ORIGINALAN NAUČNI RAD – ORIGINAL SCIENTIFIC PAPER 10.7251/ZRPIM2201476L

PREDICTION OF TIME-SERIES USING TIME-DELAY NEURAL NETWORK ON THE EXAMPLE OF TOTAL ENERGY LOAD

Luka Latinović, Vladimir Tomašević

University Union - Nikola Tesla, School of Engineering Management, Bulevar vojvode Mišića 43, 11000 Belgrade, Serbia, luka.latinovic@fim.rs

ABSTRACT

Predicting the total energy load is extremely important for all elements of the energy system for a number of reasons. This is even more pronounced when energy system includes a significant number of volatile renewable sources. This paper examines the ability of the Time-Delay Neural Network model to predict total energy load, on an hourly basis, using real-world data from Spain. The neural network was created, trained and tested in Neurosolutions 5.0 software. The dataset was obtained from Kaggle database and consists of 35,000 real-world hourly total energy load records from Spain. The results showed a high degree of linear correlation (0.975) between the observed and predicted values with satisfactory Relative Mean-absolute-error (1.239%), and relative Root-mean-square error (1.626%) values. Based on the results, TDNN model emerged as a promising method for both energy load prediction, and time-series prediction.

Keywords: TDNN, prediction, time-series, machine learning, total energy load.

INTRODUCTION

The complexity, dynamics, and nonlinearity of energy systems, regardless the scale, make them in great demand for modelling techniques. An accurate energy load prediction is of paramount importance to an energy system, especially in a situation where more and more renewable energy sources are included, characterised by the volatility of the production (Latinović, & Tomašević, 2022). Various methods have been used in the past to carry out the short-term energy load prediction task (Zhang, Wei, Li, Tan, & Zhou, 2018). There are mainly two sets of literature that are highly relevant to this work. The first one is energy load data analysis (Kermanshahi, & Iwamiya, 2002; Amarasinghe, Marino, & Manic, 2017; Dagdougui, Bagheri, Le, & Dessaint, 2019), and the other is time-series data analysis. Regarding the time-series prediction horizon, majority of research deal with the predictions on yearly or daily level, while the most common methods are based on neural networks (Šebalj, Mesarić, & Dujak, 2019; Janićijević, Petrović, & Stefanović, 2020; Šebalj, Mesarić, & Pap, 2021; Latinović, 2021). A review on neural networks as time-series prediction method shows that neural networks do represent the prediction problem mainstream (Latinović, & Gerasimović, 2021). However, the majority of research examine Recurrent Neural Networks (RNN) (Badii, Nesi, & Paoli, 2018; Camero, Toutouh, Stolfi, & Alba, 2018; Latinović, & Gerasimović, 2021) while a relatively small number is involved in Time-Delay Neural Network (TDNN) research. Nevertheless, TDNN are less resource intensive and in theory easier to train than RNN (Latinović, 2021). Moreover, several research papers showed promising TDNN prediction results such as the one from Weigend & Gershenfeld, (1993), Jha & Sinha (2014), Rafsanjani & Samareh (2016), and Latinović & Burshaid Al Dhaheri (2022). All of the papers showed that TDNN have satisfactory results in solving time-series prediction problem, such is the energy load prediction. Thus, the basic research problem this paper deals with is: is it possible to predict, with the satisfactory accuracy, total energy load on hourly base by TDNN model based only on historical data? The dataset used in this research was obtained from the Kaggle database. It consists of real-world hourly total energy load data from Spain. This article is organised as follows: Section 2 describes methods i.e., course of the research and experimental design; Results are presented in Section 3, while the discussion and conclusion are given at the Sections 4, and 5 respectively.

METHODS

Course of the research

The research was conducted in several consecutive stages as follows:

- The working dataset was created by modifying downloaded Kaggle datasets. Time and total energy load data were left while unnecessary data were removed. Date-time-hour column was divided into three columns (month, day, hour);
- In Neurosolutions for Excel, a table .xlsx file was created by tagging columns and rows as input and desired, and training, cross-validation, and testing, respectively;
- Neural network model was created by heuristically choosing architecture, topology and hyper-parameters such as number of hidden layers and neurons, learning rule and rate, and activation functions. Finally, appropriate performance metrics were chosen.
- Neural network model was trained, tested and evaluated.

The neural network model was created using the Neurosolutions 5.0 software. It is a collection of machine learning algorithms that includes methods for the main data mining problems: regression, classification, and clustering, association rules, and attribute selection.

Experimental design

Time-delay Neural Networks

Time-Delay Neural Networks represent a popular method for time-series prediction (Molina, Liang, Harley, & Venayagamoorthy, 2011; Jha, & Sinha, 2014; Gullapalli, 2018; Kalantari Khandani, & Mikhael, 2021). Past observations of a variable serve as input data, which is used to perform one-step ahead prediction. Assuming $\hat{y}(k+1)$ is the output from TDNN, it can be calculated as follows:

$$\hat{y}(k+1) = f\,(W_x \times [x(k) \cdot x(k-1) \ldots x(k-d)]^T + b_1\,) \times W_y + b_2 \quad (1)$$

where W_x and W_y are weight vectors, b_1 and b_2 represents bias value, d denotes the time delays applied to input vector $x(k)$, and f represents the activation function. TDNN will only have feed-forward connections. The TDNN modelled in this research (Figure 1) was built with three (3) input processing elements (PE's), one (1) output unit and one (1) hidden layer with eight (8) PE's. Both the hidden and output layer uses the Levenberg-Marquardt learning rule with tanh (hyperbolic tangent) activation function described below. A tap delay line represents the input vector with current and past time steps and a tapped delay represents number of samples or delay between successive taps. This model uses three (3) time delay inputs with a tapped delay of one (1). Weights vary during the training process to minimize the error and best weights are saved.

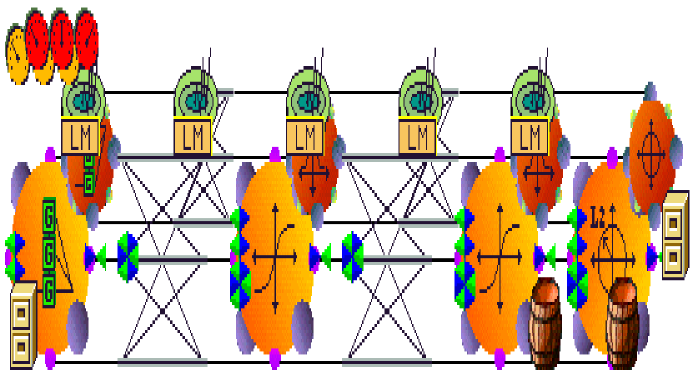

Figure 1. Neurosolutions 5 graphical representation of the experimental TDNN model

Levenberg-Marquardt Learning Algorithm (LMA)

An important step during the training process is the optimization of network parameters such as weights based on the input-output sets being trained on the neural network (Diaz, Fokoue-Nkoutche, Nannicini, & Samulowitz, 2017). There are various learning rules such as backpropagation algorithm, Momentum etc. in the time-series prediction. However, Levenberg-Marquardt algorithm (LMA) emerged as quite successful and efficient learning rule, well known for minimizing the mean square error (MSE) of the network (Kaveh, Duc Bui, & Rutschmann, 2017; Mammadli, 2017; Gullapalli, 2018; Qiao et al., 2018). It determines the best direction to move weights so as to minimize the error.

Tanh transfer/activation function

The hyperbolic tangent function maps the range of each neuron in the layer to between -1 to +1. This allows the network to apply a value to the node (negative) instead of node not having to fire at all. This provides network ability to make soft decisions. Tanh transfer function mathematical equation (2) is given as follows:

$$tanh(x) = \frac{(1-e^{-2x})}{(1+e^{-2x})} \qquad (2)$$

The Tanh transfer function is able to overcome several issues of the sigmoid function, most importantly that a larger input gives almost zero output, preventing the next nodes from learning (Menon, Mehrotra, Mohan, & Ranka, 1996). In the case of hyperbolic tangent function, -1 is given for negative values, allowing the subsequent nodes to learn from it. TDNN used in this research uses the Tanh as the activation function that adds a bias variance of 0.5 to each neuron in the hidden and output layer.

Training and testing the model

Finally, in order to make the results comparable to results from research, and based on good practices and guidelines from other authors (Latinović, & Gerasimović, 2021; Latinović, & Burshaid Al Dhaheri, 2022), the TDNN model was trained through 1000 epochs, with enabled cross-validation. Out of a total of 35,000 records in the dataset, the training set consisted of 60% of the records (21,000), while the cross-validation set consisted of the 30% of the records (10,500). Ten percent (3500) of the records were left out of the training process in order to be used for the testing purposes.

RESULTS

Standardised metrics were used, namely Mean squared error (MSE), Root-mean-square-error (RMSE), Mean absolute error (MAE), Minimum absolute error (Min Abs Err.), Maximum absolute error (Max Abs Err), and Pearson correlation coefficient (r). The results of the TDNN test are summarised in the Table 1, with the corresponding metric calculation formulae:

Table 1. TDNN model test results.

Performance		Value
MSE	$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$	238412.864
RMSE	$RMSE = \sqrt{MSE}$	488.275
MAE	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert Y_i - \hat{Y}_i\rvert$	363.877
Min Abs Err		0
Max Abs Err		1644
r	$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$	0.975

where i is a natural number variable, n is number of data points, Y_i is the actual observation of time-series, $\hat{Y}_i$ is the corresponding predicted value of time-series. The average hourly observed total energy load value of the test sample (3500 records) was 30019.96 units, while the average total energy load of the whole dataset was 28696.940 units. If the result were to be viewed as relative rather than as absolute values, and if the average value of the test sample were to be taken as the reference, then the results would be as in the Table 2.

Table 2. Relative values of the main TDNN model performance indicators.

Performance	Value
RMSE	1.626%
Relative MAE	1.239%
Max Abs Error	5.476%

DISCUSSION

The results show very high linear correlation (r) of 0.97524 between observed and predicted values, as shown in the Figure 2. This speaks in favour of the fact that TDNN model was being able to adapt with great precision to the observed values of the data set. Moreover, RMSE value of 488.275, and MAE value of 363.877 could be regarded as satisfactory. Within these values however, max. absolute error of 1644 (5.476%) would not necessarily be considered satisfactory, depending on the context of the energy system where prediction is being conducted.

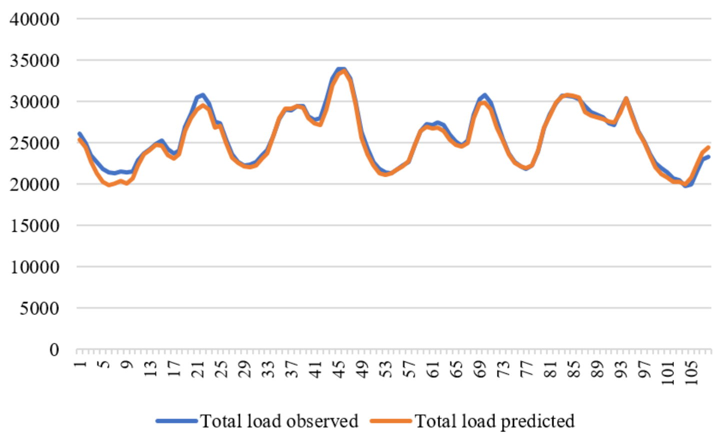

Figure 2. Difference of the predicted and observed values on the test subset.

Interestingly, in the previous research by authors Latinović & Burshaid Al Dhaheri (2022), almost the same TDNN model was experimentally tested, on the same data set. The difference was that, instead of the total energy load, natural gas (NG) consumption data were used. In the experiment, the model was trained to predict NG consumption values based only on historical hourly NG consumption data. However, the obtained results in terms of MAE and RMSE were not satisfactory and authors concluded that it was necessary to induce additional predictors (Latinović & Burshaid Al Dhaheri, 2022). Nevertheless, this difference in the quality of the results of the identical model on two different subsets of the data set can be explained by different characteristics and contexts of the subsets. The total energy load largely depends on the substantially slowly changing average consumption depending on the working time of the industry and the day / night consumption cycle etc., while NG consumption is a representative of an individual resource that can be a substitute for other energy sources when needed, which making the prediction more difficult.

CONCLUSIONS

Considering the proliferation of renewables and the growing complexity of energy systems with the volatility of energy generation and demand, this topic is increasingly relevant, especially in Serbia, where renewable energy sources are expanding. This paper examined a TDNN in energy load prediction using only one predictor, historical total energy load data. A final qualitative assessment of the results under operating conditions cannot be given without taking into account the specific conditions of the given system such as the number of energy production plants, their energy potential, etc. Still, it can be safely concluded that, within the time-series prediction problem, the results in terms of linear correlation, MAE, and RMSE could be regarded as satisfactory. The issue of the maximum absolute error remains. Given the only one predictor used for the prediction, namely historical total energy load data, TDNN showed promising results in both energy load prediction, and generally as a time-series prediction method.

LITERATURE

Amarasinghe, K., Marino, D. L., & Manic, M. (2017). *Deep neural networks for energy load forecasting. 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE).*

Badii, C., Nesi, P., & Paoli, I. (2018). Predicting available parking slots on critical and regular services by exploiting a range of open data. *IEEE Access, 6,* 44059-44071.

Camero, A., Toutouh, J., Stolfi, D. H., & Alba, E. (2018). Evolutionary Deep Learning for Car Park Occupancy Prediction in Smart Cities. *Learning and Intelligent Optimization*, 386–401.

Dagdougui, H., Bagheri, F., Le, H., & Dessaint, L. (2019). Neural Network Model for Short-term and Very-short-term Load forecasting in District Buildings. *Energy and Buildings*, 109408.

Diaz, G. I., Fokoue-Nkoutche, A., Nannicini, G., & Samulowitz, H. (2017). An effective algorithm for hyperparameter optimization of neural networks. *IBM Journal of Research and Development, 61*(4/5), 9:1–9:11.

Gullapalli, S. (2018). Learning to predict cryptocurrency price using artificial neural network models of time series. Retrieved February 10, 2022, from https://krex.k-state.edu/dspace/handle/2097/38867

Janićijević, S., Petrović, Đ., & Stefanović, M. (2020). Sales prediction on e-commerce platform, by using data mining model. *Serbian Journal of Engineering Management, 5*(2), 60-76.

Jha, G. K., & Sinha, K. (2014). Time-delay neural networks for time series prediction: An application to the monthly wholesale price of oilseeds in India. *Neural Computing and Applications*, *24*(3–4).

Kalantari Khandani, M., & Mikhael, W. B. (2021). Effect of Sparse Representation of Time Series Data on Learning Rate of Time-Delay Neural Networks. *Circuits, Systems, and Signal Processing*, *40*(6).

Kaveh, K., Duc Bui, M., & Rutschmann, P. (2017). A comparative study of three different learning algorithms applied to ANFIS for predicting daily suspended sediment concentration. *International Journal of Sediment Research*, *32*(3).

Kermanshahi, B., & Iwamiya, H. (2002). Up to year 2020 load forecasting using neural nets. *International Journal of Electrical Power & Energy Systems, 24*(9), 789–797.

Latinović, L. (2021). *Parking Occupancy Prediction by Recurrent Neural Network Model*. Master‘s thesis. School of Engineering Management, University Union – Nikola Tesla, Belgrade, Serbia.

Latinović, L., & Gerasimović, M. (2021). *Parking occupancy prediction based on artificial neural networks.* University of Novo Mesto, Faculty of Economics and Informatics and University of Novo mesto Faculty of Business and Management Sciences.

Latinović, L., & Tomašević, V. (2022). Rushing Towards Renewables in Serbia – Energy and Environmental Security, and Economic Implications. *In Proceedins 15th Annual International Scientific Conference.* Matej Bel University in Banská Bystrica, Slovakia.

Latinović, L., & Burshaid Al Dhaheri, M., S. (2022). Prediction of Hourly Natural Gas Consumption Using Time Delay Neural Networks. *In Proceedings IV International Scientific* *and Professional Conference Circular and Bioeconomics – CIBEK 22*. Belgrade: School of Engineering Management.

Mammadli, S. (2017). Financial time series prediction using artificial neural network based on Levenberg-Marquardt algorithm. *Procedia Computer Science*, 120.

Menon, A., Mehrotra, K., Mohan, C. K., & Ranka, S. (1996). Characterization of a Class of Sigmoid Functions with Applications to Neural Networks. *Neural Networks, 9*(5), 819–835.

Molina, D., Liang, J., Harley, R., & Venayagamoorthy, G. K. (2011). Comparison of TDNN and RNN performances for neuro-identification on small to medium-sized power systems. *In Proceedings 2011 IEEE Symposium on Computational Intelligence Applications In Smart Grid (CIASG)*.

Qiao, J., Wang, L., Yang, C., & Gu, K. (2018). Adaptive Levenberg-Marquardt Algorithm Based Echo State Network for Chaotic Time Series Prediction. *IEEE Access*, *6*.

Rafsanjani, M. K., & Samareh, M. (2016). Chaotic time series prediction by artificial neural networks. *Journal of Computational Methods in Sciences and Engineering, 16*(3), 599–615.

Šebalj, D., Mesarić, J., & Dujak, D. (2019). Analysis of methods and techniques for prediction of natural gas consumption: A literature review. *Journal of Information and Organizational Sciences*, *43*(1).

Šebalj, D., Mesarić, J., & Pap, A. (2021). Prediction of Natural Gas Consumption by Neural Networks. *In Proceedings Economic and Social Development,* 248-258.

Weigend, A. S., & Gershenfeld, N. A. (1993). Results of the time series prediction competition at the Santa Fe Institute. *In Proceedings IEEE International Conference on Neural Networks*.

Zhang, J., Wei, Y.-M., Li, D., Tan, Z., & Zhou, J. (2018). Short term electricity load forecasting using a hybrid model. *Energy*, 158, 774–781.