



*Supplement of*

## **A continual-learning-based multilayer perceptron for improved reconstruction of three-dimensional nitrate concentrations**

**Xiang Yu et al.**

*Correspondence to:* Jiahua Zhang (zhangjh@radi.ac.cn)

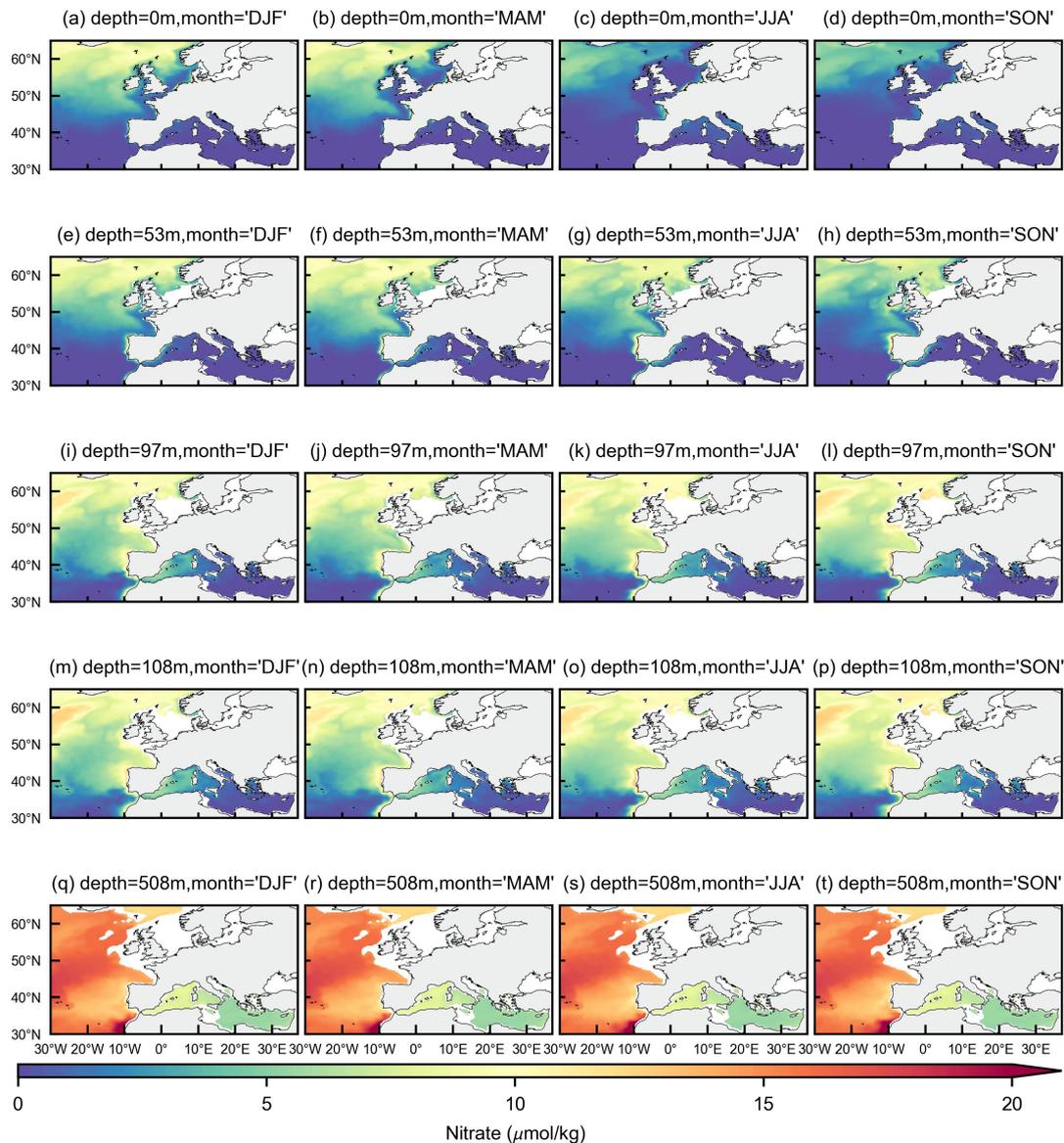
The copyright of individual parts of the supplement might differ from the article licence.

# *Supplement*

## **Table of Contents**

- Figure S1: Spatial distribution of CMEMS simulated nitrate field, with columns representing four seasons and rows representing five depth slices.
- Figure S2: Interannual trends in nitrate from multi-source datasets.
- 5 – Figure S3: Scatter plot of the model's test set performance after excluding six typical features. The red line indicates the fitted trend of the data, while the black dashed line denotes the 1:1 parity line.
- Table S1: Model's 5-fold profile-based cross-validation performance after sequentially excluding variables, with variables sorted by the increase in RMSE.

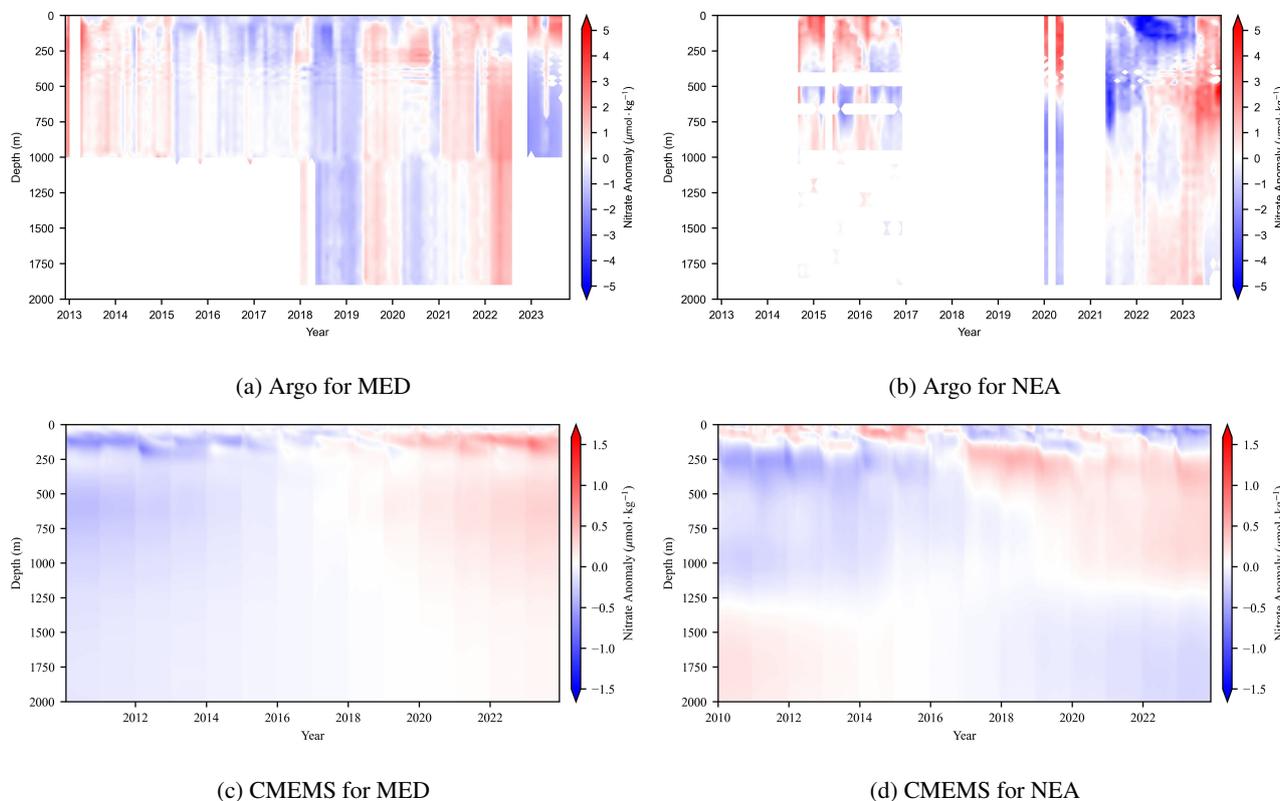
Figure S1 provides a comprehensive view of the seasonal variability in nitrate fields across various depth slices, based on 10 CMEMS dataset. This highlights distinct spatial patterns, aiding the interpretation of regional nitrate variability. Our reconstructed results (Figure 9) exhibit strong spatial consistency with CMEMS nitrate data, including in regions not covered by BGC-Argo, such as the coastal areas of the UK and Norway.



**Figure S1.** Spatial distribution of CMEMS simulated nitrate field, with columns representing four seasons and rows representing five depth slices.

Figure S2 offers a comparative analysis of interannual nitrate anomalies derived from BGC-Argo observations and CMEMS model outputs across the Mediterranean (MED) and the Northeast Atlantic (NEA), emphasizing consistency and discrepancies in temporal trends.

The most comprehensive and promising datasets for describing three-dimensional nitrate distributions are the BGC-Argo and CMEMS nitrate datasets, which we employed in our study, representing observational and modeled approaches, respectively. However, both datasets have significant limitations when depicting interannual trends. Due to the varying geographical locations of BGC-Argo observations over time, regional differences in nutrient levels introduce considerable interference and fluctuations in interannual trend calculations. As a result, BGC-Argo trends tend to be more pronounced, and in some cases, shifting sampling locations across different nutrient regimes may even lead to trend reversals. Conversely, CMEMS nitrate data exhibit a delayed response in capturing mesoscale nitrate variability. Therefore, the overly uniform trend patterns observed in Figure S2b&d are likely more conservative than actual conditions, yet they still provide a useful reference for decadal-scale trends.

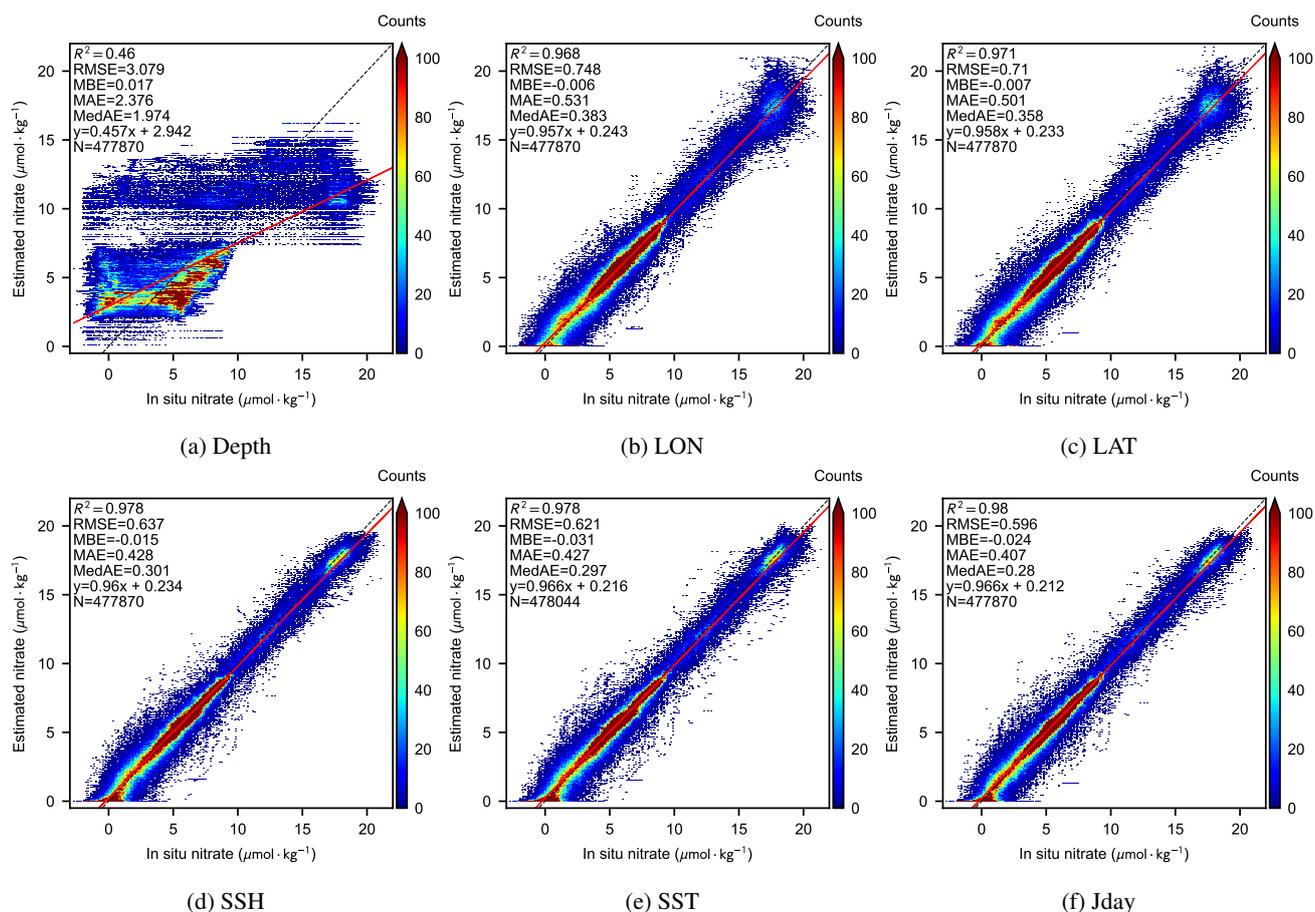


**Figure S2.** Interannual trends in nitrate from multi-source datasets

Figure S3 illustrates scatter plots highlighting the model's performance when excluding key predictive features such as depth, longitude, sea surface height (SSH), and sea surface temperature (SST).

Table S1 provides detailed insights into the 5-fold profile-based cross-validation results after sequentially excluding each variable. This table helps identify critical variables based on how their absence impacts the model's Root Mean Square Error (RMSE), emphasizing the significant role of features like depth.

30 We systematically excluded variables and compared the cross-validation simulation results based on profiles, as shown in Table S1. Table S1 is sorted by the increase RMSE after the exclusion of each variable, with RMSE being the most crucial metric as it also serves as the loss function supervising the model's training. The RMSE increased to varying degrees after the exclusion of variables, leading to the inference that the initial feature combination is one of the most optimal. Depth, latitude, and longitude, the three spatial coordinates, remain the most significant variables, with the highest RMSE increase when excluded. In particular, depth is used not only as a feature, but more importantly to map depth profiles, so exclusions will show exaggerated variation.



**Figure S3.** Scatter plot of the model's test set performance after excluding six typical features. The red line indicates the fitted trend of the data, while the black dashed line denotes the 1:1 parity line.

**Table S1.** Model’s 5-fold profile-based cross-validation performance after sequentially excluding variables, with variables sorted by the increase in RMSE.

Excluded features	$R^2$	RMSE	MBE	MAE	MedAE	Slope k	Intercept b	RMSE increase ratio
Original feature set	0.98	0.592	-0.04	0.398	0.265	0.967	0.222	
Depth	0.46	3.079	0.017	2.376	1.974	0.457	2.942	420.10%
Lat	0.973	0.686	0.003	0.48	0.342	0.954	0.245	15.88%
Lon	0.977	0.641	0.002	0.443	0.309	0.955	0.245	8.28%
SSH	0.978	0.637	-0.015	0.428	0.301	0.96	0.234	7.60%
ZSD	0.977	0.632	-0.016	0.442	0.314	0.958	0.246	6.76%
Z	0.977	0.63	-0.018	0.439	0.311	0.957	0.251	6.42%
ZEU	0.978	0.629	-0.013	0.438	0.309	0.957	0.248	6.25%
SST	0.978	0.629	-0.031	0.427	0.297	0.966	0.216	6.25%
CDM	0.978	0.628	-0.01	0.437	0.308	0.958	0.24	6.08%
ZHL	0.978	0.627	-0.007	0.438	0.31	0.957	0.241	5.91%
TP	0.978	0.625	-0.009	0.435	0.308	0.958	0.237	5.57%
V10	0.978	0.623	-0.016	0.434	0.308	0.957	0.249	5.24%
S10	0.979	0.622	-0.016	0.432	0.305	0.958	0.247	5.10%
U10	0.978	0.622	-0.017	0.432	0.303	0.958	0.248	5.07%
CF	0.978	0.621	0.002	0.431	0.304	0.96	0.218	4.90%
SP	0.978	0.621	-0.009	0.432	0.304	0.959	0.234	4.90%
Chl	0.978	0.617	-0.004	0.426	0.296	0.96	0.222	4.22%
MLD	0.978	0.616	-0.02	0.424	0.295	0.958	0.251	4.05%
NFLH	0.979	0.61	-0.023	0.419	0.289	0.961	0.237	3.04%
PAR	0.979	0.609	-0.024	0.419	0.29	0.961	0.239	2.87%
Jday	0.98	0.596	-0.024	0.407	0.28	0.966	0.212	0.68%