

지하수위 영향인자에 따른 인공신경망 기반의 지하수위 예측 성능 분석

Analysis of Groundwater Level Prediction Performance with Influencing Factors by Artificial Neural Network

김 인 철 ¹	Kim, Incheol	이 재 환 ²	Lee, Jaehwan
김 정 환 ³	Kim, Junghwan	이 형 규 ⁴	Lee, Hyoungkyu
이 준 환 ⁵	Lee, Junhwan		

Abstract

Groundwater level (GWL) causes the stress state within soil and affects the bearing capacity and the settlement of foundation. In this study, the analyses of influencing factors on GWL fluctuation were performed. From the results, river stage and moving average of precipitation were main influence components for urban near large river and rural areas, respectively. In addition, the prediction performance of GWL using artificial neural network (ANN) was conducted with respect to the influence components. As a result, the effect of main component was significant on the prediction performance of GWL.

요 지

지하수위 변동은 지반의 응력 상태에 변화를 일으켜 기초구조물의 지지력 및 침하에 직·간접적인 영향을 미칠 수 있다. 본 연구에서는 연구 대상지역을 선정하여 지하수위 영향인자 분석을 수행하였다. 그 결과 대상지역에 따라 지하수위에 미치는 영향인자들이 각각 달랐으며, 규모가 큰 하천변에 위치한 도심지역의 경우 하천수위가 지하수위 변동에 영향을 미치는 주요 인자였으며, 지표면 포장율이 낮은 도외지역의 경우는 선행강우를 고려하기 위해 도입된 강우이동평균이 주요 인자였다. 또한, 여러 입력 인자 조합을 고려하여 인공신경망을 통한 지하수위를 예측을 수행하였다. 분석결과 주요 지하수위 영향인자가 지하수위 예측 성능에 미치는 영향이 큰 것으로 나타났다. 결과적으로, 인공신경망을 이용하여 지하수위를 예측할때, 적절한 지하수위 영향인자 평가가 수행되어야 하며 이를 예측에 적용할 필요가 있는 것을 나타낸다.

Keywords : Artificial neural network, Groundwater level, Influencing factors, Moving average, Prediction performance

1 정희원, 네브라스카 주립대학교 연구위원 (Member, Research Associate, Dept. Civil and Environmental Engrg., Univ. of Nebraska-Lincoln)

2 정희원, 서울기술연구원 수석연구원 (Member, Senior Research Fellow, Dept. Urban Infrastructure Research, Seoul Institute of Technology)

3 정희원, 서울기술연구원 연구위원 (Member, Research Fellow, Dept. Smart City Research, Seoul Institute of Technology)

4 정희원, 서일대학교 토목공학과 정교수 (Member, Prof., Dept. Civil Engrg., Seoil Univ.)

5 정희원, 연세대학교 건설환경공학과 정교수 (Member, Prof., Dept. Civil and Environmental Engrg., Yonsei Univ., Tel: +82-2-2123-5796, Fax: +82-2-364-5300, junlee@yonsei.ac.kr, Corresponding author, 교신저자)

* 본 논문에 대한 토의를 원하는 회원은 2021년 11월 30일까지 그 내용을 학회로 보내주시기 바랍니다. 저자의 검토 내용과 함께 논문집에 게재하여 드립니다.

1. 서론

지하수위(Groundwater level, GWL)는 강우 발생, 하천수위 변동, 해수면 상·하강, 지하수 양수 등과 같은 자연적 및 인위적 요인에 의하여 주기적 또는 비주기적으로 변동할 수 있으며, 지하수위의 변동 양상은 지형학적(Geomorphological), 수리지질학적(Hydrogeologic) 특성 및 지반 이용 실태(land use)와 밀접한 관련이 있다(Hoque et al., 2007; Yasuhara et al., 2007; Kim et al., 2016). 지하수위변동은 지반의 응력상태(Stress state)에 직접적인 영향을 미치어 기초의 지지력을 감소 시키거나 설계단계에서 예상하지 못한 침하를 발생시킬 수 있다(Ausilio and Conte, 2005; Shaharia et al., 2014; Park et al., 2017; Park et al., 2019). 이는 결과적으로 전체 구조물의 안정성(Stability) 및 사용성(Serviceability)에 영향을 미칠 수 있으며, 구조물의 내구성 약화 및 기대 사용연한 감소로 이어질 수 있다. 따라서, 대상지역의 지하수위 변동을 사전에 평가하고, 그 영향을 기초구조물 설계 단계에서부터 반영하여 전체구조물의 운용에 대한 신뢰를 높일 필요가 있다.

지하수위를 예측하는 방법론은 크게 Knowledge-based approach와 Data-based approach로 구분 할 수 있다. Knowledge-based approach는 역학적 배경을 기반으로 하여 지하수 흐름을 수학적으로 설명하고 시뮬레이션 하는 방법이며, 대표적인 예로 analytical 및 numerical 모델들이 있다(Serrano and Workman, 1998; Batelaan et al., 2003; Kim and Lee, 2018a, 2019). 이 방법론에서는 일반적으로 지하수위 변동을 야기시키는 원인과 결과에 대한 설명이 명료하며 지형 및 지질학적 특성을 반영할 수 있지만, 계산 절차가 비교적 복잡하며 대수층의 투수특성치를 확보하는데 시간과 비용이 소요된다(Trichakis et al., 2011; Sahoo and Jha, 2013). Data-based approach는 관측된 데이터로부터의 영향인자들과 지하수위와의 통계적 상관성을 기반으로하여 지하수위를 예측하는 방법이며, 다중선형회귀법(Multiple linear regression), 인공신경망(Artificial neural network) 등을 예로 들 수 있다(Coppola et al., 2005; Sahoo and Jha, 2013). Data-based approach는 양질의 관측데이터를 필요로하지만, 비교적 간단한 절차로 신뢰할만한 예측 결과를 얻을 수 있기 때문에 많은 연구자들이 지하수위의 평가 및 예측에 이용하고 있다(Dliakopoulos et al., 2005; Suryanarayana et al., 2014; Rajae et al., 2019).

Data-based approach로써 다수의 모델들이 지하수위 예측에 사용되고 있다. 대표적인 예로 이동평균법(Moving average method, MA), 자기회귀이동평균(autoregressive moving average, ARMA), 다중선형회귀분석(Multiple linear regression, MLR), 인공신경망(artificial neural network, ANN), Support vector machine(SVM), adaptive neurofuzzy inference system approach(ANFIS) 등이 있다(Coppola et al., 2005; Daliakopoulos et al., 2005; Sahoo and Jha, 2013; Suryanarayana et al., 2014; Fallah-Mehdipour et al., 2013; Wen et al., 2015). 이 가운데 인공신경망은 입력값(input)과 결과(output)의 비선형성을 고려할 수 있는 모델로써, 지하수위를 예측하는데 가장 빈번하게 이용되고 있는 모델 중 하나이며 예측결과에 대한 신뢰성도 여러 연구자들로부터 입증되었다(Coppola et al., 2005; Daliakopoulos et al., 2005; Sahoo and Jha, 2013; Rajae et al., 2019).

Data-based approach에서는 학습 및 예측에 사용하는 데이터의 질과 양이 매우 중요하다. 지하수위를 예측함에 있어서 흔히 강우량, 하천수위, 기온 등이 입력값으로 사용되며, 1년간의 지하수위 변동을 예측하기 위하여 3~5년간 축적된 데이터를 학습한다(Sahoo and Jha, 2013; Kim and Lee, 2019). 최근에는 Wavelet decomposition 기술을 이용하여 입력값으로 사용되는 데이터를 정제하는 연구들도 진행되고 있다(Nourani and Mousavi, 2016; Rajae et al., 2019).

앞에서 언급한 것과 같이, 지하수위 변동은 여러 자연적, 인위적 요인으로 발생하고, 대상지역의 지질 및 지형적 특성, 지반이용실태 등에 따라 그 변동 특성이 다양할 수 있다. 이로 인하여 지하수위 변동에 영향을 미치는 주된 영향인자 또한 달라질 수 있다(Hoque et al., 2007; Kim et al., 2016). Kim et al.(2016)은 하천에 가까운 도심지역에서 주된 지하수위 영향인자는 하천수위이며, 지표면 포장율이 적은 도외지역의 경우 선행강우를 고려한 강우이동평균이 주된 영향인자일 수 있다고 주장했다. 이처럼 지하수위 변동에 미치는 영향인자가 지역별로 상이함에도 불구하고 대부분의 Data-based approach를 이용한 지하수위 예측 연구에서는 단순히 강우량, 하천수위, 기온 등의 데이터를 일괄적인 입력값으로 사용하고 있다. 따라서, 입력값으로 사용되는 기상 및 수문 인자들과 지하수위에 대한 상관성 분석을 수행함으로써 Data-based approach의 입력값과 예측결과 값에 대한 이해가 필요하다. 결과적으로 이러한 이해는

Data-based approach의 지하수위 예측 성능을 개선하는데 도움이 될 수 있을 것이다.

본 연구에서는, 지하수위 영향인자가 Data-based approach의 지하수위 예측에 미치는 영향을 분석하였다. 이를 위하여 서울시 마곡동, 광명시 철산동, 양평군 개군면을 연구 대상지역으로 선정하여, 강우량과 하천수위의 지하수위 변동에 대한 영향인자 분석을 수행하였다. 또한, Data-based approach 중 하나인 인공신경망을 이용하여 지하수위를 예측하여, 여러 영향인자 조합에 대한 지하수위 예측 결과를 비교 분석함으로써 지하수위 영향인자와 예측 성능과의 연관성을 고찰하였다.

2. 인공신경망(artificial neural network, ANN)과 지하수위 예측

2.1 인공신경망(ANN)

인공신경망은 Data-based approach중 한 모델로써, 지하수위 예측에 빈번하게 사용되어 왔다. 인공신경망의 원리는 생체신경세포들의 자극에 대한 정보 전달 원리를 모방한 것으로, Fig. 1과 같이 여러 종류의 입력값(input)에 각각의 가중치(w)를 적용하여 다수의 노드(node)를 가지고 있는 레이어(layer)로 전달한다. 각각의 노드에서는 가중치가 적용된 입력값들이 active function을 통하여 합산되어 결과값(output)이 된다. 일련의 과정은 다음과 같은 식으로 표현될 수 있다.

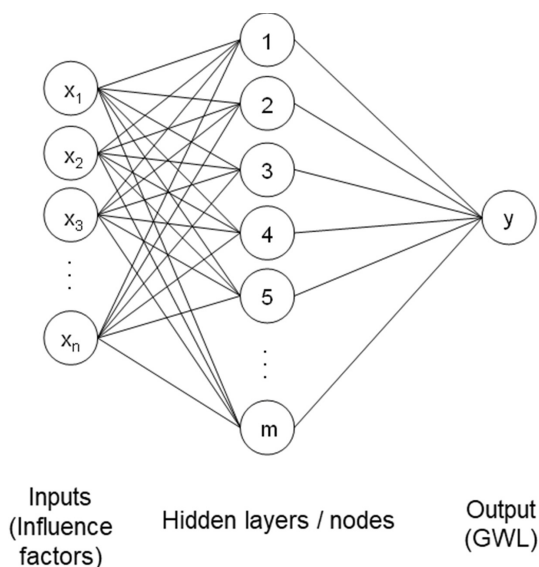


Fig. 1. Structure of artificial neural network (ANN)

$$y_i = F\left(\sum_{i=1}^n w_{ij}x_i + b\right) \quad (1)$$

여기서 x_i 와 y_i 는 각각 이전 layer의 i 번째와 현재 layer의 j 번째 node값이며, b_j 는 현재 layer의 j 번째 node의 편중 값, w_{ji} 는 x_i 와 y_i 의 연결강도(가중치), n 은 이전 layer의 node 개수, F 는 layer에 할당된 active function을 나타낸다. 얻어진 결과값은 관측된 값과의 오차가 충분히 작아질 때까지 Feed forward와 Back propagation을 반복하게 된다. 인공신경망에서는 Feed forward와 Back propagation을 위한 여러 알고리즘이 있으며, Gauss-Newton algorithm과 Gradient descent algorithm을 결합한 형태의 Levenberg-Marquardt(LM) algorithm 알고리즘이 지하수위 예측에 적절한 것으로 알려져 있다(Sahoo and Jha, 2013; Rajae et al., 2019).

인공신경망에서는 입력값과 결과값 사이의 최적 Layer와 Node 개수가 결정되어야 한다. Layer와 Node의 최적 개수는 예측하고자 하는 대상, 입력값, 결과값 특성에 따라 다르며, 일반적으로 테스트 단계에서 Trial and error를 통해 결정 할 수 있으며, Genetic algorithm(GA) 방법을 이용할 수도 있다(Gerken et al., 2006; Sahoo and Jha, 2013; Kim and Lee, 2018b).

일반적으로 인공신경망을 이용하여 지하수위 예측시 강우량, 하천수위, 해수면, 지하수 양수량 등이 입력값으로 사용되며, 경우에 따라서는 지하수위 예측에 도움이 될 수 있는 파라미터를 추가하기도 한다(Kim and Lee, 2018b; Daliakopoulos et al., 2005; Shahoo and Jha, 2013; Kim and Lee, 2019). 또한, 최근에는 Wavelet decomposition 기술을 이용하여 인풋값을 정제함으로써 지하수위 예측 성능을 향상시키는 연구도 다수 보고되고 있다(Nourani and Mousavi, 2016; Rajae et al., 2019).

2.2 지하수위 영향인자

앞에서 언급한 것과 같이 지하수위는 강우 발생, 하천수위 변동, 해수위 변동, 지하수 양수 등에 의해서 변동할 수 있다. 지하수위 변동 양상은 대상지역의 지질, 지형, 지표면 이용 현황에 따라 달라질 수 있다. Hoque et al.(2013)의 20년동안 관측된 지하수위 데이터에 의하면, 기후 및 기상 특성이 비슷할지라도 도심지역과 도외지역, 하천과의 접근성에 따라 지하수위 변동 특성이 다르다. 또한, Kim et al.(2016)은 하천변 지역을 대상으로

지하수위 영향인자 분석을 수행하였으며, 지표면 포장 정도, 하천과의 근접성, 하천의 규모에 따라 주된 영향 인자가 달라질 수 있다고 주장하였다. 또한, 모든 지역에서 일관측 강우량과 지하수위가 매우 낮은 상관관계를 보였지만, 선행강우를 고려한 강우이동평균과 지하수위는 높은 상관관계를 갖는 것을 보였다. 이처럼, 대상지역의 수문학적, 지질 및 지형학적 특성에 따라서 지하수위 주된 영향인자는 달라질 수 있으며 미치는 영향 정도 또한 달라질 수 있다. 그러나 대부분의 경우, 인공신경망을 이용한 지하수위 예측에서 영향인자에 대한 고찰 없이 데이터학습에 입력값을 일괄적으로 이용한다. 따라서, 지하수위 예측을 위한 사전 작업으로써 지하수위에 영향을 미치는 주된 인자 및 영향 정도 등을 분석한다면, 예측 결과에 대한 보다 높은 이해와 설명이 가능할 수 있다. 또한, 이를 바탕으로 인공신경망을 이용한 지하수위 예측 성능도 개선할 수 있을 것이다.

2.3 예측 성능 평가 인덱스

지하수위 예측 성능의 정량적 평가에 여러 가지 인덱스가 사용된다(Dawson and Wilby, 2001, Kim and Lee, 2019). 본 연구에서는 두 데이터의 상관성 정도를 나타내는 Pearson's correlation coefficient, PCC와 평균 에러를 나타내는 평균제곱근오차(Root mean square error, RMSE)를 예측 성능 평가 인덱스로 사용하였으며, 각각 다음과 같이 표현된다.

$$PCC = \frac{\sum_i^n (X_i - \bar{X}) \cdot \sum_i^n (Y_i - \bar{Y})}{\sqrt{\sum_i^n (X_i - \bar{X})^2} \cdot \sqrt{\sum_i^n (Y_i - \bar{Y})^2}} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_i^n (X_i - Y_i)^2}{n}} \quad (3)$$

여기서 X와 Y는 예측값과 관측값을 나타내며, \bar{X} 와 \bar{Y} 는 예측값 및 관측값의 평균을 나타낸다. n과 i는 각각 데이터 개수 및 데이터 순서를 나타낸다. PCC의 값이 1에 가까울수록 두 예측값과 관측값의 상관성이 높은 것을 나타내며, 0에 가까울수록 상관성이 적은 것을 나타낸다. 또한, RMSE의 값이 작을수록 예측값의 에러가 작다는 것을 나타낸다.

3. 지하수위 영향인자 분석

3.1 연구대상지역

본 연구에서는 서울시 마곡동(Seoul Magok, SM), 광명시 철산동(Gwangmyeong Cheolsan, GC), 양평군 개군면(Yangpyeong Gagun, YG) 지역을 연구 대상지역으로 선정하였다. 각 연구 대상지역의 위치는 Fig. 2와 같으며, 3지역 모두 하천에 인접해 있다. Fig. 3은 3연구대상지역의 지표면 상태를 나타낸다. 빨간색으로 표시된 부



Fig. 2. Location of study sites

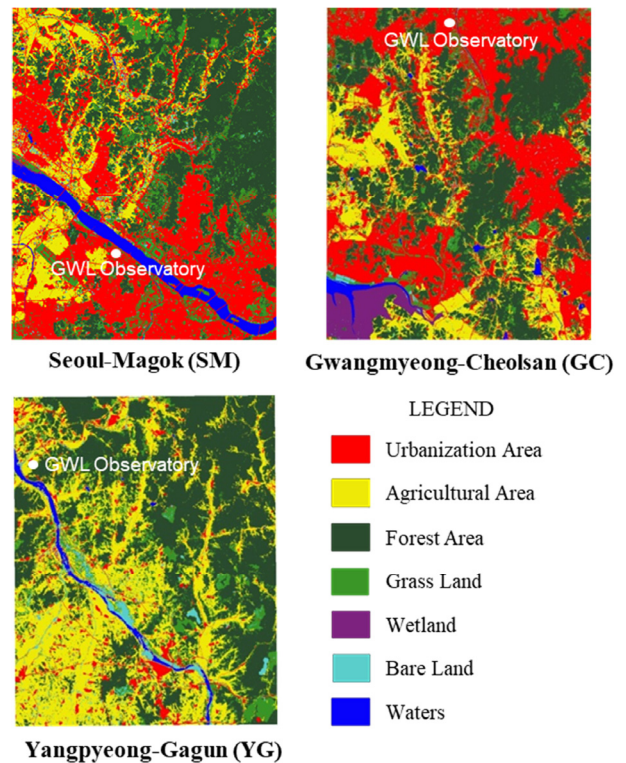


Fig. 3. Surface conditions of study sites

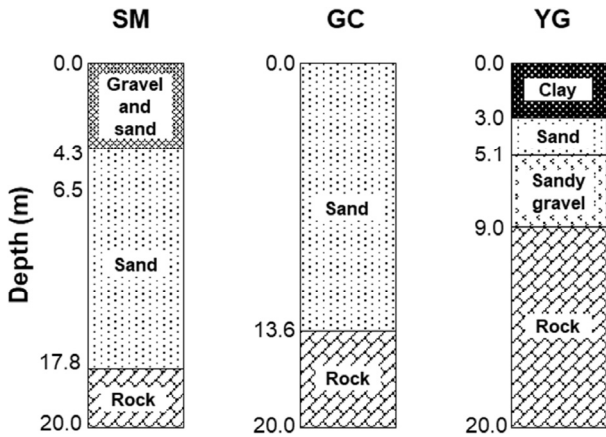


Fig. 4. Soil profile of study sites

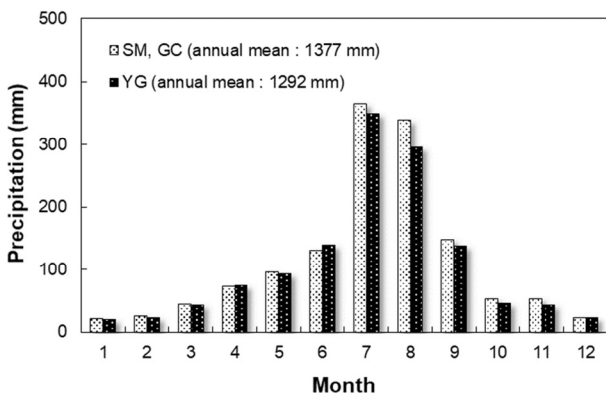


Fig. 5. Monthly mean precipitation observed over 30 years at study sites

분은 도시화된 지역으로서 강우가 직접 침투할 수 없는 불투수 지표면으로 분류된다. Fig. 3에서 확인할 수 있는 것과 같이 SM과 GC는 도심지역으로, YG는 도외지역으로 분류할 수 있다. Fig. 4는 연구 대상지역의 대표 지질 특성을 나타낸다. 대표 지질특성은 대상지역 지하수위 관측소 근방에서 관측된 다수의 지질주상도 자료들을 바탕으로 산정되었다. 3지역 모두 투수성이 높은 흙으로 구성되어 있으며, YG는 지표면이 낮은 투수성의 점토로 구성되어 있다. Fig. 5는 연구 대상지역에서 30년간 관측된 평균월강우량을 보여준다. 3지역 강우 특성이 매우 유사하며, 강우가 주로 7~8월에 집중된다. 연평균강우량은 SM과 GC가 1,377mm, YG가 1,292mm로서 SM과 GC가 YG보다 약간 많은 것을 확인할 수 있다.

3.2 관측 데이터

Fig. 6은 SM, GC, YG에서 2006년부터 2012년까지 7년간 관측된 지하수위, 하천수위, 강우량, 기온을 나타

낸다. 모든 인자들이 계절적 주기성을 가지며 변동하는 모습을 보인다. Fig. 6으로 부터 3지역의 강우량과 기온이 유사한 것을 확인할 수 있는 반면, 지하수위와 하천수위 변동은 지역마다 다른 것을 확인할 수 있다. 특히 지하수위의 경우, SM과 GC에서는 상승과 하강이 짧은 시간에 일어나지만, YG에서는 지하수위 상승과 하강에 비교적 많은 시간이 소요되는 것을 확인할 수 있다.

3.3 상관관계분석: 강우량 및 하천수위

Fig. 7은 3지역에서의 지하수위와 강우량 및 하천수위와의 상관관계를 나타낸다. 이때 사용한 데이터는 Fig. 6의 2006~2011년의 6년간 관측된 데이터를 사용하였다. SM, GC, YG의 지하수위와 강우량과의 상관계수(r)는 각각 0.416, 0.444, 0.225이며, 이 값들로부터 지하수위와 강우량과의 직접적인 상관관계는 낮다고 할 수 있다. SM, GC, YG의 지하수위와 하천수위와의 상관계수는 0.938, 0.624, 0.304이다. SM에서는 지하수위와 하천수위가 상당히 높은 상관관계를 보이나, YG의 경우는 지하수위와 하천수위의 직접적인 상관관계가 매우 낮다고 할 수 있다. GC의 경우, 상관계수가 0.6이상으로 상관관계가 높다고 할 수는 없지만 유의미한 상관성이 있다고 볼 수 있다(Yang et al., 2011).

3.4 지하수위 영향인자: 강우이동평균

강우 발생이 지하수위를 변동시키는 근본적인 원인이라고 생각하는 것이 상식적이다. 그러나 지하수위와 강우량의 직접적인 상관관계는 매우 낮다. 그 이유는 강우는 간헐적으로 발생하지만 지하수위의 경우는 연속적으로 측정되며, 발생한 강우가 지하수까지 도달하는데 시간차(Time lag)가 발생하기 때문이다. 이때의 시간차는 대상지역의 지형, 지표, 지질 특성에 영향을 받기 때문에, 지역 특성에 따라 다르게 나타난다.

이동평균(Moving average)은 간헐적인 데이터를 연속적 변동 값으로 변환 시킬 수 있으며, 시간차 성분 또한 반영할 수 있다. 이러한 특성을 강우량 데이터에 적용한 것을 강우이동평균으로 정의하며, 다음과 같은 식으로 표현할 수 있다(Kim and Lee, 2018c).

$$MA_{t,n} = \frac{PRE_t + PRE_{t-1} + \dots + PRE_{t-(n-1)}}{n} = \frac{1}{n} \sum_{i=0}^{n-1} PRE_{t-i} \quad (4)$$

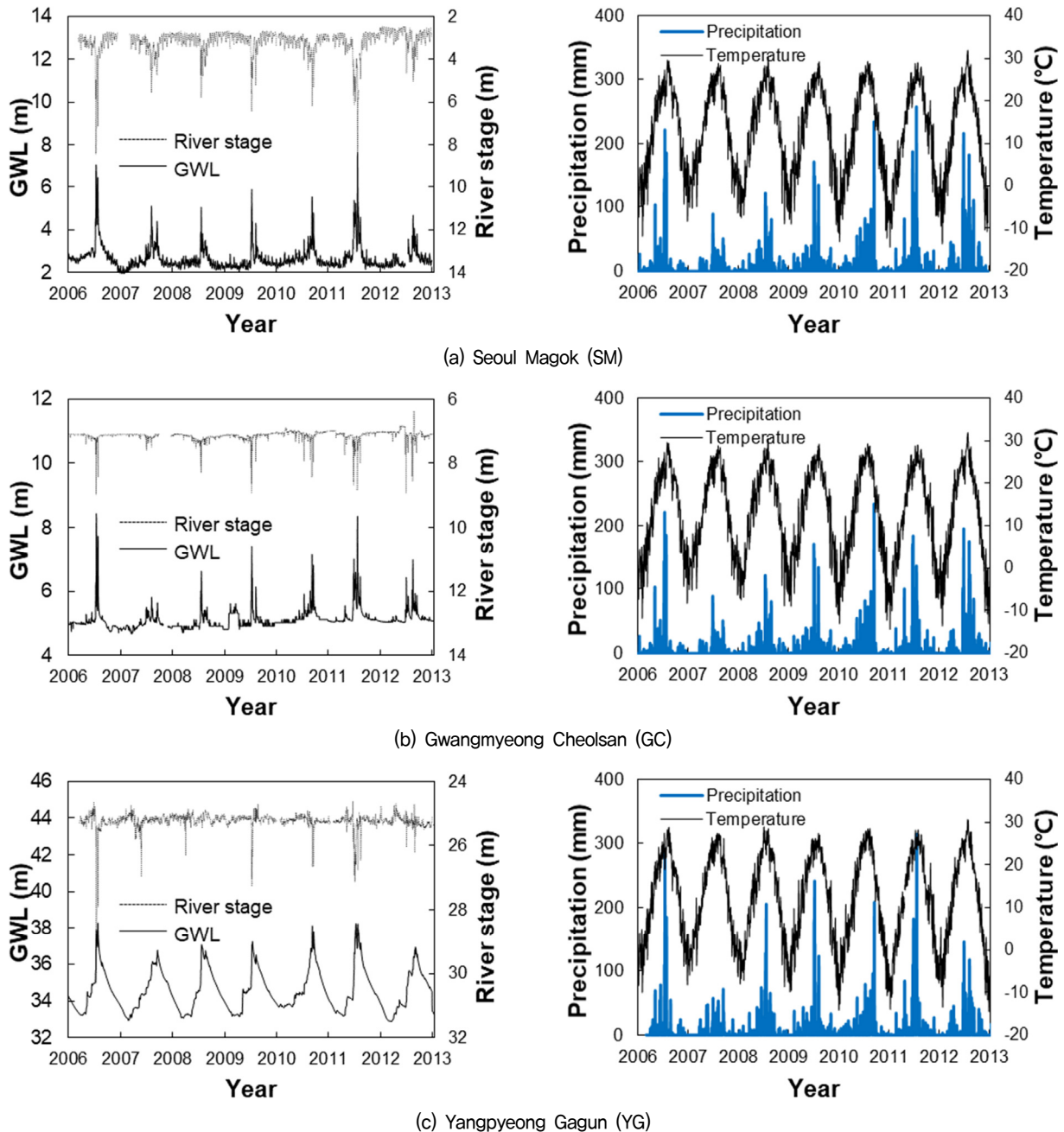
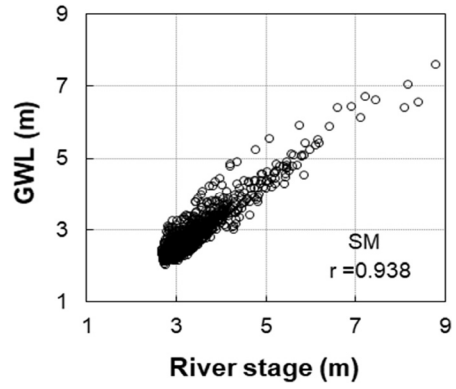
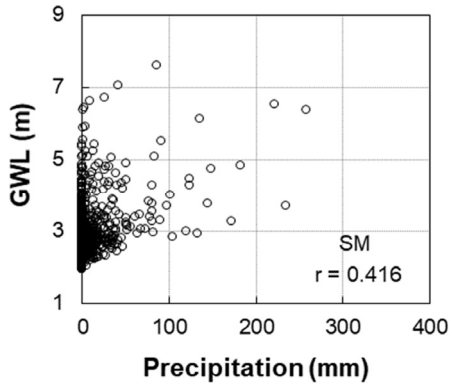


Fig. 6. Observed meteorological and hydrological datasets

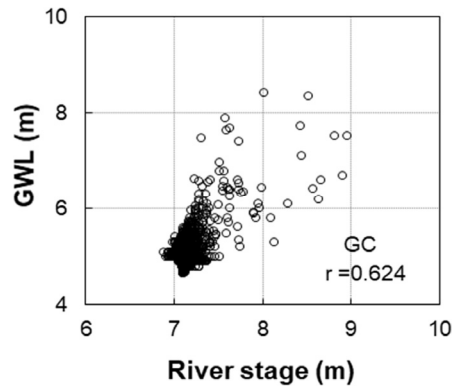
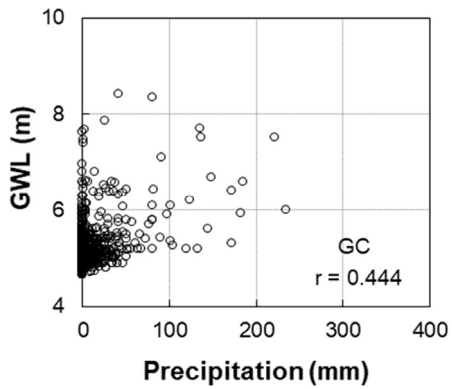
여기서 $MA_{t,n}$ 는 t 일의 n 일간 강우이동평균값을 나타내며, PRE_t 는 t 일의 강우량을 나타낸다. 강우이동평균은 선행강우를 고려하는데 빈번히 사용되어 왔으며, 지하댐 운영지표로 사용되고 있다(Wilhite and Glantz, 1985; Gutmann, 1999). 또한, Yang et al.(2011)과 Kim and Lee (2018c)은 강우량과 지하수위와의 상관성 분석에 사용하였다.

Fig. 8은 SM, GC, YG의 적용된 일수에 따른 강우이동평균과 지하수위와의 상관계수 곡선을 나타낸다. 상관계수는 SM과 GC의 경우 각각 10일과 20일에서 가장

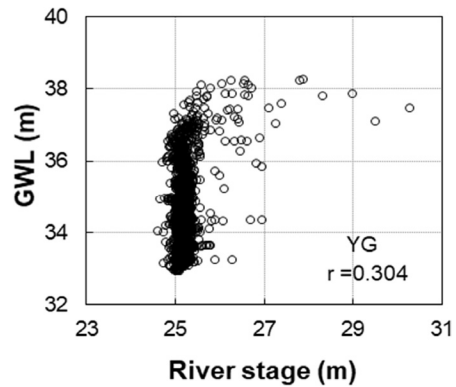
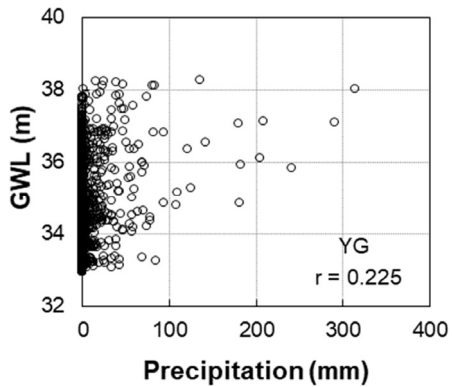
높은 상관 계수를 나타내는 반면, YG는 80일에서 가장 높은 상관계수를 보였다. 이는 발생한 강우가 지하수에 도달하는 시간과 관련이 있는 것으로 판단이 되며, SM과 GC는 발생한 강우가 하천으로 유출되어 지하수위 변동에 영향을 미치는 시간이, YG의 경우는 강우가 지표면으로 침투하여 지하수위에 영향을 미치는 시간이 반영된 것으로 판단된다. 본 연구에서 지하수위와 강우이동평균의 상관관계 분석을 위하여 가장 높은 상관계수를 보이는 일수(최적 일수)를 적용하였다. Fig. 9는 지하수위와 최적 일수가 적용된 강우이동평균과의 상관관



(a) Seoul Magok (SM)



(b) Gwangmyeong Cheolsan (GC)



(c) Yangpyeong Gagun (YG)

Fig. 7. Correlation analysis results for GWL with precipitation and river stage

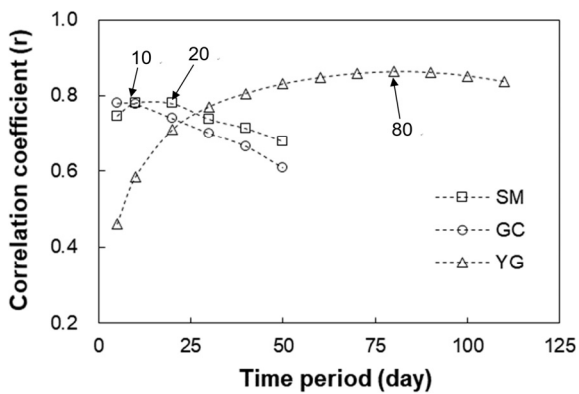
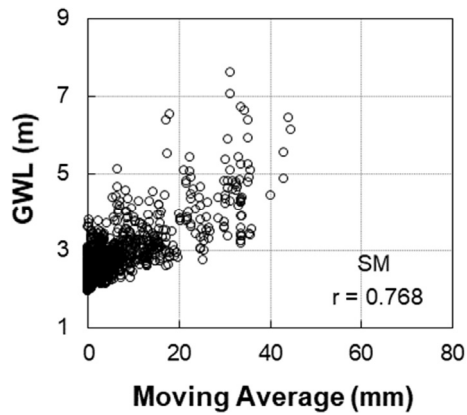
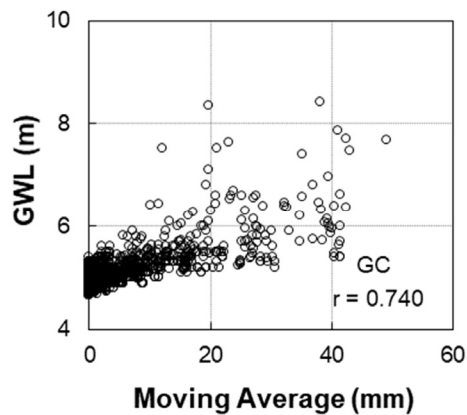


Fig. 8. Optimal days for the moving average of precipitation

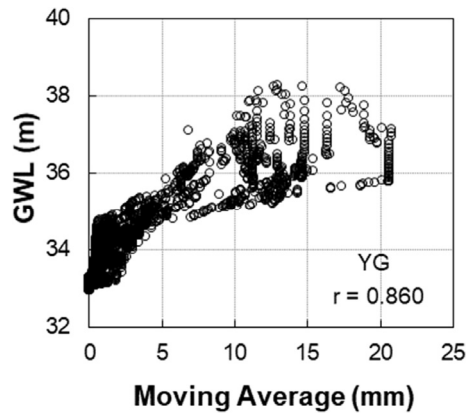
계를 나타낸다. SM, GC, YG의 상관계수는 각각 0.768, 0.740, 0.860으로 일관측 강수량 보다 높은 상관성이 관찰되었다. 특히, YG의 경우 상관계수가 0.860으로 상관성이 다른 두 지역보다 월등히 높았다. 이는 앞에서 언급된 것과 같이 이동평균 개념을 이용하여 간헐적인 강수량 데이터를 연속적 변동값으로 변환하였으며, 각 지역마다의 지질 및 지형 특성에 따른 시간차가 반영되었기 때문인 것으로 판단된다.



(a) Seoul Magok (SM)



(b) Gwangmyeong Cheolsan (GC)



(c) Yangpyeong Gagun (YG)

Fig. 9. Correlation analysis results for GWL with moving average of precipitation

3.5 영향인자 분석 결과

각 지역의 상관관계 분석 결과를 Table 1에 정리하였다. SM의 경우, 하천수위와 지하수위의 상관계수가 0.938로 이 지역의 주된 지하수위 영향인자는 하천수위인 것으로 나타났다. 그 이유는 도심지역의 높은 지표면 포장율로 인하여 강우가 지표면으로 직접 침투되지 못하고

Table 1. Description of input type and learning data length

Study site	Precipitation	River stage	Moving average
SM	0.416	0.938	0.768 (10)*
GC	0.444	0.624	0.740 (20)*
YG	0.225	0.304	0.860 (80)*

*Applied optimal days

하수관을 통하여 하천으로 유출되어, 상류에서부터 흘러오는 하천수와 함께 지하수위 변동에 직접적으로 영향을 미쳤을 것으로 판단된다. 또한 이 지역의 지반은 투수성이 높은 흙으로 되어 있기 때문에 하천수와 지하수의 수리적 연결성이 좋은 것으로 판단된다. YG의 경우, 80일 강우이동평균과 지하수위의 상관계수가 0.860으로 가장 높았으며, 이 지역의 주된 지하수위 영향인자는 선행강우를 고려한 강우량인 것으로 나타났다. 이는 발생한 강우가 지표면을 통하여 침투하여 지하수위까지 도달하지만, 지표면에 투수성이 낮은 흙이 분포하고 있어 강우 발생과 지하수위 변동 사이에 어느 정도 시간차가 발생하는 것으로 판단된다. GC의 경우, 지하수위가 하천수위와 강우이동평균과의 상관계수가 각각 0.624와 0.740이었다. 이는 하천수위 및 선행강우를 고려한 강우가 모두 어느 정도 지하수위와 상관관계가 있다고 볼 수 있다. 결과적으로 대상지역의 지질 및 지형, 지표 특성에 따라 지하수위 변동에 영향을 미치는 주된 인자는 다르다는 것을 확인할 수 있다.

4. 영향인자에 따른 지하수위 예측

4.1 인공신경망 모델 및 인풋 타입(Input type)

본 연구에서는 지하수위 영향인자에 따른 인공신경망의 지하수위 예측 성능을 분석하였다. 이를 위하여, Fig. 6의 7년 동안 관측된 지하수위(GWL), 하천수위(RS), 강우량(PRE), 기온(TEMP) 데이터를 이용하였다. 또한, Fig. 9의 상관분석에서 사용한 강우이동평균(MA)를 추가적으로 이용하였다. 간혹 지하수위 예측에 계절효과를 반영하기 위하여 Dummy 파라미터를 이용하기도 하지만 (Uddameri, 2007; Sahoo and Jha, 2013), Kim and Lee (2018b)의 연구결과에 따르면 Dummy 파라미터의 영향이 미미한 것으로 나타났으며, TEMP에 계절의 영향이 담겨 있는 것으로 판단했기 때문에 본 연구에서는 Dummy 파라미터를 사용하지 않았다. 학습데이터양(Learning data

length, LDL)에 대한 영향을 함께 분석하기 위하여, LDL을 1년, 3년, 6년으로 조절하여 지하수위 예측에 이용하였다. 한편, 일반적으로 선행 수문 이벤트를 반영하려는 목적으로 과거 몇 일간 관측된 데이터도 학습에 추가하는 경우가 많다. 본 연구에서는 특정 t일의 지하수위(i.e. GWL_t)를 위하여 입력값으로 사용된 영향인자들의 과거 3일전 데이터(Applied data time, i.e. $PRE_t, PRE_{t-1}, PRE_{t-2}, PRE_{t-3}$)까지 함께 학습 및 예측에 이용하였다. 다만, 강우이동평균의 경우 이미 선행강우가 고려되었다고 보기 때문에 Applied data time은 t일만 사용하였다(i.e. MA_t). 본 연구에서 인공신경망으로 지하수위를 예측하는데 사용한 입력값 및 LDL 조합은 Table 2에 정리하였다.

지하수위 학습을 위하여 Levenberg-Marquardt(LM) algorithm을 이용하였고, Active function으로는 지하수위 예측에 빈번히 사용되는 Sigmoid function을 적용하였다. 앞에서 언급한 것과 같이 학습에는 1년, 3년, 6년의 데이터에서 임의로 추출된 90%의 데이터를 이용하여 학습에 사용하였고, 나머지 10% 데이터를 이용하여 테스트를 수행하였다. 테스트 단계에서 Trial-and-error 최적의 Hidden later와 Hidden node 수를 결정하였다. 이를 이용하여 다음 해인 2012년의 지하수위를 예측하였다.

4.2 예측 결과 및 분석

Table 2에 정리된 입력값 및 LDL 조합으로 SM, GC, YG의 2012년도 지하수위를 예측하였다. Fig. 10은 각 조합에 대한 PCC와 RMSE를 나타낸다. SM의 경우[Fig. 10(a)], IT1과 IT2의 예측성능에 큰 차이가 있는 것을 볼 수 있다. 영향인자 분석결과에 따르면[Fig. 7(a)], SM에서 강우량은 지하수위와 상관성이 매우 낮으나 하천수위는 상관성이 매우 높기 때문인 것으로 판단된다. IT2와 IT3은 유사하다. 이는 입력값으로 사용된 하천수위

가 지하수위 예측에 충분한 역할을 하였으며, 강우량의 영향은 크지 않기 때문인 것으로 판단된다. IT4는 기온을 입력값으로 추가하였기 때문에 계절영향이 반영되었으나 계절의 영향은 크지 않은 것으로 보인다. 다만, LDL이 1년인 예측결과에서 기온 데이터가 긍정적으로 작용을 하였다. IT5와 같이 입력값에 강우이동평균의 포함은 예측 성능에 큰 영향을 미치지 않았다. Fig. 11은 SM의 관측된 지하수위와 IT1, IT2, IT5의 입력값을 이용하여 예측된 지하수위를 비교한 것을 보여준다. SM에서 지하수위 변동에 영향을 미친 주된 인자가 하천수위이기 때문에, 그의 영향이 인공신경망을 통한 예측 성능에도 반영된 것으로 보인다. 또한, LDL이 커질수록 RMSE가 작아지는 결과를 확인할 수 있었다.

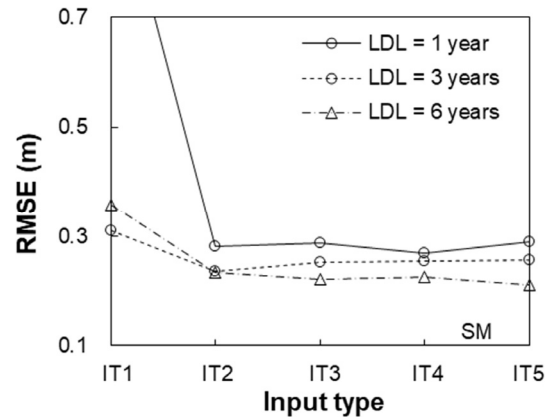
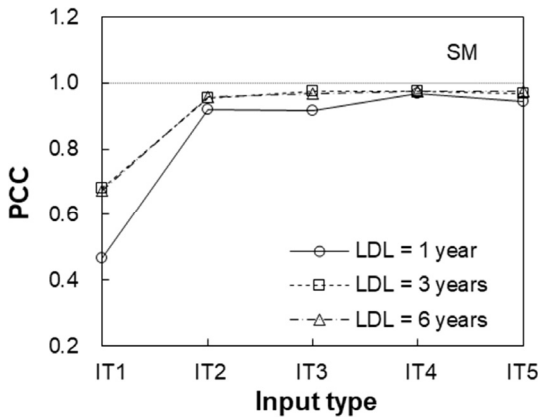
GC의 경우[Fig. 10(b)], IT1의 예측성능이 SM과는 달리 비교적 양호하였으며, 비교적 IT1~IT5의 변화가 크지 않았다. IT2는 IT1보다 약간 더 양호한 예측결과를 보였다. IT4가 IT3보다 PCC와 RMSE에서의 경우는 약간 불리해진 것을 확인할 수 있다. 이는 계절효과를 반영하기 위해 사용한 온도가 오히려 예측에 방해가 된 것으로 판단된다. IT5와 같이 강우이동평균을 입력값으로 추가한 경우 추가하지 않은 IT3과 크게 다르지 않았다. LDL에 따른 영향도 비교적 적었으며 특히 LDL이 1인 경우, LDL이 3과 5에 비하여 큰 변동성이 관찰되었다. Fig. 12는 관측된 지하수위와 IT3, IT4, IT5를 이용하여 예측한 지하수위를 보여주며, Fig. 10(b)에서 확인할 수 있는 것처럼 IT3, IT4, IT5를 이용하여 예측한 결과가 서로 크게 다르지 않은 것을 확인할 수 있다. 이는 Fig. 7과 Table 1에서 확인할 수 있는 것처럼 강우에 의해 불어난 작은 규모 하천의 수위가 지하수위 변동에 영향을 미쳤기 때문이라 판단된다.

YG의 경우[Fig. 10(c)], IT1, IT2, IT3의 PCC가 모두 0.6 보다 작으며, RMSE는 1.2m에 가깝다. 이는 인공신경망으로 인한 지하수위 예측결과가 신뢰할 수 있는 수

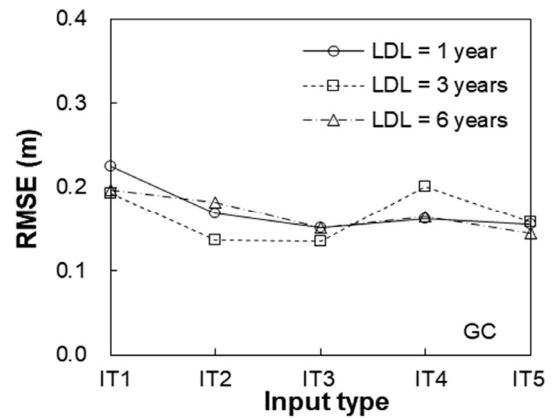
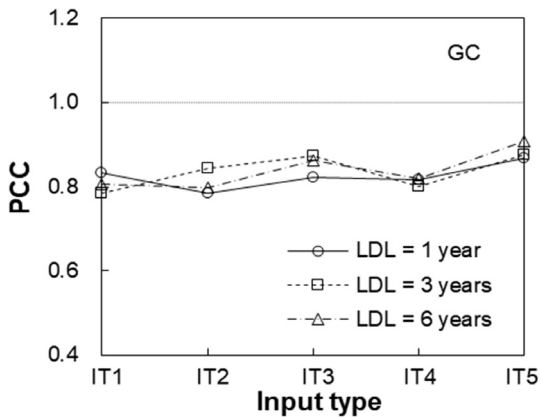
Table 2. Description of input type and learning data length

Input type	Input	Applied data time**	LDL** 1 year	LDL** 3 years	LDL** 6 years
IT1	PRE	t, t-1, t-2, t-3	2011	2009-2011	2006-2011
IT2	RS				
IT3	PRE, RS				
IT4	PRE, RS, TEMP				
IT5	PRE, RS, TEMP, MA*				

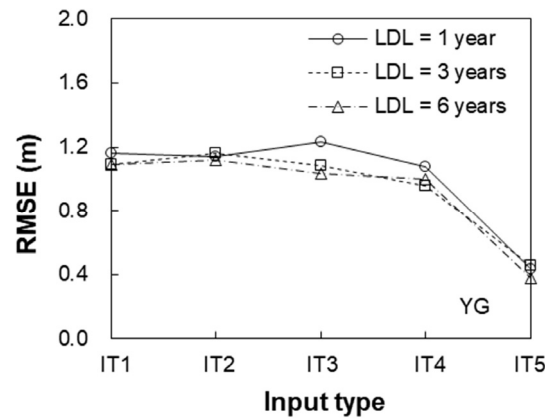
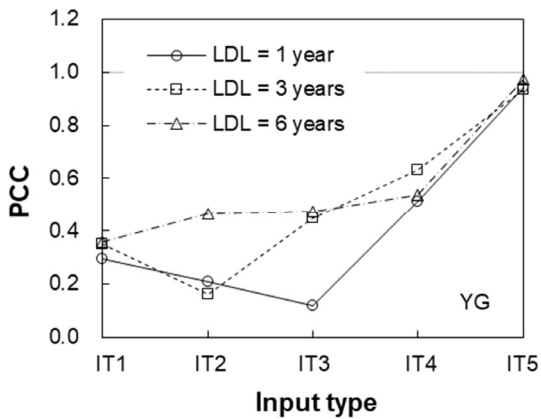
* Applied time lag for MA was only t; ** t and t-n indicate the current data and data for n days ago; *** LDL: Learning data length



(a) Seoul Magok (SM)



(b) Gwangmyeong Cheolsan (GC)

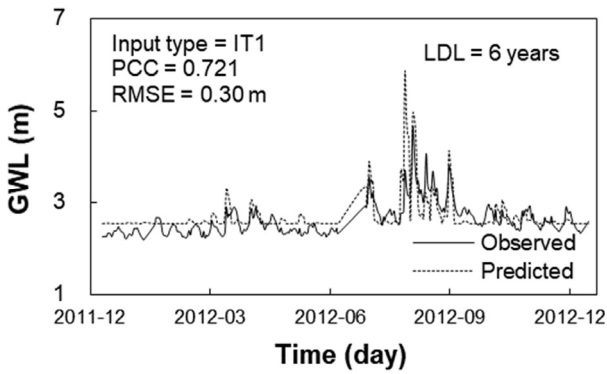


(c) Yangpyeong Gagun (YG)

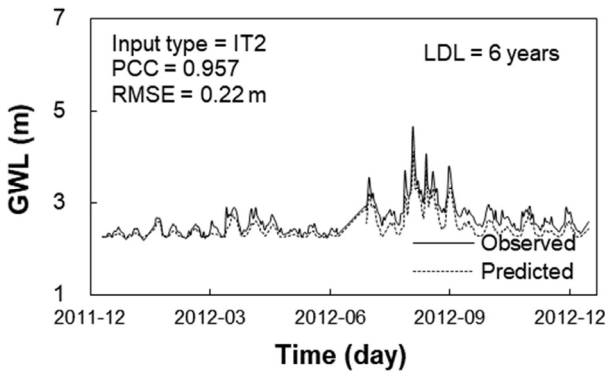
Fig. 10. Analysis results for GWL prediction performance

준이 아닌 것을 의미한다. IT4에서 계절효과를 입력값으로 추가한 결과 예측성능이 조금 개선되는 양상이 PCC와 RMSE에서 관찰된다. IT5와 같이 강우이동평균을 입력값으로 추가한 경우, 모든 LDL에서 PCC가 0.95보다 컸으며 RMSE는 0.4m에 가깝게 감소하였다. 이는 YG에서 강우이동평균이 인공지능망을 이용한 지하수위 예측성능에 긍정적으로 큰 영향을 미친 것을 의

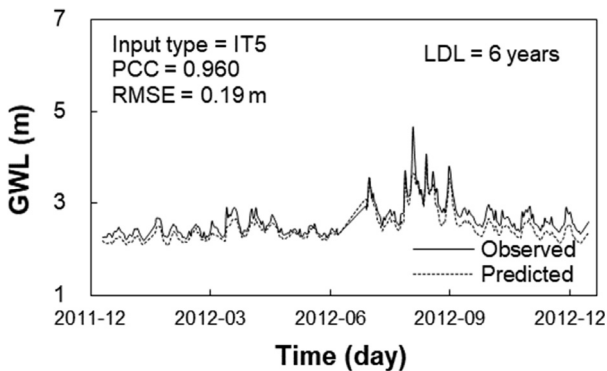
미하며, 이것은 지하수위와 강우이동평균의 상관관계가 높은 것과 밀접한 관련이 있는 것으로 판단된다. Fig. 10은 6년 LDL과 입력값 IT3, IT4, IT5를 이용하여 예측한 YG의 지하수위를 나타낸다. Fig. 13(a)에 나와 있는 것 처럼, IT3의 경우는 PCC와 RMSE가 각각 0.473, 1.04m이며, 예측 지하수위가 관측 지하수위와 전혀 다른 변동 양상을 보인다. IT4의 경우[Fig. 13(b)], PCC와



(a)



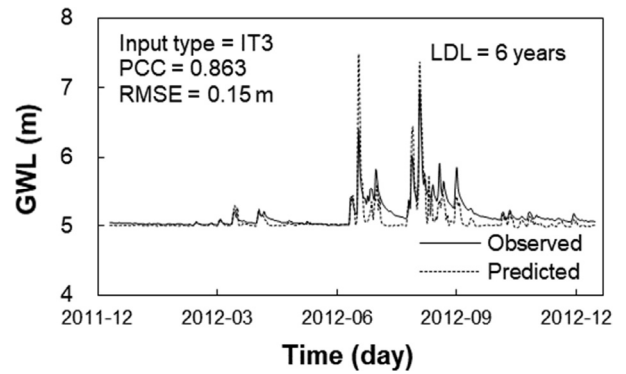
(b)



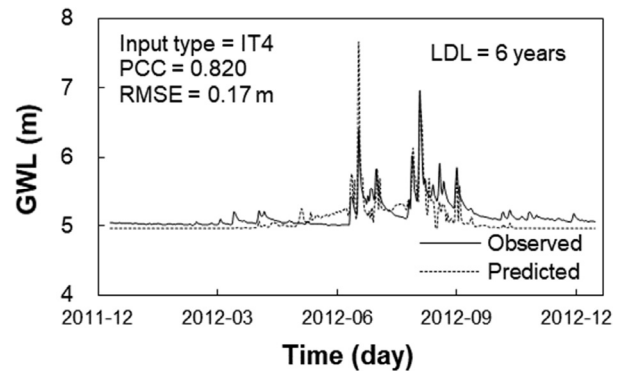
(c)

Fig. 11. Measured and predicted GWL with time for SM with (a) IT1, (b) IT2, and (c) IT5

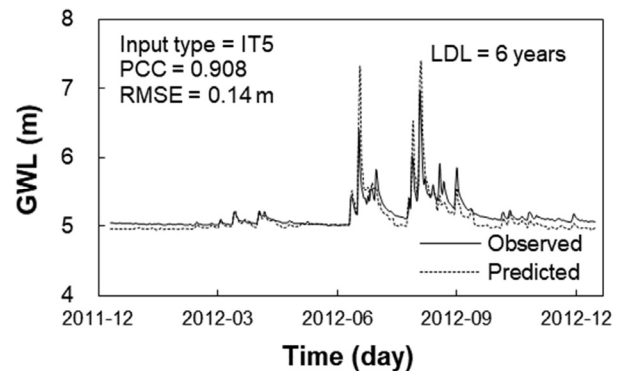
RMSE가 각각 0.539, 1.00m로 약간 개선된 것처럼 보이나 지하수위의 상승·하강하는 시기, 침투 지하수위 시기 등을 잘 예측하지 못하였다. IT5의 경우[Fig. 13(c)], PCC와 RMSE가 0.976, 0.36m로서 예측성능이 매우 좋아졌으며, 예측 지하수위의 상승·하강하는 시기 및 침투 지하수위 시기 등도 잘 예측하고 있는 것을 확인할 수 있다. 결과적으로, 인공지능망을 이용하여 YG의 지하수위 변동성을 예측할 경우 강우이동평균을 입력값으로 사용하면 양호한 예측성능을 기대할 수 있으며, 그 이유는 YG에서 선행강우를 고려한 강우가 지



(a)



(b)



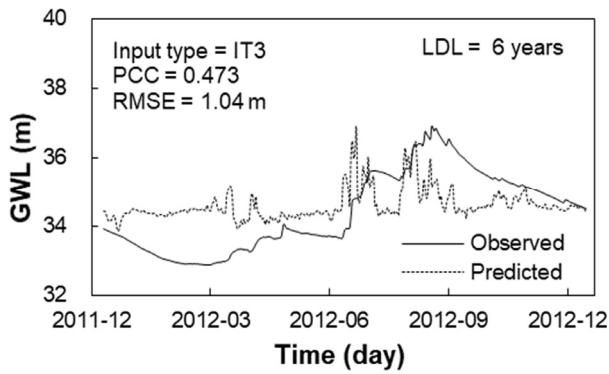
(c)

Fig. 12. Measured and predicted GWL with time for GC with (a) IT3, (b) IT4, and (c) IT5

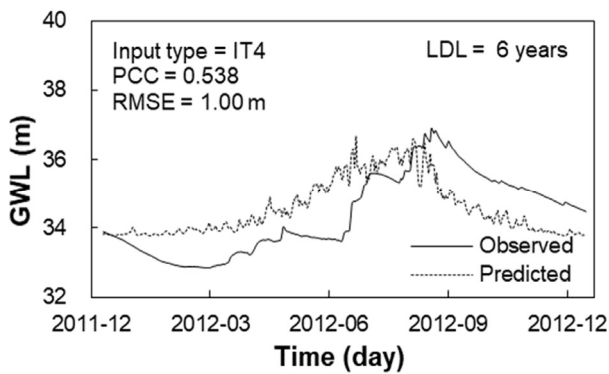
하수위 변동에 영향을 미치는 주된 인자이기 때문인 것으로 판단된다.

5. 결론

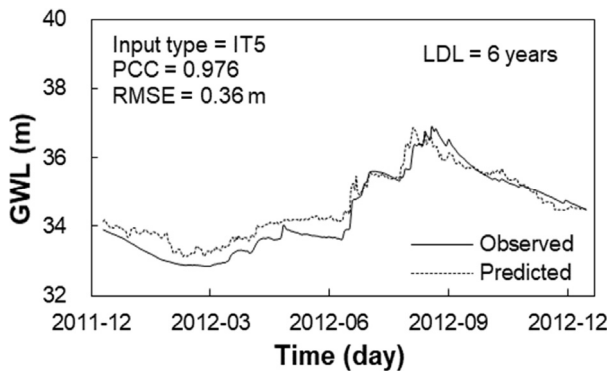
본 연구에서는 지하수위 변동에 미치는 영향인자가 Data-based approach 중 하나인 인공지능망의 지하수위 예측 성능에 미치는 영향에 대하여 분석하였다. 이를 위하여 3곳의 연구 대상지역을 선정하여 7년간의 데이터를 수집하였으며, 지하수위 영향인자 분석을 수행하였



(a)



(b)



(c)

Fig. 13. Measured and predicted GWL with time for YG with (a) IT3, (b) IT4, and (c) IT5

다. 또한, 인공지능망의 여러 입력값 조합을 이용하여 지하수위를 예측하였고, 지하수위 영향인자 분석 결과와 비교하였다.

결론적으로, 1) SM의 주된 지하수위 영향인자는 하천수위였으며, 인공지능을 이용한 지하수위 예측 시 하천수위가 포함되었을 때 양호한 예측 결과를 얻었다. 2) 또한, YG의 주된 지하수위 영향인자는 강우이동평균이었으며, 강우이동평균이 입력값으로 포함되었을 때 지하수위 예측성능이 크게 개선되었다. 3) GC의 경우는 뚜렷한 지하수위 영향인자가 관찰되지 않았으며, 지하

수위 예측결과도 입력값 조합에 따른 차이가 비교적 적었다. 4) 따라서, 인공지능망을 이용하여 지하수위를 예측할 때 사용되는 입력값이 매우 중요하며 예측성능에 미치는 영향이 매우 크다는 것을 알 수 있다. 이는 지하수위의 예측 이전에, 적절한 지하수위 영향인자 평가가 수행되어야 하며 이를 예측에 적용할 필요가 있는 것을 나타낸다.

감사의 글

본 연구는 2019년 연세대학교대학원 연구장학금 및, 한국에너지기술평가원, 한국연구재단, 국토교통과학기술진흥원의 지원(Nos. 20194030202460, 2020R1A2C2011966, 20SMIP-A158708-01)으로 수행되었으며, 이에 깊은 감사를 드립니다.

참고문헌 (References)

1. Ausilio, E. and Conte, E. (2005), "Influence of Groundwater on the Bearing Capacity of Shallow Foundations", *Canadian Geotechnical Journal*, Vol.42, pp.663-672.
2. Batelaan, O., De Smedt, F., and Triest, L. (2003), "Regional Groundwater Discharge: Phreatophyte Mapping, Groundwater Modeling and Impact Analysis of Land-use Change", *Journal of Hydrology*, Vol.275, pp.86-108.
3. Coppola, E. A., Rana, A. J., Poulton, M. M., Szidarovszky, F., and Uhl, V. W. (2005), "A Neural Network Model for Predicting Aquifer Water Level Elevations", *Groundwater*, Vol.43, No.2, pp. 231-241.
4. Daliakopoulos, I. N., Coulibaly, P., and Tsanis, I. K. (2005), "Groundwater Level Forecasting Using Artificial Neural Networks", *Journal of Hydrology*, Vol.309, pp.229-240.
5. Dawson, C. W. and Wilby, R. L. (2001), "Hydrological Modelling Using Artificial Neural Networks", *Progress in Physical Geography*, Vol.25, No.1, pp.80-108.
6. Gerken, W. C., Purvis, L. K., and Butera, R. J. (2006), "Genetic Algorithm for Optimization and Specification of a Neuron Model", *Neurocomputing*, Vol.69, pp.1039-1042.
7. Guttman, N. B. (1999), "Accepting the Standardized Precipitation Index: A Calculation Algorithm", *Journal of the American Water Resources Association*, Vol.35, No.2, pp.311-322.
8. Hoque, M., Hoque, M., and Ahmed, K. (2007), "Declining Groundwater Level and Aquifer Dewatering in Dhaka Metropolitan Area, Bangladesh: Causes and Quantification", *Hydrogeology Journal*, Vol.15, pp.1523-1534.
9. Kim, I., Park, D., Kyung, D., Kim, G., Kim, S., and Lee, J. (2016), "Comparative Influences of Precipitation and River Stage on Groundwater Levels in Near-river Areas", *Sustainability*, Vol.8, No.1, pp.1-16.
10. Kim, I. and Lee, J. (2018a), "Prediction Model for Spatial and Temporal Variation of Groundwater Level based on River Stage",

- Journal of Hydrologic Engineering*, Vol.23, No.67, DOI: 10.1061/(ASCE)HE.1943-5584.0001658, 1-8.
11. Kim, I. and Lee, J. (2018b), "The Effect of Seasonal Input on Predicting Groundwater Level Using Artificial Neural Network", *Ecology and Resilient Infrastructure*, Vol.5, No.3, pp.125-133.
 12. Kim, I. and Lee, J. (2018c), "Influencing Factor Analysis on Groundwater Level Fluctuation Near River", *Ecology and Resilient Infrastructure*, Vol.5, No.2, pp.72-81.
 13. Kim, I. and Lee, J. (2019), "Proposed Correlation Model for Groundwater Level Prediction based on River Stage Considering Changes in Hydrological and Geological Conditions", *Journal of Hydrologic Engineering*, Vol.24, No.10, DOI: 10.1061/(ASCE)HE.1943-5584.0001849, 1-14.
 14. Nourani, V. and Mousavi, S. (2016), "Spatiotemporal Groundwater Level Modeling Using Hybrid Artificial Intelligence-meshless Method", *Journal of Hydrology*, Vol.536, pp.10-25.
 15. Park, D., Kim, I., Kim, G., and Lee, J. (2017), "Groundwater Effect Factors for the Load-carrying behavior of Footings from Hydraulic Chamber Load Tests", *Geotechnical Testing Journal*, Vol.30, No.3, pp.440-451.
 16. Park, D., Kim, I., Kim, G., and Lee, J. (2019), "Effect of Groundwater Fluctuation on Load Carrying Performance of Shallow Foundation", *Geomechanics and Engineering*, Vol.18, No.6, pp.575-584.
 17. 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.
 18. Shahriar, M. A., Sivakugan, N., Das, B. M., Urquhart, A., and Tapiolas, M. (2014), "Water Table Correction Factors for Settlements of Shallow Foundations in Granular Soils", *International J. of Geomechanics*, [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0000391](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000391).
 19. Serrano, S. E. and Workman, S. R. (1998), "Modeling Transient Stream/aquifer Interaction with the Non-linear Boussinesq Equation and its Analytical Solution", *Journal of Hydrology*, Vol.206, pp. 145-255.
 20. 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, pp.1865-1887.
 21. Shamsudduha, M., Chandler, R. E., Taylor, R. G., and Ahmed, K. M. (2009), "Recent Trends in Groundwater Levels in a Highly Seasonal Hydrological System: The Ganges-Brahmaputra-Meghna Delta", *Hydrology and Earth System Sciences*, Vol.13, pp.2373-2385.
 22. Trichakis, I. C., Nikolos, I. K., and Karatzas, G. P. (2011), "Artificial Neural Network (ANN) based Modeling for Karstic Groundwater Level Simulation", *Water Resources Management*, Vol.25, pp.1143-1152.
 23. Uddameri, V. (2007), "Using Statistical and Artificial Neural Network Models to Forecast Potentiometric Levels at a Deep Well in South Texas", *Environmental Geology*, Vol.51, pp.885-895.
 24. Wen, X., Feng, Q., Yu, H., Wu, J., Si, J., Chang, Z., and Xi, C. (2015), "Wavelet and Adaptive Neuro-fuzzy Inference System Conjunction Model for Groundwater Level Predicting in a Coastal Aquifer", *Neural Computing and Applications*, Vol.26, pp.1203-1215.
 25. Wilhite, D. A. and Glantz, M. H. (1985), "Understanding the Drought Phenomenon: The Role of Definitions", *Water International*, Vol.10, pp.110-120.
 26. Yang, J. S. and Kim, N. K. (2011), "The Correlation between the Moving Average of Precipitation and Groundwater Level in Korea", *Korean Society of Civil Engineers*, Vol.31, pp.265-276.
 27. Yasuhara, K., Murakami, S., Mimura, N., Komine, H., and Recio, J. (2007), "Influence of Global Warming on Coastal Infrastructural Instability", *Sustainability Science*, 2, pp.13-25.

Received : December 1st, 2020

Revised : March 23rd, 2021

Accepted : April 8th, 2021