Scientia Iranica (2025) **32**(1): 6410

Sharif University of Technology
Scientia Iranica
Transactions on Computer Science & Engineering and Electrical Engineering
https://scientiairanica.sharif.edu

Load forecasting using two-level heterogeneous ensemble method for smart metered distribution system

Sneha Rai [a], Mala De [b, *]

a. *Department of Electrical Engineering, NIT Patna, Patna, Bihar, 800005, India, ORCID: 0000-0001-6621-9468.*

b. *Department of Electrical Engineering, NIT Patna, Patna, Bihar, 800005, India, ORCID: 0000-0002-3067-4976*

* Corresponding author: mala.de.power@gmail.com (M. De)

Received 17 January 2022; received in revised form 16 July 2022; accepted 30 January 2023

Keywords

Heterogeneous ensemble;
Load forecasting;
Machine learning;
Classical methods;
Smart metered data.

Abstract

A heterogeneous ensemble method for load forecasting (short-term and mid-term) are proposed here. The proposed approach comprises of a two-level hierarchy of machine learning based methods and classical methods to form the ensemble forecaster, where output of the first-stage forecasters are used as input in the second stage. Artificial Neural Network and Support Vector Regression methods are incorporated in the proposed approach as ML forecasters, whereas Holt's exponential smoothening and multiple linear regression techniques are included as classical forecasters. The proposed two-level ensemble approach forecasts realistic smart metered data more accurately and efficiently for multiple short-term and mid-term load forecasting scenarios with improved accuracy compared to any individual single stage forecasting methods. The prediction accuracy is shown to improve manifolds for the tested practical system. The proposed model also shows improvements compared to existing aensemble-based model.

1. Introduction

The advancement of smart electrical grids has resulted in large volumes of real-time load data. The availability of this real-time power consumption data collected from smart meters for different nodes of the system opens up new avenues for load forecasting. Maintenance scheduling, Demand-Side Management (DSM), [1], and other tasks that involve optimal planning and effective economic operation in the energy industry primarily depend on load forecasting.

Depending on the methodology used, there are three broad categories for load forecasting techniques; which are on the basis of: (i) utilization side, (ii) weather information and (iii) time horizon. For utilization side, there are two types of load forecasting (1) Utility-based forecasting, utilized in the management and planning of the energy sector, and (2) Consumer-based forecasting, useful in the optimization of energy [2]. There are numerous factors that effects load forecasting like weather factors Temperature-Humidity Index (THI), seasonal factors, past load on that particular node, day of the week, vacations and holidays, time of the day, etc [3]. According to the use of meteorological information, there are two forms of load forecasting namely (1) Univariate methods which do not need to know the weather information, and (2) Multivariate methods which need weather factors for load forecasting. The classification of load forecasting into four groups is again based on length of the forecast interval (1) Very Short-Term Load Forecasting (VSTLF), (2) Short-Term Load Forecasting (STLF), (3) Mid-Term Load Forecasting (MTLF) and (4) Long-Term Load Forecasting (LTLF) [4,5].

The energy sector has always been very interested in STLF and MTLF. For contingency analysis, hour-ahead bidding, DSM, stability studies, load flow analysis, management of ancillary services, reliability analysis, and for electric price calculation, etc., STLF is necessary [6]. Availability of the STLF plays a vital role in DSM by allowing utilities to design a proper pricing scheme so as to motivate customers to modify their demand. MTLF is usually applied for operations such as planning fuel reserves, unit commitment, maintenance scheduling, and optimal planning and decision making in energy sector [7].

To cite this article:

S. Rai and M. De "Load forecasting using two-level heterogeneous ensemble method for smart metered distribution system", *Scientia Iranica* (2025) **32**(1): 6410 https://doi.org/ 10.24200/sci.2023.59765.6410

The existing load forecasting approaches can be largely categorized as: Conventional or statistical methods including regression analysis [8,9], time series techniques such as Kalman filter [10], Autoregressive Integrated Moving Average (ARIMA) [11], exponential smoothing methods [12,13] and, Intelligent or Modern methods which includes Artificial Neural Network (ANN) [14,15], fuzzy theory [16], support vector machines [17], etc. The dependability of the selection of the many variables used for forecasting as well as the reliability of the past training data set have a considerable impact on the performance of these methods.

Load forecasting for individual nodes in any distribution network becomes more challenging compared to forecasting the sum of loads in a particular area as the variation of demand in individual nodes experience more variation with time. Demand in electrical utility is influenced by many factors like weather, time, special occasions, type of consumers, seasonal variation, etc. The accuracy of various forecasting techniques is usually limited by this significant load fluctuation and hence, the ensemble methods are attempted in [18-28] to improve the load forecasting performance further. In statistics and engineering, the concept of ensemble has come from the normal human nature to add facts from multiple resources and combine the opinions of many experts by a thought process to reach the ultimate result [21]. In recent years, ensemble methods are being widely used in the load forecasting area to improve overall performance and quality of the forecasting model. In [18,20], an exceptional load forecasting performance was obtained with the ensemble model of Wavelet Neural Network (WNN) and ANN as compared to the single load forecasting model used for STLF. The limitations of individual load forecasting methods can be overcome by combining them through various techniques like bagging, boosting [20], algebraic combiners [21], etc. Subsequently, a real-time load and price forecasting is proposed in [23] using a hybrid three-stage algorithm. In [24], the authors integrate the classical and modern ML methods for VSTLF. The works mentioned above show that the combination or ensemble of two methods improve the forecasting performance. Different ensemble methods are tried for the same. But there are limited number of works that uses heterogeneous ensemble methods [25,26] for load forecasting and none of these tried a combination of ML based and classical methods in the two different levels of ensemble. When the algorithms used in the two levels of an ensemble use different feature selection methods, then the performance improves [27]. The literature review has been summarized in Table 1 to have a clear view of references according to their approach to load forecasting.

The work proposed here is motivated by numerous advantages of application of ensemble methods in the load forecasting area. The paper mainly contributes in the development of an ensemble model for load forecasting by combining different types of forecasters in two-levels which will successfully work for both STLF and MTLF with real-time smart meter data.

This work proposes a two-stage heterogeneous ensemble model for both STLF and MTLF and tested for the residential cum academic campus of NIT Patna. Here, ML based ANN and Support Vector Regression (SVR) and classical Multiple Linear Regression (MLR) works at the first level and these are combined with classical Holt's double exponential smoothening method in the second level of the ensemble model. The detailed description of the model is given in the next section.

This paper is organized as follows: The proposed heterogeneous ensemble load forecasting method is presented in Section 2. Section 3 describes dataset of the practical system. Section 4 presents results and discussion. Finally, Section 5 concludes the paper.

2. The heterogeneous ensemble forecasting

This paper presents a two-stage heterogeneous ensemble models for STLF and MTLF. An ensemble is an amalgamation of multiple algorithms that work together to improve the prediction performance that no one algorithm could obtain individually [25]. An ensemble predictor models can be assembled in following ways by varying the choice of models used in ensemble or by varying the way the outcomes of individual predictors are combined in the ensemble model. In the simplest form of ensemble model the final prediction is averaged or the weighted sum of the outcomes of the individual predictors. In weighted sum approach, the better performing individual model gets higher weight compared to the weaker ones and hence it improves the accuracy of ensemble predictor compared to averaged ensemble model. The prediction accuracy can be improved further by stacking [27]. In stacking, the outputs of individual predictors are taken as input to the second level algorithm which produces the final output prediction. So, this will be two-level process and as shown pictorially in Figure 1.

The first level uses different classical and modern predictors to produce the forecast for all the datasets used in the analysis. However, second level uses the output of first level as input to produce the final forecasted output with reduced error and better accuracy as compared to the result produced by the first level predictors. First level or base level trains using complete training dataset and second level or meta-model takes the output of first level as input and makes final prediction. The first level often uses different predictor and therefore stacking ensembles are often called heterogeneous ensembles. The two predictors in these two levels are chosen in such a way that they use different feature selection methods using same training data.

In this paper we have proposed a two-level heterogeneous ensemble method using ML based predictor ANN, SVR and classical predictor MLR at base level and classical predictor Holt's in the second level. The Holt's exponential smoothening method is a univariate time series forecasting technique that performs direct load forecasts by using past load data as model input and is unaffected by meteorological conditions. The added advantage which Holt's method carries is the Holt's coefficients which represent level, trend and seasonal factor of load and can be optimized to achieve best forecasting performance which improves the model accuracy. Therefore, Holt's method works in a completely different approach for prediction and is a natural choice of second level predictor for heterogeneous stacked ensemble as it fulfills the basic criterion of stacking different types of predictors in two-level to achieve performance improvement. This motivates use of Holt's method with ANN, SVR and MLR techniques in this work.

The proposed ensemble is tested to forecast for multiple intervals ranging from STLF and MTLF scenarios. For all these cases the prediction error has reduced to a very small value in case of the proposed two-level ensemble.

Table 1. Summary of literature.

References/ author	Year	Contribution
Hagan and Behr [9]	1987	While load forecasting is a difficult task at any level and over any time horizon, it is particularly challenging when using fine-grained data to anticipate load at the home level. Discusses the STLF using time-series method.
Moghram and Rahman [7]	1989	Five short-term load forecasting techniques have been discussed in this paper
Alex and Timothy [8]	1990	A STLF technique based on regression approach is developed in this paper.
Polikar [21]	2006	The analysis of ensemble-based systems in decision making is done.
Filho et al. [14]	2011	A multi nodal load forecasting model is developed using general regression neural network.
Ceperic et al. [17]	2013	A strategy for short-term load forecasting is developed using support vector machine.
Zhang et al. [16]	2017	Singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm are used to develop a STLF model.
Li et al. [19]	2015	On the basis of clustering techniques, an intelligent short-term load forecasting using ANN, WNN, and KF hybrid models is proposed for the smart grid.
Hendawia and Wang [20]	2020	An ensemble method of full wavelet packet transforms and neural network is developed for STLF.
Khwajaa et al. [22]	2020	Joint bagged-boosted artificial neural networks have been analyzed using ensemble ML to improve the short-term electricity load forecasting performance.
Nazar et al. [23]	2018	Hybrid model using three-stage algorithm is developed for simultaneous load and price forecasting.
Palaninathan et al. [25] and Dudek et al [26]	2016	Discusses the heterogeneous ensemble load forecasting method.
Lee and Cho [28]	2021	Compares the performance of traditional, ML, and hybrid model for electricity peak load forecasting.
Rai and De [29]	2021	Discusses the analysis of classical and machine learning based methods for STLF and MTLF of micro grid.

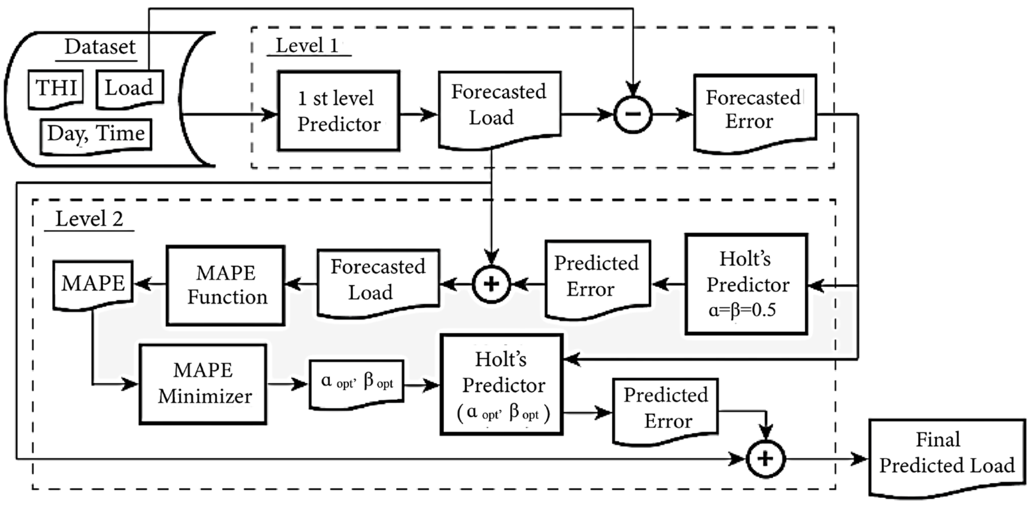

Figure 1. Two-level heterogeneous ensemble model structure.

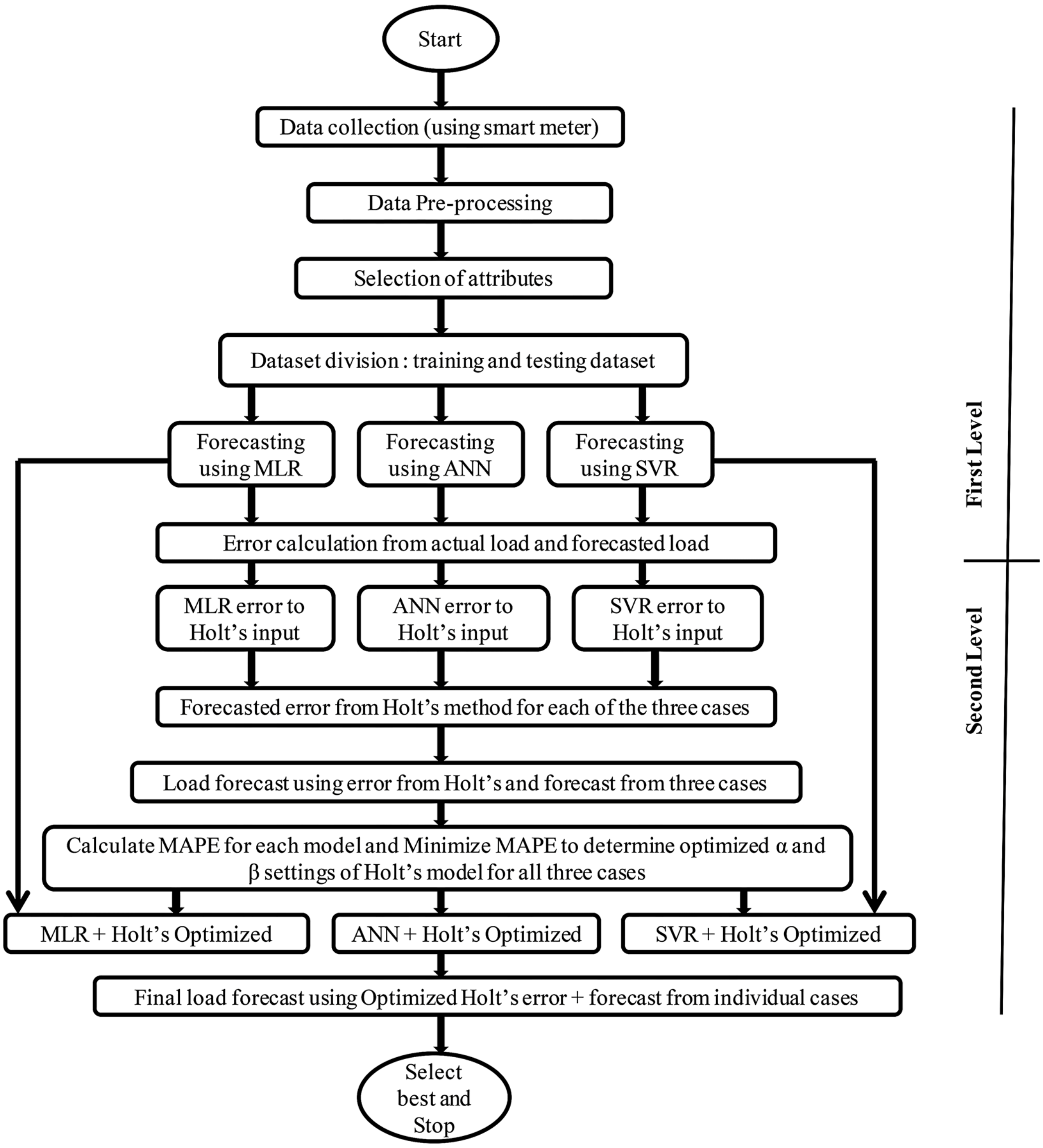

Figure 2. Block diagram of the forecasting procedure.

2.1. Methodology

Figure 2 shows the entire forecasting procedure used in this work. At various nodes, smart metres are used to collect the load data. Using a Gaussian Filter (GF), the raw data from these metres is first pre-processed to filter out outliers and normalise the data to fall within the range [0, 1].

A typical distribution system's load is significantly influenced by changes in time and day, and the weekend loads differ from weekday loads. As a result, the load forecasting model considers these qualities to be the key attributes. To determine the relationship between independent factors (weather data i.e., temperature and humidity index, THI) and the dependent variable (load), a correlation analysis is conducted. The proposed two-stage heterogeneous ensemble method, developed by the combination of MLR, ANN, and SVR models with Holt's method, is applied to forecast the loads at two different nodes of practical grid. The attributes applied to train the individual models are weather variables, time, day factor, and the output is load in kW. The error calculated by comparing real and the forecasted load from each of the individual methods is fed as an input to the Holt's model to predict the output. The final forecasted load is calculated by adding individual output of MLR, ANN, and SVR with predicted output of Holt's model. The MAPE are calculated from actual load and forecasted load of corresponding ensemble methods. Finally, the parameters of Holt's methods are optimized using a non-linear solver to achieve the best forecasting performance i.e., to minimize the MAPE error. The final forecast is calculated using error of optimized Holt's method and forecast from first level. Lastly, RMSE and MAPE are calculated and compared with different individual and ensemble models to prove its superiority. These individual forecasting methods [29] are described below.

2.1.1. MLR

The relationship between independent variable, e.g., day, time, weather elements like temperature, humidity, etc., and the dependent variable, such as load, is described by MLR. The independent and dependent variables are described with the following relationship given in Eq. (1):

$$\tilde{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots.. + \beta_n x_n + \varepsilon, \quad (1)$$

where, $\tilde{y}$ is dependent variable, x_1, x_2,….x_n are independent variables or input variables, β is regression coefficient and ε is the error term. For multiple observations, it can be written as:

$$\tilde{y}_1 = \beta_0 + \beta_1 x11 + \beta_2 x12 + \ldots\ldots.. + \beta_k x1k + \varepsilon,$$

$$\tilde{y}_2 = \beta_0 + \beta_1 x21 + \beta_2 x22 + \ldots\ldots.. + \beta_k x2k + \varepsilon,$$

$$\tilde{y}_n = \beta_0 + \beta_1 xn1 + \beta_2 xn2 + \ldots\ldots.. + \beta_k xnk + \varepsilon.$$

Eq. (2) describes the equation in general form.

$$\tilde{y}_i = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \varepsilon. \quad (2)$$

Developing a function that strongly links these attributes is the fundamental objective of MLR in order to forecast the dependent variable from the given independent variables.

2.1.2. ANN

ANN is powerful technique involving computation and machine learning for predicting the load; it has the ability of parallel processing. In this study, artificial neurons are activated using a sigmoid function. One hidden layer of ANN is trained using Bayesian Regularization (BR) algorithm with a range of 1 to 10 neurons. BR is chosen here to train ANN as it gives the best training performance for highly non-linear data with randomness.

2.1.3. Holt's exponential smoothening method

Holt's double exponential smoothening is the most suitable technique for forecasting data with trend. The technique is based on two smoothening equations, first for level component and second for the trend factor, both of them can be expressed as:

Level:

$$A_t = \alpha \times y_t + (1-\alpha) \times (A_{t-1} + T_{t-1}). \quad (3)$$

Trend:

$$T_t = \beta \times (A_t - A_{t-1}) + (1-\beta) \times T_{t-1}. \quad (4)$$

The complete forecast equation is represented as:

Forecast:

$$F_{t+m} = A_t + T_t \times m. \quad (5)$$

By using the following equation, the level and trend component's initial values can be determined:

$$A_1 = Y_1, \quad T_1 = \left((y_2 - y_1) + (y_4 - y_3)\right) \div 2, \quad (6)$$

where α and β are smoothening constants, A_t is level factor, T_t is trend component, y_t is the actual load, t is time period, F_{t+m} is the m periods ahead forecasted load.

2.1.4. SVR

A powerful machine learning procedure for regression analysis is the SVR, which is based on supervised learning method. SVR is a non-parametric method because it is kernel function-based. In a very high dimensional feature space, the dot product of two vectors is computed using a kernel. The kernel can be linear, polynomial, Radial Basis Function (RBF), or Gaussian. The model in this case was trained using a polynomial kernel. The polynomial kernel is mathematically expressed as:

$$K(x, y) = \left(c + x^T y\right)^n, \quad (7)$$

where c represents the constant term and n is the degree of polynomial.

2.2. Performance criteria

The MAPE and RMSE, which are common metrics used in the field of load forecasting, are used to measure the forecast accuracy. MAPE is the average multiplicative effect between each estimated mean and the observed output. The standard deviation of the residuals is denoted as RMSE. It indicates how tightly the data is concentrated around the line of maximum fit. It is calculated as:

$$MAPE = \frac{1}{N} \left| \sum_{i=1}^{N} \frac{\left(y_i - \tilde{y}_i\right)}{y_i} \right| \times 100, \quad (8)$$

where y_i represents real load and $\tilde{y_i}$ is forecasted load, N gives the number of samples used for testing. The RMSE between real and forecasted load is calculated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(y_i - \tilde{y}_i\right)^2}{N}}. \quad (9)$$

3. The practical system description

The dataset used to test proposed load forecasting method is described in this section. In addition, correlation analysis and data pre-processing method are also discussed here.

3.1. Dataset description

Inside the NIT Patna campus, 15 distinct nodes have smart meters installed, which log data on power usage in real time. Different load data metrics, such as voltage, current, active power, reactive power, phase angle, and frequency, can be obtained at any node at the required time interval, which can range from one minute to several hours. The practical data used for analysis consists of several datasets, viz. (a) six months daily average load at two distinct nodes, (b) load per hour at a node, (c) load at a definite time (10 a.m.), and (d) daily average load of weekend for period of six months (15th July 2018 to 15th January 2019). These datasets are used for forecasting the loads for multiple time horizons. The metrological department provides the necessary weather information, including THI, which have a substantial impact on the load.

3.2. Correlation analysis

To examine the relationship between the dependent variable and the independent variables, a correlation analysis is performed. Table 2 shows the correlation of weather parameters with the load data. The correlation coefficient of

Table 2. Weather correlation with load data.

	Temperature	Humidity	Girl's hostel load	Incomer transformer load
Temperature	1			
Humidity	0.935707	1		
Girl's Hostel load	0.811387	0.960388	1	
Incomer transformer load	0.833228	0.914292	0.912617	1

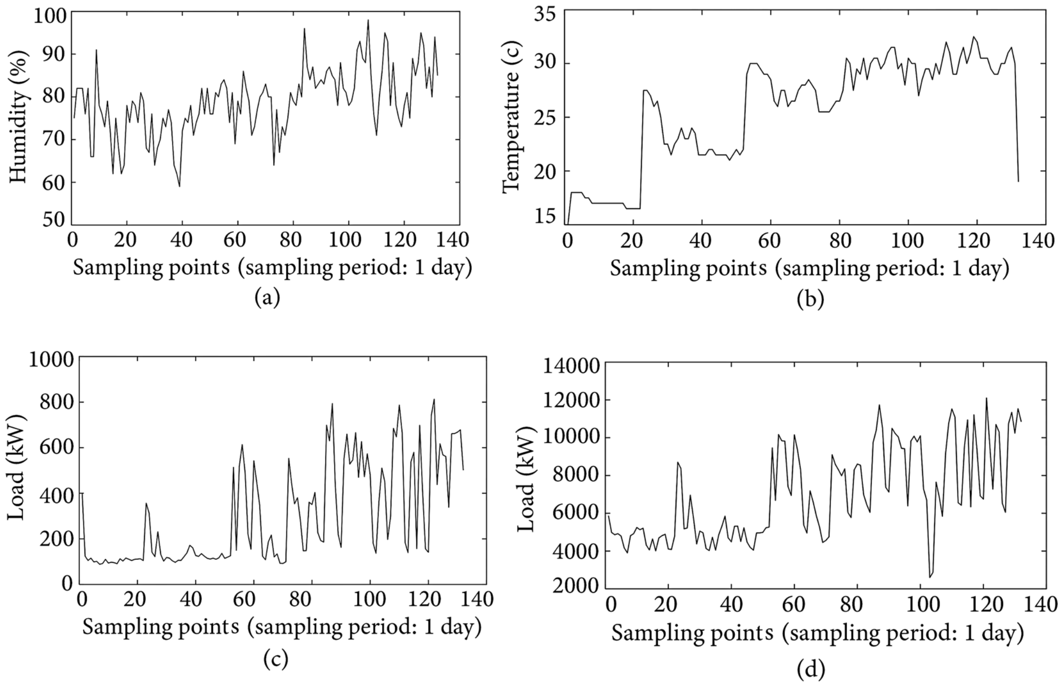

Figure 3. (a) Humidity, (b) temperature, (c), and (d) daily load curves of two nodes under consideration for six months.

Girl's Hostel load and that of Incomer transformer load with temperature and humidity is shown here.

The calculated correlation coefficients show a significant relation between load and both temperature and humidity, with very high correlation coefficient values in both cases. This signifies that both weather factors have considerable effect on the load. Also, there is high correlation among both the loads. This can also be verified from Figure 3 which illustrates how the load closely follows the temperature and humidity patterns at both nodes.

3.3. Data pre-processing

Smart meter data that is directly obtained has several anomalies like noise, incorrect or duplicate data, communication link issues, etc. To obtain precise and reliable forecasts, these inaccurate data should be filtered out. The raw data is normalized and filtered using a GF. It is a type of low pass filter whose efficiency is monitored by changing the window size. The main purpose of this filter is to eliminate noise from time series data. The window first determines the data points mean before moving one level down and repeating the calculation. To achieve the highest predicting performance, the length of window is chosen after iterative estimation; in this case, the length is 2. The processed data is used to train the load forecasting model.

The load curve per hour of GH node is shown in Figures 4 and 5 before to and following data pre-processing. Figure4 shows a very wide variation of load.

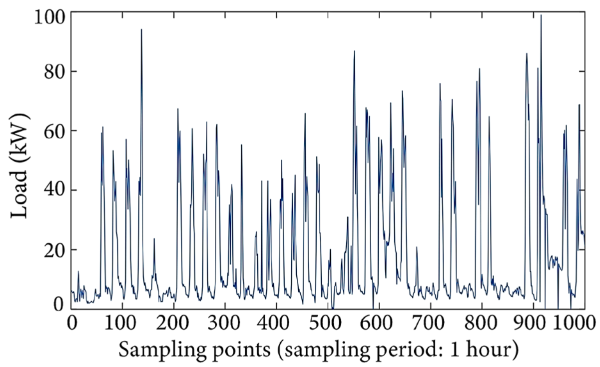

Figure 4. Hourly load (GH node).

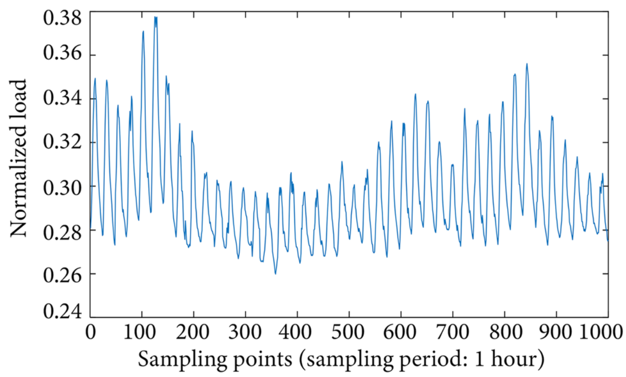

Figure 5. Hourly Load after pre-processing (GH node).

If this data is used to train the forecasters, then the lower range load values will have less effect compared to the higher values and the training will not be efficient and effective. Figure 5 demonstrates how the pre-processing retained the data trend while keeping the magnitude variation within the range [0, 1]. This allows for better predicting and training.

4. Results and discussion

The six distinct scenarios listed below are tested using the STLF and MTLF methodologies detailed in Section 2:

- Case I - Daily LF using six months data for GH node;
- Case II - Daily LF with six months data for IT node;
- Case III - Weekly LF using weekend data for GH node;
- Case IV - Weekly LF using weekend data for IT node;
- Case V - Forecasting load at a specific time (10.a.m.) with three month's data for GH node.
- Case VI - Hourly LF with six months hourly data for GH node.

The forecast models are tested with the NIT Patna campus's smart metered data. As the load pattern for an academic campus varies widely between weekdays (case I, II) and weekends (Case III, IV) so different scenarios are considered to forecast these two cases and also for each case two different node data are shown to establish that the method works for every time of node. In the same way hourly load forecasting has also been done to analyze the viability of the proposed methodology with different datasets with variable sampling period. In all the scenarios, the total dataset is partitioned in two different sets: training and testing dataset. In this case, training uses 80% of the whole dataset, while testing and validating the model uses the remaining 20%. Three different ensemble models: MLR+ Holt's, ANN+ Holt's and SVR+ Holt's are proposed for load forecasting along with the three individual (independent) prediction methods.

4.1. Load forecasting by individual forecasting techniques

In this section, the forecasting results using the individual load forecasting methods for all the six testing scenarios are presented. Table 3 shows the results of the MLR approach for various input datasets. From the table it is observed for Cases I, III, VI the calculated MAPE is below 5% which is within acceptable range but for the remaining three cases, the MAPE is more the 5%. This is because MLR is a linear regression method and the variation of load is primarily nonlinear in structure. So, the results improve when the non-linear regression methods like ANN and SVR are applied. However, the performance of the MLR is worth mentioning in Case VI where the hourly data is forecasted. In this case the MAPE is only 0.52% with an impressive RMSE value.

The feed forward neural network is used with back propagation training for the load forecasting models. The ANN has three layers: an input, a hidden, and an output. It is trained by BR algorithm as it is more suited for high non-linear data with randomness. For all cases, performance of the ANN based prediction model obtained by altering the hidden layer neurons number between 1 to 10 is shown in Table 3. The results clearly indicate that both performance parameters (MAPE and RMSE) improved significantly in case of ANN. The maximum MAPE obtained in most of the cases are comparatively less than that of MLR. Even in Case VI, the ANN outperformed the MLR and the MAPE has reduced to 1/6th value with 10 neurons. It may be noted that the results vary significantly with the variation in number of neurons and each model performs best for a particular neuron count; but overall, the forecasting with ANN is better than MLR.

Table 3. Load forecast using different forecasting methods.

SI. No.	Cases	MAPE	RMSE
1	Case I	3.64	0.0297
2	Case II	6.16	0.0607
3	Case III	4.42	0.0307
4	Case IV	5.35	0.0452
5	Case V	7.13	0.055
6	Case VI	0.52	0.0018

ANN

SI. No.	Cases	No. of neurons	MAPE	RMSE
1	Case I	2	2.85	0.0245
		5	4.39	0.0378
		10	4.51	0.042
2	Case II	2	3.91	0.048
		5	3.10	0.029
		10	3.16	0.028
3	Case III	2	4.55	0.038
		5	5.83	0.048
		10	5.09	0.047
4	Case IV	2	4.35	0.034
		5	5.63	0.043
		10	5.07	0.047
5	Case V	2	2.66	0.012
		5	2.32	0.012
		10	2.75	0.014
6	Case VI	2	0.18	0.0023
		5	0.24	0.0037
		10	0.09	0.0021

SVR

SI. No.	Cases	Order	MAPE	RMSE
1	Case I	1	3.79	0.0307
		2	3.83	0.0299
		5	4.84	0.0399
2	Case II	1	3.15	0.0288
		2	3.65	0.0390
		5	4.73	0.0691
3	Case III	1	4.02	0.0283
		2	6.07	0.0741
		5	3.76	0.0289
4	Case IV	1	4.02	0.0283
		2	6.07	0.0741
		5	3.76	0.0289
5	Case V	1	7.66	0.0329
		2	3.44	0.016
		5	2.72	0.013
6	Case VI	1	2.81	0.0031
		2	1.60	0.0018
		5	5.43	0.0070

In the context of several training and testing datasets, the load at two given nodes is forecasted using SVR using a polynomial kernel. In order to do linear separation, kernel function turns the data into a feature space with more dimensions. As the polynomial's order is changed here from 1 to 5, the corresponding RMSE and MAPE values are calculated and displayed in Table 3. Being a non-linear regressor, the SVR performs more accurate forecasting than that of MLR. Also, from Table 3, the performance of ANN seems to be superior than that of SVR, but it cannot be used for generalization purpose since the results obtained by ANN is inconsistent for a given network architecture in multiple runs. However, the average result of the multiple runs has been reported in this work.

The above-mentioned methods include the linear and the non-linear regression methods in which input features like time, day and weather factors are used to train the forecasting model, in addition to historical load data. On the other hand, the Holt's approach is a traditional time series forecasting technique which is univariate in nature. This means, the method is independent of the input attributes as it is trained only by the past load of the corresponding dataset to predict future load. Due to the univariate nature of Holt's method, it is combined with the other linear and non-linear methods to form an ensemble method for STLF and MTLF. This is done with a motive to explore the univariate nature of Holt's method and with the same time utilize the capability of other linear and non-linear regressors for better load forecasting. The results of three different combinations of ensemble models proposed in this paper are presented in Section 4.3.

4.2. Optimized Holt's method

The second level of the heterogeneous ensemble method employs Holt's exponential smoothening method having two parts: a level component and a trend component for forecasting data with a trend. The forecast is done using two smoothening coefficients α and β as parameters and the three equations Eqs. (3)-(5) mentioned in Sunsection 2.1.3. The final forecast's MAPE and RMSE values are derived using the level and trend coefficients, which are initially set at 0.5. The parameters and are tuned to obtain the minimal MAPE value in order to enhance forecasting performance. Below is the objective function.

$$\mathit{Minimize\ MAPE}\left(\alpha,\beta\right)=\frac{1}{N}\left|\sum_{i=1}^{N}\frac{\left(y_i-\tilde{y_i}\right)}{y_i}\right|\times 100, \tag{10}$$

where y_i is real load and $\tilde{y_i}$ is predicted load which can be found by adding the forecasted load from first level with error output of Holt's method, N is the total sample number used for testing. In this work, the GRG non-linear solver is applied for optimization. The coefficients α and β lie between 0 and 1. The solver first determines MAPE by setting α and β to 0.5 (initial set value). The GRG non-linear solver determines best solution when the partial derivatives equal zero and determines the objective function's slope as the input value varies. Finally, it determines the smoothening coefficients' ideal value for minimizing the MAPE.

4.3. Forecasting by proposed ensemble models

The results obtained so far show that the performance of non-linear methods like ANN and SVR is better than the linear regression methods like MLR since the load is non-linear and dynamically varying in nature. Still there is a scope for improvement in the efficacy of the load forecasting procedures used for STLF and MTLF. Therefore, this section discusses the performance of the load forecasting of the ensemble of MLR, ANN and SVR with the Holt's method explained in Section 2.1. The forecasting performance of the ensemble of MLR+ Holt's model is mentioned in Table 4. It can be observed that the MAPE obtained by the combination of MLR+ Holt's is much lower than that of individual MLR model used for load prediction. Additionally, analysis shows that the model's performance in terms of MAPE and RMSE has greatly improved with optimized Holt's model parameters (α and β) for each input scenarios. Especially in Case VI, the errors are almost negligible when MLR is ensemble with Holt's method. This is an indication that ensemble method works efficiently even with the linear regressor like MLR.

The forecasting results of the ANN+ Holt's model is given in Table 5. This table concludes that RMSE and MAPE for each case with this model are lower than that obtained with MLR+ Holt's model. MAPE is less than 1% in four out of six input scenarios which indicates consistency of the ANN+ Holt's model in load forecasting. The error values obtained by the ensemble of SVR+ Holt's model as shown in Table 6 are the lowest among all three ensemble models for every variation of input datasets. Therefore, in comparison with MLR+ Holt's model and ANN+ Holt's model, a definite improvement in prediction performance has been observed here. From the results, it can be seen that after parameter optimization of SVR+ Holt's technique, the MAPE ranges between 0–0.5% in all the cases except Cases III and IV, in which the MAPE lies around 3%.

4.4. Comparison of various forecasting techniques used

Table 7 compares the performance of load forecasting using the individual existing approaches and the proposed ensemble methods. The results show that there is an obvious advantage of the proposed heterogeneous ensemble methods over the individual models. Among the ensemble methods, MAPE for SVR+ Holt's model is least in four test cases (Cases I-V) as compared to other ensemble models. Therefore, the comparison analysis reveals that the proposed SVR+ Holt's model based ensemble method gives a best performance for both STLF and MTLF over others. However, RNN trains itself using its own forecasted output to make predictions in the future, hence it has recently proven successful as a method for time series forecasting. For each of the six cases, the MAPE generated by the proposed model is compared to that obtained by RNN to determine whether the proposed method is superior as depicted in Table 7. The obtained results demonstrate that the ensemble of SVR+Holt's technique produced better forecasts than RNN. This may be due to inaccuracies in the actual or raw load data collected from the smart meters. The study further reveals that these regression techniques rely greatly on the nature of recorded data.

Table 4. Forecasting using MLR+ Holt's.

SI. No.	Before optimization of Holt's coefficient				After optimization of Holt's coefficient			
	α	β	MAPE	RMSE	α	β	MAPE	RMSE
1	0.5	0.5	2.92	0.023	0.41	0.36	2.42	0.020
2	0.5	0.5	4.40	0.047	0.38	0.79	4.19	0.046
3	0.5	0.5	6.97	0.045	0.68	1	3.36	0.024
4	0.5	0.5	3.84	0.033	0.80	0.72	1.98	0.018
5	0.5	0.5	4.58	0.020	0.93	0.21	1.76	0.0089
6	0.5	0.5	0.0061	1.6E-05	0.99	0.018	0.000508	1.5E-06

Table 5. Forecasting using ANN+ Holt's.

SI. No	Before optimization of holt's coefficient				After optimization of holt's coefficient			
	α	β	MAPE	RMSE	α	β	MAPE	RMSE
1	0.5	0.5	1.99	0.016	0.93	0.15	0.50	0.005
2	0.5	0.5	1.65	0.015	0.97	0.09	0.29	0.003
3	0.5	0.5	7.87	0.048	0.64	1	3.29	0.028
4	0.5	0.5	6.42	0.053	1	0.60	3.45	0.032
5	0.5	0.5	1.90	0.0084	0.96	0.11	0.48	0.002
6	0.5	0.5	0.015	2.4E-05	0.99	0.006	6.29E-04	1.2E-06

Table 6. Load forecasting with SVR+ Holt's model

SI. No	Before optimization of holt's coefficient				After optimization of holt's coefficient			
	α	β	MAPE	RMSE	α	β	MAPE	RMSE
1	0.5	0.5	2.20	0.019	0.99	0.007	0.025	0.0002
2	0.5	0.5	1.83	0.018	0.99	0.015	0.043	0.0004
3	0.5	0.5	6.49	0.042	0.66	1	2.95	0.022
4	0.5	0.5	6.27	0.051	1	0.57	3.13	0.030
5	0.5	0.5	1.58	0.007	0.92	0.13	0.47	0.0023
6	0.5	0.5	0.3	5.2E-04	0.99	0.002	0.0052	0.000009

Table 7. MAPE for different forecasting models.

SI. No	Input scenarios	Individual models				Ensemble Models		
		MLR	ANN	SVR	RNN [29]	MLR + Holt's (opt)	ANN + Holt's (opt)	SVR + Holt's (opt)
1.	Case I	3.64	2.85	3.79	12.49	2.42	0.50	0.025
2.	Case II	6.16	3.39	3.15	8.03	4.19	0.29	0.043
3.	Case III	4.42	4.55	4.02	17.20	3.36	3.29	2.95
4.	Case IV	5.35	4.35	4.02	5.36	1.98	3.45	3.13
5.	Case V	7.13	2.32	2.72	6.37	1.76	0.48	0.47
6.	Case VI	0.52	0.18	1.60	15.20	0.000580	0.000629	0.00525

Figures 6 to 11 displays the graph comparing the forecasted and real load for various datasets using various ensemble approaches. These plots indicate whether the load predictions are able to follow the trend of actual load. For example, the relatively low performance of the MLR+ Holt's method in clearly indicated by the outlier red line in Figures 7, 9, and 10. Similarly, the dotted green lines for SVR+ Holt's method least deviates from the actual load with an exception in Figure 9 (Case IV) due to the reason aforementioned. In Figure 11, there is not much to difference visually between actual load and prediction as negligible error was obtained in each case.

The box plot depicted in Figure 12 can also be used to visualize the comparative MAPE values derived by various load forecasting models. The MAPEs for each of the distinct input scenarios employing multiple individual and ensemble forecasting models are displayed in the plot's boxes. The median of MAPE range is shown by the red line inside each box. The MAPE for ANN+ Holt's and SVR+ Holt's models is minimal and within a similar range, as seen in box plot.

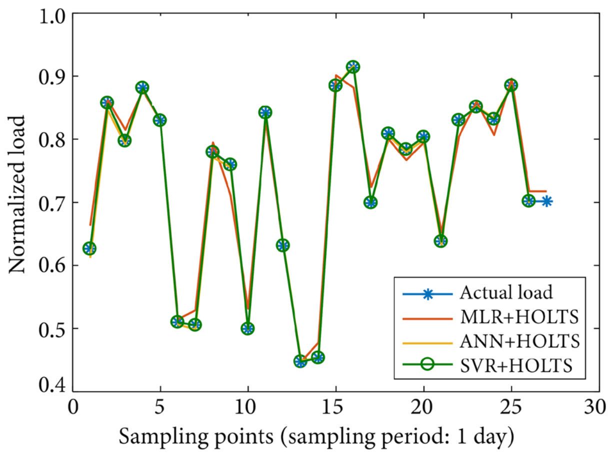

Figure 6. Six months daily average load comparison for GH (Case I).

The average MAPE calculated by ANN+ Holt's model for all the given cases is 1.33, while the average MAPE calculated by SVR+ Holt's model is 1.10.

This signifies that the SVR+ Holt's ensemble model has the lowest MAPE and, consequently, the best performance for STLF and MTLF in all the various test scenarios.

To establish the accuracy of proposed ensemble models with existing models a comparison with the wavelet-ANN based ensemble method proposed in [19] is presented in Table 8. This table presents comparison of the proposed ensemble models with existing ANN based ensemble model and proves its superiority as compared to the existing wavelet-ANN ensemble model [19] in terms of RMSE value for daily load data.

Table 8. Comparison of MAPE values of proposed ensemble models with existing ensemble models.

The ensemble model	MAPE value	
	Min	Max
Full wavelet packet transform + ANN [Ref 19]	1.03	2.14
Proposed ANN + Holt's (opt)	0.000629	0.50
Proposed SVR + Holt's (opt)	0.00525	0.47

5. Conclusion

This paper proposes a two-level heterogeneous ensemble model that can be used for Short-Term Load Forecasting (STLF) and Mid-Term Load Forecasting (MTLF). Support Vector Regression (SVR) and Holt's exponential smoothening method are the two regression models that are combined in this method.

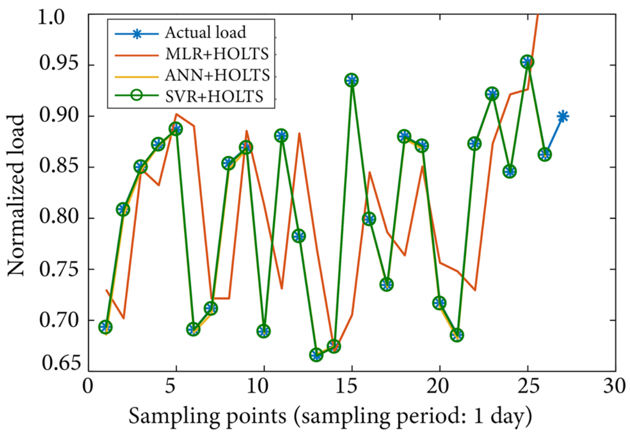

Figure 7. Six months daily average load comparison for IT (Case II).

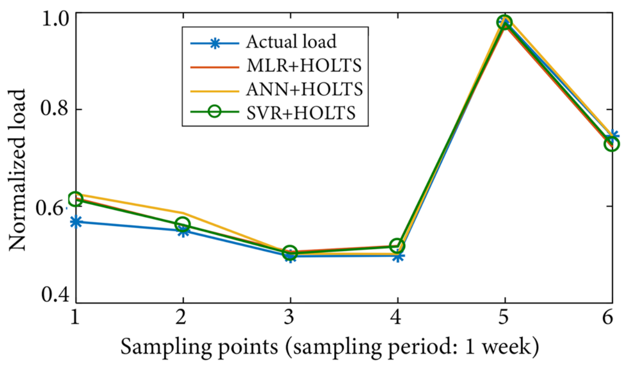

Figure 8. Weekend load comparison for (Case III).

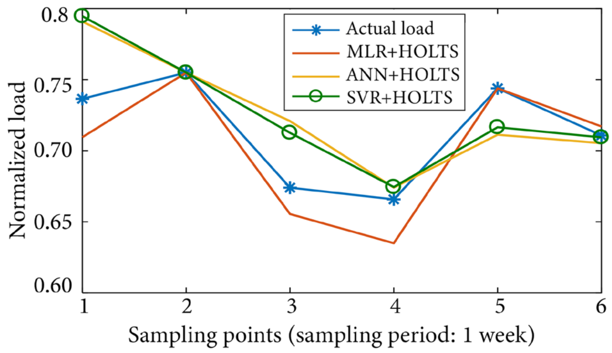

Figure 9. Weekend load comparison for IT (Case IV).

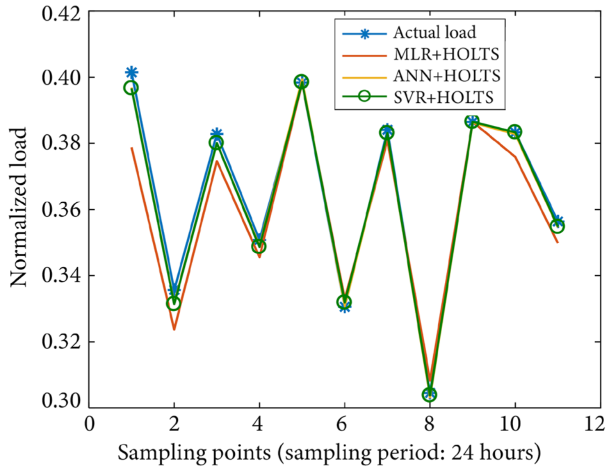

Figure 10. Hourly load comparison at 10 am for GH (Case V).

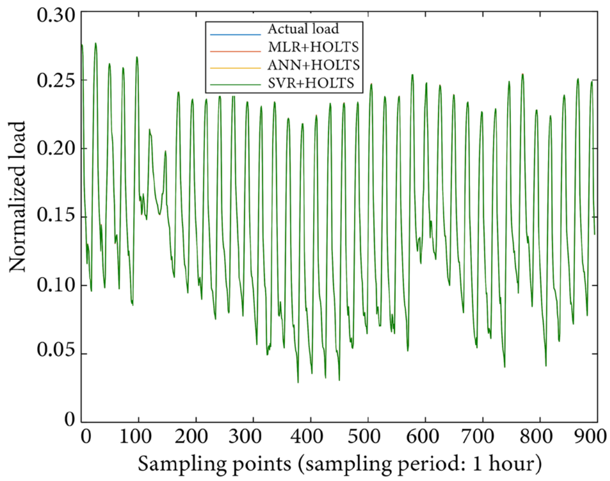

Figure 11. Six months hourly load comparison for GH (Case VI).

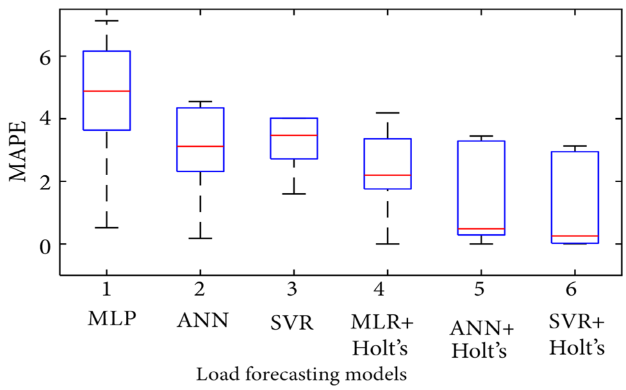

Figure 12. Box plot of MAPE values by various models.

The SVR is trained by time, day, weather factors Temperature-Humidity Index (THI) and past load as input to predict future load of the system and used in the first level. Holt's technique is a conventional time-series univariate forecasting method which only requires past load data for future load forecasting and used in second level. To achieve the optimum load forecasting performance, the coefficients of the Holt's approach are further tuned. Two other ensemble models are developed MLR+ Holt's and ANN+ Holt's to compare forecasting performance of the proposed method. The broad comparison studies using dataset availed at the NIT Patna campus demonstrates the applicability of proposed method. In comparison to various independent linear and non-linear regression models like Multiple Linear Regression (MLR), ANN, SVR and ensemble models like MLR+ Holt's and ANN+ Holt's, it was discovered that the suggested SVR+ Holt's model offers the best forecasting results for both STLF and MTLF, that can also be confirmed from the least MAPE and RMSE values. The Holt's +ANN and SVR+ANN based models show its superiority compared to existing- wavelet +ANN ensemble methods also. The same study can be expanded to a LTLF approach, but this is not done because the dataset employed for the analysis lacks long-term training data.

Abbreviatons/Symbols	**Description**
DSM	Demand-Side Management
THI	Temperature Humidity Index
VSTLF	Very Short-Term Load Forecasting
STLF	Short-Term Load Forecasting
MTLF	Mid-Term Load Forecasting
LTLF	Long-Term Load Forecasting
ARIMA	Autoregressive Integrated Moving Average
ANN	Artificial Neural Network
ML	Machine Learning
SVR	Support Vector Regression
MLR	Multiple Linear Regression
GRG	Generalized Reduced Gradient
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
RMSE	Root Mean Square Error
GH	Girls Hostel
IT	Incomer Transformer
WNN	Wavelet Neural Network
RBF	Radial Basis Function
α	Level Coefficient
β	Trend Coefficient

Data availability statement

The data used in this paper are collected from the smart meters installed at NIT Patna. The data are available on request from the authors.

Competing interest

There is no potential competing interest related to this work.

Acknowledgement

This work is supported by the Science & Engineering Research Board, Department of Science and Technology, Government of India, under grants number ECR/2017/001027.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors contribution statement

First name and last name of second author

Mala De: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Roles/Writing – original draft; Writing – review and editing.

References

1. Zhang, P., Wu, X., and Wang, X. "Short–term load forecasting based on big data technologies," *CSEE Journal of Power and Energy Systems*, **1**(3), pp. 59-67 (2015). DOI: 10.17775/CSEEJPES.2015.00036

2. Wang, Y., Chen, Q., Hong, T., et al. "Review of smart meter data analytics: applications, methodologies, and challenges", *IEEE Trans. Smart Grid*, **10**(3), pp. 3125-3148 (2018). DOI: 10.1109/TSG.2018.2818167

3. Tian, C., Ma, J., Zhang, C., et al. "A deep neural network for short- term load forecast based on LSTM and Convolution neural network", *Energies*, **11**, 3493 (2018). DOI: 10.3390/en11123493

4. Singh, N., Mohanty, S.R., and Shukla, R.D., "Short term electricity price forecast based on environmentally adapted generalized neuron", *Energy*, **125**, pp. 127–39 (2017). DOI: 10.1016/j.energy.2017.02.094

5. Kazemzadeh, M.R., Amjadian, A., and Amraee, T. "Long term electric peak load forecasting of Azarbaijan regional electricity grid", *Iranian Conference on Electrical Engineering (ICEE)* (2020). DOI: 10.1016/j.jup.2019.04.001

6. Hernandez, L., Baladron, C., AguiarCarro, J.M, et al. "A survey on electric power demand forecasting: Future trends in smart grids, microgrids, and smart buildings", *IEEE Commun. Surv. Tutorial*, **16**, pp. 1460-1495 (2014). DOI: 10.1109/SURV.2014.032014.00094

7. Moghram, I. and Rahman, S. ``Analysis and evaluation of five short-term load forecasting techniques", *IEEE Trans. Power Syst.*, **4**(4), pp. 1484-1491 (1989). DOI: 10.1109/59.41700

8. Alex, D. and Timothy, C. "A regression-based approach to short term system load forecasting", *IEEE Trans. on Power Syst.* **5**(4), pp. 1535–1550 (1990). DOI: 10.1109/59.99410

9. Hagan, M.T. and Behr, S.M. "The time series approach to short term load forecasting", *IEEE Trans. of Power Syst.* PWRS-**2**(3) (1987). DOI: 10.1109/TPWRS.1987.4335210

10. Rahman, S. and Drezga, I. "Identification of a standard for comparing short-term load forecasting techniques", *Electric Power Systems Research*, **25**(3), pp. 149-158 (1992). DOI: 10.1016/0378-7796(92)90013-Q

11. Li, Y., Han, D., and Yan, Z. "Long-term system load forecasting based on data-driven linear clustering method", *J. Mod. Power Syst. Clean Energy,* **6**, pp. 306–316 (2018). DOI: 10.1007/s40565-017-0288-x

12. Ji, P.R., Xiong, D., Wang, P., et al. "A study on exponential smoothing model for load forecasting", In *Proceedings of 2012 Power and Energy Engineering Conference*, Shanghai, pp. 1–4 (2012). DOI: 10.1109/APPEEC.2012.6307555

13. Gob, R., Lurz, K., and Pievatolo, A. "Electrical load forecasting by exponential smoothing with covariates", *Applied Stochastic Models in Business and Industry*, **29**(6), pp. 629–645 (2013). DOI: 10.1002/asmb.2008

14. Filho, K.N., Lotufo, A.D.P., and Minussi, C.R. "Multinodal load forecasting using a general regression neural network", *IEEE Trans. on Power Delivery,* **26**(4), pp. 2862-2869 (2011). DOI: 10.1109/TPWRD.2011.2166566

15. Zongying, L., Loo, C.K., and Pasupa, K. "A novel error-output recurrent two-layer extreme learning machine for multi-step time series prediction", Sustainable *Cities and Society*, 66 (2021). DOI: 10.1016/j.scs.2020.102613

16. Zhang, X., Wang, J., and Zhang, K. "Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm", *Electric Power System and Research,* **146**, pp. 270-285 (2017). DOI: 10.1016/j.epsr.2017.01.035

17. Ceperic, E., Ceperic, V., Member, S., et al. "A strategy for short-term load forecasting by support vector regression machines", *IEEE Trans. Power Syst*em, **28**(4), pp. 4356–4364 (2013). DOI: 10.1109/TPWRS.2013.2269803

18. Hamed H.H. Aly, ``A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid'', *Electric Power Systems Research*, **182** (2020). DOI: 10.1016/j.epsr.2019.106191

19. Li, S., Wang, P., and Goel, L. "Short-term load forecasting by wavelet transform and evolutionary extreme learning machine", *Electric Power Systems Research*. **122**, pp. 96–103 (2015). DOI: 10.1016/j.epsr.2015.01.002

20. Hendawia, M.E. and Wanga, Z. "An ensemble method of full wavelet packet transforms and neural network for short term electrical load forecasting", *Electric Power Systems Research*, **182**, pp. 1-13 (2020). DOI: 10.1016/j.epsr.2020.106265

21. Polikar, R. "Ensemble based systems in decision making", *IEEE Cicuits and Systems Magazine*, **6**(3), pp. 21-45 (2006). DOI: 10.1109/MCAS.2006.1688199

22. Khwajaa, A.S., Anpalagana, A., Naeemb, M., et.al. "Joint bagged-boosted artificial neural networks: Using ensemble ML to improve short-term electricity load forecasting", *Electric Power Systems Research*. **179** (2020). DOI: 10.1016/j.epsr.2019.106080

23. Nazar, M.S., Fard, A.E., Heidari, A., et al. "Hybrid model using three-stage algorithm for simultaneous load and price forecasting", *Electric Power Systems Research,* **165**, pp. 214-228 (2018). DOI: 10.1016/j.epsr.2018.09.004

24. Laouafi, A. Mordjaoui, M., Haddad, S., et al. "Online electricity demand forecasting based on an effective forecast combination methodology", *Electric Power Systems Research*, **148**, pp. 35–47 (2017). DOI: 10.1016/j.epsr.2017.03.016

25. Palaninathan, A.C., Qiu, X., and Suganthan, P.N. ``Heterogeneous ensemble for power load demand forecasting", *IEEE Region 10 Conf (TENCON)*, Singapore, pp. 2040-2045 (2016). DOI: 10.1109/EPE.2016.7521771

26. Dudek, G. "Heterogeneous ensembles for short-term electricity demand forecasting", *17th International Scientific Conference on Electric Power Engineering*, Prague, pp. 1-6 (2016). DOI: 10.1109/EPE.2016.7521771

27. Wang, L., Mao, S., Wilamowski, B.M., et al. ``Ensemble Learning for Load Forecasting", in *IEEE Transactions on Green Communications and Networking*, **4**(2), pp. 616-628 (2020). DOI: 10.1109/TGCN.2020.2987304

28. Lee, J. and Cho, Y. "National-scale electricity peak load forecasting: Traditional, machine learning, or hybrid model?", *Energy*, **239**, 122366 (2021). DOI: 10.1016/j.energy.2021.122366

29. Rai, S. and De, M. "Analysis of classical and machine learning based short-term and mid-term load forecasting for smart grid", *International Journal of Sustainable Energy*, **40**(9), pp. 821-839 (2021). DOI: 10.1080/14786451.2021.1873339

Biographies

Sneha Rai is pursuing PhD in the Department of Electrical Engineering, NIT Patna. Her area of research is Load forecasting of smart metered distribution system, application of AI techniques and optimization.

Mala De is currently working as an Assistant Professor in Electrical Engineering Department in National Institute of Technology in Patna, India. She has received her bachelors from Jalpaiguri Govt. Engg. College, MTech in Power and Energy Systems from NIT Silchar and PhD from Jadavpur University. Her Research Interest includes Power System Operation and Control; Power System Optimization; Load Forecasting for smart metered system, Demand Side Management and Power System Resilience.