

지진 전조현상 지하수 수위 예측에 대한 빅데이터 분석

이정구*

*한국과학기술정보연구원 데이터분석본부

e-mail:jglee@kisti.re.kr

Big data analysis on the prediction of groundwater level in the earthquake precursor

Jeong-Gu Lee*

*Dept. of Data Analysis, KISTI

요약

본 논문에서는 지진 전조 현상 감지를 목적으로 지하수 수위 예측에 대한 빅데이터 분석을 수행하였다. 지진 감시는 지진학적, 지구물리학적, 지구화학적 인자들의 관측을 바탕으로 하는데, 이 중 지하수 수위 및 수질 관측과 용존 가스의 관측은 지진에 민감하게 반응하여 지진 관측과 예측에 효용성이 높다고 할 수 있다. 지하수위는 센서 측정지점에서 물이 가지는 에너지를 알려주는 인자로, 지하수위 상승과 하강은 지질 매체의 응력 분포에 의한 변화, 지하수 유동 경로의 변화, 투수 계수의 변화에 기인한다. 지진 피해 방지와 인명 피해를 줄이기 위해 지진이 자주 발생하는 일본, 중국, 미국, 이탈리아 등의 국가에서는 지진의 감시와 예측을 위한 다양한 연구가 활발히 진행되고 있다. 국내에서도 2016년 경주지진과 2017년 포항지진은 지진과 관련된 지하수 연구를 활성화하는 계기가 되었다.

예측에 효용성이 높은 인자인 지하수위 데이터를 이용한 지진 전조 현상 지하수 수위 예측에 대한 빅데이터 분석을 수행하였다.

1. 서론

지진은 전 세계적으로 발생하는 일반적인 현상이지만, 지진 발생 시에 인류사회에 미치는 피해의 양적, 질적 규모가 급격히 증가하고 있으며, 인명 피해도 그 규모가 점점 증가하고 있다. 지진 피해 방지와 인명 피해를 줄이기 위해 지진 예측과 예보를 위한 다양한 연구가 전 세계적으로 시도되고 있다. 지진 피해를 최소화하기 위한 감시 및 예보 연구는 지진학적, 지구물리학적, 지구화학적 인자들의 관측을 바탕으로 이루어진다. 미소지진의 발생 패턴, 지표 변이, 지하수 이상 변동, 지자기장 변동, 응력 변화, 대기 이온층의 변동, 단층 가스의 발생 등이 지진 관측과 예측에 효용성이 높은 것으로 간주되고 있다. 특히 지진에 의한 지하수의 변동은 전 세계적으로 다양한 지역에서 보고된 바 있으며, 중국은 지하수 관측망을 이용하여 자동으로 지하수위와 수온, 전기 전도도를 측정하여 1975년 Haicheng 지진 예측을 성공한 사례가 보고된 바 있다. 국내에서도 2017년 11월 15일 규모 5.4의 포항에서 발생한 지진을 전후하여 국가지하수정보센터에 등록된 지하수 관측소 자료 분석결과 수위가 크게 낮아지거나, 높아지는 이상 변동이 관측된 바 있다. 따라서 본 연구에서는 지진 관측과

2. 지하수 수위 예측을 위한 빅데이터 분석

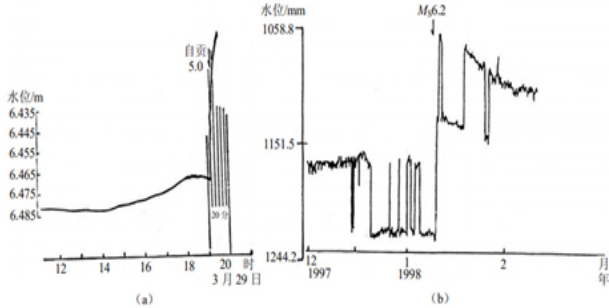
2.1 지하수의 지진 전 및 지진 동시 변동 특성

지하수는 지질 매체 내에서 일정한 연속성을 가지고 움직인다. 이러한 연속성에는 지하수가 지니는 에너지 차이에 의한 유동과 지질 매체에 작용하는 지구의 주기적인 변동이 복합적으로 작용한다. 지진에 관련한 지하수 변화는 발생 시간에 따라, 지진 전 변화와 지진 동시 변화, 지진 후 변화로 나눌 수 있다.

2.1.1 지진 전 지하수 변동

지진 전에 나타나는 지하수의 변동 특성에 대한 연구는 중국에서 가장 활발하게 진행되고 있다. 지진 전 나타나는 수위의 이상 변동은 시간에 따라 다년간의 추세 이상, 계절적 변동 이상, 단기 이상으로 나눌 수 있다. [그림 1]은 지진 전 지하수위 변동 이상을 나타낸

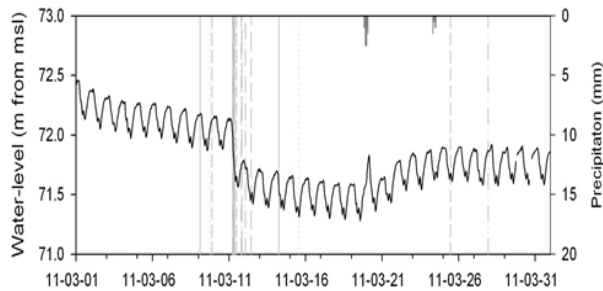
것이다. [그림 1]의 (a) 1985년 중국 쓰촨 쓰궁 청광 3 관정, (b) 베이징 올리잉 관정에서 나타난 지진 전 지하수위 변동을 나타낸 그래프이다. 이 그래프를 통해 지진 발생 전에 전조 현상으로 지하수위가 급격히 상승하거나 하강하는 것을 알 수가 있다.



[그림 1] 지진 전 지하수위 변동 특성

2.1.2 지진 동시 지하수 변동

지진과 동시에 나타나는 지하수 변동은 지진파가 지나가면서 대수층의 물을 흔들어 나타나는 진동형 변화와 대수층 사이에 균열이 생기거나 압밀되면서 발생하는 계단형 변화 및 점진적 변화가 있다. [그림 2]는 2011년 동일본 대지진에 의한 국가지하수관측망의 지진 동시 지하수위 변동 사례를 나타낸 것이다. 이 관측망의 지하수위는 동일본 대지진 발생에 따라 약 50cm 가량 감소한 것으로 조사되었다.



[그림 2] 지진 동시 지하수위 변동 특성

2.2 지하수위 예측에 대한 빅데이터 분석

2.2.1 지하수 관측 자료

지하수 자동 관측 데이터는 한국수자원공사 국가지하수정보센터를 통하여 얻을 수 있다. 이 센터에서는 지하수 관측소에 대한 이름, 위치(위도/경도), 해발고도, 관측공의 심도와 구조, 관측 요소(수위/수온/전기전도

도)의 시계열 데이터 등이 포함되어 있다. 강수 데이터는 자동기상관측소(AWS)에서 측정된 자료를 기상청의 협조를 받아 수집할 수 있다. 수집 데이터는 관측소의 위치정보와 강수 자료 등을 포함되어 있다. 지진 데이터 수집은 IRIS(Incorporated Research Institutions for Seismology)의 DB와 기상청 지진 발생 자료실에서 수집할 수 있다. 본 논문에서는 농어촌지하수관측망 데이터와 기상청 데이터를 수집하여 활용하였다.

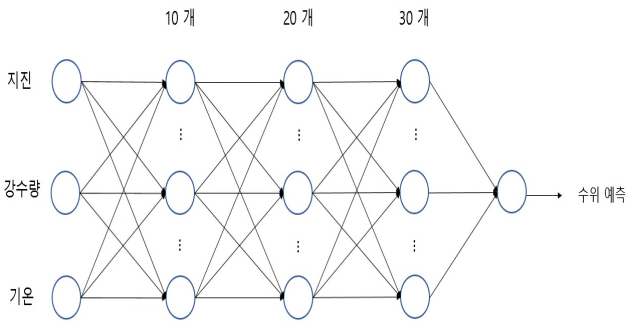
2.2.2 국내 실측 관측공 데이터를 이용한 지하수위 예측

지하수 수위를 예측하기 위해서는 모든 수위에 영향을 주는 변수에 대해 알아야 한다. Freeze and Cherry에 의하면 지하수 수위에 영향을 주는 요인은 대체로 14가지 정도로 알려져 있다.

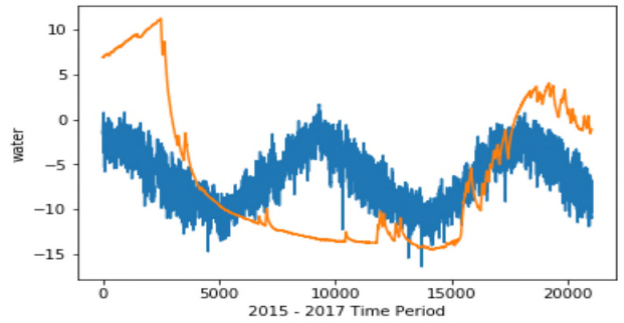
	Unconfined	Confined	Natural	Man-induced	Short-lived	Diurnal	Seasonal	Long-term	Climatic influence
Groundwater recharge	⊙		⊙				⊙		⊙
Air entrapment	⊙		⊙		⊙				⊙
Evapotranspiration and phreatophytic consumption	⊙		⊙			⊙			⊙
Bank-storage effects	⊙		⊙				⊙		⊙
Tidal effects	⊙	⊙	⊙			⊙			
Atmospheric pressure effects	⊙	⊙	⊙			⊙			⊙
External loading of confined aquifers		⊙		⊙	⊙				
Earthquakes		⊙	⊙		⊙				
Groundwater pumping	⊙	⊙		⊙					⊙
Deep-well injection		⊙		⊙					⊙
Artificial recharge	⊙			⊙					⊙
Agricultural activities	⊙			⊙					⊙
Geotechnical drainage	⊙			⊙					⊙

[그림 3] 지하수 수위에 영향을 주는 변수

지하수위에 영향을 줄 수 있는 가능한 모든 변수를 측정하여 다층 퍼셉트론(Multi-layer Perceptron, MLP) 알고리즘을 이용하여 각 지역의 수위를 예측하였다. 변수는 지진, 강수량, 기온을 사용하였고, 지역은 밀양1, 밀양2, 밀양3과 경주 지역으로 선정하였다. 은닉층의 각 개수는 10, 20, 30으로 하였고, 데이터는 2015~2017년(3년) 데이터를 활용하여 지진, 기온, 강수량 데이터를 0~1 사이로 정규화 하였다. [그림 4]는 수위 예측 MLP 알고리즘을 나타낸 것이다.



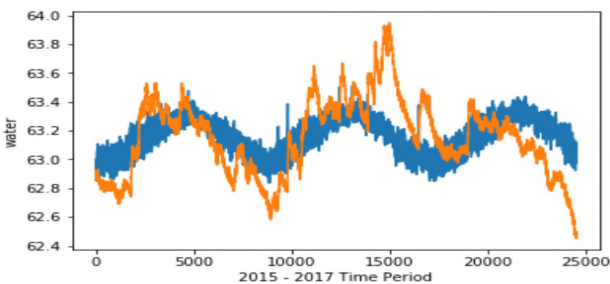
[그림 4] 지하수위 예측 MLP 알고리즘



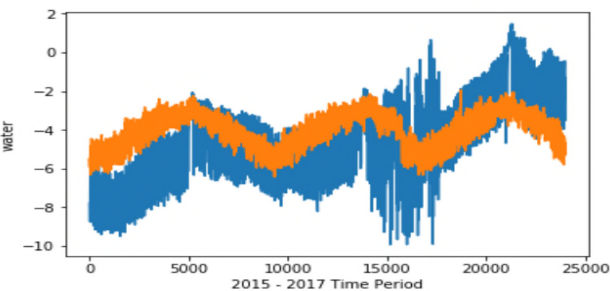
[그림 8] MLP 알고리즘 이용 수위예측(경주)

[그림 5]~[그림 8]은 MLP 알고리즘이 예측한 값과 실제 값과 비교한 그래프이다. 주황색 그래프는 실제 관측공으로부터 측정된 값이고, 파랑색 그래프는 MLP 알고리즘을 적용한 경우의 예측 값이다.

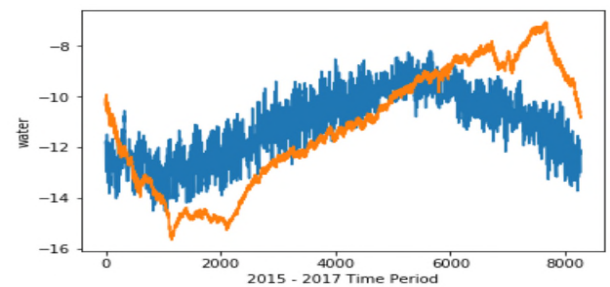
[그림 9]는 지하수 수위예측을 위한 MLP를 적용한 소스 코드를 보여준 것이다.



[그림 5] MLP 알고리즘 이용 수위예측(밀양1)



[그림 6] MLP 알고리즘 이용 수위예측(밀양2)



[그림 7] MLP 알고리즘 이용 수위예측(밀양3)

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import tensorflow as tf
data = pd.read_csv('mlp데이터_밀양2.csv', delimiter=',', dtype='object')
print(data)
#수위, 지진, 기온, 강수량

[[[62.91, 0, -1.4, 0],
 [62.91, 0, -1.4, 0],
 [62.92, 0, -1.4, 0],
 ...
 [62.47, 0, 5.2, 0],
 [62.47, 0, 6.2, 0],
 [62.47, 0, 7.5, 0]]]

x_data = data[:, 1:4] # 지진, 기온, 강수량
y_data = data[:, 0:1] # 수위
x_data = x_data.astype('float') # x_data를 실수형대로 데이터 타입을 받아드림.
y_data = y_data.astype('float')

x_data1 = x_data[:, 0:1]
x_data2 = x_data[:, 1:2]
x_data3 = x_data[:, 2:3]

x_norm1 = (x_data1 - x_data1.min()) / (x_data1.max() - x_data1.min()) # 0-1 사이로 정규화
x_norm2 = (x_data2 - x_data2.min()) / (x_data2.max() - x_data2.min())
x_norm3 = (x_data3 - x_data3.min()) / (x_data3.max() - x_data3.min())

x_norm = np.c_[x_norm1, x_norm2, x_norm3]

global_step = tf.Variable(0, trainable=False, name='global_step') # 학습을 커운트 해줄 변수 생성

X = tf.placeholder(tf.float32)
Y = tf.placeholder(tf.float32)

W1 = tf.Variable(tf.random_uniform([3,10], -1., 1.)) # 은닉층 생성, 은닉 갯수 10개
b1 = tf.Variable(tf.random_uniform([1, 10]))
L1 = tf.matmul(X, W1) + b1

W2 = tf.Variable(tf.random_uniform([10,20], -1., 1.)) # 은닉층 생성, 은닉 갯수 20개
b2 = tf.Variable(tf.random_uniform([1, 20]))
L2 = tf.matmul(L1, W2) + b2

W3 = tf.Variable(tf.random_uniform([20,30], -1., 1.)) # 은닉층 생성, 은닉 갯수 30개
b3 = tf.Variable(tf.random_uniform([1, 30]))
L3 = tf.matmul(L2, W3) + b3

W4 = tf.Variable(tf.random_uniform([30,1], -1., 1.))
b4 = tf.Variable(tf.random_uniform([1, 1]))
model = tf.matmul(L3, W4) + b4

cost = tf.reduce_mean(tf.square(model - Y)) # 오차 계산
optimizer = tf.train.AdamOptimizer(learning_rate=0.001) # 경사하강법, 학습률 0.001
train_op = optimizer.minimize(cost, global_step=global_step)

for step in range(5000): # 반복 학습 5000번 시작
    sess.run(train_op, feed_dict={X:x_data, Y:y_data})

    print("Step: %d, %s" % (step, sess.run(global_step), "Cost: %f" % sess.run(cost, feed_dict={X:x_data, Y:y_data})))

predict = sess.run(model, feed_dict={X: x_data})

plt.plot(predict.flatten()) # 예측 그래프와 실제 데이터 그래프 생성
plt.plot(y_data.flatten())
plt.xlabel("2015 - 2017 Time Period")
plt.ylabel("water")
plt.show()
```

[그림 9] 지하수위 예측 MLP 알고리즘 소스 코드

3. 결론

우리나라에서도 2016년 경주지진과 2017년 포항지진을 포함한 잇따른 지진의 발생으로 지진 예측 연구에 대한 관심과 필요성이 크게 증가하였다. 지진이 발생하기 전에 지진 예측을 통해 인명 피해와 재산 피해를 크게 줄일 수 있다.

지진 예측, 지진 예상, 지진의 영향력 평가를 위해서는 지진으로 인한 이상 변동에 대한 대규모의 분석 작업을 통해 지진의 규모, 지진과 관측점과의 거리, 지진 발생 시각과 이상 변동 발생 시각, 지진학적, 지질학적, 수리지질학적 인자들과의 상호 반응에 대한 예측이 가능하도록 하여야 한다.

이번 연구에서는 지진 관측과 예측에 효용성이 높은 인자 중 하나인 지하수 수위 예측에 대한 빅데이터 분석을 수행하였다. MLP 알고리즘을 적용하여 지하수 수위 예측을 분석한 결과, 실측값과 예측값에 오차가 있음을 알 수 있었다. 이를 통해 정제된 수집 데이터 활용, 다양한 변수 데이터 선정, 신뢰성 있는 분석 기법 등이 선행되어야 함을 알 수 있다.

향후에는 지하수 수위는 지진외에도 다양한 자연적, 인위적 영향에 민감하기 때문에 이상 변동을 지진과 관련해서 강수, 대기압, 기온, 지역의 지하수 사용량, 토목 공사 등 신뢰성 향상을 위한 다양한 자료의 수집과 비교하여 분석할 필요가 있다.

참고문헌

- [1] Liu et al, "The researches on extraction of information in the groundwater and prediction of the strong earthquakes", Dizhen Publishing House, 2010년.
- [2] H.A Lee, "Investigation of groundwater response to earthquakes using the national groundwater monitoring data of Korea", YONSEI University, 2013년.
- [3] Freeze and Cherry, "Groundwater", 1979년.