대청댐 수위 예측을 위한 Neural Networks의 적용

Application of Neural Networks to Predict Daechong Dam Water Levels

류용민* · 이의훈**

Ryu, Yong Min*, and Lee, Eui Hoon**

Abstract

The accurate prediction of water levels is essential for the efficient operation of hydraulic structures such as dams. However, such predictions are impossible when considering all factors affecting water level changes. Recently, with the development of artificial neural networks, there has been research on water level prediction using deep learning models. In this study, the deep learning technique of Multi-Layer Perceptron (MLP) was combined with Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) techniques, which afford advantages in time series data processing to predict the water level of Daechong Dam. Three data preprocessing processes were applied to improve accuracy in water level predictions by taking travel time into account, adding the outflow data of the dam, and normalizing the data. Eight cases were created with the three data preprocessing processes. When normalization was performed for the data that took travel time into account and the data were then used as input data for LSTM, the best prediction result was found for the peak water level of the dam. If various other data preprocessing techniques are applied in addition to those applied in this study, RNN and LSTM will show improved prediction results.

Key words : Deep Learning, MLP, RNN, LSTM, Data-preprocessing

요 지

정확한 수위의 예측은 댐과 같은 수공구조물의 효율적인 운영을 위해 필수적이다. 그러나 수위변화에 영향을 미치는 모든 인자들을 고려하여 예측을 하는 것은 불가능하다. 최근에는 인공지능망의 발달로 인해 딥러닝 모형을 통한 수위예측에 대한 연구가 진행되었다. 본 연구에서는 대청댐의 수위예측을 위해 딥러닝 기법 중 Multi Layer Perceptron (MLP)와 시계열 데이터 처리에 장점이 있는 Recurrent Neural Network (RNN) 및 Long Short-Term Memory (LSTM)을 사용하였다. 수위예측을 하는 과정에서 정확도 향상을 위해 세 가지 데이터 전처리 과정을 적용하였다. 적용한 데이터 전처리 과정은 지체시간의 고려, 댐의 방류량 추가, 자료의 정규화이다. 세 가지 데이터 전처리 과정을 통해 총 8개의 case를 생성하였다. 지체시간을 고려한 자료로 정규화를 실시하여 LSTM의 입력자료로 사용한 경우 댐의 첨두수위 예측에서 가장 좋은 예측결과를 나타냈다. 본 연구에서 적용한 데이터 전처리 기법뿐만이 아닌 다양한 데이터 전처리 기법을 적용하면 RNN 및 LSTM은 기존보다 좋은 예측결과를 나타낼 것이다.

핵심용어 : 딥러닝, MLP, RNN, LSTM, 데이터 전처리

1. 서 론

유량의 정확한 예측은 홍수 방지, 저수지 운영 및 가뭄 완화 등 수자원관리에 중요한 요인이다(Zhang et al., 2012;

Chen et al., 2016; Blöschl et al., 2019; Palmer and Ruhi, 2019). 국내에는 수자원 관리를 위한 다목적댐을 건설하여 운영되고 있다. 다목적댐의 경우 21개, 용수댐의 경우 14개 농업용 저수지의 경우 약 17,000개 이상이 운영되고 있다

*정회원, 충북대학교 토목공학과 석사과정(E-mail: rmfl45@naver.com)

Member, Master Student, Department of Civil Engineering, Chungbuk National University

**교신저자, 정회원, 충북대학교 토목공학부 조교수(Tel: +82-43-261-2407, Fax: +82-43-275-2377, E-mail: hydrohydro@cbnu.ac.kr)

Corresponding Author, Member, Assistant Professor, School of Civil Engineering, Chungbuk National University

(Han et al., 2021). 댐의 수문 개방 및 폐쇄 등 댐의 운영을 효율적으로 하기 위해서는 높은 신뢰도를 기반으로 댐으로 유입하는 유입량에 대한 예측이 중요하다(Han et al., 2021).

유량예측을 위해 일반적으로 사용되는 모형은 물리적 모형(Physically based models) 및 경험적 모형(Empirical models)과 같은 다양한 수문모형이다. 과거에는 물리적 모형을 통한 수위 예측이 진행되었다(Bicknell et al., 1996; Kim et al., 2007; Neitsch et al., 2011; Kang et al., 2013; Devia et al., 2015; Noh et al., 2016). 그러나 물리적 모형의 경우 수문시스템의 복잡한 특성을 단순화하는 과정에서 기상학적, 지질학적 및 지형학적인 다양한 변수를 고려하기 어렵다는 단점이 있다(Jung et al., 2018).

경험적 모형의 경우 데이터 기반 모형으로 통계적 관점에서의 입력-출력 관계를 추정하는 모형이다. 경험적 모형은 대표적으로 Autoregressive Moving Average (ARMA) 및 Autoregressive Integrated Moving Average (ARIMA) 등의 모형이 있으나, 비선형적인 요소들로 인해 유연성이 떨어졌다(Montanari et al., 1997; Kim, 2010).

과거의 경우 수위 변화에 영향을 미치는 인자들을 모두 고려한 후 물리적 모형을 구축해 모든 영향인자들과 수위 사이의 상관관계 및 인과관계를 분석하는 것은 불가능하였다. 그러나 많은 양의 정보 수집의 가능성과 함께 빅데이터의 처리가 가능한 머신러닝(Machine Learning)의 발전으로 인해 입력 및 출력자료 기반 학습 및 예측에 대한 연구가 진행되었다(Riad et al., 2004; Abrahart et al., 2004; Ghumman et al., 2011; Granata et al., 2016; Mosavi et al., 2018; Yan et al., 2018; Bae et al., 2019). 머신러닝을 통한 수위예측의 연구가 진행됨에 따라 머신러닝 기법 중 인간의 두뇌 신경망 구조를 모방해 개발된 딥러닝(Deep Learning)을 통한 수위예측의 연구가 진행되었다(Chen et al., 2013; Shoaib et al., 2016; Assem et al., 2017; Park et al., 2018; Choi et al., 2020; Kim et al., 2020; Lee et al., 2020).

본 연구에서는 구글(Google)에서 지원하는 딥러닝 오픈소스 라이브러리인 텐서플로우(Tensorflow)를 사용하여 ANN, RNN 및 LSTM을 구축해 대청댐 수위를 예측하고 각 모델의 성능을 비교하였다. 고수위를 나타내는 홍수기뿐만 아니라 저수위를 나타내는 갈수기의 변동성을 예측하기 위해 딥러닝 모델의 학습 및 예측을 위한 자료는 대청댐 상류의 수위관측소 중 옥천(산계교), 옥천(이원대교) 및 영동(심천교)의 2010년부터 2020년까지 11년의 시(hr)단위 수위 자료를 사용하였다. 학습을 위한 자료는 2010년부터 2020년 수위자료 중 2010년부터 2019년까지의 시단위 수위자료를 사용하였으며, 예측을 위한 자료는 2020년의 시단위 수위자료를 사용하였다.

수위예측을 위해 사용되는 시계열 데이터에서 지체시간, 방류량 및 정규화와 같은 데이터 전처리 과정에 대한 예측성능을 검토하였다. 예측정확도 개선을 위해 다양한 데이터

전처리를 활용한 예측결과를 비교하였다. 비교결과를 바탕으로 대청댐 수위 예측에 좋은 성능을 나타내는 데이터 전처리 과정을 제시하였다.

2. 연구방법

2.1 Artificial Neural Network (ANN)

인공신경망 구조 중 ANN은 가장 기본적인 구조이다. ANN은 입력층(Input Layer), 은닉층(Hidden Layer) 및 출력층(Output Layer)로 구성되어 있다. 각 층은 하나 이상의 뉴런들로 구성되어 있으며, 각 층마다 입력과 출력을 하기 위해 각 층의 연결강도(가중치 및 편향)가 매개변수로 설정된다. 이전 층에서 출력된 값들의 연결강도가 매개변수로 입력되며, 비선형의 함수를 통해 다음층으로 출력이 된다. ANN은 연결강도의 입력과 비선형함수를 통한 출력을 반복하여 입력층부터 출력층까지 진행하게 된다(Mahsa and Lee, 2018). ANN은 은닉층의 개수를 추가하여 Deep Neural Network (DNN)으로 구축할 수 있으나, 본 연구에서는 Multi Layer Perceptron (MLP) 중 은닉층을 1개로 설정한 모델을 사용하였다. Fig. 1은 MLP의 구조를 단순하게 나타낸 그림이다.

Fig. 1에 따르면, MLP는 입력된 자료가 입력층에서 은닉층을 거쳐 출력층으로 이동하는 것을 확인할 수 있다. Eq. (1) 및 Eq. (2)는 MLP의 연산과정을 간단하게 나타낸 식이다.

$$\hat{y}_k = f_y \left(\sum_{p=1}^j z_p w_{kp} + b_k \right) \quad (1)$$

$$z_p = f_z \left(\sum_{i=1}^q x_i v_{pi} + c_p \right) \quad (2)$$

여기서, w_{ki} 및 v_{pi} 는 연결강도를 나타내는 매개변수이며, b_k 및 c_p 는 편차를 보정하는 매개변수, f_y 및 f_z 는 활성화 함수(Activation function)이다. Eq. (1)과 Eq. (2)를 통해 산출된 결과값을 실제 값과 비교한 후 그 차이를 최소화시키는 것이 MLP에서 사용되는 목적함수이다.

2.2 Recurrent Neural Network (RNN)

순환신경망인 RNN은 MLP의 구조와 달리 과거의 데이터

Fig. 1. Structure of MLP

Fig. 2. Structure of RNN

가 임의의 시점에 영향을 주는 루프형 구조이다. Fig. 2는 RNN의 구조를 간단히 나타낸 그림이다.

Fig. 2에 따르면, RNN은 직전 데이터의 정보가 임의의 시점에 영향을 미치기 때문에 시계열 데이터를 분석하는데 좋은 성능을 나타낸다.

루프형의 구조로 인해 임의의 시점 t 에서 하나의 뉴런 객체의 출력 Y_t 는 시점 t 의 이전 출력 Y_{t-1} 을 포함한 모든 입력값 X_t 에 대한 함수이다. Eq. (3)은 임의의 시점 t 에 대한 출력 Y_t 를 식으로 나타낸 것이다.

$$Y_t = f_{act}(W_x \times X_t + W_y \times Y_{t-1} + b) \quad (3)$$

여기서, f_{act} 는 계산을 위한 활성화 함수를 나타내며, W_x 및 W_y 는 가중치, b 는 편향(Bias)을 나타낸다. Eq. (3)과 같이 가중치에 대한 업데이트를 위해 사용되는 역전파(Back-propagation) 알고리즘은 층마다 포함되어 있는 다수의 뉴런에 개별적으로 오차경사를 전파시키며 진행한다. 그러나 RNN은 오랜 기간의 자료를 입력자료로 사용함과 동시에 은닉층의 개수를 늘릴 경우 학습이 진행되는 과정에서 오차경사의 기울기가 소실되는 가중치 소실(Vanishing Gradient) 문제가 발생하며, 가까운 층에서 생성된 데이터의 정보를 기억하지 못해 학습능력이 저하된다는 단점이 있다.

2.3 Long Short-Term Memory (LSTM)

LSTM은 시계열 자료에 대해 특화된 알고리즘인 RNN의 단점을 개선한 모델이다. RNN의 단점인 가중치 소실 및 가중치 폭주(Exploding Gradient)를 개선하기 위해 제안된 모델이다(Hochreiter and Schmidhuber, 1997). Fig. 3은 LSTM의 구조를 간단하게 나타낸 그림이다.

Fig. 3에 따르면, LSTM은 RNN의 구조를 기반으로 임의의 시점에서 가중치와 편향을 업데이트하기 위해 셀(Cell)의 개념을 도입하였다. LSTM에 적용된 셀은 입력과 현재까지의 상태를 통해 모델 내부에 저장되어 있는 가중치와 편향에 대한 업데이트 여부에 대한 판단을 한다.

LSTM에 적용된 셀은 데이터의 흐름을 조절하기 위해 세 가지 게이트로 구성되어 있다. 세 가지 게이트는 입력게이트(Input Gate, i_t), 망각게이트(Forget Gate, f_t) 및 출력게이트

Fig. 3. Structure of LSTM

트(Output Gate, o_t)로 구성되어 있다. LSTM은 셀 내부에서 출력을 위해 망각게이트, 입력게이트 및 출력게이트의 순서로 진행된다. 첫 번째 과정인 망각게이트는 Fig. 3에서 주황색으로 표시된 부분이다. 망각게이트는 직전 단계에서 출력된 h_{t-1} 과 새로운 값 x_t 를 통해 0과 1사이의 값을 출력한다. 출력된 값을 바탕으로 입력된 값에 대한 정보를 유지 및 제거에 대한 판단을 하게 된다(Olah, 2015). Eq. (4)는 망각게이트인 f_t 에 사용되는 수식이다.

$$f_t = \sigma(W \times [h_{t-1}, x_t] + b) \quad (4)$$

여기서, σ 는 활성화 함수를 나타내며, W 는 가중치, b 는 편향을 나타낸다. 두 번째 단계인 입력게이트는 Fig. 3에서 녹색으로 표시된 부분이다. 입력게이트는 망각게이트 이후 과정으로 입력하고자 하는 값의 결과와 함께 새로운 셀 상태를 갱신하는 과정이다. 입력게이트인 i_t 에서는 Eq. (5)를 통해 어떤 정보를 셀에 저장할지 결정을 하게 된다.

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b) \quad (5)$$

여기서, σ 는 활성화 함수이며, h_{t-1} 은 이전 셀의 출력, x_t 는 현재의 입력, W_i 는 셀의 가중치 b 는 편향을 나타낸다. Eq. (5)를 통해 저장을 할 정보를 선택한 후, Eq. (6)을 통해 셀의 새로운 업데이트 시 사용될 후보 셀을 생성하게 된다.

$$C_n = \tanh(W_{cell} \times [h_{t-1}, x_t] + b) \quad (6)$$

여기서, C_n 은 후보 셀이며, W_{cell} 은 셀 상태의 가중치, h_{t-1}

은 이전 셀의 출력, x_t 는 현재의 입력, b 는 편향을 나타낸다. LSTM의 입력게이트는 Eq. (6)을 통해 생성된 후보 셀과 과거의 셀을 조합하여 현재의 셀 상태를 업데이트 한다. Eq. (7)은 생성된 후보 셀과 과거의 셀을 통해 현재의 셀 상태를 조합하기 위해 사용되는 식이다.

$$C_t = f_t \times C_{t-1} + i_t \times C_n \quad (7)$$

여기서, C_t 는 현재의 셀 상태를 나타내며, C_n 은 활성화 함수를 통해 생성된 새로운 셀 상태, W_{cell} 은 셀 상태의 가중치, b 는 편향을 나타낸다. 세 번째 단계인 출력게이트는 Fig. 3에서 파란색으로 표시된 부분이다. 출력게이트는 망각 게이트, 입력게이트 이후 진행되는 과정으로 셀 내에서 출력되는 값을 결정하는 과정이다. Eq. (8)은 셀 상태 중 출력하고자 하는 부분을 결정하는 식이다.

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b) \quad (8)$$

여기서, σ 는 활성화 함수이며, W_o 은 셀 상태의 가중치, h_t 는 출력되는 새로운 값, x_t 는 현재의 입력, b 는 편향을 나타낸다. 출력게이트인 o_t 는 Eq. (9)를 통해 $\tanh$ 와 활성화된 셀의 상태를 곱하여 특정 시점의 상태를 업데이트 한다.

$$h_t = o_t \times \tanh(C_t) \quad (9)$$

여기서, o_t 는 출력게이트이며, C_t 는 활성화된 셀의 상태이다.

2.4 대상유역 및 자료구축

본 연구의 대상유역은 금강에 위치한 대청댐으로 선정하였다. 대청댐의 유역면적은 금강유역의 총 42.25%인 4,189 km²이며, 만수면적은 72.8 km², 유효저수량은 790백만 m³이다. 대청댐 인근 수위관측소는 옥천(산계교), 옥천(이원대교) 및 영동(산계교) 등이 있다. 본 연구에서는 수위자료 수집을 위해 대청댐 상류의 수위관측소 중 결측치가 많이 존재하는 수위관측소를 제외하여 3개소를 선정하였다. 선정된 수위관측소는 옥천(산계교), 옥천(이원대교) 및 영동(산계교)이며, Fig. 4는 각 수위관측소의 위치를 나타낸 그림이다.

본 연구에서 사용된 대청댐의 수위자료 및 각 수위관측소의 수위자료는 국가수자원관리종합정보시스템(Water Resources Management Information System, WAMIS)에서 제공하는 자료를 사용하였다. WAMIS에서 제공하는 2010년부터 2020년까지의 대청댐에서 관측된 시간위 수위자료와 옥천(산계교), 옥천(이원대교) 및 영동(산계교)에서 관측된 시간위 수위자료를 사용하였다. 사용한 시간위 수위자료 중 2010년부터 2019년까지의 수위자료는 학습을 위해 사용하였으며, 2020년의 수위자료는 수위예측을 위해 사용되었다.

각 수위관측소와 댐에서 측정된 수위자료는 원자료이기

Fig. 4. Location of Gauging Stations

때문에 학습 및 예측에 좋은 성능을 나타내기 어렵다. 원자료를 기반으로 데이터 전처리를 하기 위해 총 세 가지 방법을 적용하였다. 첫 번째 방법으로는 각 수위관측소의 수위자료와 대청댐 수위자료의 첨두시간을 고려하여 수위관측소의 수위자료를 지체시켰다. 수위자료를 지체시키기 위해 2020년 각 수위관측소와 대청댐의 수위자료 중 첨두수위를 바탕으로 수위자료를 보정하였다. 각 수위관측소와 대청댐 사이의 지체시간은 영동(심천교)는 13시간, 옥천(산계교)는 15시간 및 옥천(이원대교)는 14시간이다. 각 수위관측소의 수위자료를 지체시켜 수위자료를 보정하였다. 두 번째 방법으로는 대청댐의 방류량(Outflow)을 통해 자료를 보정하였다. 방류량은 대청댐의 수위에 직접적인 영향을 주는 인자이다. 대청댐 방류량을 학습자료에 적용하여 수위 예측을 위한 학습자료를 보정하였다. 세 번째 방법으로는 정규화(Normalization)이다. 자료에 대한 정규화는 자료가 넓은 범주를 가지고 있을 경우 반드시 진행하는 과정이다(Mok et al., 2020). 정규화를 통해 모든 자료는 0과 1사이의 값으로 변환된다. Eq. (10)은 정규화를 위한 식이다.

$$N_i = \frac{R_i - R_{\min}}{R_{\max} - R_{\min}} \quad (10)$$

여기서, N_i 는 정규화된 변수값이며, R_i 는 실제 변수값, $R_{\min}$ 은 변수의 최솟값, $R_{\max}$ 은 변수의 최댓값이다. 세 가지 데이터 전처리 방법을 통해 생성된 Case는 총 8개이다. Table 1은 본 연구에 사용된 Case를 비교한 표이다.

MLP, RNN 및 LSTM의 예측결과를 비교분석하기 위해 Root Mean Square Error (RMSE) 및 Mean Absolute Error (MAE)의 평가지표를 사용하였다. Eqs. (11)과 (12)는 RMSE와 MAE의 식을 나타낸 것이다.

Table 1. Comparison by Cases

Case	Travel time	Outflow	Normalization
1	-	-	-
2	-	-	○
3	-	○	-
4	-	○	○
5	○	-	-
6	○	-	○
7	○	○	-
8	○	○	○

$$RMSE = \sqrt{\frac{1}{A} \sum_{i=1}^A (O_i - P_i)^2} \quad (11)$$

$$MAE = \frac{1}{A} \sum_{i=1}^A |O_i - P_i| \quad (12)$$

여기서, O_i 는 각 관측소에서 관측된 수위이며, P_i 는 각 기법을 통해 예측한 수위이다. 각 기법의 예측결과를 RMSE 및 MAE를 통해 비교하였으며, 데이터 전처리를 하지 않은 원자료 기반 예측결과와 데이터 전처리를 한 자료 기반 예측결과를 비교하였다.

3. 연구결과

MLP, RNN 및 LSTM의 결과를 비교하기 위해 각 기법의 은닉층은 1개로 설정하였으며, 은닉층의 노드(Node)는 100개, 학습을 위한 Epoch는 10, Optimizer는 Adaptive Moment Estimation (Adam)으로 설정하였다. MLP의 활성화 함수는 Swish함수로 설정하였다(Ramachandran et al., 2017). RNN 및 LSTM의 활성화함수는 Hyperbolic Tangent (tanh)로 사용하였다.

3.1 Case 1

데이터 전처리 과정을 적용하지 않은 자료인 원자료를 통해 학습 및 예측을 진행한 결과 대청댐 수위가 높았던 기간을 제외한 기간동안은 예측성능이 떨어지는 것을 확인할 수 있었다. Table 2는 전체 기간에 대한 RMSE 및 MAE를 나타낸 표이다.

Table 2에 따르면, RMSE가 가장 낮은 딥러닝 기법은 RNN이었으며, LSTM, MLP의 순으로 RMSE가 낮았다. MAE가 가장 낮은 딥러닝 기법은 RNN이었으며, LSTM, MLP의 순으로 MAE가 낮았다. RMSE 및 MAE를 통해 전체 기간에 대한 예측성능은 RNN이 가장 좋은 성능을 나타냈으며 MLP가 가장 좋지 않은 성능을 나타냈다는 것을 알 수 있었다.

본 연구에 사용된 MLP는 은닉층이 1개이기 때문에 학습

Table 2. RMSE by Prediction Method in Case 1

Prediction Method	RMSE	MAE
MLP	2.87	2.51
RNN	2.38	1.87
LSTM	2.62	2.25

Fig. 5. Comparison of the Results in Case 1

및 예측을 하는 과정에서 RNN 및 LSTM보다 낮은 성능을 나타냈다. RNN 및 LSTM의 경우 시계열 데이터 처리에 장점이 있기 때문에 시계열 데이터인 댐 수위예측에 대해 MLP보다 좋은 결과를 나타냈다. RNN은 장기간의 데이터에 대해 가중치 소실 및 가중치 폭발로 인한 기울기 소실의 단점이 있으나, 현재 적용된 학습자료의 길이가 짧기 때문에 MLP 및 LSTM보다 좋은 성능을 나타냈다. Fig. 5는 전체 예측기간에 대한 각 기법의 예측수위 및 관측수위를 나타낸 그래프이다.

Fig. 5에 따르면 대청댐의 수위가 가장 높았던 날인 2020년 7월 31일 03:00의 수위예측결과는 MLP가 가장 좋은 예측결과를 나타냈으며, RNN, LSTM의 순으로 좋은 예측결과를 나타냈다. MLP는 실제 침투수위와 예측결과 사이의 오차가 약 5.41 m였으며, RNN의 오차는 약 5.97 m, LSTM의 오차는 약 6.22 m로 나타났다.

데이터 전처리 과정을 적용하지 않은 원자료를 통한 학습 및 예측결과를 통해 전체 기간에 대해서는 순환신경망 계열인 RNN 및 LSTM은 MLP보다 낮은 RMSE를 나타냈다. 그러나 침투수위에서는 MLP가 가장 좋은 결과를 나타냈다.

3.2 Case 2

정규화를 실시한 자료를 통해 학습 및 예측을 진행한 결과 대청댐 수위가 높았던 기간을 제외한 기간동안은 예측

성능이 떨어지는 것을 확인할 수 있었다. Table 3은 전체 기간에 대한 RMSE 및 MAE를 나타낸 표이다.

Table 3에 따르면, RMSE가 가장 낮은 딥러닝 기법은 LSTM이었으며, RNN, MLP의 순으로 RMSE가 낮았다. MAE가 가장 낮은 딥러닝 기법은 LSTM이었으며, RNN, MLP의 순으로 MAE가 낮았다. RMSE 및 MAE를 통해 전체 기간에 대한 예측성능은 LSTM이 가장 좋은 성능을 나타냈으며 MLP가 가장 좋지 않은 성능을 나타냈다는 것을 알 수 있다. 데이터 전처리 과정을 적용하지 않은 Case 1과 비교하였을 경우, MLP의 RMSE와 MAE는 각각 0.41, 0.4가 증가하였으며, RNN의 RMSE와 MAE는 각각 0.86, 0.89가 증가하였다. LSTM의 RMSE는 0.1이 증가하였으나, MAE는 0.13이 감소하였다. Table 3을 통해 전체기간에 대한 예측성능이 Case 1보다 낮아졌다는 것을 알 수 있었다. Fig. 6은 전체 예측기간에 대한 각 기법의 예측수위 및 관측수위를 나타낸 그래프이다.

Fig. 6에 따르면 대청댐의 수위가 가장 높았던 날인 2020년 7월 31일 03:00의 수위예측결과는 LSTM이 가장 좋은 예측결과를 나타냈으며, RNN, MLP의 순서로 좋은 예측결과를 나타냈다. LSTM은 실제 침두수위와 예측결과 사이의 오차가 약 5.43 m였으며, RNN의 오차는 약 5.81 m, MLP의 오차는 약 7.46 m로 나타났다.

정규화를 실시한 자료를 통한 학습 및 예측결과를 통해 전체 기간에 대해서는 순환신경망 계열인 RNN 및 LSTM이

MLP보다 낮은 RMSE를 나타냈다. 그러나 침두수위에서는 MLP가 가장 좋은 결과를 나타냈다.

3.3 Case 3

방류량을 추가한 자료를 통해 학습 및 예측을 진행한 결과 대청댐 수위가 높았던 기간을 제외한 기간동안은 예측성능이 떨어지는 것을 확인할 수 있었다. Table 4는 전체 기간에 대한 RMSE 및 MAE를 나타낸 표이다.

Table 4에 따르면, RMSE가 가장 낮은 딥러닝 기법은 LSTM이었으며, RNN, MLP의 순으로 RMSE가 낮았다. MAE가 가장 낮은 딥러닝 기법은 LSTM이었으며, RNN, MLP의 순으로 MAE가 낮았다. RMSE 및 MAE를 통해 전체 기간에 대한 예측성능은 LSTM이 가장 좋은 성능을 나타냈으며 MLP가 가장 좋지않은 성능을 나타냈다는 것을 알 수 있다. Case 1과 비교하였을 경우, MLP의 RMSE와 MAE는 각각 7.36, 3.62가 증가하였으며, RNN의 RMSE와 MAE는 각각 0.51, 0.62가 증가하였다. LSTM의 RMSE와 MAE는 각각 0.09, 0.06이 증가하였다. Table 4를 통해 전체기간에 대한 예측성능이 Case 1보다 낮아졌다는 것을 알 수 있었다. Fig. 7은 전체 예측기간에 대한 각 기법의 예측수위 및 관측수위를 나타낸 그래프이다.

Fig. 7에 따르면 대청댐의 수위가 가장 높았던 날인 2020년 7월 31일 03:00의 수위예측결과는 RNN이 가장 좋은 예측결

Table 3. RMSE by Prediction Method in Case 2

Prediction Method	RMSE	MAE
MLP	3.28	2.91
RNN	3.24	2.76
LSTM	2.72	2.12

Table 4. RMSE by Prediction Method in Case 3

Prediction Method	RMSE	MAE
MLP	10.23	6.13
RNN	2.89	2.49
LSTM	2.71	2.31

Fig. 6. Comparison of the Results in Case 2

Fig. 7. Comparison of the Results in Case 3

과를 나타냈으며, LSTM, MLP의 순서로 좋은 예측결과를 나타냈다. RNN은 실제 침두수위와 예측결과 사이의 오차가 약 4.00 m였으며, LSTM의 오차는 약 6.74 m, MLP의 오차는 약 28.63 m로 나타났다. 방류량을 추가한 자료를 통한 학습 및 예측결과를 통해 전체 기간에 대해서는 순환신경망 계열인 RNN 및 LSTM이 MLP보다 낮은 RMSE를 나타냈다. MLP의 경우 방류량을 추가한 자료를 통해 학습 및 예측을 할 경우 과도한 예측을 한다는 것을 알 수 있었다.

3.4 Case 4

방류량을 추가한 자료를 이용해 정규화를 실시한 자료 통해 학습 및 예측을 진행한 결과 대청댐 수위가 높았던 기간을 제외한 기간동안은 예측성능이 떨어지는 것을 확인할 수 있었다. Table 5는 전체 기간에 대한 RMSE 및 MAE를 나타낸 표이다.

Table 5에 따르면, RMSE가 가장 낮은 딥러닝 기법은 LSTM이었으며, RNN, MLP의 순으로 RMSE가 낮았다. MAE가 가장 낮은 딥러닝 기법은 LSTM이었으며, RNN, MLP의 순으로 MAE가 낮았다.

RMSE 및 MAE를 통해 전체 기간에 대한 예측성능은 LSTM이 가장 좋은 성능을 나타냈으며 MLP가 가장 좋지 않은 성능을 나타냈다는 것을 알 수 있었다. Case 1과 비교하였을 경우, MLP의 RMSE와 MAE는 각각 0.64, 0.33이 증가하였으

며, RNN의 RMSE와 MAE는 각각 0.69, 0.74가 증가하였다. LSTM의 RMSE와 MAE는 각각 0.36, 0.26이 증가하였다. Table 5를 통해 전체기간에 대한 예측성능이 Case 1보다 낮아졌다는 것을 알 수 있었다. Fig. 8은 전체 예측기간에 대한 각 기법의 예측수위를 나타낸 그래프이다.

Fig 8에 따르면 대청댐의 수위가 가장 높았던 날인 2020년 7월 31일 03:00의 수위예측결과는 RNN이 가장 좋은 예측결과를 나타냈으며, MLP, LSTM의 순서로 좋은 예측결과를 나타냈다. RNN은 실제 침두수위와 예측결과 사이의 오차가 약 2.16 m였으며, MLP의 오차는 약 3.08 m, LSTM의 오차는 약 4.43 m로 나타났다. 방류량을 추가한 자료를 통한 학습 및 예측결과를 통해 전체 기간에 대해서는 순환신경망 계열인 RNN은 MLP보다 낮은 RMSE를 나타냈다.

MLP의 경우 방류량을 추가한 자료를 이용해 정규화를 실시한 자료를 통해 학습 및 예측을 할 경우 과도한 예측을 한다는 것을 알 수 있다.

3.5 Case 5

원자료를 기반으로 지체시간을 고려한 자료를 통해 학습 및 예측을 진행한 결과 대청댐 수위가 높았던 기간을 제외한 기간동안은 예측성능이 떨어지는 것을 확인할 수 있었다. Table 6은 전체 기간에 대한 RMSE 및 MAE를 나타낸 표이다.

Table 6에 따르면, RMSE가 가장 낮은 딥러닝 기법은 MLP

Table 5. RMSE by Prediction Method in Case 4

Prediction Method	RMSE	MAE
MLP	3.51	2.84
RNN	3.07	2.61
LSTM	2.98	2.51

Fig. 8. Comparison of the Results in Case 4

Table 6. RMSE by Prediction Method in Case 5

Prediction Method	RMSE	MAE
MLP	2.43	2.05
RNN	2.66	2.14
LSTM	2.69	2.32

Fig. 9. Comparison of the Results in Case 5

이였으며, RNN, LSTM의 순으로 RMSE가 낮았다. MAE가 가장 낮은 딥러닝 기법은 MLP이였으며, RNN, LSTM의 순으로 MAE가 낮았다. RMSE 및 MAE를 통해 전체 기간에 대한 예측성능은 MLP가 가장 좋은 성능을 나타냈으며 LSTM이 가장 좋지 않은 성능을 나타냈다는 것을 알 수 있다. Case 1과 비교하였을 경우, MLP의 RMSE와 MAE는 각각 0.44, 0.46이 감소하였으며, RNN의 RMSE와 MAE는 각각 0.28, 0.27이 증가하였다. LSTM의 RMSE와 MAE는 각각 0.07, 0.07이 증가하였다. Table 6을 통해 지체시간을 고려한 자료를 사용해 학습 및 예측을 진행할 경우 MLP의 예측성능이 개선된다는 것을 확인할 수 있었다. Fig. 9는 전체 예측기간에 대한 각 기법의 예측수위를 나타낸 그래프이다.

Fig. 9에 따르면 대청댐의 수위가 가장 높았던 날인 2020년 7월 31일 03:00의 수위예측결과는 MLP가 가장 좋은 예측결과를 나타냈으며, RNN, LSTM의 순으로 좋은 예측결과를 나타냈다. MLP는 실제 침투수위와 예측결과 사이의 오차가 약 0.45 m였으며, RNN의 오차는 약 1.91 m, LSTM의 오차는 약 5.08 m로 나타났다. 방류량을 추가한 자료를 통한 학습 및 예측결과를 통해 전체 기간에 대해서는 순환신경망 계열인 RNN은 MLP보다 낮은 RMSE를 나타냈다. MLP 및 RNN의 경우 지체시간을 고려한 자료를 통해 학습 및 예측을 할 경우 과도한 예측을 한다는 것을 알 수 있다.

3.6 Case 6

원자료를 기반으로 지체시간을 고려한 자료를 이용해 정규화를 실시한 자료를 통해 학습 및 예측을 진행한 결과 대청댐 수위가 높았던 기간을 제외한 기간동안은 예측성능이 떨어지는 것을 확인할 수 있었다. Table 7은 전체 기간에 대한 RMSE 및 MAE를 나타낸 표이다.

Table 7에 따르면, RMSE가 가장 낮은 딥러닝 기법은 RNN이였으며, MLP, LSTM의 순으로 RMSE가 낮았다. MAE가 가장 낮은 딥러닝 기법은 LSTM이였으며, RNN, MLP의 순으로 MAE가 낮았다. LSTM이 가장 낮은 MAE를 나타냈으나, RNN과 약 0.02의 작은 차이를 나타냈다. RMSE 및 MAE를 통해 전체 기간에 대한 예측성능은 RNN이 가장 좋은 성능을 나타냈으며 LSTM이 가장 좋지 않은 성능을 나타냈다는 것을 알 수 있었다.

Case 1과 비교하였을 경우, MLP의 RMSE와 MAE는 각각 0.13, 0.15가 감소하였으며, RNN의 RMSE와 MAE는 각각 0.27, 0.38이 증가하였다. LSTM의 RMSE는 0.21이 증가하였으며, MAE는 0.02가 감소하였다. Table 7을 통해 지체시간을 고려한 자료를 이용해 정규화를 실시한 자료를 통해 학습 및 예측을 진행할 경우 MLP의 예측성능이 개선된다는 것을 확인할 수 있었다. Fig. 10은 전체 예측기간에 대한 각 기법의 예측수위를 나타낸 그래프이다.

Fig. 10에 따르면 대청댐의 수위가 가장 높았던 날인 2020년 7월 31일 03:00의 수위예측결과는 LSTM이 가장 좋은

Table 7. RMSE by Prediction Method in Case 6

Prediction Method	RMSE	MAE
MLP	2.74	2.36
RNN	2.65	2.25
LSTM	2.83	2.23

Fig. 10. Comparison of the Results in Case 6

예측결과를 나타냈으며, RNN, MLP의 순으로 좋은 예측결과를 나타냈다. LSTM은 실제 침투수위와 예측결과 사이의 오차가 약 0.12 m였으며, RNN의 오차는 약 2.57 m, MLP의 오차는 약 9.98 m로 나타났다. 방류량을 추가한 자료를 통한 학습 및 예측결과를 통해 전체 기간에 대해서는 순환신경망 계열인 RNN은 MLP보다 낮은 RMSE를 나타냈다. MLP 및 LSTM의 경우 지체시간을 고려한 자료를 이용해 정규화를 실시한 자료를 통해 학습 및 예측을 할 경우 과도한 예측을 한다는 것을 알 수 있었다.

3.7 Case 7

방류량을 추가하여 지체시간을 고려한 자료를 이용해 정규화를 실시한 자료를 통해 학습 및 예측을 진행한 결과 대청댐 수위가 높았던 기간을 제외한 기간동안은 예측성능이 떨어지는 것을 확인할 수 있었다. Table 8은 전체 기간에 대한 RMSE 및 MAE를 나타낸 표이다.

Table 8에 따르면, RMSE가 가장 낮은 딥러닝 기법은 LSTM이였으며, RNN, MLP의 순으로 RMSE가 낮았다. MAE가 가장 낮은 딥러닝 기법은 LSTM이였으며, RNN, MLP의 순으로 MAE가 낮았다. RMSE 및 MAE를 통해 전체 기간에 대한 예측성능은 LSTM이 가장 좋은 성능을 나타냈으며 MLP가 가장 좋지 않은 성능을 나타냈다는 것을 알 수 있었다.

Table 8. RMSE by Prediction Method in Case 7

Prediction Method	RMSE	MAE
MLP	12.73	7.06
RNN	2.94	2.61
LSTM	2.20	1.83

Fig. 11. Comparison of the Results in Case 7

Case 1과 비교하였을 경우, MLP의 RMSE와 MAE는 각각 9.86, 4.55가 증가하였으며, RNN의 RMSE와 MAE는 각각 0.56, 0.74가 증가하였다. LSTM의 RMSE와 MAE는 각각 0.42, 0.42가 감소하였다. Table 8을 통해 방류량을 추가하여 지체시간을 고려한 자료를 이용해 학습 및 예측을 진행할 경우 LSTM의 예측성능이 개선된다는 것을 확인할 수 있었다. Fig. 11은 전체 예측기간에 대한 각 기법의 예측수위를 나타낸 그래프이다.

Fig. 11에 따르면 대청댐의 수위가 가장 높았던 날인 2020년 7월 31일 03:00의 수위예측결과는 RNN이 가장 좋은 예측결과를 나타냈으며, LSTM, MLP의 순서로 좋은 예측결과를 나타냈다. RNN은 실제 침투수위와 예측결과 사이의 오차가 약 2.60 m였으며, LSTM의 오차는 약 4.40 m, MLP의 오차는 약 75.42 m로 나타났다. 방류량을 추가하여 지체시간을 고려한 자료를 통한 학습 및 예측결과를 통해 전체 기간에 대해서는 순환신경망 계열인 RNN은 MLP보다 낮은 RMSE를 나타냈다. MLP의 경우 방류량을 추가하여 지체시간을 고려한 자료를 이용해 정규화를 실시한 자료를 통해 학습 및 예측을 할 경우 과도한 예측을 한다는 것을 알 수 있었다.

3.8 Case 8

방류량을 추가하여 지체시간을 고려한 자료를 이용해

Table 9. RMSE by Prediction Method in Case 8

Prediction Method	RMSE	MAE
MLP	3.20	2.69
RNN	3.08	2.67
LSTM	3.34	2.73

Fig. 12. Comparison of the Results in Case 8

정규화를 실시한 자료를 통해 학습 및 예측을 진행한 결과 대청댐 수위가 높았던 기간을 제외한 기간동안은 예측성능이 떨어지는 것을 확인할 수 있었다. Table 9는 전체 기간에 대한 RMSE 및 MAE를 나타낸 표이다.

Table 9에 따르면, RMSE가 가장 낮은 딥러닝 기법은 RNN이었으며, MLP, LSTM의 순으로 RMSE가 낮았다. MAE가 가장 낮은 딥러닝 기법은 RNN이었으며, MLP, LSTM의 순으로 MAE가 낮았다. RMSE 및 MAE를 통해 전체 기간에 대한 예측성능은 RNN이 가장 좋은 성능을 나타냈으며 LSTM이 가장 좋지 않은 성능을 나타냈다는 것을 알 수 있었다. Case 1과 비교하였을 경우, MLP의 RMSE와 MAE는 각각 0.33, 0.18이 증가하였으며, RNN의 RMSE와 MAE는 각각 0.70, 0.80이 증가하였다. LSTM의 RMSE와 MAE는 각각 0.72, 0.48이 증가하였다. Table 9를 통해 전체기간에 대한 예측성능이 Case 1보다 낮아졌다는 것을 알 수 있었다. Fig. 12는 전체 예측기간에 대한 각 기법의 예측수위를 나타낸 그래프이다.

Fig. 12에 따르면 대청댐의 수위가 가장 높았던 날인 2020년 7월 31일 03:00의 수위예측결과는 LSTM이 가장 좋은 예측결과를 나타냈으며, RNN, MLP의 순서로 좋은 예측결과를 나타냈다. LSTM은 실제 침투수위와 예측결과 사이의 오차가 약 4.21 m였으며, RNN의 오차는 약 4.87 m, MLP의

오차는 약 5.22 m로 나타났다. 방류량을 추가하여 지체시간을 고려한 자료를 이용해 정규화를 실시한 자료를 통한 학습 및 예측결과를 통해 전체 기간에 대해서는 순환신경망 계열인 RNN은 MLP보다 낮은 RMSE를 나타났다. MLP의 경우 방류량을 추가하여 지체시간을 고려한 자료를 이용해 정규화를 실시한 자료를 통해 학습 및 예측을 할 경우 과도한 예측을 한다는 것을 알 수 있었다.

Table 10은 Case별 가장 낮은 RMSE, MAE 및 오차(Difference)를 나타낸 기법과 RMSE, MAE 및 오차를 나타낸 표이다. 오차는 각 Case의 예측결과 중에서 첨두수위가 발생한 2020년 7월 31일 03:00에서 측정된 수위와 예측된 수위사이의 차이를 나타낸 값이다.

Table 10에 따르면, 모든 Case 중 LSTM이 4개의 Case에서 낮은 RMSE를 보여주었다. RNN의 경우 3개, MLP의 경우 1개 동안 낮은 RMSE를 보여주었다. 8개의 Case에서 가장 낮은 RMSE를 나타낸 기법은 LSTM이며, 가장 높은 RMSE를 나타낸 기법은 RNN이었다. Table 10을 통해 LSTM이 다양한 데이터 전처리에 대해 모든 기간의 RMSE가 좋은 결과를 나타낸다는 것을 알 수 있었다.

모든 Case 중 LSTM이 5개의 Case에서 낮은 MAE를 보여주었다. RNN의 경우 2개, MLP의 경우 1개 동안 낮은 MAE를 보여주었다. 8개의 Case에서 가장 낮은 MAE를 나타낸 기법은 LSTM이며, 가장 높은 MAE를 나타낸 기법은 RNN이다. Table 10을 통해 LSTM이 다양한 데이터 전처리에 대해 모든 기간의 MAE가 좋은 결과를 나타낸다는 것을 알 수 있었다.

모든 Case 중 RNN과 LSTM이 3개의 Case에서 낮은 오차

를 보여주었다. 8개의 Case 동안 가장 낮은 오차를 나타낸 기법은 LSTM이며, 가장 높은 오차를 나타낸 기법 또한 LSTM이었다. LSTM은 정규화를 실시한 자료로 학습 및 예측을 할 경우, RNN 및 MLP보다 낮은 오차를 나타내는 것을 확인할 수 있었다. RNN은 방류량을 추가한 자료로 학습 및 예측을 할 경우, MLP 및 LSTM보다 낮은 오차를 나타내는 것을 확인할 수 있었다. MLP는 원자료 및 지체시간을 고려한 원자료로 학습 및 예측을 진행할 경우 RNN 및 LSTM보다 낮은 오차를 나타내는 것을 알 수 있었다. Table 10을 통해 시계열 데이터를 처리하는데 장점이 있는 RNN 및 LSTM이 다양한 데이터 전처리에 대해 첨두수위의 오차가 좋은 결과를 나타낸다는 것을 알 수 있었다.

Table 10에 따르면 Case 7은 오차의 경우 8개의 Case 중 4번째로 좋은 결과를 나타내었다. 그러나 RMSE 및 MAE가 가장 낮은 것을 확인할 수 있다. Case 7은 데이터 전처리를 하지 않은 원자료인 Case 1과 비교하였을 경우 RMSE는 0.18, MAE는 0.07, 오차는 2.81 감소하였다. Case 1과 Case 7의 결과를 바탕으로 딥러닝을 통한 수위예측을 진행할 경우 지체시간과 방류량을 고려한 데이터 전처리 과정은 예측정확도를 개선하는데 효율적이다.

4. 결 론

일반적인 딥러닝 기법을 통한 수위예측의 경우 한 가지 방법의 데이터 전처리 과정을 적용한 수위자료를 바탕으로 수위를 예측하였다. 본 연구에서는 다양한 데이터 전처리 과정을 적용하여 수위예측에 가장 좋은 성능을 나타내는 데이터 전처리 기법에 대해 결과를 비교하였다. 또한 대칭값의 수위예측을 위해 딥러닝 기법인 MLP, RNN 및 LSTM을 사용해 원자료 및 다양한 데이터 전처리 기법을 통한 자료를 바탕으로 예측성능을 비교하였다.

RMSE를 바탕으로 결과를 비교할 경우, MLP는 지체시간을 고려한 자료를 기반으로 학습 및 예측을 할 경우 다른 Case보다 좋은 결과를 나타냈으나, 1개의 Case를 제외한 모든 Case에서 RNN 및 LSTM보다 좋지 않은 결과를 나타냈다. LSTM은 데이터 전처리를 실행한 자료를 기반으로 학습 및 예측을 할 경우, MLP 및 RNN보다 좋은 결과를 나타내는 Case가 많았다.

MAE를 바탕으로 결과를 비교할 경우, MLP는 지체시간을 고려한 자료를 기반으로 학습 및 예측을 할 경우 다른 Case보다 좋은 결과를 나타냈으나, 1개의 Case를 제외한 대부분 Case에서 RNN 및 LSTM보다 좋지 않은 결과를 나타냈다. LSTM은 데이터 전처리를 실행한 자료를 기반으로 학습 및 예측을 할 경우, MLP 및 RNN보다 좋은 결과를 나타내는 Case가 많았다.

첨두수위를 바탕으로 결과를 비교할 경우, 8개의 Case 중 Case 7에서 LSTM이 가장 낮은 오차를 나타냈으며, Case

Table 10. Comparison RMSE, MAE and Difference (between Observed Data and Predicted Data) by Each Cases

Case	RMSE (Method)	MAE (Method)	Difference (Method)
1	2.38 (RNN)	1.87 (RNN)	5.41 m (MLP)
2	2.72 (LSTM)	2.12 (LSTM)	5.43 m (LSTM)
3	2.71 (LSTM)	2.31 (LSTM)	4.00 m (RNN)
4	2.98 (LSTM)	2.51 (LSTM)	2.16 m (RNN)
5	2.43 (MLP)	2.05 (MLP)	0.45 m (MLP)
6	2.65 (RNN)	2.23 (LSTM)	0.12 m (LSTM)
7	2.20 (LSTM)	1.86 (LSTM)	2.60 m (RNN)
8	3.08 (RNN)	2.67 (RNN)	4.12 m (LSTM)

2에서 LSTM이 가장 높은 오차를 나타냈다. 그러나 8개의 Case 중 RNN 및 LSTM이 3개의 Case 동안 가장 낮은 오차를 나타냈다.

다양한 Case를 이용한 대청댐 수위예측결과를 통해 데이터 전처리 과정을 적용한 자료를 기반으로 학습 및 예측을 할 경우 RNN 및 LSTM은 MLP보다 댐 수위예측에 대해 낮은 RMSE, MAE 및 첨두수위에서는 오차를 나타낸다는 것을 알 수 있었다. 8개의 Case 중 지체시간 및 정규화를 적용한 Case 7은 다른 Case보다 좋은 결과를 나타냈다. 다양한 데이터 전처리 과정 중 지체시간 및 정규화를 적용하였을 때 예측정확도를 향상시킬 수 있었다.

향후 연구를 통해 각 기법에 대한 은닉층 및 결과에 영향을 주는 인자의 자료구축과 함께 수위자료에 대한 추가적인 데이터 전처리 과정의 적용이 이루어진다면 기존보다 향상된 수위예측결과를 나타낼 것이다.

감사의 글

본 연구는 2021년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행되었습니다(NRF-2019R111A3A01059929).

References

- Abrahart, R., Kneale, P.E., and See, L.M. (2004). *Neural networks for hydrological modeling*. CRC Press, London, UK.
- Assem, H., Ghariba, S., Makrai, G., Johnston, P., Gill, L., and Pilla, F. (2017). Urban water flow and water level prediction based on deep learning. *Joint European conference on machine learning and knowledge discovery in databases*. Springer, Skopje, Macedonia, Vol. 3, pp. 317-329.
- Bae, Y., Kim, J., Wang, W., Yoo, Y., Jung, J., and Kim, H.S. (2019). Monthly inflow forecasting of Soyang River dam using VARMA and machine learning models. *Journal of Climate Research*, Vol. 14, No. 3, pp. 183-198.
- Bicknell, B.R., Imhoff, J.C., Kittle Jr, J.L., Donigan Jr, A.S., and Johanson, R.C. (1996). *Hydrological simulation program FORTRAN*. User's manual for release 11. EPA, U.S.
- Blöschl, G., Hall, J., Viglione, A., Perdigão, R.A., Parajka, J., Merz, B., et al. (2019). Changing climate both increases and decreases European river floods. *Nature*, Vol. 573, No. 7772, pp. 108-111.
- Chen, L., Singh, V.P., Lu, W., Zhang, J., Zhou, J., and Guo, S. (2016). Streamflow forecast uncertainty evolution and its effect on real-time reservoir operation. *Journal of Hydrology*, Vol. 540, pp. 712-726.
- Chen, P.A., Chang, L.C., and Chang, F.J. (2013). Reinforced recurrent neural networks for multi-step-ahead flood forecasts. *Journal of Hydrology*, Vol. 497, pp. 71-79.
- Choi, C., Kim, J., Han, H., Han, D., and Kim, H.S. (2020). Development of water level prediction models using machine learning in wetlands: A case study of Upo Wetland in South Korea. *Water*, Vol. 12, No. 1, pp. 93-110.
- Devia, G.K., Ganasri, B.P., and Dwarakish, G.S. (2015). A review on hydrological models. *Aquatic Procedia*, Vol. 4, pp. 1001-1007.
- Ghumman, A.R., Ghazaw, Y.M., Sohail, A.R., and Watanabe, K. (2011). Runoff forecasting by artificial neural network and conventional model. *Alexandria Engineering Journal*, Vol. 50, No. 4, pp. 345-350.
- Granata, F., Gargano, R., and De Marinis, G. (2016). Support vector regression for rainfall-runoff modeling in urban drainage: A comparison with the EPA's storm water management model. *Water*, Vol. 8, No. 3, pp. 69-81.
- Han, H., Choi, C., Jung, J., and Kim, H.S. (2021). Application of sequence to sequence learning based LSTM model (LSTM-s2s) for forecasting dam inflow. *Journal of Korea Water Resources Association*, Vol. 54, No. 3, pp. 157-166.
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, Vol. 9, No. 8, pp. 1735-1780.
- Jung, S.H., Lee, D.E., and Lee, K.S. (2018). Prediction of river water level using deep-learning open library. *Journal of the Korean Society of Hazard Mitigation*, Vol. 18, No. 1, pp. 1-11.
- Kang, N.R., Noh, H.S., Lee, J.S., Lim, S.H., and Kim, H.S. (2013). Runoff simulation of an urban drainage system using radar rainfall data. *Journal of Wetlands Research*, Vol. 15, No. 3, pp. 413-422.
- Kim, B.K., Kim, S.D., Lee, E.T., and Kim, H.S. (2007). Methodology for estimating ranges of SWAT model parameters: Application to Imha Lake inflow and suspended sediments. *Journal of the Korean Society of Civil Engineers*, Vol. 27, No. 6B, pp. 661-668.
- Kim, D., Kim, J., Kwak, J., Necesito, I.V., Kim, J., and Kim, H.S. (2020). Development of water level prediction models using deep neural network in mountain wetlands. *Journal of Wetlands Research*, Vol. 22, No.

- 2, pp. 106-112.
- Kim, K.S. (2010). *A study on the real time forecasting for monthly inflow daecheong dam using hydrologic time series analyses*. Master Thesis, Seokyeong University, pp. 32-54.
- Lee, K., Choi, C., Shin, D.H., and Kim, H.S. (2020). Prediction of heavy rain damage using deep learning. *Water*, Vol. 12, No. 7, pp. 1942-1959.
- Mahsa, M., and Lee, T. (2018). Comparison of optimization algorithms in deep learning-based neural networks for hydrological forecasting: Case study of nam river daily runoff. *Journal of the Korean Society of Hazard Mitigation*, Vol. 18, No. 6, pp. 377-384.
- Mok, J.Y., Choi, J.H., and Moon, Y.I. (2020). Prediction of multipurpose dam inflow using deep learning. *Journal of Korea Water Resources Association*, Vol. 53, No. 2, pp. 97-105.
- Montanari, A., Rosso, R., and Taqqu, M.S. (1997). Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. *Water Resources Research*, Vol. 33, No. 5, pp. 1035-1044.
- Mosavi, A., Ozturk, P., and Chau, K.W. (2018). Flood prediction using machine learning models: Literature review. *Water*, Vol. 10, No. 11, p. 1536.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., and Williams, J.R. (2011). *Soil and water assessment tool theoretical documentation version 2009*. Texas Water Resources Institute, Temple, TX.
- Noh, H., Lee, J., Kang, N., Lee, D., Kim, H.S., and Kim, S. (2016). Long-term simulation of daily streamflow using radar rainfall and the SWAT model: A case study of the Gamcheon basin of the Nakdong River, Korea. *Advances in Meteorology*, Vol. 2016, pp. 431-442.
- Olah, C. (2015). *Understanding lstm networks*. GITHUB blog. Retrieved November 16, 2021, from <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Palmer, M., and Ruhi, A. (2019). Linkages between flow regime, biota, and ecosystem processes: Implications for river restoration. *Science*, Vol. 365, No. 6459, eaaw2087.
- Park, M.K., Yoon, Y.S., Lee, H.H., and Kim, J.H. (2018). Application of recurrent neural network for inflow prediction into multi-purpose dam basin. *Journal of Korea Water Resources Association*, Vol. 51, No. 12, pp. 1217-1227.
- Ramachandran, P., Zoph, B., and Le, Q.V. (2017). Searching for activation functions. arXiv preprint arXiv:1710.05941.
- Riad, S., Mania, J., Bouchaou, L., and Najjar, Y. (2004). Predicting catchment flow in a semi-arid region via an artificial neural network technique. *Hydrological Processes*, Vol. 18, No. 13, pp. 2387-2393.
- Shoaib, M., Shamseldin, A.Y., Melville, B.W., and Khan, M.M. (2016). A comparison between wavelet based static and dynamic neural network approaches for runoff prediction. *Journal of Hydrology*, Vol. 535, pp. 211-225.
- Yan, J., Jin, J., Chen, F., Yu, G., Yin, H., and Wang, W. (2018). Urban flash flood forecast using support vector machine and numerical simulation. *Journal of Hydroinformatics*, Vol. 20, No. 1, pp. 221-231.
- Zhang, Q., Xiao, M., Singh, V.P., and Li, J. (2012). Regionalization and spatial changing properties of droughts across the Pearl River basin, China. *Journal of Hydrology*, Vol. 472, pp. 355-366.

Received	December 10, 2021
Revised	December 13, 2021
Accepted	December 27, 2021