Supporting information

Data-driven modeling of dissolved iron in the global ocean

Yibin Huang[1], Alessandro Tagliabue[2], Nicolas Cassar[1,3]

[1]Division of Earth and Climate Sciences, Nicholas School of the Environment, Duke University, Durham, USA
[2]School of Environmental Sciences, University of Liverpool, Liverpool, UK
[3] CNRS, Université de Brest, IRD, Ifremer, LEMAR, Plouzané, France

Correspondence: Nicolas Cassar (Nicolas.Cassar@duke.edu)

Contents of this file

Table S1, Figure S1 to Figure S3.

Introduction

Table S1 provides the Pearson correlation coefficients between field collected dissolved Fe and environmental predictors at the different depth intervals. Figure S1 shows the performance of the depth-resolved model. Figure S2 shows the rank of predictor importance in the model construction by random forest. Figure S3 shows how ocean basins were defined.

Table S1. Pearson correlation coefficients between field collected dissolved Fe and environmental predictors at the different depth intervals. Sampling location and time coordinates are calculated according to Equations 1 and 2. POC: particulate organic carbon; NPP: net primary production; AOU: apparent oxygen utilization; N*=[NO_3^--PO_4^{3-}+2.90]*0.87; Si*=SiO_4^{2-}- NO_3^-.

Predictors	Surface (0-140 m)	Subsurface (260-340 m)	Intermediate layer (700-1200 m)	Deep depth (3200-3500 m)
Depth	0.01	*NA*	0.00	*NA*
Bottom depth	-0.16	-0.15	0.02	-0.11
Mixed layer depth	-0.13	-0.27	-0.31	-0.12
Salinity	0.20	0.13	0.14	-0.08
Temperature	0.15	0.10	0.06	0.21
O_2	-0.33	-0.61	-0.47	-0.26
NO_3^-	0.00	0.43	0.31	0.12
PO_4^{3-}	-0.01	0.41	0.28	0.13
SiO_4^{2-}	0.02	0.41	0.23	0.09
Distance to the coast	-0.24	-0.33	-0.27	0.10
Surface chlorophyll-*a*	0.14	0.36	0.20	-0.05
Surface Rrs(555)	-0.18	-0.16	-0.16	-0.18
Surface Rrs(488)	-0.32	-0.54	-0.37	-0.06
Surface POC	0.16	0.36	0.17	-0.08
Surface Rrs(442)	-0.28	-0.51	-0.32	-0.01
Surface Rrs(412)	-0.29	-0.52	-0.33	0.01
Surface Rrs(667)	-0.14	-0.06	-0.16	-0.16
Density	-0.14	0.10	0.18	0.18
10m wind speed	-0.09	-0.06	-0.11	-0.06
NPP	0.16	0.38	0.30	0.06
Aerosol optical depth	0.37	0.54	0.44	0.14
AOU	0.25	0.62	0.50	0.21
N*	0.03	0.45	0.32	0.13
S*	0.03	0.45	0.33	0.14
Distance to the bottom	-0.17	-0.15	-0.02	-0.12
Helium	0.18	0.41	0.28	0.30
Location coordiate_1	0.03	-0.08	0.01	0.28
Location coordiate_2	-0.07	0.11	0.09	0.28
Location coordiate_3	0.06	-0.11	-0.08	*NA*
Time coordiate_1	0.05	-0.02	-0.07	0.11
Time coordiate_2	-0.16	-0.29	-0.29	-0.40

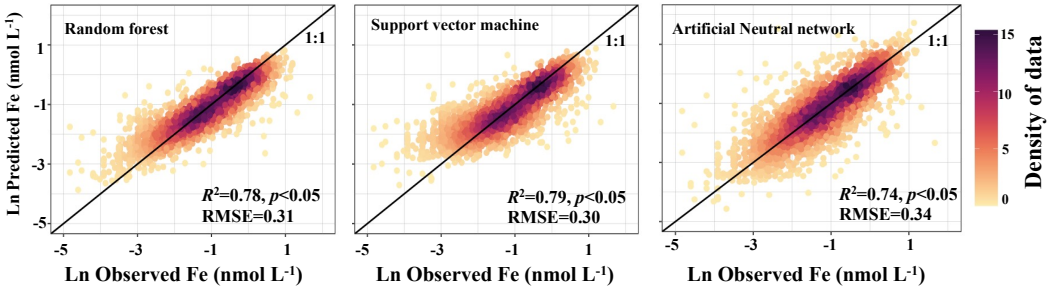

Figure S1. Performance of depth-resolved model evaluated by the validation dataset. The depth-resolved models were constructed based on datasets from different depth intervals (0-200 m, 200-1000 m, 1000-2500 m, >3200 m).

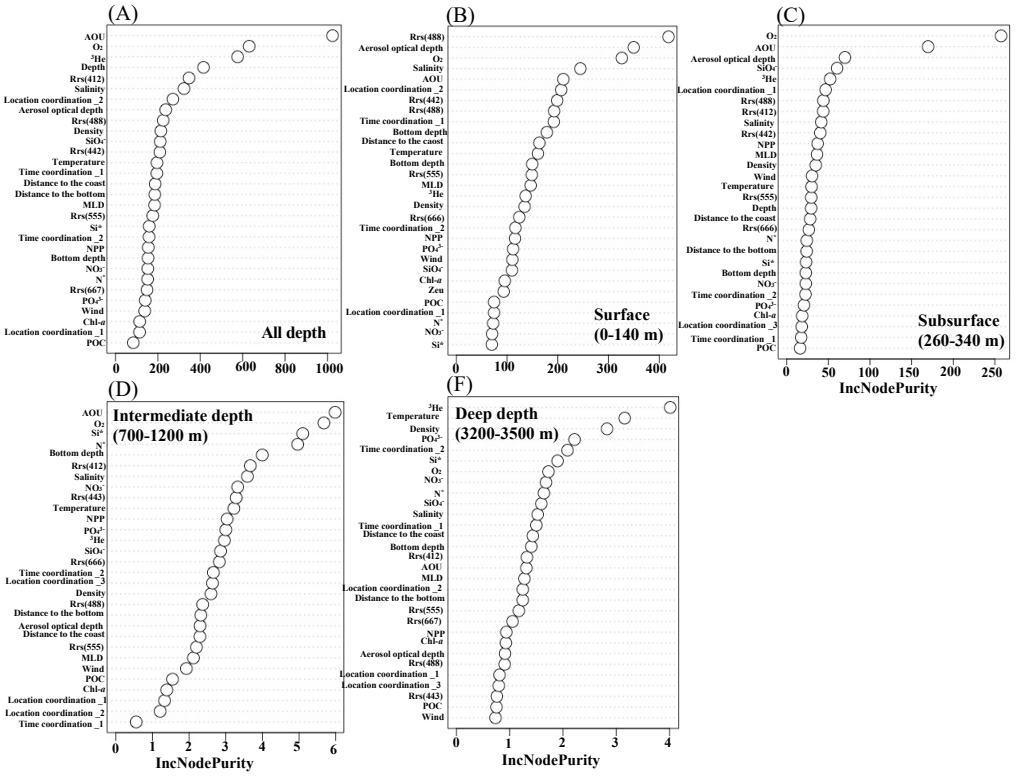

Figure S2. Ranks of predictor importance in the model construction by random forest (top 30). Sampling location and time coordinates are calculated according to Equations 1 and 2. POC: particulate organic carbon; NPP: net primary production; AOU: apparent oxygen utilization; $N^* = [NO_3^- - PO_4^{3-} + 2.90] * 0.87$; $Si^* = SiO_4^{2-} - NO_3^-$.

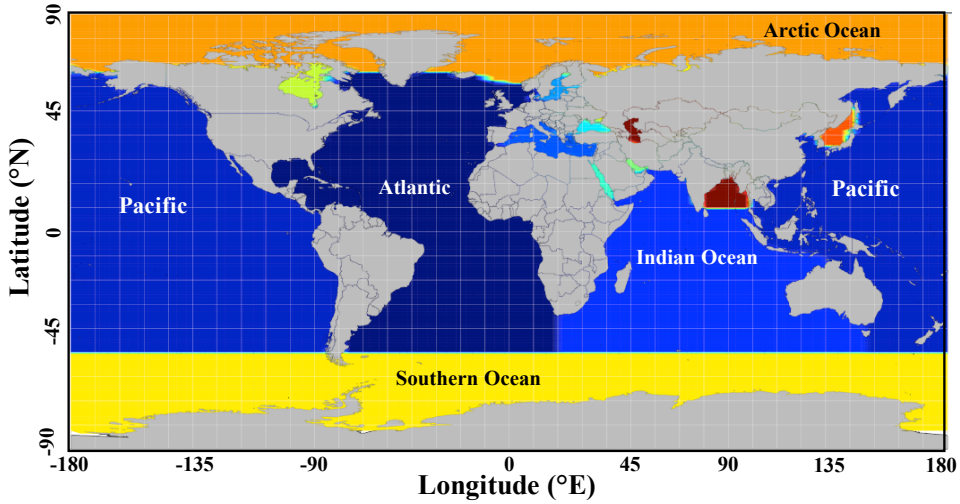

Figure S3. Ocean basins as defined in our study based on the World Ocean Atlas (2005).