

LSTM 모델 입력 조건에 따른 유의파고 예측 성능 변화 분석

Performance Analysis of Significant Wave Height Prediction under Different LSTM Model Input Conditions

이욱재* · 김태균** · 김범규*** · 고동휘****

Uk-Jae Lee*, Tae-Kyun Kim**, Bum-Kyu Kim*** and Dong-Hui Ko****

요지 : 본 연구에서는 Long Short-Term Memory(LSTM) 모델을 활용하여 국내 연안 8개 지점의 유의파고 예측 성능을 평가하고, 입력 조건에 따른 예측 성능 변화를 체계적으로 분석하였다. 이를 위해 입력 길이, 입력 변수 구성 및 예측 시간을 주요 설계 변수로 설정하고, 각 조건 변화에 따른 모델 성능을 비교·분석하였다. 분석 결과, 입력 길이는 지점별로 최적값에 차이가 있으나 전반적으로 24시간에서 안정적인 성능을 보였다. 입력 변수 구성 측면에서는 단일 변수보다 다중 변수를 사용하는 경우 예측 성능이 향상되었으며, 특히 9개 변수를 적용한 경우 전 지점에서 일관되게 가장 우수한 성능을 나타냈다. 또한 예측 시간이 증가함에 따라 모델 성능은 점진적으로 감소하는 경향을 보였으며, 1~4시간의 단기 예측 구간에서는 높은 설명력과 낮은 오차를 보였다. 아울러, 약 6시간까지는 RSR 0.60 이하 및 R^2 0.65 이상의 성능을 유지하여 실용적인 수준의 예측 성능이 확보되었으나, 10시간 이상에서는 성능 저하가 뚜렷하게 나타났고, 24시간 예측에서는 예측 신뢰성이 크게 감소하였다. 이러한 결과는 LSTM 기반 유의파고 예측에서 입력 조건이 모델 성능에 중요한 영향을 미치며, 특히 단기 예측에 적합한 특성을 가진다는 것을 보여준다.

핵심용어 : LSTM, 유의파고 예측, 입력 길이, 입력 변수, 예측 시간, 성능 평가

Abstract : In this study, a Long Short-Term Memory (LSTM) model was employed to evaluate the prediction performance of significant wave height at eight coastal stations along the Korean coast, and to analyze how prediction performance varies with input conditions. Input length, input variable configuration, and prediction lead time were defined as key design variables, and model performance was compared under different conditions. The results show that the optimal input length varies by station, while stable performance is generally achieved at 24 hours. In terms of input variables, the use of multiple variables improves prediction performance compared to a single variable, and the best performance is consistently obtained when nine variables are used across all stations. Model performance decreases with increasing lead time. High accuracy with low error is observed for short-term predictions (1~4 hours). Performance remains within $RSR \leq 0.60$ and $R^2 \geq 0.65$ up to approximately 6 hours, whereas noticeable degradation occurs beyond 10 hours. Prediction reliability is significantly reduced at the 24-hour forecast horizon. These results indicate that input conditions have a strong influence on model performance in LSTM-based significant wave height prediction and confirm that the model is well suited for short-term forecasting.

Keywords : LSTM, significant wave height prediction, input length, input variable, lead time, performance evaluation

1. 서 론

해양 및 연안에서 파랑은 해양 구조물 설계, 해상 작업 수행, 선박 운항 및 해양 에너지 등 다양한 분야에 직접적인 영향을 미치는 핵심 요소이며, 특히 유의파고는 해양 상태를 대

표하는 주요 변수로서 정확한 예측은 해상 공사, 해양 안전 및 자원 관리 측면에서 필수적이다(James et al., 2018; Meng et al., 2021; Park and Kang, 2024). 유의파고 예측의 정확도는 작업 가능 여부 판단, 항로 계획 수립 및 에너지 생산량 산정과 같은 실제 의사결정 과정에 직접적으로 영

*한국해양과학기술원 해양공간개발·에너지연구부 연수연구원(Post Doctoral Scientist, Ocean Space Development and Energy Research Department, Korea Institute of Ocean Science and Technology)

**한국해양과학기술원 해양빅데이터·AI센터 UST 학생연구원(UST Student, Marine Bigdata and AI Center, Korea Institute of Ocean Science and Technology)

***한국해양과학기술원 해양력강화·방위연구부 일용근로자(Short-term Assistant, Sea Power Reinforcement·Security Research Department, Korea Institute of Ocean Science and Technology)

****한국해양과학기술원 해양공간개발·에너지연구부 책임연구원(Corresponding author: Dong-Hui Ko, Principal Research Scientist, Ocean Space Development and Energy Research Department, Korea Institute of Ocean Science and Technology, 385, Haeyang-ro, Yeongdo-gu, Busan 49111, Korea, Tel: +82-51-664-3530, kodh02@kiost.ac.kr)

향을 미치며, 예측 오차는 작업 지연, 비용 증가 및 안전성 저하로 이어질 수 있다. 특히 해상 작업의 경우 제한된 기상 조건에서 수행되므로 단기 예측의 정확성은 작업 일정 관리와 안전 확보 측면에서 매우 중요한 요소로 작용한다.

기존의 파랑 예측 방법은 수치모델 기반 접근과 통계적 접근으로 구분된다. 수치모델은 물리 기반 방정식을 통해 파랑의 발생 및 전파를 정밀하게 표현할 수 있으나, 높은 계산 비용과 복잡한 입력자료 요구로 인해 실시간 예측 및 현장 적용에 한계가 존재한다(Jorges et al., 2021; Chen et al., 2021). 반면 통계적 방법은 계산 효율성이 높고 구현이 용이하다는 장점이 있으나, 비선형성과 시간적 의존성이 강한 해양 시계열 데이터의 특성을 충분히 반영하지 못하는 한계가 있다. 이러한 한계를 보완하기 위해 최근에는 인공지능망을 활용한 데이터 기반 접근법이 활발히 연구되고 있으며, 장단기 기억(Long Short-Term Memory, LSTM) 모델은 시계열 데이터의 장기 의존성을 효과적으로 반영할 수 있어 유의파고 예측에 널리 활용되고 있다(Fan et al., 2020; Gao et al., 2021; Jorges et al., 2021; Minuzzi and Farina, 2023). 특히, LSTM 모델은 비선형성과 시간적 패턴을 동시에 고려할 수 있어 복잡한 해양 환경 자료에 적합하며, 다양한 연구에서 기존 방법 대비 향상된 예측 성능을 보인 바 있으며(Zhou et al., 2021; Bai et al., 2022; Bethel et al., 2021; Meng et al., 2021; Lee et al., 2021), 이러한 특성으로 인해 최근 유의파고와 같은 해양 시계열 예측 분야에서 널리 활용되고 있다. 그러나, 입력 자료의 품질과 구성에 따라 성능 변동성이 크며, 고파고와 같은 극한 이벤트 재현에는 한계가 존재할 수 있다(Minuzzi and Farina, 2023).

한편, 기존 연구들은 대부분 특정 입력 조건 하에서 모델 성능을 평가하는 데 초점을 두고 있으며, 입력 조건 변화에 따른 성능 특성에 대한 체계적인 분석은 제한적으로 수행되어 왔다. 특히 입력 길이(input length), 입력 변수 구성, 예측 시간(lead time)과 같은 주요 설계 요소는 서로 독립적으로 작용하는 것이 아니라 상호작용을 통해 모델 성능에 영향을 미칠 수 있음에도 불구하고, 이를 통합적으로 고려한 분석은 충분히 이루어지지 않았다. Hao et al.(2023)은 입력 길이와 예측 시간이 예측 성능에 중요한 영향을 미치며 입력 길이 증가가 반드시 성능 향상으로 이어지지 않음을 제시하였고, Jorges et al.(2021)과 Chen et al. (2021)은 입력 변수 구성에 따라 모델 성능이 달라질 수 있음을 제시한 바 있다. 또한 Shin et al.(2024)과 Ali et al.(2021)은 입력 변수 선택의 중요성을 강조하였으며, Choi and Lee(2018)은 시계열 예측에서 입력 데이터 구성 방식이 모델 성능에 영향을 미친다는 점을 확인한 바 있다. 그럼에도 불구하고 이러한 연구들은 개별 요소에 대한 분석에 국한되어 있어 실제 모델 설계 과정에서 요구되는 최적 입력 조건을 도출하는 데에는 한계가 있다. 이러한 한계는 특정 조건에서 도출된 연구 결과를 다양한 해역 및 운영 조건에 일반화하는 데 어려움을 초래하며,

실무 적용 시 입력 조건 설정에 대한 불확실성을 증가시키는 요인으로 작용한다. 특히 입력 조건의 부적절한 경우, 예측 정확도 저하로 이어질 수 있으며, 이는 해상 작업 일정 계획, 안전 관리 및 운영 의사결정 과정에서 오류를 유발할 가능성이 있다. 따라서 입력 길이, 입력 변수 구성 및 예측 시간과 같은 주요 설계 요소를 통합적으로 고려한 체계적인 성능 분석이 필요하다. 더 나아가 기존 연구들이 주로 모델 구조 개선과 예측 정확도 향상 자체에 초점을 둔 반면, 본 연구는 입력 조건 변화에 따른 예측 성능 특성을 체계적으로 분석함으로써 실제 유의파고 예측 적용을 위한 최적 입력 조건과 활용 가능 범위를 제시하고자 하였다.

따라서, 본 연구에서는 국내 연안 8개 지점의 유의파고 관측자료를 활용하여 LSTM 기반 예측 모델을 구축하고, 다양한 입력 조건에 따른 예측 성능을 체계적으로 비교하고 분석하기 위하여 입력 길이(1 h, 4 h, 8 h, 12 h, 24 h, 36 h, 48 h, 60 h, 72 h), 입력 변수 구성(1개, 3개, 9개 변수), 그리고 예측 시간(1 h, 2 h, 4 h, 6 h, 8 h, 10 h, 12 h, 24 h)에 대한 모델 성능을 평가하였다.

2. 관측자료 및 방법

2.1 국내 파랑 관측자료

본 연구에서는 LSTM 기반 유의파고 예측 모델을 구축하고, 입력 길이, 입력 변수 구성 및 예측 시간 간격이 예측 성능에 미치는 영향을 체계적으로 평가하기 위하여, 우리나라 연안에서 관측된 유의파고 자료를 활용하였다. 분석 대상 지점은 동해, 서해, 남해 및 제주 해역을 모두 포함할 수 있도록 기상청 해양기상부이 관측 지점 중 총 8개소를 선정하였으며, 각 지점의 위치는 Fig. 1에, 지점별 세부 정보는 Table 1에 제시하였다.

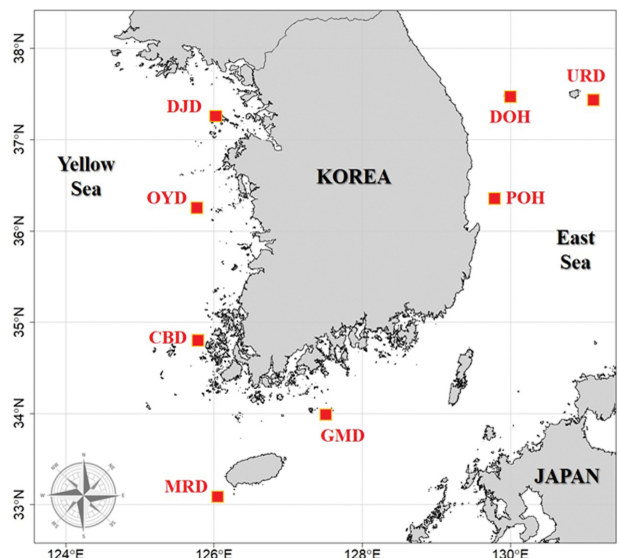


Fig. 1. Location map of analysis stations in Korea.

Table 1. Summary of significant wave height observation sites

No	Station (Code)	Longitude, Latitude (°)	Recording Period	Depth (m)	Total Data (Missing Ratio, %)	Material
1	Deokjeokdo (DJD)	126.02°E, 37.24°N	1996/07/05~2024/05/27	30.0	218,006 (1.52)	Ocean Data Buoy
2	Oeyeondo (OYD)	125.75°E, 36.25°N	2010/01/01~2025/12/31	45.0	125,958 (1.40)	
3	Chilbaldo (CBD)	125.78°E, 34.79°N	1996/05/29~2022/04/22	33.0	209,119 (1.79)	
4	Marado (MRD)	126.03°E, 33.08°N	2008/11/14~2025/12/31	104.0	138,906 (1.96)	
5	Geomundo (GMD)	127.50°E, 34.00°N	1998/01/07~2025/12/31	80.0	214,184 (1.03)	
6	Pohang (POH)	129.78°E, 36.35°N	2008/11/14~2025/12/31	310.0	143,073 (1.10)	
7	Donghae (DOH)	129.95°E, 37.48°N	2001/01/28~2025/12/31	1,518.0	184,299 (1.37)	
8	Ulleungdo (URD)	131.11°E, 37.46°N	2011/12/28~2025/12/31	2,200.0	119,751 (2.36)	

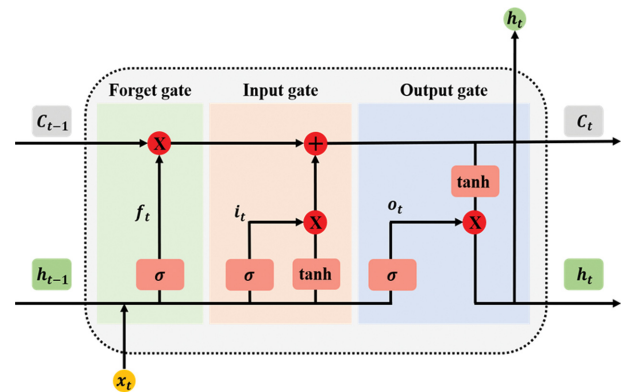
선정된 8개 지점은 덕적도(DJD), 외연도(OYD), 칠발도(CBD), 마라도(MRD), 거문도(GMD), 포항(POH), 동해(DOH), 울릉도(URD)이며, 수심은 약 30 m에서 2,200 m까지 분포하여 연안역부터 외해에 이르는 다양한 해양환경 조건을 포함한다. 관측 기간은 지점별로 상이하게 나타났으며, DJD와 GMD 지점은 약 28년에 달하는 장기 관측 자료를 포함하는 반면, URD 지점은 약 14년으로 상대적으로 짧은 관측 기간을 보였다.

기상청 해양기상부이는 해양기상 관측을 목적으로 설치된 장비로, 풍속, 풍향, GUST 풍속, 기압, 습도, 수온 등의 기상 요소와 최대파고, 평균파고, 유의파고, 파주기, 파향 등의 해양 정보를 1시간 간격으로 관측한다. 아울러, 수집된 관측자료에는 결측값이 포함되어 있어 시계열 학습에 적합한 자료 구성을 위해 전처리를 수행하였다. 단기 결측 구간에 대해서는 시계열의 연속성을 유지하기 위해 인접 관측값의 이동평균을 적용하여 보간하였으며, 일정 시간 이상 지속되는 연속 결측 구간은 보간에 따른 왜곡 가능성을 고려하여 분석에서 제외하였다(Kong et al., 2017; Hewamalage et al., 2021).

2.2 LSTM 모델 개요

LSTM(Long Short-Term Memory)은 순환신경망(Recurrent Neural Network, RNN) 기반의 시계열 예측 모델로, 기존 RNN에서 발생하는 기울기 소실 및 기울기 발산 문제를 효과적으로 해결하기 위해 제안되었다(Hochreiter and Schmidhuber, 1997). 기존 RNN은 시계열 길이가 증가함에 따라 장기 의존성을 학습하는 데 한계가 있으나, LSTM은 cell state와 gate 기반의 정보 제어 메커니즘을 도입함으로써 장기적인 시계열 의존성을 안정적으로 학습할 수 있도록 설계되었다.

LSTM은 cell state와 세 가지 gate(forget gate, input gate, output gate)로 구성되며, 각 gate는 입력 벡터와 이전 시점의 hidden state를 기반으로 정보의 유지, 갱신, 출력 여부를 결정한다. forget gate는 이전 시점의 cell state에서 유지할 정보를 선택하고, input gate는 현재 시점의 입력으로부터 새로운 정보를 선택적으로 반영한다. 이후 cell state는 이전 상태와 새로운 정보의 결합을 통해 갱신되며, output gate는 최종적으로 hidden state를 생성하여 다음 시점으로 전달한다. 이러한 구조를 통해 LSTM은 시계열 데이터 내 중요한 정보를

**Fig. 2.** Schematic diagram of LSTM unit.

선택적으로 유지하면서 불필요한 정보는 효과적으로 제거할 수 있으며(Yu et al., 2019), 해당 모델은 다음 식(1)~(6) 및 Fig. 2로 나타낼 수 있다.

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (3)$$

$$g_t = \sigma(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \quad (4)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot g_t \quad (5)$$

$$h_t = o_t \odot \tanh(C_t) \quad (6)$$

여기서, x_t 는 현재 시점의 입력벡터, h_{t-1} 은 이전 시점의 hidden state, C_{t-1} 은 이전 시점의 cell state를 의미한다. f_t , i_t , o_t 는 각각 forget, input, output gate의 활성화값이며, g_t 는 후보 cell state이다. 또한, σ 는 sigmoid 함수, $\tanh$ 는 탄젠트 하이퍼볼릭 함수, $\odot$ 는 원소별 곱셈을 나타낸다.

2.3 모델 성능평가

본 연구에서는 LSTM 기반 유의파고 예측 모델의 성능을 정량적으로 평가하기 위하여 평균절대오차(MAE, Mean Absolute Error), 결정계수(R^2), 그리고 RMSE를 표준편차로 정규화한 지표인 RSR(RMSE-observations Standard deviation

Ratio)을 사용하였다(Willmott and Matsuura, 2005; Chai and Draxler, 2014).

RSR은 관측자료의 변동성을 고려하여 모델의 오차를 상대적으로 평가할 수 있는 지표로서, 값이 0에 가까울수록 예측 성능이 우수함을 의미하며(Moriasi et al., 2007), 다음 식(7)~(8)과 같이 표현할 수 있다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (H_{s,i}^{obs} - H_{s,i}^{pred})^2} \quad (7)$$

$$RSR = \frac{RMSE}{\sigma_{obs}} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (H_{s,i}^{obs} - H_{s,i}^{pred})^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^N (H_{s,i}^{obs} - \overline{H_s^{obs}})^2}} \quad (8)$$

여기서, $H_{s,i}^{obs}$ 와 $H_{s,i}^{pred}$ 는 각 i 번째 시점에서의 관측 및 예측 유의파고, $\overline{H_s^{obs}}$ 는 관측 유의파고의 평균값을 나타내며, N 은 전체 자료 개수를 의미한다.

RSR 결과의 해석을 위해 Moriasi et al.(2007)에서 제시한 평가 기준을 참고하여 상대적인 성능 평가 기준으로 활용하였다(Table 2 참조). 해당 기준에 따르면, RSR 값이 0.6 이하인 경우 모델 성능이 양호한 수준으로 평가된다.

MAE는 예측값과 관측값 간 절대 오차의 평균으로 정의됨에 따라 이상치에 상대적으로 덜 민감하여 전체적인 오차 수준을 직관적으로 나타내며, 결정계수(R^2)는 모델이 관측값의 변동성을 얼마나 잘 설명하는지를 나타내는 지표로, 1.0에 가까울수록 높은 설명력을 의미한다(Willmott and Matsuura, 2005; Nash and Sutcliffe, 1970). 이러한 지표들을 종합적으

로 활용함으로써 모델의 평균적 성능과 변동성 재현 능력을 동시에 평가하였다. 성능평가 방법은 다음 식(9)~(10)으로 표현할 수 있다.

$$MAE = \frac{1}{N} \sum_{i=1}^N |H_{s,i}^{obs} - H_{s,i}^{pred}| \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (H_{s,i}^{obs} - H_{s,i}^{pred})^2}{\sum_{i=1}^N (H_{s,i}^{obs} - \overline{H_s^{obs}})^2} \quad (10)$$

2.4 모델 구성 및 계산 조건

LSTM 기반 유의파고 예측 성능을 체계적으로 평가하기 위해 세 가지 분석 케이스를 구성하였다. Case 1은 입력 길이의 영향을 평가하기 위해 유의파고, 파주기, 파향의 3개 입력변수를 고정하고, 예측 시간 1시간 및 4시간 조건에서 1, 4, 8, 12, 24, 36, 48, 60, 72시간의 입력 길이를 비교하였다. Case 2는 입력변수 구성의 영향을 평가하기 위해 Case 1에서 도출된 최적 입력 길이를 고정하고, 예측 시간 4시간 조건에서 1) 입력변수 1개(유의파고), 2) 입력변수 3개(유의파고, 파주기, 파향), 3) 입력변수 9개(유의파고, 최대파고, 평균파고, 파주기, 파향, 풍속, GUST 풍속, 풍향, 기압)의 세 가지 입력변수 조합을 비교하였다. Case 3은 Case 1과 Case 2에서 선정된 최적 입력 길이 및 변수 구성 조건을 적용하여 예측 시간 1, 2, 4, 6, 8, 10, 12, 24시간에 대한 예측 성능을 평가하였다. 세 가지 케이스의 상세 사항은 Table 3에 제시하였다.

분석 자료는 시간 순서를 유지한 상태에서 학습, 검증, 테스트 데이터로 분할하였다. 전체 자료 중 64%를 학습 데이터, 16%를 검증 데이터, 20%를 테스트 데이터로 사용하였으며, 시계열 자료의 시간적 의존성을 유지하며 미래 정보가 과거 학습에 반영되는 데이터 누수를 방지하기 위해 무작위 분할이 아닌 순차적 분할 방식을 적용하였다(Bergmeir and Benitez, 2012; Bergmeir et al., 2018). 아울러, 각 관측 지점별 자료 길이는 상이하나, 본 연구에서는 각 지점을 독립

Table 2. Classification of model performance based on RSR

Performance Rating	RSR
Very good	$0.0 \leq RSR \leq 0.5$
Good	$0.5 < RSR \leq 0.6$
Satisfactory	$0.6 < RSR \leq 0.7$
Unsatisfactory	$RSR > 0.7$

Table 3. Summary of analysis design for LSTM-based wave height prediction

Case	Input Variables (IV)	Input Length	Lead Time (hour)
Case 1	3 (Significant Wave Height, Peak Wave Period, Wave Direction)	1, 4, 8, 12, 24, 36, 48, 60, 72 h	1, 4
Case 2	1 (Significant Wave Height)	Fixed (optimal from Case 1)	4
	3 (Significant Wave Height, Peak Wave Period, Wave Direction)		
Case 2	9 (Wind Speed, Wind Direction, GUST Wind Speed, Pressure, Maximum Wave Height, Significant Wave Height, Mean Wave Height, Wave Period, Wave Direction)	Fixed (optimal from Case 1)	4
	3 (Significant Wave Height, Peak Wave Period, Wave Direction)		
Case 3	Selected configuration (optimal input variables and input length)	Fixed (optimal from Case 1&2)	1, 2, 4, 6, 8, 10, 12, 24

Table 4. Summary of training, validation, and test datasets for each observation sites (St. = Station, Tr-Sa = Train Samples, Va-Sa = Validation Samples, Te-Sa = Test Samples)

St.	Tr-Sa	Train Period	Va-Sa	Validation Period	Te-Sa	Test Period
DJD	139,523	1996/07/05 01:00~ 2015/03/23 05:00	34,881	2015-03-23 07:00~ 2019-03-28 12:00	43,602	2019-03-28 13:00~ 2024-05-27 3:00
OYD	80,613	2010/01/01 00:00~ 2019/11/03 16:00	20,153	2019-11-03 17:00~ 2022-02-28 23:00	25,192	2022-03-01 00:00~ 2025-12-31 23:00
CBD	133,836	1996-05-29 16:00~ 2013-04-26 15:00	33,459	2013-04-26 16:00~ 2017-06-12 22:00	41,824	2017-06-12 23:00~ 2022-04-22 23:00
MRD	88,899	2008-11-14 19:00~ 2020-02-03 14:00	22,225	2020-02-03 15:00~ 2022-08-26 20:00	27,782	2022-08-26 21:00~ 2025-12-31 23:00
GMD	137,077	1998-01-07 12:00~ 2016-03-14 19:00	34,270	2016-03-14 20:00~ 2021-01-24 05:00	42,837	2021-01-24 06:00~ 2025-12-31 23:00
POH	91,566	2008-11-14 19:00~ 2019-09-20 20:00	22,892	2019-09-20 21:00~ 2022-05-29 23:00	28,615	2022-05-30 00:00~ 2025-12-31 23:00
DOH	11,7951	2001-01-28 15:00~ 2017-08-06 21:00	29,488	2017-08-06 22:00~ 2020-12-31 13:00	36,860	2020-12-31 14:00~ 2025-12-31 23:00
URD	76,640	2011-12-28 00:00~ 2020-10-27 07:00	19,160	2020-10-27 08:00~ 2023-01-13 05:00	23,951	2023-01-13 06:00~ 2025-12-31 23:00

Table 5. Configuration and hyperparameter settings of the LSTM model

Category	Parameter	Setting	Description
Output	Target variable	Hs	Significant wave height
Model	LSTM layers	1	Single-layer LSTM structure
Model	Hidden units	32	Number of neurons in the LSTM layer
Model	Activation function	tanh, sigmoid	Default activation functions in LSTM gates
Training	Optimizer	Adam	Adaptive moment estimation optimizer
Training	Learning rate	0.001	Step size for weight update
Training	Batch size	32	Number of samples per batch
Training	Epochs	100	Maximum number of training iterations
Regularization	Early stopping	Applied	Training stops when validation loss does not improve
Data processing	Normalization	Min-Max scaling	Applied based on training data

적인 시계열 자료로 간주하고 동일한 분할 기준을 적용하여 학습 및 평가를 수행하였으며, 각 지점별 학습, 검증 및 테스트 데이터의 정보는 다음 Table 4와 같다.

한편, LSTM 모델은 단일 LSTM 층, hidden unit 32개로 구성하였으며, Adam optimizer(학습률 0.001)를 사용하여 모델을 학습하였다. 손실함수는 평균제곱오차를 적용하였고, 과적합 방지를 위해 early stopping 기법을 적용하여 검증 데이터의 손실이 일정 epoch 동안 개선되지 않을 경우 학습을 종료하도록 하였다. 입력 변수 간의 스케일 차이를 고려하여 Min-Max 정규화를 적용하였으며, 학습 데이터로부터 산정한 정규화 파라미터를 검증 및 테스트 데이터에 동일하게 적용하였다(Fan et al., 2020; Hao et al., 2023). 모델 구성에 사용된 주요 하이퍼파라미터는 Table 5에 제시하였다.

3. 분석 결과

3.1 입력 길이에 따른 유의파고 예측 성능 평가 분석 결과

Fig. 3에는 각 지점별 입력 길이 변화에 따른 RSR, MAE, R^2 의 변화를 예측 시간 1시간 및 4시간에 대해 제시하였다.

R^2 의 변화를 예측 시간 1시간 및 4시간에 대해 제시하였다.

예측 시간 1시간의 경우, 전 지점에서 RSR은 0.16~0.26, MAE는 0.05~0.12 m, R^2 는 0.94~0.97 범위로 나타나 전반적으로 높은 예측 성능을 보였다. 입력 길이에 따른 성능 변화는 전 지점에 걸쳐 매우 제한적으로 나타났으며, 입력 길이 증가에 따른 뚜렷한 성능 향상 경향은 확인되지 않음에 따라 (Fig. 3a, 3c, 3e). 입력 길이의 영향이 상대적으로 작게 나타난 것으로 판단된다.

예측 시간 4시간의 경우, 전반적인 예측 성능은 1시간 예측 대비 뚜렷하게 저하되었다. RSR은 0.30~0.50, MAE는 0.13~0.25 m, R^2 는 0.78~0.88 범위로 나타났으며, 예측 시간 1시간 대비 RSR과 MAE는 약 1.8~2.1배 증가하고 R^2 는 약 0.08~0.10 감소하였다. 입력 길이에 따른 성능 변화는 예측 시간 1시간에 비해 보다 뚜렷하게 나타났으며, 초기 구간 (1~24시간)에서는 대부분의 지점에서 입력 길이 증가에 따라 RSR과 MAE가 감소하여 예측 성능이 향상되는 경향이 확인되었다. 반면 입력 길이 36시간 이상에서는 성능 개선이 정체되거나 일부 지점에서 RSR과 MAE가 오히려 증가하여 성

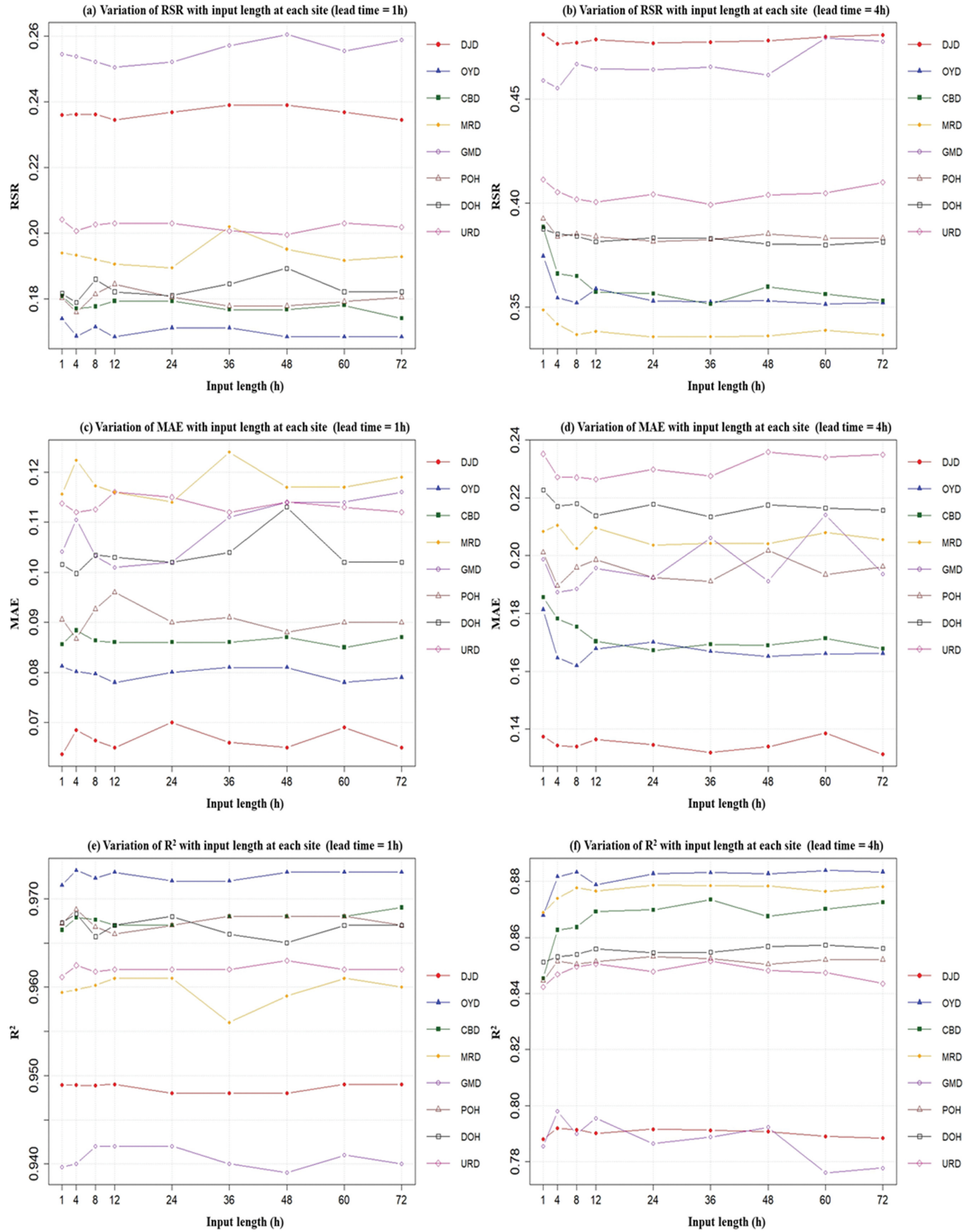


Fig. 3. Variation of RSR, MAE, and R^2 with input length at each site for different lead times.

능이 저하되는 양상이 나타났다(Fig. 3b, 3d, 3f).

지점별 RSR 최솟값을 기준으로 최적 입력 길이를 살펴보면, DOH 지점과 MRD 지점은 12시간, OYD 지점은 24시

간에서 최적 성능을 나타냈으며, URD 지점과 POH 지점은 36시간, CBD 지점은 60시간, DJD 지점은 72시간에서 각각 최솟값을 기록하였다. GMD 지점은 8시간에서 RSR 최솟값

(0.447)을 나타냈으나, 이후 입력 길이가 증가함에 따라 RSR이 증가하고 R^2 가 감소하는 경향을 보여 타 지점과 상이한 거동을 나타냈다.

이와 같이 지점별 최적 입력 길이는 12~72시간으로 다양하게 나타났다. DOH, MRD, OYD 지점은 12~24시간 구간에서 최적 성능을 나타냈으며, URD, CBD, POH, DJD 지점의 경우에도 각 지점의 최적 입력 길이와 24시간에서의 RSR 차이는 URD 지점은 1.2%, CBD 지점은 0.1%, POH 지점은 0.2%, DJD 지점은 0.6%로 모두 1.5% 이내로 매우 작은 수준이었다. GMD 지점 또한 최적 시간(8시간)과 24시간의 RSR 차이는 3.9%로 상대적으로 크게 나타났으나, 절대적인 성능 차이는 제한적인 수준이었다.

한편, 지점별 최적 입력 길이는 상이하게 나타났으나, 다수의 지점에서 12~24시간 구간이 최적이거나 이에 근접한 성능을 보였으며, 36시간 이상에서는 추가적인 성능 개선이 제한적인 것으로 확인되었다. 따라서 계산 효율성과 다지점 적용 가능성을 함께 고려할 때, 입력 길이 24시간이 다양한 지점에 대해 안정적으로 적용 가능한 실용적인 최적값으로 판단된다.

3.2 입력 변수 차이에 따른 유의파고 예측 성능 평가 분석 결과

Fig. 4에는 입력 변수 조합(IV1, IV3, IV9)에 따른 각 지점별 RSR, MAE, R^2 를 비교하여 제시하였다.

RSR을 기준으로 살펴보면, 전 지점에서 IV1 대비 IV9 적용 시 RSR이 감소하여 예측 성능이 향상되는 경향이 확인되었다. IV1의 경우 RSR은 0.414(MRD)~0.617(GMD) 범위로 나타났으며, IV9 적용 시 0.379(MRD)~0.560(DJD) 범위로 감소하였다. 지점별로 살펴보면, GMD 지점은 IV1에서 RSR 0.617로 전 지점 중 가장 높은 오차를 보였으나, IV9 적용 시 RSR 0.436으로 약 29% 감소하여 가장 큰 성능 향상을 나타냈다. DJD 지점 또한 IV1에서 RSR 0.586이었으나 IV9에서 RSR 0.559로 약 5% 감소하였다. URD, POH, DOH 지점은 IV1 대비 IV9에서 RSR이 각각 약 15%, 9%, 7% 감소하였다. MRD와 OYD 지점은 IV9 적용 시 RSR이 각각 0.378, 0.418로 감소하여 비교적 우수한 성능을 나타냈다.

IV3와 IV1을 비교하면, 대부분의 지점에서 IV3 적용 시 RSR과 MAE가 소폭 감소하였으나 그 차이는 IV9에 비해 제한적이었다. URD 지점의 경우 IV1 RSR 0.515에서 IV3 RSR 0.508로 약 1% 감소에 그쳤으나, IV9에서는 0.440으로 약 15% 감소하여 변수 수 증가에 따른 성능 향상이 IV9에서 두드러지게 나타났다.

MAE를 기준으로 살펴보면, IV9 적용 시 전 지점에서 MAE가 감소하여 절대 예측 오차가 개선되었다. GMD 지점은 IV1에서 MAE 0.267 m이었으나 IV9에서 0.188 m로 약 30% 감소하였으며, URD 지점은 IV1 MAE 0.297 m에서 IV9 MAE 0.254 m로 약 14% 감소하였다. DJD 지점은

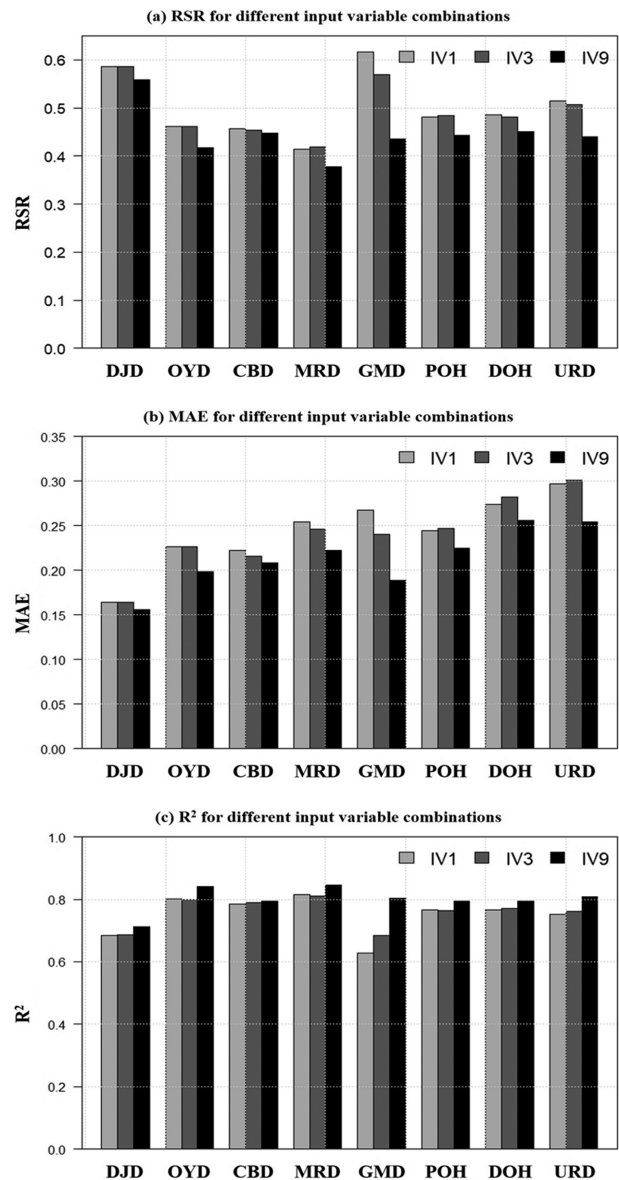


Fig. 4. Variation of RSR, MAE, and R^2 with input variable combinations at each site for Case 2.

IV9 적용 시 MAE 0.156 m로 전 지점 중 가장 낮은 절대 오차를 나타냈다.

한편, DJD 지점은 MAE가 전 지점 중 가장 낮은 수준임에도 불구하고, RSR은 0.559~0.586, R^2 는 0.685~0.711로 타 지점 대비 상대적으로 낮은 값을 나타냈다. 이는 해당 지점의 유의파고 변동 범위가 상대적으로 작아 절대 오차는 작게 나타나지만, 변동성 대비 오차 비율을 나타내는 RSR과 R^2 에서는 불리하게 평가된 것으로 판단된다.

R^2 를 기준으로 살펴보면, 전 지점에서 IV9 적용 시 R^2 가 가장 높게 나타났다. MRD 지점은 IV9에서 R^2 0.846으로 전 지점 중 가장 높은 설명력을 보였으며, OYD 지점 또한 IV9에서 R^2 0.841로 높은 수준을 유지하였다. 반면 DJD 지점은 IV9 적용 시에도 R^2 0.711로 타 지점 대비 낮은 수준에 머물렀다.

이에 따라 모든 지점에서 입력 변수 수가 증가할수록 RSR과 MAE가 감소하고 R^2 가 향상되는 경향이 확인되었으며, IV9가 모든 지점에서 가장 우수한 예측 성능을 나타냈다. 따라서 이후 분석에서는 IV9를 최적 입력 변수 조합으로 적용하였다.

3.3 예측시간에 따른 유의파고 예측 성능 평가 분석 결과

3.1절 및 3.2절에서 도출된 최적 입력 조건인 입력 길이 24

시간과 입력 변수 조합 IV9를 적용하여, 예측 시간 1, 2, 4, 6, 8, 10, 12, 24시간에 대한 유의파고 예측 성능을 분석하였다. 동해, 서해, 남해 및 제주 해역의 대표 지점에서 산출된 시계열 비교 결과와 산점도는 각각 Fig. 5와 Fig. 6에 제시하였으며, 나머지 지점에 대한 결과는 Appendix A1 및 A2에 수록하였다. 전 지점에 대한 예측 시간별 정량적 성능 평가 결과는 Table 6에 제시하였다.

Fig. 5의 시계열 비교 결과를 통해 예측 시간이 증가함에

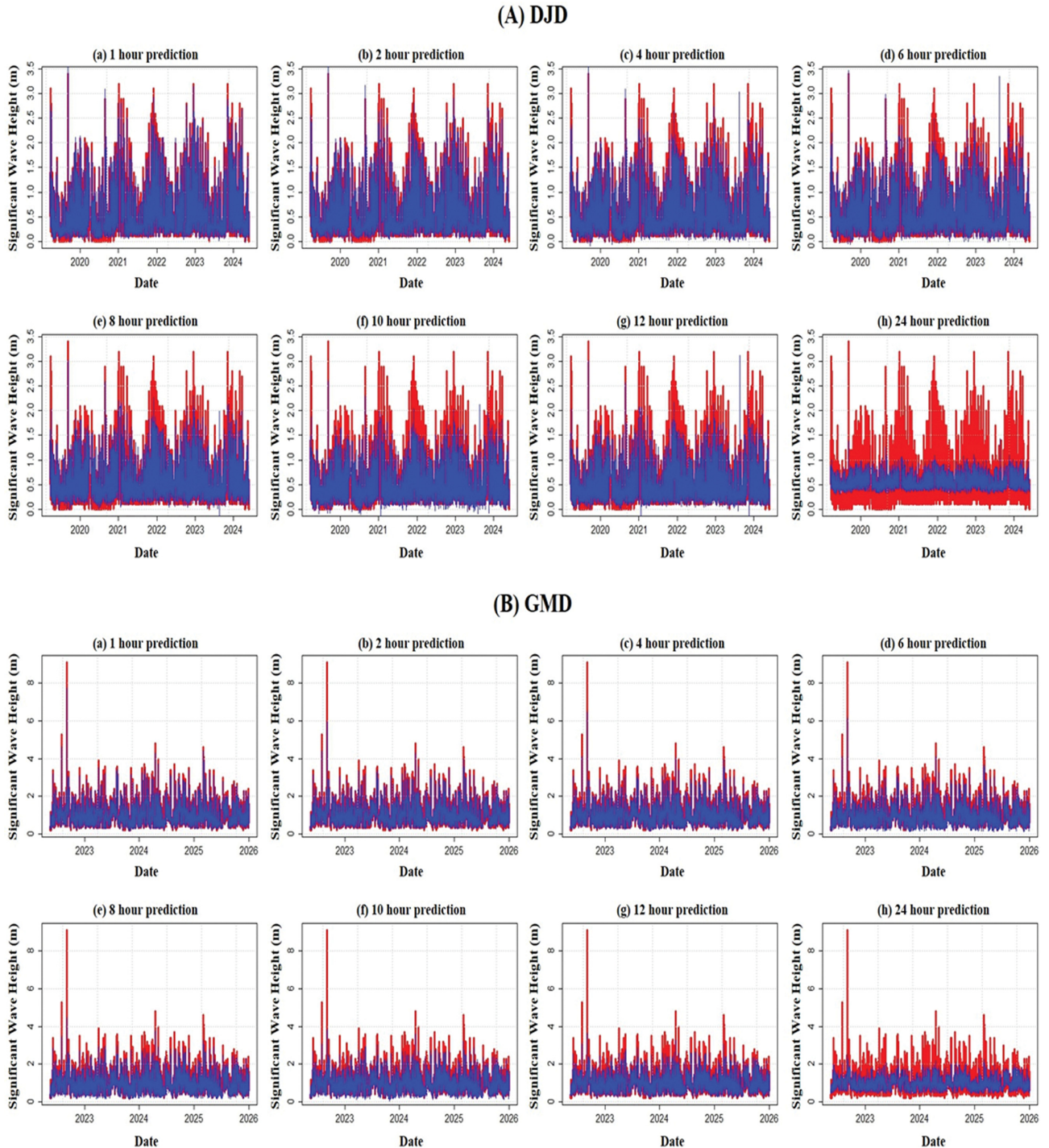
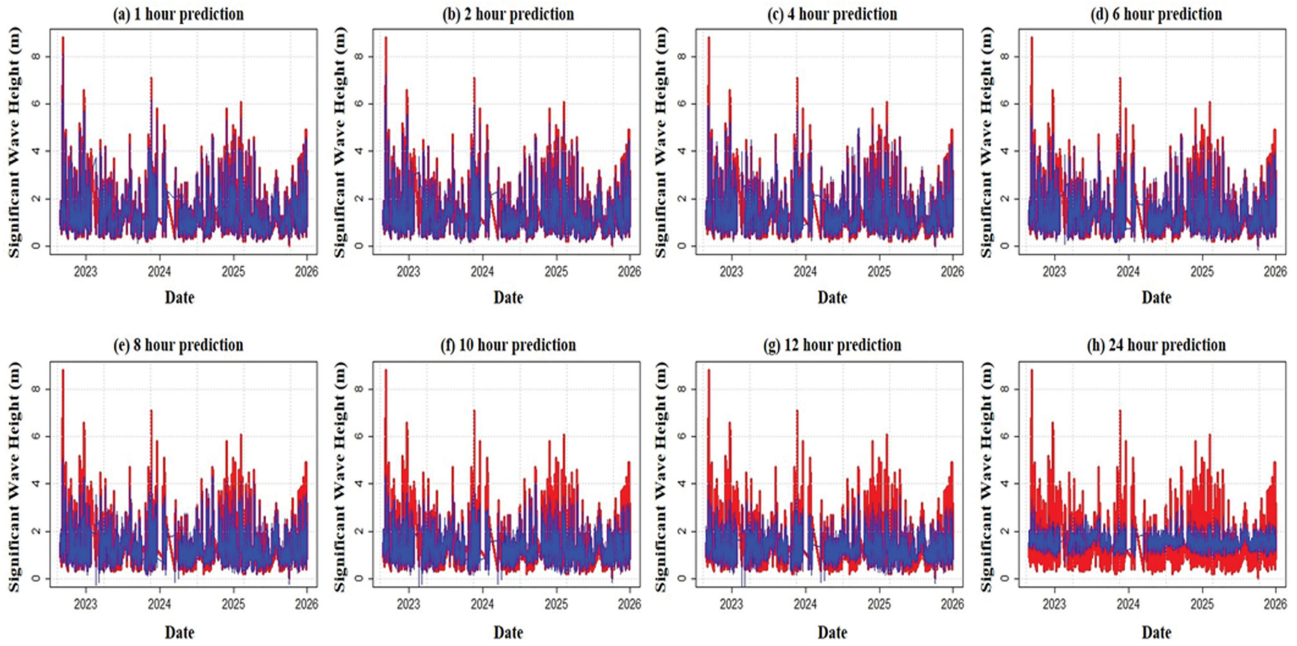


Fig. 5. Time series of observed and LSTM-predicted significant wave height at selected stations for different forecast lead times (red: observed, blue: predicted).

(C) MRD



(D) URD

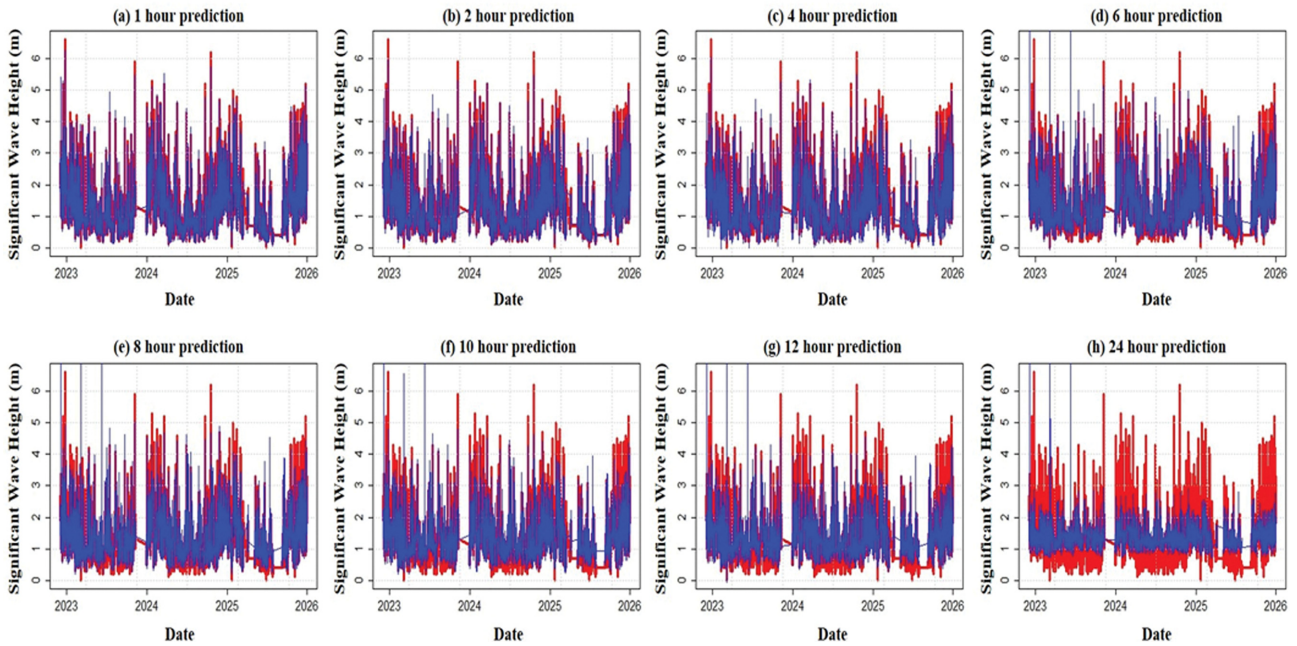


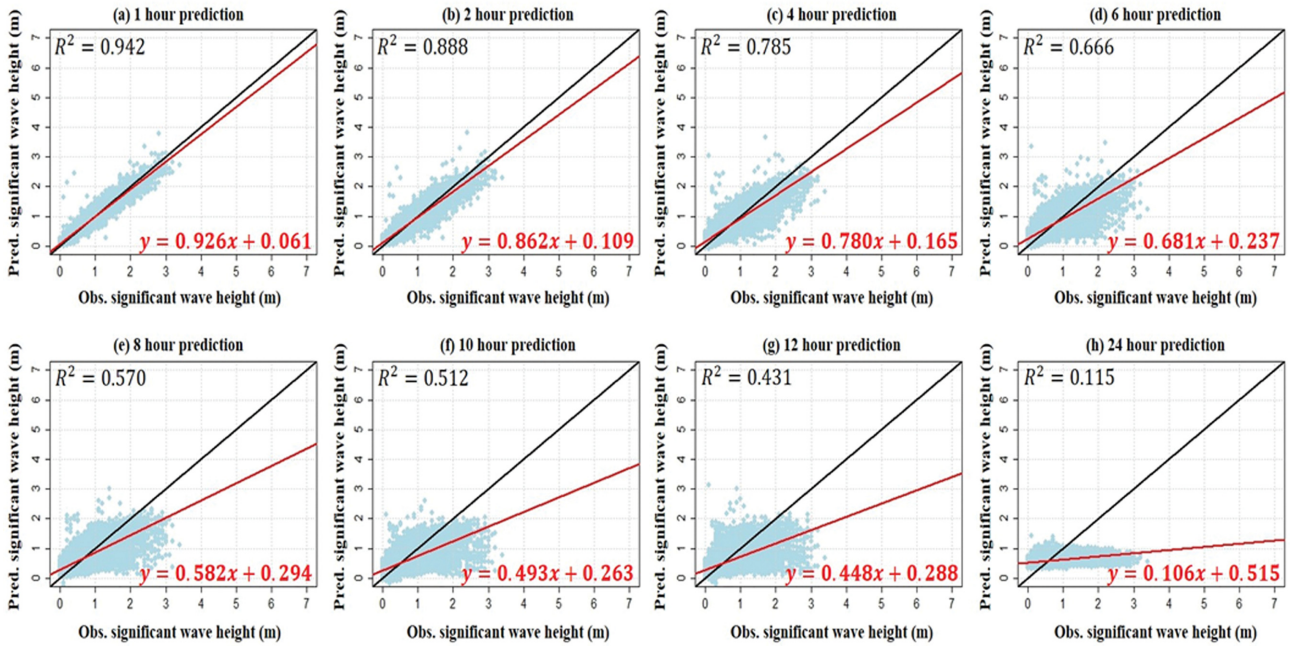
Fig. 5. Continued.

따라 관측값과 예측값 간의 차이가 점차 증가하는 경향이 확인되었다. 또한, Fig. 6의 산점도에서는 예측 시간이 증가할수록 회귀 기울기 및 결정계수가 감소하는 경향이 나타났으며, 특히 고파고 구간에서 과소추정이 발생하는 것을 확인할 수 있다. 특히 Fig. 6의 산점도에서는 예측 시간이 증가할수록 고파고 구간에서 관측값 대비 예측값이 과소추정되는 경향이 점차 뚜렷하게 나타났다. 이러한 경향은 상대적으로 출현 빈도가 낮은 고파고 조건이 학습 과정에서 충분히 반영되지 못한 데 기인하는 것으로 판단되며, 평상파 조건에서는 전

반적으로 관측값과 예측값이 유사하게 나타남에 따라 모델이 일반적인 파랑 변동 특성은 비교적 안정적으로 재현하는 것으로 나타났다.

Table 6을 기준으로 예측 시간에 따른 성능 변화를 살펴보면, 전반적으로 예측 시간이 증가함에 따라 RSR과 MAE가 증가하고 R^2 가 감소하여 예측 성능이 저하되는 경향이 나타났다. 단기 예측 구간인 1~4시간에서는 DJD 지점을 제외한 대부분의 지점에서 R^2 는 0.85 이상, RSR은 0.40 이하로 비교적 우수한 예측 성능을 유지하였다. 특히 1시간 예측의 경

(A) DJD



(B) GMD

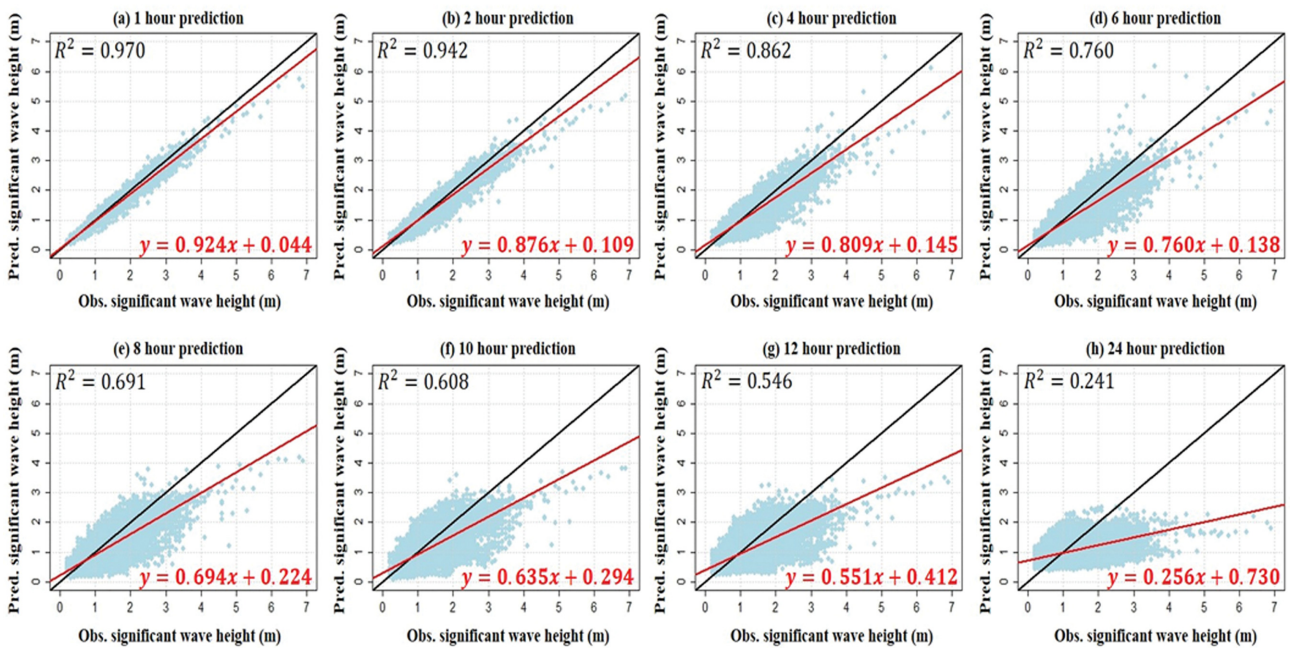


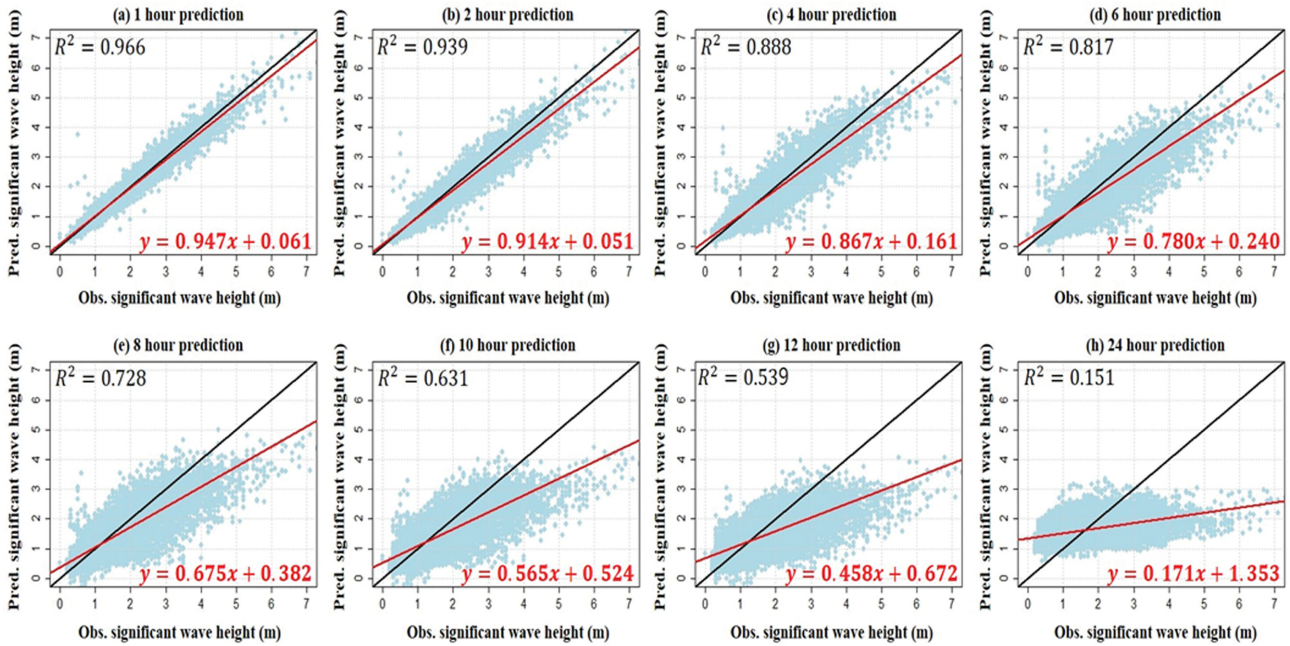
Fig. 6. Scatter plots of observed versus LSTM-predicted significant wave height for different forecast lead times at selected stations.

우, DOH 지점은 RSR이 0.166, R^2 는 0.972, OYD 지점은 RSR이 0.156, R^2 는 0.976, MRD 지점은 RSR이 0.184, R^2 가 0.966으로 전 지점에 걸쳐 높은 예측 성능을 나타냈다. 반면, DJD 지점은 1시간 예측에서 RSR이 0.240, R^2 가 0.942로 타 지점 대비 상대적으로 낮은 성능을 보였으며, 4시간 예측에서는 RSR은 0.463, R^2 는 0.785로 성능 저하가 다른 지점에 비해 더 크게 나타났다.

예측 시간이 증가함에 따라 성능 저하는 점진적으로 나타

났으며, 대부분의 지점에서 6시간 예측까지는 RSR 0.60 이하를 유지하여 Table 2에 제시된 기준과 같이 Good 수준 이상의 예측 성능이 유지되는 것으로 나타났다. 8시간 예측의 경우, DJD 지점을 제외한 대부분의 지점에서 RSR이 0.60 이하, R^2 는 0.65 이상을 유지하였다. 구체적으로 DOH 지점은 RSR이 0.546, R^2 는 0.701, POH 지점은 RSR이 0.595, R^2 는 0.646, MRD 지점의 RSR은 0.522, R^2 는 0.723, OYD 지점은 RSR이 0.525, R^2 는 0.724, CBD 지점의 RSR은 0.565,

(C) MRD



(D) URD

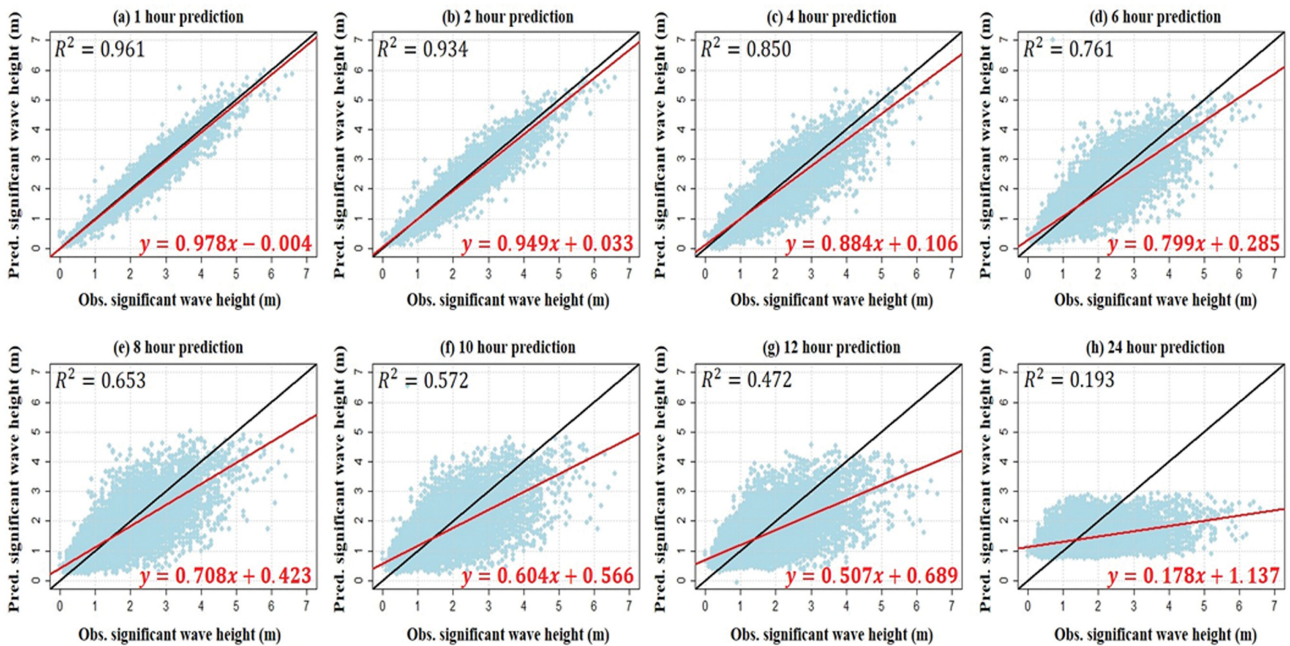


Fig. 6. Continued.

R^2 는 0.681로 나타났다. 반면 DJD 지점은 8시간 예측에서 RSR 0.656, R^2 0.570으로 타 지점 대비 성능 저하가 뚜렷하게 나타났으며, 6시간 예측에서 이미 RSR이 0.578로 0.60에 근접하여 상대적으로 이른 시점에서 성능 저하가 발생하는 것으로 확인되었다.

10시간 이상의 예측 구간에서는 대부분의 지점에서 RSR의 증가와 R^2 의 감소가 동시에 나타나며 성능이 급격히 저하되는 경향이 확인되었다. 24시간 예측에서는 URD 지점에서 RSR이 0.898, R^2 는 0.193, DOH 지점의 RSR은 0.869, R^2

는 0.244, POH 지점의 RSR은 0.947, R^2 는 0.102, GMD 지점에서 RSR은 0.871, R^2 가 0.241, MRD 지점의 RSR은 0.921, R^2 는 0.151, OYD 지점에서의 RSR은 0.930, R^2 는 0.136, CBD 지점 RSR은 0.901, R^2 는 0.188로 나타나, 전 지점에서 R^2 가 0.25 이하로 급감하고 RSR이 0.85 이상으로 크게 증가하여 모델의 예측 신뢰성이 크게 저하되는 것으로 나타났다.

이에 따라, 예측 시간이 증가함에 따라 전 지점에서 성능이 점진적으로 저하되는 경향이 확인되었으며, 대부분의 지

Table 6. Evaluation results of LSTM-based significant wave height prediction using RSR, MAE, and R^2 across different forecast lead times and stations

Station	Lead time (hour)	RSR	MAE	R^2
East Sea	URD	1	0.197	0.961
		2	0.257	0.934
		4	0.388	0.850
		6	0.489	0.761
		8	0.589	0.653
		10	0.654	0.572
		12	0.727	0.472
		24	0.898	0.193
	DOH	1	0.166	0.972
		2	0.230	0.947
		4	0.354	0.875
		6	0.461	0.787
		8	0.546	0.701
		10	0.619	0.616
		12	0.678	0.540
		24	0.869	0.244
	POH	1	0.179	0.968
		2	0.256	0.935
		4	0.402	0.858
		6	0.511	0.738
		8	0.595	0.646
		10	0.675	0.545
		12	0.751	0.436
		24	0.947	0.102
South Sea	GMD	1	0.173	0.970
		2	0.240	0.942
		4	0.371	0.862
		6	0.490	0.760
		8	0.556	0.690
		10	0.626	0.608
		12	0.674	0.546
		24	0.871	0.241
Jeju	MRD	1	0.184	0.966
		2	0.246	0.939
		4	0.335	0.888
		6	0.427	0.817
		8	0.522	0.723
		10	0.607	0.631
		12	0.679	0.538
		24	0.921	0.151
West Sea	DJD	1	0.240	0.942
		2	0.335	0.888
		4	0.463	0.785
		6	0.578	0.666
		8	0.656	0.570
		10	0.698	0.512
		12	0.754	0.432
		24	0.940	0.116
	OYD	1	0.156	0.976
		2	0.221	0.951
		4	0.340	0.884
		6	0.439	0.808
		8	0.525	0.724
		10	0.605	0.634
		12	0.671	0.550
		24	0.930	0.136

West Sea	CBD	1	0.174	0.090	0.970
		2	0.236	0.114	0.944
		4	0.371	0.177	0.862
		6	0.466	0.224	0.783
		8	0.565	0.271	0.681
		10	0.624	0.306	0.611
		12	0.691	0.334	0.523
		24	0.901	0.469	0.188

점에서 6~8시간 예측 구간까지는 일정 수준의 예측 성능이 유지되는 것으로 나타났다. DJD 지점은 전 예측 시간에 걸쳐 타 지점 대비 상대적으로 낮은 성능을 보여 지점 고유의 특성이 예측 성능에 영향을 미치는 것으로 판단된다.

4. 결론 및 토의

본 연구에서는 LSTM 모델을 이용하여 입력 길이, 입력 변수 조건, 예측 시간에 따른 유의파고 예측 성능 변화를 우리나라 연안 8개 지점을 대상으로 체계적으로 분석하였다. 다양한 입력 조건에 따른 모델 성능을 정량적으로 비교함으로써 최적 입력 설정을 도출하였으며, 도출된 주요 결론은 다음과 같다.

첫째, 입력 길이에 따른 예측 성능 분석 결과, 지점별 최적 입력 길이는 12~72시간으로 상이하게 나타났으나, 최적값과 입력 길이 24시간 간의 RSR 차이가 대부분의 지점에서 2% 이내로 매우 작은 수준이었다. 특히 예측 시간 4시간에서는 입력 길이 증가에 따라 일부 성능 개선 경향이 나타났으나, 36시간 이상에서는 추가적인 성능 개선이 제한적이었다. 계산 효율성과 다지점 적용 가능성을 함께 고려할 때, 입력 길이 24시간이 실용적인 최적값으로 판단된다.

둘째, 입력 변수 조합에 따른 예측 성능 분석 결과, 예측에 사용된 입력 변수의 수가 증가할수록 전 지점에서 RSR과 MAE가 감소하고 R^2 가 향상되는 경향이 뚜렷하게 나타났다. 단일 변수(IV1) 조건에서는 일부 지점에서 RSR이 0.6을 초과하는 등 예측 정확도가 크게 저하된 반면, 다변수 조건인 IV9에서는 전 지점에서 가장 우수한 성능을 나타냈다. 이는 유의파고가 복합적인 기상 및 해양 환경 요인의 영향을 받는 특성을 반영한 결과로, 다변수 입력 조건의 적용이 예측 성능 향상에 효과적임을 확인하였다.

셋째, 예측 시간에 따른 성능 분석 결과, 예측 시간이 증가함에 따라 전 지점에서 성능이 점진적으로 저하되는 경향이 확인되었다. 1~4시간 단기 예측 구간에서는 대부분의 지점에서 R^2 0.85 이상, RSR 0.40 이하로 우수한 성능을 유지하였으며, 6~8시간 구간까지는 일정 수준의 예측 성능이 유지되는 것으로 나타났다. 반면 10시간 이상에서는 성능 저하가 뚜렷하게 나타났으며, 24시간 예측에서는 전 지점에서 R^2 가 0.25 이하로 감소하여 예측 신뢰성이 크게 저하되었다. 이러한 경향은 예측 시간이 증가할수록 입력 시점과 예측 대상

시점 간 시간 간격이 확대되면서 시계열 자기상관성이 점차 약화되고, 급격한 파랑 변동에 따른 누적 오차가 증가하기 때문으로 판단된다. 한편, DJD 지점은 다른 지점 대비 상대적으로 낮은 R^2 및 높은 RSR 값을 나타내는 경향이 확인되었다. 이는 MAE가 유사한 수준으로 나타난 점을 고려할 때 절대적인 예측 오차 증가보다는 해당 지점의 시계열 변동 구조 차이에 기인한 것으로 판단된다. 서해 해역은 천해역 특성 및 계절풍의 복합적 영향으로 파랑 변동성이 비정상적으로 나타날 가능성이 있으며, 이러한 특성이 예측 성능 저하에 일부 영향을 미쳤을 가능성이 있다. 따라서 실용적 활용 측면에서는 6시간 이내의 단기 예측 구간에 적용하는 것이 적절하며, 특히, 1~4시간 수준의 짧은 시간 예측은 항만 및 연안에서 이루어지는 해상작업의 작업 가능 여부 판단, 소형 선박 운항 지원, 해양관측 장비의 유지관리 작업 여부 결정과 같은 분야에서 활용될 수 있을 것으로 판단된다.

본 연구는 관측자료 기반 LSTM 모델을 적용하여 유의파고 예측 성능을 분석하였으나, 몇 가지 한계를 가진다. 입력 길이 및 입력 변수 조합에 따른 예측 성능 변화를 중점적으로 분석하였으나, LSTM 모델의 hidden state dimension, number of units 등 내부 하이퍼파라미터는 고정된 조건에서 적용하였다. 향후 체계적인 하이퍼파라미터 최적화를 통해 모델 성능을 추가적으로 개선할 수 있을 것으로 판단된다. 또한 예측 시간이 증가함에 따라 고파고 구간에서 관측값 대비 예측값이 과소추정되는 경향이 확인되었다. 이러한 경향은 시계열 기반 유의파고 예측 연구에서 공통적으로 나타나는 한계로서, 자료 내 출현 빈도가 상대적으로 낮은 고파고 조건이 학습 과정에서 충분히 반영되지 못한 데 기인하는 것으로 판단된다. 이는 단일 LSTM 모델이 전반적인 시계열 경향 및 평상파 조건에 대해서는 비교적 안정적인 예측 성능을 보이지만, 급격한 파고 증가와 같은 극한 현상을 동시에 재현하는 데에는 한계가 있음을 의미한다. 이러한 결과는 Minuzzi and Farina(2023)에서도 유사하게 나타난 바 있으며, 특히 약 6시간 수준까지는 전체적인 시계열 경향을 비교적 잘 재현하였으나, 고파고와 같은 극한 조건에서는 일부 과소추정 경향이 나타났다. 따라서 향후에는 하이브리드 모델 또는 앙상블 기반 접근 기법을 적용하여 극한 고파고 조건에 대한 예측 성능을 추가적으로 개선할 필요가 있다.

아울러, 본 연구에서는 단일 LSTM 모델을 대상으로 분석을 수행하였으므로, RNN, GRU, CNN 등 다양한 딥러닝 모델과의 비교 평가는 수행하지 않았으며, 향후 연구에서는 체계적인 하이퍼파라미터 최적화와 다양한 딥러닝 모델 간 비교를 통해 예측 범위 확장 및 극한 조건 대응 성능을 추가적으로 검토할 필요가 있다.

감사의 글

본 연구는 기후에너지환경부(MCEE)와 한국에너지기술평

가원(KETEP)의 지원을 받아 수행한 연구 과제입니다(No. RS-2025-02313959).

References

- Ali, M., Prasad, R., Xiang, Y., Sankaran, A., Deo, R.C., Xiao, F. and Zhu, S. (2021). Advanced extreme learning machines vs. deep learning models for peak wave energy period forecasting: A case study in Queensland, Australia. *Renewable Energy*, 177, 1031-1044.
- Bai, G., Wang, Z., Zhu, X. and Feng, Y. (2022). Development of a 2-D deep learning regional wave field forecast model based on convolutional neural network and the application in South China Sea. *Applied Ocean Research*, 118, 103012.
- Bergmeir, C. and Benitez, J.M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 192-213.
- Bergmeir, C., Hyndman, R.J. and Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, 120, 70-83.
- Bethel, B.J., Dong, C., Zhou, S. and Cao, Y. (2021). Bidirectional modeling of surface winds and significant wave heights in the Caribbean Sea. *Journal of Marine Science and Engineering*, 9(5), 547.
- Chai, T. and Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific Model Development Discussions*, 7(1), 1525-1534.
- Chen, J., Pillai, A.C., Johanning, L. and Ashton, I. (2021). Using machine learning to derive spatial wave data: A case study for a marine energy site. *Environmental Modelling & Software*, 142, 105066.
- Choi, J.Y. and Lee, B. (2018). Combining LSTM network ensemble via adaptive weighting for improved time series forecasting. *Mathematical Problems in Engineering*, 2018(1), 2470171.
- Fan, S., Xiao, N. and Dong, S. (2020). A novel model to predict significant wave height based on long short-term memory network. *Ocean Engineering*, 205, 107298.
- Gao, S., Huang, J., Li, Y., Liu, G., Bi, F. and Bai, Z. (2021). A forecasting model for wave heights based on a long short-term memory neural network. *Acta Oceanologica Sinica*, 40(1), 62-69.
- Hao, P., Li, S. and Gao, Y. (2023). Significant wave height prediction based on deep learning in the South China Sea. *Frontiers in Marine Science*, 9, 1113788.
- Hewamalage, H., Bergmeir, C. and Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1), 388-427.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
- James, S.C., Zhang, Y. and O'Donncha, F. (2018). A machine learning framework to forecast wave conditions. *Coastal Engineer-*

- ing, 137, 1-10.
- Jorges, C., Berkenbrink, C. and Stumpe, B. (2021). Prediction and reconstruction of ocean wave heights based on bathymetric data using LSTM neural networks. *Ocean Engineering*, 232, 109046.
- Kong, W., Dong, Z.Y., Jia, Y., Hill, D.J., Xu, Y. and Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*, 10(1), 841-851.
- Lee, G.S., Jeong, D.H., Moon, Y.H., Park, W.K. and Chae, J.W. (2021). Wave prediction in a harbour using deep learning with offshore data. *Journal of Korean Society of Coastal and Ocean Engineers*, 33(6), 367-373 (in Korean).
- Meng, F., Song, T., Xu, D., Xie, P. and Li, Y. (2021). Forecasting tropical cyclones wave height using bidirectional gated recurrent unit. *Ocean Engineering*, 234, 108795.
- Minuzzi, F.C. and Farina, L. (2023). A deep learning approach to predict significant wave height using long short-term memory. *Ocean Modelling*, 181, 102151.
- 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*, 50(3), 885-900.
- Nash, J.E. and Sutcliffe, J.V. (1970). River flow forecasting through conceptual models. *Journal of Hydrology*, 10(3), 282-290.
- Park, H.J. and Kang, Y.J. (2024). Comparison of wave prediction and performance evaluation in Korea waters based on machine learning. *Journal of Ocean Engineering and Technology*, 38(1), 18-29.
- Shin, S., Park, S., Jung, K.H. and Park, S.B. (2024). Study for filling missing wave data in geomundo ocean buoy using artificial neural networks. *Journal of Ocean Engineering and Technology*, 38(6), 426-437.
- Willmott, C.J. and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79-82.
- Yu, Y., Si, X., Hu, C. and Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, 31(7), 1235-1270.
- Zhou, S., Xie, W., Lu, Y., Wang, Y., Zhou, Y., Hui, N. and Dong, C. (2021). ConvLSTM-based wave forecasts in the South and East China seas. *Frontiers in Marine Science*, 8, 680079.

Received 30 April, 2026

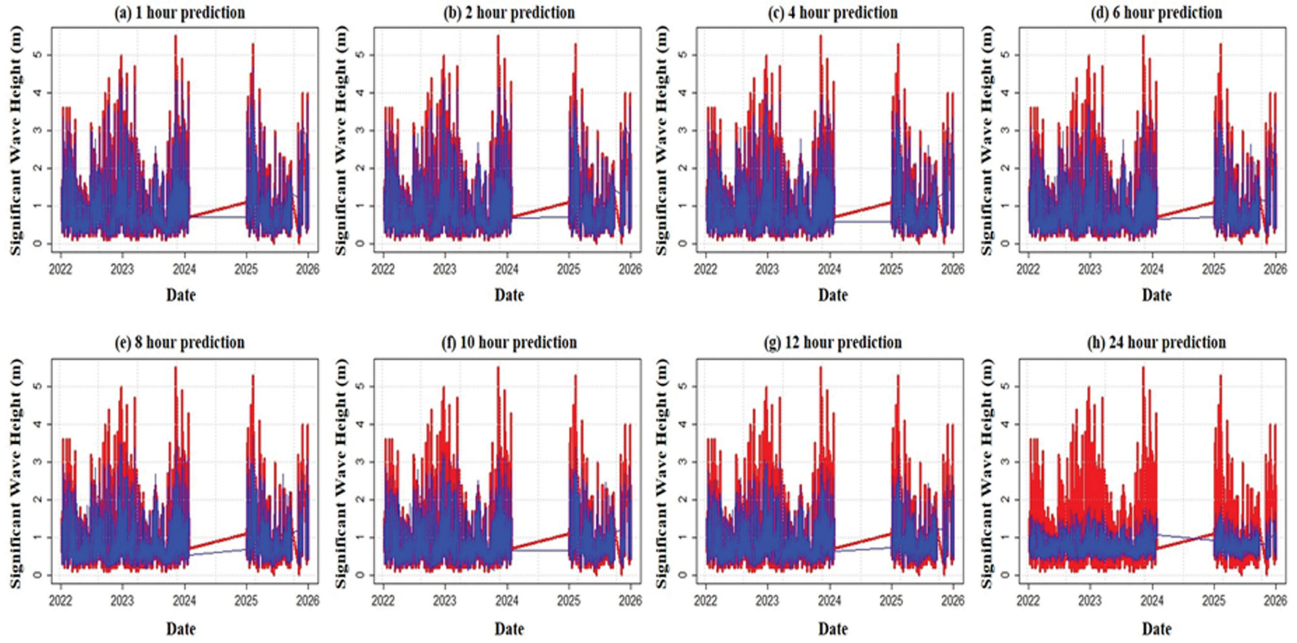
Revised 8 June, 2026

Accepted 9 June, 2026

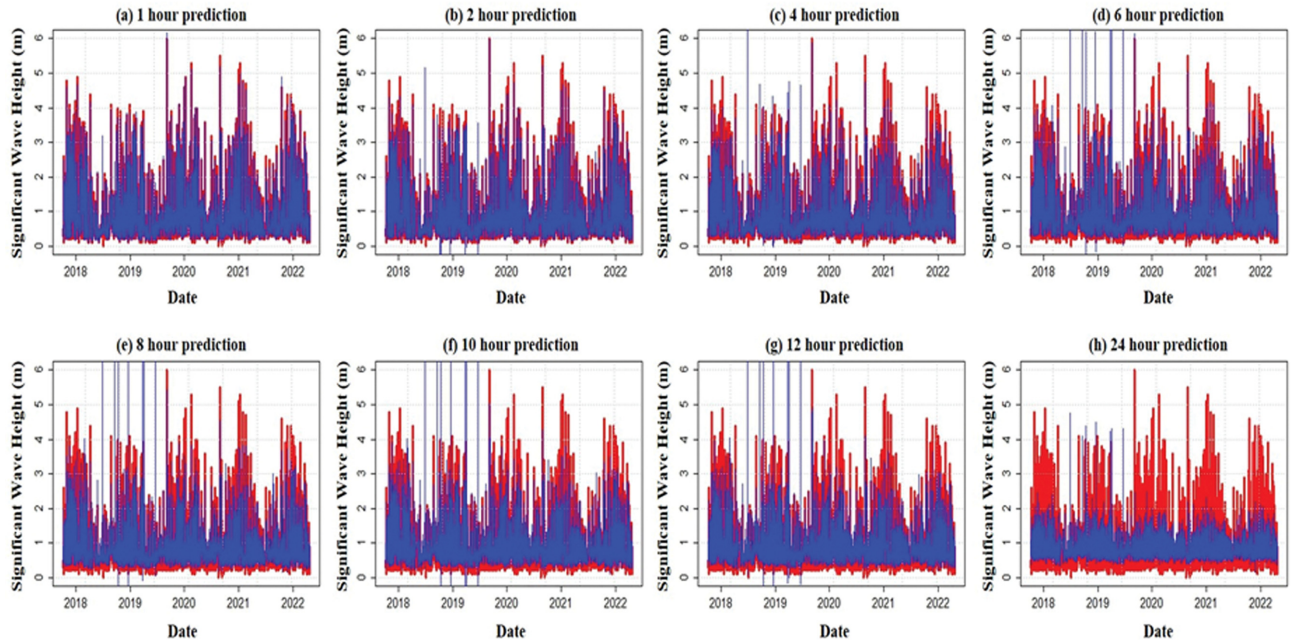
Appendix

Appendix A1. Time Series of Significant Wave Height Prediction at Other Stations for Different Lead Times

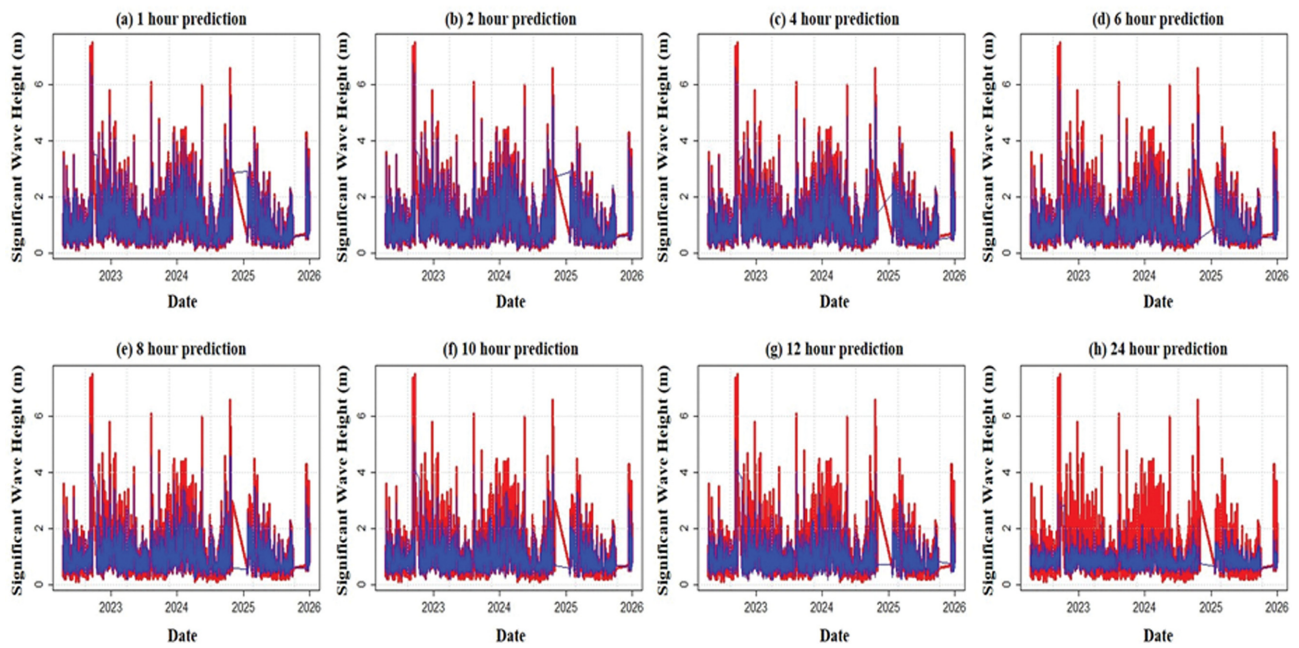
(A) OYD



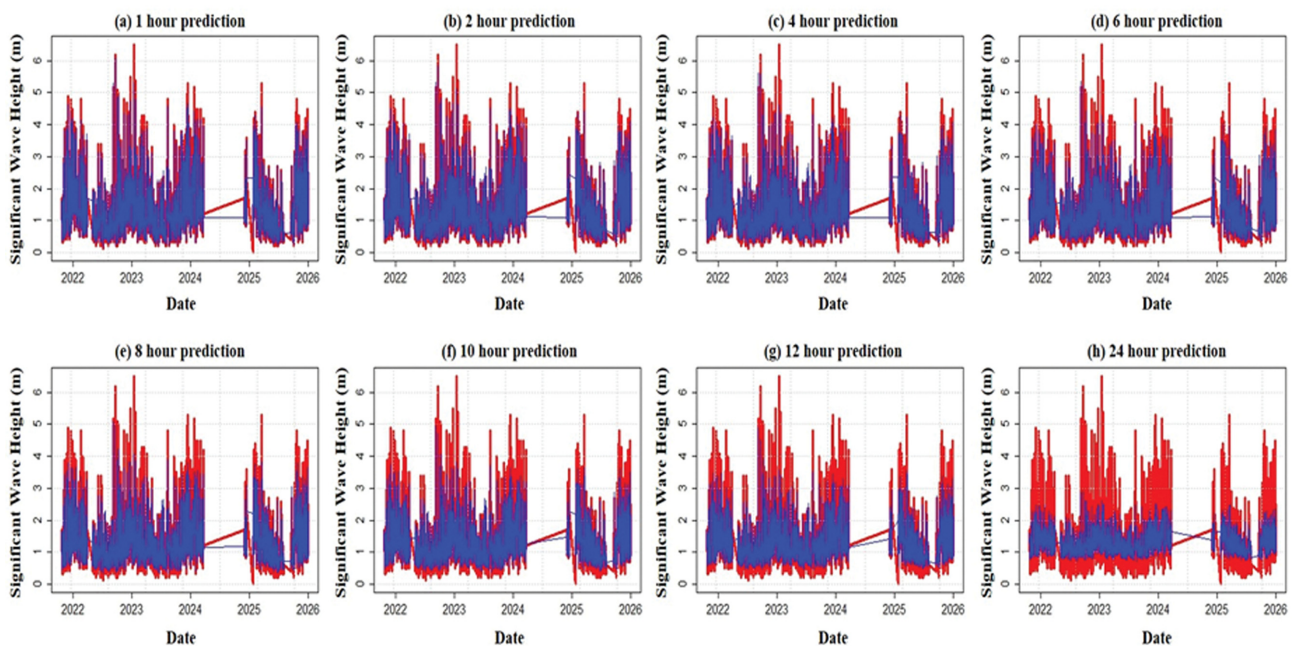
(B) CBD



(C) POH

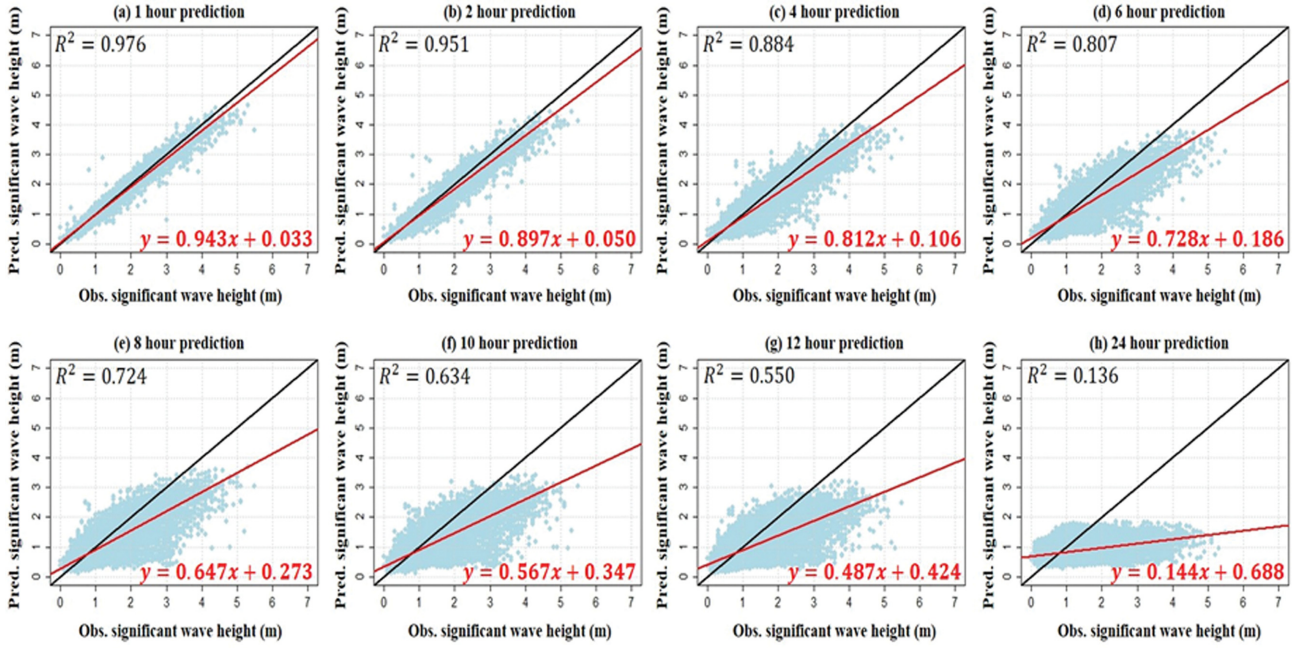


(D) DOH

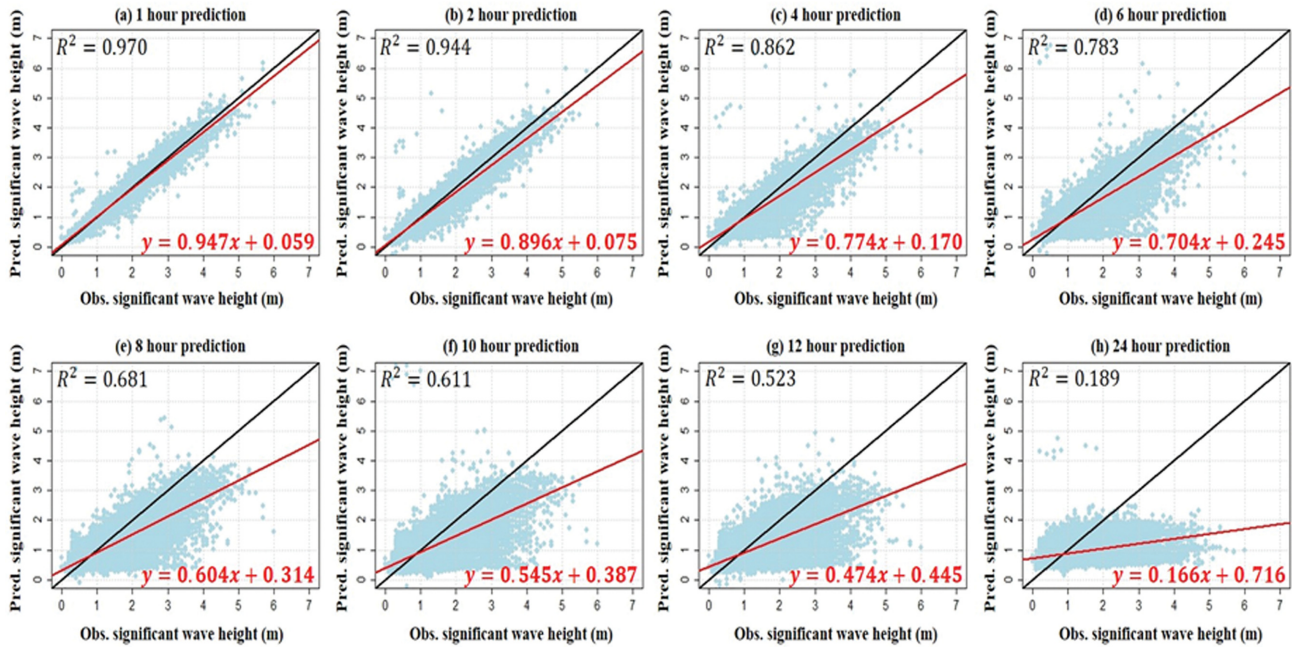


Appendix A2. Scatter Plots of Significant Wave Height Prediction at Other Stations for Different Lead Times

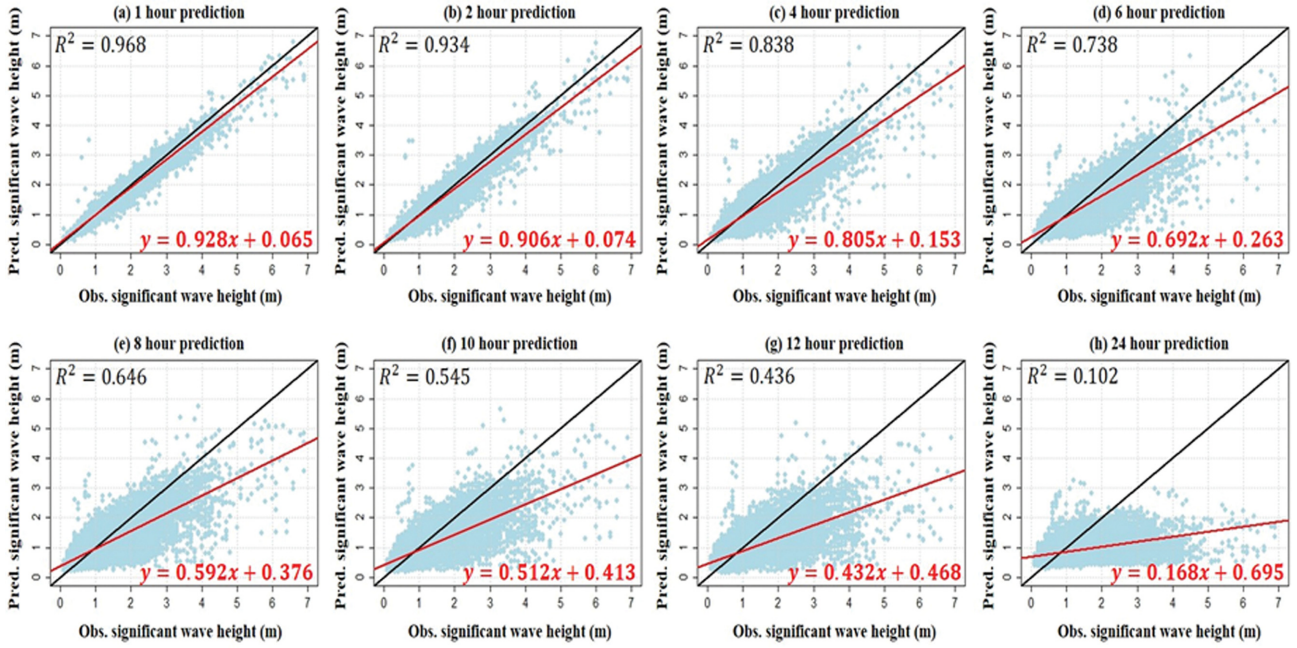
(A) OYD



(B) CBD



(C) POH



(D) DOH

