Multi-Step Load Demand Forecasting Using Neural Network

Sonu Kumar Jha, Chaman Lal Dewangan, *Student Member, IEEE*, Nishchal K. Verma, *Senior Member, IEEE*
Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, UP, India, 208016
Email: {ksonu, chaman, nishchal}@iitk.ac.in

***Abstract*—The accuracy of load demand forecasting plays a vital role in economic operation and planning in the power sector. Therefore, many techniques and approaches have been proposed in the literature for forecasting. However, there is still an essential need to develop more accurate load forecast method. In this paper, three different strategies of Multi-Step-Ahead Load Forecasting (MSALF), i.e. Direct Strategy (DS), Recursive Strategy (RS) and DirRec Strategy (Direct-Recursive Strategy or DRS) have been used for electricity load demand forecasting by using the Artificial Neural Network (ANN) with Levenberg-Marquardt (LM) training algorithm. The performance evaluation for three different strategies of MSALF has been analysed on two different substations of NE-ISO data sets. Each data sets is analysed for four different cases. The performance of the DRS is better than DS and RS.**

Index Terms*: *Artificial Neural Network, Direct Strategy, DirRec Strategy, Levenberg Marquardt, Load demand, Multi-step-ahead load forecasting, Recursive Strategy.

I. Introduction

The prediction of load demand is beneficial in the power system for the minimization of the operational cost, power management sector, and power security [1]. Therefore, accurate load demand forecasting model need to be prepared for the proper power system operations and plannings.

Load forecasting can be done for different time horizons such as year, month, week, day, hour, and minute depending on its applications. Load forecasting for less than a week is called short-term load forecasting. The single-step-ahead load forecasting is used for dispatching on-load tap changer transformer, the available transmission capacity of transmission lines. The multi-step-ahead are employed for the intraday electricity market, electrical grid planning.

Different load forecasting models are available in the literature [2], [3]. Load forecasting can be modelled by using statistical techniques, soft computing techniques and a hybrid of soft computing. The soft computing methods such as ANN, support vector regression (SVR), and fuzzy logic are found better in approximating the non-linear relations between the inputs and output of load forecasting model [4], [5].

The inputs of load demand forecasting are previously recorded load demand, forecasted ambient temperature, holidays, and day of the week. There are two types of strategies to predict the load demand for time horizons more than one-hour is single and multi-step strategy.

In [6] multi-step flood forecasting was implemented using neural network for three hours ahead prediction. The multi-input single-output and serial-propagated were found more accurate than multi-input multi-output. The different multi-step-ahead river flow forecasting methods such as direct and recursive method were compared [7]. Recursive methods were modelled as parallel and series-parallel method. In comparison, the direct method was outperformed the compared methods for the long time horizon (12-months) forecasts. Similarly, multi-step forecasting strategies were applied for stock price forecasting [8]. The performance of the direct recursive method was found the most accurate. More details of the multi-step forecasting are available in [9], [10], and [11]. From the literature review, it can be noted that different methods perform inconsistently for different data sets.

In this work, different multi-step forecasting strategies have been implemented for load demand forecasting for time horizons from one hour to five hour-ahead for the two data sets of New England independent system operator (NE-ISO). ANN with LM training algorithm is used for preparing the forecasting models. Before applying ANN with LM training algorithm for MSALF, the data sets have been analysed for selection of most effective input features and size of training data sets. In the analysis, DirRec based MSALF outperformed the compared strategies.

The organisation of the paper is as follows: Section II explains a brief overview of the ANN architecture and training algorithms. Section III describes the selection of input features. Section IV describes three different multi-step ahead load forecasting strategies and criteria to compare the performance of the forecasting models. Section V details of data sets, size of training data, and analysis of results. Section VI gives the conclusion of the work.

II. Artificial Neural Network (ANN)

ANN has played a major role in the enhancement of forecasting accuracy. ANN are mathematical tools originally inspired by the way the human brain processes information. ANN have generally three layers are input, hidden, and output. The hidden layer is between input and output layer [12]. ANN approximates the non-linear function between inputs and output for preparing forecasting models. The effective

978-1-7281-3192-4/19/$31.00 © 2019 IEEE

TABLE I
DATA-SET SELECTION ANALYSIS

Set of Data	Case-I(3-Months)		Case-II(4-Months)		Case-III(6-Months)		Case-IV(8-Months)		Case-V(10-Months)		Case-VI(12-Months)	
Month	**MAE(%)**	**RMSE(%)**	**MAE(%)**	**RMSE(%)**	**MAE(%)**	**RMSE(%)**	**MAE(%)**	**RMSE(%)**	**MAE(%)**	**RMSE(%)**	**MAE(%)**	**RMSE(%)**
Jan	0.797	1.117	0.925	1.276	0.785	1.047	1.068	1.479	0.790	1.085	0.642	0.877
Feb	0.629	0.872	0.634	0.848	0.649	0.877	0.692	0.947	0.751	0.996	0.661	0.908
Mar	0.446	0.586	0.459	0.608	0.488	0.672	0.522	0.673	0.463	0.605	0.466	0.610
Apr	0.625	0.895	0.547	0.731	0.732	0.915	0.576	0.742	0.621	0.847	0.598	0.781
May	0.625	0.790	0.476	0.669	0.492	0.703	0.525	0.709	0.568	0.723	0.523	0.679
Jun	0.670	1.098	0.672	1.040	0.583	0.887	0.728	1.167	0.613	0.983	0.637	1.049
Jul	1.085	1.528	1.456	2.128	1.082	1.512	1.126	1.480	0.930	1.350	0.875	1.168
Aug	0.526	0.693	0.547	0.686	0.486	0.633	0.481	0.623	0.543	0.696	0.488	0.625
Sep	0.813	1.087	0.789	1.038	0.819	1.070	0.755	1.015	0.800	1.035	0.843	1.116
Oct	0.526	0.704	0.515	0.684	0.478	0.648	0.519	0.678	0.501	0.685	0.516	0.689
Nov	0.844	1.179	0.706	1.035	0.726	1.023	0.688	1.044	0.586	0.874	0.579	0.879
Dec	0.763	1.007	0.792	1.168	0.781	1.128	0.800	1.161	0.674	0.983	0.652	0.893
Average	0.696	0.963	0.710	0.992	0.675	0.927	0.707	0.977	0.653	0.905	**0.623**	**0.856**

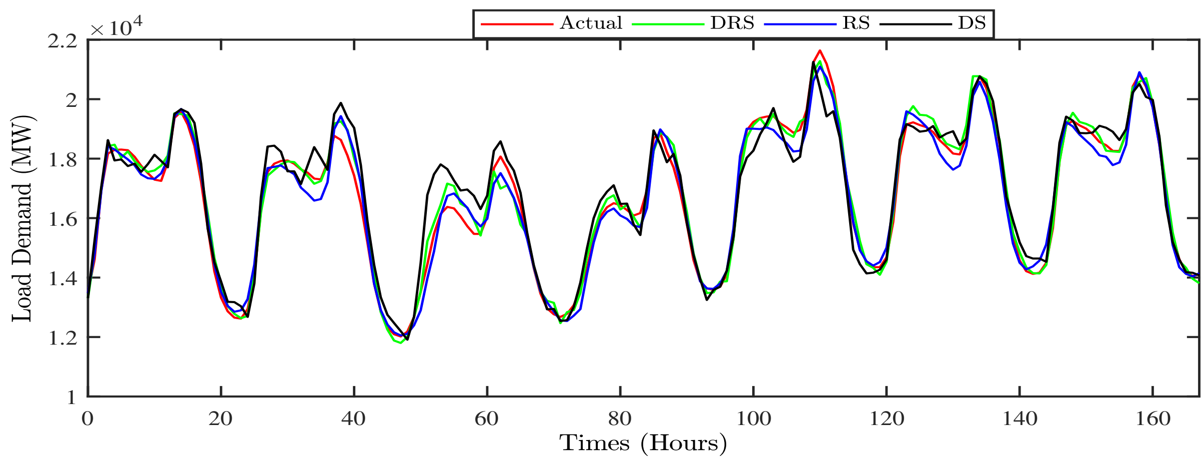

Fig. 1. The performance comparison graph for Case-I of NE-ISO (substation 1) data sets for 5^{th} hour MSALF

TABLE II
PERFORMANCE COMPARISON OF MULTI-STEP STRATEGIES OF NE-ISO SUBSTATION 1 (CASE-I)

Time	Direct Strategy		Recursive Strategy		DirRec Strategy	
Hours	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)
1	0.490	0.682	0.490	0.682	**0.490**	**0.682**
2	1.101	1.528	0.959	1.302	**0.949**	**1.332**
3	1.518	2.136	1.339	1.814	**1.269**	**1.706**
4	2.145	2.962	1.630	2.225	**1.477**	**2.017**
5	2.392	3.256	1.841	2.524	**1.574**	**2.120**

TABLE III
PERFORMANCE COMPARISON OF MULTI-STEP STRATEGIES OF NE-ISO SUBSTATION 1 (CASE-II)

Time	Direct Strategy		Recursive Strategy		DirRec Strategy	
Hours	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)
1	0.543	0.717	0.543	0.717	**0.543**	**0.717**
2	1.121	1.438	0.986	1.269	**0.938**	**1.205**
3	1.406	1.904	1.263	1.632	**1.173**	**1.498**
4	1.754	2.290	1.504	1.932	**1.352**	**1.725**
5	1.867	2.385	1.698	2.184	**1.465**	**1.885**

inputs and number of hidden neurons should be selected appropriately for the accurate prediction [13].

After the selection of neural network structure, training of a neural network is done to update the weights of the network such that it minimizes the prediction error. There are different types of training algorithm such as back-propagation training algorithm based on gradient descent, Levenberg-Marquardt, or particle swarm optimization. The whole data sets is divided into three parts training, validation, and testing. Validation step check the overfitting of the neural network. Finally, the performance of the prepared models is tested on test data sets [14].

Levenberg-Marquardt (LM) training algorithm was used to train the neural network since the LM algorithm is fast in convergence and gives accurate results [15], [16].

III. INPUT FEATURES SELECTION

ANN has been implemented for forecasting of load demand. Temperature and previously recorded load demand have been taken as inputs to the ANN model. To decide the effective lags of temperature and load demand, cross-correlation coefficients

TABLE IV
PERFORMANCE COMPARISON OF MULTI-STEP STRATEGIES OF NE-ISO SUBSTATION 1 (CASE-III)

Time	**Direct Strategy**		**Recursive Strategy**		**DirRec Strategy**	
Hours	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)
1	0.728	0.955	0.728	0.955	**0.728**	**0.955**
2	1.437	1.858	1.502	1.968	**1.455**	**1.794**
3	2.011	2.655	2.220	2.997	**1.932**	**2.516**
4	2.532	3.349	2.892	3.981	**2.498**	**3.146**
5	3.289	4.326	3.480	4.884	**2.656**	**3.336**

TABLE V
PERFORMANCE COMPARISON OF MULTI-STEP STRATEGIES OF NE-ISO SUBSTATION 1 (CASE-IV)

Time	**Direct Strategy**		**Recursive Strategy**		**DirRec Strategy**	
Hours	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)
1	0.731	1.056	0.731	1.056	**0.731**	**1.056**
2	1.189	1.707	1.345	1.898	**1.181**	**1.697**
3	2.072	2.826	1.722	2.387	**1.551**	**2.139**
4	2.235	2.881	1.961	2.704	**1.822**	**2.440**
5	2.144	2.874	2.097	2.912	**1.814**	**2.522**

and auto-correlation coefficients have been obtained, respectively [17], [18]. Twelve input features were found effective to prepare the hour-ahead forecasting models of load demand. The following set of input variables has been used to forecast the load demand, Ld_{t_1}, at time t_1.

$$
\begin{aligned}
(Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)}, Ld_{(t_1-23)}, Ld_{(t_1-24)}, \\
Ld_{(t_1-25)}, Ld_{(t_1-167)}, Ld_{(t_1-168)}, Ld_{(t_1-169)}, \\
Temp_{(t_1)}, Temp_{(t_1-1)}, Temp_{(t_1-24)})
\end{aligned}
\tag{1}
$$

Where, $Ld_{(t_1-1)}$ is actual load demand at time (t_1-1) and $Temp_{(t_1)}$ is forecasted ambient temperature at time (t_1).

IV. LOAD FORECASTING STRATEGIES

Load forecasting is considered to be an essential area of power systems for predicting future load demand. Load forecasting has been done by single-step and multi-step methods. For one-hour-ahead, both methods have the same accuracy of prediction. For the prediction of the second hour, the generally multi-step method uses the first hour forecasted load into input model, then train the forecasting model and predict the load demand for the second hour. Similarly, it can be done for more than two hours ahead. The different forecasting strategies are explained in the following subsections.

A. Single-step-ahead Forecasting

Single-step is the simplest method for forecasting the load. In this method predicted values of previous hours load are not used for further load forecasting, so this method is known as one-step or open-loop forecasting. One-step-ahead load forecasting value at time t_1 is achieved by passing the current and past actual demand of load at time (t_1-1), (t_1-2), (t_1-3), (t_1-23), (t_1-24), (t_1-25), (t_1-167), (t_1-168), (t_1-169) and forecasted ambient temperature at time (t_1), (t_1-1), (t_1-24) to a model [19].

TABLE VI
PERFORMANCE COMPARISON OF MULTI-STEP STRATEGIES OF NE-ISO SUBSTATION 2 (CASE-I)

Time	**Direct Strategy**		**Recursive Strategy**		**DirRec Strategy**	
Hours	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)
1	0.597	0.809	0.597	0.809	**0.597**	**0.809**
2	1.242	1.688	1.274	1.713	**1.233**	**1.658**
3	2.298	3.032	1.842	2.462	**1.721**	**2.342**
4	2.830	3.661	2.274	3.017	**2.095**	**2.747**
5	2.928	3.790	2.572	3.416	**2.430**	**3.211**

TABLE VII
PERFORMANCE COMPARISON OF MULTI-STEP STRATEGIES OF NE-ISO SUBSTATION 2 (CASE-II)

Time	**Direct Strategy**		**Recursive Strategy**		**DirRec Strategy**	
Hours	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)
1	0.628	0.808	0.628	0.808	**0.628**	**0.808**
2	1.102	1.422	1.271	1.662	**1.059**	**1.395**
3	1.406	1.862	1.786	2.365	**1.289**	**1.657**
4	1.922	2.479	2.212	2.924	**1.517**	**1.966**
5	2.160	2.735	2.602	3.413	**1.650**	**2.115**

$$
\begin{aligned}
F(t_1) = M(&Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)}, Ld_{(t_1-23)}, \\
&Ld_{(t_1-24)}, Ld_{(t_1-25)}, Ld_{(t_1-167)}, Ld_{(t_1-168)}, \\
&Ld_{(t_1-169)}, Temp_{(t_1)}, Temp_{(t_1-1)}, Temp_{(t_1-24)})
\end{aligned}
\tag{2}
$$

where $F(t_1)$ is the forecasted demand of load in single-step-ahead, $Ld_{(t_1-1)}$ is the actual load demand at time (t_1-1), $Temp_{(t_1)}$ is forecasted ambient temperature at time (t_1), and M is model for single step ahead.

B. Multi-step-ahead Forecasting

In this method, after the first hour ahead, generally forecasted values of previous hours load are used in the input model to predicting the load demand in next hour ahead. Since forecasted values are used in the input model, so this method is also known as a closed-loop forecasting method. Ben Taieb.et.al [9] described different strategies of multiple-step-ahead forecasting, among them DS, RS, and DRS have been considered in this work.

1) Direct Strategy: In direct strategy for forecasting the 1^{st}, 2^{nd}, 3^{rd},...,k^{th} hour-ahead load demand, M_1, M_2, M_3,......,M_k models are formed, respectively. In this strategy, forecasted values of previous hours load demand are not used for forecasting the next hour-ahead load demand. The different hour-ahead load forecasting models are independent of each other. The algorithm has been used for 1^{st}, 2^{nd}, 3^{rd},...,k^{th} hour-ahead is similar to the single-step-ahead method, which is shown in below equations.

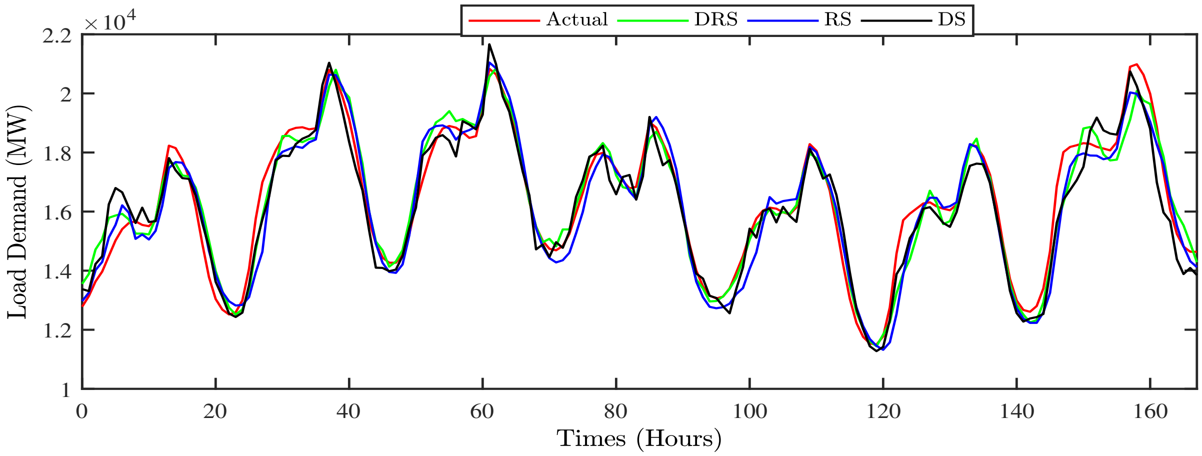

Fig. 2. The performance comparison graph for Case-I of NE-ISO (substation 2) data sets for 5th hour MSALF

$$
\begin{aligned}
F_1(t_1) = M_1(&Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)},\\
&Ld_{(t_1-23)}, Ld_{(t_1-24)}, Ld_{(t_1-25)}, Ld_{(t_1-167)},\\
&Ld_{(t_1-168)}, Ld_{(t_1-169)}, Temp_{(t_1)},\\
&Temp_{(t_1-1)}, Temp_{(t_1-24)}) \qquad (3)\\
F_1(t_1+1) = M_2(&Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)},\\
&Ld_{(t_1+1-23)}, Ld_{(t_1+1-24)}, Ld_{(t_1+1-25)},\\
&Ld_{(t_1+1-167)}, Ld_{(t_1+1-168)}, Ld_{(t_1+1-169)},\\
&Temp_{(t_1+1)}, Temp_{(t_1+1-1)}, Temp_{(t_1+1-24)}) \qquad (4)\\
&\vdots\\
F_1(t_1+k-1) = M_k(&Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)},\\
&Ld_{(t_1+k-1-23)}, Ld_{(t_1+k-1-24)},\\
&Ld_{(t_1+k-1-25)}, Ld_{(t_1+k-1-167)},\\
&Ld_{(t_1+k-1-168)}, Ld_{(t_1+k-1-169)},\\
&Temp_{(t_1+k-1)}, Temp_{(t_1+k-1-1)},\\
&Temp_{(t_1+k-1-24)}) \qquad (5)
\end{aligned}
$$

where $F_1(t_1)$, $F_1(t_1+1)$,...,$F_1(t_1+k-1)$ are forecasted values of load demand for 1^{st}, 2^{nd},...,k^{th} hour-ahead, $Ld_{(t_1-1)}$ is the actual load demand at time (t_1-1), $Temp_{(t_1)}$ is forecasted ambient temperature at time (t_1), and M_1, M_2,...,M_k are models for 1^{st}, 2^{nd},...,k^{th} hour-ahead, respectively. The inputs such as $(Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)})$ are common in all the forecasting models since the forecasted values are not used as input in DS. These three inputs are recently available load demand, which can be used as inputs.

2) Recursive Strategy: In recursive strategy for forecasting the 1^{st}, 2^{nd},...,k^{th} hour-ahead load demand, only one model i.e., M is needed to prepare. It can be obtained from (2) which is just one hour-head forecasting model. In this strategy, forecasted values of previous hours load have been used as input variables to forecast the load for 2^{nd}, 3^{rd},...,k^{th} hour-ahead, which is clearly shown in the below equations. For example, to forecast the k_{th} hour-ahead of load using a recursive strategy, at first find the 1^{st}, 2^{nd}, 3^{rd},...,$(k-1)^{th}$ hour-ahead forecasted load i.e $F_2(t_1)$, $F_2(t_1+1)$, $F_2(t_1+2)$,..., $F_2(t_1+k-2)$ by using a model M and then by using these forecasted load as input variables into the same input model M to forecast the load demand in the k^{th} hour-ahead. In this strategy number of input variables have been same in each model.

$$
\begin{aligned}
F_2(t_1) = M(&Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)}, Ld_{(t_1-23)},\\
&Ld_{(t_1-24)}, Ld_{(t_1-25)}, Ld_{(t_1-167)}, Ld_{(t_1-168)},\\
&Ld_{(t_1-169)}, Temp_{(t_1)}, Temp_{(t_1-1)},\\
&Temp_{(t_1-24)}) \qquad (6)\\
F_2(t_1+1) = M(&F_2(t_1), Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1+1-23)},\\
&Ld_{(t_1+1-24)}, Ld_{(t_1+1-25)}, Ld_{(t_1+1-167)},\\
&Ld_{(t_1+1-168)}, Ld_{(t_1+1-169)}, Temp_{(t_1+1)},\\
&Temp_{(t_1+1-1)}, Temp_{(t_1+1-24)}) \qquad (7)\\
&\vdots\\
F_2(t_1+k-1) = M(&F_2(t_1+k-2), F_2(t_1+k-3),\\
&F_2(t_1+k-4), Ld_{(t_1+k-1-23)},\\
&Ld_{(t_1+k-1-24)}, Ld_{(t_1+k-1-25)},\\
&Ld_{(t_1+k-1-167)}, Ld_{(t_1+k-1-168)},\\
&Ld_{(t_1+k-1-169)}, Temp_{(t_1+k-1)},\\
&Temp_{(t_1+k-1-1)}, Temp_{(t_1+k-1-24)}) \qquad (8)
\end{aligned}
$$

where $F_2(t_1)$, $F_2(t_1+1)$,...,$F_2(t_1+k-1)$ are forecasted values of load demand for 1^{st}, 2^{nd},...,k^{th} hour-ahead, $Ld_{(t_1-1)}$ is the actual load demand at time (t_1-1), $Temp_{(t_1)}$ is forecasted ambient temperature at time (t_1),and M is model for 1^{st}, 2^{nd},...,k^{th} hour-ahead, respectively.

3) DirRec Strategy: DirRec strategy is the combination of direct and recursive strategy. In DirRec strategy for forecasting the 1^{st}, 2^{nd}, 3^{rd},...,k^{th} hour-ahead load demand, M_1, M_2, M_3,...,M_k models are formed respectively. In this strategy, forecasted values of previous hours load have been used as input variables to forecast the load for 2^{nd}, 3^{rd},...,k^{th} hour-ahead, which is clearly shown in below equations.

Thus, in this strategy number of input variables have been changed in each model. For example, to forecast the k^{th} hour-ahead of load using a DirRec strategy, at first find the 1^{st}, 2^{nd}, 3^{rd},...,$(k-1)^{th}$ hour-ahead forecasted load i.e $F_3(t_1)$, $F_3(t_1+1)$, $F_3(t_1+2)$,..., $F_3(t_1+k-2)$ by using M_1, M_2, M_3,..., $M_{(k-1)}$ models and then a new model M_k has been formed by using 1^{st} to $(k-1)^{th}$ hour-ahead forecasted load, actual load demand and forecast ambient temperature, which is used in first hour ahead model formation. Second hour-ahead model is dependent on 1^{st} hour-ahead model, Third hour-ahead model is dependent on $1st$ and $2nd$ hour-ahead model,..., k^{th} hour-ahead model is dependent on 1^{st}, 2^{nd}, 3^{rd},...,$(k-1)^{th}$ hour-ahead model.

$$\begin{aligned} F_3(t_1) = M_1(&Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)}, Ld_{(t_1-23)}, \\ &Ld_{(t_1-24)}, Ld_{(t_1-25)}, Ld_{(t_1-167)}, Ld_{(t_1-168)}, \\ &Ld_{(t_1-169)}, Temp_{(t_1)}, Temp_{(t_1-1)}, \\ &Temp_{(t_1-24)}) \end{aligned} \tag{9}$$

$$\begin{aligned} F_3(t_1+1) = M_2(&F_3(t_1), Ld_{(t_1-1)}, Ld_{(t_1-2)}, Ld_{(t_1-3)}, \\ &Ld_{(t_1+1-23)}, Ld_{(t_1+1-24)}, Ld_{(t_1+1-25)}, \\ &Ld_{(t_1+1-167)}, Ld_{(t_1+1-168)}, Ld_{(t_1+1-169)}, \\ &Temp_{(t_1+1)}, Temp_{(t_1+1-1)}, Temp_{(t_1+1-24)}) \end{aligned} \tag{10}$$

$$\vdots$$

$$\begin{aligned} F_3(t_1+k-1) = M_k(&F_3(t_1+k-2), F_3(t_1+k-3), \ldots, \\ &F_3(t_1+1), F_3(t_1), Ld_{(t_1-1)}, Ld_{(t_1-2)}, \\ &Ld_{(t_1-3)}, Ld_{(t_1+k-1-23)}, Ld_{(t_1+k-1-24)}, \\ &Ld_{(t_1+k-1-25)}, Ld_{(t_1+k-1-167)}, \\ &Ld_{(t_1+k-1-168)}, Ld_{(t_1+k-1-169)}, \\ &Temp_{(t_1+k-1)}, Temp_{(t_1+k-1-1)}, \\ &Temp_{(t_1+k-1-24)}) \end{aligned} \tag{11}$$

where $F_3(t_1)$, $F_3(t_1+1)$,...,$F_3(t_1+k-1)$ are forecasted values of load demand for 1^{st}, 2^{nd},...,k^{th} hour-ahead, $Ld_{(t_1-1)}$ is the actual load demand at time (t_1-1), $Temp_{(t_1)}$ is forecasted ambient temperature at time (t_1), and M_1, M_2,...,M_k are models for 1^{st}, 2^{nd},...,k^{th} hour-ahead, respectively.

C. Forecasting Evaluation Criteria

The performance of above discussed strategies is compared by calculating the mean absolute error, (MAE) & root mean square error, $(RMSE)$. The MAE & $RMSE$ are calculated in percentage with respect to the actual value of the load. They are expressed as follows:

$$|e_t| = |a_t - \hat{a}_t(\mathbf{x})| \tag{12}$$

$$MAE = \left(\frac{1}{K}\sum_{t=1}^{K}(|e_t|/|a_t|)\right) \times 100\% \tag{13}$$

TABLE VIII
PERFORMANCE COMPARISON OF MULTI-STEP STRATEGIES OF NE-ISO SUBSTATION 2 (CASE-III)

Time	Direct Strategy		Recursive Strategy		DirRec Strategy	
Hours	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)
1	0.551	0.749	0.551	0.749	**0.551**	**0.749**
2	1.035	1.434	1.158	1.651	**1.065**	**1.544**
3	1.622	2.156	1.630	2.434	**1.415**	**2.151**
4	2.405	3.294	2.035	3.057	**1.851**	**2.729**
5	2.621	3.513	2.419	3.560	**2.188**	**3.004**

TABLE IX
PERFORMANCE COMPARISON OF MULTI-STEP STRATEGIES OF NE-ISO SUBSTATION 2 (CASE-IV)

Time	Direct Strategy		Recursive Strategy		DirRec Strategy	
Hours	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)	MAE(%)	RMSE(%)
1	0.488	0.637	0.488	0.637	**0.488**	**0.637**
2	1.282	1.716	0.968	1.291	**0.843**	**1.152**
3	1.503	1.934	1.351	1.877	**1.142**	**1.610**
4	1.770	2.388	1.632	2.350	**1.330**	**1.867**
5	1.794	2.417	1.858	2.704	**1.370**	**1.942**

$$RMSE = \sqrt{\frac{1}{K}\sum_{t=1}^{K}(|e_t|/|a_t|)^2} \times 100\% \tag{14}$$

where K is the total number of observations. e_t, a_t and $\hat{a}_t(\mathbf{x})$ are the error, actual and forecasted load demand for observation t, respectively.

V. SIMULATION RESULTS AND ANALYSIS

A. Description of Data

data sets of two substations of NE-ISO has been used for the analysis of MSALF. The duration of the recorded data of substation 1 is (01-01-2006 to 31-12-2007) [20] and substation 2 is (01-01-2013 to 31-12-2014) [21]. These two data sets contain load and temperature data of one-hour time resolution. For forecasting of load demand, the forecasted temperature is used in the model, but due to unavailability of forecasted temperature, the actual temperature has been used to prepare the models. The data sets are divided by the maximum value of load and temperature to obtain the normalised values. This converts the values of the data sets in between 0 and 1. Normalisation prevents the neural network activation functions from saturation. The NE-ISO (substation 1 and substation 2) data sets are divided into four cases. For substation 1 : Case-I (01-02-2006 to 14-02-2007), Case-II (01-05-2006 to 14-05-2007), Case-III (01-08-2006 to 14-08-2007), and Case-IV (01-11-2006 to 14-11-2007). For substation 2 : Case-I (01-01-2013 to 14-01-2014), Case-II (01-04-2013 to 14-04-2014), Case-III (01-07-2013 to 14-07-2014), and Case-IV (01-10-2013 to 14-10-2014). For each case, last 14 days are used for testing, and remaining data are used for training. The performance of each data sets has been analysed for four Case in next section.

B. Size of data sets

The primary purpose of load forecasting is to obtain better prediction accuracy, that depends on the size of data sets for training the ANN models. The size of data sets can be three months, four months, six months, or 1 year. The six different size of training data sets such as three, four, six, eight, ten, and twelve months have been compared for twelve months of a year. The forecasting models have been prepared for six cases of each month separately. Then, the ANN with LM training algorithm has been applied for forecasting of one-hour ahead load demand for the first two-week, i.e. (14 days) of each month. The performance of each month for six cases are shown in Table-I. The average performance for the case-VI data sets, which considered the size of data sets of twelve months, is better as compared to other cases. Therefore, in this paper, the forecasting models have been prepared by twelve months of training data sets.

C. Results and Discussions

The load demand forecasting models have been implemented by using the ANN with LM training algorithm with one hidden layer consist of 25 neurons. Two different substations of NE-ISO data sets are analyzed for four different cases. The performance of different MSALF strategies has been compared by forecasting one to five hours-ahead load demand for a two week period of each case. The performance evaluation of two different substations of NE-ISO data sets from 1^{st} to 5^{th} hour-ahead is shown in Table II-V and Table VI-IX, respectively. It can be analysed from Table II-V and Table VI-IX that the performance of DirRec strategy is better than direct and recursive strategy.

The actual and predicted load demand for three different strategies of MSALF for one case of two different substations of NE-ISO data sets are shown in Fig. 1 and Fig. 2, respectively. It can be analysed from Figs. 1-2 that the forecasted load demand curve for DirRec strategy is close to the actual load demand curve.

For two different data sets, DirRec strategy accuracy is superior as compared to direct and recursive strategy in all the four cases. The performance of the direct and recursive strategy is dependent on the data sets, as shown in Table-II to Table-IX.

VI. Conclusion

The importance of multi-step-ahead short-term load forecasting is utilized for renewable energy usage, electricity market pricing, and enhanced energy saving. The twelve months of the training data sets was found the best accuracy as compared to the smaller size of the training data sets. The performance of three different strategies for MSALF, i.e. direct strategy, recursive strategy, and DirRec strategy were compared using artificial neural network with LM training algorithm. The performance of the DirRec strategy gave better result than direct and recursive strategy. The performance of DirRec strategy is better since it combines the advantages of the direct and recursive strategies, which is validated by the obtained results in this paper.

It can not be concluded that DirRec strategy always gives better results than other multi-step strategies, but it may be a good option for forecasting of multi-step load demand because machine learning models may have different behaviour for different data sets.

References

[1] M. Shahidehpour, H. Yamin, and Z. Li, *Market operations in electric power systems: forecasting, scheduling, and risk management*. John Wiley & Sons, 2003.

[2] M. A. Abu-El-Magd and N. K. Sinha, "Short-term load demand modeling and forecasting: a review," *IEEE transactions on systems, man, and cybernetics*, vol. 12, no. 3, pp. 370–382, 1982.

[3] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558–1573, 1987.

[4] T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, "One-hour-ahead load forecasting using neural network," *IEEE Transactions on power systems*, vol. 17, no. 1, pp. 113–118, 2002.

[5] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks:: The state of the art," *International journal of forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[6] F.-J. Chang, Y.-M. Chiang, and L.-C. Chang, "Multi-step-ahead neural networks for flood forecasting," *Hydrological sciences journal*, vol. 52, no. 1, pp. 114–130, 2007.

[7] A. F. Atiya, S. M. El-Shoura, S. I. Shaheen, and M. S. El-Sherif, "A comparison between neural-network forecasting techniques-case study: river flow forecasting," *IEEE Transactions on neural networks*, vol. 10, no. 2, pp. 402–409, 1999.

[8] N. H. An and D. T. Anh, "Comparison of strategies for multi-step-ahead prediction of time series using neural network," in *2015 International Conference on Advanced Computing and Applications (ACOMP)*. IEEE, 2015, pp. 142–149.

[9] S. B. Taieb, G. Bontempi, A. F. Atiya, and A. Sorjamaa, "A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition," *Expert systems with applications*, vol. 39, no. 8, pp. 7067–7083, 2012.

[10] T. Hill, L. Marquez, M. O'Connor, and W. Remus, "Artificial neural network models for forecasting and decision making," *International journal of forecasting*, vol. 10, no. 1, pp. 5–15, 1994.

[11] X. Zhang, "Time series analysis and prediction by neural networks," *Optimization Methods and Software*, vol. 4, no. 2, pp. 151–170, 1994.

[12] A. K. Alexandridis and A. D. Zapranis, *Wavelet neural networks: with applications in financial engineering, chaos, and classification*. John Wiley & Sons, 2014.

[13] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Transactions on power systems*, vol. 16, no. 1, pp. 44–55, 2001.

[14] C. L. Dewangan, S. Singh, and S. Chakrabarti, "Solar irradiance forecasting using wavelet neural network," in *2017 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*. IEEE, 2017, pp. 1–6.

[15] Y. Ning, Y. Liu, H. Zhang, and Q. Ji, "Comparison of different bp neural network models for short-term load forecasting," in *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*, vol. 3. IEEE, 2010, pp. 435–438.

[16] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the marquardt algorithm," *IEEE transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994.

[17] MathWorks, "Matlab 2018a," https://www.mathworks.com/help/, accessed June, 2018.

[18] N. Pindoriya, S. Singh, and S. Singh, "One-step-ahead hourly load forecasting using artificial neural network," in *2009 International Conference on Power Systems*. IEEE, 2009, pp. 1–6.

[19] V. M. Landassuri-Moreno, C. L. Bustillo-Hernández, J. J. Carbajal-Hernández, and L. P. S. Fernández, "Single-step-ahead and multi-step-ahead prediction with evolutionary artificial neural networks," in *Iberoamerican Congress on Pattern Recognition*. Springer, 2013, pp. 65–72.

[20] "Ne-iso-substation-1," https://in.mathworks.com/electricity-load-and-price-forecasting.

[21] "Ne-iso-substation-2," https://www.iso-ne.com/isoexpres/web/reports/load-and-demand.