IOP Conf. Series: Journal of Physics: Conf. Series **881** (2017) 012036 doi:10.1088/1742-6596/881/1/012036

Evaluation and prediction of solar radiation for energy management based on neural networks

O V Aldoshina[1,2] and Dinh Van Tai[2]

[1] Professor, Institute of Non-Destructive Testing, Tomsk Polytechnic University, Tomsk, Russia
[2] Postgraduate Student, Institute of Non-Destructive Testing, Tomsk Polytechnic University, Tomsk, Russia

E-mail: niipp@inbox.ru

Abstract. Currently, there is a high rate of distribution of renewable energy sources and distributed power generation based on intelligent networks; therefore, meteorological forecasts are particularly useful for planning and managing the energy system in order to increase its overall efficiency and productivity. The application of artificial neural networks (ANN) in the field of photovoltaic energy is presented in this article. Implemented in this study, two periodically repeating dynamic ANS, that are the concentration of the time delay of a neural network (CTDNN) and the non-linear autoregression of a network with exogenous inputs of the NAEI, are used in the development of a model for estimating and daily forecasting of solar radiation. ANN show good productivity, as reliable and accurate models of daily solar radiation are obtained. This allows to successfully predict the photovoltaic output power for this installation. The potential of the proposed method for controlling the energy of the electrical network is shown using the example of the application of the NAEI network for predicting the electric load.

1. Introduction

Estimation and forecasting of renewable energy sources is of great importance for ensuring optimal exploitation of available resources and estimating the capacity of the potential output of power plants. This is especially advantageous in the intellectual context of the network, where optimal energy management can be achieved through a complete set of data and forecast data [1,2].

The forecast of methodological time series is particularly advantageous, for example, for energy suppliers who can use such tool to plan and manage the energy system in order to improve its overall efficiency and productivity. However, the forecast is based on a long data-sampling period. The data set may be corrupted due to lack of data or data errors in some time intervals. For this reason, a program for restoring data (estimates) is advantageous.

Therefore, photoelectric modeling of solar radiation by means of estimating time series and forecasting methods is becoming increasingly popular. The ANN forecasting models are used for estimating the average monthly and daily total solar radiation on a horizontal surface and predicting hourly solar radiation, respectively. Ikdur and Zeroual proposed the use of a Multilayer Perceptron (MLP) of the neural network to predict daily solar radiation data. Yona and the others proposed to predict the output power of photovoltaic systems based on day-ahead solar radiation forecasting using three different models based on ANN: direct transmission of the neural network (DTNN), the radial basis of neural network functions (RBNNF), and the current neural network (CNN). They decided to demonstrate that it is possible to predict the results using only meteorological data in a short period of time[3].

The authors proposed the use of dynamic recurrent ANN for forecasting of solar radiation in order to evaluate the methodology. When comparing ANN algorithms, based on linear and non-linear models, to solve the problems of filling missing data and time series, the latter approach turned out to be more profitable.

As a result, ANN proved to be more effective than its analogs. On this basis, this article proposes the use of two dynamic neural networks, i.e., the focused time delay of a neural network (FTDNN) and the nonlinear autoregression of a network with exogenous inputs (NAEI of the network) that receive an estimate and a daily forecast of solar radiation. The ANNs, implemented in this study, are particularly suitable for estimating time series, and show higher productivity for classical direct ANN communication in a similar application.

The geographical location considered in this study is located in Central Kazakhstan. The obtained models are experimentally confirmed. As a result, an example of an application, where the NAEV network is used to perform an electric load forecast, is presented to highlight the utility of the proposed approach to the net of the optimal design and energy management. In particular, the forecast for the production of photovoltaic energy and loads can be used to optimize the choice of the size of renewable sources / storage systems and energy development.

2. Selected intelligent neural networks

The selected INN are focused on the time delay of the neural network (CTDNN) and the non-linear autoregression of the network with exogenous inputs (NAEI of the network). They can be classified as periodic and dynamic INN. The specificity of such neural networks of observing the static direct connection of networks, as the reverse propagation of the error (RPE) of the INN or cascade forward INN is their ability to prepare the dynamics of the ratio of time series. In particular, in dynamic INN, the output depends not only on the current input, but also on the current and previous inputs, outputs or network status [4,5].

The mechanism of adaptive adjustment, on which dynamic INN are based, ensures high accuracy of the model. The description of the INN used is described below.

2.1. Focused time delay of a neural network (FTDNN)

FTDNN is the simplest dynamic Network; it consists of a direct network with a delay line that has taps at the input. This is a part of the general class of dynamic networks, called oriented networks, in which dynamics appear only at the input of a layer of a static multilayer network with a direct link.

This network is well suited for estimating time series. FTDNN can be prepared in order to perform an assessment either one-step ahead or a few steps ahead. In the latter case, evaluations are fed back to the network entrance, continuing to iterate. The structure of the FTDNN is presented in Figure 1.

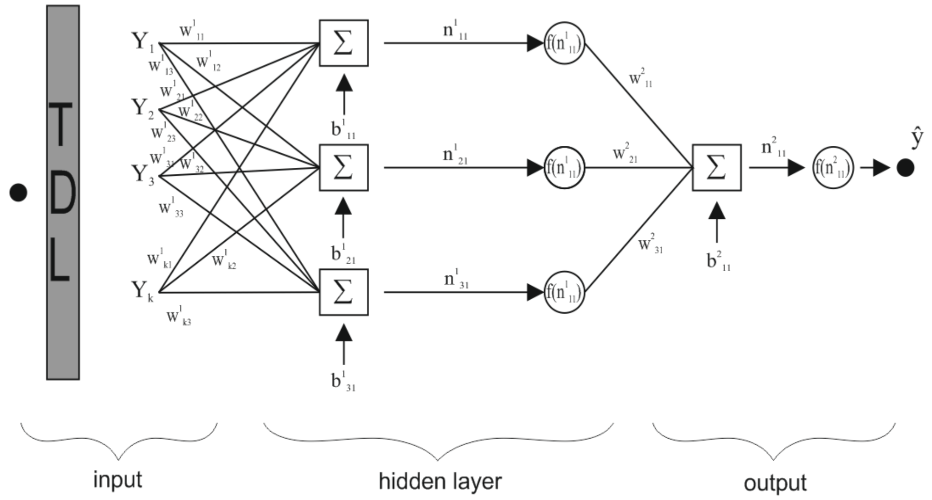

Figure 1. Structure of the FTDNN.

The network was prepared by a conjugate spread with updates of the Polak-Ribiera (described in Section 2.3) and the initial weight was randomly extracted from a standardized normal distribution. The value of the obtained velocity, initially equal to 0.01, and decreases in accordance with the exponential law. The preparation and the recall of the phase were compared using the statistical indicators in order to determine the best model.

2.2 Nonlinear autoregression with exogenous network inputs (NAENI)

NAEI network, in contrast to a focused network, is a constant dynamic network with feedback connection of several network layers. The NAEI model is based on the linear model (LM), it is well suited for modeling nonlinear dynamic systems and is usually used in the form of a temporal sequence of modeling, due to its adaptive preparation process at the small scales of meteorological data, collected, for example, in less than one year. This can be mathematically represented as

$$y(t) = f(y(t-1),...,y(t-ny); u(t-1),...,u(t-nu)) + \varepsilon(t) \tag{2.1}$$

where: y (t) and u (t) the output and input of the model to a discrete step in time t, respectively, while ny ≥ 1, nu ≥ 1, nu ≤ ny, are the input and output memory elements and ε (T) - noise, as a rule, is assumed to be Gaussian or white.

In this network, the following value: the dependent output signal y (t) is the regression on the previous values of the output signal and the previous values of the independent (exogenous) input signal. Neural network with direct connection, such as a standard multi-layer perceptron (SMP), can be used to display the approximate nonlinear function f.

The output of the NAEI network is fed back to the input of the direct neural network as part of the standard NAEI architecture. Since the corresponding output is available during network preparation, it is possible to create a sequential-parallel structure in which the output is used instead of the feeding on the reverse output. This gives two advantages. First, the input of the direct network is more accurate. Secondly, as a result, the network has a structure with a direct link, and static reverse propagation can be used for preparation.

The application of the NAEI network, considered in this article, where the exogenous input is a temperature, is realized with the help of two configurations: a sequential-parallel and parallel structure. The former is used to study and evaluate the behavior of solar radiation, and the latter is used for prediction of solar radiation (multistage prediction).

In other words, the first configuration is used for the same purposes as the FTDNN and reports an estimate of the radiation in the absence of the data. The second configuration is used in the cascade with the first one to perform the radiation forecast. These configurations are shown in Figure 2:

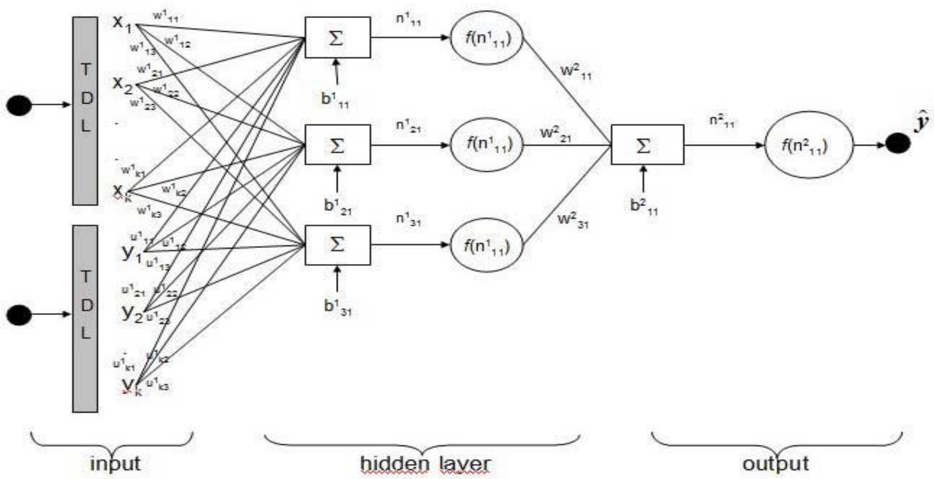

Figure 2. Structures of the NAEI network.

In addition, in this case, an algorithm of obtaining the conjugate gradients of backward propagation with Polac-Ribiera updates and the value of receiving rate, initially equal to 0.01, are used.

2.3 The Algorithm of Preparation

The method of preparation for both networks is the conjugate gradient method of backward propagation with the Polak-Ribiera algorithm. This algorithm was chosen because it converges faster than other more common algorithms, such as gradient descent, and gradient descent with momentum.

The basic algorithm of the error's reverse propagation regulates the weight of the steepest descent (negative gradient). In this direction, the productivity function decreases faster. However, it is not necessary to produce fast convergence.

In conjugate gradient algorithms, the search is performed along the conjugate directions, which are usually made faster than the convergence of the steepest descent of the direction.

The reverse propagation algorithm of the error is used to calculate the derivatives of the error's square (E) with regard to the weight and the displacement variable x (formula 2).

$$x_{n+1} = x_n + \alpha_n \Delta x_n \qquad (2.2)$$

Where the function f (x) given by n variables, it can be minimized Δx = -∇x f (x) and indicates the direction of the maximum magnification, i.e. the direction of the search.

The parameter α (k) is chosen to minimize the productivity in the direction of the search. Thus, the first direction of search is opposite to the gradient of the representation:

$$S_0 = \Delta x_0 = -\nabla x F(x_0) \qquad (2.3)$$

In subsequent repetitions, the search direction is calculated from the new gradient of the previous direction search, according to the formula

$$s_n = \Delta x_n + \beta n s_n - 1 \quad (n \geq 1) \qquad (2.4)$$

The parameter βn can be calculated in several different ways.

The change of the conjugate gradient of Polac-Ribiere is calculated by the formula:

$$\beta_n^{PR} = \frac{\Delta_x^T - (\Delta x_n - \Delta x_{n-1})}{\Delta_x^T \Delta x_{n-1}} \quad (2.5)$$

3. Data Measurement and Pre-processing

Vertical solar-powered facades will produce relatively more power during winter (and less during summer) and more during the hours of dawn and dusk when the sun is low. As a rule, a building will have at least two if not four open facades with opposing orientation and that is why different sides will give the maximum power output during different times of day. This effect leads to an increase of peak energy production during the day which allows to adapt for variable energy demand and thus saves electric energy, as energy generated by non-renewable resources is used less. As an illustration of hourly function of solar-powered facades: Fig. 3 compares typical energy distribution with hourly output of optimized complex of solar powered roofs and facades during winter and summer periods in Almaty. Table 2 shows the results of relative power distribution on different surfaces. The daily solar radiation and daily maximum and minimum temperatures recorded between 2007 and 2010 in Kazakhstan weather stations are used as the data set for the preparation of INN [6].

As for the temperature used to determine the model based on the NAEI network, the corresponding time series was pre-treated to fill the missing values on the characteristic area of the reverse distance algorithm. Figure 3 shows the time series of the daily solar radiation data recorded for 1 year.

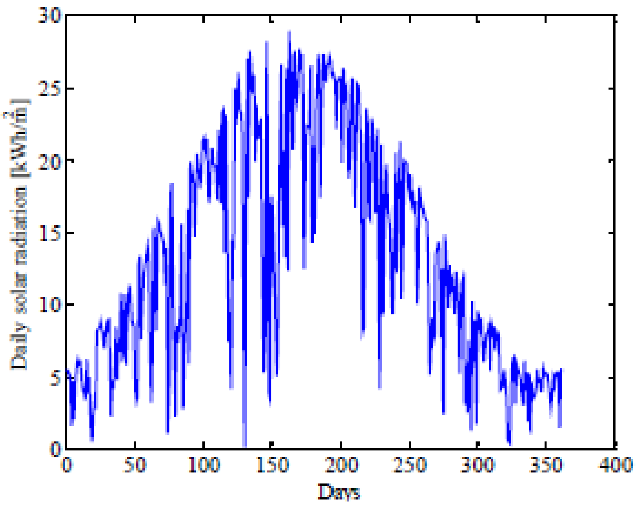

Figure 3. Daily solar radiation recorded during the year.

Two sets of data were used to prepare the two INNs. In particular, for FTDNN, a set of preparation and validation systems are built from the measurements over 2 years, while measurements for the third year remain at the testing stage[7,8].

In the context of the NAEI network, the preparation set is formed from the time series of daily solar radiation and time series of the temperature in the period 2007-2010. The data of the last 150 days was excluded from the prepared data set. This data was used to test the network as a forecast model.

In order to make the preparation of INN more efficient, a pre-treatment procedure was applied to the temperature and solar radiation data.

In particular, the normalization step was applied to both input vectors and target vectors in the data set and network output, then back transformed into the original output units (post-processing procedure).

IOP Conf. Series: Journal of Physics: Conf. Series **881** (2017) 012036 doi:10.1088/1742-6596/881/1/012036

4. Evaluation and forecasting of the solar radiation: results

The Normalized Mean-Square Error (NMSE) and the coefficient of variation of Mean-Square Error CV (MSE) are used to determine the deviation between observed and estimated values. They are defined as:

$$\text{NMSE} = \frac{1}{Y_{\max} - Y_{\min}} \sqrt{\sum_{i=1}^{n} \frac{(Y_i - \bar{Y})^2}{N}} \tag{4.1}$$

$$\text{CV(NMSE)} = \frac{1}{Y} \sqrt{\sum_{i=1}^{n} \frac{(Y_i - \bar{Y})^2}{N}} \tag{4.2}$$

Where Y is the original time series, Yi are the calculated time series, Ymax, min are the maximum and minimum observable values and represent the average observed values. With regard to the first use of the FTDNN, three network structures were compared, with different number of neurons in the hidden layer, for different values of the input delays. The chosen criteria for measuring the productivity of neural networks is represented by NMSE, CV (MSE), as shown in Table 1. In particular, in Table 1 are shown the values of MSE and CV both in the preparation phase (MSEP, CVP) and inverse phases (RMSEr, CVR) [9].

The best network is the one in which there is a minimum deviation between the indices of the two phases. From the point of view of the generalization of the ability, the more similar the indexes, the more executive is the network. Based on this consideration, the structure 1-3-1 with a time delay of 4 days is the best for the neural network. For this configuration, the following metrics have values:

MemSet = 0.17, nEMSEr = 0.17, CVT = 0.38 u CVR = 0.37

The comparison between the measured (initial targets) and the calculated data (network forecasts) is shown in Figure 4. The NAEI network is implemented and previously described by two configurations (sequential-parallel and parallel). The sequential-parallel configuration is used to perform a series of daily estimates of solar radiation, and a parallel is used to perform a forecast of the same data. In this case the MSE, CV indicators, the deviation between the indicators in the preparation and forecast phases are taken into account for estimation of the method.

As can be seen from Table 2, the best configuration in this case is the 2-10-1 configuration with a time delay of 4 days with the following values of the statistical indices: nRMSEt = 0.14, nRMSEf = 0.18, CVT = 0.31 and CVF = 0.48

Figures 5 and 6 show the estimation productivity of the daily solar radiation on the basis of the NAEI network model for forecasting its future trends.

In addition, the NAEI network provides a further advantage for ensuring the extraction of missing data of time series of solar radiation to predict future trends from the same values.

Figure 7 shows the error of the autocorrelation function of the section for the model based on the NAEI network. It is used for further confirmation of the network productivity. In particular, it shows that the network was correctly prepared from the moment of correlation, except that with the zero delay, in 95% of cases the confidence intervals drop to zero.

Table 1. Evaluation of the radiation model based on the FTDNN model.

ANN	Deley	nRMSEt	nRMSEf	CVt	CVf
1-3-1	2	0.19	0.18	0.42	0.39
1-3-1	4	0.17	0.17	0.38	0.37
1-3-1	6	0.16	0.18	0.35	0.40
1-3-1	8	0.16	0.18	0.36	0.39

1-5-1	2	0.18	0.19	0.41	0.41
1-5-1	4	0.17	0.18	0.36	0.39
1-5-1	6	0.15	0.21	0.32	0.46
1-5-1	8	0.15	0.20	0.33	0.44
1-10-1	2	0.18	0.19	0.40	0.41
1-10-1	4	0.15	0.18	0.32	0.40
1-10-1	6	0.13	0.21	0.29	0.46
1-10-1	8	0.12	0.29	0.26	0.63

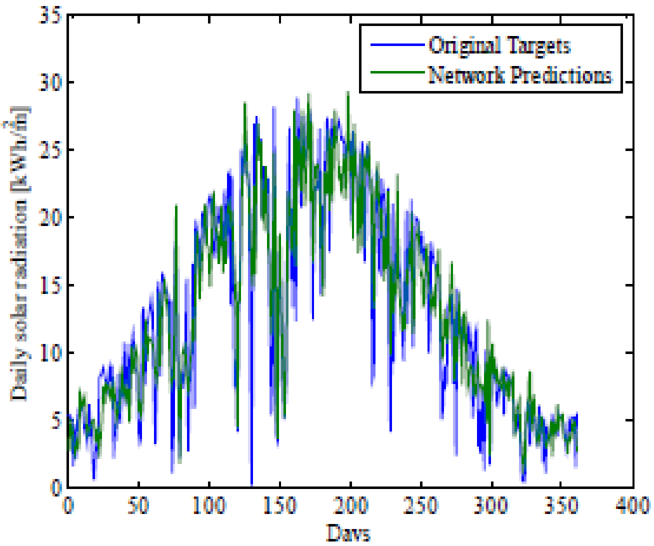

Figure 4. Measured (initial indicators) and calculated (network forecasts) data based on the FTDNN model.

Table 2. Evaluation of the radiation model based on the model of the NAEI network.

ANN	Delay	nRMSEt	nRMSEf	CVt	CVf
2-3-1	2	0.15	0.21	0.34	0.56
2-3-1	4	0.14	0.22	0.33	0.58
2-3-1	6	1.14	0.20	0.32	0.54
2-3-1	8	0.14	0.33	0.33	0.88
2-5-1	2	0.15	0.20	0.34	0.54
2-5-1	4	0.14	0.22	0.33	0.59
2-5-1	6	0.14	0.23	0.32	0.61
2-5-1	8	0.14	0.22	0.32	0.60
2-10-1	2	0.15	0.19	0.33	0.51
2-10-1	4	0.14	0.18	0.31	0.48
2-10-1	6	0.13	0.19	0.30	0.52
2-10-1	8	0.13	0.23	0.30	0.61

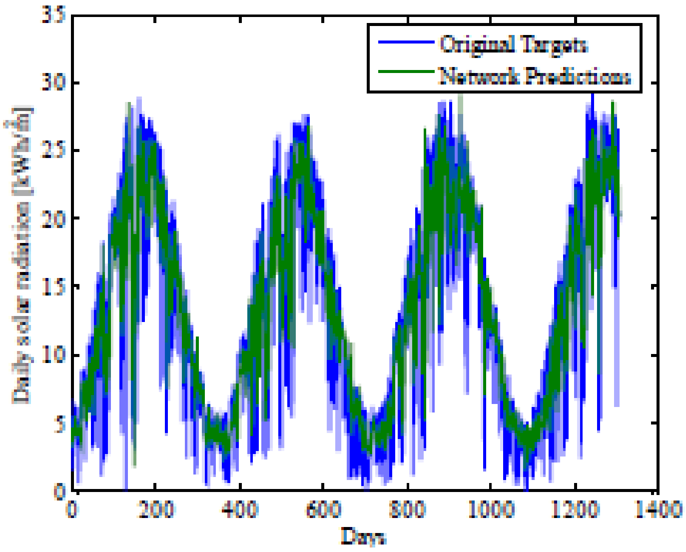

Figure 5. Evaluation of solar radiation based on the model of the NAEI network.

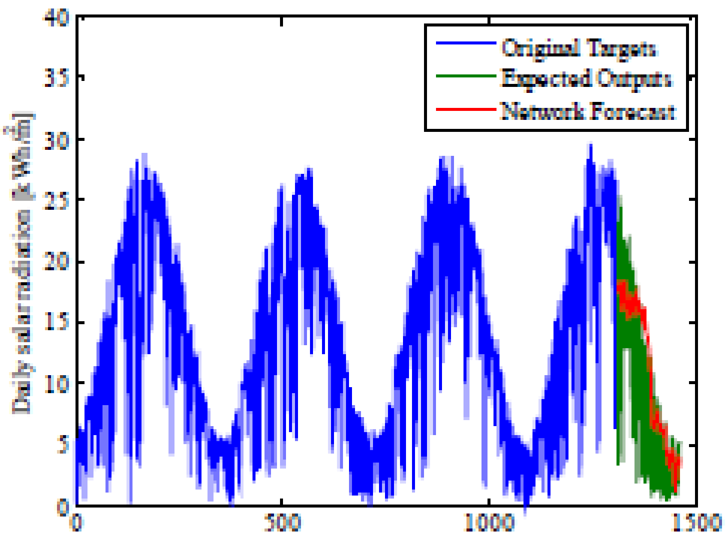

Figure 6. Forecasting of solar radiation based on the model of the NAEI network.

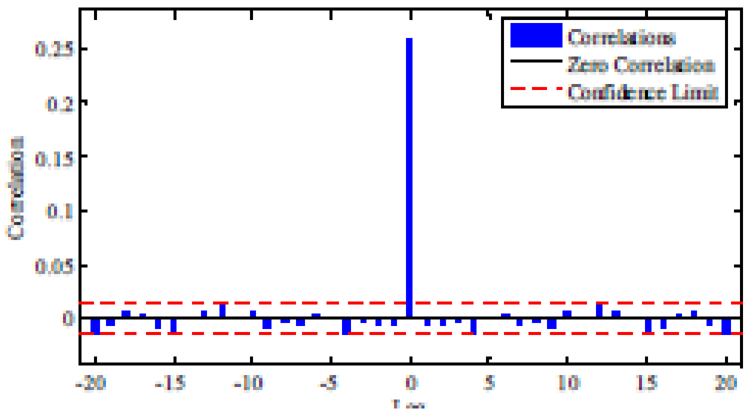

Figure 7. The error of the autocorrelation function of the model based on the NAEI network.

IOP Conf. Series: Journal of Physics: Conf. Series **881** (2017) 012036 doi:10.1088/1742-6596/881/1/012036

5. Example of the application based on network power management

In order to emphasize the potential of the proposed method of optimal size / design of energy management of electrical networks, the NAEI network was used to evaluate and predict the electrical load of a residential block. In this case, the daily temperature and the daily energy consumption during the year are considered the input variables.

For these purposes, on the whole, the daily energy consumption of a residential unit within a year make up the entire electrical load. It should be noted that the load profile, as shown in the Figure 8, takes into account the seasonal variations with maxima in the winter and summer periods. The small peaks, imposed on the annual trend, are determined by higher load power consumption over the weekend.

The best model supplied by the NAEI network was selected on the basis of statistical indices, in accordance with the values recorded in Table III that were received during the preparatory phase. In particular, model 2-10-1 with a time delay of 8 days provides better performance. Parallel configuration of the NAEI is used to obtain a load forecast for 15 days. Figure 8 shows the load power estimate applied to the observed profile. Figure 9 shows the forecast of the power load. It can be noted that the NAEI network can effectively perform both the estimation and the forecast of the residential power unit of the load.

As previously assumed, in the case of a subject of even illuminated photoelectric field, solar radiation can be associated with electricity produced by a photovoltaic system if its characteristics are known. Starting with observing and evaluating data (solar radiation and load consumption), it is possible to estimate the difference between the output power of the photovoltaic system and the power consumption of electrical loads (Delta P = PPV-PL) [10].

Functions for photovoltaic supply station (residential unit is considered):
- Rated power, P = 1.5 kW;
- The surface of the photoelectric area = 25.92 m2;
- Photoelectric efficiency, η = 0.1.

The output power of the photovoltaic system receives solar radiation using the following relation:

$$P_{pv} = \frac{1000}{24} GS\eta[1 - k(T_C - 25)] \tag{5.1}$$

Where, G solar radiation is expressed in kWh / m2,

TS - Operating temperature of the photoelectric array in ° C,

K = 0.005

Temperature coefficient and PPV output power of the photovoltaic system is expressed in W [14-15]. Trends in the variation of Delta P received from observed and calculated values of solar radiation and load power are shown in Figure 10. When calculating Delta P, the effect of the operating temperature of the array can be neglected.

A good coordination between the two curves is observed in Figure 10. This indicates that the proposed method is an effective tool for predicting the energy balance in electrical networks. Therefore, it can be useful for optimizing the size of renewable sources / storage systems and developing an energy management system (EMS) strategy.

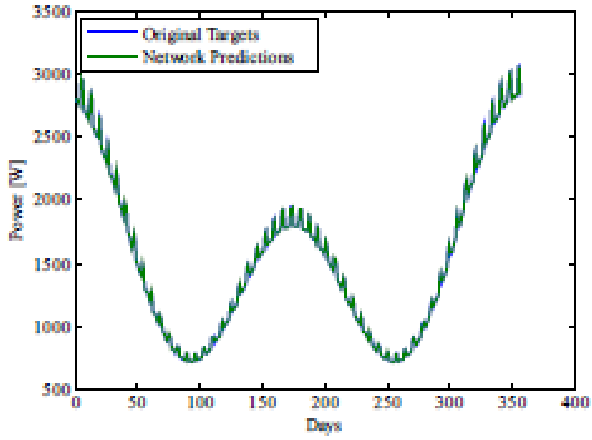

Figure 8. Estimating the load capacity of the model based on the NAEI network.

Table 3. Load estimation based on the NAEI network model.

ANN	Delay	nRMSEt	CVt
2-3-1	2	0.0309	0.0460
2-3-1	4	0.0275	0.0410
2-3-1	6	0.0118	0.0175
2-3-1	8	0.0040	0.0059
2-5-1	2	0.0305	0.0405
2-5-1	4	0.0213	0.0317
2-5-1	6	0.0244	0.0363
2-5-1	8	0.0027	0.0040
2-10-1	2	0.0189	0.0430
2-10-1	4	0.0213	0.0317
2-10-1	6	0.0127	0.0189
2-10-1	8	0.0025	0.0037

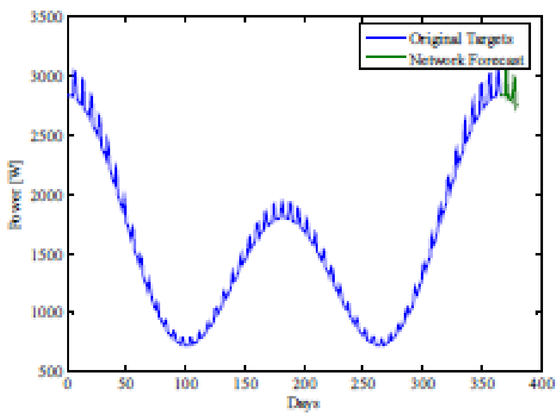

Figure 9. Forecast of the power load of the model based on the NAEI network.

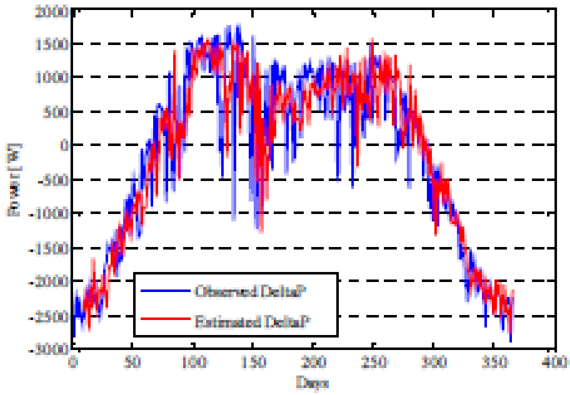

Figure 10. The comparison of the observed and calculated difference between the output power of the photovoltaic system and the power of consumed electrical load.

6. Conclusions

In this article, two appropriately prepared dynamic periodic INNs, in particular FTDNN and the NAEI network, are used to model daily solar radiation, as well as for estimation and prediction. The considered geographical area is located in Central Kazakhstan. The models based on INN have been experimentally confirmed: they both show good productivity, since the reliable and accurate representations of daily solar radiation have been obtained.

In the subsequent approach, a complete set of data can be achieved, even in the absence of data due to sensor failures or maintenance. In addition, due to the use of forecasting capabilities, it is possible to determine a suitable energy schedule along with the optimal calibration at various sections of the electrical network.

References

[1] Paoli C, Voyant C, Muselli M and Nivet M L 2010 Forecasting of preprocessed daily solar radiation time series using neural networks *Solar Energy* **84(12)** 2146-60
[2] Mellit A, Kalogirou S A, Hontoria L and Shaari S 2009 Artificial intelligence techniques for sizing photovoltaic systems: A review *Renewable and Sustainable Energy Reviews* **13-2** 406-19
[3] Mubiru J and Banda E 2008 Estimation of monthly average daily global solar irradiation using artificial neural networks *Solar Energy* **82-2** 181-7
[4] Navon I M and Legler D M 1987 Conjugate-Gradient Methods for Large-Scale Minimization in Meterorology *American Meteorological Society* **115** 1479-1502
[5] Eccel E, Cau P and Renzi R 2012 Data reconstruction and homogenization for reducing uncertainties in high-resolution climate analysis in Alpine regions *Theoretical and Applied Climatology* DOI 10.1007/s00704-012-0624-z
[6] Yurchenko A, Zotov L, Jugaj V, Tatkeeva G and Mekhtiev A 2015 Power supply of autonomous systems using solar modules *IOP Conference Series: Materials Science and Engineering* **81** 012112
[7] Di Piazza A and Eccel E 2012 *Analisi di serie giornaliere di Temperatura e Precipitazioni in Trentino nel periodo 1958-2010* (Provincia Autonoma di Trento)
[8] Di Piazza M C and Vitale G 2010 Photovoltaic field emulation including dynamic and partial shadow conditions *Applied Energy* **87** 814-23
[9] Di Piazza M C and Vitale G 2013 *Photovoltaic Sources: Modeling and Emulation* (Springer, London)
[10] Yurchenko A, Syrjamkin V, Okhorzhina A and Kurkan N 2015 PV effectiveness under natural conditions *IOP Conference Series: Materials Science and Engineering* 81 012097