



Utilizing deep learning algorithm and high-resolution precipitation product to predict water level variability

Han, Heechan^a · Kang, Narae^b · Yoon, Jungsoo^c · Hwang, Seokhwan^{d*}

^aAssistant Professor, Department of Civil Engineering, Chosun University, Gwangju, Korea

^bSenior Researcher, Department of Hydro Science and Engineering Research, Korea Institute of Civil Engineering and Building Technology, Goyang, Korea

^cSenior Researcher, Department of Hydro Science and Engineering Research, Korea Institute of Civil Engineering and Building Technology, Goyang, Korea

^dResearch Fellow, Department of Hydro Science and Engineering Research, Korea Institute of Civil Engineering and Building Technology, Goyang, Korea

Paper number: 24-020

Received: 4 April 2024; Revised: 21 June 2024; Accepted: 25 June 2024

Abstract

Flood damage is becoming more serious due to the heavy rainfall caused by climate change. Physically based hydrological models have been utilized to predict stream water level variability and provide flood forecasting. Recently, hydrological simulations using machine learning and deep learning algorithms based on nonlinear relationships between hydrological data have been getting attention. In this study, the Long Short-Term Memory (LSTM) algorithm is used to predict the water level of the Seomjin River watershed. In addition, Climate Prediction Center morphing method (CMORPH)-based gridded precipitation data is applied as input data for the algorithm to overcome for the limitations of ground data. The water level prediction results of the LSTM algorithm coupling with the CMORPH data showed that the mean CC was 0.98, RMSE was 0.07 m, and NSE was 0.97. It is expected that deep learning and remote data can be used together to overcome for the shortcomings of ground observation data and to obtain reliable prediction results.

Keywords: CMORPH, Deep learning, LSTM algorithm, Seomjin River, Water level variability prediction

고해상도 강우자료와 딥러닝 알고리즘을 활용한 수위 변동성 예측

한희찬^a · 강나래^b · 윤정수^c · 황석환^{d*}

^a조선대학교 토목공학과 조교수, ^b한국건설기술연구원 수자원하천연구본부 수석연구원,

^c한국건설기술연구원 수자원하천연구본부 수석연구원, ^d한국건설기술연구원 수자원하천연구본부 연구위원

요 지

기후변화로 인한 집중호우의 발생으로 홍수 피해가 심각해지고 있다. 하천의 수위 변동성을 예측하고 신속한 홍수 예·경보를 위해 물리적 기반의 수문 모형이 활용됐다. 최근에는 수문 데이터 간의 비선형적인 관계를 기반으로 머신러닝, 딥러닝 알고리즘을 활용한 수문 모의가 주목받고 있다. 본 연구에서는 Long Short-Term Memory (LSTM) 알고리즘을 활용하여 섬진강 수계의 하천 수위를 예측하고자 한다. 또한 Climate Prediction Center morphing method (CMORPH) 기반의 격자형 강우 자료를 알고리즘의 입력자료로 적용하여 지상 데이터의 한계를 보완하고자 한다. CMORPH 데이터와 LSTM 알고리즘을 결합한 모형의 수위 예측 결과는 평균 CC가 0.98, RMSE는 0.07 m, 그리고 NSE는 0.97로 나타났다. 향후 딥러닝과 원격자료를 활용하여 수위 예측을 수행한다면 지상 관측 데이터의 단점을 보완하고, 신뢰도 높은 예측 결과를 얻을 수 있을 것으로 기대 되는 바이다.

핵심용어: CMORPH, 딥러닝, LSTM 알고리즘, 섬진강, 수위 변동성 예측

*Corresponding Author. Tel: +82-31-910-0241

E-mail: sukany@kict.re.kr (Hwang, Seokhwan)

1. 서론

최근 기후변화와 이상기후에 따른 집중호우의 발생빈도가 증가함에 따라 홍수 피해의 규모가 증가하고 있다(Schreider *et al.*, 2000; Arora and Boer, 2001; Kay *et al.*, 2021). 2020년 여름 섬진강 지역에서는 집중호우로 인해 하천이 범람하면서 주변 도심과 농경지에 심각한 침수피해가 발생하였다(Lee *et al.*, 2021). 이러한 기후변동과 인간 활동이 홍수 발생에 미치는 영향을 해석하기 위해서 다양한 물리학적 기반의 수문 모형이 활용되고 있다(Tessema *et al.*, 2021; Tarekegn *et al.*, 2022).

최근 컴퓨터 시스템과 데이터 관측 기술이 동시에 발전하면서 많은 분야에서 다양한 종류의 머신러닝 알고리즘의 적용성이 입증되고 있다(Xiang *et al.*, 2020). 특히 많은 데이터를 필요로 하는 수문 분야에서도 빅데이터와 머신러닝, 딥러닝 알고리즘을 활용함으로써 강우-유출 모의, 수위 예측, 홍수 예-경보 시스템 개발 등을 위한 새로운 방법으로 주목받고 있다(Han and Morrison, 2022).

또한 데이터 관측 기술이 진보하면서 레이더, 위성 데이터 기반의 수문 데이터 분석 및 활용 범위가 증가하고 있다. 특히 지상 관측이 어려운 산악지형이나 관측소의 공간적 밀도가 낮은 지역을 대상으로 데이터를 제공함으로써 수문학적 분석의 신뢰도를 높이는 데 기여하고 있다(Han *et al.*, 2023). 대표적으로 Climate Prediction Center morphing method (CMORPH; Joyce *et al.*, 2004), Tropical Rain-fall Measuring Mission (TRMM; Huffman *et al.*, 2007), and Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN; Sorooshian *et al.*, 2000)와 같은 위성 기반 자료에서 생성하는 다양한 기상 정보를 활용하여 수문학적 분석을 수행하는 다양한 연구가 진행 중이다(Tobin and Bennett, 2010; Le *et al.*, 2020).

이전 연구를 통해 수위를 포함한 다양한 수문학적 요소들의 예측을 위한 기계학습 알고리즘의 적용성이 평가되었다. 많은 연구에서는 기계학습 알고리즘이 수문학적 예측에 높은 성능을 보이는 것을 확인하였다(Assem *et al.*, 2017; Baek *et al.*, 2020; Park *et al.*, 2022). 예를 들어, Li *et al.* (2022)는 Convolutional Neural Network와 Long-Short Term Memory가 결합된(CNN-LSTM) 알고리즘을 활용하여 3시간 단위의 강우 발생을 예측하는 모형 개발에 관한 연구를 수행하여 기존 머신러닝 알고리즘보다 우수한 예측 정확도를 나타내는 것을 확인하였다. 또한 Hu *et al.* (2018)은 LSTM 알고리즘을 활용하여 강우-유출 사상을 모의하였고 그 결과NSE가 0.9이상으로 나타내는 결과를 도출하였다.

Li *et al.* (2024)는 3가지 딥러닝 알고리즘을 활용하여 최대 7일 이후의 수위 변동을 예측하는 모형을 개발하였다. 개발된 모형을 적용한 결과 80% 이상의 예측 정확도를 나타내었다. 또한 Wang and Wang (2020)은 5가지 머신러닝 알고리즘 기반의 수위 예측 모형을 제안하였다. 이 연구에서는 수위 변동에 영향을 미치는 다양한 기상학적 요인들과 과거 수위 자료의 통계학적 관계를 이용하여 알고리즘을 학습하였고, 그 결과 기존의 물리적 기반 수문 모형보다 더 나은 예측 성능을 나타내었다.

수자원 분야에서 활용성이 주목받고 있는 위성 기반의 원격자료를 활용하여 하천 수위 변동을 모의하는 다양한 연구 역시 활발히 진행 중이다(Harris *et al.*, 2007; Bitew and Gebremichael, 2011; Velpuri *et al.*, 2012). 선행 연구들에서는 수문 모형의 입력자료로 기존의 지상 관측 데이터를 대체하기 위해 TRMM, CMORPH, PERSIANN 등의 다양한 종류의 원격자료를 활용하였다. 이를 통해 원격자료의 적용 가능성을 평가하고, 지상 관측 시스템이 부재한 지역에 대한 보다 정확한 수문학적 분석 가능성을 확인하였다.

다만, 수위 예측을 위해 기계학습과 위성 기반의 원격자료를 함께 사용한 연구는 아직 부족한 상황이다. 특히, 국내의 경우 원격자료를 기계학습의 입력자료로 사용하여 하천에 대한 수위를 예측한 연구사례는 매우 부족한 상황이기 때문에 본 연구에서는 섬진강 유역의 대표 수위 지점 및 주변 하천을 대상으로 시간별 수위 변동 예측을 위한 원격자료와 기계학습의 적용성을 평가하고자 한다. 이를 위해 시계열 예측을 위한 대표적인 딥러닝 알고리즘의 하나인 Long-Short Term Memory (LSTM) 알고리즘을 적용하였다.

2. 연구방법 및 자료

2.1 연구 대상 지역 선정

본 연구에서는 섬진강 유역을 대상으로 수위 예측을 수행하였다(Fig. 1). 유역 내에 위치한 순천시(유적교), 곡성군(고달교), 그리고 구례군(구례교)를 대상 지점으로 선정하였다. 한반도의 남해안 중서부에 있는 섬진강 유역의 행정구역은 경상남도와 전라남-북도를 모두 포함하고 있다. 섬진강 유역은 위도 34.5 - 36.0도, 경도 126.5 - 128.0도 사이에 있고, 고도는 최저 5 m부터 최고 1,500 m이며 산지형 및 평활한 지형이 혼재되어 있다. 섬진강 유역의 연평균 강우량은 순천시(유적교) 지점 기준으로 약 1,264.5 mm이고, 계절적 특성에 따라 여름철인 6월부터 8월 사이에 연 강우량의 약 58% 이상의 강

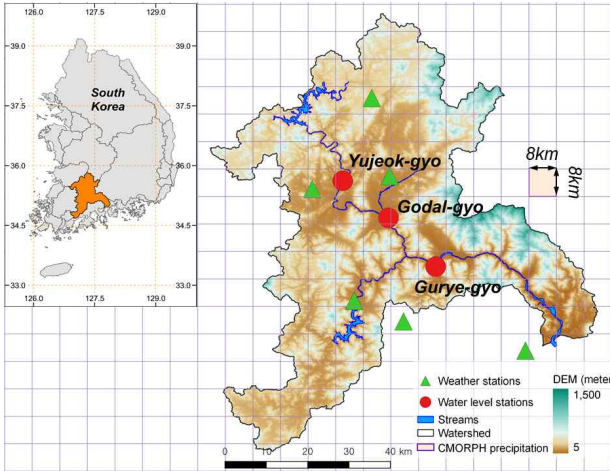


Fig. 1. Seomjin river watershed and stations. Each grid includes precipitation data provided from CMORPH product

우가 관측되는 기상 특성을 보인다.

2.2 연구 자료

2.2.1 수위 데이터

본 연구에서는 섬진강 유역에 위치한 순천시(유적교), 곡성군(고달교), 그리고 구례군(구례교)를 대상으로 수위 데이터를 수집하였다. 지점별 시(hourly) 수위 자료를 수집하기 위해 2000년 1월 1일 1:00부터 2021년 12월 31일 23:00로 총 22년간의 자료를 고려하였다. 지점별로 관측된 수위 자료는 전체 기간 중 95%이상 데이터가 확보된 것으로 확인되었다.

2.2.2 지상 강우 데이터

섬진강 유역 주변에 있는 기상관측소는 총 6개이며, 그중 데이터의 양이 충분하고, 각 수위관측소와 가장 근접한 3군데 기상관측소(순천; 174, 남원; 247, 순창; 254)의 시간별 강우량 자료를 활용하였다. 각 관측소에서 강우량 관측 기간이 다르므로 관측소별 가용한 자료를 활용하였다. 순천관측소의 경우 2011년 4월부터, 남원 관측소와 순창 관측소의 경우 2000년 1월, 2009년 1월부터 관측된 강우량 자료를 수집하였다.

2.2.3 CMORPH 강우 데이터

국립해양대기청(National Oceanic and Atmospheric Administration, NOAA)에서 제공하는 CMORPH은 전 지구를 대상으로 지상관측소보다 장기간의 강수 정보를 제공하고 있다(Kim and Han, 2021). CMORPH 데이터는 1998년 1월부터 현재까지 20년 이상의 자료를 보유하고 있으며, NOAA 자료 제공 포털(<https://www.ncei.noaa.gov/>)을 통해 자료의 수

집이 가능하다. CMORPH은 저궤도 위성(Low Earth Orbit-satellite)의 수동 마이크로웨이브(passive microwave)를 기반으로 강우 자료를 생성하고 있다. 또한 CMORPH은 보정과정을 거쳐 제공되고 있다(Xie *et al.*, 2017). CMORPH의 원자료는 NOAA Climate Prediction Center (CPC)의 지상 관측 강우 및 Global Precipitation Climatology Project (GPCP)의 강우 자료와의 확률밀도함수(probability density function) 비교를 통해 보정이 이루어진다. 이를 통해 보정된 CMORPH은 전 세계 지역을 대상으로 다양한 해상도의 강우 자료를 제공하고 있다. 가장 고해상도의 자료는 8km X 8km / 30 min 가 있으며, 0.25° × 0.25° / 3 hour, 그리고 0.25° × 0.25° / 24 hour 해상도의 자료도 이용이 가능하다. 본 연구에서는 고해상도의 자료(8 km × 8 km / 30 min)를 활용하였고, 각 수위관측소 및 주변 지역과 중첩되는 CMORPH 격자로부터 강수량 자료를 수집하였다. CMORPH으로부터 수집된 강우 데이터의 기간은 2000년 1월 1일 1:00부터 2021년 12월 31일 23:00로 총 22년에 해당한다.

2.3 Long-Short Term Memory algorithm

본 연구에서는 섬진강 유역 수위 지점별 예측을 위해 LSTM 알고리즘을 이용하였다. LSTM 알고리즘은 순환신경망(Recurrent Neural Network, RNN) 알고리즘과 함께 연속적인 시간 간격으로 배치된 데이터를 처리하는데 효과적인 알고리즘으로 알려져 있다. 하지만, RNN의 경우 짧은 시간 단위의 데이터를 처리하는 경우 유리하지만, 시간 단위가 길어질수록 장기 의존성 학습 능력이 현저하게 저하된다는 단점이 있다. 이러한 단점을 개선하기 위해 Hochreiter and Schmidhuber (1997)에 의해 제안된 알고리즘이 LSTM이다(Fig. 2).

LSTM 알고리즘은 3가지 게이트(gate)와 셀 상태(cell state)라는 구조를 기반으로 각 데이터의 상태를 메모리 cell에 저장 및 제거함과 동시에 각 게이트에서 데이터를 계산하는 과정에서 발생하는 불필요한 연산, 오차 등을 최소화시켜 장기 의존성을 처리할 수 있도록 고안되었다. 3가지 게이트는 입력(Input gate, i_t), 망각(Forget gate, f_t), 그리고 출력 게이트(Output gate, O_t)로 구성되어 있다. 입력 게이트에서는 새로운 정보를 가지고 있는 데이터를 cell에 저장 여부를 결정하는 역할을 하고, 망각 게이트는 과거 정보의 저장 및 제거를 결정하는 역할을 하고 있다. 그리고 출력 게이트에서는 최종적으로 어느 정보를 출력으로 내보낼지 결정하는 역할을 한다.

2.4 예측 성능 평가 지표

본 연구에서는 LSTM 알고리즘과 CMORPH 데이터를 활

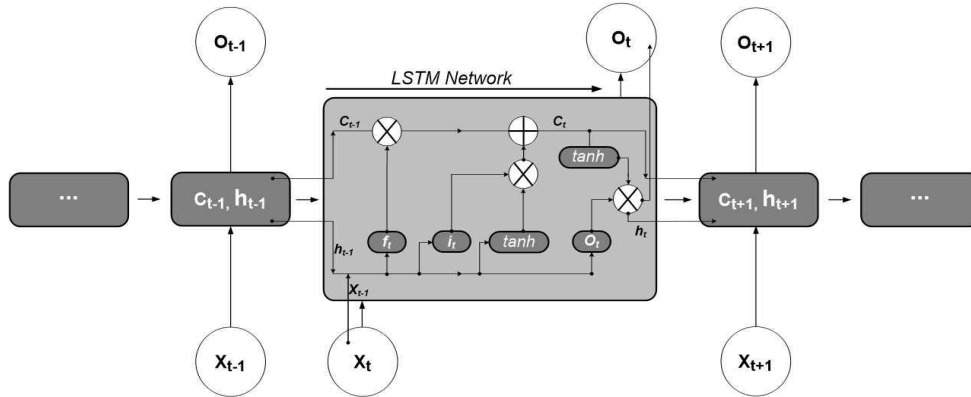


Fig. 2. Conceptual diagram of LSTM algorithm (Han *et al.*, 2021)

용하여 예측한 수위 결과의 신뢰성을 평가하기 위해 상관계수 (Correlation Coefficient, CC), 평균 제곱근 오차(Root Mean Square Error, RMSE), Nash-Sutcliffe efficiency (NSE)를 포함한 3가지 평가 지표를 사용하였다. 각 지표에 대한 수식은 아래 Eqs. (1)~(3)과 같다.

$$CC = \frac{\sum_{i=1}^n (S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (S_i - \bar{S})^2} \sqrt{\sum_{i=1}^n (O_i - \bar{O})^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - S_i)^2}{n}} \quad (2)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (S_i - O_i)^2}{\sum_{i=1}^n (S_i - \bar{S})^2} \quad (3)$$

여기서 S_i 와 O_i 는 예측값과 관측값, \bar{S} 와 \bar{O} 는 예측값과 관측값의 평균값, 그리고 n 은 자료의 개수를 의미한다.

3. 모형 적용 및 분석

3.1 수위 예측을 위한 LSTM 알고리즘 구축

본 연구에서는 수위 예측을 위한 LSTM 알고리즘의 입력 자료로 CMORPH 자료의 적용성을 평가하는 것이 목적이므로, 수위관측소 위치 및 주변에 해당하는 격자의 CMORPH 강우를 전처리하여 알고리즘의 입력자료로 활용하였다. 또

한 지상 강우를 입력자료로 활용한 수위 예측 결과와 비교를 통해 수위 예측을 위한 원격자료의 적용 가능성을 평가하였다.

Fig. 3은 3개 관측소 중 유적교 수위관측소에 대한 LSTM 알고리즘의 입력자료 구축 과정을 도식화하고 있다. 입력자료 구축 과정은 다음과 같다. 1) 유적교 수위관측소와 주변 지역에 해당하는 총 9개 격자의 CMORPH 기반의 시 강우 데이터와 수위 데이터를 2000년 1월 1일부터 2021년 12월 31일까지 수집하였다. 2) 9개의 격자에서 수집된 강우 자료와 수위 자료는 2시간부터 최대 12시간 이후의 수위 변동을 예측하기 위한 LSTM 알고리즘의 입력자료로 적용되었다. 3) 이와 동시에 수위관측소와 가장 가까운 곳에 위치한 지상 강우관측소 (순창 기상관측소)에서 관측된 시 강우 데이터를 함께 수집하였다. 다만, 지상관측소의 데이터 수집 일자가 2009년 이후부터 이루어진 관계로 지상 기반 강우 데이터는 2009년 1월 1일부터 2021년 12월 31일까지에 해당하는 데이터를 이용하였다. 4) 최종적으로, 지상 강우를 입력자료로 활용한 또 하나의 LSTM 알고리즘을 함께 구축하여 CMORPH 강우를 입력자료로 활용한 알고리즘의 예측 성능을 비교 분석 하였다.

3.2 LSTM 알고리즘의 매개변수 추정

본 연구에서는 LSTM 알고리즘 입력자료의 다른 lead-time을 고려하여 수위를 예측하기 위해 2, 4, 8, 12시간으로 lead-time으로 설정하고 수위를 예측하였다. 알고리즘의 레이어 (Layer)는 총 3개로 구성하였으며, 시행착오법을 통해 각 레이어의 유닛을 결정하였다. Unit number과 batch size의 수를 32-512사이의 다양한 값을 적용하여 매개변수에 따른 예측 성능을 확인하였다. 그 결과, 각 레이어의 unit number가 128, 64, 1의 조합에서 비교적 뛰어난 성능(CC = 0.98, NSE = 0.97)을 나타내는 것을 확인하였다. 또한, 알고리즘의 활성화 함수 (activation function)과 최적화 기법(optimizer)은 ‘Rectified

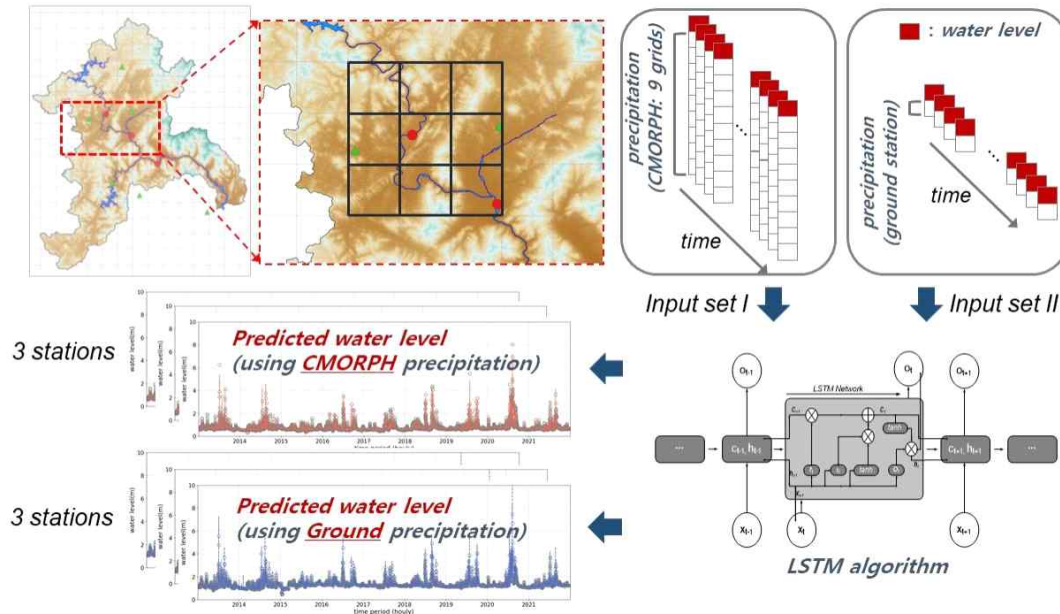


Fig. 3. Flowchart of this study

Linear Unit (ReLU)와 ‘Adaptive Moment Estimation (Adam)’를 적용하였다. ‘ReLU’는 학습 과정에서 데이터 간의 비선형적 관계를 효과적으로 고려할 수 있다는 장점이 있으며, LSTM 알고리즘의 대표적인 활성화 함수인 ‘tanh’와 함께 가장 널리 쓰이고 있는 활성화 함수 중 하나이다(Godin *et al.*, 2018; Szandala, 2021). 본 연구의 대상 지점 중 하나인 구례교에서 수위 예측을 위해 ‘ReLU’와 ‘tanh’를 비교한 결과 ‘ReLU’를 활성화 함수로 적용한 알고리즘(CC = 0.99, NSE = 0.98, RMSE = 0.09 m) ‘tanh’를 사용한 경우(CC = 0.97, NSE = 0.96, RMSE = 0.11 m)보다 비교적 높은 성능을 나타내었기에 본 연구에서는 ‘ReLU’를 활성화 함수로 활용하였다. ‘Adam’ 역시 딥러닝 알고리즘에서 가장 많이 사용되고 있는 최적화 기법 중 하나이다. 모형의 학습(training)을 위해서 2000 - 2015년까지의 자료(16년)를 이용하였고, 학습된 알고리즘의 평가(test)는 2016 - 2021년까지의 자료(6년)를 활용하였다. 또한 본 연구에서는 LSTM 알고리즘의 과적합 방지를 위해 dropout rate를 0.1로 적용하였다.

3.3 CMORPH 강우와 LSTM 알고리즘을 이용한 하천 수위 변동성 예측 결과

수위 변동성 예측을 위해 LSTM 알고리즘에 CMORPH 강우와 수위 자료를 입력자료로 활용하였고, 입력자료의 lead-time에 따른 수위 예측 성능을 평가하였다. 예측 성능에 대한 정량적인 평가를 위해 3가지 평가 지표를 활용하였으며, Tables 1~3은 3군데 수위관측소에 대한 lead-time별 수위 예

Table 1. Predictive performance for different lead-time (Yujeok-gyo station)

Lead-time (hr)	CC	RMSE (m)	NSE
2	0.981	0.076	0.962
4	0.979	0.080	0.958
8	0.977	0.086	0.951
12	0.980	0.081	0.957

Table 2. Predictive performance for different lead-time (Godal-gyo station)

Lead-time (hr)	CC	RMSE (m)	NSE
2	0.976	0.087	0.951
4	0.979	0.080	0.959
8	0.980	0.078	0.961
12	0.972	0.093	0.945

Table 3. Predictive performance for different lead-time (Gurye-gyo station)

Lead-time (hr)	CC	RMSE (m)	NSE
2	0.988	0.134	0.962
4	0.988	0.152	0.950
8	0.984	0.138	0.959
12	0.981	0.150	0.951

측 결과를 나타내고 있다. 각 관측소와 lead-time마다 근소하게 차이는 있었지만, 전반적으로 매우 우수한 예측 성능을 확

인하였다. 관측 데이터 대비 예측 수위의 시간적 경향성을 나타내는 CC는 0.98 이상으로 나타났고, NSE의 경우 0.95 이상의 결과를 보였다. Lead-time에 대한 평가는 유적교와 구례교에서는 2시간인 경우 CC와 NSE가 0.98, 0.96이상, 그리고 RMSE는 0.076 m, 0.134 m로 가장 좋은 성능을 나타냈고, 고달교의 경우 8시간의 lead-time에서 예측 성능이 가장 높게 나타났다. 다만, 2시간 4시간과의 예측 성능 차이는 크지 않은 것으로 확인되었다. 또한, 세 지점 모두 공통적으로 12시간의 lead-time에서는 비교적 낮은 예측 성능이 확인되었다.

Fig. 4는 CMORPH 강우량, 예측된 시간별 수위와 관측 데이터와의 시계열 그래프와 산점도를 나타내고 있다. 각 그래프는 각 지점에서 가장 좋은 예측 성능을 나타내는 lead-time에 대한 예측 결과를 보여주고 있다. Tables 1~3에 제시된 평가 지표 중 CC가 0.9 이상으로 나타났듯이 예측된 수위는 실제 수위의 변동성을 매우 잘 반영하는 것을 알 수 있었다. 다만,

침투 수위를 예측하는 성능은 변동성 예측 성능에 비해 다소 아쉬운 결과를 나타냈다. 3개 지점에서 침투 수위의 예측값과 관측값 사이에 평균 5 - 8%의 오차가 발생하였다. 이는 딥러닝 알고리즘 학습 과정에서 침투 수위에 대한 데이터의 부족으로 인해 발생하는 것으로 판단된다. 이러한 결과는 딥러닝 알고리즘을 활용한 수위 및 유량 예측을 수행한 선행 연구에서도 비슷한 결과를 확인할 수 있다(Xie *et al.*, 2021; Kim *et al.*, 2022).

3.4 지상 강우와 LSTM 알고리즘을 이용한 하천 수위 변동성 예측 결과

3.4.1 지상 강우와 CMORPH 강우 자료 비교

수위 변동성 예측을 위한 LSTM 알고리즘의 입력자료로 CMORPH 기반 격자형 강우 데이터의 적용성을 평가하기 위해 본 연구에서는 지상 강우 데이터 및 지상 강우를 입력자료로 활용한 하천 수위 변동성 예측 결과를 비교 자료로 이용하

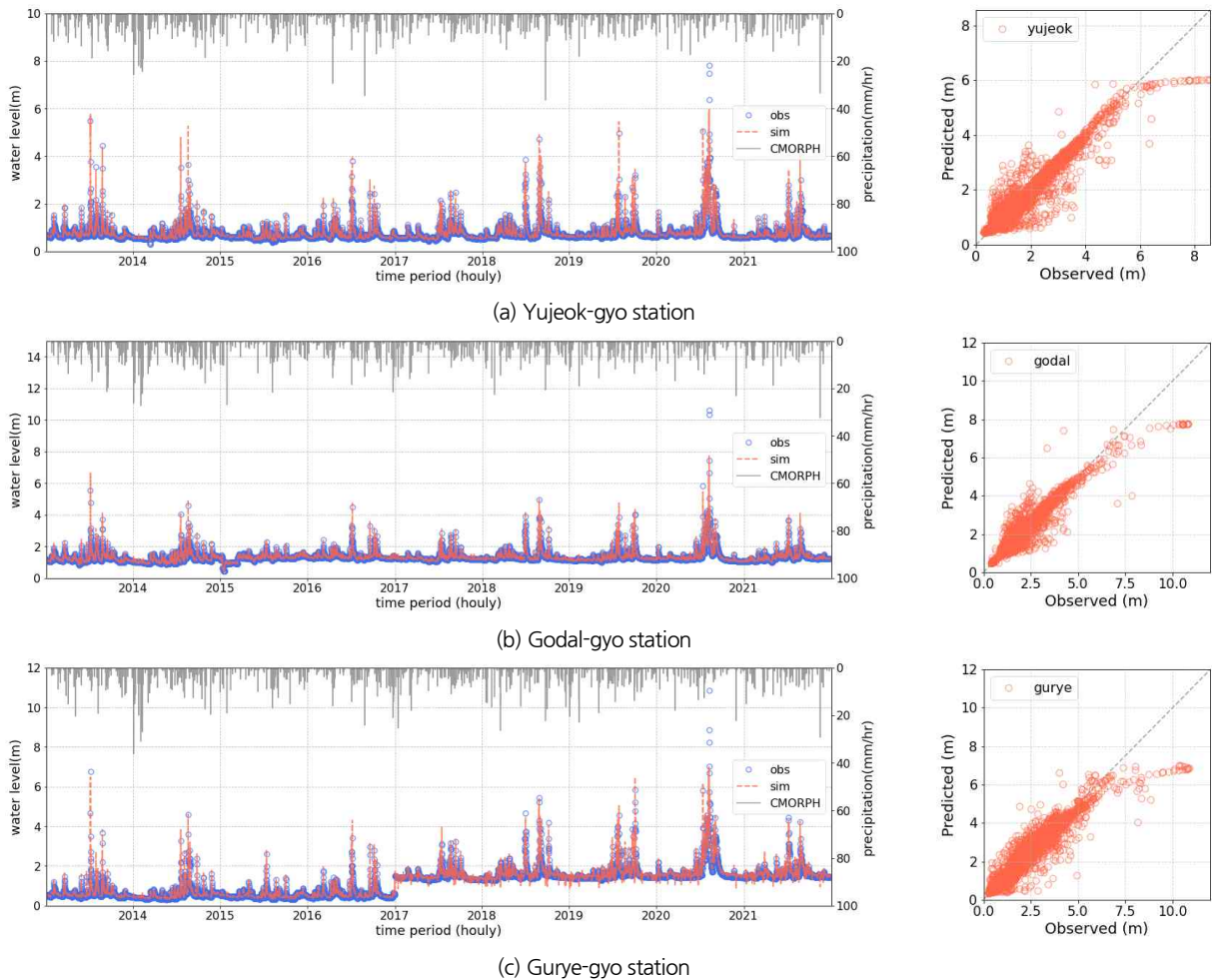


Fig. 4. Time-series and scatter plots of predicted and observed water level (2hr of lead time at Yujeok-gyo and Gurye-gyo stations and 8hr of lead time at Godal-gyo station)

였다. 본 절에서는 CMORPH 강우와 지상 강우의 정량적 평가를 수행하였다. 관련 연구로 Kim and Han (2021)는 전국 48개 강우 관측 지점에 대해 CMORPH 강우의 정량적 평가를 수행한 바 있다. 그 결과 시 단위 CMORPH 강우의 평균 CC와 RMSE는 0.5, 1.2 mm/hr로 나타났다. 본 연구에서는 연구 대상 지역의 3개 강우 관측 지점에 대해 CMORPH 강우와의 비교를 수행하였다. 그 결과 3개 지점에서 RMSE는 1.5 mm/hr로 나타났고, CC는 다소 낮은 0.1로 나타났다. 이는 CMORPH 강우가

지상 관측 강우량은 유사한 반면에 시간적 경향성은 다소 차이가 있는 것을 알 수 있었다.

3.4.2 지상 강우와 CMORPH 강우를 이용한 하천 수위 변동성 예측 결과 비교

LSTM 알고리즘의 입력변수로 CMORPH 강우를 사용하여 예측한 수위 결과와 지상 강우를 사용하여 예측한 결과를 비교하였다(Fig. 5). 지상 강우의 경우 데이터의 확보 기간이

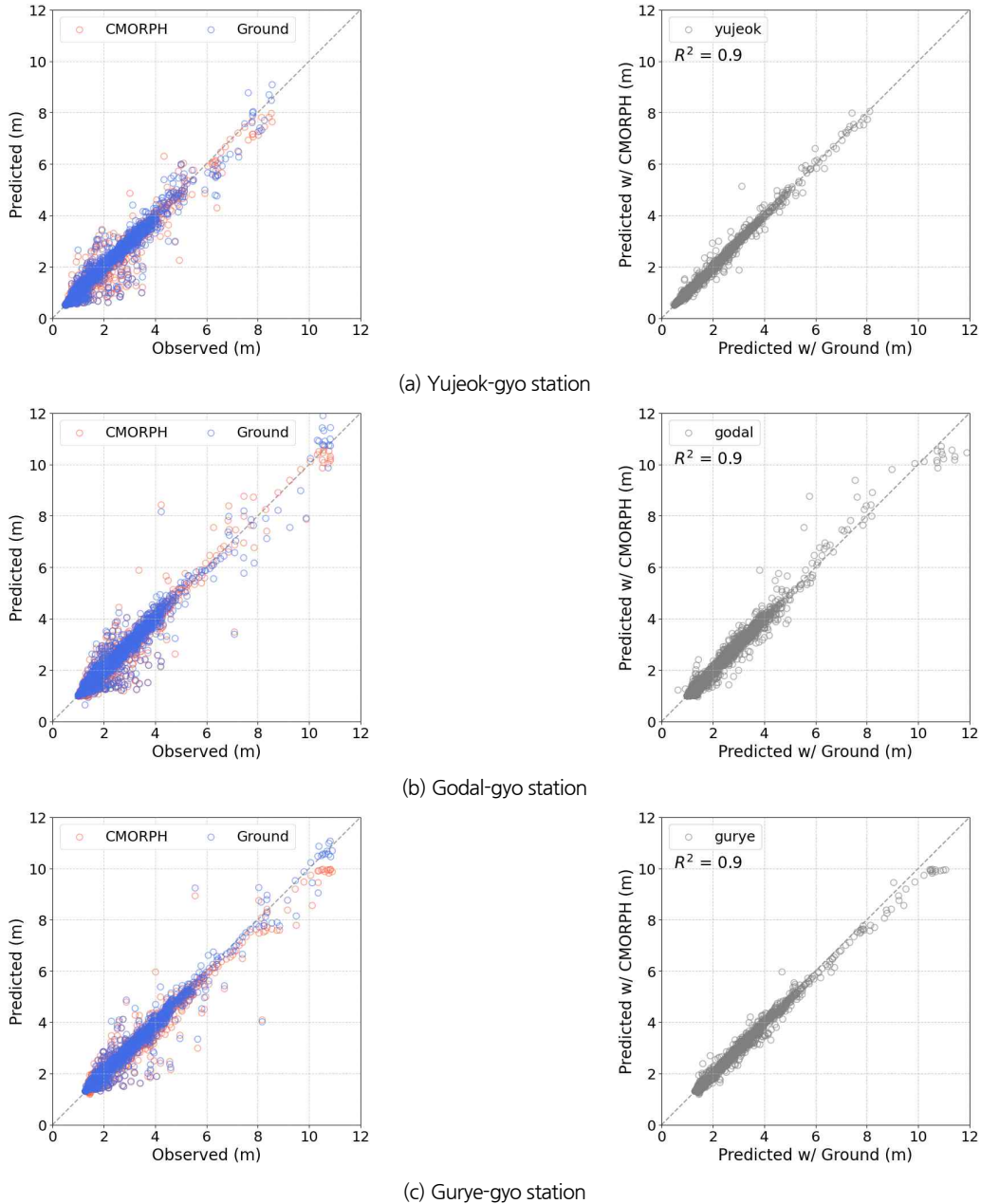


Fig. 5. Scatter plots of the predicted water levels using CMORPH (Red), ground (Blue) precipitation and reference data (Left panel). Comparison of water level predicted from the CMORPH and ground precipitation (Right panel)

지점별로 다르므로 지상 강우 데이터 기간(순천관측소: 2011/4-2021/12, 남원 관측소: 2000/1-2021/12, 순창 관측소: 2000/1-2021/12)과 동일한 CMORPH 강우를 적용하였다.

CMORPH 강우를 입력변수로 사용하여 예측한 결과와 지상 강우를 입력변수로 활용한 예측 결과가 모든 관측 지점에서 매우 유사한 것을 확인할 수 있다. 또한 두 결과 데이터의 R^2 값 역시 0.9 이상으로 나타났다. 즉, 하천 수위 예측을 위한 딥러닝 알고리즘의 입력자료로 지상 강우 대신 CMORPH 기반 격자형 강우 데이터를 활용할 수 있음을 알 수 있다. 다만, CMORPH 강우를 이용한 예측 수위의 침투 깊이가 지상 강우를 이용한 예측 수위의 침투 깊보다 다소 과소 추정되는 결과를 확인할 수 있었다.

4. 요약 및 결론

본 연구에서는 CMORPH 기반 격자형 강우와 지점별 수위 자료를 딥러닝 알고리즘의 입력자료로 활용하여 섬진강 유역의 대표 수위관측소(유적교, 고달교, 구례교)의 수위 변동성을 예측하였다. 예측 모형으로는 LSTM 알고리즘을 활용하였으며, 2, 4, 8, 12시간의 lead-time을 고려하여 수위 예측 성능을 평가하였다.

CMORPH 강우와 LSTM 알고리즘을 함께 사용한 예측 결과는 매우 우수한 것으로 확인되었다. 각 지점의 평균 CC가 0.98, RMSE는 0.07 m, 그리고 NSE는 0.97로 나타났다. 또한, 지점별로 우수한 예측 성능을 보이는 lead-time이 다르게 나타났다. 유적교의 경우 4시간 lead-time, 고달교와 구례교의 경우 12시간 lead-time이 가장 좋은 예측 정확도를 나타냈다.

CMORPH 기반 강우와 지상 강우를 사용한 경우의 예측 결과를 비교 분석하였다. CMORPH 강우를 LSTM 알고리즘의 입력변수로 활용한 예측 결과와 지상 강우를 입력변수로 활용한 경우의 예측 결과가 상당히 유사한 것을 확인할 수 있었다. 지상 강우를 이용한 예측 결과가 비교적 좋은 성능을 나타냈지만, CMORPH 강우를 고려한 알고리즘 역시 매우 우수한 예측 정확도를 나타냈다.

본 연구는 하천 수위 변동성 예측을 위한 딥러닝 알고리즘의 입력변수로 CMORPH 기반 강우 데이터의 적용성을 평가하였다. 지상관측소의 경우 정확한 강우 데이터를 제공한다는 분명한 장점이 있다. 하지만, 관측소의 설치 위치에 따라 하천 수위 변동에 영향을 미치는 강우 특성을 정확히 반영하기 어렵고, 강우 데이터의 수집 기간이 지점별로 상이하다는 단점이 있다. 특히, 1980-1990년대 강우 데이터에 대한 정보가 부족하다는 한계가 있다. 이러한 한계를 보완하기 위해

CMORPH와 같은 원격자료가 지상 강우 자료의 부분적 대체 방안이 될 수 있을 것으로 생각된다. 향후 수위 예측뿐만 아니라 강우-유출 모의, 홍수예보와 같은 수문 분야의 다양한 연구에서도 지상 강우 대신 원격 강우 자료를 충분히 활용할 수 있을 것으로 기대하는 바이다.

감사의 글

이 논문은 행정안전부 기후변화대응 AI 기반 풍수해 위험도 예측기술개발 사업의 지원을 받아 수행된 연구임(2022-MOIS61-002).

Conflicts of Interest

The authors declare no conflict of interest.

References

- Arora, V.K., and Boer, G.J. (2001). "Effects of simulated climate change on the hydrology of major river basins." *Journal of Geophysical Research: Atmospheres*, Vol. 106, No. D4, pp. 3335-3348.
- 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." *In Machine Learning and Knowledge Discovery in Databases: European Conference*, Springer, Skopje, Macedonia, Vol. 10536, pp. 317-329.
- Baek, S.S., Pyo, J., and Chun, J.A. (2020). "Prediction of water level and water quality using a CNN-LSTM combined deep learning approach." *Water*, Vol. 12, No. 12, 3399.
- Bitew, M.M., and Gebremichael, M. (2011). "Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model." *Water Resources Research*, Vol. 47, No. 6. doi: 10.1029/2010WR009917.
- Godin, F., Degraeve, J., Dambre, J., De Neve, W. (2018). "Dual rectified linear units (DReLU): A replacement for tanh activation functions in quasi-recurrent neural networks." *Pattern Recognition Letters*, Vol. 116, pp. 8-14.
- Han, H., Abitew, T.A., Park, S., Green, C.H., and Jeong, J. (2023). "Spatiotemporal evaluation of satellite-based precipitation products in the Colorado river basin." *Journal of Hydro-meteorology*, Vol. 24, No. 10, pp. 1739-1754.
- Han, H., and Morrison, R.R. (2022). "Improved runoff forecasting performance through error predictions using a deep-learning approach." *Journal of Hydrology*, Vol. 608, 127653.

- Han, H., Choi, C., Kim, J., Morrison, R.R., Jung, J., and Kim, H.S. (2021). "Multiple-depth soil moisture estimates using artificial neural network and long short-term memory models." *Water*, Vol. 13, No. 18, 2584.
- Harris, A., Rahman, S., Hossain, F., Yarborough, L., Bagtzoglou, A. C., and Eason, G. (2007). "Satellite-based flood modeling using TRMM-based rainfall products." *Sensors*, Vol. 7, No. 12, pp. 3416-3427.
- Hochreiter, S., and Schmidhuber, J. (1997). "Long short-term memory." *Neural Computation*, Vol. 9, No. 8, pp. 1735-1780.
- Hu, C., Wu, Q., Li, H., Jian, S., Li, N., and Lou, Z. (2018). "Deep learning with a long short-term memory networks approach for rainfall-runoff simulation." *Water*, Vol. 10, No. 11, 1543. doi: 10.3390/w10111543.
- Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P., and Stocker, E.F. (2007). "The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales." *Journal of Hydrometeorology*, Vol. 8, No. 1, pp. 38-55.
- Joyce, R.J., Janowiak, J.E., Arkin, P.A., and Xie, P. (2004). "CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution." *Journal of Hydrometeorology*, Vol. 5, No. 3, pp. 487-503.
- Kay, A.L., Rudd, A.C., Fry, M., Nash, G., and Allen, S. (2021). "Climate change impacts on peak river flows: Combining national-scale hydrological modelling and probabilistic projections." *Climate Risk Management*, Vol. 31, 100263.
- Kim, D., Han, H., Wang, W. and Kim, H. S. (2022). "Improvement of deep learning models for river water level prediction using complex network method." *Water*, Vol. 14, 466.
- Kim, J., and Han, H. (2021). "Evaluation of the CMORPH high-resolution precipitation product for hydrological applications over South Korea." *Atmospheric Research*, Vol. 258, 105650.
- Le, M.H., Lakshmi, V., Bolten, J., and Du Bui, D. (2020). "Adequacy of satellite-derived precipitation estimate for hydrological modeling in Vietnam basins." *Journal of Hydrology*, Vol. 586, 124820.
- Lee, M., Kim, J., Yoo, Y., Kim, H.S., Kim, S.E., and Kim, S. (2021). "Water level prediction in Taehwa River basin using deep learning model based on DNN and LSTM." *Journal of Korea Water Resources Association*, Vol. 54, No. spc1, pp. 1061-1069.
- Li, G., Liu, Z., Zhang, J., Han, H., and Shu, Z. (2024). "Bayesian model averaging by combining deep learning models to improve lake water level prediction." *Science of The Total Environment*, Vol. 906, 167718.
- Li, W., Gao, X., Hao, Z. and Sun, R. (2022). "Using deep learning for precipitation forecasting based on spatio-temporal information: A case study." *Climate Dynamics*, Vol. 58, pp. 443-457.
- Park, K., Jung, Y., Seong, Y., and Lee, S. (2022). "Development of deep learning models to improve the accuracy of water levels time series prediction through multivariate hydrological data." *Water*, Vol. 14, No. 3, 469.
- Schreider, S.Y., Smith, D.I., and Jakeman, A.J. (2000). "Climate change impacts on urban flooding." *Climatic Change*, Vol. 47, pp. 91-115.
- Sorooshian, S., Hsu, K.L., Gao, X., Gupta, H.V., Imam, B., and Braithwaite, D. (2000). "Evaluation of PERSIANN system satellite-based estimates of tropical rainfall." *Bulletin of the American Meteorological Society*, Vol. 81, No. 9, pp. 2035-2046.
- Szandała, T. (2021). *Review and comparison of commonly used activation functions for deep neural networks*. Bio-inspired Neurocomputing, Springer, Singapore, pp. 203-224.
- Tarekegn, N., Abate, B., Muluneh, A., and Dile, Y. (2022). "Modeling the impact of climate change on the hydrology of Andasa watershed." *Modeling Earth Systems and Environment*, Vol. 8, No. 1, pp. 103-119.
- Tessema, N., Kebede, A., and Yadeta, D. (2021). "Modelling the effects of climate change on streamflow using climate and hydrological models: the case of the Kesem sub-basin of the Awash River basin, Ethiopia." *International Journal of River Basin Management*, Vol. 19, No. 4, pp. 469-480.
- Tobin, K.J., and Bennett, M.E. (2010). "Adjusting satellite precipitation data to facilitate hydrologic modeling." *Journal of Hydrometeorology*, Vol. 11, No. 4, pp. 966-978.
- Velpuri, N.M., Senay, G.B., and Asante, K.O. (2012). "A multi-source satellite data approach for modelling Lake Turkana water level: calibration and validation using satellite altimetry data." *Hydrology and Earth System Sciences*, Vol. 16, No. 1, pp. 1-18.
- Wang, Q., and Wang, S. (2020). "Machine learning-based water level prediction in Lake Erie." *Water*, Vol. 12, No. 10, 2654.
- Xiang, Z., Yan, J., and Demir, I. (2020). "A rainfall-runoff model with LSTM-based sequence-to-sequence learning." *Water Resources Research*, Vol. 56, No. 1, e2019WR025326.
- Xie, P., Joyce, R., Wu, S., Yoo, S. H., Yarosh, Y., Sun, F. Lin, R. (2017). "Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998." *Journal of Hydrometeorology*, Vol. 18, pp. 1617-1641.
- Xie, Z., Liu, Q. and Cao, Y. (2021). "Hybrid deep learning modeling for water level prediction in Yangtze River." *Intelligent Automation & Soft Computing*, Vol. 28, pp. 153-166.