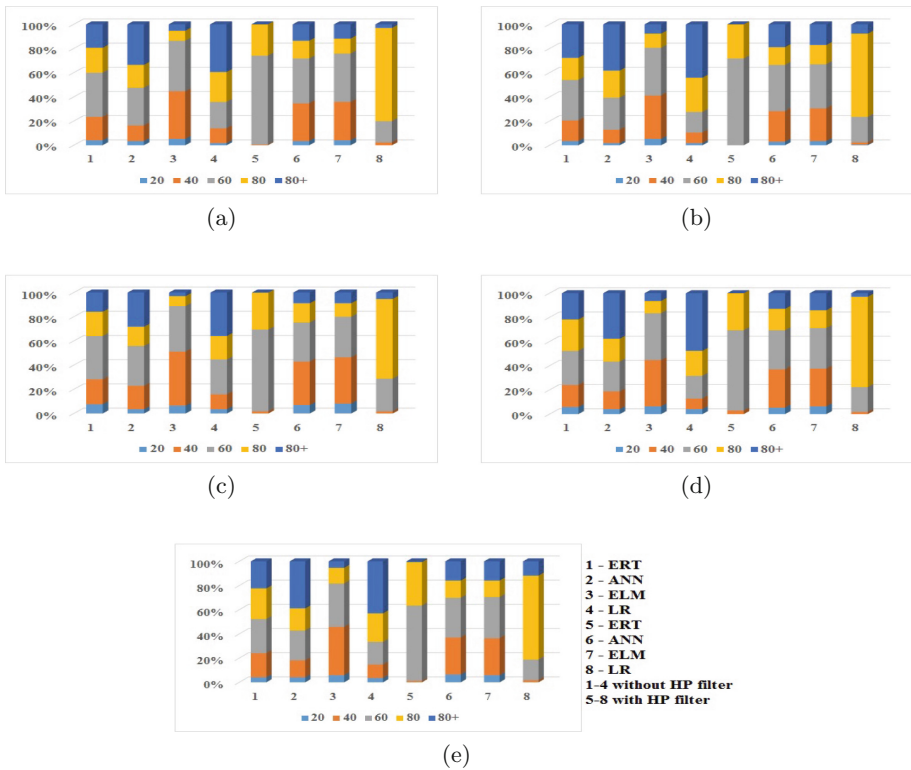
ERTs are forecast models which are generally weighted combinations of several regression trees which contributes to increase in predictive performance of regression tree. ERT models in the present study are developed using bagging [19] and number of trees is 5 (selected using trial and error).

### 2.2.4 Linear Regression (LR)

This are the simplest curve fitting algorithms available. The coefficients for the LR models considered in the present study were identified using simple Ordinary Least Squares technique.

## 3 Results and Analysis

Due to the large amount of data to be processed, a representative subset from the full set of users was selected for each of the three categories SME, RES and OTH. All the models were first validated on these subsets and only the best



**Fig. 5.** MdAPE plots for all models with OTH users - (a) all seasons together (b) Autumn (c) Winter (d) Spring and (e) Summer.

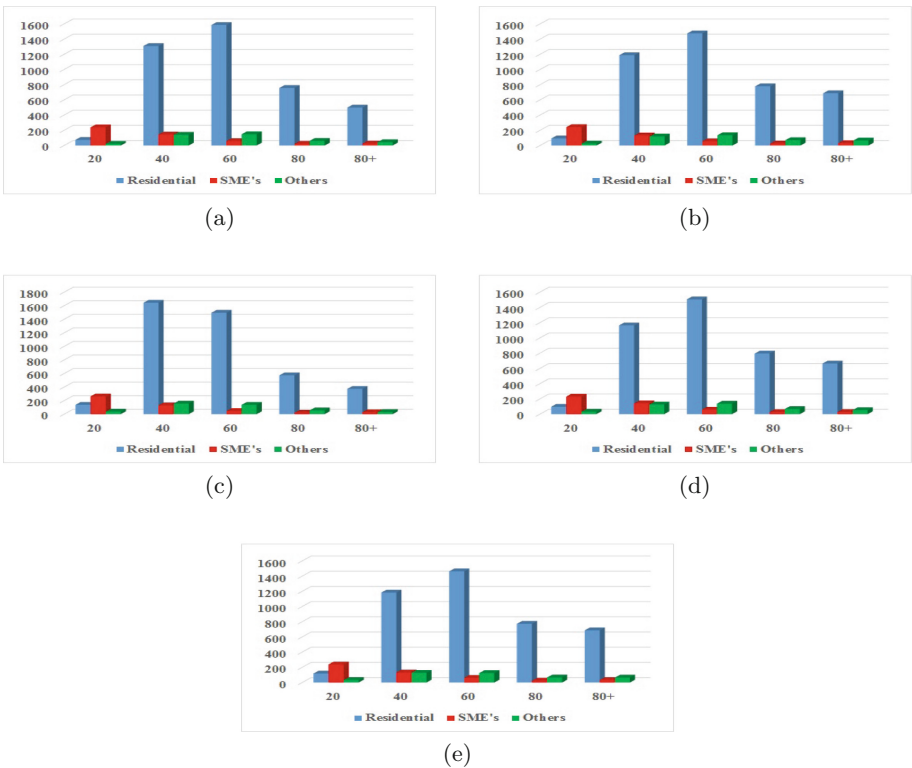
performing models were considered for evaluation on the complete datasets. The subset selection criterion, as proposed in [20] is given below:

$$\text{sample size} = \frac{(Z^2(p)(1 - p))}{C^2} \tag{3}$$

where:  $z = 1.96$  for 95% confidence level,  $p = 0.5$  (variability, maximum is 0.5) and  $c = 0.05$  (i.e. 5%) is the confidence interval.

$$\text{For Finite population sample size} = \frac{\text{samplesize}}{1 + \frac{\text{samplesize}-1}{\text{TotalPopulation}}} \tag{4}$$

The nominal sample size required is found to be 215 for ‘Small and Medium Enterprise users’, 352 for ‘residential’ users and 315 for ‘others’ with a confidence level as 95% and confidence interval as ‘5’. Performance metric considered for the present study is Median Absolute Percentage Error (MdAPE) [21]. MdAPE has also been used as the performance measure in [2].



**Fig. 6.** MdAPE Histogram for RES, SME and OTH consumers using HP-ELM 2-hour ahead forecasting system for (a) all seasons together (b) Autumn (c) Winter (d) Spring and (e) Summer.

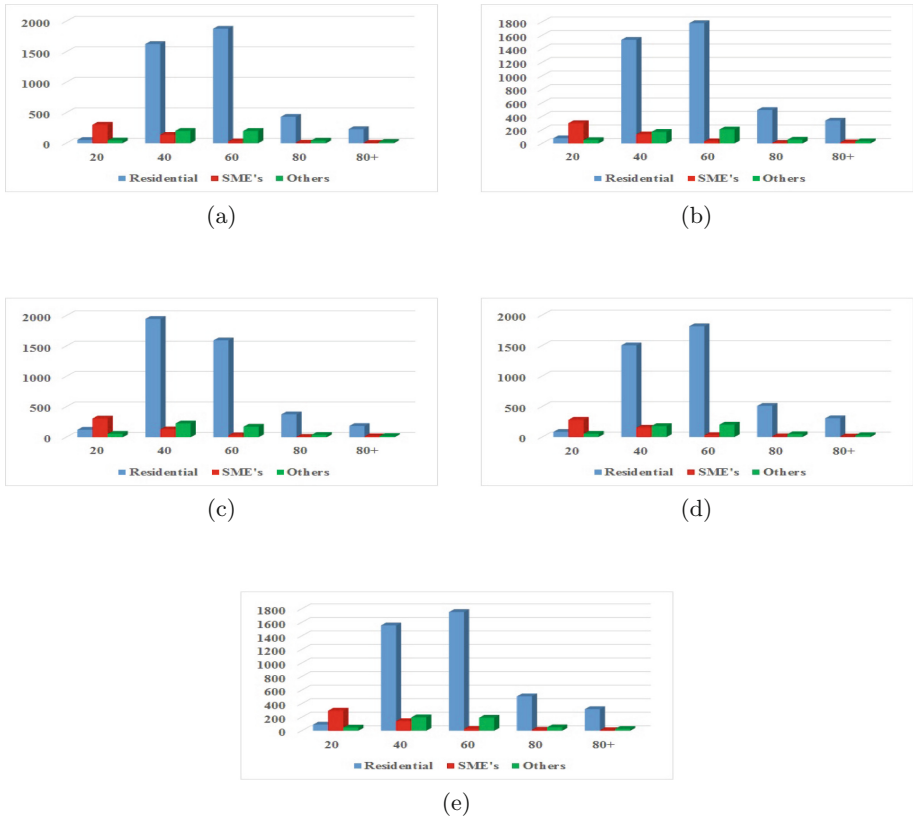
All the systems under consideration are validated for their 2, 4, 6 and 8 - hour ahead forecast capabilities. Since data is available at half hourly intervals, this translates to 4, 8, 12 and 16 step-ahead forecasts. Sliding window technique was used to re-train the forecasting systems with window length of 10. i.e.,

$$\hat{y}(t + 4) = f(y(t), y(t - 1), \dots, y(t - 9)) \tag{5}$$

where:  $\hat{y}(t+4)$  = estimated two hour ahead consumption,  $f(\cdot)$  = machine learning based forecasting technique and  $y(t)$  = consumption at time instant ‘t’.

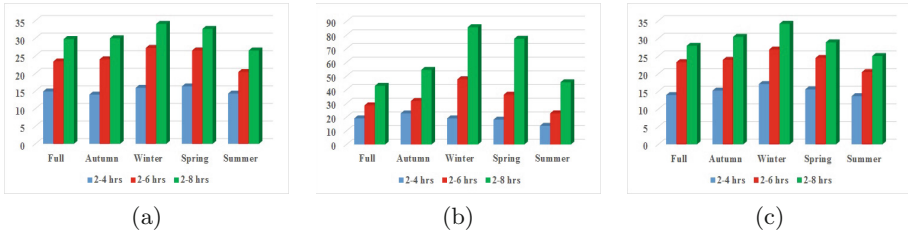
Forecasting results for all the eight models (1–8 in Figs. 3, 4 and 5) are presented in the form of stacked column charts in figures below. The five colors in the charts represent the MdAPE ranges 0–20%, between 20 and 40%, 40 to 60%, 60 to 80% and >80% (legends 20, 40, 60, 80 and 80+ in Figs. 3, 4 and 5).

It can be observed from the results presented in Figs. 3, 4 and 5 that ELM based models tend to outperform all the other models considered. Hence, the



**Fig. 7.** MdAPE Histogram for RES, SME and OTH category consumers using ELM (without using HP filter) 2-hour ahead forecasting system for (a) all seasons together (b) Autumn (c) Winter (d) Spring and (e) Summer.





**Fig. 8.** Percentage increase in MdAPE between different forecast periods (a–c) forecasting system trained using ELMs without splitting the time series using HP filter for Residential, SMEs and Others category users respectively

two ELM based models (HP-ELM and ELM without HP preprocessing) were evaluated on the entire set of 5275 users. The results are presented in Figs. 6 and 7. It can be observed from the Fig. 8 that the 2-hour ahead forecasts were consistently better than the longer horizon forecasts of 6 and 8 h-ahead. It was also observed that forecasting system trained with ELMs and without using HP filter produce better result with consistency than the rest of systems considered. Employing HP filter for separation of trend and cyclic components and trained using ELMs and ANN also generated better forecasts, albeit slightly worse than the without HP-ELM system discussed earlier. It was also evident that linear regression models were consistently worse off compared to the other three techniques considered.

### 4 Conclusion

Based on the results presented above, it can be concluded that shorter term horizon forecasts tend to be more accurate while employing the proposed forecasting systems. It is also observed that ELM based forecasting systems are able to generate better forecasts when compared to ANN or ERT based systems. Due to the forecasting performance exhibited by ELM based systems in the present study, ELMs can be considered to be good candidates for generating electricity consumption forecasts as a part of a larger AMI system.

### References

1. Siano, P.: Demand response and smart gridsA survey. *Renew. Sustain. Energy Rev.* **30**, 461–478 (2014)
2. Jayanth Balaji, A., Harish Ram, D.S., Nair, B.B.: Modeling of consumption data for forecasting in automated metering infrastructure (AMI) systems. In: Silhavy, R., Senkerik, R., Oplatkova, Z.K., Silhavy, P., Prokopova, Z. (eds.) *Automation Control Theory Perspectives in Intelligent Systems*. AISC, vol. 466, pp. 165–173. Springer, Cham (2016). doi:10.1007/978-3-319-33389-2\_16
3. Chan, S., et al.: Load/Price forecasting and managing demand response for smart grids: methodologies and challenges. *IEEE Sig. Process. Mag.* **29**(5), 68–85 (2012)

4. Zhao, H., Tang, Z.: The review of demand side management and load forecasting in smart grid. In: 2016 12th World Congress on Intelligent Control and Automation (WCICA) (2016)
5. Tasdighi, M., et al.: Residential microgrid scheduling based on smart meters data and temperature dependent thermal load modeling. *IEEE Trans. Smart Grid* **5**(1), 349–357 (2014)
6. Hong, T., et al.: Long term probabilistic load forecasting and normalization with hourly information. *IEEE Trans. Smart Grid* **5**(1), 456–462 (2014)
7. Kwac, J., et al.: Household energy consumption segmentation using hourly data. *IEEE Trans. Smart Grid* **5**(1), 420–430 (2014)
8. ISSDA. <http://www.ucd.ie/issda/data/commissionforenergyregulationcer/>
9. Hodrick, R., Prescott, E.: Postwar U.S. business cycles. In: *Real Business Cycles A Reader*, pp. 593–608 (1998)
10. Nair, B.B., et al.: Clustering stock price time series data to generate stock trading recommendations: an empirical study. *Expert Syst. Appl.* **70**, 20–36 (2017)
11. Nair, B.B., Mohandas, V.: An intelligent recommender system for stock trading. *Intell. Decis. Technol. IDT* **9**(3), 243–269 (2015)
12. Nair, B.B., Mohandas, V.: Artificial intelligence applications in financial forecasting—a survey and some empirical results. *Intell. Decis. Technol. IDT* **9**(2), 99–140 (2015)
13. Nair, B.B., et al.: A stock trading recommender system based on temporal association rule mining. *SAGE Open* **5**, 2 (2015)
14. Huang, G.-B., et al.: Extreme learning machine: a new learning scheme of feed-forward neural networks. In: 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)
15. Huang, G.-B., et al.: Extreme learning machine: theory and applications. *Neurocomputing* **70**(1–3), 489–501 (2006)
16. Huang, G.-B., et al.: Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **42**(2), 513–529 (2012)
17. Huang, G.-B.: An insight into extreme learning machines: random neurons, random features and kernels. *Cogn. Comput.* **6**(3), 376–390 (2014)
18. Specht, D.F.: The general regression neural network—Rediscovered. *Neural Netw.* **6**(7), 1033–1034 (1993)
19. Breiman, L.: Bagging predictors. *Mach. Learn.* **24**(2), 123–140 (1996)
20. Israel, G.D.: Determining sample size. University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS (1992)
21. Gooijer, J.G.D., Hyndman, R.J.: 25 years of time series forecasting. *Int. J. Forecast.* **22**(3), 443–473 (2006)