



Prediction of water level in a tidal river using a deep-learning based LSTM model

Jung, Sungho^a · Cho, Hyoseob^b · Kim, Jeongyup^b · Lee, Giha^{a*}

^aDepartment of Disaster Prevention and Environmental Engineering, Kyungpook National University

^bWater Resources information center of Han River Flood Control Office, Ministry of Environment

Paper number: 18-085

Received: 11 September 2018; Revised: 23 October 2018; Accepted: 23 October 2018

Abstract

Discharge or water level predictions at tidally affected river reaches are currently still a great challenge in hydrological practices. This research aims to predict water level of the tide dominated site, Jamsu bridge in the Han River downstream. Physics-based hydrodynamic approaches are sometimes not applicable for water level prediction in such a tidal river due to uncertainty sources like rainfall forecasting data. In this study, TensorFlow deep learning framework was used to build a deep neural network based LSTM model and its applications. The LSTM model was trained based on 3 data sets having 10-min temporal resolution: Paldang dam release, Jamsu bridge water level, predicted tidal level for 6 years (2011~2016) and then predict the water level time series given the six lead times: 1, 3, 6, 9, 12, 24 hours. The optimal hyper-parameters of LSTM model were set up as follows: 6 hidden layers number, 0.01 learning rate, 3000 iterations. In addition, we changed the key parameter of LSTM model, sequence length, ranging from 1 to 6 hours to test its affect to prediction results. The LSTM model with the 1 hr sequence length led to the best performing prediction results for the all cases. In particular, it resulted in very accurate prediction: RMSE (0.065 cm) and NSE (0.99) for the 1 hr lead time prediction case. However, as the lead time became longer, the RMSE increased from 0.08 m (1 hr lead time) to 0.28 m (24 hrs lead time) and the NSE decreased from 0.99 (1 hr lead time) to 0.74 (24 hrs lead time), respectively.

Keywords: Tidal river, Deep learning, lead time, Jamsu bridge, LSTM model

딥러닝 기반 LSTM 모형을 이용한 감조하천 수위 예측

정성호^a · 조효섭^b · 김정엽^b · 이기하^{a*}

^a경북대학교 과학기술대학 건설방재공학과, ^b환경부 한강홍수통제소 수자원정보센터

요 지

본 연구는 물리적 수리-수문모형의 적용이 제한적인 감조하천에서의 수위예측을 목적으로 하고 있으며, 이를 위해 한강 잠수교를 대상으로 딥러닝 오픈소스 소프트웨어 라이브러리인 TensorFlow를 활용하여 LSTM 모형을 구성하고 2011년부터 2017년까지의 10분 단위의 잠수교 수위, 팔당 댐 방류량과 한강하구 강화대교지점의 예측조위 자료를 이용하여 모형학습(2011~2016) 및 수위예측(2017)을 수행하였다. 모형 매개변수는 민감도 분석을 통해 은닉층의 개수는 6개, 학습속도는 0.01, 학습횟수는 3000번로 결정하였으며, 모형 학습 시 학습정보의 시간적 양을 결정하는 중요한 매개변수인 시퀀스길이는 1시간, 3시간, 6시간으로 변화시키며 모의하였다. 최종적으로 선행시간에 따른 모의 예측능력을 평가하기 위해 LSTM 모형의 예측 선행시간을 6개(1 ~ 24시간)로 구분하여 실측수위와 예측수위와의 비교·분석을 수행한 결과, LSTM 모형의 최적의 성능을 내는 결과는 시퀀스길이를 1시간으로 하였을 때로 분석되었으며, 특히 선행시간 1시간에 대한 예측정확도는 RMSE는 0.065 m, NSE는 0.99로 실측수위에 매우 근접한 예측 결과를 나타내었다. 또한 시퀀스길이에 상관없이 선행시간이 길어질수록 모형의 예측 정확도는 2017년 전기간에 걸쳐 평균적으로 RMSE 0.08 m에서 0.28 m로 오차가 증가하였으며, NSE는 0.99에서 0.74로 감소하였다.

핵심용어: 감조하천, 딥러닝, 선행시간, 잠수교, LSTM 모형

*Corresponding Author. Tel: +82-54-530-1259
E-mail: leegiha@knu.ac.kr (G. Lee)

1. 서론

해수와 만나 기수역을 형성하는 하천 하구는 수자원 및 환경생태 관리의 목적으로 하굿둑을 건설하여 운영하고 있으나 한강은 4대강 중 유일하게 하굿둑이 없어 서해의 조위에 따라 한강분류의 중·하류부의 수위가 변동하는 대표적인 감조하천이다. Park and Baek (2017)은 이러한 감조하천의 경우 조위의 영향으로 인해 하천수위의 낙폭이 크고, 하상변동이 심하기 때문에 홍수위 분석 및 예측에 있어 일반하천에 비해 보다 세심한 주의가 필요하다고 지적한 바 있다.

현재 한강홍수통제소에서는 한강 14개 및 안성천 2개 총 16개 주요지점을 대상으로 주위보 수위 및 경보 수위를 사전에 결정·고시하고, 이를 기준으로 홍수특보를 발령하고 있다. 이상의 홍수예보지점 이외에도 잠수교의 경우, 한강분류 서울시 구간에 위치하고 있어 통행량이 많고, 잦은 침수로 인하여 보행자 및 통행제한이 발생하는 지점으로 잠수교의 수위가 완전잠수(6.5 m) 뿐만 아니라 차량제한(6.2 m), 보행제한(5.5 m) 등에 도달할 것으로 예상되는 경우 수위예측정보를 유관기관에 실시간으로 제공하여 차량 및 보행자 통행제한을 시행함으로써 인명 및 재산피해가 발생하지 않도록 조치하고 있다.

상기 주요지점의 홍수위 예측시스템은 수공구조물에 의한 배수영향이나 조위영향이 있는 지점의 경우는 수리학적 홍수예측, 이외의 경우는 수문학적 홍수예측을 수행하고 있다 (MOLIT, 2017). Shin and Yoon (2005)은 서해 조위변화가 한강분류에서는 잠실수중보(잠실대교)까지 영향을 미칠 수 있음을 분석한 바 있으며, 이는 잠실대교 보다 하류에 위치한 홍수특보 지점인 한강대교와 관심지점인 잠수교의 경우, 홍수위 예측을 위한 수리모형 기반 부정류 모의 시 반드시 조위의 영향을 고려해야 함을 의미한다.

감조하천에서의 수리·수문 특성 변화를 위한 국내연구는 대부분 물리적 기반의 수리모형이 활용되었으며, 주요연구는 다음과 같다. Lee *et al.* (2005)은 1차원 홍수추적모형 DWOPER을 적용하여 하상변화와 지류 및 본류의 유입유량 규모와 서해 조위 등의 변화요인들에 따른 한강 수위거동을 분석하였으며, Seo *et al.* (2008)은 한강 신곡수중보에서 월곡면 유도지점까지의 구간을 대상으로 RMA2 및 RAM4모형을 사용하여 하천흐름 및 오염물질의 혼합거동을 분석하였다. 또한 Lee and Lee (2010)은 1차원 부정류모형인 FLDWAV를 이용하여 조석과 팔당댐 방류유형에 따른 수위상승도달시간의 관계를 2차 다항식으로 유도하여 나타내어 분석하였으며, Song *et al.* (2014)은 유량규모에 따른 배수(backwater) 흐름을 해석하기 위해 천수방정식을 SU/PG 기법으로 이산화한 수치모형을 개발하고 한강분류구간(영동대교~행주대교)

을 대상으로 최대 배수거리를 산정한 바 있다. 수리모형 적용에 관한 연구 이외에도 Lee *et al.* (2007)은 감조하천에서 저·평수 위 수위-유량관계곡선식 개발의 한계를 극복하기 위해 조위를 고려한 수위-수면차-유량관계를 통해 지수형 곡선식을 개발하여 자동유량측정 자료와의 비교·검토를 수행한 바 있다.

전술한 연구동향에서 알 수 있듯이 일반적으로 하천 홍수위 예측을 위해서는 기상학적, 지형학적, 지질학적 매개변수 등을 이용한 강우-유출모형과 하도에서의 부정류모의를 위한 수리모형을 결합한 모의기법이 요구되며, 하천 상·하류의 수위 또는 유량자료를 모의 경계조건으로 활용한다. 그러나 Lee *et al.* (2007)은 실제 감조하천에서는 배수효과로 인해 수위와 유량간의 뚜렷한 상관관계가 나타나지 않아 정확한 유량 정보를 획득하기 어렵다고 지적한 바 있으며, 이는 매개변수 보정 및 모형 검증에 어려움이 있음을 의미한다. 실제로 수리·수문모형을 이용하여 감조하천의 홍수위 모의에 관한 모든 물리적 과정을 반영하는 것은 한계가 있으며, 예보를 위한 예측강우의 시공간적 불확실성으로 인해 예측결과에 불확실성을 초래할 수 있고 방대한 컴퓨팅 자원(computing resources)와 수치계산시간(numerical computation time) 등의 제약으로 후속조치를 위한 선행시간(lead time) 확보가 필수인 홍수예보에서 그 활용에 제한적일 수 있다(Oh *et al.*, 2014).

반면에 인공신경망(artificial neural network, ANN) 모형 등과 같은 입·출력자료의 인과(상관)관계를 기반으로 학습·예측하는 black-box 모형의 경우, 유역 및 하도에서 발생하는 물리적인 수리·수문학적 특성을 반영할 수는 없지만, 양질의 대용량 자료가 획득될 경우 특정 수문시계열 예측에서 우수한 결과를 도출하는 것으로 알려져 있으며 이미 다양한 수자원 분야에서 활용되고 있다(Maier and Dandy, 2000). Chiang *et al.* (2010), Chen *et al.* (2012), Kim and Tachikawa. (2017)은 일반하천에서 ANN 기반의 모형을 구성하여 수위예측을 수행하고 그 적용성을 검토하였으며, 딥러닝기반의 LSTM(Long and Short Term Memory) 모형의 경우, Jung *et al.* (2018)은 금강상류에 위치한 상류(2개소) 및 지류(1개소) 수위만을 이용하여 옥천수위관측지점의 수위예측을 수행하였으며, Lee *et al.* (2018)은 물리적 수문모형인 SWAT과 LSTM 모형을 이용하여 메콩강 하류 Kratie 관측소의 장기 유출해석을 수행하고, 두 모형의 정확도를 비교·검토한 바 있다. Oh *et al.* (2014)는 현행 홍수예보시스템의 보조적 수단으로 Web기반의 인공신경망 모형을 구축하고, 국가표준수문 DB를 활용하여 영상강 홍수예보지점인 선암과 남평지점에 대해 홍수위를 예측한 바 있다. 이와 달리 조위영향을 받는 감조하천에서의 수위예측에 관한 연구는 제한적으로 수행되었으며 특히 다양한 선행예보시간에 따른 예측 정확도 분석에 관한 연구는 미

비한 것으로 조사되었다. 다만, 본 연구목적과 유사한 주요연구로써 Supharatid (2003)은 다층신경망(multi-layer neural network, MNN)을 사용하여 태국 Chao Phraya 강 하구에서의 조위예측을 수행하고, 해당 조위자료를 이용하여 감소하천 내 수위-유량 관계곡선식을 생성하였으며, Hidayat et al. (2014)은 인도네시아 Mahkam 강을 대상으로 ANN 모형을 구축하고 지류 수위 및 조위 정보를 이용하여 실시간 수위예측 시스템을 구축하여 정확도를 검증한 바 있다.

본 연구에서는 심층신경망(deep neural network) 구조를 갖는 딥러닝 기반의 순환신경망(recurrent neural network, RNN) 모형을 구축하여 한강 잠수교의 수위예측을 수행하고 선행시간에 따른 예측 정확도를 분석한다. 이를 위해 RNN은 구글에서 개발하고 배포한 딥러닝 오픈소스 소프트웨어 라이브러리 TensorFlow (<https://www.TensorFlow.org>)의 LSTM 알고리즘을 적용하며, LSTM 기반의 신경망 모형의 학습과 검증에 위해 총 7년 기간(2011~2017년)의 10분 단위 잠수교 수위자료, 팔당댐 방류량과 강하대교의 예측 조위자료를 활용한다. 2011년부터 2016년까지의 자료는 LSTM 모형의 학습(training), 2017년 자료는 LSTM 모형의 검증(validation)에 활용하고, 민감도 분석(sensitivity analysis)을 통해 LSTM 모형의 최적 매개변수(optimal hyper-parameters)를 추정하여 선행시간 1시간, 3시간, 6시간, 9시간, 12시간, 24시간(6 Cases)에 대한 잠수교 수위를 예측하고 그 정확도를 분석한다.

2. 적용모형

2.1 딥러닝

심층 학습을 뜻하는 딥러닝(deep learning)은 인공지능 분야에서 가장 많은 연구가 이루어지고 있으며 자연언어처리, 음성 인식 및 이미지 인식 등의 연구에서 높은 정확도를 보이는 등 여러 분야에서 다양하게 쓰이고 있다. 딥러닝의 기초가 되는 단일 인공신경망은 뇌세포의 뉴런의 동작을 흡사하게 구현하여 입력값(input data)에 대한 가중치(weight) 신호를 주어 정보를 전달하도록 만든 구조이다. 일반적으로 인공신경망의 구조는 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)의 세 개의 층으로 구성되어 있으며, 입력층에서 은닉층을 통하여 출력층까지 각 층(layer)사이의 가중치를 순차적으로 업데이트하며 예측값과 실제값과의 오차를 최소화하면서 학습을 하는 신경망을 피드포워드신경망(feed forward neural network FFNN)이라고 한다. 이러한 구조에서 2개 이상의 은닉층을 가지면 다층신경망, 그 이상일 경우를 심층신경망(deep neural network) 또는 딥러닝으로 구분한

다. 그리고 딥러닝 기반 모형개발을 위한 다양한 오픈소스 프레임워크가 존재하며, 이러한 프레임워크는 여러 소프트웨어 라이브러리와 API (Application Programming Interface)를 지원함으로써 개발자로 하여금 모형 개발과 구현까지의 소요시간을 줄여줄 뿐만 아니라 성능향상에 도움을 준다. 본 연구에서는 구글의 브레인 프로젝트팀에서 개발하여 공개 이후 많은 프로그램 개발자, 연구자들이 사용하고 있으며, 개발진들과 전세계적인 오픈라인 피드백을 통해 지속적으로 업그레이드 되고 있는 딥러닝 오픈 프레임워크인 TensorFlow (<https://www.tensorflow.org/>)를 사용한다.

2.2 순환신경망(RNN)

Fig. 1은 일반적인 RNN 구조를 나타내고 있으며, 활성화 신호가 입력층에서 출력층으로만 이동되는 기본적인 FFNN과 달리 Fig. 1 좌측의 RNN 구조를 펼치게 되면 우측처럼 과거의 데이터의 정보가 임의의 시점에 영향을 주는 루프형의 구조를 가지고 있으며, 데이터의 가변길이인 시퀀스(sequence)를 활용함으로써 시계열 데이터를 분석하는데 유용하다.

임의의 시점 t 에서 순환뉴런(recurrent neuron)의 출력 Y_t 는 이전 타임스텝 출력 Y_{t-1} 을 포함한 모든 입력 X_t 에 대한 함수이며, 순환뉴런의 출력은 Eq. (1)과 같이 입력에 대한 가중치 W_x 와 이전 타임스텝의 출력에 대한 가중치 W_y 로 표현된다.

$$Y_t = \phi(W_x^T \cdot X_t + W_y^T \cdot Y_{t-1} + b) \quad (1)$$

여기서, b 는 편향(bias)을 나타내며, ϕ 는 계산을 위한 활성화 함수(activation function)를 나타낸다.

학습을 통한 Eq. (1)의 가중치 업데이트를 위해 일반적으로 사용되는 역전파(back-propagation) 알고리즘은 각 층에 포함되어 있는 개별 뉴런마다 오차경사를 전파시키면서 진행한다. 다만 긴 시퀀스의 자료를 사용한 깊게 쌓은 심층 RNN의 경우 학습이 진행되다 오차경사의 기울기가 소실되고 입력층

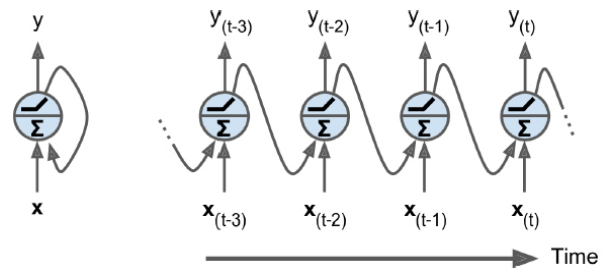


Fig. 1. Recurrent neural network structure (Geron, 2017)

과 가까운 층에서의 데이터의 정보를 기억하지 못하고 학습에 활용할 수 없는 단점이 있다.

2.3 LSTM

LSTM (Long Short-Term Memory)은 RNN에서 발생하는 오차경사의 기울기 소실(*gradient vanishing*) 및 최적화 오류(*optimization hurdle*)를 해결하고 장기적인 시간의 종속성을 포착하는데 효과적이므로 여러 가지 시계열 데이터와 관련된 학습문제에 효과적이고 확장 가능한 모형으로 이미 음성인식, 언어모델링, 번역뿐만 아니라 다른 신경망과의 결합을 통한 다양한 분야에서 활용되고 있다.

LSTM 구조는 Fig. 2와 같으며, 핵심요소로는 시간에 따라 상태(*state*)를 유지할 수 있는 메모리동 셀(*cell*)과 셀 안팎으로 데이터의 흐름을 조절하는 세 개의 비선형 게이트 구조이다(Klaus *et al.*, 2015). 즉, LSTM은 특정 시점의 상태(h_t)를 업데이트하기 위해 셀(*cell*, C_t)이라는 개념을 도입하여 입력과 현재까지의 상태를 이용하여 내부에 가지고 있는 정보를 업데이트할 것인지 아닌지를 판단한다. 그리고 이 셀의 데이터 흐름을 조절하기 위한 게이트의 종류로는 입력게이트 (*input gate*, i_t), 망각게이트(*forget gate*, f_t) 그리고 출력게이트(*output gate*, o_t)로 구성되어 있다.

망각게이트 f_t 는 Eq. (2)으로 나타낼 수 있으며 이전 셀의 출력인 h_{t-1} 과 현재의 입력인 x_t 를 시그모이드(*sigmoid function*) 활성화레이어에 적용해 0과 1사이의 값을 얻어 이 값을 현재 상태와 요소끼리 곱하게 되며, 그 과정에서 이 정보를 유지할지 제거할지를 선택하게 된다.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

여기서, σ 는 활성화함수를 나타내며 W_f 는 망각게이트의 가중치, b_f 는 편향을 나타낸다.

입력게이트 i_t 에서는 Eq. (3)과 같이 어떤 정보를 셀에 저장할 것인지를 결정하게 되는데 이는 크게 다음과 같은 두 단계로 구분된다. 1) 시그모이드 함수를 이용해 무엇을 업데이트 할지 결정하고, 2) 하이퍼탄젠트(*hypertangent*, *tanh*) 함수를 이용해 새로운 셀 상태 업데이트 시 사용되는 후보 셀(\tilde{C}_t)을 Eq. (4)를 통해 생성하게 된다.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

여기서, W_i , W_c 는 각각 입력게이트 및 후보 셀의 가중치, b_i , b_c 는 입력게이트와 후보 셀의 편향을 나타낸다.

다음으로 과거의 셀 상태(C_{t-1})와 후보 셀(\tilde{C}_t)을 Eq. (5)와 같이 조합하여 현재의 셀 상태(C_t)를 업데이트하게 된다.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{5}$$

마지막으로 출력게이트 o_t 는 Eq. (6)과 같이 시그모이드 함수를 이용해서 셀 상태의 어느 부분을 출력하고자 하는지 결정하고, 마지막으로 Eq. (7)과 같이 하이퍼탄젠트 함수를 이용하여 활성화된 셀 상태(C_t)와의 곱을 통해 특정 시점의 상태(h_t)를 업데이트하게 된다.

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \tag{6}$$

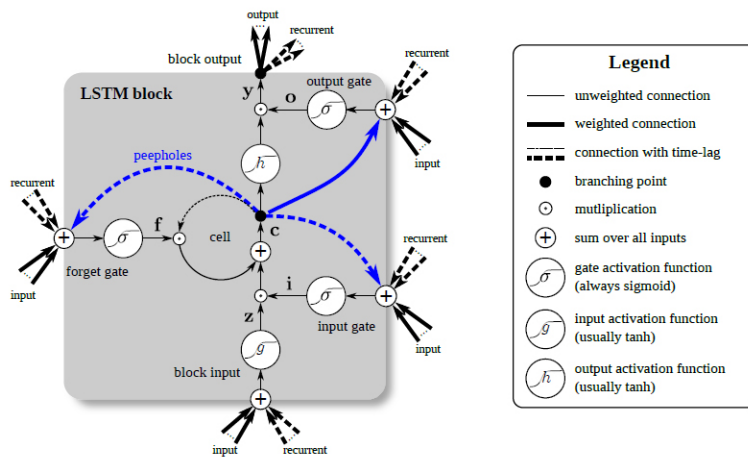


Fig. 2. LSTM Structure (Greff, 2017)

$$h_t = o_t \cdot \tanh(C_t) \quad (7)$$

여기서, W_o 는 출력게이트의 가중치, b_o 는 편향을 나타낸다. LSTM 구조와 관련된 자세한 사항은 Olah (2015)를 참조할 수 있다.

3. 대상 유역 및 모형 구축

3.1 대상 유역

한강은 유역면적이 34,428 km^2 이며 유로연장은 483 km 인 전 국토면적의 약 23%를 차지하고 있는 한반도 중앙부의 최대 하천이자 4대강 중 유일하게 하굿둑이 없어 조위의 영향을 받는 감조하천이다.

한강하류에서 감조구간의 상한은 잠실수중보이며 수위

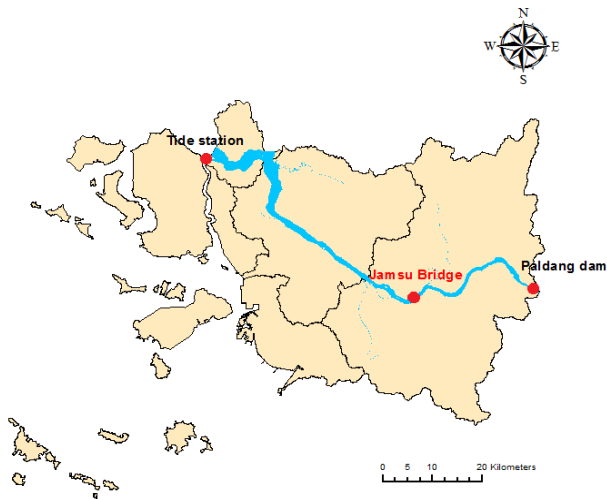


Fig. 3. Jamsu bridge location

를 기록하는 관측소는 잠실수중보 하류에 위치한 영동대교, 잠수교, 한강대교, 행주대교, 전류리 등 5개소이다(Shin and Yoon, 2005). Fig. 3은 한강유역에서 잠수교의 위치를 보여주고 있으며, 잠수교는 한강 하구에서 약 50 km 떨어진 지점에 위치하고 있으며, 수위가 상승하여 5.5 m 일 경우 보행제한, 6.2 m 일 경우 차량제한이 이뤄지며, 6.5 m 일 경우 완전잠수로 인해 교량이 통제된다. 2017년의 잠수교의 실측 수위자료는 Fig. 4와 같으며 4번의 보행자제한, 2번의 차량통제 및 완전침수가 발생한 것을 확인할 수 있다. 이와 같이 잠수교는 한강본류 서울시에 위치하고 있어 차량뿐만 아니라 자전거와 보행자의 통행이 잦으므로 홍수예보지점은 아니지만 시민들의 안전과 재산보호를 위하여 정확한 수위 예측이 요구된다.

3.2 LSTM 모형 구축

본 연구에서는 심층신경망 구조의 LSTM모형을 구축하여 감조하천에서 조위영향에 따른 한강 잠수교의 수위예측을 수행하고 선행시간(lead time)별 예측 정확도를 분석하는 것을 목적으로 한다. Fig. 5는 LSTM 모형의 학습과 예측을 위해 사용되는 2011년부터 2017년 11월 30일까지의 10분단위의 팔당댐의 방류량과 한강하구의 강화대교의 예측조위 및 잠수교의 수위 시계열 자료이다. 본 연구에서는 2011년부터 2016년의 자료는 신경망 학습을 위해 활용하고 2017년은 검증 및 시험데이터로 활용하여 예측 정확도를 비교·분석한다.

우선, MinMaxscaler 변환을 이용하여 Fig. 5의 시계열 자료를 정규화(normalization) 하였으며, 학습에 사용되는 입력데이터의 길이를 조정하는 시퀀스길이(sequence_length)는 1시간($t-5 \sim t$), 3시간($t-17 \sim t$), 6시간($t-35 \sim t$)으로 변화하며 신경망 학습을 수행하였다. 즉, 기본자료가 10분단위로 구성되어 있기 때문에 현재시간이 t 일 때 시퀀스길이가 6일 경우 $t-5 \sim t$ 까지 60분(1시간) 자료가 학습으로 활용, 시퀀스길이가

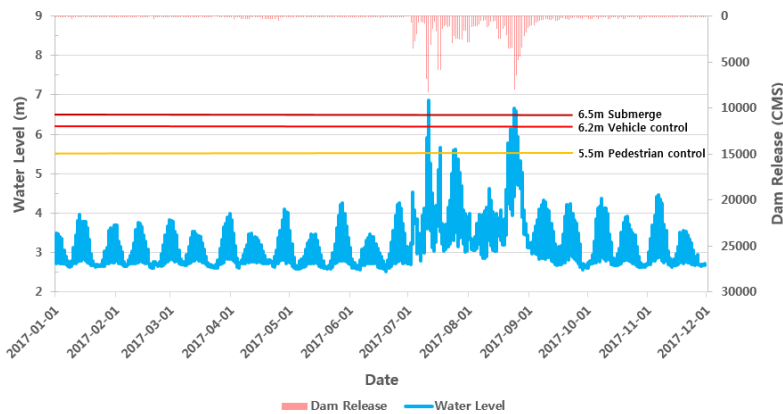


Fig. 4. Time series of water level at the jamsu bridge in 2017

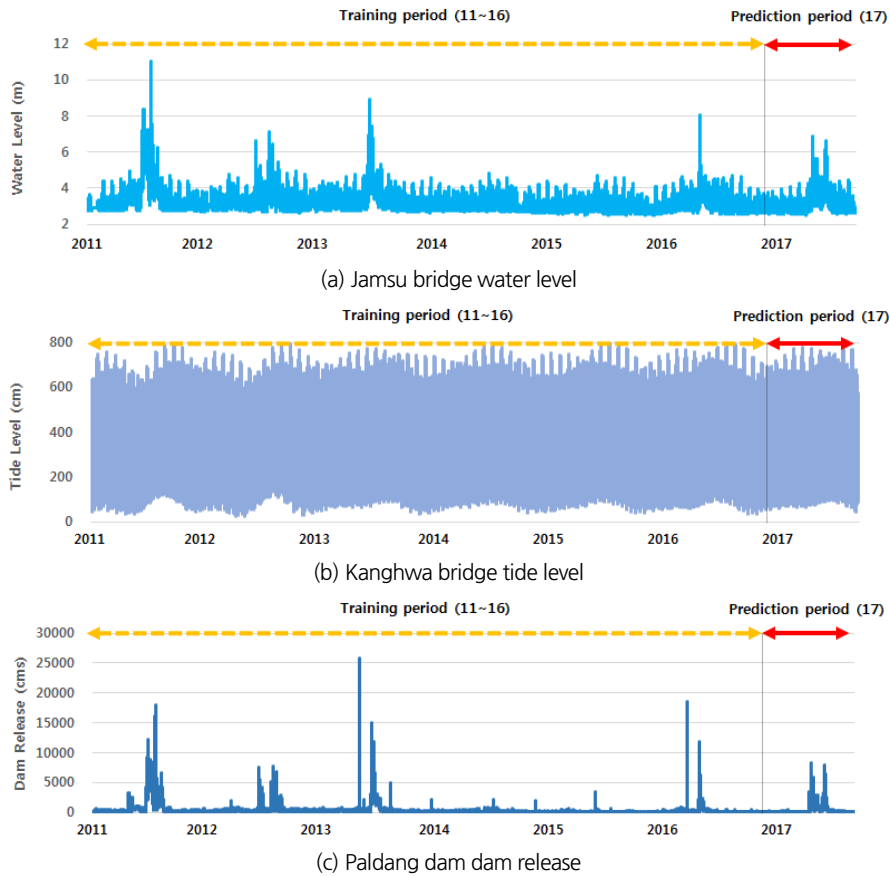


Fig. 5. Time series data sets for LSTM applications

18일 경우 $t-17 \sim t$ 까지 180분(3시간) 자료가 학습에 활용됨을 의미한다. 또한 민감도 분석을 통해 학습속도(learning_rate)는 0.01, 반복시행횟수(iteration)는 3000번, 은닉층 개수는 6으로 구성하였으며, 이상의 LSTM 모형 매개변수에 따른 선행시간 1시간($t+6$), 3시간($t+18$), 6시간($t+36$), 9시간($t+54$), 12시간($t+72$), 24시간($t+144$)의 잠수교 수위예측을 수행하고, 예측정확도의 정량화를 위해 RMSE (Root Mean Square Error)와 NSE (Nash-Sutcliffe Efficiency) 두 개의 평가지수를 사용하여 비교·분석하였다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (O_t - P_t)^2} \quad (8)$$

$$NSE = 1 - \frac{\sum_{t=1}^N (O_t - P_t)^2}{\sum_{t=1}^N (O_t - \bar{O})^2} \quad (9)$$

여기서, N 은 자료의 개수, O_t 와 P_t 는 시간 t 에서의 실제수위와 예측수위, \bar{O} 는 실제수위의 평균값을 나타낸다.

4. 모의 결과 및 분석

전술한 바와 같이 본 연구에서는 강우자료 및 지류 유입량 등 수문기상 정보를 배제하고, 10분단위의 각 잠수교의 수위 자료와 팔당댐의 방류량 자료 및 강화대교의 예측조위 세 가지 최소한의 자료만을 사용하여 LSTM 모형을 학습하고, 6개의 선행시간(1, 3, 6, 9, 12, 24시간)에 대한 잠수교 수위를 예측하고 그 성능을 평가하였다.

Tables 1 and 2는 LSTM 모형의 각 시퀀스길이에 따른 RMSE와 NSE 분석결과를 나타내고 있다. 그 결과, 학습에 사용되는 시퀀스길이가 길어질수록 전반적으로 예측 정확도는 낮아지고 있으나 12시간 및 24시간 예측 결과는 유사하거나 미세하게 향상되는 것으로 나타났다.

또한, LSTM 모형의 선행시간이 길어질수록 평균적으로

Table 1. RMSE results of different sequence length cases

Sequence length	Lead Time					
	1 hour	3 hours	6 hours	9 hours	12 hours	24 hours
1 hr (t-5 ~ t)	0.065	0.090	0.147	0.214	0.235	0.279
3 hr (t-17 ~ t)	0.079	0.118	0.208	0.216	0.237	0.282
6 hr (t-35 ~ t)	0.092	0.113	0.203	0.218	0.224	0.271
Average	0.079	0.107	0.186	0.216	0.232	0.277

Table 2. NSE results of different sequence length cases

Sequence length	Lead Time					
	1 hours	3 hours	6 hours	9 hours	12 hours	24 hours
1 hr (t-5 ~ t)	0.986	0.973	0.928	0.846	0.815	0.738
3 hr (t-17 ~ t)	0.979	0.953	0.855	0.843	0.811	0.732
6 hr (t-35 ~ t)	0.971	0.957	0.862	0.841	0.831	0.753
Average	0.979	0.961	0.882	0.843	0.819	0.741

RMSE는 0.08 m에서 0.28 m로 0.2 m 오차가 증가하였으며, NSE는 0.99에서 0.74로 모형의 예측 정확도가 감소하였다. 예측 정확도의 경우 시퀀스길이를 1시간으로 주었을 때 가장 높게 나타났으며, 이는 해당 선행시간에 대한 예측을 위해 LSTM 학습 진행 시 시퀀스길이를 선행시간과 동일하게 조정하는 것이 예측의 정확도를 보장하는 것은 아닌 것을 의미한다. 선행시간 1시간의 경우 RMSE는 0.065 m, NSE는 0.99로 매우 우수한 것으로 분석되었다.

Figs. 6 and 7은 시퀀스길이 1시간(t-5 ~ t)의 LSTM 모형을 이용한 선행시간별 2017년 잠수교 수위 예측결과를 나타내고 있으며, 수위비교 결과에서도 나타났듯이 선행시간이 짧을수록 2017년 예측기간에 대해 저수위뿐만 아니라 고수위까지 안정적으로 예측하는 것을 확인할 수 있다. 반면에 선행시간이 길어질수록 수위의 시간적 패턴은 유사하게 모의하고 있으나 우측 산포도에서 볼 수 있듯이 중고수위에서 실측수위의 오차가 크게 발생하는 것을 확인할 수 있다.

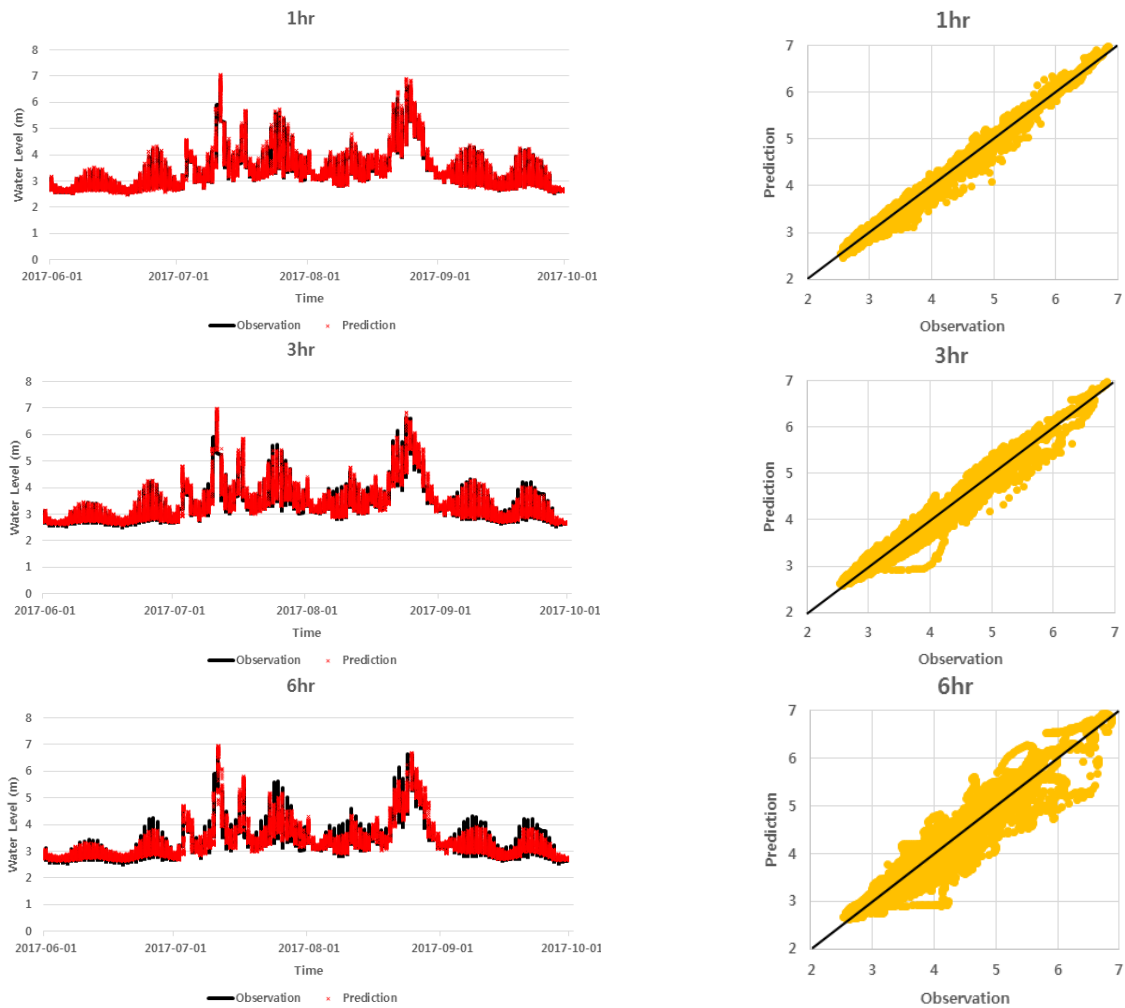


Fig. 6. Water level time series and scatter plots for lead time of 1 hr ~ 6 hrs

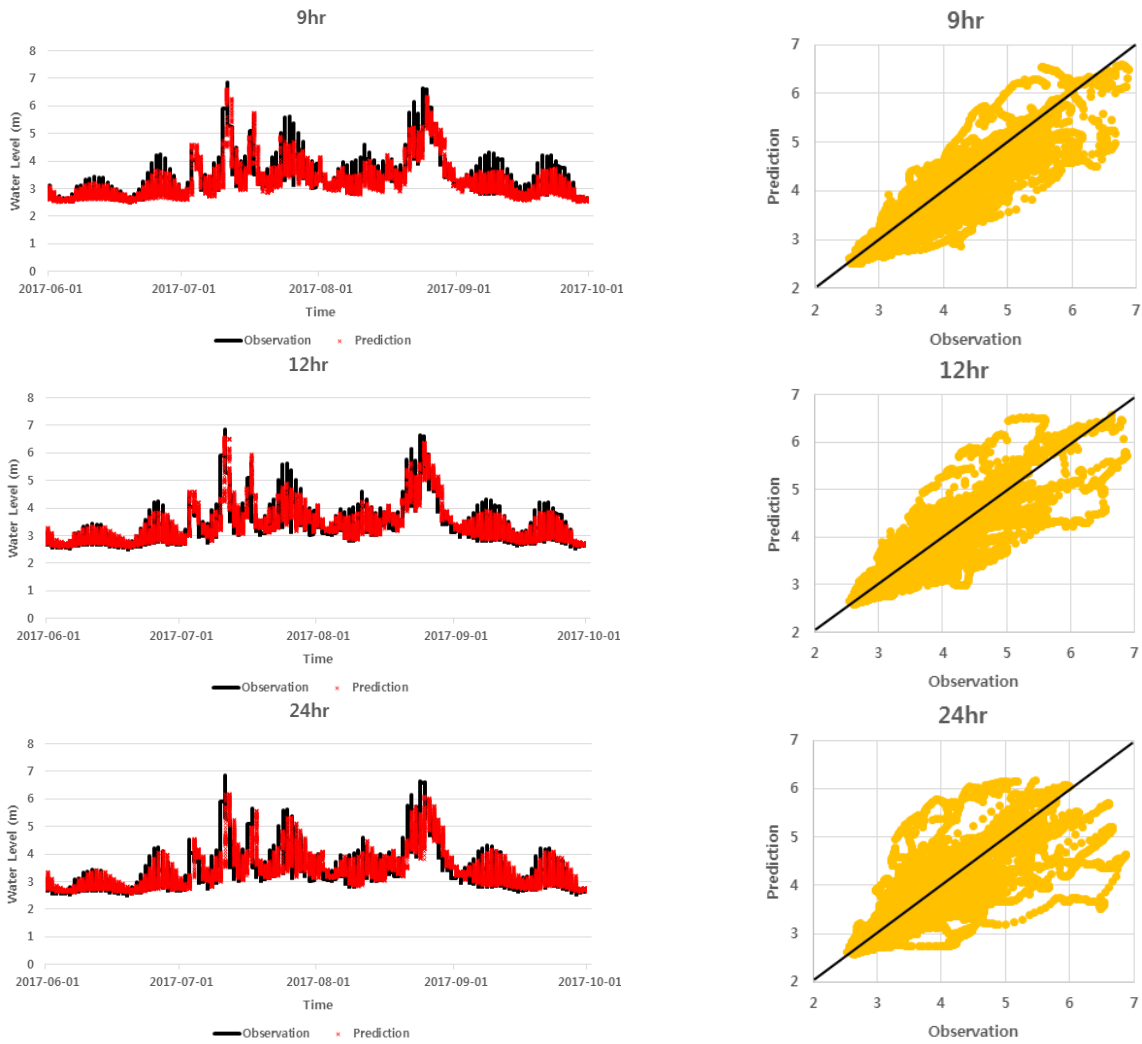


Fig. 7. Water level time series and scatter plots for lead time of 9 hrs ~ 24 hrs

2017년 모의기간 중 집중호우로 인하여 실제 잠수교 통제가 있었던 7월과 8월 사상(events)에 대한 시퀀스길이 1시간($t-5 \sim t$) 기반 LSTM 모형의 선행시간별 수위 예측결과는 Fig. 8과 같다. Fig. 8에서 각각의 색깔별 곡선은 선행시간별 수위 예측결과를 나타내고 있으며, 색깔별 점은 임의 시점 t 에서 선행시간별로 예측된 수위를 나타내고 있다. 예를 들어 Fig. 8(a)의 7월 10일 12시를 기점으로 선행시간 1시간에 해당하는 7월 10일 13시에 대한 예측결과는 노란색 점으로 표시되었으며, 선행시간 24시간에 해당하는 7월 11일 12시에 대한 예측결과는 파란색 점으로 표시되어 있다. 즉, Fig. 8의 색깔별로 표시된 곡선은 7월 및 8월 사상에 대한 선행시간별 예측수위(점)를 연결한 것이며, 모의기간 전기간(2017년)에 대한 개별 선행시간별 예측수위와 실측수위와의 비교결과가 Figs. 6 and 7이다.

실제로 Fig. 8에서 알 수 있듯이 선행시간이 길어질수록 예측

수위곡선이 전반적으로 과소추정되는 것을 확인할 수 있었으며, Fig. 8(a)의 실제 침두수위 (6.87 m)가 발생한 일시(2017. 7월 11일 10시 20분)에서 선행시간 1시간(6.93 m), 3시간(6.73 m), 6시간(6.64 m)의 침두수위는 비교적 높은 정확도를 나타내고 있으나 9시간부터 예측 침두수위가 과소추정되는 것으로 나타났다. Fig. 8(b)의 8월 사상의 경우는 비교적 고수위가 5 m ~ 6.7 m 사이에서 지속되어 과거 입력정보를 기억하여 예측에 활용하는 LSTM의 모형특성 상 각 선행시간별 모의결과의 재현성이 높게 나타난 것으로 나타났으나 7월 사상과 마찬가지로 선행시간 9시간 이후 예측수위는 전반적으로 과소추정되었다.

5. 결론

본 연구는 물리적 수리·수문모형의 적용이 제한적인 조위영

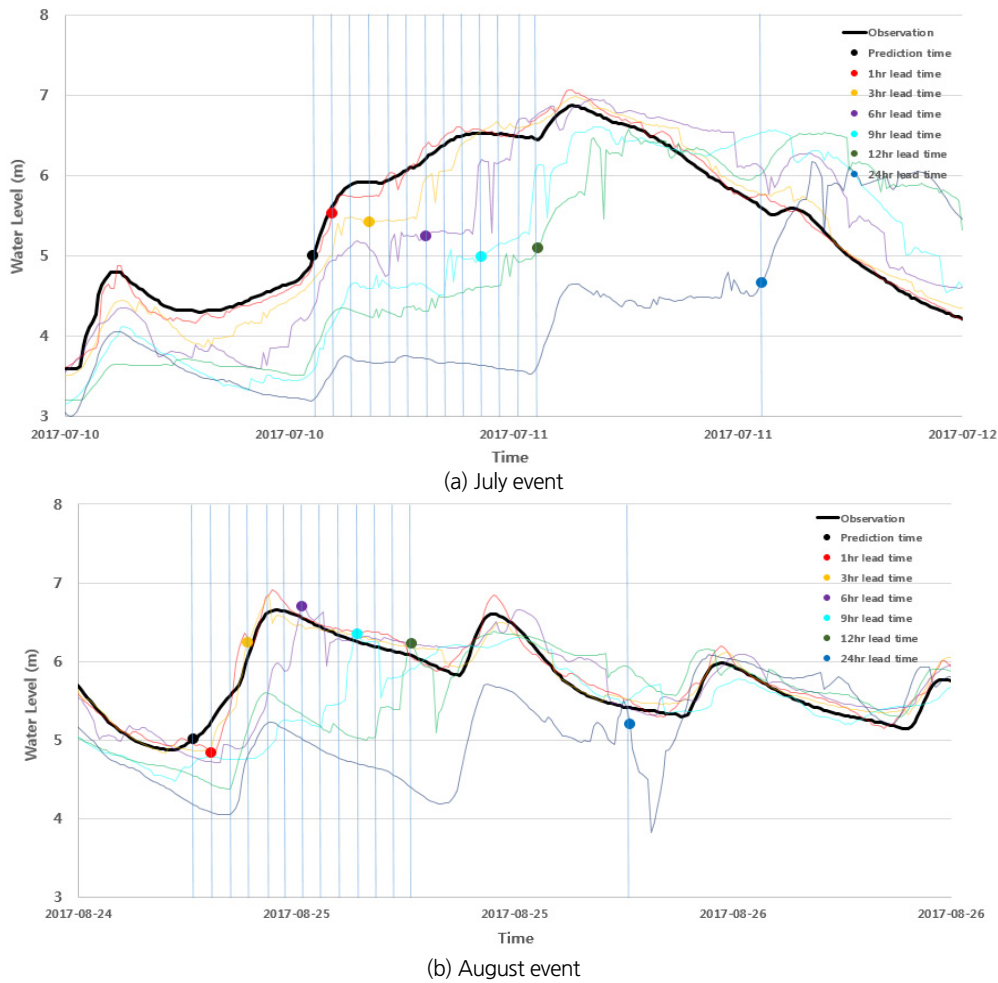


Fig. 8. Comparison of water levels for the specific rainfall events

향을 받는 감조하천에서의 수위예측을 목적으로 하고 있으며, 이를 위해 한강 잠수교를 대상으로 딥러닝 오픈소스 소프트웨어 라이브러리인 TensorFlow를 활용하여 LSTM 모델을 구성하고 2011년부터 2017년까지의 10분 단위의 잠수교 수위, 팔당댐 방류량과 한강하구 강화대교지점의 예측조위 자료를 이용하여 수위예측을 수행하였다. 모형 매개변수는 민감도 분석을 통해 은닉층의 개수는 6개, 학습속도는 0.01, 학습횟수는 3000번로 결정하였으며, 모형 학습 시 학습정보의 시간적 양을 결정하는 중요한 매개변수인 시퀀스길이는 1시간, 3시간, 6시간으로 변화시키며 모의하였다. 최종적으로 선행시간에 따른 모의 예측능력을 평가하기 위해 LSTM 모형의 예측 선행시간을 6개 (1~24시간)로 구분하여 실측수위와 예측수위와의 비교·분석을 수행한 후, 다음과 같은 결과를 도출하였다.

첫째, LSTM 모형의 최적의 성능을 내는 결과는 시퀀스길이를 1시간($t-5 \sim t$)으로 하였을 때로 분석되었으며, 특히 선행시간 1시간에 대한 예측정확도는 RMSE는 0.065 m, NSE는 0.99

로 실측수위에 매우 근접한 예측 결과를 나타내었다. 다만, 모형 학습 시 시퀀스길이가 길어질수록 전 기간에 걸쳐 동일한 선행시간에 대해 정확도가 미세하게 감소하는 것으로 나타났으며, 시퀀스길이에 상관없이 선행시간이 길어질수록 모형의 예측 정확도는 평균적으로 RMSE 0.08 m에서 0.28 m로 오차가 증가하였으며, NSE는 0.99에서 0.74로 감소한 것으로 분석되었다.

둘째, 실제 홍수기간 중 완전잠수(6.5 m)를 초과한 사상에 대한 분석한 결과, 7월 집중호우로 인한 사상에서는 선행시간 6시간이하의 경우에서는 침두수위에 대해 어느 정도의 정확도가 확보되었으나 9시간이후에서는 침두수위 및 전체적인 수위곡선이 과소모의된 것으로 분석되었다. 8월 사상의 경우 완전잠수가 발생한 침두수위 전·후로 고수위가 지속되어 선행시간의 정보를 기억하여 학습에 반영하는 LSTM의 특성으로 7월에 비해 선행시간별 침두수위 모의결과의 재현성이 높게 나타났으나 마찬가지로 선행시간 9시간이후의 경우, 수위곡선이 전반적으로 과소추정된 것으로 확인되었다. 일반

적으로 수위예측을 위한 선행시간이 길어질수록 예측정확도가 감소하는 것을 확인하였으나 시계열 분석 시 현재시점을 기준으로 과거의 자료가 어느 정도 구간까지 영향을 미치는지에 대한 자료구조의 시계열 특성(자기상관성 등)을 분석하고, LSTM의 중요 매개변수인 시퀀스 길이의 객관적인 설정을 위한 후속 연구가 필요하다.

현재 운영중인 홍수예보시스템의 경우, 예측강우를 활용하여 홍수예보를 수행하고 있어 충분한 선행시간에 대한 예측강우에 대한 신뢰성이 보장되지 않으면 그 결과 역시 신뢰할 수 없다. 이에 본 연구에서는 조위영향을 받는 감소하천에서의 수위예측을 위한 물리적 수리-수문모형기반 시스템의 보완적 수단으로 딥러닝 기반의 LSTM 모형을 제안하였으며, 향후 딥러닝 관련 알고리즘과 컴퓨팅 능력의 지속적인 향상으로 인하여 다양한 인자를 기반으로 하는 시계열 예측모의가 가능할 것으로 판단된다. 특히 한강하류부와 같이 자연적, 인위적 구속(regulated)을 동시에 받는 특정지점의 비선형 수문 시계열 분석이 목적인 경우, 방대한 자료의 장기적인 정보를 기억하며 이를 조정하여 예측에 반영하는 LSTM 모형은 홍수예보뿐만 아니라 수문-기상학적 분석에서의 활용 및 응용이 가능할 것으로 판단된다.

감사의 글

본 연구는 국토교통부/국토교통과학기술진흥원의 지원으로 수행되었습니다(과제번호 18AWMP-B127555-02).

References

- Chen, W. B., Liu, W. C., and Hsu, M. H. (2012). "Comparison of ANN approach with 2D and 3D hydrodynamic models for simulating estuary water stage." *Advances in Engineering Software*, Vol. 45, No. 1, pp. 69-79.
- Chiang, Y. M., Chang, L. C., Tsai, M. J., Wang, Y.F., and Chang, F.J. (2010). "Dynamic neural networks for real-time water level predictions of sewerage systems-covering gauged and ungauged sites." *Hydrology and Earth System Sciences*, Vol. 14, No. 7, pp. 1309-1319.
- Géron, A. (2017). "Hands-on machine learning with Scikit-Learn and TensorFlow." O'Reilly Media.
- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B.R., and Schmidhuber, J. (2017). "LSTM: A search space odyssey." *IEEE transactions on neural networks and learning systems*, Vol. 28, No. 10, pp. 2222-2232.
- Hidayat, H., Hoitink, A. J. F., Sassi, M.G., and Torfs, P. J. J. F. (2014). "Prediction of discharge in a tidal river using artificial neural networks." *Journal of Hydrologic Engineering*, ASCE, Vol. 19, No. 8, 04014006.
- 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 Korean Society of Hazard Mitigation*, Vol. 18, No. 1, pp. 1-11.
- Kim, S., and Tachikawa, Y. (2017). "Real-time river-stage prediction with artificial neural network based on only upstream observation data." *Annual Journal of Hydraulic Engineering*, JSCE, Vol. 62, pp. 1375-1380.
- Lee, E. R., Kim, W., and Kim, S. H. (2005). "Effect of Flood Stage by Hydraulic Factors in Han River." *Journal of Korea Water Resources Association*, Vol. 38, No. 2, pp. 121-131.
- Lee, G. H., Jung, S. H., and Lee, D. E. (2018). "Comparison of physics-based and data-driven models for streamflow simulation of the Mekong river." *Journal of Korea Water Resources Association*, Vol. 51, No. 6, pp. 503-514.
- Lee, J. K., and Lee, J. H. (2010). "A Study on Water Level Rising Travel Time due to Discharge of Paldang Dam and Tide of Yellow Sea in Downstream Part of Paldang Dam." *Journal of Korean Society of Hazard Mitigation*, Vol. 10, No. 2, pp. 111-122.
- Lee, S. J., Ryoo, K. S., Lee, B. S., and Yoon, J. S. (2007). "Development of Regression Equation for Water Quantity Estimation in a Tidal River." *Journal of Korean Society on Water Quality*, Vol. 23, No. 3, pp. 385-390.
- Maier, H. R., and Dandy, G. C. (2000). "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications." *Environmental modelling & software*, Vol. 15, No. 1, pp. 101-124.
- Ministry of Land Infrastructure and Transport, Han River Flood Control Office (2017). Han River Forecast Yearbook.
- Oh, C. Y., Yang, D. H., Lee, J. H., and Won, Y. S. (2014). "Application of Artificial Neural Network Model for Flood Forecasting." *Water for Future*, Vol. 47, No. 5, pp. 50-62.
- Olah, C. (2015). "Understanding lstm networks." GITHUB blog, <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- Park, C. G., and Baek, K. O. (2017). "Reconsideration of evaluating design flood level at Imjin River estuary." *Journal of Korea Water Resources Association*, Vol. 50, No. 9, pp. 617-625.
- Song, C. G., Kim, H. J., and Lee, D. S. (2014). "Analysis of Flow Reversal by Tidal Elevation and Discharge Conditions in a Tidal River." *Journal of the Korean society of Safety*, Vol. 29, No. 6, pp. 104-110.
- Seo, I. W., Song, C. G., and Lee, M. E. (2008). "Flow and Mixing behavior at the Tidal Reach of Han River." *Journal of the Korean Society of Civil Engineers*, KSCE, Vol. 28, No. 6B, pp. 731-741.
- Supharatid, S. (2003). "Application of a neural network model in establishing a stage-discharge relationship for a tidal river." *Hydrological processes*, Vol. 17, No. 15, pp. 3085-3099.
- Shin, Y. K., and Yoon, K. S. (2005). "The Spatial Distribution of Water Quality and Sediments Characteristics in the Han River Estuary." *Journal of the Korean Geomorphological Association*, Vol. 12, No. 4, pp. 13-23.