



Water level prediction in Taehwa River basin using deep learning model based on DNN and LSTM

Lee, Myungjin^a · Kim, Jongsung^b · Yoo, Younghoon^c · Kim, Hung Soo^d · Kim, Sam Eun^e · Kim, Soojun^{f*}

^aPost-Doctoral Researcher, Institute of Water Resources System, Inha University, Incheon, Korea

^bPost-Doctoral Researcher, Institute of Water Resources System, Inha University, Incheon, Korea

^cPh.D candidate, Program in Smart City Engineering, Inha University, Incheon, Korea

^dProfessor, Department of Civil Engineering, Inha University, Incheon, Korea

^ePh.D candidate, Department of Civil Engineering, Inha University, Incheon, Korea

^fAssociate Professor, Department of Civil Engineering, Inha University, Incheon, Korea

Paper number: 21-078

Received: 27 September 2021; Revised: 25 October 2021; Accepted: 26 October 2021

Abstract

Recently, the magnitude and frequency of extreme heavy rains and localized heavy rains have increased due to abnormal climate, which caused increased flood damage in river basin. As a result, the nonlinearity of the hydrological system of rivers or basins is increasing, and there is a limitation in that the lead time is insufficient to predict the water level using the existing physical-based hydrological model. This study predicted the water level at Ulsan (Taehwagyo) with a lead time of 0, 1, 2, 3, 6, 12 hours by applying deep learning techniques based on Deep Neural Network (DNN) and Long Short-Term Memory (LSTM) and evaluated the prediction accuracy. As a result, DNN model using the sliding window concept showed the highest accuracy with a correlation coefficient of 0.97 and RMSE of 0.82 m. If deep learning-based water level prediction using a DNN model is performed in the future, high prediction accuracy and sufficient lead time can be secured than water level prediction using existing physical-based hydrological models.

Keywords: Deep learning, Deep neural network, Long short-term memory, Water level prediction

DNN 및 LSTM 기반 딥러닝 모형을 활용한 태화강 유역의 수위 예측

이명진^a · 김종성^b · 유영훈^c · 김형수^d · 김삼은^e · 김수진^{f*}

^a인하대학교 수자원시스템 연구소 박사후 연구원, ^b인하대학교 수자원시스템 연구소 박사후 연구원, ^c인하대학교 스마트시티공학과 박사과정,

^d인하대학교 사회인프라공학과 교수, ^e인하대학교 토목공학과 박사과정, ^f인하대학교 사회인프라공학과 부교수

요 지

최근 이상 기후로 인해 극한 호우 및 국지성 호우의 규모 및 빈도가 증가하여 하천 주변의 홍수 피해가 증가하고 있다. 이에 따라 하천 또는 유역 내 수문학적 시스템의 비선형성이 증가하고 있으며, 기존의 물리적 기반의 수문 모형을 활용하여 홍수위를 예측하기에는 선행시간이 부족한 한계점이 존재한다. 본 연구에서는 Deep Neural Network (DNN) 및 Long Short-Term Memory (LSTM) 기반의 딥러닝 기법을 적용하여 울산시(태화교) 지점의 수위를 0, 1, 2, 3, 6, 12시간에 대해 선행 예측을 수행하였고 예측 정확도를 비교 분석하였다. 그 결과 sliding window 개념을 적용한 DNN 모형이 선행시간 12시간까지 상관계수 0.97, RMSE 0.82 m로 가장 높은 정확도를 보이고 있음을 확인하였다. 향후 DNN 모형을 활용하여 딥러닝 기반의 수위 예측을 수행한다면 기존의 물리적 모형을 통한 홍수위 예측보다 향상된 예측 정확도와 충분한 선행시간을 확보할 수 있을 것으로 판단된다.

핵심용어: 딥러닝, DNN, LSTM, 수위 예측

*Corresponding Author. Tel: +82-32-860-7563

E-mail: sk325@inha.ac.kr (S. Kim)

1. 서론

최근 전 세계적으로 기후변화로 인해 이상기후가 발생하여 높은 강도의 재해가 빈번하게 발생하고 있다. 2020년 여름에는 54일이라는 역대 최장 장마와 연속된 3개의 태풍이 발생하여 전국적으로 큰 호우 피해가 발생하였다. 특히 2020년 8월 7일에 발생한 집중호우로 인해 영산강 및 섬진강 권역에서는 200년 빈도의 계획 홍수위를 초과한 수위가 발생하여 도심과 농경지가 침수되고 하천 호안, 제방 등의 시설물이 유실되는 심각한 피해가 발생하였다(Lee *et al.*, 2020). 이와 같은 재해를 저감하기 위해 국제적으로 재난에 대한 경각심 및 방재 대책의 관심이 증가하고 있다(Kim *et al.*, 2019a).

침수를 유발하는 호우 사상은 다양한 형태와 빈도로 발생할 수 있으며, 이에 대한 특징을 사전에 파악하는 것은 방재 대책 수립에 중요한 사항이다(Kim *et al.*, 2020b). 최근 인공지능(Artificial intelligence, AI)의 발전으로 인해 다양한 분야에서 그 적용성과 활용성이 입증되고 있으며, 수자원 분야에서도 홍수 및 재난 예측을 위한 새로운 모델로 널리 사용되고 있다(Bae *et al.*, 2019; Kim *et al.*, 2019b, 2020b). 기존에는 물리적 기반의 모형을 사용하여 유출량 및 수위를 산정하는 분석이 주로 적용되어 왔으나, 관측 자료의 축적으로 인해 많은 양의 자료를 활용한 기계 학습 분야가 새로운 모델링 방법으로 부각되고 있다.

최근 기계 학습을 이용한 수위 예측에 대한 국외 연구로는 Dawson and Wilby (1998)은 역전파 알고리즘 신경망을 이용하여 강우-유출 모델링을 수행한 결과, 데이터 최적화에 따라 홍수예측의 정확도가 향상되는 것을 입증하였다. Coulibaly and Ancil (1999)은 효율적 저수지 운영을 위해 순환신경망(Recurrent Neural Network, RNN)을 적용한 단기 유입량 예측을 수행하였고, 예측 결과를 통해 실시간으로 수문학적 예측이 가능한 것을 확인하였다. Chang *et al.* (2014)는 정적 인공신경망인 BPNN (Back- Propagation Neural Network)와 강우에 대한 저류지 수위 예측에 대해 동적 인공신경망인 Elman NN과 NARX 인공신경망의 비교를 실시하였고, NARX 인공신경망의 우수한 성능을 검증하였다. Adnan *et al.* (2017)은 지하 수위 예측을 위해 기계학습 알고리즘 중 SVM (support vector machine)과 ANN (Artificial Neural Network)을 사용하여 분석을 수행하였고, 두 알고리즘 모두 높은 정확도로 지하수위를 예측할 수 있는 것으로 평가되었다. Pan *et al.* (2020)은 미국 Erie 호수의 수위를 예측하기 위해 가우시안 프로세스(GP), 다중 선형 회귀(MLR), 다층 퍼셉트론(MLP), M5P 모델 트리, 랜덤 포레스트(RF), k-최근접 이웃

(KNN)을 포함한 머신 러닝(ML) 알고리즘을 적용하였으며, MLR과 M5P가 가장 높은 정확도를 보임을 확인하였다. Huang *et al.* (2021)은 ANN, LSTM, Adaptive Neuro-Fuzzy Inference System (ANFIS) 모형과 물리적 기반 모델(SOBEK) 모형을 활용하여 조석의 영향을 받는 유량을 예측하였으며, 물리적 모형은 첨두 유량을 잘 모의하는 반면 LSTM 모형은 평균 유량을 잘 모의하는 것을 확인하였다.

수위 예측에 대한 국내 연구로는 Kim *et al.* (2007)은 저류함 수모형과 SSARR 모형을 이용해 금강 미호천 유역에 대하여 홍수모의를 수행하여 통계적 기법인 다중선형회귀분석과 인공신경망모형을 이용하여 홍수예측을 수행하였으며, 다중선형회귀 모형이 더 적합함을 보였다. Jung *et al.* (2018)은 감조하천인 한강 잠수교를 대상으로 LSTM 모형을 활용하여 수위를 예측한 결과 선행시간 1시간이 가장 높은 정확도를 보이고 있음을 제시하였다. Yoo *et al.* (2019)은 기존 ANN 모형을 활용하여 수위예측을 한 연구들이 시간적 매개변수를 고려하지 못하였다는 한계를 지적하였고, 이를 보완할 수 있는 NARX 모형을 적용하여 한강대교의 수위를 예측하였다. Kim *et al.* (2020a)은 금정산 산지습지의 수위를 예측하기 위해 다중회귀분석, 주성분회귀분석, 인공신경망, DNN을 활용하여 수위 예측모형을 개발하였고, DNN을 활용한 수위 예측모형의 예측력이 가장 우수한 것을 확인하였다. Jung *et al.* (2021)은 Long Short-Term Memory (LSTM) 기반의 딥러닝 기법을 적용하여 섬진강 수계 구례 지점의 하천 수위를 1, 3, 6시간 선행 예측하였으며, 서포트벡터머신(SVM) 모형, 다층퍼셉트론(MLP) 모형을 적용한 결과와의 비교한 결과 LSTM 모형이 다른 모형보다 우수한 예측 성능을 보이고 있음을 제시하였다.

위의 많은 연구들은 주로 자료 기반의 모형으로 유역 및 하도에서 발생하는 수문학적 특성을 반영하지 못하고, 자료의 비선형적 특성을 반영하여 수위를 예측하는 한계점이 존재한다. 실제의 하천 수위는 강우량에 의한 유출 특성 이외에 댐 방류량, 강우의 분포에 따라 유출 특성이 변하기 때문에 이와 같은 특성을 반영하여 모형을 구축할 필요가 있다. 이에 본 연구에서는 다양한 수문학적 인자에 영향을 받는 태화강 유역을 대상으로 DNN 및 LSTM 모형을 활용하여 울산시(태화교) 지점의 수위를 예측하고자 한다. 유역 및 하도에서 발생하는 수문학적 특성을 반영할 수 있도록 태화강 유역의 임계지속기간이 12시간인 점을 반영하여 선행시간 1, 2, 3, 6, 12시간을 설정하고, 해당하는 선행시간 이전의 상류 수위, 댐 방류량 및 강우량 자료를 활용하여 하루에 위치한 울산시(태화교) 지점의 수위를 예측하고, DNN 및 LSTM 모형의 정확도를 비교 분석하고자 한다.

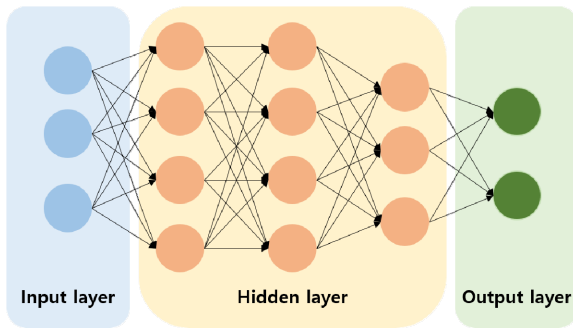


Fig. 1. Deep neural network

2. 연구 방법

2.1 Deep Neural Network (DNN)

인공 지능은 학습, 패턴 인식 등의 문제를 해결할 수 있는 컴퓨터 공학 분야이며, 시스템에 의해 만들어진 지능, 즉 인공적인 지능을 뜻한다. 인공지능의 발달로 인해 수자원 분야에서도 시계열 예측에 대한 패러다임이 변화하고 있으며, 물리적 모형에만 의존하지 않고 과거의 데이터를 활용하여 모형을 학습시키고, 예측 값을 추정하는 심층 학습(Deep learning, DL)의 연구가 활발하게 진행되고 있다(Hinton et al., 2006; Kim et al., 2019b, 2020b; Seo et al., 2017).

심층신경망(Deep Neural Network, DNN)은 일반적인 인공신경망(Artificial Neural Network, ANN)에서 은닉층을 2개 이상 사용하여 더욱 고도화된 학습이 가능도록 구성된 가장 기본적인 딥 러닝 기법이다(Kim et al., 2020b). 데이터 양이 많아 질수록 과적합(Over Fitting)되는 문제가 되는 기존의 인공신경망과 달리, DNN은 은닉층이 2개 이상으로 구성되어 이러한 문제를 최소화할 수 있는 장점이 있다. Fig. 1은 DNN의 구조를 보여주고 있으며, 입력층, 출력층 및 은닉층으로 구성되어 있다.

2.2 Long Short-Term Memory (LSTM)

순환신경망(Recurrent Neural Network, RNN)은 인공신경망 알고리즘의 한 종류로, 스스로 반복하면서 이전 단계에서 얻은 정보가 지속되는 특징이 있지만, 자료의 처음과 끝이 너무 먼 경우 지난 정보를 계속 이어가는데 문제가 발생한다. 이러한 장기 의존성에 의한 문제를 해결하기 위해 LSTM이 제시되었다(Hochreiter and Schmidhuber, 1997). LSTM은 RNN에서 발생하는 오차 경사의 기울기 소실 및 최적화 오류를 해결하고 입력 자료에 대한 정보를 더욱 장기적으로 기억하기 위해 은닉층에 셀 상태(cell state) 구조를 추가한 모형으로, 이에 대한 구조는 Fig. 2와 같다(Jung et al., 2018).

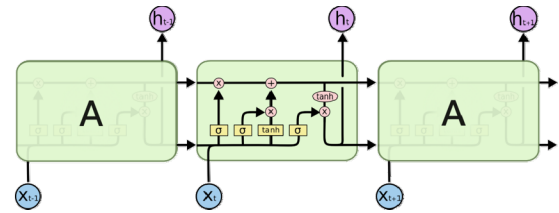


Fig. 2. Structure of LSTM (Olah, 2015)

LSTM에서는 입력게이트(Input gate, i_t), 망각 게이트(forget gate, f_t), 출력 게이트(output gate, o_t)가 존재하여 전 단계의 정보 중에서 버릴 정보와 업데이트 할 정보를 결정하여 다음 단계로 전달한다. 먼저 망각게이트(f_t)는 어떤 정보를 버릴 것인지에 대해 Eq. (1)과 같이 계산하며, 이전 단계의 h_{t-1} 과 현재 단계의 x_t 를 시그모이드 함수(활성 함수)에 적용하여 0과 1사이의 값을 산정한다. 이를 현재 상태와 곱해지게 되며, 이 과정에서 정보를 사용할지 제거할 지를 선택한다.

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

여기서, σ 는 활성화 함수, W_f 는 망각 게이트의 가중치, h_{t-1} 은 이전값의 출력값, x_t 는 현재 입력값, b_f 는 편향을 의미한다. 다음으로, 입력 게이트(i_t)라 불리는 시그모이드 함수가 어떤 데이터를 업데이트 할지를 결정한다. 하이퍼 탄젠트 함수(\tanh)를 통해 새로운 후보 값들인 \bar{C}_t 벡터를 만들게 되고, i_t 값과 조합하여 셀 상태에 더한다.

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\bar{C}_t = \tanh(W_C \times [h_{t-1}, x_t] + b_C) \quad (3)$$

여기서, W_i , W_C 는 각각 함수의 가중치, b_i , b_C 는 각각 함수의 편향을 의미하며, Eqs. (1)~(3)을 이용하여, 과거 상태인 C_{t-1} 을 업데이트해서 새로운 상태 C_t 를 만든다. Eq. (4)를 이용하여 현재 단계에 대한 정보 상태를 업데이트 할 수 있다.

$$C_t = f_t \times C_{t-1} + i_t \times \bar{C}_t \quad (4)$$

2.3 예측력 평가 지표

DNN 및 LSTM 기법을 활용하여 예측한 수위의 정확도를 평가하기 위해 상관계수(Correlation Coefficient, CC), 평균제곱근오차(Root Mean Square Error, RMSE) 및 평균절대오차(Mean Absolute Error, MAE)를 사용하였으며, 이는 Eqs.

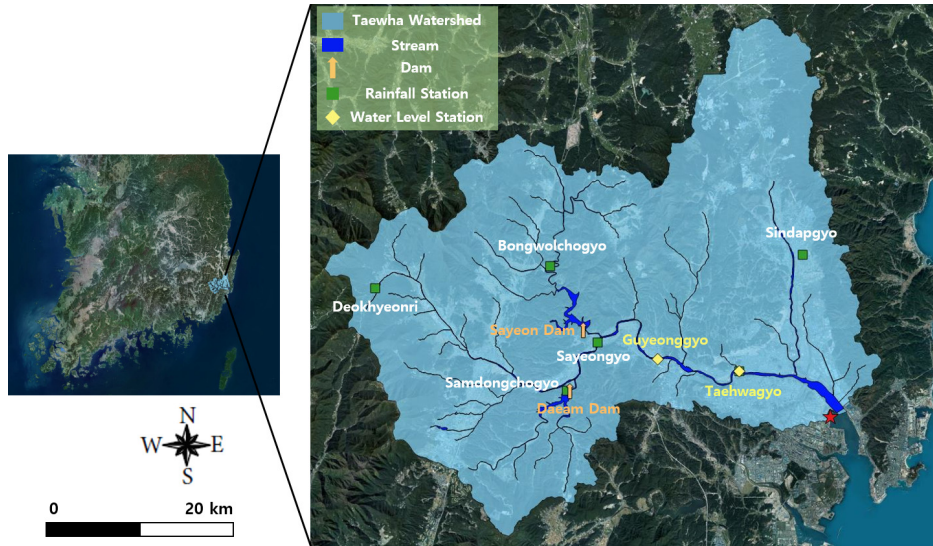


Fig. 3. Taehwa River basin and stations

(5)-(7)과 같다. 상관계수는 1과 가까울수록, RMSE 및 MAE는 0에 가까울수록 예측값이 실제값과 유사하다는 것을 의미한다.

$$CC = \frac{\sum(O_t - \bar{O})(P_t - \bar{P})}{\sqrt{\sum(O_t - \bar{O})^2 \times \sum(P_t - \bar{P})^2}} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (O_t - P_t)^2} \quad (6)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |O_t - P_t| \quad (7)$$

여기서, O_t 와 P_t 는 시간 t 에서의 관측 수위와 예측 수위, \bar{O} 와 \bar{P} 는 관측 수위와 예측 수위의 평균값, N 은 자료의 개수를 의미한다.

3. 연구 대상 및 결과

3.1 대상유역 선정

대상 유역은 조석 및 강우로 인해 복잡한 수문학적 특성을 보이고 있는 태화강 유역에 위치한 울산시(태화교) 지점의 수위를 예측하였다. 울산시(태화교) 지점은 감소하천의 특성을 보이며, 태화강 하류 말단에는 울산 조위관측소가 위치하며 상류에는 조위의 영향을 받지 않는 구영 수위관측소가 위치하고 있다(Fig. 3).

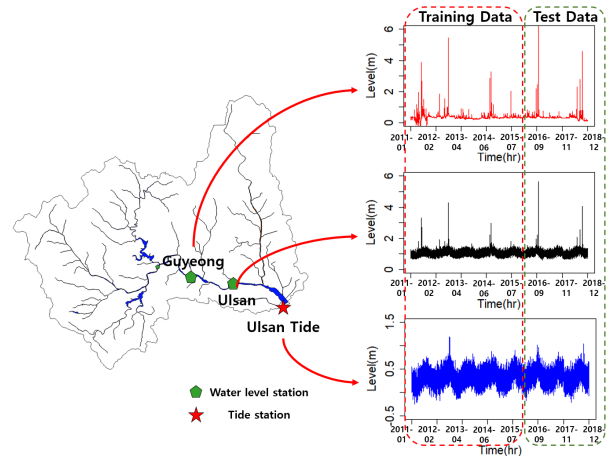


Fig. 4. Time series of water level and tide

태화강 유역은 조석의 영향에 따라 수위 및 유량이 변동하는 특성을 가지고 있으며, 만조시에는 배수효과로 인하여 홍수피해가 가중되기도 한다(Lee *et al.*, 2018). 2016년에 발생한 태풍 ‘차바’ 내습시 태화강 유역 내에 위치한 울산광역시 중구는 만조 시기와 태풍 내습시기가 겹쳐 하천으로 유출량이 배수되지 못하여 홍수 피해가 가중되는 현상이 발생하였다. 수위 예측의 대상 지점은 울산시(태화교) 수위 관측소로, Fig. 4와 같이 울산시(태화교) 수위 관측소의 수위는 저수위에서 조석의 영향을 받아 변동하고 있으며, 강우 발생 시 유출량에 따라 수위가 상승하는 특성을 보이고 있다. 상류에 위치한 울산시(구영교) 수위 관측소는 감소 구간이 아닌 일반적인 하천으로, 강우에 의해 수위가 상승하는 특성이 있어 상류의 수위 자료를 통해 감소 구간의 수위를 예측할 수 있다.

울산시(태화교)의 수위를 예측하기 위해 상류에 위치한 울산시(구영교) 수위관측소와 태화강 중류에 위치한 울산시(사연교), 하류에 위치한 울산시(태화교) 지점에 대해 2011년부터 2018년까지 총 8년간의 시단위 자료를 수집하였다. 또한, 태화강 유역은 중류에 사연댐 및 대암댐이 위치하고 있어 댐의 방류량에 따라 수위가 변동하는 특성을 보이고 있으므로, 대암댐의 방류량도 입력 자료로 사용하였다. 모형의 학습을 위한 훈련 자료(Training data)는 2011년부터 2015년까지 5년의 자료를 사용하였으며, 2016년부터 2018년까지 3년의 자료를 평가 자료(Test data)로 사용하여 울산시(태화교) 지점의 수위를 예측하고, 관측 수위와 비교하여 모형의 정확도를 평가하였다.

3.2 DNN 모형 구축

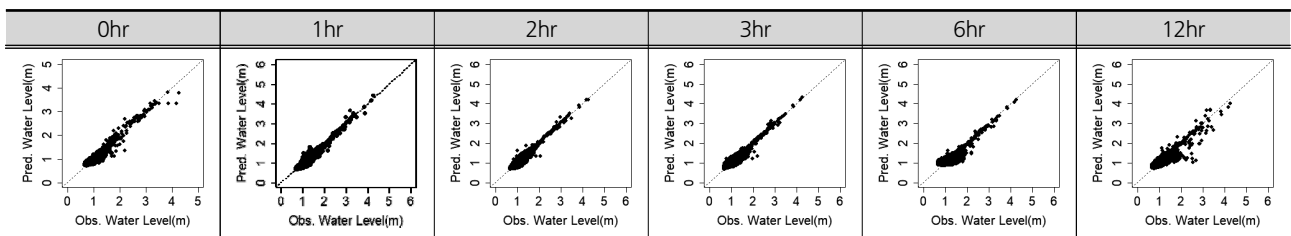
감조하천의 특성을 보이고 있는 울산시(태화교)의 수위를 예측하기 위해 Table 1과 같이 DNN 모형의 입력자료, 은닉층 (Hidden layer), 활성화 함수(Activation function) 및 손실 함수(Loss function)를 구축하였다. DNN의 은닉층은 3개의 layer로 구성하였으며, 각 레이어의 유닛은 시행착오법을 통해 가장 적합한 유닛의 개수를 선정하였으며, 그 결과 512, 256 및 128개로 구축하였다. 활성화 함수는 비선형성을 고려할 수 있고, gradient vanishing 문제를 해결할 수 있는 ReLU 함수를 사용하였으며, 손실함수는 평균제곱근오차(Root Mean Square Error, RMSE)를 사용하여 학습을 위한 기준을 지정하였다. 또한, 학습 데이터의 과잉학습에 의해 실제 데이터에 대한 오차가 증가하는 현상인 과적합(Overfitting)이 발생하여 예측값의 정도가 하락하는 문제점을 해결하기 위해 Drop out 기법을 적용하였다(Kim et al., 2019b). DNN의 학습 및 예측을 위

한 입력 자료는 상류지점인 구영관측소의 수위 및 강우량, 대암댐의 방류량, 하류에 위치한 울산 조위관측소 및 하류의 울산 강우량을 사용하였다. 이때, 선행시간을 0, 1, 2, 3, 6, 12시간으로 설정하여 울산 수위관측소의 수위를 예측하였으며, 각 시계열 자료는 Min-Max 표준화를 통해 정규화를 수행하였다. 시계열 예측에서 데이터를 연속적으로 인식하기 위해서 슬라이딩 윈도우 기법(Sliding window technique)이 주로 사용되고 있으며, 이 기법은 데이터를 순차적으로 윈도우 사이즈에 맞춰서 이동시키는 형태로 사용된다(Koç, 1995). 예를 들어 3시간 선행시간에 대해 울산관측소의 수위를 예측할 때, -1hr의 종속변수, -2hr의 종속변수, -3hr의 종속변수를 각각 입력 자료로 사용하여 모형을 학습하였다.

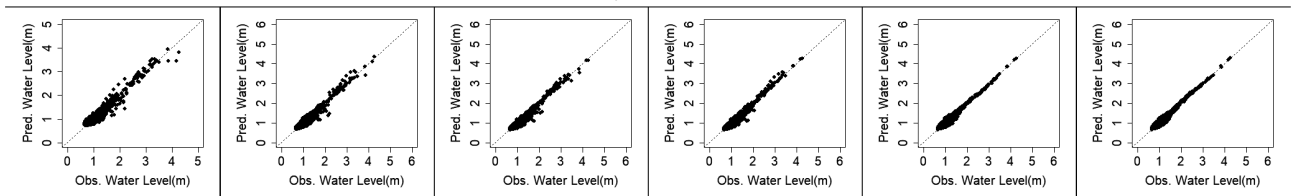
이에 대한 결과는 Fig. 5와 같으며, Fig. 5(a)는 단순하게 선

Table 1. Parameters for DNN and LSTM Model

Model	Deep Neural Network (DNN)	Long Short-Term Memory (LSTM)
Hidden layer	3 layers (Units : 512, 256, 128)	4 layers (Units : 128, 64, 32, 1)
Activation function	ReLU	ReLU
Dropout	O	O
Loss function	RMSE	RMSE
Lead Time	0hr, 1hr, 2hr, 3hr, 6hr, 12hr	
Input data	Guyeong water level, Ulsan tide level, Guyeong rainfall, Ulsan rainfall, Daeam dam's discharge	
Output data	Ulsan water level	
Training data	48324 (2011~2015)	
Test data	21084 (2016~2018)	



(a) Simple DNN result



(b) DNN result using sliding window technique

Fig. 5. Scatter plot using DNN model

행시간 1시간~12시간 전의 종속변수를 사용하여 모형을 학습한 결과이며, Fig. 5(b)는 target time 이전의 데이터를 time window의 개념을 사용하여 모형을 학습한 결과이다. 그림에서 확인할 수 있듯이 time window의 개념을 사용하였을 때, 과거의 데이터를 모두 사용하여 학습한 결과가 선형성을 더 잘 보여주고 있음을 확인할 수 있으며, 강우가 발생하였을 때 발생하는 고수위도 관측 수위와 유사한 값을 보이고 있다.

3.3 LSTM 모형 구축

본 연구에서는 선행시간을 고려하여 울산 수위관측소의 수위를 예측하는 것이 목적이므로, DNN 모형과 동일하게 선행시간을 0, 1, 2, 3, 6, 12시간으로 설정하고 수위를 예측하였다. LSTM 모형의 경우 은닉층(Hidden layer)은 시행착오법을 통해 가장 적합한 4개의 layer를 사용하였으며, 각 레이어의 유닛은 128개, 64개, 32개 및 1개로 설정하였다. 활성화 함수는 DNN 모형과 동일하게 비선형성을 고려할 수 있고, gradient vanishing 문제를 해결할 수 있는 ReLU 함수를 사용하였으며, 손실함수는 평균제곱근오차(Root Mean Square Error, RMSE)를 사용하여 학습을 위한 기준을 설정하였다. DNN 모형과 동일하게 LSTM 모형 역시 n시간 전의 선행시간 자료만 사용하는 simple LSTM 모형과 n시간까지의 선행시간에 대한 자료를 모두 사용하는 time window LSTM 모형을 구축하여 모형을 학습하였다. 이에 대한 결과는 Fig. 6과 같으며 Fig. 6(a)는 simple LSTM 모형에 대한 scatter plot을, Fig. 6(b)는 time window LSTM 모형에 대한 scatter plot을 나타내고 있다. DNN 모형의 결과와 동일하게 LSTM 모형의 경우에도 단순히 n시간 전의 입력 자료만 활용하여 학습한 결과보다 n시간

전까지의 모든 자료를 입력 자료로 활용하여 학습한 결과가 더 좋은 결과를 나타내고 있음을 확인할 수 있다.

3.4 모형 결과 및 분석

구축한 4개의 모형에 대한 정확도를 평가하기 위해 2.3절에서 제시한 3개의 평가지표를 활용하여 정량적인 비교를 수행하였으며, 이에 대한 결과는 Table 2와 같다. 선행시간이 없이 실시간 데이터로 구축한 모형을 제외하고는 모든 경우에서 time window를 고려하여 학습한 모형이 좋은 성능을 보이고 있는 것을 확인할 수 있으며, LSTM 모형보다 DNN 모형이 더 우수한 성능을 보이고 있다. 이는 Figs. 5 and 6에서 확인할 수 있듯이 LSTM 모형의 경우 DNN 모형에 비해 고수위에 대한 예측값이 관측값에 비해 과소 추정되고 있어 RMSE 및 MAE는 크게 나타나는 반면, 상관계수는 낮게 산정된 것으로 판단된다.

구축한 4개의 모형의 정확도를 평가하기 위해 평가 기간인 2016년부터 2018년까지의 입력 데이터를 활용하여 울산 수위관측소의 수위를 예측하였으며, Figs. 7 and 8과 같이 시계열 그래프로 도시하였다. Fig. 7에서 확인할 수 있듯이 Simple DNN 모형의 경우 선행시간이 길어질수록 강우가 발생하였을 때 상승하는 고수위에 대한 예측도가 떨어지는 반면, time window를 고려하여 예측한 DNN 모형은 고수위에 대한 정확도가 높게 산정되는 것을 확인할 수 있다.

LSTM 모형의 경우에도 simple LSTM 모형은 선행시간이 길어질수록 고수위에서의 예측값이 과소 추정되는 문제가 있었지만, time window를 고려하여 예측한 모형은 고수위를 과대 추정하는 경향을 Fig. 8에서 확인할 수 있다. 즉, Fig. 8의 산

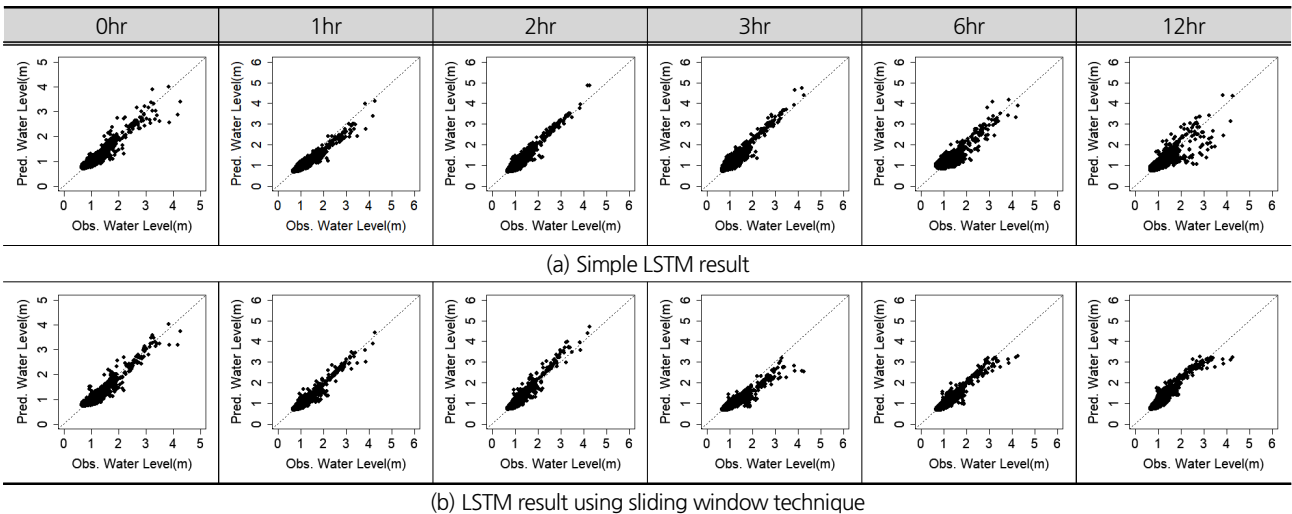
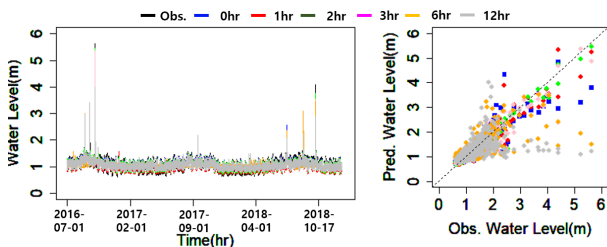


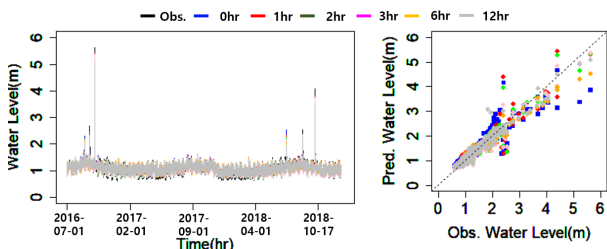
Fig. 6. Scatter plot using LSTM model

Table 2. Comparisons of DNN and LSTM model

Lead Time	Model	Model performance			Lead Time	Model	Model performance		
		CC	RMSE (m)	MAE (m)			CC	RMSE (m)	MAE (m)
0hr	DNN_SIM	0.953	0.057	0.061	3hr	DNN_SIM	0.676	0.987	0.998
	DNN_WD	0.953	0.060	0.066		DNN_WD	0.972	0.991	1.005
	LSTM_SIM	0.949	0.061	0.078		LSTM_SIM	0.675	0.985	0.993
	LSTM_WD	0.948	0.061	0.076		LSTM_WD	0.955	0.996	1.009
1hr	DNN_SIM	0.956	0.992	1.002	6hr	DNN_SIM	0.524	0.994	1.013
	DNN_WD	0.956	0.981	0.991		DNN_WD	0.973	0.981	0.998
	LSTM_SIM	0.952	0.993	1.001		LSTM_SIM	0.585	0.989	0.998
	LSTM_WD	0.952	0.991	1.001		LSTM_WD	0.966	0.992	1.021
2hr	DNN_SIM	0.849	0.985	0.995	12hr	DNN_SIM	0.859	0.992	1.003
	DNN_WD	0.972	0.992	1.004		DNN_WD	0.973	0.982	0.993
	LSTM_SIM	0.848	0.988	0.996		LSTM_SIM	0.844	0.988	0.997
	LSTM_WD	0.967	0.979	0.990		LSTM_WD	0.965	0.978	2.941

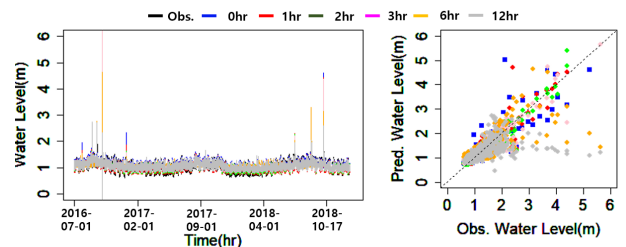


(a) Prediction results of simple DNN model

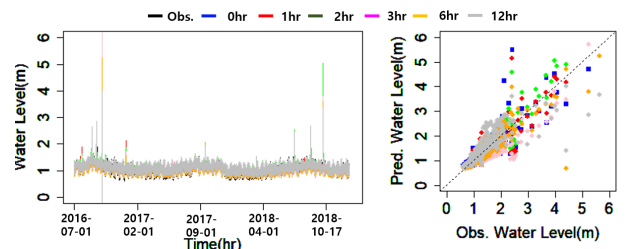


(b) Prediction results of time window DNN model

Fig. 7. Time series of DNN model



(a) Prediction results of simple LSTM model



(b) Prediction results of time window LSTM model

Fig. 8. Time series of LSTM model

점도에서 선행시간 12시간에 대한 simple LSTM 모형은 관측 값에 비해 과소 추정되는 경향을 보이고 있지만, time window 를 고려한 LSTM 모형은 어느 정도의 선형 특성을 보이는 것을 확인할 수 있다. 이는 강수량 및 댐 방류량의 경우 연속적인 시 계열로 상류에서 하류로 영향을 미치고 있으나, Simple LSTM 모형은 단순하게 1, 2, 3, 6, 12시간 전의 데이터만 사용하기 때문에 유역의 도달시간 및 유출 특성을 반영하지 못하기 때문에 실제 관측 수위보다 과소 추정되는 문제가 발생한다.

4개의 모형에 대한 정확도는 Table 3과 같이 정리할 수 있으며, 이는 선행시간 0, 1, 2, 3, 6, 12시간에 대한 통계 지표의 평균

Table 3. Performance of DNN and LSTM models

Model	Model performance		
	CC	RMSE (m)	MAE (m)
Simple DNN Model	0.738	0.977	0.987
Time Window DNN Model	0.966	0.821	0.824
Simple LSTM Model	0.750	0.974	0.982
Time Window LSTM Model	0.633	0.821	1.001

값을 보여주고 있다. 학습 데이터를 통한 결과와 동일하게 평가 데이터를 사용한 결과에서도 time window를 고려한 DNN

모형이 가장 우수한 성능을 나타내고 있다. 이는 time window를 사용할 경우 강수량, 유출량 및 댐 방류량이 하루로 도달하는 유출 특성을 반영할 수 있으므로, 관측 수위와 유사한 값으로 예측할 수 있기 때문이다.

선행시간별 정확도를 검토할 경우 선행시간 3시간까지는 4개의 모형 모두 상관계수가 0.95 이상, RMSE 및 MAE는 0.97 m 이하로 높은 정확도를 보이고 있으나 선행시간 6시간에는 상관계수가 0.5 이하로 정확도가 현저하게 떨어지는 것을 확인할 수 있다. 하지만, 선행시간 12시간에서는 상관계수가 0.8 이상, RMSE 및 MAE는 약 0.98 m 정도로 선행시간 3시간까지의 결과와 거의 유사한 값을 보이는 것을 확인할 수 있다(Table 2). 본 연구 결과를 토대로 12시간의 선행시간까지 태화강 하류의 수위를 예측할 수 있으며, 정확도 있는 홍수 예경보는 3시간의 선행시간을 확보하여 수행할 수 있을 것으로 판단된다.

4. 결론

본 연구에서는 복잡한 수문특성을 보이는 태화강 하류 지역을 대상으로 상류의 강수량, 수위 및 댐 방류량, 하류의 조위 자료 및 하류의 강수량 자료를 활용하여 울산시(태화교) 지점의 수위를 예측하였다. 수위 예측을 위한 모형으로는 최근 수 자원 분야에서 활발하게 사용 중인 DNN 모형 및 LSTM 모형을 활용하였으며, 선행시간(1, 2, 3, 6, 12시간)을 고려하여 수위 예측의 정확도를 평가하였다. 모형의 학습을 위해 2011년부터 2015년까지의 시자료를 활용하여 모형을 구축하였으며, 2016년부터 2018년까지의 시자료를 활용하여 구축한 모형에 대한 정확도 평가를 수행하였다. 수위 예측을 위한 모형은 n시간(n = 1, 2, 3, 6, 12)전의 자료를 활용하는 Simple DNN 모형 및 LSTM 모형과 1~n시간 전까지의 자료를 sliding window 기법을 통해 적용한 Time window DNN 모형 및 LSTM 모형을 구축하였다. 수위 예측 결과는 단순히 n시간 전의 자료만 사용한 Simple DNN 모형 및 LSTM 모형보다 Time window DNN 모형 및 LSTM 모형이 상관계수 약 0.30, RMSE 약 0.01 m 및 MAE 0.50 m 정도의 향상된 성능을 보이는 것을 확인하였다. 이에 대한 오차의 원인으로는 n시간 이전의 자료만 사용한 모형의 경우 연속적인 시계열 특성을 보이는 강우 및 유출 특성을 반영하지 못하기 때문에 낮은 정확도를 보이는 반면, time window를 적용한 모형의 경우 상류에서 발생한 유출량이 하루로 도달할 때까지 발생하는 수문 특성이 데이터의 입력 자료로 반영이 되기 때문에 하류에서 발생하는 고수위도 높은 정확도로 예측할 수 있다.

향후 연구에서는 상류의 강수량, 댐 방류량 등의 자료를 활용하여 태화강 유역의 수위 예측의 선행시간 확보에 대한 분석을 수행하고자 한다. 본 연구 결과를 활용한다면 기존의 물리적 기반의 모형을 통해 준실시간으로 수행되었던 수문 모형의 단점을 보완할 수 있을 것으로 판단되며, 충분한 선행시간을 활용하여 치수 대책을 수립한다면 태화강 하류의 홍수 피해를 저감할 수 있을 것으로 기대된다.

감사의 글

본 결과물은 환경부의 재원으로 한국환경산업기술원 물관리연구사업의 지원을 받아 연구되었습니다. 이에 감사드립니다(127570).

References

- Adnan, R.M., Yuan, X., Kisi, O., and Yuan, Y. (2017). "Streamflow forecasting using artificial neural network and support vector machine models." *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, Vol. 29, No. 1, pp. 286-294.
- 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.
- Chang, F.J., Chen, P.A., Lu, Y.R., Huang, E. and Chang, K.Y. (2014). "Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control." *Journal of Hydrology*, Vol. 517, pp. 836-846.
- Coulibaly, P., and Anctil, F. (1999). "Real-time short-term natural water inflows forecasting using recurrent neural networks." *International Joint Conference on IEEE*, Washington DC, U.S., pp. 3802-3805.
- Dawson, C.W., and Wilby, R. (1998). "An artificial neural network approach to rainfall-runoff modelling". *Hydrological Sciences Journal*, Vol. 43, No. 1, pp. 47-66.
- Hinton, G.E., Osindero, S., and Teh, Y.W. (2006). "A fast learning algorithm for deep belief nets." *Neural Computation*, Vol. 18, No. 7, pp. 1527-1554.
- Hochreiter, S., and Schmidhuber, J. (1997). "Long short-term memory." *Neural Computation*, Vol. 9, No. 8, pp. 1735-1780.
- Huang, X., Li, Y., Tian, Z., Ye, Q., Ke, Q., Fan, D., Mao, G., Chen, A. Liu, J. (2021). "Evaluation of short-term streamflow prediction methods in Urban river basins." *Physics and Chemistry of the Earth, Parts A/B/C*, Vol. 123.
- Jung, J., Mo, H., Lee, J., Yoo, Y., and Kim, H.S. (2021). "Flood stage forecasting at the Gurye-Gyo station in Sumjin River using

- LSTM-based deep learning models.” *Journal of the Korean Society of Hazard Mitigation*, Vol. 21, No. 3, pp. 193-201.
- Jung, S., Cho, H., Kim, J., and Lee, G. (2018). “Prediction of water level in a tidal river using a deep-learning based LSTM model.” *Journal of Korea Water Resources Association*, Vol. 51, No. 12, pp. 1207-1216.
- Kim, B.J. (2007). *Comparative study of storage function and SSARR models for the flood hydrograph forecasting of a Miho Stream*. Ph. D. dissertation, Inha University,
- Kim, D., Kim, J., Kwak, J., Necesito, I.V., Kim, J., and Kim, H.S. (2020a). “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, H.I., Lee, J.Y., Han, K.Y. and Cho, J.W. (2020b). “Applying observed rainfall and deep neural network for urban flood analysis.” *Journal of the Korean Society of Hazard Mitigation*, Vol. 20, No. 1, pp. 339-350.
- Kim, Y.J., Kim, T.W., Yoon, J.S., and Kim, I.H. (2019a). “Study on prediction of similar typhoons through neural network optimization.” *Journal of Ocean Engineering and Technology*, Vol. 33, No. 5, pp. 427-434.
- Kim, Y.J., Kim, T.W., Yoon, J.S., and Kim, M.K. (2019b). “Study of the construction of a coastal disaster prevention system using deep learning.” *Journal of Ocean Engineering and Technology*, Vol. 33, No. 6, pp. 590-596.
- Koç, C.K. (1995). “Analysis of sliding window techniques for exponentiation.” *Computers & Mathematics with Applications*, Vol. 30, No. 10, pp. 17-24.
- Lee, G.H., Ryu, Y.U., and Park, J.S. (2020). “Investigation and analysis of causes of flood damage in the Yeongsan River and Seomjin River basins in August 2020.” *Water for Future*, Vol. 53, No. 11, pp. 21-48.
- Lee, M., You, Y., Kim, S., Kim, K., and Kim, H. (2018). “Decomposition of water level time series of a tidal river into tide, wave and rainfall-runoff components.” *Water*, Vol. 10, No. 11, pp. 1568.
- Olah, C. (2015). Understanding lstm networks, Accessed on August, 2011, <<http://https://kyeongsoo.github.io/>>.
- Pan, M., Zhou, H., Cao, J., Liu, Y., Hao, J., Li, S., and Chen, C.H. (2020). “Water level prediction model based on GRU and CNN.” *Ieee Access*, Vol. 8, pp. 60090-60100.
- Seo, Y., Choi, E., and Yeo, W. (2017). “Reservoir water level forecasting using machine learning models.” *Journal of the Korean Society of Agricultural Engineers*, Vol. 59, No. 3, pp. 97-110.
- Yoo, H.J., Lee, S.O., Choi, S.H., and Park, M.H. (2019). “A study on the data driven neural network model for the prediction of time series data: Application of water surface elevation forecasting in Hangang River bridge.” *Korean Society of Disaster & Security*, Vol. 12, No. 2, pp. 73-82.