Is the Deep-Learning Technique a Completely Alternative for the Hydrological Model?: A Case Study on Hyeongsan River Basin, Korea

Jaewon Kwak
Han River Flood Control Office

Heechan Han (hanheechan11@gmail.com)
Texas A&M AgriLife Research Blackland Research and Extension Center: Texas A&M University Blackland Research and Extension Center https://orcid.org/0000-0002-4078-0967

Soojun Kim
Inha University

Hung Soo Kim
Inha University

Research Article

Keywords: Deep-learning, LSTM model, Runoff simulation, TANK model

Posted Date: July 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-669773/v1

Is the deep-learning technique a completely alternative for the hydrological model?: A case study on Hyeongsan River Basin, Korea

Jaewon Kwak[a], Heechan Han[b*], Soojun Kim[c], Hung Soo Kim[c]

[a] Han River Flood Control Office, Ministry of Environment, Seoul, Republic of Korea

[b] Blackland Extension and Research Center, Texas A&M AgriLife, USA

[c] Department of Civil Engineering, Inha University, Incheon, Republic of Korea

Corresponding author: Heechan Han (hanheechan11@gmail.com)

Abstract

It is no doubt that the reliable runoff simulation for proper water resources management is essential. In the past, the runoff was generally modeled from hydrologic models that analyze the rainfall-runoff relationship of the basin. However, since techniques have developed rapidly, it has been attempted to apply especially deep-learning technique for hydrological studies as an alternative to the hydrologic model. The objective of the study is to examine whether the deep-learning technique can completely replace the hydrologic model and show how to improve the performance of runoff simulation using deep-learning technique. The runoff in the Hyeongsan River basin, South Korea from 2013 to 2020 were simulated using two models, 1) Long Short-Term Memory model that is a deep learning technique widely used in the hydrological study and 2) TANK model, and then we compared the runoff modeling results from both models. The results suggested that it is hard to completely replace the hydrological model with the deep-learning technique due to its simulating behavior and discussed how to improve the reliability of runoff simulation results. Also, a method to improve the efficiency of runoff simulation through a hybrid model which is a combination of two approaches, deep-learning technique and hydrologic model was presented.

Keywords: Deep-learning; LSTM model; Runoff simulation; TANK model

Declarations

Funding

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2017R1A2B3005695).

Conflicts of interest/Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and material (data transparency)

Not applicable

Code availability (software application or custom code)

Not applicable

Authors' contributions (optional: please review the submission guidelines from the journal whether statements are mandatory)

All authors contributed to the study conception and design. Conceptualization: [Jaewon Kwak], [Heechan Han], [Hung Soo Kim]; Methodology: [Jaewon Kwak], [Heechan Han]; Data collection: [Jaewon Kwak], [Soojun Kim], Formal analysis and investigation: [Heechan Han], [Soojun Kim]; Writing - original draft preparation: [Jaewon Kwak], [Heechan Han], [Soojun Kim]; Writing - review and editing: [Jaewon Kwak], [Hung Soo Kim]; Supervision: [Jaewon Kwak], [Hung Soo Kim]; All authors read and approved the final manuscript.

1. Introduction

One of the biggest issues in engineering field in the late 2010s is deep learning based data driven model. In the past, the use of water-related information is the key to securing the sustainability and resilience of water resources and has been recognized as an opportunity to innovate water governance in the future (Grossman et al., 2015). In addition, recently, numerous studies have attempted to apply data-driven models such as machine leaning or deep learning techniques for various purposes from water information analysis to hydrological, hydraulic analysis based on vast amount of water-related data (Sit et al., 2020). General hydrologic models conceptualize the physical characteristics of rainfall-runoff relationship by using mathematical equations (Singh, 1998). However, since a system conceptualized by human cannot clearly represent the complexity of natural systems (Marçais and Dreuzy, 2017), it cannot be guaranteed that the hydrologic model always simulates the rainfall-runoff relationships well. Also, a statistical model for simulating rainfall-runoff based on hydrological data has been used, but it also has limitations that indicate uncertainty and excessive computational power (Ardabili et al., 2019). Therefore, many hydrologic models have been used in consideration of different purpose and characteristics of the watershed (Lyubchich et al., 2019). However, in the late 2010s, the technology related to deep learning, a branch of machine learning, has been rapidly developed and it is showing a different trend than ever.

Deep learning technique is a type of black-box technique that can derive information desired by users from multidimensional datasets (Sengupta et al., 2020), and is known to be suitable for simulating complex rainfall-runoff relationship in hydrological study (Zhang et al., 2018). Accordingly, the number of cases of applying machine learning and deep learning techniques to the hydrology field increased more than 2.5 times from about 1,700 in 2010 to about 4,400 in 2018 (Ardabili et al., 2019). The types of applied techniques are gradually diversifying. For instance, Sit et al. (2020) introduced numerous deep learning technique applied to the hydrological studies such as Convolutional neural networks (CNN; LeCun, 1989), Generative adversarial networks (GAN; Goodfellow et al., 2014), Recurrent neural networks (RNN; Pollack, 1990), Long short-term memory (LSTM; Hochreiter and Schmidhuber, 1997), Gated recurrent unit networks (GRU; Cho et al., 2014), Nonlinear autoregressive models (NAR; Lin et al., 1996), Elman Network (Elman, 1990), Autoencoders (Rumelhart et al., 1985), Restricted boltzmann machines (RBM; Hinton, 2002) and deep belief networks (DBN; Hinton, 2009), Extreme learning machines (ELM; Huang et al., 2006), Deep Q networks (DQN; Mnih et al., 2013). Here, noteworthy is that there are various research results showing that the deep learning technique applied

to the rainfall-runoff simulation shows better results than the traditional approach. For example, in the case of LSTM model used in this study, Kratzert et al (2018) simulated daily runoff in the US watersheds and analyzed that LSTM showed better performance than the Sacramento Soil Moisture Accounting (SAC-SMA) model. Also, Damavandi et al.,(2019) found that the LSTM model was better than CaMa-Flood model for daily runoff simulation. In addition, the LSTM model shows better simulation accuracy than the Auto-regression model, which has been widely used to predict the hydrological time series (Qin et al., 2019), and it also showed better results than the Soil and Water Assessment Tool (SWAT) model which is widely used in the water resource field (Fan et al., 2020). Moreover, other previous studies showed that the LSTM model provides better performance of runoff simulation compared to the SIMHYD model, GR4J model (Bai et al., 2021), M5 Cubist model (Shortridge et al. 2016), Xinanjing model (Yin et al., 2019). However, in the hydrological field, even though many deep learning related studies have been conducted, a few studies showed that there is still limitation in deep learning models, so deep learning technique has been evaluated as an early stage (Ardabili et al., 2019). Also, other studies suggested that the deep learning model can provide higher reliability through ensemble approach. Recently, a hybrid model combining two or more deep learning techniques is also introduced to reduce the uncertainty in modeling process. For example, the inflow of the reservoir was predicted through the combined hybrid model of DBN and LSTM techniques (Luo et al., 2020), and a study was also conducted to build a hybrid model and improve the prediction efficiency by using the runoff data simulated by the VIC-CaMa-Flood model as the input data of the LSTM model (Yang et al., 2019). Based on these results provided from the previous studies, it seems that is it possible to replace the hydrologic model through the deep learning technique as suggested in numerous studies. However, since this possibility is an analysis through simple simulation efficiency, it is time to consider the actual hydrology simulation.

Therefore, in this study, the possibility of hydrological simulation through the deep learning technique is considered, and its strengths and limitations for simulating hydrological process are discussed. Also, a method to improve the simulation performance using deep learning technique was studied. For this, the runoff from 2013 to 2020 was simulated for the Hyeongsan River basin using LSTM, a deep learning technique widely used for hydrological simulation, and the TANK model. Through comparative analysis of observed runoff and simulated runoff from both models, the features of deep learning and hydrologic models were analyzed, and whether deep learning technique is able to completely replace hydrologic model was examined. Moreover, a method for

improving the runoff simulation through a hybrid model combining deep learning and hydrologic models was considered.

2. Method and Data Description

2.1 Study Area

Large scale basin is generally to be preferred for hydrological or another modeling due to the chance to prove their applicability. But, it also has limitation that there are many water related infrastructures such as dams that have a significant influence on the trend of outflow on basins. In this case, medium or small-scale basin is to be rather appropriate to clearly evaluate hydrological model and deep-learning technique for the simulation efficiency of rainfall-runoff process. Therefore, in this study, the Hyeongsan River basin was selected as the study area, where there are no large infrastructures and there are reliable observation data for a long period of time. The Hyeongsan River is located in the southeast of the Korean Peninsula, and has a basin area of 1,139.9 km^2, a boundary of 226.2 km, and a length of 57.4 km. The upstream area of the basin is mountainous and forested, and the main stream of the river forms an alluvial plain, which is mostly used as agricultural land, and a part of the urban area is located downstream. In terms of land use, forested areas accounted for 65.5%, agricultural land accounted for 29.1%, and urban areas accounted for 4.1% (Ministry of Land, Infrastructure and Transport, 2013). The hydro-meteorological data used in this study was collected from the observatories of the Ministry of Environment, and the precipitation data were converted into areal average values by Thiessen method. The runoff data was collected from the Hydrological Survey Annual Report (Ministry of Environment, 2019) and Water Resources Management Information System (Han River Flood Control Office, 2021). In the case of evaporation data, it was collected from the Korea Meteorological Administration. Figures 1 and 2 represent the information of study area and data used in this study.

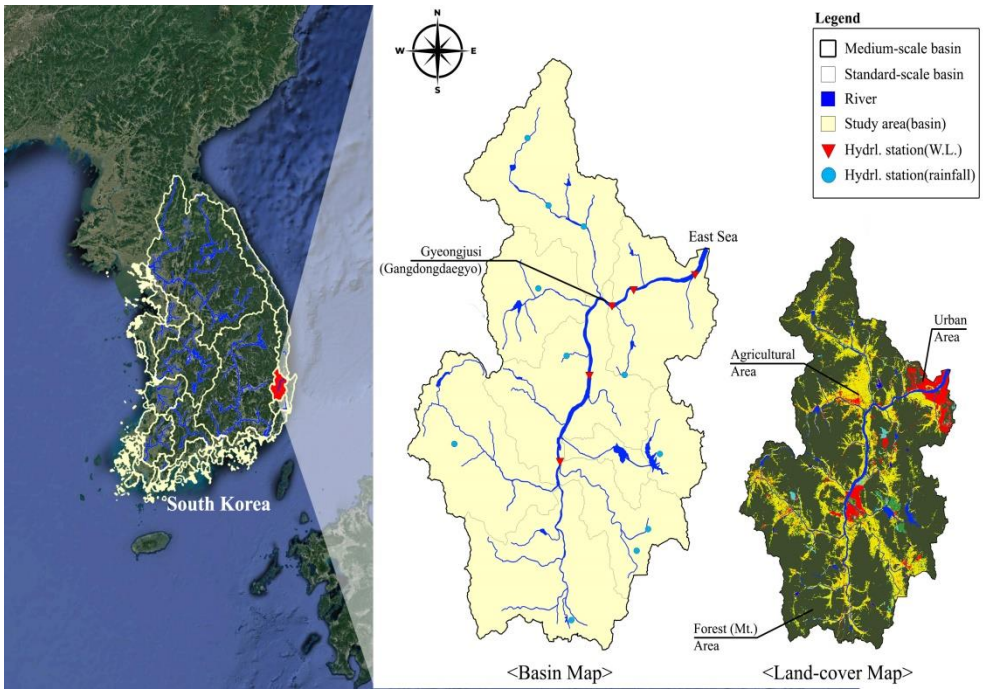


Fig. 1 Hyeongsan River basin area in South Korea. Blue and red points represent the observatories for hydrological data.

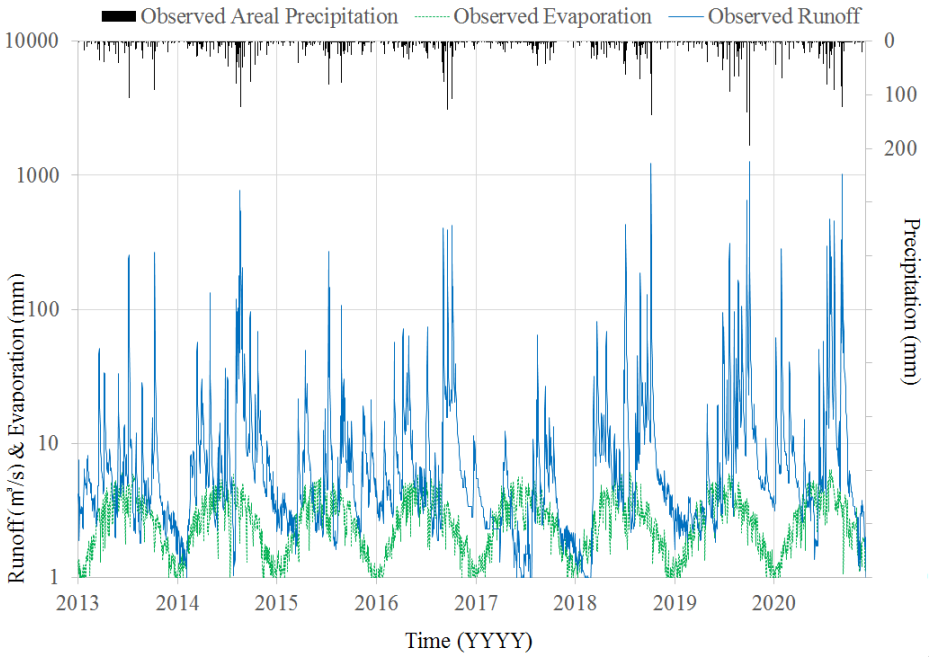


Fig. 2 Time series of obtained hydro-meteorological data including precipitation, evaporation, and runoff.

2.2 Models for Runoff Simulation

2.2.1 Hydrologic Model: TANK model

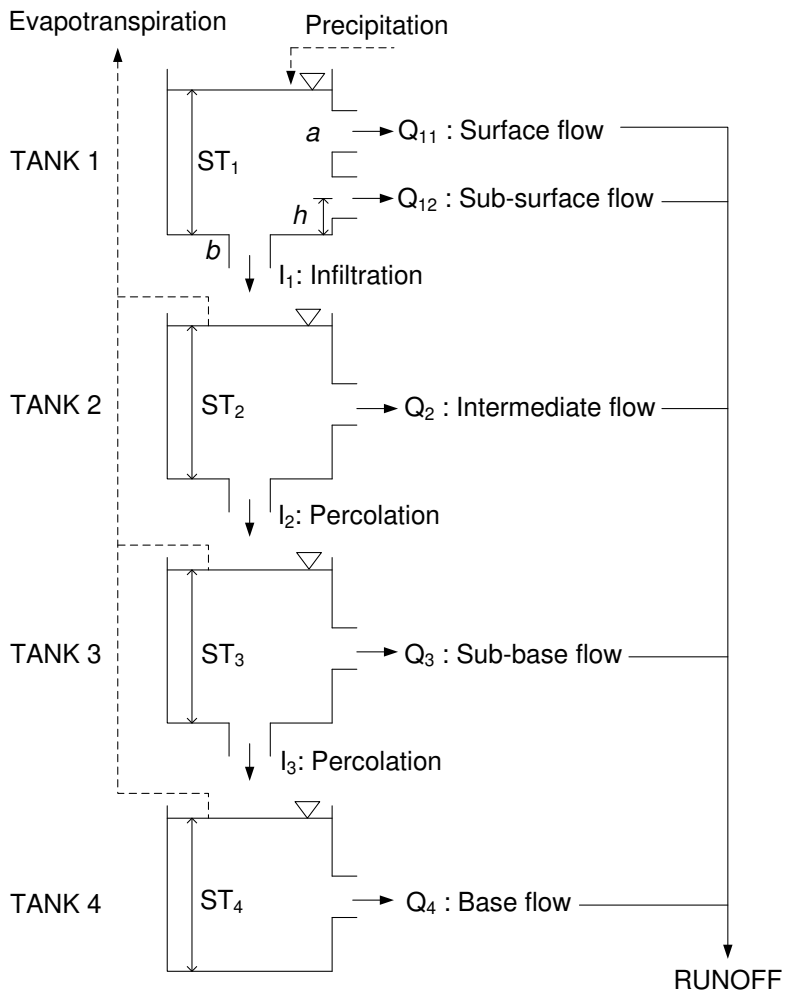


Fig. 3 Conceptual diagram of TANK model.

In order to simulate and understand the hydrological process based on rainfall-runoff relationship, it can be expressed as a vertical water flow. The model developed based on this process is the TANK model. The TANK model can simulate the runoff by considering the amount of water storage with a series of water storages (i.e. TANK). The TANK model is known to have simple algorithm and high accuracy compared to other hydrologic models for simulating rainfall-runoff process of watershed (Yokoo et al., 2001). In general, the TANK model is

based on a process of simulating evapotranspiration, infiltration, and outflow from each tank after precipitation enter to first tank. The upper two tanks simulate surface flow and intermediate flow generated from precipitation, and the lower two tanks are mainly used to simulate sub-base flow and base flow (Figure 3). The total runoff (Q_{total}) of the watershed simulated by the tank model is the sum of the runoff of each tank, and it is formulated as follows.

$$Q_{total} = Q_{11} + Q_{12} + Q_2 + Q_3 + Q_4 \tag{1}$$

The model contains three types of main parameters related to the runoff (*a*), height of the tank (*h*), and infiltration (*b*). The runoff and infiltration from each tank can be calculated using equations (2) and (3) as follows.

$$Q_n = a \times (ST_n - h) \tag{2}$$

$$I_n = b \times ST_n \tag{3}$$

Here, Q_n, ST_n denote the runoff and height of water in the nth tank. I_n indicates the amount of infiltration from the upper tank to the lower tank. Also, the height of water for the next time step (i.e. $t+1$) is calculated as follows (Song et al., 2017).

$$ST_{n,t+1} = ST_{n,t} + P_{t+1} - ET_{n,t+1} - I_{n,t} - Q_{n,t} \quad \text{for } n = 1 \tag{4}$$

$$ST_{n,t+1} = ST_{n,t} + I_{n-1,t} - ET_{n,t+1} - I_{n,t} - Q_{n,t} \quad \text{for } n = 2,3 \tag{5}$$

$$ST_{n,t+1} = ST_{n,t} + I_{n-1,t} - ET_{n,t+1} - Q_{n,t} \quad \text{for } n = 4 \tag{6}$$

Where, $ET_{n,t}$ is the evapotranspiration in the nth tank at time step *t* and is calculated by subtracting the sum of the evapotranspiration in the upper tank from the actual evapotranspiration ($ET_{a,t}$) and it is calculated as follows.

$$ET_{n,t} = ET_{a,t} - \sum_{n=1}^{i-1} ET_{n,t} \quad for\ ET_{n,t} < ST_{n,t} \tag{7}$$

$$ET_{n,t} = ST_{n,t} \quad for \ ET_{n,t} \geq ST_{n,t} \tag{8}$$

The actual evapotranspiration ($ET_{a,t}$) is calculated as combination of crop coefficient, soil water stress coefficient, and potential evapotranspiration. More detailed information about the process of TANK model is described in Song et al. (2017).

2.2.2 Deep Learning Model: LSTM model

The LSTM model, introduced by Hochreiter and Schmidhuber (1997), is one of the popular deep learning models and is based on a RNN model. Generally, the RNN model was used to examine the time sequence issues and predict the target variable considering past information and patterns. However, the RNN model showed some problems such as gradient vanishing and gradient exploding of the error slope when considering long-term dataset. In order to solve these limitations, the LSTM model was proposed and it was used for learning continuously composed data, mainly for purposes such as language translation and speech pattern recognition. In the hydrological studies, the LSTM model was applied for prediction of various hydrological variables such as runoff (Hu et al., 2018; Fan et al., 2020), water level (Zhang et al., 2020), soil properties (Adeyemi et al., 2018), and precipitation (Akbari Asanjan et al., 2018). These previous studies using LSTM model provided desirable predictive performances for hydrological forecasting in various regions.

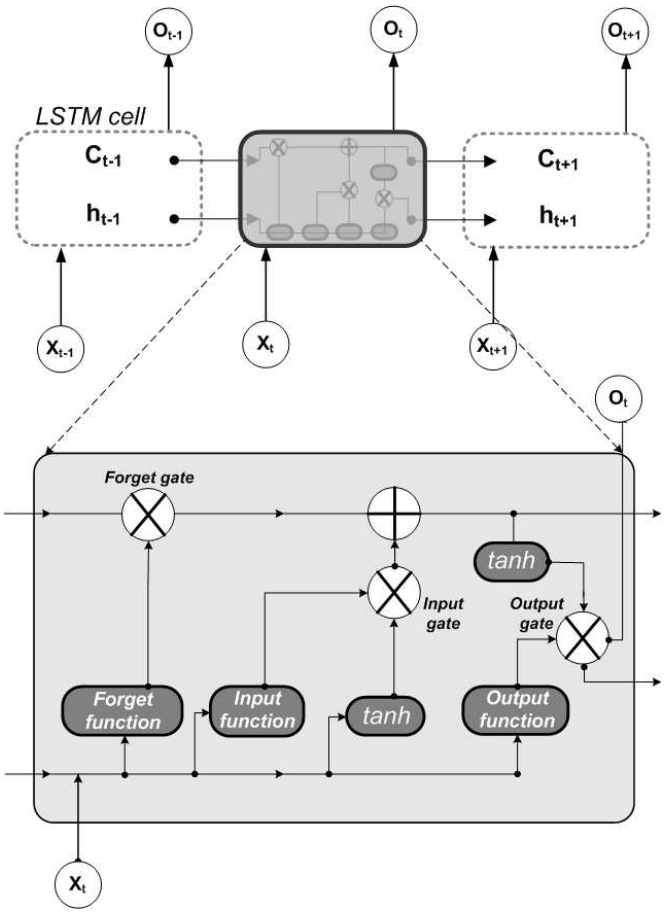


Fig. 4 Conceptual diagram of LSTM model.

The LSTM model contains multiple cells that can maintain information comes from previous state with time and nonlinear gates that can control the data flow. As shown in Figure. 4, the LSTM model contains three main gates such as forget gate, input gate, and output gate. These gates can determine how much of the information can be stored, removed, and give to the next cell. The LSTM algorithm is operated from an input sequence data (x_t) to final outcome (O_t) by looping through equations (9) – (14).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{9}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{10}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{11}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{12}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (13)$$

$$h_t = O_t \times \tanh(C_t) \quad (14)$$

Here, σ denotes the non-linear activation function. W_f, W_i, W_o and W_c are weight values of forget gate, input gate, output gate and memory cells, h_{t-1} is an output data from previous cell, x_t is current input data and b_f, b_i and b_o are bias vectors of each gate, respectively. Moreover, $\tilde{C}_t$ is the state of any cell generated from the activation function. In this study, Rectified Linear Unit (ReLU) is used as activation function.

3. Results

3.1 Performance of Runoff Simulation Using Two Models

This study simulated daily runoff using LSTM and TANK models and compared simulation performance of both models. The collected data for 8 years from 2013 to 2020 were applied for runoff simulation, and the data for period from 2013 to 2019 was used for the calibration, and data for 2020 was considered as validation of the models. In particular, in 2020, the observed precipitation during the rainy season in the Korean Peninsula was about 858 mm, which is about 2.5 times higher than average amount of precipitation, so it is suitable for verifying the modeling performance for runoff simulation.

In the case of the TANK model, additional parameter optimization was performed using a genetic algorithm (GA; Whitley, 1994) based on the model parameters presented in the Hyeongsan River Basic Plan Report (Ministry of Land, Infrastructure and Transport, 2013). For LSTM model, since it is necessary to determine proper hyperparameters such as number of layers and epoch size, this study referred to related studies (Kratzert et al., 2018; Han et al., 2021). In both model, precipitation and evaporation were used as input variables, and parameters were corrected based on the observed runoff data from the outlet point of the study area. Figure 5 represents the runoff resulted from TANK and LSTM models and observations in calibration period from 2013 to 2019.

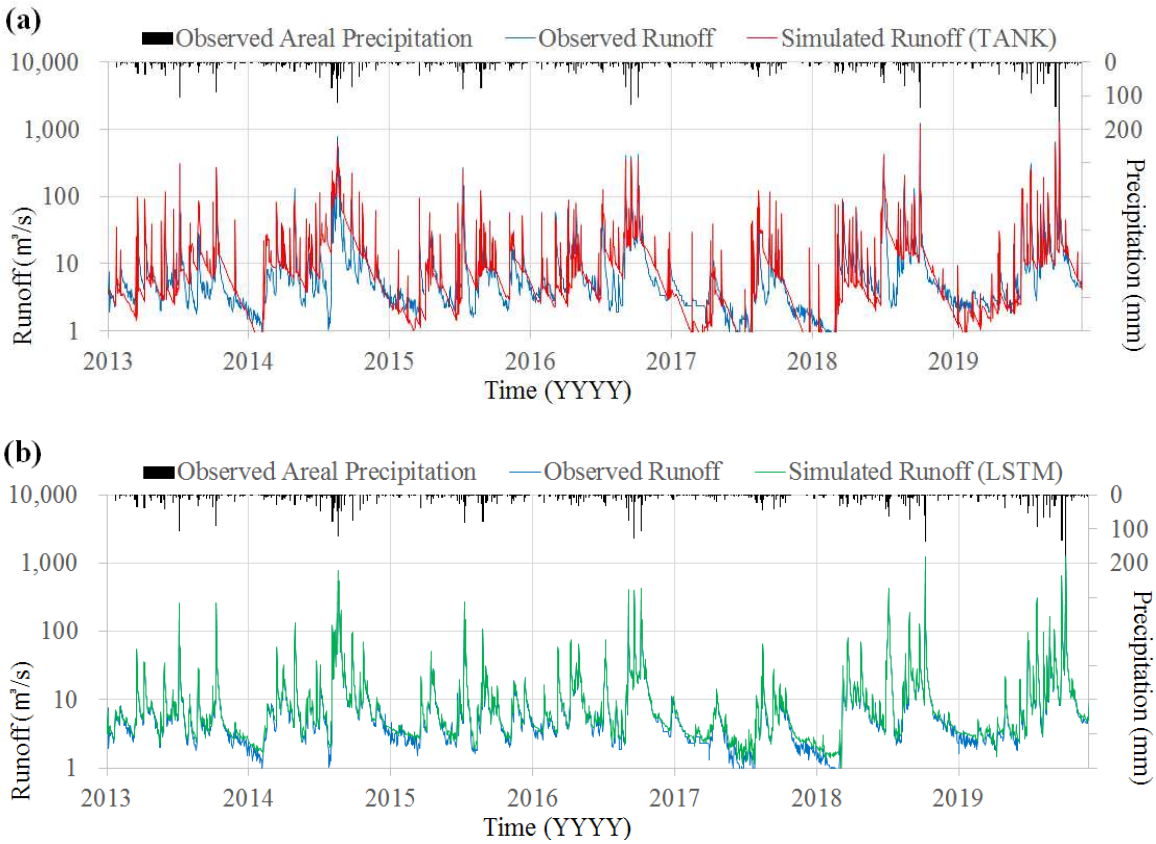


Fig. 5 Simulated runoffs in calibration period with both models: (a) TANK, (b) LSTM models.

From the results presented from both models, the LSTM model shows better simulation results than the TANK model. In addition, for quantitative comparison of models, R^2 and RMSE, which are widely used for the efficiency of hydrology simulation, were used as evaluation functions (Hyndman and Koehler, 2006). As a result of evaluation metrics, the TANK model showed R^2 = 0.83 and RMSE = 19.9m^3/s, whereas the LSTM showed very good results with R^2 = 0.98 and RMSE = 0.7m^3/s. In particular, the LSTM model simulated at a very similar level to the observed runoff time series, showing the simulation efficiency that is difficult to reach with the traditional hydrologic model. Therefore, although problems such as overfitting are pointed out (Horenko, 2020), the LSTM model seems to have strength over the hydrologic model in terms of simple simulation efficiency. To test how reasonably the LSTM and TANK models simulate the rainfall-runoff process in the watershed, the two models that have been calibrated were applied to the runoff simulation for 2020 (Figure 6). For validation (i.e. test) period, the TANK model showed R^2 = 0.65 and RMSE = 41.3m^3/s, LSTM model showed R^2 = 0.73 and RMSE =

39.3m^3/s, indicating the LSTM model provided better simulation performance. Therefore, for both periods for calibration and verification, the deep learning technique showed better efficiency than the hydrologic model.

3.2 Uncertainty in Runoff Simulation Using Deep Learning Technique

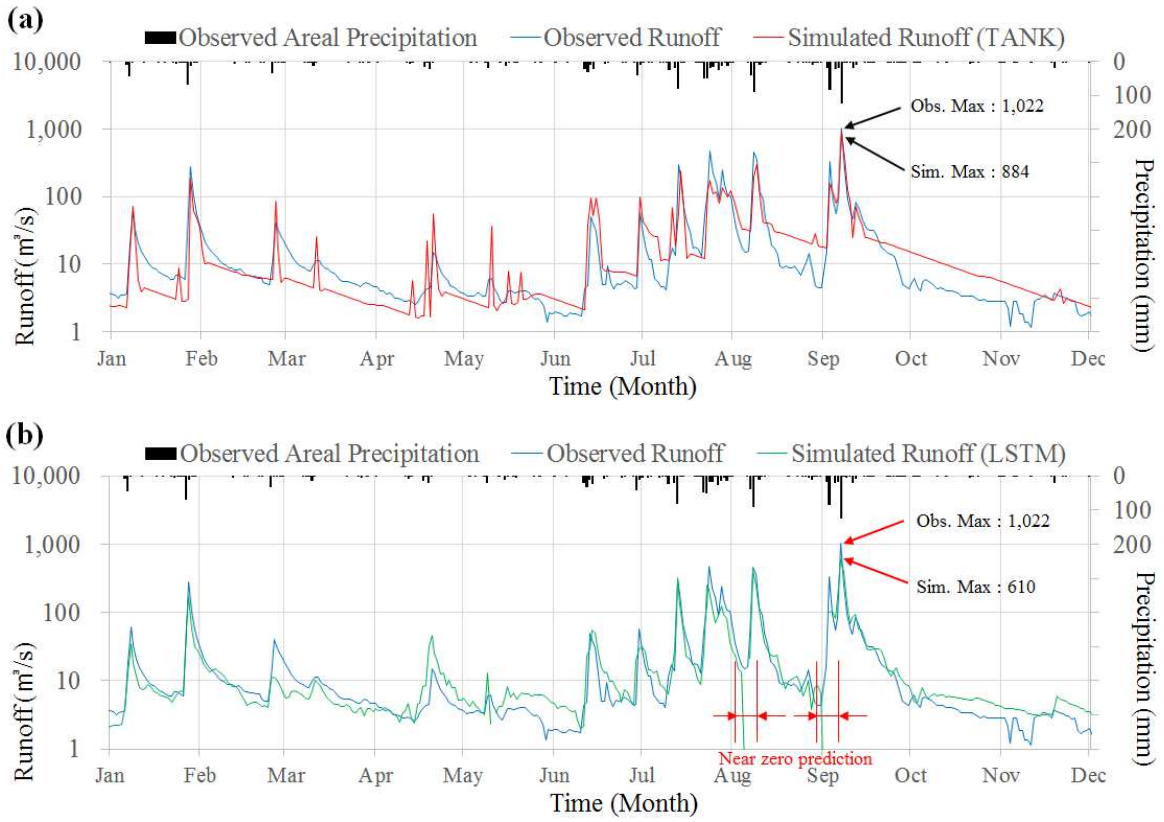
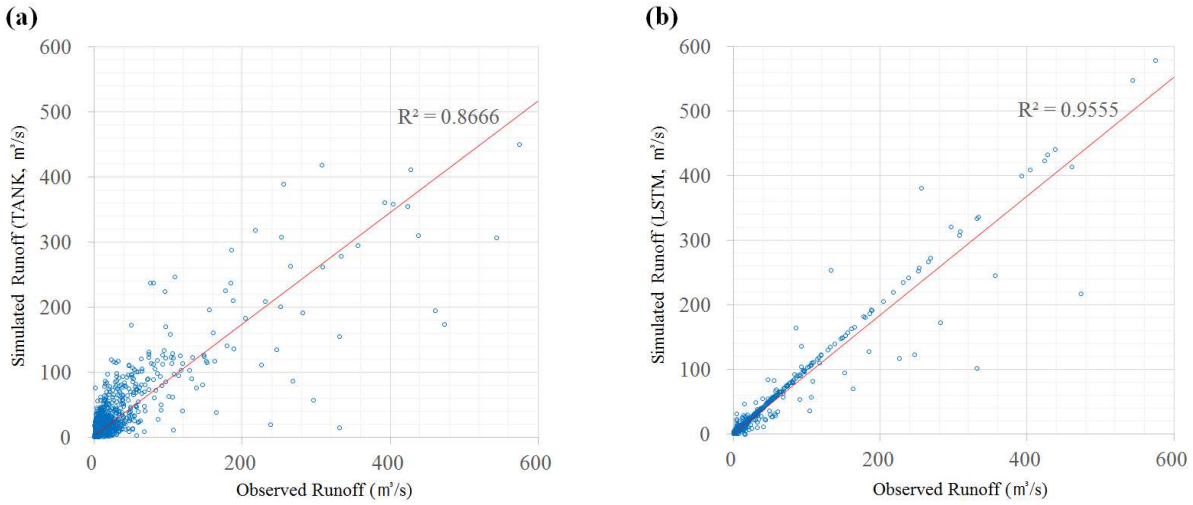


Fig. 6 Simulated runoff in validation period with each model: (a) TANK, (b) LSTM models.

In the previous section, runoff from 2013 to 2020 was simulated through TANK and LSTM models, and the overall performance of the deep learning technique (i.e. LSTM model) showed better than the hydrologic model (i.e. TANK model). Unlike the LSTM model that provided very good results in the calibration period (2013 to 2019), the results for validation period showed relatively lower efficiency. Nevertheless, it shows better simulation efficiency compared to the TANK model, so it is considered that the rainfall-runoff simulation of the deep learning technique has sufficient applicability. However, simulation results of the LSTM for validation period, there were two limitations that cannot be seen in the simulation results of the TANK model. First, the overall LSTM showed an efficient simulation trend, but for the peak runoff simulation, the accuracy was lower

than that of the TANK model (Figure 6). The second limitation is that in some periods, the runoff was simulated to decrease to zero even through the actual runoff increases rapidly (Figure 6). This is a characteristic that cannot occur in the conceptual model of the rainfall-runoff process. The first cause of these limitations is the overfitting problem. In order to avoid the overfitting problem, a dropout ratio of 0.2 was set by referring to existing studies (Shah et al., 2018). Nevertheless, compared to the high simulation performance of the calibration period, overfitting may have caused to lower accuracy for validation period. Another cause may be that the rainfall pattern in 2020 is different from the rainfall pattern in the calibration periods (2013-2019), so it may be a result of an unlearned event. In fact, during the rainy season from August to September 2020, unusual precipitation was observed, and it was exceeding the past maximum precipitation in the past 30 years (Kim et al., 2020). However, since it is difficult to understand the learning process of deep learning technique (Sengupta et al., 2020), it is hard to determine which cause is the dominant influence in uncertainty.



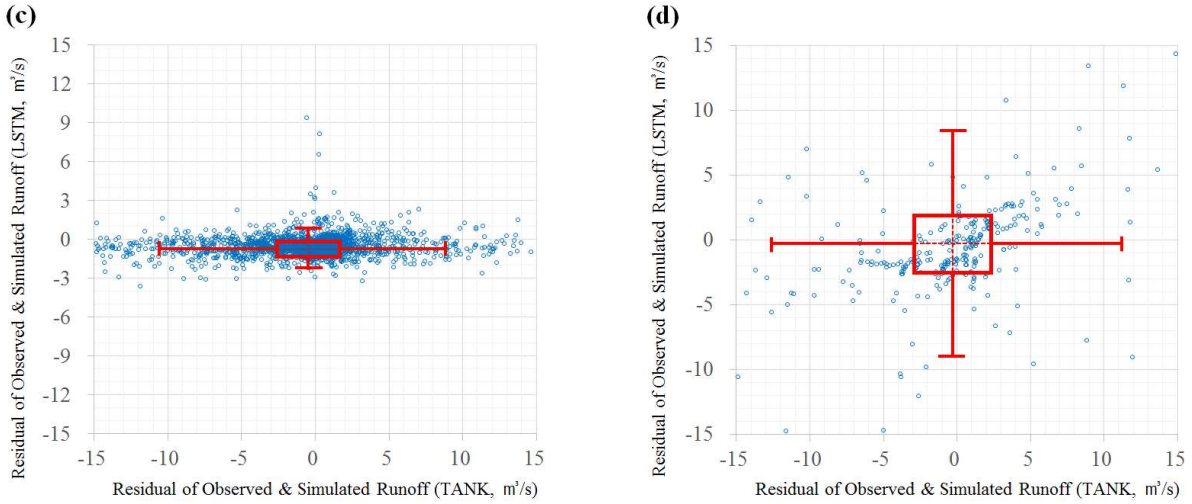


Fig. 7 Scatter plots of observed runoff against simulated runoff for whole period (2013~2020), and crossed box-plot of residual between observed and simulated runoff with TANK and LSTM; (a) Scatter plot of runoff with TANK model; (b) Scatter plot of runoff with LSTM; (c) Crossed box-plot of residual in calibration period (2013~2019); (d) Crossed box-plot of residual in validation period (2020).

Deep learning is called as a black-box technique that allows users to extract or derive information they want from multi-dimensional complex datasets through autonomous learning without specifying them (Sengupta et al., 2020). In other words, it means that it is impossible to know what patterns and structures the deep learning technique has learned about the rainfall-runoff relationship during calibration period. In contrast, in the case of the TANK model, runoff can be simulated through the conceptual rainfall-runoff process with clear equations and mathematical theories. As shown in Figure 7, the observed and simulated flow scatter plots and crossed boxplots show the characteristics of the simulated runoff more clearly. As shown in the Figure 7(a) and (b), it was found that the LSTM model shows better simulation performance than the TANK model. Therefore, in terms of overall simulation efficiency, it is obvious that LSTM is better than TANK model. However, the crossed boxplots for the errors of the two models shown in Figure 7(c) and (d) show a problem in another aspect. The TANK model shows an error distribution (range of boxes) within ± 12.0 for both the calibration and verification period. As described in the section 2.2.1, the TANK model simulates the runoff generation considering the runoff height and runoff coefficient in the four connected storage tanks, and it has a structure that conceptualizes the rainfall-runoff relationship of the entire watershed. Therefore, although the overall performance of the rainfall simulation may

decrease, a constant runoff pattern is shown within the conceptualized structure. In the case of the LSTM model, it was found that the residual range was between -2.0 and 1.0 during the calibration period, but the distribution range reached ± 9.0 during the verification period. This irregular distribution range suggests that the rainfall-runoff structure learned by the LSTM method through the calibration period (2013-2019) is not partially suitable for the verification event in 2020. However, since users cannot conceptually understand the learning structure of deep learning, it is difficult to find the main cause of the problem about the suitability and uncertainty in deep learning model process. In fact, many studies have suggested that "Deep learning shows higher simulation efficiency than hydrologic models", but did not suggest clearly "Reasons why deep learning shows higher simulation efficiency than hydrologic models".

4. Discussions

4.1 Is the deep-learning technique a completely alternative for the hydrological model?

Based on features of deep learning technique presented in previous sections, when flood or drought forecasting is performed through a deep learning, "Critical malfunction" that performs an incorrect prediction even though a dangerous level of flood or drought will occur due to unlearned hydrological events" can occurred. For example, as shown in Figure 6 (b), it is the result of simulating a different trend of runoff decrease from actual trend or a lower peak runoff value compared to the actual peak runoff. In fact, these problems have already been pointed out in previous studies. For instance, Sit et al. (2020) commented that "ethics in disaster management and public planning may arise due to the automation of hydrologic modeling with deep learning" through the title "Ethics in deep learning applications" in the research.

In addition, it was suggested that the impact on individuals, the environment, and society should be prioritized according to decision making in the water management field (Ewing and Demir, 2020). Moreover, Nearing et al. (2020) also claimed that "Most hydrologists would agree that the calibrated conceptual model would provide better performance than the data-based model for simulations under changing conditions". It becomes clearer when these research cases are applied to the field of water management. For example, suppose that a flood or drought forecast is performed using deep learning, and the forecasted result is incorrect due to unusual weather conditions such as extreme precipitation. Incorrect flood forecasting may result in damage to human life or property, or personal water use may be unreasonably restricted due to incorrect prediction results

(Ewing and Demir, 2020).

Also, if the user cannot capture the structure learned by deep learning, it would be hard to understand or evaluate the results of the deep learning technique or its uncertainty for runoff simulation. Therefore, if it is not possible to predict and examine the suitability of simulated runoff by deep learning, the responsibilities will also be unclear (Campolo and Crawford, 2020; Orr and Davis, 2020). Jobin et al. (2019) suggested that five ethics containing transparency, justice and fairness, non-maliciousness, responsibility, and privacy are necessary when applying artificial intelligence technology such as deep learning. If a "Critical malfunction" occurs due to incorrect prediction as shown in Figure 6 (b), it is also necessary to consider how to deal with ethical issues based on transparency and responsibility for error occurrence in simulation results.

Therefore, in the field of hydrological study, it will be difficult for deep learning to completely replace the hydrologic model, and solutions to the logical and ethical problems considered above or social consensus must be prioritized. In conclusion, this study proposes the following three points in order to replace the hydrologic model or to perform hydrologic modeling using deep learning so far. 1) Application to simple research purposes or risk analysis, etc., not to fields that directly affect people's lives, such as flood and drought forecasting or water resource management, 2) Building an effective decision-making system that allows users to review or evaluate the simulation results of deep learning, and 3) Prioritizing social consensus on who is responsible for ethical review of simulated results.

4.2 Alternatives to Deep Learning for Hydrological Study: Hybrid Model Combining Deep Learning and Hydrologic Models

In this study, we discussed that the deep learning technique cannot completely replace the hydrologic model, even though it shows superior simulation efficiency compared to the hydrologic model. However, the superior efficiency of deep learning compared to the hydrological model is clearly an offer we cannot refuse. Therefore, this study proposes a hybrid model and methodology to utilize and combine both the conceptualized rainfall-runoff process of the hydrologic model (i.e. TANK model) and the accuracy and effectiveness of deep learning technique (i.e. LSTM). For this, it is necessary to discuss the uncertainty of the hydrologic model first. One of the main reasons for the uncertainty in the hydrologic model is that the rainfall-runoff process is not fixed in time and space (Montanari and Grossi, 2008). Therefore, it is difficult to eliminate entire uncertainty in the simulation

results of the hydrologic model, and it is important to quantify it as much as possible (Krzysztofowicz, 2001).

In the past, quantification or analysis of uncertainty was mainly performed through statistical methods, but we already have "deep learning technique" in our hands, which is a powerful analysis tool for complex and high-dimensional data. If the uncertainty generated from the hydrologic model is caused by temporal and spatial changes in the rainfall-runoff process according to the weather or watershed characteristics, it would be a very suitable material for deep learning technique pursuing the "black-box method that derives user-desired information from complex multidimensional data (Sengupta et al., 2020)". Therefore, assuming that the hydrologic model simulates the rainfall-runoff process of the watershed, the residuals occurred in the model can be caused by temporal and spatial characteristics in the physical response of the watershed, and it will be possible to supplement using deep learning technique. It is formulated as follows.

$$runoff_{basin} = runoff_{model} + runoff_{uncertainty} \tag{15}$$

Here, $runoff_{basin}$ is simulated runoff of the watershed, $runoff_{model}$ denotes runoff simulated from the hydrologic model considering fixed rainfall-runoff relationship in the watershed, and $runoff_{uncertainty}$ is the residual between observation and results from the model estimated from the deep learning technique that takes account of patterns of residual in the modeling results. Therefore, the runoff generation in the watershed can be simulated by mixing the simulation result of the rainfall-runoff process through the hydrologic model and the estimation of the spatiotemporal uncertainty of the modeling process using the deep learning technique. These two methods can be used as a complementary hybrid model. This study applied this hybrid model for the Hyeongsan River basin and evaluated its applicability. In order to examine the uncertainty of each model, the time series of residual, which is difference between the simulated and observed runoff, was analyzed for calibration and validation periods (Figure 8).

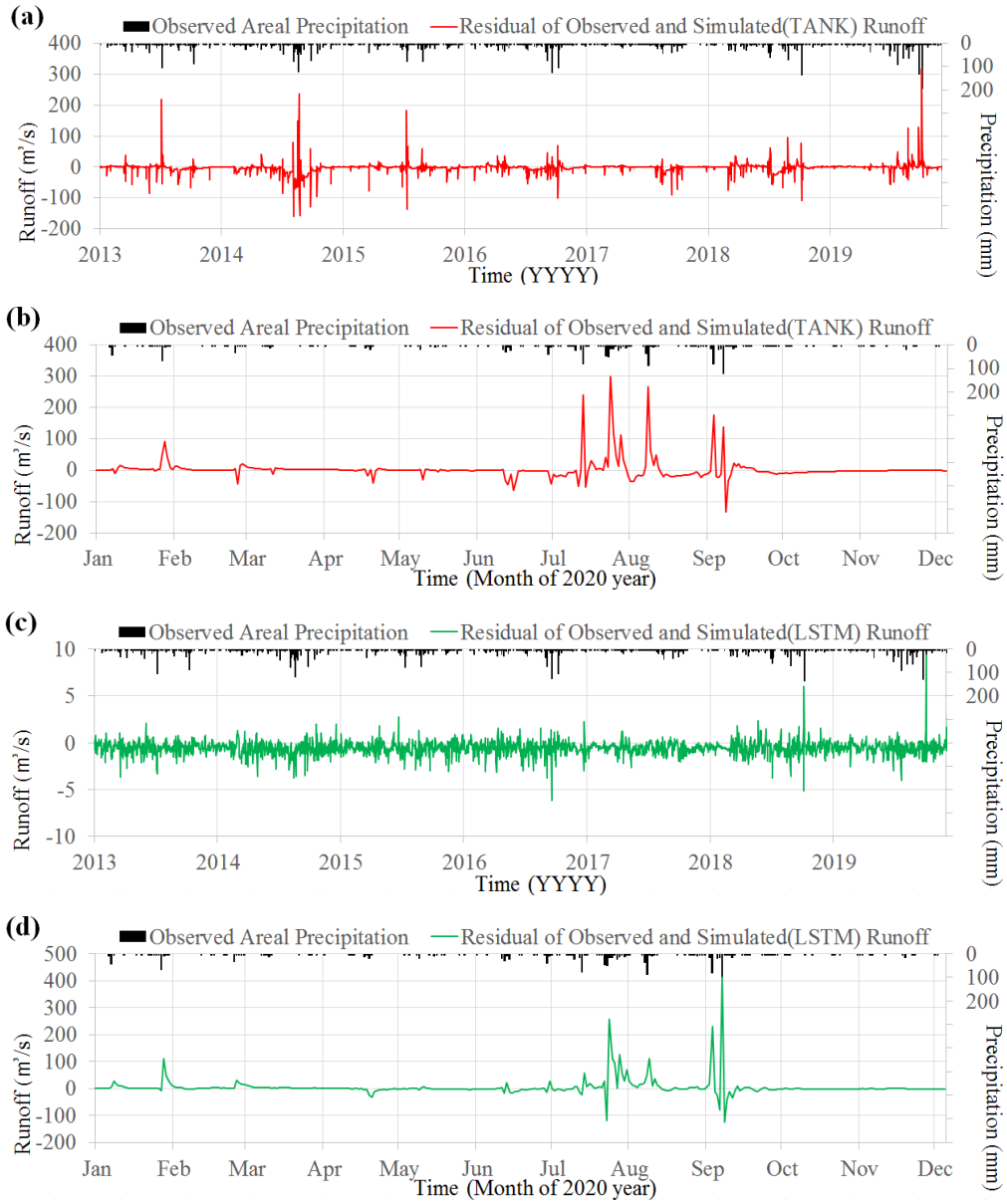


Fig. 8 Residual time series between observed and simulated runoff; (a) TANK model in calibration period; (b) TANK model in validation period; (c) LSTM model in calibration period; and (d) LSTM model in validation period.

As shown in Figure 8 (a) and (b), it was found that the residuals of the TANK model show a similar trend as described in Figure 7. On the other hand, the residual time series of the LSTM model represented in Figure 8 (c) and (d) showed a different trend between the calibration and verification periods. The residuals for calibration

period showed high accuracy within ±5 m^3/s, but for the verification period, it provided a large residual, showing a similar trend to the TANK model. Therefore, the residual of the TANK model cannot be conceptualized due to its non-linearity and complexity, but it allows us to grasp certain patterns and trends, and these residuals can be estimated using the deep learning technique (i.e. LSTM model). In this study, for the residual time series of the TANK model during the calibration period (2013-2019), the LSTM model under the same conditions was trained, and the verification was also performed in 2020 (Figure 9).

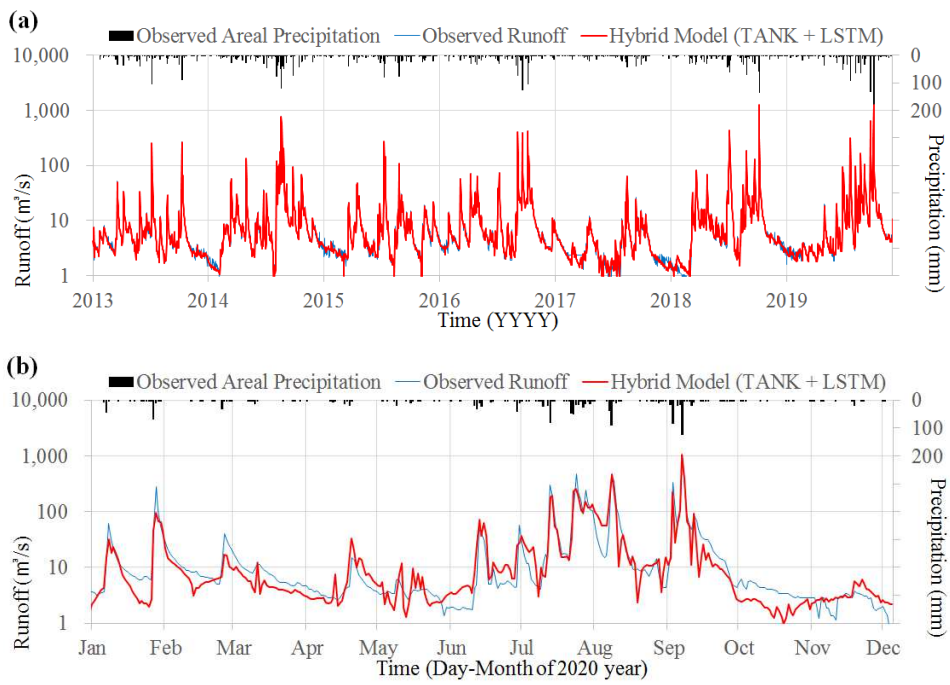


Fig. 9 Simulated runoff result with hybrid (TANK and LSTM) model; (a) calibration period (2013~2019); (b) validation period (2020).

Figure 9 indicated that the hybrid model combining simulated runoff from the TANK model and estimated residuals from the LSTM model provided desirable performance for both of calibration and verification periods. The evaluation metrics also showed good performance with R^2 = 0.98, RMSE = 0.6 m^3/s for calibration period, and R^2 = 0.90, RMSE = 25.1 m^3/s for verification period compared to the results when two models are used independently (R^2 = 0.65, RMSE = 41.3 m^3/s for TANK model and R^2 = 0.73, RMSE = 39.3 m^3/s for LSTM model). For the calibration period, the simulation efficiency was almost the same as the result when using only

the LSTM model, but for the verification period, it showed better simulation efficiency compared to two models. This proves that the hybrid model proposed in this study can be applied to the Hyeongsan River basin for runoff simulation.

The limitation of the proposed hybrid model is that the uncertainty still exists in runoff simulation results. The hybrid model tried to reduce the uncertainty in the runoff simulation by combining the runoff simulated from the TANK model and the residuals estimated from the LSTM model. However, uncertainty in the runoff generation still exists even though the magnitude of error has been reduced. In addition, the "Critical malfunction" of the deep learning technique can occurred when using LSTM model as presented previous section. Therefore, additional processes are needed to minimize the overall uncertainty in residual estimation from deep learning when using the hybrid model.

For example, if the uncertainty estimated from the deep learning exceeds the range of the runoff simulation result from the hydrologic model, incorrect errors can be ignored. In addition, there are essential alternatives to solve all the problems considered in this paper. If deep learning can conceptually understand and learn rainfall-runoff process, it can provide higher simulation efficiency and show a similar structure to the hydrologic model, and this will be the ultimate answer.

5. Conclusions

This study conducted the runoff simulation using a deep learning technique (i.e. LSTM model) that has recently been spotlighted, examined whether it could replace the hydrological model that used in numerous studies, and suggested a novel approach how to improve the performance of runoff simulation with minimum errors. For this, we simulated and compared the runoff generation from 2013 to 2020 in the Hyeongsan River basin in Korea peninsula using LSTM model and the TANK model and it was considered that it is impossible to completely replace the hydrologic model by deep learning technique. Also, a method for improving the overall performance of runoff simulation through a hybrid model of deep learning and hydrologic model was proposed. The results obtained in this study are briefly summarized as follows.

1. For the Hyeongsan River basin, runoff simulation was performed using hydrological data from 2013 to 2020, LSTM and TANK models. The result of the simulation showed that overall performance of runoff modeling through the deep learning (i.e. LSTM model) is better than the hydrologic model (i.e.

TANK model).

2. Despite the high simulation result, unlike the hydrologic model that conceptually simulate the rainfall-runoff process of the watershed, the deep learning technique is built through process that users cannot understand, so prediction errors may occur due to input data with a different features. This is one of the limitations the deep learning contains. Therefore, it would be difficult to completely replace the hydrologic model in the field of hydrological modeling using deep learning technique.
3. This study proposes the following three points in order to replace the hydrologic model or to perform hydrologic modeling using deep learning so far. 1) Application to simple research purposes or risk analysis, etc., not to fields that directly affect people's lives, such as flood and drought forecasting or water resource management, 2) Building an effective decision-making system that allows users to review or evaluate the simulation results of deep learning, and 3) Prioritizing social consensus on who is responsible for ethical review of simulated results.
4. This study also proposed a hybrid model to utilize both the conceptualized rainfall-runoff simulation of the hydrologic model and the high accuracy of the deep learning. The proposed model showed desirable performance of runoff simulation than LSTM and TANK models.

In the discussion section of the IAHS "Unsolved Problems in Hydrology" (Blöschl et al., 2019), the question "Will deep learning play a real role in hydrology modeling in the future?" was proposed (Nearing et al., 2021). As conclusion, this study would like to answer this as follows. "Deep learning will be an effective solution that hydrologist cannot refuse. However, 'at least for now' it is hard to be a complete alternative for hydrological model".

Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2017R1A2B3005695).

References

Adeyemi O, Grove I, Peets S, Domun Y, Norton T (2018) Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling. Sensors 18(10):3408. https://doi.org/10.3390/s18103408

Akbari Asanjan A, Yang T, Hsu K, Sorooshian S, Lin J, Peng Q (2018) Short-term precipitation forecast based on the PERSIANN system and LSTM recurrent neural networks. J Geophys Res: Atmos 123(22):12-543. https://doi.org/10.1029/2018JD028375

Ardabili S, Mosavi A, Dehghani M, Várkonyi-Kóczy A R (2019) Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review. In International Conference on Global Research and Education, 52-62, Springer, Cham.

Bai P, Liu X, Xie J (2021) Simulating runoff under changing climatic conditions: a comparison of the long short-term memory network with two conceptual hydrologic models. J Hydrol 592:125779. https://doi.org/10.1016/j.jhydrol.2020.125779

Blöschl G, Bierkens M F, Chambel A, Cudennec C, Destouni G, Fiori A, ..., Renner M (2019) Twenty-three unsolved problems in hydrology (UPH)–a community perspective. Hydrol Sci J 64(10):1141-1158. https://doi.org/10.1080/02626667.2019.1620507

Campolo A, Crawford K (2020) Enchanted determinism: Power without responsibility in artificial intelligence. Engaging Sci Technol and Soc 6:1-19. https://doi.org/10.17351/ests2020.277

Cho K., Van Merriënboer B, Bahdanau D, Bengio Y (2014) On the Properties of Neural Machine Translation: Encoder Decoder Approaches. arXiv preprint arXiv:1409.1259.

Damavandi H G, Shah R, Stampoulis D, Wei Y, Boscovic D, Sabo J (2020) Accurate prediction of streamflow using long short-term memory network: a case study in the Brazos River Basin in Texas. Int J Environ Sci and Dev 10(10):294–300.

Elman J L, (1990) Finding structure in time. Cognit Sci 14(2):179–211. https://doi.org/10.1207/s15516709cog1402_1

Ewing G, Demir I (2020) An ethical decision-making framework with serious gaming: smart water case study on flooding. EarthArXiv. doi:10.31223/osf.io/y873 g.

Fan H, Jiang M, Xu L, Zhu H, Cheng J, Jiang J (2020) Comparison of long short term memory networks and the hydrological model in runoff simulation. Water 12(1):175. https://doi.org/10.3390/w12010175

Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. MIT Press, Cambridge, MA.

Grossman D, Buckley N, Doyle M (2015) Data Intelligence for 21st Century Water Management: A Report From the 2015 Aspen-Nicholas Water Forum. Aspen-Nicholas Water Forum. Available from

https://www.aspeninstitute.org/publications/data-intelligence-21st-century-water-management-report-2015-aspen-nicholas-water-forum.

Han H, Choi C, Jung J, Kim H S (2021) Application of sequence to sequence learning based LSTM model (LSTM-s2s) for forecasting dam inflow. J Korean Water Resour Assoc 54(3):157-166. https://doi.org/10.3741/JKWRA.2021.54.3.157

Han River Flood Control Office (HRFCO) (2021) Ministry of Environment, http://www.hrfco.go.kr/main.do.

Hinton G E (2002) Training products of experts by minimizing contrastive divergence. Neural Computation 14(8):1771–1800. https://doi.org/10.1162/089976602760128018

Hinton G E (2009) Deep belief networks. Scholarpedia 4(5):5947.

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Horenko I (2020) On a Scalable Entropic Breaching of the Overfitting Barrier for Small Data Problems in Machine Learning. Neural Comput 32(8):1563-1579. https://doi.org/10.1162/neco_a_01296

Hu C, Wu Q, Li H, Jian S, Li N, Lou Z (2018) Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. Water 10(11):1543. https://doi.org/10.3390/w10111543

Huang G B, Zhu Q Y, Siew C K (2006) Extreme learning machine: theory and applications. Neurocomputing 70(1–3):489–501. https://doi.org/10.1016/j.neucom.2005.12.126

Hyndman R J, Koehler A B (2006) Another look at measures of forecast accuracy. Int J Forecasting 22(4):679-688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nat Mach Intell 1(9):389-399. https://doi.org/10.1038/s42256-019-0088-2

Kim Y T, Park M, Kwon H H (2020) Spatio-Temporal Summer Rainfall Pattern in 2020 from a Rainfall Frequency Perspective. J Korean Soc Disaster Secur 13(4):93-104, https://doi.org/10.21729/ksds.2020.13.4.93

Kratzert F, Klotz D, Brenner C, Schulz K, Herrnegger M (2018) Rainfall–runoff modelling using long short-term memory (LSTM) networks. Hydrol Earth Syst Sci 22(11):6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Krzysztofowicz R (2001) Integrator of uncertainties for probabilistic river stage forecasting: precipitation-dependent model. J Hydrol 249(1-4):69-85. https://doi.org/10.1016/S0022-1694(01)00413-9

LeCun Y (1989) Generalization and network design strategies. Connectionism in Perspective 19:143–155.

Lin T, Horne B G, Tino P, Giles C L (1996) Learning longterm dependencies in NARX recurrent neural networks. IEEE Trans. Neural Networks 7(6):1329–1338. doi: 10.1109/72.548162.

Luo B, Fang Y, Wang H, Zang D (2020) Reservoir inflow prediction using a hybrid model based on deep learning. In IOP Conference Series: Materials Science and Engineering. 715(1):012044.

Lyubchich V, Newlands N K, Ghahari A, Mahdi T, Gel Y R (2019) Insurance risk assessment in the face of climate change: Integrating data science and statistics. Wiley Interdisciplinary Reviews: Computational Statistics 11(4):e1462. https://doi.org/10.1002/wics.1462

Marçais J, de Dreuzy J R (2017) Prospective interest of deep learning for hydrological inference. Groundwater 55 (5):688–692.

Ministry of Environment (2019) Hydrological Annual Report in Korea. Sejong: Ministry of Environment.

Ministry of Land, Infrastructure and Transport (MOLIT) (2013) Basic river plan for Hyeongsan River, Publication No. DM00034815, Busan Construction and Management Administration, Busan, Korea, 25-50.

Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M (2013) Playing Atari with Deep Reinforcement Learning. arXiv preprint arXiv:1312.5602.

Montanari A, Grossi G (2008) Estimating the uncertainty of hydrological forecasts: A statistical approach. Water Resour Res 44(12). https://doi.org/10.1029/2008WR006897

Nearing G, Kratzert F, Pelissier C, Klotz D, Frame J, Gupta H (2020) Machine Learning is Central to the Future of Hydrological Modeling. EGU General Assembly, Online, 4–8 May 2020, EGU2020-6111, https://doi.org/10.5194/egusphere-egu2020-6111, 2020

Nearing G S, Kratzert F, Sampson A K, Pelissier C S, Klotz D, Frame J M, ..., Gupta H V (2021) What role does hydrological science play in the age of machine learning?. Water Resour Res 57(3):e2020WR028091. https://doi.org/10.1029/2020WR028091

Orr W, Davis J L (2020) Attributions of ethical responsibility by Artificial Intelligence practitioners. Inf Commun and Soc 23(5):719-735. https://doi.org/10.1080/1369118X.2020.1713842

Pollack J B (1990) Recursive distributed representations. Artif Intell 46(1–2):77–105. https://doi.org/10.1016/0004-3702(90)90005-K

Qin J, Liang J, Chen T, Lei X, Kang A (2019) Simulating and predicting of hydrological time series based on

tensorFlow deep learning. Polish Int J Environ Stud 28(2):796–802. https://doi.org/10.15244/pjoes/81557

Rumelhart D E, Hinton G E, Williams R J (1985) Learning representations by back-propagating errors. Nature 323(6088):533–536. https://doi.org/10.1038/323533a0

Sengupta S, Basak S, Saikia P, Paul S, Tsalavoutis V, Atiah F, Ravi V, Peters (2020) A review of deep learning with special emphasis on architectures, applications and recent trends. Knowledge-Based Systems 194:105596. https://doi.org/10.1016/j.knosys.2020.105596

Shah D, Campbell W, Zulkernine F H (2018) A comparative study of LSTM and DNN for stock market forecasting. In 2018 IEEE International Conference on Big Data (Big Data). 4148-4155. doi: https://doi.org/10.1109/BigData.2018.8622462

Singh V P (1988) Hydrologic systems. Volume I: Rainfall-runoff modeling. Prentice Hall, Englewood Cliffs New Jersey. 480.

Sit M, Demiray B Z, Xiang Z, Ewing G J, Sermet Y, Demir I (2020) A comprehensive review of deep learning applications in hydrology and water resources. Water Sci Technol 82(12):2635-2670. https://doi.org/10.2166/wst.2020.369

Shortridge J E, Guikema S D, Zaitchik B F (2016) Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. Hydrol Earth Syst Sci 20(7):2611-2628. https://doi.org/10.5194/hess-20-2611-2016

Song J H, Her Y, Park J, Lee K D, Kang M S (2017) Simulink implementation of a hydrologic model: A Tank model case study. Water, 9(9):639. https://doi.org/10.3390/w9090639

Whitley D (1994) A genetic algorithm tutorial. Stat and Comput 4(2):65-85. https://doi.org/10.1007/BF00175354

Yang T, Sun F, Gentine P, Liu W, Wang H, Yin J, ..., Liu C (2019) Evaluation and machine learning improvement of global hydrological model-based flood simulations. Environ Res Lett 14(11):114027. https://doi.org/10.1088/1748-9326/ab4d5e

Yin Z, Liao W, Wang R, Lei X (2019) Rainfall-runoff modelling and forecasting based on long short-term memory (LSTM). SN Water Transf Water Sci Technol 6:1-9.

Yokoo Y, Kazama S, Sawamoto M, Nishimura H (2001) Regionalization of lumped water balance model parameters based on multiple regression. J Hydrol 246(1-4):209-222. https://doi.org/10.1016/S0022-1694(01)00372-9

Zhang D, Lindholm G, Ratnaweera H (2018) Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. J Hydrol 556:409-418. https://doi.org/10.1016/j.jhydrol.2017.11.018

Zhang J, Wang X, Zhao C, Bai W, Shen J, Li Y, ..., Duan Y (2020) Application of cost-sensitive LSTM in water level prediction for nuclear reactor pressurizer. Nucl Eng Technol 52(7):1429-1435. https://doi.org/10.1016/j.net.2019.12.025