

Article

다중 위성자료와 기계학습을 이용한 고해상도 표층 염분 추정

성태준¹⁾ · 심성문¹⁾ · 장은나²⁾ · 임정호³⁾†

Estimation of High Resolution Sea Surface Salinity Using Multi Satellite Data and Machine Learning

Taejun Sung¹⁾ · Seongmun Sim¹⁾ · Eunna Jang²⁾ · Jungho Im³⁾†

Abstract: Ocean salinity affects ocean circulation on a global scale and low salinity water around coastal areas often has an impact on aquaculture and fisheries. Microwave satellite sensors (e.g., Soil Moisture Active Passive [SMAP]) have provided sea surface salinity (SSS) based on the dielectric characteristics of water associated with SSS and sea surface temperature (SST). In this study, a Light Gradient Boosting Machine (LGBM)-based model for generating high resolution SSS from Geostationary Ocean Color Imager (GOCI) data was proposed, having machine learning-based improved SMAP SSS by Jang *et al.* (2022) as reference data (SMAP SSS (Jang)). Three schemes with different input variables were tested, and scheme 3 with all variables including Multi-scale Ultra-high Resolution SST yielded the best performance (coefficient of determination = 0.60, root mean square error = 0.91 psu). The proposed LGBM-based GOCI SSS had a similar spatiotemporal pattern with SMAP SSS (Jang), with much higher spatial resolution even in coastal areas, where SMAP SSS (Jang) was not available. In addition, when tested for the great flood occurred in Southern China in August 2020, GOCI SSS well simulated the spatial and temporal change of Changjiang Diluted Water. This research provided a potential that optical satellite data can be used to generate high resolution SSS associated with the improved microwave-based SSS especially in coastal areas.

Key Words: Sea surface salinity, SMAP, GOCI, Changjiang diluted water, Machine learning

요약: 해양 염분은 전 지구 규모에서 해수 순환에 영향을 미칠 뿐만 아니라, 연·근해 지역 저염분수가 어족자원 및 수산업에 피해를 줄 수 있는 등 해양 식생환경의 변화를 줄 수 있다. 해수의 표면 특성인 sea surface salinity

Received October 4, 2022; Revised October 18, 2022; Accepted October 22, 2022; Published online October 31, 2022

¹⁾ 울산과학기술원 도시환경공학과 석·박사과정생(Combined MS/PhD Student, Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea)

²⁾ 한국해양과학기술원 해양위성센터 선임연구원(Senior Researcher, Korea Ocean Satellite Center, Korea Institute of Ocean Science and Technology, Busan, Republic of Korea)

³⁾ 울산과학기술원 도시환경공학과 교수(Professor, Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea)

† Corresponding Author: Jungho Im (ersgis@unist.ac.kr)

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

(SSS)에 따라 마이크로웨이브 영역의 방사율이 달라지며, 이를 통해 Soil Moisture Active Passive (SMAP) 등 위성 센서를 활용한 SSS 산출물이 제공되고 있다. 하지만 마이크로파 위성 센서 기반의 SSS 산출물은 낮은 시공간 해상도로 자료를 생산하며, 연안지역과 고위도 지역에서 정확도가 낮다. 이러한 이유로 연·근해 지역 SSS의 상세한 시공간적 변화를 관측하기에는 적합하지 않다. 본 연구에서는 Jang *et al.* (2022)에서 제시한 기계학습 기반의 개선된 SMAP SSS (SMAP SSS (Jang))를 참조자료로 활용하여, 정지궤도해색센서(Geostationary Ocean Color Imager, GOCI) 영상으로부터 고해상도 SSS를 추정하는 Light Gradient Boosting Machine (LGBM) 기반의 모델을 개발하였다. 3가지 입력변수 조합을 테스트하였고, Multi-scale Ultra-high Resolution Sea Surface Temperature (SST) 자료가 추가된 scheme 3가 가장 높은 정확도를 보였다($R^2 = 0.60$, RMSE = 0.91 psu). 이를 바탕으로 본 연구영역에서 SST가 SSS 모의에 효과적인 환경변수로 작용함을 보였다. 본 연구에서 제시한 LGBM 기반의 GOCI SSS는 SMAP SSS (Jang)와 비슷한 시공간적 패턴을 보였지만, 더 높은 공간해상도를 바탕으로 SSS의 보다 상세한 공간적 분포와 더불어 SMAP SSS (Jang)에서 산출하지 않는 연안 지역의 정보까지 모의하였다. 또한, 중국 남방지역에 대홍수가 발생하였던 2020년 8월을 대상으로 양자강 유출수(Changjiang Diluted Water)의 거동을 분석한 결과, GOCI SSS는 한국 해양수산연구원의 보도자료와 비교하여 일관성 있는 시공간적 변화를 보였다. 본 연구의 결과로 연안 지역의 저염수 뿐 아니라, 원해 지역에서 광학위성 신호를 활용한 고해상도 SSS 산출의 가능성을 제시하였다.

1. 서론

해양 염분은 해양 온도와 더불어 해양환경 및 해양물리 변화 관측에 필수인 인자이다. 강물의 유입, 강수, 증·발산 등에 의해 변화되는 해양 염분은 작은 규모의 해양 식생환경의 변화 뿐 아니라 큰 규모의 해수 순환에도 영향을 미칠 수 있다(NOAA Climate.gov, 2014; Jang *et al.*, 2022; Hoffman *et al.*, 2022). 따라서, 해양 염분의 지속적이고 광역적인 관측이 필요하다. 해수는 표면 특성인 sea surface temperature (SST), sea surface salinity (SSS)에 따라 마이크로웨이브 영역의 방사율이 달라지는데, 이를 통해 인공위성 센서를 활용한 SSS 관측이 가능하다(Klein and Swift, 1977). 현재, L-밴드(1.4 GHz) 마이크로파 센서를 가진 극 궤도 위성들(Soil Moisture and Ocean Salinity; Soil Moisture Active Passive, SMAP)이 현업에서 SSS 산출물을 제공하고 있다. 하지만, 해수 표면에서 위성까지 도달하는 마이크로파 방사량은 매우 적어, 해당 위성들은 주로 저궤도 비행을 하며 0.25° - 1° 의 공간해상도와 3일 이상의 시간해상도로 자료를 생산한다. 마이크로파 기반의 SSS 산출은 SST가 낮아질수록 SSS에 대한 마이크로파의 반응성이 약해져, 고위도 지역에 대한 산출 정확도가 떨어진다는(Klein and Swift, 1977; Barale *et al.*, 2010; Qin *et al.*, 2020). 또한, 육상 근처의 radio frequency interference (RFI)와 land contamination (LC)에 취약하여, 연안 지역에 대한 SSS 산출을 하지

않거나 낮은 품질 플래그와 함께 제공을 하고 있다(Olmedo *et al.*, 2019; Jang *et al.*, 2021; Jang *et al.*, 2022). 이러한 이유로 마이크로파 위성 기반의 SSS 모니터링은 글로벌 규모에서의 염분 거동과 해수 순환 관측에는 효과적일 수 있으나, 연·근해 지역 SSS의 상세한 시공간적 변화를 관측하기에는 적합하지 않다(Chen and Hu, 2017; Olmedo *et al.*, 2018). 하지만 강하구에 근접한 연안 지역 SSS의 경우 육지에서의 담수 유출에 영향을 받으며, 다량의 담수 유출이 있을 시 발생하는 저염분수가 어족자원 및 수산업에 피해를 끼칠 수 있어, 지속적인 모니터링이 필요하다(Moon *et al.*, 2019).

이러한 마이크로웨이브 위성센서의 한계를 극복하고자 해색위성을 활용한 SSS 추정 연구들이 수행되어 왔다. 해색위성을 이용한 SSS 추정은 청색 파장영역의 흡수율이 높은 유색용존유기물(colored dissolved organic matter, CDOM)이 고염분수로 갈수록 적게 분포하는 경향에 기반하여 발전되어왔다(Geiger *et al.*, 2013; Kim *et al.*, 2020). CDOM은 주로 육지에서 발생하여 강을 통해 해수로 유입되기 때문에, 강물에서 유출된 담수는 염분이 낮을 뿐 아니라 많은 양의 CDOM을 포함하고 있다(Kim *et al.*, 2020; Kim *et al.*, 2021). 하지만 해색 위성 신호는 CDOM 뿐만 아니라 용존부유물질(suspended particulate matter, SPM), 클로로필(chlorophyll) 등의 영향도 받을 수 있다. 이러한 인자들의 복합적인 상관성 모의를 위해, 단순 선형회귀식이 아닌 비선형 모델, 특히 인공지능 기

법들을 활용한 연구가 많이 시도되어 오고 있다(Chen and Hu, 2017; Kim *et al.*, 2020; Kim *et al.*, 2021). 인공지능 기법을 활용한 모델링에는 해석위성영상과 매칭이 되는 높은 신뢰도의 검증자료가 필요하며, 부이나 정선관측을 통해 구축된 현장관측자료나 마이크로웨이브 위성자료가 주로 쓰였다(Chen and Hu, 2017; Kim *et al.*, 2020; Kim *et al.*, 2021).

Chen and Hu (2017)에서는 멕시코만 북부해역에서 취득한 약 20년간의 부이 및 선박관측 자료를 훈련 및 검증자료로 활용하여, Sea-Viewing Wide Field-of-View Sensor와 Moderate Resolution Imaging Spectroradiometer (MODIS)의 해석영상 및 SST 자료를 입력자료로 하는 Multilayer Perceptron Neural Network 모델을 개발하였다. 해당 지역은 미시시피강 담수 유출이 활발한 지역으로, 다양한 염분 값(3–37 psu)의 현장관측자료를 통해 모델 학습이 수행되었다. 학습된 모델은 높은 성능을 보였으나, 시공간 검증 및 일반화에 대한 어려움이 있었다. 이에 마이크로웨이브 센서에서 산출되는 격자형태 SSS 산출물을 검증자료로 활용하는 접근 방법도 시도되었다(Kim *et al.*, 2020; Kim *et al.*, 2021). 하지만 현업 산출물의 경우 현장관측자료에 비해 불확실성이 커 8일 합성 영상을 검증자료로 활용했고, 이로 인해 SSS 추정모델의 시간당위성 확보가 어려웠다(Kim *et al.*, 2020; Kim *et al.*, 2021). Jang *et al.* (2022)은 기계학습 모델을 기반으로 SMAP 위성의 SSS 산출물과 해양수치모델인 HYbrid Coordinate Ocean Model (HYCOM)의 SSS 자료를 상호 보완적으로 융합하여, 개별 자료가 가지는 불확실성의 보정을 시도하였다. 그 결과 글로벌 해역에서 안정적으로 높은 정확도를 가진 swath 형태의 SSS를 제공할 수 있게 되었다.

이에 본 연구에서는 Jang *et al.* (2022)에서 제시한 기계학습 기반의 개선된 SMAP SSS를 참조자료로 활용하여, 정지궤도해색센서인 천리안해양관측위성(Geostationary Ocean Color Imager, GOCI) 영상으로부터 고해상도 SSS를 추정하는 Light Gradient Boosting Machine (LGBM) 기반의 모델을 개발하였다. 개선된 SMAP SSS와 GOCI 위성영상을 활용하여, 높은 정확도의 검증자료를 이용한 모델링을 수행하였으며, 고해상도의 실시간 SSS를 추정하였다. 이후 개발한 고해상도 SSS를 활용하여 동아시아 해역에 대한 SSS의 공간적 분포 및 계절적 특성 변

화와, 양자강 유출수(Changjiang Diluted Water, CDW)의 시공간적 거동을 분석하였다. 또한, SHapley Additive exPlanations (SHAP) 기법을 활용하여, 각 입력변수와 SSS 사이의 상관관계 및 변수 중요도를 분석하였다.

2. 연구지역 및 연구자료

1) 연구지역 및 연구기간

연구지역은 동중국해(East China Sea), 황해(Yellow Sea), 동해(East Sea), 그리고 북서태평양(Northwestern Pacific) 일부를 포함하는 영역(위도: 30°N–42°N, 경도: 120°E–144°E)으로 설정하였다. 해당지역은 CDW의 영향을 직접적으로 받는 저염분수 지역과, 동해와 북서태평양의 고염분수 지역을 모두 포함하는 영역으로써, 원·근해 모두에서의 광학위성자료를 이용한 염분 추정 가능성을 모의하기에 적합하다. 2018년과 2019년을 대상으로 모델 구축 및 검증을 진행하였으며, 중국 남방 지역에 대홍수가 발생하였던 2020년 여름을 대상으로 SSS의 시공간적 거동을 분석하였다.

2) 현장자료

Jang *et al.* (2022)에서 제시한 기계학습 기반의 개선된 SMAP SSS (이하 SMAP SSS (Jang))와 본 연구에서 산출한 SSS (이하 GOCI SSS)의 검증을 위해 Coriolis Data Center (<http://www.coriolis.eu.org/>)에서 제공되는 thermosalinograph (TSG) 자료를 사용하였다. TSG는 Global Ocean Surface Underway Data 프로그램의 일환으로 수집되는 선박 관측 기반의 현장자료로써, Coriolis에서 제공되는 여러 현장자료 플랫폼 중 근해 지역에서 비교적 많은 자료를 제공한다. 본 연구지역 내에서는 동해 원해와 북서태평양 지역에 관한 자료를 사용하였다(Fig. 1). Quality control flag를 활용하여 신뢰도가 낮은 자료를 제거하였으며, flag 1과 2 (good, probably good), 관측 수심 5 m 이내의 자료만을 활용하였다(Kolodziejczyk *et al.*, 2021; Jang *et al.*, 2022). Coriolis에서 제공되는 자료 중 대표적인 현장관측 플랫폼인 Argo 자료는 연구지역 내 샘플 수가 부족하여 사용하지 않았다.

TSG 자료가 동해 원해와 북서태평양의 고염분수 지역에서만 제공되므로, 한반도 근해와 동중국해의 상대

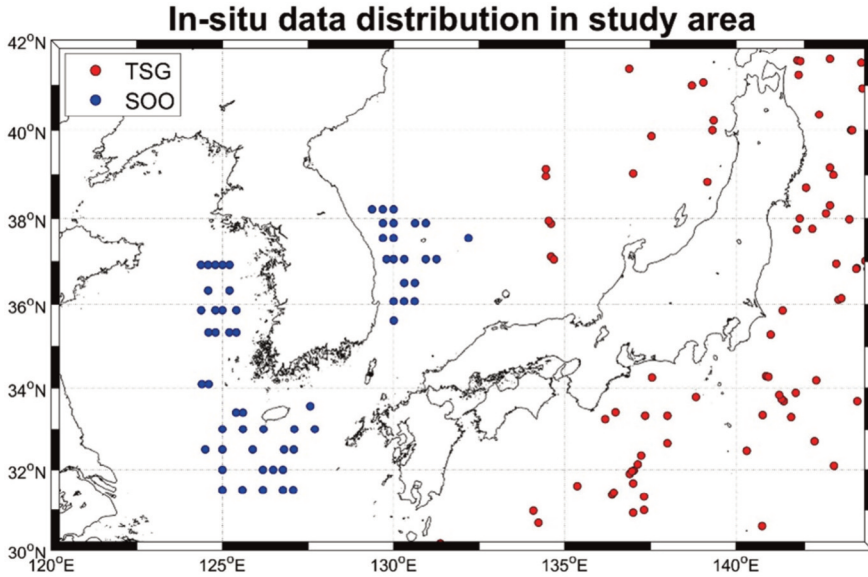


Fig. 1. Spatial distribution of two types of in-situ data (i.e., thermosalinograph and serial oceanographic observation) measured in the study area from 2018 to 2019.

적으로 저염분수인 지역에 관한 검증을 위해 정선해양 관측(serial oceanographic observation, SOO) 자료를 사용하였다. 정선해양관측 자료는 국립수산물과학원(Korean National Institute of Fisheries Science; https://www.nifs.go.kr/page?id=en_index)에서 제공되는 정기은행선박 자료로서, 동해, 서해, 남해, 동중국해역의 207개 정점에 관하여 연 6회(2, 4, 6, 8, 10, 12월) 조사 및 제공되고 있다. 본 연구에서는 flag 1과 2 (good, not evaluated), 관측 수심 0 m의 자료를 활용하였다(Kim *et al.*, 2021).

3) 위성기반자료

(1) 참조자료

SMAP SSS (Jang) 자료를 기계학습 모델의 훈련을 위한 종속변수로 사용하였다. 해당 자료는 기계학습 모델을 기반으로 SMAP과 HYCOM 자료의 상호보완적 융합을 통해 생성된 글로벌 염분 자료이다. SMAP SSS (Jang)은 기존의 SMAP과 HYCOM 자료의 단점을 효과적으로 보완하여, 전세계 해역에서 개별 자료보다 높은 성능을 검증하였다(Jang *et al.*, 2022). 해당 자료의 생성을 위해 National Aeronautics and Space Administration (NASA)와 Jet Propulsion Laboratory (JPL)에서 제공하는 SMAP Level 2B v5 자료와 HYCOM 해양 일반 순환 모델의 SSS 자료를 주요 입력변수로 사용하였다. 또한, National

Oceanic and Atmospheric Administration (NOAA)에서 제공되는 Optimal Interpolation Sea Surface Temperatures, NOAA National Centers for Environmental Prediction Global Forecast System의 대기 순환 모델에서 산출되는 풍속, NOAA WaveWatch III의 significant wave height, 그리고 Global Precipitation Measurement (GPM) Level 3 v.6 Integrated Multi-satellite Retrievals for GPM (IMERG)의 일일 강수(GPM_3IMERGDL) 자료를 보조 입력변수로 사용하였다. SMAP Level 2B 자료는 25 km의 공간해상도와 글로벌 커버리지 2-3일의 시간해상도로 제공되며, 최종 SMAP SSS (Jang) 산출물 또한 SMAP Level 2B 자료와 같은 시공간 해상도로 산출하였다.

검증자료로 사용된 SMAP SSS (Jang)는 전 지구 영역에 대한 보정 및 검증이 수행되었지만(Jang *et al.*, 2022), 한반도 근해역을 대상으로 기존 SMAP Level 2B SSS(이하 SMAP SSS (original)) 자료와 비교 분석된 적은 없다. 이에, 본 연구지역에서 SMAP SSS (Jang)와 SMAP SSS (original)의 정확도 평가 및 비교를 위해 현장관측자료인 TSG와 SOO를 이용한 검증을 수행하였다. 2018-2019년 기간동안 SMAP 위성과 매칭된 총 77개의 TSG 자료와 79개의 SOO 자료를 활용하여 산점도 평가를 진행하였다(Fig. 3, 4). 북서태평양 및 동해 원해역에 고루 분포하는 TSG의 경우 약 32에서 35 psu의 SSS 분포를 가

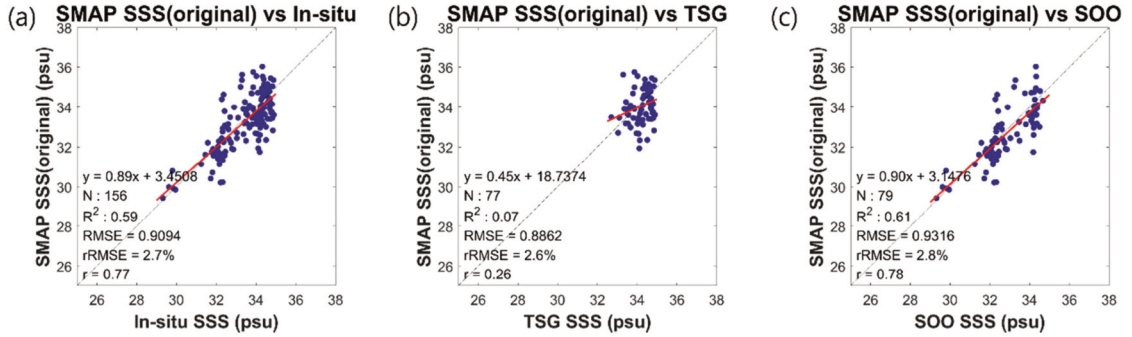


Fig. 2. Scatterplots between SMAP SSS (original) and (a) entire in-situ, (b) thermosalinograph, and (c) serial oceanographic observation data.

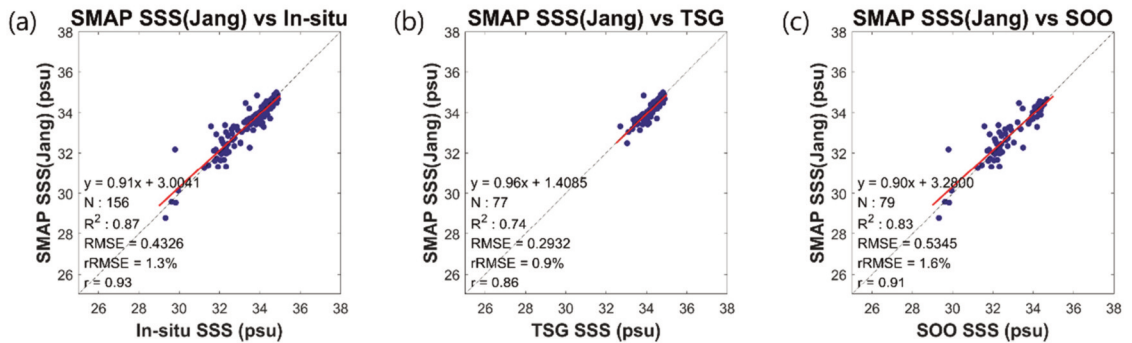


Fig. 3. Scatterplots between SMAP SSS (Jang) and (a) entire in-situ, (b) thermosalinograph, and (c) serial oceanographic observation data.

지고 있다. TSG와 SMAP SSS (Jang)와의 산점평가 결과, 0.74의 R^2 , 0.29 psu의 RMSE와 0.9%의 rRMSE를 확인하였다(Fig. 4(b)). 이는 SMAP SSS (original)의 중고염분 영역에 대한 정량평가에 비해 우수함을 확인할 수 있다 (Fig. 3(b)). 한반도 주변 정선관측으로 취득이 되는 SOO의 경우 약 28에서 34 psu의 SSS 분포를 가지고 있다. SOO와 SMAP SSS (Jang)와의 산점평가 결과, 0.83의 R^2 , 0.53 psu의 RMSE와 1.6%의 rRMSE를 보였다(Fig. 4(c)). 이를 통해, TSG에 비해 저염분 사례가 많은 SOO에서도 기존 현업 산출물에 비해 우수한 성능을 보임을 확인하였다(Fig. 3(c)).

(2) 입력자료

SSS가 가지는 간접적인 광학 신호의 모의를 위해 GOCI 원격반사도(remote sensing reflectance, R_{rs}) 자료를 기계학습 모델의 입력변수로 사용하였다. GOCI는 6개의 가시광역 밴드(412 nm, 443 nm, 490 nm, 555 nm, 660 nm, 680 nm)와 2개의 근적외선 밴드(745 nm, 865 nm)를

가지고 있다. 또한, 500 m의 공간해상도로 coordinated universal time (UTC) 기준 0시부터 7시까지 매시간마다 하루 8번의 영상을 취득한다. GOCI 자료의 경우 한국해양위성센터(Korea Ocean Satellite Center, KOSC; <http://kosc.kiost.ac.kr/>)에서 Level 1B 형태로 제공되며, 천리안 해양관측위성 자료처리시스템(GOCI Data Processing System, GDPS) v2.0을 활용하여 대기보정을 수행하여 Level 2A R_{rs} 자료로 가공이 가능하다. GDPS를 이용한 대기보정은 KOSC standard 알고리즘을 사용하였으며, Level 2A 변환 후 원격반사도 값에 음수가 있는 픽셀을 제거한 후 사용하였다. 본 연구에서는 GOCI의 6개 가시광역 밴드 반사도(412 nm, 443 nm, 490 nm, 555 nm, 660 nm, 680 nm; 이하 R_{rs} 412, R_{rs} 443, R_{rs} 490, R_{rs} 555, R_{rs} 660, R_{rs} 680)와 3개 반사도 비(412 nm와 555 nm, 443 nm와 555 nm, 490 nm와 555 nm; 이하 R_{rs} 412/555, R_{rs} 443/555, R_{rs} 490/555)를 사용하였다.

(3) 보조자료

Multi-scale Ultra-high Resolution Sea Surface Temperature (MURSST) 자료를 기계학습 모델의 보조 입력변수로 사용하였다. MURSST 자료의 경우 Multi-Resolution Variational Analysis 기법을 활용한 다중위성 합성자료로써, MODIS, Advanced Very High-Resolution Radiometer, Advanced Microwave Scanning Radiometer - Earth Observing System 등 적외선과 마이크로파 위성 기반의 SST 자료를 합성하여 생성된다(Chin *et al.*, 2017). MURSST 자료는 글로벌 1 km의 공간해상도로 하루 한번 제공되며, 본 연구에서는 NASA JPL의 Physical Oceanography Distributed Active Archive Center (<https://podaac.jpl.nasa.gov/dataset/MUR-JPL-L4-GLOB-v4.1>)에서 제공되는 4.1 버전(MUR-JPL-L4-GLOB-v4.1) 자료를 사용하였다.

3. 연구방법

본 연구에서는 기계학습 기법인 LGBM을 사용하여, GOCI 위성과 같은 시공간 해상도로 SSS를 산출하는 알고리즘을 개발하였다(Fig. 2). 모델의 훈련을 위한 독립변수로는 GOCI R_{rs} 와 MURSST가 사용되었고, 종속 변수로는 SMAP SSS (Jang)가 사용되었다. 훈련된 모델로 산출된 고해상도 GOCI SSS는, 훈련에 사용되지 않은 SMAP SSS (Jang)를 활용하여 검증한 후, 현장자료를

활용하여 추가로 검증하였다. 이후 2019년의 월 평균 GOCI SSS를 활용하여 동아시아 주요해역에서 SSS의 시공간적, 계절적 패턴을 분석하였다. 또한, 중국 남방 지역에 대홍수가 발생하였던 2020년 8월을 대상으로 CDW의 거동을 분석하였다. 마지막으로, SHAP 알고리즘을 활용하여 각 입력변수가 예측 값에 미치는 영향을 분석하였다.

1) 자료 전처리 및 준비

앞서 언급하였듯 SMAP SSS (Jang)가 SMAP Level 2B 자료와 같은 공간 해상도로 산출되기 때문에, 기계학습 모델의 입력 및 보조자료 또한 이에 맞추기 위해 선형보간법을 사용해 25 km로 업스케일(upscale)하였다. SMAP 위성이 연구지역을 촬영하는 주요 시간대는 UTC 기준 07-09시와 19-21시인 반면, GOCI의 촬영시간은 00-07시이다. 따라서, SSS의 일변화에 따른 오차를 최소화하기 위해, SMAP과 GOCI의 관측시간의 차가 3시간 이내인 경우에만 훈련에 사용하였다. MURSST의 경우 매일 UTC 09시의 자료가 제공되기 때문에, GOCI 관측시간 기준 전날의 자료를 매칭하여 활용하였다. SMAP SSS (Jang)의 검증에 사용된 현장자료의 경우, SMAP 위성을 기준으로 관측 위치의 차가 12.5 km 이내, 관측시간의 차가 3시간 이내인 경우만 매칭하여, 총 156개의 샘플을 활용하였다. 기계학습 모델로 생성된 GOCI SSS의 경우, GOCI 위성과 같은 시공간 해상도로 산출되기 때

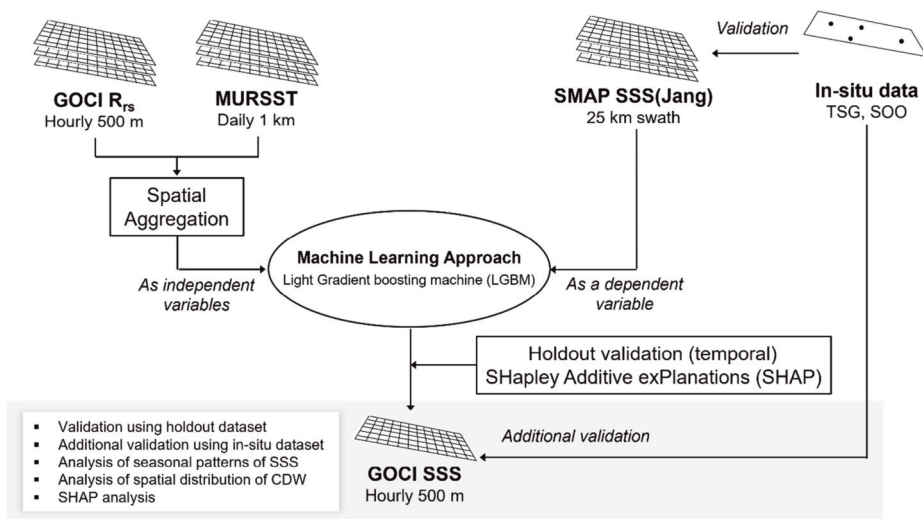


Fig. 4. The process flow diagram of this study.

문에 관측 위치의 차가 250 m 이내, 관측시간의 차가 30 분 이내인 현장자료 샘플 총 352개를 추가 검증에 사용하였다. 본 연구에서 사용된 매칭 기준은 기존의 위성 기반 SSS와 현장자료를 활용한 선행연구에서 주로 사용된 시간 매칭 기준보다 짧다(Fournier *et al.*, 2019; Tang *et al.*, 2018; Boutin *et al.*, 2018). 이는, 담수 유출의 영향을 많이 받는 본 연구지역의 특성상, 원해에 비해 염분의 일변화가 큰 것을 감안하여 위와 같이 설정하였다.

2) 기계학습 알고리즘

본 연구에서는 고해상도 SSS 추정을 위한 기계학습 알고리즘으로 트리기반(tree-based) 모델인 LGBM을 사용하였다. LGBM은 결정나무(Decision Tree, DT)의 부스팅(boosting) 앙상블 기법 중 하나인 Gradient Boosting Machine (GBM)의 개선된 모델이다. GBM은 이전에 생성된 DT에서 발생한 잔차를 예측하는 새로운 DT를 학습하는 과정을 반복하며, 이후 형성된 모든 DT를 결합하여 최종 결과를 산출한다(Friedman, 2001). LGBM은 손실(loss)이 가장 큰 노드(node)에서 분할을 계속하는 leaf-wise 방식을 사용하여, GBM보다 적은 계산시간과 비용을 요구한다(Ke *et al.*, 2017). LGBM은 빠른 계산 시간과 높은 정확도를 장점으로 많은 해양원격탐사 분야에 활용되었다(Jang *et al.*, 2022; Werther *et al.*, 2021; Taggio *et al.*, 2022). 본 연구에서는 Python 3.8의 LGBMRegressor 패키지가 사용되었으며, 격자 탐색(gridsearch) 기법을 통해 하이퍼파라미터(hyperparameter)를 최적화하였다. 부스팅 방식으로 Gradient-based One-Side Sampling 기법이 사용되었으며, 트리 수 500개, 부트스트랩(bootstrap) 훈련샘플의 샘플링 비율 0.2, 학습률(learning-rate) 0.01의 LGBM 모델을 구축하였다. 또한 본 연구에서는 각 변수의 유무에 따른 모델의 성능 변화를 분석하기 위해, 입력변수를 선행연구에서 주로 사용된 3개 조합으로 나누어 모델 훈련을 진행하였다(Chen and Hu, 2017; Sun *et al.*, 2019; Kim *et al.*, 2020). 3개 입력변수 조합은 다음과 같다: GOCI의 6개 R_{rs} (R_{rs} 412, R_{rs} 443, R_{rs} 490, R_{rs} 555, R_{rs} 660, R_{rs} 680)로 구성된 scheme 1, scheme 1에 사용된 변수와 3개 R_{rs} 비(R_{rs} 412/555, R_{rs} 443/555, R_{rs} 490/555)로 구성된 scheme 2, scheme 1과 2에 사용된 변수와 MURSST로 구성된 scheme 3의 총 3개 scheme으로 구성하였다.

3) 검증 방법

모델 검증에는 홀드아웃(holdout) 검증이 사용되었다. 홀드아웃 검증은 전체 훈련자료 중 일부를 훈련에 사용하고 나머지를 이용해 검증을 수행하는 방식이다. SMAP 위성의 촬영일자를 기준으로 5일마다 1일의 샘플들을 검증에 사용하여, 훈련과 검증자료 모두에 시간적인 다양성이 반영되도록 하였다. 모델의 평가 지표로는 산점도의 기울기(slope), 결정계수(coefficient of determination; R^2), 평균 제곱근 오차(root mean square error, RMSE), 평균 제곱근 상대오차(relative root mean squared error, rRMSE)를 사용하였으며, 각 지표의 계산 방법은 아래 식(1)~(4)과 같다. 아래 식에서 x 는 기계학습 모델로 추정된 GOCI SSS, y 는 참조자료로 사용된 SMAP SSS (Jang)와 현장자료 기반 SSS를 의미한다.

$$slope = \frac{\Delta y}{\Delta x} \quad (1)$$

$$R^2 = \frac{\sum_i [(x_i - \bar{x})(y_i - \bar{y})]}{[(\sum_i x_i - \bar{x})^2 (\sum_i y_i - \bar{y})^2]^{1/2}} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (3)$$

$$rRMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n \sum_{i=1}^n (y_i)^2}} \quad (4)$$

4) SHAP

SHAP은 복잡한 블랙박스 형태의 인공지능 모델의 해석을 위해 고안된 eXplainable Artificial Intelligence 방법 중 하나이다(Shapley, 2016). SHAP은 게임 이론을 바탕으로 게임에서 각 참가자의 기여분을 계산하는 방식으로 작동하며, 각 샘플과 특성별로 최종 예측에 대한 기여도를 제공한다(Lundberg and Lee, 2017). 이를 통해 각 입력변수의 상대적 중요도 뿐 아니라, 최종 결과에 미치는 영향의 방향성과 크기까지 확인할 수 있다(Park *et al.*, 2021). 따라서, 입력변수와 출력변수의 관계가 간접적이거나 비선형적인 경우에 대한 해석도 용이할 것으로 판단하여, 본 연구에서 각 입력변수의 변화에 따른 SSS 추정 값 변화 분석 및 모델 해석을 위해 사용하였다.

4. 연구결과 및 토의

1) LGBM 모델 정확도 검증

본 연구에서는 GOCI Level 2A R_s 와 R_{rs} 비, 그리고 MURSST를 입력자료로 하고, Jang *et al.* (2022)에서 제시한 SMAP SSS (Jang)를 타겟으로 하는 LGBM 모델을 제시하였다. 모델 평가를 위해 먼저 SMAP SSS (Jang)와의 산포분석을 수행하였고, 모델 학습에 사용되지 않은 날짜에 대해 총 10,497개의 샘플을 사용하였다. 입력변수의 조합을 기준으로 세 개의 scheme으로 구분하여 LGBM 모델을 생성 후 검증을 수행하였다. Scheme 3이 0.60의 R^2 , 0.91 psu의 RMSE로 가장 높은 성능을 보였고, 뒤이어 scheme 2가 0.30의 R^2 , 1.20 psu의 RMSE, 마지막으로 scheme 1이 0.29의 R^2 , 1.22 psu의 RMSE를 보였다 (Fig. 5). SMAP SSS (Jang)와 함께 현장관측자료와의 추가 검증도 수행을 하였는데, 추가 검증 역시 scheme 3이 0.45의 R^2 , 0.79 psu의 RMSE로 가장 높은 성능을 보였고, 뒤이어 scheme 1이 0.10의 R^2 , 1.18 psu의 RMSE, 마지막으로 scheme 2가 0.09의 R^2 , 1.19 psu의 RMSE를 보였다

(Fig. 6).

Scheme 1과 입력자료로 밴드비를 추가한 scheme 2의 경우 검증자료별 성능차이가 0.01의 R^2 와 0.01 psu의 RMSE로 매우 적는데 반해, MURSST를 추가한 scheme 3과 scheme 2의 차이는 0.3의 R^2 와 0.29 psu의 RMSE로 매우 크게 향상이 되었다. 이를 통해, 동아시아 해역 전체에서 광학위성신호만으로 SSS를 추정하는 것에 한계가 있으며, SST 자료가 모델을 효과적으로 보조하는 것을 알 수 있다. 특히 SSS가 30 psu보다 낮은 영역에 대해 scheme 1과 scheme 2에서 과대모의 경향이 있던 것이, MURSST를 입력변수로 추가함으로써 완화되었다. 이를 통해, 고염분수에 치중된 형태로 SSS를 모의하던 구조에서 저염분수 유형을 분리 후 모의하는 형태로 변화했다는 것을 유추할 수 있다. 광학위성신호로 SSS를 추정하는 방법은 CDOM과 SSS의 상관성이 각 저염분수 영역에 따라 상이하며, 외해로 갈수록 상관성이 사라지는 한계가 있다(Kowalczyk *et al.*, 2003). SST는 SSS와 직접적인 상관성은 없지만, 많은 선행연구에서 SSS 추정을 위한 보조변수로 사용되었으며, 본 연구영역에서도

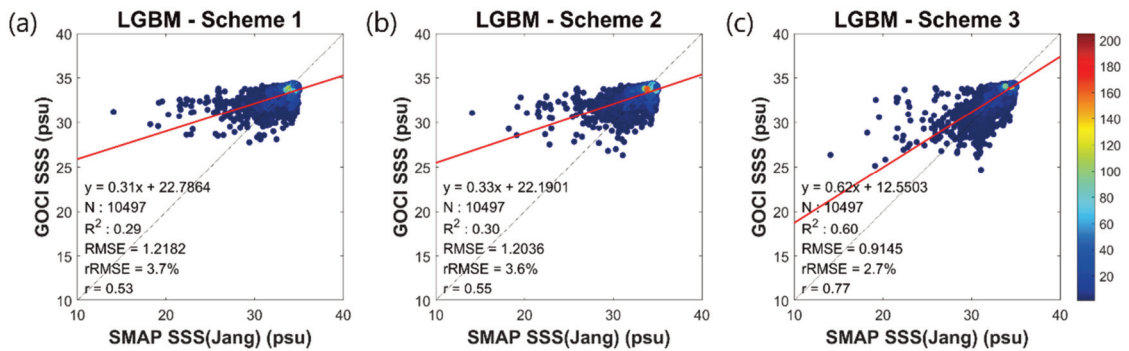


Fig. 5. Validation results of GOCI SSS using SMAP SSS (Jang): (a) scheme 1, (b) scheme 2, and (c) scheme 3.

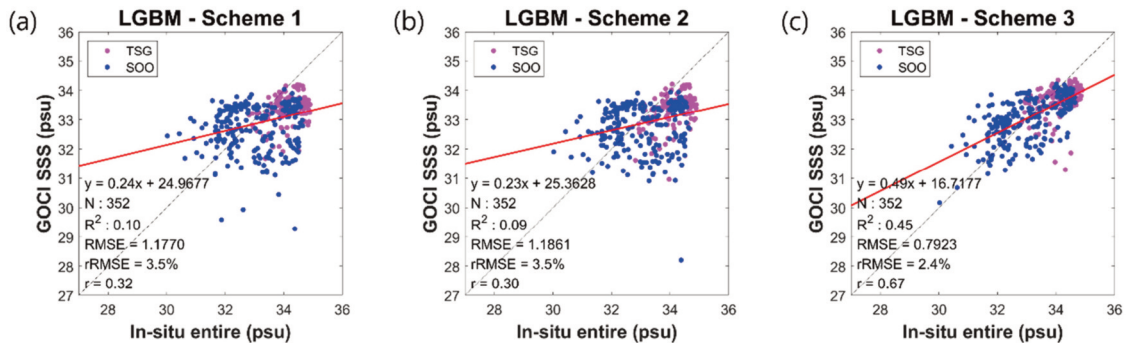


Fig. 6. Validation results of GOCI SSS using the entire in-situ data: (a) scheme 1, (b) scheme 2, and (c) scheme 3.

효과적인 환경변수로 작용하였다(Chen and Hu, 2017; Kim *et al.*, 2020).

2) 계절에 따른 동아시아 해역의 SSS 시공간적 분포 특성

SMAP SSS (Jang)와 GOCI SSS를 활용하여 지난 2019년 동아시아의 계절별 평균 SSS 분포 특성을 분석하였다. GOCI SSS의 경우 본 연구에서 제시한 3개의 scheme 중 가장 추정 성능이 높았던 scheme 3 모델을 활용하였으며, SMAP SSS (Jang)와 GOCI SSS 모두 훈련에 사용되지 않은 일자만 사용하였다.

SMAP SSS (Jang)와 GOCI SSS는 동아시아 해역에서 비슷한 계절적 분포 특성을 보였다(Fig. 7, 8). 담수의 영향을 비교적 적게 받는 동해와 북서태평양의 경우 연중 일정한 고염분수를 유지하는 것으로 나타났다. 반면 CDW의 영향을 직접적으로 받는 동중국해는 상대적으

로 저염분수로 나타났으며, 황해 또한 그 영향을 일부 받아 동해에 비해 낮은 염분 특성을 보였다. 특히 CDW의 양이 증가하는 7-8월 기간에는 동중국해에서 저염수의 영향 범위가 가장 넓었으며, 그 영향이 동해 근해 일부까지 미치는 것으로 나타났다.

SMAP SSS (Jang)는 25 km의 낮은 공간해상도 때문에, CDW 지역 저염분수를 분석한 다른 선행연구들과 같은 상세한 SSS 공간 패턴을 추정하기엔 한계가 있다(Ahn *et al.*, 2008; Beardsley *et al.*, 1985; Kim *et al.*, 2020; Fig. 7). 반면, GOCI SSS는 SMAP SSS (Jang)에 비해 높은 공간해상도(500 m) 덕분에, SMAP SSS (Jang)에서는 모의되지 않는 보다 상세한 공간적 패턴을 나타냈다(Fig. 8). 또한, SMAP SSS (Jang)가 RFI와 LC 등의 이유로 인해 연안 근처에서의 정보를 제공하지 않는 반면, GOCI SSS는 연안 해역의 상대적인 저염수까지도 모의하였다. 하지만, GOCI SSS는 SMAP SSS (Jang)에 비해 CDW 지역의 SSS

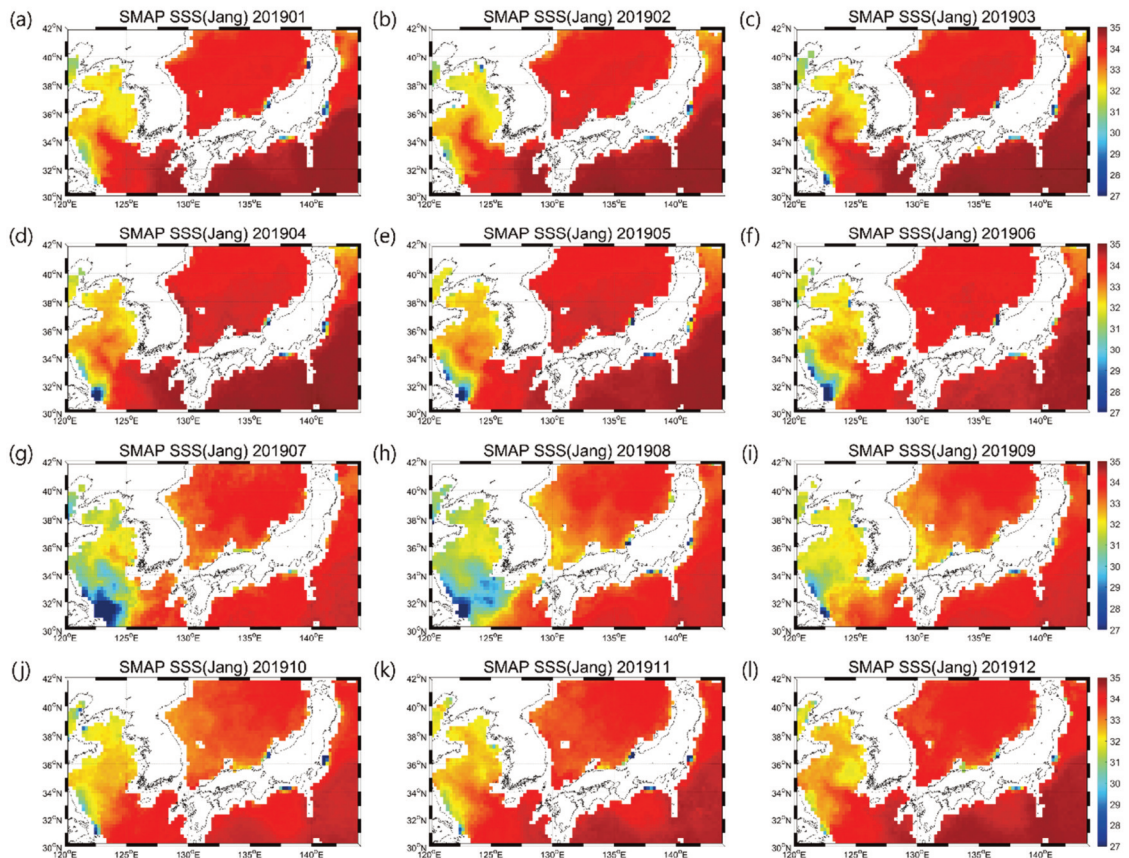


Fig. 7. Spatial distribution of monthly averaged SMAP SSS (Jang) in 2019: (a) January, (b) February, (c) March, (d) April, (e) May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December.

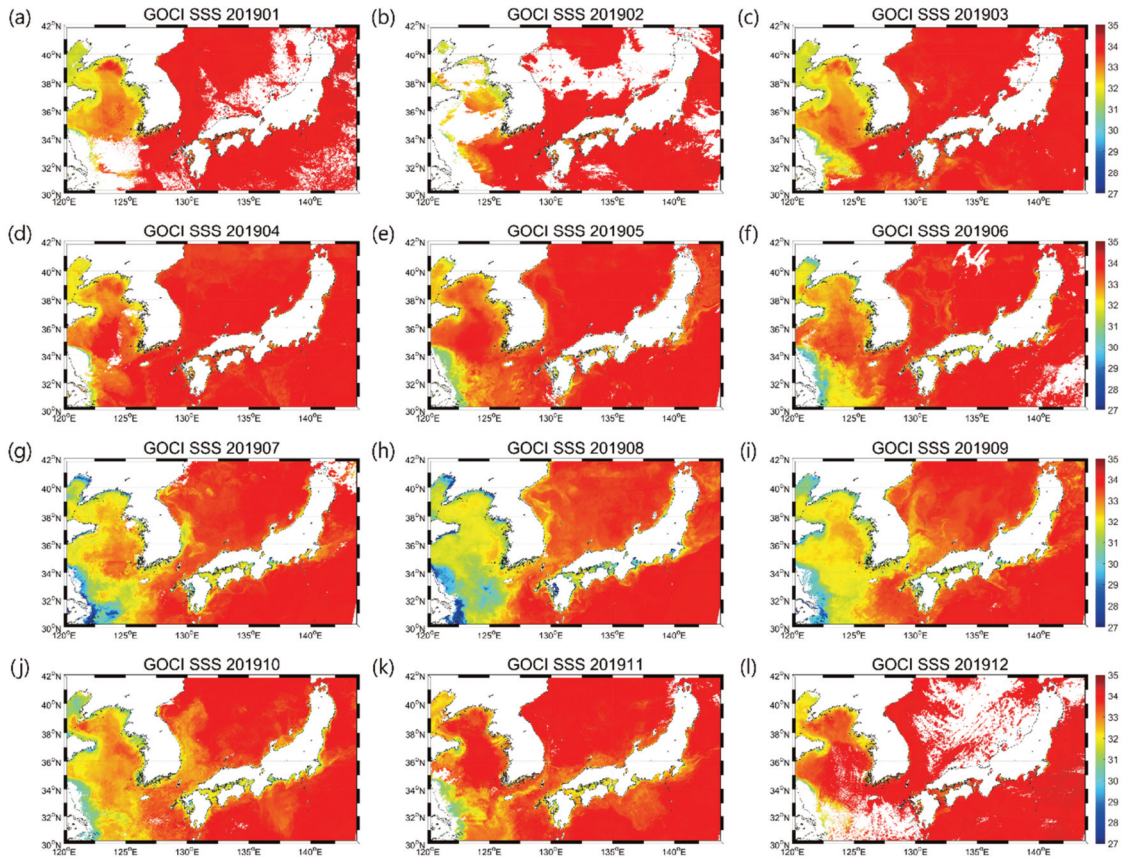


Fig. 8. Spatial distribution of monthly averaged GOCI SSS in 2019: (a) January, (b) February, (c) March, (d) April, (e) May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December.

를 다소 과대추정하는 것으로 보이며(Fig. 7, 8), 이는 저염수가 주로 발생하는 여름철에 GOCI 자료가 구름으로 인해 많이 소실되었기 때문으로 보인다(Fig. 9).

3) CDW의 거동 특성 분석

GOCI SSS를 활용하여 중국 남방지역에 대홍수가 발생하였던 2020년 여름을 대상으로 SSS의 시공간적 거동을 분석하였다. 앞서 4-2)절과 마찬가지로 scheme 3 모델을 활용하였으며, 구름의 영향을 비교적 적게 받았던 4일을 선정하여 일평균 자료를 생성하여 분석하였다.

2020년 여름은 중국 남방지역에 유래 없는 폭우가 발생하였으며, 6월 2일 호우 경보가 발령된 이후 많은 양의 CDW가 동중국해로 유입되었다. 2020년 7월 14일 기준으로 CDW의 유입량은 초당 82,000 톤으로 2003년 해양수산연구원 관측 이후 최대치를 기록하였다(MarineCom,

2020a). 이러한 특성은 8월 1일의 GOCI SSS에 잘 나타난다(Fig. 10(a)). 이후, 추가로 유입된 저염분수로 인해 8월 13일의 GOCI SSS에는 CDW의 영향 범위가 더욱 넓어진 것을 확인할 수 있다. 하지만 8월 20일 이후 8호 태풍 마이삭(Maysak)과 9호 태풍 바비(Bavi)가 동중국해와 한반도 영해를 통과하면서 표층의 저염분수를 심층 해수와 혼합하였고, 이로 인해 양자강 저염분수가 소멸한 것으로 보도되었다(MarineCom, 2020b). 8월 24일의 GOCI SSS는 이전 일자와 비교하여 CDW의 영향이 약해졌으며, 31일에는 거의 소멸하여 양자강 근처의 일부 영역에만 저염분수가 남아있는 것을 확인할 수 있다. 위의 분석 결과로, GOCI SSS가 CDW의 시공간적 거동을 효과적으로 모의하는 것으로 보인다.

4) SHAP 분석

제시한 scheme들 중 성능이 가장 높은 scheme 3에 대

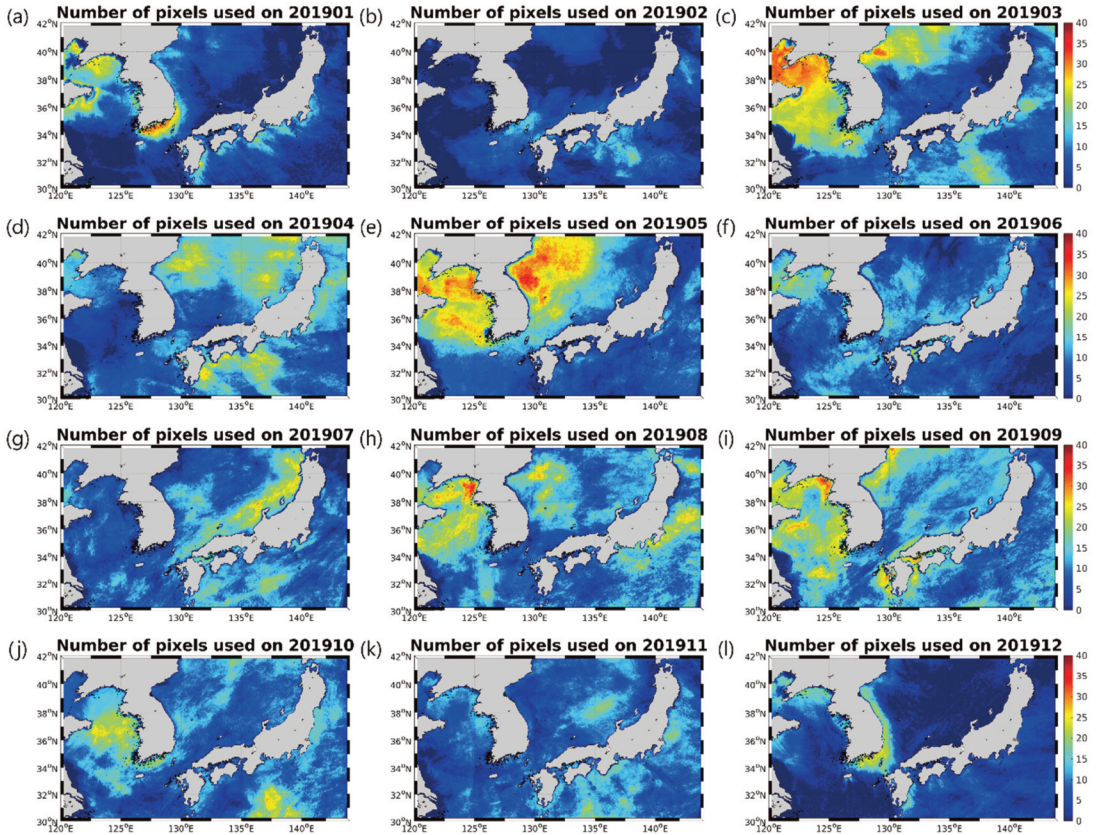


Fig. 9. Number of pixels used to produce monthly averaged GOCI SSS in 2019: (a) January, (b) February, (c) March, (d) April, (e) May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December. Note that the number of pixels for a month can reach up to 248 (31×8) with no cloud contamination.

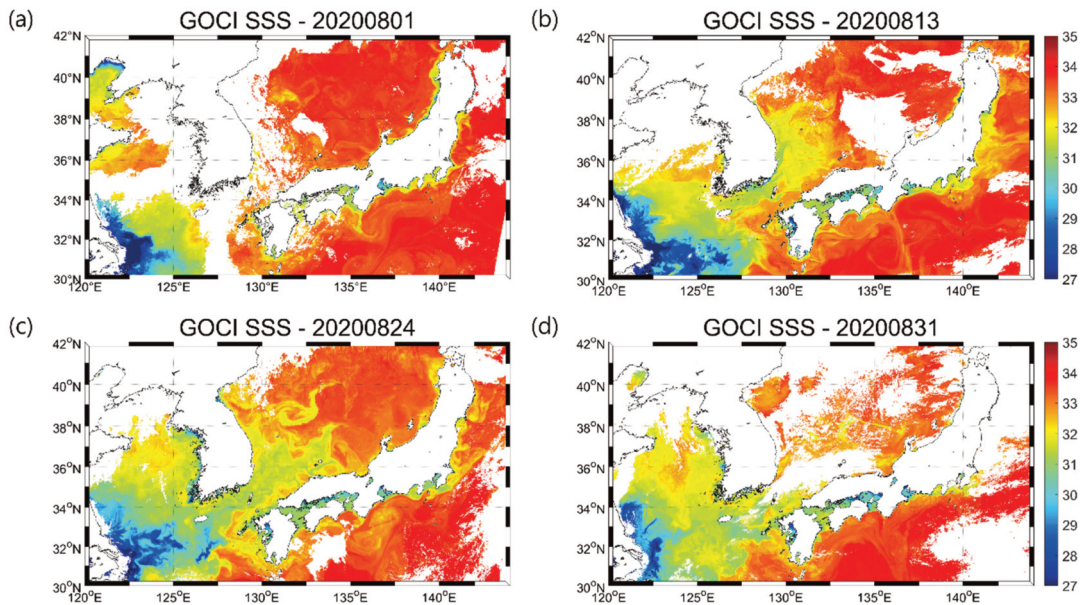


Fig. 10. Daily distribution of averaged GOCI SSS proposed in this study: August (a) 01, (b) 13, (c) 24, (d) 31 in 2019.

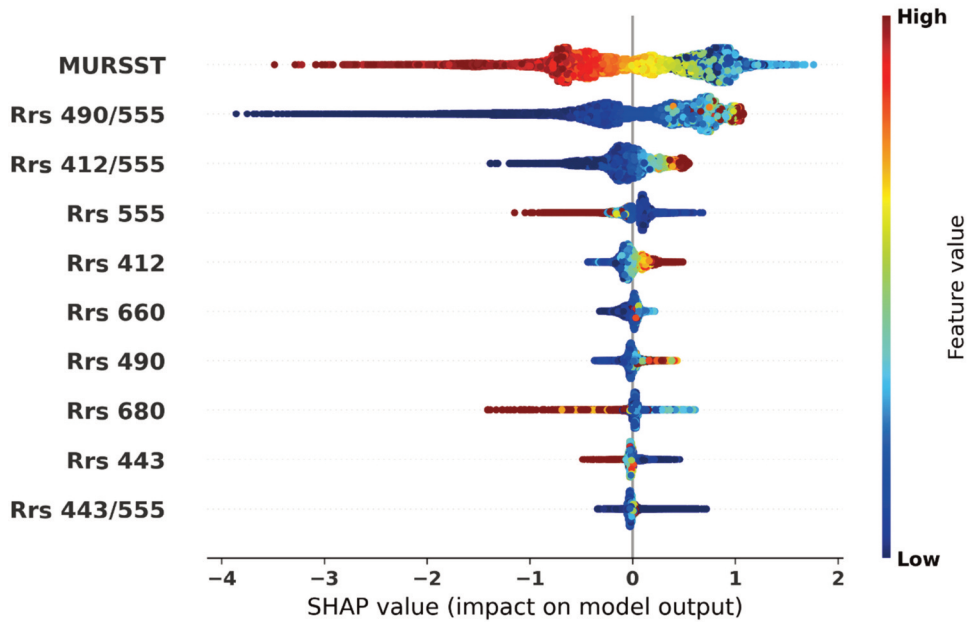


Fig. 11. Summary plot of SHapley Additive exPlanation analysis of the LGBM model to generate high resolution GOCI SSS.

해 SHAP 분석을 수행하였다. 앞서 4-1)절에서 Scheme 2의 입력변수에 MURSST를 추가한 후 R^2 0.3, RMSE 0.29 psu의 성능향상이 있었던 것에서 유추할 수 있듯이, scheme 3에서 MURSST가 가장 높은 변수기여도를 보였다(Fig. 11). 또한, 낮은 MURSST 값을 가지는 경우 SSS 예측 값을 증가시키는 방향으로, 높은 MURSST 값을 가지는 경우 SSS 예측 값을 감소시키는 방향으로 기여하였다. MURSST에 뒤따라, 낮은 SSS에서 낮은 값을 가지는 R_{rs} 490/555와 R_{rs} 412/555가 SSS와의 이론적 경향성을 따르며 두번째, 세번째로 높은 기여를 하였다(Park *et al.*, 2014; Chen *et al.*, 2017). R_{rs} 555는 주로 SPM의 농도와 양의 상관관계를 가지고 있는 변수로, CDOM과도 어느 정도의 상관성이 있어 SSS와 음의 상관성을 띤다(Chen *et al.*, 2017). 하지만, 그 정도가 크지 않아 SSS 산출에 활용성이 낮음을 알 수 있다. 대신 R_{rs} 555는 반사도비를 통해 단파장의 SPM 반사보정에 기여를 하였다. 기여도가 낮은 변수 중 R_{rs} 680의 경우 해수 탁도에 민감한 밴드로 광학적 해수 유형(예: case-1, case-2)을 구분하는데 주로 사용되는 변수이다(Feng *et al.*, 2020). 탁도 경향이 다른 해역을 구분하여 지역적으로 조금 더 정밀한 모델링이 되도록 기대를 하며 R_{rs} 680 변수를 활용하였다. 하지만, MURSST가 해당 역할을 대체하며 변수기여도가 낮게

나왔음을 예측할 수 있다.

SSS와 SST는 기상요인(예: 바람, 강수 등) 및 지리적 요인에 영향을 많이 받는 변수들이기 때문에(Kido *et al.*, 2021), 연구지역 내 특성이 다른 네 개의 해역에 대해 각각 산포분석을 수행하였다(Fig. 12). 동중국해, 황해, 동해, 북서태평양 지역에 대해 각각 32°N 124°E, 35°N 125°E, 38°N 134°E, 32°N 140°E을 중심으로 한 2°×2°영역을 대표지역으로 선정 후, 해당지역 내 샘플을 통해 산포분석을 수행하였다. 이를 통해, 각 지역별 평균 MURSST 및 SMAP SSS (Jang)와의 추세 경향이 다름을 확인하였다. 연안에서 멀리 떨어진 동해(Fig. 12 - green circles)와 대양에 포함된 북서태평양(Fig. 12 - blue circles)에서는 MURSST와 SMAP SSS (Jang) 사이의 상관성이 거의 없는 반면, CDW의 영향을 많이 받는 동중국해와 황해(Fig. 12 - red and yellow circles)에서는 음의 상관관계가 나타났다. 이를 통해 특정 상관성이 있는 지역에 대해 MURSST가 강하게 기여하였음을 알 수 있다.

5) 한계점 및 활용방안

본 연구에서는 현장관측자료를 이용한 모델 학습의 시공간 제약성과 기존 SMAP SSS 산출물의 불안정한 정확도의 한계를 극복하고자 기계학습을 활용해 보정

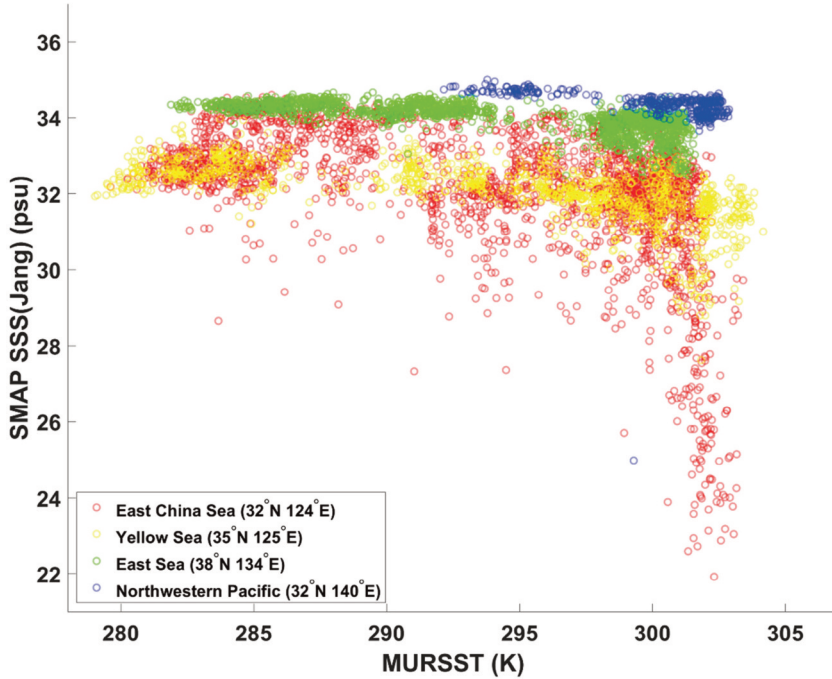


Fig. 12. Correlation plots between Multi-scale Ultra-high Resolution Sea Surface Temperature and SMAP SSS (Jang) in four different regions of $2^\circ \times 2^\circ$ window centered on 32°N 124°E (East China Sea), 35°N 125°E (Yellow Sea), 38°N 134°E (East Sea), 32°N 140°E (Northwestern Pacific).

된 격자 형태의 SSS를 참조자료로 하여 모델을 생성하였다. 격자형태의 검증자료를 통해 다량의 샘플이 확보되었으나, 마이크로웨이브 위성의 RFI 및 LC 한계로 인해 연안 근처의 샘플이 많이 부족하였다. 이로 인해, SSS가 30 psu 이하인 영역에 대한 모의력이 상대적으로 떨어지는 것이 확인되었다. 또한 광학위성의 특성상 구름으로 가려진 영역에 대해서는 SSS 산출이 불가능한데, 이에 따라 구름 발생 빈도가 상대적으로 높은 황해지역에 대한 학습샘플이 동해와 북서태평양 해역에 비해 떨어져, 모델의 시공간적 일반화에 한계가 있었다. 2020년 8월의 CDW 거동 분석에서도, 많은 자료가 비구름에 가려져 보다 상세한 시간적 변화 분석이 불가능했다. 위성기반 자료를 참조자료로 하여 학습한 모델의 추가 검증을 위해 현장자료를 활용하였지만, 동중국해 및 연안 해역에 관한 샘플이 부족하여 저염수에 관한 충분한 검증을 수행하지 못하였다. 해석위성을 이용한 SSS 모의는 SSS와 CDOM의 관계에 기반하여 간접적으로 유추를 해내는 만큼, 비선형적 상관성과 더불어 많은 오차 요인들이 발생할 수 있다. 이에 인공지능 기법을 도입

하여 오차 완화와 비선형적 모의를 시도하였지만, 인공지능 기법의 특성상 모델 구조 해석에 한계가 있었다. SHAP 알고리즘을 사용하여 입력 변수와 SSS 사이의 상관성을 분석하였지만, 그럼에도 완벽한 모델 해석에는 어려움이 있었으며, 이론과 더불어 간접적 해석만이 수행되었다.

기존의 마이크로웨이브 위성 기반의 SSS 자료는 낮은 시공간 해상도로 인해 글로벌 스케일의 해수 이동 및 기후 변화 연구에는 적합하지만, 연·근해 지역 SSS 모니터링에는 적합하지 않다. SSS의 현장자료 수집에는 많은 비용이 들며, 동아시아 해역이 CDW 지역의 저염분수에 많은 영향을 받음에도 불구하고, 해당 지역의 현장자료를 획득할 수 없다는 단점이 있다. 반면, 본 연구에서 제시한 LGBM 기반 모델은 CDW 지역의 저염분수와 더불어 동아시아 해역 전체의 고해상도 SSS 영상을 높은 시간해상도로 산출할 수 있다. 또한 기존의 광학위성을 활용한 SSS 산출 연구들은 시공간적인 제약성과 마이크로웨이브 SSS 산출물의 불확실성 등의 한계를 가지고 있지만, 본 연구에서 제시한 모델은 전 지

구 영역의 현장자료를 참조로 보정된 SMAP SSS (Jang)를 타겟으로 학습함으로써 이를 극복하였다(Chen and Hu, 2017; Kim *et al.*, 2020; Kim *et al.*, 2021). 기존의 연구에서는 광학신호와 SSS의 관계가 분석되지 않았거나, 입력변수들 사이의 상대적인 중요도만 분석되었다. 이에 반해, 본 연구에서는 SHAP 알고리즘을 이용해 각 입력변수의 범위에 따라 SSS 예측에 미치는 영향의 크기와 방향성까지 분석하였다. 이를 바탕으로, 10분의 짧은 시간해상도로 SST 자료를 제공하는 천리안위성 2A/2B호(GEO-KOMPSAT-2A/2B, GK-2A/2B) 위성영상과, 250 m의 높은 공간해상도 및 보다 많은 가시광 밴드를 가진 GOCI-2 위성센서자료 등 차세대 위성정보를 활용한다면, 동아시아 영역에서 보다 수준 높은 염분 모니터링이 가능할 것으로 보인다.

5. 결론

본 연구에서는 GOCI R_{rs} , MURSST와 SMAP SSS (Jang) 자료를 활용하여 고해상도 SSS를 추정하는 LGBM 기반의 알고리즘을 개발하였다. 3가지 입력변수 조합을 테스트하였고, 각 scheme별로 산출된 고해상도 SSS를 훈련에 사용되지 않은 SMAP SSS (Jang) 및 현장자료를 활용하여 검증하였다. MURSST 자료가 추가된 scheme 3가 GOCI 광학위성신호만을 사용한 scheme 1과 2에 정확도가 높았고, 이를 통해 본 연구영역에서 SST가 SSS 모의에 효과적인 환경변수로 작용함을 보였다. 이후 scheme 3의 LGBM 모델로 산출한 2019년의 월 평균 고해상도 SSS를 활용하여 동아시아 주요해역에서 SSS의 시공간적, 계절적 패턴을 분석하였다. 본 연구에서 제시한 GOCI SSS는 SMAP SSS (Jang)와 비슷한 시공간적 패턴을 보였지만, 더 높은 공간해상도를 바탕으로 SSS의 보다 상세한 공간적 분포와 더불어 SMAP SSS (Jang)에서 산출하지 않는 연안 지역의 정보까지 모의하였다. 또한 중국 남방지역에 대홍수가 발생하였던 2020년 8월을 대상으로 CDW의 거동을 분석한 결과, GOCI SSS는 한국 해양수산연구원의 보도자료와 비교하여 일관성 있는 시공간적 변화를 보였다. 마지막으로, SHAP 알고리즘을 활용하여 각 입력변수가 예측 값에 미치는 영향을 분석하였다. 분석결과 MURSST의 기여도가 가장 높았

으며, 그 외 R_{rs} 490/555와 R_{rs} 412/555가 이론적 경향성에 잘 따르며 높은 기여도를 보였다. 본 연구의 결과로 동아시아 해역에서 광학위성 신호를 활용한 고해상도 SSS 산출의 가능성을 제시하였다.

사사

본 연구는 해양수산부의 ‘해양위성영상 분석 활용 기술 개발(KIMST-20210046)’과 ‘과학기술기반 해양환경영향평가 기술개발(KIMST-20210427)’ 및 행정안전부의 ‘재난안전 부처협력 기술개발사업(20009742)’의 지원을 받아 수행된 연구임.

References

- Ahn, Y., P. Shanmugam, J. Moon, and J.-H. Ryu, 2008. Satellite remote sensing of a low-salinity water plume in the East China Sea, *Annales Geophysicae*, 26(7): 2019-2035. <https://doi.org/10.5194/angeo-26-2019-2008>
- Bao, S., H. Wang, R. Zhang, H. Yan, and J. Chen, 2019. Comparison of satellite-derived sea surface salinity products from SMOS, Aquarius, and SMAP, *Journal of Geophysical Research: Oceans*, 124(3): 1932-1944. <https://doi.org/10.1029/2019JC014937>
- Barale, V., J. Gower, and L. Alberotanza, 2010. *Oceanography from space: revisited*, Springer Science & Business Media, Berlin, Heidelberg, Germany.
- Beardsley, R., R. Limeburner, H. Yu, and G. Cannon, 1985. Discharge of the Changjiang (Yangtze river) into the East China sea, *Continental Shelf Research*, 4(1-2): 57-76. [https://doi.org/10.1016/0278-4343\(85\)90022-6](https://doi.org/10.1016/0278-4343(85)90022-6)
- Boutin, J., J.-L. Vergely, S. Marchand, F. d’Amico, A. Hasson, N. Kolodziejczyk, N. Reul, G. Reverdin, and J. Vialard, 2018. New SMOS Sea Surface Salinity with reduced systematic errors and improved

- variability, *Remote Sensing of Environment*, 214: 115-134. <https://doi.org/10.1016/j.rse.2018.05.022>
- Breiman, L., 2001. Random forests, *Machine Learning*, 45(1): 5-32.
- Chen, S. and C. Hu, 2017. Estimating sea surface salinity in the northern Gulf of Mexico from satellite ocean color measurements, *Remote Sensing of Environment*, 201: 115-132. <https://doi.org/10.1016/j.rse.2017.09.004>
- Chin, T.M., J. Vazquez-Cuervo, and E.M. Armstrong, 2017. A multi-scale high-resolution analysis of global sea surface temperature, *Remote Sensing of Environment*, 200: 154-169. <https://doi.org/10.1016/j.rse.2017.07.029>
- Feng, J., H. Chen, H. Zhang, Z. Li, Y. Yu, Y. Zhang, M. Bilal, and Z. Qiu, 2020. Turbidity estimation from GOCI satellite data in the turbid estuaries of China's coast, *Remote Sensing*, 12(22): 3770. <https://doi.org/10.3390/rs12223770>
- Fournier, S., T. Lee, W. Tang, M. Steele, and E. Olmedo, 2019. Evaluation and intercomparison of SMOS, Aquarius, and SMAP sea surface salinity products in the Arctic Ocean, *Remote Sensing*, 11(24): 3043. <https://doi.org/10.3390/rs11243043>
- Friedman, J. H., 2001. Greedy function approximation: a gradient boosting machine, *Annals of Statistics*, 29(5): 1189-1232. <http://www.jstor.org/stable/2699986>
- Geiger, E.F., M.D. Grossi, A.C. Trembanis, J.T. Kohut, and M. J. Oliver, 2013. Satellite-derived coastal ocean and estuarine salinity in the Mid-Atlantic, *Continental Shelf Research*, 63: S235-S242. <https://doi.org/10.1016/j.csr.2011.12.001>
- Hoffman, L., M.R. Mazloff, S.T. Gille, D. Giglio, and A. Varadarajan, 2022. Ocean Surface Salinity Response to Atmospheric River Precipitation in the California Current System, *Journal of Physical Oceanography*. 52(8): 1867-1885. <https://doi.org/10.1175/JPO-D-21-0272.1>
- Jang, E., Y.J. Kim, J. Im, Y.-G. Park, and T. Sung, 2022. Global sea surface salinity via the synergistic use of SMAP satellite and HYCOM data based on machine learning, *Remote Sensing of Environment*, 273: 112980. <https://doi.org/10.1016/j.rse.2022.112980>
- Ke, G., Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, 2017. LightGBM: A highly efficient gradient boosting decision tree, *Proc. of Neural Information Processing Systems*, Long Beach, CA, Dec. 4-9, vol. 30, pp. 1-9.
- Kido, S., M. Nonaka, and Y. Tanimoto, 2021. Sea Surface Temperature-Salinity Covariability and Its Scale-Dependent Characteristics, *Geophysical Research Letters*, 48(24): e2021GL096010. <https://doi.org/10.1029/2021GL096010>
- Kim, D.-W., S.-H. Kim, and Y.-H. Jo, 2021. A Development for Sea Surface Salinity Algorithm Using GOCI in the East China Sea, *Korean Journal of Remote Sensing*, 37(5-2): 1307-1315 (in Korean with English abstract). <https://doi.org/10.7780/kjrs.2021.37.5.2.8>
- Kim, D.-W., Y.-J. Park, J.-Y. Jeong, and Y.-H. Jo, 2020. Estimation of hourly sea surface salinity in the east China sea using geostationary ocean color imager measurements, *Remote Sensing*, 12(5): 755. <https://doi.org/10.3390/rs12050755>
- Klein, L. and C. Swift, 1977. An improved model for the dielectric constant of sea water at microwave frequencies, *IEEE Transactions on Antennas and Propagation*, 25(1): 104-111. <https://doi.org/10.1109/TAP.1977.1141539>
- Kolodziejczyk, N., M. Hamon, J. Boutin, J.-L. Vergely, G. Reverdin, A. Supply, and N. Reul, 2021. Objective analysis of SMOS and SMAP sea surface salinity to reduce large-scale and time-dependent biases from low to high latitudes, *Journal of Atmospheric and Oceanic Technology*, 38(3): 405-421. <https://doi.org/10.1175/JTECH-D-20-0093.1>
- Kowalczyk, P., W.J. Cooper, R.F. Whitehead, M.J. Durako, and W. Sheldon, 2003. Characterization of CDOM in an organic-rich river and surrounding

- coastal ocean in the South Atlantic Bight, *Aquatic Sciences*, 65(4): 384-401. <https://doi.org/10.1007/s00027-003-0678-1>
- Le Vine, D. M., 2019. RFI and remote sensing of the earth from space, *Journal of Astronomical Instrumentation*, 8(01): 1940001. <https://doi.org/10.1142/S2251171719400014>
- Lundberg, S.M. and S.-I. Lee, 2017. A unified approach to interpreting model predictions, *Proc. of Neural Information Processing Systems*, Long Beach, CA, Dec. 4-9, vol. 30, pp. 1-10.
- MarineCom, 2020a. 2020 High temperature-low salinity water weekly monitoring information (No.4), <http://www.jeju.go.kr/jori/monitor/oceanmonitor/monitor.htm?page=2&act=view&seq=1251829#english>, Accessed on Oct. 23, 2022.
- MarineCom, 2020b. 2020 High temperature-low salinity water weekly monitoring information (No.7), <http://www.jeju.go.kr/jori/monitor/oceanmonitor/monitor.htm?page=2&act=view&seq=1255564>, Accessed on Oct. 23, 2022.
- Moon, J.-H., T. Kim, Y.B. Son, J.-S. Hong, J.-H. Lee, P.-H. Chang, and S.-K. Kim, 2019. Contribution of low-salinity water to sea surface warming of the East China Sea in the summer of 2016, *Progress in Oceanography*, 175: 68-80. <https://doi.org/10.1016/j.pocean.2019.03.012>
- NOAA Climate.gov, 2014. 2013 State of the Climate: Ocean Salinity, <https://www.climate.gov/news-features/understanding-climate/2013-state-climate-ocean-salinity>, Accessed on Oct. 23, 2022.
- Olmedo, E., I. Taupier-Letage, A. Turiel, and A. Alvera-Azcárate, 2018. Improving SMOS sea surface salinity in the western Mediterranean sea through multivariate and multifractal analysis, *Remote Sensing*, 10(3): 485. <https://doi.org/10.3390/rs10030485>
- Olmedo, E., V. González-Gambau, A. Turiel, J. Martínez, C. Gabarró, M. Portabella, J. Ballabrera-Poy, M. Arias, R. Sabia, and R. Oliva, 2019. Empirical characterization of the SMOS brightness temperature bias and uncertainty for improving sea surface salinity retrieval, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7): 2486-2503. <https://doi.org/10.1109/JSTARS.2019.2904947>
- Park, S., Y. Noh, S. Jung, and E. Hwang, 2021. SHAP-based Explainable Photovoltaic Power Forecasting Scheme Using LSTM, *Proc. of the Korea Information Processing Society Conference*, Yeosu, Nov. 4-6, vol. 28, pp. 846-848. <https://doi.org/10.3745/PKIPS.y2021m11a.845>
- Park, Y., Y. Ahn, H. Han, H. Yang, J. Moon, J. Ahn, B. Lee, J. Min, S. Lee, and K. Kim, 2014. *GOCI Level 2 Ocean Color Products (GDPS 1.3) Brief Algorithm Description*, Korea Ocean Satellite Center (KOSC), Ansan, Republic of Korea, pp. 24-40.
- Qin, S., H. Wang, J. Zhu, L. Wan, Y. Zhang, and H. Wang, 2020. Validation and correction of sea surface salinity retrieval from SMAP, *Acta Oceanologica Sinica*, 39(3): 148-158. <https://doi.org/10.1007/s13131-020-1533-0>
- Shapley, L.S., 2016. 17. A Value for n-Person Games, In: Kuhn, H.W., Tucker, A.W. (eds), *Contributions to the Theory of Games (AM-28), Volume II*, Princeton University Press, Princeton, NJ, USA, pp. 307-318. <https://doi.org/10.1515/9781400881970-018>
- Sun, D., X. Su, Z. Qiu, S. Wang, Z. Mao, and Y. He, 2019. Remote Sensing Estimation of Sea Surface Salinity from GOCI Measurements in the Southern Yellow Sea, *Remote Sensing*, 11(7): 775. <https://doi.org/10.3390/rs11070775>
- Taggio, N., A. Aiello, G. Ceriola, M. Kremezi, V. Kristollari, P. Kolokoussis, V. Karathanassi, and E. Barbone, 2022. A Combination of machine learning algorithms for marine plastic litter detection exploiting hyperspectral PRISMA data, *Remote Sensing*, 14(15): 3606. <https://doi.org/>

- 10.3390/rs14153606
- Tang, W., S. Yueh, D. Yang, A. Fore, A. Hayashi, T. Lee, S. Fournier, and B. Holt, 2018. The potential and challenges of using Soil Moisture Active Passive (SMAP) sea surface salinity to monitor Arctic Ocean freshwater changes, *Remote Sensing*, 10(6): 869. <https://doi.org/10.3390/rs10060869>
- Werther, M., E. Spyrakos, S. G. Simis, D. Odermatt, K. Stelzer, H. Krawczyk, O. Berlage, P. Hunter, and A. Tyler, 2021. Meta-classification of remote sensing reflectance to estimate trophic status of inland and nearshore waters, *ISPRS Journal of Photogrammetry and Remote Sensing*, 176: 109-126. <https://doi.org/10.1016/j.isprsjprs.2021.04.003>