



Received: 2026/07/22
Revised: 2026/08/02
Accepted: 2026/08/26
Published: 2026/09/30

***Corresponding Author:**

Sunghoon Jung
Tel: +82-55-907-5320
E-mail: shjung@navy.ac.kr

표층정보를 활용한 연직 해양환경 예측 및 음속장 구축

Surface Information-Based Prediction of the Subsurface Ocean Environment and Sound-Speed Profiles

정성훈^{1*}, 이호준²

¹해군 소령/해군사관학교 기계시스템공학과 부교수

²해군 소령/해군사관학교 해양학과 조교수

Sunghoon Jung^{1*}, Hojun Lee²

¹LCDR, ROK Navy/Associate professor, Dept. of Mechanical System Engineering, Republic of Korea Naval Academy

²LCDR, ROK Navy/Assistant professor, Dept. of Oceanography, Republic of Korea Naval Academy

Abstract

본 연구는 한국해양자료센터(KODC) 정선해양조사자료를 이용하여 표층수온 및 표층 염분을 활용한 연직 해양환경 예측 및 음속장 구축 기법을 제안한다. 위도, 경도, 연도, 월, 수심 등의 기본입력자료와 표층수온 및 표층염분을 활용하여 연직 수온 및 염분분포를 예측하고, 예측 결과를 기반으로 Mackenzie 음속식을 이용하여 음속을 산출하였다. 실험 결과, 표층수온 및 표층염분 정보를 추가한 경우 연직 수온 및 염분분포와 음속의 예측 성능이 모두 향상되었으며, 특히 연직 염분분포의 예측 정확도가 크게 개선되었다. 제안한 기법은 다양한 해역의 연직 해양환경 및 음속장 구축에 활용될 수 있을 것으로 기대된다.

This study proposes a method for predicting the subsurface ocean environment and constructing sound-speed profiles using sea-surface temperature and salinity based on the hydrographic observation dataset provided by the Korea Oceanographic Data Center (KODC). The vertical temperature and salinity profiles were predicted using latitude, longitude, year, month, and depth as the basic input variables, together with sea-surface temperature and sea-surface salinity. The predicted profiles were then used to calculate the sound-speed profiles based on the Mackenzie sound-speed equation. The experimental results show that incorporating sea-surface temperature and sea-surface salinity improved the prediction performance of vertical temperature and salinity profiles, as well as sound-speed profiles, with the greatest improvement observed in the prediction accuracy of the vertical salinity profiles. The proposed method is expected to facilitate the reconstruction of subsurface ocean environments and sound-speed profiles in various marine regions.

Keywords

연직 해양환경(Subsurface Ocean Environment), 수온 프로파일(Temperature Profile), 염분 프로파일(Salinity Profile), 음속 프로파일(Sound-Speed Profile), 기계학습(Machine Learning)

Acknowledgement

이 논문은 2026년 해군사관학교 해양연구소 학술연구과제 연구비의 지원으로 수행된 연구임.

1. 서론

수중에서 음파는 전자기파보다 장거리까지 효과적으로 전파될 수 있어 잠수함 탐지, 수중 감시, 기뢰전 및 대잠전 등 다양한 해군 작전에서 핵심적인 수단으로 활용된다. 특히 음파의 전파 경로는 해수의 음속 분포에 의해 크게 좌우되므로, 작전 해역의 음속장을 정확하게 구축하는 것은 수중환경 분석과 음향체계 운용을 위한 필수 요소이다. 따라서 작전 해역의 음속장을 신속하고 정확하게 확보하는 것은 해군 전력 운용의 효율성과 직결되는 중요한 문제이다.

해수의 음속은 수온, 염분 및 압력(수심)에 따라 달라지는데 이 중 수온이 음속 변화에 가장 큰 영향을 미친다. 압력은 수심으로부터 산출할 수 있으므로 정확한 음속장을 구축하기 위해서는 연직 방향의 수온과 염분 정보가 반드시 필요하다. 연직 수온과 염분을 직접 관측하기 위해서 일반적으로 CTD(conductivity-temperature-depth), XBT(expendable bathy thermograph) 등의 해양관측 장비를 이용한 현장 관측이 필요하지만, 넓은 해역의 모든 위치와 수심에 대해 이러한 자료를 지속적으로 확보하는 것은 시간적·공간적·경제적 제약으로 인해 현실적으로 어렵다. 이로 인해 해양환경 정보의 시·공간

적 공백이 발생하며, 음속장의 정확한 구축에도 제약이 따른다.

최근에는 이러한 한계를 극복하기 위하여 머신러닝 기법을 활용한 해양환경 예측 연구가 활발히 수행되고 있다. [1]에서는 동해 수직 수온 예측을 위해 random forest(RF), XGBoost 및 LightGBM을 적용하였으며, [2]에서는 연안 해역에서 표층수온을 이용한 저층 수온 예측에 딥러닝 기법을 적용하였다. 이러한 연구들은 머신러닝을 이용한 해양환경 예측의 활용 가능성을 제시하였으나, 대부분 특정 변수인 수온을 대상으로 하거나 제한된 수심 및 해역에서 수행되었으며, 예측모델의 성능 향상에 초점을 두고 연구가 이루어졌다.

한편, 연직 해양환경을 직접 관측하는 데에는 공간적·시간적 제약이 존재하지만, 표층수온과 표층염분은 위성 기반 원격탐사를 통해 지속적으로 제공되고 있다[3,4]. 따라서 예측 대상 시점에서 확보 가능한 표층수온 및 표층염분(이하 표층정보)을 입력자료로 활용한다면 직접 관측하지 않은 수심의 수온 및 염분을 보다 효과적으로 추정할 수 있을 것으로 기대된다.

이에 본 연구에서는 실제 운용 환경에서 확보 가능한 표층정보를 활용하여 동해 연직 수온 및 염분을 예측하는 방법을 제안한다. 이를 위해 국립수산물과학원 한국해양자료센터(Korea Oceanographic Data Center, KODC)의 정선해양조사자료(이하 정선관측자료)[5]를 이용하여 위치, 시간 및 수심으로 구성된 기본입력자료와 여기에 표층수온 및 표층염분을 추가한 표층정보를 각각 구축하였다. 또한, linear regression(LR), RF, XGBoost 및 multi-layer perceptron(MLP)을 적용하여 입력자료 구성에 따른 예측 성능을 비교·분석하였다. 마지막으로 예측된 수온과 염분으로부터 음속을 계산하여, 제안한 방법의 수중 음향환경 예측에 대한 활용 가능성을 평가하였다.

2. 연구방법

2.1 연구자료

본 연구에서는 KODC에서 제공하는 정선관측자료를 이용하였다. KODC는 우리나라 주변 해역에서 수행된 해양관측을 통해 생산된 자료를 수집, 관리하는 국가 해양자료센터로서, 장기간 축적된 해양환경 자

료를 제공하고 있다.

정선해양조사는 미리 설정된 정선과 정점에서 동일한 관측항목을 2~3개월 간격 주기로 반복 조사하는 방식이다. 동일 해역을 장기간 관측하기 때문에 계절변동뿐만 아니라 해양환경의 장기적인 변화를 분석하기 위한 기초자료로 활용할 수 있다. 정선은 여러 관측정점으로 구성되며, 각 정점에서는 표층부터 해당 정점의 최대 관측 수심까지 수심별 해양환경 자료를 관측한다.

동해 정선관측망 위치는 Fig. 1과 같이 102, 103, 104, 105, 106, 107, 208 및 209의 8개 정선과 총 69개 정점으로 구성되어 있다. 관측은 매년 짝수 달에 연 6회 수행되며, 이를 통해 동해 해양환경의 계절적·장기적 변화를 지속적으로 추적하고 있다. 본 연구에서는 이 가운데 1970년부터 2025년까지의 동해 정선관측자료를 대상으로 하였다.

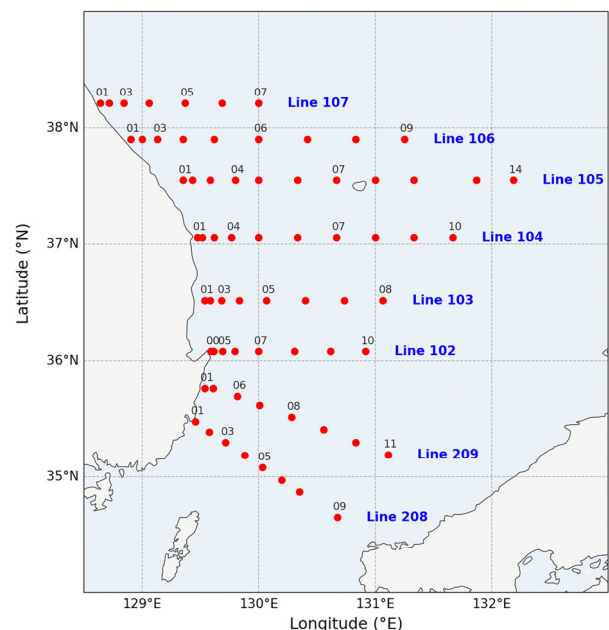


Fig. 1. Locations of KODC observation lines and stations in the East Sea

정선관측자료에는 관측연도와 관측 위치를 나타내는 위도·경도, 수심별 수온 및 염분 자료가 포함되어 있으며, 이와 함께 용존산소, 영양염류, pH, 투명도 등의 해양환경 항목과 기압 및 조사선 등의 관측정보도 제공된다. 본 연구에서는 이들 관측항목 중 연직 수온과 염분을 예측하기 위해 위도, 경도, 관측연도, 관측월, 수심, 수온 및 염분 자료를 활용하였다.

2.2 데이터셋 구축 및 전처리

본 연구는 장기간 축적된 정선관측자료를 이용하여 동해 해양환경의 시·공간적 특성을 학습하고, 음속 산출의 기반이 되는 특정 시점과 위치 및 수심에서의 수온과 염분 예측을 목적으로 한다. 이를 위해 KODC 원자료를 기계학습에 적용 가능한 형태의 데이터셋으로 재구성하였다. 하나의 관측 프로파일에는 동일한 시점과 정점에서 여러 수심의 수온 및 염분자료가 포함되어 있으므로, 각 수심의 관측자료를 하나의 독립적인 학습 표본으로 구성하였다.

연직 수온과 염분의 변화에 영향을 미칠 수 있는 위치, 시간 및 수심 정보를 기본입력자료로 선정하였으며, 각 표본의 출력자료는 해당 수심에서 관측된 수온과 염분으로 구성하였다. 그러나 동일한 위치와 시기에서도 실제 해양 상태에 따라 연직 수온 및 염분의 분포가 달라질 수 있다. 한편, 서론에서 언급한 바와 같이 표층수온과 표층염분은 위성 기반 원격탐사를 통해 예측 시점에도 확보 가능한 정보이다. 따라서 예측 시점의 표층수온과 표층염분을 추가 입력자료로 활용하면 해당 시점의 해양환경을 보다 효과적으로 반영할 수 있을 것으로 기대하였다. 이에 기본입력자료만을 사용하는 경우와 기본입력자료에 표층수온 및 표층염분을 추가하는 경우를 각각 구성하여, 표층정보가 연직 수온 및 염분 예측 성능에 미치는 영향을 분석하였다.

원자료 중 수온 또는 염분이 하나라도 결측된 표본은 분석에서 제외하였다. 표층정보를 사용하는 경우에는 동일한 정선-정점, 연도 및 월에서 관측된 0 m 수온과 염분을 표층정보로 정의하고, 이를 동일한 관측 프로파일에 포함된 각 수심 자료에 결합하였다. 표층수온 또는 표층염분이 존재하지 않는 프로파일은 표층정보를 이용한 데이터셋에서 제외하였다. 또한, 입력변수와 출력변수에 동일한 관측값이 포함되지 않도록 표층정보를 사용하는 데이터셋에서는 0 m 자료를 예측 대상에서 제외하였다.

자료는 시간 순서를 고려하여 1970~2018년을 학습 세트, 2019~2021년을 검증 세트, 2022~2025년을 시험 세트로 분할하였다. 월은 계절적 주기성을 반영하기 위해 사인과 코사인 함수로 변환하였으며, 수심은 표층과 심층 간 값의 범위 차이를 완화하기 위해 로그 변환을 적용하였다. 위도, 경도, 연도, 로그 변환

수심, 수온 및 염분에는 Min-Max 정규화를 수행하였으며, 표층정보를 사용하는 경우에는 표층수온과 표층염분에도 같은 방법을 적용하였다. 정규화에 필요한 최솟값과 최댓값은 학습자료에서만 산출하고, 동일한 기준을 검증자료와 시험자료에 적용하여 데이터 누출을 방지하였다.

최종적으로 기본입력자료는 위도, 경도, 정규화된 연도, 월의 사인·코사인 성분 및 로그 변환·정규화된 수심의 6개 입력변수로 구성하였다. 표층정보를 사용하는 경우에는 표층수온과 표층염분을 추가하여 총 8개의 입력변수로 구성하였다. 두 경우 모두 출력자료는 해당 수심의 수온과 염분으로 동일하게 설정하였다. 데이터셋 구성과 전처리 과정은 Table 1에 요약하였다.

Table 1. Dataset composition and data preprocessing

	Basic Inputs	Basic Inputs + Surface Information
Input Variables	Latitude, Longitude, Year, Month (sin, cos), Depth	Basic Inputs + Sea Surface Temperature + Sea Surface Salinity
Target Variables	Temperature, Salinity	
Dataset Split	Train: 212,140 Validation: 12,693 Test: 15,574	Train: 193,872 Validation: 11,559 Test: 14,184
Month Encoding	$\sin\left(\frac{2\pi}{12}\text{month}\right), \cos\left(\frac{2\pi}{12}\text{month}\right)$	
Depth Transformation	$\log(1 + D)$	
Normalization	$\frac{x - x_{\min}}{x_{\max} - x_{\min}}$ (Based on the Training Set)	

2.3 예측모델 구성 및 성능평가

본 연구에서는 동해 해양환경의 시·공간적 특성을 학습하여 특정 시점과 위치에서 연직 수온 및 염분을 예측하는 모델을 구축하고, 입력자료 구성에 따른 성능 변화를 비교하기 위하여 선형 회귀모델부터 전통적인 머신러닝 모델, 딥러닝 모델까지 서로 다른 특성을 갖는 대표적인 예측모델을 선정하였다. 이를 위해 linear regression(LR), random forest(RF), extreme gradient boosting(XGBoost) 및 multi-layer perceptron(MLP)을 적용하였다[6-8]. LR은 입력변수와 출력변수 간의 선형 관계를 기반으로 예측하는

대표적인 회귀모델이며, RF는 여러 개의 결정트리를 이용하여 복잡한 비선형 관계를 안정적으로 학습할 수 있는 모델이다. XGBoost는 이전 학습 결과를 반복적으로 보완하면서 예측 성능을 향상시키는 앙상블 모델로 다양한 회귀 문제에서 우수한 성능을 보이는 것으로 알려져 있다. 또한, MLP는 다층 신경망을 이용하여 입력변수 간의 복잡한 비선형 특성을 효과적으로 학습할 수 있는 대표적인 딥러닝 모델이다. 본 연구에서는 서로 다른 특성을 갖는 네 가지 예측모델을 동일한 데이터셋과 학습 조건에서 비교함으로써 입력자료 구성에 따른 예측 성능의 변화를 종합적으로 분석하였다.

각 예측모델의 하이퍼파라미터는 예비 실험과 성능 비교를 통해 최적화하였다. RF와 XGBoost는 주요 하이퍼파라미터를 조정하여 가장 우수한 성능을 나타내는 조합을 선정하였으며, MLP는 은닉층 수, 노드 수 및 학습률 등을 대상으로 grid search를 이용한 하이퍼파라미터 최적화를 수행하였다. 또한, 기본입력자료 모델과 표층정보 모델은 입력변수 구성이 서로 다르므로, 각 데이터셋에 대해 독립적으로 grid search를 수행하여 최적의 신경망 구조를 선정하였다.

모델의 예측 성능은 아래와 같이 평균절대오차(mean absolute error, MAE), 평균제곱근오차(root mean squared error, RMSE) 및 결정계수(coefficient of determination, R^2)를 이용하여 평가하였다.

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

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

$$R^2 = 1 - \left\{ \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right\} \quad (3)$$

여기서 y_i 는 실제값, $\hat{y}_i$ 는 예측값, $\bar{y}$ 는 실제값의 평균, n 은 전체 표본 수를 의미한다. MAE는 예측값과 실제값 사이의 평균 절대오차를 나타내며 평균적으로 발생하는 오차의 크기를 실제 단위(°C, psu 및 m/s)로 직관적으로 해석할 수 있는 지표이다. RMSE는 오차를 제곱하여 평균한 후 제곱근을 취한 값으로 큰 오차에 더 큰 가중치를 부여하므로 모델의 안정성과 예측 오차의 분산을 평가하는 데 적합하다. 또한, R^2 는 모

델이 실제 자료의 변동성을 얼마나 설명하는지를 나타내는 결정계수로, 1에 가까울수록 예측 성능이 우수함을 의미한다.

본 연구에서는 수온과 염분의 예측 성능뿐만 아니라, 예측된 수온과 염분을 Mackenzie가 제안한 음속식[9]에 적용하여 계산한 음속에 대해서도 동일한 평가지표를 적용하였다.

$$C = 1448.96 + 4.591T - 5.304 \times 10^{-2}T^2 + 2.374 \times 10^{-4}T^3 + 1.340(S - 35) + 1.630 \times 10^{-2}D + 1.675 \times 10^{-7}D^2 - 1.025 \times 10^{-2}T(S - 35) - 7.139 \times 10^{-13}TD^3 \quad (4)$$

여기서 C 는 음속(m/s), T 는 수온(°C), S 는 염분(psu), D 는 수심(m)을 의미한다. 수온과 염분의 예측 성능뿐만 아니라 최종적으로 산출되는 음속의 예측 성능을 함께 평가함으로써, 입력자료 구성과 예측모델이 실제 음속장 구축에 미치는 영향을 종합적으로 분석하였다.

3. 실험 결과 및 분석

3.1 기본입력자료를 이용한 예측 성능

기본입력자료를 이용하여 LR, RF, XGBoost 및 MLP의 예측 성능을 비교하였다. MLP는 기본입력자료에 대해 독립적으로 수행한 grid search 결과 중 음속 MAE가 가장 낮은 은닉층 7개, 층별 128개 노드 구조를 최종 모델로 선정하였다. 본 연구에서는 수온과 염분을 음속 산출의 입력자료로 활용하므로, 각 모델의 수온 및 염분 예측 성능과 함께 최종 음속 예측 성능을 종합적으로 비교하였다.

Table 2는 기본입력자료를 이용한 모델별 예측 성능을 나타낸다. 모든 비선형 모델은 LR보다 우수한 성능을 나타냈으며, 특히 MLP와 XGBoost가 전반적으로 높은 예측 정확도를 보였다. 수온 예측에서는 MLP가 MAE 1.5381°C, RMSE 2.2417°C 및 R^2 0.9058로 가장 우수한 성능을 나타냈으며, 염분 예측에서는 XGBoost와 MLP가 유사한 수준의 성능을 보였다. 이는 기본입력자료만으로도 비선형 모델이 동해 해양환경의 시·공간적 특성을 효과적으로 학습할 수 있음을 보여준다. 최종적으로 예측된 수온과 염

분을 Mackenzie 음속식에 적용하여 음속을 계산한 결과, MLP가 음속 MAE 5.4199 m/s, RMSE 7.9665 m/s 및 R^2 0.8905로 가장 우수한 성능을 나타냈다. 반면 LR은 모든 평가항목에서 가장 큰 오차를 보였으며, RF는 LR보다는 우수하였으나 XGBoost와 MLP에는 미치지 못하였다.

Table 2. Prediction performance of different models using basic inputs

Model	Metric	Temperature (°C)	Salinity (psu)	Sound Speed (m/s)
LR	MAE	3.6209	0.2502	12.8842
	RMSE	4.3792	0.4049	15.2292
	R^2	0.6404	0.2151	0.5999
RF	MAE	1.8628	0.1884	6.5085
	RMSE	2.8551	0.3421	10.0637
	R^2	0.8472	0.4399	0.8253
XG Boost	MAE	1.6757	0.1624	5.9665
	RMSE	2.3946	0.2834	8.5945
	R^2	0.8925	0.6157	0.8726
MLP	MAE	1.5381	0.1675	5.4199
	RMSE	2.2417	0.2856	7.9665
	R^2	0.9058	0.6096	0.8905

3.2 표층정보를 포함한 예측 성능

기본입력자료에 표층수온과 표층염분을 추가하여 LR, RF, XGBoost 및 MLP의 예측 성능을 비교하였다. MLP는 표층정보 데이터셋에 대해 독립적으로 수행한 grid search 결과 중 음속 MAE가 가장 낮은 은닉층 6개, 층별 128개 노드 구조를 최종 모델로 선정하였다. 기본입력자료 모델과 표층정보 모델은 입력 변수 구성이 서로 다르므로, 각 데이터셋에 대해 독립적으로 최적 신경망 구조를 선정하였다.

Table 3는 표층정보를 포함한 모델별 예측 성능을 나타낸다. 기본입력자료만을 이용한 경우와 마찬가지로 비선형 모델이 LR보다 우수한 성능을 보였으며, 표층정보를 추가함에 따라 전체적으로 예측 오차가 감소하는 경향을 나타냈다. 수온 예측에서는 MLP가 MAE 1.2973°C, RMSE 2.0778°C 및 R^2 0.9088로 가장 우수한 성능을 나타냈으며, 염분 예측에서는 XGBoost가 가장 높은 예측 정확도를 보였다. 이는 표층정보가 추가됨에 따라 수온과 염분 모두에서 실제 해양환경을 보다 효과적으로 반영할 수 있었음을 의

미한다. 예측된 수온과 염분을 이용하여 음속을 계산한 결과, MLP가 음속 MAE 4.7355 m/s, RMSE 7.6327 m/s 및 R^2 0.8910으로 가장 우수한 성능을 나타냈다. RF와 XGBoost 역시 기본입력자료를 이용한 경우보다 향상된 성능을 보였으며, MLP와의 성능 차이도 전반적으로 감소하였다. 이는 표층수온과 표층염분이 연직 구조를 설명하는 중요한 정보를 제공함으로써 모델의 복잡성보다 입력자료의 품질이 예측 성능에 더 큰 영향을 미쳤기 때문으로 판단된다.

Table 3. Prediction performance of different models using surface information

Model	Metric	Temperature (°C)	Salinity (psu)	Sound Speed (m/s)
LR	MAE	2.8641	0.2089	10.4675
	RMSE	3.4712	0.3226	12.5835
	R^2	0.7456	0.1981	0.7038
RF	MAE	1.3449	0.0905	4.8801
	RMSE	2.2219	0.1836	8.1566
	R^2	0.8957	0.7404	0.8756
XG Boost	MAE	1.3957	0.0890	5.1059
	RMSE	2.1451	0.1689	7.8674
	R^2	0.9028	0.7803	0.8842
MLP	MAE	1.2973	0.1121	4.7355
	RMSE	2.0778	0.1896	7.6327
	R^2	0.9088	0.7231	0.8910

3.3 표층정보 활용에 따른 성능 향상 분석

표층정보의 활용 효과를 정량적으로 분석하기 위해 기본입력자료와 표층정보 조건의 MAE를 비교하고, 오차 감소율을 다음 식과 같이 산출하였다.

$$\text{MAE 감소율 [\%]} = \frac{\text{MAE}_{\text{base}} - \text{MAE}_{\text{surface}}}{\text{MAE}_{\text{base}}} \quad (5)$$

Table 4는 각 모델의 수온, 염분 및 음속에 대한 MAE 감소율을 나타낸다. 표층정보를 활용한 경우 모든 모델에서 수온, 염분 및 음속의 예측 오차가 감소하였다. 특히 염분에서 가장 큰 성능 향상이 나타났으며, RF와 XGBoost는 기존 대비 약 50% 수준의 MAE 감소를 보였다. MLP 역시 수온과 염분의 예측 오차가 모두 감소하였으며, 최종 음속 MAE는 5.4199 m/s에서 4.7355 m/s로 약 12.6% 감소하였다. 이러

한 결과는 표층정보가 연직 해양환경 예측의 정확도를 향상시키는 데 효과적으로 활용될 수 있음을 보여준다.

Table 4. MAE reduction (%) achieved by using surface information

Model	MAE Reduction (%)		
	Temperature (°C)	Salinity (psu)	Sound Speed (m/s)
LR	20.90	16.49	18.76
RF	27.80	52.00	25.02
XGBoost	16.71	45.22	14.42
MLP	15.66	33.09	12.63

표층정보를 추가하였을 때 성능이 향상된 이유는 예측 시점의 실제 해양환경을 입력자료에 직접 반영할 수 있기 때문으로 판단된다. 기본입력자료는 위치, 시간 및 수심 정보를 이용하여 장기간 축적된 해양환경의 시·공간적 특성을 학습하는 반면, 동일한 위치와 계절이라도 해마다 달라질 수 있는 실제 해양상태는 충분히 반영하기 어렵다. 반면 표층수온과 표층염분은 위성 원격탐사를 통해 예측 시점에서도 확보 가능한 정보로, 해당 시점의 해양환경을 직접 반영할 수 있다. 따라서 표층정보를 함께 활용함으로써 위치와 시간 정보만으로 설명하기 어려운 연직 수온 및 염분 구조를 보다 정확하게 추정할 수 있었던 것으로 판단된다.

또한, 기본입력자료만을 이용한 경우에는 모델 간 성능 차이가 비교적 크게 나타났으나, 표층정보를 추가한 이후에는 RF, XGBoost 및 MLP의 성능 차이가 전반적으로 감소하였다. 이는 표층정보 자체가 연직 구조를 설명하는 중요한 정보를 포함하고 있어 모델의 복잡성보다 입력자료의 품질이 예측 성능에 더 큰 영향을 미쳤기 때문으로 해석된다. 즉, 표층정보는 특정 모델에만 효과적인 입력변수가 아니라 다양한 기계학습 모델에서 공통적으로 활용 가능한 유용한 입력자료임을 확인할 수 있었다.

이러한 결과는 표층정보 활용이 개별 변수인 수온과 염분의 예측 정확도 향상에 그치지 않고, 최종 활용목적인 음속장 구축의 정확도 향상에도 직접적으로 기여할 수 있음을 의미한다. 따라서 본 연구에서 제안한 표층정보 활용 방법은 향후 동해 음속장 구축과 수중 음향환경 예측을 위한 효과적인 입력자료 구

성 방법으로 활용될 수 있을 것으로 판단된다.

4. 결론

해양환경은 해군의 수중작전 수행과 음향체계 운용에 직접적인 영향을 미치는 핵심 환경요소이다. 그러나 현장 해양관측은 시·공간적 제약으로 인해 모든 해역의 해양환경정보를 필요한 시기에 확보하는 데 한계가 있다. 따라서 제한된 관측자료를 활용하여 관측되지 않은 시·공간의 해양환경을 추정하고 이를 기반으로 음속장을 산출하는 기술이 중요하다. 이에 본 연구에서는 장기간 축적된 KODC 정선관측자료를 이용하여 머신러닝 기반의 해양환경 예측모델을 구축하고, 시공간 및 수심 정보를 이용한 수온·염분 예측과 함께 표층수온 및 표층염분 정보가 연직 해양환경과 음속 예측에 미치는 영향을 분석함으로써 표층정보 활용의 가능성을 검증하였다.

기본입력자료와 표층정보를 포함한 입력자료를 각각 이용하여 LR, RF, XGBoost 및 MLP 모델의 성능을 비교한 결과, 모든 입력조건에서 MLP가 가장 우수한 음속 예측 성능을 나타내었다. 또한, 표층수온과 표층염분을 입력자료에 추가한 경우 모든 모델에서 예측 정확도가 향상되었으며, 특히 모든 입력조건에서 가장 우수한 음속 예측 성능을 나타낸 MLP의 음속 MAE는 5.4199 m/s에서 4.7355 m/s로 약 12.6% 감소하였다. 이러한 결과는 표층정보가 연직 수온 및 염분 구조를 설명하는 중요한 입력자료이며, 음속장 예측 정확도 향상에도 효과적으로 활용될 수 있음을 확인하였다.

본 연구의 의의는 단순히 머신러닝 모델의 예측 성능을 비교하는 데 있는 것이 아니라, 표층수온과 표층염분이 연직 해양환경을 설명하는 핵심 정보로 활용될 수 있음을 정량적으로 확인하였다는 점에 있다. 이는 GHRSSST 및 SMAP과 같은 위성 원격탐사 자료를 활용하여 직접 해양관측이 어려운 시·공간에서도 연직 수온·염분 구조와 음속장을 보다 정확하게 추정할 수 있는 가능성을 제시한다. 또한, 장기간 축적된 해양관측자료와 머신러닝을 결합하여 4차원 해양환경을 복원하기 위한 기반기술을 제시하였다는 점에서도 의미가 있다.

본 연구에서 제안한 방법은 해군의 대잠전 환경분석, 음향전파 예측, 작전 해역 환경평가 및 해양환경

정보 지원체계 구축 등 다양한 분야에 활용될 수 있을 것으로 기대된다. 특히 직접 관측이 제한되는 해역에서도 위성 기반 표층정보를 활용하여 연직 음속장을 추정할 수 있으므로 해양환경정보 생산의 효율성과 활용성을 동시에 향상시킬 수 있을 것으로 판단된다. 향후 연구에서는 시간 예측을 넘어 관측 정점 간 공간 내삽, 수심 방향 보간, 항차 단위 해양장 복원 등 다양한 해양환경 재구성 문제로 연구를 확장하여 보다 완전한 4차원 해양환경 복원 기술을 구축하고자 한다. 또한, 실제 GHR SST 및 SMAP 위성 자료를 적용하여 관측 공백 해역의 해양환경 및 음속장을 추정하고, 이를 활용한 수중 음향환경 예측으로 연구를 확장할 계획이다.

참고문헌

- [1] Young-Joo Kim, Soojin Lee, & Young-won Kim, 'Prediction of Vertical Sea Water Temperature Profile in the East Sea Based on Machine Learning and XBT Data,' Journal of the Korea Society of Computer and Information, VOL. 27, NO. 11, 2022, pp. 47-55.
- [2] Kwan-Seob Ko, Seong-Hyeon Byeon, & Young-Won Kim, 'Prediction of Sea Water Temperature by Using Deep Learning Technology Based on Ocean Buoy,' Korean Journal of Remote Sensing, VOL. 38, NO. 3, 2022, pp. 299-309.
- [3] GHR SST Science Team, Group for High Resolution Sea Surface Temperature (GHR SST). <https://www.ghrsst.org>
- [4] NASA Jet Propulsion Laboratory, Soil Moisture Active Passive (SMAP) Sea Surface Salinity. <https://podaac.jpl.nasa.gov/SMAP> (accessed 2026.07.22.)
- [5] 해양자료센터(KODC), 정션해양조사자료. <https://www.nifs.go.kr/kodc/observe/line/data> (accessed 2026.07.22.)
- [6] Leo Breiman, 'Random Forests,' Machine Learning, VOL. 45, 2001, pp. 5-32.
- [7] Tianqi Chen & Carlos Guestrin, 'XGBoost: A Scalable Tree Boosting System,' in proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785-794.
- [8] David E. Rumelhart, Geoffrey E. Hinton, & Ronald J. Williams, 'Learning Representations by Back-Propagating Errors,' Nature, VOL. 323, 1986, pp. 533-536.
- [9] Kenneth V. Mackenzie, 'Nine-Term Equation for Sound Speed in the Oceans,' Journal of the Acoustical Society of America, VOL. 70, NO. 3, 1981, pp. 807-812.