

Article

<https://doi.org/10.7850/jkso.2024.29.2.101>
pISSN : 1226-2978 eISSN : 2671-8820

한반도 고수온 예측 시스템의 수온 과소모의 보정을 위한 LSTM 모델 구축 및 예측성 평가

임나경¹ · 진현근² · 박균도³ · 박영규⁴ · 김경옥⁴ · 최용한⁵ · 김영호^{6*}¹국립부경대학교 지구환경시스템과학부 해양학전공 학생, ²한국해양과학기술원 해양순환기후연구부 무기계약직기술원,³한국해양과학기술원 해양순환기후연구부 선임연구원, ⁴한국해양과학기술원 해양순환기후연구부 책임연구원,⁵극지연구소 해양대기연구본부 선임연구원, ⁶국립부경대학교 지구환경시스템과학부 해양학전공 조교수

Development and Assessment of LSTM Model for Correcting Underestimation of Water Temperature in Korean Marine Heatwave Prediction System

NA KYOUNG IM¹, HYUNKEUN JIN², GYUNDO PAK³, YOUNG-GYU PARK⁴, KYEONG OK KIM⁴, YONGHAN CHOI⁵ AND YOUNG HO KIM^{6*}¹Undergraduate Student, Division of Earth Environmental System Science, Pukyong National University, Busan 48513, Korea²Full-time Research Specialist, Ocean Circulation & Climate Research Department, Korea Institute of Ocean Science & Technology, Busan 49111, Korea³Senior Research Scientist, Ocean Circulation & Climate Research Department, Korea Institute of Ocean Science & Technology, Busan 49111, Korea⁴Principal Research Scientist, Ocean Circulation & Climate Research Department, Korea Institute of Ocean Science & Technology, Busan 49111, Korea⁵Senior Research Scientist, Korea Polar Research Institute, Division of Ocean & Atmosphere Sciences, Incheon 21990, Korea⁶Assistant Professor, Division of Earth Environmental System Science, Pukyong National University, Busan 48513, Korea*Corresponding author: yhokim@pknu.ac.kr

Editor Byoung-Ju Choi

Received 16 April 2024; Revised 28 May 2024; Accepted 29 May 2024

ABSTRACT

해양의 고수온 현상은 지구온난화로 인한 주요 문제 중 하나로, 식량 자원의 감소와 해양 탄소 흡수력의 저하 등, 해양 생태계와 인류에게 직접적인 위협으로 부상하고 있다. 따라서, 한반도 주변 해역에서의 고수온 예측은 해양 환경 모니터링 및 관리에 중요하다. 본 연구에서는 역학 모델 기반 한반도 고수온 예측 시스템의 성긴 해양의 수직격자체계로 인한 고수온 예측의 과소모의를 개선하기 위해 LSTM 모델을 개발하였다. 2023년에 대해 수행된 한반도 고수온 예측 시스템의 고수온 예측 결과와 LSTM 모델의 결과를 기반으로 한반도 주변의 동해 해역, 황해 해역 그리고 남해 해역에서의 고수온 예측 성능을 평가했다. 본 연구에서 개발된 LSTM 모델이 세 영역 모두에서 수온이 상승하는 시기에 수온 예측 성능을 크게 개선하는 것으로 나타났으며, 수온 상승이 시작되기 전이나 하강하는 시기에는 예측 성능의 개선 효과가 미미했다. 이는 LSTM 모델이 성층이 강화되는 환경에서 성긴 수직격자로 인해 발생하는 고수온 예측의 과소모의를 개선할 수 있는 가능성을 보여준다. 향후 역학 모델의 예측 성능 개선이나 역학 모델의 대체에 자료 기반 인공지능 모델의 활용성이 확대될 것으로 기대한다.

The ocean heatwave is emerging as a major issue due to global warming, posing a direct threat to marine ecosystems and humanity through decreased food resources and reduced carbon absorption capacity of the oceans. Consequently, the prediction of ocean heatwaves in the vicinity of the Korean Peninsula is becoming increasingly important for marine environmental monitoring and management. In this study, an LSTM model was developed to improve the underestimated prediction of ocean heatwaves caused by the coarse vertical grid system of the Korean Peninsula Ocean Prediction System. Based on the results of ocean heatwave predictions for the Korean Peninsula conducted in 2023, as well as those generated by the LSTM model, the performance of heatwave predictions in the East Sea, Yellow Sea, and South Sea areas surrounding the Korean Peninsula was evaluated. The LSTM model developed in this

study significantly improved the prediction performance of sea surface temperatures during periods of temperature increase in all three regions. However, its effectiveness in improving prediction performance during periods of temperature decrease or before temperature rise initiation was limited. This demonstrates the potential of the LSTM model to address the underestimated prediction of ocean heatwaves caused by the coarse vertical grid system during periods of enhanced stratification. It is anticipated that the utility of data-driven artificial intelligence models will expand in the future to improve the prediction performance of dynamical models or even replace them.

Keywords: Marine heatwave, Machine learning, Artificial intelligence, LSTM

1. 소 개

지구온난화로 인해 전 세계적으로 발생하고 있는 해양의 고수온 문제는 해양 생태계에 심각한 영향을 끼치고 있으며, 이로 인한 경제적 손실 또한 증가하고 있다(Oliver *et al.*, 2018). 한반도 해역에서도 2010년대 이후 거의 매년 고수온이 발생하고 있으며(Han and Lee, 2020), 2018년에 발생한 고수온은 그 정도가 강하여 600억 원 규모의 피해를 야기하였다(KMA, 2018). 한반도 해역에서 고수온 문제가 심각해짐에 따라 고수온의 원인을 분석하기 위한 연구가 활발히 진행되고 있다. 2021년 남서부 일부 해역을 포함한 동중국해에서 고수온이 장기간 지속되었고, 원인을 파키스탄 폭우로 인해 강하게 발달된 지구 순환 원격상관으로 지목하고 있다(Oh *et al.*, 2023). 또한, 서태평양의 고수온 원인은 엘니뇨 쇠퇴 등으로 인한 해양기인 고수온과 대기압과 태양 복사열로 인한 대기기인 고수온이 있음을 확인하였다(Lee *et al.*, 2020). 이렇듯 고수온은 강우, 폭염 등 다양한 기작에 의해 발생, 지속기간, 강도가 달라지기 때문에 이를 전문적으로 예측하는 예측 시스템 개발이 필요하다(Huang *et al.*, 2021; Tan *et al.*, 2022; Lee *et al.*, 2022).

한국해양과학기술원, 극지연구소, 부경대학교 등 관련 기관들은 2023년 해양-대기 결합 모형을 기반으로 한 한반도 고수온 예측 시스템(KMHPS; Korea Marine Heatwave Prediction System)을 구축하였다(Fig. 1). 이를 통해 2016년, 2017년, 2018년, 2021년의 고수온 현상을 재현한 결과 2016년, 2017년, 2018년에는 수온의 상승과 고수온 정점 시기, 하강 시기를 잘 모의하는 것을 확인할 수 있었다. 하지만, 전반적으로 수온을 과소모의하고 있으며 2018년과 2021년에는 2°C 이상 과소모의하였다.

일반적으로 역학 모델의 예측 성능은 수직격자의 해상도에 영향을 받는다. 특히, 성층이 강하게 발달한 해양 환경에서 저 해상도의 수직격자는 고수온을 예측함에 있어 오차를 발생시킬 수 있다(Lee *et al.*, 2011). KMHPS의 해양 모형 상층 수직격자 두께는 5 m로 가장 상층의 격자가 상층 5 m의 수온을 대표하게 된다. 그에 반해 GLORYS (Global Ocean Reanalysis and Simulation) 재분석자료를 생산한 모형의 상층 수직격자 두께는 약 1 m로 KMHPS의 수직격자보다 조밀하다. GLORYS 재분석자료는 고수온 현상을 재현한 모든 실험에서 KMHPS보다 성층이 강화되는 시기의 수온 수직 구배를 보다 현실적으로 표현하는 것으로 나타났다. 이에 본 연구는 KMHPS 과소모의의 원인이 해양 모형 상층 수직격자의 해상도로부터 기인했다는 가능성을 두고, 이를 가정하여 진행하였다.

기상 및 해양 분야의 역학 모델은 복잡한 계산 과정을 가지고 있어 예측하는데 많은 시간과 비용이 소모되고, 해상도를 높이는 경우 계산량은 기하급수적으로 증가하게 된다(Ahn *et al.*, 2012). 그에 반해 기계학습 모형은 훈련 과정에서는 많은 전산자원을 요구하지만, 예측 단계에서는 비교적 적은 자원을 사용한다(Hu *et al.*, 2021). 이에 기상 및 해양 분야에서는 수치예보모델의 정확도 및 예측 성능 향상을 위해 기계학습을 활용하는 연구가 활발히 이뤄지고 있다(Hu *et al.*, 2021; Jeong *et al.*, 2023). Hu *et al.*(2021)은 XGBoost (Extreme Gradient Boosting)와 LSTM (Long Short-Term Memory)을 사용하여 호수의 파도 높이를 예측하였고, 그 결과 두 기계학습 모두 WW3 (WaveWatch III)에 비해 유의파고의 과소모의가 줄어들었다. Jeong *et al.*(2023)은 수치예보모델의 후처리 과정에 기계학습 방법을 활용하여 수치 모델 예측값의 편차를 보정하는 모형을 개발하였고, 그 결과 36시간 예측에 대하여 예측 오차가 평균적으로 편차 보정 전 0.2%, 편차 보정 후 21.0% 감소하였다.

본 논문에서는 성긴 수직 해상도에서 기인한 KMHPS의 과소모의를 보정하는 LSTM 모델을 구축하였다. 이는 LSTM 모델을 활용하여 비교적 적은 자원으로 고수온의 예측 정확도를 향상시키는 것이 목적이다. 고수온은 지속적인 폭염과 강우 등에 영향을 받기 때문에(Lee *et al.*, 2022) 시계열 데이터 처리에 뛰어난 성능을 가진 LSTM 모델을 사용하였다(Hochreiter and Schmidhuber, 1997; Zhang *et al.*, 2017; Siami-Namin *et al.*, 2018). 또한, 고해상도 수직격자 자료를 제공하는 GLORYS 재분석자료를 훈련 및 테스트 자료로 활용하였다. 구축한 LSTM 모델을 통해 예측된 수온은 OSTIA 표층 수온과 비교하였고, 이를 통해 LSTM의 고수온 예측 성능을 평가하였다.

2. 자 료

2.1 한반도 고수온 예측 시스템(KMHPS)

KMHPS는 해양-대기 결합 자료동화 시스템과 해양-대기 결합 중기예측 시스템으로 구성된다. 해양-대기 결합 모형은 해양순환모형(GFDL-MOM5)과 대기순환모형(WRF-ARW Version 4.1.5)으로 구성되며, 대기모형-해양모형 간 변수를 교환해주는 결합자는 OASIS3-MCT를 바탕으로 개발되었다. 대기모형은 총 5개의 변수(단파복사, 현열, 잠열, x응력, y응력)를 해양모형으로 전달하고, 해양모형은 총 3개 변수(SST, SSU, SSV)를 대기모형으로 전달하며, 변수의 교환은 1시간 간격으로 이뤄진다. 또한, 해양모형의 수직격자는 Z* 격자체계로 0-5,000 m까지 가변격자의 51개층으로 구성되었으며, 표층으로부터 첫 층의 수직격자 두께는 5 m이다.

KMHPS은 2023년 7월 6일부터 9월 6일까지 7일 간격으로 매 예측에 대해 14일 동안 총 8차례 고수온 예측을 수행하였고, 본 연구에서는 Lee *et al.*(2023)를 참조하여 동해(ES), 황해(YS), 남해(SS) 영역으로 나누어 예측 성능을 평가하고, 예측 성능을 향상시키고자 인공지능 LSTM 모델을 개발하였다(Fig. 2 참조).

Fig. 3은 14일의 예측 기간에 대해 각 영역에서 공간 평균한 OSTIA의 관측 SST와 KMHPS으로 예측된 SST의 특징을 나타낸다. 구체적으로, 이 그림은 관측된 SST(적색)와 예측된 SST(청색)에서 예측 기간의 첫 날에 해당하는 관측된 SST를 뺀 값으로, 즉 첫날의 관측 SST를 0으로 두고, 그에 대비 관측 및 예측된 SST를 나타낸 것이다. 얇은 실선은 7월 6일부터 9월 6일까지 8번에 걸쳐 예측된 SST와 같은 시간의 관측 SST를 각 영역에 대해 평균하여 예측 기간에 대해 나타낸 것이며, 모든 경우에 대해 평균하여 굵은 선으로 나타내었다. Fig. 3에서 볼 수 있듯이, KMHPS은 동해(ES), 황해(YS), 그리고 남해(SS) 영역에서 예측된 SST가 일관되게 과소 평가되고 있음을 드러낸다. 동해(ES)의 경우, 예측 4일차까지는 KMHPS의 평균 SST가 OSTIA의 평균 SST보다 높지만, 예측 기간이 진행됨에 따라 과소 평가가 점점 심해지는 경향을 보인다. 황해(YS)와 남해(SS)에서는 예측된 SST가 일관되게 과소 평가되고 있으며, 특히 남해에서는 예측 14일차에 2°C 이상 과소 평가되는 현상이 나타난다. 이러한 발견을 토대로, KMHPS의 고수온 과소모의를 보정하기 위한 인공지능 기반의 LSTM 모델을 개발하였다.

2.2 GLORYS 재분석자료

본 연구에서는 KMHPS에서 해양 모형의 성긴 수직해상도에 의해 고수온이 과소모의되는 문제점을 평가하기 위한 비교 자료로, 그리고 고수온 과소모의를 보정하기 위한 LSTM 모델의 훈련자료로 GLORYS (Global Ocean Reanalysis and Simulation) 재분석자료를 활용(CMEMS, 2023)하였다. GLORYS 재분석자료는 1/12°(약 8 km)의 수평 해상도와 50개의 수직 층을 갖는 전지구 고해상도 재분석자료이다. NEMO (Nucleus for European Modeling of the Ocean) 해양모형을 사용

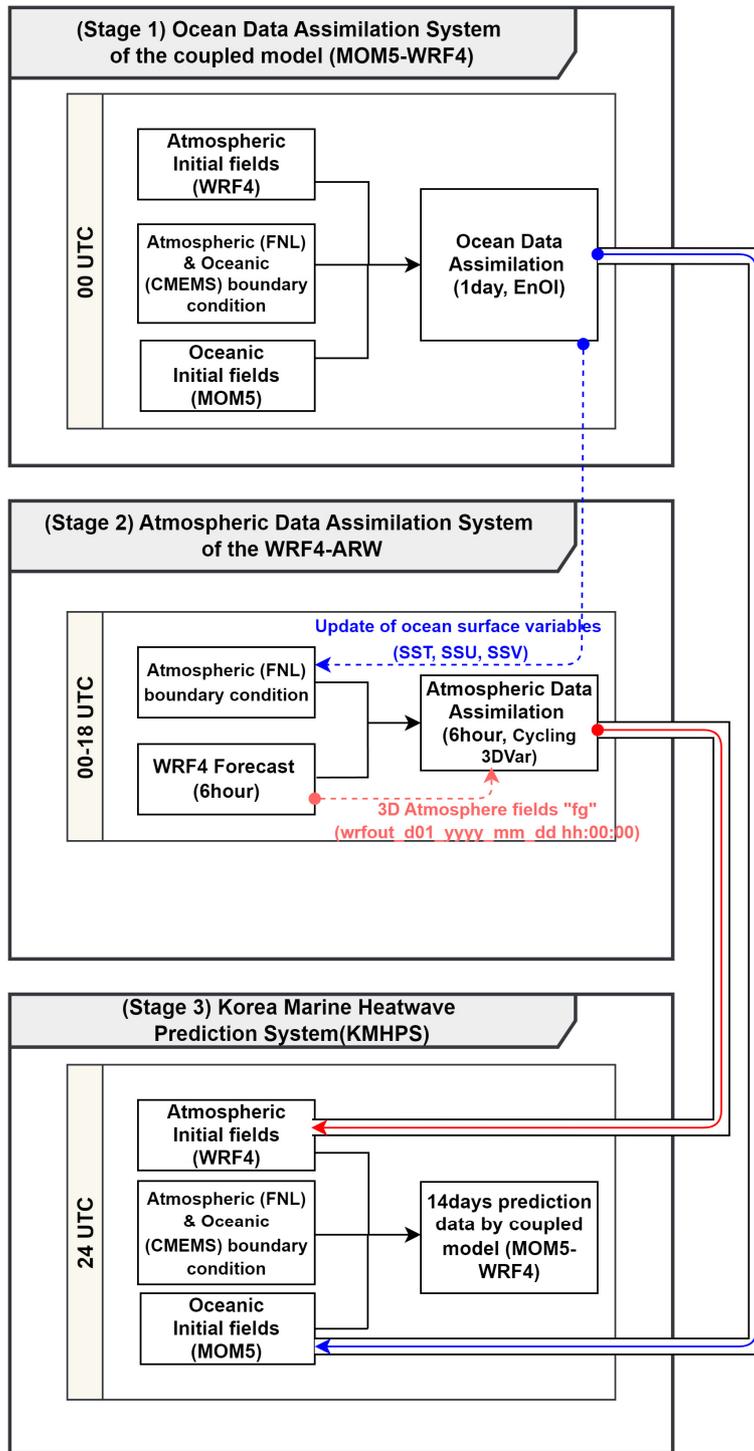


Fig. 1. Schematic diagram of Korea Marine Heatwave Prediction System operation (KIOST, 2024).

하여 생성되었으며, ECMWF (European Centre for Medium-Range Weather Forecasts)의 ERA-Interim자료로부터 해면 경계 조건을 처방하였다. NOAA로부터 위성관측 해면수온, CMEMS로부터 위성관측 해면고도, CMEMS로부터 현장관측 수온과 염분 프로파일 그리고 Centre ERS d'Archivage et de Traitement로부터 얻어진 해빙농도를 Reduced-order Kalman

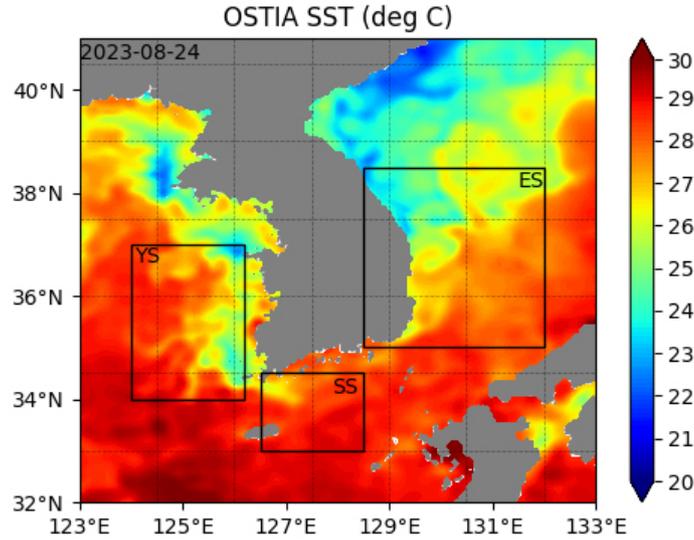


Fig. 2. Observed SST of OSTIA on August 24, 2023, and Study area in the East Sea (ES, 128.5-132°E, 35-38.5°N), Yellow Sea (YS, 124-126.2°E, 34-37°N) and South Sea (SS, 126.5-128.5°E, 33-34.5°N).

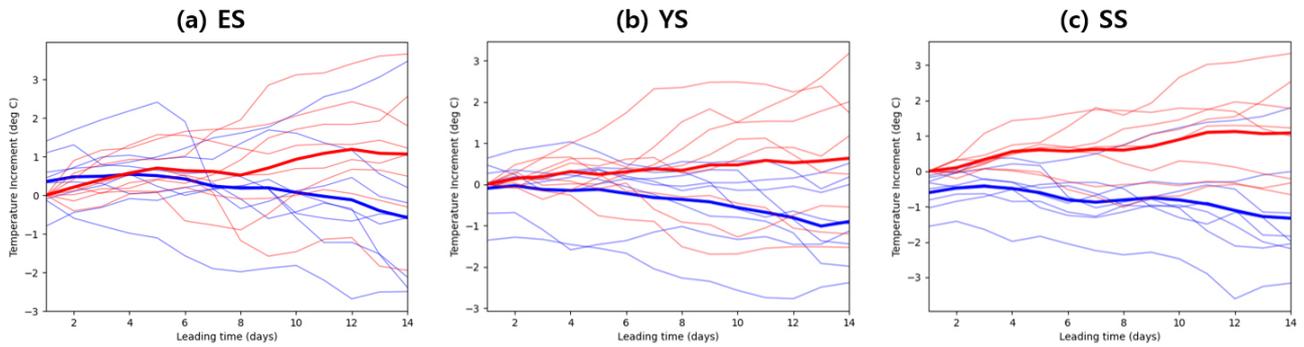


Fig. 3. SST changes in OSTIA observations compared to predictions made by the Korea Marine Heatwave Prediction System are analyzed relative to the observed SST on the first day of each prediction periods in the ES (a), YS (b), and SS (c) regions. The red (blue) line represents the observed SST (predicted SST), with thicker lines indicating the averages.

Filter를 적용하여 해양모델에 동화하였다. 제공되는 자료는 1993년부터 2023년까지의 일일 및 월별 평균 해양 변수이며, 해면고도와 3차원 수온, 염분, 속도 장을 포함한다(CMEMS, 2023).

GLORYS 재분석자료는 약 1 m의 상층 수직격자를 가지고 있어 5 m 상층 수직격자를 가진 KMHP가 표현하지 못하는 상층의 수직 변화를 표현할 수 있다. 이에 KMHP의 과소모의를 보정하는 LSTM을 구축하기 위한 훈련 및 테스트 자료로 고수온이 강하게 발달했던 2018년부터 2021년까지 4년의 GLORYS재분석자료를 사용했다.

2.3 태양 단파 복사

SSR (Surface net Shortwave Radiation), 즉 지표면 순단파 복사량은 태양으로부터 오는 단파 복사 에너지 중 지표면에 의해 반사되어 다시 우주로 돌아가는 양을 제외한, 실질적으로 지구 표면에 도달하여 흡수되는 에너지의 양을 의미한다. SSR은 해양 표층의 온도를 상승시키는 주요 요인으로, 해양 열파 형성에 결정적인 영향을 미친다. 특히 고수온이 발생하는 해역은 대체로 고기압의 영향으로 인한 맑은 날씨와 강한 태양 복사가 지속되는 조건이 일치한다(Lee *et al.*, 2022). 증가한 SSR

은 해수면 온도의 급격한 상승을 유발하며, 이는 해양 표층의 성층을 강화하고, 상층부의 수직 혼합을 감소시켜 고수온을 유발한다. 또한, 바람이 약하고 해류가 느릴 때는 해양 깊은 곳으로의 열 전달이 저해되어, 표층의 온도 상승이 더욱 두드러지게 된다(Tan *et al.*, 2022).

따라서, 고수온 현상의 예측 정확도를 개선하기 위해 학습과정에서 ECMWF (European Centre for Medium-Range Weather Forecasts) 에서 제공하는 ERA5 (ECMWF Reanalysis v5)의 SSR 데이터를 LSTM 모델의 입력 변수로 사용하였고(Hersbach *et al.*, 2023), 예측과정에서는 KMHPS에서 예측된 SSR을 입력변수로 사용하여 동일한 형태의 입력 데이터로 구성하였다. 이는 모델이 해양 표층 온도 변화를 보다 정확하게 예측하는데 기여할 것으로 기대된다.

3. 모 델

Fig. 4(a)는 GLORYS 재분석자료를 기반으로 2021년 동해 (ES) 영역의 공간 평균 수온 자료를 나타내며, 시간에 따른 원시 자료의 수심별 변화를 5월부터 8월까지 보여준다. 이와 대조적으로, Fig. 4(b)는 표층의 첫 격자와 표층 5 m의 평균 수온 간의 격차를 보여준다. 겨울철에는 일반적으로 표층 혼합층이 깊게 형성되어 표층 5 m 내에서 수직적인 수온 변화가 크지 않지만, 여름철에는 혼합층이 얇게 형성되거나 표층에서부터 급격한 수온 변화가 나타날 수 있다. 그렇기 때문에 GLORYS 재분석자료가 수직격자 간격이 표층부터 약 1 m로 조밀하게 나뉘어 있어 수온의 수직 변화를 정확하게 표현하기에 비교적 적합하다. 그러나 성긴 수직격자 체계를 사용하는 모델은 수온의 상승을 과소모의할 가능성이 있다. 따라서, 여름철에는 수온의 수직 구배가 큰 환경에서 고수온 현상이 발생할 경우, 예측 성능을 향상시키기 위해서 더 조밀한 표층 수직격자 체계가 필요하다. Fig. 4(a)에서 확인할 수 있듯이, 수온이 상승하는 시기에는 표층 5 m 내에서 수온의 수직 구배가 증가하고, 수온이 감소하는 시기에는 오히려 수온의 수직 구배 또한 감소하는 것을 확인할 수 있다. KMHPS에서 표층의 첫 격자가 5 m에 위치하기 때문에 수온이 상승하는 시기에는 표층 수온의 상승이 과소모의될 수 있다. Fig. 4(b)는 실제로 표층의 격자가 5 m인 경우를 가정하여 표층 5 m의 수온을 평균한 후, 원시 자료의 표층(0.49 m) 수온과의 차이를 나타내었다. 이 결과는 표층의 성긴 격자로 인해 고수온 예측이 최대 약 1.2 °C까지 과소모의될 수 있음을 시사한다.

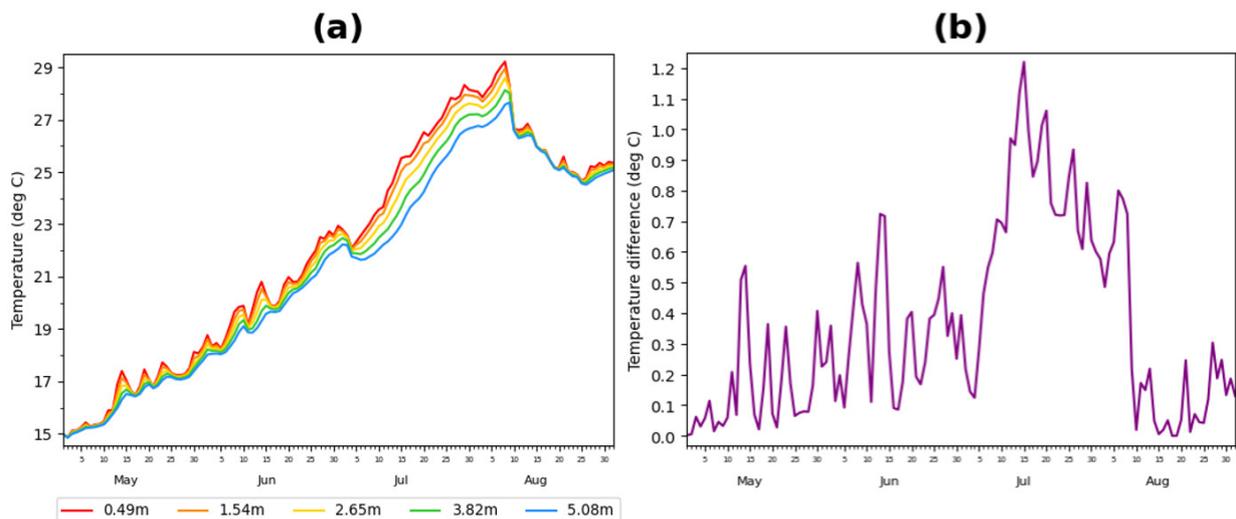


Fig. 4. (a) Temperature profile by grid and (b) difference between SST and mean temperature at 0-5 m depth in East Sea (128.5-132.0°E, 35.0-38.5°N) using GLORYS reanalysis data for 2021. At depths of 0-5 m, there are five grids, and the difference is the value by subtracting the mean temperature from the SST.

따라서, KMHPS의 표층의 성긴 수직격자 체계로 인한 고수온 과소모의를 보정할 필요가 있다. 본 장에서는 조밀한 수직 격자를 가진 GLORYS 재분석자료를 훈련 및 테스트 자료로 활용했다. 입력 자료로 사용할 KMHPS의 첫 번째 격자에 해당하는 표층 5 m의 수직 평균 수온을 GLORYS 자료로부터 얻었으며, SSR을 함께 입력 자료로 사용하였다. 이 연구의 주요 목적은 표층 5 m 격자의 예측 수온을 기반으로 표층 1 m의 수온을 예측하는 인공지능 LSTM 모델을 개발하는 것이다. 본 장에서는 구체적인 인공지능 LSTM 모델에 대해 설명할 것이다.

3.1 LSTM

KMHPS에 의한 고수온 예측 성능을 향상시키기 위해 LSTM (Long Short-Term Memory)을 적용한 인공지능 기계학습 모델을 개발하였다. LSTM은 순환 신경망(RNN; Recurrent Neural Network)의 한 종류로, 시계열 데이터나 자연어 처리 등의 문제에 특히 효과적인 것으로 알려져 있다(Hochreiter and Schmidhuber, 1997). LSTM은 게이트 메커니즘을 활용함으로써 단기 및 장기 메모리를 유지하고 갱신하는 능력을 갖추고 있어, 시계열 데이터의 장기적인 의존 관계를 효과적으로 모델링할 수 있다. 이러한 특성으로 인해 LSTM은 시간에 따라 변화하는 특성을 갖는 수온 예측과 같은 문제에 적합한 것으로 판단하였다.

LSTM의 핵심은 세 가지 주요 게이트를 통한 정보 흐름의 정교한 조절에 있다. 삭제 게이트(forget gate)는 현재 입력과 이전 셀 상태를 바탕으로 중요하지 않은 정보를 ‘잊게’ 하는 기능을 담당한다. 시그모이드 함수(sigmoid function)가 적용되어 0(완전히 잊음)에서 1(완전히 기억) 사이의 값으로 이전 정보의 보존 수준을 결정한다. 다음으로, 입력 게이트(input gate)는 현재 정보가 새로운 셀 상태에 얼마나 기여할지를 조정한다. 이를 위해 두 가지 작업을 수행하는데, 탄젠트 함수를 통해 새로운 후보 상태를 생성하고, 시그모이드 함수를 적용한 가중치를 곱하여 이 후보 상태의 중요성을 결정한다. 마지막으로, 출력 게이트(output gate)는 결정된 셀 상태를 바탕으로 다음 타임 스텝에 출력할 정보를 결정한다. 시그모이드 함수로 셀 상태의 어느 부분을 전달할지 결정하고, 탄젠트 함수로 정규화된 셀 상태와 결합하여 최종 출력을 생성한다(Fig. 5). 이러한 복합적인 메커니즘을 통해 LSTM은 시계열 데이터의 중요한 정보를 장기간 유지하면서, 불필요한 정보는 제거한다(Siami-Namin *et al.*, 2018). 이는 LSTM이 순차적인 데이터 처리에서 장기적 의존성을 효과적으로 모델링하고, 시계열 예측, 자연어 처리 등의 분야에서 뛰어난 성능을 발휘하는 이유이다.

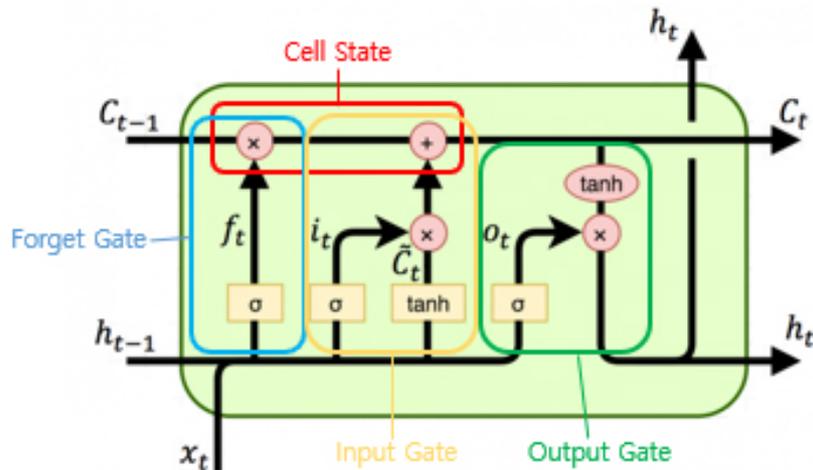


Fig. 5. Architecture of LSTM cell.

3.2 학습 데이터 구조

Fig. 6(a)는 학습 데이터의 구조이다(Kim *et al.*, 2021). Dataset_I와 Dataset_O는 각각 입력과 출력되는 Dataset이다. 본 연구에서 Dataset은 (t,s,n)구조를 갖게 된다. t는 시계열 순서이다. 이는 Fig. 6(a)의 아래 첨자로 표시되어 있다. 7월 6일부터 7월 19일까지의 데이터를 학습시킨다고 가정하였을 때, t는 14가 된다. s는 목표 날짜를 기준으로 전후에 포함되는 날에 따라 결정된다. 예를 들어 7월 6일을 목표 날짜로 지정하고 전후로 하루씩 포함하여 7월 5일부터 7월 7일까지의 자료를 입력하고 목표 날짜인 7월 6일의 자료만 출력하도록 학습시킨다면, Dataset_I는 (t,3,n), Dataset_O는 (t,1,n)의 구조를 갖게 되는 것이다. n은 n가지의 특징 데이터로, Dataset_I와 Dataset_O는 여러 특징 데이터를 가질 수 있다. 이는 Fig. 6(a)의 위 첨자로 표시되어 있다. 본 연구에서 사용한 학습 데이터 구조를 설정하는 과정은 다음과 같다.

1. GLORYS 재분석자료의 2018년부터 2021년까지의 일별 자료를 연구 영역에 대해 공간 평균하여 이용한다. 이에 Dataset은 (1461, s, n)의 구조를 갖게 된다.
2. 목표 날짜의 입력 자료로부터 목표 날짜의 고해상도 수온 자료를 예측하도록 설정한다. 이에 Dataset_I과 Dataset_O의 구조 모두 (1461,1,n)이 된다.
3. GLORYS 재분석자료의 0~5 m의 수온 자료를 평균하여 KMHPs의 해양 상층 격자와 유사한 해상도의 수온 자료를 얻는다.
4. GLORYS의 평균으로 가공한 수온 자료와 ERA5의 SSR를 공간 평균하여 Dataset_I의 특징 데이터로 지정한다. 이에 Dataset_I는 (1461,1,2)의 구조를 갖게 된다.

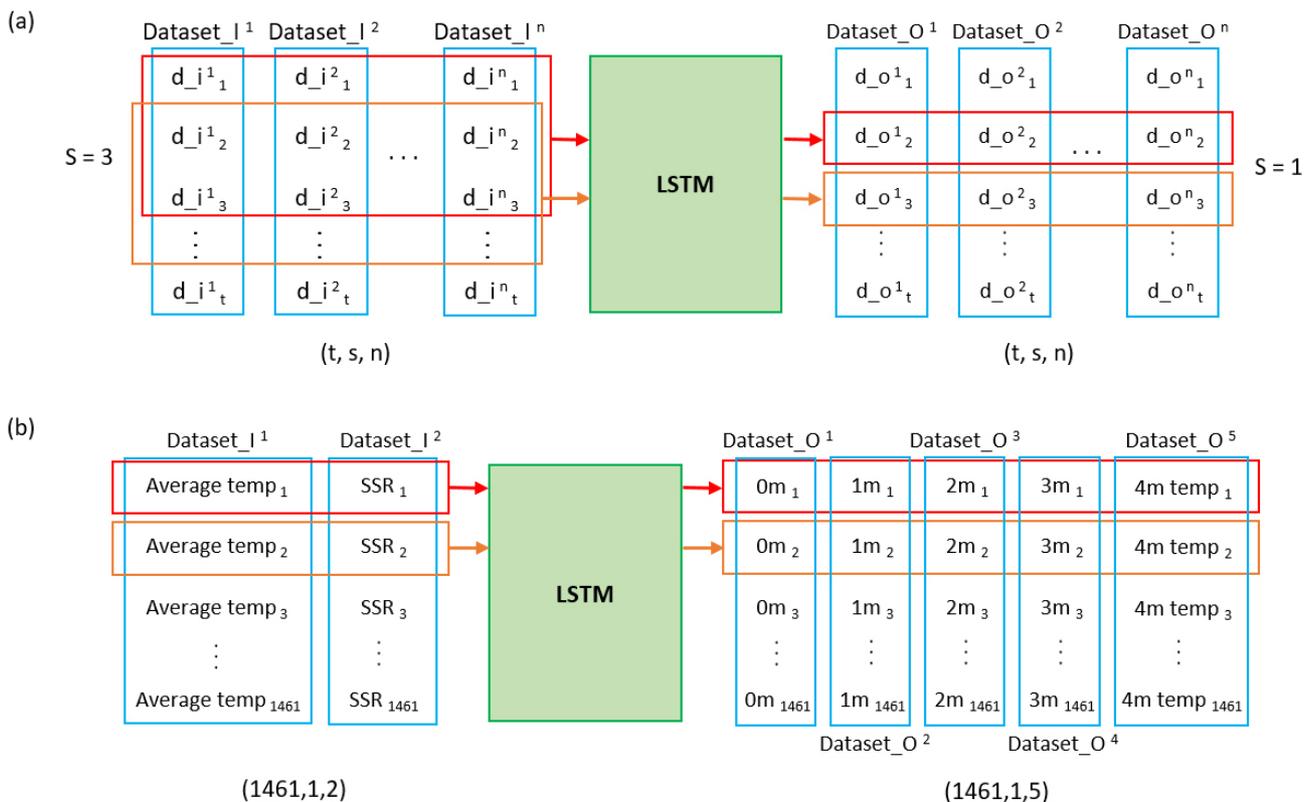


Fig. 6. Architecture of training data (a) and its application in this study (b).

5. 가공하지 않은 GLORYS 재분석자료의 0~5 m의 각 층 공간 평균한 고해상도 수온 자료를 Dataset_O의 특징 데이터로 지정한다. 이에 Dataset_O는 (1461,1,5)의 구조를 갖게 된다.

3.3 전략

설정된 학습 데이터 구조를 통해 고해상도의 수온 값을 예측하는 LSTM 모델을 구축하였다. 100개의 unit를 갖는 LSTM 셀이 3개의 은닉층으로 구성되어 있으며, DENSE 층에서 정보를 받아 고해상도 수온 값을 최종 출력한다. 학습과 예측 과정은 다음과 같다.

1. Dataset_I와 Dataset_O를 훈련 데이터와 테스트 데이터로 분리한다. 본 논문에서는 70%를 훈련 데이터로 사용하였고, 30%를 테스트 데이터로 사용하였다. 이는 과적합을 감지하기 위한 데이터를 설정하는 단계이다.
2. LSTM의 은닉층과 함께 신경망의 25%를 드롭아웃 레이어로 설정한다. 이는 과적합을 방지하기 위한 것으로 신경망의 25%가 무작위로 선택된다.
3. 훈련 데이터를 통해 LSTM을 훈련시킨다. LSTM은 손실함수가 적어지는 최적의 가중치와 편향을 찾는 방향으로 훈련되고, metrics를 통해 모델의 성능 평가도 이뤄진다. 본 연구에서는 MSE (Mean Squared Error)를 손실함수로 설정하였고, 최적화 알고리즘으로 ADAM (Adaptive Moment Estimation) (Kingma and Ba, 2014)을 적용하였다. 모델의 성

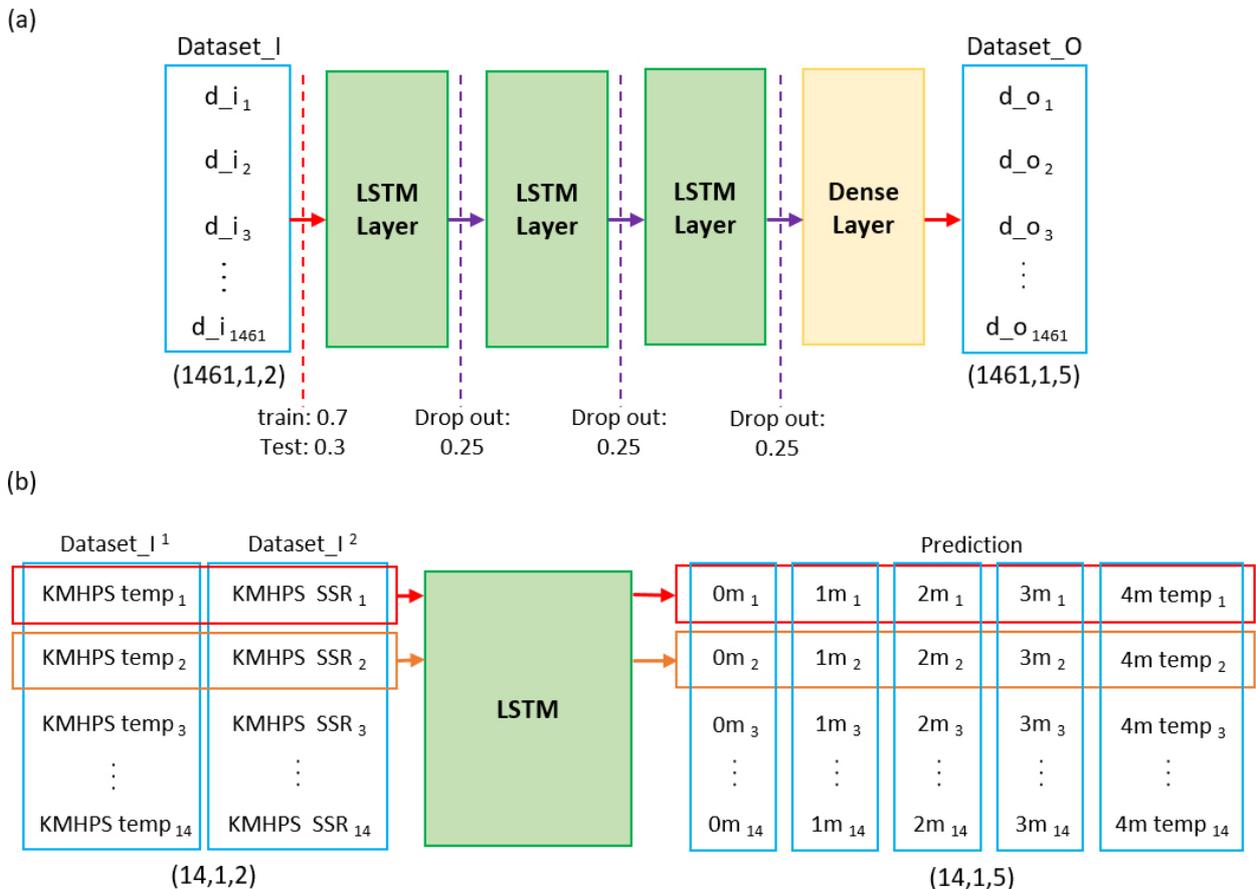


Fig. 7. Architecture of LSTM training (a) and prediction (b) process.

능은 MAE (Mean Absolute Error)로 평가하였다. 또한, 훈련과 동시에 테스트 데이터의 손실함수와 성능도 함께 평가되며 과적합을 감지하였다.

4. 구축된 LSTM 모델을 이용하여 예측을 수행한다(Fig. 7(b) 참조). 본 논문에서는 KMHPs의 상층 수온 자료와 해양-대기 결합 모형의 SSR를 공간 평균하여 입력 자료로 사용하였다.
5. 2023년 7월 6일부터 9월 6일까지 7일 간격으로 매 예측에 대해 14일 동안의 예측을 진행한다. LSTM 모델을 통해 예측된 고해상도 수온 자료 중 가장 표층에 존재하는 약 0 m의 수온 자료와 OSTIA의 표층 수온 자료를 공간 평균하여 비교하고, 그 결과에 따라 예측 성능을 평가하였다.

4. 결 과

KMHPs은 7일 간격으로 매 예측에 대해 14일 동안의 한반도 고수온 예측 자료를 생산한다. 2023년 7월 6일부터 9월 6일까지 수행된 총 8차례의 고수온 예측 결과를 ES, YS, SS의 각 영역별로 나눠 구축된 LSTM 모델에 적용하여 예측 성능을 향상시키는 연구를 수행하였다. 수행된 실험 결과, KMHPs에 의한 고수온 예측의 과소모의가 개선되었음을 확인할 수 있었다. 그러나, 시기에 따라 그리고 영역에 따라 LSTM 모델의 성능이 달라지는 것을 발견하였다. 이러한 차이는 여름철과 겨울철의 수온 패턴의 차이나 각 영역의 해양 환경의 특성에 기인할 수 있다.

ES 영역에서의 수온은 7월 18일부터 상승하기 시작하여 8월 7일에 약 28°C로 정점에 이르렀으며, 이후 8월 17일까지 25.5°C까지 감소한 후에 다시 상승하여 9월 6일까지 26°C에서 27°C 사이를 유지하였다(Fig. 8).

KMHPs에 의한 고수온 예측은 ES 영역에서 7월 20일 - 8월 2일(Fig. 8(c)) 그리고 7월 27일 - 8월 9일(Fig. 8(d)) 예측 실험에서 수온을 2°C 이상으로 과소모의하였다. 7월 20일 - 8월 2일(Fig. 8(c)) 예측에서는 예측된 값과 실제 수온 간의 오차가 예측 4일 후부터 증가하기 시작하여 9일 후부터는 오차가 2°C 이상으로 크게 증가하였다. 7월 27일 - 8월 9일(Fig. 8(d)) 예측에

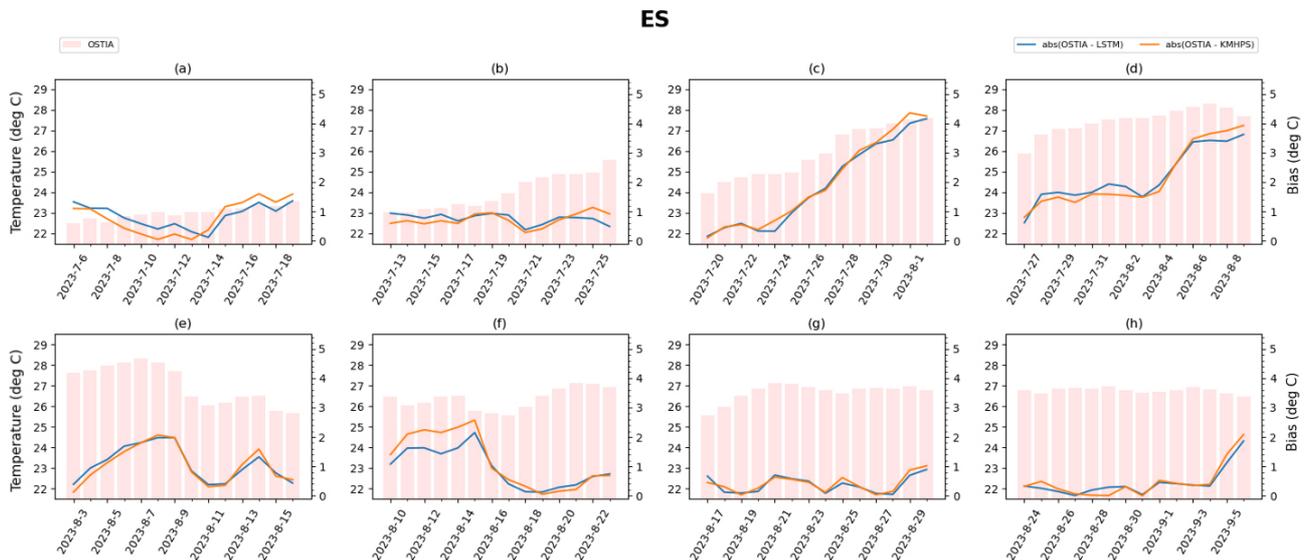


Fig. 8. Observed sea surface temperature (SST) from OSTIA is depicted in light red bar, and while biases of predicted SST by KMHPs and Machine Learning Model are shown in orange and blue, respectively, in the ES region. The light red bar corresponds to the left y-axis, while the blue and orange lines correspond to the right y-axis.

서는 예측된 값과 실제 수온 간의 오차가 예측 9일 후까지는 1.5°C 이하로 유지되다가 10일 후에는 2°C 이상으로 증가하였다. 8월 3일 - 8월 16일(Fig. 8(e)) 그리고 8월 10일 - 8월 23일(Fig. 8(f)) 예측에서는 예측 9일 이내에서 오차가 2°C에 이를 정도로 크게 나타나다가 예측 10일 후부터는 오히려 오차가 감소하는 특징을 보였다. 그 외의 예측 실험에서는 오차가 1°C에서 1.5°C 이내로 유지되었다.

ES 영역에서 LSTM 모델은 7월 6일 - 7월 19일(Fig. 8(a)) 그리고 7월 13일 - 7월 26일(Fig. 8(b)) 예측에서는 KMHPS에 의한 수온의 과소모의를 개선하지 못했을 뿐만 아니라, 오히려 예측 성능을 악화시킨 것으로 나타났다. 그러나, 수온이 상승하는 시기인 7월 20일 - 8월 2일(Fig. 8(c)) 예측에서는 전 기간에서 그리고 7월 27일 - 8월 9일(Fig. 8(d)) 예측에서는 예측 10일 이후에 LSTM 모델에 의해서 예측 성능이 개선되었다. 또한, 8월 3일 - 8월 16일(Fig. 8(e)) 예측에서는 예측 5일 이후부터, 그리고 8월 10일 - 8월 23일(Fig. 8(f)) 예측에서는 예측 초기부터 예측 6일까지 예측 성능이 개선되었다. 한편, 8월 17일 이후 수온 변화가 정체되는 기간에는 LSTM 모델에 의한 예측 성능의 개선은 미미했다.

YS 영역에서의 수온 변화는 ES 영역보다 약간 큰 상승폭과 약간 다른 상승 및 감소 시점을 보이지만, 전반적인 추이는 일관되게 나타났다. 구체적으로, 7월 18일부터 수온이 상승을 시작하여 8월 2일부터 7월 동안 약 28°C로 정점에 이르러, 이후 8월 12일까지 26.5°C까지 감소한 후 다시 상승하여 8월 19일에 28.5°C까지 상승한 뒤 9월 6일까지 감소하였다(Fig. 9).

YS 영역에서 KMHPS에 의한 고수온 예측은 7월 6일 - 7월 19일(Fig. 9(a)) 예측에서 14일간의 예측에서 오차가 1°C 이하로 유지되었다가, 7월 13일 - 7월 26일(Fig. 9(b)) 예측에서 수온이 상승하기 시작하는 7월 18일부터 오차가 증가하여 예측 14일 후에 1.5°C까지 증가하였다. 또한 수온이 증가하거나 높은 수온이 유지되는 기간에 해당하는 7월 20일 - 8월 2일(Fig. 9(c))과 7월 27일 - 8월 9일(Fig. 9(d)) 예측에서 초기부터 예측 14일까지 예측 오차가 급격하게 증가하여 약 4°C까지 증가하였다. 수온의 상승이 정체되는 기간에 해당하는 8월 3일 - 8월 16일(Fig. 9(e)) 예측에서는 1.5 - 2°C의 오차가 유지되고, 수온이 감소하는 기간에는 점차 예측 오차가 감소하여 대체로 오차가 1°C 이하로 유지되었다.

YS 영역에서 LSTM 모델은 대부분의 예측 시기에서 KMHPS에 의한 수온의 과소모의를 성공적으로 개선하는 것으로 나

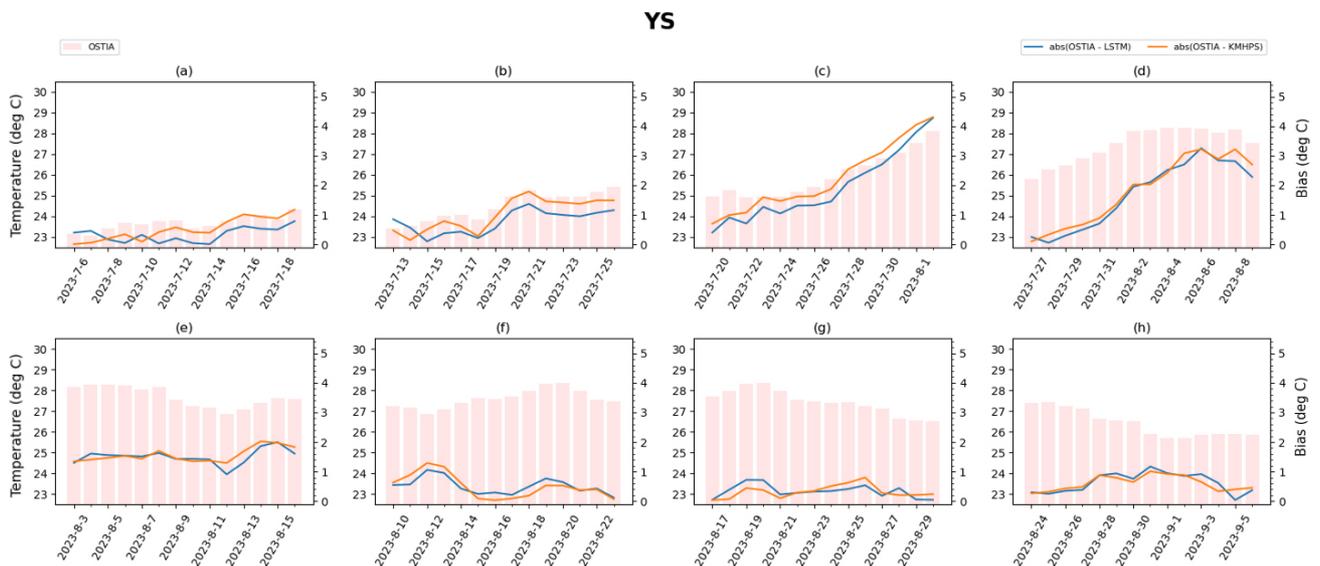


Fig. 9. Observed sea surface temperature (SST) from OSTIA is depicted in light red bar, and while biases of predicted SST by KMHPS and Machine Learning Model are shown in orange and blue, respectively, in the YS region. The light red bar corresponds to the left y-axis, while the blue and orange lines correspond to the right y-axis.

타났다. 특히 7월 6일 - 7월 19일(Fig. 9(a)) 예측 기간부터 수온이 증가하여 고수온이 유지되는 기간인 7월 27일 - 8월 9일 (Fig. 9(d))까지 LSTM 모델은 KMHPs에 의한 수온의 과소모의를 최대 0.5°C까지 개선하는 것으로 나타났다. 다만, 수온이 감소하는 기간인 8월 3일부터 9월 6일 사이의 예측 기간에는 개선 정도가 미미했다.

SS 영역에서의 수온 변화는 7월 6일부터 지속적으로 상승하여 7월 30일부터 8월 27일까지 다른 영역에 비해 오랫동안 28°C에서 29°C 사이의 수온을 유지하는 것이 특징적이다. 8월 27일 이후에도 수온이 다소 감소하지만, 9월 6일까지 27°C 이상의 수온을 유지하였다(Fig. 10).

SS 영역에서 KMHPs에 의한 수온의 예측 오차는 다른 영역보다 큰 것으로 나타났다. 대부분의 기간에서 초기에는 오차가 대체로 1°C보다 작았지만, 예측 기간이 증가할수록 오차가 커졌다. 특히, 수온이 증가하는 기간인 7월 20일 - 8월 2일(Fig. 10(c)) 예측 기간에는 예측 14일 후 수온의 예측 오차가 5°C 이상으로 크게 증가했다. 또한, 7월 27일 - 8월 9일(Fig. 10(d))과 8월 3일 - 8월 16일(Fig. 10(e)) 예측 기간에는 예측 14일 후 수온의 예측 오차가 약 3°C까지 증가했다. 나머지 기간에서도 약 2°C까지 오차가 증가하는 것으로 나타났다.

SS 영역에서 LSTM 모델은 대부분의 예측 시기에서 KMHPs에 의한 수온의 과소모의를 다른 영역에 비해 가장 크게 개선하는 것으로 나타났다. 다만, 7월 13일 - 7월 26일(Fig. 10(b)) 예측 기간에는 LSTM 모델에 의한 예측 성능이 오히려 악화되었다. 또한, 수온이 감소하는 기간인 8월 24일 - 9월 6일(Fig. 10(h)) 예측 기간에는 초기부터 예측 7일까지는 예측 성능의 개선이 미미한 것으로 나타났다. 그 외의 기간에는 대부분의 예측일에 1°C 이상 수온의 예측 오차를 감소시킨 것으로 나타났다.

Fig. 11은 ES, YS 및 SS 영역에 대한 KMHPs에 의한 예측 수온과 LSTM 모델에 의한 예측 수온의 평균 편향(bias) 및 RMSE (Root-Mean-Square-Error)를 보여준다. 평균 편향과 RMSE는 예측 초기로부터 14일까지 일별로 나타났다.

ES 영역에서는 KMHPs에 의한 예측 수온의 평균 편향과 RMSE가 예측 1일부터 9일까지 크게 증가하지 않고 약 1°C 정도로 유지되었다. 그러나, 예측 10일 이후에는 두 값이 모두 증가하여 평균 편향은 1.9°C, RMSE는 2.3°C까지 증가하였다. LSTM 모델은 KMHPs의 예측 오차가 크지 않은 예측 1일부터 9일까지는 예측 성능을 개선하지 못하였으나, 오차가 증가하

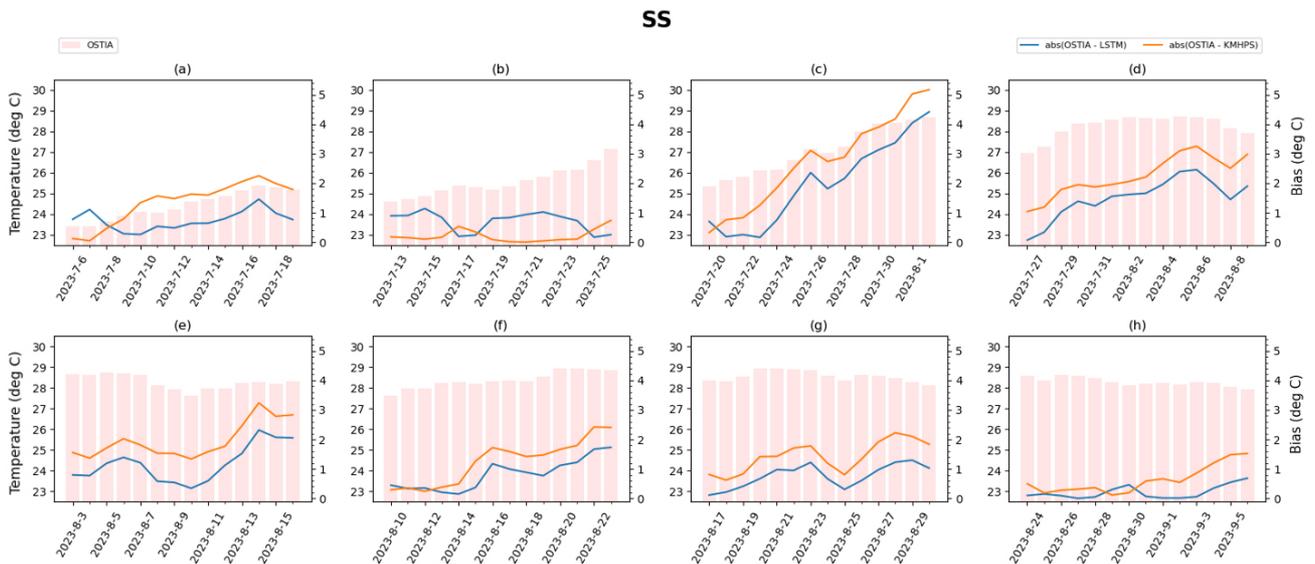


Fig. 10. Observed sea surface temperature (SST) from OSTIA is depicted in light red bar, and while biases of predicted SST by KMHPs and Machine Learning Model are shown in orange and blue, respectively, in the SS region. The light red bar corresponds to the left y-axis, while the blue and orange lines correspond to the right y-axis.

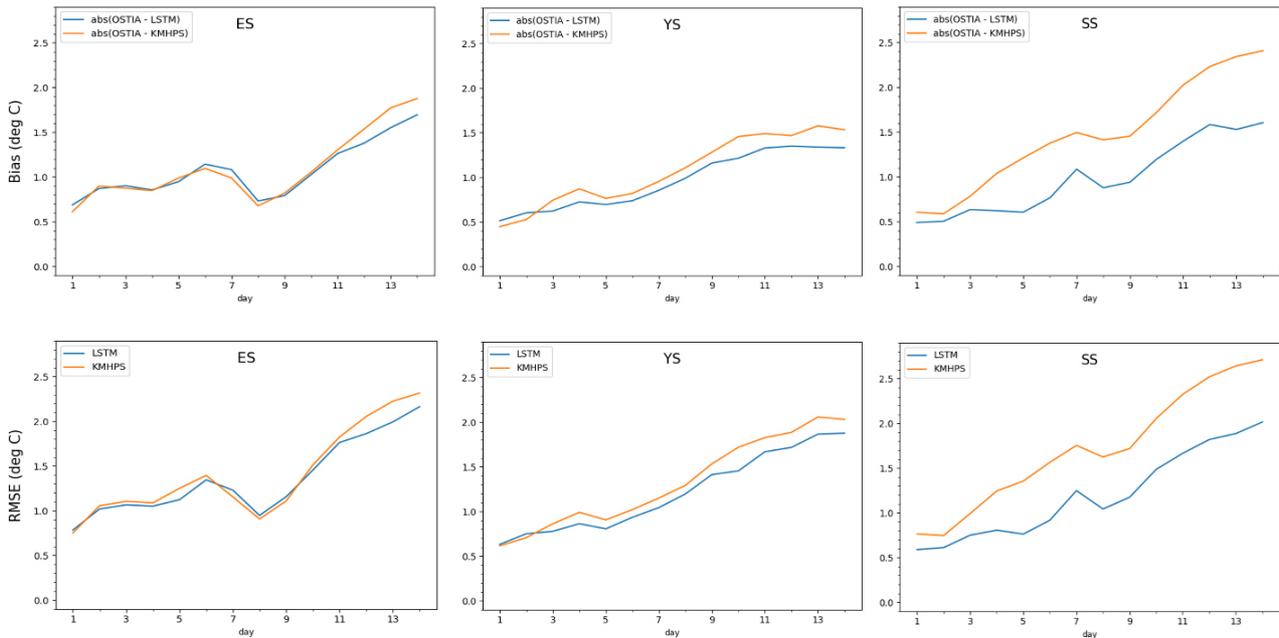


Fig. 11. Daily mean biases and RMSEs of predicted SST by KMHPs and the Machine Learning Model in the ES, YS, and SS regions. The orange line represents KMHPs, while the blue line represents the Machine Learning Model.

는 예측 10일 후부터는 예측 성능을 개선하였다. LSTM 모델에 의해 예측 14일 후의 평균 편향과 RMSE는 각각 1.7°C와 2.2°C로 감소하였다.

YS 영역에서는 KMHPs에 의한 예측 수온의 평균 편향과 RMSE가 예측 1일부터 14일까지 증가하는 것을 확인할 수 있다. 예측 14일 후에 평균 편향과 RMSE는 각각 1.5°C와 2.0°C였다. YS 영역에서 LSTM 모델은 초기부터 예측 2일까지의 개선이 미미했지만, 예측 3일 후부터 14일 후까지 전반적으로 수온 예측 성능을 향상시켰다. 예측 14일 후에 LSTM 모델에 의한 예측 수온의 평균 편향과 RMSE는 각각 1.3°C와 1.9°C로 개선되었다.

SS 영역에서는 KMHPs에 의한 예측 수온의 오차가 다른 영역에 비해 매우 높게 나타났다. KMHPs에 의한 예측 수온의 평균 편향과 RMSE는 각각 2.4°C와 2.7°C였다. 오차가 가장 큰 SS 영역에서는 LSTM 모델에 의한 수온 예측 성능의 향상이 두드러진다. 예측 1일부터 14일까지 전반적으로 예측 성능을 향상시키고, 예측 14일 후의 예측 수온의 평균 편향과 RMSE는 각각 1.6°C와 2.0°C로 나타났다.

5. 요약 및 제언

지구온난화로 인한 해양의 고온 현상은 해양 생태계에 대한 지속적인 위협으로 작용할 것으로 예상되며, 이는 식량 자원의 감소와 해양 탄소 흡수력의 저하 등을 통해 인류에게 직접적인 영향을 미칠 것으로 우려된다(Oliver *et al.*, 2018). 이에 따라, 고수온 예측은 한반도 주변 해역의 해양 환경 모니터링 및 관리에 중요한 역할을 할 것으로 기대된다(Han *et al.*, 2023). 본 연구에서는 KMHPs의 성긴 수직격자로 인한 고수온 예측의 과소모의를 개선하기 위해 LSTM 모델을 개발하였다. 또한 2023년도에 대해 수행된 KMHPs의 한반도 고수온 예측 성능을 진단하고, LSTM 모델을 적용하여 ES, YS 및 SS 영역에서 고수온 예측의 과소모의가 개선되었음을 확인하였다.

ES 영역에서는 KMHPS에 의한 고수온 예측의 오차는 다른 지역에 비해 가장 작았다. 8번의 예측에 대해 평균 편향과 RMSE가 상대적으로 작았던 예측 초기부터 예측 9일까지 LSTM에 의한 예측 성능 개선은 미비했으나, 예측 10일 이후로 KMHPS에 의한 예측 수온의 평균 편향과 RMSE가 커짐에 따라 LSTM 모델에 의해 예측 성능이 개선되었다. YS 영역에서 LSTM 모델은 수온이 증가하는 시기에 KMHPS에 의한 예측 수온의 오차가 클 때 예측 성능 개선 효과가 뚜렷한 것으로 나타났다. YS 영역에서 평균 편향과 RMSE는, ES 영역에서와는 달리, 비교적 예측 초반인 예측 3일 이후부터 예측 14일까지 LSTM 모델에 의한 예측 성능 개선 효과가 나타났다. SS 영역에서 KMHPS의 고수온 예측 오차는 다른 영역에 비해 높게 나타났다. 그만큼 LSTM 모델에 의한 예측 성능 개선 효과도 크게 나타났다.

본 연구에서 개발된 LSTM 모델이 세 영역 모두에서 수온이 상승하는 시기에 수온 예측 성능을 크게 개선하는 것으로 나타났으며, 수온 상승이 시작되기 전이나 하강하는 시기에는 예측 성능의 개선 효과가 미미했다. 수온이 상승하는 시기에는 성층이 강화되고, 이에 따라 KMHPS가 성긴 수직격자체제로 인해 고수온을 과소모의할 가능성이 큰 시기라고 할 수 있다. 따라서, LSTM 모델이 본 연구의 목적에 맞게 성층이 강화되는 환경에서 KMHPS의 과소모의를 개선하는데 성공적으로 작동하는 것으로 판단하였다.

이 논문에서는 태양 단파 복사에 대해 자세히 다루지 않았지만, 수온 데이터만 사용하여 학습하였을 때보다 태양 단파 복사를 추가하여 학습한 모델이 고수온 예측 성능을 개선하는데 보다 효과적인 것으로 나타났다. 본 연구에서는 고수온 예측 모형의 성긴 수직해상도에 기인한 고수온의 과소모의를 개선하기 위한 목적으로 LSTM 모델을 개발하였기 때문에, 역학 모형을 대체하는 고수온 예측 인공지능 모형의 개발과는 거리가 멀다. 또한, 본 연구에서는 고수온 현상에 미치는 다양한 요소들을 고려하지 않았다. 예를 들어, 운량, 풍속, 혼합층의 두께 등을 고려하지 않았다. 향후 고수온을 보다 정확하게 예측하기 위해서는 고수온에 영향을 미치는 다양한 요소들을 고려해야 할 필요가 있다(Huang *et al.*, 2021; Lee *et al.*, 2022; Tan *et al.*, 2022). 또한 최근 인공지능 기계학습을 적용하여 역학 모형을 대체하려는 노력들이 시도되고 있다(Hu *et al.*, 2021; Jeong *et al.*, 2023). 향후 인공지능 고수온 예측 시스템 모델에 대한 연구 및 개발이 필요할 것이다.

사 사

이 논문은 해양수산과학기술진흥원의 지원으로 수행된 “아북극-서태평양 기인 한반도 주변 고수온 현상 규명 및 예측시스템 구축(20190344)” 연구과제의 결과임.

참고문헌(References)

- Ahn, J., I. Park, S. Lyu and Y. Hur, 2012. Evaluation of Optimal Grid Resolution for Hydrodynamic Proper Simulation. *Journal of Korean Society for Geospatial Information Science*, **20**(1): 109-116. DOI: 10.7319/kogsis.2012.20.1.109.
- EU Copernicus Marine Environment Monitoring Service (CMEMS), 2023. PRODUCT USER MANUAL For the Global Ocean Physical Multi Year product GLOBAL_MULTIYEAR_PHY_001_030, EU Copernicus Marine Service – Public, no.1.5, DOI: 10.48670/moi-00021.
- Han, I. and J. Lee, 2020. Change the Annual Amplitude of Sea Surface Temperature due to Climate Change in a Recent Decade around the Korean Peninsula. *Journal of the Korean Society of Marine Environment & Safety*, **26**(3): 233-241. DOI: 10.7837/kosomes.2020.26.3.233.
- Han, I., J. Lee, C. Kim and J. Yaong, 2023. Impacts, projections and assessments related to climate change in ocean and fisheries. *Journal of Climate Change Research*, **14**(6-2): 965-972: DOI: 10.15531/KSCCR.2023.14.6.965.

- Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee and J.-N. Thépaut, 2023. ERA5 hourly data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.adbb2d47.
- Hochreiter, S. and J. Schmidhuber, 1997. Long short-term memory. *Neural computation*, **9**(8): 1735-1780. DOI: 10.1162/neco.1997.9.8.1735.
- Hu, H., A. Westhuysen, P. Chu and A. Fujisaki-Manome, 2021. Predicting Lake Erie wave heights and periods using XGBoost and LSTM. *Ocean Modelling*, **164**. DOI: 10.1016/j.ocemod.2021.101832.
- Huang, B., Z. Wang, X. Yin, A. Arguez, G. Graham, C. Liu, T. Smith and H. Zhang, 2021. Prolonged Marine Heatwaves in the Arctic: 1982-2020. *Geophysical Research Letters*, **48**(24). DOI: 10.1029/2021GL095590.
- Jeong, H., S. Lee, Y. Lee and Y. Seo, 2023. Development of bias correction model for downward short wave radiation in South Korea using high-resolution numerical model prediction information. *Journal of Korean Institute of Intelligent Systems*, **33**(2): 162-172. DOI: 10.5391/JKIS.2023.33.2.162.
- Kim, M., H. Choi and H. Yang, 2021. Method for Improving Sea Surface Temperature and High Water Temperature Prediction Performance using LSTM. *Korean Institute of Information Scientists and Engineers*, 744-746.
- Kingma, D. and J. Ba, 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, DOI: 10.48550/arXiv.1412.6980.
- Korea Institute of Ocean Science & Technology (KIOST). 2024. Investigation and prediction system development of marine heatwave around the Korean Peninsula originated from the subarctic and western Pacific, 696 pp.
- Korea Meteorological Administration (KMA), 2018. Abnormal climate report 2018, Korea Meteorological Administration.
- Lee, K., I. Zo, J. Jee and Y. Choi, 2011. Temporal and Spatial Distributions of the Surface Solar Radiation by Spatial Resolutions on Korea Peninsula. *New & Renewable Energy*, **7**(1): 22-28. DOI: 10.7849/ksnre.2011.7.1.022.
- Lee, K., M. Kwon and H. Kang, 2022. Record-breaking High Temperature in July 2021 over East Sea and Possible Mechanism. *Atmosphere*, **32**(1): 17-25. DOI: 10.14191/Atmos.2022.32.1.017.
- Lee, S., M. Park, M. Kwon, Y. Kim and Y. Park, 2020. Two major modes of East Asian marine heatwaves. *Environmental Research Letters*, **15**(7). DOI: 10.1088/1748-9326/ab8527.
- Lee, S., M. Park, M. Kwon, Y. Park, Y. Kim and N. Choi, 2023. Rapidly changing East Asian marine heatwaves under a warming climate. *Journal of Geophysical Research - Oceans*, **128**(6). DOI: 10.1029/2023JC019761.
- Oh, H., G. Kim, J. Chu, K. Lee and J. Jeong, 2023. The record-breaking 2022 long-lasting marine heatwaves in the East China Sea. *Environmental Research Letters*, **18**(6). DOI: 10.1088/1748-9326/acd267.
- Oliver, E., M. Donat, M. Burrows, P. Moore, D. Smale, L. Alexander, J. Benthuisen, M. Feng, A. Gupta, A. Hobday, N. Holbrook, S. Perkins-Kirkpatrick, H. Scannell, S. Straub and T. Wernberg, 2018. Longer and more frequent marine heatwaves over the past century. *Nature Communications*, **9**(1): 1-12.
- Siami-Namin, S., N. Tavakoli and A.S. Namin, 2018. A Comparison of ARIMA and LSTM in Forecasting Time Series, 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 1394-1401, DOI: 10.1109/ICMLA.2018.00227.
- Tan, H., R. Cai and R. Wu, 2022. Summer marine heatwaves in the South China Sea: Trend, variability and possible causes. *Advances in Climate Change Research*, **13**(3): 323-332. DOI: 10.1016/j.accre.2022.04.003.
- Zhang, Q., H. Wang, J. Dong, G. Zhong and X. Sun, 2017. Prediction of sea surface temperature using long short-term memory. *IEEE geoscience and remote sensing letters*, **14**(10): 1745-1749. DOI: 10.1109/LGRS.2017.2733548.