Echo State Network Performance in Electrical and Industrial Applications

Muhammad Mansoor, *, Francesco Grimaccia*, Marco Mussetta*

*Politecnico di Milano, Dipartimento di Energia Via La Masa 34, 20156 Milano, Italy muhammad.mansoor@polimi.it, francesco.grimaccia@polimi.it, marco.mussetta@polimi.it

Abstract—Echo State Network (ESN) attracted significant interest in the research activities in last years. In this paper some application to industrial cases are presented, considering in particular the energy and manufacturing sectors. In particular, load forecasting is crucial for penetrations of renewable energy sources and extension of programs in the paradigm of smart grids. Feed-Forward Neural Network (FFNN) based techniques have been widely used in recent years and applied to predict the electric load with high accuracy. This research work is focused on the use and comparison of neural network approaches, *i.e.* FFNN and ESN, on a dataset related to industrial application. The results of both models are compared based on their accuracy through experimental measurements and suitably defined metrics.

Index Terms—Load forecasting; Neural Network; Echo State Network; Demand Response programs

I. INTRODUCTION

Based on the feedback connection, neural networks can be divided into two main classes, *i.e.* Feed-Forward Neural Networks (FFNN) and Recurrent Neural Networks (RNN).

FFNN is the neural network which does not have any feedback connection from the output to the input. The activation function f is chosen to satisfy the specifications of the problem the neuron is attempting to solve and then the parameter w and b are adjusted by learning rule in order to match the input and output relationship.

Recurrent Neural Network (RNN) is the neural network with at least one feedback loop, i.e. when the output of an element of the system influences the input applied to that particular element of the system. The concept was presented by Hopfield [1] with use of statistical machines to describe the working of recurrent networks and used them as memory.

In the basic model of RNN there is only one hidden layer. The output of the neurons in hidden layer is given as feedback to the same layer as input but it is processed through unit-time delay block, to have the output delayed by one time step.

The hidden neurons have the record of their previous activations which enable them to perform the tasks of learning that extend over time and due to this cyclic feedback of information, these neurons discover the abstract representation of the time [2]. Like the FFN, the RNN can also have multiple layers but with feedback in each layer, which makes them Recurrent Multilayer Perceptron (RMLP), whose basic model is shown in Figure 1. Reservoir Computing (RC) is a paradigm of training and understanding the RNN where an RNN is generated on random basis and remains unchanged during the training and only the readout is trained. This randomly generated RNN is called the Reservoir which is a collection of recurrently connected neurons having a nonlinear transformation of the past input; the desired output signal is obtained as the linear combination of the signals coming from the reservoir, as described in detail in [3].

In this research work, the electrical load forecasting is discussed with respect to theory and mathematical modelling. Two specific machine learning techniques based on artificial neural networks i.e. FFNN and Echo State Network (ESN) are considered for implementation by using the data after clustering by K-Means technique, and compared considering newly introduced error metrics.

The paper is organized as follows: in Section II, the ESN concept is presented, while its application to electrical load forecasting is presented in Section III, with a detailed description on the proposed methodology in terms of data analysis and clustering, model definition, proposed forecasting accuracy measures and resulting performance. In Section IV, the industrial application in presented. In Section V conclusions are drawn and possible future perspective are discussed.

II. ECHO STATE NETWORKS

Echo State Networks (ESN) was proposed by Jaeger [4] as a simple approach of training RNN with inexpensive computational capability. The training of RNN through ESN requires experience as the randomly generated reservoir depends upon global parameters which have to be set correctly for successful results.

The fundamental concept of ESN is that the activation states of the randomly generated reservoir contain the input history and these activation states are called the *echoes* of the input history hence the name of the method is echo state network. If u(n), u(n-1), u(n-2), ... represents the input history, then the activation state x(n) of an ESN can be written in the form given by:

$$x(n) = E(u(n), u(n-1), u(n-2), ...)$$
(1)

where E is termed as Echo Function and is essential for the Echo State Property which states that the activations x(n) of



Fig. 1. Recurrent Multilayer Perceptron (RMLP).



Fig. 2. Echo State Network Model.

the neurons in the reservoir are the systematic variations of the input driver signal u(n) [5]. The basic model of ESN is shown in Figure 2.

If there are N input units, M internal network units in the reservoir and L outputs units with their corresponding activations at time step n to be $u(n) = (u_1(n), ..., u_K(n))$, $x(n) = (x_1(n), ..., x_M(n))$ and $y(n) = (y_1(n), ..., y_K(n))$ respectively, then the state update equation of the ESN reservoir shown in Figure 2 is written as:

$$x(n) = f\left(W_{in} \cdot u(n) + W \cdot x(n-1) + W_{back} \cdot y(n-1)\right)$$
(2)

where n is the current time step, x(n) is the vector of neuron reservoir activations, f is the activation function of the neurons in the reservoir, W_{in} is the input weight matrix of order $M \times$ N, u(n) is the vector of inputs, W is the reservoir weight matrix of order $M \times M$ of internal states, x(n-1) is the one time step back activation state of the reservoir, W_{back} is the output feedback weight matrix of order $M \times L$ and y(n-1)is the one time step back output generated by the network. The ESN training algorithm following the supervised learning approach can be summarised in four steps as follows.

- 1) the parameters N, M, L for the network and generate a large reservoir RNN randomly in terms of W_{in} , W_{back} (optional), W and α (if required by the task).
- Use the training input u(n) to collect the activation states x(n) of the reservoir corresponding to the inputs.
- 3) Calculate the linear readout weights W_{out} from the reservoir by the use of linear regression and minimizing the mean square error between the network output y(n) and $y_{target}(n)$.

4) With help of the trained readout weights, calculate the network output y(n) by using novel input data u(n).

III. ELECTRICAL LOAD FORECASTING

Load forecasting is relevant not only for long-term strategic planning and operational decision making: Recently, this area is getting interest due to deregulation of the electric power systems, requiring the analysis of different geographical, meteorological, social and load type aspects related to the energy demand [7].

In particular, there are three types of inputs data used in STLF. These are weather variables (temperature, humidity, wind and cloud covers), seasonal variations related to heating, ventilation and air conditioning (HVAC) and measured load data samples (hourly loads for the previous hour, the previous day, and the same day of the previous week). The output of this STLF is the predicted hourly load in a specific day.

As mentioned by Jindal et al. in [8] effective DR is dependent critically on load peak management and data analytics: [9] highlights the need for new multi-dimensional load forecasting model in the view of requirements set by DR programs, for encouraging the participation of consumers and at the same time for increasing the penetration of renewable energy sources.

In stochastic time series model, the load series is modeled as the output from a linear system that usually has a white noise as the input series. Based on the features of the linear filter, several models can be used, such as [10]: Autoregressive (AR) Model, Autoregressive Moving-Average (ARMA) Model, Autoregressive Integrated Moving-Average (ARIMA) Model [11], Autoregressive Moving Average with Exogenous Variables (ARMAX) Model [12], Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) Model [13].

In fact, neural networks are usually capable of general purpose nonlinear time series forecasting [14]. The first investigations done in the field of STLF by using ANN approach in the literature include the work of Lee et al. [15].

A recent study by Wollsen et al. [16] has shown the effect of influential factors for load prediction in a demand response (DR) context. By comparing the forecasting algorithms, they conclude that FFNN and Echo State Network (ESN) have almost similar results which is interesting in comparative analysis. ESN is claimed to be suitable for load forecasting because of its temporal memory. In more recent years several studies analyzed ESN performance on STLF application [17], considering the choice of its key parameters [18], its prediction accuracy [19], and a comparison with other widely used approaches [20], [21].

All the approaches described above are generally reliable, however, often they fail to adapt to unexpected weather conditions or unusual holidays: these present a highly nonlinear relationship with the daily load. Load forecast in these cases may result less accurate than desired, therefore different approaches must be used in order to correctly link the variables.

On the other hand, computational intelligence-based forecasting techniques are commonly adopted for the ability to deal with non-linearity in time series modelling [22].

This section explains the methodology adopted to obtain the results of the electrical load forecasting by using FFNN and ESN techniques of ANN. The implementation of load forecasting models is shown in Figure 3.



Fig. 3. Implementation of Load Forecasting Models.

A. Data Analysis and Clustering

The data considered for load forecasting in this paper are obtained from Bovisa Campus of Politecnico di Milano, Italy. The measurements of the hourly load demand of the campus are obtained for 4 sub-zones. These sub-zones represent different buildings in the campus which are categorized in the commercial buildings. The data length for each sub-zone is from 1st April 2016 to 28th February 2017 which constitutes 334 load profiles each having 24 variables as hourly readings. The individual load data of the sub-zones is aggregated to get the aggregated load demand.

The load measurements of the 4 different sub-zones are shown in Figures 4, 5, 6 and 7, respectively. As it can be seen from these trends of the loads that the load has not a uniform periodic shape mainly because it is associated with the university buildings where there can be different types of loads, such as those in laboratories, auditoriums and classrooms have different consumption patterns.

Moreover, in order to show the relation among electrical consumption and the seasonal variability, Figures 8, 9 and 10 show, respectively, the aggregate load measurements and



Fig. 4. Load measurements of sub-zone 1 in the considered period.



Fig. 5. Load measurements of sub-zone 2 in the considered period.



Fig. 6. Load measurements of sub-zone 3 in the considered period.



Fig. 7. Load measurements of sub-zone 4 in the considered period.



Fig. 8. Aggregate load measurements and corresponding temperature in the considered period.



Fig. 9. Global irradiance measured in the considered period.



Fig. 10. Relative humidity measurements in the considered period.

corresponding temperature, global irradiance and relative humidity in the considered period.

The consumption also depends upon the nature of the day, like weekdays, weekends and holidays. Therefore, it is required to process this data before applying to the load forecasting models. In this work, the clustering technique is used to process the data and cluster it into different types of days. The clustering in different types of the days will help in managing the demand response events effectively for each day type. The clustering method used is the popular K-Means Clustering [23].

The results of the K-Means clustering including the clustered profiles, and Aggregated Load respectively, are shown in Table I, which summarizes the number of profiles in all the clusters out of 334 total profiles.

TABLE I DISTRIBUTION OF AVAILABLE DATA SAMPLES IN CLUSTERS

Data	Cluster 1	Cluster 2	Cluster 3	
Sub-zone 1	3624	2784	1608	
Sub-zone 2	3840	2520	1656	
Sub-zone 3	3744	2784	1488	
Sub-zone 4	3792	2760	1464	
Aggregated	3864	2712	1440	

Each of the samples indicated in the Tab. 1 corresponds tyo sequential hourly load measurements, so the number of daily profiles for each cluster can be obtained by dividing by 24 the number of data samples in each cluster. The data related to external weather parameters are linked with these clustered profiles for completing the clustered data sets.

After obtaining the complete clustered data sets, each individual data set is divided into Training Set and Testing Set for load forecasting models. The division is based on the total dimension of the data which has a length from 1st April 2016 to 28th February 2017, as shown in Table II.

TABLE II DISTRIBUTION OF AVAILABLE DAILY PROFILES IN TRAINING AND TEST SETS

	Cluster 1		Cluster 2		Cluster 3	
Data	train	test	train	test	train	test
Sub-zone 1	131	20	108	8	41	26
Sub-zone 2	143	17	97	8	45	24
Sub-zone 3	135	21	109	7	50	12
Sub-zone 4	134	24	111	4	49	12
Aggregated	141	20	105	8	47	13

The training set indicated for each data set is fully utilized for the training of the load forecasting models. However, the first 24 data points i.e. 24 hours (One day) are considered from testing set for generating the load forecast from these models because the day-ahead forecasting is only considered in the scope of this work which lies under STLF but all the testing set can also be used to generate load forecast.

B. Model Definitions

After the data analysis and processing step, the load forecasting models are implemented both for FFNN and ESN. MATLAB software is used for both of these models and specifically, Neural Network ToolboxTM is used for FFNN model and Very Simple ESN Toolbox by Jaeger [24] is used for the ESN model.

The data division for ESN model is same as defined in FFNN model but here the training data and testing data are combined together into a single set separately for input and output and then fed into the ESN. The reservoir is considered to be large and is varied from 100 to 1000 network dimensions.

The number of the inputs are 4 and the number of output is 1. The global parameters of ESN are manually changed to get better and consistent results. The reservoir updates are considered with output feedbacks, scaled white noise, teacher forcing and without leaky integrator co-efficient. The readout is trained through direct pseudoinverse solution.

C. Forecasting Accuracy Measures

There are several types of error metrics in literature, each one differently identifying the distance between the actual measured value and the forecasted one. These error measures produce different performance ranking when comparing different forecasting methods on the same data. Each error measure has its advantages and disadvantages which can lead to inaccurate evaluation of forecasting results so it is impossible to choose only one error measure [25].

MAPE is the percentage based error measure which is the most commonly used error measure in the forecasting domain:

MAPE =
$$\frac{1}{n} \sum_{i=1}^{n} \frac{|y_{m,i} - y_{f,i}|}{y_{m,i}} \cdot 100$$
 (3)

where, n is the forecast horizon, $y_{m,i}$ is the actual measured value at time i and $y_{f,i}$ is the forecasted value for the same time.

WMAE is another commonly used error metric, based on the definition of the scale-dependent MAE. It measures the forecasting performance with respect to total energy consumption actually measured:

WMAE =
$$\frac{\sum_{i=1}^{n} |y_{m,i} - y_{f,i}|}{\sum_{i=1}^{n} y_{m,i}} \cdot 100$$
 (4)

A novel accuracy measure (EMAE) was proposed in [26] to avoid the shortcomings of (4). In fact, EMAE provides forecasting accuracy value defined in the range $0 \div 100\%$:

$$\text{EMAE} = \frac{\sum_{i=1}^{n} |y_{m,i} - y_{f,i}|}{\sum_{i=1}^{n} \max(y_{m,i}, y_{f,i})} \cdot 100$$
(5)

The normalized RMSE is the scale-dependent error measure which is also popular in the forecasting domain. It is defined as:

NRMSE =
$$\sqrt{\frac{1}{n} \sum_{i=1}^{n} \frac{(y_{m,i} - y_{f,i})^2}{f_{norm}} \cdot 100}$$
 (6)

where f_{norm} is the normalizing factor: in this study, it is the maximum of the actual measured value $y_{m,i}$.

D. Load Forecasting Results

The results of day ahead 24 hour load forecasting for each cluster of the sub-zones and aggregated load from both of FFNN and ESN models are presented for sub-zones 1-4 and aggregated load in Figure 11, in terms of EMAE indicator. It is to be noted that the choosing of the day for load forecasting and including it in the testing set is done through deterministic approach. For example, the cluster 1 represents the load profiles for normal weekdays where air-conditioning



Fig. 11. EMAE indicator of aggregated and sub-zones load forecasting.

is off so the training data for this cluster is considered from 1st April 2016 to the 31st January 2017 and the very first profile occurring in this cluster in the period between 1st February 2017 and 28th February 2017 is considered as testing set for load forecasting purpose. This very first profile from the testing set shows the similar type of the day as identified by the cluster and the type of the day is deterministically known. The same applies to cluster 2 and cluster 3 of all the sub-zones and aggregated load.

The comparison can be summarized by ranking the model with respect to each cluster taking into account all subzones and aggregated load and also with respect to the subzones taking into account all clusters together. The comparison given below is based on considering the MAPE only as it is considered to be useful in literature for ranking different models with different data sets.

With respect to each cluster taking into account all subzones and aggregated load, FFNN performs better than ESN in all clusters with minimum of MAPE to be 4.761% for cluster 1 (weekdays with low load), 2.937% for cluster 2 (weekends and holidays) and 7.542% for cluster 3 (weekdays with high load).

With respect to each sub-zone and aggregated load taking into account all clusters together, FFNN performs well in subzone 1 with minimum of MAPE to be 2.937% for cluster 2, FFNN performs well in sub-zone 2 with minimum of MAPE to be 4.761% for cluster 1, ESN performs well in sub-zone 3 with minimum of MAPE to be 3.653% for cluster 2, ESN performs well in sub-zone 4 with minimum of MAPE to be 6.789% for cluster 1 and FFNN performs well in aggregated load with minimum of MAPE to be 5.576% for cluster 1.

IV. INDUSTRIAL APPLICATION

Fluorescent Penetrant Inspection (FPI) is a well assessed non-destructive test method used in manufacturing for detecting cracks and other flaws of the product under test. The purpose, in this case, is to apply the ESN model presented in previous sections to an automated inspection system, based on a vision-based expert system for automating the inspection phase of the FPI process in an aerospace manufacturing line. The aim of the process is the identification of the defectiveness status of some mechanical pieces by mean of images.

In particular, ESN was applied to different possible system architecture and to perform some preliminary testing on a basic dataset. The system has been designed for increasing the reliability of the evaluations performed on the part, by implementing an ad-hoc designed human-machine interface to gather feedback from expert operators, thus validating the neural network results in terms of detected flaws in the parts under test.

V. CONCLUSIONS

Load forecasting is an essential tool for planning of the electric power system. It can also be used for demand response programs in the paradigm of smart grids. In this work, data-set related to hourly load profiles together with weather parameters (humidity, ambient temperature and global irradiance) of commercial buildings of a university campus are considered which are divided into four sub-zones. The load profiles of each sub-zone and aggregated load of the commercial buildings are first analyzed and processed through K-Means clustering technique in order to differentiate them with respect to different types of day: three clusters are then considered based on this clustering evaluation.

Two types of load forecasting models have been considered: *i.e.* the feed-forward neural network and echo-state network, a special type of recurrent neural networks. The results from both of the models for each cluster of sub-zones and aggregated load are obtained and compared. MAPE, NRMSE and the recently introduced EMAE metrics are considered as load forecasting accuracy measures: the promising results of ESN, confirm this as a suitable method for load forecasting because.

Future work on this subject will be aimed to use computational optimization techniques to optimize the global parameters and reservoir size of the echo state model, for increasing the overall accuracy, in particular for the industrial case of automated process of flaws detection in manufacturing.

ACKNOWLEDGMENTS

This work is part of the AUTOFPI project, a research program funded by the MANUNET initiative, in the framework of the European Union's Horizon 2020 research and innovation programme. In particular, this work has been funded by: Regione Lombardia with Decreto DD. n.3705 (15.03.2018). MANUNET Transnational Call 2017 has received funding under grant agreement No.721267, in the framework of the ERA-LEARN 2020 support action.

REFERENCES

- J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities". *Proceedings of the National Academy* of Sciences, Vol. 79, No. 8, pp. 2554–2558.
- [2] Haykin, S.O, Neural networks and learning machines (3rd Edition), Pearson, Upper Saddle River, NJ, 2009.
- [3] M. Lukosevicius, H. Jaeger, "Reservoir computing approaches to recurrent neural network training", *Computer Science Review*, Vol. 3, No. 3, pp. 127–149, 2009.

- [4] H. Jaeger, "The *echo state* approach to analysing and training recurrent neural networks-with an erratum note", *German National Research Center for Information Technology GMD Technical Report*, Bonn, Germany, Vol. 148, No. 34, pp. 13, 2001.
- [5] M. Lukosevicius, "Echo state networks with trained feedbacks", *Technical report*, No. 4, 2007.
- [6] G.S. Aujla, S. Garg, S. Batra, N. Kumar, I. You, V. Sharma, "DROPS: A demand response optimization scheme in SDN-enabled smart energy ecosystem", *Information Sciences*, Vol. 476, pp. 453–473, 2019.
- [7] Kyriakides, E. and Polycarpou, M., "Short term electric load forecasting: A tutorial", *Trends in Neural Computation*, Springer Berlin Heidelberg, pp. 391–418, 2007.
- [8] A. Jindal, M. Singh, N. Kumar, "Consumption-Aware Data Analytical Demand Response Scheme for Peak Load Reduction in Smart Grid", *IEEE Transactions on Industrial Electronics*, Vol. 65, No. 11, pp. 8993– 9004, 2018.
- [9] S.C. Chan, K.M. Tsui, H.C. Wu, Y. Hou, Y.C. Wu, F.F. Wu, "Load/price forecasting and managing demand response for smart grids: Methodologies and challenges", *IEEE Signal Processing Magazine*, Vol. 29, No. 5, pp. 68-85, 2012.
- [10] G.E. Box, G.M. Jenkins, G.C. Reinsel, G.M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- [11] M. Hagan, R. Klein, "Identification techniques of Box and Jenkins applied to the problem of short term load forecasting", 1977 Summer Meeting of the IEEE Power Engineering Society, pp. 618–612, 1997.
- [12] H.T. Yang, C.M. Huang, C.L. Huang, "Identification of ARMAX model for short term load forecasting: An evolutionary programming approach", *IEEE Power Industry Computer Application Conference*, pp. 325–330, 1995.
- [13] M.Y. Cho, J.C. Hwang, C.S. Chen, "Customer short term load forecasting by using ARIMA transfer function model", *International Conference on Energy Management and Power Delivery (EMPD'95)*, Vol. 1, pp. 317–322, 1995.
- [14] S. Dwijayanti, M. Hagan, "Short Term Load Forecasting Using a Neural Network Based Time Series Approach", 2013 1st International Conference on Artificial Intelligence, Modelling and Simulation (AIMS), pp. 17–22, 2013.
- [15] K.Y. Lee, Y.T. Cha, J.H. Park, "Short-term load forecasting using an artificial neural network", *IEEE Transactions on Power Systems*, Vol. 7, No. 1, pp. 124–132, 1992.
- [16] M.G. Wollsen, M.B. Kjærgaard, B.N. Jørgensen, "Influential factors for accurate load prediction in a Demand Response context", 2016 IEEE Conference on Technologies for Sustainability (SusTech), (pp. 9-13, 2016.
- [17] I. Jayawardene, G. K. Venayagamoorthy, "Comparison of echo state network and extreme learning machine for PV power prediction", *IEEE Symposium on Computational Intelligence Applications in Smart Grid* (CIASG), Orlando, FL, pp. 1–8, 2014.
- [18] L. Wei, L. Haitian, "Electrical Load Forecasting Using Echo State Network and Optimizing by PSO Algorithm", 10th International Conference on Intelligent Computation Technology and Automation (ICICTA), Changsha, pp. 394–397, 2017.
- [19] Y. Litao, H. Aoyang, W. Li, J. Xu, Z. Zhisheng, "Short-term load forecasting model for metro power supply system based on echo state neural network", 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, pp. 906–909, 2016. doi: 10.1109/ICSESS.2016.7883212
- [20] M.J.A. Rabin, M.S. Hossain, M.S. Ahsan, M.A.S. Mollah, M.T. Rahman, "Sensitivity learning oriented nonmonotonic multi reservoir echo state network for short-term load forecasting", 2013 International Conference on Informatics, Electronics and Vision (ICIEV), Dhaka, pp. 1–6, 2013.
- [21] Q. Song, X. Zhao, Z. Feng, Y. An, B. Song, "Hourly electric load forecasting algorithm based on echo state neural network", 2011 Chinese Control and Decision Conference (CCDC), Mianyang, pp. 3893–3897, 2011.
- [22] M. Louzazni, A. Khouya, K. Amechnoue, A. Gandelli, M. Mussetta, A. Craciunescu, "Metaheuristic Algorithm for Photovoltaic Parameters: Comparative Study and Prediction with a Firefly Algorithm", *Applied Sciences*, Vol. 8, No. 3, art no. 339, 2018.
- [23] J. MacQueen, "Some methods for classification and analysis of multivariate observations", *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Vol. 1, No. 14, pp. 281-297, 1967.

- [24] H. Jaeger, "Simple and very simple Matlab toolbox for Echo State Networks", 2009.
- [25] M.V. Shcherbakov, A. Brebels, N.L. Shcherbakova, A.P. Tyukov, T.A. Janovsky, V.A.E. Kamaev, "A survey of forecast error measures", *World Applied Sciences Journal*, Vol. 24, pp. 171–176, 2013.
 [26] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, E. Ogliari, "Comparison for Provide for Physical Sciences of Machine For Physical Sciences of Physical Sciences of Machine For Physical Sciences of Physical S
- [26] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, E. Ogliari, "Comparison of Training Approaches for Photovoltaic Forecasts by Means of Machine Learning", *Applied Sciences*, Vol. 8, No. 2, pp. 228, 2018.