FORECASTING SUSPENDED SEDIMENT LOAD USING REGULARIZED NEURAL NETWORK: CASE OF THE SEBAOU RIVER (ALGERIA)

Salah Eddine Tachi [1], Lahbassi Ouerdachi [2], Mohamed Remaoun [1]

[1] Laboratoire Eau et Energie, université Hassiba Ben Bouali, Hay Salem route nationale N°19, 02000, Chlef, Algeria;
[2] Laboratoire hydraulique et construction hydraulique, université Badji Mokhtar, Bloc de rechereche BP12 - 23000-Annaba, Algeria ;
Corresponding author : E-mail: salah008@hotmail.fr

Abstract

In the management of water resources in different hydro- systems it is important to evaluate and predict the sediment load in rivers. It is difficult to obtain an effective and fast estimation of sediment load by artificial neural network without avoiding over-fitting of the training data. The present paper comprises the comparison of a multi-layer perception network using the regularized neural network using the Early Stopping technique to estimate and forecast suspended sediment load in the Sebaou River, northern Algeria. The study was carried out on daily sediment discharge and water discharge data of 9 years (1978–1987).the data sets were divided into three sets; training, validation and testing. The validation set were used only for cross validation. The results of the Back Propagation based models were evaluated in terms of the coefficient of determination (R^2) and the root mean square error (*RMSE*). The comparison results indicate that the regularizing ANN using the Early Stopping technique to avoid over-fitting stops in the first 50 epochs from 200 epochs. and the results show that the overtraining in the back propagation occurs because of the complexity of the data introduced to the network.
Keywords: artificial neural network, sediment discharge, early stopping, river, beni amrane, Algeria.

INTRODUCTION

Among the most critical environmental hazards that hydrologists are dealing with nowadays is sediment load in watersheds. An effective and fast estimation of flow and flux in watersheds are ones of great interests for large number of engineering applications to protect hydraulic infrastructure from different disasters such as: stability problems, the loss of water storage in reservoir and the deterioration of water quality.

The processes of flow and sediment load are complex in Algeria, due to rainfall regime which is infrequent, intense and occurs in the coastal band, as well as the shortage of data and the difficulty of daily direct measurement. According to (REMINI, 2004; REMINI et al., 2009) the erosion rate is between 2000 and 4000 t/km2/year. The average annual amount of deposited sediment in dams passed from 20

million m3 in the 80's to 35 million m3 in the 90's and reached 45 million m3 in 2000 (SERBAH, 2011).

The increasing of suspended sediment load and its sedimentation in Algeria led hydrologists to research the phenomenon of suspended sediment discharge and its relation with some of hydro-climatic parameters, such as rainfall, runoff, land cover, and sediment concentration in different rivers.
Recently hydrologists compared different artificial intelligence techniques encouraging the search for novel methods to improve the ANN training and avoid the over-fitting that occurs in the networks. The over-training of the used data may result in deterioration of generalization properties of the model and when applied to novel measurements lead to its unreliable performance (PIOTROWSKI and NAPIORKOWSKI, 2013). The early stopping criterion is one of the most common methods used in artificial neural network to avoid over-fitting because of its simplicity of understanding and implementation (YINYIN LIU et al., 2008), (PRECHLET, 1998).

This paper tries to apply the early stopping technique to estimate and forecast the sediment discharge, and presents the comparison results of regularized neural networks using different input combination in the case of the Sebaou River, situated in northern Algeria.

STUDY AREA:

The study area comprises the watershed of the Sebaou that is Located north central Algeria, it is limited to the north by the sea Mediterranean, to the east by the watershed of the coastal Constantine, the south by the basin of the Soummam, the Isser and Cheliff which borders the west also. The southern boundary is formed mainly by the mountain range the Tell Atlas. The Algiers Coastal Basin is contained between latitudes 36 ° 1 ' 0and 36 ° 58 ' 0 North and between longitudes 5 ° 6' 0 and 7 ° West. It covers an area of 11 972 km², and is of the twenty hydrological basins that are divided by ANRH.

METHOD TO PREVENT OVER-FITTING:

To avoid over-fitting, in the present paper we used one of the common techniques for regularising errors in ANN, so-called 'Early Stopping' (PRECHLET, 1998) (HAYKIN, 1999). We applied the Early Stopping criterion based on Prechlet's criterion (PRECHLET, 1998) and (PIOTROWSKI and NAPIORKOWSKI, 2013), where we divided our data set into three parts: training, validation and testing. The validation set was used only to evaluate the error during training set once in a while, knowing that we used only training set for training. To avoid Early Stopping on the validation error that may still go further down after it has begun to increase, we let the training iterations finish with a condition given to our model to save the network with

the lowest generalized error that was evaluated and compared in every epochs.

RESULTS AND DISCUSSION

The sediment load was predicted using the "Early Stopping" technique which depended on the best network during the same iterations. A different numbers of input combinations were tried by BP model and the performances were compared to each other for the best input combination that gave the best values of RMSE and R^2.

ANN	Input Combination	Epoch	Training		Validation		Testing	
			RMSE	R^2 (%)	RMSE	R^2 (%)	RMSE	R^2 (%)
ANN_1	WD_t	05	8.21	50.8	0.30	88.5	0.82	69.8
ANN_2	SSD_{t-1}	36	8.35	32.5	3.52	19.2	0.91	07.2
ANN_3	WD_t, WD_{t-1}	08	6.47	59.1	8.35	05.3	3.73	70.0
ANN_4	WD_t, SSD_{t-1}	08	6.82	54.3	1.74	80.3	3.00	70.7
ANN_5	SSD_{t-1}, SSD_{t-2}	19	8.21	34.6	3.34	27.2	0.90	07.4
ANN_6	WD_t, SSD_{t-1}, SSD_{t-2}	03	7.15	51.5	0.89	94.8	1.96	62.9
ANN_7	WD_t, WD_{t-1}, SSD_{t-1}	06	4.82	77.4	1.15	91.4	3.53	70.1
ANN_8	WD_t,WD_{t-1}, SSD_{t-1}, SSD_{t-2}	09	8.50	29.4	1.27	89.6	0.25	75.7

The RMSE and R^2 values of the testing period (Table 1) were significantly improved as well, the ANN_8 gave the best results with RMSE (0.25) and R^2 (75.7%). The networks with current and previous water discharge ANN_3, the network with current water discharge and previous sediment discharges ANN_4 and the network with current and previous water discharge and previous sediment discharge showed well goodness of fit, and close values to ANN_8 with RMSE (3.73, 3.00 and 3.53) and R^2 (70.0%, 70.7% and 70.1%) respectively. We can notice from all networks, regularized and non regularized, that the predicted values over-predict the observed values during small events. Contrary, during the largest events the values are under-predicted. We can also notice from our results that the predicted sediment discharge showed high goodness of fit in ANN_1, _3, _4, _7and _8, opposite to the other networks that showed poor values during flood period. The results presented in Table 1 showed that over-fitting occurred in our MLP model. We can see that most of our networks were improved using cross validation technique.

CONCLUSION:

The processes of flow and sediment load are complex in Algeria. The present study investigates the comparison between non regularized and regularized ANN using the Early Stopping technique for estimating suspended sediment load on a daily scale in the case of the Sebaou River.

Different input combinations including daily current and previous water discharge and previous sediment discharge were used in the ANN models to obtain the optimal input combination. The results obtained in this study indicate that over-fitting occurred in our networks, and the use of the Early Stopping technique performs better results of our predictive model comparing to non regularizing networks. The major over-fitting that occurred in our ANN model was when we used previous values of sediment discharge, which means the complexity and the big size of data that were introduced to our network. In conclusion, we have shown that forecasting suspended sediment load using the early stopping criterion in ANN training is very robust and effective.

In the future we aim to use the noise injection and optimized approximation algorithm in artificial neural network to avoid over-fitting and to improve our networks for better results.

REFERENCES

Haykin, S., 1999. Neural Networks, A Comprehensive Foundation. Macmillan College Publishing Co., New York, USA.

LIU Y., Starzyk J.A., Zhu Z. 2008. Optimized approximation algorithm in neural networks without overfitting. IEEE Trans. Neural Networks 19 (6). p. 983–995.

PRECHLET L.1998. Automatic early stopping using cross-validation: quantifying the criteria. Neural Netw.11(4). p.761–777.

Piotrowski, A. P., Napiorkowski, J. J. 2013. A comparison of methods to avoid over- FITTING IN NEURAL NETWORKS TRAINING IN THE CASE OF CATCHMENT RUNOFF MODELING.J. HYDROL. 476, 97–111.

REMINI B. 2004. Sédimentation des barrages en algérie [The sedimentation in Algerian dams]. Revue Internationale La Houille Blanche. No1, janvier. PP.1-5.

REMINI B., AALOUCH W., ACHOUR B. 2009. L'algerie: plus d'un siecle d'envasement des barrages (Chapitre 8) [Algeria : more than a century of siltation dam (Chaptre 8)]. Ouvrage Intitule: Etat des resources en eau au maghreb en 2009. Rabat. UNESCO p.123-142.

SERBAH B. 2011. Etude et valorisation des sédiments de dragage du barrage Bakhadda Tiaret. These de Magister. Telemcen. Univ.Abou Baker Belkaid pp44.