



## Comparative analysis of activation functions of artificial neural network for prediction of optimal groundwater level in the middle mountainous area of Pyoseon watershed in Jeju Island

Shin, Mun-Ju<sup>a\*</sup> · Kim, Jin-Woo<sup>b</sup> · Moon, Duk-Chul<sup>c</sup> · Lee, Jeong-Han<sup>d</sup> · Kang, Kyung Goo<sup>e</sup>

<sup>a</sup>Senior researcher, Water Resources Research Team, Jeju Province Development Corporation, Jeju, Korea

<sup>b</sup>Senior researcher, Water Resources Research Team, Jeju Province Development Corporation, Jeju, Korea

<sup>c</sup>Senior researcher, Water Resources Research Team, Jeju Province Development Corporation, Jeju, Korea

<sup>d</sup>Principal researcher, Water Resources Research Team, Jeju Province Development Corporation, Jeju, Korea

<sup>e</sup>Director, R&D Innovation Center, Jeju Province Development Corporation, Jeju, Korea

Paper number: 21-075

Received: 27 September 2021; Revised: 9 November 2021; Accepted: 18 November 2021

### Abstract

The selection of activation function has a great influence on the groundwater level prediction performance of artificial neural network (ANN) model. In this study, five activation functions were applied to ANN model for two groundwater level observation wells in the middle mountainous area of the Pyoseon watershed in Jeju Island. The results of the prediction of the groundwater level were compared and analyzed, and the optimal activation function was derived. In addition, the results of LSTM model, which is a widely used recurrent neural network model, were compared and analyzed with the results of the ANN models with each activation function. As a result, ELU and Leaky ReLU functions were derived as the optimal activation functions for the prediction of the groundwater level for observation well with relatively large fluctuations in groundwater level and for observation well with relatively small fluctuations, respectively. On the other hand, sigmoid function had the lowest predictive performance among the five activation functions for training period, and produced inappropriate results in peak and lowest groundwater level prediction. The ANN-ELU and ANN-Leaky ReLU models showed groundwater level prediction performance comparable to that of the LSTM model, and thus had sufficient potential for application. The methods and results of this study can be usefully used in other studies.

**Keywords:** Activation function, ANN, LSTM, Groundwater level prediction, Pyoseon watershed in Jeju Island

## 제주도 표선유역 중산간지역의 최적 지하수위 예측을 위한 인공신경망의 활성화함수 비교분석

신문주<sup>a\*</sup> · 김진우<sup>b</sup> · 문덕철<sup>c</sup> · 이정환<sup>d</sup> · 강경구<sup>e</sup>

<sup>a</sup>제주특별자치도개발공사 수자원연구팀 선임연구원, <sup>b</sup>제주특별자치도개발공사 수자원연구팀 선임연구원,

<sup>c</sup>제주특별자치도개발공사 수자원연구팀 선임연구원, <sup>d</sup>제주특별자치도개발공사 수자원연구팀 책임연구원, <sup>e</sup>제주특별자치도개발공사 R&D 혁신센터 센터장

### 요 지

활성화함수의 선택은 인공신경망(Artificial Neural Network, ANN) 모델의 지하수위 예측성능에 큰 영향을 미친다. 특히 제주도의 중산간 지역과 같이 지하수위의 변동폭이 크고 변동양상이 복잡한 경우 적절한 지하수위 예측을 위해서는 다양한 활성화함수의 비교분석을 통한 최적의 활성화함수 선택이 반드시 필요하다. 본 연구에서는 지하수위의 변동폭이 크고 변동양상이 복잡한 제주도 표선유역 중산간지역 2개 지하수위 관측정을 대상으로 5개의 활성화함수(sigmoid, hyperbolic tangent (tanh), Rectified Linear Unit (ReLU), Leaky Rectified Linear Unit (Leaky ReLU), Exponential Linear Unit (ELU))를 ANN 모델에 적용하여 지하수위 예측결과를 비교 및 분석하고 최적 활성화함수를 도출하였다. 그리고 최근 널리 사용되고 있는 순환신경망 모델인 Long Short-Term Memory (LSTM) 모델의 결과와 비교분석하였다. 분석결과 지하수위 변동폭이 상대적으로 큰 관측정과 상대적으로 작은 관측정에 대한 지하수위 예측에 대해서는 각각 ELU와 Leaky ReLU 함수가 최적의 활성화함수로 도출되었다. 반면 sigmoid 함수는 학습기간에 대해 5개 활성화함수 중 예측성능이 가장 낮았으며 침투 및 최저 지하수위 예측에서 적절하지 못한 결과를 도출하였다. 따라서 ANN-sigmoid 모델은 가뭄기간의 지하수위 예측을 통한 지하수자원 관리목적으로 사용할 경우 주의가 필요하다. ANN-ELU와 ANN-Leaky ReLU 모델은 LSTM 모델과 대등한 지하수위 예측성능을 보여 활용가능성이 충분히 있으며 LSTM 모델은 ANN 모델들 보다 예측성능이 높아 인공지능 모델의 예측성능 비교분석 시 참고 모델로 활용될 수 있다. 마지막으로 학습기간의 정보량에 따라 학습기간의 지하수위 예측성능이 검증 및 테스트 기간의 예측성능보다 낮을 수 있다는 것을 확인하였으며, 관측지하수위의 변동폭이 크고 변동양상이 복잡할수록 인공지능 모델별 지하수위 예측능력의 차이는 커졌다. 본 연구에서 제시한 5개의 활성화함수를 적용한 연구방법 및 비교분석 결과는 지하수위 예측뿐만 아니라 일단위 하천유출량 및 시간단위 홍수량 등 지표수 예측을 포함한 다양한 연구에 유용하게 사용될 수 있다.

**핵심용어:** 활성화함수, ANN, LSTM, 지하수위 예측, 제주도 표선유역

\*Corresponding Author. Tel: +82-64-780-3751

E-mail: munjushin@jpdco.kr (M.-J. Shin)

### 1. 서론

지하수는 지표수와 함께 공업, 농업 및 생활용수 등으로 사용할 수 있는 중요한 수자원이다. 특히 제주도의 경우 지하수가 전체 용수의 81%를 차지하므로(JSSGP, 2018) 지속가능한 이용을 위해 지하수의 관리는 매우 중요하다. 수량적인 관점에서 볼 때 안정적으로 지하수를 이용하기 위해서는 취수지역의 기상변화 등을 고려한 정확한 지하수위 변동예측 연구가 반드시 필요하다.

지하수위 예측을 위해 유한차분법 등 연산방법과 Darcy 법칙 등 물리적 법칙, 그리고 수치표고모델(Digital Elevation Model) 등 다양한 물리적 자료를 사용하여 지하수 흐름을 모의하고 대상지점의 지하수위를 계산하는(Todd and Larry, 2004) 지하수 수치모델인 MODFLOW (McDonald and Harbaugh, 1988)를 사용할 수 있다. 이 지하수 수치모델은 지하수위 모의에 널리 사용되고 있다(Mohanty *et al.*, 2013). 지하수 수치모델을 사용하여 공간적으로 균질하지 않은 지하수 시스템을 모의하기 위해서는 시간 및 공간적으로 다양한 자료가 필요하며 (Barthel and Banzhaf, 2016), 만약 양질의 시공간적 자료의 취득이 어려울 경우 지하수위를 적절히 모의하는데 어려움이 있다(Sun *et al.*, 2016). 또한 복잡한 유동프로세스의 단순화에 따른 불확실성(White *et al.*, 2014), 수치모델 모의격자의 크기(discretization) 설정에 따른 불확실성(White *et al.*, 2020), 그리고 많은 비용과 긴 모의시간(Maxwell *et al.*, 2015) 등의 단점이 있다. 대안적인 방법으로 목표(반응)변수(지하수위 등)와 목표변수에 관련된 설명변수(강수 및 취수량 등) 간의 관련정도(연결강도)를 복수의 은닉층 내의 셀들로 구성된 망을 사용하여 학습함으로써 지하수위를 예측하는 데이터기반 모델인 인공지능 모델을 활용할 수 있다. 인공지능망(Artificial Neural Network, ANN)을 포함한 인공지능 모델은 다양한 공간자료 이용의 부담이 없다는 장점으로 인해 다수의 연구에 성공적으로 사용되었다(Rajae *et al.*, 2019). ANN 모델은 장단기기억(Long Short-Term Memory, LSTM) (Hochreiter and Schmidhuber, 1997) 모델과 더불어 수문학 및 수자원분야 연구에 최근까지 널리 사용되었다(Sit *et al.*, 2020). ANN 모델을 활용한 지하수위 예측은 수자원관리에 필요한 유용한 정보를 제공할 수 있다.

지하수위 예측을 위한 ANN 모델 개발 시 활성화함수는 ANN 모델의 학습 및 예측성능에 큰 영향을 미친다(Kim *et al.*, 2021a). 일반적으로 널리 사용되는 활성화함수는 sigmoid 함수와 hyperbolic tangent (tanh) 함수이다(Haykin, 2009). 본 연구의 조사를 기준으로 지하수위 예측 연구의 경우, 과거 20년 동안 ANN 모델을 활용한 연구를 조사한 결과 sigmoid를

사용한 연구가 16편으로 가장 많았고 tanh를 사용한 연구가 3편 그리고 Rectified Linear Unit (ReLU) (Hahnloser *et al.*, 2000)를 사용한 연구가 2편으로써 주로 sigmoid 함수를 사용하였다(Table 1). 주목할 점으로써 최근의 연구에서는 sigmoid 함수의 단점을 보완하기 위해 ReLU 함수를 사용하였다는 점이다(Table 1). 그리고 이 활성화함수들 이외에 Leaky Rectified Linear Unit (Leaky ReLU) (Maas *et al.*, 2013)와 Exponential Linear Unit (ELU) (Clevert *et al.*, 2016) 등 다양한 활성화함수들이 사용되고 있다. 따라서 연구대상지역의 지하수위 예측에 적합한 최적의 ANN 모델을 개발하기 위해서는 다양한 활성화함수를 사용한 지하수위 예측결과의 비교분석이 필요하다. 댐 유입량 및 홍수량 예측 등을 위해 다양한 활성화함수를 비교분석한 연구가 수행되었지만(Ukkonen and Mäkelä, 2019; Worland *et al.*, 2019; Kim *et al.*, 2021a, 2021b) 지하수위 예측에 대해 활성화함수를 비교분석한 연구는 충분하지 않다. 특히 제주도의 중산간 지역과 같이 지하수위의 변동폭이 크고 변동양상이 복잡한 경우 ANN 모델을 활용하여 적절한 지하수위 예측을 위해서는 다양한 활성화함수의 비교분석을 통한 최적의 활성화함수 선택이 반드시 필요하다.

본 연구의 목적은 ANN 모델에 5개의 다양한 활성화함수

Table 1. Groundwater level analysis study applying various activation functions to ANN model

No.	Authors (year)	Activation Function (s)
1	Coulibaly <i>et al.</i> (2001)	sigmoid
2	Nayak <i>et al.</i> (2006)	sigmoid
3	Krishna <i>et al.</i> (2008)	sigmoid
4	Yoon <i>et al.</i> (2011)	sigmoid
5	Rakhshandehroo <i>et al.</i> (2012)	sigmoid
6	Taormina <i>et al.</i> (2012)	tanh
7	Sahoo and Jha (2013)	sigmoid
8	Emamgholizadeh <i>et al.</i> (2014)	sigmoid, tanh
9	Jha and Sahoo (2014)	sigmoid
10	Chang <i>et al.</i> (2015)	sigmoid
11	Hosseini <i>et al.</i> (2016)	sigmoid
12	Sun <i>et al.</i> (2016)	sigmoid
13	Yoon <i>et al.</i> (2016)	sigmoid
14	Sahoo <i>et al.</i> (2017)	sigmoid
15	Wen <i>et al.</i> (2017)	sigmoid
16	Kim and Oh (2018)	tanh
17	Yu <i>et al.</i> (2018)	sigmoid
18	Lee <i>et al.</i> (2019)	sigmoid
19	Afzaal <i>et al.</i> (2020)	ReLU
20	Müller <i>et al.</i> (2021)	ReLU

를 적용하여 제주도내 표선유역 중산간지역 2개 관측정의 지하수위 예측결과를 비교분석하고 최적의 활성화함수를 선택하는데 있다. 적용된 활성화함수는 인공지능 모델에 널리 사용되고 있는 sigmoid, tanh, ReLU, Leaky ReLU 및 ELU 함수이다. 또한 2개 관측정의 관측지하수위의 변동특성과 활성화함수별 ANN 모델의 예측성능의 관계를 분석하였다. 그리고 최근 널리 사용되고 있는 대표적인 순환신경망 모델인 LSTM 모델을 참고모델로 사용하여 ANN의 결과와 비교분석하였으며 이 점에서 기존연구와 차이점이 있다. 본 연구에서 사용한 연구방법은 2장에 상세히 기술하였으며 결과 및 결론은 각각 3장과 4장에 제시하였다.

## 2. 자료 및 방법

### 2.1 인공지능 모델

#### 2.1.1 인공신경망 모델

인공신경망(ANN) 모델은 뉴런(노드)과 시냅스로 이루어진 인간 뇌의 병렬적 학습 프로세스로부터 영감을 받아 개발한 기계학습 방법 중 하나이다(Haykin, 2009). 이 모델은 연구대상 지역에 대한 물리적 특성자료의 사용 없이 다양한 입력 시계열자료(관측 강수량 및 취수량 등)와 목적 시계열자료(관측 지하수위 등) 간의 상관관계를 이용하여 목적 시계열자료를 모사한다(Jha and Sahoo, 2014). ANN 모델은 입력층(input layer), 단수 또는 복수의 은닉층(hidden layer) 그리고 출력층(output layer) 순서의 연결구조를 가지며, 각 층은 단수 또는 복수의 뉴런으로 구성된다(Fig. 1). ANN 모델의 출력값 계산식은 다음과 나타낼 수 있다(Kim and Valdés, 2003).

$$\hat{y}_k = f_o \left[ \sum_{j=1}^m w_{kj} \cdot f_h \left( \sum_{i=1}^n w_{ji} x_i + w_{jo} \right) + w_{ko} \right] \quad (1)$$

여기에서  $w_{ji}$ 는 입력층의  $i$ 번째 뉴런에서 은닉층의  $j$ 번째 뉴런으로의 연결강도(가중치),  $x_i$ 는 입력층의  $i$ 번째 입력변수,  $w_{jo}$ 는  $j$ 번째 은닉뉴런의 편이,  $f_h$ 는 은닉뉴런의 활성화함수,  $w_{kj}$ 는 은닉층의  $j$ 번째 뉴런에서 출력층의  $k$ 번째 뉴런으로의 연결강도,  $w_{ko}$ 는  $k$ 번째 출력뉴런의 편이,  $f_o$ 는 출력뉴런의 활성화함수,  $\hat{y}_k$ 는 출력변수로서 모의결과를 나타낸다. 이 가중치들은 매개변수로서 모의결과와 관측자료 간의 오차를 역전파 알고리즘(backpropagation algorithm)을 사용하여 최소화 시키는 학습과정을 통해 최적의 값으로 업데이트된다(Haykin, 2009).

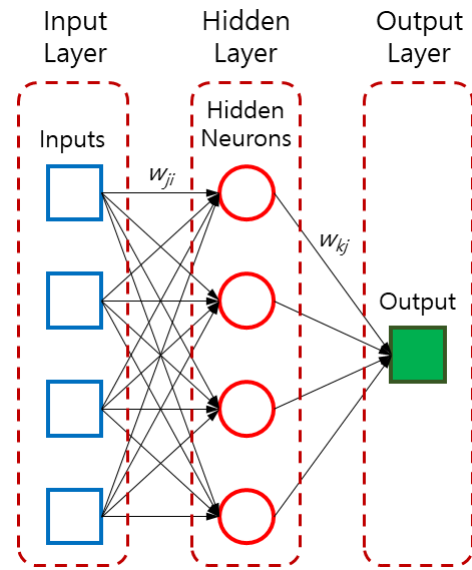


Fig. 1. The structure of ANN model

#### 2.1.2 장단기기억 모델

장단기기억(LSTM) 모델은 재귀신경망 모델의 변형으로써 순환신경망 모델의 장기간 기억능력(장기의존성, long-term dependencies)을 저해하는 경사소멸(vanishing gradients) 문제(Bengio et al., 1994)를 해결하기 위해 개발되었다. LSTM 모델은 각 계산시간(time step,  $t$ )에서 추출한 정보를 장기간 동안 기억하기 위해 carry track 이라는 컨베이어벨트를 사용한다(Fig. 2). 이 carry track은 시계열자료의 시간대별 처리과정과 평행하게 배치하여 추출한 정보를 이동시키고 각 계산시간대에서 필요할 때마다 추출한 정보를 재사용하므로 시계열자료의 계산과정에서 오래된 정보가 사라지는 문제가 발생하지 않는다(Chollet and Allaire, 2018). LSTM 모델은 네 가지의 독특한 변환을 사용하여 장기간의 정보를 학습한다는 점에서 기본적인 순환신경망 모델과 차이점이 있다. 뉴런에서  $t$  시간대에 대한 결과( $output_t$ )는 다음과 같이 계산된다.

$$output_t = activation(Wo \cdot input_t + Uo \cdot state_t + Vo \cdot c_t + bo) \quad (2)$$

여기에서  $input_t$ 는  $t$  시간대의 입력자료,  $state_t$ 는  $t$  시간대의 상태로써  $t-1$  시간대의 결과의 상태,  $c_t$ 는  $t$  시간대의 carry값,  $Wo$ ,  $Uo$ ,  $Vo$ 는 결과계산을 위한  $input_t$ ,  $state_t$ ,  $c_t$  각각의 가중치(매개변수) 행렬,  $\cdot$ 은 내적(dot product),  $bo$ 는 뉴런의 편이,  $activation$ 은 sigmoid 함수와 tanh 함수를 사용하는 활성화함수이다.

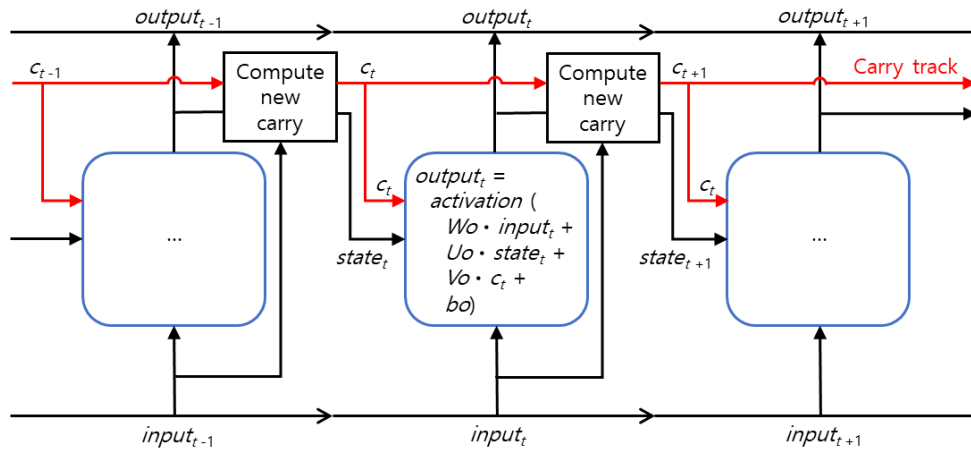


Fig. 2. The structure of LSTM model. Reproduced from Chollet and Allaire (2018)

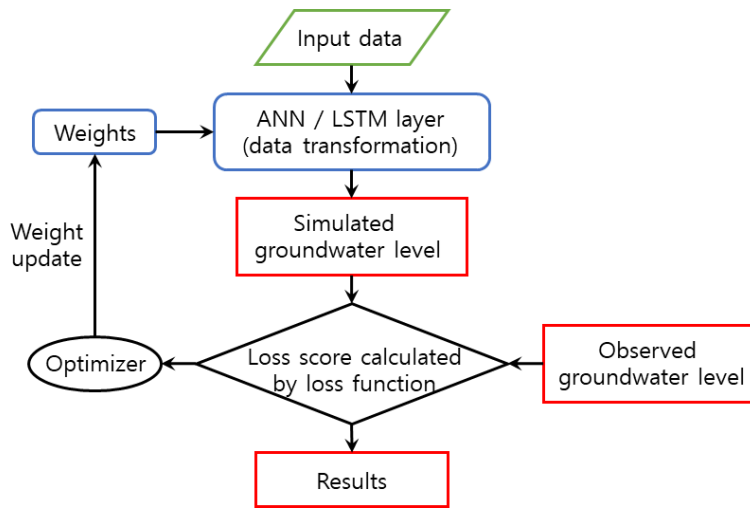


Fig. 3. Groundwater level modeling process of artificial intelligence model. Reproduced from Chollet and Allaire (2018)

carry값은 세 가지의 개별적인 변환을 통해 업데이트되며 그 변환식은 아래와 같다.

$$i_t = sigmoid(W_i \cdot input_t + U_i \cdot state_t + b_i) \quad (3)$$

$$f_t = sigmoid(W_f \cdot input_t + U_f \cdot state_t + b_f) \quad (4)$$

$$k_t = tanh(W_k \cdot input_t + U_k \cdot state_t + b_k) \quad (5)$$

$$c_{t+1} = i_t k_t + c_t f_t \quad (6)$$

여기에서  $i_t$ 는 sigmoid 함수를 통해 새롭게 추가된 정보([0, 1]의 범위),  $f_t$ 는 sigmoid 함수를 통해 삭제된 정보([0, 1]의 범위),  $k_t$ 는 tanh 함수를 통한 정보의 중요도([-1, 1]의 범위)를 의

미한다. 즉,  $i_t$ 와  $k_t$ 를 곱하여 새로운 정보를 얻고,  $c_t$ 와  $f_t$ 를 곱하여 관련성 없는 carry의 정보를 제거하며, 마지막으로  $i_t k_t$ 와  $c_t f_t$ 를 더하여 새로운 carry값을 생성한다. LSTM 모델은 지하수위 예측, 지하수 취수영향 분석, 결측데이터 처리 등 다양한 지하수위 예측연구에 성공적으로 사용되었다(Jeong and Park, 2019; Shin *et al.*, 2020; Vu *et al.*, 2021).

인공지능 모델의 지하수위 모델링을 위한 지도학습(supervised learning) 절차는 Fig. 3과 같다. 입력 시계열자료들은 인공지능 모델의 가중치와 활성화함수에 의해 모의지하수위 시계열자료로 변환되며, 최적화기(optimizer)를 사용하여 모의 지하수위 시계열자료가 관측지하수위 시계열자료와 가장 근사하도록 가중치를 업데이트하는 과정을 반복한다. 본 연구에서는 R 언어 기반 딥러닝 프레임워크인 Keras 패키지(Falbel *et al.*, 2019)에서 제공하는 ANN 및 LSTM 모델을 사용하였다.

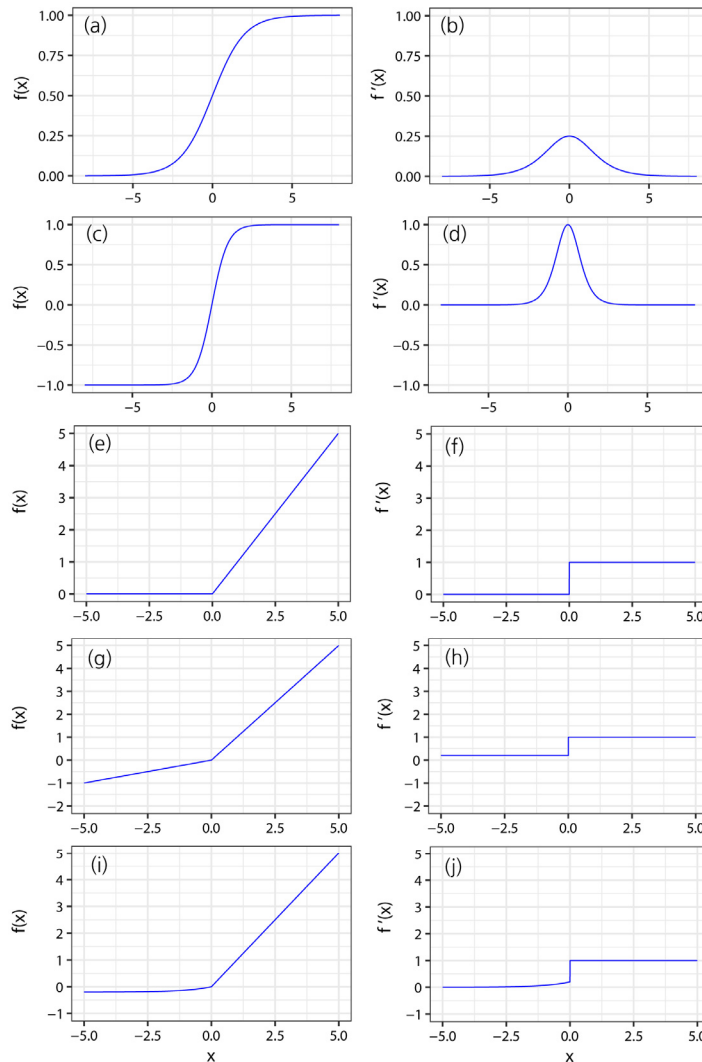
### 2.2 활성화함수

활성화함수는 뉴런에 수신되는 모든 입력신호를 더한 후 출력할 때 출력의 크기를 결정하는 압축함수(compressing function)의 역할을 하며(Sahoo *et al.*, 2017) 인공지능 모델의 학습 성능에 큰 영향을 미친다. 본 연구에서는 ANN 모델의 은닉층에 있는 은닉뉴런에 sigmoid, tanh, ReLU, Leaky ReLU 및 ELU의 5개 활성화함수를 적용하였다(Table 2, Fig. 4).

sigmoid 함수는 [0, 1]의 범위를 가지는 단조롭고 감소하지 않는 비선형함수이다(Fig. 4(a)). 매개변수인 가중치를 업데이트하기 위해 사용되는 역전파 알고리즘은 편미분을 활용한 경사하강법을 사용하며, 입력값이 매우 크거나 매우 작으면 활성화함수의 미분 값은 0에 근사하게 된다(Fig. 4(b)). 또한

**Table 2.** Equations of activation function and differentiation of the activation function

Activation function	Activation function equation	Differentiation of activation function equation
sigmoid	$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = \frac{e^{-x}}{(1 + e^{-x})^2}$
tanh	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$f'(x) = 1 - \left(\frac{e^x - e^{-x}}{e^x + e^{-x}}\right)^2$
ReLU	$f(x) = \max(0, x)$	$f'(x) = \begin{cases} 1 & (x \geq 0) \\ 0 & (x < 0) \end{cases}$
Leaky ReLU	$f(x) = \max(\alpha x, x)$	$f'(x) = \begin{cases} 1 & (x \geq 0) \\ \alpha & (x < 0) \end{cases}$
ELU	$f(x) = \begin{cases} x & (x \geq 0) \\ \alpha(e^x - 1) & (x < 0) \end{cases}$	$f'(x) = \begin{cases} 1 & (x \geq 0) \\ \alpha e^x & (x < 0) \end{cases}$



**Fig. 4.** Graph of activation functions: (a) sigmoid function, (b) differentiation of sigmoid function, (c) tanh function, (d) differentiation of tanh function, (e) ReLU function, (f) differentiation of ReLU function, (g) Leaky ReLU function ( $\alpha = 0.2$ ), (h) differentiation of Leaky ReLU function ( $\alpha = 0.2$ ), (i) ELU function ( $\alpha = 0.2$ ), (j) differentiation of ELU function ( $\alpha = 0.2$ )

은닉층수가 늘어나 0.25 보다 작은 편미분 값을 여러번 곱하게 되면 제일 상단 은닉층의 편미분 값은 0에 가까워지게 된다. 이러한 경우 매개변수값이 개선되지 않는 경사소멸 문제가 발생한다. 따라서 sigmoid 함수는 은닉뉴런의 활성화 함수로 사용하는데 한계가 있다. 역전파 알고리즘에 대한 설명은 Kim *et al.* (2021a)의 연구를 참고할 수 있다.

tanh 함수는 [-1, 1]의 범위를 가지며 sigmoid 함수와 유사한 곡선의 형태를 나타낸다(Fig. 4(c)). tanh 함수는 sigmoid 보다 넓은 범위의 함수 값 및 함수의 미분 값을 갖는다(Fig. 4(d)). 하지만 tanh 함수도 sigmoid 함수처럼 입력값이 매우 크거나 매우 작으면 활성화함수의 편미분 값은 0에 근사하게 되어 경사소멸 문제가 발생한다. 그리고 sigmoid 함수와 마찬가지로 tanh 함수에 지수함수가 포함되어 있으므로 계산속도가 느릴 수 있다.

ReLU 함수는 입력값이 0보다 작으면 0의 값을, 입력값이 0보다 크면 입력값과 동일한 값을 출력하는 활성화함수이다(Fig. 4(e)). ReLU 함수의 미분 값은 입력값이 0보다 작으면 0의 값을, 입력값이 0보다 크면 1의 값을 갖는다(Fig. 4(f)). 따라서 입력값이 0보다 큰 경우 기울기가 1로 일정하기 때문에 역전파 연산과정 시 경사소멸 문제가 발생하지 않는다. 또한 함수의 형태가 단순하여 지수함수를 포함하는 sigmoid 함수와 tanh 함수보다 연산속도가 빠르다는 장점이 있다. 하지만 입력값이 0보다 작은 경우 기울기가 0이 되기 때문에 매개변수가 업데이트되지 않으며 또한 ReLU 함수값이 0이 되므로 해당 뉴런이 비활성화 되는 문제점(dying ReLU)이 발생하게 된다.

Leaky ReLU 함수는 입력값이 0보다 큰 경우 ReLU 함수와 동일한 결과를 출력하며, 입력값이 0보다 작으면 기울기인  $\alpha$ 에 입력값을 곱하여 활성화함수의 출력값을 결정한다(Table 2, Fig. 4(g)). 그리고 Leaky ReLU 함수의 미분 값은 입력값이 0보다 작으면  $\alpha$ 를 갖는다(Table 2, Fig. 4(h)). 따라서 이 활성화함수는 경사소멸 문제와 ReLU 함수에서 입력값이 0보다 작을 때 발생하는 뉴런의 비활성화 문제를 해결한다. Leaky ReLU 함수는 ReLU 함수를 개선한 활성화함수이며 ReLU

함수보다 나은 결과를 도출한다고 알려져있다(Xu *et al.*, 2015).

ELU 함수는 입력값이 0보다 큰 경우 ReLU 함수의 식과 같으며, 입력값이 0보다 작은 경우  $\alpha$ 에 지수함수를 곱하여 활성화함수의 출력값을 결정한다(Table 2, Fig. 4(i)). ELU 함수는 입력값이 0보다 작은 경우 활성화함수값이 0이 아니라는 점에서 Leaky ReLU 함수와 유사하지만 지수함수를 사용함으로써 활성화함수값이 비선형이라는 점에서 Leaky ReLU 함수와 차이가 있다(Fig. 4(j)). ELU 함수는 경사소멸 문제와 뉴런의 비활성화 문제가 발생하지 않으며, ReLU 함수나 Leaky ReLU 함수보다 뛰어난 성능을 나타낸다고 알려져 있다(Clevert *et al.*, 2016). 하지만 이 활성화함수는 지수함수를 포함하고 있기 때문에 ReLU 함수나 Leaky ReLU 함수보다 계산속도가 느리다.

2.3 연구대상 지역 및 자료

연구대상지역은 제주도 남동쪽 표선유역의 중산간에 위치한 2개 지하수위 관측정 지점이다(Fig. 5). 본 연구에서는 일단위의 관측정 지하수위 자료, 인근의 2개 강우관측소(성판악, 교래) 강수량자료 및 2개 지하수 취수정 취수량 자료를 사용하였다(Table 3). 성판악 강우관측소는 기상청(<http://www.weather.go.kr>)에서 운영중에 있는 자동기상관측소(Automatic Weather Station)이고, 교래 강우관측소는 제주도 재난안전대책본부(<http://bangjae.jeu119.go.kr>)에서 운영하고 있으며 웹사이트에서 강수량 자료를 다운받아 사용할 수 있다. 지하수위 및 취수량 자료는 제주특별자치도개발공사에서 관측 및 관리하고 있으며 공개적으로 제공되지 않는 자료이다.

성판악과 교래 강우관측소의 일단위 강수량의 경우, 성판악 강우관측소의 강수량이 교래 강우관측소의 강수량보다 전반적으로 많은 것을 확인할 수 있다(Fig. 6). 교래 강우관측소(El. 400 m, Fig. 5)보다 고지대에 위치하고 있는 성판악 강우관측소(El. 763 m, Fig. 5)는 산지효과로 인해 강수량이 더 많은 것으로 판단된다. 관측정 A (MW-A)의 지하수위 변동폭(22.6 m)은 관측정 B (MW-B)의 지하수위 변동폭(17.5 m)보다 약 5 m 크다(Fig. 7). 그리고 관측정 A의 표고는 관측정 B의 표고보다 낮으나(Fig. 5) 관측정 A의 지하수위는 관측정 B의

Table 3. Data period of rainfall station, groundwater withdrawal well and groundwater monitoring well

Classification	Station name	Data period	Remarks
Rainfall station	Seongpanak	1992. 01. 01. ~ 2020. 12. 31.	Precipitation (mm/day)
	Gyoraе	1992. 01. 01. ~ 2020. 12. 31.	
Pumping well	PW1	2001. 01. 01. ~ 2020. 12. 31.	Pumping rate (m <sup>3</sup> /day)
	PW2	2013. 07. 31. ~ 2020. 12. 31.	
Monitoring well	MW-A	2001. 02. 11. ~ 2020. 12. 31.	Groundwater level (m)
	MW-B	2012. 03. 13. ~ 2020. 12. 31.	

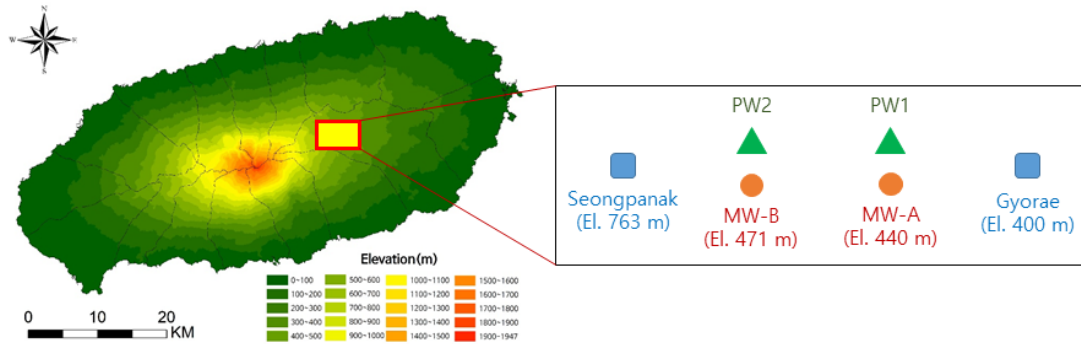


Fig. 5. Schematic diagram of the location of rainfall stations, groundwater withdrawal wells and groundwater monitoring wells

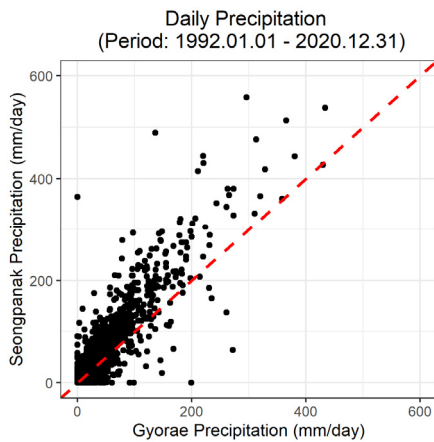


Fig. 6. Comparison of precipitation data from Seongpanak and Gyorae rainfall stations

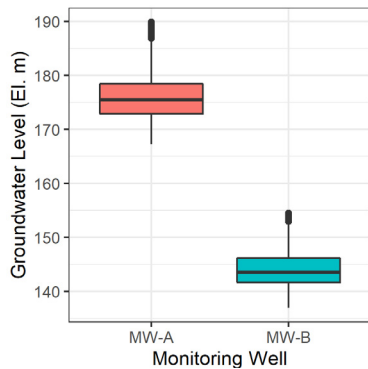


Fig. 7. Variation range of the groundwater level in the groundwater monitoring wells

지하수위보다 높게 형성되어 있다(Fig. 7). 이러한 현상은 관정별 복잡한 지하지질의 차이가 기인한 것으로 판단된다.

### 2.4 연구방법

ANN 모델에 5개 활성화함수를 각각 적용하여 지하수위 변동폭이 다른 2개 지하수위 관측정의 1일 후 지하수위를 예측하고 예측성능을 비교분석 하였다. 또한 LSTM 모델을 참고 모델로 사용하여 1일 후 지하수위를 예측하고 ANN 모델의 결과와 비교분석 하였다. 1일 후 지하수위를 예측한 이유는 인공지능 모델의 1개 스텝(time step) 미래 예측성능이 가장 좋아 각 인공지능 모델의 예측특성을 적절히 반영하기 때문이다. 인공지능 모델의 학습기간에 대한 과적합(overfitting)을 방지하기 위해 매개변수 보정과과정 중에 콜백(callback) 기능을 사용하여 추정된 매개변수를 검증기간에 대해 검증하였다. 그리고 split sample test (Klemeš, 1986)를 위하여 테스트 기간에 대해 인공지능 모델의 예측성능을 평가하였다. 학습, 검증 및 테스트 기간은 서로 독립된 기간으로써 Table 4와 같다.

인공지능 모델의 학습을 위해 가중치 이외의 매개변수인 하이퍼매개변수(hyper-parameter) 값의 설정이 필요하다. 하이퍼매개변수 값의 설정방법에 대한 명확한 기준은 없으며 (Chollet and Allaire, 2018) 본 연구에서는 시행착오 후 Table 5와 같이 설정하였다. 각 은닉층의 은닉뉴런의 개수( $n_{units}$ )는 100개로 설정하였으며 ANN 및 LSTM 모델 모두 2개의 은닉층을 사용하였다. 은닉뉴런이 많을수록 학습기간에 대한 과적합 문제가 발생할 수 있으나 본 연구에서는 이 문제를

Table 4. Duration of training, validation and testing of artificial intelligence models<sup>a</sup>

Monitoring Well	Training Period	Validation Period	Testing Period
MW-A	2007. 01. 01. ~ 2020. 12. 31.	2004. 01. 01. ~ 2006. 12. 31.	2001. 02. 11. ~ 2003. 12. 31.
MW-B	2016. 01. 01. ~ 2020. 12. 31.	2014. 01. 01. ~ 2015. 12. 31.	2012. 03. 13. ~ 2013. 12. 31.

<sup>a</sup>The training, validation, and testing period of the precipitation and groundwater withdrawal used for simulating the groundwater level of each groundwater monitoring well is the same as the training, validation and testing period of each groundwater monitoring well.

Table 5. Hyper-parameters of artificial intelligence models

Hyper-Parameter	Range	Setting Value	Description
n_units	-	100	Number of hidden units in hidden layer
batch_size	-	50	Number of samples fed to LSTM in one sub-simulation
dropout	0 - 1	0.5	Fraction of the units to drop for the linear transformation of the inputs
recurrent_dropout	0 - 1	0.5	Fraction of the units to drop for the linear transformation of the recurrent state in LSTM
learning_rate	float >= 1	0.001	Learning rate of Adam optimizer
n_epochs	-	50	Number of iterations
patience	-	10	Number of epochs for early termination of training when simulation values do not improve
$\alpha$	0.0001 - 1	0.2	Gradient of Leaky ReLU or ELU activation function

해결하기 위해 드롭아웃(dropout)과 콜백 기능을 사용하기 때문에 충분히 많은 수의 은닉뉴런을 사용하였다. 학습을 위해 한번 처리되는 소규모 자료의 개수(batch\_size)는 일단위 데이터의 길이를 고려하여 50개로 설정하였으며, 과적합을 방지하기 위해 무작위로 은닉뉴런을 비활성화 시키는 비율인 드롭아웃과 재귀적 드롭아웃(recurrent\_dropout)은 모두 0.5로 설정하였다. 인공지능의 학습을 위한 최적화기는 최근의 딥러닝 분야에서 널리 사용되고 있는(Le *et al.*, 2019) Adam (Kingma and Ba, 2014)을 사용하였다. 이 최적화기의 학습률(learning\_rate)은 0.001로 설정하였으며 매개변수의 최적화에 대한 목적함수는 평균절대오차(mean absolute error)를 사용하였다. 매개변수 최적화 과정 중 전체 데이터에 대해 매개변수 업데이트 최대 반복 횟수인 n\_epochs는 50번으로 설정하였다. 인공지능 모델의 과적합을 방지하기 위해 임의로 설정한 반복횟수(patience)만큼 매개변수를 업데이트 하더라도 검증기간의 모의결과가 더 이상 개선되지 않을 때 학습을 조기종료(early-stopping) (Prechelt, 2012) 하는 콜백 기능의 patience는 10으로 설정하였다. ANN의 은닉층에 적용된 Leaky ReLU 함수와 ELU 함수의  $\alpha$  값은 0.2로 설정하였다. 마지막으로 ANN 모델과 LSTM 모델의 출력층에 대한 활성화 함수는 Maier and Dandy (2000)의 제안에 따라 선형함수를 사용하였으며, 이 함수는 출력층의 뉴런에 수신되는 입력신호를 단순히 그대로 출력신호로 넘기는 역할을 한다.

인공지능 모델의 예측능력을 비교분석하기 위해 수문학 분야에서 널리 사용되고 있는 평가지수인 Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970)와 Root Mean Square Error (RMSE)를 사용하였다. NSE는 모의결과에 대한 전반적인 정보를 제공하며(Moriasi *et al.*, 2007) RMSE는 모의값이 관측값을 얼마나 가까이 일치시키는지 나타낸다(Le *et al.*, 2019). NSE와 RMSE의 정의는 아래와 같다.

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2} \quad (8)$$

여기에서  $n$ 은 시계열자료의 개수,  $Q_{obs,i}$ 와  $Q_{sim,i}$ 는 각각  $i$  시간대의 관측 및 모의 지하수위,  $\overline{Q_{obs}}$ 는 관측 지하수위의 평균값이다. NSE의 범위는  $-\infty$ 에서 1이며 1은 모의값과 관측값이 정확히 일치하는 것을 의미하고, 0은 모의값이 관측값의 평균값과 동일함을 의미한다. 그리고 RMSE가 0일 때 관측값과 모의값이 정확히 일치하는 것을 의미한다.

### 3. 결과 및 고찰

#### 3.1 활성화함수 비교 및 ANN 모델과 LSTM 모델 비교

5개의 활성화함수를 적용한 ANN 모델의 1일 후 지하수위 예측결과는 Tables 6 and 7과 같다. Tables 6 and 7은 각각 관측 정 A와 B에 대한 결과이며 기준이 되는 모델의 결과와 비교를 위해 LSTM 모델의 결과 또한 제시하였다. 상세한 비교를 위해 NSE와 RMSE의 적절성 통계지수를 학습(training)기간과 검증 및 테스트(validation+testing)기간의 결과로 분리하여 나타내었다. 주의할 점으로, 인공지능 모델의 학습(training) 시 검증(validation)기간을 사용하여 모델의 과적합(overfitting)을 방지한 후 추정된 매개변수를 사용하여 테스트(testing)기간에 대한 모델의 성능을 평가하였다. 하지만 검증 기간과 테스트기간의 모의결과는 모두 학습을 통해 도출된 매개변수를 활용하여 학습 이외의 기간에 대해 인공지능 모델



**Table 6.** Results of activation functions for groundwater monitoring well A

Model	NSE		RMSE	
	Training	Validation +Testing	Training	Validation +Testing
ANN-ELU	0.998	0.998	0.182	0.149
ANN-Leaky ReLU	0.998	0.998	0.181	0.162
ANN-ReLU	0.997	0.997	0.227	0.181
ANN-tanh	0.996	0.997	0.292	0.175
ANN-sigmoid	0.993	0.997	0.357	0.170
LSTM	0.998	0.998	0.178	0.130

**Table 7.** Results of activation functions for groundwater monitoring well B

Model	NSE		RMSE	
	Training	Validation +Testing	Training	Validation +Testing
ANN-ELU	0.997	0.997	0.179	0.155
ANN-Leaky ReLU	0.997	0.997	0.167	0.139
ANN-ReLU	0.997	0.996	0.196	0.163
ANN-tanh	0.996	0.997	0.199	0.132
ANN-sigmoid	0.994	0.997	0.267	0.139
LSTM	0.998	0.998	0.158	0.124

의 성능을 검증한 결과이므로, 결과비교의 편리성을 위해 검증기간과 테스트기간의 모의시계열 결과를 합하여(validation +testing) 적절성 통계지수를 계산한 결과를 Tables 6 and 7에 나타내었다.

지하수위 변동폭이 상대적으로 큰 관측정 A에 대한 지하수위 예측성능은 ELU 함수 사용 시 가장 높은 것으로 나타났으며 학습기간 기준 Leaky ReLU > ReLU > tanh > sigmoid 순으로 예측성능이 높은 것으로 나타났다. ELU 함수는 Leaky ReLU 보다 학습기간의 RMSE가 약간 크지만(0.182 m) 검증 및 테스트 기간에 대해 작은 RMSE (0.149 m) 결과를 도출함으로 인해 관측정 A에 대한 지하수위 예측에 대해 가장 적절한 활성화함수로 판단된다. 또한 비록 ANN-ELU 모델은 LSTM 모델보다 RMSE가 커 상대적으로 예측성능이 낮지만 NSE를 고려할 경우 LSTM 모델과 대등한 예측성능을 나타내었다.

지하수위 변동폭이 상대적으로 작은 관측정 B에 대한 지하수위 예측성능은 Leaky ReLU 함수 사용 시 가장 높은 것으로 나타났으며 학습기간 기준 ELU > ReLU > tanh > sigmoid 순으로 예측성능이 높은 것으로 나타났다. 따라서 Leaky ReLU 가 관측정 B에 대한 지하수위 예측에 대해 가장 적절한 활성화함수로 판단된다. ANN-Leaky ReLU 모델은 LSTM 모델보

다 NSE가 작고 RMSE가 커 상대적으로 예측성능이 낮지만 그 차이가 크지 않아 LSTM 모델과 대등한 예측성능을 보인다고 판단된다. 또한 이 두 가지 경우의 결과를 종합하여 볼 때 LSTM 모델은 5개의 활성화 함수를 사용한 ANN 모델보다 예측성능이 대등하거나 높아 인공지능 모델의 예측성능 비교 분석 시 참고(reference) 모델로 활용될 수 있다. 참고로, 관측정 A와 B에 대한 모든 인공지능 모델의 지하수위 예측성능은 NSE가 0.99 이상을 보여 높은 예측성능을 나타내었다.

흥미로운 점은 검증 및 테스트 기간에 대한 예측성능이 학습기간의 예측성능보다 높다는 점이다. 그 이유는 학습기간 선택 시 관측지하수위의 변동폭이 상대적으로 크고 변동양상이 상대적으로 복잡한 기간을 선택했기 때문에 학습기간에 대한 예측오차가 검증 및 테스트 기간에 대한 예측보다 큰 것으로 판단된다. 이것은 학습기간의 예측성능이 검증 및 테스트 기간에 대한 예측성능보다 항상 좋은 것은 아니라는 것을 의미한다. 또한 학습기간에 대한 예측결과를 기준으로 활성화함수들의 NSE 및 RMSE의 차이를 분석한 결과 관측정 A에 대한 차이(NSE: 0.998 - 0.993 = 0.005; RMSE: 0.357 - 0.181 = 0.176 m)가 관측정 B에 대한 차이(NSE: 0.997 - 0.994 = 0.003; RMSE: 0.267 - 0.167 = 0.1 m)보다 컸다. 즉 관측정 A에 대한 NSE와 RMSE의 차이는 관측정 B에 대한 경우보다 각각 67% 및 76% 가 컸다. 이것은 관측지하수위의 변동폭이 크고 변동양상이 복잡할수록 인공지능 모델별 지하수위 예측능력의 차이는 커진다는 것을 의미한다.

### 3.2 최적 활성화 함수와 sigmoid 함수 비교

Tables 6 and 7에서 보여준 바와 같이 활성화함수 별 ANN 모델의 지하수위 예측성능의 차이는 학습기간에 가장 크게 나타났다. 따라서 학습기간에 대한 지하수위 예측결과 중 가장 큰 차이를 보이는 두 ANN 모델의 결과를 Figs. 8 and 9와 같이 도시하였다. 관측지하수위 변동폭이 상대적으로 큰 관측정 A의 경우 ANN-sigmoid 모델은 ANN-ELU 모델에 비해 침투 및 최저 지하수위 예측성능이 낮은 것을 확인할 수 있다(Fig. 8). 이러한 현상은 관측지하수위 변동폭이 상대적으로 작은 관측정 B의 경우에서도 확인할 수 있는데 ANN-Leaky ReLU 모델은 고 및 저 지하수위 모두 관측 및 모의 지하수위의 1:1 선에 근사하지만 ANN-sigmoid 모델의 경우 고 지하수위는 과소예측, 저 지하수위는 과대예측을 하는 것으로 나타났다(Fig. 9). 지하수자원의 관리측면을 고려할 경우 최고 지하수위 보다 최저 지하수위의 정확한 예측이 중요하다. 또한 최저 지하수위의 예측의 경우 과대예측 보다 과소예측이 나옴에 그 이유는 최저 지하수위의 과소예측이 가뭄 대비에 효과적이

Monitoring Well A

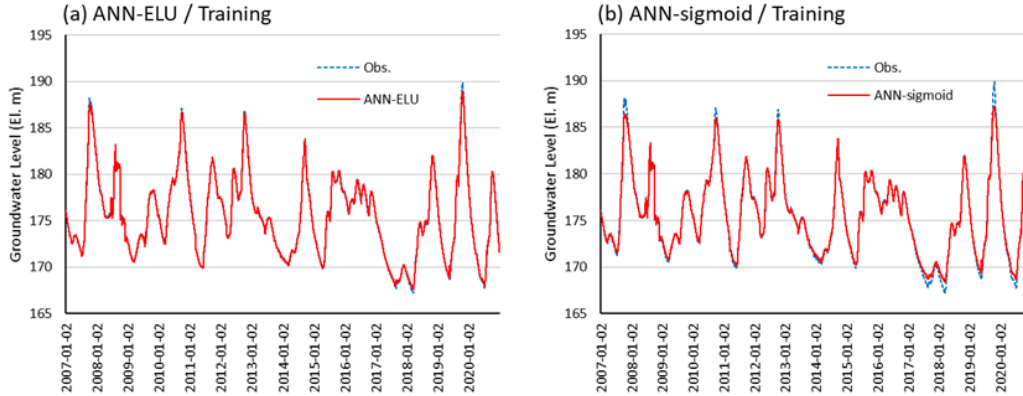


Fig. 8. Comparison of observed groundwater level and simulated groundwater level of ANN-ELU and ANN-sigmoid for the training period of monitoring well A

Monitoring Well B

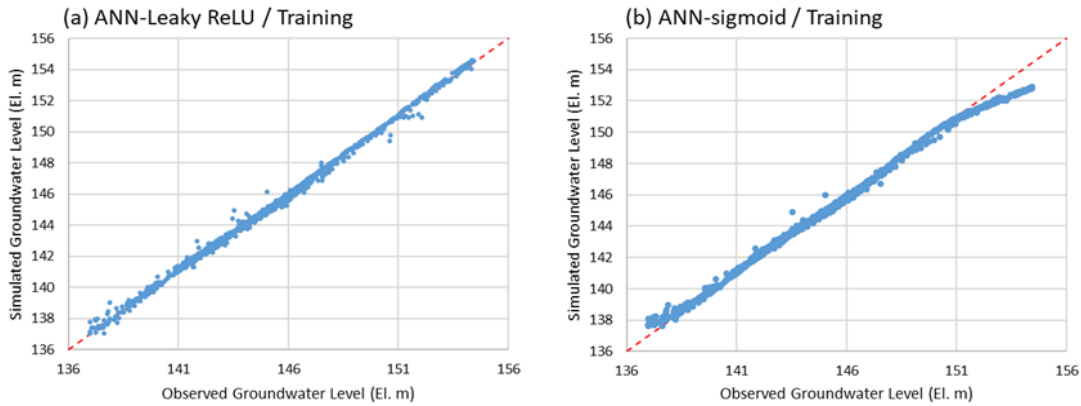


Fig. 9. Comparison of observed groundwater level and simulated groundwater level of ANN-leaky ReLU and ANN-sigmoid for the training period of monitoring well B

기 때문이다. 본 대상지역의 경우 sigmoid 함수는 최저 지하수위에 대해 과대예측을 하므로 사용에 주의가 필요하며 따라서 ANN 모델을 활용한 지하수위 예측에 ELU 또는 Leaky ReLU 함수를 사용하는 것이 적절하다.

4. 결론

본 연구에서는 지하수위의 변동폭이 크고 변동양상이 복잡한 제주도 표선유역 중산간지역 2개 지하수위 관측정을 대상으로 적절한 지하수위 예측을 위해 5개의 활성화함수를 ANN 모델에 적용하여 지하수위 예측결과를 비교 및 분석하고 최적 활성화함수를 도출하였다. 사용된 활성화함수들은 ELU, Leaky ReLU, ReLU, tanh, sigmoid 함수들이며 최근 널리 사

용되고 있는 순환신경망 모델인 LSTM 모델의 결과와 비교분석하였다. 분석결과 지하수위 변동폭이 상대적으로 큰 관측정과 상대적으로 작은 관측정에 대한 지하수위 예측성능은 각각 ELU와 Leaky ReLU 함수가 최적의 활성화함수로 도출되었다. 반면 sigmoid 함수는 학습기간에 대해 5개 활성화함수 중 예측성능이 가장 낮았으며 침투 및 최저 지하수위 예측에서 적절하지 못한 결과를 도출하였다. 따라서 ANN-sigmoid 모델은 가뭄기간의 지하수위 예측을 통한 지하수자원 관리목적으로 사용할 경우 주의가 필요하다. ANN-ELU와 ANN-Leaky ReLU 모델은 LSTM 모델과 대등한 지하수위 예측성능을 보여 활용가능성이 충분히 있으며 LSTM 모델은 ANN 모델들보다 예측성능이 높아 인공지능 모델의 예측성능 비교분석 시 참고 모델로 활용될 수 있다. 또한 학습기간의 정보량에 따라 학습기간의 지하수위 예측성능이 검증 및 테스트 기간의 예측

성능보다 낮을 수 있다는 것을 확인하였으며, 관측지하수위의 변동폭이 크고 변동양상이 복잡할수록 인공지능 모델별 지하수위 예측능력의 차이는 커졌다. 본 연구에서 제시한 5개의 활성화함수를 적용한 연구방법 및 비교분석 결과는 지하수위 예측뿐만 아니라 일단위 하천유출량 및 시간단위 홍수량 등 지표수 예측을 포함한 다양한 연구에 유용하게 사용될 수 있다.

## References

- Afzaal, H., Farooque, A.A., Abbas, F., Acharya, B., and Esau, T. (2020). "Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning." *Water*, Vol. 12, No. 1, p. 5.
- Barthel, R., and Banzhaf, S. (2016). "Groundwater and surface water interaction at the regional-scale - a review with focus on regional integrated models." *Water Resources Management*, Vol. 30, No. 1, pp. 1-32.
- Bengio, Y., Simard, P., and Frasconi, P. (1994). "Learning long-term dependencies with gradient descent is difficult." *IEEE Transactions on Neural Networks*, Vol. 5, No. 2, pp. 157-166.
- Chang, J., Wang, G., and Mao, T. (2015). "Simulation and prediction of suprapermafrost groundwater level variation in response to climate change using a neural network model." *Journal of Hydrology*, Vol. 529, pp. 1211-1220.
- Chollet, F., Allaire, J.J. (2018). *Deep learning with R*, Manning Publications, Shelter Island, NY, U.S., p. 360.
- Clevert, D.A., Unterthiner, T., and Hochreiter, S. (2016). "Fast and accurate deep network learning by exponential linear units (ELUs)." *arXiv preprint arXiv:1511.07289*.
- Coulibaly, P., Anctil, F., Aravena, R., and Bobée, B. (2001). "Artificial neural network modeling of water table depth fluctuations." *Water Resources Research*, Vol. 37, No. 4, pp. 885-896.
- Emamgholizadeh, S., Moslemi, K., and Karami, G. (2014). "Prediction the groundwater level of bastam plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS)." *Water Resources Management*, Vol. 28, No. 15, pp. 5433-5446.
- Falbel, D., Allaire, J.J., Chollet, F., Tang, Y., Van Der Bijl, W., Studer, M., Keydana, S. (2019). R interface to 'Keras'. R package version 2.2.4.1, accessed on 5 April 2019, <<https://CRAN.R-project.org/package=keras>>.
- Hahnloser, R.H., Sarpeshkar, R., Mahowald, M.A., Douglas, R.J., and Seung, H.S. (2000). "Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit." *Nature*, Vol. 405, No. 6789, pp. 947-951.
- Haykin, S. (2009). *Neural networks and learning machines*, Pearson Prentice Hall, Upper Saddle River, NJ, U.S.
- Hochreiter, S., and Schmidhuber, J. (1997). "Long short-term memory." *Neural Computation*, Vol. 9, No. 8, pp. 1735-1780.
- Hosseini, Z., Gharechelou, S., Nakhaei, M., and Gharechelou, S. (2016). "Optimal design of BP algorithm by ACO<sub>R</sub> model for groundwater-level forecasting: A case study on Shabestar plain, Iran." *Arabian Journal of Geosciences*, Vol. 9, No. 6, p. 436.
- Jeju Special Self-Governing Province (JSSGP) (2018). *Comprehensive water resources management plan in Jeju Island*. pp. 1-328.
- Jeong, J., and Park, E. (2019). "Comparative applications of data-driven models representing water table fluctuations." *Journal of Hydrology*, Vol. 572, pp. 261-273.
- Jha, M.K., and Sahoo, S. (2014). "Efficacy of neural network and genetic algorithm techniques in simulating spatio-temporal fluctuations of groundwater." *Hydrological Processes*, Vol. 29, No. 5, pp. 671-691.
- Kim, G.B., and Oh, D.H. (2018). "Determination of the groundwater yield of horizontal wells using an artificial neural network model incorporating riverside groundwater level data." *The Journal of Engineering Geology*, Vol. 28, No. 4, pp. 583-592.
- Kim, J., Jun, S.M., Hwang, S., Kim, H.K., Heo, J., and Kang, M.S. (2021a). "Impact of activation functions on flood forecasting model based on artificial neural networks." *Journal of The Korean Society of Agricultural Engineers*, Vol. 63, No. 1, pp. 11-25.
- Kim, M., Choi, J.Y., Bang, J., Yoon, P., and Kim, K. (2021b). "Comparison of artificial neural network model capability for runoff estimation about activation functions." *Journal of The Korean Society of Agricultural Engineers*, Vol. 63, No. 1, pp. 103-116.
- Kim, T.W., and Valdés, J.B. (2003). "Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks." *Journal of Hydrologic Engineering*, Vol. 8, No. 6, pp. 319-328.
- Kingma, D.P., and Ba, J. (2014). "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980*.
- Klemeš, V. (1986). "Operational testing of hydrological simulation models." *Hydrological Sciences Journal*, Vol. 31, No. 1, pp. 13-24.
- Krishna, B., Satyaji Rao, Y.R., and Vijaya, T. (2008). "Modelling groundwater levels in an urban coastal aquifer using artificial neural networks." *Hydrological Processes*, Vol. 22, No. 8, pp. 1180-1188.
- Le, X.H., Ho, H.V., Lee, G., and Jung, S. (2019). "Application of long short-term memory (LSTM) neural network for flood forecasting." *Water*, Vol. 11, No. 7, p. 1387.
- Lee, S., Lee, K.K., and Yoon, H. (2019). "Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors." *Hydrogeology Journal*, Vol. 27, No. 2, pp. 567-579.
- Maas, A.L., Hannun, A.Y., and Ng, A.Y. (2013). "Rectifier nonlinearities improve neural network acoustic models." *Proceedings of the 30th International Conference on Machine Learning*, Atlanta, GA, U.S., Vol. 30, No. 1, p. 3.
- 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.
- Maxwell, R.M., Condon, L.E., and Kollet, S.J. (2015). "A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3." *Geoscientific Model Development*, Vol. 8, No. 3, pp. 923-937.
- McDonald, M.G., and Harbaugh, A.W. (1988). *A modular three-dimensional finite-difference ground-water flow model. Vol. 6*. US Geological Survey, Reston, VA, U.S.
- Mohanty, S., Jha, M.K., Kumar, A., and Panda, D.K. (2013). "Comparative evaluation of numerical model and artificial neural network for simulating groundwater flow in Kathajodi - Surua Inter-basin of Odisha, India." *Journal of Hydrology*, Vol. 495, pp. 38-51.
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., and Veith, T.L. (2007). "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations." *Transactions of the ASABE*, Vol. 50, No. 3, pp. 885-900.
- Müller, J., Park, J., Sahu, R., Varadharajan, C., Arora, B., Faybishenko, B., and Agarwal, D. (2021). "Surrogate optimization of deep neural networks for groundwater predictions." *Journal of Global Optimization*, Vol. 81, No. 1, pp. 203-231.
- Nash, J.E., and Sutcliffe, J.V. (1970). "River flow forecasting through conceptual models part I - A discussion of principles." *Journal of Hydrology*, Vol. 10, No. 3, pp. 282-290.
- Nayak, P.C., Rao, Y.S., and Sudheer, K.P. (2006). "Groundwater level forecasting in a shallow aquifer using artificial neural network approach." *Water Resources Management*, Vol. 20, No. 1, pp. 77-90.
- Prechelt, L. (2012) "Early stopping - But when?." *Neural networks: Tricks of the trade*, Edited by Montavon G., Orr G.B., and Müller K.R., Springer, Berlin Heidelberg, pp. 53-67.
- Rajaei, T., Ebrahimi, H., and Nourani, V. (2019). "A review of the artificial intelligence methods in groundwater level modeling." *Journal of Hydrology*, Vol. 572, pp. 336-351.
- Rakhshandehroo, G.R., Vaghefi, M., and Aghbolaghi, M.A. (2012). "Forecasting groundwater level in Shiraz plain using artificial neural networks." *Arabian Journal for Science and Engineering*, Vol. 37, No. 7, pp. 1871-1883.
- Sahoo, S., and Jha, M.K. (2013). "Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment." *Hydrogeology Journal*, Vol. 21, No. 8, pp. 1865-1887.
- Sahoo, S., Russo, T.A., Elliott, J., and Foster, I. (2017). "Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US." *Water Resources Research*, Vol. 53, No. 5, pp. 3878-3895.
- Shin, M.J., Moon, S.H., Kang, K.G., Moon, D.C., and Koh, H.J. (2020). "Analysis of groundwater level variations caused by the changes in groundwater withdrawals using long short-term memory network." *Hydrology*, Vol. 7, No. 3, p. 64.
- Sit, M., Demiray, B.Z., Xiang, Z., Ewing, G.J., Sermet, Y., and Demir, I. (2020). "A comprehensive review of deep learning applications in hydrology and water resources." *Water Science and Technology*, Vol. 82, No. 12, pp. 2635-2670.
- Sun, Y., Wendi, D., Kim, D.E., and Liang, S.Y. (2016). "Application of artificial neural networks in groundwater table forecasting-a case study in a Singapore swamp forest." *Hydrology and Earth System Sciences*, Vol. 20, No. 4, pp. 1405-1412.
- Taormina, R., Chau, K.W., and Sethi, R. (2012). "Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon." *Engineering Applications of Artificial Intelligence*, Vol. 25, No. 8, pp. 1670-1676.
- Todd, D.K. and Larry, W.M. (2004). *Groundwater hydrology, Third edition*. John Wiley & Sons Inc., Hoboken, NJ, U.S., pp. 1-656.
- Ukkonen, P., and Mäkelä, A. (2019). "Evaluation of machine learning classifiers for predicting deep convection." *Journal of Advances in Modeling Earth Systems*, Vol. 11, No. 6, pp. 1784-1802.
- Vu, M.T., Jardani, A., Massei, N., and Fournier, M. (2021). "Reconstruction of missing groundwater level data by using Long Short-Term Memory (LSTM) deep neural network." *Journal of Hydrology*, Vol. 597, p. 125776.
- Wen, X., Feng, Q., Deo, R.C., Wu, M., and Si, J. (2017). "Wavelet analysis - artificial neural network conjunction models for multi-scale monthly groundwater level predicting in an arid inland river basin, northwestern China." *Hydrology Research*, Vol. 48, No. 6, pp. 1710-1729.
- White, J.T., Doherty, J.E., and Hughes, J.D. (2014). "Quantifying the predictive consequences of model error with linear subspace analysis." *Water Resources Research*, Vol. 50, No. 2, pp. 1152-1173.
- White, J.T., Knowling, M.J., and Moore, C.R. (2020). "Consequences of groundwater-Model vertical discretization in risk-Based decision-Making." *Groundwater*, Vol. 58, No. 5, pp. 695-709.
- Worland, S.C., Steinschneider, S., Asquith, W., Knight, R., and Wiczorek, M. (2019). "Prediction and inference of flow duration curves using multioutput neural networks." *Water Resources Research*, Vol. 55, No. 8, pp. 6850-6868.
- Xu, B., Wang, N., Chen, T., and Li, M. (2015). "Empirical evaluation of rectified activations in convolutional network." *arXiv preprint arXiv:1505.00853*.
- Yoon, H., Hyun, Y., Ha, K., Lee, K.K., and Kim, G.B. (2016). "A method to improve the stability and accuracy of ANN- and SVM-based time series models for long-term groundwater level predictions." *Computers & Geosciences*, Vol. 90, pp. 144-155.
- Yoon, H., Jun, S.C., Hyun, Y., Bae, G.O., and Lee, K.K. (2011). "A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer." *Journal of Hydrology*, Vol. 396, No. 1-2, pp. 128-138.
- Yu, H., Wen, X., Feng, Q., Deo, R.C., Si, J., and Wu, M. (2018). "Comparative study of hybrid-wavelet artificial intelligence models for monthly groundwater depth forecasting in extreme arid regions, Northwest China." *Water Resources Management*, Vol. 32, No. 1, pp. 301-323.