DAILY INFLOW FORECASTING USING ARTIFICIAL NEURAL NETWORKS FOR THE LARGEST ROMANIAN RESERVOIR

Angela NEAGOE [1], Eliza-Isabela TICĂ [2, *], Florica POPA [3], Bogdan POPA [4]

For water resources planning and management, it is important to develop a forecasting model with high accuracy, based on the historical inflows in the reservoir and other environmental factors including temperature and precipitation. The relationship that shapes the dependence between these variables in the model is non-linear, non-stationary and stochastic and improving inflows prediction is an intense topic approached by researchers. Artificial Neural Networks (ANNs) have shown strong potential in capturing complex patterns within these data, enabling highly accurate forecasts.

In this paper, a neural network three-layer feedforward back-propagation with Bayesian regularization learning algorithms was developed in MATLAB language programming. An ANN with eight input parameters, trained over the period 2005–2019, was selected to test its ability to forecast daily reservoir inflows for the following two years (2020–2021) using observed meteorological and hydrological inputs that were excluded from training. The aim of this test was to evaluate the model's predictive performance and generalization capability beyond the training period. The forecasted inflows showed very good agreement with the observed data.

The best-performing configuration (M8), which included lagged inflows, precipitation, temperature, and day index as inputs, achieved an R^2 of 0.9224, RMSE of 11.83 m^3/s, during the validation period. In contrast, a simpler model (M1), which used only the inflow from the previous day (lag 1), reached an R^2 of 0.8954 and RMSE of 13.69 m^3/s. These results confirm the effectiveness of the proposed ANN approach in forecasting reservoir inflows with high accuracy.

Keywords: ANN, inflows, reservoir, one day ahead forecasting.

1. Introduction

Reservoirs can have numerous uses, such as water supply for the population and industry, irrigation, fish farming, recreation, flood prevention, energy

[1] Lect., Dept. of Hydraulics, Hydraulic Machinery and Environmental Engineering, NUST Politehnica Bucharest, Romania, e-mail angela.neagoe@upb.ro
[2] * Lect., Dept. of Hydraulics, Hydraulic Machinery and Environmental Engineering, NUST Politehnica Bucharest, Romania, corresponding author, e-mail: eliza.tica@upb.ro
[3] Assoc. Prof., Dept. of Electronics, Telecommunication and Energy Engineering, University Valahia of Târgoviște, Romania; icapopa10@gmail.com
[4] Prof., Dept. of Hydraulics, Hydraulic Machinery and Environmental Engineering, NUST Politehnica Bucharest, Romania, e-mail bogdan.popa@upb.ro

generation. Optimal management of water resources requires strategic planning, which involves forecasting the inflows into the reservoir [1, 2]. The temporal variation of the water volumes in the reservoir is primarily driven by meteorological factors.

The classical statistical models dedicated to forecasting are developed for stationary time series and assume a linear dependence between the current variable and its past values. The time series of the daily or monthly average flows are non-stationary and must be processed to be used in an autoregressive forecasting model. Thus, models such as linear regression (LR), autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA) [3] have been developed. Among the statistical models, ARIMA is the most suitable for modeling non-stationary time series because it includes differentiation (once or several times, as the case may be) for transforming the series into a stationary series [4, 5, 6]. These models provide very good results especially for short and medium-term forecasts. In addition, to better model the time series that present periodicity due to the alternation of the seasons, seasonal ARIMA was developed, for captures both short-term and long-term dependencies within the data [7].

Artificial Neural Networks (ANNs) are part of the broader category of artificial intelligence techniques. These models can identify nonlinear relationships between a variable and its historical values within data series, which can be leveraged to forecast future variables. Their effectiveness has been demonstrated in various hydrological applications [8, 9, 10]. In order to improve the accuracy of short-term load forecasting, different neural network architectures such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU) were explored in [11].

The performance of forecasting models based on ANNs depends not only on the architecture of the network [12], but also on the selection of input variables and the specific time delays associated with them, which influence the predicted values [13, 14, 15, 16]. Reference [17] provides a review of studies that have applied ANNs, Support Vector Machine, Fuzzy and Evolutionary computation, in predictions in hydrology as well as studies that have made predictions according to different input variables: streamflow, precipitation, evaporation, temperature and time scale in streamflow forecasting.

Comparisons between ANN and traditional statistical models are made in many papers [8, 10, 18, 19, 20], with the conclusion that artificial intelligence provides better results for daily inflows forecasting using an artificial neural network based on past patterns and meteorological data.

The predictions of a certain ANN model are very sensitive to each parameter of the network structure. For example, by training several times a network in which only the initial weights are modified, slightly different results are obtained [21].

The network training method is also very important, Levenberg-Marquardt being the commonly used optimization method. Also, the Bayesian approach involves a probability distribution of network weights can make the algorithm robust for prediction [22, 23].

A very large number of neurons in the network or too many epochs can lead to the phenomenon called overfitting, which means that during the training period the network provides very good results, instead it generalizes poorly. On the other hand, too few epochs may not be enough for the network to find the optimal weights that provide good predictions. Setting these parameters is a challenge in developing forecasting models.

Most previous works have focused on optimizing either the network architecture (such as the number of neurons or hidden layers) or the selection of input variables, while few have investigated the combined influence of these two aspects on model accuracy and generalization. The originality of this study lies in the joint analysis of the number of input parameters and the number of neurons in the hidden layer, aiming to improve the predictive performance of ANN-based inflow forecasting models. Moreover, the model is applied to the Izvorul Muntelui–Bicaz reservoir, one of Romania's most important hydropower reservoirs, for which few machine learning–based forecasting studies have been reported in the literature.

The objective of this research is therefore to develop and evaluate an ANN model capable of accurately forecasting daily inflows based on historical inflows and meteorological data, and to assess the effect of model configuration parameters on forecasting accuracy.

2. Materials and Methods

ANNs were inspired by the functioning of neurons in the human brain. Mainly, an ANN consists of an input layer, one or more hidden layers and the output layer, each of these layers containing a fixed number of neurons. The information is transmitted between the layers by means of weights. The vectors containing these weights are adjusted during the training process when the network receives significant input-output data. Backpropagation is a method used to adjust the connection weights thus the error between response of the ANN and the expected one must be minimal.

The architecture of an ANN with an input layer, a hidden layer and an output layer is shown in Fig. 1.

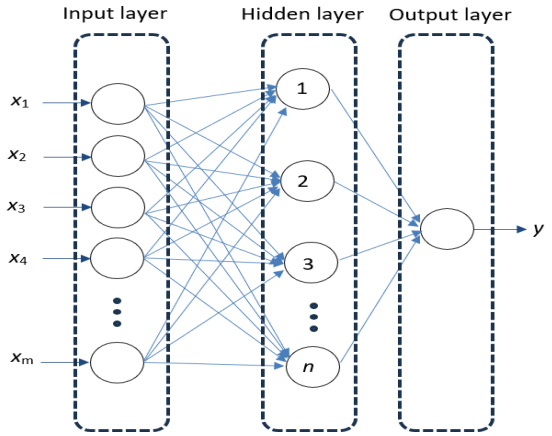

Fig. 1. ANN architecture with *m* inputs (*x*), one hidden layer with *n* neurons and one output (*y*)

Network architecture determines the number of connection weights. Each neuron on the input layer is interconnected with each neuron on the hidden layer.

It is known that the parameters that can be set, such as the initial weights and biases, the number of hidden layers, the number of neurons on each hidden layer, the optimization method used to adjust the weights, the learning constant or the number of training epochs are chosen by trials. ANN's performance depends on the right choice of these parameters.

When some input data has the order of magnitude much different from the rest of the data, they have a greater influence on the output variable. To avoid this situation, all the data used in ANN are normalized with the relation:

$$X_n = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \tag{1}$$

where X_n is the normalized value, $X_{\min}$ and $X_{\max}$ are the minimum, respectively the maximum values of variable time series.

Performance evaluation criteria are an important aspect of validating the prediction model. The most used indicators are coefficient of determination (R^2), root mean square error (*RMSE*) and Nash-Sutcliffe efficiency (*NSE*).

Coefficient of determination, denoted as R^2, ranges from 0 to 1. When R^2 is higher than 0.85 the predictions for daily inflow are considered very good. This coefficient is calculated by relation:

$$R^2 = \left\{ \frac{\sum_{i=1}^{N} \left[\left(Q_{i,obs} - \overline{Q}_{obs} \right) \left(Q_{i,sim} - \overline{Q}_{sim} \right) \right]}{\sqrt{\sum_{i=1}^{N} \left(Q_{i,obs} - \overline{Q}_{obs} \right)^2 \sum_{i=1}^{N} \left(Q_{i,sim} - \overline{Q}_{sim} \right)^2}} \right\}^2, \quad (2)$$

where $Q_{i,obs}$ is the *i*-th observed value; $Q_{i,sim}$ is the *i*-th simulated value; $\overline{Q}_{obs}$ is the mean of observed values; $\overline{Q}_{sim}$ is the mean of simulated values and N represents the total number of observations.

Root mean square error, *RMSE*, may be considered acceptable if is below half of the standard deviation of the observed values. This metric is computed according to the following formula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(Q_{i,sim} - \overline{Q}_{obs} \right)^2}{N}}. \quad (3)$$

Nash–Sutcliffe efficiency coefficient (*NSE*) is calculated using the following equation:

$$NSE = 1 - \frac{\sum_{i=1}^{N} \left(Q_{i,sim} - Q_{i,obs} \right)^2}{\sum_{i=1}^{N} \left(Q_{i,obs} - \overline{Q}_{obs} \right)^2}. \quad (4)$$

A *NSE* value greater than 0.85 reflects excellent model performance, while negative values suggest that the model performs worse than simply using the mean of the observed data.

3. Case study and data

The case study is the largest artificial reservoir on interior rivers in Romania, Izvorul Muntelui-Bicaz. The image from the study area (Fig. 2) was obtained from Google Maps Street view [24]. The primary function of this reservoir is to supply water to Dimitrie Leonida (Stejaru) hydropower plant for electricity generation. Being the first in a series of reservoirs located along the Bistrița River, it benefits

from inflows that are unaffected by upstream human interventions and follow a natural hydrological pattern, making them more predictable.

Fig. 2. Izvorul Muntelui-Bicaz reservoir (street view image [24])

Data used are daily recorded values in the period 2005-2021, divided to be used in the ANN as: the daily inflows for the period 2005-2019 are all used for training (80%) and validation (20%); the daily inflows for 2020 and 2021 are used for testing. As input data into the neural network were used average daily inflows in the reservoir, precipitation, atmospheric temperature and the day index.

The meteorological data for Izvorul Muntelui-Bicaz reservoir area, recorded with daily time step, were downloaded from the https://open-meteo.com/ website. The statistical characteristics of the time series consisting of temperatures, precipitations and average daily inflows used in the analysis are listed in Table 1. For each variable, the minimum, maximum, average values (Mean) as well as the standard deviation (Std) are listed.

Table 1

Statistical features of data recorded for the period 2005-2019 and for 2020 and 2021

Period	Period 2005-2019	Minimum	Maximum	Mean	Std
2005-2019	Temperature [°C]	-21.00	26.50	8.72	8.91
	Precipitation [mm]	0.00	67.20	2.05	4.49
	Inflow [m^3/s]	2.54	362.42	45.39	42.22
Year 2020	Temperature [°C]	-4.70	22.70	9.41	7.40
	Precipitation [mm]	0.00	27.10	2.41	4.60
	Inflow [m^3/s]	5.43	343.70	39.58	42.26
Year 2021	Temperature [°C]	-11.10	23.60	8.37	8.11
	Precipitation [mm]	0.00	28.40	2.18	4.32
	Inflow [m^3/s]	6.61	259.73	46.31	38.82

The analyzed variables show no significant differences in mean and standard deviation across the three considered time periods.

To determine the dependence between the inflow and its past values from previous time steps, the time series was analyzed, and the Partial Autocorrelation Function (PACF) was represented using the econometric modeler toolbox in MATLAB (Fig. 3).

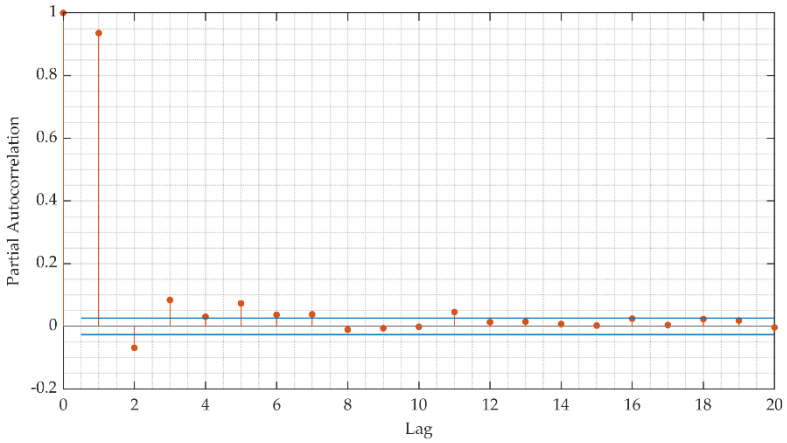

Fig. 3. Partial Autocorrelation function of the inflow series

Fig. 3 shows the Partial Autocorrelation Function (PACF) of the inflow time series, where the red dots represent the PACF coefficients at each lag, and the blue horizontal lines indicate the 95% confidence bounds—values outside these bounds are considered statistically significant. It can be observed that the strongest correlation is between the current inflow and the inflow from the previous day. These links exist up to the 11th order, the first 7 being more important, as observed from the figure when PACF values are outside the range limited by the two blue lines. Trials were made and $Q(t–1)$, $Q(t–2)$, $Q(t–3)$, $Q(t–7)$ were kept as first four input variables in the ANN, where $Q(t–i)$ represents the inflow at the i-th lag number.

Next three variables chosen as inputs in the ANN are meteorological data: average daily precipitations, $P(t–1)$ and $P(t–3)$, and atmospheric temperature, $T(t–1)$, where $P(t–i)$ and $T(t–i)$ represent the precipitation and atmospheric temperature at the i-th lag.

The eighth value of network entry data is the day index. Like the other data, the day index, $i = 1, ..., 365$, has been normalized.

The output of the model is the inflow at the time t, $Q(t)$.

There have been built 8 different models, M1 to M8, as follows:
M1: $Q(t–1)$;
M2: $Q(t–1)$, day index;
M3: $Q(t–1)$, $Q(t–2)$, day index;
M4: $Q(t–1)$, $Q(t–2)$, $Q(t–3)$, day index;
M5: $Q(t–1)$, $Q(t–2)$, $Q(t–3)$, $Q(t–7)$, day index;
M6: $Q(t–1)$, $Q(t–2)$, $Q(t–3)$, $Q(t–7)$, $P(t–1)$, day index;
M7: Q(t–1), $Q(t–2)$, $Q(t–3)$, $Q(t–7)$, $P(t–1)$, $P(t–3)$, day index;
M8: $Q(t–1)$, $Q(t–2)$, $Q(t–3)$, $Q(t–7)$, $P(t–1)$, $P(t–3)$, $T(t–1)$, day index.

The influence of both the number of inputs in the network and the number of neurons on the hidden layer on the performance of the prediction model was studied.

4. Results and discussions

A three-layer feedforward back-propagation artificial neural network (ANN) model was designed, in which both the number of inputs and the number of neurons in the hidden layer were variable. Initially, the case with a single input (denoted M1 for ease of annotation in figure legends) and a variable number of neurons ranging from 1 to 12 was analyzed. It was used a small value for the learning constant, 0.01 and the number of epochs was set to 2000. There were made 20 runs of the program for each model (number of inputs) and a fixed number of neurons and those weights that lead to the minimum *RMSE* and the maximum R^2 for the validation data set were kept. Bayesian regularization was chosen as optimization algorithm to minimize a combination of squared errors and weights that generalize well in this case. For the analyzed case study M1 to M8 the results are represented in Fig. 4.

The one-day prediction was made for 365 days, updating the input data daily with measured values. Daily predictions obtained at each iteration are not used for future values.

It can be seen from Fig. 4 that during training period, the allure of the R^2 curve depending on the number of neurons on the hidden layer increases continuously. Increasing the number of neurons enhances the network's ability to capture nonlinear patterns and improve learning performance. The M1 model, in which the input is only 1 lagged inflow, the network has the worst performance, regardless of the number of neurons on the hidden layer. At a small number of neurons all models learn equally poorly, but it can still be said that the best of them is the model with the highest number of inputs, M8. When the number of neurons increases to the n=12 value, during the training period, the M8 model is able to learn very well.

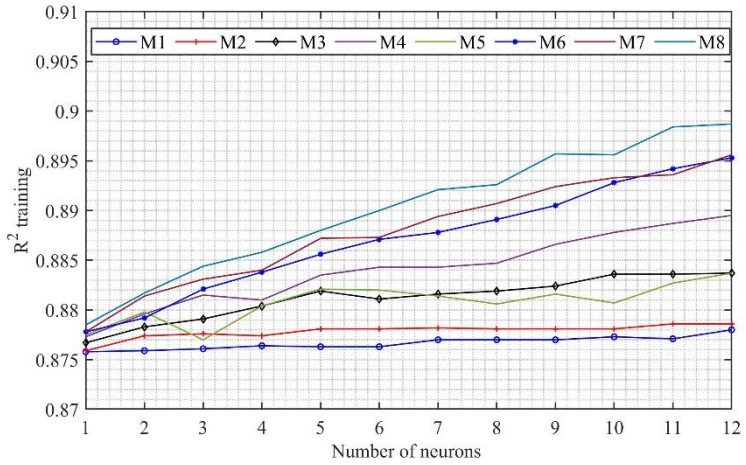

Fig. 4. Coefficient of determination, R^2, for training data set

For the eight models, M1 to M8, the performance evaluation criteria indicators, R^2 for validation data set (Fig. 5), *RMSE* for training and validation dataset (Fig. 6 and Fig. 7) and *NSE* for training and validation dataset were determined. Values are presented in Table 2.

High performance during the training phase does not guarantee good generalization on validation or test datasets. In terms of prediction, the most important thing is that for the validation and testing dataset the correlation between the network responses and the expected values (target) should be as close as possible.

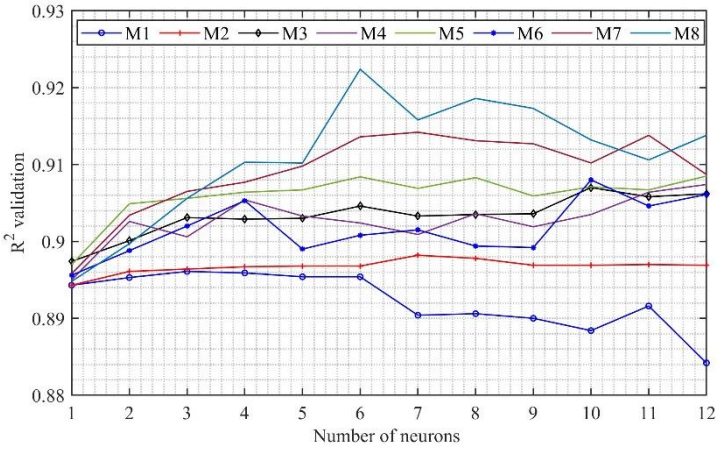

Fig. 5. Coefficient of determination, R^2, for validation data set

From Fig. 5 it can be seen that R^2 as a function of the number of neurons has an ascending tendency, after which it begins to decrease. A good prediction of the model is obtained when this curve reaches its maximum. It is observed for the analyzed case that for M8, a maximum coefficient of determination, R^2=0.922, is obtained.

An important indicator of the performance of an ANN is the root mean square error (*RMSE*). In Fig.6 and Fig.7 are represented *RMSE* for training, respectively validating dataset.

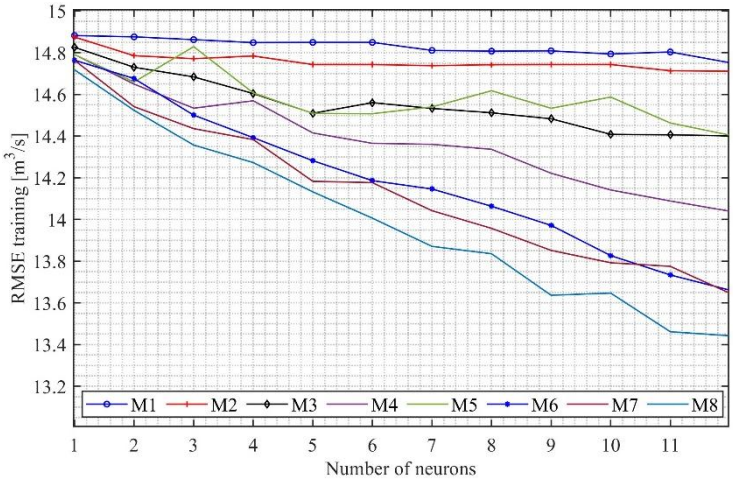

Fig. 6. Coefficient RMSE variation during the training period

As shown in Fig. 5, model M8 demonstrates improved response accuracy with an increasing number of neurons during training, as anticipated. In models with few inputs (M1, M2) there is a slight decrease in *RMSE* as the number of neurons increases, but from a certain number onwards (n=5) the curve reaches a horizontal level, and the errors remain the same. For M6, M7 and M8, the increase in the number of neurons means the continuous decrease in errors between ANN responses and measured values.

As can be seen from Fig. 7, each ANN model/architecture responds differently depending on the number of neurons in the hidden layer, but in general, as the number of network input variables increases, the network performs better.

For the M7 model, the optimal architecture was obtained with 7 neurons on the hidden layer (*RMSE* =12.7295 m^3/s), and for M8, which was also the best of the analyzed models, the optimal number of neurons on the hidden layer was 6 (*RMSE* =11.8276 m^3/s).

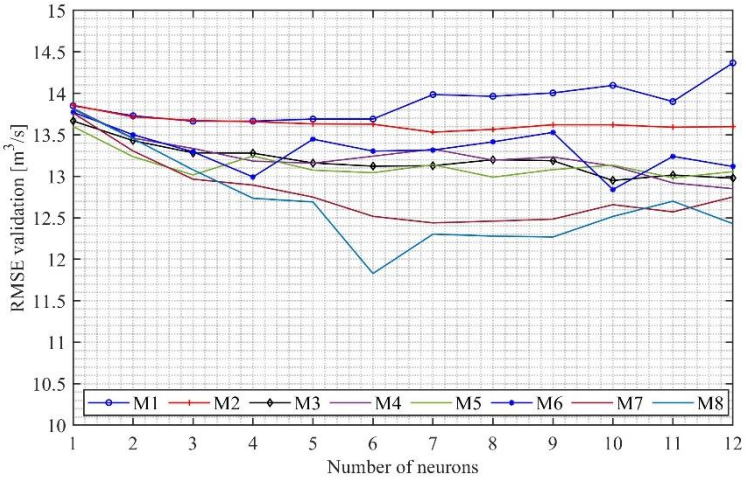

Fig. 7. Coefficient RMSE variation during the validation period

Table 2 lists the performance indices calculated for the optimal alternative found with each of the models. Starting with the M3 model, all the others are very suitable for forecasting, with R^2 corresponding to the validation period greater than 0.9. The last 2 columns represent the *NSE* coefficient, which is also greater than 0.85, which indicates a very good forecast model, as specified in the specialized literature. Since the values of the Nash–Sutcliffe Efficiency (NSE) coefficient are very close to those of the R^2, the *NSE* graph was not included in the paper.

Table 2

Performance indices for each model at the optimal number of neurons

Model	Number of neurons	R^2		*RMSE* [m³/s]		*NSE*	
		training	validation	training	validation	training	validation
M1	5	0.8762	0.8954	14.8561	13.6923	0.8762	0.8946
M2	6	0.8781	0.8984	14.7420	13.5264	0.8781	0.8971
M3	5	0.8801	0.9049	14.6193	13.1390	0.8801	0.9030
M4	6	0.8843	0.9024	14.3654	13.2421	0.8843	0.9014
M5	9	0.8801	0.9043	14.6245	13.1888	0.8801	0.9022
M6	9	0.8895	0.9028	14.0378	13.2300	0.8895	0.9016
M7	6	0.8880	0.9114	14.1288	12.7295	0.8880	0.9089
M8	6	0.8900	0.9224	14.0065	11.8276	0.8900	0.9214

From Table 2, it can be observed that the R^2 values are close to 1, the RMSE is lower than half the standard deviation (i.e., 42.22 m³/s for training and 42.26 m³/s for validation), and the NSE exceeds 0.85. These performance indicators demonstrate that all the models are suitable for daily inflow forecasting.

The results given by the most appropriate model (M8) were represented on the same graph, along with the measured values, Fig. 8.

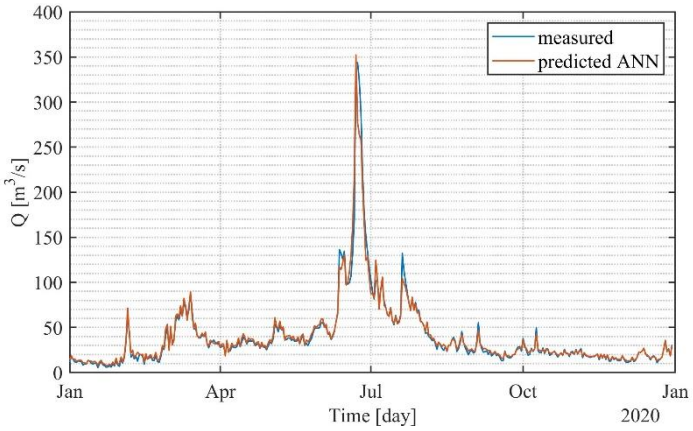

Fig. 8. Observed inflows and M8 model predicted inflows for the testing period (2020)

It should be noted that the values of inflows, precipitations and temperatures for 2020 were not used during the training period. More important differences occur in the case of extreme maximum values. Although the two curves in Fig. 8 seem very close, there are differences between the measured values and those predicted by the M8 model, highlighted in Fig. 9.

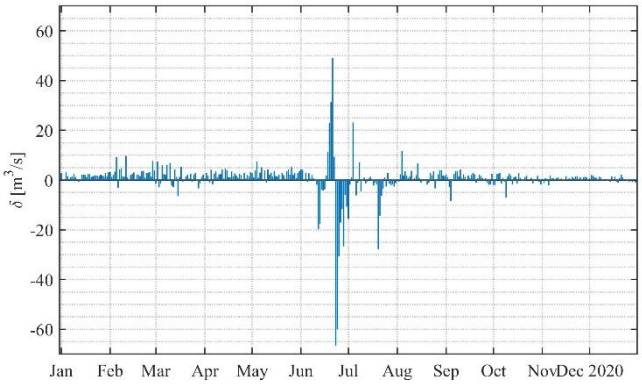

Fig. 9. Differences between observed inflows and M8 model predicted inflows during the validation period (2020)

The variable shown in Fig. 9 corresponds to the prediction error of model M8, calculated as the difference between predicted and measured values for 2020 dataset, denoted by δ. Analyzing these differences, it can be observed that there are

maximum in the summer period, when the flow values are high and in general there are floods whose duration and time of occurrence are less predictable. An analysis of the prediction errors shows that in 28.61% of the cases, the difference between predicted and measured values was less than 1 m³/s. In contrast, only 2.46% of the cases exceeded 20 m³/s, and just 0.82% were greater than 40 m³/s.

Further testing was performed using the M8 model for the following year. The comparison between model M8 predictions and daily measured data for 2021 is illustrated in Fig. 10. The obtained values of the performance indices are R^2=0.84; *RMSE* =15.77 m^3/s; *NSE*=0.83. Similar to the forecast for 2020, in 2021 there are considerable differences for the summer period. The sets of inflows and meteorological data for 2021 were not used during the network training period.

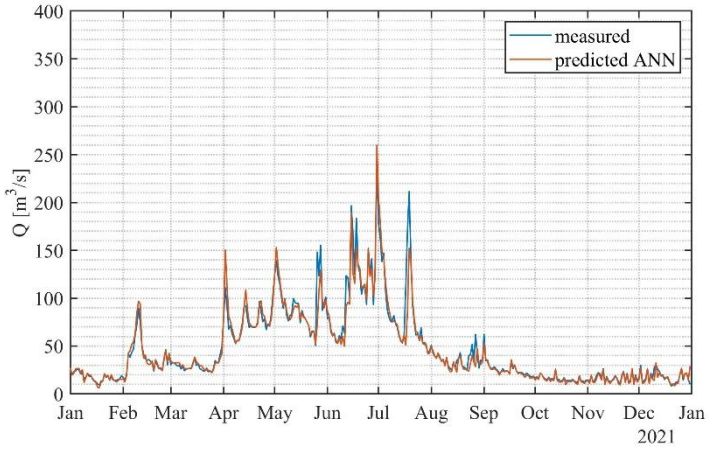

Fig. 10. Observed inflows, M8 model predicted inflows during the testing period (2021)

Fig. 11 represents the difference between the measured values and those forecasted with the M8 model (with 8 inputs), calibrated for the period 2005-2019 and tested with real inputs from 2021. These differences, denoted by δ, were analyzed, and in 32.6% of the cases, they were found to be less than 1 m³/s. Considering that only 3.8% of the cases exhibit differences greater than 20 m³/s, and just 1.09% exceed 40 m³/s, it can be concluded that the model provides highly accurate forecasts of inflows into the analyzed reservoir.

Although the differences between the forecasted and measured values are slightly higher in 2021 compared to 2020, this can be attributed to the increased daily variability of the observed inflows during that year. The model maintained good predictive performance despite the more dynamic hydrological conditions, demonstrating its robustness.

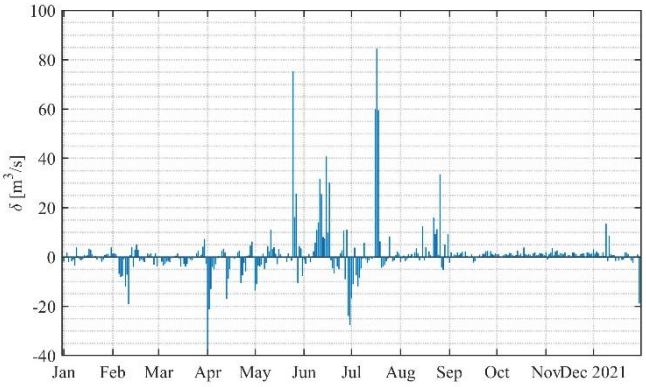

Fig. 11. Differences between observed inflows and M8 model predicted inflows during the testing period (2021)

4. Conclusions

This paper proposes an artificial neural network (ANN)–based method for forecasting reservoir inflows, with a focus on improving prediction accuracy. The study investigates the influence of two key factors: the number of input parameters used in the model and the number of neurons in the hidden layer for each network architecture. A few aspects were observed during the simulations:

- for the training period, as the number of neurons increases, the network learns better, but with the disadvantage of a poor generalization performance, especially in the case of models with few input parameters;
- for the validation period, for a chosen number of inputs, when the number of neurons on the hidden layer increases, the RMSE decreases to a certain value, after which the RMSE starts to increase, the allure being similar for all 8 models.

Similar to our observation regarding the decrease in RMSE up to an optimal number of neurons followed by an increase due to overfitting, study [25] graphically demonstrates how varying the number of neurons from 2 to 25 leads to a minimum error, after which the error rises significantly. This trend confirms the importance of fine-tuning the network architecture to avoid overfitting and to preserve generalization capability.

Furthermore, the analysis of exogenous variables such as precipitation and temperature aligns with the findings of study [26], which highlights that input selection is a critical factor in the performance of neural networks applied to streamflow prediction. Careful selection of input variables significantly contributes to improving both the accuracy and robustness of the models. The model's robustness was demonstrated through validation using inflow data from two successive years (2020–2021), which were excluded from training. The consistent performance across both years confirms the model's ability to generalize well under

varying hydrological conditions. These results highlight the novelty of the proposed approach, which combines a carefully selected input structure with optimized network architecture to achieve reliable and high-resolution daily inflow forecasts.

Acknowledgements:

This study was done within the framework of the project "Intelligent asset Management Platform for Hydropower operation & maintenance" (iAMP-Hydro) funded by the European Union's Horizon Europe research and innovation programme under grant agreement N° 101122167.

The authors thank HIDROELECTRICA, the owner of the data for the inflows, for agreeing with the use of the data.

R E F E R E N C E S

[1] *M. Akbarian, B. Saghafian, S. Golian*, Monthly streamflow forecasting by machine learning methods using dynamic weather prediction model outputs over Iran, Journal of Hydrology, Vol. **620**, Part B, 129480, 2023.

[2] *B. Haznedar, H.C. Kilinc, F. Ozkan, A. Yurtsever*, Streamflow forecasting using a hybrid LSTM-PSO approach: the case of Seyhan Basin, Natural Hazards, Vol. **117**, pp. 681–701, 2023.

[3] *W. L. Young*, The Box-Jenkins approach to time series analysis and forecasting: principles and applications, Rairo-operations Research, Vol.**11**, pp.129-143, 1997.

[4] *A. Katimon, S. Shahid, M. Mohsenipour*, Modeling water quality and hydrological variables using ARIMA: A case study of Johor River, Malaysia. Sustainable Water Resources Management, 4, pp. 991–998, 2017.

[5] *B. N. Ghimire*, Application of ARIMA Model for River Discharges Analysis. J. Nepal Phys. Soc., Vol. 4, 27–32, 2017.

[6] *V.I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, G. K. Matsopoulos*, A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks, Future Internet, Vol.**15**, Iss.8, 2023.

[7] *K. Gharde, M. Kothari, D. Mahale*, Developed seasonal ARIMA model to forecast streamflow for Savitri Basin in Konkan Region of Maharshtra on daily basis. J. Indian Soc. Coast. Agric. Res. 2016.

[8] *R.J. Abrahart, L. See*, Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments, *Hydrol Process*, Vol. **14**, pp.2157–2172, 2000.

[9] ASCE Task Committee, Artificial neural networks in hydrology. II: Hydrological applications. J. Hydrol. Engng ASCE Vol. **5**, Iss. 2, pp. 124–137, 2000.

[10] *H. Feizi, H. Apaydin, M. Taghi Sattari, M.S. Colak, M. Sibtain*, Improving reservoir inflow prediction via rolling window and deep learning-based multi-model approach: case study from Ermenek Dam, Turkey, Stochastic Environmental Research and Risk Assessment, Vol. **36**, pp.3149–3169, 2022.

[11] *A. Tudose*, *I. Picioroaga*, *D. Sidea*, *C. Bulac*, Neural networks application in short-term load forecasting, U.P.B. Sci. Bull., Series C, Vol. **83**, Iss. 2, 2021.

[12] *D. Zhang, Q. Peng, J. Lin, D. Wang, X. Liu, J. Zhuang*, Simulating Reservoir Operation Using a Recurrent Neural Network Algorithm, Water, Vol.**11**, pp. 865-883, 2019.

[13] *G.J. Bowden, G.C. Dandy, H.R. Maier*, Input determination for neural network models in water resources applications: Part 1 - Background and methodology, Journal of Hydrology, Vol.**301**(1-4), pp.75-92, 2005.
[14] *S. Ghimire, Z.M. Yaseen, A.A. Farooque, R.C. Deo, J. Zhang, X. Tao*, Streamflow prediction using an integrated methodology based on convolutional neural network and long short-term memory networks, Sci Rep., Vol.**11**, Iss. 1:17497, 2021.
[15] *Z.M. Yaseen, W.H.M.W. Mohtar, A.M.S. Ameen, I. Ebtehaj, S.F.M. Razali, H. Bonakdari, S.Q. Salih, N. Al-Ansari, S. Shahid*, Implementation of Univariate Paradigm for Streamflow Simulation Using Hybrid Data-Driven Model: Case Study in Tropical Region, IEEE Access, Vol. **7**, pp. 74471-74481, 2019.
[16] *S. Karimi-Googhari, H.Y. Feng, A.H. Ghazali, L.T. Shui*, Neural Networks for Forecasting Daily Reservoir Inflows, Pertanika J. Sci. & Technol., Vol. **18**, Iss.1, pp. 33 – 41, 2010.
[17] *Z.M. Yaseen, A. El-Shafie, O. Jaafar, H.A. Afan, K.N. Sayl*, Artificial intelligence based models for stream-flow forecasting: 2000–2015, Journal of Hydrology, Vol.**530**, pp. 829-844, 2015.
[18] *H. Maier, G. Dandy*, Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, Environmental Modelling & Software, Vol. **15**, Iss. 1, pp. 101-124, 2000.
[19] *J. Lin, C. Cheng, K. Chau*, Using support vector machines for long-term discharge prediction, Hydrological Sciences Journal, Vol. **51**, pp.599–612, 2006.
[20] *Z. Wang, J. Qiu, F. Li*, Hybrid models combining EMD/EEMD and ARIMA for long-term streamfow forecasting. Water, Vol. **10**, Iss. 7, 853, 2018.
[21] *A. Lee, Z. W. Geem, K.-D Suh*, Determination of Optimal Initial Weights of an Artificial Neural Network by Using the Harmony Search Algorithm: Application to Breakwater Armor Stones, Appl. Sci., Vol. **6**, 2016.
[22] *L.V. Jospin, W. Buntine, F. Boussaid, H. Laga, M. Bennamoun*, Hands-on Bayesian neural networks—a tutorial for deep learning users, IEEE Comput Intell Mag., Vol **17**, Iss.2, pp.29–48, 2022.
[23] *M. Magris, A. Iosifdis*, Bayesian learning for neural networks: an algorithmic survey, Artificial Intelligence Review, Vol. **56**, pp.11773–11823, 2023.
[24] Google Maps, https://maps.app.goo.gl/zvULddXz14ZHvyWCA (accessed 17/09/2025).
[25] *C.-T. Cheng*, *Z.-K. Feng*, *W.-J. Niu*, *S.-L. Liao*, Heuristic methods for reservoir monthly inflow forecasting: A case study of Xinfengjiang Reservoir in Pearl River, China, Water, Vol. 7, No. 8, pp. 4477–4495, 2015.
[26] *C. Wannasin*, *C. C. Brauer*, *R. Uijlenhoet*, *P. J. J. F. Torfs*, *A. H. Weerts*, Machine learning for real-time reservoir operation simulation: Comparing input variables and algorithms for the Sirikit Reservoir, Thailand, Journal of Hydroinformatics, Vol. 26, No. 12, pp. 3151–3171, 2024.