energies

MDPI

Article

A Day-Ahead Short-Term Load Forecasting Using M5P Machine Learning Algorithm along with Elitist Genetic Algorithm (EGA) and Random Forest-Based Hybrid Feature Selection

Ankit Kumar Srivastava [1], Ajay Shekhar Pandey [2], Mohamad Abou Houran [3], Varun Kumar [2], Dinesh Kumar [1], Saurabh Mani Tripathi [2], Sivasankar Gangatharan [4] and Rajvikram Madurai Elavarasan [5,*]

1 Electrical Engineering Department, Dr. Rammanohar Lohia Avadh University, Ayodhya 224001, India
2 Department of Electrical Engineering, Kamla Nehru Institute of Technology, Sultanpur 228118, India
3 School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, China
4 Electrical & Electronics Engineering Department, Thiagarajar College of Engineering, Madurai 625015, India
5 School of Information Technology and Electrical Engineering, The University of Queensland, St. Lucia, QLD 4072, Australia
* Correspondence: r.maduraielavarasan@uq.edu.au or rajvikram787@gmail.com

Abstract: A hybrid feature selection (HFS) algorithm to obtain the optimal feature set to attain optimal forecast accuracy for short-term load forecasting (STLF) problems is proposed in this paper. The HFS employs an elitist genetic algorithm (EGA) and random forest method, which is embedded in the load forecasting algorithm for online feature selection (FS). Using selected features, the performance of the forecaster was tested to signify the utility of the proposed methodology. For this, a day-ahead STLF using the M5P forecaster (a comprehensive forecasting approach using the regression tree concept) was implemented with FS and without FS (WoFS). The performance of the proposed forecaster (with FS and WoFS) was compared with the forecasters based on J48 and Bagging. The simulation was carried out in MATLAB and WEKA software. Through analyzing short-term load forecasts for the Australian electricity markets, evaluation of the proposed approach indicates that the input feature selected by the HFS approach consistently outperforms forecasters with larger feature sets.

Keywords: confidence interval; elitist genetic algorithm; feature selection; short-term load forecasting; M5P forecaster; machine learning

check for updates

Citation: Srivastava, A.K.; Pandey, A.S.; Houran, M.A.; Kumar, V.; Kumar, D.; Tripathi, S.M.; Gangatharan, S.; Elavarasan, R.M. A Day-Ahead Short-Term Load Forecasting Using M5P Machine Learning Algorithm along with Elitist Genetic Algorithm (EGA) and Random Forest-Based Hybrid Feature Selection. *Energies* **2023**, *16*, 867. https://doi.org/10.3390/en16020867

Academic Editors: Mohan Lal Kolhe, Sathyajith Mathew and Axel Sikora

Received: 15 December 2022
Revised: 3 January 2023
Accepted: 10 January 2023
Published: 12 January 2023

1. Introduction

Population growth and technology advancements are the primary factors fueling the historical changes incurred in the electricity demand across the world. Electric power plays a key role in the overall sustainable development of a region or a country. Due to the increase in power consumption and rapid electrification across various regions, establishing a robust framework that could manage the price and consumption pattern of electricity is of the utmost importance [1]. Since the mid-1980s, the electrical business has been witnessing a consistent transformation. The electricity market is a client-driven market and thus forecasting of demand load and cost of electricity serves as a crucial planning tool for the market players [2]. In the current power sector scenario, new rules and tariff schemes are being put forward to encourage competitiveness among every generation station, transmission companies and distribution companies. The aforementioned energy market players are not bound to sell and purchase the electricity in realtime to the buyer and seller, respectively, as per their choice. Therefore, it becomes fundamental to perceive the accurate load demand and electricity prices of a particular region and if this accurate information is predicted in advance, then the companies can make a substantial profit in their bid [3]. Predicting the demand accurately and obtaining its pattern well in advance can also help to optimize available generation efficiently. The seamless connectivity that underpins power exchanges encourages the power generation station to sell their electricity at a higher price whenever energy demand peaks [4]. Transmission and distribution companies also avoid long-term contracts because current infrastructure flexibility allows them to purchase the required electricity during peak demand. This ultimately saves a humongous sum of money that is required to pay generation companies (GENCOs) for their fixed electricity cost. Another impact on the power system is influenced by the utilization of power generated from renewable resources whose cost of electricity is now competitive with respect to their conventional counterpart or even lower than them in some cases. This scenario induces difficulties for the thermal stations in terms of selling their generated electricity price at higher margins. Solar energy is abundantly available throughout the daytime in most habitable regions, which has the potential to force thermal stations to run at a technical minimum or even to reserve energy by shutting down [5]. Thus, accurately forecasting the load demand for a given region is vital for the sustainable energy business.

STLF is one of the fields receiving more attention among research scholars, and technological development has refined STLF to improve its forecasting accuracy. Load forecasting largely depends on many seasonal factors (such as temperature, relative humidity and sun availability), economical parameters (such as availability of fuels, i.e., coal, naphtha, etc.) and availability of other types of generating stations [6]. Electric load demand needs to be predicted diligently to accommodate for aberrant climatic conditions including extreme cold conditions or blistering climate [7]. The power sector also needs to precisely monitor industrial and consumer energy consumption pattern to identify whether an unprecedented raise or drop in demand that can potentially put the overall power system security at risk occurs [8]. So, it becomes quite important to forecast the electricity demand block-wise accurately to minimize the generation demand gap. Thus, due to the above-mentioned reasons, STLF emerged as one of the attractive research areas and is of great interest for researchers in the power system domain. The artificial neural network (ANN) method, the time-series method, the regression method, the semi-parametric method and the non-parametric method [9–16] are some of the commonly used approaches for forecasting electricity loads. Grzegorz [17] used a stepwise lasso regression (LR) method and introduced a model which decreases the desired result in the predictor's number. This model of STLF uses the LR and daily cycle basis load pattern, and is based on a univariate model which considers a selection of variables in relationship with local current input. Cecati et al. [18] emphasize that the calculation provided by decay radial basis function neural networks (DRBFNNs), extreme learning machines (ELMs) and support vector regression (SVR) machine enhanced the performance with better error adjustments and with more improved second-order helpful outcomes for forecasting for a whole day. The study conducted by Zhai et al. [19] utilized the self-closeness of electrical load recorded data, which yielded a multi-resolution wavelet and then the Hurst parameter values were included to evaluate the vertical scaling factors in function systems (IFS). The study used this model to forecast the electricity load in two scenarios: fractal extrapolation and fractal interpolation. Arora et al. [20] demonstrated ANN-based triple-seasonal auto regressive moving average (ARMA), exponential smoothing and triple-seasonal Holt–Winters–Taylor (HWT). The authors also discussed the triple-seasonal intraweek singular value decomposition (SVD), which was based on exponential smoothing methods. Further, the method proposed in [20] can be used to predict the model load for particular days. Zeng et al. [21] proposed an STLF approach based on the cross multi-model and second decision mechanism to improve the stability and forecasting accuracy. Nose-Filho et al. [22] elaborated on a method that minimized the input to ANN to perform forecasting with a modified general regression neural network. They also introduced two methods: one for short-term multimodal load forecasting for a local load and another one for short-term multimodal load forecasting for a global load. Zhang et al. [23] proposed a method to integrate the hierarchical structure and the forecasting model via a novel closed-loop clustering (CLC) algorithm. Rafi et al. [24] developed a new method for STLF based on a long short-term memory (LSTM) network and convolutional neural network (CNN). Li et al. [25] proposed a novel model that utilizes the theory of extreme learning machines (ELMs), wavelet transform (WT) and multi-species artificial bee colony (MABC) algorithms. In [25], the authors utilized the changes in wavelet to break down the load series to obtain complex parameters at different frequencies which are then estimated independently with a hybrid model based on MABC and ELM. Kouhi et al. [26] discussed a model for forecasting using the differential evolutionary (DE) feature with a new multi layer perception (MLP) neural network based on the hybrid Levenberg–Marquardt (LM) method. In [26], input data reconstruction is accomplished by employing the Taken embedding theorem using the chaotic intelligent FS method. Ungureanu et al. [27] developed a new approach for STLF for non-residential consumers based on market-oriented machine learning (ML) models. Several past research studies on STLF are summarized in Table 1.

Table 1. Review of several past research studies on STLF.

Sr. No.	Year	Author [Ref.]	Methodology Used	Feature Selection	Performemnce Measure			
					MAPE	MAE	RMSE	EV
1.	2018	Luo et al. [28]	Dynamic Regression Model (DRM)-based detection method	No	√	X	X	X
2.	2018	Jiao et al. [29]	Multiple Sequence LSTM Recurrent Neural Network	No	√	√	√	X
3.	2019	Haq et al. [30]	T-Copula-IEMD-DBN Method	No	√	X	√	X
4.	2019	Deng et al. [31]	TCMS-CNN Algorithm	Yes	√	√	√	X
5.	2020	Hong et al. [32]	Iterative Resblocks-Based Deep Neural Network (IRBDNN)	No	√	√	√	X
6.	2020	Ahmad et al. [33]	SVM-GS, ELM-GA	Yes	√	√	√	X
7.	2020	Pei et al. [34]	ILSTM network	Yes	√	√	√	X
8.	2021	Rafi et al. [24]	CNN-LSTM-based hybrid Network	Yes	√	√	√	√
9.	2021	Ungureanu et al. [27]	LSTM, LSTMed, GRU, CNN-LSTM	Yes	√	√	√	X
10.	2021	Xuan et al. [35]	CNN-BiGRU Algorithm	Yes	√	X	√	X
11.	2022	Ijaz et al. [36]	Artificial Neural Network (ANN) layer and LSTM	Yes	√	√	√	X
12.	2022	Zhang et al. [37]	Improved Seagull Optimization Algorithm and SVM (ISOA-SVM) Method	No	√	√	√	X
13.	2022	Liu et al. [38]	DenseNet-iTCN)	Yes	√	√	√	√

One can see that different classifiers based on support vector machine (SVM), ANN, fuzzy logic, etc., have been employed in the previously published studies. The M5P method has been used in many problems; however, it has not commonly been used in load forecasting problems. In this work, the applicability and utility of the M5P forecaster has been studied using the proposed HFS algorithm (which employs an EGA and random forest method) to address the load forecasting problem. The key contributions made in this work are as follows:

1. Proposal of a novel HFS employing an EGA and random forest method for FS meant for the load forecasting problem;

2. Implementation of the M5P forecaster with FS and WoFS to analyze the short-term load forecasts for the Australian electricity markets;
3. Application of confidence interval to fix the margins of error in the forecasted load;
4. Drawing certain insights on the number as well as type of features that affect the load in different seasons;
5. Comparing the performance of the proposed forecaster (with FS and WoFS) to the performance of forecasters based on J48 and Bagging.

The remaining sections of the paper are structured as follows. The methodology adopted for comparison of forecasts with FS and WoFS is discussed in Section 2. Section 3 explains the STLF using the M5P forecaster. Section 4 elucidates the methodology used for input feature selection using the novel HFS algorithm. The results and performance of the proposed methodology are presented in Section 5. Conclusions and remarks concerning findings are provided in Section 6.

2. Methodology Adopted for Comparison of Forecasts with FS and WoFS

In this work, STLF with FS and WoFS for next day was considered. For STLF, the proposed HFS algorithm is based on EGA and the random forest method. STLF was implemented using the concept of a similar week for each day and for all seasons on a half-hourly basis. STLF was performed using the M5P forecaster employing the full input feature set as well as with the selected (reduced) input feature set. The forecast accuracy was compared between using the full input feature set and using the reduced input feature set, validating the utility of STLF with a reduced input feature set.

The performance results obtained from M5P (with FS and WoFS) were also compared with the forecasts based on J48 and Bagging.

Accordingly, to ascertain the superiority of FS, the methodology adopted meant for comparison of forecasts with FS and WoFS is shown in Figure 1.

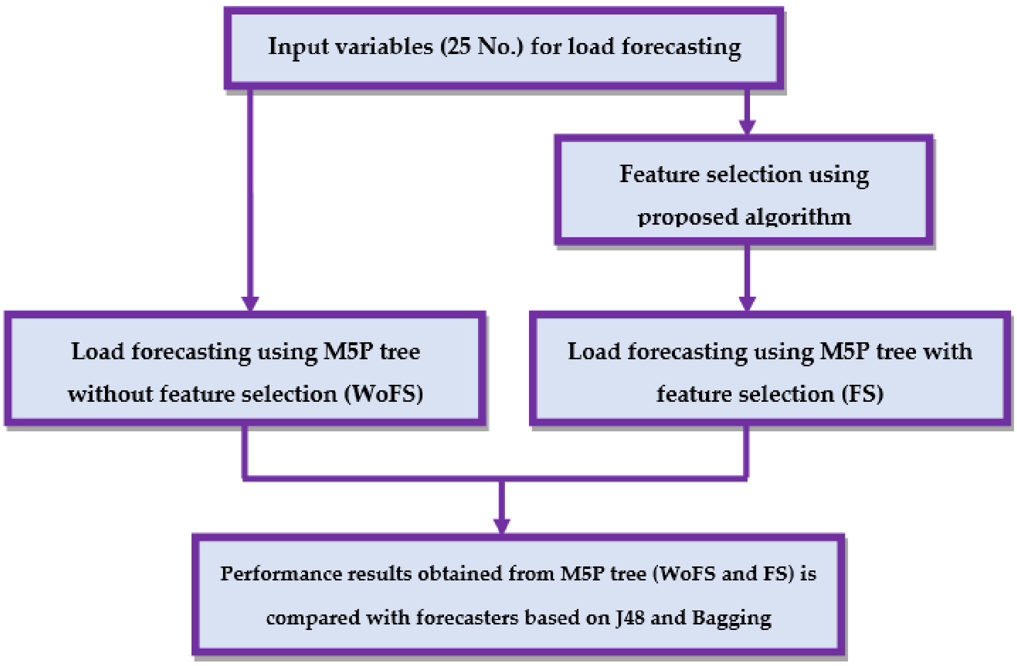


Figure 1. Methodology adopted for comparison of forecasts with FS and WoFS.

3. STLF Using M5P Forecaster Model

M5P is an ML algorithm which is a modified version of the M5 tree algorithm [39] and is used for both classification as well as regression problems. This modification allows it to deal with attribute missing values and enumerated attributes. M5P gives better results with longer data series as input since it is more sensitive to data splitting. The M5 tree was developed for the prediction of continuous variables and it serves as a flexible prediction tool since the construction of the tree is based on a multivariate linear model [40]. Generally, the M5 tree is a three-step process, i.e., construction of tree using input data, tree pruning and tree smoothing, whereas the M5P model consists of five important steps. M5P is a binary regression tree that stores a linear regression model at every leaf (last node), which predicts the class value of incoming instances. It uses the splitting criterion for the best split of a portion of training data that reaches any node. In the M5P tree, the standard deviation of the portion is used as a measure of error at that node. The tree of the M5P model is shown in Figure 2. The five steps of the M5P forecaster model are elaborated as follows:

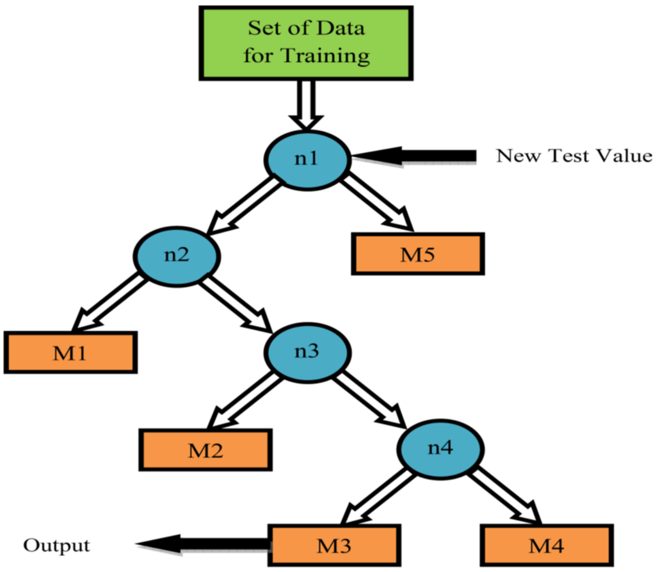


Figure 2. M5P Tree.

Step 1:For the algorithm to maximize the standard deviation reduction (SDR), the input data (enumerated attributes) are taken and converted to binary variables using expression (1).

$$\mathrm{SDR} = \sigma(\mathrm{Cs}) - \sum_{k} \frac{|\mathrm{Cs_k}|}{|\mathrm{Cs}|} \times \sigma(\mathrm{Cs_k}) \tag{1}$$

where Cs is the data set of STLF, $\mathrm{Cs_k}$ is the kth subset of STLF, $\sigma(\mathrm{Cs_k})$ is the standard deviation of the kth subset of STLF as a measure of error and $\sigma(\mathrm{Cs})$ is the standard deviation of Cs.

Step 2: The tree is constructed with these binary variables. Overfitting increases as the size of the tree grows. Here, the data overfitting problem will be overcome.

Step 3: To reduce the problem of overfitting, there is a pruning process and discontinuities are compensated.

Step 4: The tree smoothing process is included to balance sharp discontinuities that take place between linear adjacent models at the end nodes (leaf) of the pruned tree.

Step 5: Output is produced as a tree model.

4. Input Feature Selection Using the Proposed HFS Algorithm

Based on Darwin's theory of natural evolution and the genetics of survival of the fittest, the genetic algorithm (GA) is one of the elite and heuristic search techniques and is used to produce useful solutions to optimization problems. The assumption of the relationships between characteristics involved was not considered in this approach when searching the space for FS. GA can easily encode decisions as Boolean value sequences, permitting the feature space to be explored by retaining the choices that support the classification task. Due to its inherent randomness, it also prevents local optimums concurrently. To solve optimization problems, it also makes use of operators inspired by natural evolution viz. selection, mutation and crossover.

Regression trees are traditionally used to predict data provided by the values of the function. A novel HFS algorithm based on EGA and the random forest method was used in this work. In EGA, 20% of the elite population is transferred to the next generation, through which the next generation has a feature set population whose classification accuracy is no less than that of the previous generation. The stratified 10-fold cross-validation (10-FCV) classification accuracy of a given dataset is used as the fitness function. In the present problem, strings of 1 s and 0 s are taken as chromosome segments, with 0 signifying that the feature corresponding to the index is not selected and 1 signifying that the relevant feature is selected. The length of the string is equal to the number of features in the dataset. The stratified 10-FCV classification accuracy (which is measured by means of the WEKA data mining workbench), using a random forest classifier, is all the fitness function computation. Classification accuracy refers to an approximation of the correctly identified number of instances. Roulette wheel selection is used here, and then a single-site crossover is carried out with a probability of 0.7 at each step. Mutation also occurs with a probability of 0.005. In addition, 20% of the elite population is transferred to the next generation. The best collection of features selected through final combination or encoding of chromosomes is given for STLF. The flowchart of the proposed novel HFS algorithm is depicted in Figure 3.

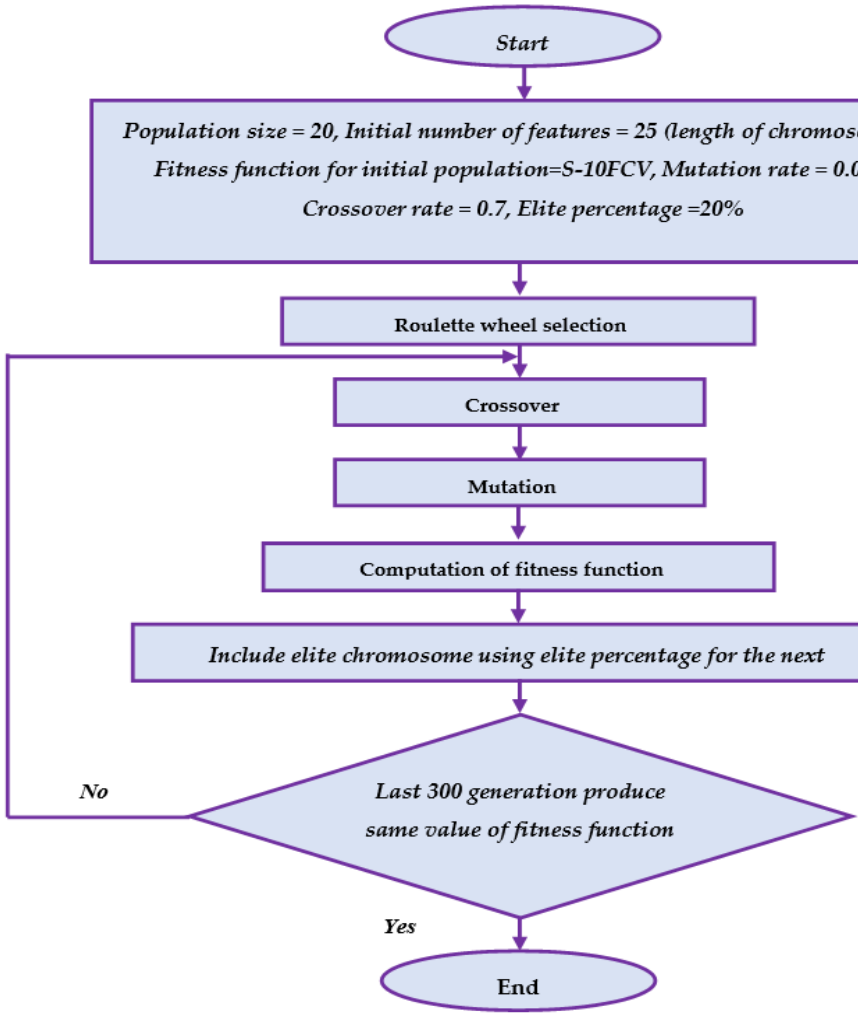


Figure 3. Flow chart for feature selection using the proposed novel hybrid algorithm.

5. Results and Discussions

For STLF, half-hourly historical load data for New South Wales, Australia, and weather data for Sydney for the period January 2014 to June 2016 were obtained from the Australian Energy Market Operator (AEMO) and weatherzone.com.au, respectively. Humidity, wind speed and temperature have been considered as weather data. The EGA algorithm is run in MATLAB, while the formation of optimal tree is carried out in WEKA software. WEKA software is interfaced with MATLAB to perform all regression tree calculations. Final forecasting is carried out using MATLAB. Table 2 lists the input variables affecting the half-hourly STLF.

Table 2. Input features affecting half-hourly STLF.

Class of Input Feature	Timing of Input Feature	Name of Input Feature
Load (Ld)	$Ld_{(K\text{-}00.30)}$	Ld_1
	$Ld_{(K\text{-}01:00)}$	Ld_2
	$Ld_{(K\text{-}01:30)}$	Ld_3
	$Ld_{(K\text{-}24:00)}$	Ld_4
	$Ld_{(K\text{-}23:30)}$	Ld_5
	$Ld_{(K\text{-}23:00)}$	Ld_6
Wind speed (Ws)	$Ws_{(K\text{-}00.30)}$	Ws_1
	$Ws_{(K\text{-}01:00)}$	Ws_2
	$Ws_{(K\text{-}01:30)}$	Ws_3
	$Ws_{(K\text{-}24:00)}$	Ws_4
	$Ws_{(K\text{-}23:30)}$	Ws_5
	$Ws_{(K\text{-}23:00)}$	Ws_6
Temperature (Tem)	$Tem_{(K\text{-}00.30)}$	Tem_1
	$Tem_{(K\text{-}01:00)}$	Tem_2
	$Tem_{(K\text{-}01:30)}$	Tem_3
	$Tem_{(K\text{-}24:00)}$	Tem_4
	$Tem_{(K\text{-}23:30)}$	Tem_5
	$Tem_{(K\text{-}23:00)}$	Tem_6
Humidity (Hy)	$Hy_{(K\text{-}00.30)}$	Hy_1
	$Hy_{(K\text{-}01:00)}$	Hy_2
	$Hy_{(K\text{-}01:30)}$	Hy_3
	$Hy_{(K\text{-}24:00)}$	Hy_4
	$Hy_{(K\text{-}23:30)}$	Hy_5
	$Hy_{(K\text{-}23:00)}$	Hy_6
Hour timing (HTo)	$HT_{o(K\text{-}00.00)}$	HT_o

Each data set for STLF has 25 input features and a total of 2016 data sets were used in a training set to forecast the electricity load. The results derived from the proposed study are explained in two parts—the importance of FS is discussed in the first part and the various performance measures to compute the forecast accuracy is presented in the second part. The input features set for FS, which affects STLF, are taken from Table 2 and forecasting is carried out on the concept of similar week. Data sets were considered on the basis of similar weeks. Each data set consists of six weeks, i.e., for a given week to be forecasted, the preceding and successive two weeks along with the same week corresponding to the previous year were considered, while one preceding week of the same year was also included. For instance, if the input FS or forecasting of the electricity load is to be done for the week of 15–21 January 2016, the training set would consist of the data corresponding to 8–14 January 2016, 15–21 January 2015, 8–14 January 2015, 1–7 January 2015, 22–28 January 2015 and 29 January–4 February 2015. To obtain the FS, the accuracy of the data set for the proposed HFS was computed using 10-FCV. All the data were tested with this algorithm at least once. Thus, FS is the only algorithm that can be used to perform feature analysis.

Table 3 shows the number of times a particular input feature was selected out of the total 36 times for which the input FS was made.

Table 3. Input feature selected (year-wise) for STLF.

Name of Input Feature	Number of Times Input Feature Selected	Name of Input Feature	Number of Times Input Feature Selected
Ld_6	17	Tem_6	11
Ld_5	12	Tem_5	08
Ld_4	12	Tem_4	12
Ld_3	22	Tem_3	18
Ld_2	29	Tem_2	12
Ld_1	36	Tem_1	13
Ws_6	02	Hy_6	07
Ws_5	07	Hy_5	10
Ws_4	09	Hy_4	12
Ws_3	12	Hy_3	12
Ws_2	11	Hy_2	14
Ws_1	07	Hy_1	12
		HT_o	36

It is clear from Table 3 that the input feature load of the present-day (Ld_1), (Ld_2) and (Ld_3) are significant variables and are selected 36, 29 and 22 times, respectively. The variable wind speed of the present day (Ws_3) is selected more often than the previous day. Moreover, the input feature temperature of the present day (Tem_3) is often selected as compared to the previous day of the present day. On the other hand, the humidity of the present day (Hy_2) was found to be selected more often than the previous day. The input feature hour type (HT_o) is selected in all the runs, i.e., 36 times.

The effects of features can also be analyzed according to seasons. Table 4 shows the season-wise significance of different features. From Table 4, it can be seen that Ld_2, Ld_1, Ws_2, etc., are the features that assume more significance during the winter season. Ld_2, Ls_3, etc., are the features that assume higher priority during the spring season. Ld_2, Tem_3, etc., are the features that assume higher priority during the summer season. The input feature load of the present day Ld_1 and hour type HT_o seems to be a feature, regardless of the season. These analyses point out the relative significance of the feature in terms of seasonal variations.

Table 4. Input feature is selected (season-wise) for STLF.

Name of Input Feature	Summer	Winter	Spring
Ld_6	06	06	05
Ld_5	03	03	06
Ld_4	03	03	06
Ld_3	06	07	09
Ld_2	09	09	11
Ld_1	12	12	12
Ws_6	00	01	01
Ws_5	03	04	00
Ws_4	02	05	02
Ws_3	02	06	04
Ws_2	03	07	01
Ws_1	02	02	03

Table 4. *Cont.*

Name of Input Feature	Summer	Winter	Spring
Tem_6	03	03	05
Tem_5	02	05	01
Tem_4	04	06	02
Tem_3	08	06	04
Tem_2	02	05	05
Tem_1	06	02	05
Hy_6	01	04	02
Hy_5	03	02	05
Hy_4	05	04	03
Hy_3	04	05	03
Hy_2	06	05	03
Hy_1	02	06	04
HT_o	12	12	12

Performance measures viz. mean absolute percentage error (MAPE), error variance (EV), root mean square error (RMSE) and mean absolute error (MAE) were used to assess the numerical accuracy of the load forecasting [41].

The average error of each method was calculated week-wise for all seasons. Table 5 depicts the comparison between the M5P + FS approach and five other approaches (J48, Bagging, J48 + FS, Bagging + FS, and M5P) in terms of various performance measures viz. MAPE, MAE, EV and RMSE. The overall average performance for each method is also summarized in the last column. The results show that the M5P + FS method performs better than the rest of the methods used for comparison.

The error of electricity load is evaluated for the four prior weeks and at regular intervals of half an hour to calculate the confidence interval for one day. Afterward, half-hourly standard deviations (Δ) and (2Δ) were computed for 95% confidence interval. The lower and upper limits are computed as follows:

$$Lower\ Limit = Forecast\ value - 2\Delta$$

$$Upper\ Limit = Forecast\ value + 2\Delta$$

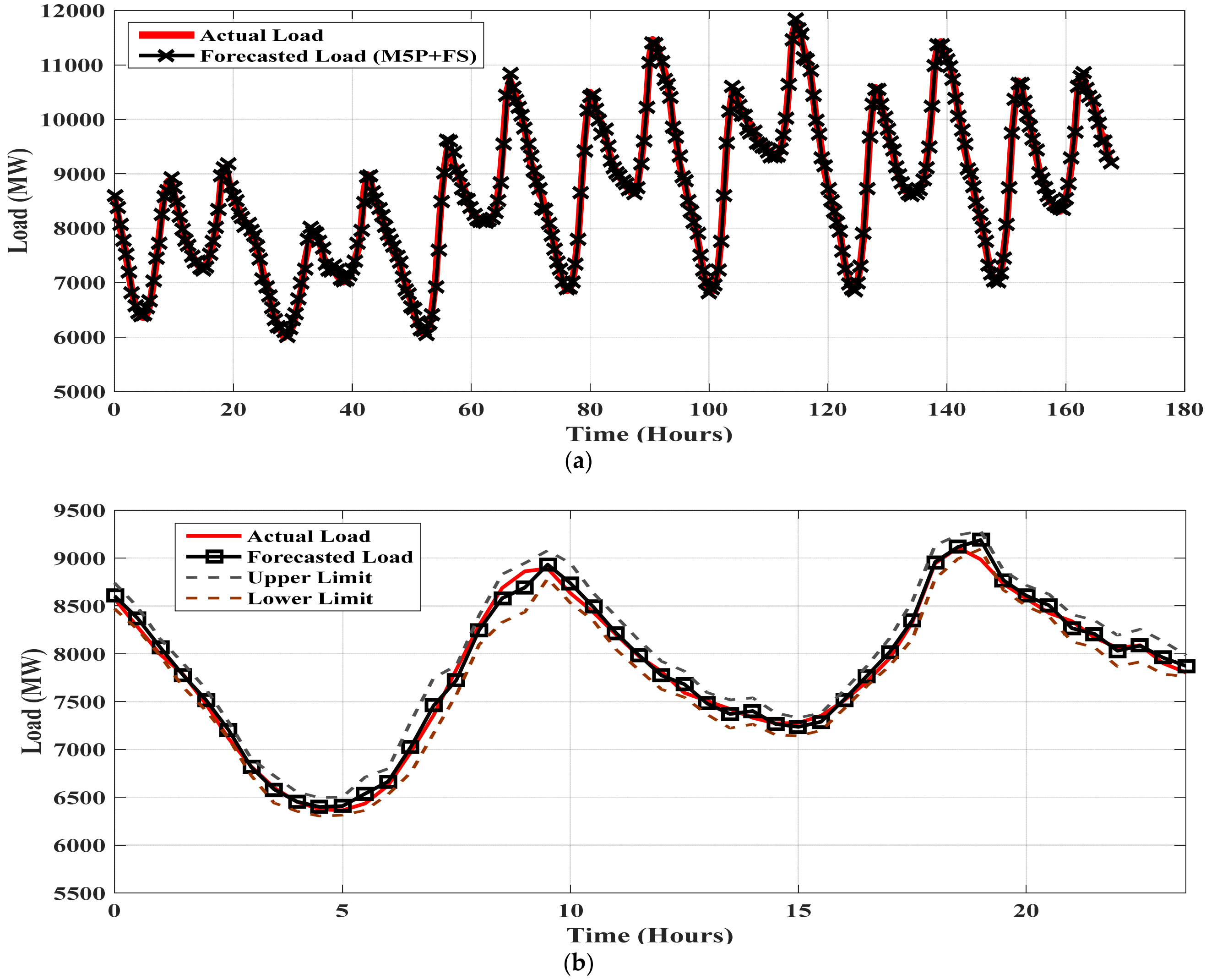
Results corresponding to the proposed methodology for the winter, spring and summer for the Australian electricity market are shown in Figures 4–6, respectively. Table 5 clearly shows that the proposed method (M5P + FS) performs better than other methods in terms of all performance measures in all seasons.

Table 5. Performance of proposed methodology in terms of various performance measures.

Sr. No.	Methodology	Name of Performance Measures	Season			Mean
			Winter (1–7 August 2015)	Spring (1–7 September 2015)	Summer (1–7 February 2016)	
1	J48	MAPE	1.66	1.82	1.42	1.63
	J48 + FS		1.37	1.53	0.95	1.28
	Bagging		1.21	0.98	0.83	1.01
	Bagging + FS		1.16	0.93	0.80	0.96
	M5P		1.07	0.99	0.64	0.90
	M5P + FS		0.67	0.70	0.61	0.66
2	J48	MAE	147.39	138.43	108.40	131.41
	J48 + FS		120.43	114.05	73.79	102.76
	Bagging		106.51	78.28	64.05	82.95
	Bagging + FS		102.43	74.52	62.02	79.66
	M5P		93.73	80.15	49.50	74.46
	M5P + FS		56.54	55.42	47.76	53.24

Table 5. *Cont.*

Sr. No.	Methodology	Name of Performance Measures	Season			Mean
			Winter (1–7 August 2015)	Spring (1–7 September 2015)	Summer (1–7 February 2016)	
3	J48	RMSE	215.63	221.77	140.64	192.68
	J48 + FS		190.14	164.95	95.13	150.07
	Bagging		151.66	108.72	81.76	114.05
	Bagging + FS		147.97	100.75	78.41	109.04
	M5P		131.17	107.06	63.91	100.71
	M5P + FS		73.00	73.75	60.33	69.03
4	J48	EV	0.00033	0.00055	0.00013	0.00034
	J48 + FS		0.00029	0.00026	0.00006	0.00020
	Bagging		0.00016	0.00009	0.00004	0.00010
	Bagging + FS		0.00015	0.00007	0.00004	0.00009
	M5P		0.00011	0.00008	0.00003	0.00007
	M5P + FS		0.00003	0.00004	0.00002	0.00003



(**a**)

(**b**)

Figure 4. *Cont.*

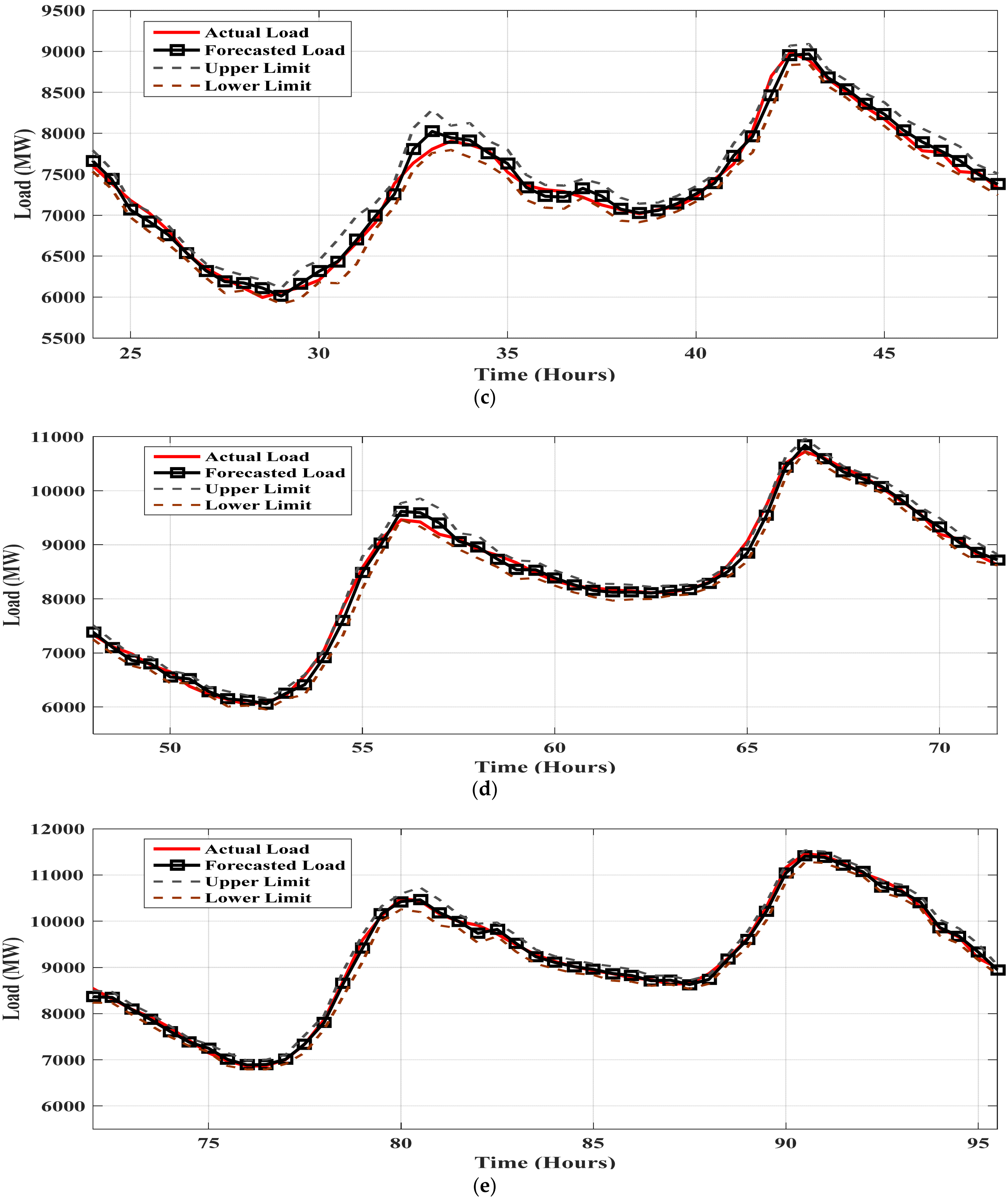
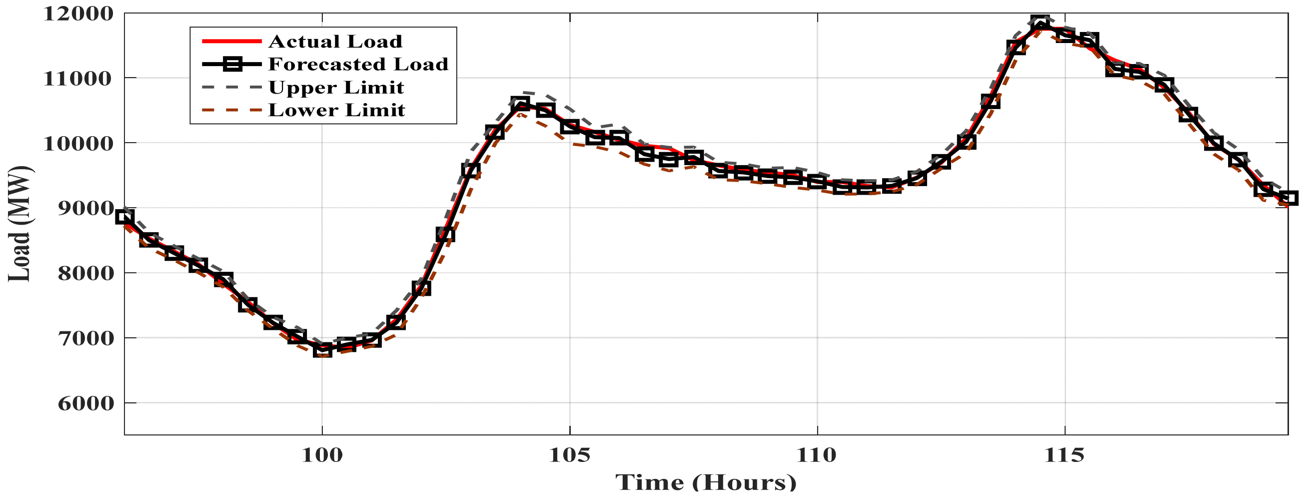
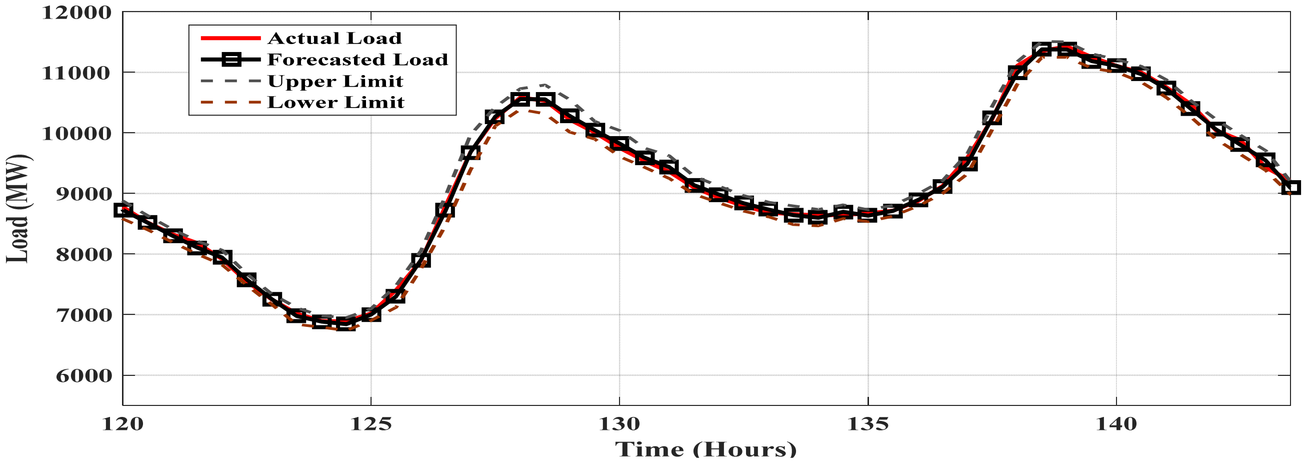


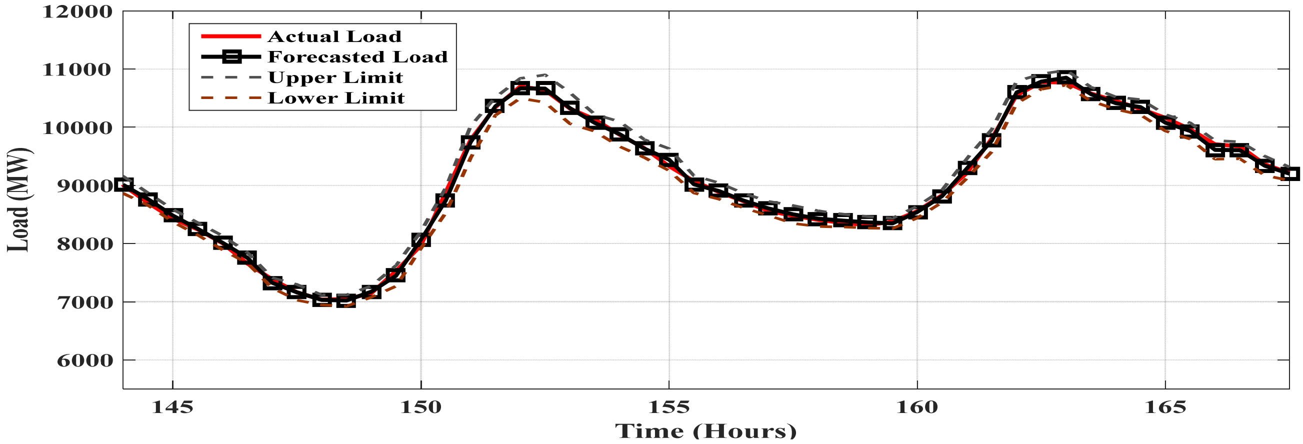
Figure 4. *Cont.*



(f)



(g)



(h)

Figure 4. Forecasting in winter 1–7 August 2015 with M5P +FS, (**a**) entire week, (**b**–**h**) daily forecast.

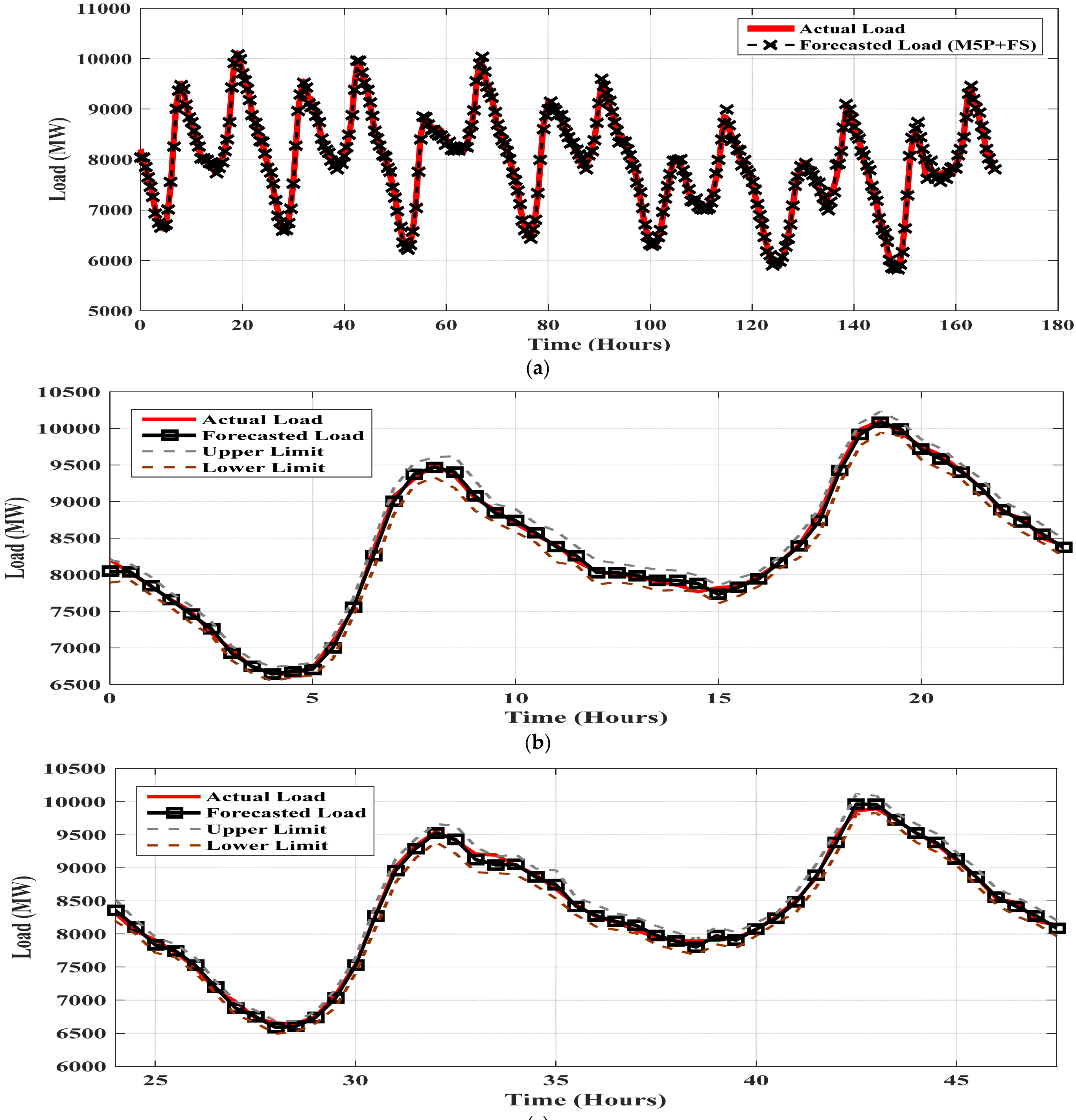


Figure 5. *Cont.*

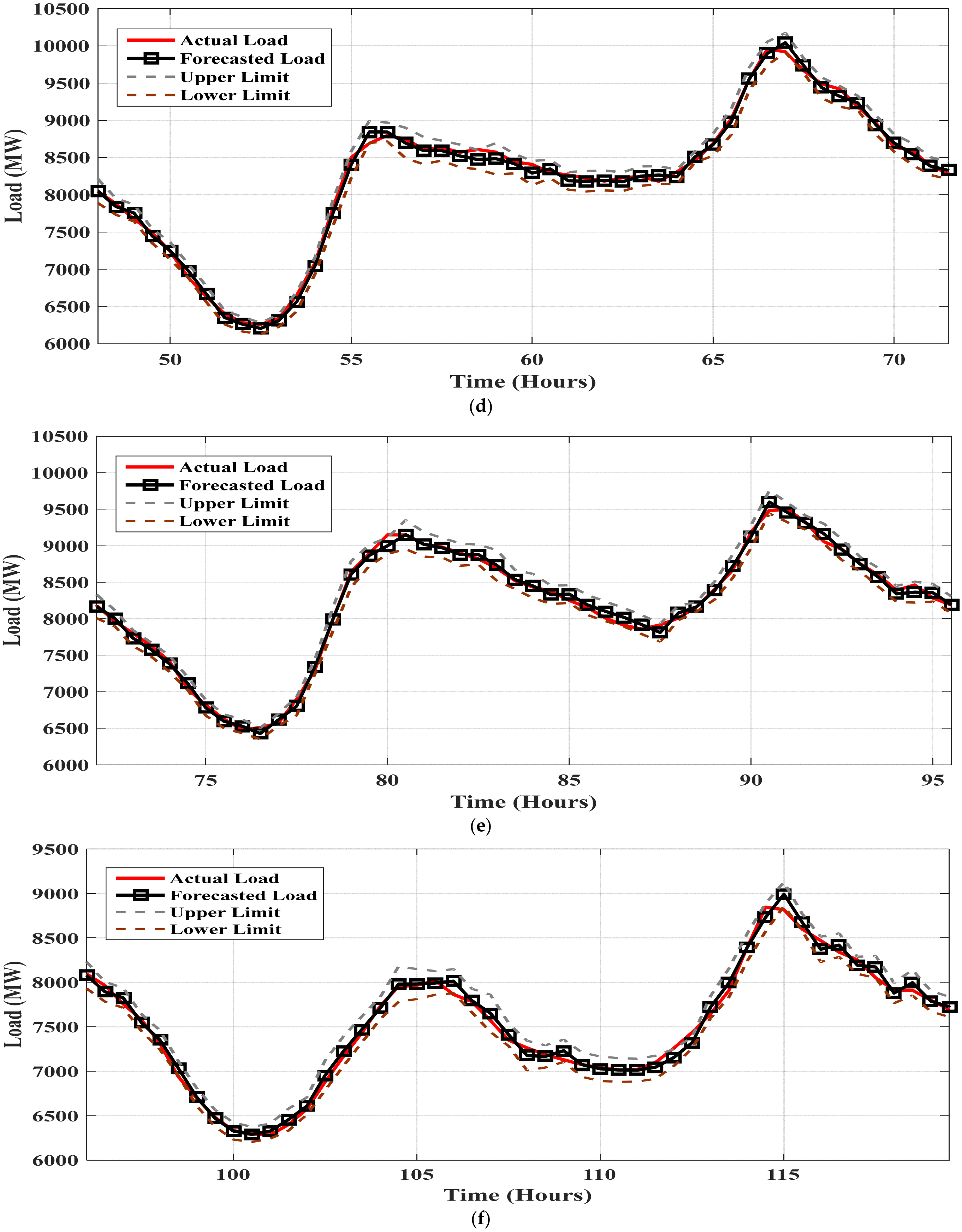
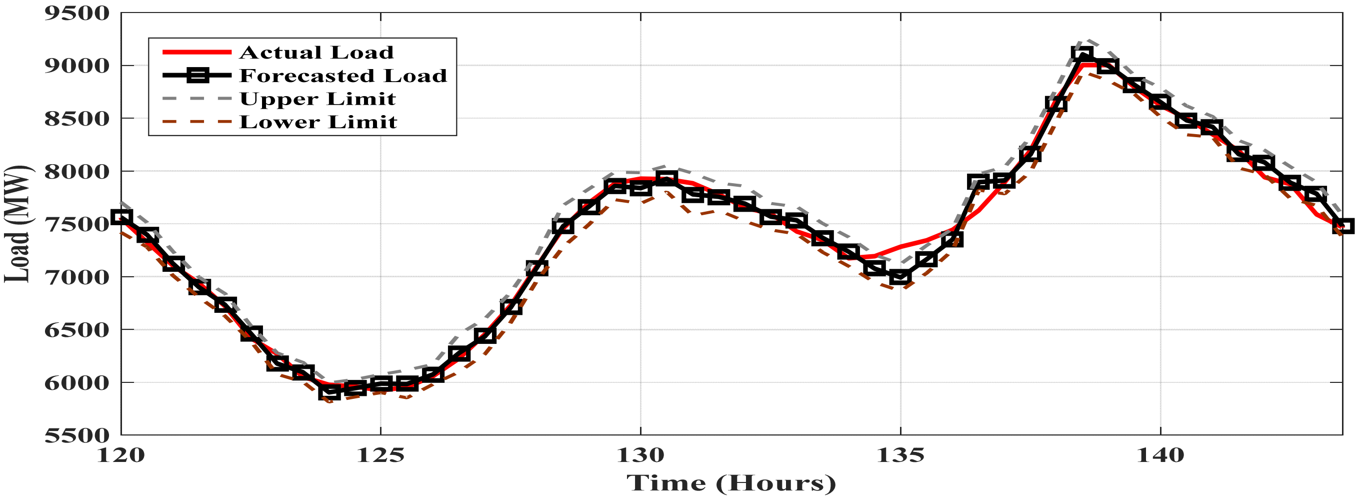
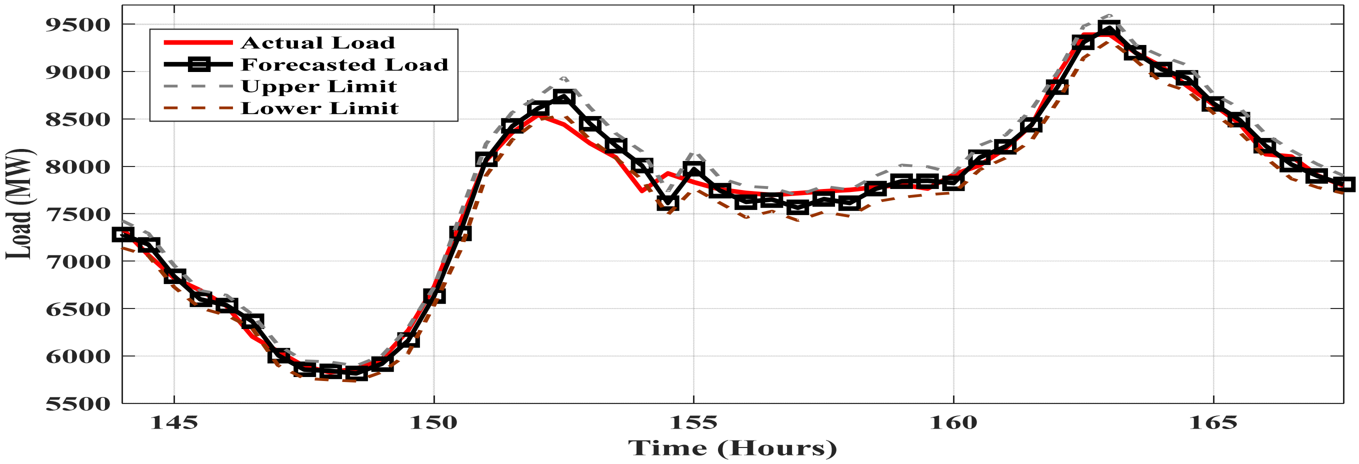


Figure 5. *Cont.*

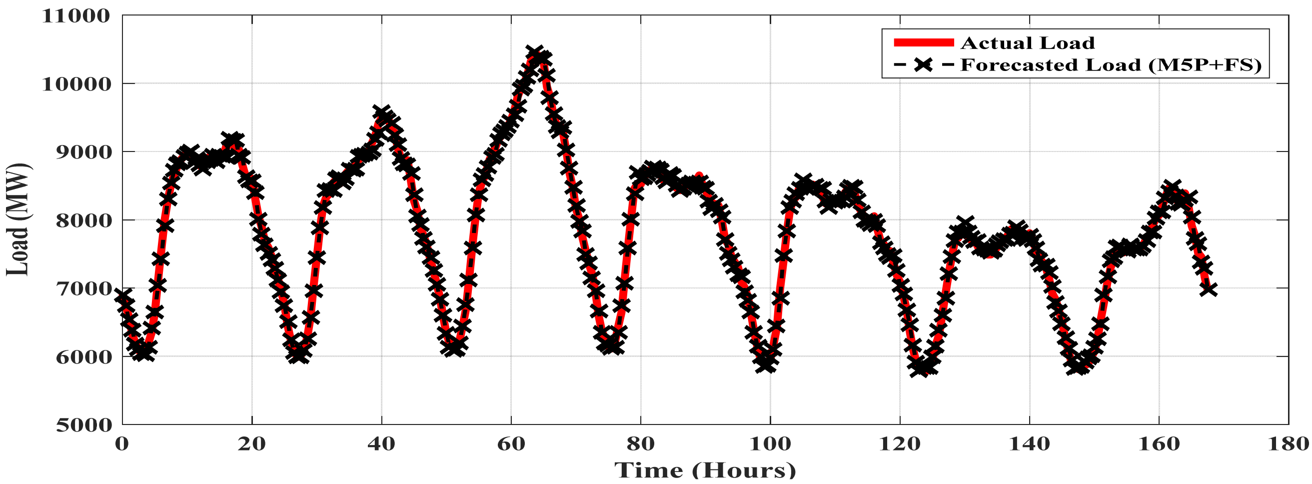


(**g**)



(**h**)

Figure 5. Forecasting in spring 1–7 September 2015 with M5P + FS, (**a**) entire week, (**b**–**h**) daily forecast.



(**a**)

Figure 6. *Cont.*

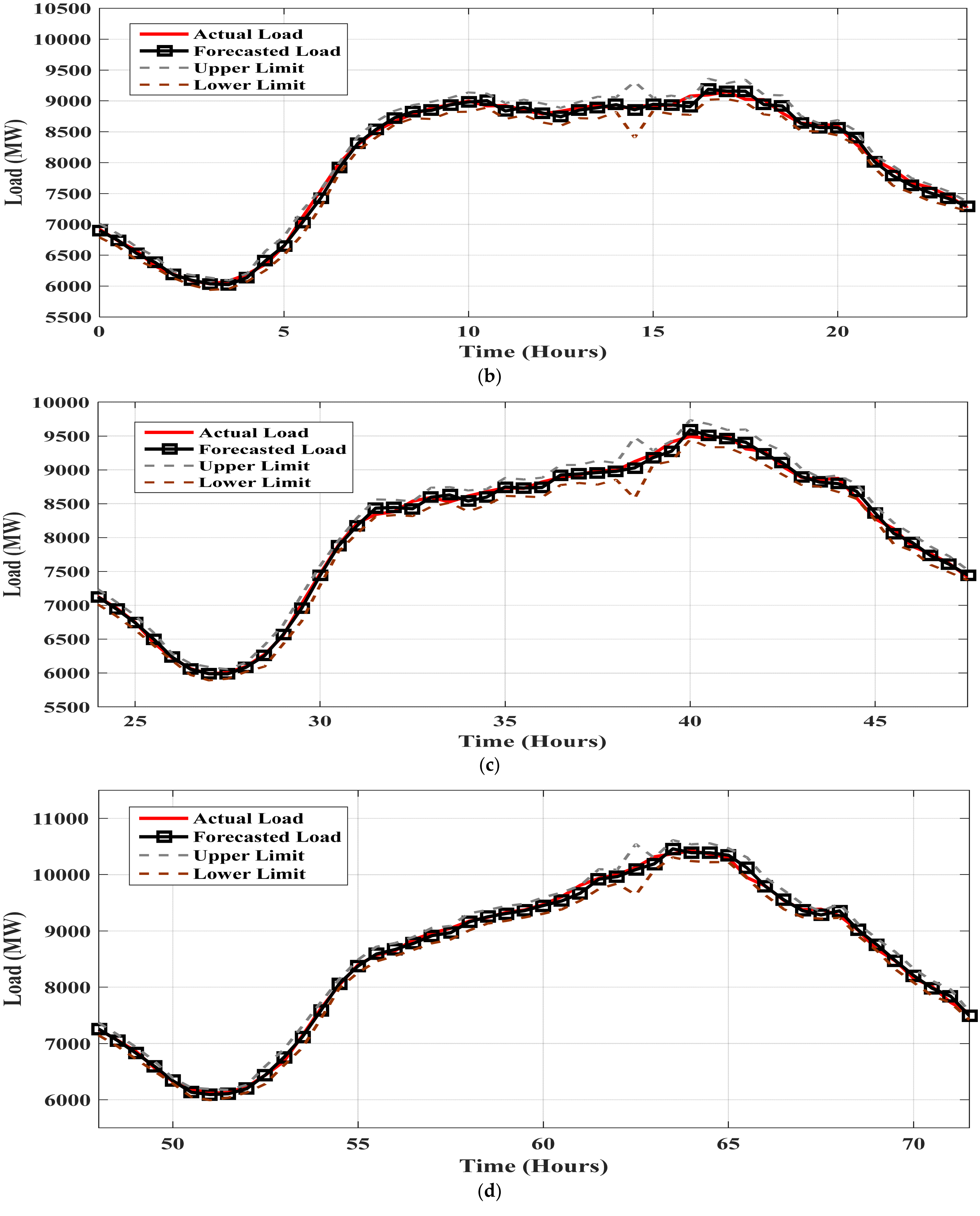
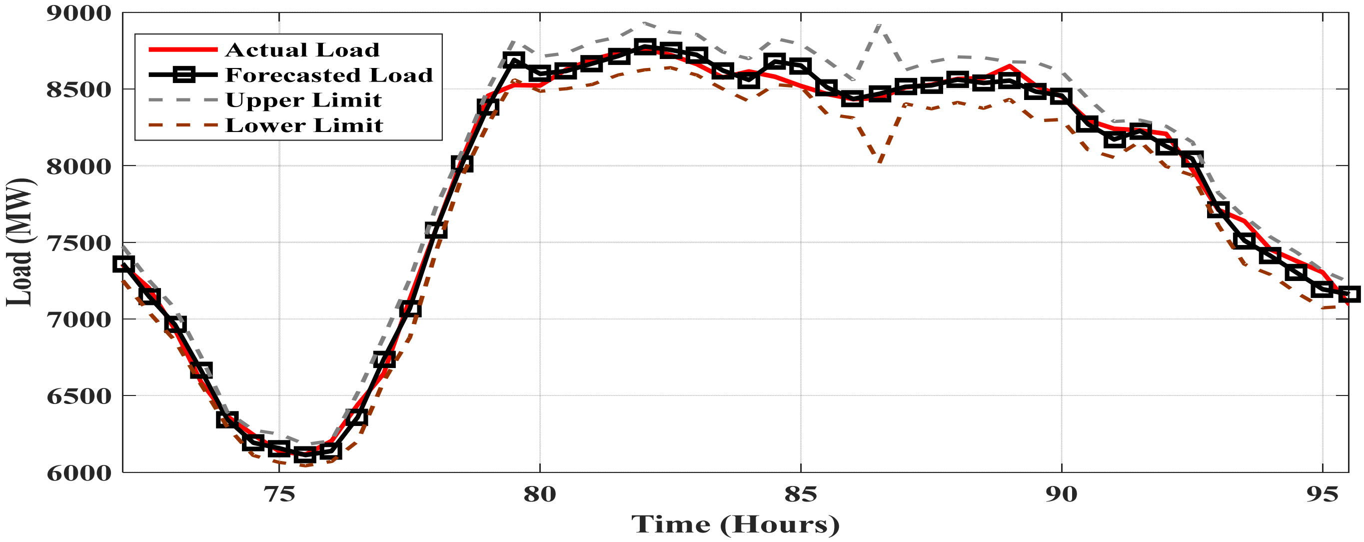
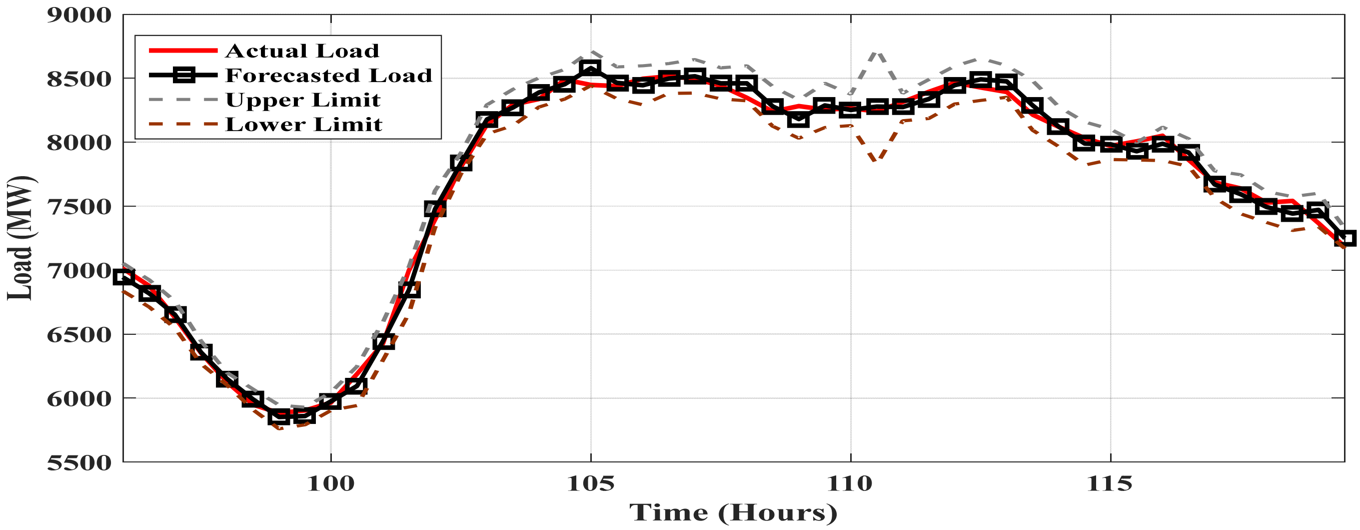


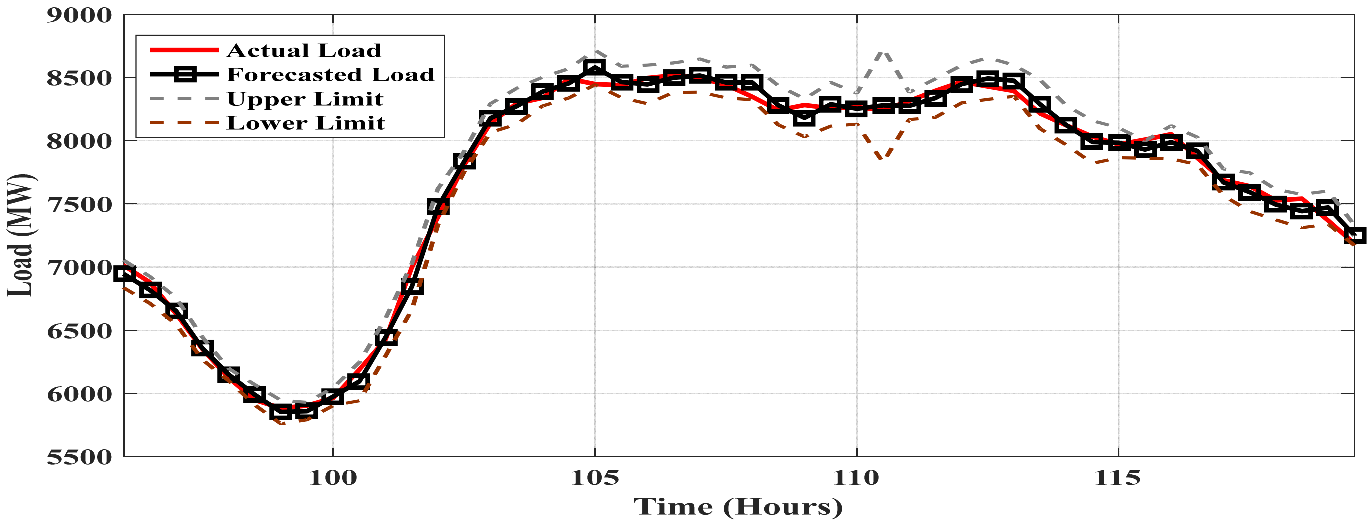
Figure 6. *Cont.*



(e)

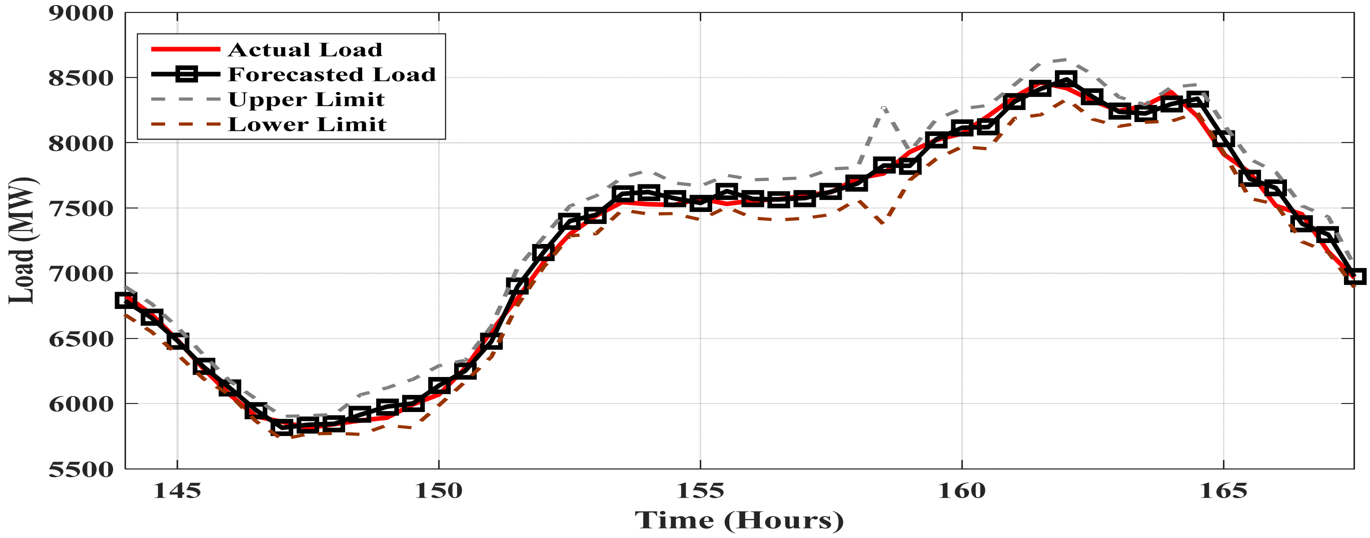


(f)



(g)

Figure 6. *Cont.*



(h)

Figure 6. Forecasting in summer 1–7 Feb. 2016 with M5P + FS, (**a**) entire week, (**b**–**h**) daily forecast.

Table 6 shows the percentage improvement attained with the proposed method (M5P + FS) over the other approaches. It is noted that the proposed methodology resulted in a 59.51% improvement compared to J48. It is also worth noting that M5P + FS has enhanced forecast accuracy over all considered methods.

Table 6. Improvement of different performance measures using M5P + FS compared to other approaches for STLF.

Sr. No.	Methodology	Mean MAPE	Percentage Improvement (%)
1.	M5P +FS	0.66	-
2.	J48	1.63	59.51
3.	J48 + FS	1.28	48.44
4.	Bagging	1.01	34.65
5.	Bagging + FS	0.96	31.25
6.	M5P	0.90	26.67
Sr. No.	**Methodology**	**Mean MAE**	**Percentage Improvement (%)**
1.	M5P +FS	53.24	-
2.	J48	131.41	59.48
3.	J48 + FS	102.76	48.19
4.	Bagging	82.95	35.81
5.	Bagging + FS	79.66	33.16
6.	M5P	74.46	28.50
Sr. No.	**Methodology**	**Mean RMSE**	**Percentage Improvement (%)**
1.	M5P +FS	69.03	-
2.	J48	192.68	64.18
3.	J48 + FS	150.07	54.01
4.	Bagging	114.05	39.48
5.	Bagging + FS	109.04	36.70
6.	M5P	100.71	31.46

The daily MAPEs corresponding to M5P and M5P + FS are calculated in Table 7. The graphical representations of daily MAPE for the all seasons are depicted in Figures 7–9. These results show that the performance of M5P+ FS is completely superior to the performance of M5P.

Table 7. Daily MAPE for all seasons corresponding to M5P and M5P + FS.

Sr. No.	(1–7 Aug 2015) Winter		(1–7 Sep 2015) Spring		(1–7 Feb 2016) Summer	
	M5P	M5P + FS	M5P	M5P + FS	M5P	M5P + FS
1	0.95	0.69	1.12	0.53	0.66	0.63
2	1.60	0.92	0.98	0.54	0.61	0.54
3	1.26	0.94	1.00	0.66	0.48	0.49
4	0.92	0.60	0.92	0.60	0.58	0.63
5	1.09	0.58	0.80	0.71	0.71	0.65
6	0.90	0.48	0.83	0.82	0.65	0.61
7	0.79	0.46	1.26	1.07	0.75	0.73
	5.25	1.07	0.67	0.99	0.70	0.64

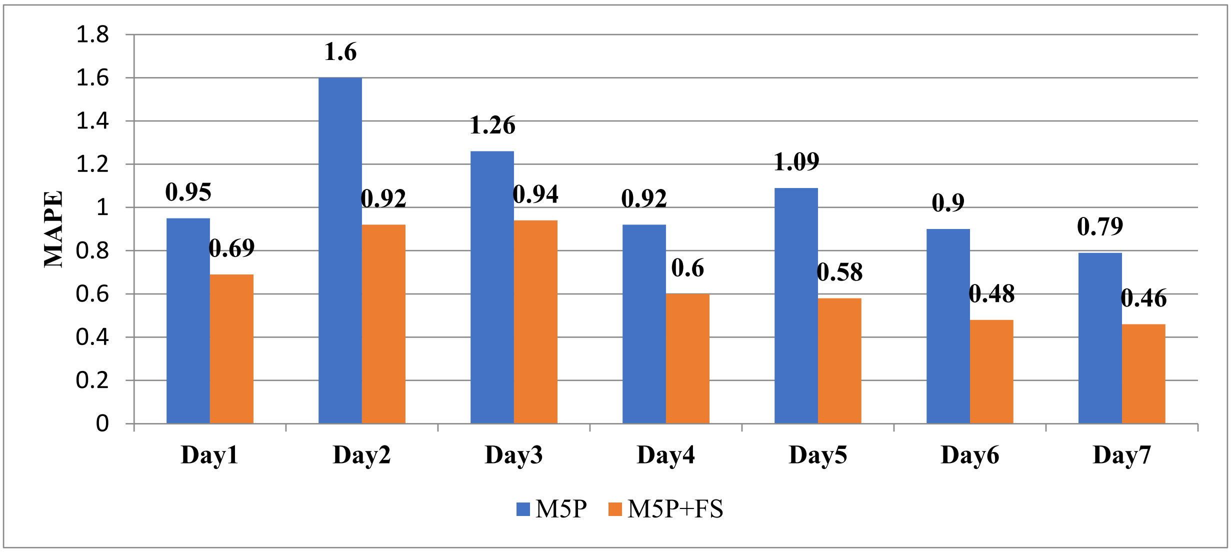


Figure 7. Representation of daily MAPE for winter season corresponding to M5P and M5P + FS.

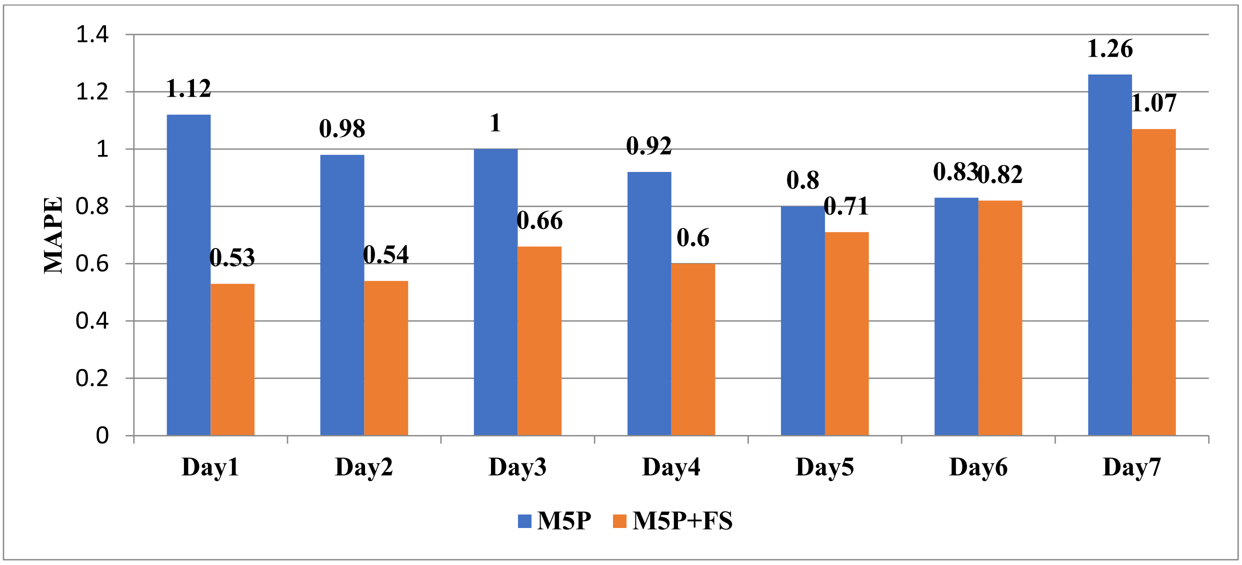


Figure 8. Representation of daily MAPE for spring season corresponding to M5P and M5P + FS.

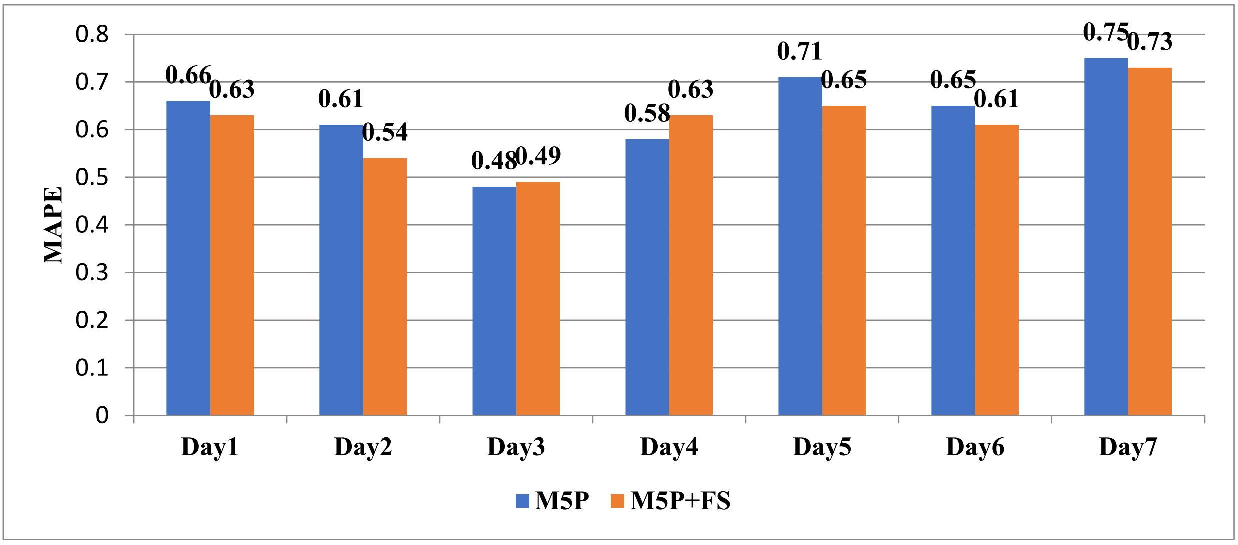


Figure 9. Representation of daily MAPE for summer season corresponding to M5P and M5P + FS.

To validate the proposed model, the MAPE values presented in this work were compared with those reported in [42] using another method and considering similar data sets. This will certainly help readers to enhance their understanding of this work since MAPE is one of the most commonly used key performance indicators to measure forecast accuracy (i.e., the lower the MAPE, the higher the forecast accuracy). The results listed in Table 8 clearly show that the proposed method performed better than previously reported methods.

Table 8. Validation of proposed method.

Sr. No.	Duration	Methodology	MAPE
1	1–7 December 2015	Random Forest [42]	1.02
2		Proposed Algorithm (M5P + FS)	0.70

6. Conclusions

In this paper, a day-ahead STLF employing M5P and a novel HFS approach based on EGA and the random forest method was presented. STLF was implemented for a whole year (in a week-wise manner for each day and for all seasons) with FS and WoFS. Performance measures such as MAPE, MAE, EV and RMSE were computed season-wise, week-wise and day-wise. The proposed methodology (M5P + FS) consists of two stages; i.e., in the first stage, FS is performed using the HFS algorithm and then in the second stage, forecasting is carried out by forecasters (M5P, Bagging and J48). For STLF with FS and WoFS, the results obtained with the M5P forecaster model were been compared with those obtained with J48 and Bagging. It is evident from the simulation results that the FS approach provides better short-term load forecasts over the WoFS approach and M5P outperforms J48 and Bagging. It is also evident that M5P + FS can offer improved accuracy (MAPE) in the range of 34.65 (for Bagging) to 59.51 (for J48).

Author Contributions: Software, methodology, conceptualization, investigation, validation and data curation, A.S.P. and A.K.S.; formal analysis, writing—original draft, supervision and visualization, A.K.S., V.K., S.M.T., A.S.P. and R.M.E.; editing and writing—review S.M.T., A.K.S., D.K., S.G. and M.A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The half hourly historical load data of New South Wales, Australia taken from Australian Energy Market Operator (AEMO)) and weather data of Sydney City (www.weatherzone.com.au, accessed on 14 December 2022) has been taken for the STLF.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

AEMO	Australian Energy Market Operator
ANN	Artificial Neural Network
ARMA	Auto Regressive Moving Average
CLC	Closed Loop Clustering
CNN	Convolutional Neural Network
DE	Differential Evolution
DRBFNNs	Decay Radial-Basis Function Neural Networks
DRM	Dynamic Regression Model
ELM	Extreme Learning Machine
EGA	Elitist Genetic Algorithm
EV	Error Variance
FCV	Fold Cross-Validation
FS	Feature Selection
GA	Genetic Algorithm
GENCOs	Generation Companies
HFS	Hybrid Feature Selection
HWT	Holt Winters Taylor
IEMD	Improved Empirical Mode Decomposition
IFS	In Function Systems
LM	Levenberg Marquardt
LR	Lasso Regression
LSTM	Long Short Term Memory
MABC	Multi-Species Artificial Bee Colony
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MLP	Multi Layer Perception
ML	Machine Learning
NSW	New South Wales
RMSE	Root-Mean Square Error
SDR	Standard Deviation Reduction
STLF	Short-Term Load Forecasting
SVD	Singular Value Decomposition
SVR	Support Vector Regression
SVM	Support Vector Machine
WoFS	Without Features Selection
WT	Wavelet Transform
10-FCV	10 Fold Cross Validation

References

1. Muratori, M.; Rizzoni, G. Residential demand response: Dynamic energy management and time-varying electricity pricing. *IEEE Trans. Power Syst.* **2016**, *31*, 1108–1117. [CrossRef]
2. Bunn, D.W. Forecasting loads and prices in competitive powermarkets. *Proc. IEEE* **2000**, *88*, 163–169. [CrossRef]
3. Voronin, S.; Partanen, J. A Hybrid electricity price forecasting model for the Finnish electricity spot market. In Proceedings of the 32st Annual International Symposium on Forecasting, Boston, MA, USA, 24–27 June 2012.
4. Yun, Z.; Quan, Z.; Caixin, S.; Shaolan, L.; Yuming, L.; Yang, S. RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment. *IEEE Trans. Power Syst.* **2008**, *23*, 853–858.
5. Mund, C.; Rathore, S.K.; Sahoo, R.K. A review of solar air collectors about various modifications for performance enhancement. *Sol. Energy* **2021**, *228*, 140–167. [CrossRef]
6. Sobhani, M.; Campbell, A.; Sangamwar, S.; Li, C.; Hong, T. Combining weather stations for electric load forecasting. *Energies* **2019**, *12*, 1510. [CrossRef]
7. Melissa, D.; Binita, K.; Colin, C.I. *Extreme Weather and Climate Vulnerabilities of the Electric Grid: A Summary of Environmental Sensitivity Quantification Methods*; No. ORNL/TM-2019/1252; Oak Ridge National Lab. (ORNL): Oak Ridge, TN, USA, 2019.
8. Dar-Mousa, R.N.; Makhamreh, Z. Analysis of the pattern of energy consumptions and its impact on urban environmental sustainability in Jordan: Amman City as a case study. *Energy Sustain. Soc.* **2019**, *9*, 15. [CrossRef]

9. Tso, G.K.; Yau, K.K. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy* **2007**, *32*, 1761–1768. [CrossRef]
10. Alghandoor, A.; Phelan, P.; Villalobos, R.; Phelan, B. US manufacturing aggregate energy intensity decomposition: The application of multivariate regression analysis. *Int. J. Energy Res.* **2008**, *32*, 91–106. [CrossRef]
11. Zaza, F.; Paoletti, C.; LoPresti, R.; Simonetti, E.; Pasquali, M. Multiple regression analysis of hydrogen sulphide poisoning in molten carbonate fuel cells used for waste-to-energy conversions. *Int. J. Hydrogen Energy* **2011**, *36*, 8119–8125. [CrossRef]
12. Kumar, U.; Jain, V. Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. *Energy* **2010**, *35*, 1709–1716. [CrossRef]
13. Rumbayan, M.; Abudureyimu, A.; Nagasaka, K. Mapping of solar energy potential in Indonesia using artificial neural network and geographical information system. *Renew. Sustain. Energy Rev.* **2012**, *16*, 1437–1449. [CrossRef]
14. Matallanas, E.; Castillo-Cagigal, M.; Gutiérrez, A.; Monasterio-Huelin, F.; Caamaño-Martín, E.; Masa, D.; Jiménez-Leube, J. Neural network controller for active demand-side management with PV energy in the residential sector. *Appl. Energy* **2012**, *91*, 90–97. [CrossRef]
15. Cheong, C.W. Parametric and non-parametric approaches in evaluating martingale hypothesis of energy spot markets. *Math. Comput. Model.* **2011**, *54*, 1499–1509. [CrossRef]
16. Wesseh, P.K., Jr.; Zoumara, B. Causal independence between energy consumption and economic growth in Liberia: Evidence from a non-parametric bootstrapped causality test. *Energy Policy* **2012**, *50*, 518–527. [CrossRef]
17. Grzegorz, D. Pattern-based local linear regression models for short-term load forecasting. *Electr. Power Syst. Res.* **2016**, *130*, 139–147.
18. Cecati, C.; Kolbusz, J.; Rozycki, P.; Siano, P.; Wilamowski, B.M. A Novel RBF Training Algorithm for Short-Term Electric Load Forecasting and Comparative Studies. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6519–6529. [CrossRef]
19. Zhai, M.-Y. A new method for short-term load forecasting based on fractal interpretation and wavelet analysis. *Int. J. Electr. Power Energy Syst.* **2015**, *69*, 241–245. [CrossRef]
20. Arora, S.; Taylor, J.W. Short-Term Forecasting of Anomalous Load Using Rule-Based Triple Seasonal Methods. *IEEE Trans. Power Syst.* **2013**, *28*, 3235–3242. [CrossRef]
21. Zeng, P.; Jin, M.; Elahe, M.F. Short-Term Power Load Forecasting Based on Cross Multi-Model and Second Decision Mechanism. *IEEE Access* **2020**, *8*, 184061–184072. [CrossRef]
22. Nose-Filho, K.; Lotufo, A.D.P.; Minussi, C.R. Short-Term Multinodal Load Forecasting Using a Modified General Regression Neural Network. *IEEE Trans. Power Deliv.* **2011**, *26*, 2862–2869. [CrossRef]
23. Zhang, C.; Li, R. A Novel Closed-Loop Clustering Algorithm for Hierarchical Load Forecasting. *IEEE Trans. Smart Grid* **2021**, *12*, 432–441. [CrossRef]
24. Rafi, S.H.; Nahid-Al-Masood; Deeba, S.R.; Hossain, E. A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network. *IEEE Access* **2021**, *9*, 32436–32448. [CrossRef]
25. Song, L.; Peng, W.; Lalit, G. Short-term load forecasting by wavelet transform and evolutionary extreme learning machine. *Electr. Power Syst. Res.* **2015**, *122*, 96–103.
26. Kouhi, S.; Keynia, F.; Ravadanegh, S.N. A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection. *Int. J. Electr. Power Energy Syst.* **2014**, *62*, 862–867. [CrossRef]
27. Ungureanu, S.; Topa, V.; Cziker, A.C. Analysis for Non-Residential Short-Term Load Forecasting Using Machine Learning and Statistical Methods with Financial Impact on the Power Market. *Energies* **2021**, *14*, 6966. [CrossRef]
28. Luo, J.; Hong, T.; Yue, M. Real-time anomaly detection for very short-term load forecasting. *J. Mod. Power Syst. Clean Energy* **2018**, *6*, 235–243. [CrossRef]
29. Jiao, R.; Zhang, T.; Jiang, Y.; He, H. Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network. *IEEE Access* **2018**, *6*, 59438–59448. [CrossRef]
30. Haq, M.R.; Ni, Z. A new hybrid model for short-term electricity load forecasting. *IEEE Access* **2019**, *7*, 125413–125423. [CrossRef]
31. Deng, Z.; Wang, B.; Xu, Y.; Xu, T.; Liu, C.; Zhu, Z. Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting. *IEEE Access* **2019**, *7*, 88058–88071. [CrossRef]
32. Hong, Y.; Zhou, Y.; Li, Q.; Xu, W.; Zheng, X. A deep learning method for short-term residential load forecasting in smart grid. *IEEE Access* **2020**, *8*, 55785–55797. [CrossRef]
33. Ahmad, W.; Ayub, N.; Ali, T.; Irfan, M.; Awais, M.; Shiraz, M.; Glowacz, A. Towards short term electricity load forecasting using improved support vector machine and extreme learning machine. *Energies* **2020**, *13*, 2907. [CrossRef]
34. Pei, S.; Qin, H.; Yao, L.; Liu, Y.; Wang, C.; Zhou, J. Multi-step ahead short-term load forecasting using hybrid feature selection and improved long short-term memory network. *Energies* **2020**, *13*, 4121. [CrossRef]
35. Xuan, Y.; Si, W.; Zhu, J.; Sun, Z.; Zhao, J.; Xu, M.; Xu, S. Multi-model fusion short-term load forecasting based on random forest feature selection and hybrid neural network. *IEEE Access* **2021**, *9*, 69002–69009. [CrossRef]
36. Ijaz, K.; Hussain, Z.; Ahmad, J.; Ali, S.F.; Adnan, M.; Khosa, I. A Novel Temporal Feature Selection Based LSTM Model for Electrical Short-Term Load Forecasting. *IEEE Access* **2022**, *10*, 82596–82613. [CrossRef]
37. Zhang, S.; Zhang, N.; Zhang, Z.; Chen, Y. Electric Power Load Forecasting Method Based on a Support Vector Machine Optimized by the Improved Seagull Optimization Algorithm. *Energies* **2022**, *15*, 9197. [CrossRef]

38. Liu, M.; Qin, H.; Cao, R.; Deng, S. Short-Term Load Forecasting Based on Improved TCN and DenseNet. *IEEE Access* **2022**, *10*, 115945–115957. [CrossRef]
39. Yi, H.S.; Lee, B.; Park, S.; Kwak, K.C.; An, K.G. Prediction of short-term algal bloom using the M5P model-tree and extreme learning machine. *Environ. Eng. Res.* **2019**, *24*, 404–411. [CrossRef]
40. Kisi, O.; Shiri, J.; Demir, V. Hydrological time series forecasting using three different heuristic regression techniques. In *Handbook of Neural Computation*; Academic Press: Cambridge, MA, USA, 2017; pp. 45–65.
41. Srivastava, A.K.; Pandey, A.S.; Elavarasan, R.M.; Subramaniam, U.; Mekhilef, S.; Mihet-Popa, L. A Novel Hybrid Feature Selection Method for Day-Ahead Electricity Price Forecasting. *Energies* **2021**, *14*, 8455. [CrossRef]
42. Dudek, G. A Comprehensive Study of Random Forest for Short-Term Load Forecasting. *Energies* **2022**, *15*, 7547. [CrossRef]