

Multi-Output Neural Networks for Estimation of Synthetic Unit Hydrograph Parameters: A Case Study of a Catchment in Turkey

A. GÜVEN, A.Y. GÜNAL* AND M. GÜNAL

University of Gaziantep, Department of Civil Engineering, Gaziantep, Turkey

For developing unit hydrographs of catchments, the detailed information about the rainfall and the resulting flood hydrographs are needed. Such information, however, is available only for a few locations and for the remote locations such information is normally very scanty. In this study, Snyder based synthetic unit hydrographs were developed by using both, the digitized map and the digital elevation model of a case study of a small catchment in Turkey. Multi-output neural network technique was applied to predict three unit hydrograph parameters: peak discharge q_p , time to peak t_p and time base t_b of a number of unit hydrographs observed in the catchment, based on most relevant geomorphological and meteorological parameters. Multi-output neural network was observed to outperform the conventional synthetic unit hydrograph methods. The advantage of the proposed multi-output neural network is based on the fact that it predicts the three parameters of the unit hydrograph, based on a single model, compared to the conventional neural network technique, which utilizes a model for each parameter.

DOI: [10.12693/APhysPolA.132.591](https://doi.org/10.12693/APhysPolA.132.591)

PACS/topics: 92.40.FB, 92.40.QP, 07.05.MH

1. Introduction

Synthetic unit hydrographs try to relate the unit hydrograph shape to watershed characteristics, such as basin length and area [1]. Snyder [2] has proposed the first unit hydrograph technique, that was applicable to ungauged areas. His approach was to estimate the time to peak from watershed length, the distance from the outlet to the watershed centroid, and a regional coefficient, whereas the predicted peak flow rate is calculated using the watershed area, the time to peak, and a storage coefficient.

In the last two decades, neural networks (NNs) have been proposed as an alternative to conventional regression analysis and to numerical methods, used in estimation of hydrological and engineering data [3–10]. NNs are relatively stable with respect to noise in data and have a good generalization potential, to represent input-output relationships [11]. Once a NN model is trained for its generalization properties, it can be demonstrated that the trained model represents the physical process of the system. The knowledge acquired for the problem domain during the training process is encoded within the NN in two forms: (a) in the network architecture itself (through number of hidden units), and: (b) in a set of constants, or weights [12]. Although there are several attempts in other scientific branches that have shown, that useful information could be obtained from trained neural networks (see Yao [12] for relevant references), limited research has been done for water-engineering applications [11, 13].

This study focuses on modelling of three unit hydrograph parameters, namely, the peak discharge q_p , the time to peak discharge t_p , and the time base of unit hydrograph t_b , based on the most relevant hydrological and geomorphological variables. These three parameters are the characteristics of a unit hydrograph, as they are known. The general shape of the corresponding unit hydrograph can be practically determined by drawing a smooth curve from origin to the peak discharge point (rising limb) and from peak point till the end of time base of the hydrograph (falling limb). Therefore estimation of these three parameters will give good enough information of the time-variation of the unit hydrograph. This study directly aims to derive the knowledge from a well trained and tested NN model. The experimental measurements were taken from small-scale catchment research reports from Turkey.

2. Overview of multi-output neural networks

The main difference between the multi-output neural network MONN technique and the conventional NNs is that MONN utilizes multiple inputs to predict multiple outputs, whereas conventional NNs consider a single output. Although the basics of NNs have been sufficiently described in previous studies, the authors find it necessary to remind, especially to non-specialist readers, of most important elements of NNs, used in this study, in order to comprehend the explicit neural networks formulation.

A multi-layer neural network model is considered in this study, which consists of three layers: input, hidden, and output layers. The input layer consists of input nodes, representing input variables. The outputs of the input nodes are normalized and transferred to the hidden layer in which they are processed through a transfer

*corresponding author; e-mail: agunal@gantep.edu.tr

function. The output layer consists of output variables. The most commonly used sigmoid transfer function ($y = 1/(1 + e^{-x})$) is utilised in this study. The proposed NN architecture is trained, till the mean square error between the model prediction and the observed values gets minimal and remains unchanged.

Readers who are interested in more information on MONN technique can refer to previously published works in related journals [3, 12].

3. Case study: Damlica catchment in Turkey

Damlica catchment lies in Marmara Region of Turkey. The area of the Damlica catchment A_w is 7.63 km² and there are 2 recording precipitation gauges and 3 non recording precipitation gauges in this catchment. The runoff of Damlica catchment is measured by one stage recording gauge, installed on a triangular weir, over the main waterway at the outlet of the catchment. The location of the centroid of catchment is at the 41° 06' 04'

North latitude and the 20° 25' 00" East longitude. Its average altitude is 184 m, the length of main stream L_r is 4.8 km and the distance of the catchment outlet from centroid L_{ca} is 1.9 km. The DEM of Damlica catchment, which is at the north west of Turkey, was downloaded from the Aster Global Digital Elevation Model (Aster GDEM) web site. The boundary of Damlica catchment is extracted from the DEM map.

4. Results of training and testing of NN model

The data of Damlica catchment were randomly split into training and testing sets. Namely, 8 of the total 29 data sets were reserved for testing (30%), and 21 (70%) of sets were used for model training. Statistical values of data sets are given in Table I. First, different NN models were developed and their performance was evaluated, based on estimations to these data sets. The MONN modelling is described in the next section.

TABLE I

Statistical values of input and output variables of Damlica basin (29 data sets).

Variable	Min. value		Max. value		Mean		Variance		Std. dev.	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
SR	0.05	0.14	9.38	3.13	1.25	0.76	3.91	0.92	1.98	0.90
F	1.27	1.84	127.68	35.40	12.78	8.88	716.82	112.16	26.77	9.98
t_r	0.08	0.17	1.50	2.55	0.60	1.09	0.16	0.59	0.40	0.73
API	0.00	6.78	42.16	57.94	13.56	20.16	123.94	235.55	11.13	14.47
t_b	4.83	7.90	28.00	23.00	13.64	14.92	34.49	26.12	5.87	4.82
t_p	2.00	2.60	11.00	7.00	4.40	4.29	2.48	6.65	1.57	2.43
q_p	0.42	0.42	0.66	0.65	0.51	0.49	0.01	0.01	0.09	0.07

The model parameters of the multi-layer feed forward neural networks with back-propagation learning algorithm were optimized by Levenberg and Marquardt algorithm, which is one of the most common and successful back-propagation algorithm.

Another important issue, is to find the optimal architecture of the NN model. Most studies in the literature have used a trial approach, which generally leads to local maxima or minima. In this study, this issue is eliminated by using a genetic algorithm in order to find the optimal architecture of the proposed NN model. Namely, a fitness function was chosen based on MSE of the testing set and the program searched for optimal architecture with least MSE for testing set. The optimal architecture of the proposed NN was found to be 4-9-3 (no. of inputs, no. of hidden neurons, and no. of outputs), as shown in Fig. 1. The correlation matrix among the inputs and the outputs is shown in Table II.

The overall performance of the training and testing sets was evaluated by MSE and the R^2 . The training

TABLE II

Correlation matrix for the field data (29 data sets).

	SR	F	t_r	API	t_b	t_p	q_p
SR	1						
F	-0.003	1					
t_r	0.168	-0.312	1				
API	-0.234	-0.049	-0.090	1			
t_b	0.124	-0.241	-0.051	0.180	1		
t_p	0.160	-0.283	-0.048	0.203	0.603	1	
q_p	-0.172	-0.028	-0.115	-0.004	0.326	-0.032	1

results of the proposed models show that the NN has learned the highly non-linear relationship between the input parameters and the unit hydrograph parameters and the UH parameters, with high correlations ($R^2 = 0.990, 0.991, 0.989$ for $q_p, t_p,$ and $t_b,$ respectively) and low errors (MSE = 0.990, 0.991, 0.989 for $q_p, t_p,$ and $t_b,$ respectively).

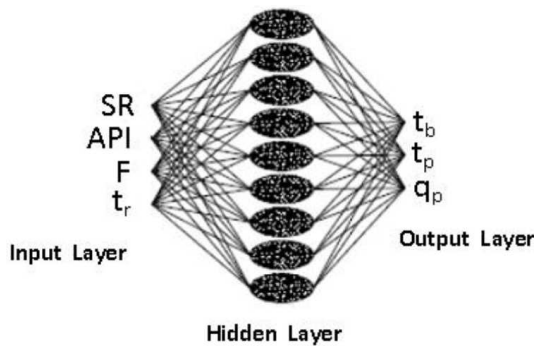
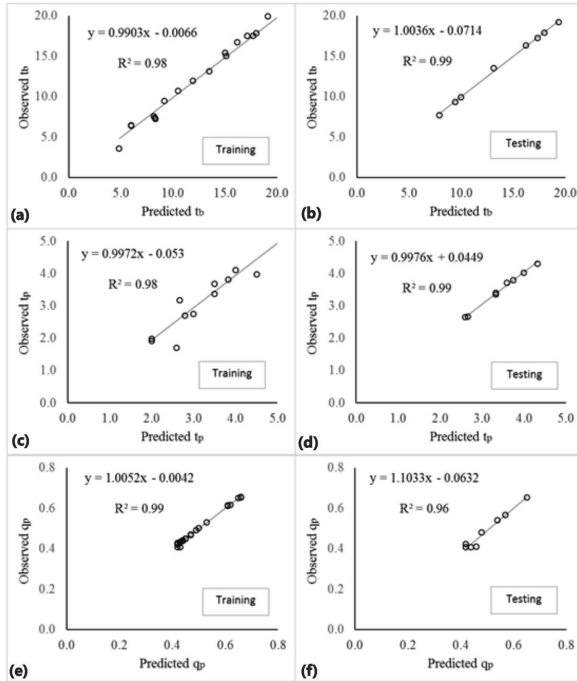


Fig. 1. Architecture of the proposed NN.

Fig. 2. Observed and predicted t_b , t_p and q_p , for the training and testing sets.

Validation of the trained model has proven the high generalization capacity of the proposed model, with a high correlation and low error ($R^2 = 0.990, 0.991, 0.989$ for q_p, t_p , and t_b , respectively). As it is seen in Fig. 2, almost all estimations of the NN fall on the line of perfect agreement, which shows strictly good agreement with the observed values. The overall results show that the proposed NN architecture is well applicable for the next steps of the MONN methodology.

5. Comparison of the MONN with the Snyder's technique

5.1. Performance criteria

It is important to define the criteria by which the performance of a model and its estimation accuracy will be evaluated in the model-development process. Various statistical measures have been developed and used to assess the performance of a model. In our case we have considered a number of error measures that would evaluate the performance of the compared models with respect to model size. The coefficient of determination R^2 , the mean square error (MSE), the variance (VAR), and the standard deviation (Std. Dev.) were used in the comparison of the proposed MONN with the TM-based and the DEM-based Snyder's methods, using the testing set.

5.2. Peak discharge estimation

In this section, the performance of MONN and of the other methods are compared based on the testing data set. Table III shows the overall best performance of MONN with the lowest values of error measures (MSE = 0.077, VAR = 0.00, Std.Dev. = 0.012) and highest correlation ($R^2 = 0.99$). On the other hand, TM-based Snyder estimated q_p is slightly better than the DEM-based Snyder estimation, considering the AIC and the MSE criteria.

TABLE III

Observed and predicted UH parameters for the testing set.

Measure	Models								
	MONN			Snyder			DEM-Snyder		
	t_b	t_p	q_p	t_b	t_p	q_p	t_b	t_p	q_p
R^2	0.989	0.975	0.986	0.374	0.960	0.005	0.374	0.960	0.007
MSE	0.254	0.462	0.077	71.666	0.614	0.034	133.946	0.614	0.024
VAR	24.702	5.539	0.000	3085.106	6.499	0.001	9809.892	6.499	0.000
Std.Dev.	4.830	2.287	0.012	53.979	6.855	8.328	96.254	2.478	0.019

5.3. Time to peak estimation

Table III indicates the performance of MONN in estimation of t_p , compared to the two Snyder's methods. MONN-estimated t_p values have the lowest errors

(MSE = 0.462, VAR = 5.539, Std.Dev. = 2.287) and the highest correlation ($R^2 = 0.975$). The TM-based Snyder's model has shown, in overall, a better performance, compared to the DEM-based Snyder's model. It should

be noted that the TM-based Snyder's model gives over-predicted q_p , while the DEM-based Snyder's model gives generally the under-predicted values. The predictions of the MONN are found to be the closest ones to the observed values.

5.4. Time base of unit hydrograph estimation

Table III indicates the same performance of MONN in estimation of t_b . Namely, MONN estimated the t_b values with the lowest errors (MSE = 0.254, VAR = 24.702, Std.Dev. = 4.83) and the highest correlation ($R^2 = 0.989$). It should be noted that DEM-based-Snyder's model has underestimated all t_b values with considerable deviation, while those of TM-based Snyder's model are relatively close to experimental values, compared to those of TM-based Snyder's model.

6. Conclusions

Three parameters of a synthetic unit hydrograph were estimated by a multi-output neural network model. MONN has estimated the time base and the peak discharge of the corresponding unit hydrographs with very good agreement with the field data, obtained from Damlica Basin in Turkey. MONN is found to outperform the Snyder's synthetic unit hydrograph technique, which is the most widely used in its subject area.

The significance of this study is in the fact, that the proposed MONN model estimates three unit hydrograph parameters q_p , t_p , and t_b in a single NN model, instead of developing three separate NN models, as is done in conventional NN technique. By this, the proposed MONN model can be refined/improved by using more experimental or field observations, and also it can be used in comparison with conventional methods.

The robustness of the proposed MONN is validated in estimation of field data and the results are very promising. MONN estimated the observed q_p , t_p and t_b with quite good accuracy ($R^2 = 0.989, 0.975, 0.986$, respectively).

The results of this study show that conventional synthetic unit hydrograph prediction techniques could be replaced by MONN technique.

References

- [1] B.C. Yen, K.T. Lee, *J. Hydrological Eng.* **2**, 1 (1997).
- [2] W.M. Snyder, *Trans. Amer. Geophys. Union* **19**, 447 (1938).
- [3] A. Aytekin, A. Guven, M.İ. Yuce, H. Aksoy, *Hydrological Sci. J.* **53**, 893 (2008).
- [4] H.R. Maier, G.C. Dandy, *Environmental Modell. Software* **15**, 101 (2000).
- [5] H.Md. Azamathulla, *A Review on Application of Soft Computing Methods in Water Resources Engineering*, in: *Metaheuristics in Water, Geotechnical and Transport Engineering*, chapter 2, Elsevier, 2013, p. 27.
- [6] M. Davraz, Ş. Kiliçarslan, H. Ceylan, *Acta Phys. Pol. A* **128**, B-184 (2015).
- [7] E. Boutalbi, L. Ait Gougam, F. Mekideche-Chafa, *Acta Phys. Pol. A* **128**, 271 (2015).
- [8] İ.H. Karahan, R. Özdemir, *Acta Phys. Pol. A* **128**, B-427 (2015).
- [9] İ.S. Üncü, A. Arisoy, B. Büyükarikan, *Acta Phys. Pol. A* **128**, B-474 (2015).
- [10] Y. Özcanlı, F. Kosovali Çavuş, M. Beken, *Acta Phys. Pol. A* **130**, 444 (2016).
- [11] A. Guven, M. Gunal, A.K. Cevik, *Can. J. Civ. Eng.* **33**, 1379 (2006).
- [12] J.T. Yao, *Lect. Notes Comput. Sci.* **2639**, 430 (2003).
- [13] X. Zhang, F. Liang, B. Yu, Z. Zong, *J. Hydrology* **409**, 696 (2011).