ESTIMATION OF DAILY RUNOFF USING WATER LEVEL DATA AND OBSERVED FLOWRATE DATA

Maga Kim[1], Jin-Yong Choi[2] and Jehong Bang[3]

ABSTRACT

Runoff is a vital factor in water resource planning and water quality management. Rainfall-runoff processes are affected by the complex relation of watershed factors such as geographical features, vegetation distribution, land use, soil properties, and weather conditions. These relationships make the rainfall-runoff processes nonlinear. Artificial Neural Network (ANN) model is the computational model created by simplifying the brain structure of a human. ANN models have been applied to modeling and recreating non-linear natural phenomena such as weather or hydrological data. Once an ANN model has been trained it can be utilized to predict the outcome of a process. In Korea, the Korea Rural Community Corporation (KRC) has been collecting the reservoir water level data every 10 minutes, and a flow observation net has been created to collect runoff flow data. Knowing the volume of runoff from a watershed means that the inflow into the reservoir can be estimated using the reservoir water level and evaporation data in the non-irrigated period via a water balance method. Observed flow data can be used as runoff from a watershed at the observation point. Therefore in this study, the ANN model was utilized to predict daily runoff. Reservoir water level data and observed flow data were used for training the ANN model. The ANN model consists of three layers including input, hidden, and output layer. Runoff data obtained from reservoir water level, the observed flow, weather data, watershed characteristics, and others applied as input data to the ANN model. This dataset was divided into training dataset and test dataset. These two sub-sets of the data were used to train and verify the ANN model. The sigmoid function, which is commonly used for the ANN models, was applied as the activation function of the hidden nodes. The softplus function, which gives only positive results, was applied as the activation function of the output nodes. The results of each model run were evaluated by comparing with the target data. They were also compared with the results of the TANK model which can estimate the runoff of the watershed with input data of the ANN model. This study is useful for estimating and predicting the daily inflow to the reservoir. It is also helpful for managing water resources and operating the reservoirs.

Keywords : Runoff, Artificial neural network model, TANK model;

1. INTRODUCTION

Runoff is the crucial factor for hydrological analysis to manage water resources (Ahn et al., 2002). Hydrological analysis of runoff is divided into long-term runoff analysis and flood analysis. Especially, long-term runoff, continuous hydrological event, is used for a determination of reservoir capacity, using irrigation water for multipurpose, development and management of water resources and establishment of water resource plans (Yoon et al., 1998).

[1] Ph.D. Course in Department of Rural Systems Engineering, Agriculture and Life Sciences, Seooul National University, Seoul 08826, Republic of Korea, E-mail: mnkm53@snu.ac.kr

[2] Professor in Department of Rural Systems Engineering, Agriculture and Life Sciences, Seoul National University, Seoul 08826, Republic of Korea, Email: iamchoi@snu.ac.kr

[3] Ph.D. Course in Department of Rural Systems Engineering, Agriculture and Life Sciences, Seooul National University, Seoul 08826, Republic of Korea, E-mail: jaehong999@snu.ac.kr

The reaction of the watershed, which is the unit of water resources production, causes runoff. To interpret and predict these rainfall-runoff phenomena, deep understanding of the rainfall-runoff process should be necessary. But because the watershed, which is the connector of the rainfall-runoff process, is made up of complex characteristic factors such as topography, vegetation, land use, soil, and weather condition, the watershed induces a non-linear change. The existing linear models have disregarded interpretation of non-linear data in the past, but it can be possible because of advances in computation recently (Oh et al., 2008). The artificial neural network (ANN), which imitates the calculation process of human (Lee and Park, 2003), can model natural phenomena that have a non-linear characteristic such as meteorological or hydrological event. Because of these features, the ANN models are utilized in many realms of water resources and environment, rainfall-runoff interpretation, rainfall distribution, drought, groundwater flow (Choi and Kang, 2000).
In this study, the prediction of daily runoff using the ANN model is carried out. To train the ANN model, reservoir watershed runoff data, which are generated from reservoir water level data in the non-irrigation period, and measured flow data were used. Also, the predicted runoff is compared with results of the TANK model.

2. METHODS

2.1 Artificial Neural Network (Ann)

The ANN model, which simplifies the neuron and synapse of the human brain, is the computational model, made up of simple operators combinations. The ANN model can solve the problem only with given data, even if there's no algorithm or direct solving method (Yeo et al., 2010). Also, ANN models are valuable for modeling complex non-linear and multi-dimension In-Out relation.

The processing nodes, which are arranged on each layer, are the fundamental components of the ANN model. The outputs of the nodes from a layer deliver to nodes of the next layer through connection strength (Ryu et al., 2002). Multi-layer perceptron (MLP), which is the typical structure of ANN models, consist of more than three layers mostly (Seo et al., 2017). The layer where information are entered called the input layer and the layer on which processed information appears called the output layer and all of the layer between the input layer and output layer called the hidden layer. The hidden layers can be many depending on the structure of the ANN model (Atkinson and Tatnall, 1997).MLP with hidden nodes, which is consist of three layers, performs the same operation as following equation.

$$\begin{aligned} \mathrm{o}(\mathbf{x}) &= f\left(w_0 + \sum_{j=1}^{J} w_j \cdot f\left(w_{0j} + \sum_{i=1}^{n} w_{ij} x_i\right)\right) \\ &= f\left(w_0 + \sum_{j=1}^{J} w_j \cdot f\left(w_{0j} + \mathbf{w}_j^T \mathbf{x}\right)\right) \end{aligned}$$

where $\mathbf{x} = \{x_i, i = 1, 2, \cdots, n\}$: input vector, f: activationfunction, $\mathrm{o}(\mathbf{x})$: output vector, w_0: bias for output nodes, w_j: connection strength, $\mathbf{w}_{ij} = \{w_{ij}, i = 1, 2, \cdots, n\}$: connection strength vector of j^{th} hidden node, w_{0j}: bias for j^{th} hidden node (Günther and Fritsch, 2010) Fig 1 is the general structure of MLP which has No. of n_i input nodes, n_h hidden nodes, and n_k output nodes.

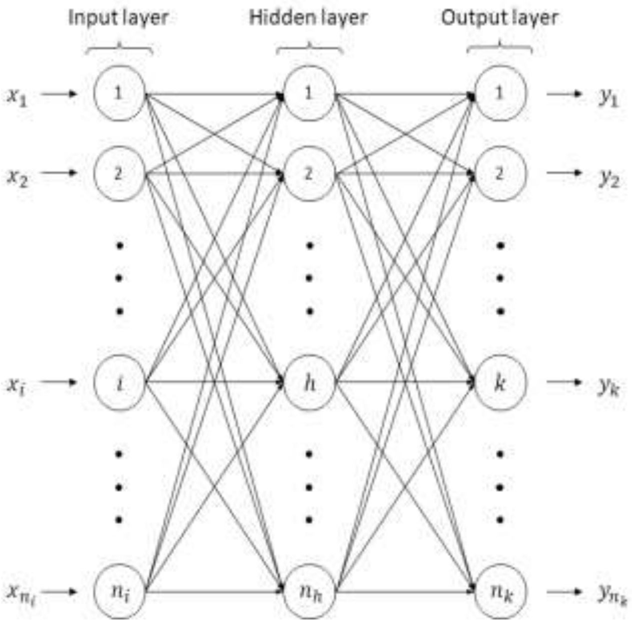


Figure 1 General structure of multi-layer perceptron

2.2 TANK Model

TANK model, which is developed by Sugawara, is a conceptual rainfall-runoff model (Sugawara, 1972). TANK model has a comparatively simple algorithm for interpretation of complex rainfall-runoff phenomena and needs less input data and parameters (Koo et al., 2006). TANK model consists of three or four conceptual tanks, and each tank has two or three discharge outlets. TANK model assumes the watershed as three or four stages of tanks, and carry out the runoff interpretation of watershed by optimizing parameters. TANK model simulates rainfall, evapotranspiration and surface runoff in the upper tank, and simulates interflow and infiltration conceptually by water transfer between each tank. TANK model has advantages that can be applied to watershed lacking in observational data because necessary input data and parameters for runoff interpretation are few, but it cannot ensure runoff interpretation for a watershed where parameters are not known (Kim and Kim, 2012).

Kim and Park (1988) suggested a modified 3-TANK model, which has three tanks and four discharge outlet, for upper stream watershed of the agricultural reservoir of Korea. In this study, the modified 3-TANK model is used for the estimation of runoff. The parameters of the TANK model are calculated according to the regression equations suggested by Kim and Park (1988). The modified 3-TANK model estimate runoff according to the following equations.

$$Q_t = \sum_{i=1}^{n}\sum_{j=1}^{m}(ST_{i,t} - h_{ij})a_{ij}$$

where Q_t: total runoff on day t (mm), i: No. of tanks, j: No. of discharge outlet, $ST_{i,t}$: storage water level of i^{th}tank according to unit time t (mm), h_{ij}: height of j^{th}discharge outlet in i^{th}tank (mm), a_{ij}: coefficient of j^{th}discharge outlet in i^{th}tank

$$ST_{1,t} = ST_{1,t-1} + R_t - ET_t - I_{1,t} - Q_{1,t-1}$$

$$ST_{i,t} = ST_{i,t-1} + I_{i-1,t-1} - I_{i,t} - Q_{i,t-1}$$

where R_t: precipitation on day t (mm), ET_t: evapotranspiration on day t (mm), $Q_{i,t}$: discharge of i^{th} tank on day t (mm), $I_{i,t}$: infiltration of i^{th} tank on day t (mm)

$$Q_{i,t} = \left(ST_{i,t} - h_{ij}\right) \times a_{ij}$$
$$I_{i,t} = ST_{i,t} \times b_i$$

where b_i: infiltration coefficient of i^{th} tank (Ahn et al., 2015)

Figure 2 is a schema of the estimation process of the modified 3-TANK model.

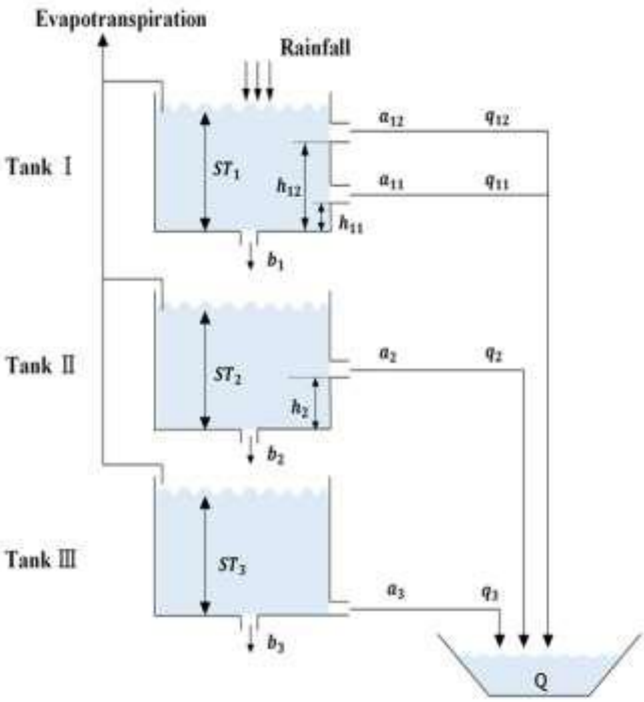


Figure 2. schematic of the modified 3-TANK model (Kim and Park 1988; Song, 2017)

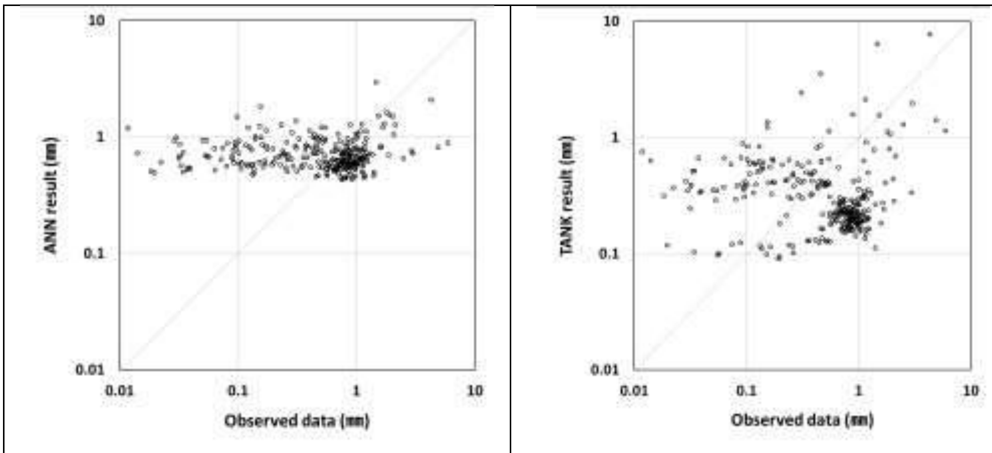
3. RESULTS AND DISCUSSION

In this study, the ANN model estimated runoff of watershed WS, TG, BW1, and BW3 and the results of each watershed were evaluated comparing with generated runoff data using reservoir water level and measured runoff data. Table 1 presents the statistical parameters for each verification period.

Watershed		model	R^2	$RMSE$ (mm)	MAE (mm)
Wangsong (WS)		TANK	0.11	0.82	0.58
		ANN	0.05	0.67	0.46
Togyo (TG)		TANK	0.20	4.44	1.32
		ANN	0.44	3.21	0.83
Banwol	BW1	TANK	0.47	5.38	1.97
		ANN	0.51	2.98	1.58
	BW3	TANK	0.71	5.54	2.29
		ANN	0.59	6.14	2.50

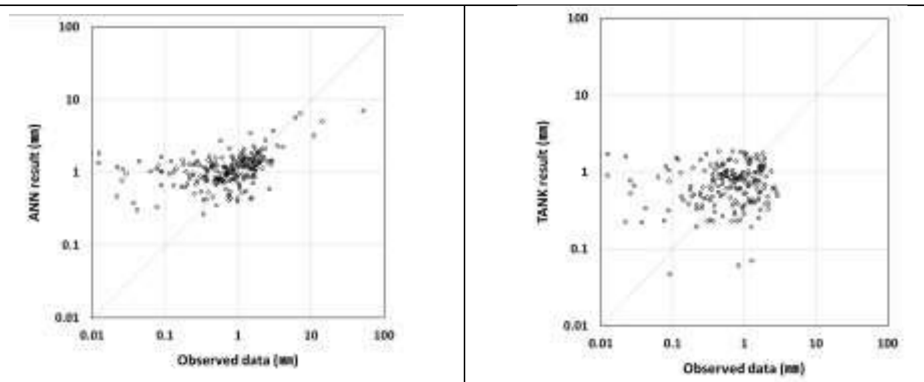
In the case of WS, RMSE and MAE are less than 1 mm for both ANN and TANK model, and R^2 is 0.05 for the ANN model and 0.11 for TANK model. Both ANN model and TANK model seem to unsuitable for WS watershed. In the case of TG, RMSE ans MAE are 3.21 mm and 0.83 mm for the ANN model, 4.44 mm and 1.32 mm for TANK model. R^2 is 0.44 for the ANN model and 0.20 for TANK model. The ANN model reflects the tendency of generated runoff better than TANK model. WS and TG applied generated runoff data for training ANN model. But water level data cannot represent only runoff because it is affected by inflow (watershed runoff) and outflow (release). Even in the non-irrigation period, there is outflow for flood control, canal management, environmental water. It seems that WS reservoir was more affected these outflow factors because WS reservoir has much less effective storage than TG reservoir. It makes generated runoff data, which means training dataset, inappropriate.

In the case of BW1, RMSE and MAE are 2.98 mm and 1.58 mm for the ANN model, 5.38 mm and 1.97 mm for TANK model. R^2 is 0.51 for the ANN model and 0.47 for TANK model. In the case of BW3, RMSE and MAE are 6.14 mm and 2.50 mm for the ANN model, 5.54 mm and 2.29 mm for TANK model. R^2 is 0.59 for the ANN model and 0.71 for TANK model. Both ANN model and TANK models show similar performance. In the BW1 case, the ANN model is slightly better than TANK model, and in the BW3 case, ANN model is worse than TANK model. The ANN model needs sufficient training of dataset to have appropriate performance. But measured flow data for BW1 and BW3 are restricted in aspects of No. of data. It seems to disturb training of the ANN model. In the case of BW3, precipitation was much higher in July especially, and in the other months, which had respectively lower precipitation, had zero or near the zero measured runoff values. This runoff data distribution also can disturb training of the ANN model. Also, because TANK model tends to estimate runoff value smaller than measured runoff data, it can affect statistical parameters. Fig 3-6 are scatter plot of the ANN model and TANK model compared with generated runoff data and measuring runoff data.



(a) ANN model | (b) TANK model

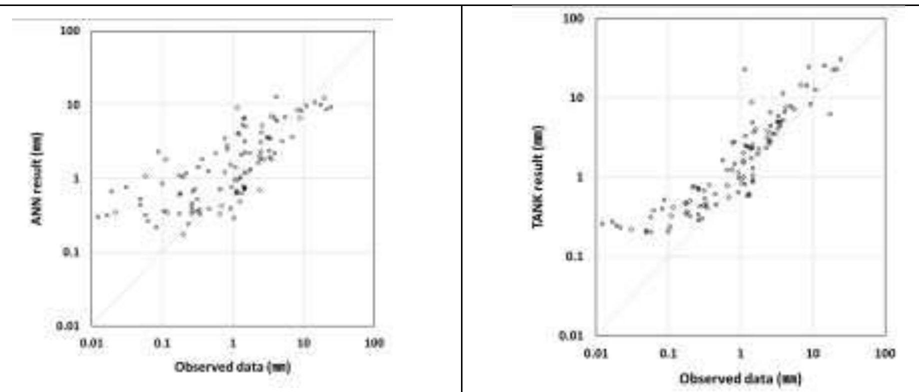
Figure 3. Scatter plot of (a) ANN model result, (b) TANK model result compared with observed runoff data for WS



(a) ANN model

(b) TANK model

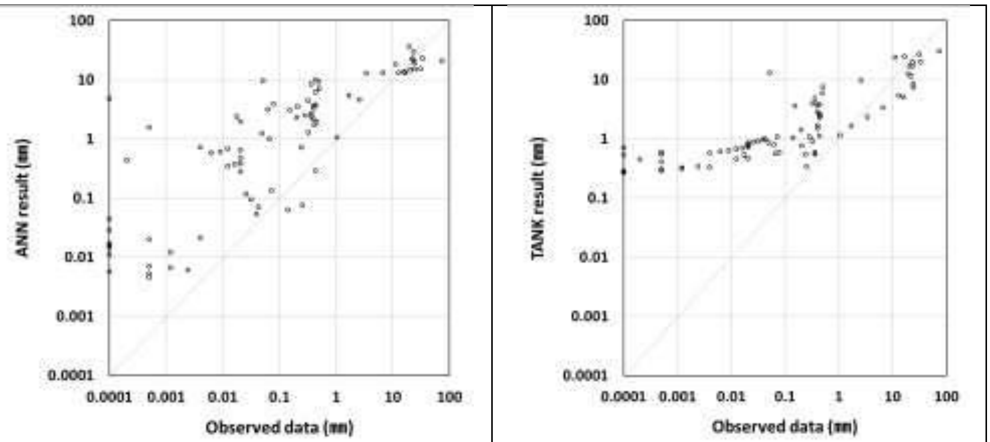
Figure 3. Scatter plot of (a) ANN model result, (b) TANK model result compared with observed runoff data for TG



(a) ANN model

(b) TANK model

Figure 3. Scatter plot of (a) ANN model result, (b) TANK model result compared with observed runoff data for BW1



(a) ANN model

(b) TANK model

Figure 3. Scatter plot of (a) ANN model result, (b) TANK model result compared with observed runoff data for BW3

4. CONCLUSIONS

In this study, the ANN model and TANK model were capplied to estimate runoff of Wangsong watershed (WS), Togyo watershed (TG) and sub watershed BW1 and BW3 of Banwol reservoir. The results of each model was evaluated by comparing estimation runoff data with generated runoff data and measured runoff data.

In the case of WS, both the ANN model and TANK model showed low R2, that means both models cannot reflect the tendency of the generated runoff data in WS watershed. In the case of TG, the ANN model show better performance than TANK model. The ANN models for WS and TG generated runoff data as training data. Because the reliability of the generated data is lower than well-measured data, it can be a hindrance in training the ANN model. Also, although water level data of the non-irrigation period were used to generated runoff data, there are factors except inflow (reservoir watershed runoff) affecting the reservoir water level. It makes the reliability of generated runoff data lesser. Because Wangsong reservoir has a much smaller capacity than Togyo reservoir, it seems Wangsong reservoir is more affected by such factors.

In the case of BW1, the ANN model shows slightly better performance than TANK model. In the case of BW3, TANK model shows better performance. In the case of BW1 and BW3, the ANN model used measured data. But because No. of measured data being less, ANN model cannot have sufficient training dataset. In the BW3 case, measured data has many zero or near the zero values, especially in the verification period of data. A significant difference in the training dataset and testing dataset can cause lower performance in the ANN model. That can be one of the causes that the ANN model shows worse performance than TANK model in the case of BW3.

The variation of TANK model was more significant than the ANN model at some of the points. This tendency is intensified as runoff value becomes higher. Although the training data of the ANN model has restriction such as No. of data, reliability, gap between training data and testing data, the ANN model shows similar or better performance as compared to the TANK model. It seems to be possible to get better performance by utilizing appropriate dataset.

5. REFERENCES

Allen, R., Gichuki, F. & Rosenzweig, C. 1991 CO2-Induced climatic changes and irrigation-water requirements. J. Water Resour. Plann. Manage. 117, 157-178.

Ahn, J. S., Song,J. H., Kang,M. S., Song,I. H., Jun,S. M., Park, J. H. 2015. Regression equations for estimating the TANK model parameters. Journal of the Korean Society of Agricultural Engineers 57(4), 121-133.

Ahn, S. J., Lee, J. K., Han, Y. S., Jun,K. W. 2002. Flood ruoff analysisusing runoff forecasting model. Journal of The Korean Society of Civil Engineers 22(3-B), 311-319.

Atkinson, P. M. & Tatnall,A. R. L. 1997. Introduction neural networks in remote sensing. International Journal of Remote Sensing 18(4), 699-709.

Choi, J. K. & Kang,M. S. 2000. Application of neural network to water resources. Journal of Korean National Comittee on Irrigation and Drainage 7(2), 248-258.

Günther, F.& Fritsch,S. 2010. Neuralnet: training of neural networks. The R Journal. 2(1), 30-38.

Kim, C. G. &Kim, N. W. 2012. Comparison of natural flow estimates for the Han river basin using TANK. Journal of Korean Water Resources Association 45(3), 301-316.

Kim, H. Y. & Park,S. W. 1988. Simulating daily inflow and release rates for irrigation reservoirs(Ⅰ): modeling inflow rates by a linear reservoir model. Journal of the Korean Society of Agricultural Engineers 30(1), 50-62.

Koo, B. Y., Jung,I. W., Bae, D. H. 2006. A study on baseflow parameters estimation of TANK model. In Proceedings of the Korean Water Resources Association Conefrence, 1970-1974

Lee, J. K. & Park,S. J. 2003. Compartive study on eak flow prediction using mulitlayer neural network. Journal of The Korean Society of Civil Engineers (2B), 105-114.

Oh, C. R., Jin,Y. H., Kim,D. R., Park,S. C. 2008. Study on development of artificial neural network foecasting model using runoff, water quality data. Journal of Korean Water Resources Association 41(10), 1035-1044.

Ryu, J. H., Lee,S., Won,J. S., 2002. Weight determination of landslide factors using artificial neural networks. Economic and Environmental Geology 35(1), 67-74.

Seo, Y. M., Choi,E. H., Yeo, W. K. 2017. Reservoir water level forecasting using machine learning models. Journal of the Korean Society of Agricultural Engineers 59(3), 97-110.

Song, J. H., 2017. Hydrologic analysis system with multi-objective optimization dor agricultural watersheds. Ph.D. diss., Seoul, Ind.:Seoul National University.

Sugawara, E. O., 1972. Method of rainfall-runoff analysis, Kyouritsu Shuppan Co., Ltd., Tokyo, Japan.

Yeo, W. K., Seo,Y. M., Lee,S. Y., Ji,H. K. 2010. Study on water stage prediction using hybrid model of artificial neural network and genetic algorithm. Journal of Korean Water Resources Association 43(8), 721-731.

Yoon, Y. N., Yoo,C. S., Ahn,J. H., Yang,I. T. Koh,D. K. 1998. An analysis on the variation of long-term runoff characteristics of basin before and after dam construction: Ⅰ.Estimation of NWS-PC model parameters using GIS. Journal of The Korean Society of Civil Engineers 18(Ⅱ-5), 449-459.