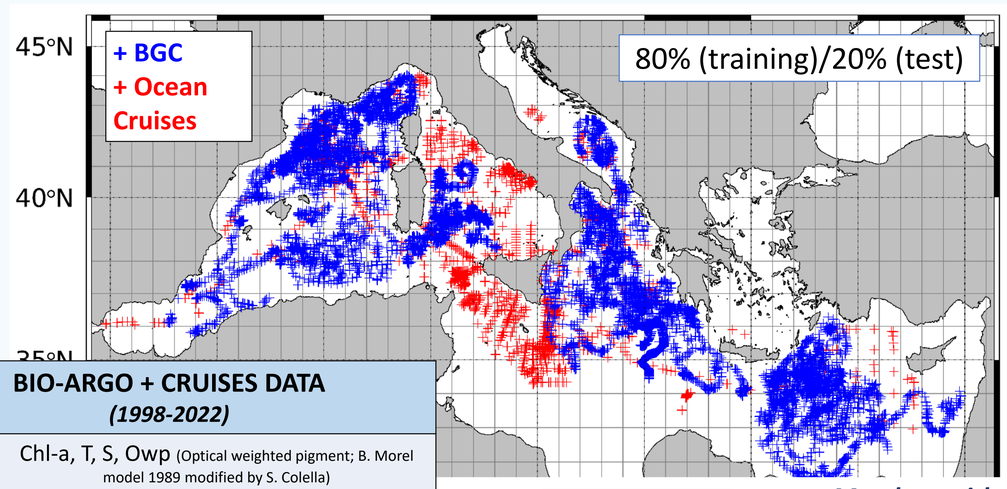
Advancing Ocean Insights: AI-Driven 4D joint Reconstruction of Physical and Biological fields in the Mediterranean Sea within the ESA 4DMED-Sea Project

Sammartino M.[1], Colella S.[2], Della Cioppa L.[1] and Buongiorno Nardelli B.[1]

[1]Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine, (CNR-ISMAR), 80133, Naples, IT
[2] Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine, (CNR-ISMAR), 00133, Rome, IT

INTRODUCTION | *At the sea surface, satellite data provide synoptic observations of multiple essential variables with high spatio-temporal resolution. However, satellite cannot observe deeper ocean layers, where data acquisition relies entirely on accurate but sparse in situ sensor measurements. Within the 4DMED-Sea project we developed a data-driven approach to combine surface observations and vertical profiles. We implemented an algorithm based on artificial intelligence (AI) to project surface values at depth and provide a 4D-reconstruction of key physical and biological variables (temperature T, salinity S, density D and chlorophyll-a Chl). Then, the reconstructed 4D physical tracers have been integrated with surface geostrophic currents to derive the 4D geostrophic velocities. 4D Chl, T and atmospheric-ocean model outputs were finally used to estimate 4D primary production (PP) fields. Scientific analysis of 4D data is ongoing.*

DATA | In situ and Satellite data



BIO-ARGO + CRUISES DATA (1998-2022)	
#Var	Chl-a, T, S, Owp (Optical weighted pigment; B. Morel model 1989 modified by S. Colella)
#Casts	14000
depth	3-150 m

→ *#Tot_Profiles 8624* → *Matchup with altimetric satellite data*

METHODS.1 | Deep Neural Network model (DNN)

INPUT
• Day of year (sin,cos)
• Latitude
• Longitude
• SSS/S(0)
• SSD/D(0)
• SST/T(0)
• CHL/OWP
• ADT
• Ugos
• Vgos

CONV1D layer
↓
LSTM layer
Dropout
LSTM layer
Dropout
CONV1D layer
↓
TimeDistributed (Dense) layer

OUTPUT
T profile
S profile
Chl profile
D profile

METHODS.2 | Physically-informed DNN

The total loss function was a sum of three different customized loss functions:

1.
$$Loss_{depth} = w_{surf}\frac{1}{N}\sum_{n=1}^{N}(y_{true}[z_1:z_{10}] - y_{predicted}[z_1:z_{10}])^2 + w_{profile}\frac{1}{N}\sum_{n=1}^{N}(y_{true} - y_{predicted})^2 + w_{deep}\frac{1}{N}\sum_{n=1}^{N}(y_{true}[z_{141}:z_{150}] - y_{predicted}[z_{141}:z_{150}])^2$$

- different weighting factors to prioritize specific regions of the profile to reduce edge effects

2.
$$Loss_{eq.state} = \frac{1}{N}\sum_{n=1}^{N}(\sigma_{true} - \sigma(T_{pred}, S_{pred}))^2$$

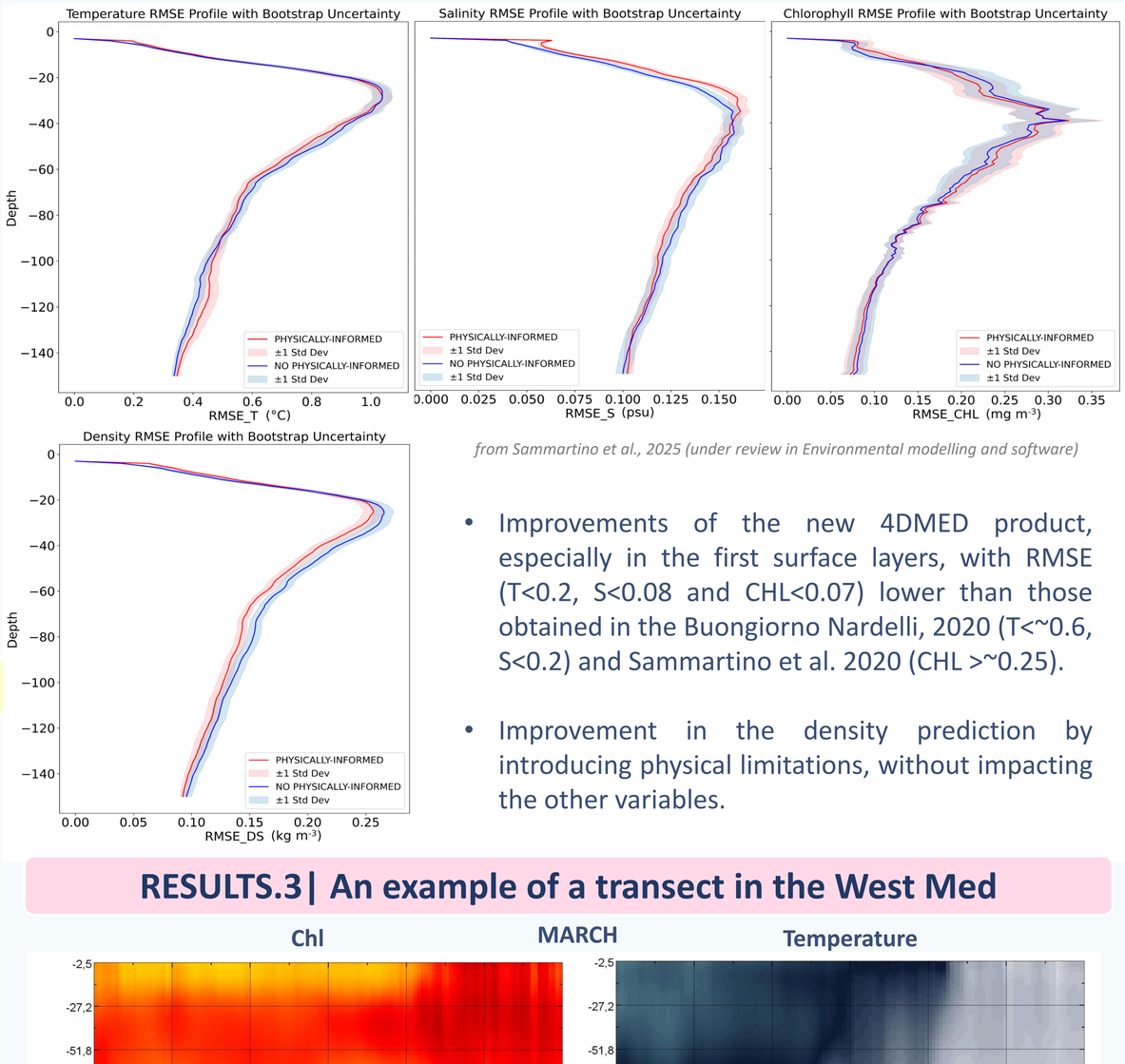
- enforcing consistency between density estimated from predicted T/S (through the **equation of state**) and true density profiles

3.
$$Loss_{inv} = \frac{1}{N}\sum_{n=1}^{N} RELU\left(\frac{\Delta\sigma pred}{\Delta z}\right)$$

- constraining the monotony of the density profile by penalizing the predictions estimated from predicted T/S that contain **density inversions**

from Sammartino et al., 2025 (under review in Environmental modelling and software)

RESULTS.1| Statistical Results

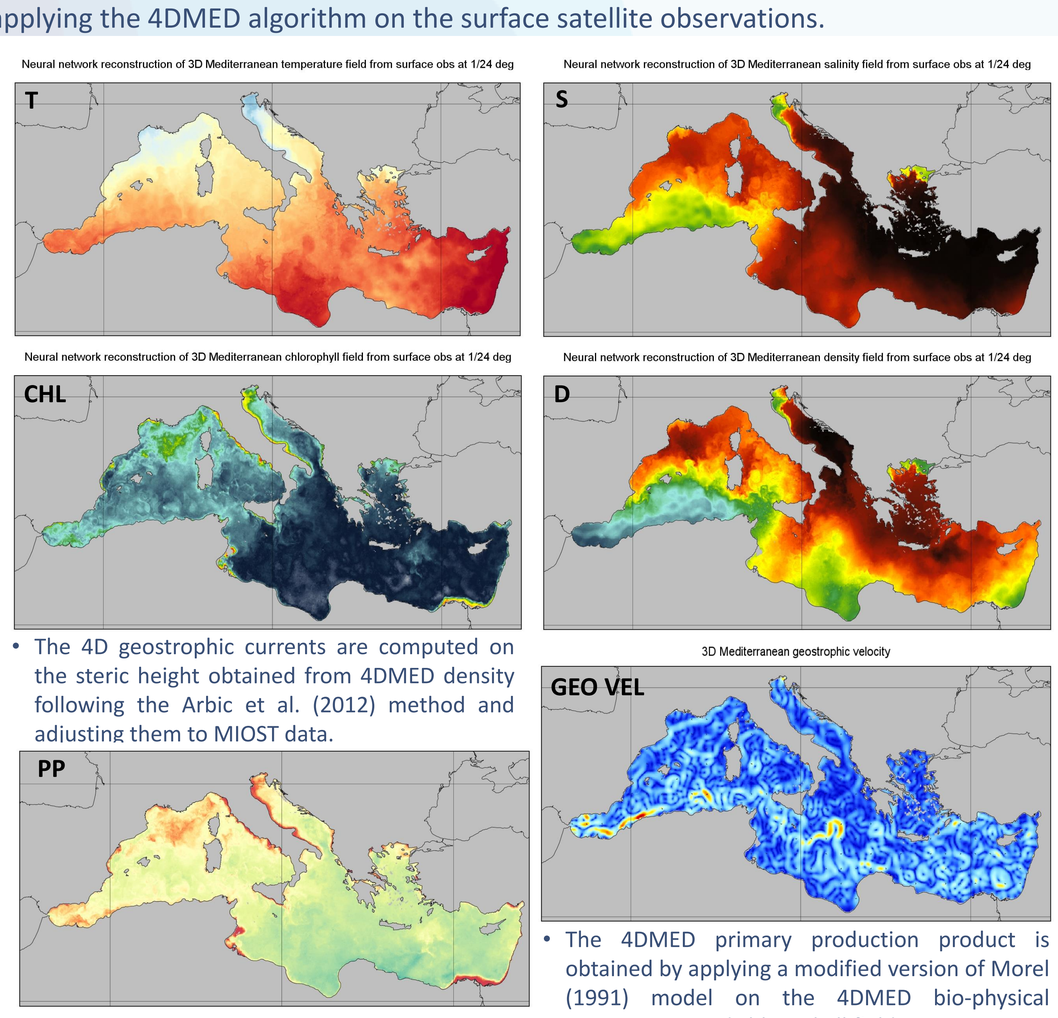


from Sammartino et al., 2025 (under review in Environmental modelling and software)

- Improvements of the new 4DMED product, especially in the first surface layers, with RMSE (T<0.2, S<0.08 and CHL<0.07) lower than those obtained in the Buongiorno Nardelli, 2020 (T<~0.6, S<0.2) and Sammartino et al. 2020 (CHL >~0.25).
- Improvement in the density prediction by introducing physical limitations, without impacting the other variables.

RESULTS.2| Generation of the 4DMED-SEA bio-physical fields

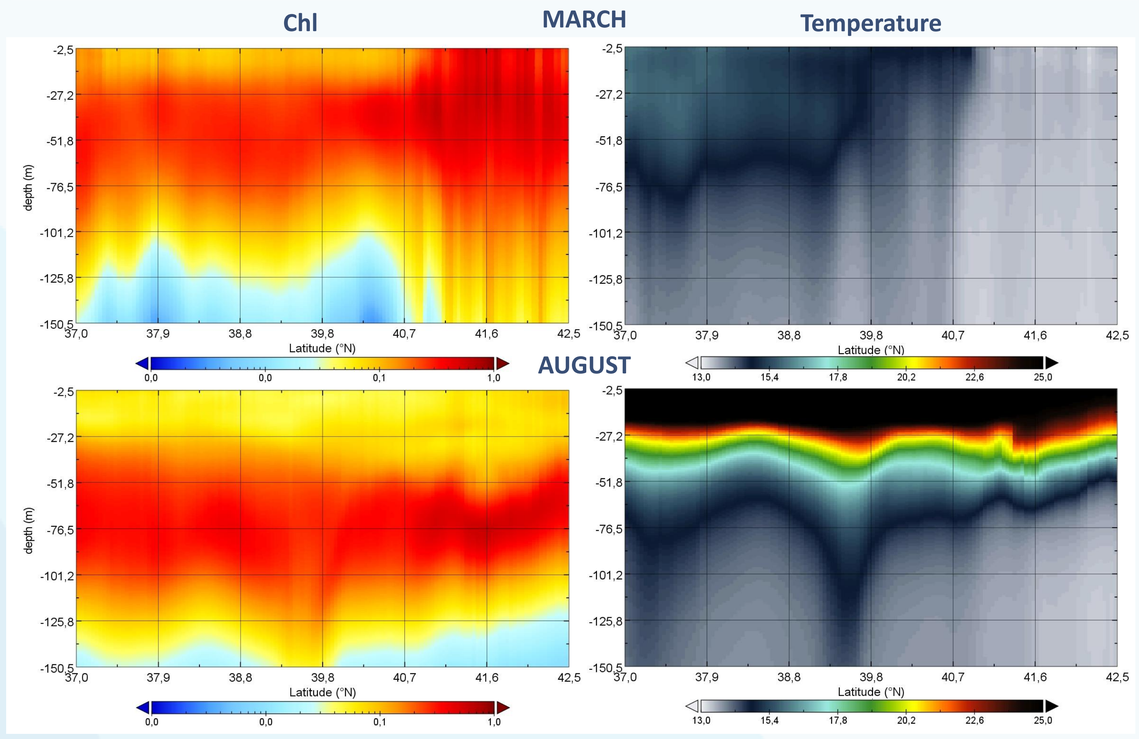
- The final combined physical-biological 4DMED experimental product are obtained by applying the 4DMED algorithm on the surface satellite observations.



- The 4D geostrophic currents are computed on the steric height obtained from 4DMED density following the Arbic et al. (2012) method and adjusting them to MIOST data.

- The 4DMED primary production product is obtained by applying a modified version of Morel (1991) model on the 4DMED bio-physical temperature and chlorophyll fields.

RESULTS.3| An example of a transect in the West Med



The final 4DMED-Sea database is freely distributed:
https://doi.org/10.25423/CMCC/4DMEDSEA_BIOPHYS_REP_3D

CONCLUSION & FUTURE PERSPECTIVES|

- **Data combination**: better coverage and understanding of ocean variability, which would be difficult to achieve using just one data type
- **AI Model Development and Refinement**: further adjustments to improve accuracy and efficiency. Expanding to incorporate/retrieve additional variables (e.g., dissolved oxygen).
- **Broader Geographical Application**: Future applications could involve expanding this model to other regions.

REFERENCES|
- Arbic, B. K., Scott, R. B., Chelton, D. B., Richman, J. G., and Shriver, J. F.: Effects of stencil width on surface ocean geostrophic velocity and vorticity estimation from gridded satellite altimeter data: OCEAN GEOSTROPHIC VELOCITY ESTIMATION, J. Geophys. Res. Oceans, 117, https://doi.org/10.1029/2011JC007367, 2012.
- Buongiorno Nardelli, B. A Deep Learning Network to Retrieve Ocean Hydrographic Profiles from Combined Satellite and In Situ Measurements. Remote Sens. 2020, 12, 3151.
- Morel, A. (1991). Light and marine photosynthesis: a spectral model with geochemical and climatological implications. Progress in oceanography, 26(3), 263-306.
- Sammartino, M., B. B. Nardelli, S. Marullo, and R. Santoleri (2020), An artificial neural network to infer the Mediterranean 3D chlorophyll-a and temperature fields from remote sensing observations, Remote Sens., 12(24), 1–27, doi:10.3390/rs12244123.

ACKNOWLEDGEMENTS| This work has been carried out as part of and the ESA-4DMED-Sea project (European Space Agency – 4DMED-Sea, ESA Contract No. 4000141547/23/I-DT).