BULETINUL INSTITUTULUI POLITEHNIC DIN IAȘI Publicat de Universitatea Tehnică "Gheorghe Asachi" din Iași Volumul 65 (69), Numărul 2, 2019 Secția ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ

# SHORT TERM LOAD FORECAST APPROACH BASED ON ARTIFICIAL NEURAL NETWORKS

ΒY

## BOGDAN-CONSTANTIN NEAGU<sup>\*</sup>, OVIDIU IVANOV, GHEORGHE GRIGORAȘ and MIHAI GAVRILAȘ

"Gheorghe Asachi" Technical University of Iaşi, Faculty of Electrical Engineering, Department of Power Engineering

Received: April 8, 2019 Accepted for publication: May 15, 2019

**Abstract.** The short-term load forecast (STLF) is crucial for distribution network efficient operation, being extensively used in decisions regarding generation and reserve power planning at system level, security monitoring and offline analysis, or market transaction planning. The paper proposes a comparative analysis regarding the performance of the two most used artificial neural networks (ANN) approaches for STLF – the Multilayer Perceptron (MLP) and backpropagation through time (BPTT). The two types of ANN were tested for a real MV/LV substation from Iasi. The results validate the superiority of the BPTT method for power demand forecast and the advantage of the MLP method for the energy load forecast.

Keywords: load forecast; artificial neural networks.

## 1. Introduction

The load forecast accuracy plays a significant role for electricity utilities in ensuring efficient operation, quality of supply and rational development.

<sup>\*</sup>Corresponding author: *e-mail*: bogdan.neagu@tuiasi.ro

Short term forecast usually concern the daily peak load, the load in specific day intervals, the daily or weekly electricity demand. An accurate STLF is crucial for distribution network efficient operation, being extensively used in decisions regarding generation and reserve power planning at system level, security monitoring and offline analysis, or market transaction planning respectively.

The STLF is carried out using an accurate mathematical model, a database that contains measured and recorded load values, and weather data, combined in a user interface application tool. According to the information available in the literature (Bhandari *et al.*, 2018; Kumar & Dixit, 2018; Moon *et al.*, 2018; Gavrilaş, 2002; Neagu, 2014; Olegario *et al.*, 2019; Tian *et al.*, 2018), STLF studies have shown a significant dependence on the electricity demand and the weather conditions, of which the most important being the temperature. Other factors include the non-steady patterns of time series (measurements) and the specific electricity demand occurring in special event days (public or religious celebrations, worker strikes and so on).

In this paper, a comparative study regarding the performance of the two ANN approaches for STLF is performed. For this reason, in the following sections the theoretical aspects and a study case were presented. Thus, in section two are synthetically presented the analysis of time series using statistical methods and artificial neural networks. Section three deals with the STLF problem statement using ANN, and in section four a case study is proposed, where the two ANN methods are explained – Multilayer Perceptron (MLP) and backpropagation through time (BPTT). The simulation results are compared using STLF parameters and a critical discussion are made. Finally, several conclusions are drawn.

## 2. Time Series Analysis Using Statistical Methods and Artificial Neural Networks

The analogy between the statistical models and the artificial neural networks (ANN), showing at the same time how a backpropagation network without hidden layers using linear activation functions leads to results similar to those obtained with an autoregressive moving average (ARMA) model, and, by using sigmoid activation functions and a larger number of neurons, a nonlinear ARMA model (Panapongpakorn & Banjerdpongchai, 2019).

The autoregressive (AR) statistical model can be described with an equation written as

$$y(t) = \phi_1 \cdot y(t-1) + \phi_2 \cdot y(t-2) + \dots + \phi_n \cdot y(t-p) + a(t).$$
(1)

If a third order model is used, then:

$$y(t) = \phi_1 \cdot y(t-1) + \phi_2 \cdot y(t-2) + \phi_3 \cdot y(t-3) + a(t).$$
(2)

The  $\phi_i$  coefficients are estimated using the well-known least squares method, used also for computing weights in backpropagation ANNs. Equation (2) can be represented using an ANN having the structure depicted in Fig. 1.



Fig. 1 – An ANN describing a third order autoregressive model.

The ANN from Fig. 1 does not have hidden layers, and its output neuron uses a linear activation function, combining the inputs. Given the simplifying assumption used to represent an AR process with an ANN, the latter model can be more powerful than a statistical approach. The only difference is that, while the statistical model finds the exact solution using the pseudoinverse matrix, the ANN model needs an iterative algorithm, called backpropagation (Gavrilaş, 2002; Neagu, 2014; Olegario *et al.*, 2019). The weights  $w_1$ ,  $w_2$ ,  $w_3$  will have, after the training stage, the values of  $\phi_1$ ,  $\phi_2$  and  $\phi_3$ .

Mathematically, the moving average model can be formulated as:

$$y(t) = a(t) - \theta_1 \cdot \alpha (t-1) - \theta_2 \cdot \alpha (t-2) - \dots - \theta_q \cdot \alpha (t-q),$$
(3)

and, for a third-order model:

$$y(t) = a(t) - \theta_1 \cdot \alpha(t-1) - \theta_2 \cdot \alpha(t-2) - \theta_3 \cdot \alpha(t-3).$$
(4)

The  $\theta_i$  coefficients from (3) and (4) are estimated using the least squares method used by the ANN backpropagation algorithm, but the recurrence caused by using the error from the previous step as input can lead to major difficulties. For modeling of this process, a recurrent ANN must be used, as illustrated in Fig. 2, using an algorithm that minimizes the squared deviations between the actual and the desired values. This type of algorithm varies sequentially each unknown variable, until the minimal error, value is reached and the iterative process continues until no new solution is found. For obtaining the weights  $w_i$ , they are incremented with values depending on the sign of the first derivative of the error with respect to these weights and on the desired precision of the forecast. The desired solution is found after a much higher number of iterations, while the statistical method uses the faster Seidel-Gauss algorithm.



Fig. 2 - An ANN describing a third order moving average process.

ANNs allow an easy modelling of exogenous inputs, by increasing the number of input neurons. For instance, if an exogenous input variable x(t) delayed with a period must be used, which would lead to a simplified AR process written as:

$$y(t) = \phi_1 \cdot y(t-1) + \phi_2 \cdot y(t-2) + \phi_3 \cdot y(t-3) + w_1 \cdot x(t-1) + a(t)$$
(5)

then the corresponding ANN would be the one depicted in Fig. 2, which shows that exogenous inputs do not make the ANN structure and training algorithm more complicated.

In order to be analyzed using the Box-Jenkins method, the time series must be stationary (Abu & Ismail, 2019). Non-stationary time series must be initially processed in this regard, a task usually performed offline. The previously presented ANNs are reproductions of the Box-Jenkins method and can be used only on stationary time series. It is known that an ANN with hidden layers and sigmoid activation functions can approximate non-linear functions, thus skipping this data-preprocessing step (Farhadi, 2017). The equivalent statistical model is the nonlinear regression with moving average.

# 3. Short Term Load Forecast using Artificial Neural Networks

The last decades of research (Georgescu *et al.*, 2004; Gavrilaş, 2002; Moon *et al.*, 2018; Raza & Khosravi, 2015) show a significant shift towards load forecast methods that use artificial intelligence elements. Artificial intelligence techniques, especially ANN are powerful and flexible tools. MLP-ANNs can be used for the approximated of multivariable time series, which are harder to approach using traditional statistical methods (Fernandes *et al.*, 2019). More than 50% of the ANN models use the backpropagation algorithm, which takes advantage of the ANN ability to discover data correlations without the help of human expertise (Neagu, 2014). The main input information needed for STLF is illustrated in Fig. 3.



Fig. 3 - An ANN describing a third order moving average process.

ANNs are easy to build and train and can memorize complex interdependencies even in the absence of a functional model, being able to find solutions for highly complex nonlinear models (which many of the typical electrical network analysis problems are). For STLF, MLP require an extended database with demand history recordings, from which the training data is extracted. The training data must be constantly updated, for future ANN retraining. This approach can have two disadvantages (Gavrilaş, 2002):

- The initial training data is insufficient.
- The training data does not fully cover the range of the input data, as new models replace old recordings.

Recurrent neural networks (RNN) can handle better the sequential nature of time-dependent series, including the electricity demand (Peng *et al.*, 2019). By using RNNs, the load can be considered as a collection of recordings in which time is irrelevant. Such ANNs are capable of exploiting all the information hidden in the time variation function.

Particular types of RNNs use multilayer architectures with static and dynamic (feedback) neurons. These ANNs proved to be very efficient, because they use concurrently the classification capabilities of multiple layer structures and the time characteristics as inputs. They exhibit other remarkable time-related capabilities, such as: they 'remember' sequences of records from the recent past and can 'forget' information too old to bring useful input information.

#### 4. Study case. Results and Discussions

In this section, two types of ANNs were tested for STLF of the electricity demand at the LV bus of a MV/LV substation belonging to the distribution network that supplies the city of Iasi. First, some considerations regarding the essential steps for building an MLP or Recurrent BPTT for STLF model are given. Any ANN, regardless of its type, requires an input database

13

with recordings taken for an extended time period (years, if possible), which should contain data modeling the load behavior in time and other useful data, especially weather related. The database used in the case study was obtained by continuous monitoring and recording the total electricity demand obtained from the Smart Meters mounted at the LV side of the analyzed MV/LV substation. The substation supplies residential and tertiary (textile factory, hotel, school) consumers. The measurement showed a non-linear load behavior in time. The instantaneous power at time t, P(t), can be represented as a linear combination of four independent components:

$$P(t) = P_{\text{base}}(t) + P_{\text{meteo}}(t) + P_{\text{sp.ev}}(t) + P_{\varepsilon}(t), \qquad (6)$$

where:  $P_{\text{base}}(t)$  is the base power, independent of external factors;  $P_{\text{meteo}}(t)$  represent the power influenced by the weather;  $P_{\text{sp.ev.}}(t)$  is the power associated to special events;  $P_{\varepsilon}(t)$  is a random small component.

Moreover, the weather-dependent power can be separated into other four components (Neagu, 2014):

$$P_{\text{meteo}}(t) = \alpha_1 \cdot T(t) + \alpha_2 \cdot N(t) + \alpha_3 \cdot V(t) + \alpha_4 \cdot H(t)$$
(7)

where: where: T – temperature; N – cloud density; V – wind speed, H – humidity;  $\alpha_i$ , (i = 1...4) – conversion coefficients.

Equations (6) and (7) show that weather conditions have a significant influence on the electricity demand. For choosing the ANN inputs for load forecast, a statistical self-correlation analysis must be performed for load values at different moments in time. It is also of interest the correlation between the load and the temperature or other indices such as the peak load duration, maximum apparent power, loss duration. In order to establish the Pearson correlation coefficients (cc) for two statistical variables x and y, the following equation can be used (Georgescu *et al.*, 2004):

$$cc = \frac{\operatorname{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{xy} - \overline{xy}}{\sqrt{\overline{x^2} - (\overline{x})^2} \cdot \sqrt{\overline{y^2} - (\overline{y})^2}}$$
(8)

where: cov(x, y) is the covariance between x and y, and  $\sigma_x^2$  and  $\sigma_y^2$  is the variance of x and y.

The load recordings database was analyzed considering only two typical scenarios in a year (cold season – winter and warm season – summer), and the weekdays were considered working (Monday – Friday) and weekend (Saturday and Sunday). The correlation coefficients were computed between the active power demand and other quantities of interest, at the LV substation level. Table 1 shows the average self-correlation coefficients between the power to be forecasted P(0h) and other power readings from previous intervals, and the

average correlation coefficients between P(0h) and earlier temperatures and other quantities of interest.

Table 1
Self-correlation Between P(0h) and Another Earlier Power Measurements,
Temperatures and Other Variables

Power	СС	Temperatures	СС	Other	СС
P(-1h)	0.9157	T(-1h)	0.7324	$T_{\rm med \ prev.}$	-0.0016
P(-2h)	0.8206	T(-2h)	0.6905	S <sub>max</sub>	0.0335
P(-3h)	0.6472	T(-3h)	0.6346	$T_P$	0.2165
P(-4h)	0.3017	T(-4h)	0.4201	$ au_P$	0.1981
P(-5h)	0.1823	T(-5h)	0.3897	•	
P(-6h)	-0.1986	T(-6h)	0.2365		
P(-7h)	-0.3092	T(-7h)	0.2117		
P(-8h)	-0.3441	T(-8h)	0.3152		
P(-9h)	-0.3227	T(-9h)	0.3614		
P(-10h)	-0.3165	T(-10h)	0.2498		
P(-11h)	-0.3014	T(-11h)	0.0932		
P(-12h)	-0.2921	T(-12h)	-0.1175		
P(-24h)	0.9972				
P(-168h)	0.9899				
P(-192h)	0.9812				

The analysis of the correlation coefficients from Table 1 yields the following important conclusions:

- The strongest correlations of P(0h) are with the power from the same moment of the previous day P(-24h), same day from the previous week P(-168h), and previous day of the previous week P(-192h).
- There is a strong correlation of P(0h) with the powers for the previous three hours P(-1h), P(-2h) and P(-3h). Beyond this window, starting from hour h 4, the correlation coefficient drops, having a good self-correlation (0.3) on 7,...,12 hours before the forecast time.
- The best correlations between the power and the temperature occur for h-1, h-2 and h-3.

The correlation coefficients with other quantities of interest (average temperature of the forecast day, peak load and peak load duration from the previous day) are relatively small and should not be included in the ANN training data. Such a correlation analysis can easily decide the information included in the ANN training data. Other useful information can be: the week day name (Monday – Sunday) or type (working – weekend) and the forecast hour (1,...,24). The ANN training data used in the case study included the weekday type and the forecast hour.

Based on the results of the statistical correlation analysis, the load data used in the ANN training set consisted, for each input model, of the following values for power: : P(-1h), P(-2h), P(-3h), P(-24h), P(-168h), P(-192h) and

temperature: T(-1h), T(-2h), T(-3h). The day index used a single neuron, with input one for working days (without significand differences) and 0.5 for weekend days. The hour was binary encoded, with five neurons for the 24 hours of a day. Thus, an input model can be described formally using:

$$P(t) = f \left[ P(t - \tau_1), T(t - \tau_2), z, h \right], \tag{9}$$

where:  $\tau_1$  is the load delay ( $\tau_1 = 1, 2, 3, 24, 168, 192$  hours);  $\tau_2$  – the temperature delay ( $\tau_2 = 1, 2, 3$  hours); z – the day of the week; h – the hour of the day.

Each training model contained 15 inputs, as described above, and one output value, which is the forecasted load. The number of neurons of the hidden layer was determined by trial. The hidden and output neurons use the logistic sigmoid activation function. The training data set had 336 input-output pairs, corresponding to 2 weeks of data. The values were scaled in the (0, 1) interval. Data scaling is a very important step, because is leveling the magnitude orders of different measurement units (MWh, Celsius degrees, day index). The Resilient Backpropagation algorithm (RBA) proposed by (Riedmiller & Brown, 1993) was used for training. Fig. 4 shows the PMS architecture used in the case study. Trials determined the optimal number of hidden neurons as 12. The average training SSE for each model dropped after 20,000 training cycles to  $1.8 \times 10^{-5}$ . The forecasting capabilities of the trained ANN were verified with a test data set, for which the absolute mean square error (MSE) was computed as:



The test data set used 168 input-output pairs, corresponding to the week following the interval used for training. For each input model, the

corresponding output was computed through forward propagation and the result was compared to the desired value. The maximum percent error was 3.24%, only two of the 168 models exceeding the 3% threshold.

For the proposed forecast model, the only scenario in which the MLP inputs can use real measured data corresponds to the load forecast for the hour that follows the last available measurement. If a larger forecast window is desired (a day, up to a week), forecasted values need to be used for the ahead intervals, for which no measurements are available. This approach has the risk of accumulating the forecast errors, hour by hour. The tests showed that the accumulated errors do not lead to significant forecast errors. The maximum percentage deviation was 2.78%, only 6 models exceeding the 2% threshold.

Fig. 5 presents the distribution of the percentage MSE for the entire test week, while Figs. 6 show comparisons between the real and the forecasted load for Friday and Sunday, the days with the highest obtained errors. The forecast error is always lower than the 3% considered satisfactory in practice (Abedinia & Amjady, 2016).



Fig. 5 – The percentage error distribution for the test data.



The STLF is a dynamic process. The layered architecture of MLP show

sometimes limited performance in modeling dynamic behaviors which require for the forecast at moment t+1 data from the previous moment, t. Also, the STLF can benefit from using certain connections suggested by previous experience regarding the studied problem, because the active and reactive load is undoubtably a dynamic process. Thus, a feedforward ANN should include feedback connections when the neural network is time-delayed. In this manner, feedforward ANNs can simulate dynamic systems, but only for short time intervals (Costa *et al.*, 1999).

For solving complex problems, previous knowledge must be used in the input data. For daily load forecast (both power and energy) on 24 hourly intervals (the next day), the input data should contain not only the previous day peak load, but also exogenous inputs regarding the type of the day (working, weekend), weather conditions and the hourly load for the previous three hours, which is an example of previous knowledge implementation. Following the aforementioned considerations, a recurrent BPTT was used next for STLF (Fig. 7). This ANN has a single output, for the forecasted load at hour t. For 24-hour forecast, the base ANN from Fig. 7 was unfolded with 24 copies, one for each load profile interval. The result is depicted in Fig. 8. A BPTT is a feedforward network, but it is also modelling a relation between events occurring in different moments in time.



It is very important to define different types, of connections between neurons, with or without time delay. In this way, regardless of the problem complexity, the information is stored in the neuron weights during the training process. Also, in BPTT networks, the user can specify different outputs for each ANN copy. Thus, a general ANN architecture can be built, which can solve complex problems overcoming the limitations of the multilayered structure. In these case study, the same database and statistical correlation analysis were used to build the BPTT from Fig. 9, which uses 10 inputs, as follows:

- The load recorded for given previous hours: P(t-24), P(t-168), P(t-192).
- The type of day (1 working; 0.5 weekend).
- The temperature in the previous three hours T(t-1), T(t-2), T(t-3).
- The load in the previous three hours P(t-1), P(t-2), P(t-3).



Fig. 9 – The base ANN for hourly load forecast.

Based on trial testing, the base ANN had 14 hidden neurons and was unfolded in 24 copies, one for each hour of the day. The same RBA was used for training. For the daily load forecast at LV substation side, the average MSE per model, reached the value of  $2.81 \times 10^{-5}$  after 20,000 training cycles. Although it requires longer training times, the main advantage of the BPTT method is the possibility of making 24-hour forecasts, while the simple MLP can provide results for only 1 hour. The BPTT was tested on the same data and achieved a maximum percent error of 1.72%, smaller than for the MLP error achieved earlier. The error distribution for the entire test set is given in Fig. 10. The real and forecasted 24-hour load curves are compared in Fig. 11 for Sunday, the day with the highest forecast errors.





#### **5.** Conclusions

The short-term load forecast for distribution network represent an important task in the transition to future super grids, as an efficient tool, extensively used in decisions making process regarding the power networks operation, security monitoring and market transaction planning. In the paper, a comparison between the MLP and BPTT approaches, showing the superiority of the BPTT for power demand forecast and the advantage of the MLP method for the energy load forecast was tested. The proposed ANN approach shows that both methods have prediction errors in the accepted range, and con be used with success for STLF of active power or energy demand or consumptions.

#### REFERENCES

- Abedinia O., Amjady N., Short-Term Load Forecast of Electrical Power System by Radial Basis Function Neural Network and New Stochastic Search Algorithm. Int. Trans. on Elec. En. Syst., **26**, 7, 1511-1525 (2016).
- Abu N., Ismail Z., A Study on Private Vehicle Demand Forecasting Based on Box-Jenkins Method, In AIP Conf. Proc., 2059, 1, 020005, 2019.
- Bhandari B., Shakya S.R., Jha A.K., *Short Term Electric Load Forecasting of Kathmandu Valley of Nepal using Artificial Neural Network*, Kathford Journal of Engineering and Management, **1**, *1*, 43-48 (2018).
- Costa M., Pasero E., Piglione F., Radasanu D., *Short Term Using a Synchronously Operated Recurrent Neural Network*, Int. Joint Conf. Neural Networks, IJCNN '99, 5, 1999, 3478-3482.
- Farhadi F., *Learning Activation Functions in Deep Neural Networks*, Ph.D. Thesis, École Polytechnique de Montréal, 2017.
- Fernandes K.C., Sardinha R., Rebelo S., Singh R., *Electric Load Analysis and Forecasting Using Artificial Neural Networks*, In 2019 3<sup>rd</sup> Int. Conf. on Trends in Electr. and Inf. (ICOEI), pp. 1274-1278, 2019.

- Gavrilaș M., Artificial Intelligence and Application in Power Engineering (in romanian), Ed. "Gheorghe Asachi", Iași, 2002.
- Georgescu G., Varvara V., Curteanu C., *Utilizarea rețelelor neuronale artificiale pentru prognoza sarcinilor electrice*, Analele Universității din Oradea, Energetică, **9**, 2004.
- Kumar V.R., Dixit, P., Artificial Neural Network Model for Hourly Peak Load Forecast, Int. J. of Energ. Ec. and Policy, **8**, 5, 155-160 (2018).
- Moon J., Kim Y., Son M., Hwang E., *Hybrid Short-Term Load Forecasting Scheme* Using Random Forest and Multilayer Perceptron, Energies, **11**, 12, 3283 (2018).
- Neagu B., Contributions Regarding the Optimization of the Structure and Operating Regimes of Electricity Distribution and Distribution Systems, Ph.D. Diss., "Gh. Asachi" Technical University, Iaşi, 2014.
- Olegario C.C., Coronel A.D., Gerardo B.D., Medina R.P., *Towards an Enhanced Backpropagation Network for Short-Term Load Demand Forecasting*, Int. J. of El. Eng. and App. Sci., **2**, *1*, 65-70 (2019).
- Panapongpakorn T., Banjerdpongchai D., Short-Term Load Forecast for Energy Management Systems Using Time Series Analysis and Neural Network Method with Average True Range. In 2019 First Int. Symp. on Instrum., Control, Artif. Intell., and Robotics (ICA-SYMP), pp. 86-89, 2019.
- Peng Y., Wang Y., Lu X., Li H., Shi D., Wang Z., Li J., Short-Term Load Forecasting at Different Aggregation Levels with Predictability Analysis, arXiv preprint arXiv:1903.10679, 2019.
- Raza M. Q., Khosravi A., A Review on Artificial Intelligence-Based Load Demand Forecasting Techniques for Smart Grid and Buildings, Renew. and Sust. En. Rev., 50, 1352-1372 (2015).
- Riedmiller M., Braun H., A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm, Proc. of Int. Conf. on Neural Networks (ICNN 93), 1993, 586-591.
- Tian C., Ma J., Zhang C., Zhan P., A Deep Neural Network Model for Short-Term Load Forecast Based on Long Short-Term Memory Network and Convolutional Neural Network. Energies, **11**, 12, 3493 (2018).

#### PROGNOZA PE TERMEN SCURT A SARCINII UTILIZÂND REȚELE NEURONALE ARTIFICIALE

#### (Rezumat)

Prognoza sarcinii pe termen scurt este o problemă esențială pentru funcționarea eficientă a rețelei electrice de distribuție, fiind utilizată pe scară largă în deciziile privind planificarea puterii de generare și a rezervei la nivel de sistem, la monitorizarea securității și analiza offline, precum și la planificarea tranzacțiilor de piață de energie. Lucrarea propune o analiză comparativă cu privire la performanța a două din cele mai utilizate rețele neuronale artificiale (RNA) pentru prognoza pe termen scurt a sarcinii, și anume Perceptronul multistrat (PMS) și a rețelelor recurente cu propagare înapoi în timp (BPTT). Cele două tipuri de RNA au fost testate pentru un post real de transformare (MV/LV) din Iași. Rezultatele obținute arată superioritatea metodei BPTT pentru prognoza cererii de energie și validează avantajul utilizării metodei PMS pentru prognoza energiei consumate.