

시계열 분석을 이용한 흙막이 벽체 변형 예측

Time Series Analysis for Predicting Deformation of Earth Retaining Walls

서 승 환¹ Seo, Seunghwan

정 문 경² Chung, Moonkyung

Abstract

This study employs traditional statistical auto-regressive integrated moving average (ARIMA) and deep learning-based long short-term memory (LSTM) models to predict the deformation of earth retaining walls using inclinometer data from excavation sites. It compares the predictive capabilities of both models. The ARIMA model excels in analyzing linear patterns as time progresses, while the LSTM model is adept at handling complex nonlinear patterns and long-term dependencies in the data. This research includes preprocessing of inclinometer measurement data, performance evaluation across various data lengths and input conditions, and demonstrates that the LSTM model provides statistically significant improvements in prediction accuracy over the ARIMA model. The findings suggest that LSTM models can effectively assess the stability of retaining walls at excavation sites. Additionally, this study is expected to contribute to the development of safety monitoring systems at excavation sites and the advancement of time series prediction models.

요 지

본 연구는 전통적인 통계기반 ARIMA(Auto-Regressive Integrated Moving Average) 모델과 딥러닝 기반 LSTM(Long Short-Term Memory) 모델을 활용하여 굴착 현장의 지중경사계 데이터를 통한 흙막이 벽체 변형을 예측하고, 두 모델의 예측 성능을 비교 분석하였다. ARIMA 모델은 시간의 흐름에 따른 시계열 데이터의 선형적 패턴을 분석하는 데 강점을 보이는 반면, LSTM은 데이터의 복잡한 비선형 패턴과 장기 의존성을 포착하는 데 우수한 능력을 보여주었다. 본 연구는 흙막이 벽체 변형 예측을 위해 지중경사계 계측 데이터에 대한 전처리, 다양한 시계열 데이터 길이 및 입력변수 조건 등에 따른 성능 평가를 포함하였으며, LSTM 모델이 ARIMA 모델에 비해 통계적으로 유의미한 예측 성능 향상을 확인하였다. 본 연구의 결과는 굴착 현장에서의 지중경사계 데이터를 활용한 흙막이 벽체의 안정성 평가에 LSTM 모델을 효과적으로 적용할 수 있음을 보여준다. 또한 이를 바탕으로 향후 굴착 현장 전체에 대한 안전 모니터링 시스템 구축과 시계열 예측 모델 발전에 기여할 것으로 기대된다.

Keywords : ARIMA, Deep learning, Earth retaining wall, Excavation, LSTM

1 정회원, 한국건설기술연구원 지반연구본부 수석연구원 (Member, Senior Researcher, Korea Institute of Civil Engr. and Building Tech.)

2 정회원, 한국건설기술연구원 지반연구본부 선임연구위원 (Member, Senior Research Fellow, Korea Institute of Civil Engr. and Building Tech., Tel: +82-31-910-0219, mkchung@kict.re.kr, Corresponding author, 교신저자)

* 본 논문에 대한 토의를 원하는 회원은 2024년 10월 31일까지 그 내용을 학회로 보내주시기 바랍니다. 저자의 검토 내용과 함께 논문집에 게재하여 드립니다.

1. 서론

지하 공간 개발에 따른 굴착작업은 필연적으로 흙막이 벽체에 변형을 발생시킨다. 굴착 과정에서 벽체 변형은 주변지반 및 인접구조물에 심각한 영향을 미칠 수 있다. 굴착작업이 주변 환경에 미치는 영향을 최소화하기 위해서는 흙막이 벽체의 변형을 정확하게 예측하는 것이 중요하다.

굴착에 의한 흙막이 벽체의 변형은 오래전부터 광범위하게 연구되어 왔다. 현장 계측은 벽체 변형을 이해하는 가장 간단한 방법이며, Peck의 연구(Peck, 1969)가 발표된 1969년부터 현재까지 실측 결과를 토대로 한 굴착현장 데이터가 많은 양으로 문서화되어 있다(Hsiung, 2020; Lim et al., 2018; Long, 2001; Meng et al., 2020; Moormann, 2004). 예를 들어 현장 계측 데이터를 기반으로 토양 유형에 따라 흙막이 벽체 움직임 패턴을 분류하고 예측하기 위한 경험적 방법을 확립하였다(Clough and O'Rourke, 1990). 국내에서도 계측 데이터를 활용하여 굴착 공정별 벽체 변위 및 주변지반에 대한 거동분석 및 지반침하와 인접 구조물의 관계에 관한 연구가 다수 수행되었다(Seo et al., 2006; Lee et al., 2007; Son and Cording, 2005; Son and Cording, 2007, Lee and Kim, 2008).

현장 계측 외에 수치해석 방법 또한 굴착으로 인한 변형 문제를 평가하기 위해 널리 사용된다(Li et al., 2019; Hou et al., 2009, Hsieh and Ou, 1998). 굴착공사에 의한 지반거동을 평가하기 위해 흙의 종류, 굴착형상, 벽체의 종류, 지지 시스템, 시공기법 등에 따른 수치해석 연구(Wong and Poh, 2000; Zhang et al., 2015)가 수행되었다. 또한 벽체의 길이, 심도, 지지층의 심도, 지보재의 강성 등에 따라 지반거동의 영향을 분석하였다(Goh et al., 2020; Seo et al., 2023). 수치해석은 지반-구조물의 상호작용을 고려하여 이론적으로 더 정확하게 예측할 수 있지만, 거동에 대한 모든 요인을 고려하기 어렵고 해석결과가 현장 계측 결과와 불일치하는 문제가 자주 발생한다(Goh et al., 2017).

근래 지반공학에서는 데이터 기반의 머신러닝 기법이 다수 활용되고 있다. 흙막이 벽체의 변형 예측을 위해 머신러닝 기법을 적용한 다양한 연구를 통해 ML(Machine Learning)방법(ANN(Artificial Neural Network), RF(Random Forest), GRNN(General Regression Neural Network), SVM(Support Vector Machine), MARS(Multivariate Adaptive

Regression Splines), XGB(eXtreme Gradient Boosting))의 적용 가능성이 입증되었다(Kim et al., 2000; Kung et al., 2007; Goh et al., 1995; Zhang et al., 2020a; Zhang et al., 2021). 대부분의 연구는 수치해석 결과 또는 최대변위값을 예측변수로 하여 예측시스템을 구축하였다. Seo and Chung(2023)은 심도별 계측결과를 예측변수로 하여 수치해석 결과와 실제 계측결과 사이의 관계를 학습한 앙상블 모델을 구축하여 현장 적용성을 향상시켰다.

머신러닝 방법뿐만 아니라 딥러닝(Deep Learning) 기법 또한 주목받고 있다. 대량의 복잡한 데이터를 처리할 수 있는 다양한 딥러닝 방법 중 순환신경망(RNN, Recurrent Neural Network)은 지반 변형과 같은 시계열 문제를 처리하는 데 고유한 이점을 지니고 있다(Yang et al., 2019; Mahmoodzadeh et al., 2020; Zhang et al., 2020b). 딥러닝을 활용한 흙막이 벽체의 변형 예측에 관한 최근 연구는 다음과 같다. 굴착현장 및 흙막이 벽체의 공간적 특성을 반영하기 위해 Zhao et al.(2021)은 심도에 따른 벽체 변형 예측에 CNN(Convolutional Neural Network)을 적용했으며, CNN이 시계열 예측에서 일반적으로 사용되는 LSTM(Long Short Term Memory) 및 ARIMA(Auto-Regressive Integrated Moving Average)방법보다 우수한 성능을 보여주는 것으로 보고하였다. 이 연구는 7일 이내의 단기예측과 흙막이 벽체의 공간적 변화에 초점을 두고 있다. Seo and Chung(2022)은 흙막이 벽체 변형에 대한 1D CNN-LSTM 모델을 제안하였으며, 1D CNN과 LSTM을 결합하여 흙막이 벽체의 공간 및 시간적 특성을 모두 반영한 예측이 가능함을 보여주었다. 1D CNN과 LSTM을 결합하는 경우 모델의 파라미터가 증가하게 되어 예측모델의 정확도를 높이기 위해 시간이 소요되는 단점이 있다. Shan et al.(2024)는 공간 클러스터링과 LSTM기반의 EMD(Empirical Mode Decomposition) 방법을 이용하여 굴착현장의 배면침하, 주변 건물 및 벽체 변형을 예측하였다. 이 경우도 모델 학습에 사용되는 수많은 파라미터로 인해 예측 비용이 높고, 7일 이내의 단기예측에 초점을 두고 있는 특징이 있다.

굴착현장의 데이터는 프로젝트마다 데이터의 분포 및 특성이 다르기 때문에 수집 데이터의 전처리 방법 등에 따라 최적 예측 모형이 달라진다. 또한 지반계측은 계측 항목에 따라 측정주기가 2~3일 또는 그 이상이며, 굴착공사 기간에 따라 시계열 데이터의 길이가 다양하게 존재한다. 따라서 굴착현장의 시계열 데이터는 짧은 시퀀스(sequence)를 가진 스몰데이터의 형태로 딥러닝 기반

시계열 학습에 제한이 따른다. 본 연구는 이러한 기반계측 데이터의 특성을 고려하여 딥러닝과 전통적인 통계기반 모델을 통해 효과적인 시계열 분석 방법을 다룬다. 선행 연구에서는 흠막이 벽체 변형 예측을 위해 딥러닝 알고리즘의 조합과 단기예측 성능에 초점을 두었으며, 장기예측 및 전통적인 통계기법과의 비교 분석에 대해서는 중점적으로 다루어지지 않았다.

전통적인 통계기반 모형인 ARIMA모형은 다양한 분야에서 장단기 예측에 활용되는 대표적인 시계열 모델이다. 이후 딥러닝 방법이 발전하면서 LSTM은 시계열 예측에 최적화된 인공지능망 네트워크로 자리 잡고 있다. ARIMA는 시계열의 시차 값 간의 관계에 의존하는 반면 LSTM은 데이터 시퀀스의 패턴을 인식하는 인공지능망이다. ARIMA와 LSTM은 통계 및 딥러닝 기반으로 시계열 데이터를 예측하는 대표적인 도구이다. 본 연구에서는 흠막이 벽체 변형을 예측하기 위해 통계기반 모형과 딥러닝 기법을 활용하여 시계열 분석을 수행하고, 두 가지 접근 방법에 대한 장단점을 비교하였다. 굴착현장의 지중경사계 계측데이터를 기반으로 데이터를 전처리 하고 단변량 및 다변량 조건, 시계열 데이터 길이 변화 등에 따른 예측모델의 성능을 확인하였다. 또한 LSTM 모델을 효과적으로 활용하여 흠막이 벽체 변형을 예측하는 방법을 제안하였다.

2. 분석기법

본 연구에서는 굴착현장 흠막이 벽체 변형을 나타내는 심도별 지중경사계 데이터의 시계열 분석을 위해 자기회귀누적이동평균(ARIMA)과 여기에 외생변수를 포함할 수 있도록 변형한 ARIMAX(Autoregressive Integrated Moving Average Exogenous)가 통계기반 모델로 활용된다. 딥러닝을 활용한 알고리즘으로는 단변량 및 다변량 LSTM이 활용된다. ARIMA와 LSTM 모두 단변량 및 다변량 조건 모두 적용 가능한 모델이며, 다변량인 경우 동일 대표단면의 다른 심도에 위치한 경사계 센서값을 활용하여 흠막이 벽체 전체 변화를 고려할 수 있도록 하였다.

2.1 ARIMA 모델

ARIMA모형은 Box와 Jenkins에 의해 처음 제안되었으며, 확률론적으로 시계열 데이터를 예측하기 위해

자기회귀(Autoregressive, AR) 모델과 이동평균(Moving Average, MA)모형을 결합한 통계 모델이다(Box et al., 2015). 즉, ARIMA는 자신의 과거 정보를 활용하는 AR 모델과 과거 오류정보를 활용하는 MA모형을 결합한 ARMA에 추세까지 고려한 모델이다. 이 모델은 과거 값과 현재 값 사이의 본질적인 관계를 식별하여 미래를 예측한다.

AR모형은 현재 시간(t)과 이전 시간(p)값 사이의 관계인 시계열 자기상관관계를 통해 나타낸다. p 차 자기회귀 모델은 AR(p)로 표현되며 식 (1)과 같다.

$$y_t = c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \epsilon_t \quad (1)$$

MA모형은 이동평균모델로서 white noise의 가중 평균 합계로 나타낸다. q 차 이동평균모델은 식 (2)와 같이 MA(q)로 나타낸다.

$$y_t = c - \beta_1 \epsilon_{t-1} - \beta_2 \epsilon_{t-2} - \dots - \beta_q \epsilon_{t-q} - \epsilon_t \quad (2)$$

AR모형과 MA모형을 결합하면 식 (3)과 같이 ARMA 모델이 된다.

$$y_t = c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \epsilon_t - \beta_1 \epsilon_{t-1} - \beta_2 \epsilon_{t-2} - \dots - \beta_q \epsilon_{t-q} \quad (3)$$

ARIMA모형은 ARMA모형을 사용하여 데이터를 d 번 차분하여 계산한 모델로 ARIMA(p, d, q)로 표현된다. ARIMA 예측 모델은 모델을 구축하기 전 데이터의 정상성(stationary)을 평가하는 것이 필수적이다. 본 연구에서는 데이터의 정상성을 추정하기 위해 확장된 디카플러 검증(Augmented Dickey-Fuller test, ADF)이 사용되었다. 데이터가 정상성을 나타낼 때까지 비정상 데이터에 차분을 적용하고, 데이터가 안정성을 나타내는 경우 p-value값이 0.05미만인 차분 횟수 d 값을 적용한다. 추가적으로 시계열 데이터의 자기상관계수(ACF, Autocorrelation Function)와 부분자기상관계수(PACF, Partial Autocorrelation Function)를 통해 데이터의 정상성을 확인한다. 그 외 파라미터 p 와 q 는 다양한 조합의 모델을 통해 가장 낮은 AIC(Akaike Information Criterion) 및 BIC(Bayesian Information Criterion)값을 생성하는 모델의 p 와 q 를 사용하였다. 이 경우 python auto.arima 함수를 이용하여 p 와 q 의 최적 조합 결정이 가능하다. ARIMA

모델은 단변량 시계열을 표현하는데 적절한 모형이며, ARIMAX 모델은 추가적인 외생변수(explanatory variable)를 포함함으로써 다변량 시계열 데이터 처리가 가능하다. ARIMAX는 다양한 외생변수 조건을 고려하고 종속 변수와 관계성을 반영하여 모형을 개선하는 특징이 있고, 다수 입력변수들의 영향력을 고려하므로 장기예측 성능을 기대할 수 있다(Suhermi et al., 2019). 하지만 외생변수를 고려하는 경우 모델 내에서 외생변수의 중요도 평가와 미래 가용성 등에 대한 분석을 통해 적용할 필요가 있다.

2.2 LSTM

LSTM은 순환신경망(RNN, Recurrent Neural Network)의 장기 의존성 문제를 개선하기 위한 신경망 모델이다(Hochreiter and Schmidhoubler, 1997). LSTM cell의 내부구조는 Fig. 1과 같으며, 시간 t 에서의 은닉층을 나타낸다. x_t 는 시간단계 t 에서의 입력벡터를 의미한다. h_{t-1} 과 h_t 는 각각 은닉층에서 시간단계 $t-1$ 과 t 에서의 은닉상태벡터(hidden state vector)이며, 단기기억(short term memory)정보를 처리한다. C_{t-1} 과 C_t 는 각각 은닉층에서 시간단계 $t-1$ 과 t 에서의 셀상태벡터(cell state vector)이고 장기기억(long term memory)정보를 담당한다. LSTM 셀 내부에는 입력벡터(input vector)와 은닉상태벡터(hidden state vector)를 처리하는 다양한 기능이 있으며, 식 (4)~(9)와 같이 정의한다.

$$f_t = \sigma(U_f h_{t-1} + b_{u_f} + W_f h_t + b_{w_f}) \quad (4)$$

$$i_t = \sigma(U_i h_{t-1} + b_{u_i} + W_i h_t + b_{w_i}) \quad (5)$$

$$o_t = \sigma(U_o h_{t-1} + b_{u_o} + W_o h_t + b_{w_o}) \quad (6)$$

$$\tilde{C}_t = \tanh(U_{\tilde{C}} h_{t-1} + b_{u_{\tilde{C}}} + W_{\tilde{C}} h_t + b_{w_{\tilde{C}}}) \quad (7)$$

$$C_t = (f_t \otimes C_{t-1}) \oplus (i_t \otimes \tilde{C}_t) \quad (8)$$

$$h_t = o_t \otimes \tanh(C_t) \quad (9)$$

시간단계 $t-1$ 에서 은닉상태벡터 h_{t-1} 은 데이터를 forget gate(f_t), input/update gate(i_t), output gate(o_t), cell input(\tilde{C}_t)으로 순전파할 때 각각 $U_f, U_i, U_o, U_{\tilde{C}}$ 텐서를 곱한 후 전달된다. 입력벡터 x_t 가 각 게이트로 전달될 때 각각 $W_f, W_i, W_o, W_{\tilde{C}}$ 텐서를 곱한 후 전달된다. 여기서 $b_{u_f}, b_{u_i}, b_{u_o}, b_{u_{\tilde{C}}}, b_{w_f}, b_{w_i}, b_{w_o}, b_{w_{\tilde{C}}}$ 는 벡터형태의 편향값(bias)이다. σ 와 \tanh 는 각각 시그모이드 및 \tanh 활성화 함수를 뜻한다. \oplus, \otimes 는 벡터의 element-wise 덧셈 및 곱셈 연산자를 의미한다. Forget gate vector는 시그모이드 활성화 함수가 적용된 값이기 때문에 $0 < f_t < 1$ 의 값을 갖는다. 따라서 이전 시간단계 $t-1$ 에서 전달된 셀상태 벡터 C_{t-1} 이 기억될 필요가 있다고 판단되면 1에 가까운 값이 되도록 학습된다. 반대로 셀상태 벡터가 기억될 필요가 없다고 판단되면 0에 가까운 값이 되도록 학습된다.

장기기억을 담당하는 셀상태 벡터 C_t 는 이전 시간단계 셀상태 벡터 C_{t-1} 에 forget gate vector를 요소별로 곱한 벡터와 input gate vector와 cell input vector를 요소

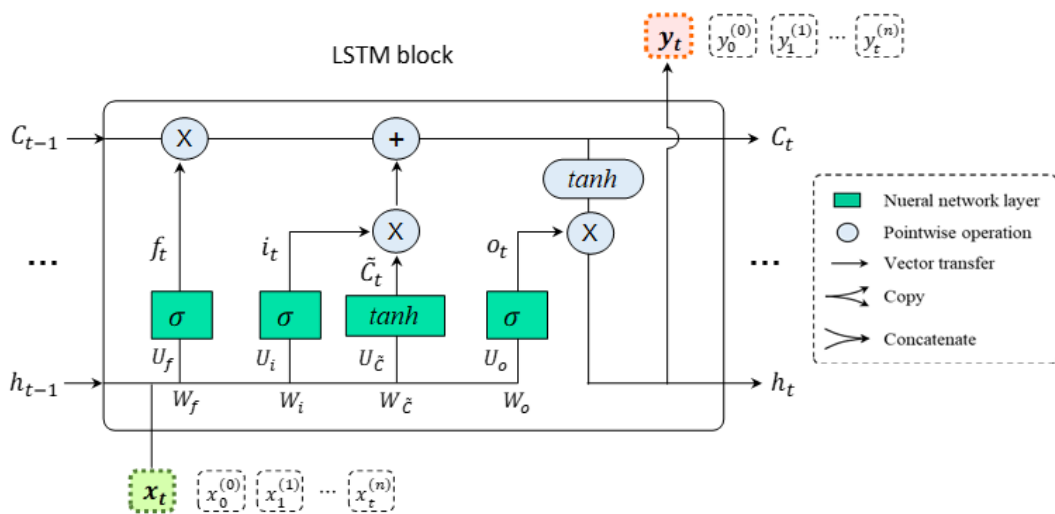


Fig. 1. LSTM cell structure

별로 곱한 벡터를 각각 요소별로 합해서 구한다. 수직전달과 수평전달 할 은닉상태벡터 h_t 는 셀상태 벡터에 활성화 함수 \tanh 를 적용한 것과 output gate vector의 요소별로 곱해서 구한다.

3. 데이터 전처리 및 성능평가 방법

3.1 데이터 수집 및 전처리

본 연구에 사용한 시계열 데이터는 청주시에 위치한

○○현장의 지중경사계 측정값이다. 해당 현장의 굴착공사 기간은 2021.6.24.~2022.1.5.이며, 굴착심도가 각각 15m, 20m인 2개 대표단면(Cross Section, CS) CS-1, 2를 분석 데이터로 사용하였다. 지중경사계 데이터는 심도 1m 간격으로 측정값이 존재하며, 약 3~7일 측정 주기로 각 심도별 1개의 측정값으로 구성되어 있다. 따라서 전체 굴착공사 기간은 195일이지만, 긴 측정주기로 인해 원시데이터의 시퀀스 길이(sequence length)는 35 steps이다. CS-1, 2의 전체 데이터는 Table 1, 2에 나타나 있다. 본 연구에서 지중경사계 원시데이터는 Fig. 2에 나

Table 1. Time series dataset for cross section-1

Depth Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2021-06-24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2021-06-28	0.23	0.56	0.71	0.62	0.51	0.21	0.23	0.22	0.20	0.15	0.10	0.05	0.02	-0.01	-0.03
2021-07-01	0.22	0.50	0.72	0.83	0.89	0.98	0.91	0.76	0.60	0.40	0.47	0.24	0.15	0.09	0.04
2021-07-05	0.38	0.89	1.22	1.26	1.25	1.12	1.07	0.92	0.74	0.50	0.54	0.27	0.16	0.08	0.02
2021-07-08	0.46	1.07	1.48	1.59	1.63	1.58	1.49	1.27	1.01	0.68	0.75	0.38	0.23	0.13	0.05
2021-07-12	0.52	1.21	1.70	1.85	1.93	1.95	1.83	1.55	1.23	0.83	0.93	0.47	0.29	0.16	0.07
2021-07-15	0.53	1.35	1.91	2.12	2.23	2.30	2.16	1.82	1.44	0.97	1.10	0.56	0.34	0.20	0.09
2021-07-19	0.71	1.75	2.44	2.64	2.73	2.70	2.54	2.15	1.71	1.15	1.29	0.65	0.40	0.22	0.09
2021-07-22	0.75	1.85	2.61	2.88	3.01	3.09	2.90	2.44	1.94	1.30	1.47	0.75	0.46	0.26	0.12
2021-07-26	0.84	2.02	2.85	3.13	3.28	3.35	3.14	2.65	2.10	1.67	1.60	0.81	0.50	0.29	0.13
2021-07-29	0.96	2.35	3.31	3.62	3.77	3.81	3.58	3.02	2.40	2.03	1.82	0.92	0.57	0.32	0.14
2021-08-02	1.03	2.53	3.55	3.87	4.02	4.05	3.81	3.21	2.55	2.05	1.93	0.98	0.60	0.34	0.14
2021-08-05	1.11	2.70	3.80	4.16	4.34	4.40	4.14	3.49	2.77	2.29	2.10	1.07	0.66	0.37	0.16
2021-08-09	1.16	2.85	4.01	4.38	4.57	4.62	4.34	3.66	2.91	2.39	2.21	1.12	0.69	0.39	0.17
2021-08-12	1.22	2.98	4.19	4.58	4.78	4.84	4.54	3.84	3.05	2.51	2.31	1.17	0.72	0.41	0.18
2021-08-19	1.26	3.08	4.33	4.74	4.94	5.00	4.70	3.96	3.15	2.59	2.39	1.21	0.74	0.42	0.18
2021-08-26	1.31	3.20	4.49	4.92	5.13	5.19	4.88	4.11	3.27	2.69	2.48	1.26	0.77	0.44	0.19
2021-09-02	1.35	3.31	4.66	5.09	5.31	5.38	5.05	4.26	3.39	2.78	2.57	1.30	0.80	0.45	0.20
2021-09-09	1.40	3.44	4.83	5.28	5.51	5.58	5.24	4.42	3.51	2.89	2.66	1.35	0.83	0.47	0.20
2021-09-16	1.46	3.56	5.01	5.48	5.71	5.78	5.43	4.58	3.64	2.99	2.76	1.40	0.86	0.49	0.21
2021-09-23	1.51	3.69	5.19	5.68	5.92	6.00	5.63	4.75	3.77	3.10	2.86	1.45	0.89	0.51	0.22
2021-09-30	1.56	3.83	5.39	5.89	6.14	6.22	5.84	4.93	3.91	3.22	2.97	1.51	0.93	0.52	0.23
2021-10-07	1.58	3.86	5.42	5.93	6.18	6.26	5.88	4.96	3.94	3.24	2.99	1.52	0.93	0.53	0.23
2021-10-14	1.59	3.88	5.46	5.97	6.22	6.30	5.92	5.00	3.97	3.26	3.01	1.53	0.94	0.53	0.23
2021-10-21	1.60	3.91	5.50	6.01	6.27	6.35	5.96	5.03	4.00	3.28	3.03	1.54	0.95	0.54	0.23
2021-10-28	1.62	3.96	5.57	6.09	6.35	6.43	6.04	5.09	4.05	3.33	3.07	1.56	0.96	0.54	0.24
2021-11-04	1.63	4.00	5.62	6.15	6.41	6.49	6.10	5.15	4.09	3.36	3.10	1.57	0.97	0.55	0.24
2021-11-11	1.65	4.03	5.66	6.19	6.46	6.54	6.14	5.19	4.12	3.38	3.12	1.58	0.97	0.55	0.24
2021-11-18	1.67	4.10	5.76	6.30	6.57	6.65	6.25	5.28	4.19	3.44	3.18	1.61	0.99	0.56	0.24
2021-11-25	1.70	4.16	5.84	6.39	6.66	6.74	6.34	5.35	4.25	3.49	3.22	1.63	1.00	0.57	0.25
2021-12-02	1.72	4.21	5.92	6.47	6.75	6.83	6.42	5.42	4.30	3.54	3.26	1.65	1.02	0.58	0.25
2021-12-16	1.71	4.18	5.88	6.43	6.70	6.79	6.38	5.38	4.27	3.51	3.24	1.64	1.01	0.57	0.25
2021-12-22	1.65	4.03	5.67	6.20	6.46	6.55	6.15	5.19	4.12	3.39	3.12	1.71	1.05	0.60	0.26
2021-12-29	1.62	3.98	5.59	6.11	6.37	6.45	6.06	5.12	4.06	3.34	3.08	1.64	1.01	0.57	0.25
2022-01-05	1.62	3.96	5.57	6.09	6.35	6.43	6.04	5.10	4.05	3.33	3.07	1.65	1.02	0.58	0.25

Table 2. Time series dataset for cross section-2

Depth Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2021-06-24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2021-06-28	0.89	0.86	0.84	0.79	0.73	0.66	0.61	0.60	0.59	0.52	0.46	0.39	0.31	0.25	0.19	0.18	0.17	0.14	0.10	0.08	0.05	0.04
2021-07-01	1.46	1.42	1.37	1.29	1.20	1.09	1.00	0.98	0.96	0.86	0.75	0.64	0.52	0.41	0.31	0.29	0.28	0.22	0.17	0.12	0.08	0.06
2021-07-05	2.26	2.19	2.13	2.00	1.85	1.68	1.54	1.52	1.49	1.33	1.17	0.98	0.80	0.64	0.49	0.46	0.43	0.34	0.26	0.19	0.12	0.10
2021-07-08	2.54	2.46	2.39	2.24	2.08	1.89	1.73	1.70	1.67	1.49	1.31	1.10	0.90	0.72	0.55	0.51	0.48	0.39	0.29	0.22	0.14	0.11
2021-07-12	2.76	2.68	2.60	2.44	2.26	2.06	1.88	1.85	1.82	1.62	1.43	1.20	0.98	0.78	0.59	0.56	0.52	0.42	0.32	0.23	0.15	0.12
2021-07-15	3.12	3.03	2.94	2.76	2.56	2.33	2.13	2.09	2.06	1.84	1.62	1.36	1.10	0.89	0.67	0.63	0.59	0.47	0.36	0.27	0.17	0.13
2021-07-19	3.34	3.24	3.14	2.95	2.74	2.49	2.28	2.24	2.20	1.96	1.73	1.45	1.18	0.95	0.72	0.67	0.63	0.51	0.39	0.28	0.18	0.14
2021-07-22	3.48	3.38	3.28	3.08	2.86	2.59	2.38	2.34	2.30	2.05	1.80	1.52	1.23	0.99	0.75	0.70	0.66	0.53	0.40	0.30	0.19	0.15
2021-07-26	3.55	3.45	3.34	3.14	2.92	2.65	2.42	2.38	2.34	2.09	1.84	1.55	1.26	1.01	0.77	0.72	0.67	0.54	0.41	0.30	0.19	0.15
2021-07-29	3.84	3.73	3.61	3.39	3.15	2.86	2.62	2.58	2.53	2.26	1.99	1.67	1.36	1.09	0.83	0.78	0.72	0.58	0.44	0.33	0.21	0.17
2021-08-02	3.91	3.80	3.68	3.46	3.21	2.91	2.67	2.62	2.58	2.30	2.02	1.70	1.38	1.11	0.84	0.79	0.74	0.59	0.45	0.33	0.21	0.17
2021-08-05	4.07	4.06	3.93	3.69	3.43	3.11	2.85	2.80	2.76	2.46	2.16	1.82	1.48	1.19	0.90	0.84	0.79	0.64	0.48	0.36	0.23	0.18
2021-08-09	4.34	4.30	4.17	3.91	3.63	3.30	3.02	2.97	2.92	2.60	2.29	1.93	1.57	1.26	0.95	0.89	0.83	0.67	0.51	0.38	0.24	0.19
2021-08-12	4.56	4.52	4.38	4.12	3.82	3.47	3.18	3.12	3.07	2.74	2.41	2.03	1.65	1.32	1.00	0.94	0.88	0.71	0.54	0.40	0.25	0.20
2021-08-19	4.83	4.79	4.64	4.36	4.05	3.67	3.36	3.31	3.25	2.90	2.55	2.15	1.74	1.40	1.06	1.00	0.93	0.75	0.57	0.42	0.27	0.21
2021-08-26	5.02	4.97	4.82	4.52	4.20	3.81	3.49	3.43	3.37	3.01	2.65	2.23	1.81	1.46	1.10	1.03	0.97	0.78	0.59	0.44	0.28	0.22
2021-09-02	5.33	5.28	5.12	4.81	4.47	4.05	3.71	3.65	3.59	3.20	2.82	2.37	1.92	1.55	1.17	1.10	1.03	0.83	0.63	0.46	0.30	0.23
2021-09-09	5.54	5.49	5.32	5.00	4.64	4.21	3.86	3.79	3.73	3.33	2.93	2.46	2.00	1.61	1.22	1.14	1.07	0.86	0.65	0.48	0.31	0.24
2021-09-16	5.89	5.83	5.66	5.31	4.93	4.48	4.10	4.03	3.96	3.53	3.11	2.62	2.12	1.71	1.29	1.21	1.13	0.91	0.69	0.51	0.33	0.26
2021-09-23	6.22	6.16	5.97	5.61	5.21	4.73	4.33	4.25	4.18	3.73	3.28	2.76	2.24	1.80	1.37	1.28	1.20	0.96	0.73	0.54	0.35	0.27
2021-09-30	6.58	6.51	6.32	5.93	5.51	5.00	4.58	4.50	4.42	3.95	3.47	2.92	2.37	1.91	1.45	1.36	1.27	1.02	0.77	0.57	0.37	0.29
2021-10-07	6.62	6.56	6.36	5.98	5.55	5.04	4.61	4.53	4.46	3.98	3.50	2.94	2.39	1.92	1.46	1.37	1.27	1.03	0.78	0.57	0.37	0.29
2021-10-14	6.66	6.60	6.40	6.01	5.58	5.06	4.64	4.56	4.48	4.00	3.52	2.96	2.40	1.93	1.46	1.37	1.28	1.03	0.78	0.58	0.37	0.29
2021-10-21	6.69	6.63	6.43	6.04	5.61	5.09	4.66	4.58	4.50	4.02	3.54	2.98	2.42	1.94	1.47	1.38	1.29	1.04	0.79	0.58	0.37	0.30
2021-10-28	6.73	6.67	6.47	6.07	5.64	5.12	4.69	4.61	4.53	4.04	3.55	2.99	2.43	1.95	1.48	1.39	1.30	1.04	0.79	0.58	0.38	0.30
2021-11-04	6.79	6.72	6.52	6.12	5.69	5.16	4.72	4.65	4.57	4.07	3.58	3.02	2.45	1.97	1.49	1.40	1.31	1.05	0.80	0.59	0.38	0.30
2021-11-11	6.83	6.77	6.56	6.16	5.72	5.19	4.75	4.67	4.59	4.10	3.61	3.04	2.46	1.98	1.50	1.41	1.31	1.06	0.80	0.59	0.38	0.30
2021-11-18	6.88	6.81	6.61	6.20	5.76	5.23	4.79	4.71	4.63	4.13	3.63	3.06	2.48	2.00	1.51	1.42	1.32	1.07	0.81	0.60	0.38	0.30
2021-11-25	6.91	6.85	6.64	6.23	5.79	5.25	4.81	4.73	4.65	4.15	3.65	3.07	2.49	2.01	1.52	1.42	1.33	1.07	0.81	0.60	0.39	0.30
2021-12-02	6.92	6.86	6.65	6.24	5.80	5.26	4.82	4.74	4.66	4.15	3.65	3.08	2.50	2.01	1.52	1.43	1.33	1.07	0.82	0.60	0.39	0.30
2021-12-16	6.80	6.74	6.53	6.13	5.70	5.17	4.73	4.65	4.57	4.08	3.59	3.02	2.45	1.97	1.49	1.40	1.31	1.05	0.80	0.59	0.38	0.30
2021-12-22	6.83	6.77	6.56	6.16	5.72	5.19	4.75	4.67	4.59	4.10	3.61	3.04	2.46	1.98	1.50	1.41	1.31	1.06	0.80	0.59	0.38	0.30
2021-12-29	6.88	6.81	6.61	6.20	5.76	5.23	4.79	4.71	4.63	4.13	3.63	3.06	2.48	2.00	1.51	1.42	1.32	1.07	0.81	0.60	0.38	0.30
2022-01-05	6.91	6.85	6.64	6.23	5.79	5.25	4.81	4.73	4.65	4.15	3.65	3.07	2.49	2.01	1.52	1.42	1.33	1.07	0.81	0.60	0.39	0.30

타낸 순서로 데이터 전처리를 수행하였다. 대표단면에 따라 굴착심도가 다르기 때문에 각 단면의 최대 굴착심도를 기준으로 정규화하였다(Fig. 3). 본 연구의 다변량 시계열 분석에서는 다른 심도의 측정값이 외생변수 및 feature로 사용되므로 심도 정규화를 통해 사용자가 원하는 변수 선택할 수 있도록 하였다. 전체 굴착심도를 0~1 범위의 값으로 스케일링하고, 정규화된 굴착심도를 hN으로 표현하였다. 지중경사계 측정값은 식 (10)의 최소-최대값을 기준으로 0~1 범위 값으로 정규화하였다.

$$y_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (10)$$

본 연구에서 심도별 흙막이 벽체의 변위 발생 정도에 따른 예측 모델의 영향을 확인하기 위해 벽체 상단(Surface), 중간(Middle), 하단(Bottom)의 분석 지점 설정하였으며, 각 지점은 hN이 0, 0.4, 1 인 위치를 나타낸다.

지중경사계 측정값의 공간적 스케일을 정규화한 후 시간간격(time step)에 대한 전처리를 수행하였다. 우선

원시데이터에서 굴착 시작과 종료 일자를 기준으로 중복된 날짜를 검색하고, 중복된 데이터 중 하나는 제외한 후 시간을 index로 하는 데이터를 생성한다. 시퀀스 길이는 전체 굴착기간에 대한 일(day)기준의 등간격 데이터로 재구성하였으며, 이 때 측정주기로 인한 결측치는 선형보간(linear interpolation)을 통해 데이터를 전처리하였다. 선형보간을 통한 내삽으로 전체 시계열 데이터의 구조는 바뀌지 않는다. 분석 현장의 굴착공사 기간이 195일이므로 데이터의 시퀀스 길이 또한 195 steps로 구성하였다. 기본적으로 굴착기간과 동일한 일 기준 시계

열 데이터를 분석에 사용하며, 시퀀스 길이에 따른 예측 모델의 성능을 비교하는 경우 일 기준 데이터에 업샘플링(upsampling) 방법을 적용하였다. 업샘플링을 통해 시계열 데이터의 시간 간격을 조밀하게 추출하는 것이 가능하다. 따라서 시퀀스 길이를 조절하는 경우 업샘플링을 통해 195 steps의 일 기준 시퀀스 데이터를 300, 500, 1,000으로 time step의 해상도를 조정하였다. LSTM 모델의 경우 시퀀스 길이가 1,000 steps보다 길어질 경우 성능저하가 발생할 우려가 있으므로 최대값을 1,000으로 설정하고 성능을 비교하였다(Li et al., 2018). 또한 LSTM 모델에서는 모델이 고려하는 과거 정보의 범위로서 최적의 window size(W)를 설정할 필요가 있다. 즉 LSTM이 기억할 수 있는 과거 정보의 양을 W로 설명할 수 있으며, 긴 W는 오래된 정보를 포함할 수 있으므로 장기 의존성을 모델링하는데 이점이 있다. 하지만 W가 증가하면 모델이 처리해야 하는 계산량이 증가하고 그만큼 시퀀스 데이터의 더 넓은 컨텍스트(context)를 파악하게 되므로 과적합을 유발할 수 있다. 따라서 분석 시계열 데이터 특성을 고려하여 시퀀스 길이에 맞는 적절한 W를 사용해야 한다. 본 연구에서는 W를 10~50 범위 내에서 반복 실험을 통해 최적의 값을 선정하였으며, 단변량의 경우 10, 다변량의 경우 30의 window size를 사용하였다.

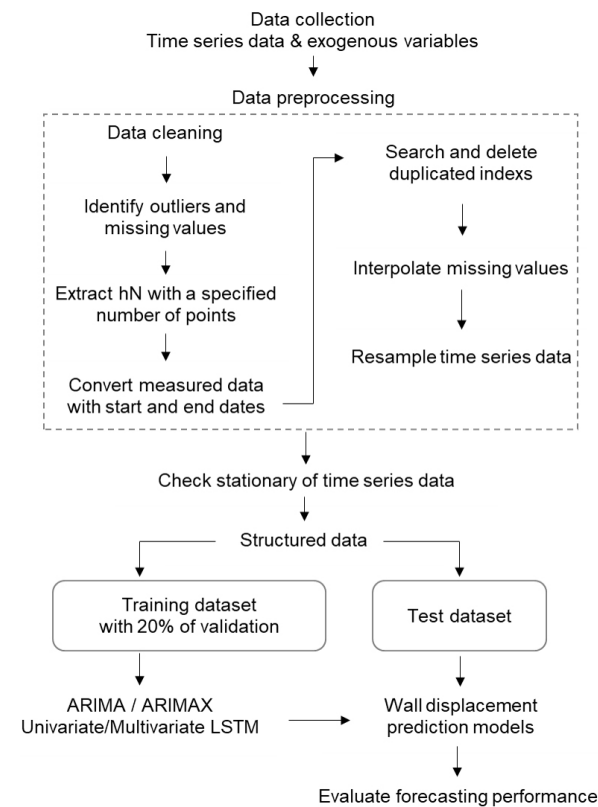


Fig. 2. Analytical procedure for predicting earth-retaining wall deformation

3.2 입력변수 및 모델 학습

기존 연구(Zhang et al., 2020a; Zhang et al., 2021; Zhao et al., 2021; Seo and Chung, 2023)에서는 흙막이 벽체의 변형을 예측하기 위해 지반정보, 수치해석 결과 등 다양한 종류의 변수가 사용되었다. 본 연구에서는 지반계측 데이터의 시계열 분석에 중점을 두었으며, 예측

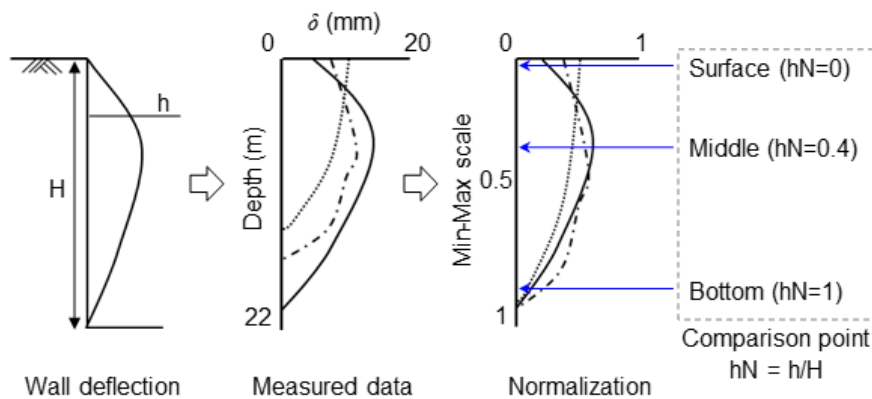


Fig. 3. Normalization schematic of inclinometer time series data

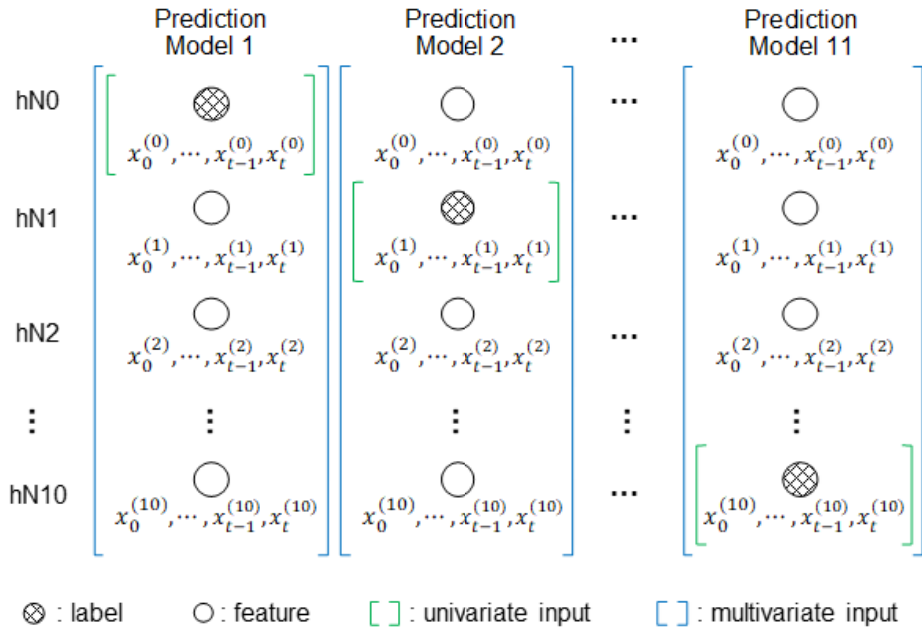


Fig. 4. Description of input variables for univariate and multivariate models

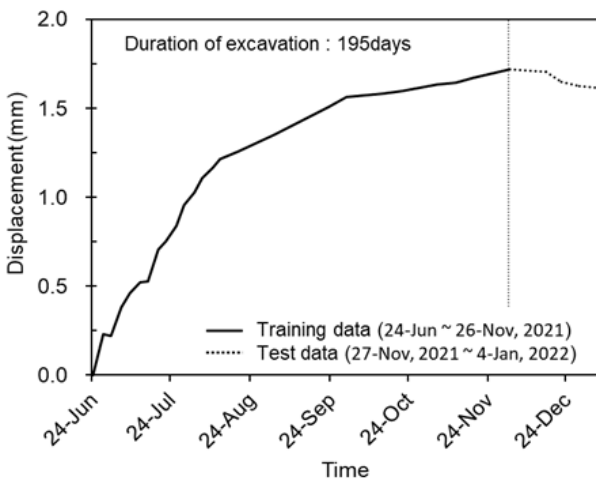


Fig. 5. Training and testing dataset for time series data

모델 해석에 대한 불확실성을 줄이기 위해 지반조사 및 부재정보, 수치해석 결과 등은 입력변수로 사용하지 않았다. 즉 굴착현장의 심도별 지중경사계 계측값만 사용하였다. 단변량 모델의 경우 분석 대상 심도의 계측값을 label로 하였으며, 다변량인 경우 분석 대상 심도는 label, 그 외 심도는 외생변수 또는 feature로 구성하였다. Fig. 4는 입력변수(hN 0~10)에 대한 단변량 및 다변량 예측 모델에 사용되는 값을 나타낸 것이다. 따라서 입력변수로 설정한 정규화된 심도값 hN의 갯수만큼 예측 모델을 생성하고, 흙막이 벽체 전체 심도에 대한 변형 예측을 수행하였다.

모델 학습에는 전체 데이터 중 20%는 테스트 데이터 세

트(test data)로 하고 나머지는 학습 데이터 세트(training data)로 설정하였다. 따라서 2021년 6월 24일부터 11월 26일까지는 학습 데이터로 2021년 11월 27일부터 2022년 1월 4일까지는 테스트 데이터가 된다(Fig. 5). 모델 학습 시 과적합(over fitting)을 방지하기 위해 학습데이터 중 20%는 검증(validation) 데이터로 사용하였다. 또한, LSTM 모델을 구성하는 다양한 파라미터에 대한 최적화를 위해 LSTM 셀 수는 10~50, 레이어는 1~4개, 최적화 함수는 SGD(Standard Gradient Decent), ADAM (Ada-Momentum)을 적용하여 최적의 값을 사용하였다.

3.3 성능평가방법

예측 모델의 성능을 평가하기 위해 본 연구에서는 평가 지표로 평균제곱근오차(Root Mean Square Error, RMSE)와 평균절대오차(Mean Absolute Error, MAE) 및 평균절대비율오차(Mean Absolute Percentage Error, MAPE)를 사용하였다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (12)$$

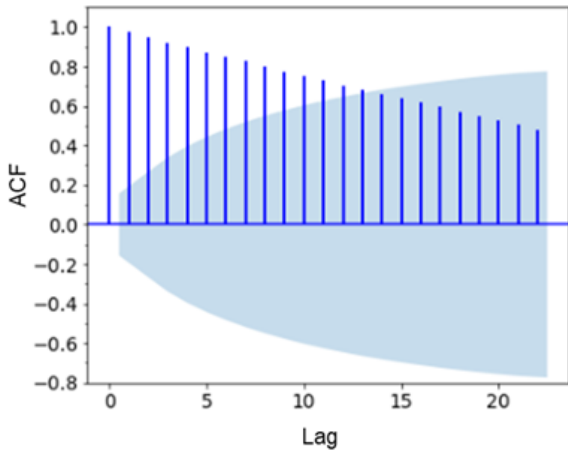
$$MAPE = \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \frac{100\%}{n} \quad (13)$$

여기서 y_i 는 실제 계측값이고, \hat{y}_i 는 모델에 의한 예측값을 나타낸다. MAE는 절대 오차값의 평균으로 오차에 대한 직관적 평가가 가능하지만 오차의 크기에 대한 민감성이 부족하다. RMSE를 통해 과소평가된 항목의 왜곡을 줄여 평가할 수 있으며, 값이 작을수록 정밀도가 높다. MAPE는 MAE를 비율로 표현한 것으로 오차에 대한 직관적 평가가 용이하고 모형 간 예측 성능 비교에 사용된다.

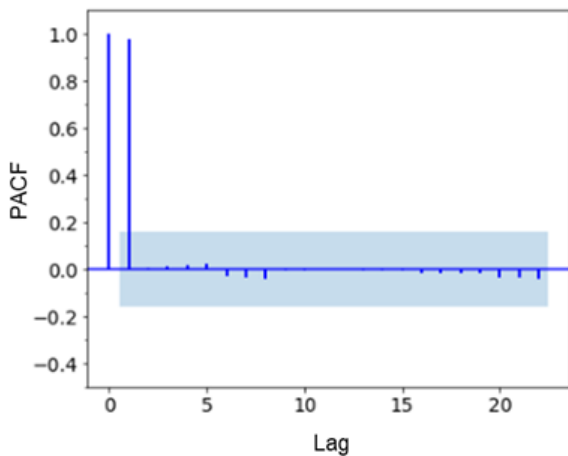
4. 실험결과

4.1 ARIMA 및 단변량 LSTM 모델의 예측 결과

ARIMA 모형과 단변량 LSTM 모델을 사용하여 심도별 지중경사계 시계열 분석을 수행하였다. ARIMA모형을 사용하기 위해서는 우선 데이터의 정상성을 확인해야



(a) ACF with respect to the lag number



(b) PACF with respect to the lag number

Fig. 6. Correlation analysis using ACF and PACF

한다. 시차가 13이상부터 자기상관계수(ACF)가 허용 범위를 초과하고, 시차가 커질수록 ACF가 점진적으로 줄어드는 형태를 보이고 있다(Fig. 6(a)). 이를 통해 지중경사계 데이터는 시간이 지날수록 데이터 간 자기 상관이 감소하는 비정상적(non-stationary) 데이터임을 확인할 수 있다. 또한 부분자기상관함수(PACF)에서 3번째 LAG 이후에 0으로 수렴하는 것을 확인하였으며, 지중경사계 데이터의 정상성을 확보하기 위해 차분 횟수 d (Lag-1)는 2로 추정하고 차분을 진행하였다. 차분 후 디카-풀러 검증 결과 p-value값이 0.000006으로 도출되어 d 는 2로 하는 경우 정상성 변환이 가능할 것으로 판단하였다. p , q 는 다양한 조합에 대해 반복적으로 AIC를 확인하였으며, 최종적으로 ARIMA모형의 최적의 p, d, q 값을 각각 4, 2, 3으로 선정하였다.

Fig. 7은 테스트 데이터 세트에 대한 ARIMA와 LSTM 모델의 예측결과를 나타낸 것이다. ARIMA와 LSTM의 MAPE는 각각 5.6042%, 1.5259%로 LSTM의 성능이 전반적으로 우수한 것을 알 수 있다. LSTM의 경우 RMSE, MAE지표 또한 ARIMA 대비 낮게 나타나 예측값의 오차가 작음을 알 수 있다. 하지만 단기 예측에 대해서는 ARIMA가 LSTM보다 유리한 점이 있다. Fig. 8은 최대 변위가 발생한 심도의 계측값과 예측결과를 보여준다. 테스트 데이터의 시작 기간인 11월 27일부터 약 4일 이내 기간에서는 ARIMA 모델의 예측성능이 높게 나타났다. 하지만 전체 약 36일 간의 장기 예측에 대해서는 LSTM은 실제 계측값과 유사한 패턴으로 예측하였고, MAPE 지표로 ARIMA 대비 높은 장기 예측 성능을 확인할 수

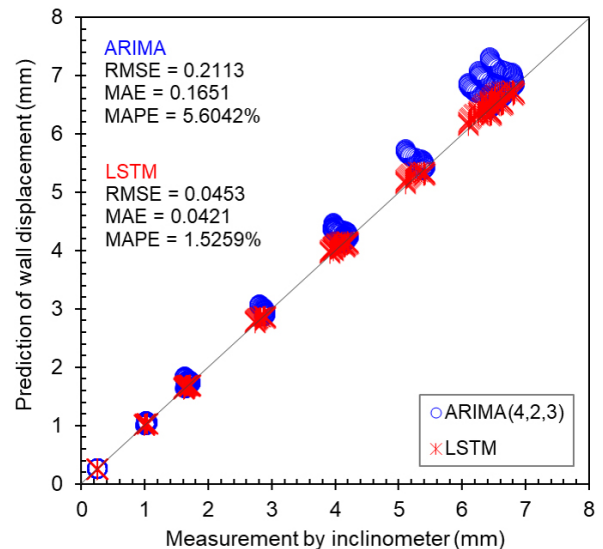


Fig. 7. Comparison of actual and predicted values for ARIMA and LSTM

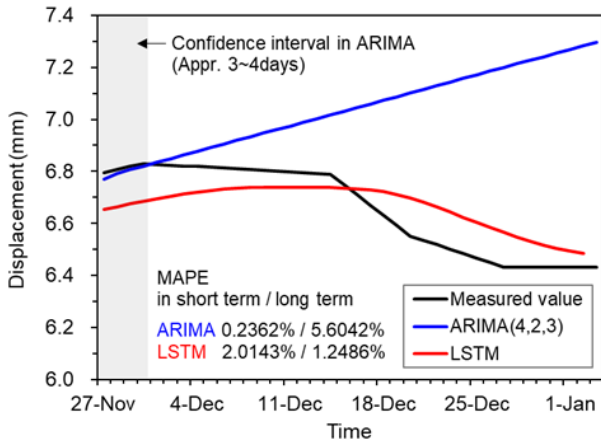


Fig. 8. Prediction results for ARIMA and LSTM (hN=0.4)

있다. 따라서 ARIMA 모델의 예측값은 약 3~4일 정도 단기 예측에 효과적인 것을 알 수 있으며, 장기에측 시 오차가 커지게 된다. ARIMA는 과거의 패턴이 미래에 반복되는 정상성(stationary)을 가정으로 하는 모델이다. 흠막이 벽체 변위 계측 데이터는 기본적으로 비정상성(non-stationary) 데이터로서 과거의 변위 패턴이 미래에 반복되어 발생하기보다 근래의 측정값이 현재 측정값과 더 큰 상관성을 가지는 특성을 가지고 있다. 따라서

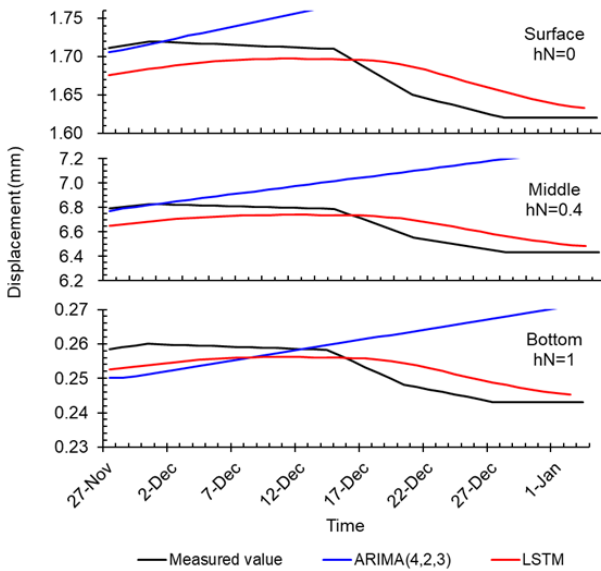


Fig. 9. Depth-specific prediction results for ARIMA and LSTM

ARIMA모델을 사용하는 경우 시계열 데이터의 차분을 통해 정상성 시계열 데이터로 변환하여 단기 예측에 보다 효과적으로 사용할 수 있다. 하지만 측정값이 갑자기 변하는 구간에서는 오차가 크게 나타나므로 사용에 주의가 필요할 것으로 판단된다. LSTM은 장단기 예측 성능이 전반적으로 우수한 편이며, 특히 흠막이 벽체 변위와 같은 비정상성 데이터에 성공적으로 적용 가능한 것을 알 수 있다.

Fig. 9는 다양한 심도에 대한 예측결과를 보여주며, 예측 성능지표는 Table 3과 같다. Table 3은 Fig. 3에서 표시한 흠막이 벽체의 비교 지점에 대해 예측결과를 나타낸 것이며, 심도별로 모델별 예측결과와의 차이를 확인할 수 있다. 전반적으로 흠막이 벽체 전체에 대해 LSTM의 오차가 모두 낮게 나타났다. 예측 모델 개별적으로 확인하였을 때, ARIMA 모델에서는 bottom에서 오차가 낮게 나타났고, LSTM은 bottom에서 오차가 비교적 크게 나타났다. ARIMA모델에서는 bottom이 전체적인 변위 변화량이 작아 상대적으로 surface와 middle 부분 대비 시계열 데이터의 비정상성이 약한 것이 원인으로 판단된다. LSTM의 경우 bottom에서 성능이 낮은 원인은 심도가 깊은 곳은 비교적 변위값의 변화가 작아 시계열 데이터의 변화 패턴이 상대적으로 확실하지 않은 것이 원인으로 판단된다. 종합적으로 ARIMA 모델은 심도별 데이터의 특성에 영향을 많이 받는 반면, LSTM은 심도에 관계없이 전반적으로 실측값의 패턴을 잘 예측하는 것으로 볼 수 있다.

4.2 시퀀스 길이에 따른 예측 성능 비교

시퀀스 길이(sequence length, L)는 시계열 예측 모델이 학습할 수 있는 정보의 양을 의미하므로 예측 모델의 성능에 영향을 줄 수 있다. 본 연구에서는 일(day) 기준 데이터를 기반으로 업샘플링(upsampling)을 통해 L을 300, 500, 1000으로 시계열 데이터의 해상도를 조정하여 예측 모델에 미치는 영향을 조사하였다. L값의 증가에 따라 ARIMA 모델의 파라미터는 최적의 값을 재선

Table 3. Depth-specific performance indicators for ARIMA and LSTM prediction models

Layer	ARIMA			LSTM		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Surface (hN=0)	0.1243	0.0968	5.8943	0.0220	0.0202	1.2011
Middle (hN=0.4)	0.4943	0.3860	5.9172	0.0874	0.0800	1.2011
Bottom (hN=1)	0.0155	0.0124	5.0011	0.0037	0.0034	1.3437

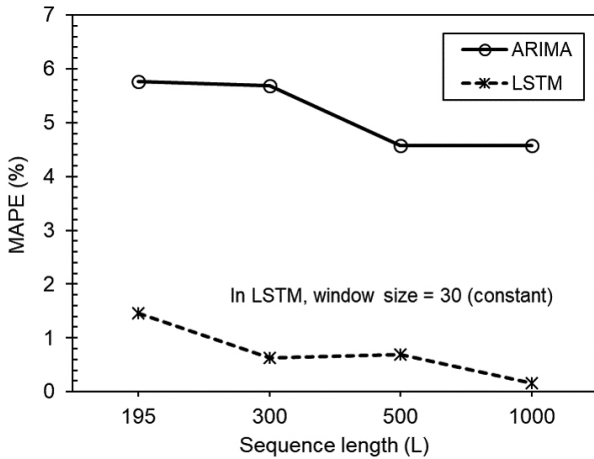


Fig. 10. MAPE results by available sequence length

정하여 사용하였고, LSTM에서는 window size를 30으로 고정하였다. Fig. 10은 L값의 증가에 따른 MAPE 변화를 보여준다. 두 예측 모델 모두 L이 증가에 따라 오차가 낮아지며 전반적인 성능이 향상되는 것을 알 수 있다. Fig. 11은 테스트 데이터에 대한 실측값과 예측값을 나타낸 것이다. ARIMA의 경우 L 증가에 따른 전반적으로 미소한 예측성능 향상은 있으나, 장기예측 성능은 개선되지 않았다. 또한 5일 이내의 단기예측 성능이 감소하는 것을 확인할 수 있었다. 반면 LSTM은 L이 증가할수록 장단기 예측성능이 모두 개선되는 것을 확인할 수 있었다. 단기예측 성능은 L이 500까지 개선되는 경향을 보이며, L이 1000이 되는 경우 단기예측 대비 장기예측 성능의 향상이 크게 나타남을 알 수 있다. 이를 통해 LSTM을 사용하는 경우 시계열 데이터의 해상도도 조절을 통해 장단기 예측성능을 향상시킬 수 있음을 알 수 있다. 따라서 ARIMA의 경우 시계열 예측 모델을 구축할 때 시계열 데이터의 원래 시간 간격을 유지하는 것이 중요하며, 업샘플링 보다 원본 데이터를 그대로 사용하고 파라미터 설정을 통해 예측 성능을 향상시키는 것이 바람직한 것으로 판단된다. 반면 LSTM의 경우 시계열 데이터의 원래 시간 간격을 조정하여 데이터를 증강하는 것이 예측 성능 향상에 효과적인 것을 알 수 있었다. 이는 데이터의 패턴 정보를 세밀하게 제공하는 효과로 인해 예측 성능이 향상된 것으로 판단된다.

시계열 데이터를 리샘플링(resampling)하는 다양한 방법 중 업샘플링이 예측 성능 향상에 가장 효과적인 것으로 알려져 있다(Semenoglou et al., 2023). 본 연구에서도 시퀀스 길이를 조정하는 방법으로 업샘플링을 적용하였으며, 이 방법은 시계열 데이터의 해상도를 조정하

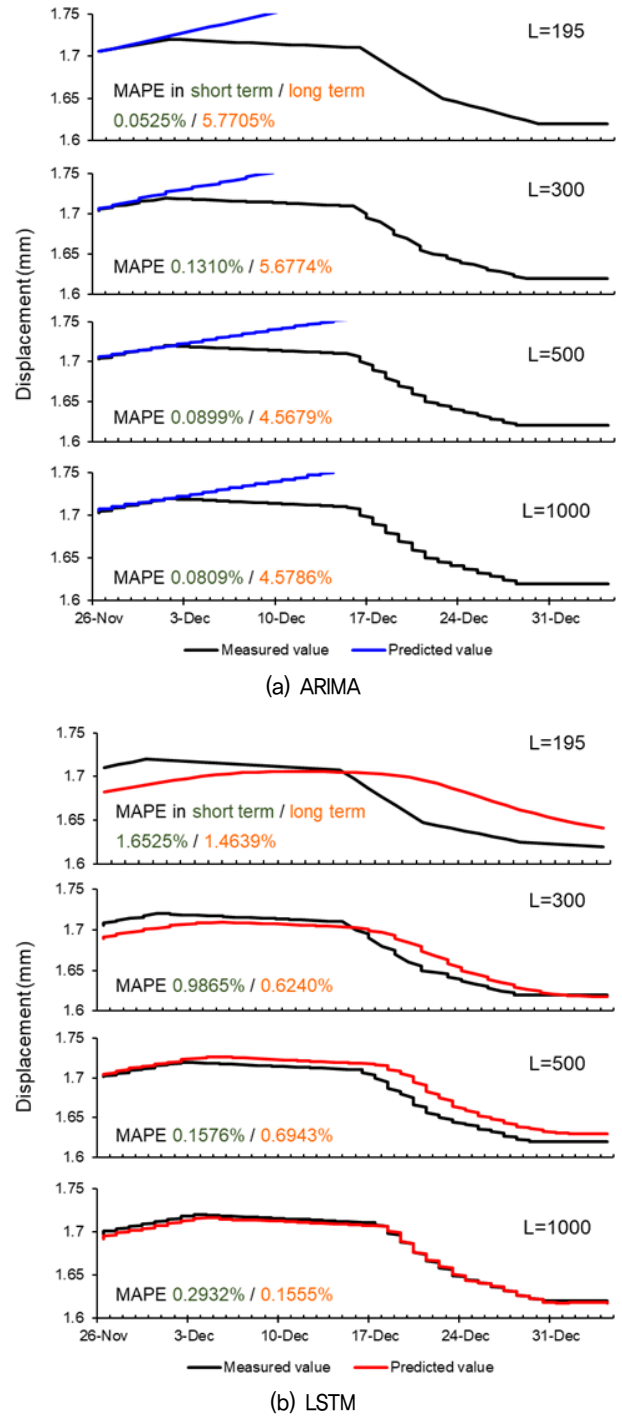
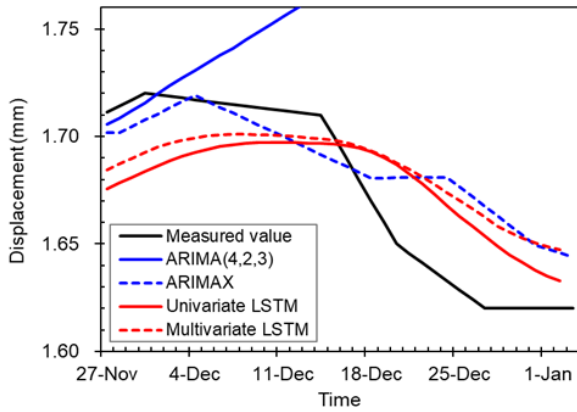
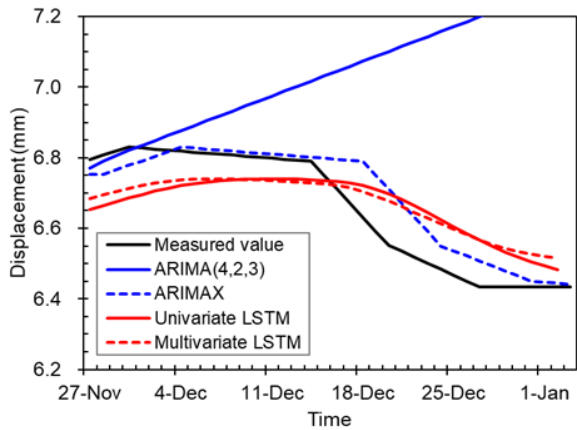


Fig. 11. Comparison of prediction results by sequence length

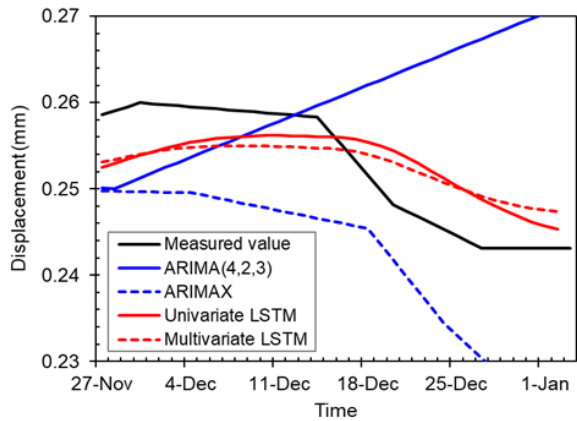
는 것으로 데이터의 구조는 유지된다. 따라서 사용 데이터에 대한 예측성능은 향상시킬 수 있으나, 다양한 특성의 데이터에 모두 효과적인 일반화 성능의 향상은 기대하기 어려운 것으로 판단된다. 일반화 성능을 향상시키기 위해서는 다양한 현장의 계측 데이터 또는 노이즈를 추가하여 시계열 데이터의 구조를 변형시켜 학습 자료로 활용하는 것이 효과적일 것으로 판단된다.



(a) Surface (N=0)



(b) Middle (hN=0.4)



(c) Bottom (hN=1)

Fig. 12. Comparison of prediction results for ARIMAX and multivariate LSTM

4.3 ARIMAX 및 다변량 LSTM 모델의 예측 결과

본 연구에서 사용한 두 가지 모델은 다변량 해석이 가능하다. 다른 심도의 지중경사계 측정값을 외생변수 또는 feature로 하여 다변량 해석 시 예측 성능을 비교하였다. Fig. 12는 비교 심도별 실측값과 예측결과를 나타내고, Table 4는 예측 모델의 성능지표를 보여준다. 단변량 모델 ARIMA와 다변량 모델 ARIMAX의 경우 예측 결과에 큰 차이를 보인다. ARIMAX 예측 결과 hN=1을 제외한 경우 예측 결과가 대폭 향상되는 것을 알 수 있다. ARIMA 모델은 외생변수로 사용한 다른 심도의 측정값에 과도한 영향을 받는 것으로 나타났다. 즉 ARIMAX 모델이 실제로 시계열의 동적 특성을 학습한 것이 아니라 단순히 외생변수인 주변 심도의 변위값 패턴을 따라가고 있음을 의미한다. 이는 외생변수의 미래 변화가 모델에 과도하게 의존적임을 의미하며, 외생변수에 지나치게 민감한 모델은 실제 환경에서 좋은 성능을 유지하기 어렵다.

LSTM의 경우 단변량과 다변량의 결과 차이가 크지 않았으며, 이는 각 심도별 변위값이 전반적으로 비슷한 패턴으로 움직이고 있으므로 예측에 큰 영향을 주지 않는 것을 알 수 있다. 다변량 LSTM의 경우 여러 입력변수를 사용하여 모델 구성이 더욱 복잡해지고 예측력이 향상되는 것이 일반적이다. 하지만 지중경사계의 변위값과 같이 시간에 따른 변화가 크지 않고 패턴이 비교적 단순한 경우, 심도별 측정값은 예측 모델에 추가적인 정보를 제공하지 않는 것으로 판단된다. 따라서 단변량 LSTM을 통해 각 심도별 예측결과를 종합하여 흠막이 전체 변형을 예측하는 것이 효과적일 것으로 판단된다. 반면 다변량의 경우 다른 심도의 측정값을 추가 변수로 사용하여 공간적 특성을 고려할 수 있는 장점이 있으므로 분석 목적에 따라 선택적으로 사용할 필요가 있다.

4.4 흠막이 벽체 변형 예측

다양한 실험결과에 따라 ARIMA와 단변량 LSTM을

Table 4. Depth-specific performance indicators for ARIMAX and multivariate LSTM prediction models

Model	Surface (hN=0)			Middle (hN=0.4)			Bottom (hN=1)		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
ARIMA	0.1243	0.0968	5.8943	0.4943	0.3860	5.9172	0.0155	0.0124	5.0011
ARIMAX	0.0211	0.0174	1.0489	0.0175	0.0119	0.1792	0.0155	0.0112	4.422
Univariate LSTM	0.0220	0.0202	1.2011	0.0874	0.0800	1.2011	0.0037	0.0034	1.3437
Multivariate LSTM (W=30)	0.0273	0.0248	1.4913	0.0923	0.0880	1.3256	0.0040	0.0037	1.4776

기반으로 흙막이 벽체 전체 변형을 예측하였다. 이 때 사용 데이터는 업샘플링 하지 않은 일 기준 데이터(L=195)를 적용하였다. Fig. 13은 CS-1, 2의 실측값과 각 모델의 예측결과를 보여주며, 테스트 데이터 세트에 대해 t+1, 17, 36의 세 가지 시점에 대한 결과를 나타낸다.

ARIMA의 경우 단기예측에 효과적인 반면 장기예측 성능은 부족한 것을 알 수 있다. 특히 변위가 크게 발생하는 지점에서는 시간이 지날수록 오차가 더욱 증가하는 것을 알 수 있다. CS-1, 2 모두 동일한 예측 양상을 보이며, CS-1에서는 $hN=0.4$ 인 지점, CS-2에서는 $hN=0$ 인 지점에서 오차가 가장 크게 나타났다. 최대변위가 크게 발생하는 벽체의 경우 시간에 따른 예측오차는 더욱

커질 것으로 예상된다.

LSTM의 경우 CS-1, 2 모두 실측값과 동일한 패턴으로 예측하며, 장단기 예측성능이 우수한 것을 확인할 수 있다. 추가적으로 Fig. 11과 같이 시계열 데이터를 업샘플링하는 경우 정확도를 크게 향상시킬 수 있을 것으로 판단된다.

흙막이 벽체 변위를 측정하는 지중경사계 측정값은 비정상성 데이터 특성이 강하다. 즉, 어제의 데이터가 오늘을 가장 잘 설명하고, 일주일, 한 달 등 과거로 갈수록 상관성이 낮아짐을 의미한다. 이러한 경우 ARIMA는 선형성 가정으로 인해 예측 정확도가 떨어지며, 한계를 극복하기 위해 데이터 차분을 통한 변환 등을 적용할

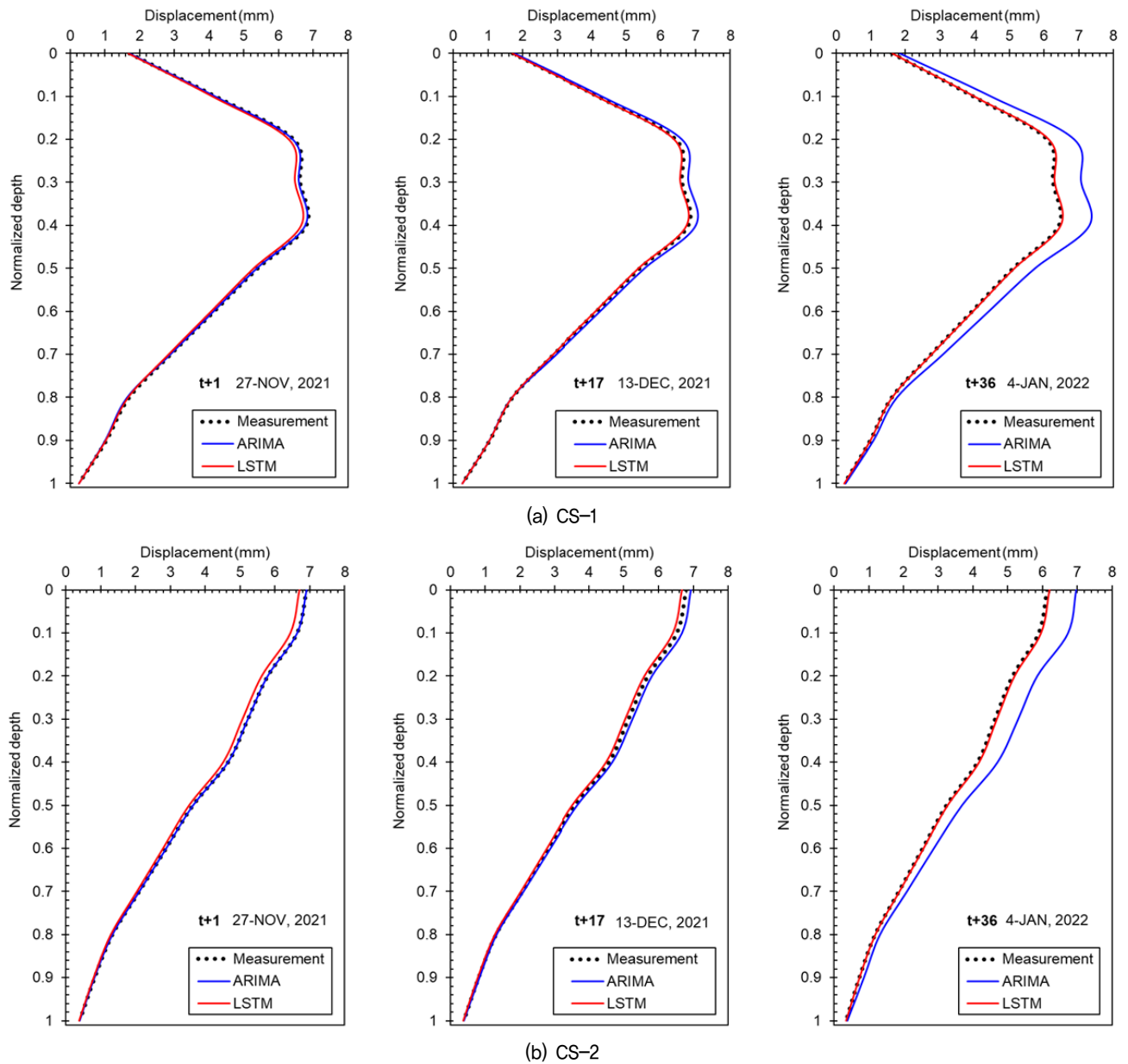


Fig. 13. Prediction results for earth retaining wall deformation using ARIMA and LSTM

수 있다. 하지만 장기적인 관점의 정확성을 보장하기 위해서는 실시간으로 파라미터를 조정하고 업데이트하여 모델의 지속적인 수정이 필요하다. 반면 LSTM의 경우 시퀀스 데이터의 복잡성과 비정상성 처리에 효과적이며, 특히 장기예측에 높은 성능을 보여주고 있다. 또한 LSTM 모델은 여러 굴착현장을 동시에 학습할 수 있으므로 다양한 시계열 데이터 특성을 고려하여 일반화 성능을 향상시킬 수 있는 장점이 있다. ARIMA와 LSTM 모두 시계열 데이터 분석에 장단점이 존재하며, 데이터의 특성을 명확히 파악한 후 적용할 필요가 있다. 본 연구의 실험결과에 따라 단변량 LSTM으로 약 98.5%의 정확도로 흙막이 벽체 변형 예측이 가능함을 확인하였다. 추가적으로 정확도를 향상시키기 위해 ARIMA의 초단기 예측과 LSTM의 중장기 예측을 결합하여 사용하는 방법 또는 증강 데이터기반의 LSTM 사용을 통해 예측 모델의 성능을 향상시킬 수 있을 것으로 판단된다.

5. 결론

본 연구에서는 굴착 현장의 지중경사계 측정 데이터를 이용하여 ARIMA와 LSTM 두 가지 예측 모델을 구축하고 비교 분석하였다. 특히 현장 데이터의 시계열 분석을 통해 흙막이 벽체의 안정성 평가와 미래 변위 예측에 있어서 전통적인 통계기반 모형과 딥러닝 방법의 차이를 비교하고, 그 정확성을 향상시키는 것을 목표로 하였다. 본 연구의 주요 결론을 다음과 같이 도출할 수 있었다.

- (1) 다양한 실험 결과 LSTM 모델이 ARIMA 모델 대비 전반적으로 우수한 예측 성능을 보였으며, 장기 예측에서 뚜렷한 성능 차이를 보였다. 이는 LSTM 모델이 시계열 데이터의 비선형 패턴과 장기 의존성을 효과적으로 모델링할 수 있음을 의미한다.
- (2) 시퀀스 길이의 증가가 두 모델의 예측 성능에 미치는 영향을 분석한 결과, 시퀀스 길이가 길어질수록 LSTM 모델의 예측 성능 향상이 더욱 두드러졌다. 반면, ARIMA 모델은 시퀀스 길이의 증가에 따른 장기 예측 성능 향상이 제한적이었다. 따라서 장기적인 지중경사계 변화 예측을 위해서는 LSTM 모델이 더 적합할 것으로 판단된다.
- (3) LSTM 모델은 단기 예측에서 다소 아쉬운 성능을 보였으나, 업샘플링과 같은 데이터 증강을 통해 이

- 를 개선할 수 있었다. 이를 통해 LSTM 모델은 장단기 예측 모두에서 높은 유연성과 적응성을 보였으며, 굴착현장의 다양한 변화에 대응할 수 있는 강력한 예측 도구로 활용될 수 있을 것으로 판단된다.
- (4) ARIMA 모델은 단기 예측에 강점을 보이나, 비정상성 데이터 처리와 모델의 지속적인 업데이트가 필요하다라는 한계가 있다. 반면, LSTM 모델은 비정상성 데이터를 효과적으로 처리할 수 있고, 데이터 증강을 통해 보다 일반화된 예측 모델을 구축할 수 있다. 따라서 굴착현장 모니터링과 같은 실무적 환경에서는 LSTM 모델의 활용이 더 효과적일 것으로 판단되며, 실시간 데이터의 모니터링 및 예측을 통해 굴착 작업의 효율성을 높이고, 예측 정보를 기반으로 한 신속한 의사 결정을 통해 재산 손실 및 인명 피해를 최소화할 수 있을 것으로 예상된다.

감사의 글

본 연구는 과학기술정보통신부 한국건설기술연구원 주요사업으로 수행되었습니다(과제번호 20240133-001, 지반분야 재난재해 대응과 미래 건설산업 신성장을 위한 지반 기술 연구).

참고문헌 (References)

1. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015), *Time Series Analysis: Forecasting and Control* (5th edition), Hoboken, New Jersey, USA, John Wiley & Sons.
2. Clough, G. W. and O'Rourke, T. D. (1990), *Construction Induced Movements of in Situ Walls, Proceedings of the Specialty Conference on Design and Performance of Earth Retaining Structures*, ASCE, Reston, VA, USA, pp.439-470.
3. Goh, A. T. C., Zhang, R. H., Wang, W., Wang, L., Liu, H. L., and Zhang, W. G. (2020), Numerical Study of the Effects of Groundwater Drawdown on Ground Settlement for Excavation in Residual Soils, *Acta Geotechnica*, Vol.15, pp.1259-1272.
4. Goh, A. T. C., Zhang, F., Zhang, W., Zhang, Y., and Liu, H. (2017), A Simple Estimation Model for 3D Braced Excavation Wall Deflection, *Computers and Geotechnics*, Vol.83, No.3, pp.106-113.
5. Goh, A. T. C., Wong, K. S., and Broms, B. B. (1995), Estimation of Lateral Wall Movements in Braced Excavations Using Neural Networks, *Canadian Geotechnical Journal*, Vol.32, No.6, pp.1059-1064.
6. Hsieh, P. G. and Ou, C. Y. (1998), Shape of Ground Surface Settlement Profiles Caused by Excavation, *Can. geotechnical J.*, Vol.35, No.6, pp.1004-1017.
7. Hochreiter, S. and Schmidhuber, J. (1997), Long Short-Term Memory, *Neural Comput.*, Vol.9, No.8, pp.1735-1780.

8. Hou, Y. M., Wang, J. H., and Zhang, L. L. (2009), Finite-element Modeling of a Complex Deep Excavation in Shanghai, *Acta Geotechnica*, Vol.4, pp.7-16.
9. Hsiung, B.C. B. (2020), Observations of the Ground and Structural Behaviors Induced by a Deep Excavation in Loose sands, *Acta Geotechnica*, Vol.15, pp.1577-1593.
10. Kim, H. T., Park, S. W., Kwon, Y. H., and Kim, J. H. (2000), Development of a System Predicting Maximum Displacement of Earth Retaining Walls at Various Excavation Stages Using Artificial Neural Network, *J. of the Korean Geotechnical Society*, Vol.16, No.1, pp.83-97.
11. Kung, G. T. C., Hsiao, E. C. L., Schuster, M., and Juang, C. H. (2007), A Neural Network Approach to Estimating Deflection of Diaphragm Walls Caused by Excavation in Clays, *Computers and Geotechnics*, Vol.34, No.5, pp.385-396.
12. Lee, S. J., Song, T. W., Lee, Y. S., Song, Y. H., and Kim, H. K. (2007), A Case Study of Building Damage Risk Assessment Due to the Multi-propped Deep Excavation in Deep Soft Soil, *Proceedings of the 4th Int. Conf. Soft Soil Eng.*, Vancouver, London, Taylor Francis.
13. Lee, S. and Kim, S. K. (2008), A Study on Deformation Analysis of the Earth Retaining Wall, *J. of the Korean Geotechnical Society*, Vol.24, No.2, pp.27-36.
14. Li, M. G., Xiao, X., Wang, J. H., and Chen, J. J. (2019), Numerical Study on Responses of an Existing Metro Line to Staged Deep Excavations, *Tunn. Undergr. Space Technol.*, Vol.85, pp.268-281.
15. Li, S., Li, W., Cook, C., Zhu, C., and Gao, Y. (2018), Independently Recurrent Neural Network (IndRNN): Building a Longer and Deeper RNN, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp.5457-5466.
16. Lim, A., Ou, C. Y., and Hsieh, P. G. (2018), Investigation of the Integrated Retaining System to Limit Deformations Induced by Deep Excavation, *Acta Geotechnica*, Vol.13, pp.973-995.
17. Long, M. (2001), Database for Retaining Wall and Ground Movements Due to Deep Excavations, *J. Geotech. Geoenviron. Eng.*, Vol.127, No.3, pp.203-224.
18. Mahmoodzadeh, A., Mohammadi, M., Daraei, A., Ali, H.F.H., Al-Salihi, N.K., and Omer, R.M.D. (2020), Forecasting Maximum Surface Settlement Caused by Urban Tunneling, *Automation in Construction*, Vol.120, 103375.
19. Meng, F. F., Chen, R. P., Wu, H.N., Xie, S.W., and Liu, Y. (2020), Observed Behaviors of a Long and Deep Excavation and Collinear Underlying Tunnels in Shenzhen Granite Residual Soil, *Tunn. Undergr. Space Technol.*, 103:103504.
20. Moormann, C. (2004), Analysis of Wall and Ground Movements Due to Deep Excavation in Soft Soil Based on a New Worldwide Database, *Soils and Foundations*, Vol.44, No.10, pp.87-98.
21. Peck, R. B. (1969), Deep Excavations and Tunneling in Soft Ground, *Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering*, Mexico City, pp.225-290.
22. Seo, M. W., Seok, J. W., Yang, K. S., and Kim, M. M. (2006), Sequential Analysis of Adjacent Ground Behavior Caused by Deep Excavations, *J. of the Korean Geotechnical Society*, Vol.22, No.2, pp.19-28.
23. Seo, S. and Chung, M. (2023), Development of an Ensemble Prediction Model for Lateral Deformation of Retaining Wall under Construction, *J. of the Korean Geotechnical Society*, Vol.39, No.4, pp.5-17.
24. Seo, S., Park, J., Ko, Y., Kim, G., and Chung M. (2023), Geotechnical Factors Influencing Earth Retaining Wall Deformation During Excavations, *Front. Earth Sci.*, 11:1263997.
25. Seo, S. and Chung, M. (2022), Evaluation of Applicability of 1D-CNN and LSTM to Predict Horizontal Displacement of Retaining Wall According to Excavation Work, *Int. J. Adv. Comput. Sci. Appl.*, Vol.13, No.2, pp.86-91.
26. Semenoglou, A. A., Spiliotis, E., and Assimakopoulos, V. (2023), Data Augmentation for Univariate Time Series Forecasting with Neural Networks, *Pattern Recognition*, Vol.134, 109132.
27. Shan, J., Zhang, X., Liu, Y., Zhang, C., and Zhou, J. (2024), Deformation prediction of large-scale civil structures using spatio-temporal clustering and empirical mode decomposition-based long short-term memory network, *Automation in Construction*, 158, 105222,.
28. Son, M. and Cording, E. J. (2005), Estimation of Building Damage Due to Excavation-induced Ground Movements, *Geotech. Geoenvironmental. Eng.*, Vol.131, No.2, pp.162-177.
29. Son, M. and Cording, E. J. (2007), Evaluation of Building Stiffness for Building Response Analysis to Excavation-induced Ground Movements, *Geotech. Geoenvironmental. Eng.*, Vol.133, No.8, pp. 995-1002.
30. Suhermi, N., Permata, R. P., and Rahayu, S. P. (2019), Forecasting the Search Trend of Muslim Clothing in Indonesia on Google Trends Data Using ARIMAX and Neural Network, *Proceedings of Soft Computing in Data Science: 5th International Conference*, Lizuka, Japan, pp.272-286.
31. Wong, I. H. and Poh, T. Y. (2000), Effects of Jet Grouting on Adjacent Ground and Structures, *J. Geotechnical Geoenvironmental Eng.* Vol.126, No.3, pp.247-256.
32. Yang, B.B., Yin, K.L., Lacasse, S., and Liu, Z.Q., (2019), Time Series Analysis and Long Short-term Memory Neural Network to Predict Landslide Displacement, *Landslides*, 16, pp.677-694.
33. Zhao, H. J., Liu, W., Shi, P. X, Du, J. T., and Chen, X. M. (2021), Spatiotemporal Deep Learning Approach on Estimation of Diaphragm Wall Deformation Induced by Excavation, *Acta Geotechnica*, Vol.16, pp.3631-3645.
34. Zhang, W., Zhang, R., Wu C., Goh, A.T.C., Lacasse, S., Liu, Z., and Liu H. (2020a), State-of-the-art Review of Soft Computing Application in Underground Excavations, *Geoscience Frontiers*, Vol.11, No.4, pp.1095-1106.
35. Zhang, R., Wu, C., Goh, A. T. C., B'ohlke, T., and Zhang, W. (2021), Estimation of Diaphragm Wall Deflections for Deep Braced Excavation in Anisotropic Clays Using Ensemble Learning, *Geoscience Frontiers*, Vol.12, No.1, pp.365-373.
36. Zhang, P., Wu, H.N., Chen, R.P., Dai, T., Meng, F.Y., and Wang, H.B. (2020b), A Critical Evaluation of Machine Learning and Deep Learning in Shield-ground Interaction Prediction, *Tunn. Undergr. Space Technol.*, Vol.106, 103593.
37. Zhang, W., Goh, A. T., and Xuan, F. (2015), A Simple Prediction Model for Wall Deflection Caused by Braced Excavation in Clays, *Comput. Geotechnics*, Vol.63, pp.67-72.

Received : March 8th, 2024

Revised : March 28th, 2024

Accepted : March 28th, 2024