

# 수온 데이터 예측 연구를 위한 통계적 방법과 딥러닝 모델 적용 연구

조문원\* · 최흥배\*\* · 한명수\*\*\* · 정은송\*\*\*\* · 강태순\*\*\*\*\*

\* (주)지오시스템리서치 주임, \*\* (주)지오시스템리서치 상무이사, \*\*\* (주)지오시스템리서치 선임,  
\*\*\*\* (주)지오시스템리서치 주임, \*\*\*\*\* (주)지오시스템리서치 전무이사

## Statistical Method and Deep Learning Model for Sea Surface Temperature Prediction

Moon-Won Cho\* · Heung-Bae Choi\*\* · Myeong-Soo Han\*\*\* · Eun-Song Jung\*\*\*\* · Tae-Soon Kang\*\*\*\*\*

\* Assistant Research Engineer, GeoSystem Research Corp., Gyeonggi 15870, Korea

\*\* Executive Director, GeoSystem Research Corp., Gyeonggi 15870, Korea

\*\*\* Research Engineer, GeoSystem Research Corp., Gyeonggi 15870, Korea

\*\*\*\* Assistant Research Engineer, GeoSystem Research Corp., Gyeonggi 15870, Korea

\*\*\*\*\* Executive Vice President, GeoSystem Research Corp., Gyeonggi 15870, Korea

**요약** : 기후변화 영향으로 이상고수온, 태풍, 홍수, 가뭄 등 재난 및 안전 관리기술은 지속적으로 고도화를 요구받고 있으며, 특히 해수면 온도는 한반도 주변에서 발생하는 여름철 적조 발생과 동해안 냉수대 출현, 소멸 등에 영향을 신속하게 분석할 수 있는 중요한 인자이다. 따라서, 본 연구에서는 해수면 온도 자료를 해양 이상현상 및 연구에 적극 활용되기 위해 통계적 방법과 딥러닝 알고리즘을 적용하여 예측성능을 평가하였다. 예측에 사용된 해수면 수온자료는 흑산도 조위관측소의 2018년부터 2022년까지 자료이며, 기존 통계적 ARIMA 방법과 Long Short-Term Memory(LSTM), Gated Recurrent Unit(GRU)을 사용하였고, LSTM의 성능을 더욱 향상할 수 있는 Sequence-to-Sequence(s2s) 구조에 Attention 기법을 추가한 Attention Long Short-Term Memory (LSTM)기법을 사용하여 예측 성능 평가를 진행하였다. 평가 결과 Attention LSTM 모델이 타 모델과 비교하여 더 좋은 성능을 보였으며, Hyper parameter 튜닝을 통해 해수면 수온 성능을 개선할 수 있었다.

**핵심용어** : ARIMA, LSTM, GRU, Attention LSTM, 해수면 수온

**Abstract** : As climate change continues to prompt an increasing demand for advancements in disaster and safety management technologies to address abnormal high water temperatures, typhoons, floods, and droughts, sea surface temperature has emerged as a pivotal factor for swiftly assessing the impacts of summer harmful algal blooms in the seas surrounding Korean Peninsula and the formation and dissipation of cold water along the East Coast of Korea. Therefore, this study sought to gauge predictive performance by leveraging statistical methods and deep learning algorithms to harness sea surface temperature data effectively for marine anomaly research. The sea surface temperature data employed in the predictions spans from 2018 to 2022 and originates from the Heuksando Tidal Observatory. Both traditional statistical ARIMA methods and advanced deep learning models, including long short-term memory (LSTM) and gated recurrent unit (GRU), were employed. Furthermore, prediction performance was evaluated using the attention LSTM technique. The technique integrated an attention mechanism into the sequence-to-sequence (s2s), further augmenting the performance of LSTM. The results showed that the attention LSTM model outperformed the other models, signifying its superior predictive performance. Additionally, fine-tuning hyperparameters can improve sea surface temperature performance.

**Key Words** : ARIMA, LSTM, GRU, Attention LSTM, Sea Surface Temperature (SST)

\* First Author : mwcho@geosr.com, 031-5180-5861

† Corresponding Author : hbchoi@geosr.com, 031-5180-5862

## 1. 서론

해수면 온도(Sea Surface Temperature, SST) 예측은 해양생태계 및 수자원 관리 분야와 이상고수온, 태풍, 홍수 및 가뭄 등 자연재해 대응 분야에서 필수적으로 요구되고 있다(Chen et al., 2011; Saunders and Lea, 2008; Cassola et al., 2016; Shabbar and Skinner, 2004). 특히, 한반도에서는 여름철 적조 발생과 동해안 냉수대 출현 및 소멸과 같이 환경·경제적으로 막대한 손실을 입히는 해양 현상이 SST의 변동과 밀접하게 연관되어 있다(Lee et al., 2001; Cho and Kim, 1996). 따라서 SST의 변동을 정확히 예측하고 이상 변동 현상을 조기에 파악하는 기술을 확보하는 것이 필요하다.

시계열 예측의 정확도는 모델 피팅에 크게 의존한다. 예측 정확도의 개선을 위해 수십 년간 많은 시계열 모델이 개발되어 왔다. Autoregressive Integrated Moving Average(ARIMA)는 가장 보편적인 확률론적 시계열 예측 모형이며, 장기 추세와 계절성을 동시에 가지는 SST 시계열 예측에 효과적으로 적용되었다(Breaker, 2006). 이후로 다양한 SST 예측의 연구 진행됨에 따라 데이터셋의 통계적 분포를 가정하지 않는 인공신경망 기법이 활용 빈도가 증가하였다(Thomson et al., 2006). 특히, 시계열 예측에 특화된 Long-Short Term Memory (LSTM), Gated Recurrent Unit(GRU), Attention 기법은 다양한 해역의 SST 예측에 성공적으로 적용되었으며, 해역에 따라 적합한 알고리즘과 모델 변수가 변화할 수 있음을 시사하였다(Verdin et al., 2005; Zhang et al., 2020; Xie et al., 2019). 시계열 모형은 공간 인자를 포함하지 않고 시계열 데이터에 의존하여 모델 피팅이 이루어지므로, 알고리즘과 모델 변수의 선택이 최종 예측 성능을 크게 좌우하며, 예측 시 모형 선택과 변수 최적화가 필수적으로 선행되어야 한다.

SST 시계열을 대상으로 여러 모형 간 비교를 연구한 사례는 아래와 같다. Xie et al.(2019)과 Sarkar et al.(2020)은 ARIMA 및 LSTM, GRU, Attention 등의 모형의 SST 예측 성능을 비교하였으며, 예측 기간의 증가에 따른 성능 감소 정도는 알고리즘 간 유의한 차이가 나타내는 것으로 보고하였다. Kartal (2023)은 통계모형 및 인공신경망에서 일부 모델 변수를 대상으로 실험하여 성능 변화를 추적하였다. 인공신경망은 많은 모델 변수로 구성되나, 이를 고려한 체계적인 모델 실험을 수행한 연구 자료는 매우 부족하다. 따라서 본 논문에서는, SST 시계열 예측의 최적 기법을 도출하기 위한 기초자료로서 선형 통계 및 비선형 인공지능 모델의 예측 정확도를 비교하였으며, 추가적으로 인공지능 모델의 하이퍼파라미터에 따른 예측 정확도 변화를 비교하였다.

## 2. 시계열 예측 모델 비교

본 연구에서는 시계열 데이터 분석 기법 중 자주 사용되는 알고리즘인 ARIMA와 인공지능 모델 LSTM, GRU, Attention LSTM을 사용하여 학습 및 예측을 진행하였다. 본 연구 내용과 같이 이전에 비트코인 가격을 통한 ARIMA와 LSTM, GRU에 대하여 정확도 비교 검증 연구가 진행된 사례가 있었다(Yamak et al., 2019). 현재 LSTM에서 고도화된 알고리즘이 다수 등장하였으며, 이러한 알고리즘을 활용한 다양한 분야의 시계열 예측에 활용되어 기존 통계적 방법과 비교하여 인공지능을 활용한 알고리즘에서 시계열 예측의 성능이 향상되었다. 또한 시계열 데이터의 데이터 특성에 따라 같은 인공지능 알고리즘이라도 성능이 다르게 나타난다. 본 연구는 기존 기계학습(ARIMA)과 인공지능 알고리즘을 활용하여 수온 시계열 예측성능을 평가하고 시계열 예측에 적합한 알고리즘을 제시하고자 한다. 다음은 각 알고리즘에 대한 설명이다.

### 2.1 ARIMA(Autoregressive Integrated Moving Average)

ARIMA 모델은 자기회귀(Autoregressive) 모형과 이동평균(Moving Average) 모형을 결합하여 만들어진 모델이다. 먼저 자기회귀 기법은 시계열 데이터가 가지고 있는 데이터가 시간 순서에 따라 이전 데이터값이 다음 데이터값에 영향을 준다는 것을 기반으로 제작된 모델이다. 자기회귀 기법은 식(1)과 같이 표현할 수 있다.

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + \Phi_p y_{t-p} + \epsilon_t \quad (1)$$

이동평균 모델은 자기회귀 모델과 다르게 과거의 데이터 원본을 이용하는 것이 아닌 과거 데이터의 연속적인 오차를 이용하여 미래 데이터를 예측하는 것이다. 이동평균 모델은 식(2)와 같이 나타낼 수 있다

$$y_t = \theta_0 + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad (2)$$

위와 같은 자기회귀 모형과 이동평균 모형을 단순 결합하여 식(3)과 같이 ARMA 모델을 구축할 수 있다.

$$y_t = c + \Phi_1 y_{t-1} + \dots + \Phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t \quad (3)$$

하지만 자기회귀 모델, 이동평균 모델, ARMA 모델을 사용하기 위해서는 데이터가 정상성(stationary)을 지닌 데이터를 이용해야 한다. 하지만 대부분의 시계열 데이터는 비정상성(non-stationary)의 특징을 지닌 데이터이기에 해당 모델

들을 적용하기에는 어려움이 있다. 이러한 문제를 해결하기 위하여 비정상성을 지닌 데이터를 정상성의 데이터로 변환하기 위한 차분을 사용한 모델이 ARIMA 모델이다. 차분을 통하여 비정상성의 데이터를 정상성 데이터로 변환 후에 ARIMA 모델을 적용하여 시계열 미래 예측을 진행하는 모델이다.

### 2.2 LSTM(Long-short term memory)

기존 시계열 미래 예측을 위한 인공지능 모델로 RNN (Recurrent Neural Network) 모델을 주로 이용하였다. RNN은 Fig. 1과 같은 모델 구조를 지니고 있으며, RNN 모델의 경우 학습이 데이터의 길이가 길어짐에 따라 예측 기간과 먼 이전 학습 결과의 영향력이 작아지는 기울기 소실 문제(vanishing gradient problem)가 발생한다.

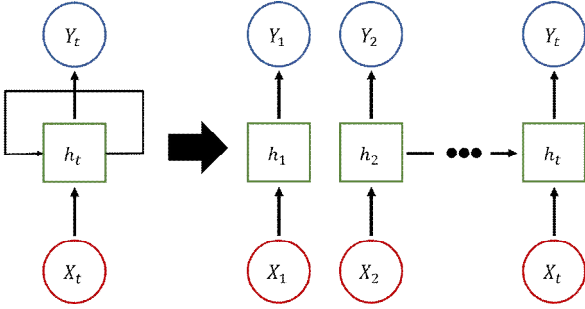


Fig. 1. Typical RNN network structure with RNN cells, here  $X_t$ : input time sequence,  $h_t$ : output time sequence.

LSTM은 이러한 RNN의 기울기 소실 문제를 해결하기 위해 고안된 모델이며, LSTM의 구조는 Fig. 2와 같다. 해당 알고리즘은 input gate, output gate, forget gate를 이용하여 해당 gate를 지나는 학습 정보에 대한 필요·불필요 여부를 판단하여 Cell-state에 정보를 전달한다.

gate에 따라 input gate는 식(4)와 식(7)과 같은 계산 과정을 통해 현재 입력 정보를 Cell-state에 얼마나 반영하는지를 판단하는 역할을 수행하며, forget gate는 식(5)와 같은 계산 과정을 통해 이전에 학습된 정보에서 불필요한 정보를 판단하여 제거하는 역할을 수행한다. output gate는 식(6)과 같은 계산 과정을 통해 현재 Cell에서 학습된 정보에 각각 다른 가중치를 곱하여 하나는 Cell-state에 반영, 다른 하나는 다음 학습에 사용될 hidden state 값을 얻을 수 있다.

Cell-state는 각 정보를 식(8)을 이용하여 처리하여 다음 Cell에 전달된다.

이러한 Cell-state를 이용하여 장기적으로 학습 정보를 저장하며, hidden state를 통하여 식(9)와 같이 계산된 단기적 학습 정보를 다음 학습에 반영하여 학습을 진행하게 된다.

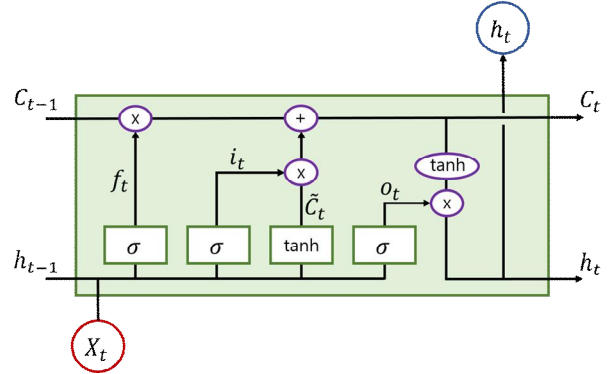


Fig. 2. LSTM cell architecture, here  $X_t$ : input time step,  $h_t$ : output,  $C_t$ : cell state,  $f_t$ : forget gate,  $i_t$ : input gate,  $o_t$ : output gate,  $\hat{C}_t$ : internal cell state. Operations inside light red circle are pointwise.

$$i_t = \sigma(X_t U^i + h_{t-1} W^i) \quad (4)$$

$$f_t = \sigma(X_t U^f + h_{t-1} W^f) \quad (5)$$

$$o_t = \sigma(X_t U^o + h_{t-1} W^o) \quad (6)$$

$$\tilde{C}_t = \tanh(X_t U^g + h_{t-1} W^g) \quad (7)$$

$$C_t = \sigma(f_t C_{t-1} + i_t \times \tilde{C}_t) \quad (8)$$

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

### 2.3 GRU(Gated Recurrent Unit)

GRU는 기존 LSTM의 경우 장·단기 학습 메모리를 모두 사용하기에 학습 시간이 길어짐에 따라 학습 과정을 단순화하여 학습 시간 단축을 목적으로 제작된 모델이다. Fig. 3과 같이 기존 gate가 3개로 구성된 LSTM에서 gate의 수가 2개로 줄어들었으며, Cell-state와 hidden state가 결합된 형태로 구성되어 있다.

GRU의 2개의 gate는 reset gate와 update gate로 구성된다. reset gate는 LSTM의 Forget gate와 같이 이전에 학습된 정보에서 필요한 정보를 판단하여 불필요한 정보를 제거하는 역할을 수행하며, 식(10)과 같이 나타낼 수 있다. update gate는 현재 학습에 입력되는 입력 자료와 이전 학습 결과를 이용하여 새로운 학습 정보를 생성, hidden state에 반영하는 역할을 수행하며, 식(11)과 같이 표현된다. 생성된 데이터를 이용하여 식(12)와 식(13)을 이용하여 다음 Cell로 이동될 학습 정보를 계산하게 된다.

$$r_t = \sigma(W_{xr} X_t + W_{hr} h_{t-1} + b_r) \quad (10)$$

$$z_t = \sigma(W_{xz} X_t + W_{hz} h_{t-1} + b_z) \quad (11)$$

$$g_t = \tanh(W_{hg}(r_t \circ h_{t-1}) + W_{xg} X_t + b_g) \quad (12)$$

$$h_t = (1 - z_t) \circ g_t + z_t \circ h_{t-1} \quad (13)$$

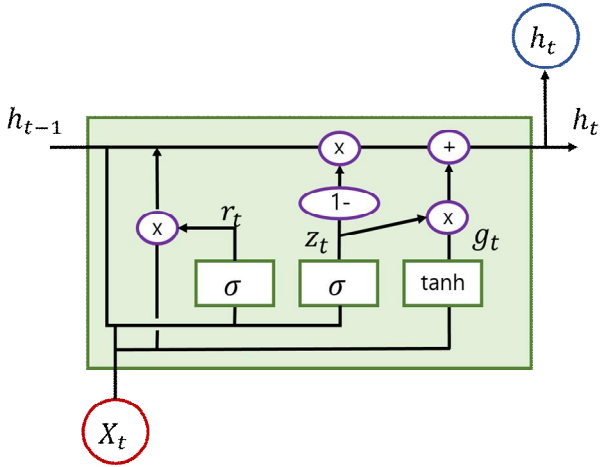


Fig. 3. GRU cell architecture, here  $X_t$ : input time step,  $h_t$ : output,  $r_t$ : reset gate,  $z_t$ : update gate,  $\hat{h}_t$ : internal cell state. Operations inside light red circle are pointwise.

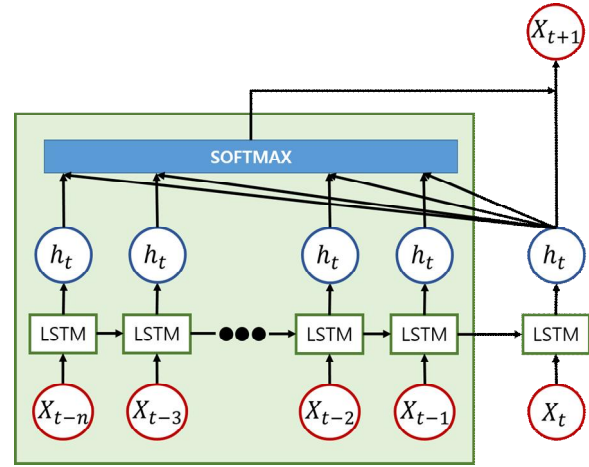


Fig. 5. Attention LSTM network structure with LSTM cells, here  $X_t$ : input time sequence,  $h_t$ : output time sequence.

### 2.4 Attention LSTM(Attention Long-short term memory)

기존 시계열 예측 모델의 경우 학습이 순차적으로 진행됨에 따라 이전에 생성된 학습 정보를 새로운 학습 정보를 활용하여 변경하는 과정을 반복하며 학습이 진행되었다. 이러한 경우 현재 학습되는 학습 데이터와 다른 학습 데이터와의 관계성이 고려되지 않는다. Attention은 각 Cell 별로 학습된 학습 정보들을 종합하여 벡터데이터 형식으로 학습 정보를 처리하게 된다. 이러한 벡터 형식의 학습 정보는 현재 학습 데이터와 다른 학습 데이터와 관계성을 고려하여 예측을 진행하게 해준다.

$X_t$	Query * Key <sup>T</sup>	Score	Softmax	Value
$X_t$	$X_t$	= 150	0.68	$X_t$
$X_t$	$X_{t-1}$	= 23	0.11	$X_{t-1}$
$X_t$	$X_{t-2}$	= 20	0.09	$X_{t-2}$
$X_t$	$X_{t-3}$	= 27	0.12	$X_{t-3}$

} Softmax \* Value  
 $X_t$

Fig. 4. Computing the output of attention pooling as a weighted average of values, where weights are computed with the attention scoring function and the softmax operation.

Self Attention Layer를 구성하는 함수는 Fig. 4와 같이 크게 Query, Key, Value 3가지로 분류할 수 있다. 먼저 Query는 입력데이터에 대한 정보를 나타내는 벡터이며, Key는 입력데이터 외의 다른 학습 데이터의 정보를 나타내는 벡터이며, Value는 학습 데이터의 각 위치정보를 나타내는 벡터이다. Query와 Key 두 벡터는 내적을 통하여 Query와 Key와의 유사

도를 측정하여 score로 계산되며, 계산된 score와 Value 벡터를 이용하여 학습 데이터의 각 위치에 해당하는 데이터와 입력 데이터 사이의 상대적 중요도를 계산하게 된다. 위와 같은 Self Attention Layer를 이용하여 Fig. 5와 같은 구조의 Self Attention 알고리즘을 구축하였다.

## 3. 실험 데이터 및 학습 환경 설정

### 3.1 실험 데이터 설정

본 연구 사용된 국립해양조사원 조위관측소 흑산도 1시간 간격 수온 데이터는 학습에 사용하기 전 자료의 결측값을 liner를 통하여 제거한 후, 일 단위 데이터 자료로 변환하여 사용하였다. Fig. 6과 같이 데이터 수집 기간은 2018년 01월 01일부터 2022년 11월 29일까지 약 5년간의 데이터를 수집하였다. 수집된 데이터를 각 연 단위로 데이터를 분류하여 그 래프로 표현하여 수온의 계절 경향성이 잘 나타나는지 검토를 진행하였으며, Fig. 6에 해당 결과를 도시하였다. 분석 결과 학습에 사용된 모든 연도에서 계절별 수온의 변화가 잘 나타내는 것을 확인하였다. 인공지능 학습 시 학습 데이터에 편향된 학습이 진행되는 과적합을 방지하기 위해 데이터를 학습(Train), 검증(Validation), 시험(Test) 3가지로 분류하여 진행하게 된다. 학습 데이터는 학습에 직접적으로 사용되는 데이터이며, 검증 데이터는 학습 과정에서 학습된 네트워크를 평가하여 학습 정확도 향상을 위해 사용된다. 마지막으로 시험 데이터는 학습이 완료된 네트워크를 평가하기 위한 데이터이며, 검증 데이터를 이용하여 학습 데이터에 편향된 학습을 방지하며, 학습에 관여하지 않는 시험 데이터를 이용하여 학습 모델의 정확도 평가를 진행한다.

## 수온 데이터 예측 연구를 위한 통계적 방법과 딥러닝 모델 적용 연구

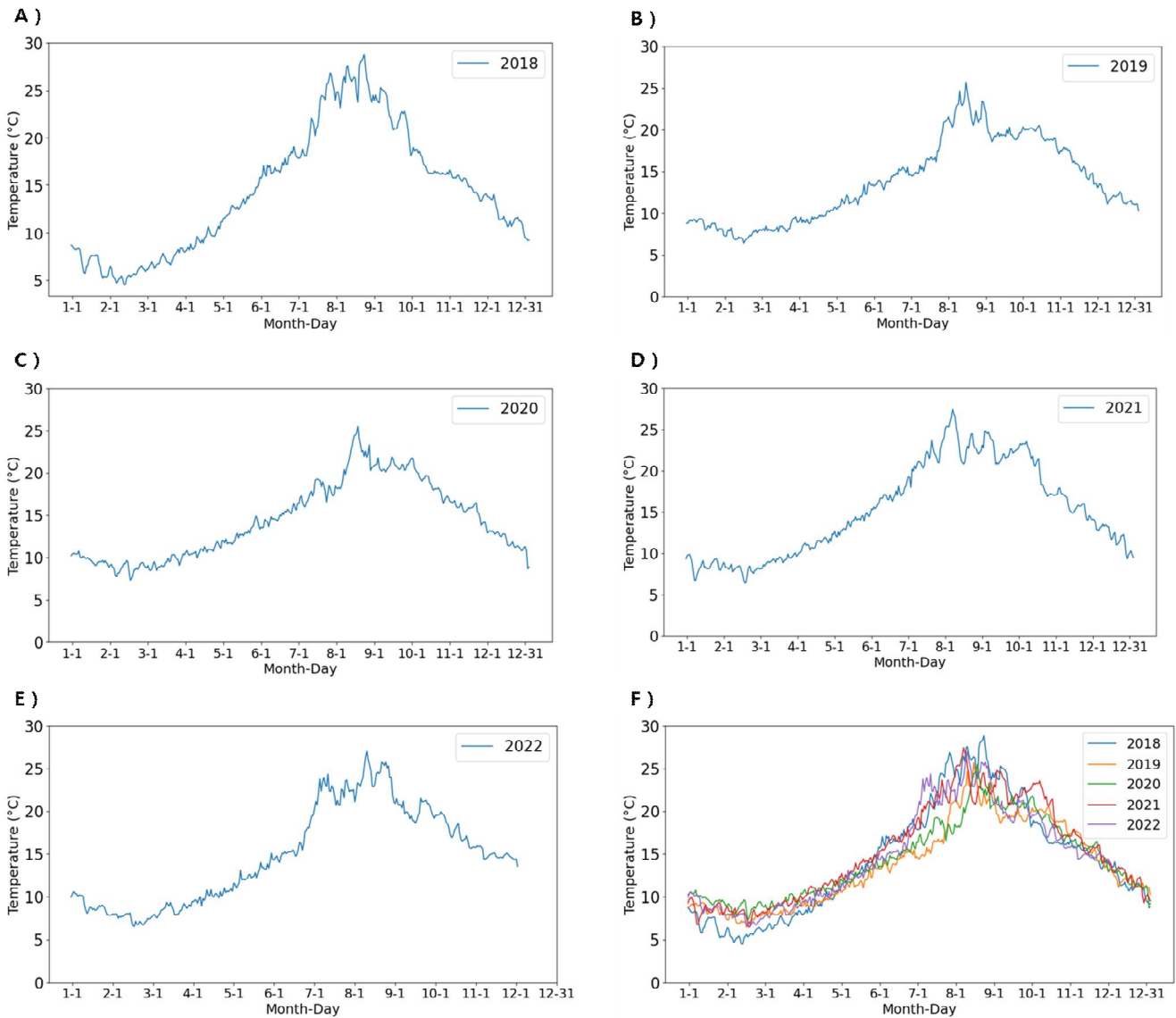


Fig. 6. Inter-annual variation of the daily water temperature time series in the years 2018 (a), 2019 (b), 2020 (c), 2021 (d), 2022 (e), and the entire period (f) of Heuk-san Island, respectively.

본 시계열 예측 연구는 국립해양조사원 흑산도 조위관측소 2018-2020년까지의 수온데이터를 학습자료로 사용하고 2021년, 2022년 자료는 각각 데이터검증, 테스트에 사용하였다. 따라서 시계열 예측에 사용된 학습 : 검증 : 테스트의 비율은 6 : 2 : 2 설정하였으며, 수온의 계절 주기성을 반영하기 위해 최소한의 데이터 자료를 1년으로 구성하였다. 예측에 사용된 학습 데이터는 과거 수온자료 7일을 입력하여 미래 수온자료 T+1의 데이터를 예측하는 형태로 구성하여 실험을 진행하였다.

### 3.2 학습 환경 설정

ARIMA 알고리즘의 경우 AR과 MA의 계수와 차수(I)에 따라 정확도가 달라진다. 그에 따라 적합한 학습 parameter를 설정하는 것이 매우 중요하다. 본 연구에서 사용된 데이터를 이용하여 p-value 분석을 진행하였다. p-value 분석 결과 Table 1과 같은 결과를 확인 할 수 있었다. 해당 데이터를 통하여 본 연구에 사용된 데이터의 경우  $d=1$ 로 설정하였을 때,  $p\text{-value} < 0.05$ 이기에 정상성을 보이는 데이터로 변환이 된 것으로 판단하였다. 따라서  $d = 1$ 로 설정하였으며, Fig. 7과 같은 ACF 그래프와 Fig. 8과 같은 PACF 그래프를 이용하여 AR과 MA의 계수는 (1,0)으로 설정하였다.

Table 1. Augmented Dickey-Fuller test on residuals of ARIMA model

ADF Statistic	p-value	Critical Values	
		1%	5%
-3.007162	0.034231	1%	-3.434
		5%	-2.863
		10%	-2.568

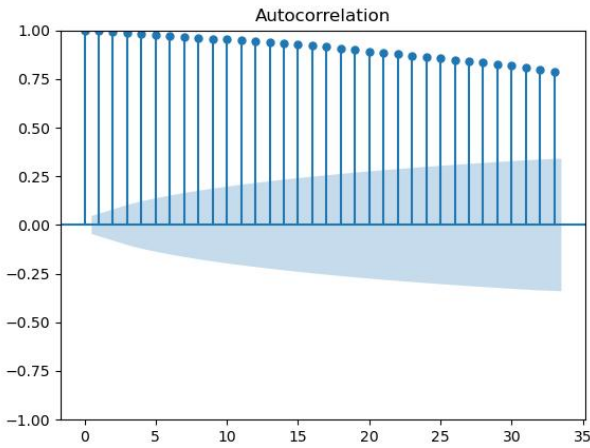


Fig. 7. Autocorrelation against the time-lag with 5% significance limits.

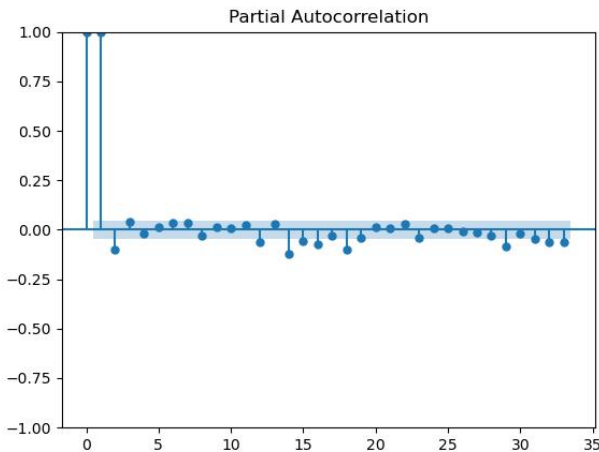


Fig. 8. Partial autocorrelation against the time-lag with 5% significance limits.

인공지능 정확도 검증을 위해 동일한 Hyper parameter를 적용하여 학습 네트워크를 구성하였다. 학습 네트워크의 Hyper parameter는 Table 2와 같이 설정하였으며, Epoch = 500, Learning rate = 0.001, Dropout = 0.2 (20%), batch size = 16으로 설정하였다. 네트워크별 Cell의 개수는 16과 32로 설정하여 총 8개의

학습 결과에 대하여 비교 분석을 진행하였다.

Table 2. Experimental design of deep learning algorithms and related hyper-parameters including cells, learning rate, and drop-out.

Model	Cell	Learning rate	Dropout
LSTM	16	0.001	0.2
GRU			
Attention LSTM	32		

인공지능 예측 결과 비교를 위하여 MAE(Mean absolute error) 지표로 평가하였으며, MAE는 식(14)와 같이 예측 결과와 관측값에 대한 차를 구한 후, 절댓값의 평균을 이용하여 오차를 계산하는 기법으로 인공지능 네트워크 정확도 평가에 대중적으로 이용되는 평가 기법이다.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{x}_i - x_i| \quad (14)$$

#### 4. 학습 진행 및 정확도 평가

ARIMA는 앞선 학습 설정 단계에서 설정된 p, d, q의 값을 (1, 1, 0)로 예측을 진행하였다. 예측을 진행한 결과 Fig. 9와 같은 예측 성능을 보여주었으며, MAE 계산 결과 0.396의 결과를 확인 할 수 있었다.

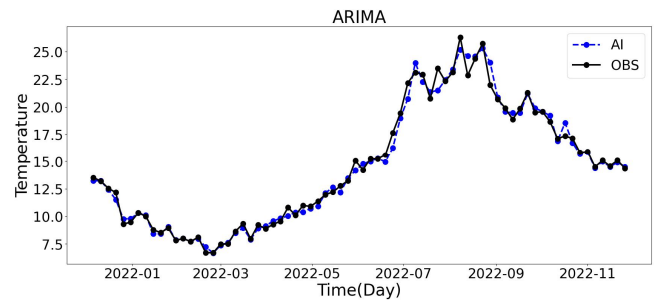


Fig. 9. Daily forecast of water temperature from the ARIMA model.

인공지능 학습은 앞서 설정한 학습 설정 및 정확도 검증 방법을 이용하여 학습 및 예측을 진행하여 Fig. 10 ~ Fig. 15와 같은 결과를 얻을 수 있었다. 이를 앞서 설정한 오차 검증 기법인 MAE를 이용하여 정확도 검증을 진행한 결과 모델별 성능은 Table 3과 같다. 결과를 통해 알 수 있듯이 Cell의 수

수온 데이터 예측 연구를 위한 통계적 방법과 딥러닝 모델 적용 연구

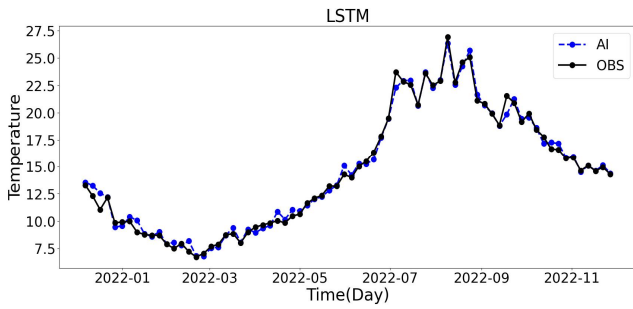


Fig. 10. Daily forecast of water temperature from the LSTM (Cell 16) model.

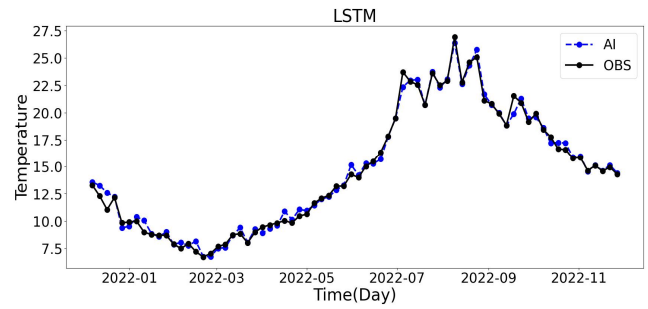


Fig. 11. Daily forecast of water temperature from the LSTM (Cell 32) model.

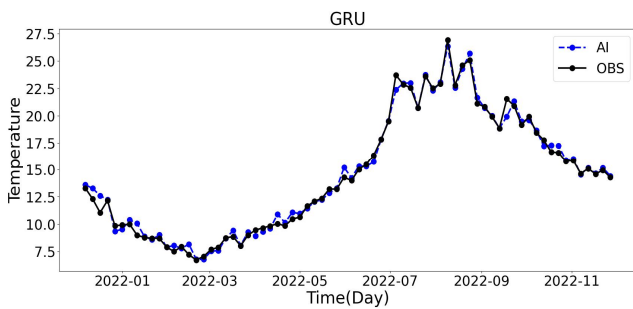


Fig. 12. Daily forecast of water temperature from the GRU (Cell 16) model.

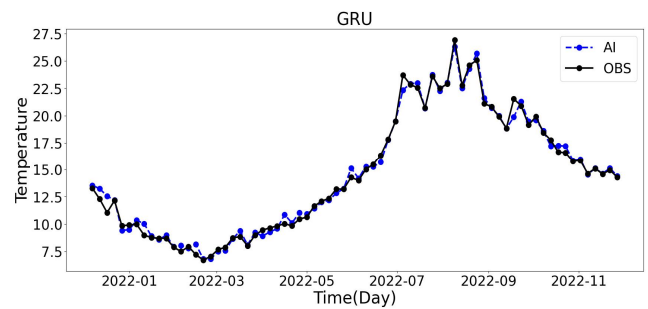


Fig. 13. Daily forecast of water temperature from the GRU (Cell 32) model.

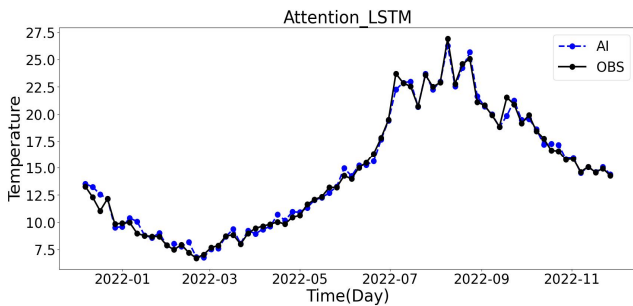


Fig. 14. Daily forecast of water temperature from the Attention LSTM (Cell 16) model.

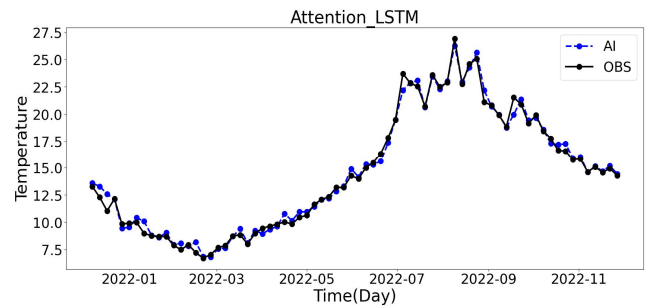


Fig. 15. Daily forecast of water temperature from the Attention LSTM (Cell 32) model.

Table 3. Comparison of performance of the experimental models

Model	Cell	Learning rate	Dropout	MAE
LSTM	16	0.001	0.2	0.392
GRU				0.394
Attention LSTM				0.389
LSTM	32	0.001	0.2	0.393
GRU				0.392
Attention LSTM				0.393

Table 4. Variations in prediction accuracy of the experimental models with an increment in prediction time

Model	T+1	T+2	T+3	T+4	T+5	T+6
LSTM(16)	0.392	0.553	0.667	0.792	0.889	0.980
LSTM (32)	0.393	0.557	0.672	0.802	0.901	0.992
GRU (16)	0.394	0.556	0.673	0.802	0.900	0.991
GRU (32)	0.392	0.552	0.666	0.789	0.893	0.981
Attention LSTM (16)	0.389	0.549	0.666	0.791	0.890	0.976
Attention LSTM (32)	0.393	0.560	0.685	0.815	0.925	1.022

가 16일 경우 Attention LSTM, LSTM, GRU 순서로 정확도가 높음을 확인할 수 있었다. 하지만 Cell의 수가 32로 증가할 경우 다른 알고리즘의 정확도가 향상되어 GRU, Attention LSTM, LSTM 순서로 정확도가 높은 결과를 보여주었다.

추가적으로 학습 기간에 따른 정확도 평가를 위해, T+1 ~ T+6에 대한 예측을 진행하여 정확도 검증을 진행하였다. 해당 기법은 T+1 예측을 진행한 후 예측된 데이터를 학습 데이터에 결합하며, 이렇게 생성된 새로운 학습 데이터로 입력하여 T+2를 예측 진행하게 된다. 이와 같은 입력과 예측을 반복하여 T+6까지 예측을 진행하였다.

예측 결과를 Table 4에 제시하였으며, Cell 16의 경우 기존 T+1 예측 결과와 같이 예측 기간이 증가하여도 Attention LSTM의 성능이 높은 것으로 확인되었다. Cell의 수가 32인 경우 T+1 예측 결과에서는 Attention LSTM, LSTM, GRU가 비슷한 성능을 보이는 것과 다르게 Attention LSTM의 예측 기간이 길어짐에 따라 성능이 낮아짐을 확인할 수 있었다. 최종적으로 Attention LSTM(Cell 16)의 정확도가 가장 높은 것을 확인할 수 있었다.

### 5. Attention LSTM 고도화

앞선 인공지능 알고리즘별 성능평가를 통해 Cell의 변경으로 Attention LSTM의 정확도가 달라지는 것을 보여주었으며, 이를 통하여 Hyper parameter의 설정에 따라서 정확도의 변화가 생기는 것을 확인할 수 있었다. 본 연구에서는 Hyper parameter의 설정에 따른 각기 다른 Attention LSTM 모델의 정확도를 비교하여 수은 예측 모델에 최적화된 Hyper parameter의 값을 제시하고자 한다.

Table 5. Experimental design of Attention LSTM and its hyper-parameters including cells, learning rate, and drop-out.

Cell	Learning rate	Dropout
16	0.01	0.1
		0.2
		0.3
	0.001	0.1
		0.2
		0.3
32	0.0001	0.1
		0.2
		0.3

Table 6. Comparison of performance of the Attention LSTM models

Cell	Learning rate	Dropout	MAE
16	0.01	0.1	0.391
		0.2	0.389
		0.3	0.398
	0.001	0.1	0.391
		0.2	0.389
		0.3	0.398
	0.0001	0.1	0.391
		0.2	0.389
		0.3	0.398
32	0.01	0.1	0.395
		0.2	0.393
		0.3	0.394
	0.001	0.1	0.395
		0.2	0.393
		0.3	0.394
	0.0001	0.1	0.395
		0.2	0.393
		0.3	0.394

학습 데이터의 Train, validation, test의 비율은 이전에 알고리즘별 정확도 검증에서 사용된 것과 동일한 비율로 진행하였으며, 알고리즘 고도화를 위해 변경된 Hyper parameter는 Cell 수, Dropout, Learning rate로 Table 5에 제시된 실험안에 대해 시계열 예측성능 고도화 평가를 수행하였다.

각 모델별 학습 및 예측 평가 결과를 Table 6에 제시하였으며, 평가 결과 Hyper parameter의 Cell 수 = 16, Dropout = 0.2로 설정 시 시계열 예측에 가장 높은 정확도를 보여준을 알 수 있었으며, Learning rate 설정에 따른 예측성능은 결과에 미미한 영향을 주는 것을 확인할 수 있었다.

### 6. 결론 및 향후 연구 방향

본 통계적 방법과 인공지능을 활용한 수은 시계열 데이터 예측 실험 결과 통계적인 방법보다 인공지능을 활용한 결과에서 예측성능이 다소 우수한 것으로 나타났으며, 인공지능 학습 시 Cell 수가 16일 경우 Attention LSTM이 가장 높은 정확도를 나타냈으며, 32일 경우 GRU가 가장 높은 정확도를 보여주었다. 장기 예측의 경우 Cell 수를 16으로 설정한 Attention LSTM이 타 알고리즘과 비교하여 높은 정확도를 보여주었으며, 이는 LSTM과 GRU와 다르게 Attention LSTM



알고리즘은 긴 시계열 데이터의 관계성을 학습하고 특징을 추출할 수 있기 때문인 것으로 판단된다(Sarkar et al., 2020). 따라서 시계열 예측 성능이 우수한 Attention LSTM을 활용하여 예측 성능 고도화를 수행하였으며, 해당 알고리즘의 Hyper parameter Cell 수=16, Dropout=0.2로 설정 시 가장 높은 정확도를 나타내었다. 다만 성능 평가에 사용된 Learning rate에 따른 학습 정확도에 미치는 영향은 미미한 것으로 나타났다.

향후 Attention LSTM을 활용한 시계열자료 예측 시 추가적인 기상 및 다른 관측소의 수온자료를 함께 활용한다면 예측 성능 개선뿐만 아니라 고수온, 저수온과 같은 이상 해양 현상 탐지 활용에 가능할 것으로 기대된다.

### 감사의 글

이 논문은 해양수산과학기술진흥원의 ‘해양위성영상 분석 활용 기술 개발(20210046)’ 연구사업의 지원을 받아 수행되었습니다.

### References

- [1] Breaker, L. C.(2006), Nonlinear aspects of sea surface temperature in Monterey Bay. *Progress in Oceanography*, 69(1), pp. 61-89.
- [2] Cassola, F., F. Ferrari, A. Mazzino, and M. M. Miglietta (2016), The role of the sea on the flash floods events over Liguria (northwestern Italy). *Geophysical Research Letters*, 43(7), pp. 3534-3542.
- [3] Chen, Y., J. T. Randerson, D. C. Morton, R. S. DeFries, G. J. Collatz, P. S. Kasibhatla, L. Giglio, Y. Jin, and M. E. Marlier(2011), Forecasting fire season severity in South America using sea surface temperature anomalies. *Science*, 334(6057), pp. 787-791.
- [4] Cho, Y. K. and K. Kim(1996), Seasonal variation of the East Korea Warm Current and its relation with the cold water. *La mer*, 34, pp. 172-182
- [5] Kartal, S.(2023), Assessment of the spatiotemporal prediction capabilities of machine learning algorithms on Sea Surface Temperature data: A comprehensive study. *Engineering Applications of Artificial Intelligence*, 118, 105675.
- [6] Lee, C. K., H. C. Kim, S. G. Lee, C. S. Jing, H. G. Kim, and W. A. Lim(2001), Abundance of Harmful Algae, *Cochlodinium polykrikoides*, *Gyrodinium impudicum* and *Gymnodinium catenatum* in the Coastal Area of South Sea of Korea and Their Effects of Temperature, Salinity, Irradiance and Nutrient on the Growth in Culture. *Korean Journal of Fisheries and Aquatic Sciences*, 34(5), pp. 536-544.
- [7] Saunders, M. A. and A. S. Lea(2008), Large contribution of sea surface warming to recent increase in Atlantic hurricane activity. *Nature*, 451(7178), pp. 557-560.
- [8] Sarkar, P. P., P. Janardhan, and P. Roy(2020), Prediction of sea surface temperatures using deep learning neural networks. *SN Applied Sciences*, 2(8), 1458.
- [9] Shabbar, A. and W. Skinner(2004), Summer drought patterns in Canada and the relationship to global sea surface temperatures. *Journal of Climate*, 17(14), pp. 2866-2880.
- [10] Thomson, M. C., F. J. Doblas-Reyes, S. J. Mason, R. Hagedorn, S. J. Connor, T. Phindela, A. P. Morse, and T. N. Palmer(2006), Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature*, 439(7076), pp. 576-579.
- [11] Verdin, J., C. Funk, G. Senay, and R. Choulaton(2005), Climate science and famine early warning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1463), pp. 2155-2168.
- [12] Xie, J., J. Zhang, J. Yu, and L. Xu(2019), An adaptive scale sea surface temperature predicting method based on deep learning with attention mechanism. *IEEE Geoscience and Remote Sensing Letters*, 17(5), pp. 740-744.
- [13] Yamak, P. T., L. Yujian, and P. K. Gadosey(2019), A comparison between arima, lstm, and gru for time series forecasting, *Proceedings of the 2019 2nd international conference on algorithms, computing and artificial intelligence*, pp. 49-55.
- [14] Zhang, Z., X. Pan, T. Jiang, B. Sui, C. Liu, and W. Sun (2020), Monthly and quarterly sea surface temperature prediction based on gated recurrent unit neural network. *Journal of Marine Science and Engineering*, 8(4), p. 249.

Received : 2023. 09. 08.

Revised : 2023. 09. 27.

Accepted : 2023. 10. 27.