SHORT TERM LOAD FORECASTING TECHNIQUE BASED ON INTEGRATION OF CONVOLUTIONAL NEURAL NETWORK AND LONG-SHORT-TERM MEMORY NETWORK

SHAFIUL HASAN RAFI

(*BSc Engg., RUET*)

A THESIS SUBMITTED FOR THE DEGREE OF

MASTER OF SCIENCE IN
ELECTRICAL, ELECTRONIC AND COMMUNICATION ENGINEERING

DEPARTMENT OF ELECTRICAL, ELECTRONIC AND COMMUNICATION ENGINEERING

MILITARY INSTITUTE OF SCIENCE AND TECHNOLOGY

2020

APPROVAL CERTIFICATE

The thesis titled **"Short Term Load Forecasting Technique Based on Integration of Convolutional Neural Network and Long-Short-Term Memory Network"** submitted by Shafiul Hasan Rafi, Student No. - 1017160007 (P), Session: 2016-2017 has been accepted as satisfactory in partial fulfillment of the requirements for the degree of Master of Science in Electrical, Electronic and Communication Engineering on 26 August 2020.

BOARD OF EXAMINERS

1. __________________________________ Chairman (Supervisor)
   Dr. Nahid-Al-Masood
   Associate Professor
   Dept. of EEE, BUET, Dhaka - 1205

2. __________________________________ Member (Co-Supervisor)
   Major Md. Ali Azam Khan, EME
   Instructor Class B
   Dept. of EECE, MIST, Dhaka - 1216

3. __________________________________ Member (Ex-officio)
   Brigadier General A K M Nazrul Islam, PhD
   Senior Instructor and Head
   Dept. of EECE, MIST, Dhaka - 1216

4. __________________________________ Member (Internal)
   Lieutenant Colonel Md. Tawfiq Amin, PhD, EME
   Instructor Class A
   Dept. of EECE, MIST, Dhaka - 1216

5. __________________________________ Member (External)
   Dr. Tareq Aziz
   Professor
   Dept. of EEE, AUST, Dhaka – 1208

CANDIDATE'S DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. The thesis has not been submitted partially or fully for any degree or diploma in any university or institute previously.

Signature ____________________

(Shafiul Hasan Rafi)

Date: 26 Aug 2020

DEDICATION

To My Parents and Family

ABSTRACT

In this thesis work, a new technique is proposed to forecast short term electrical load. Load forecasting is an integral part of power system planning and operation. Precise forecasting of load is essential for unit commitment, capacity planning, network augmentation and demand side management. It is significantly imperative for energy providers and other members in electric energy generation, transmission, distribution and markets. Forecasting of load demand is a complex problem as it is to solve nonlinearity with influenced external factors. Load forecasting can be generally categorized into three classes such as sort term, midterm and long term. Short term forecasting is usually done to predict load for next few hours to few weeks. In the literature various methodologies such as regression analysis, machine learning approaches, deep learning methods and artificial intelligence systems have been used for short term load forecasting. Existing forecasting techniques may not always provide higher accuracy in short term load forecasting. To overcome this challenge, a new approach is proposed in this thesis for short term load forecasting. The developed method is based on the integration of convolutional neural network and long short-term memory network. The method is applied to Bangladesh power system to provide day ahead forecasting to month ahead. It is found that in the field of short-term load forecasting, the proposed strategy results in higher precision and accuracy in terms of Mean average error (MAE), Mean average percentage error (MAPE) and root mean square error (RMSE).

ACKNOWLEDGEMENTS

This thesis is the most significant accomplishment in my life. I would like to thank my parents for their continuous support, encouragement and sacrifice throughout the period and I will be indebted to them forever for all they have done. Similarly, I could not find words that express my gratefulness to my wife Sharmin Sultana, for her supportive and understanding attitude.

This is my immense pleasure to express my sincere and profound gratitude to my supervisor Dr. Nahid-Al-Masood, Associate Professor, EEE Dept, BUET for the patient guidance, encouragement and advice he has provided throughout my research time. I am grateful to him for giving me directions, confidence and valuable advice, and for everything that I have learned from him. He has had a tremendous impact on the way I perceive research and life in general. Dr. Nahid-Al-Masood has also been very helpful in many ways to me and for my research. His feedback from different critical angles has always been encouraging and useful. He also has taught me many practical aspects of research from the elementary level. Once again I thank my supervisor, with whom I feel that I genuinely have been privileged to work.

I profoundly thank Brigadier General A K M Nazrul Islam, PhD, Head, Dept. of EECE, MIST for his valuable and constructive feedback, assistance and being on my final defense panel. I thank Lieutenant Colonel Md Tawfiq Amin, PhD, EME for kind coordination and cooperation as the postgraduate course coordinator. Also, thanks to all the reviewers, for their appreciation and feedback on my thesis and publications

I would be pleased to extend my sincere thanks to all of my course teachers and staffs of EECE department, MIST for their cordial help and adequate support for successful completion of my research works. Special thanks to Major Wahed, Major Mohazzem, Assistant Professor Soyaeb and Mr. Hafizur Rahman for their cordial assistance and official support to complete the administrative procedures.

Although many of you may not understand, nor ever want to understand, much of what is written beyond this page, you should at least know that it could not have been done without you.

SHAFIUL HASAN RAFI

Military Institute of Science and Technology

Dhaka, Bangladesh

Aug 2020

TABLE OF CONTENTS

LIST OF SYMBOLS

D	Similar Day approaches
F_W	The recursive function.
f_t	Forget
f	Activation function.
h_t	The hidden state
i_t	Input gate
N	Total number of the observation periods
O_t	Output gate
P_t	Actual load of a system
S_t	State at time t
x_t	Input at time
$\breve{Y}_L$	Actual load value
Y_L	Predicted value
σ	Sigmoid activation function
ϵ_t	Random factor

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Normal
ARMAX	Autoregressive Integrated Moving Average with Exogenous Variables
BPS	Bangladesh Power System
CNN	Convolutional Neural Network
C-LSTM	Cycle based Long Short Term Memory Network
DNN	Deep Neural Network
GA	Genetic Algorithms
GRUN	Gated Recurrent Unit Network
LSTM	Long Short Term Memory Network
LTLF	Long Term Load Forecasting
MAPE	Mean Absolute Percentage Error
MTLF	Midterm Load Forecasting
MAE	Mean Average Error
PGCB	Power Grid Company Bangladesh
RNN	Recurrent Neural Network
RELU	Rectified Linear Unit
RMSE	Root Mean Square Error
STLF	Short Term Load Forecasting
SVM	Support Vector Regression Machine
TD-CNN	Time Division Convolutional Neural Network

CHAPTER 1
INTRODUCTION

1.1 Introduction

In the field of the electrical power system network, load forecasting is essential for appropriate planning, operation and control. Without precise prediction of load, power scheduling and unit commitment can’t be done in a proper way. Load forecasting is particularly very important research issue for the achievement of getting higher efficiency and reliability in power system operation, transmission and control. Price forecasting of load can minimize the operational cost by giving correct input to the prior days scheduling. Power system reliability, load flow analysis, planning for the power system operation, transmission and distribution facilities also depend on precise load forecasting.

1.2 Electric Load Forecasting

In our daily life electricity plays a very important role. The economic development of any country is subjected to the electricity consumption. Designing, planning and operation of electric utilities are greatly depending on precise load forecasting procedure. In addition, it is indispensable to distribute substantial extent of electricity to increase the economy considerably by electric power suppliers. Based on the existing system information, load forecasting is the way of assessing the future electric load trend that is the calculation of future demand for a given perspective. Increment of power consumption in any country is greatly depending on the population growth, economic infrastructure development, fastest transportation system and human facilities. According to the interval period of time, load forecasting procedure is mostly classified into three main categories. Furthermore, it can also be divided in to the subjective and objective load forecasting methods. Fist one is the subjective methods that are usually used to measure either individual or group opinion. These forecasting methods involve:

a. Scales force aggregates.
b. Customer assessment.
c. Jury of supervisory judgment.
d. The Delphi method.

The second one is the objective forecasting methods that can be divided in two bodies such as time series methods that are computed using past history data and regression analysis. Regression models often integrate the previous history of other series. On the contrary, time series forecasting methodology is used to predict the future load based on previously observed electric load values which can easily be integrated into a computer program such that machine learning algorithm for getting automatic forecasting outcomes. The ultimate objective of applying time series forecasting method is to find predictable and repeatable patterns including increasing or decreasing linear trend, curvilinear trend and seasonal change in historical past data. However, now-a-days electricity load forecasting is important research issue because it is exceptionally critical task to achieve highly efficient and reliable the power system of a country with regards to demand response and the allocation of available resources.

1.3 Types of Load Forecasting

Load forecasting is generally partitioned into three categories. These are short term load forecasting (STLF), which predicts the load of next few hours to few weeks; midterm load forecasting (MTLF), which basically covers a week to a year; and long term load forecasting (LTLF), which predicts load for more than a year [1]- [2]. Some factors are usually considered for short term load forecasting such as input load flow study, energy transfer scheduling, and demand-side management in the daily operation [3]. Day-ahead load scheduling, which is imperative for power system operation and control, is done through short term load forecasting. In order to correlate the predicted advancement in demand, the medium and long-term forecasting are applied for enlargement of capacity of generation, transmission and distribution. Different kinds of load forecasting are illustrated in Table 1.1

Table 1.1: Categories of Load Forecasting

	Time Intervals	**Forecasted Outcomes**	**Accuracy**	**Operation**	**Planning**
STLF	24h-1week	Load curves	Fixed load curves	Economics load dispatch	Unit commitment
MTLF	1 week-1 year	Load curves	Capacity>> Error	Unit Commitment	Reserve planning
LTLF	>1 year	Energy needed	Fixed energy	Planning of the power system	Future capacity extension

Precise load forecasting is very much needed for mitigating the challenge with the upward trend in electricity market prices. Energy producer or suppliers are always dependent on the future demand to reduce power generation cost. Thereby, it is imperative to make a balance between power consumption and power supply. With the development of power market over the past decades, it has been appeared that the short term load forecasting takes part a vital role in Energy management system (EMS).

1.4 Important Factors for Load Forecasting

There are some factors which must be considered for short-term load forecasting like time factors, weather changing data, and classes of the consumers [4]. The medium- and long-term load forecasting consider the historical weather and load data, the number of different consumers, the appliances used in a particular place and their characteristics in different periods, the economic and demographic data and their forecasts, sales data of electric home appliances, and other factors. The most prominent factor to be specific the time factors take account of the time of the year, the day of the week, and the hour of the day [5]. Electric load between weekdays and weekends are not similar. Moreover, the load on different weekdays has different characteristics. Therefore, to forecast the future demand a mathematical model is needed using the proper selection of appropriate forecasting variables. Quality of the input variable signifies the accuracy of the model which would predict the future load. There are two types of variables such as deterministic and stochastic which affects the load forecasting. On the contrary, electric energy consumption is greatly depended on the human activities. Similarly, population and their economic status signify the human activities. Therefore, the factor affecting variables are correlated to each other. Sometimes the change of the comfort feeling of the consumers may changes the condition of the weather. For example, the necessity of the heaters (water, room) and air conditioner affect the weather conditions.

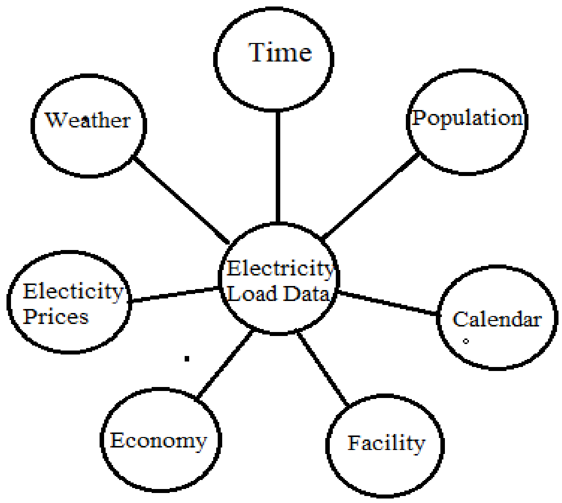

Fig. 1.1. Various factors affecting load forecasting

In Fig. 1.1, it is clear that there are many influencing factors affected on the electricity load demand. Firstly, time is one of the most popular factors that can be separated into midnight, morning, evening, night, lunch time and so on when it is considered for one day. Thus, it can be forecasted next day electricity demand because of available different past data for one day. Similarly, it can be considered seasonally, yearly, monthly, weekly and daily etc. Next, weather is also one important thing affected on load, for example, it can be diverged temperature, cloud cover or sunshine, humidity and so on. Moreover, calendar can also be considered like seasonal variation, daily variation, weekly cyclic and holidays, hence it can be get different data and predict different results. Likewise, the remaining things such as population, human facilities, economic for business and electricity prices are influenced on forecasting electricity load demand.

1.5 Literature Review

There are a very few reports to determine the load of Bangladesh Power System (BPS). A long term electricity demand of Bangladesh was projected from 2005 to 2035 considering the economic growth scenario. This study did not incorporate any methodology for short term load forecasting procedure. Some regional basis reports have already been focused for the issue of the load forecasting of BPS where different horizon load forecasting like daily, weekly, monthly, yearly is not presented [6]-[7]. In spite of the fact that there are a few studies of whole country which incorporates load forecasting of BPS by applying a conventional linear regression strategy using temperature variations [8].

The STLF issue has been dealt with different procedures [9]-[10]. These procedures can be approximately classified into two gatherings, to be specific traditional and artificial intelligence systems. In the early literature, the most regularly used methodology is statistical method [11], which includes multiple linear regression [12]-[14], exponential smoothing [15], and the autoregressive coordinates moving normal (ARIMA) [16]. In any case, due to the characteristics of non-linear features of the time series univariate load data, above mentioned techniques perform ineffectively in terms of mean square error (MAE), mean average percentage error (MAPE) and root mean square error (RMSE) to address the STLF [17]-[18].

In recent days , Intelligence based forecasting scheme made extraordinary progress to solve the STLF because of having the non-linear learning capabilities which includes clustering strategies [19] , fuzzy logic framework [20], support vector machine (SVM) [21] , artificial neural networks (ANN) [22]-[28], recurrent neural network (RNN) [29] and hybrid methods [30]-[31]. In [32], an effective method based on ANN strategies fortified by a wavelet denoising algorithm is developed to predict the short-term load demand. The obtained results from the proposed approach exhibit that it considerably increases the accurateness of prediction. In recent times, deep learning approaches have drawn a special concern because of having more number of hidden layers which enables this model to deal with the complicated non-linear patterns [33]-[34]. Due to have the effective learning ability to capture the non-stationary and long term dependent load data pattern, RNN are most frequently used in the field of load prediction [35]. In [36], a RNN algorithm is implemented for household load forecasting which gives better performance in terms of root mean square error (RMSE). A modern load forecasting methodology using RNN is introduced in [37] which employments a concept of one-step-ahead. The proposed method gives outstanding performance in low power demand and high power demand region. It also exhibits the tinier fluctuations in different region as compared to the other models. After all, vanishing gradient and exploding gradient problem arise in RNN which reduce the prediction accuracy. To solve this issue, LSTM network has been implemented based on past information and obtained satisfactory outcomes in the long-term load forecasting arena [38]-[39]. In [40], an effective methodology using LSTM network is developed to make a precise forecasting procedure that is capable of handling more complex time series load data with long-term dependencies. The proposed technique outstrips any other model in the complex electrical load forecasting arena. Moreover, recurrent unit neural network (GRUN) has been extensively used in the recent years due to the absence of vanishing gradient problem. In [41], a GRUN based algorithm is proposed for STLF procedure with multi-source data. The mean average percentage error (MAPE) obtained from this network is minimum, which outperforms other current methods.

In conjunction with the above illustrative techniques, the convolutional neural network (CNN) has been frequently used in the field of load prediction because of the ability to capture the electric demand local trend features when the two adjacent data points have strong relationship [42]. Local trend of the load data pattern in adjacent hours can be extracted by CNN. In [43], a modern load forecasting methodology based on CNN model is introduced and given significant comparison with various artificial neural networks. Obtained outcomes appeared that the error of the proposed network is smallest among all models. From this experiment it tends to be reasoned that the CNN is especially viable in the field of load prediction and its hidden feature can be extricated by designing one dimensional convolution layer.

Moreover, for improving the forecasting accuracy, time-dependency convolutional neural network (TD-CNN) and cycle-based LSTM (C-LSTM) network has been implemented in the domain of STLF [44]. In [45], time-cognition CNN (TCMS-CNN) based multi-step STLF procedure is proposed which significantly improves the prediction ability after extracting substantial and complex features from the electric load sequences. It serves additional precision of the forecasted results and appears excellent constancy in multi-step time series and probabilistic forecasting, giving robust simplification in electricity market bidding and spot price calculation.

In light of the above writing, LSTM and CNN are both illustrated to afford high precision results in STLF because of their preferences to extract hidden features from load sequences. Hence, it is preferred to implement a hybrid neural network based on CNN-LSTM architecture for addressing the issue of STLF for BPS that can be able to capture and integrate that kind of hidden features to give a stable performance.

1.6 Thesis Objective

The specific aims of the research work are as follows:

a. To develop CNN-LSTM based hybrid forecasting framework for STLF.

b. To increase the prediction accuracy with the integration of the hidden features of CNN model and LSTM model.

c. To achieve better and stable performance in STLF.

d. To apply the developed model on the historical load data of Bangladesh power system.

1.7 Thesis Outcome

The outcome of the research work is a CNN-LSTM network, which predicts the load demand of BPS with the help of the historical load data. Load forecasting results demonstrate the superiority of the proposed methodology by comparing with existing techniques.

1.8 Thesis Outlines

The remainder of this thesis is organized as follows.

Chapter 2 highlights the brief overview of short term load forecasting techniques and their significance.

Chapter 3 mainly describes the electric load characteristic of BPS. How the load data is formulated for problem design is also discussed in this chapter.

Chapter 4 mainly focused on the problem formulation and proposed methodologies for short term load forecasting. Details architecture of the CNN-LSTM model design is explained in this chapter.

Chapter 5 contains the results obtained from the LSTM and CNN-LSTM based forecasting models. An extensive comparison in terms of certain performance metrics and in terms of detailed plot of predicted and actual values is presented in this section. Observations from the result as well as an analysis of the results observed are also included.

Chapter 6 serves as the concluding remarks of the thesis. This chapter details the future works and improvements that can be carried out on these models as well as suggests a way in which these models can be used in tandem with one another for a power system.

1.9 Summary

This chapter addresses the introduction, details of the research objectives and motivation in general. Types of electric load forecasting, important factors of affecting load forecasting are also analyzed in this chapter. This chapter covers the recent previous works done by several researchers all over the world and describes the most modern method of load forecasting. Then the performance evaluation matrices of discussed methodology are stated clearly here. In the next chapter, various types of the short-term load forecasting techniques and their significance will be discussed.

CHAPTER 2
SHORT-TERM LOAD FORECASTING TECHNIQUES

2.1 Introduction

Various strategies, which incorporate similar day approach, different regression approaches, time series forecasting, neural network, fuzzy logic, expert system, and statistical approaches, hybrid models are utilized for STLF. Time series component can be separated into four parts: trend, seasonal, cyclic and random. In trend component, it can be persistent and it indicates upward or downward trends. It can change due to technologies, population, age, culture, etc. and it can occur generally within several years' duration. A brief description of these techniques is presented below.

2.2 Forecasting Model Classification

Forecasting techniques are categorized into two main groups, i.e., traditional statistic models and artificial intelligence (AI) based models for STLF. Regression analysis, moving average, exponential smoothing, and stochastic time series models etc. are included in traditional statistical models. AI-based models include machine learning, data mining, genetic algorithms, artificial neural networks (ANN), fuzzy logic and knowledge based expert systems. Various models of load forecasting are illustrated as follows.

2.2.1 Similar day approach

Based on the historical data searching similar day approached is used. It usually searches the data for days within one, two, or three years with similar natures to the forecast day. Some indexes such as date, day of the particular week and weather are considered as comparative individualities. Rather than considering a load of similar single day, this approach can include a regression analysis or a linear combination strategy with several similar days. This is usually clarified by the condition [3].

$$D = \sqrt{\widehat{W_1}(\Delta L_1)^2 + \widehat{W_2}(\Delta H_h)^2 + \widehat{W_3}(\Delta T_t)^2} \tag{2.1}$$

Where, ΔL_1 is denoted for the load deviation of the forecast day and historical days, ΔT_t temperature deviation on forecast day and historical days and ΔH_h is the humidity deviation between historical days and forecast day. Temperature and load data basically defines the weighted factor $\widehat{W_i}$ (i = 1, 2, 3,..). This factor is calculated using least square method which is constructed using regression analysis using historical load data and temperature. Therefore, similar days for load forecasting are selected by considering trends of load, humidity and temperature.

2.2.2 Statistical techniques

Statistical techniques rely on mathematical expressions, which give the relationship between load and several input factors. This kind of techniques can be divided into following groups

a. Multiple regression method
b. Exponential smoothing
c. Simple moving average
d. Time series approaches

2.2.2.1 Multiple regression method

It is a one kind of statistical techniques which is most frequently used in the field of STLF. In order to forecast electrical load, it is needed to build a relationship between load consumption and other factors such as day type index, weather condition, and lastly the classes of the customer. For making such relationship, various regression methods are generally used. Moreover, independent variables are selected based on the correlation coefficients between load and candidate independent variables [26]. A multiple regression equation capturing the load pattern is developed as follows [14].

$$y = c_0x_0 + c_1x_1 + \cdots + c_nx_n + e = \vec{c}.\,\vec{x}+e \tag{2.2}$$

Where, the electrical load is defined by y, x_j denotes the independent variable, regression coefficient is c_j for $j \in [0, \mathrm{n}]$, and e is the modeling error. The row vector $\vec{x} = [x_0, x_1, \ldots, x_n]$ contains independent variables which are affecting the load, and the column vector $\vec{c} = [c_0, c_1, \ldots, c_n]^t$ comprises of unknown regression coefficients, where a^t refers to a transposition of a.

2.2.2.2 Exponential smoothing

It is a traditional methodology which is regularly used for the issue of STLF. This models can be mathematically shown as follows.

$$Y_{t+1} = aX_t + (1-a)Y_t \tag{2.3}$$

Where, the forecast value at time t and t + 1 is Y_t and Y_{t+1}, respectively. Actual value at time t is X_t. The weight of a defines the most recent observation and a weight of (1- a) signifies the most recent forecast. $a = \frac{1}{N}$ is the smoothing factor, where N is the number of observations involved in the average operation. Using (3), it can be noticed that the value obtained from the forecast at time $(t+1)^{th}$ instantly depends on the actual value observed at time ‘t’ and the forecasted value of the time period ‘t’. In this way, the exponential smoothing strategy deploys previous observations and requires less computational memory.

2.2.2.3 Simple moving average (S.M.A)

This method utilizes the concept of moving average (i.e. rolling window based technique) to predict electrical load. Mathematically, the S.M.A. strategy can be represented using the following equation.

$$Y_{t+1} = \frac{1}{N}\sum_{i=t-N+1}^{t} X_i \tag{2.4}$$

Where, Y_{t+1} is the forecast value at time t+1, X_i is the observation at time t and N is the number of terms included in the average. This strategy can be utilized for forecasting up to one, two, or three time periods. However, it is generally utilized for forecasting for one time period in progress because forecasting error increases with increases in time period [9].

2.2.2.4 Time series approach

Time series approaches are basically assumption techniques which comprises of the internal factors such as seasonal trend and auto correlation. These strategies identify and make a relation between such internal factors. Over the past decades, time series approach has been commonly used in the field of prediction such as electric load forecasting. In specific, classical time series strategies include ARMA (autoregressive moving normal), ARIMA (autoregressive integrated moving average), ARMAX (autoregressive moving average with exogenous variables), and ARIMAX (autoregressive integrated moving average with exogenous variables) models are most frequently used in STLF. ARMA model stands for stationary forms whereas an ARIMA is an expansion of ARMA which stands for non-stationary forms. Moreover, time and electrical load as input are used by both the model ARMA and ARIMA. Among all the classical models, ARIMAX is the most effective approach for STLF because of having the dependency of on weather and time of a day.

2.2.3 Fuzzy logic technique

Fuzzy logic is considered as a non-conventional load forecasting technique. The complete Fuzzy logic process is described below.

a. Fuzzification: Different inputs and output parameters which are often called fuzzy set numbers are defined by the variables of fuzzy logic. There may have nonlinearity between input and output parameters which can be simplified as linear by the membership functions.

b. Fuzzy rule base: "IF-THEN" rules are used for calculating the forecasting value.

c. Fuzzy inference system: This rule is employed to carry the data or information to fuzzy inference framework. It evaluates the data scrambled within the fuzzy rule base to create the output.

d. Defuzzification, Output obtained from the fuzzy inference system is considered as input to make an output of non-fuzzy.

2.2.4 Support vector machine (SVM)

SVM is a method which is fruitfully used to solve the regression and classification problems. This technique was originated from Vapnik's statistical learning theory. The most common reasons for the wide utilization of SVM in STLF are: hypothetical guarantees around their execution, lower susceptibility to local minima and higher resistance to expanded complexity of the model, which stands for adding extra extents to input parameters. So that, SVM method can develops a fast algorithm for forecasting problem and offer good outcomes for many kind of tasks. A SVM generally used quadratic or linear algorithm in the training process. Sometimes asymmetric loss function is used for training process. A hyperplane or set of hyperplane is used to setup SVM with an infinite dimensional space. A support vector machine usually sets up a hyperplane or set of hyperplanes in a high- or infinite dimensional space. In order to get good separation, a hyper plane with the largest distance to the nearest training data points of functional margin is used. The sets of hyper planes basically separate the nonlinearly in a finite dimensional space.

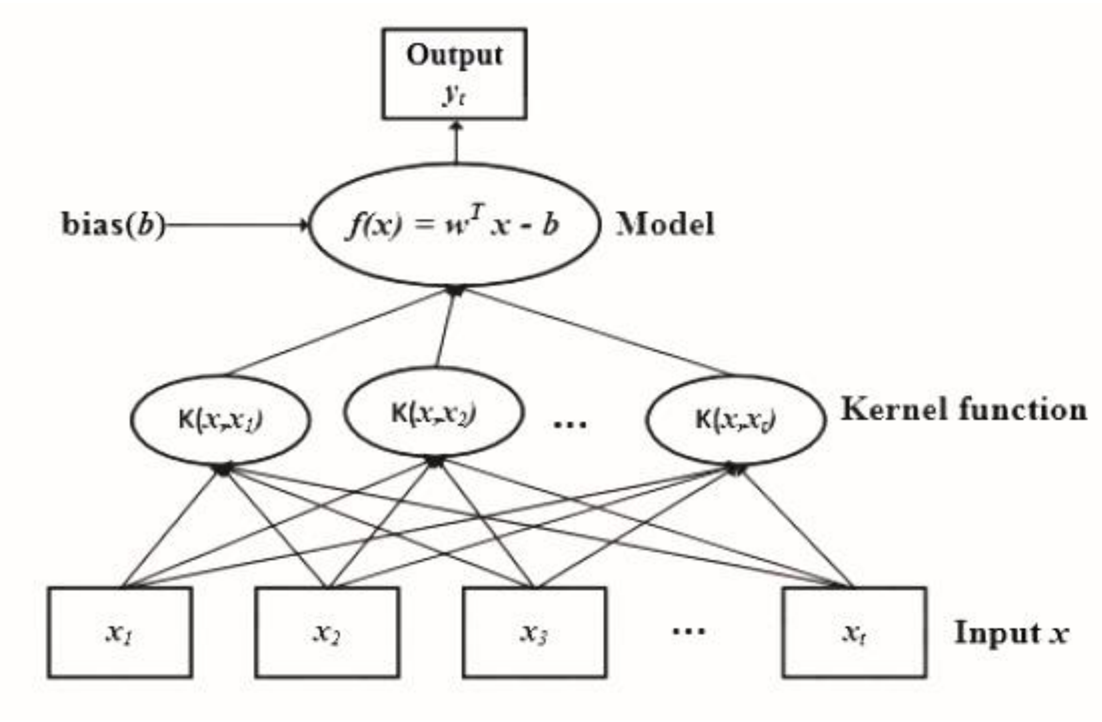

Fig. 2.1. Support vector machine structure. [10]

Fig. 2.1 demonstrates a basic support vector machine structure. Firstly, a kernel function is introduced to sole the large amount of solution function. In the whole training process an additional quadratic function is used to achieve the requirement of the SVM parameters. In the Figure 2.1, input *x* represents the input of training data and output *y* refers to a corresponding output value. The prime objective function of the SVM is to features a map with the high dimensional input from a nonlinear mapping and proposing a linear regression.

The objective concept of SVM is to map the input data into a high dimensional space from a non-linear mapping and conduct a linear regression.

2.2.5 Artificial neural network (ANN)

Biological nervous systems can be represented a as computational based neural network which is called Artificial neural network. A neural network has a basic component which is called neuron. The fundamental operation of the neuron is the handling the information such as historical data or inputs. The basic components of an ANN are as follows:

a. Weighting function.

b. Summation of the input signals

c. Various kinds of activation functions

An artificial neuron structure is illustrated in Fig. 2.2. The output of the neuron is mathematically presented by the following equation:

$$y = f(\sum_{n=1}^{N} X_n W_n + b_n) \quad (2.5)$$

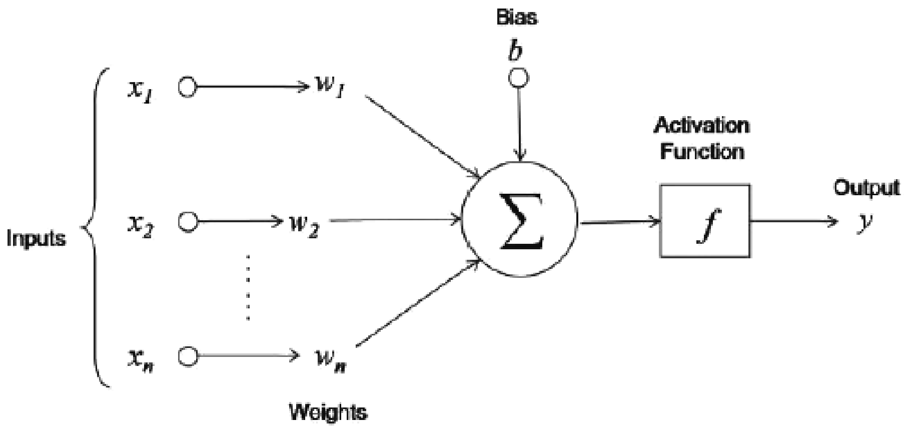

Fig. 2.2. An artificial neuron unit [28] .

Where, y is the output, x is the input, w is the weight and b is a numerical bias. Activation function f is usually restricted the output to value ranging from 0 to 1 or -1 to 1. The choice and inclusion of activation function in ANN is extremely important. The main contribution of the activation function is to introduce non-linear function mapping between response variable and inputs. Alternatively, it can be thought that, a NN without having non-linear activation function it is considered as like as a single-layer perceptron. Input to networks are usually linear transformation (input x weight), but real world problems are non-linear. A particular neural network function is determined by the presence of an activation function. So that it is a called a decision making function of a particular neural structure. This feature mapping is usually laid between 0 and 1, where 0 represent the absent of features whereas 1 signifies the presence of the feature. As a result, small change in the weighting function cannot be affected the activation values because it only takes the value either 0 or 1.

Therefore, necessity of having continuous nonlinear functions arises between this ranges which must be differentiable. A neural network must have the capabilities of taking any input value from ∞ to ∞. There are several different activation functions that are used in a neural network. For the sake of brevity, we will only look at three activation functions that are used in the LSTM architecture we used:

Sigmoid: The sigmoid function is defined as:

$$f(x) = \frac{1}{e^{x} - e^{-x}} \tag{2.6}$$

The function is illustrated in the Fig. 2.3 where it is evident that the output varies from 0 to 1. This function although used in many Neural Network calculations, it is slow to compute when used in a computationally demanding task. In many cases the high-precision exponential results aren't needed, and an approximation will suffice. Such is the case in many forms of gradient-descent/optimization neural networks: the exact values aren't as important as the "ballpark" values, insofar as the results are comparable with small error. Fig. 2.4 shows two variations of this function: ultra-fast sigmoid and hard sigmoid. There are different definitions of these [27], [46] and implementation depends on the software being used.

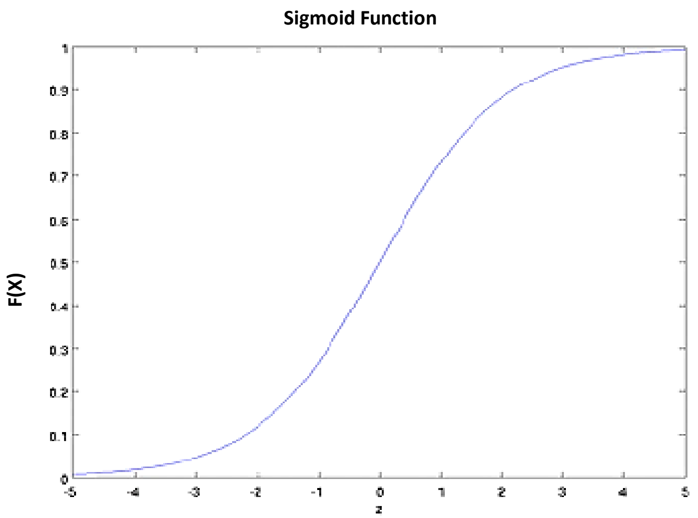

Fig. 2.3. Sigmoid function [47]

The sigmoid (blue) is smooth, while the ultra-fast (green) and hard (red) sigmoids are linear piece-wise. In fact, these approximations are computed as linear interpolations between pairs of cut-points. The green line plot, that touches the blue one at a few points forming a set of line segments. Computing the results of this approximation is significantly faster than calling a routine implementing the sigmoid via exponential and division: all it requires is determining in which linear segment x lies and doing a simple interpolation. The approximation is just that: approximate, but the errors are low enough that many ANN algorithms run fine with the approximation. More information regarding the applicability of the sigmoid function can be found in [47].

Linear: The linear activation function is simplified as

$$f(x) = x \tag{2.7}$$

tanh: The-tanh activation function seen in Fig. 2.5 is defined as follows:

$$f(x) = \frac{e^{x}+e^{-x}}{e^{x}-e^{-x}} \tag{2.8}$$

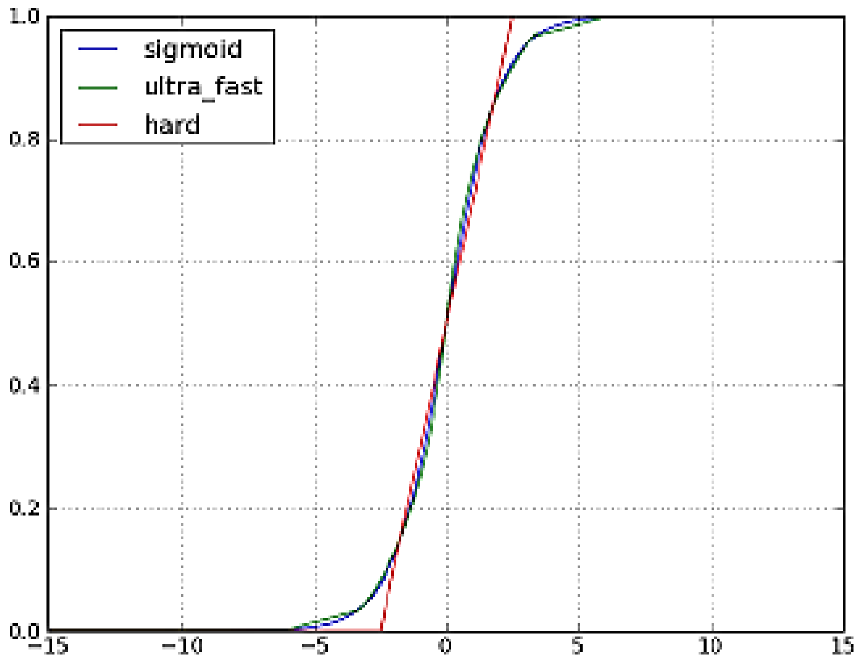

Fig. 2.4. Comparison of sigmoid, hard sigmoid and ultra-fast sigmoid [46].

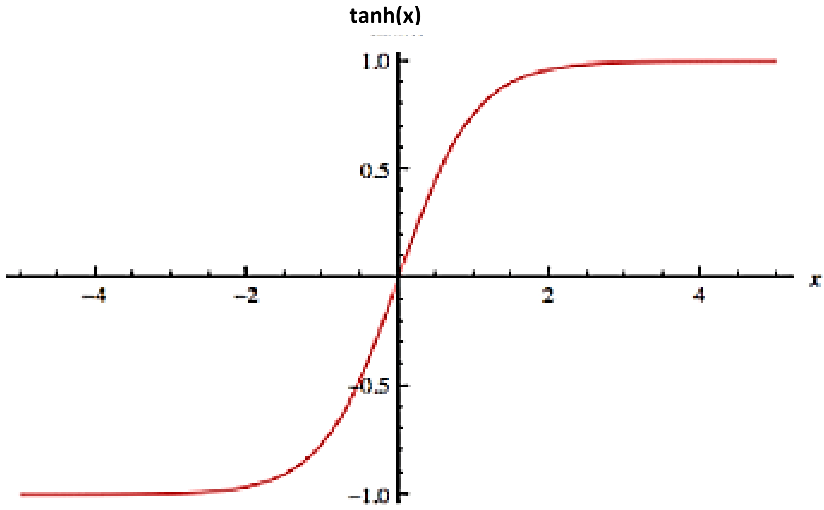

Fig. 2.5. tanh activation function [32].

In load forecasting, various input/ target sets are required to prepare a neural network. Fig. 2.6 shows the basic diagram of an ANN.

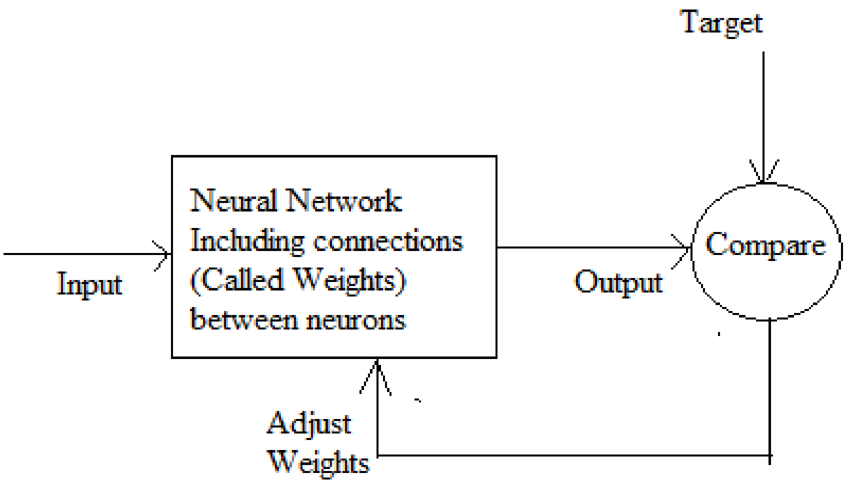

Fig. 2.6. A model of ANN [28].

For solving the issue of load forecasting model shown in figure 2.6 outlines time histories of actual load values. Then, it is signified by a non-linear function as given by

$$P_t = f(P_{t-1,}\epsilon_t) \tag{2.9}$$

Where, the actual load of a system at an instant t is P_t, the previous values of load P_{t-n}, n is the retrospective index review and ϵ_t is a random factor.

2.2.6 Deep neural network (DNN)

In this study, deep learning algorithm is approached for training the model comprising of a large number of processing layers. Deep neural network is one kind of neural network model consisting of an input layer, hidden layer and an output layer. In Fig. 2.7, deep neural network structure is given, where $x_{1,}$ $x_{2\ldots}$ x_n is represented as input in input layer and $y_{1\ldots}$ y_2 is the output in output layer. The hidden layers (h) between input and output represents a black box of the network.

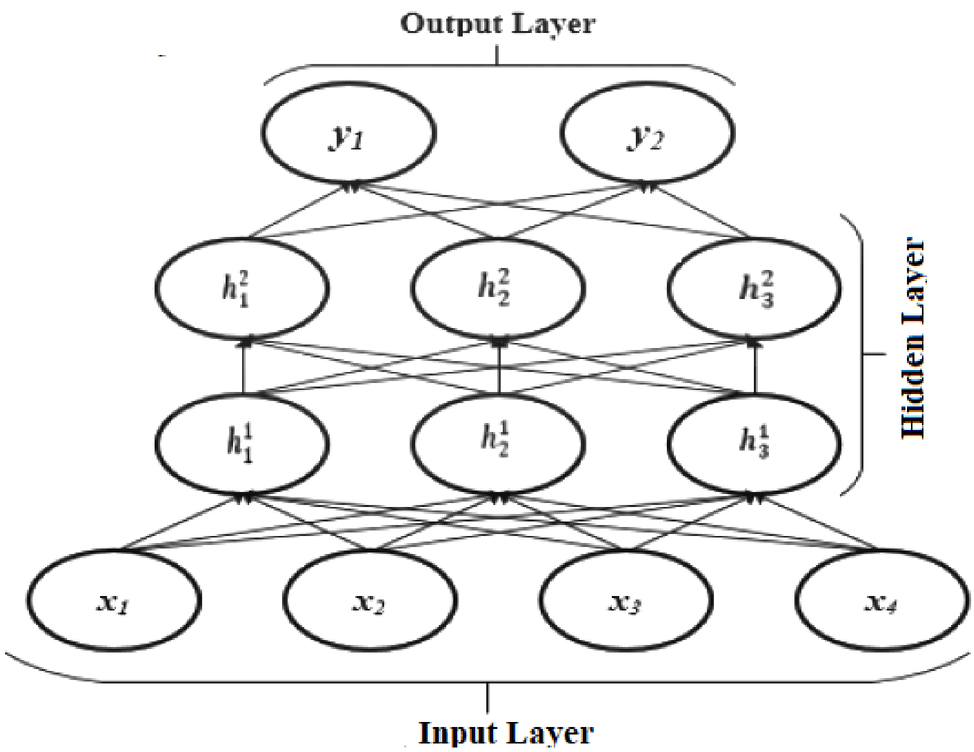

Fig. 2.7. Deep neural network [10].

In the forward propagation, a feed-forward network is trained by the DNN model to execute the corresponding output values throughout all hidden layers and neuron nodes. The following equation directs a nonlinear function indicated by f is calculated as follows:

$$y = f(\sum_{i=1}^{k} W_{ij}^{l}) = f(W^t x) \tag{2.10}$$

where,

y = scalar output,

x_i = = the i th input,

W_{ij}^{l}= the weight between nodes,

f = activation function.

The main core is to minimize an error term for the output layer after producing the corresponding output values. Thus, this network likens predicted output values with the actual existing values.

2.2.7 Recurrent neural network (RNN)

Recurrent neural networks are a special type of artificial neural network where previous step output is fed to the present input step. It has a memory unit which can able to remember all the information which has been calculated earlier. Theoretically, RNN generally handles all the available information in arbitrarily long sequences, however practically; they take into account a few previous steps. A basic RNN model is presented in the Fig. 2.8.

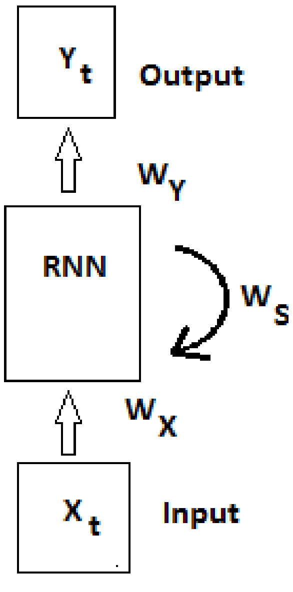

Fig. 2.8. A model of RNN

Recursive formula that gives the definition of RNN is as follows:

$$S_t = F_w(S_{t-1}, X_t) \tag{2.11}$$

$$S_t = \tanh(W_s S_{t-1} + W_x X_t) \tag{2.12}$$

$$Y_t = W_y S_t \tag{2.13}$$

Where, input at time step t is X_t , the state at time t is S_t and F_W is the recursive function.

Moreover, a basic RNN endures from the gradient vanishing and explosion problem when modeling long sequences data in the back propagation step of training process. To overcome this problem, a long short term memory networks (LSTM) is introduced in the field of load forecasting problems.

2.2.8 Long short term memory network (LSTM)

In the field of the time series prediction, a special type of recurrent neural network i.e. LSTM is used due to have the properties of remembering past data in its memory. This network is very much effective for processing, classifying and making time series based forecast. A LSTM network comprised of four units such as cell, input gate, output gate and forget gate. The cell transfers the information over random time periods. The gates are different neural networks which keep the data flow in to and out of the cell state. During training the standard RNNs, it experiences gradient vanishing and exploding problems which can easily be solved by LSTM network. A graphical representation of LSTM network is shown in Fig 2.9. The outputs of each node in LSTM block is computed as [34]:

$$i_t = \sigma(W_i . [h_{t-1,} x_t] + b_i) \tag{2.14}$$

$$f_t = \sigma(W_f . [h_{t-1,} x_t] + b_f) \tag{2.15}$$

$$\widehat{C_t} = \tanh(W_c . [h_{t-1,} x_t] + b_c \tag{2.16}$$

$$C_t = \mathrm{f_t} * C_{t-1} + i_t * \widehat{C_t} \tag{2.17}$$

$$o_t = \sigma(W_o . [h_{t-1}, x_t] + b_0 \tag{2.18}$$

$$h_t = o_t * \tanh(c_t) \tag{2.19}$$

Where, x_t is the input at time *t*. i_t, f_t, O_t represent input gate, forget and output gate respectively. Weight matrices are denoted as: W_i, W_f, W_c, W_o. σ is called sigmoid activation function and h_t signifies the hidden state of all inputs at time t in an input vector representation form.

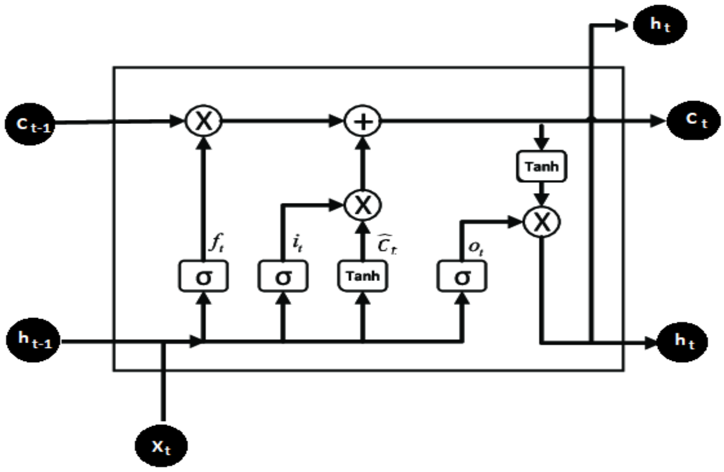

Fig. 2.9. A model of LSTM

2.2.9 Encoder decoder LSTM model

The encoder-decoder approach to sequence prediction has proven much more effective than outputting a vector directly and is the preferred approach. An encoder-decoder LSTM model consists of three sections such as: encoder block, intermediate vector section and decoder section which is shown in Fig. 2.10. Encoder is used to read the input sequences of the load data and the hidden state of this section is defined as:

$$h_t = f(W^{hh}h_{t-1} + W^{hx}x_t) \tag{2.20}$$

In order to get the accurate prediction, encoder vector outlines all the values for all input elements. Internal representation and prediction of the output sequence is interpreted by decoder section. Hidden layer of the decoder is calculated using (2.21).

$$h_t = f(W^{hh}h_{t-1}) \tag{2.21}$$

The output load values are calculated using $softmax$ function which can be written as:

$$y_t = softmax(W^S h_t) \tag{2.22}$$

The encoder-decoder LSTM model used for STLF is presented in Fig. 2.10

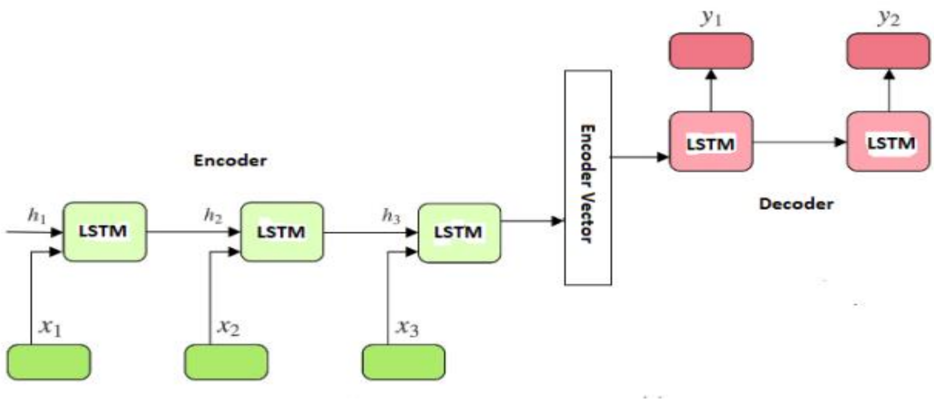

Fig. 2.10. Encoder-Decoder LSTM model

2.2.10 Convolutional neural network (CNN)

CNN is one of the most promising neural networks for addressing the issue of STLF. CNNs are basically a special sort of artificial neural network that is capable of handling time series data using an organized network. Convolutional operation combines two functions on real valued arguments. This convolutional operation can be expressed as follows.

$$s = (x * w) \tag{2.23}$$

Where, x is called input function, w represents the weighting function i.e. a kernel function. Convolution operation's output is often called “feature map”. Convolutional operation in two-dimensional axis can be expressed as

$$s = (I * K)(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n) \tag{2.24}$$

Where, I is defined as input function and K is denoted as two dimensional kernel.

A convolutional network comprises of three layers such as convolutional layer, pooling layer, dense layer. Convolutional layer has three stages. A few convolution operations with a linear activation are performed in the first stage. Second stage i.e. detector stage detects each linear activation and rectified it as rectified linear function. A pooling operation is used in the third stage. Pooling function reduces the dimensionality of the time series data and quickens the training time of the model. Pooling function is also used to down sample the data feature map autonomously without changing the depth. Pooling layer of maximum pooling operation is most commonly used to select the maximum values from obtained the sequence of the convolution layer and project it on the maximum pooling window. It doesn't have the parameters like convolution layers. Some other pooling operation such as average pooling, minimum pooling is often used where needed.

After getting the value from the convolutional layers it is sent to the pooling layers for any kinds of pooling operation. At the end of this stage, the output is flatten and sent it to a dense layer for creating one dimensional output sequence. A back propagation algorithm is used for continuation of model learning scheme. A CNN Architecture is shown in Fig. 2.11.

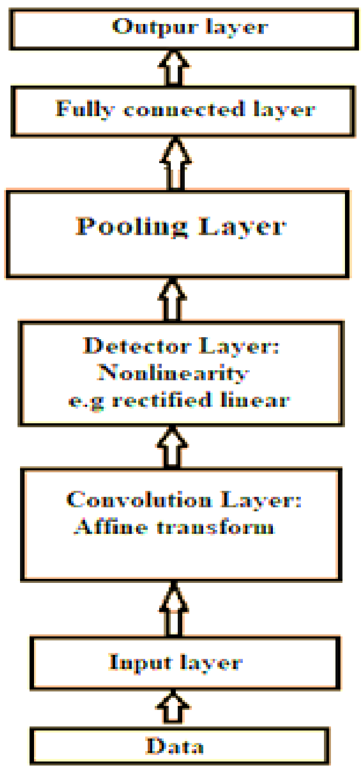

Fig. 2.11. A model of CNN

2.2.11 Genetic algorithm

Genetic algorithms (GA) represent an effective and vigorous approach for load forecasting. Some optimization problems such as constrained and unconstrained complications are solved by GA using natural selection and driven by the biological evolution. It adjusts a population of different solutions in numerous steps. In every step, this algorithm arbitrarily chooses individuals from the current population called as parents and generally produces children using parents for the next generation. Optimal solution is found over progressive generations according to their fitness. The GA has some fundamental rules to a new solution produced from the current solution. Steps that are involving to GA is represented as:

a. Initialisation of the individuals, called parents that contribute to the population of the subsequent generation.
b. Introducing Fitness function.
c. Crossover such that the combining rules of producing next generation
d. Mutation rules make children by applying arbitrary changes to individual parents

2.2.12 Knowledge based expert system

Knowledge based expert frameworks are most effective modern methods that have been technologically advanced in the field of artificial intelligence (AI) algorithm. This framework is basically introduced with a new problem using a computer based programming language which have the capacity of identifying, clarifying a new information available to it. To use this method in the field of STLF, electric load informations are to be provided for extracting the features with the help of the “skilled engineer” called knowledge based expert system. This information is represented as facts and IF-THEN rules, and comprises of the set of relationships. Based on such relationship, load is eventually forecasted.

2.3 Summary

This chapter reviews the recent improvement works in the region of Electrical load forecasting. Emphasis has been given to categorizing different short term electric load forecasting methods. In addition, this chapter presented notable feature of the different short term electrical load forecasting strategies. The next chapter describes the data analysis and electric load characteristic of BPS.

CHAPTER 3
DATA ANALYSIS

3.1 Introduction

Forecasting of electrical load mainly depends on the past data. That is, the future load is usually forecasted on the basis of previous historical load data. Before we explain the methodology of our work, we would like to present the historical load data of BPS and explain some characteristic features of it which forms the basis of our methods.

3.2 Prerequisites of a Good STLF System

The radical changes in the electricity usage owing to the innovations in the technology that generates and delivers it, has set the start of a complete overhauling of the power grid and the entire energy landscape. Need for energy efficiency has become very crucial due to factors like the distributed generations, behind the meter solar photo voltaic resources, frequent peak demand variations and few others. Growth of the open and competitive energy markets sustain reliability and enhance efficiency. Appropriate load evaluation at various stages of power systems is essentially required for such energy markets' planning and operation. With this perspective, an accurate load forecasting solution could play a significant role in the optimization of capital, efficient utilization of distribution networks while maintaining system reliability intact. An efficient, reliable and robust STLF paradigm shall definitely shoulder the smart grid implementations of the power systems. A thorough study of literature in the contemporary perspective led to the following important requirements to be considered, analyzed and addressed while developing the proposed STLF technique in this research:

a. Nonstationary nature of load profiles

b. Adaptiveness of the STLF model

c. Robustness of the STLF technique.

3.3 Load Characteristics of BPS

Time series load data of BPS features a few interesting characteristics. Time series historical electric load data of BPS over the last five years (year of 2014 to 2019) have been collected from the Power Grid Company Bangladesh Limited (PGCB). The time series data were stored over a complete year, every week days from Saturday to Friday at 30 minute intervals in whole Bangladesh. The major load affecting variables are considered as the temperature variation, seasonal effect and day type index. Apparently, temperature variation is considered for the preliminary cause of electric load variation. In specific, the range of temperature variation frequently decides the electricity load variation range. Average temperature variation and load demand variation of Bangladesh is presented in the Fig. 3.1 and Fig. 3.2 respectively.

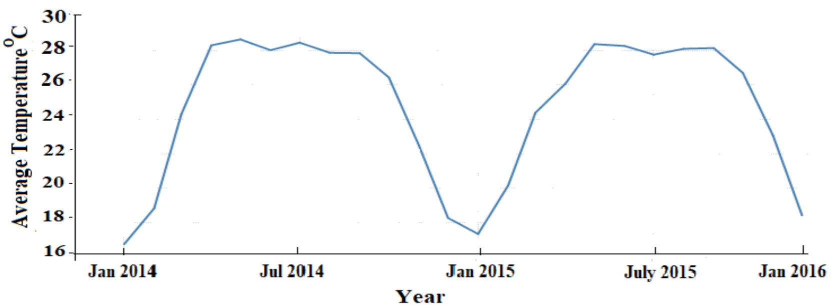

Fig. 3.1. Average Temperature variation

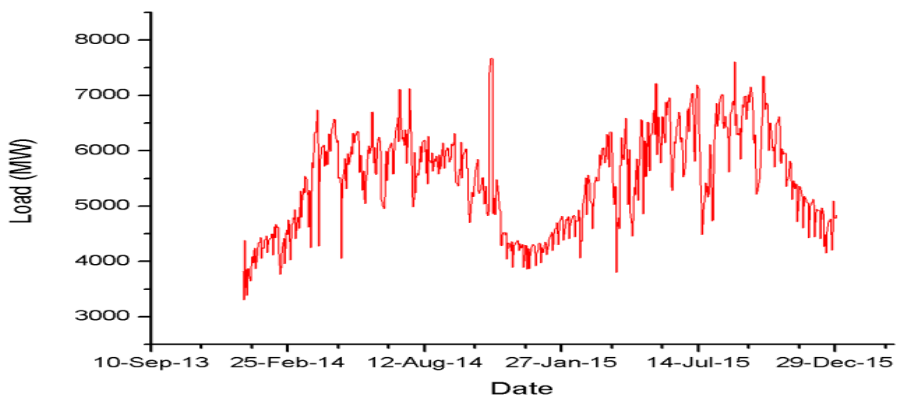

Fig. 3.2. Average load demand variation

A day of 24 hours, in the early morning before sun rising (4.00 A.M. to 6.00 A.M.), the power demand is very low. The demand comes to minimum point around 6 A.M-6.30 A.M. A while later, power demand starts to increase. It is slightly decreasing in the late morning (11.30 A.M.). In the following hours up to 5.00 P.M., it keeps up the demand apparently constant. The electricity consumption rapidly increases after 5.00 P.M. and comes to a peak value around 6.00 P.M. After reaching the peak value, it starts to decrease again. Up to 9.00 P.M. power demand increases a very little, but it starts to decrease from 12.00 A.M. In the summer, peak load demand is around 20% of the average demand while it may reach around 50% of the average load, particularly in the winter season. It is very difficult for any load forecasting framework to make a half hourly forecasting. So seasonal information is needed and month of that season should be identified to make a price forecasting otherwise it would fail to follow the load peak. Moreover, it is observed that within a week, consumption of power is higher on weekdays as compared to weekend or Govt. holidays. Sometimes, the rate of power consumption may have a few distinctions, but in general load pattern trend shows similarity. Besides, the load time series of the similar season shows the similar characteristics, to be specific December and January month exhibit similar load pattern which is shown in Fig 3.3.

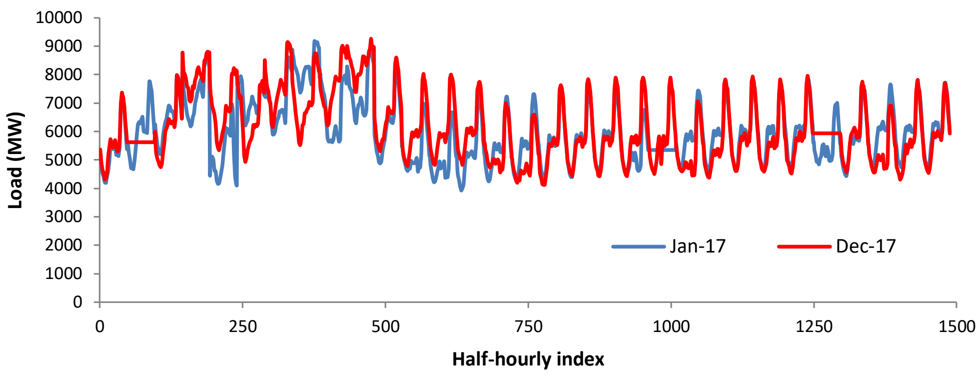

Fig. 3.3. Load demand variation of January 2017 and December 2017

When a summer and winter month, namely January 2017 and May 2017 is plotted together Fig. 3.4 we notice that the wave shape is completely different insofar as the peak load doesn't exceed the average load by similar amounts in these two months, neither is the minimum demand characterized by drooping troughs. Two summer months, May and June in Fig 3.5 exhibit similar wave shape characteristics. The second half of June is distinctly different compared to the pattern of the first half with high difference between peak and average demands. This is due to the fact that the second half of June is when the rainy season sets in and the pattern of the load time series changes. The load time series for the first half of June, however, is similar to the May wave shape. The similarity of load time series in the rainy season is exhibited in Fig. 3.6. The first half of August and the July wave shape are extremely similar.

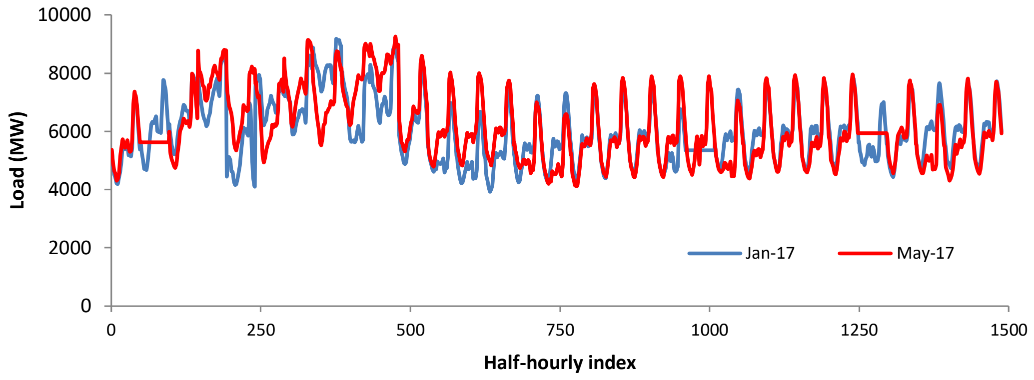

Fig. 3.4. Load demand variation of January 2017 and May 2017

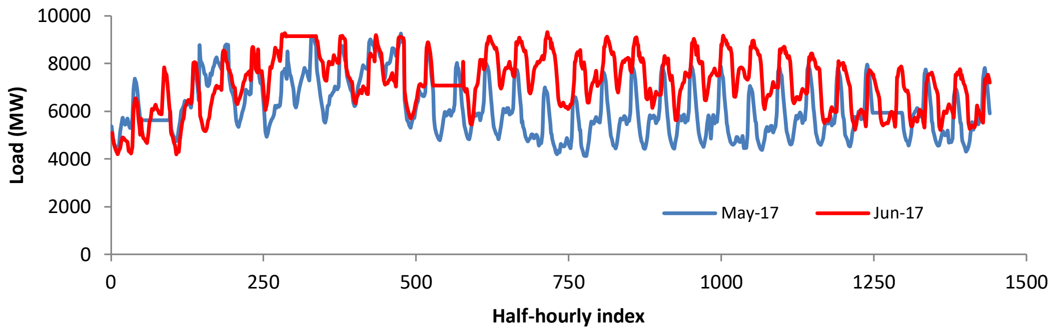

Fig. 3.5. Load demand variation of May 2017 and June 2017

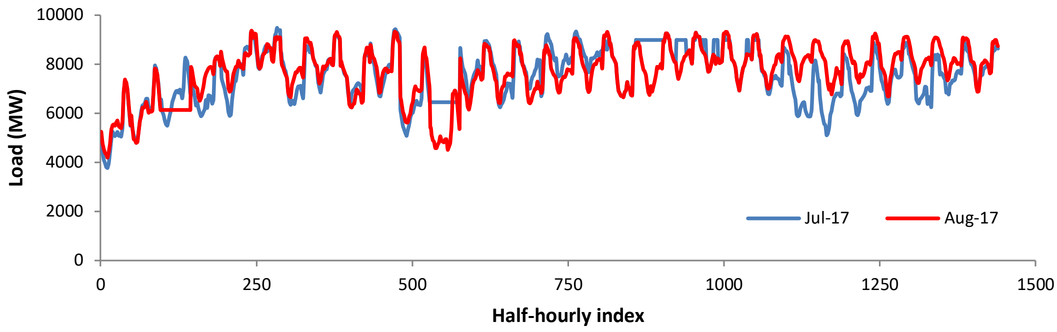

Fig. 3.6. Load demand variation of July 2017 and August 2017

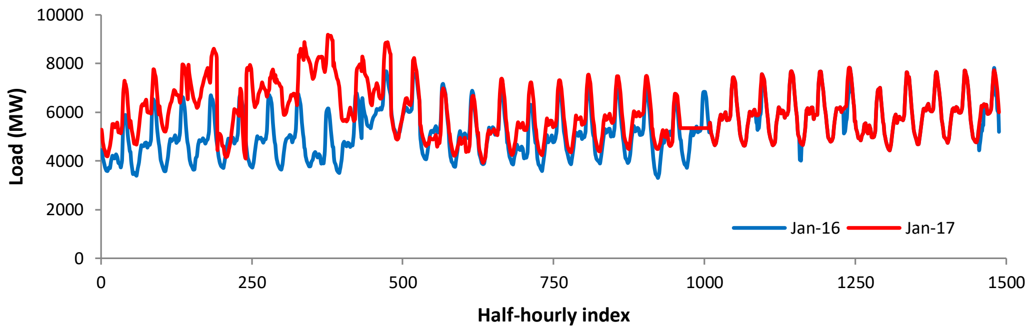

Fig. 3.7. Load demand variation of January 2016 and January 2017

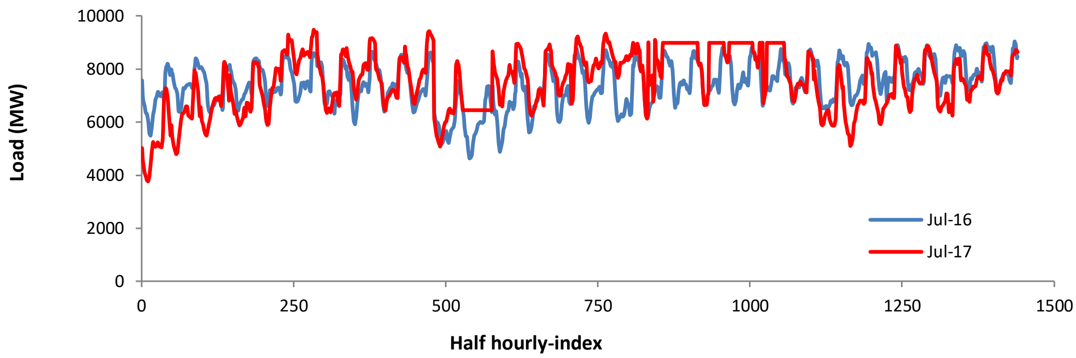

Fig. 3.8. Load demand variation of July 2016 and July 2017

Furthermore, the same months of different years exhibit almost identical wave shapes which are presented in Fig. 3.7 and Fig. 3.8. The peaks and troughs values with respect to the average load value remains similar from year to year. In the light of the above discussion, it can be determined that a fruitful load forecasting model is very much necessary which can precisely forecast the future load demand based on the historical data of the previous year. It also has the capabilities of following the load peaks and crest in different period of the days and months to construct a time series.

3.4 Summary

This chapter represents the electric load characteristics of Bangladesh. From the analysis of the load characteristics it is observed that load pattern has some nonlinear and linear pattern depending on seasonal effect and temperature variation. Electricity demand is higher in the month of May to July(summer) than the month of Aug to December (Winter). Moreover, in a single day demand peak is changing rapidly in the peak hours such that from 5.00 P.M. to 9.00 P.M. This analysis indicates the necessity of constructing a multiple time series for precise short term load forecasting. The next chapter describes the details procedure of STLF using proposed CNN-LSTM model.

CHAPTER 4
PROBLEM FORMULATION AND PROPOSED METHODOLOGY

4.1 Introduction

The main challenge of the short term load forecasting issue is the diversity and volatility. It is greatly affected by input variables of the load datasets. Before going to the modelling part, the data need to be prepared appropriately and time series ought to be decomposed and stationarized. Toward this determination, it has been received a well ordered methodology for analysis and decomposition of the data to enable the flexibility of picking-up and using the time series at any degree of processing as per the prerequisite of modelling strategy and the analysis.

4.2 Proposed Methodology

Step 1: Data Framing

In this step historical load data need to be collected from a particular region and null values have to be checked. Then load data set has to be divided in to training and test set for evaluating the proposed model. Collected data set ought to split into different standard weeks. Reformation of this data frame is very much effective for defining the model which can predict the power consumption for the week ahead and month ahead.

Step 2: Constructing Multistep Time Series

In the proposed model electric load data set has to transform in the shape of [sample, time steps, features]. At first, per sample, seven time steps have to take having one feature for total daily power consumption of seven days. The information of this pattern is not sufficient to train the network. So, it is needed to create more training information by changing the problem to predict the next seven days given the prior seven days, irrespective of the standard week. Secondly, data set is needed to flatten at first and make eight-time series sequences. Then it is obligatory to iterate over the time steps and divide the data set into overlapping windows where it moves along one-time step and predicts the subsequent seven days. However, the test information from the data set is remained same in every case.

Step 3: Building Forecasting Model

It is needed to implement an encoder-decoder CNN-LSTM model which basically deals with the one dimensional data in the three dimensional pattern. The CNN block of the proposed model is defined two convolutional layers where convolution is taken place with the help of the kernel filter. The first convolution layer read across the input series and projects its sequences on to the features windows where second convolution layer is operated for amplifying the features obtained from the first layer. In our model, number of feature maps are 64 per each convolutional layer with three-time step kernel filter. A maximum pooling layer is generally used for getting the values after two times convolutions in the convolution layers. It is actually used for simplifying the input features. In the proposed model, maximum pooling operation has to be done by taking ¼ of the values with the original sequence. The results obtained from this operation are then flattened in to a long vector that is used as input to the decoding process of LSTM unit followed by dense layer which is used to provide the output. The developed model for load forecasting is shown in the Fig. 4.1.

Step 4: Training the Proposed Model

Proposed CNN-LSTM architecture has been build using Keras, an open source neural network library which is written in Python. Then the network is needed to train with the following layer specifications.

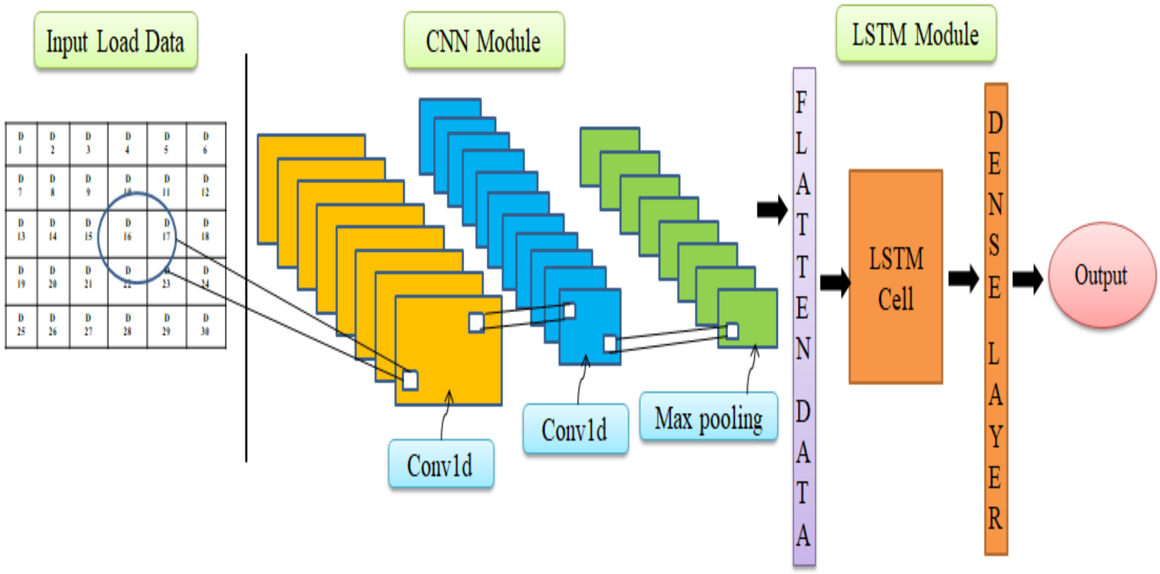

Fig. 4.1. Proposed CNN-LSTM model

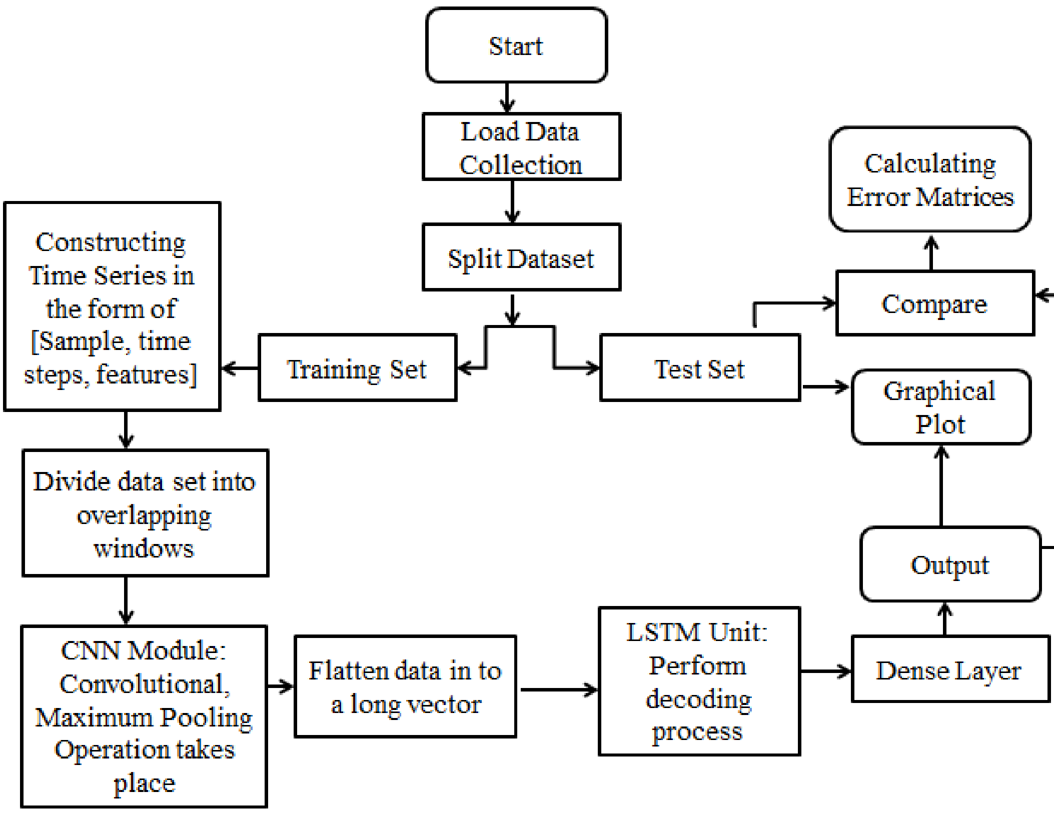

Fig 4.2. Flow chart of the proposed technique.

a. Convolution type : 1D

b. No of filter: 64 with kernel size 3

c. Activation function : Rectified linear unit (RELU) for CNN, LSTM and Dense layer

d. Optimizer : Adam

e. No of hidden layer : 200 for LSTM

f. No. of training iterations (epochs) : 20

g. Batch size : 16

Flow chart of the proposed technique is shown in Fig. 4.2 and the detail algorithm for formulation of proposed scheme is as follows:

Step 1: Read the electric load data.

Step 2: Split the data set into training and test set.

Step 3: Construct a time series using training in the shape of [sample, time step, features].

Step 4: Divide processed data into overlapping windows.

Step 5: Give input to the CNN from step 4.

Step 6: Perform one dimensional convolution using two convolutional layer.

Step 7: Pool maximum value from the result of each convolution.

Step 8: Flatten the data from step 7 in to a long vector.

Step 9: Decode the values from step 8 using LSTM.

Step 10: Insert the decode values in to a Dense layer.

Step 11: Take the output from dense layer and compare with the test data set which is partitioned in step 2.

Step 12: Plot the output values against test values from step 2.

Step 13: Calculate error matrices using error matrices formula.

Step 14: Stop the process.

4.3 Application of the Proposed Methodology

Step 1:

Historical time series half hourly Electric load data of BPS over the last five years has been collected from PGCB. First four years data is used for model training process and the data of final year is used for evaluating the performance of the models. Collected data set is spilt in to different standard weeks which starts from Saturday and ends on Friday. The daily data starts in late 2014 where the first Saturday in the dataset is 4th January. The final year of the data is in 2018 and the first Saturday for 2018 was January 6th. The data ends in last December 2018 and the closest final Friday in the data is 28th December. This gives 52 weeks of test data. After organizing all the split dataset, it gives 208 full standard weeks to train the developed CNN-LSTM model. This process can be examined by taking test data from the year of 2019 and testing data from 2014 to 2018.

Step 2:

The training data set in standard week is now provided with eight variables having the shape of [208, 7, 8].

Step 3:

Processed data from step 2 is applied in the proposed CNN-LSTM model is shown in Fig. 4.1.

Step 4:

Proposed model is then trained with the processed electric load data of BPS with the layer specification described in step 4 in the previous section.

4.6 Summary

Electric load profiles and the step by step proposed methodology is discussed in this chapter. Usually, use of the proposed CNN-LSTM scheme is verified with the real data collected from PGCB. In next chapter, the results obtained are presented and analyzed in terms of tables and graphs. A detailed discussion about the obtained results for different test is also synthesized on the chapter.

CHAPTER 5
RESULT AND DISCUSSION

5.1 Introduction

This chapter presents the evaluation of the proposed CNN-LSTM scheme presented in Chapter 4. The performance of the proposed scheme is validated with the real electric load data set recorded in 2018 and 2019 separately. In section 5.4, we displayed the forecasting outcomes with the graphical representation in different periods of the year of 2018 and 2019 respectively and in section. 5.5, evaluation metrics are analyzed in different horizons by comparing it with conventional LSTM network.

The proposed CNN-LSTM model was implemented using Keras, an open neural network library with specified layer specification described in section 4.2. However, it is to note that, with the programming skills and platforms, the forecasting outcomes may vary without significantly changing the evaluation matrices reported in this thesis.

5.2 Evaluation Metrics

For the purpose of evaluating the effectiveness and accuracy of the developed model three type of matrices is defined such as mean average error (MAE), root mean squared error (RMSE) and mean absolute percentage error (MAPE). More accurate forecasting can be evaluated if the value of evaluation index is closer to zero.

5.2.1 Mean absolute error

Mean absolute error (MAE) is an evaluation metric by which difference between two observations or variables is calculated. Without considering any direction, MAE computes the average magnitude of the errors in a set of predictions. Average of the absolute error is calculated by MAE which can be expressed as follows:

$$MAE = \frac{1}{N}\sum_{L=1}^{N}\left|(\breve{Y}_L - Y_L)\right| \quad (5.1)$$

Where N is the total number of the observation periods, $\breve{Y}_L$ is the actual load value and Y_L denotes the predicted value at time t. Moreover, MAE is a negatively oriented score which means that the lower the value the more accurate the prediction.

5.2.2 Mean absolute percentage error

The mean absolute percentage error (MAPE deviation (MAPD), is a measurement of prediction accuracy of a forecasting method. For example, it is also known as mean absolute percentage error in the estimation of a trend. It generally states accuracy in terms of percentage calculation which is stimulated by the formula:

$$MAPE = \frac{\sum_{L=1}^{N}\left|\frac{(\breve{Y}_L - Y_L)}{\breve{Y}_L}\right|}{N} \times 100 \quad (5.2)$$

Like the previous metric, this is also a negatively oriented score.

5.2.3 Root mean square error

A quadratic scoring rule that measures the average magnitude of the error is called root mean square error (RMSE). It performs the square root of the average of squared differences between prediction and actual values. With symbols having the meanings as defined in 23 we can express RMSE by the equation:

$$RMSE = \sqrt{\frac{1}{N}\sum_{L=1}^{N}(\breve{Y}_L - Y_L)^2} \quad (5.3)$$

RMSE is also a negatively oriented score. In all cases, the value of RMSE will greater than or equal to MAE.

5.3 Comparison of the Metrics

None of the three metrics described so far is universally applicable in forecasting tasks. Some interesting implications for RMSE are the square root of the average squared errors. Before performing average operation, error values are squared so that the RMSE gives a relatively high weight to large errors. When large error values are unwanted, the RMSE would be more useful in the field of load forecasting. MAE, on the other hand, has been suggested to be better in assessing average model performance [17]. This claim, however, has been contested and a combination of the two has been proposed in [18]. Regardless, both of these methods have a glaring weakness in that their values can't be compared between two models without taking the base load level into consideration. For example, if two models predicting loads of two different power systems have a RMSE of say 2000 KW and 3000 KW respectively, we cannot say with certainty that the second model is performing worse than the first one. If the first power system has a base load of 4 KW and the second 4 MW in which case the second model would be performing significantly better.

MAPE overcomes this limitation by expressing the error in mean percentage of actual values. This enables straightforward comparison of different models since just by observing the value a decision can be made. MAPE however, is also not without its limitations. The MAPE is expressed in terms of percentage which makes sense for values where divisions and ratios make a significant contribution. For temperature forecasting problems, it doesn't have any fruitful contribution so that MAPE are not considering an evaluation metrics. Even in load forecasts, the percentage expressed is a mean percentage of each actual value which makes it difficult to interpret the result physically; only an intuitive mathematical value can be discerned. Secondly, if just a single actual is zero such that $\breve{Y}_L(t)$ = 0, then the value of MAPE comes undefined. Furthermore, if it is wished make a forecast with positive data (MAPE doesn't make any sensible operation), then it is not able to make any forecast below the value of zero. In MAPE calculation, inappropriately overforecast problem is treated differently than underforecast where it is not possible for an underforecast to contribute more than 100% (e.g., if $Y_L(t)$ = 0 and $\breve{Y}_L(t)$ = 1), whereas an overforecast contribution is unbounded (e.g., if $Y_L(t)$ = 5 and $\breve{Y}_L(t)$ = 1). This implies that the value of MAPE can be lower for biased than for unbiased forecasting procedure. Minimizing it may lead to forecasts that are biased low. It is quite clear from the preceding discussion then that none of the metrics can be considered appropriate or self-sufficient to compare performances of different forecasting models. The comparison can be summarized as follows (shown in table 5.1):

Table 5.1: Comparison of MAE, RMSE and MAPE

RMSE	MAE	MAPE
Error values are squared so that the RMSE gives a relatively high weight to large errors.	Errors values are relatively low compared to RMSE because error values are not squared	It gives high errors during low-demand periods
Applied where large load prediction error values are unwanted	It can be applied for both low demand periods and high demand periods	Applied where higher load demand forecasting is required.
It exhibits good Key performance indicator (KPI)	It exhibits relatively high KPI.	It exhibits poor KPI.

Therefore, in this research work, all three of the metrics described above have been used to gauge the performance of the proposed model.

5.4 Forecasting Outcomes

The trained proposed CNN-LSTM model and LSTM model is able to forecast over the different time periods of 24 hours, 7 days and 30 days of the year of 2018 and 2019 with half hourly information. The time series load data set used in our experiment is collected from 1st January 2014 to 31st December 2019. Sampling of the load information was taken with half hourly interval. The electrical load data set contains total 105167 tests. For the first experiment, the first four years load data set was chosen for training the CNN-LSTM model where test information was selected from different periods of 2018. For the second experiment, first five years data was chosen for training the network and the data of the last year i.e. 2019 was selected for test information. Data from the forecasting plots shown in Fig. 5.1 to Fig. 5.76, it has been observed that both the models can predict the actual load demand trend. However, the prediction trend obtained from CNN-LSTM based forecasts is very closer to the actual load pattern as compared to the prediction trend of LSTM model.

5.4.1 Monthly prediction for the year of 2018

The proposed models are used to forecast load for each month in 2018. The summarized results for each month are shown in the following figures (Fig. 5.1 to Fig. 5.2) individually. From all figures, horizontal axis and vertical axis represents time periods in half hourly index and load demand in MW. From the forecasting plots, it has been seen that the forecasting outcomes using proposed CNN-LSTM model follows the actual load trend precisely than the forecasting outcomes of the LSTM model. Moreover, prediction curves obtained from proposed CNN-LSTM model has minimum deviation (shown in table 5.1) from actual load demands as compared to LSTM model.

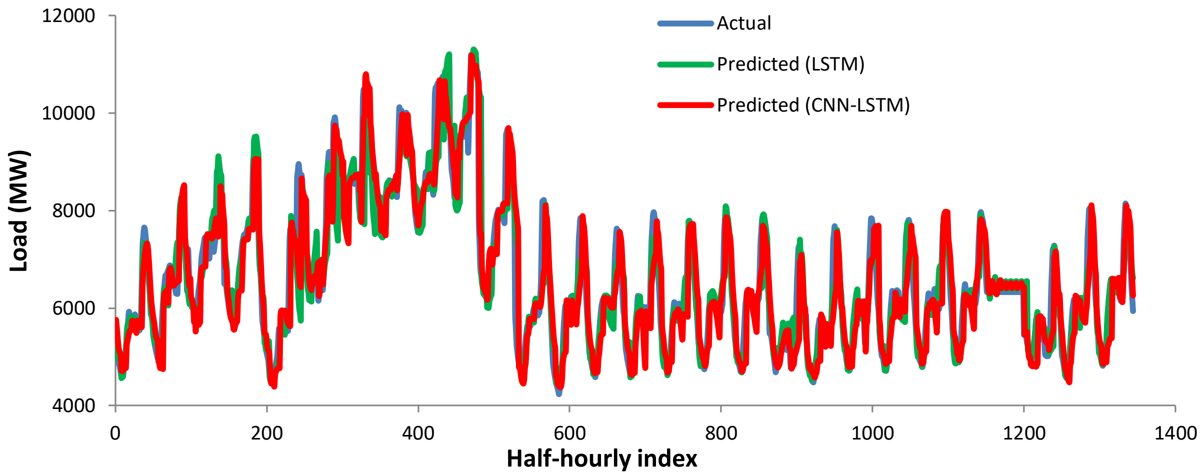

Fig. 5.1. Load forecasting of BPS for January 2018

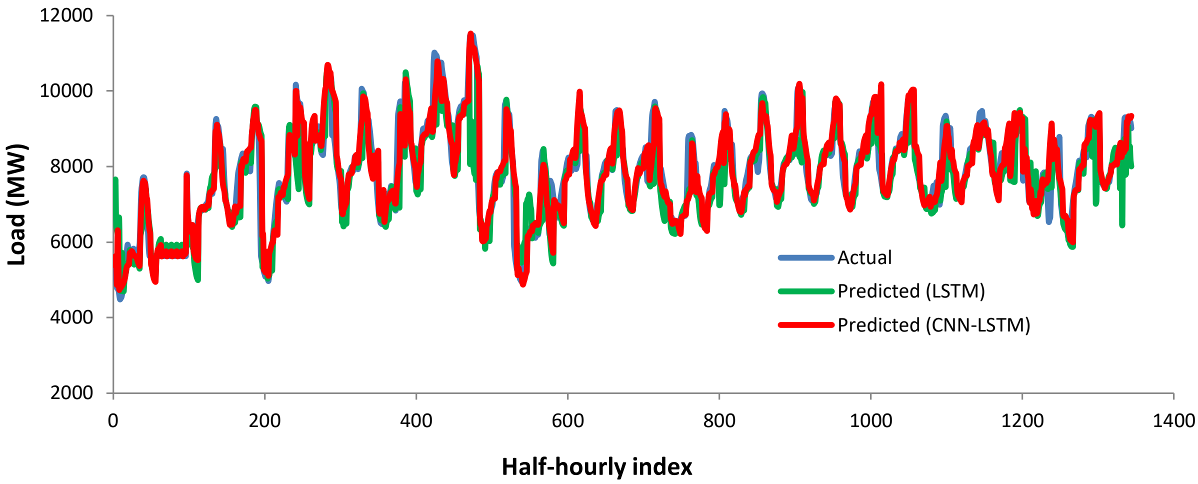

Fig. 5.2. Load forecasting of BPS for March 2018

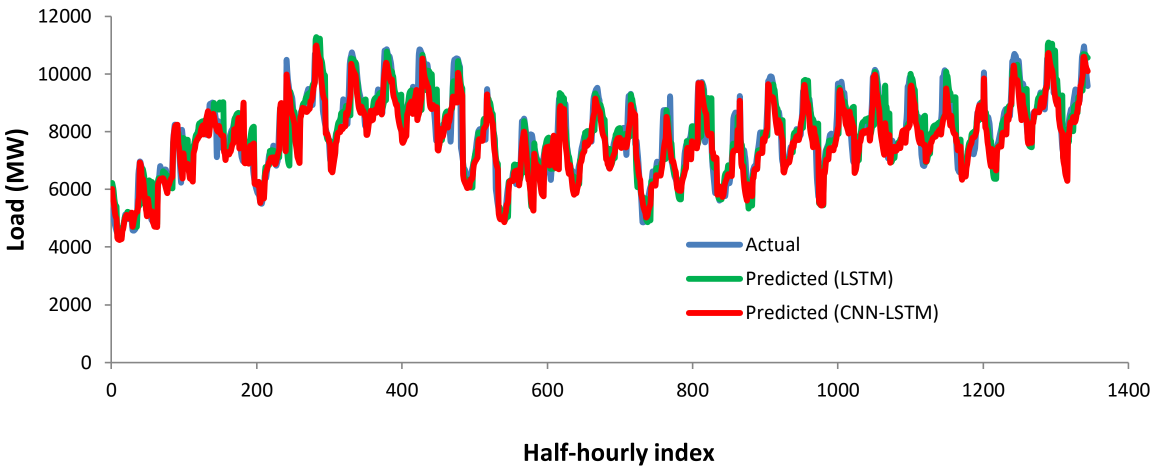

Fig. 5.3. Load forecasting of BPS for May 2018

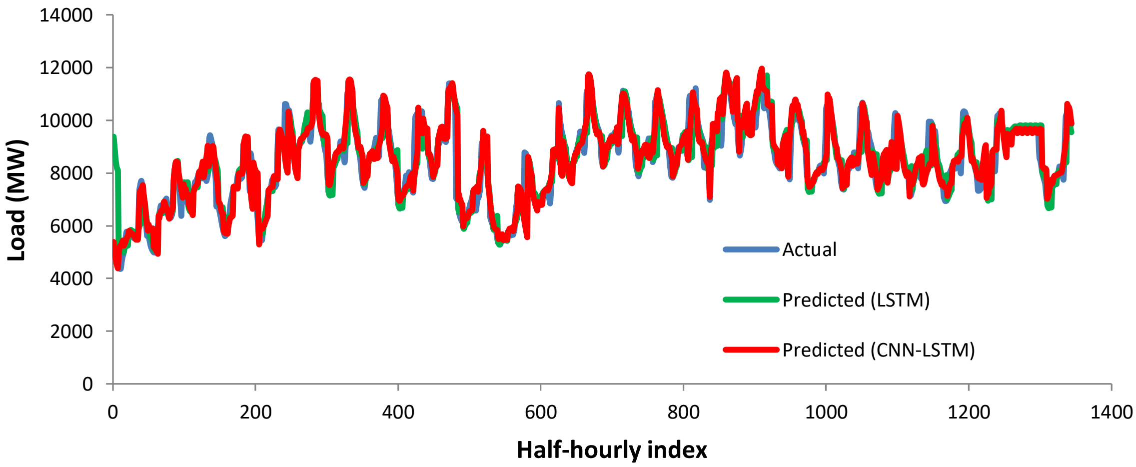

Fig. 5.4. Load forecasting of BPS for July 2018

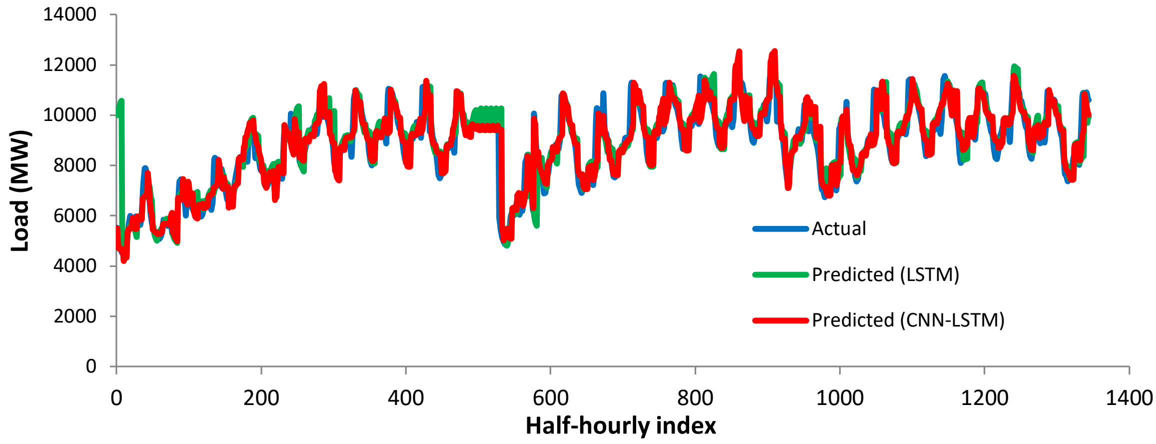

Fig. 5.5. Load forecasting of BPS for October 2018

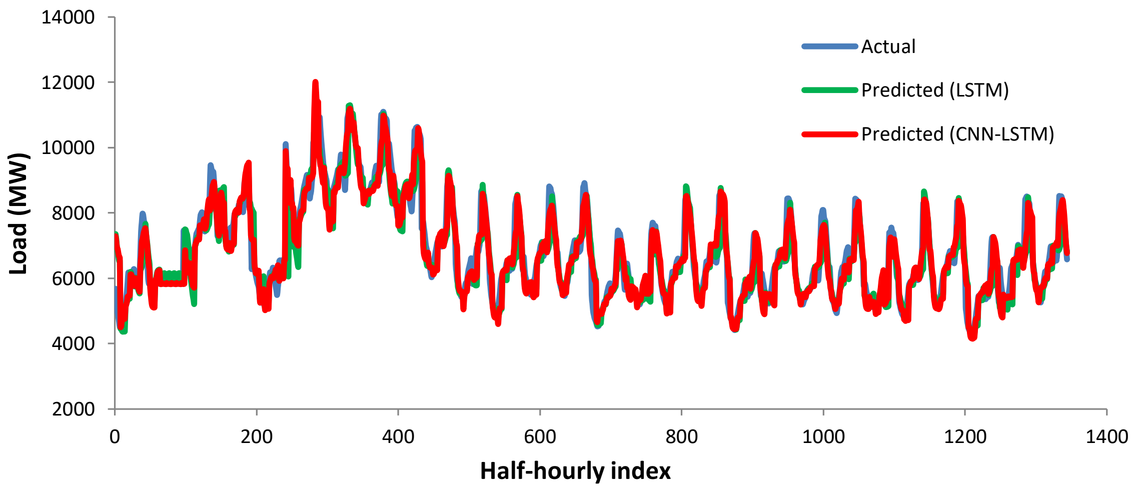

Fig. 5.6. Load forecasting of BPS for November 2018

5.4.2 Weekly prediction for the year of 2018.

The proposed models are used to forecast load for each week in every month in 2018. The forecasted results for some weeks in differents month are shown in the following figures (Fig. 5.7 to Fig. 5.18) individually. From all figures, horizontal axis and vertical axis represents time periods in half hourly index and load demand in MW. From the forecasting plots, it has been seen that the forecasting outcomes using proposed CNN-LSTM model follows the actual load trend precisely than the forecasting outcomes of the LSTM model. Prediction curves obtained from proposed CNN-LSTM model has minimum deviation (shown in table 5.2) from actual load demands as compared to LSTM model.

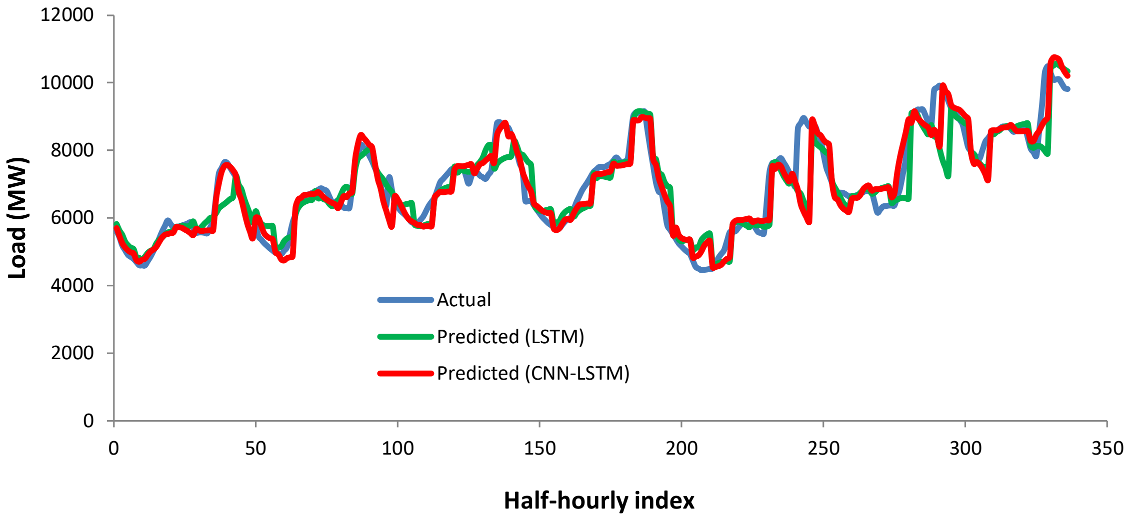

Fig. 5.7. Load forecasting of BPS for 01-07 January 2018

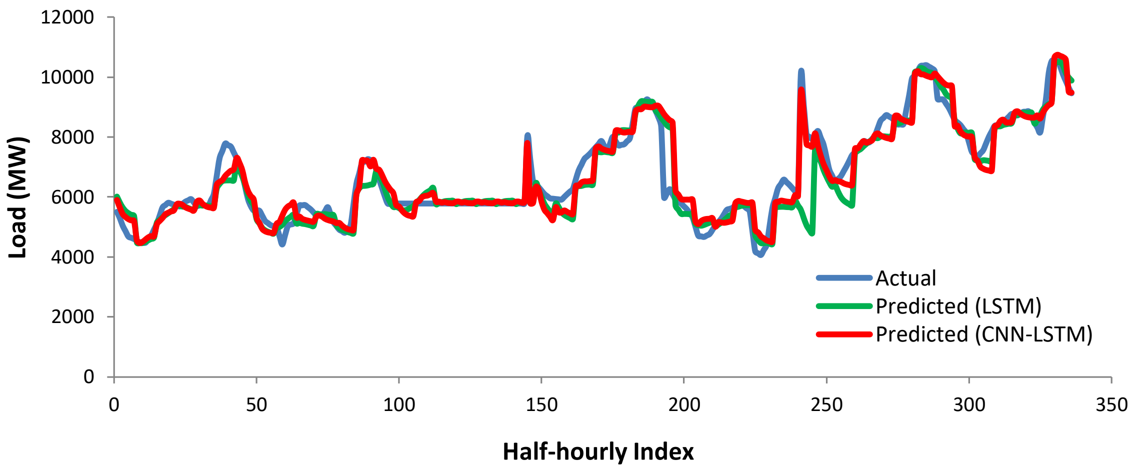

Fig. 5.8. Load forecasting of BPS for 01-07 February 2018

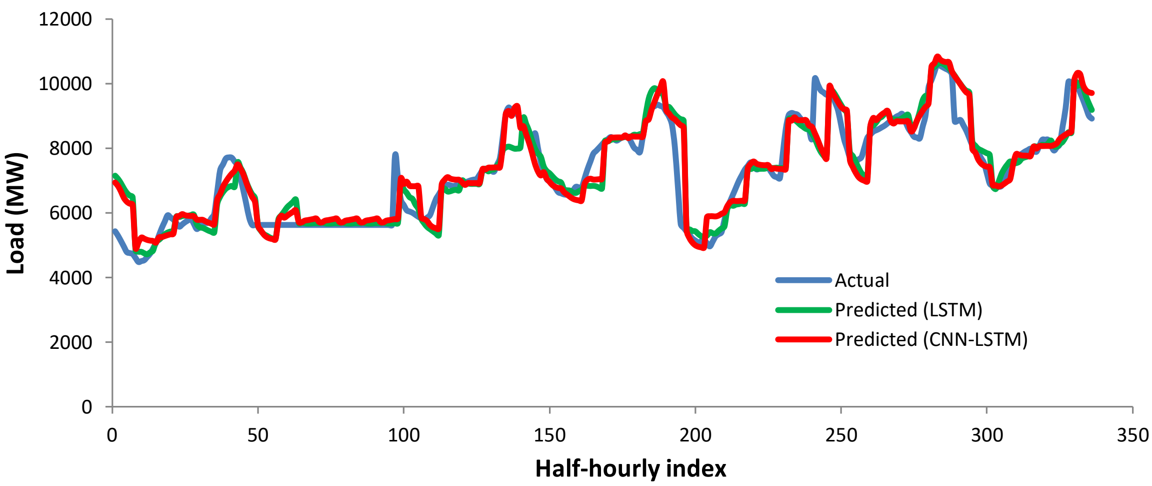

Fig. 5.9. Load forecasting of BPS for 01-07 March 2018

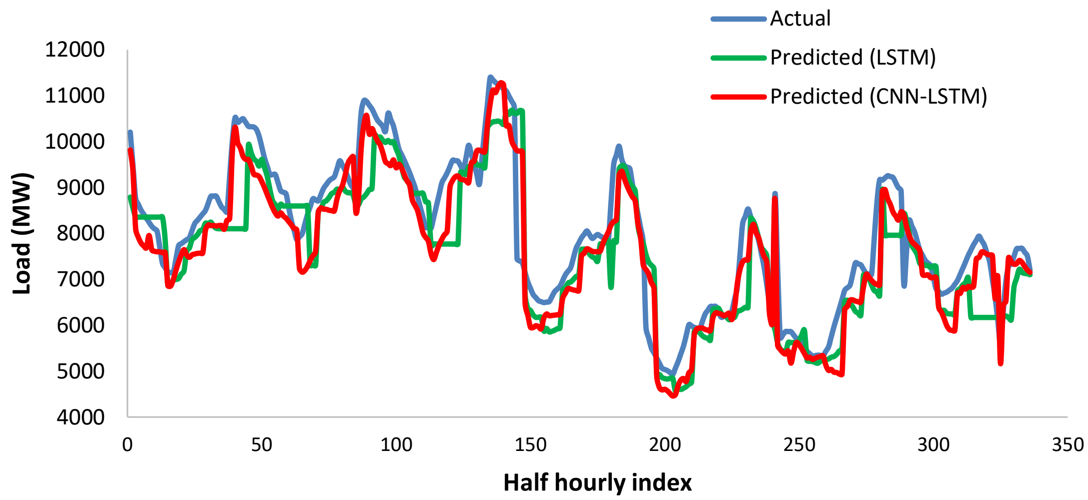

Fig. 5.10. Load forecasting of BPS for 07-14 April 2018

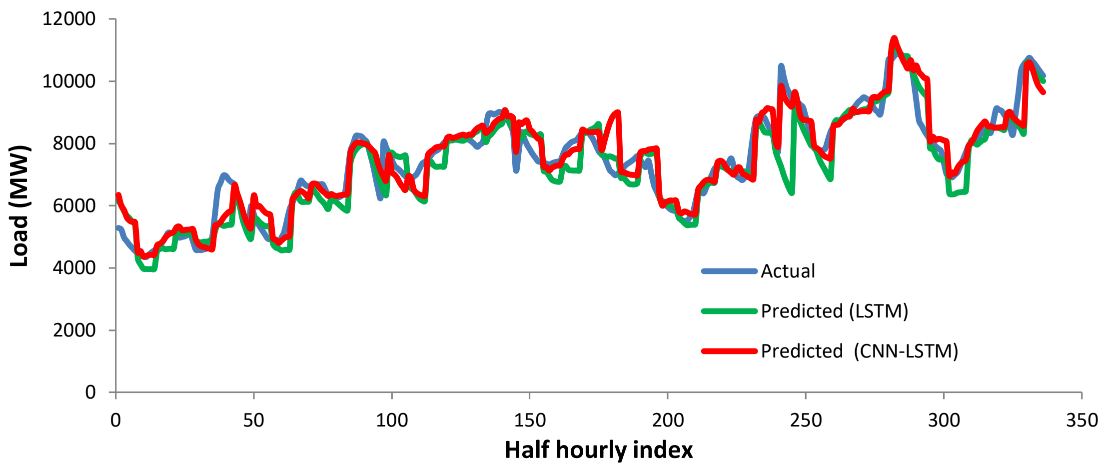

Fig. 5.11. Load forecasting of BPS for 01-07 May 2018

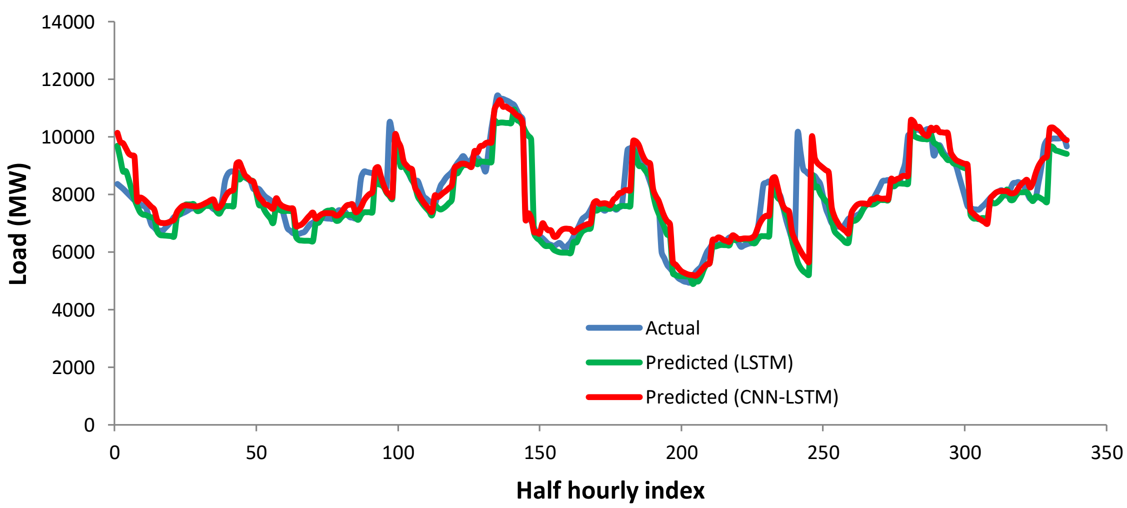

Fig. 5.12. Load forecasting of BPS for 08-14 Jun 2018

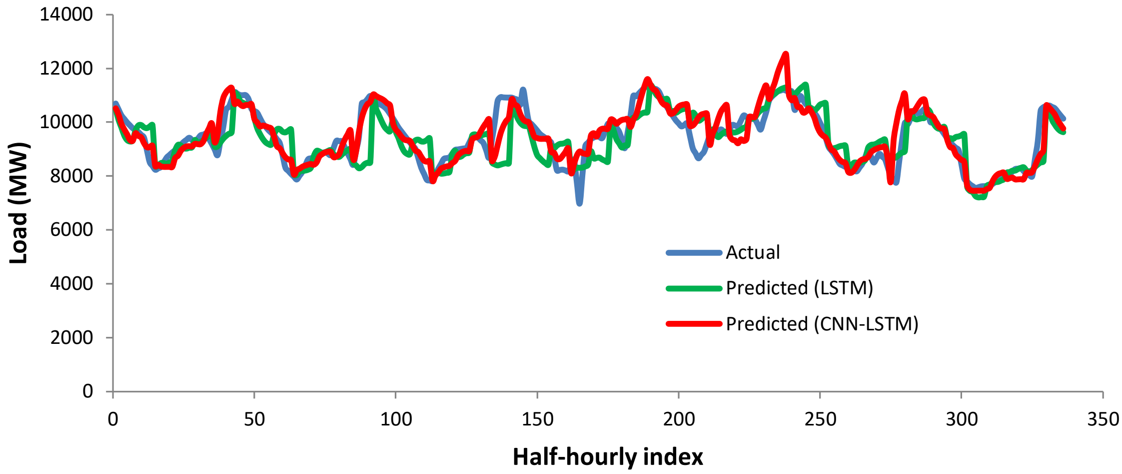

Fig. 5.13. Load forecasting of BPS for 15-22 July 2018

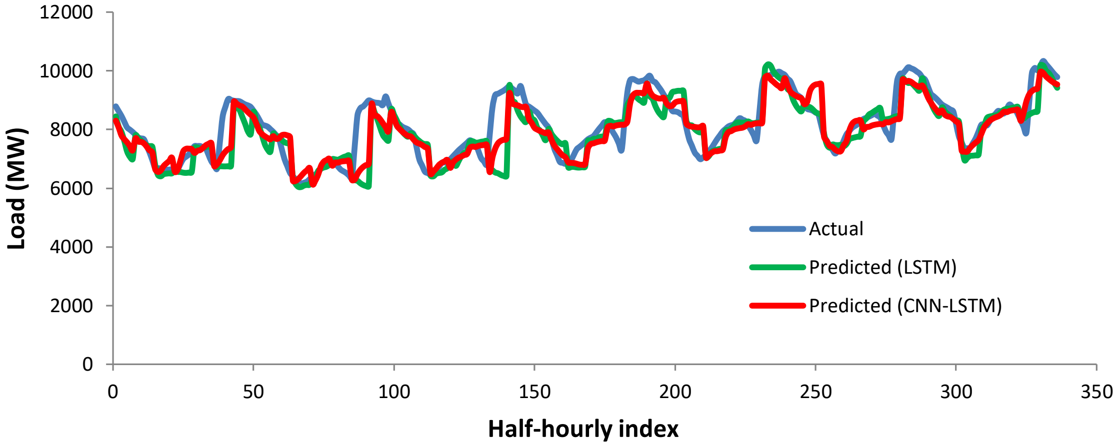

Fig. 5.14. Load forecasting of BPS for 23-29 August 2018

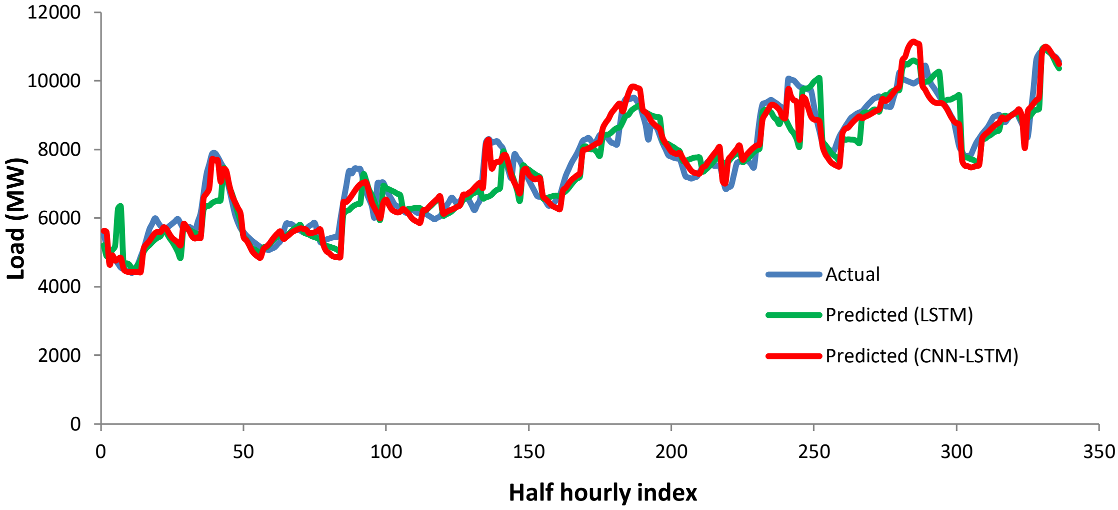

Fig. 5.15. Load forecasting of BPS for 01-07 September 2018

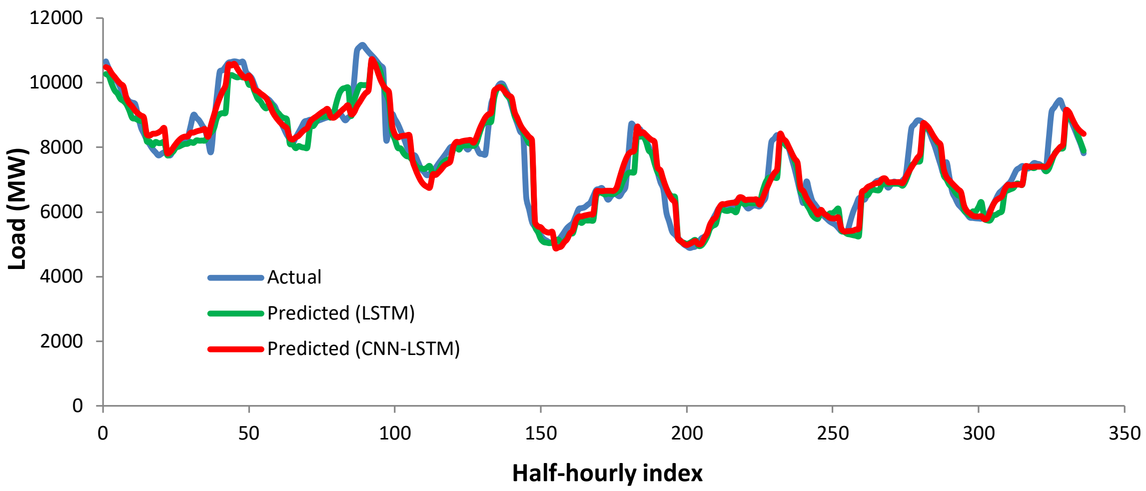

Fig. 5.16. Load forecasting of BPS for 08-14 October 2018

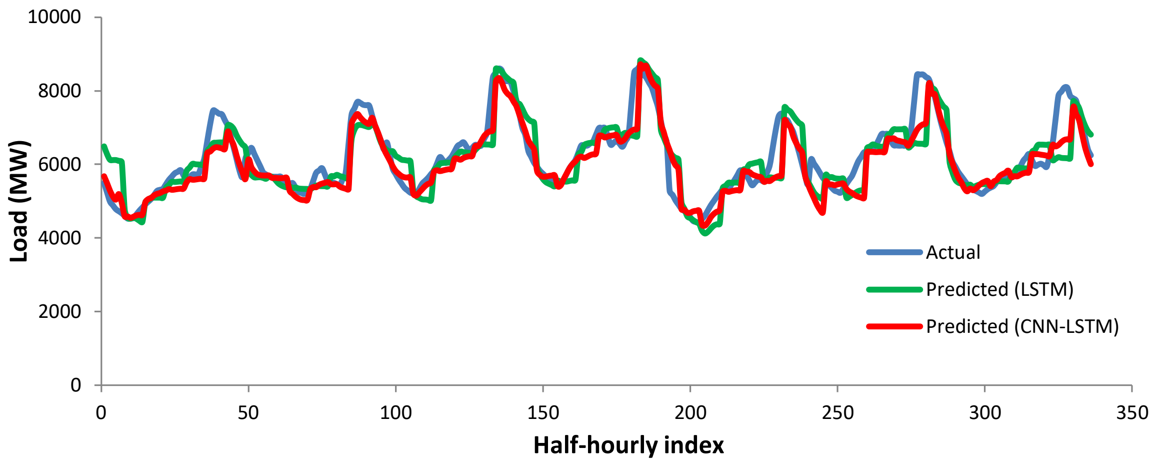

Fig. 5.17. Load forecasting of BPS for 15-21 November 2018

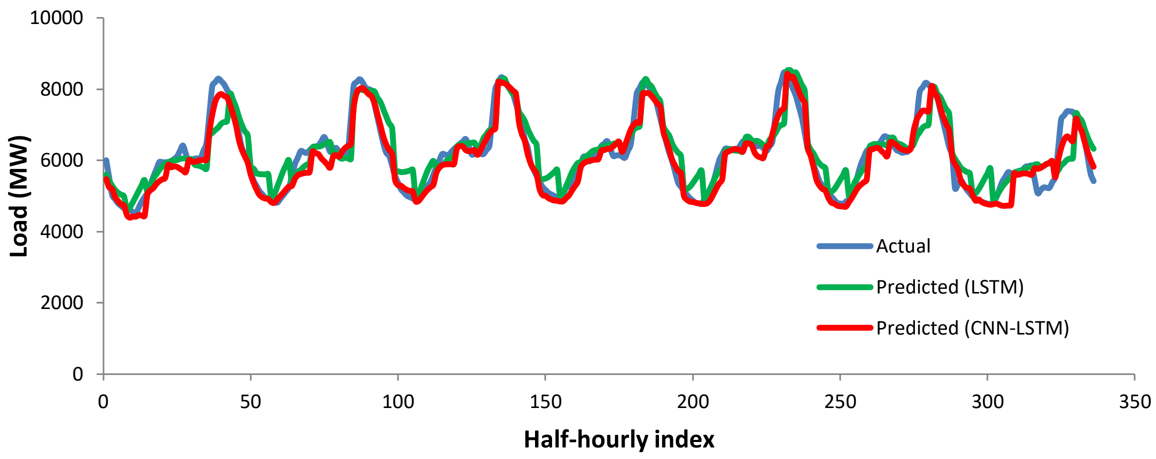

Fig. 5.18. Load forecasting of BPS for 23-29 December 2018

5.4.3 48 hours prediction for the year of 2018

It is used to forecast the load for 48 hours at a time such that prior two days using proposed CNN-LSTM model in every month of 2018. The forecasted results for 48 hours in differents couple of a days are shown in the following figures (Fig. 5.19 to Fig. 5.24) individually. From the forecasting plots, it has been seen that the forecasting outcomes using proposed CNN-LSTM model follows the actual load trend precisely than the forecasting outcomes of the LSTM model. From the prdiction pltots it is seen that proposed CNN-LSTM model exibits minimum deviation (shown in table 5.3) from actual load demands as compared to LSTM model.

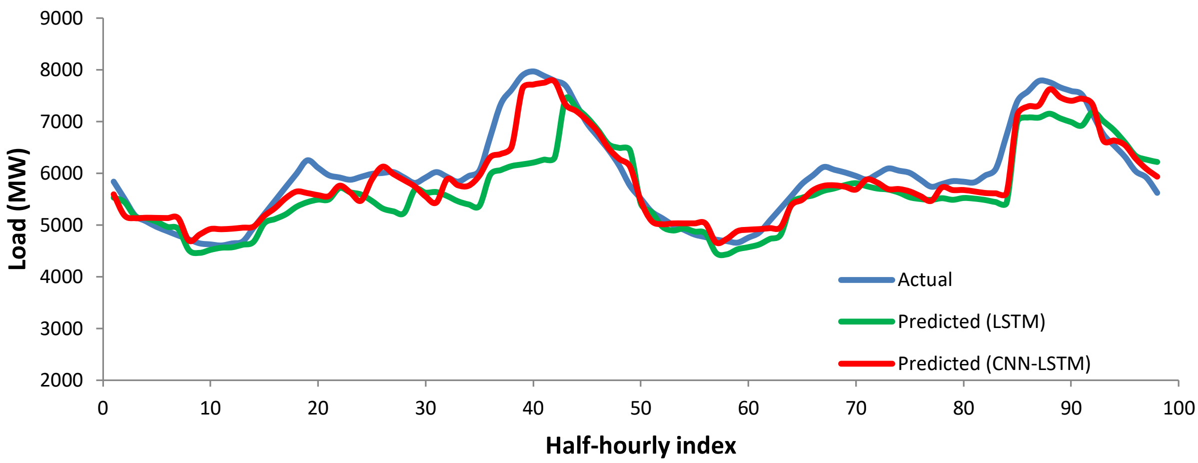

Fig: 5.19. Load forecasting of BPS for 15-16 January 2018

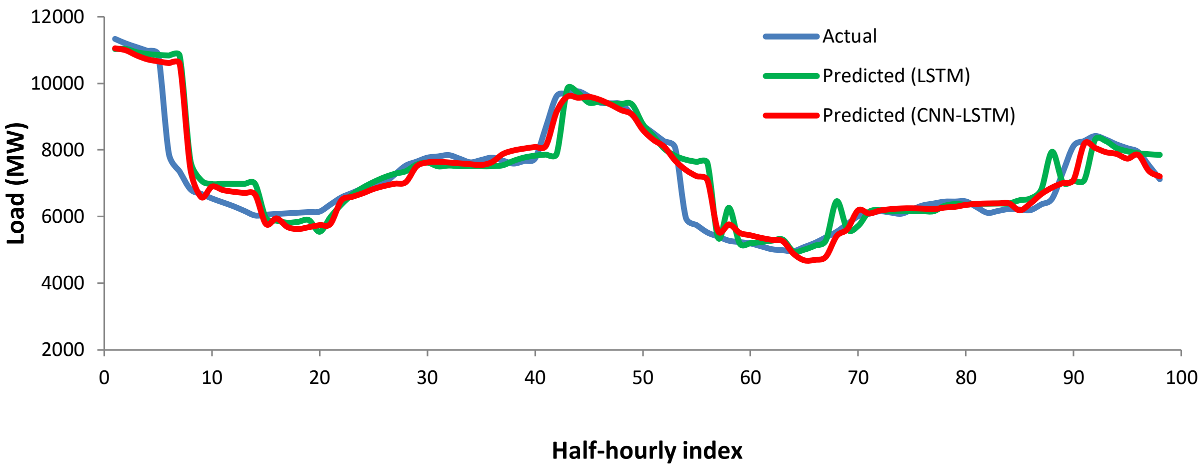

Fig: 5.20. Load forecasting of BPS for 11-12 March 2018

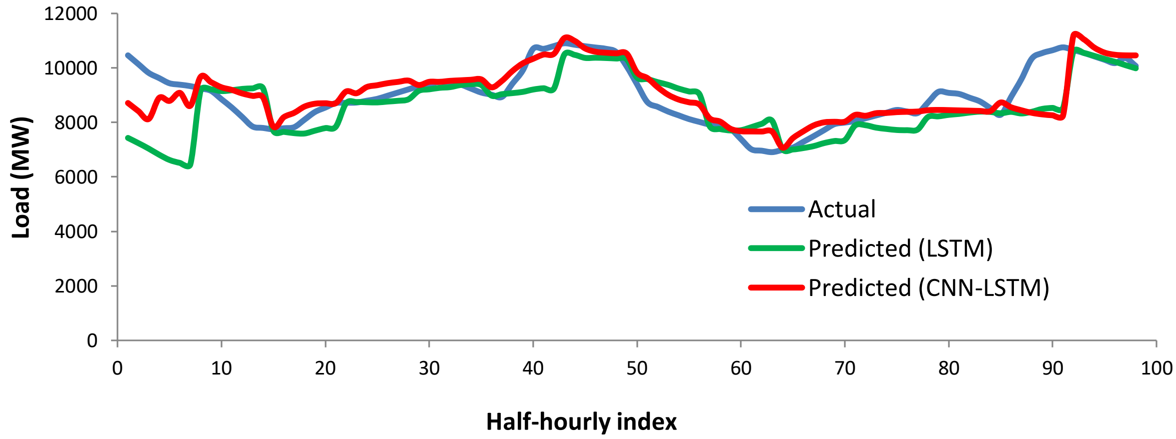

Fig: 5.21. Load forecasting of BPS for 06-07 May 2018

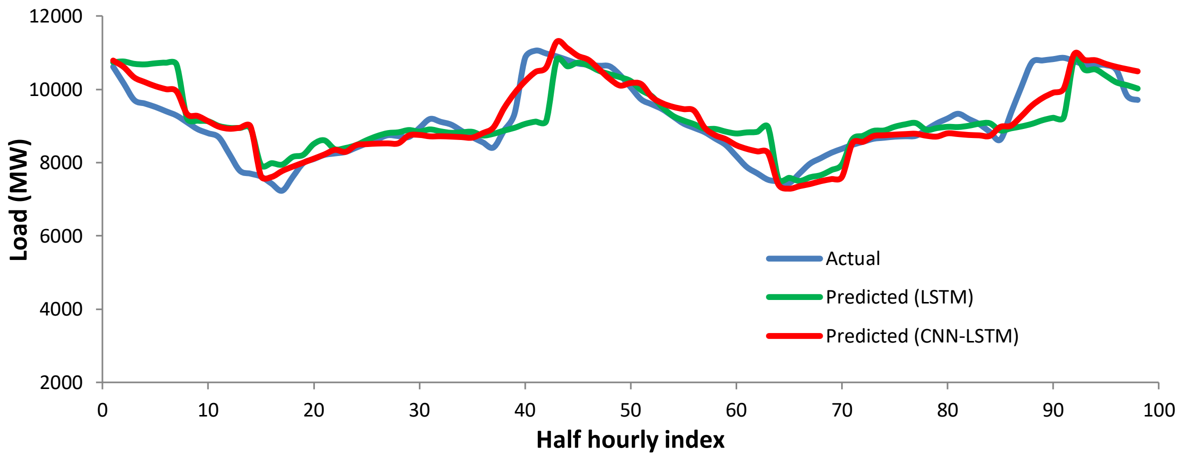

Fig: 5.22. Load forecasting of BPS for 07-08 July 2018

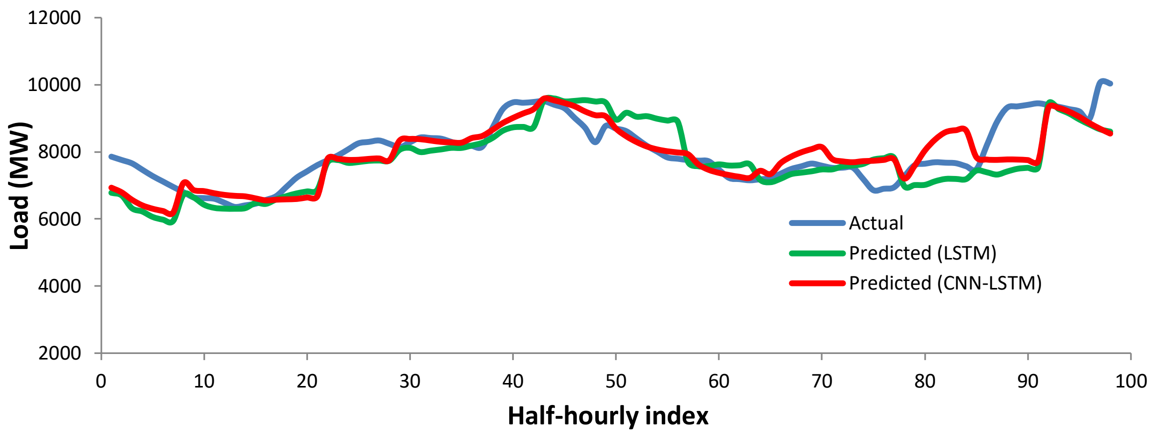

Fig: 5.23. Load forecasting of BPS for 04-05 September 2018

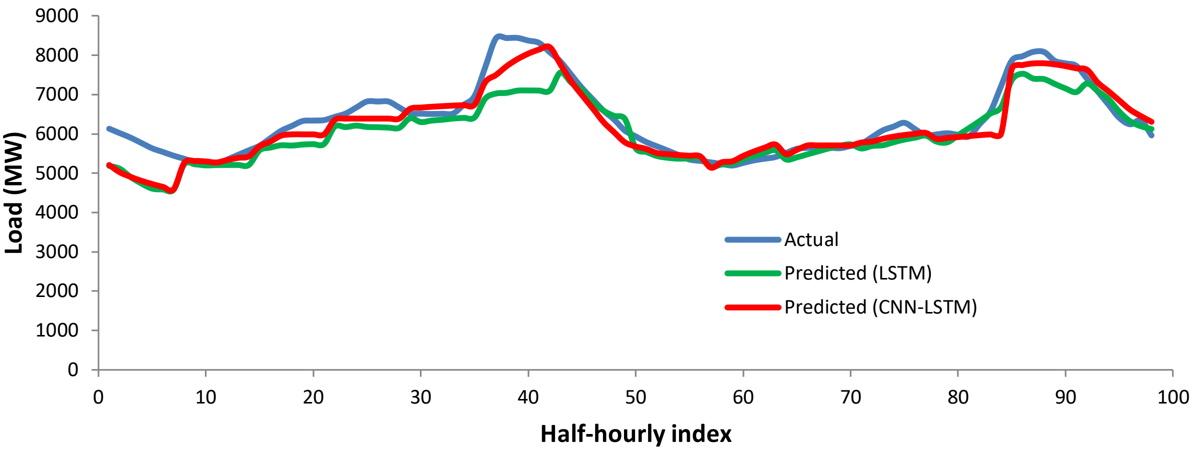

Fig 5.24. Load forecasting of BPS for 20-21 November 2018

5.4.4 24 hours prediction for the year of 2018

It is used to forecast the load for 24 hours at a time such that day ahead using proposed CNN-LSTM model in every month of 2018. The forecasted results for 24 hours in differents days are shown in the following figures (Fig. 5.25 to Fig. 5.26) individually. From the forecasting plots, it has been seen that the forecasting outcomes using proposed CNN-LSTM model follows the actual load trend precisely than the forecasting outcomes of the LSTM model. From the prdiction pltots it is seen that proposed CNN-LSTM model exibits minimum deviation (shown in table 5.4) from actual load demands as compared to LSTM model.

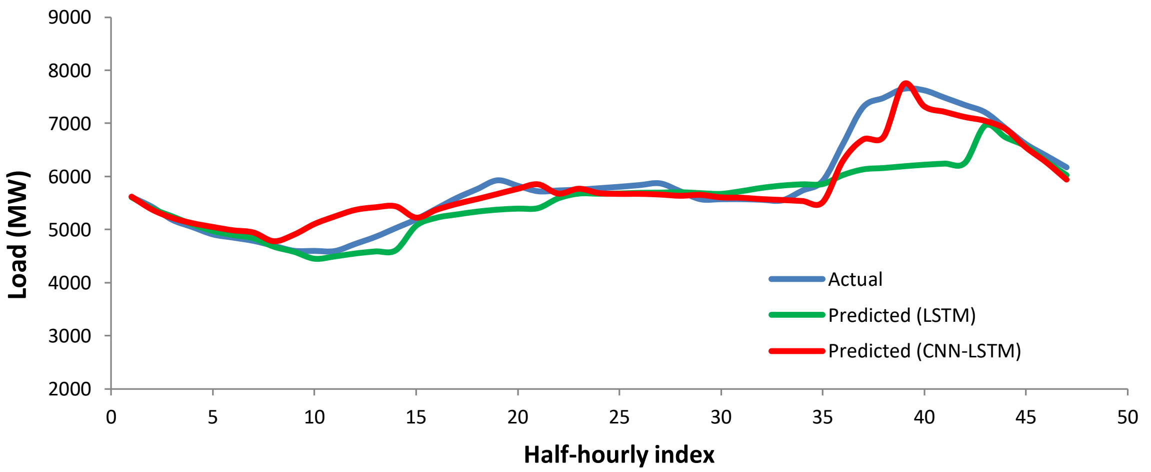

Fig. 5.25. Load forecasting of BPS for 01 January 2018.

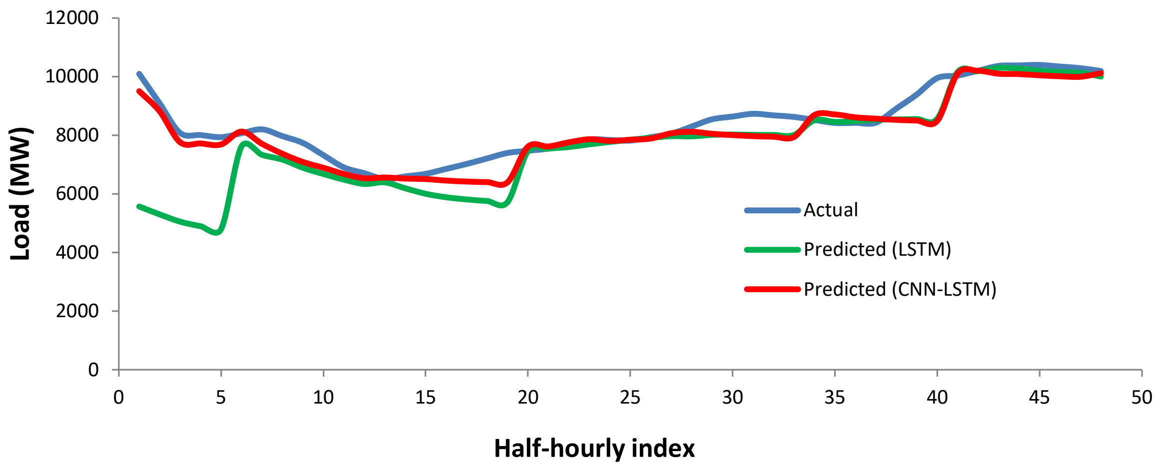

Fig. 5.26. Load forecasting of BPS for 06 February 2018.

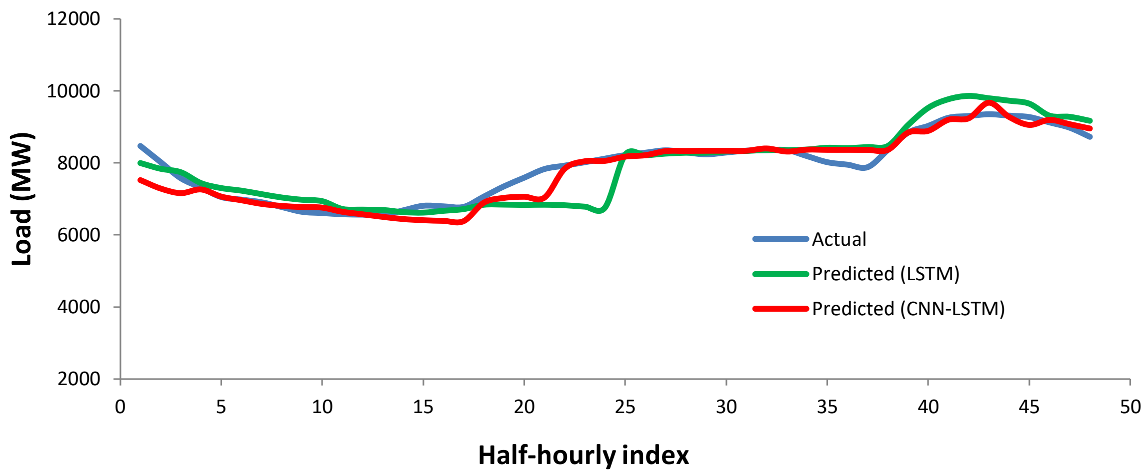

Fig. 5.27. Load forecasting of BPS for 04 March 2018

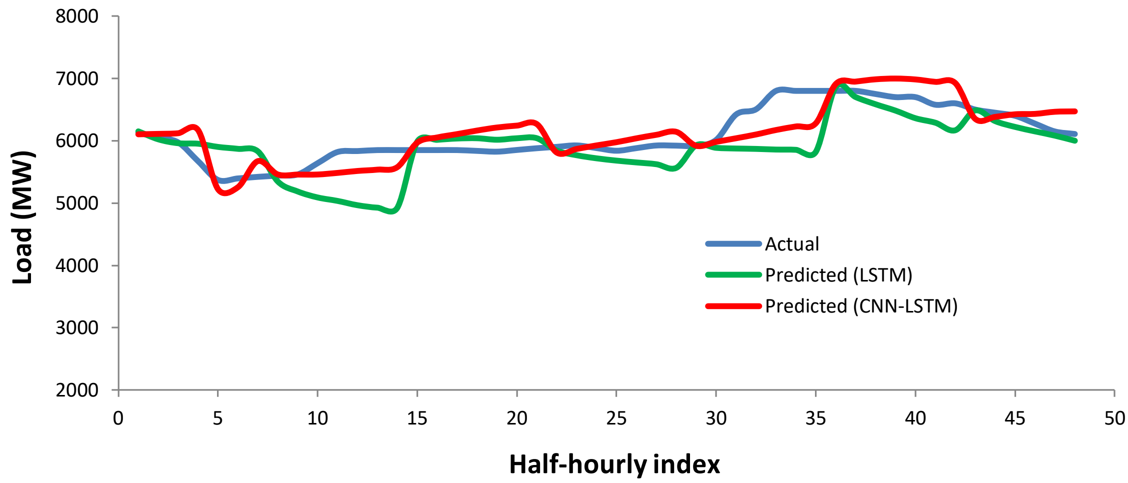

Fig. 5.28. Load forecasting of BPS for 01 April 2018

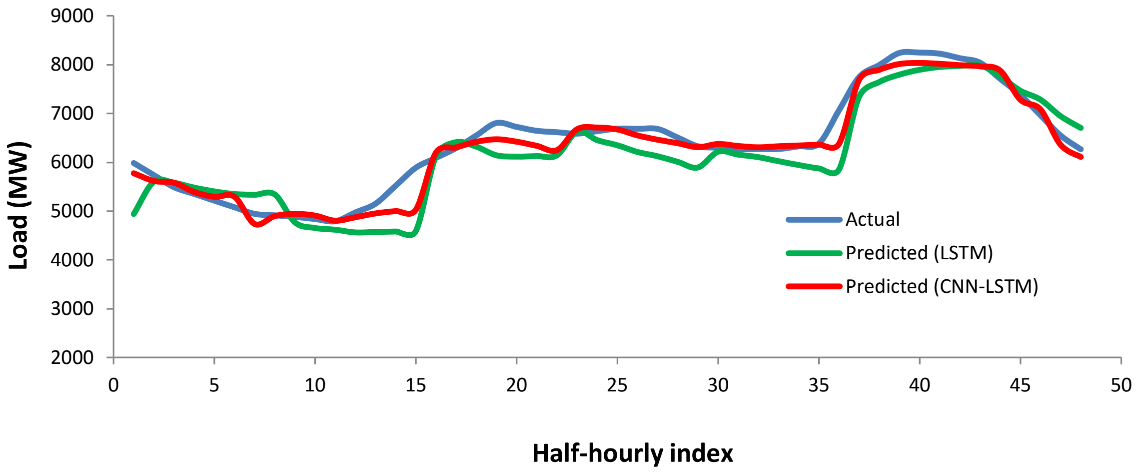

Fig. 5.29. Load forecasting of BPS for 02 May 2018

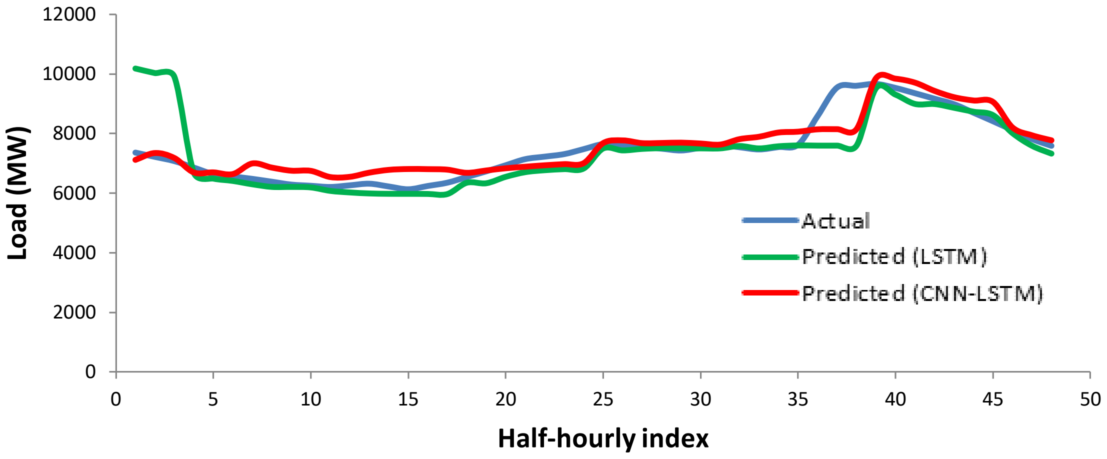

Fig. 5.30. Load forecasting of BPS for 11 June 2018

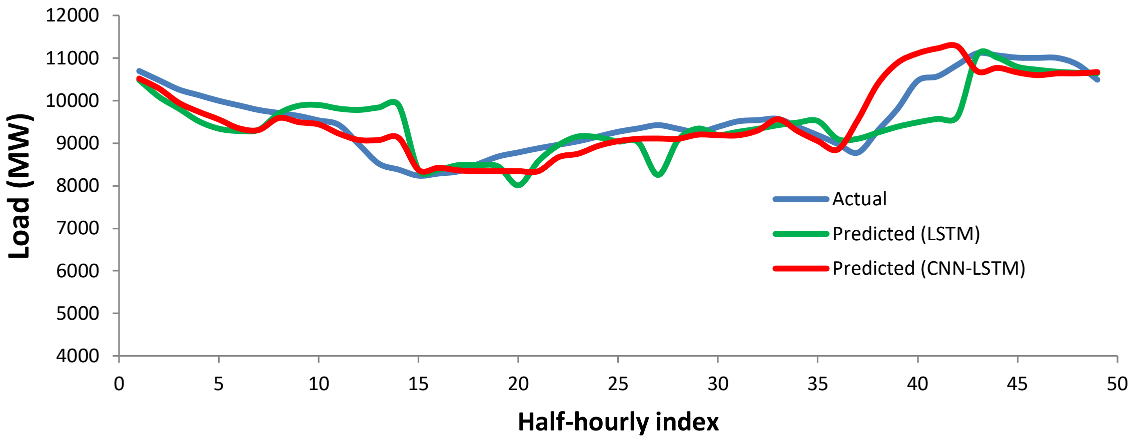

Fig. 5.31. Load forecasting of BPS for 15 July 2018

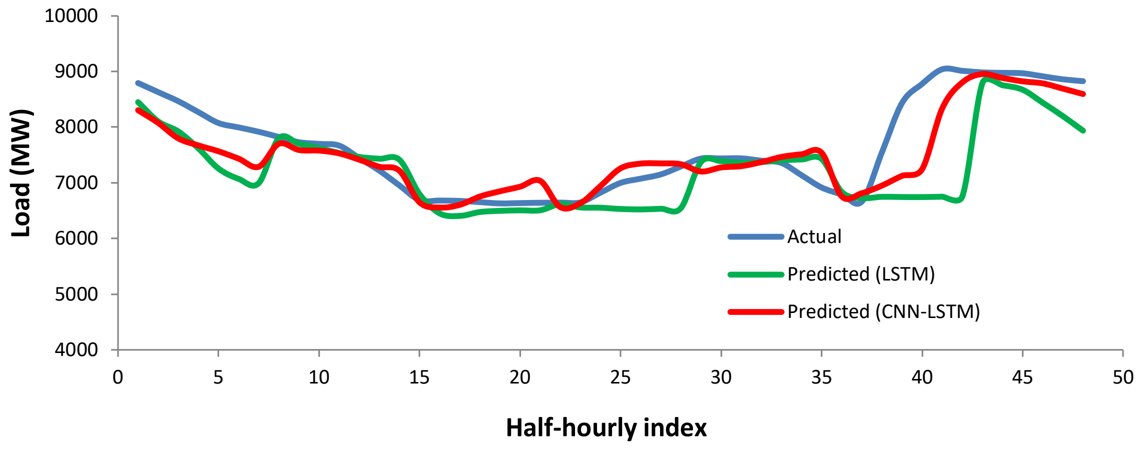

Fig. 5.32. Load forecasting of BPS for 23 August 2018

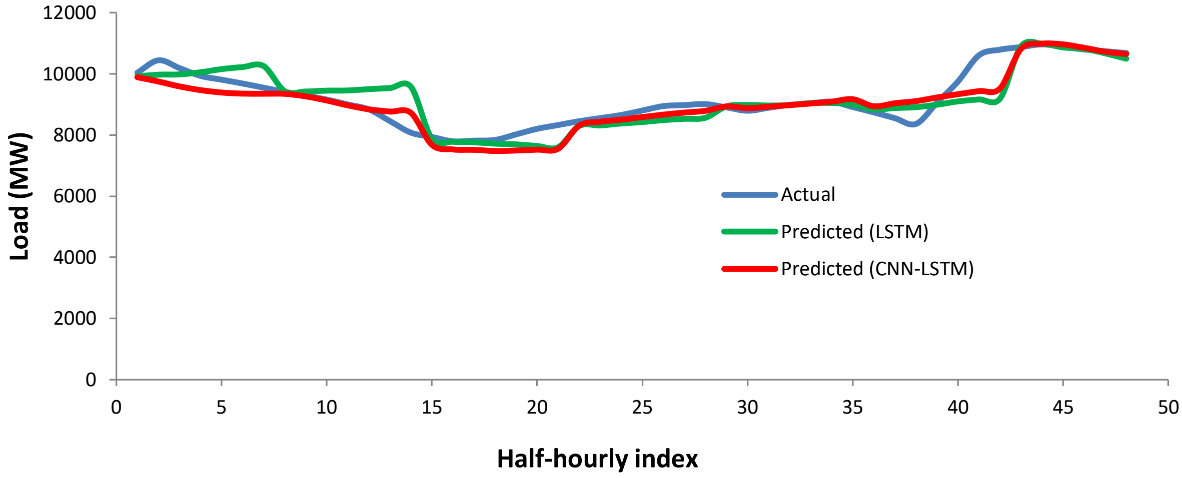

Fig. 5.33. Load forecasting of BPS for 07 September 2018

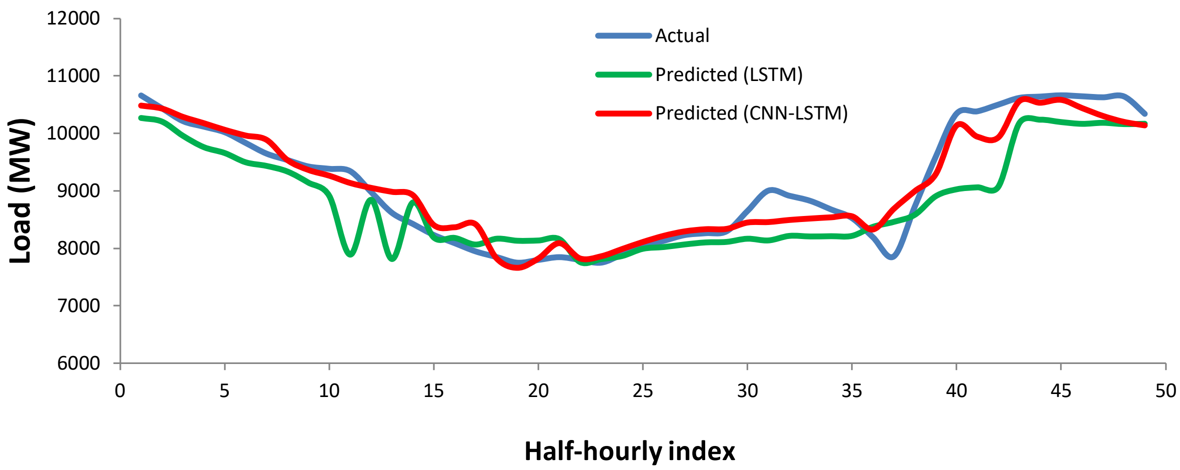

Fig. 5.34. Load forecasting of BPS for 08 October 2018

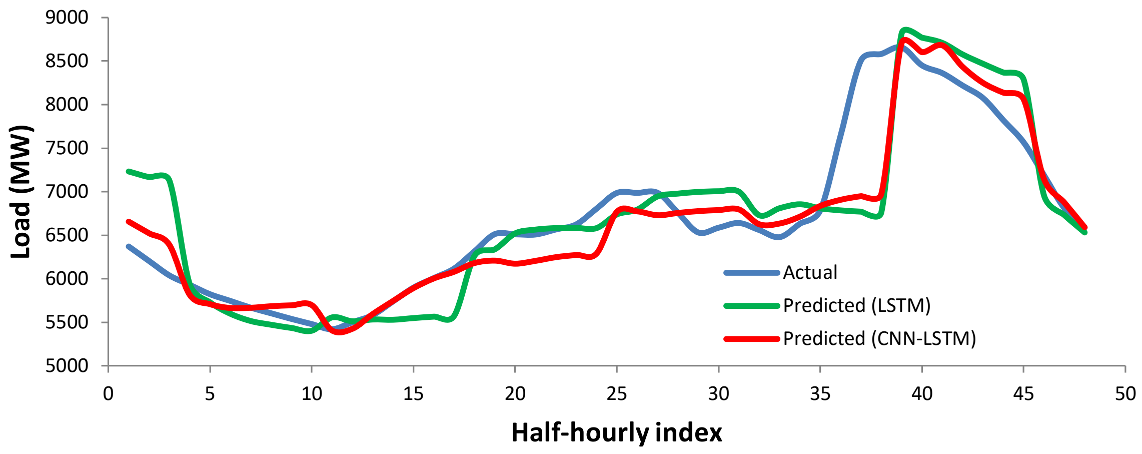

Fig. 5.35. Load forecasting of BPS for 18 November 2018

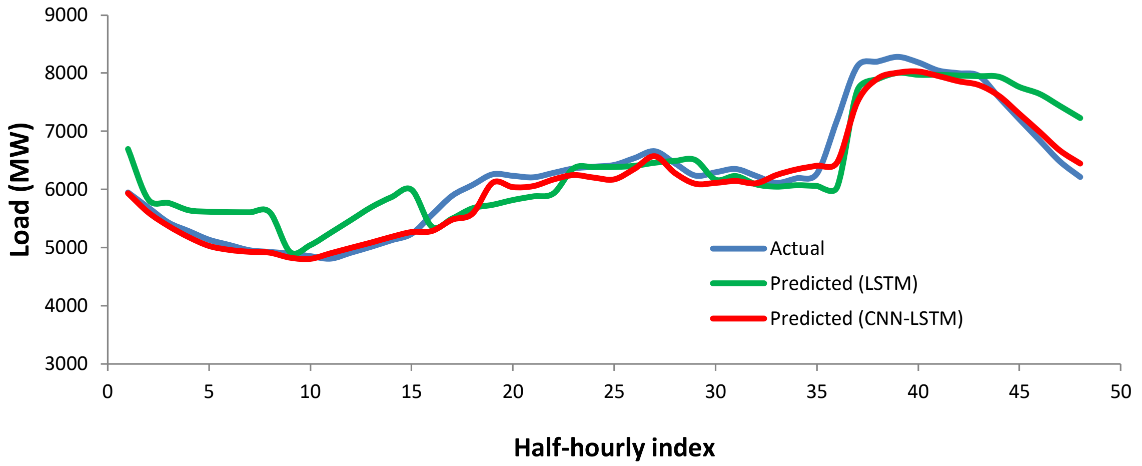

Fig. 5.36. Load forecasting of BPS for 24 December 2018

5.4.5 Monthly prediction for the year of 2019

The proposed models are used to forecast load for each month in 2019. The forecasted results for each month are shown in the following figures (Fig. 5.37 to Fig. 5.48) individually. From all figures, horizontal axis and vertical axis represents time periods in half hourly index and load demand in MW. From the forecasting plots, it has been seen that the forecasting outcomes using proposed CNN-LSTM model follows the actual load trend precisely than the forecasting outcomes of the LSTM model. Prediction curves obtained from proposed CNN-LSTM model has minimum deviation (shown in table 5.5) from actual load demands as compared to LSTM model.

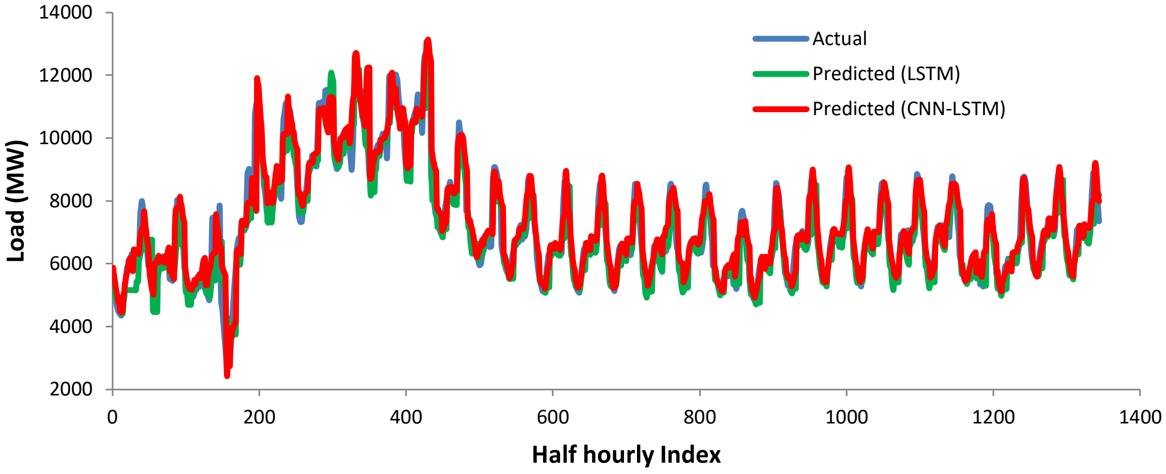

Fig. 5.37. Load forecasting of BPS for January 2019

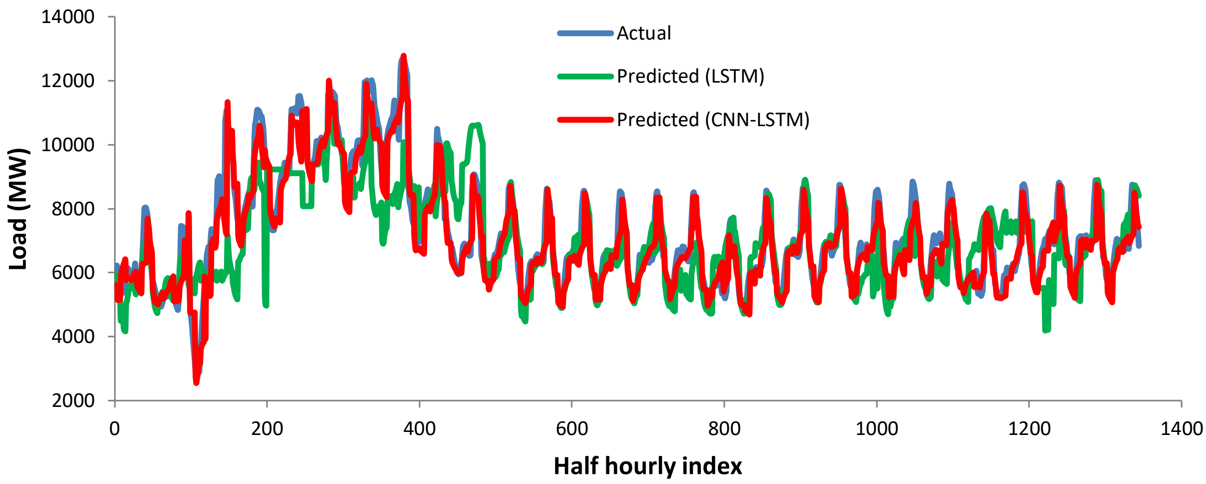

Fig. 5.38. Load forecasting of BPS for February 2019

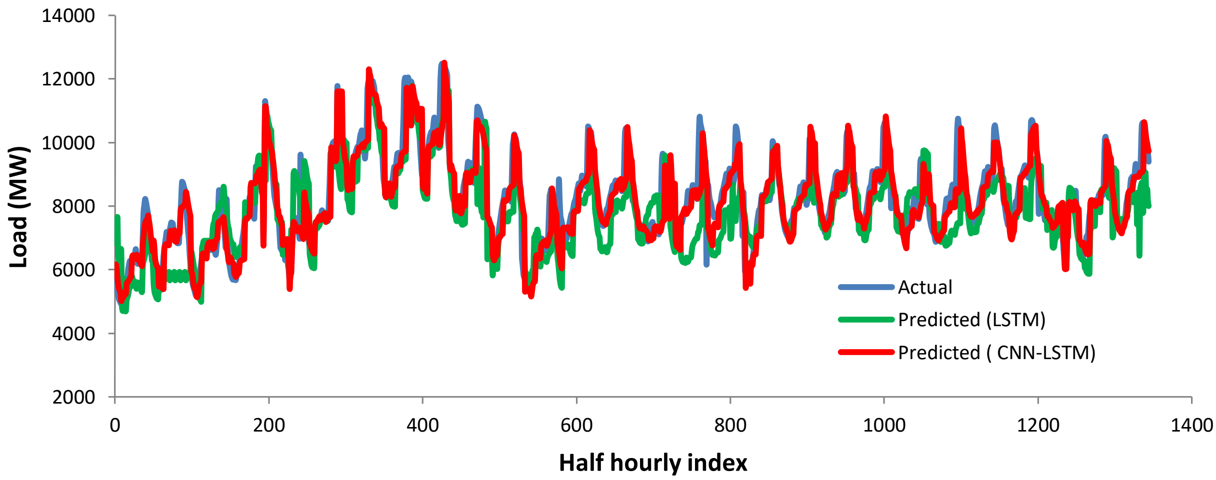

Fig. 5.39. Load forecasting of BPS for March 2019

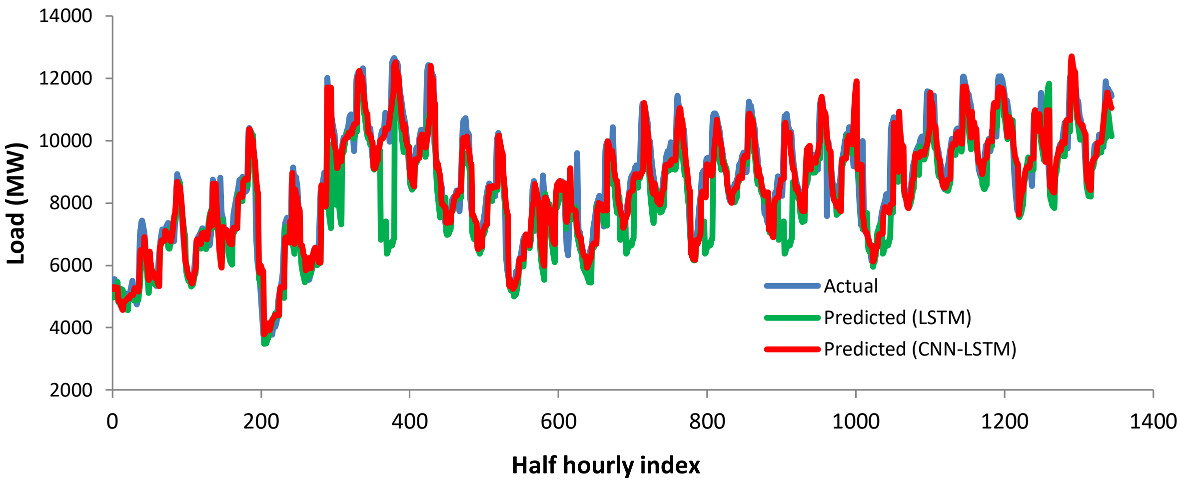

Fig. 5.40. Load forecasting of BPS for April 2019

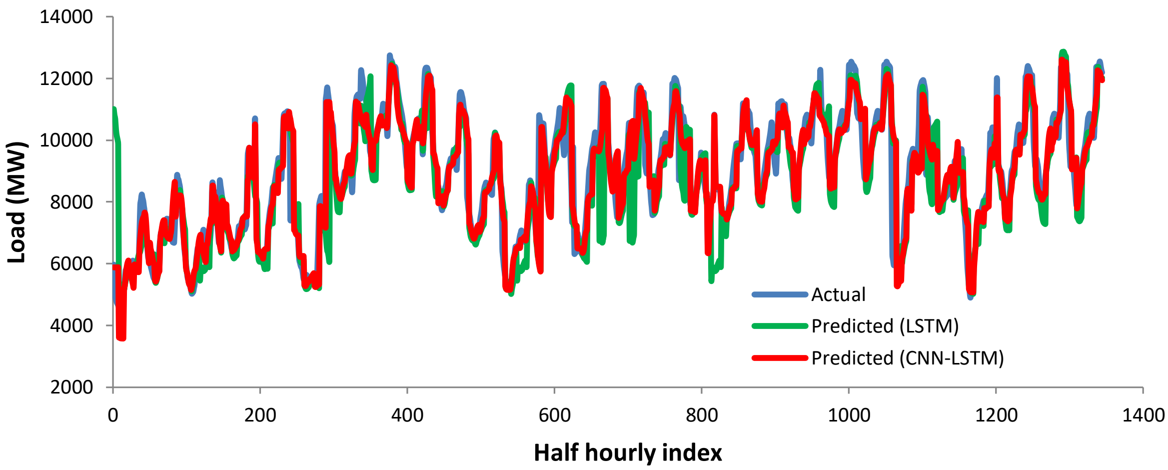

Fig. 5.41. Load forecasting of BPS for May 2019

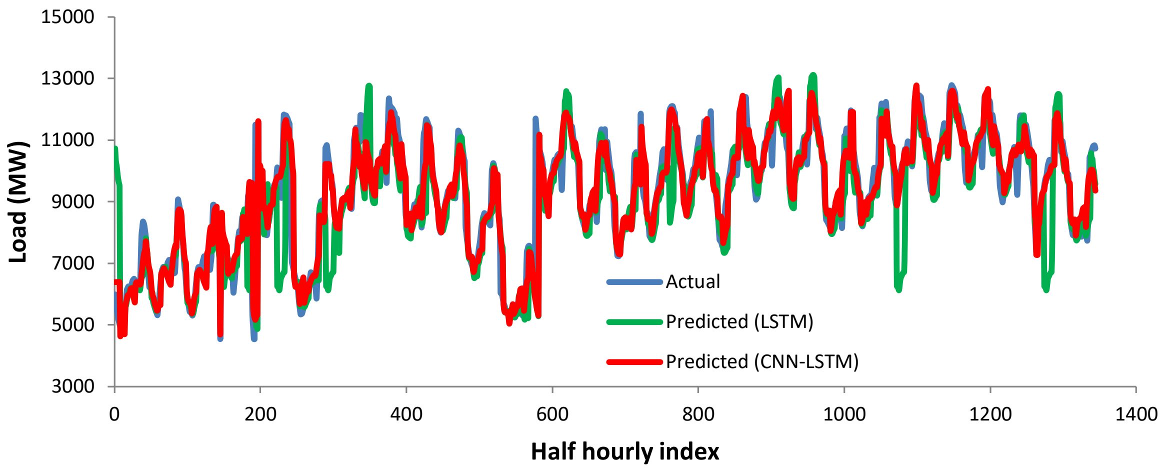

Fig. 5.42. Load forecasting of BPS for June 2019

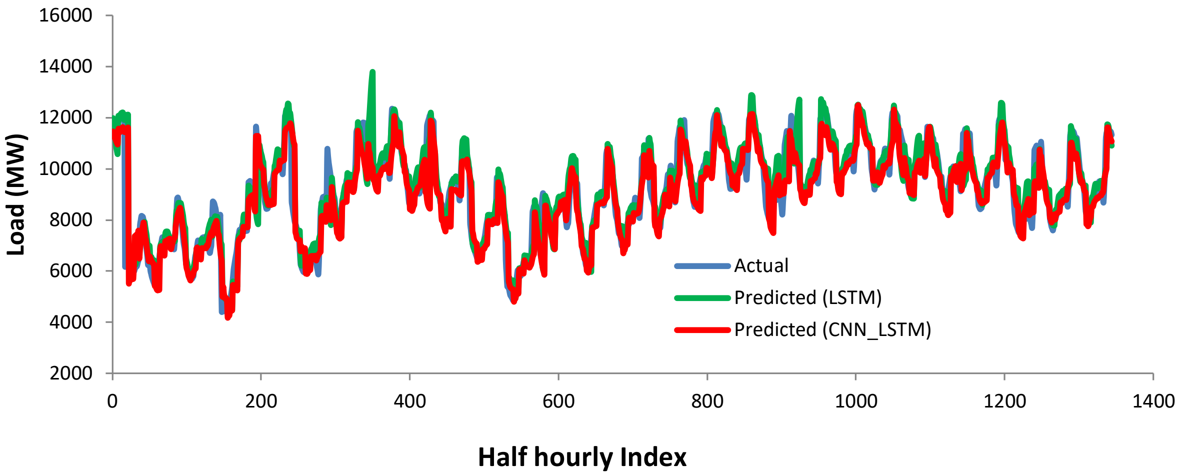

Fig. 5.43. Load forecasting of BPS for July 2019

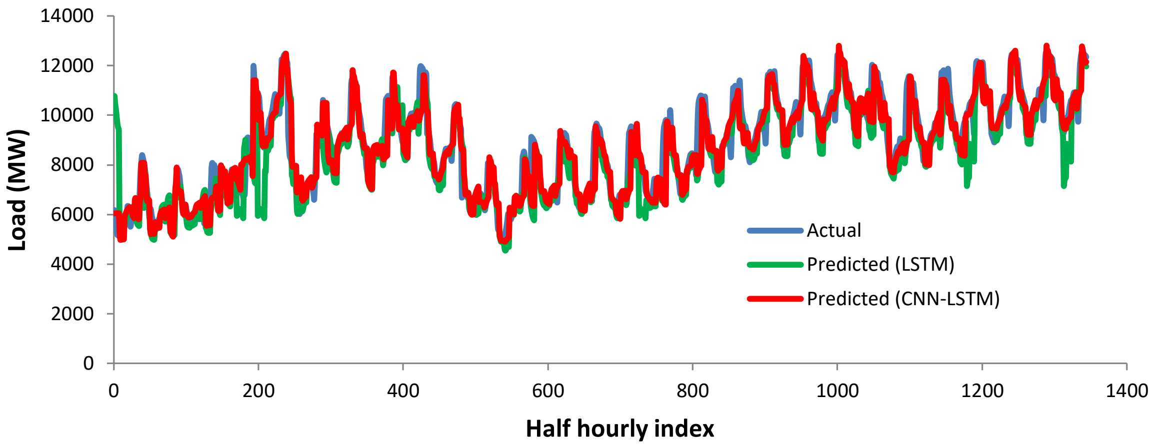

Fig. 5.44. Load forecasting of BPS for August 2019

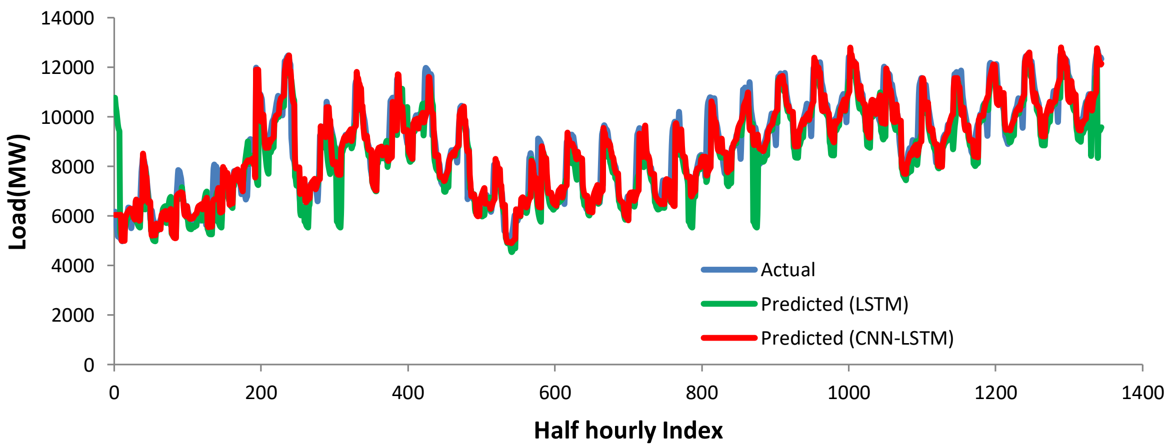

Fig. 5.45. Load forecasting of BPS for September 2019

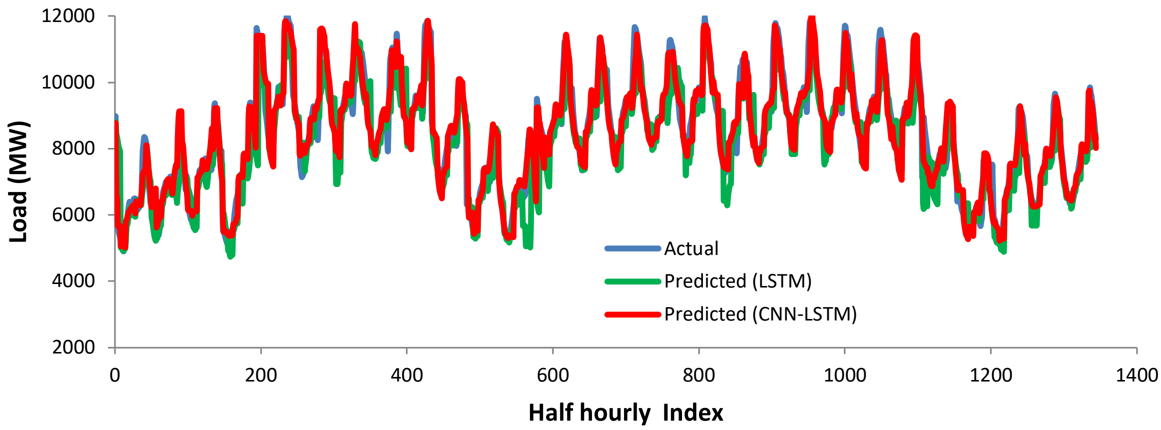

Fig. 5.46. Load forecasting of BPS for October 2019

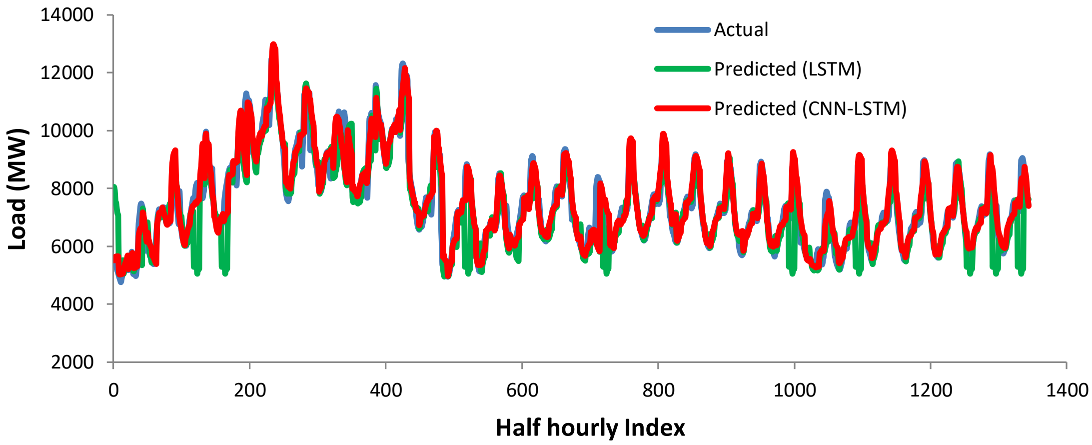

Fig. 5.47. Load forecasting of BPS for November 2019

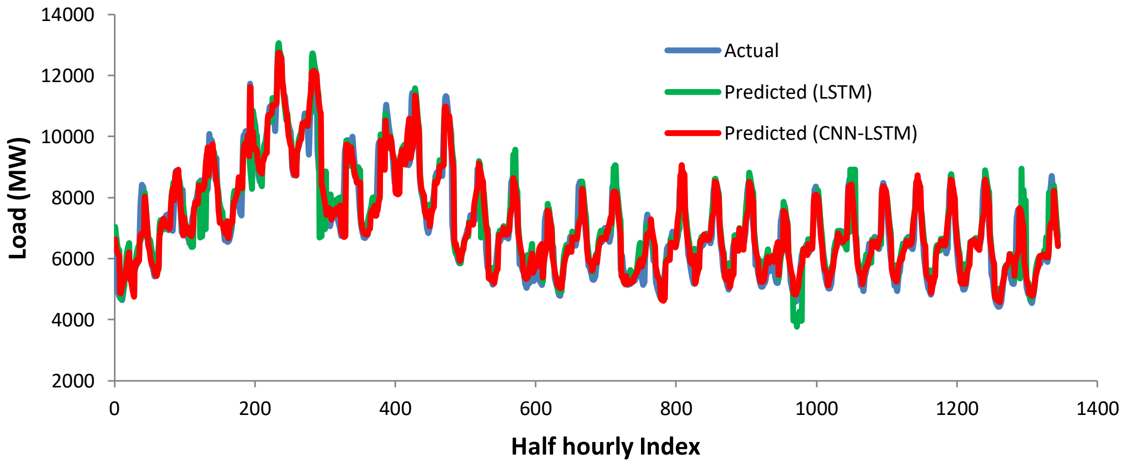

Fig. 5.48. Load forecasting of BPS for December 2019

5.4.6 Weekly prediction for the year of 2019

The proposed models are used to forecast load for each week in every month in 2019. The forecasted results for some weeks in differents month are shown in the following figures (Fig. 5.49 to Fig. 5.60) individually. From all figures, horizontal axis and vertical axis represents time periods in half hourly index and load demand in MW. From the forecasting plots, it has been seen that the forecasting outcomes using proposed CNN-LSTM model follows the actual load trend precisely than the forecasting outcomes of the LSTM model. Prediction curves obtained from proposed CNN-LSTM model has minimum (shown in table 5.6) deviation from actual load demands as compared to LSTM model.

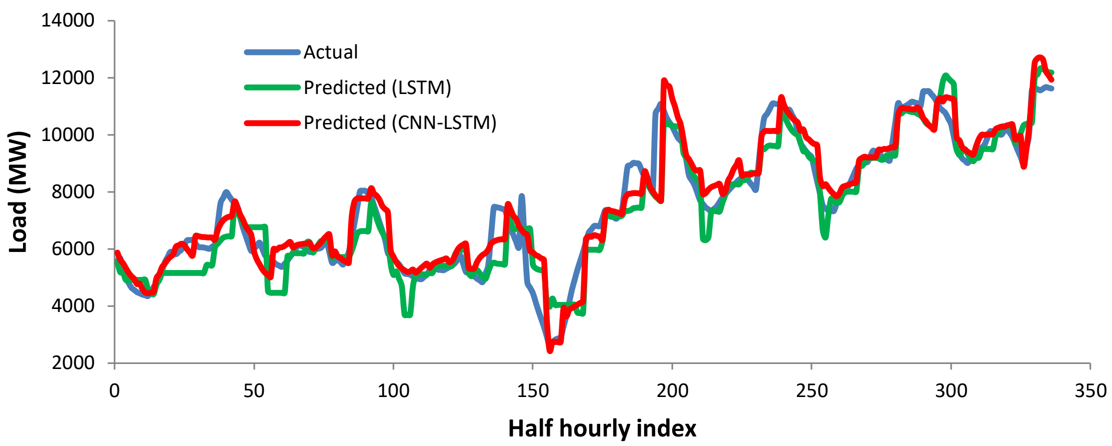

Fig. 5.49. Load forecasting of BPS for 01-07 January 2019

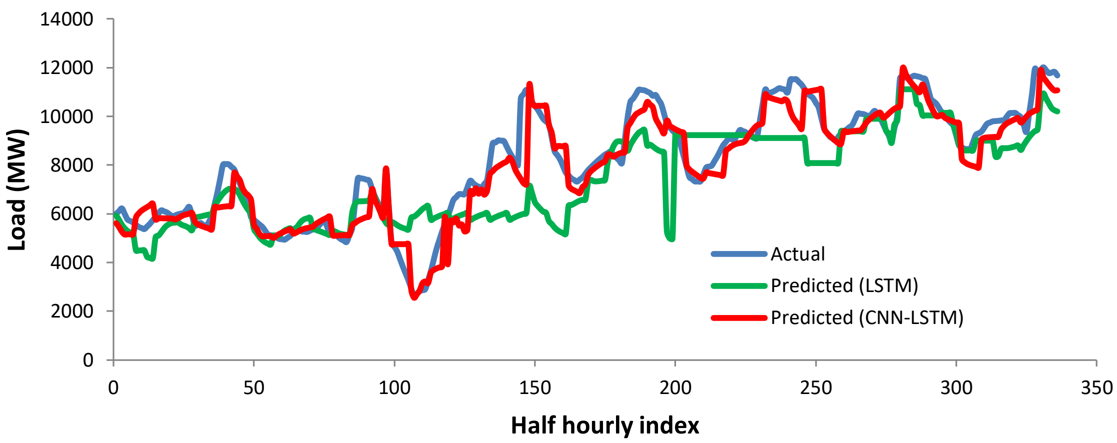

Fig. 5.50. Load forecasting of BPS for 01-07 February 2019

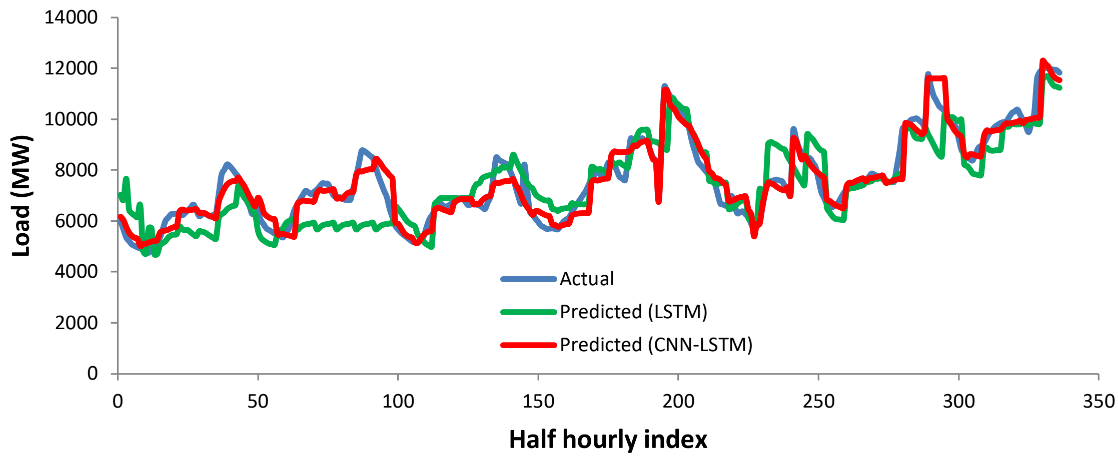

Fig. 5.51. Load forecasting of BPS for 01-07 March 2019

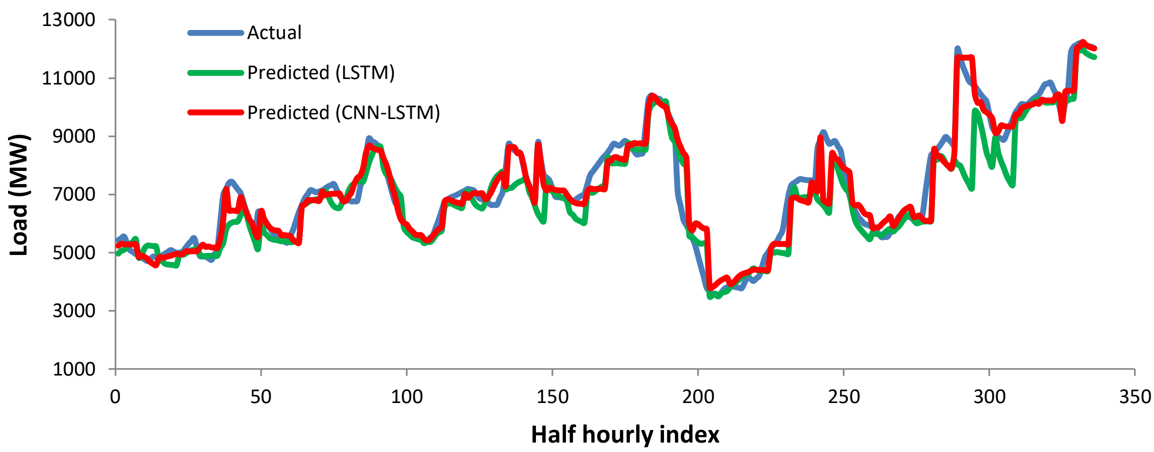

Fig. 5.52. Load forecasting of BPS for 01-07 April 2019

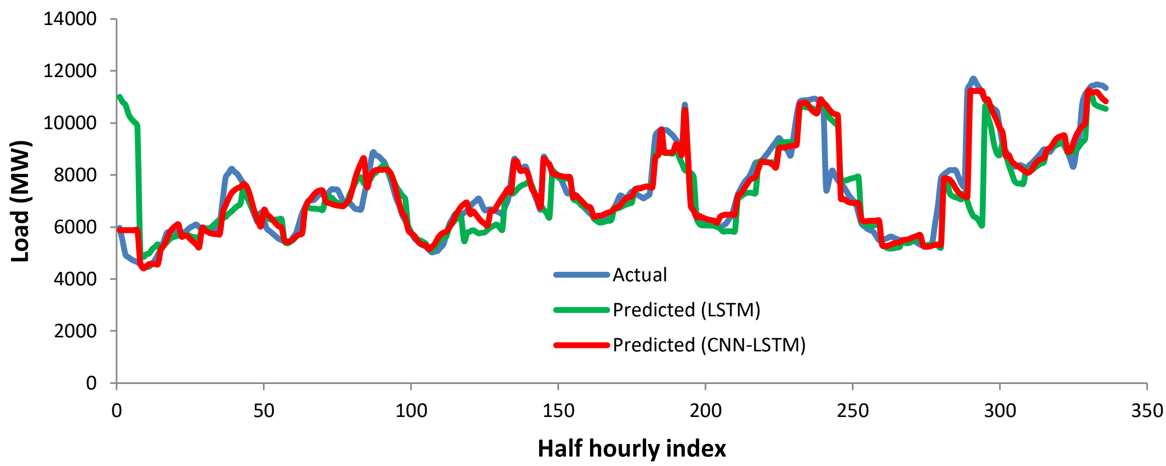

Fig. 5.53. Load forecasting of BPS for 01-07 May 2019

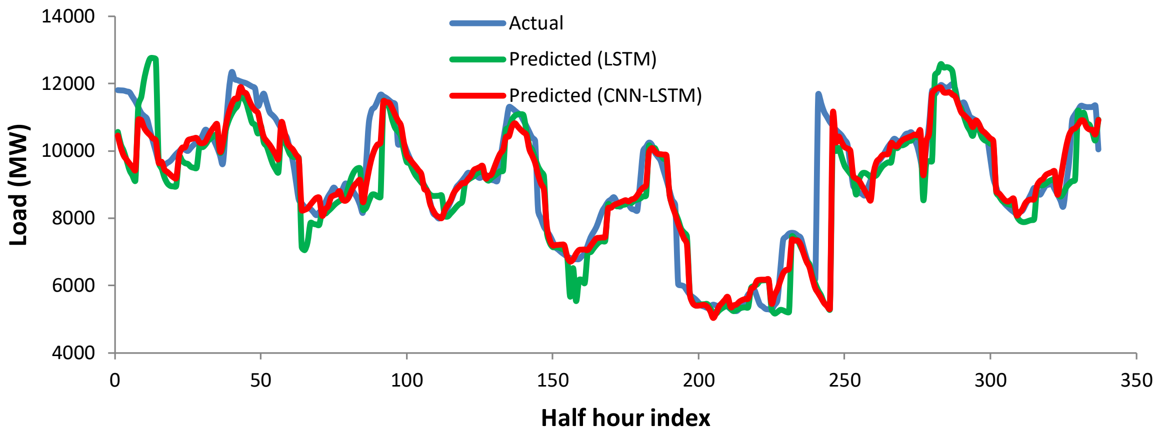

Fig. 5.54. Load forecasting of BPS for 08-14 June 2019

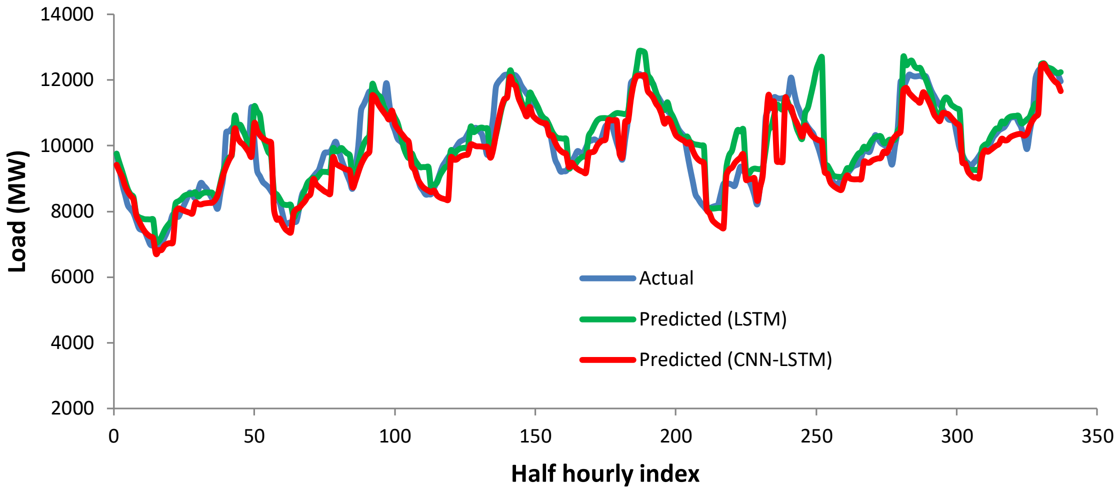

Fig. 5.55. Load forecasting of BPS for 15-22 July 2019

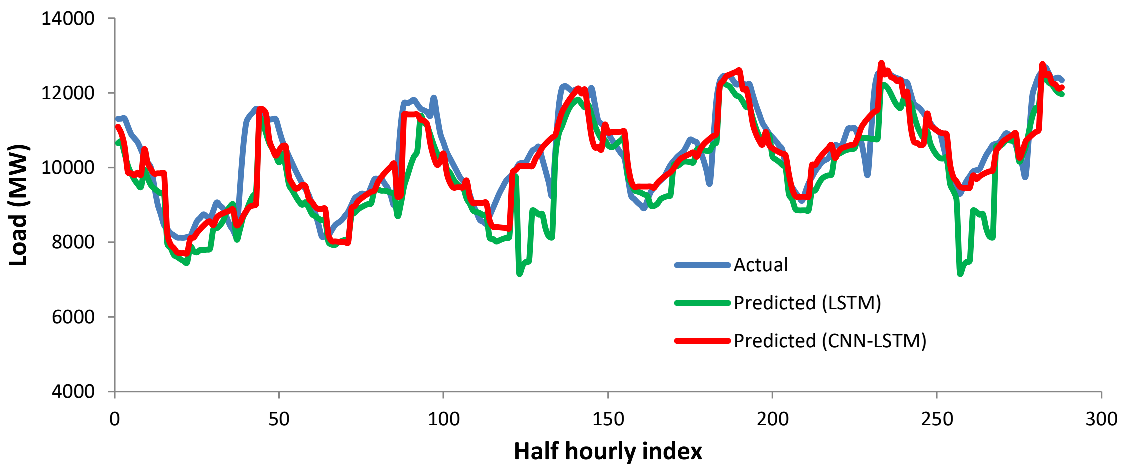

Fig. 5.56. Load forecasting of BPS for 23-29 August 2019

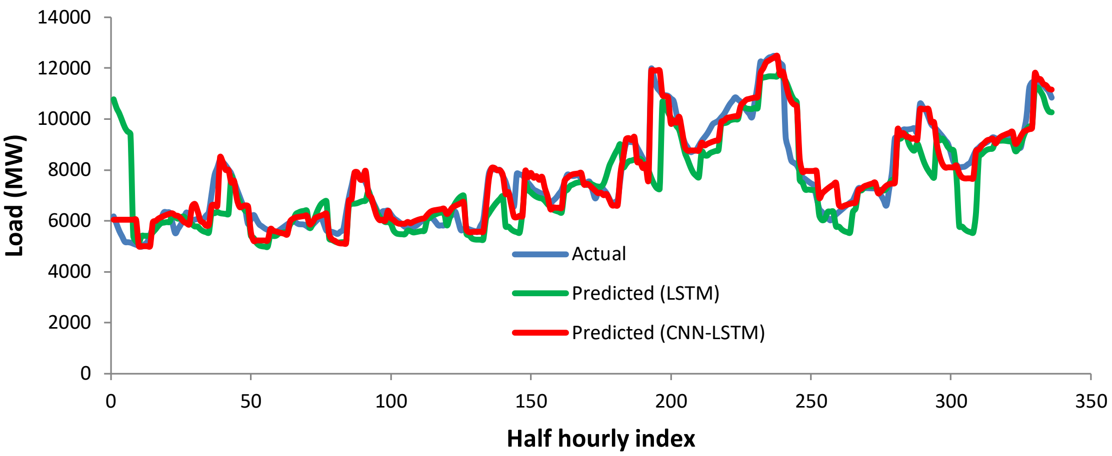

Fig. 5.57. Load forecasting of BPS for 01-07 September 2019

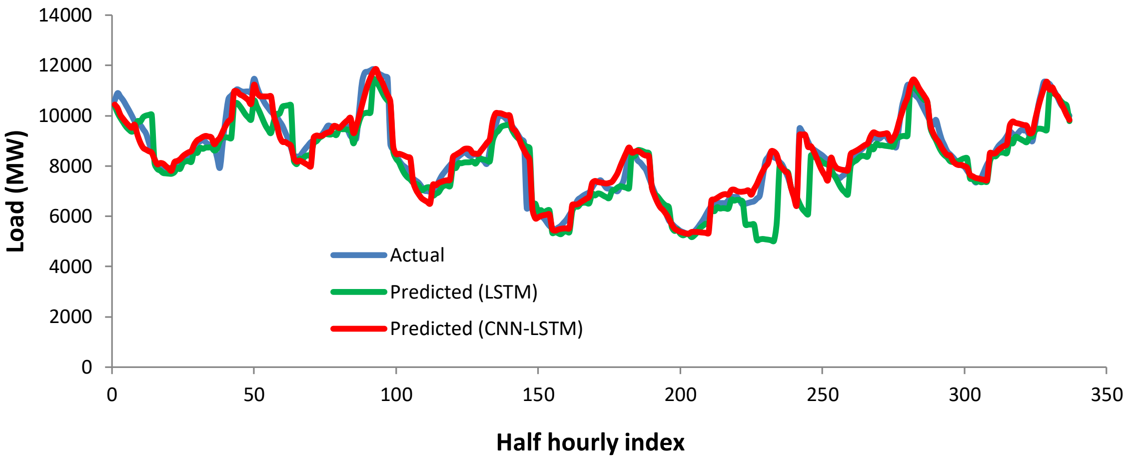

Fig. 5.58. Load forecasting of BPS for 08-14 October 2019

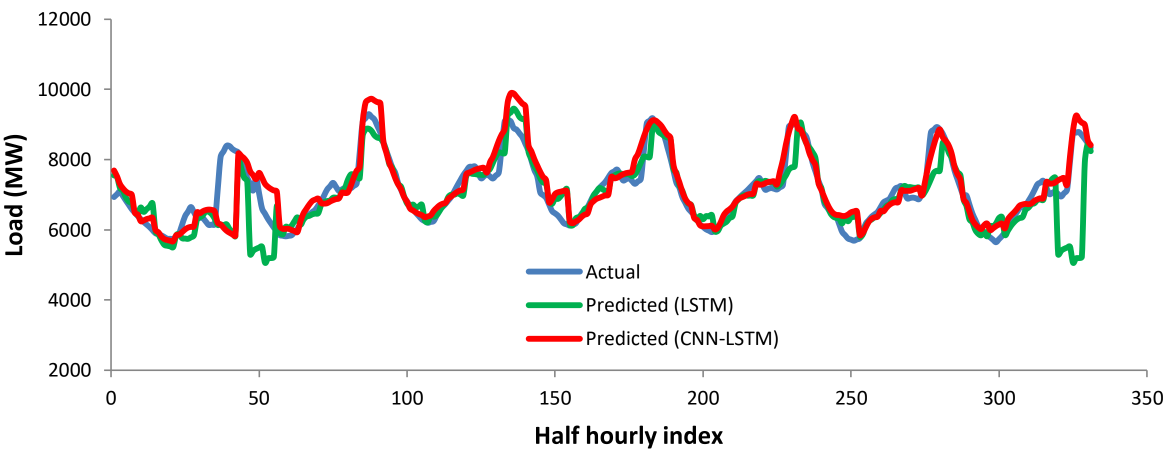

Fig. 5.59. Load forecasting of BPS for 15-21 November 2019

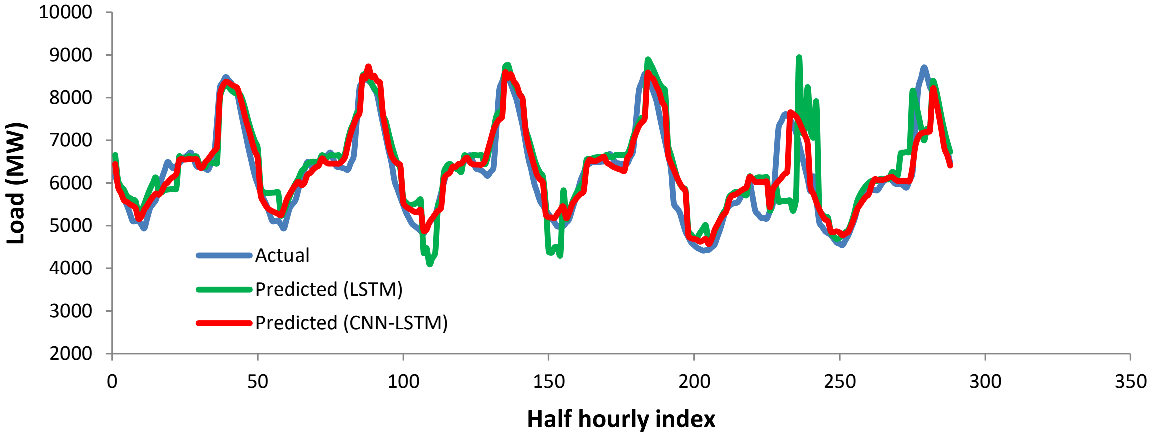

Fig. 5.60. Load forecasting of BPS for 23-29 December 2019

5.4.7 48 hours prediction for the year of 2019

It is used to forecast the load for 48 hours at a time such that prior two days using proposed CNN-LSTM model in every month of 2019. The forecasted results for 48 hours in differents couple of a days are shown in the following figures (Fig. 5.61 to Fig. 5.66) individually. From the forecasting plots, it has been seen that the forecasting outcomes using proposed CNN-LSTM model follows the actual load trend precisely than the forecasting outcomes of the LSTM model. From the prdiction pltots it is seen that proposed CNN-LSTM model exibits minimum (shown in table 5.7) deviation from actual load demands as compared to LSTM model.

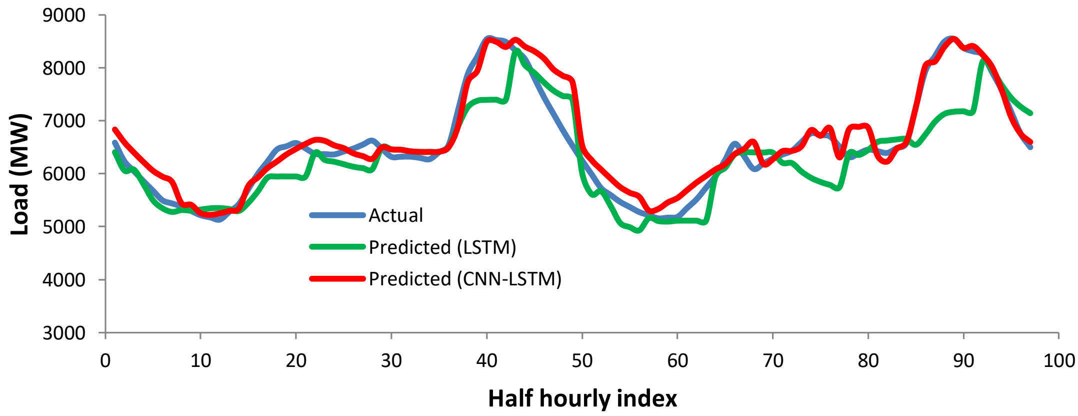

Fig. 5.61. Load forecasting of BPS for 15-16 January 2019

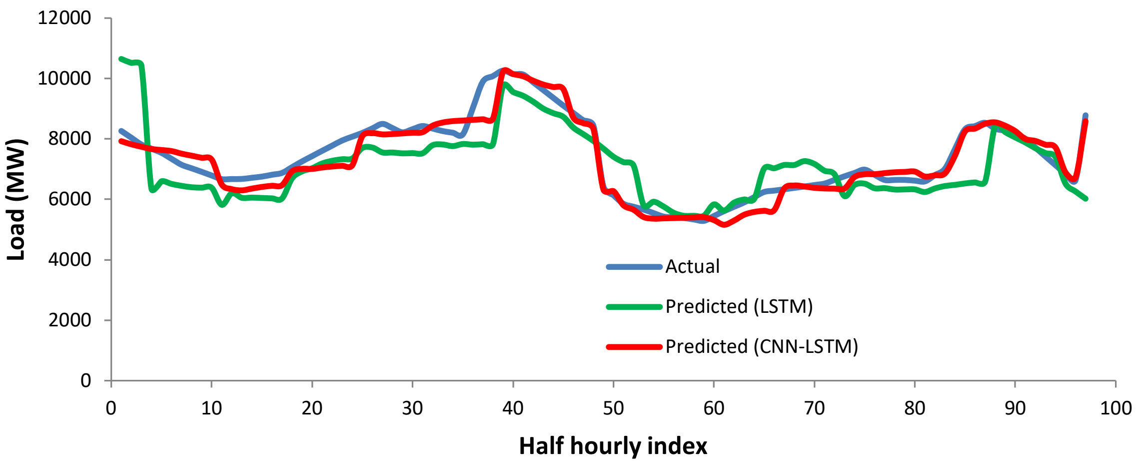

Fig. 5.62. Load forecasting of BPS for 11-12 March 2019

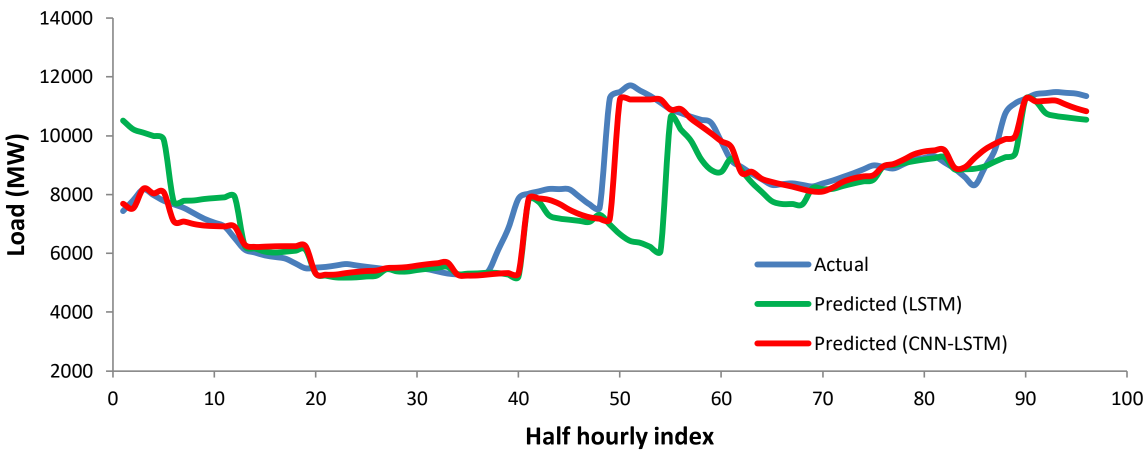

Fig. 5.63. Load forecasting of BPS for 06-07 May 2019

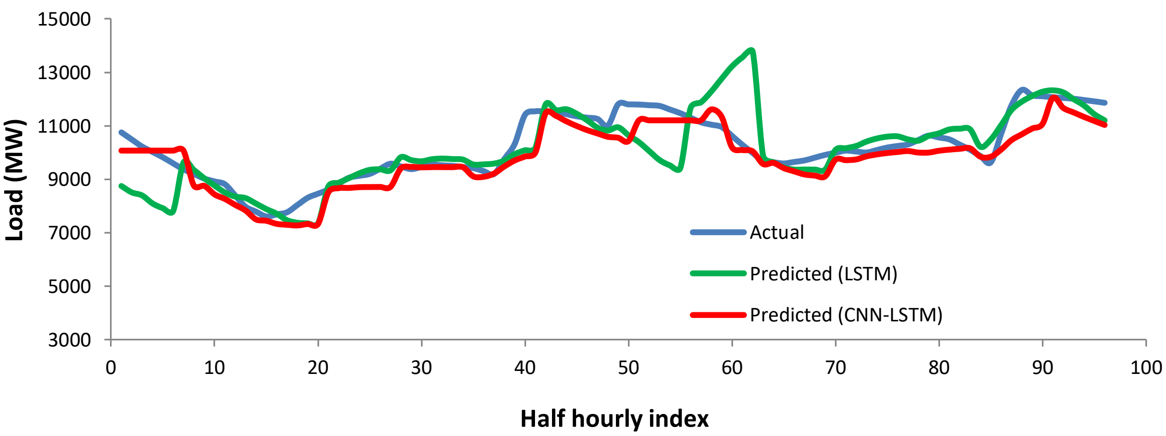

Fig. 5.64. Load forecasting of BPS for 07-08 July 2019

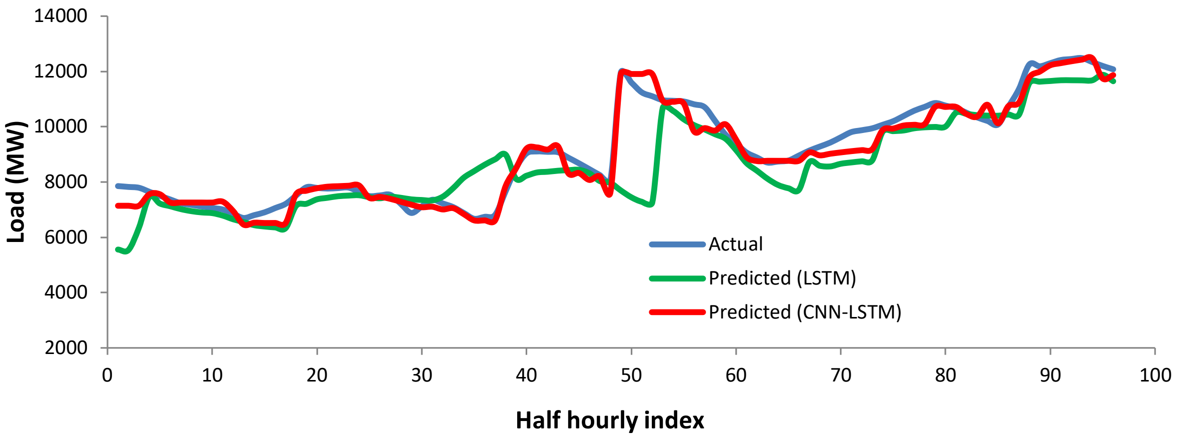

Fig. 5.65. Load forecasting of BPS for 04-05 September 2019

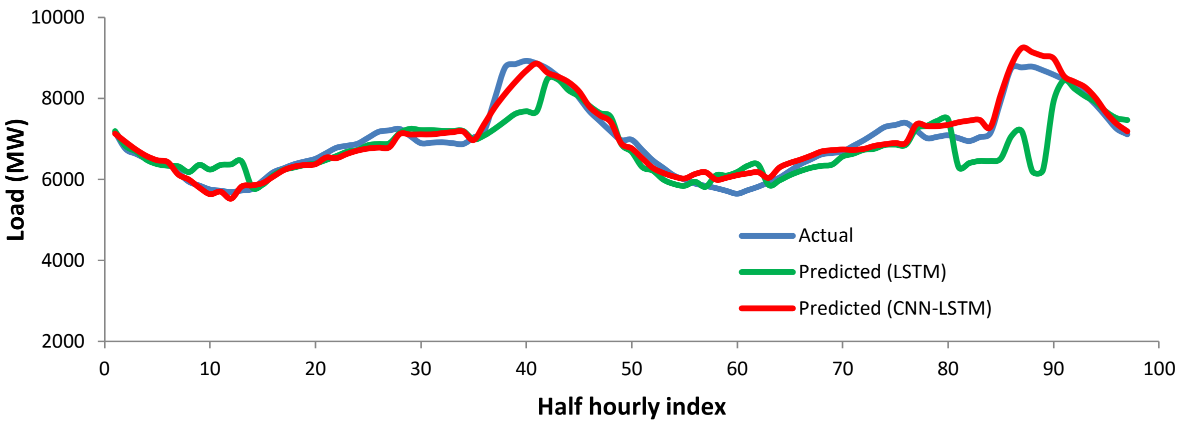

Fig. 5.66. Load forecasting of BPS for 20-21 November 2019

5.4.8 24 hours prediction for the year of 2019

It is used to forecast the load for 24 hours at a time such that day ahead using proposed CNN-LSTM model in every month of 2019. The forecasted results for 24 hours in differents days are shown in the following figures (Fig. 5.67 to Fig. 5.76) individually. From the forecasting plots, it has been seen that the forecasting outcomes using proposed CNN-LSTM model follows the actual load trend precisely than the forecasting outcomes of the LSTM model. From the prdiction pltots it is seen that proposed CNN-LSTM model exibits minimum deviation (shown in table 5.8) from actual load demands as compared to LSTM model.

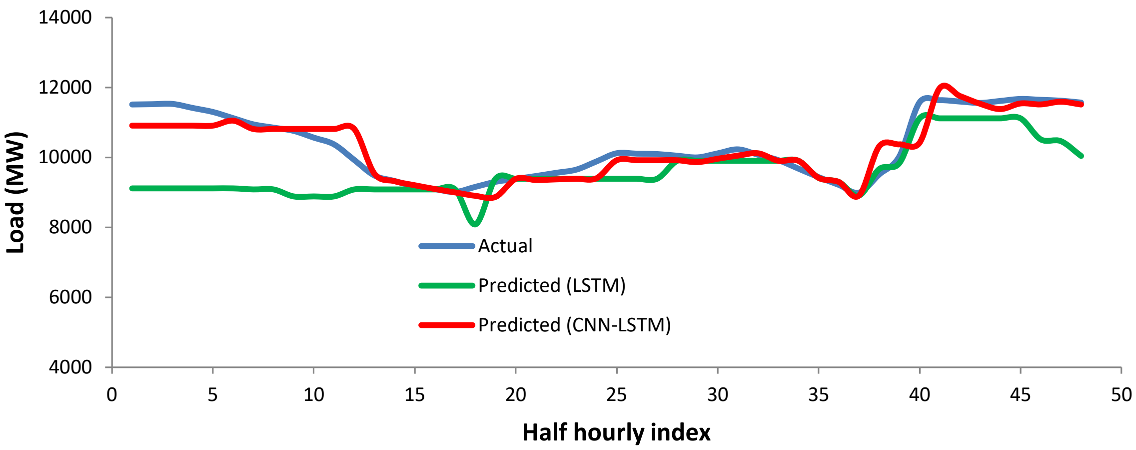

Fig. 5.67. Load forecasting of BPS for 06 February 2019

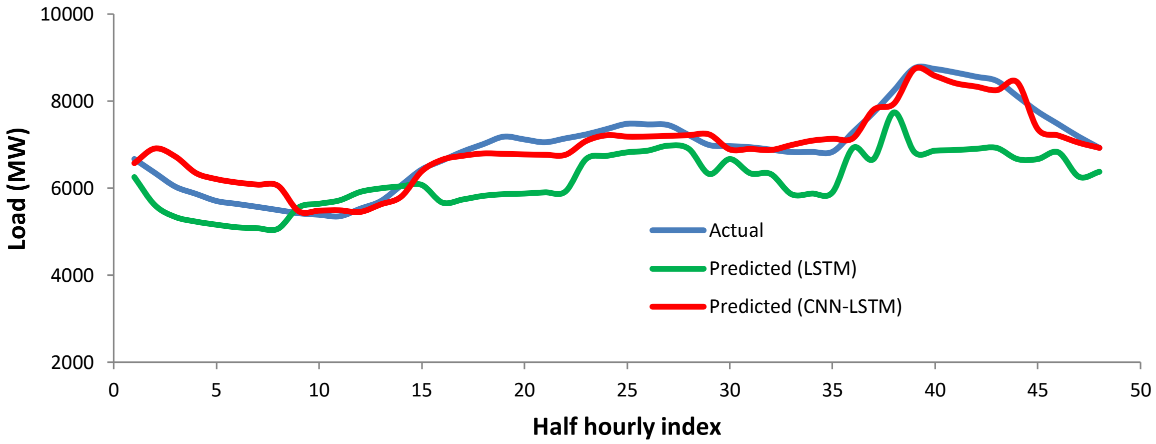

Fig. 5.68. Load forecasting of BPS for 02 March 2019

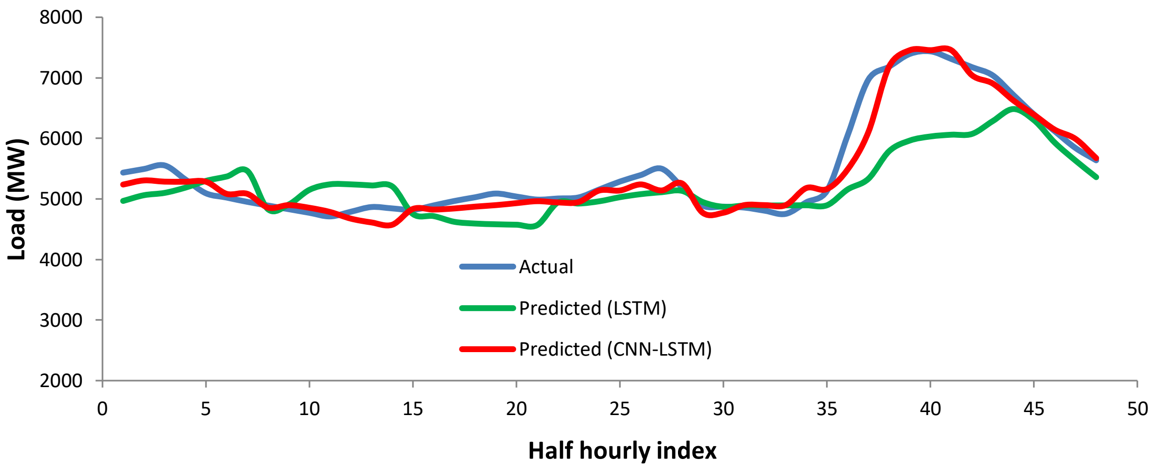

Fig. 5.69. Load forecasting of BPS for 01 April 2019

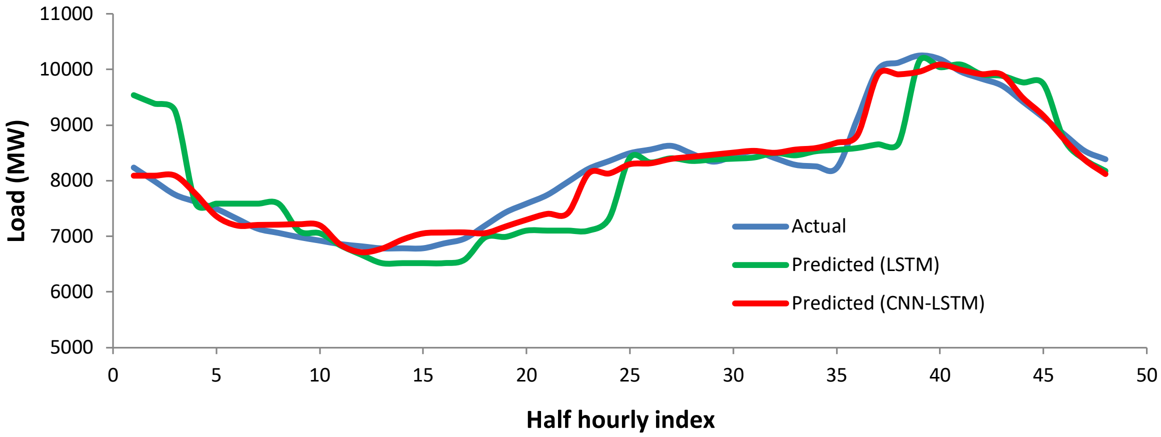

Fig. 5.70. Load forecasting of BPS for 01 June 2019

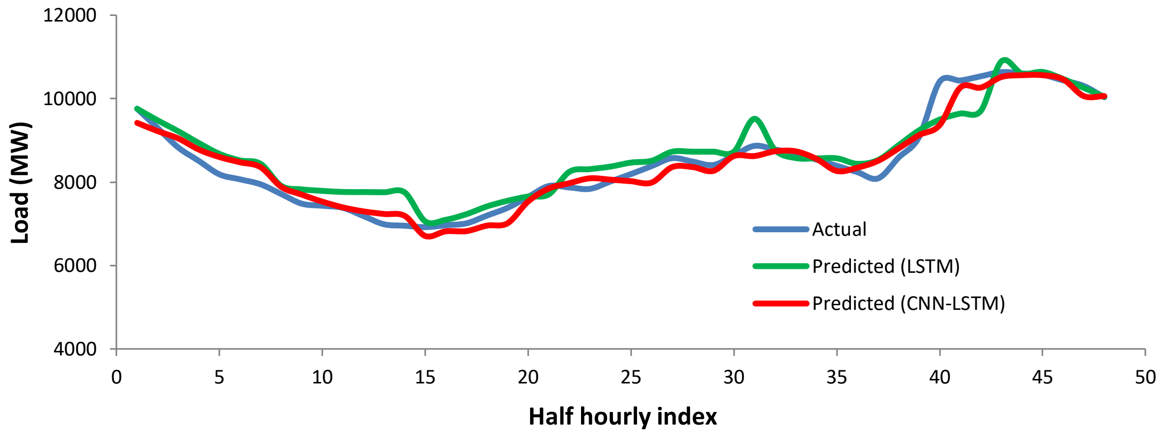

Fig 5.71. Load forecasting of BPS for 15 July 2019

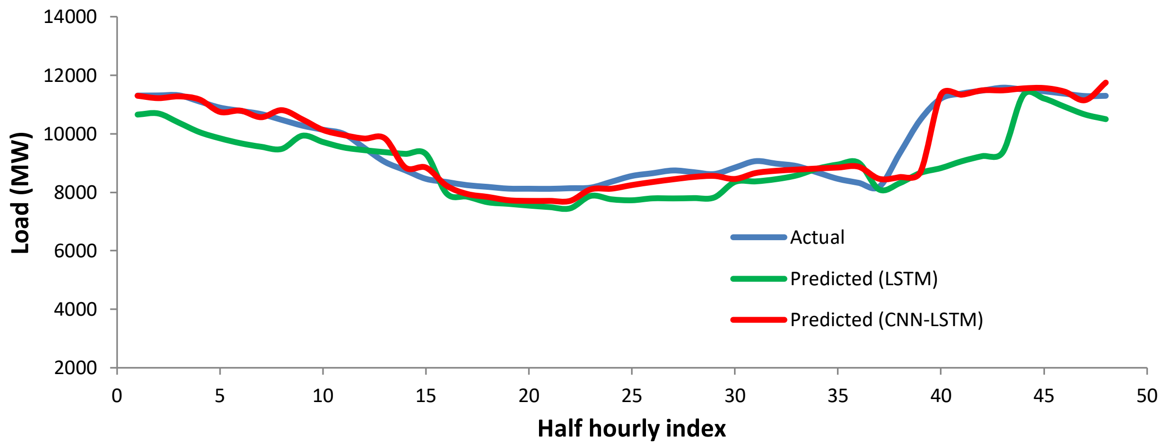

Fig 5.72. Load forecasting of BPS for 23 August 2019

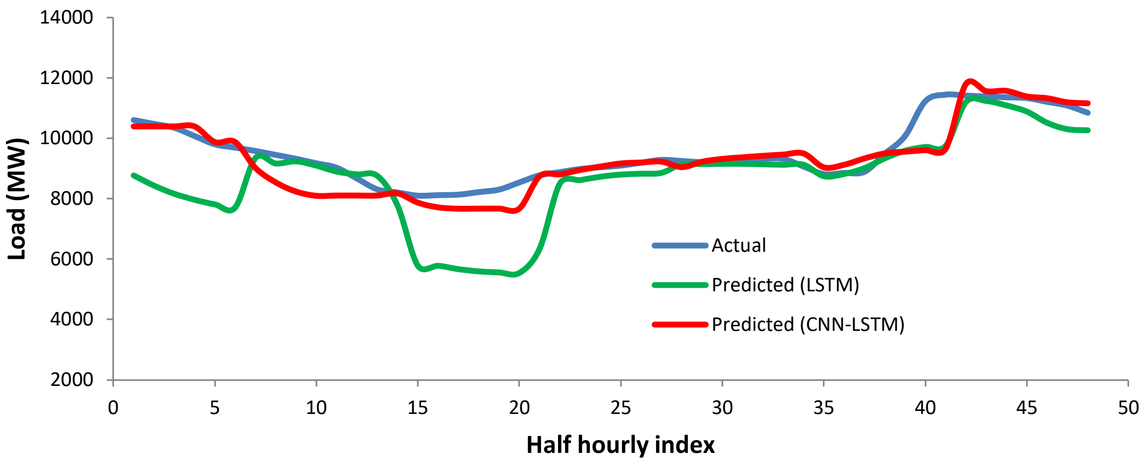

Fig. 5.73. Load forecasting of BPS for 07 September 2019

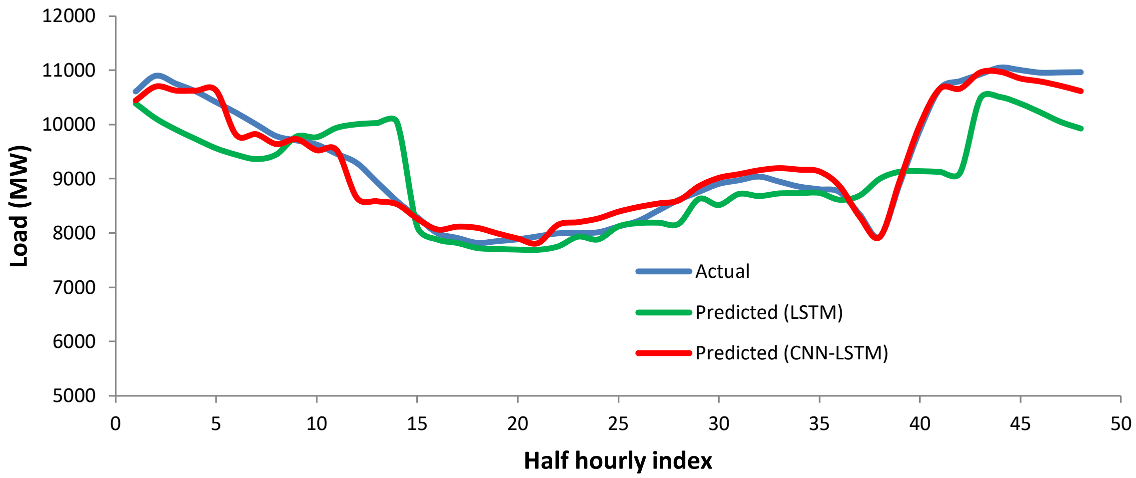

Fig. 5.74. Load forecasting of BPS for 08 October 2019

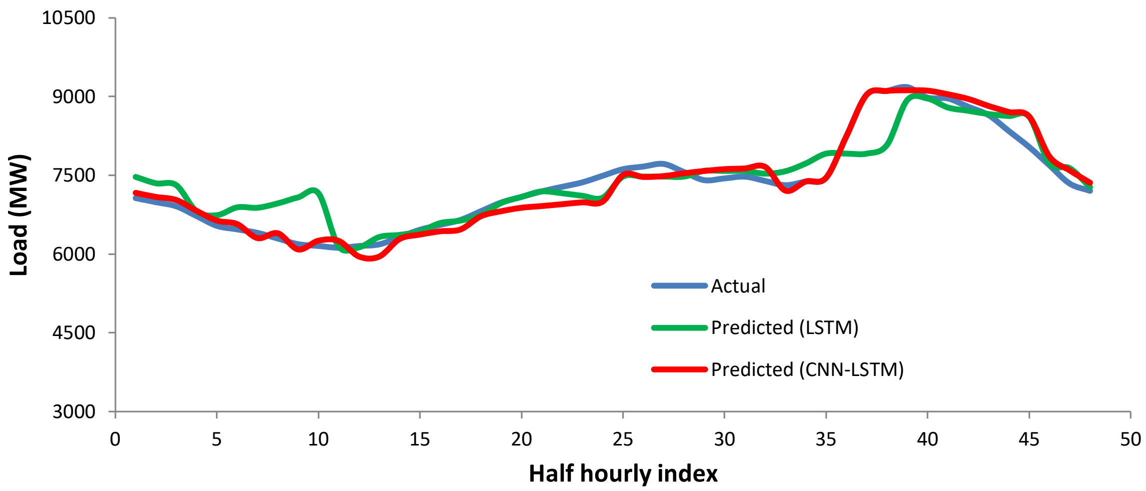

Fig 5.75. Load forecasting of BPS for 18 November 2019

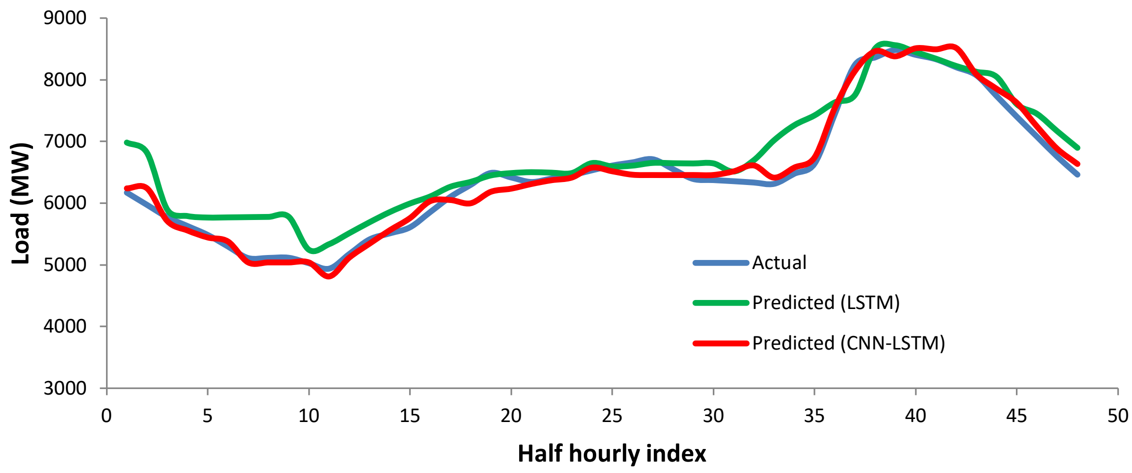

Fig. 5.76. Load forecasting of BPS for 23 December 2019

5.5 Performance Calculation Using Evaluation Metrics

The trained proposed CNN-LSTM and LSTM model provide good performance for different horizons on new load data. Observing the outcomes of MAE, RMSE and MAPE of the forecasted data, we can represent that the load forecasting of CNN-LSTM network outperforms the forecasting results of the LSTM model all the times. The significant difference during certain days, weeks and months ensures the selection of CNN-LSTM is superior to the LSTM model. The proposed method clearly demonstrates that the forecasting results in terms of error matrices are reduced. The performance of the proposed method is now explained with detail comparisons in terms of the error metrics (MAE, RMSE, MAPE) considering different time periods.

5.5.1 Comparison of evaluation matrices for 30 days for the year of 2018

The performance of the proposed CNN-LSTM network is compared with the LSTM network in terms of monthly MAE, RMSE, MAPE in Table 5.2. For example, with the forecasted data in Jan 2018, the LSTM network provides an MAE of 336.48, while the proposed CNN-LSTM offers an MAE of 265.67. Similarly, the proposed method offers less MAE with the data forecasted data in Dec 2018. The similar trend is also evident from the data available in other months. In average, the proposed method provides 47, 119.52 and

0.7% less MAE, RMSE and MAPE than the LSTM network respectively. The improvement of all metrics is also graphically illustrated in the Fig. 5.77 to Fig. 5.79.

Table 5.2: Monthly MAE, RMSE and MAPE for LSTM and CNN-LSTM

Observation Period	MAE		RMSE		MAPE	
	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM
Jan 18	336.48	265.67	552.83	416.09	4.94	4.10
Feb 18	421.68	376.91	621.92	578.48	6.60	5.77
Mar 18	364.66	314.41	570.72	492.41	4.84	4.01
Apr 18	575.43	526.17	817.86	676.36	8.09	7.46
May 18	432.37	420.92	646.28	564.69	5.66	5.59
Jun 18	431.96	391.46	717.09	554.90	5.78	5.30
July 18	388.03	361.63	657.26	551.80	4.79	4.34
Aug 18	491.58	388.62	840.20	706.53	6.31	4.94
Sep 18	471.79	386.76	745.05	562.88	5.71	4.41
Oct 18	394.06	336.45	713.56	514.08	5.29	4.53
Nov 18	328.40	314.23	578.29	481.74	4.90	4.69
Dec 18	344.71	325.11	577.13	504.02	5.40	4.96
Average	415.1	367.36	669.85	550.33	5.69	5.01

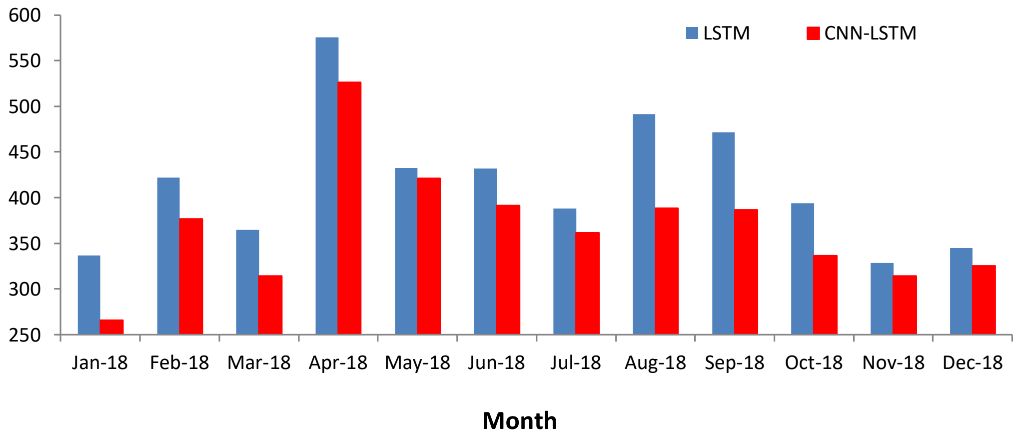

Fig. 5.77. Monthly MAE for 2018

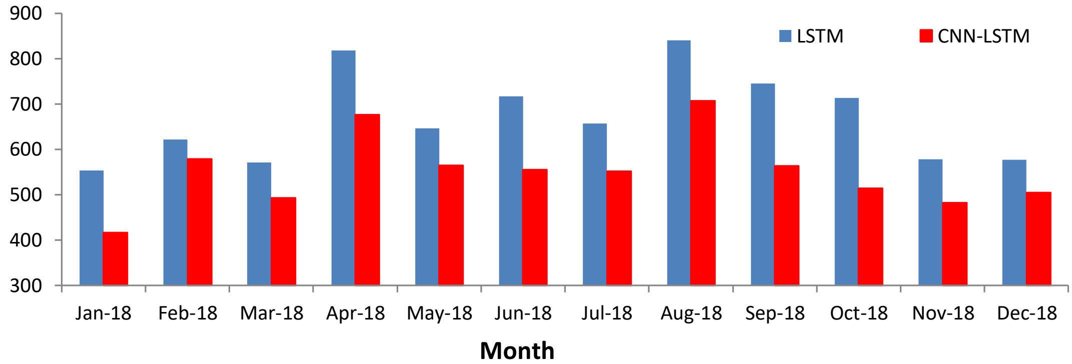

Fig. 5.78. Monthly RMSE for 2018

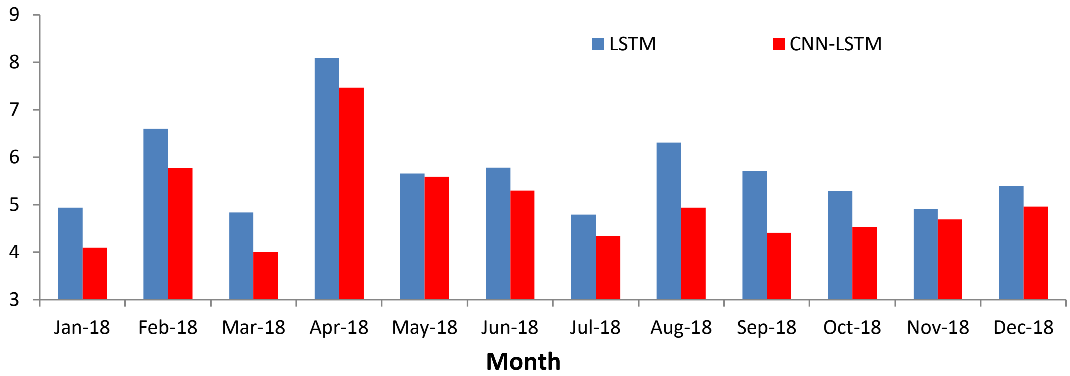

Fig. 5.79. Monthly MAPE for 2018

5.5.2 Comparison of evaluation metrics for 07 days for the year of 2018

The weekly comparison of MAE, RMSE, MAPE between the proposed CNN-LSTM and LSTM network is described in Table 5.3. For example, with the forecasted data in 01-07 Jan 2018, the LSTM network provides an RMSE of 612.70 while the proposed CNN-LSTM offers a RMSE of 549.94 Similarly, the proposed method offers less RMSE with the data forecasted data in 22-28 Dec 2018. The similar trend is also evident from the data available in other weeks. In average, the proposed method provides 54.66, 99.12 and 0.93% less MAE, MAPE and RMSE than the LSTM network respectively. The improvement of all metrics is also graphically illustrated in the Fig. 5.80 to Fig. 5.82.

Table 5.3: Weekly MAE, RMSE and MAPE for LSTM and CNN-LSTM

Observation Period	MAE		RMSE		MAPE	
	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM
01-07 Jan 18	393.83	347.11	612.70	549.94	5.81	5.12
01-07 Feb 18	396.44	368.74	679.23	538.87	6.44	5.64
01-07 Mar18	405.23	391.59	647.73	605.66	5.57	5.46
07-14 Apr 18	262.83	253.22	416.95	359.29	4.64	4.08
01-07May 18	451.88	401.30	667.95	615.01	6.68	5.43
08-14 Jun 18	443.62	417.72	773.26	667.27	6.09	5.41
15-22 July 18	483.56	376.51	699.04	531.03	5.26	3.91
23-29 Aug 18	460.19	393.01	712.63	566.49	6.15	5.02
01-07 Sep 18	390.63	335.34	547.63	450.38	5.43	4.59
08-14Oct 18	340.48	311.50	516.18	490.41	4.51	4.08
15-21Nov 18	399.76	296.51	580.65	453.58	6.55	4.96
22-28 Dec 18	363.16	243.11	496.43	333.01	5.80	4.04
Average	399.3	344.64	612.53	513.41	5.74	4.81

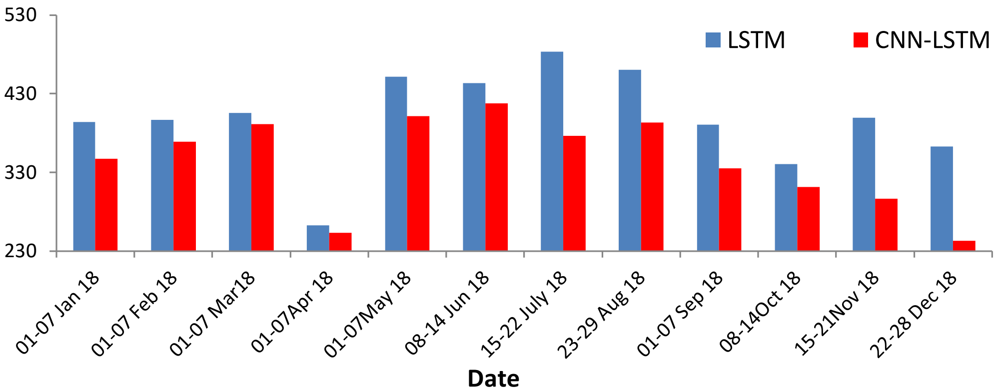

Fig. 5.80. Weekly MAE for 2018

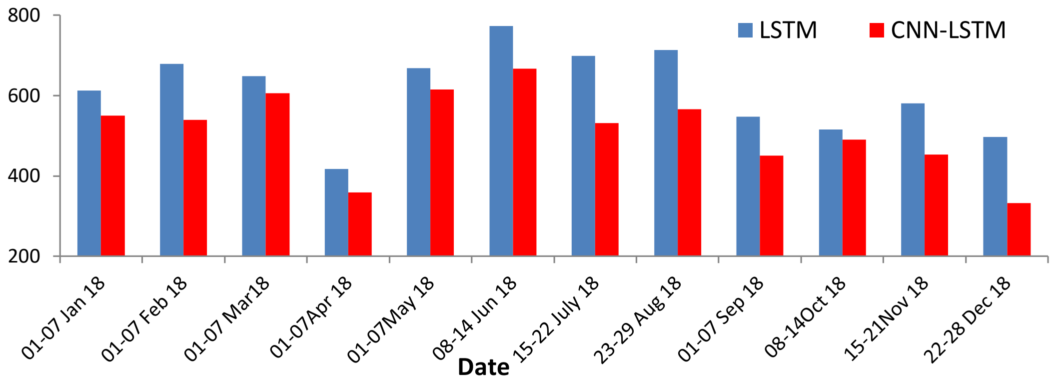

Fig. 5.81. Weekly RMSE for 2018

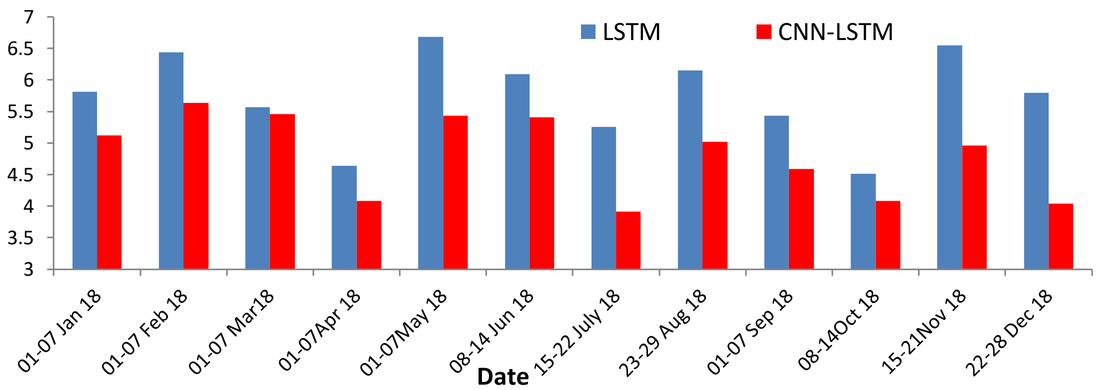

Fig 5.82. Weekly MAPE for 2018

5.5.3 Comparison of evaluation metrics for 48 hours for the year of 2018

Table 5.4 demonstrate the selected 48 hours comparison of MAE, RMSE, MAPE for the proposed CNN-LSTM and LSTM network respectively. For example, with the forecasted data in 01-07 Jan 2018, the LSTM network provides an MAPE of 7.07 while the proposed CNN-LSTM offers a RMSE of 5.42. Similarly, the proposed method offers less MAPE with the data forecasted data in 06-07 May 2018. The similar trend is also evident from the data available in other a couple of days. In average, the proposed method provides 106.81, 180.5 and 1.52% less MAE, MAPE and RMSE than the LSTM network respectively. The graphical improvement of all metrics is also displayed in the Fig. 5.83 to Fig 5.85.

Table 5.4: Weekly MAE, RMSE and MAPE for LSTM and CNN-LSTM

	MAE		RMSE		MAPE	
Observation Period	**LSTM**	**Proposed CNN-LSTM**	**LSTM**	**Proposed CNN-LSTM**	**LSTM**	**Proposed CNN-LSTM**
15-16 Jan 18	407.74	308.53	551.53	391.58	7.07	5.42
11-12 Mar 18	387.04	341.22	705.26	581.14	5.12	4.71
06-07 May 18	702.71	493.87	1052.01	706.32	8.78	5.64
07-08 July 18	446.11	354.88	663.68	457.28	4.87	3.87
04-05 Sep18	521.17	432.16	709.73	602.57	6.86	5.65
20-21 Nov 18	368.46	261.66	510.63	370.96	6.12	4.41
Average	472.21	365.39	698.81	518.31	6.47	4.95

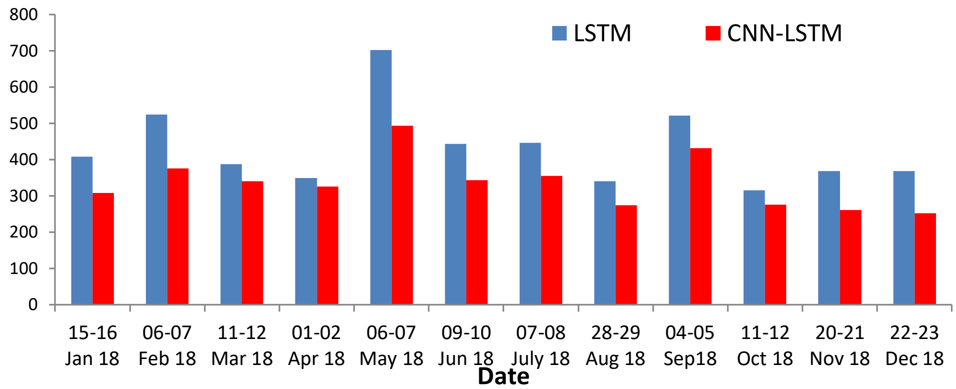

Fig 5.83. MAE of 48 hours for the different dates of 2018

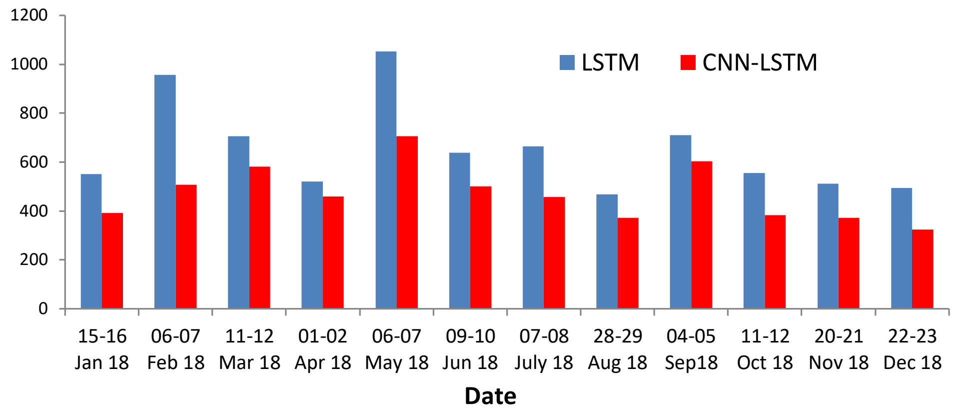

Fig 5.84. RMSE of 48 hours for the different dates of 2018

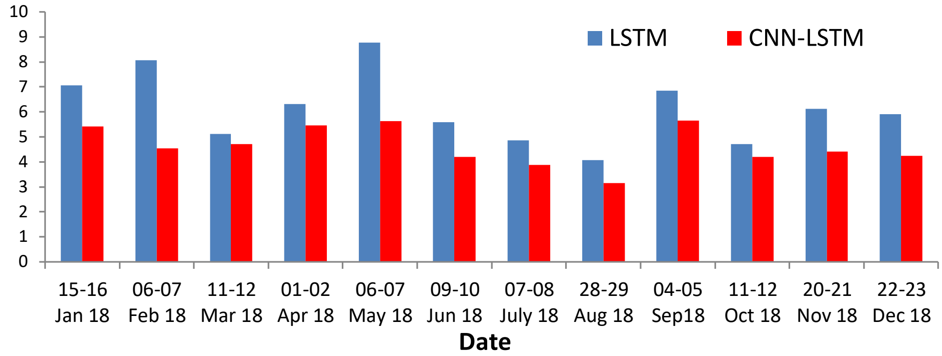

Fig 5.85. MAPE of 48 hours for the different dates of 2018

5.5.4 Comparison of evaluation metrics for 24 hours for the year of 2018

Daily comparison of MAE, RMSE, MAPE for the proposed CNN-LSTM and LSTM network are shown in Table 5.5. To give an example, with the forecasted data in 01 Jan 2018, the LSTM network provides an MAE of 305.75 while the proposed CNN-LSTM offers a RMSE of 243.43. Similarly, the proposed method offers less MAE with the data forecasted data in 18 Nov 2018. The similar trend is also evident from the data available in other day of 24 hours. In average, the proposed method provides 107.48, 195.88 and 1.79% less MAE, RMSE and MAPE than the LSTM network respectively. The graphical improvement of all metrics is also displayed in the Fig. 5.86 to Fig. 5.88.

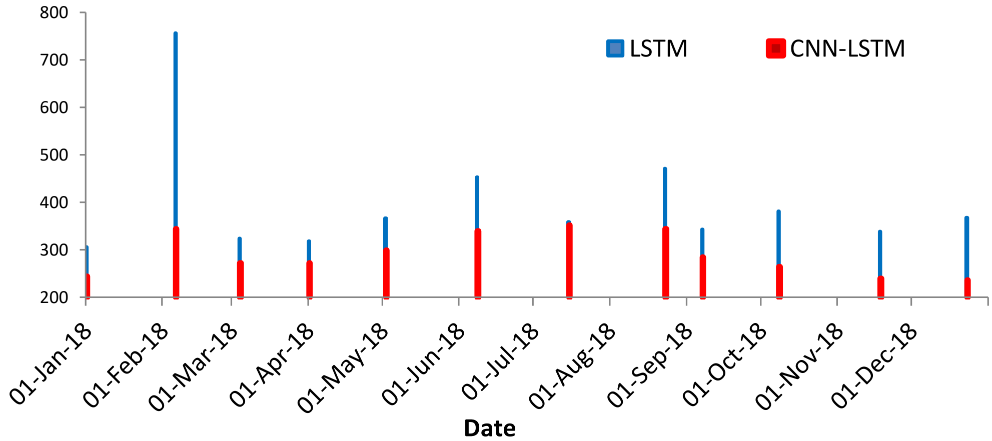

Fig 5.86. MAE of 24 hours for the different dates of 2018

Table 5.5: Weekly MAE for LSTM and CNN-LSTM

Observation Period	MAE		RMSE		MAPE	
	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM
01 Jan 18	305.75	243.46	497.91	356.34	5.27	4.12
06 Feb 18	755.71	344.21	1287.22	460.56	12.68	4.42
04 Mar 18	324.18	272.24	451.26	403.67	4.25	3.61
01 Apr 18	317.56	272.30	428.13	347.16	5.64	4.38
02 May 18	366.22	299.24	464.66	410.13	6.39	4.82
08 Jun 18	452.67	339.19	857.19	442.01	5.64	4.48
15 July 18	358.50	351.99	496.76	426.26	3.73	3.62
23 Aug 18	471.06	343.89	728.52	533.58	6.67	4.64
07 Sep 18	342.25	284.72	517.38	409.23	3.77	3.24
08 Oct 18	380.64	263.62	491.98	354.28	4.26	2.93
18 Nov 18	338.80	239.95	521.23	399.50	4.93	3.54
23 Dec 18	367.18	236.03	465.97	314.83	5.89	3.89
Average	398.38	290.9	600.68	404.8	5.76	3.97

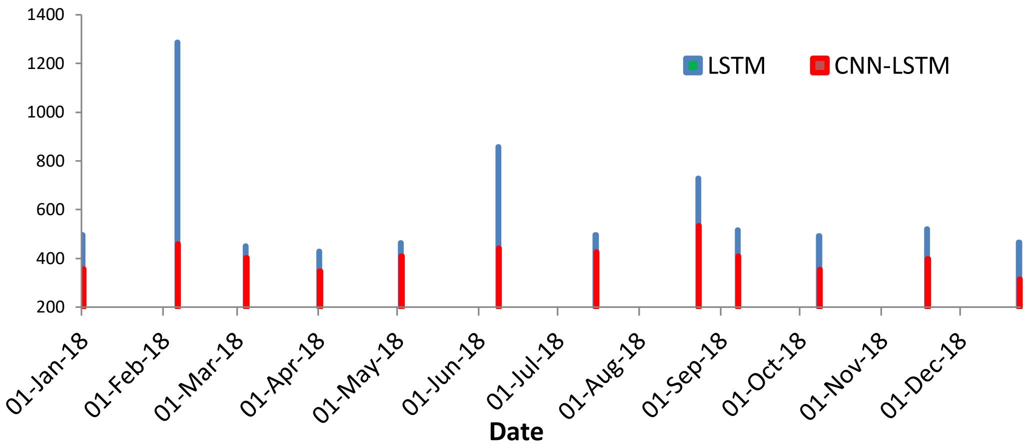

Fig. 5.87. RMSE of 24 hours for the different dates of 2018

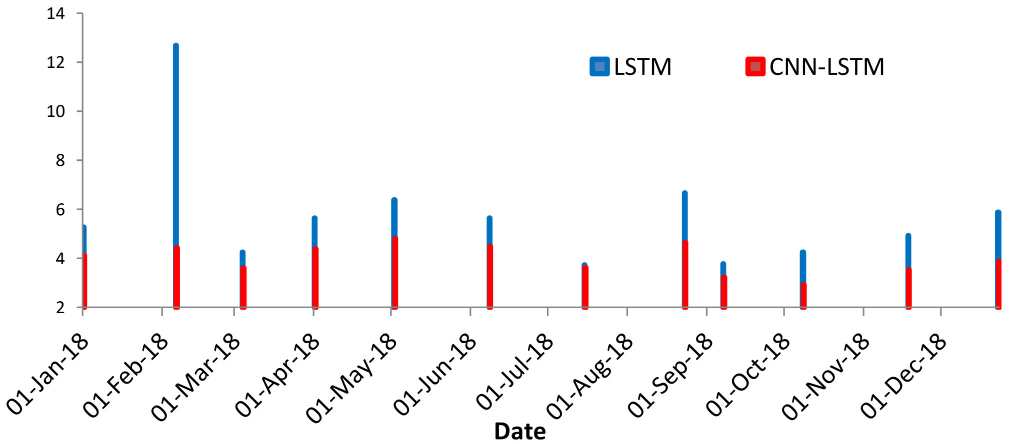

Fig. 5.88. MAPE of 24 hours for the different dates of 2018.

5.5.5 Comparison of evaluation metrics for 30 days for the year of 2019

The performance of the proposed CNN-LSTM network is compared with the LSTM network in terms of monthly MAE, RMSE, and MAPE in Table 5.6. For instance, with the forecasted data in Jan 2019, the LSTM network provides a MAE of 486.49, while the proposed CNN-LSTM offers a MAE of 304.97. For the month of Feb 2019, LSTM offers a MAE of. 889.05, where MAE obtained from CNN-LSTM is 370.18. LSTM Similarly, the proposed method offers less MAE with the data forecasted data in October 2019. The similar trend is also evident from the data available in other months. In average, the proposed method provides 173.76, 330.2, 3.07% less MAE, RMSE and MAPE respectively than the LSTM network. The improvement of all metrics is also graphically illustrated in the Fig. 5.89 to Fig.5.91.

Table 5.6: Monthly MAE, RMSE and MAPE for LSTM and CNN-LSTM

Observation Period	MAE		RMSE		MAPE	
	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM
Jan 19	486.49	304.97	577.94	429.79	5.71	4.23
Feb 19	889.05	370.18	1260.03	538.77	12.81	5.34
Mar 19	734.98	381.50	962.59	573.09	9.76	4.73
Apr 19	664.96	402.77	1036.38	610.37	8.70	4.88
May 19	692.62	471.63	1110.09	733.13	8.43	5.47
Jun 19	649.99	426.56	1122.29	736.35	8.05	4.94
July 19	559.02	489.24	859.10	645.65	6.58	5.65
Aug 19	621.80	437.53	939.49	668.87	7.81	5.31
Sep 19	614.79	438.94	929.80	669.94	7.71	5.35
Oct 19	476.99	291.16	713.112	406.45	6.09	3.49
Nov 19	474.16	298.62	821.09	464.95	7.13	3.93
Dec 19	430.14	335.55	606.02	498.14	6.13	4.72
Average	561.15	387.39	911.49	581.29	7.91	4.84

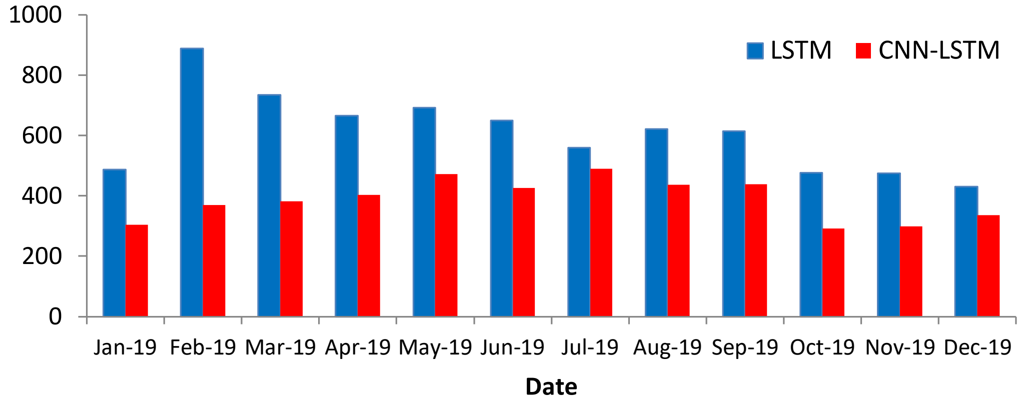

Fig. 5.89. Monthly MAE for 2019

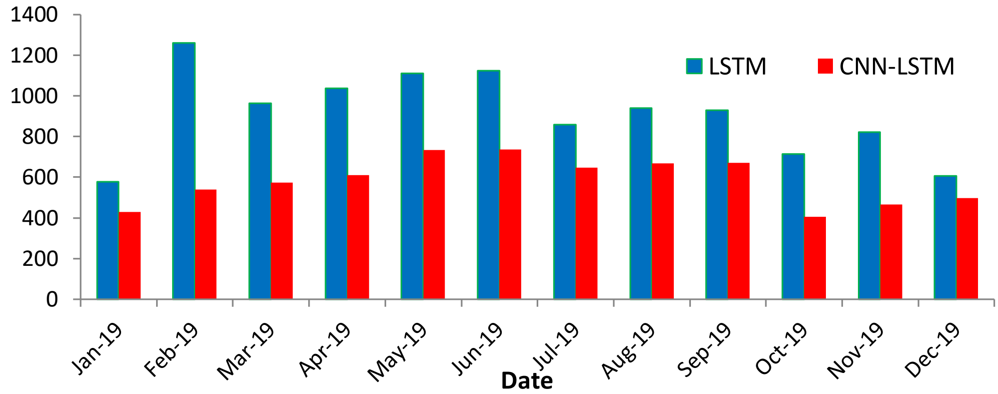

Fig. 5.90. Monthly RMSE for 2019

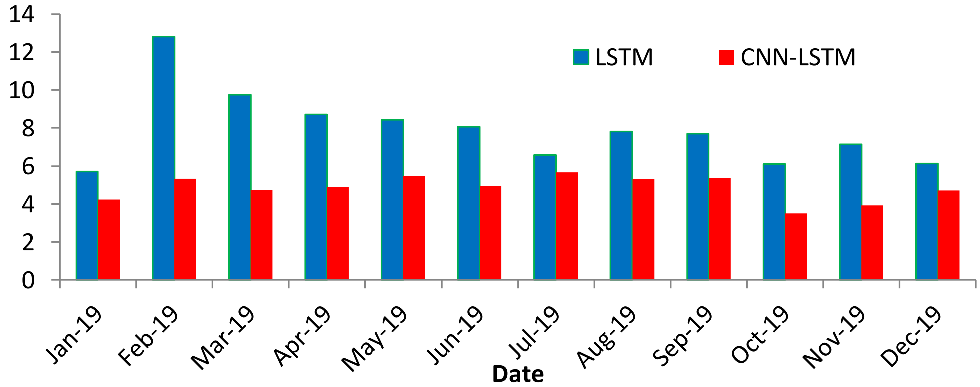

Fig. 5.91. Monthly MAPE for 2019

5.5.6 Comparison of evaluation metrics for 07 days for the year of 2019

The weekly comparison of MAE, RMSE, MAPE between the proposed CNN-LSTM and LSTM network is described in Table 5.7. For example, with the forecasted data in 01-07 Jan 2019, the LSTM network provides an RMSE of 803.53 while the proposed CNN-LSTM offers a RMSE of 331.91. Similarly, the proposed method offers less RMSE with the data forecasted data in 08-14 October 2019. The similar trend is also evident from the data available in other weeks. In average, the proposed method provides 271, 400.77, 3.63% less MAE, RMSE and MAPE than the LSTM network. The improvement of all metrics is also graphically illustrated in the Fig.5.92 to Fig 5.94.

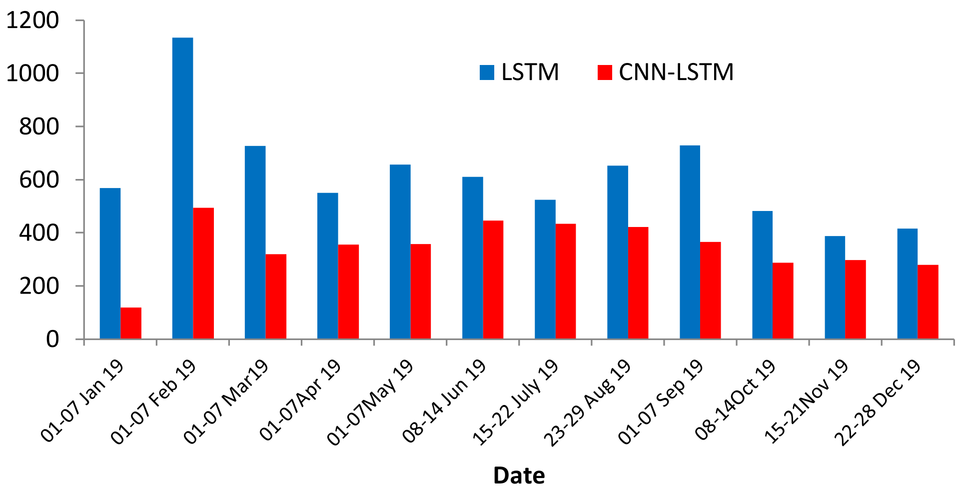

Fig. 5.92. Weekly MAE for 2019

Table 5.7: Weekly MAE, RMSE and MAPE for LSTM and CNN-LSTM

Observation Period	MAE		RMSE		MAPE	
	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM
01-07 Jan 19	567.99	119.03	803.53	331.91	8.99	6.65
01-07 Feb 19	1133.49	494.83	1552.89	718.33	16.64	6.86
01-07 Mar19	727.63	318.64	944.99	445.34	10.48	4.26
01-07 Apr 19	549.86	356.17	858.77	552.55	8.12	5.26
01-07 May 19	657.59	358.08	1220.78	591.73	8.82	4.88
08-14 Jun 19	609.90	446.49	1049.54	882.45	7.56	5.55
15-22 July 19	523.35	433.68	716.04	577.57	5.52	4.14
23-29 Aug 19	652.43	422.81	968.67	607.37	7.13	4.28
01-07 Sep 19	728.19	365.88	1155.09	548.76	9.92	4.92
08-14 Oct 19	482.92	286.49	774.84	394.61	6.36	3.42
15-21 Nov 19	386.92	297.45	690.64	470.34	6.08	4.18
22-28 Dec 19	416.52	279.59	594.52	400.14	6.64	4.38
Average	619.73	348.26	944.19	543.42	8.52	4.89

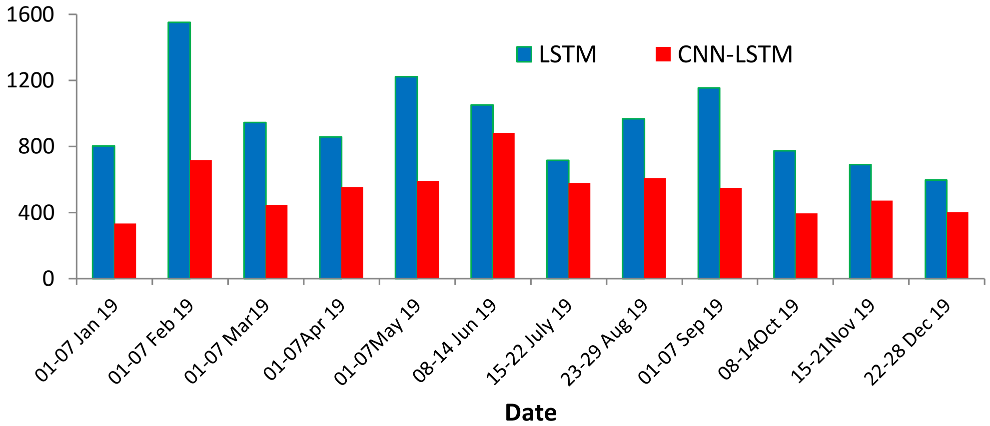

Fig. 5.93. Weekly RMSE for 2019

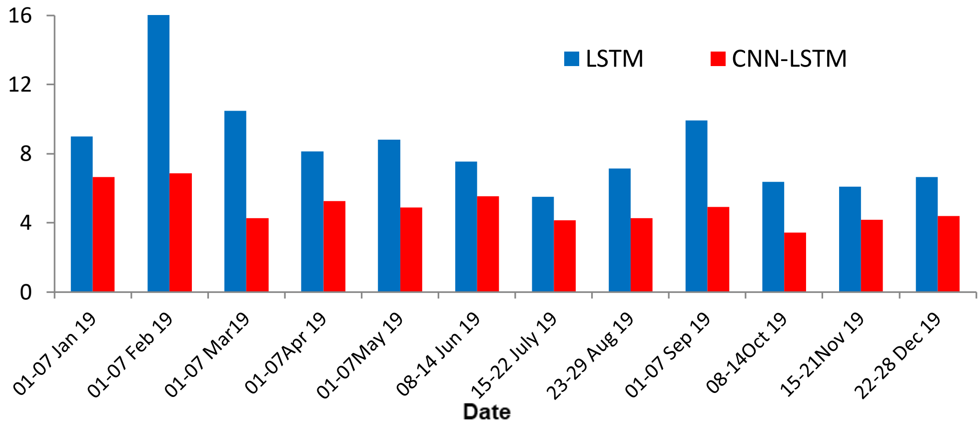

Fig. 5.94. Weekly MAPE for 2019

5.5.7 Comparison of evaluation matrices of 48 hours for the year of 2019

Table 5.8 demonstrates the selected 48 hours comparison of MAE, RMSE, and MAPE for the proposed CNN-LSTM and LSTM network respectively. For example, with the forecasted data in 15-16 Jan 2019, the LSTM network provides RMSE of 503.85 while the proposed CNN-LSTM offers a RMSE of 290.17. Similarly, the proposed method offers less MAPE with the data forecasted data in 04-05 September 2019. The similar trend is also evident from the data available in other a couple of days. In average, the proposed method provides 324.22, 529.96, 4.37 less MAE, MAPE and RMSE respectively than the LSTM network. The graphical improvement of all metrics is also displayed in the Fig 5.95 to Fig 5.97.

Table 5.8: Weekly MAE, RMSE and MAPE for LSTM and CNN-LSTM

Observation Period	MAE		RMSE		MAPE	
	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM
15-16 Jan 19	355.97	201.21	503.85	290.17	5.54	3.03
11-12 Mar 19	692.20	262.17	910.44	360.40	9.58	3.69
06-07 May 19	848.83	347.19	1493.41	619.53	11.63	4.77
07-08 July 19	618.83	439.01	976.78	575.28	6.19	4.53
04-05 Sep19	732.71	249.17	1112.42	341.84	9.15	2.86
20-21 Nov 19	386.50	190.95	606.73	236.63	5.72	2.71
Average	605.84	281.62	933.94	403.98	7.97	3.59

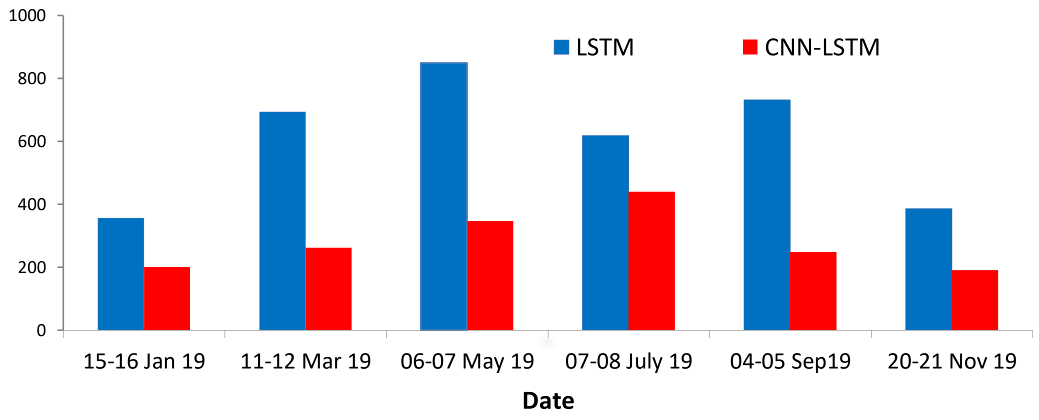

Fig. 5.95. MAE for 48 hours of the different month of 2019

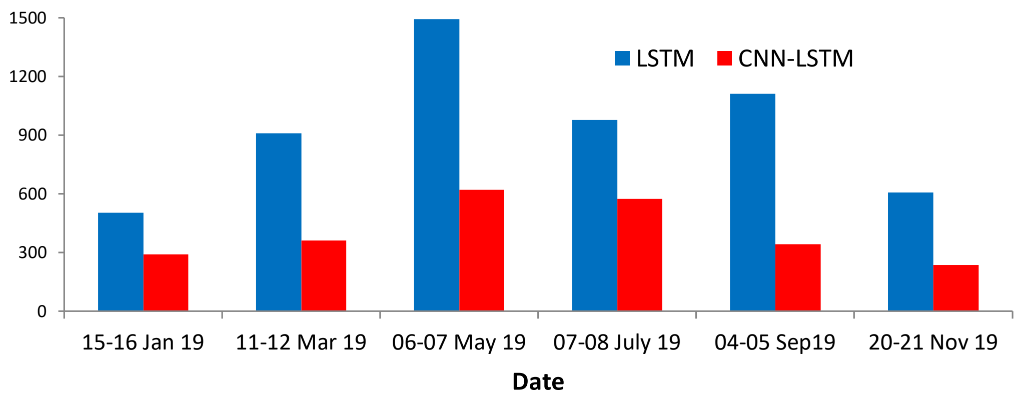

Fig. 5.96. RMSE for 48 hours of the different month of 2019

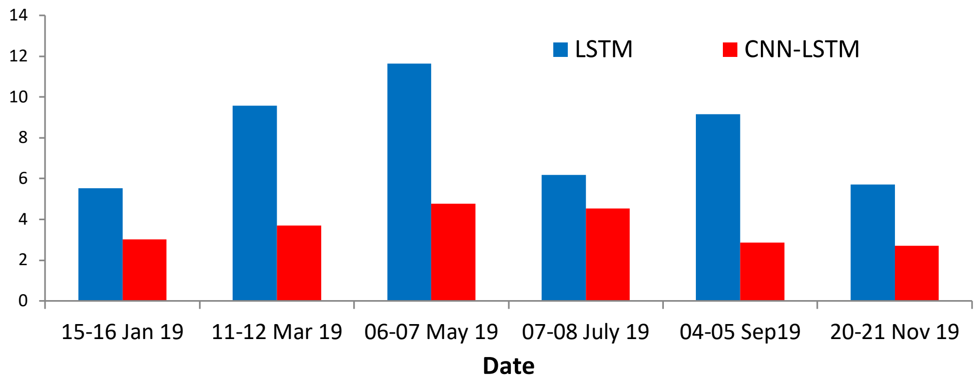

Fig. 5.97. MAPE for 48 hours of the different month of 2019

5.5.8 Comparison of evaluation matrices of 24 hours for the year of 2019

Daily comparison of MAE, RMSE, MAPE for the proposed CNN-LSTM and LSTM network are shown in Table 5.9. To give an example, with the forecasted data in 01 Jan 2019, the LSTM network provides an RMSE of 652.79 while the proposed CNN-LSTM offers a RMSE of 249.90. Similarly, the proposed method offers less RMSE with the data forecasted data in 07 September 2019. The similar trend is also evident from the data available in other day of 24 hours. In average, the proposed method provides 324.77, 434.49, 4.33 less MAE, RMSE and MAPE than the LSTM network respectively. The graphical improvement of all metrics is also displayed in the Fig. 5.98 to Fig 5.100.

Table 5.9: Weekly MAE for LSTM and CNN-LSTM

Observation Period	MAE		RMSE		MAPE	
	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM	LSTM	Proposed CNN-LSTM
01 Jan 19	493.71	211.92	652.79	249.90	8.82	3.65
06 Feb 19	642.76	239.56	865.10	347.95	6.67	2.31
02 Mar 19	793.11	230.04	918.56	286.09	12.72	3.38
01 Apr 19	418.05	136.18	585.16	200.50	7.75	2.56
02 May 19	407.96	150.39	502.86	206.25	5.82	5.21
11 Jun 19	403.92	175.74	579.69	210.82	4.99	2.19
15 July 19	775.62	192.31	954.05	258.19	8.72	2.31
23 Aug 19	775.62	265.20	954.05	394.45	8.72	2.99
07 Sep 19	878.13	353.07	1287.83	353.07	12.29	3.99
08 Oct 19	480.82	335.58	636.21	435.76	5.12	3.62
18 Nov 19	312.11	252.61	406.64	370.54	4.16	3.37
23 Dec 19	300.40	242.38	377.38	192.81	4.76	3.03
Average	556.85	232.08	726.69	292.19	7.55	3.22

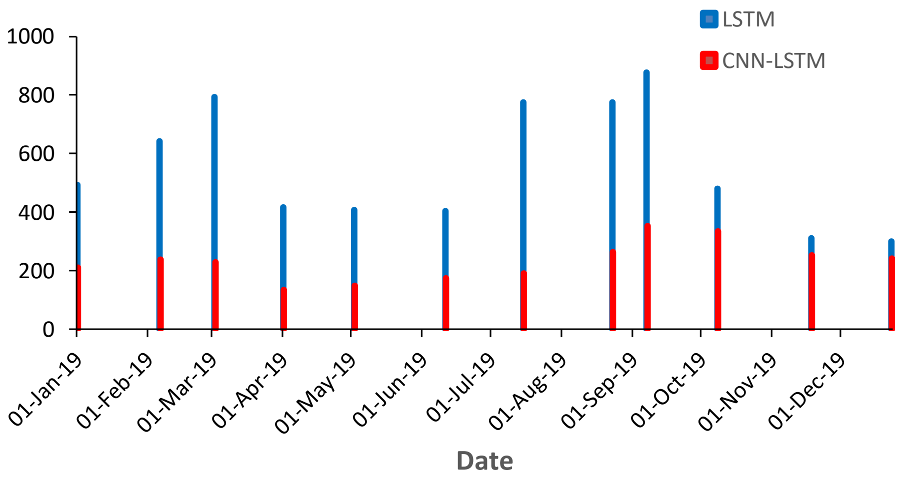

Fig. 5.98. MAE for 24 hours of different dates of 2019

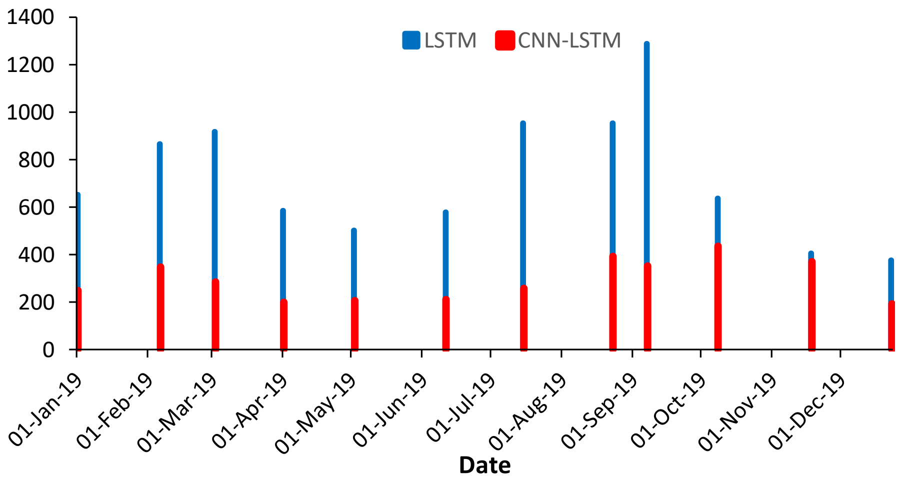

Fig. 5.99. RMSE for 24 hours of different dates of 2019

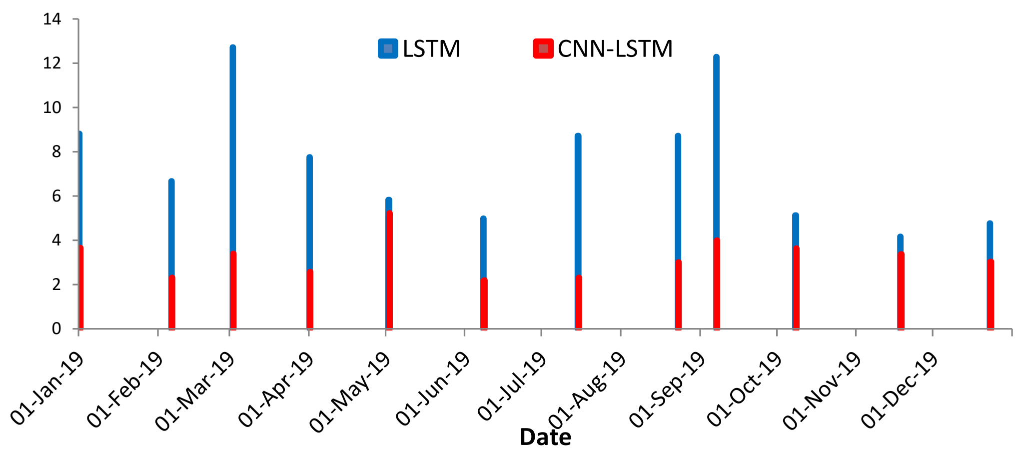

Fig. 5.100. MAPE for 24 hours of different dates of 2019

5.6 Summary

In this chapter, the performance of the proposed CNN-LSTM based load forecasting methodology us evaluated with real case historical load data recorded from 2014 to 2019 in Bangladesh power system network. The performance evaluation matrices are considered as MAE, MAPE and RMSE. This proposed scheme offered low value of MAE, MAPE and RMSE as compared to the others conventional forecasting model like LSTM method. Based on the observations presented in this chapter, concluding remarks and future works of the research are presented in the following chapter.

CHAPTER 6
CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

This research work proposes a noble approach of CNN-LSTM model for STLF. LSTM and CNN are both outlined to provide high precision forecast in STLF due to their advantages to capture hidden features. In this way, it has been developed a hybrid neural network CNN-LSTM framework that can capture and facilitated such various hidden features to provide better execution. CNN-LSTM system basically comprises of a Convolution Neural Network (CNN) module, a Long Short Term Memory (LSTM) module and a feature-fusion module. The original electric load data from 2014 to 2019 is obtained from Power Grid Company Bangladesh. Within the data preparation step, null values are being checked and the load data is being part into training and test sets. At that point, the original dataset is transferred into two different datasets. Firstly, it is considered training set from 2014 to 2017 and test set is chosen from 2018. Secondly, training set is selected from 2014 to 2018 and test set is chosen from 2019. The CNN module is utilized to capture the local trend and the LSTM module is utilized to learn the long-term dependency. The two hidden feature is being concatenated within the feature-fusion module. The final forecast is generally created after a completely connected layer.

The performance calculation of the developed model was evaluated by exploring the electrical load forecast of Bangladesh. To verify the networks stability and effectiveness, a various segment of the data set was executed properly. The proposed model with everyday training dataset obtains better forecasting performance compared to LSTM model for every month in 2018 (test set was taken 2018) and 2019 (when test set was taken 2019). Furthermore, proposed model also performs better for weekdays, weekends, and holidays.

This empirical result shows that CNN-LSTM provides a promising model for electricity load forecasting in electric power system network.

In addition, proposed CNN-LSTM model gives the minimum value of MAE, RMSE and MAPE as compared to LSTM model. In every validation test predicted load values using CNN-LSTM model outperforms LSTM model all the times. Moreover, it also gives higher accuracy for March and July where they normally have lowest accuracy comparing to other months of the year. It also provides better accuracy for months with minimum temperature variation. In view of short term forecast model the output of proposed model is acceptable since mean absolute percentage error (MAPE) lies within 7%. Towards this end, it can be illustrated that proposed CNN-LSTM model can deal with the long sequence time series electric load data and predict the future load demand over a long period.

6.2 Future Work

The outcomes obtained from the proposed model indicate a future work where highly precise load forecasting framework can be developed by using CNN network integrated with a Gated Recurrent Unit (GRU). Another possible test what would be given satisfactory load forecasting outcomes by using convolutional long short term memory network (Conv-LSTM) which is a variant of LSTM containing a convolution operation inside the LSTM cell. It basically replaces matrix multiplication with convolution operation at each gate in the LSTM cell. By doing so, it captures underlying spatial features by convolution operations in multiple-dimensional data.

REFERENCES

[1] R. M. Qamar and K. Abbas, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews.*, vol. 50, pp. 1352-1372, 2015.

[2] A.K. Sing, Ibraheem, S. Khatoon, M. Muazzam, and D. K. Chaturvedi, "Load forecasting techniques and methodologies: A review," *2nd International Conference on Power*, India, 2012.

[3] M. Joshia and R. Singh, "Short-term load forecasting approaches: A review," *International Journal of Recent Engineering Research and Development (IJRERD),* vol. 1, no. 3, pp. 9-17.

[4] Girraj Singh, D. S. Chauhan, Aseem Chandel, Deepak Parashar, and Girijapati Sharma, "Factor affecting elements and short term load forecasting based on multiple linear regression method," *International Journal of Engineering Research & Technology,* vol. 03, no. 12, 2014.

[5] M. U. Fahad and N. Arnab,, "Factors affecting short term load forecasting," *Journal of Clean Energy Technologies,* vol. 2, no. 4, 2014.

[6] M. A. Mahmud, "Isolated area load forecasting using linear regression analysis: practical approach," *Energy and Power Engineering,* vol. 3, no. 4, pp. 547-550, 2011.

[7] A. Islam, S. R. Hasib, and M. S. Islam, "Short term electricity demand forecasting for an isolated area using two different approaches," *Journal of Power Technologies,* vol. 93, no. 4, pp. 96-101, 2013.

[8] N.A. Masood, M. Z. Sadi, S. R. Deeba and R. H. Siddique, "Temperature sensitivity forecasting of electrical load," *4th International Power Engineering and Optimization Conference (PEOCO),* Malaysia, 2010.

[9] H. K.Temraz, M. M. A.Salama, and A. Y. Chikhani, "Review of electric load forecasting methods," *IEEE Canadian Conference on Electrical and Computer Engineering*, Canada, May 1997.

[10] P. P. PHYO, "Deep learning for short-term electricity," *Master of Science Thesis, Sirindhorn International Institute of Engineering and Technology,* Mayanmar, 2017.

[11] A. Karsaz, H.R. Mashhadi, and M.M. Mirsalehi, "Market clearing price and load forecasting using cooperative co-evolutionary approach," *Int. J. of El. Power & Energy*

Systems, vol. 32, no. 5, pp. 408-415, Jun 2010.

[12] W. Charytoniuk, M. S. Chen and P. V. Olinda, "Nonparametric regression based short-term load forecasting," *IEEE Trans. Power Syst.,* vol. 13, pp. 725-730, 1998.

[13] K. B. Song, Y. S. Baek, D. H. Hong and G. Jang, "Short-term load forecasting for the holidays using fuzzy linear regression method," *IEEE Trans. Power Syst.,* pp. 96-101, 2005.

[14] J. Kim, S. Cho, K. Ko and R. R. Rao,, "Short-term electric load prediction using multiple linear regression method," *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grid*, Aalborg, 2018.

[15] W. R. Christiaanse, "Short-term load forecasting using general exponential smoothing," *IEEE Trans. Power Appl. Syst,* vol. 2, pp. 900-911, 1971.

[16] C. M. Lee and C.N. Ko, "Short-term load forecasting using lifting scheme and ARIMA models," *Expert Syst. Appl.,* vol. 38, pp. 5902-5911, 2011.

[17] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research,* vol. 30, p. 79, Dec 2005.

[18] T. Chai and R.R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE) Arguments against avoiding RMSE in the literature," *Geoscientific Model Development,* vol. 7, pp. 1247-1250, Feb 2014.

[19] H. H Çevik and M. Çunka¸, "“Short-term load forecasting using fuzzy logic and ANFIS," *Neural Comput. Appl.,* vol. 26, pp. 1355-1367, 2015.

[20] M. Rejc and M. Pantos, "Short-term transmission-loss forecast for the Slovenian transmission power system based on a fuzzy-logic decision approach," *Neural Comput. Appl.,* vol. 26, pp. 1511-1521, 2011.

[21] Y. Chen, P. Xu, Y. Chu, W. Li, Y. Wu, L. Ni, Y. Bao, and K. Wang, "Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office building," *Appl. Energy,* vol. 195, pp. 659-670, 2017.

[22] P. H. Kuo and C. J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies, vol. 11,* p. 213, 2018.

[23] B. S. Bisht and R. M. Holmukhe, "Electricity load forecasting by artificial neural network model using weather data," *IJEET Trans. Power Syst.,*, vol. 4, no. 1, pp. 91-99, Jan 2013.

[24] P. Ray, D. P. Mishra, and R. K. Lenka, "Short term load forecasting by artificial neural network," *International Conference on Next Generation Intelligent systems (ICNGIS)*, India, 2016.

[25] K. B. Sahay and M. M. Triapthi, "Day ahead hourly load and price forecast in ISO new England market by using ANN," *IEEE INDICON*, India, 2013.

[26] Islam and D. Badar, "Comparison of conventional and modern load forecasting techniques based on artificial intelligence and expert systems," *IJCSI International Journal of Computer Science*, vol. 8, p. 1694, Sep 2011.

[27] M. Courbariaux, I. Hubara, D. Soudry, R. El-Yaniv, and Y. Bengio, "M. Courbariaux, I. Hubara, D. Soudry, R. El-Yaniv, and Y. Bengio, “Binarized Neural Networks: Training Deep Neural Networks with Weights and Activations Constrained to +1 or -1.," ArXiv e-prints , Feb. 2016.

[28] A.K Pandey, K. B. Sahay, M. M Tripathi, and D. Chandra, "Short-Term Load Forecasting of UPPCL using ANN," *6th IEEE Power India International Conference (PIICON)*, India, 2014.

[29] M. A. Yahya, S. P Hadi, and L. M. Putranto, "Short-term electric load forecasting using recurrent neural network," *4th International Conference on Science and Technology*, Yogyakarta, Indonesia, 2018.

[30] S. Fan and L. Chen, "Short-term load forecasting based on an adaptive hybrid method," *IEEE Trans. Power Syst*, vol. 21, no. 1, pp. 392-401, 2006.

[31] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation," *Int. J. Elect. Power Energy Syst,* vol. 31, no. 1, pp. 13-22, 2009.

[32] L. Ekonomou, C. A. Christodoulou, and V. A. Mladenov, "Short-term load forecasting method using artificialneural networks and wavelet analysis,," *Int. J. Power Syst,* vol. 1, pp. 64-68, 2016.

[33] G. Merkel, R. Povinelli, and R. Brown, "Short-term load forecasting of natural gas with deep neural network regression," *Energies,* vol. 11, p. 2008, 2018.

[34] S. Bouktif, A. Fiaz, A. Ouni and M. Serhani, "Optimal Deep Learning LSTM Model for

Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches.," *Energies,* vol. 11, p. 1636, 2018.

[35] H. Zheng, J. Yuan, and L. Chen, "Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation," *Energies,* vol. 10, p. 1168, 2017.

[36] H. Shi, M. Xu, and R. Li, "Deep learning for household load forecasting—a novel pooling deep RNN," *IEEE Trans. Smart Grid,* vol. 9, pp. 5271-5280, 2018.

[37] L. Y. Wei, C. H. Tsai, Y. C. Chung, K. H. Liao, H. E. Chueh, and J. S. Lin, "A study of the hybrid recurrent neural network model for electricity loads forecasting," *Int. J. Acad. Res. Account. Financ. Manag. Sci,* vol. 7, pp. 21-29, 2017.

[38] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks," *IEEE International Conference on Systems, Hungary,* Oct 2016.

[39] L. Kuan, Z. Yan, W. Xin, C. Yan and P. Xiangkun, S. Wenxue, J. Zhe, Z. Yong, X. Nan and Z. Xin, "Short-term electricity load forecasting method based on multilayered self-normalizing GRU network," *IEEE Conference on Energy Internet and Energy System Integration,* China, 2017.

[40] S. Muzaffar and A. Afshari, "Short-term load forecasts using LSTM networks," *10th International Conference on Applied Energy,* China, 2018.

[41] Y. Wang, M. Liu, Z. Bao, and S. Zhang, "Short-term load forecasting with multi-source data using gated recurrent unit neural Networks," *Energies,* vol. 11, p. 1138, 2018.

[42] S. Du, T. Li, X. Gong, Y. Yang and S. J. Horng, "Traffic flow forecasting based on hybrid deep learning framework," *12th International Conference on Intelligent Systems and Knowledge Engineering,* China, 2017.

[43] P. H. Kuo and C. J. Huang, "A high precision artificial neural networks model forsShort-term energy load forecasting," *Energies,* vol. 11, p. 213, 2018.

[44] L. Han, Y. Peng, Y. Li, B. Yong, Q. Zhou, and L. Shu, "Enhanced deep networks for short-term and medium-term load forecasting," *IEEE Access,* vol. 7, pp. 4045-4055, Dec 2018.

[45] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu and Z. Zhu, "Multi-Scale Convolutional Neural Network With Time-Cognition for Multi-Step Short-Term Load Forecasting," *IEEE Access,* vol. 7, pp. 88058-88071, Jun 2019.

[46] M. J. Silbermann and H.D. Tagare, "Curves and surfaces in computer vision and graphics II," *Society of Photo Optical*, 1992.

[47] Y. Ito, "Representation of functions by super positions of a step or sigmoid function and their applications to neural network theory," *Neural Networks*, vol. 4, Jan 1991.