A Deep Feed-Forward Neural Network to reconstruct the Mediterranean 3D chlorophyll-a and temperature fields from satellite measurements

Michela Sammartino[1], Bruno Buongiorno Nardelli[2], Salvatore Marullo[3,1] Rosalia Santoleri[1]

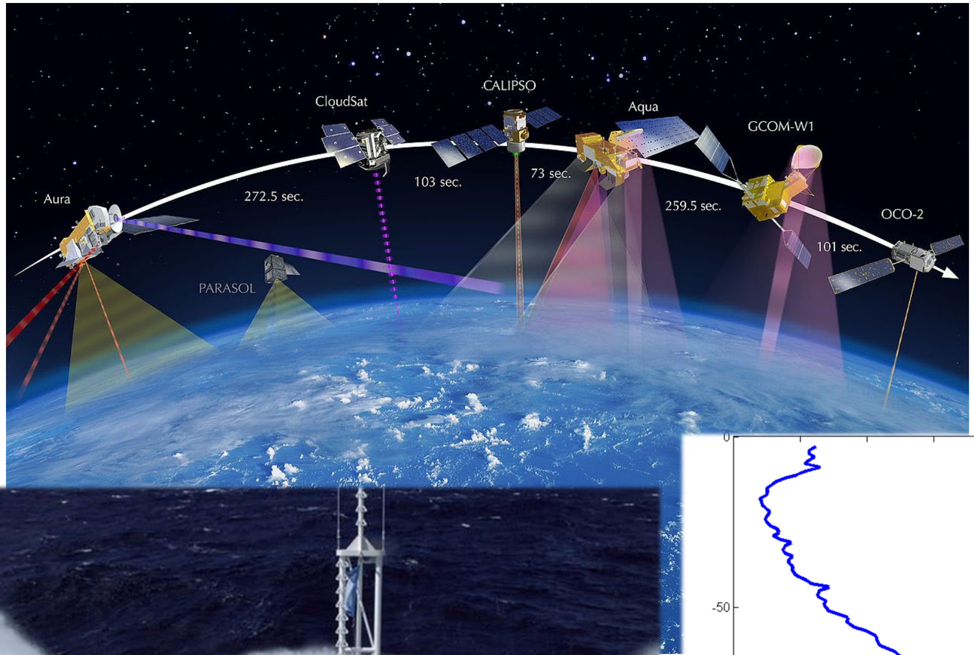
[1] Institute of Marine Sciences - National Research Council (ISMAR-CNR), Rome, Italy
[2] [1]Institute of Marine Sciences - National Research Council (ISMAR-CNR), Naples, Italy
[3] ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Frascati, Italy

(michela.sammartino@artov.ismar.cnr.it)

ECMWF–ESA Workshop on Machine Learning for Earth Observation and Prediction 14-17 NOVEMBER 2022

...at surface

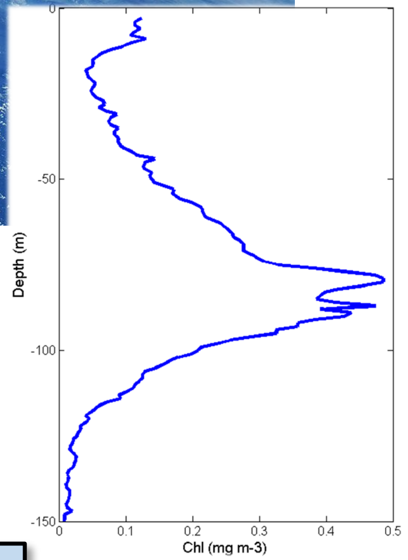
Satellite provides data at high resolution with synoptic coverage and high repetitiveness but cannot sense the ocean deep layers



...along the vertical component

In situ data provide accurate information but with:

- Practical difficulties
- Punctual data/discrete samples
- Time consuming/economic effort
- Low coverage

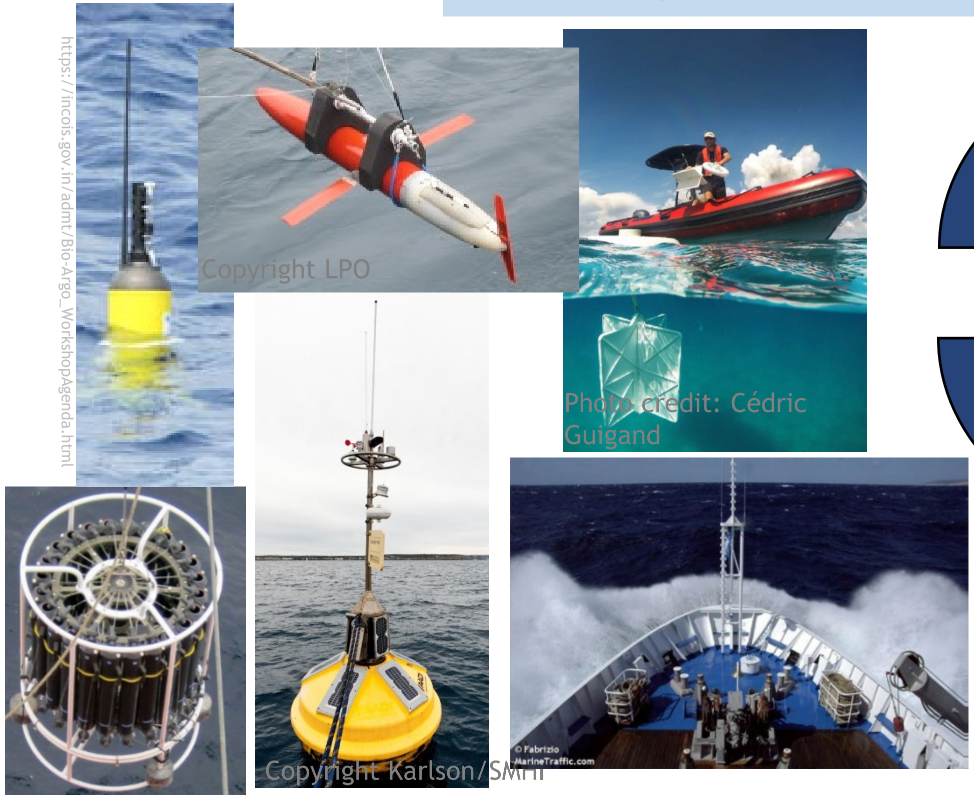


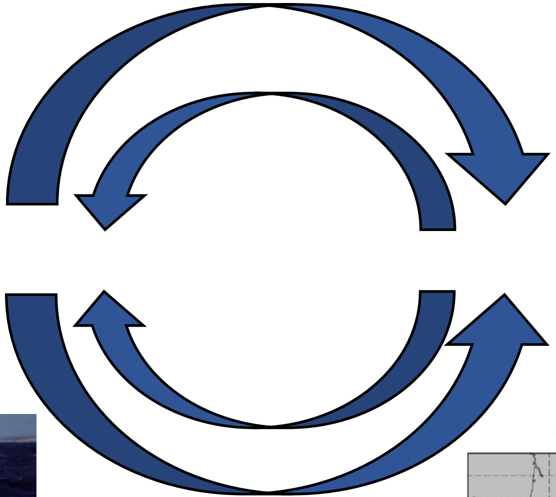
3D ocean interior structure remain undersampled

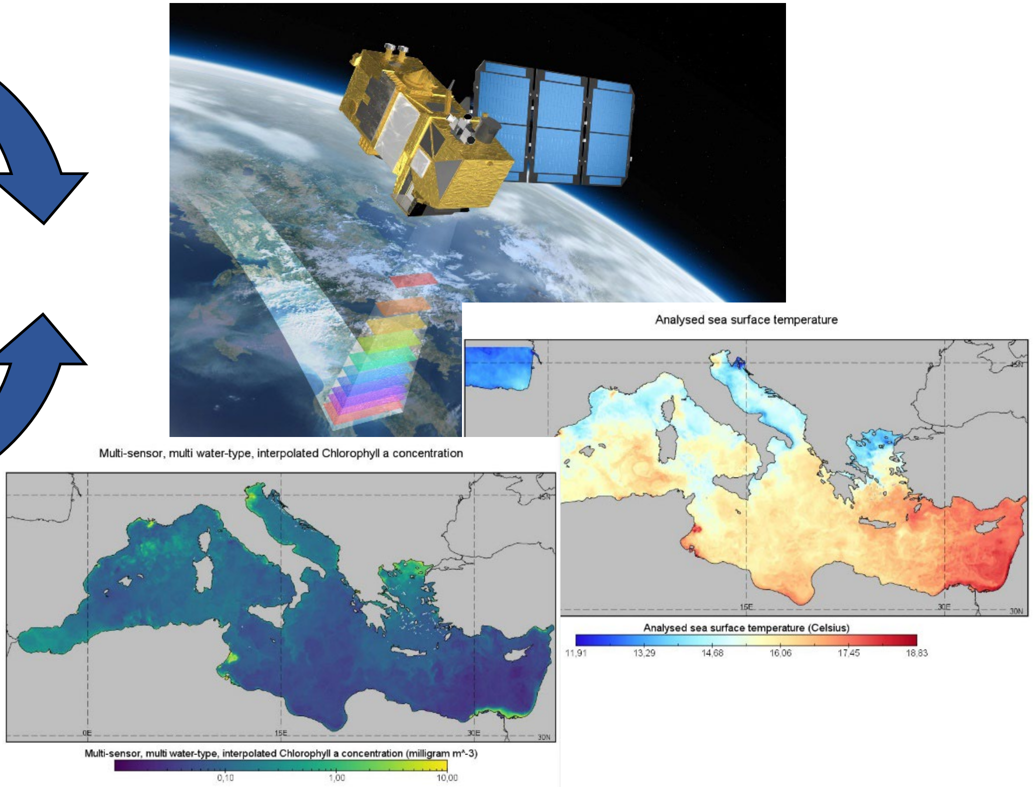
⇩

Information GAP between 2D and 3D!

Synergic use of data coming from different platforms







Integrated with new modelling approaches to improve the ocean observing system

Aim of the work

Reconstruction of the 3D chlorophyll and temperature fields from satellite data through the application of a Multi-Layer Perceptron (MLP)

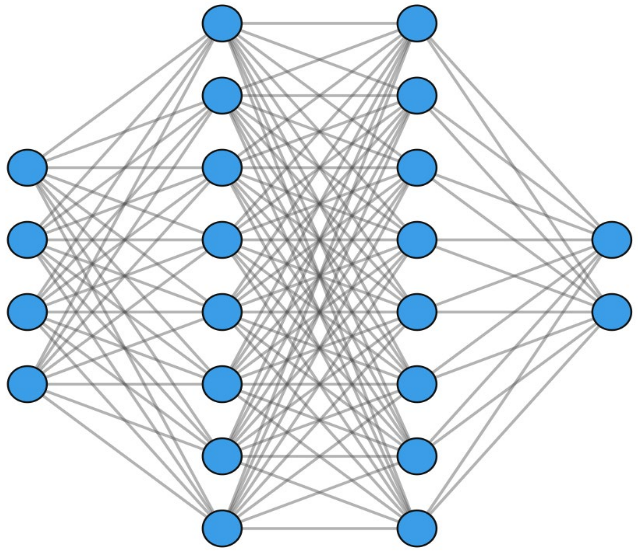
- To exploit the resolution of **satellite imagery** to project chlorophyll and temperature surface data at deeper layers.
- To overcome the **discontinuous nature of *in situ* datasets**.
- To demonstrate one of the possible applications of **artificial intelligence** to the ocean data for an improved **ocean monitoring**.

Machine Learning model

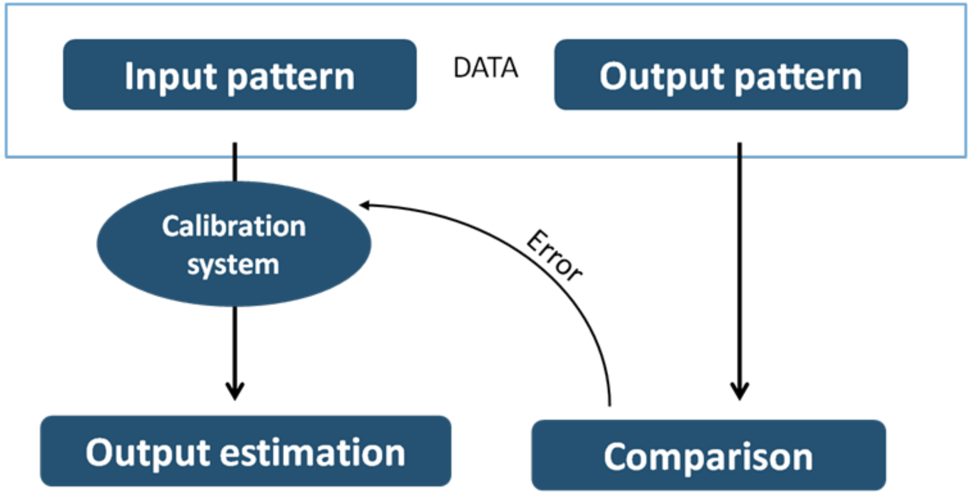
- Computational tools that combine many simple processing units to obtain complex responses (**Machine Learning**).
- A different approach from the conventional modelling, able **to find the non-linear functional relationship among the variables**, without *a priori* assumption of suitable function or algorithm.

… in our work

Deep Feed-Forward Neural Network (FFNN)



Error Backpropagation algorithm



Mean-square error minimisation

..adapted from Lek et al. (2000)

Method development

A large *in situ* dataset of Chl*a* and Temp profiles with concurrent remotely-sensed variables was used for the **training** and **test** of the network

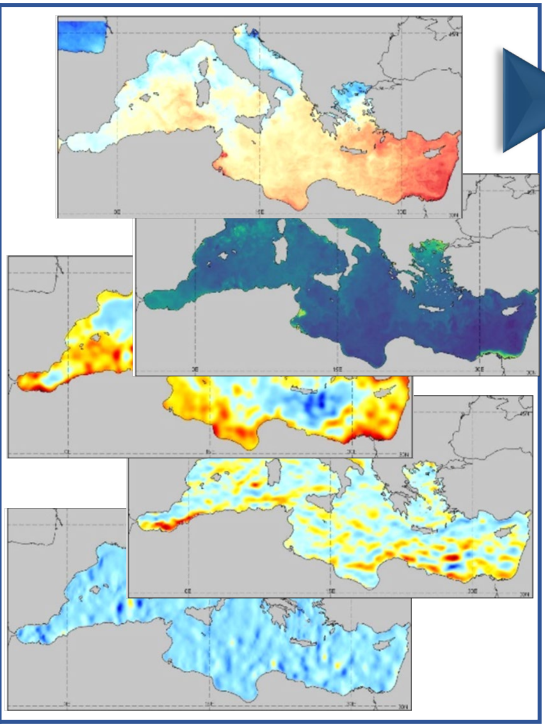
In situ data + **Surface observations**

Training/Test of the Neural Network

Application

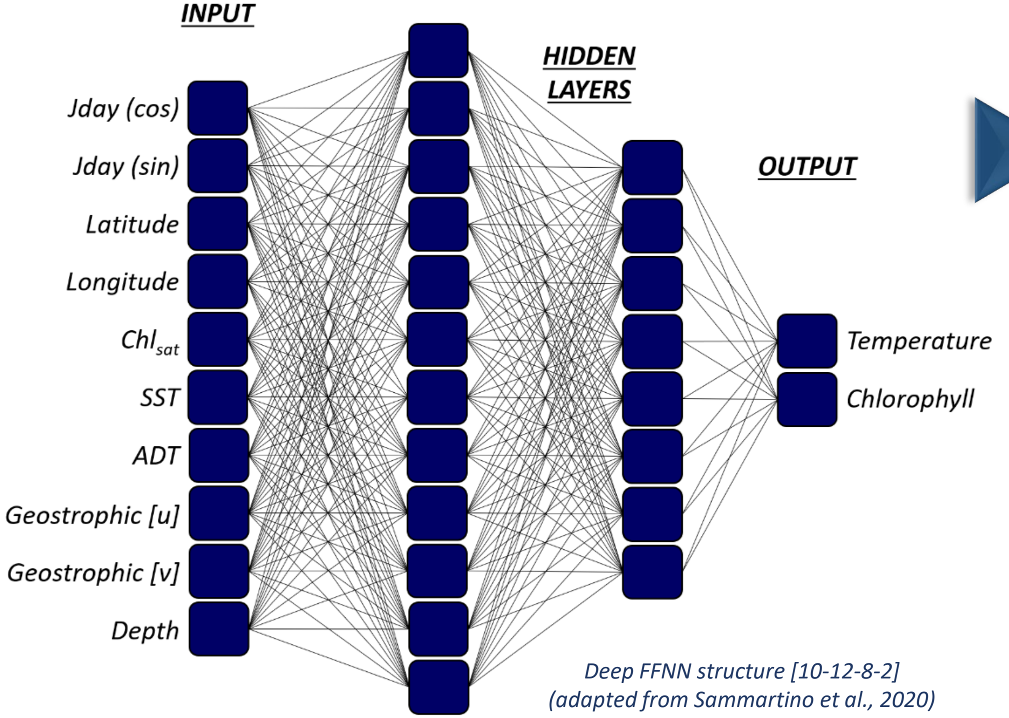
Surface observations

INPUT



Deep Feed-Forward Neural Network

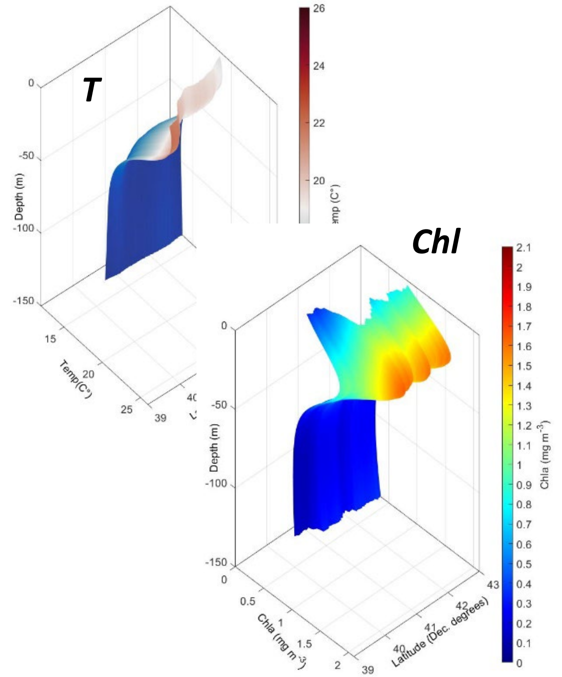
MACHINE LEARNING APPROACH



Deep FFNN structure [10-12-8-2]
(adapted from Sammartino et al., 2020)

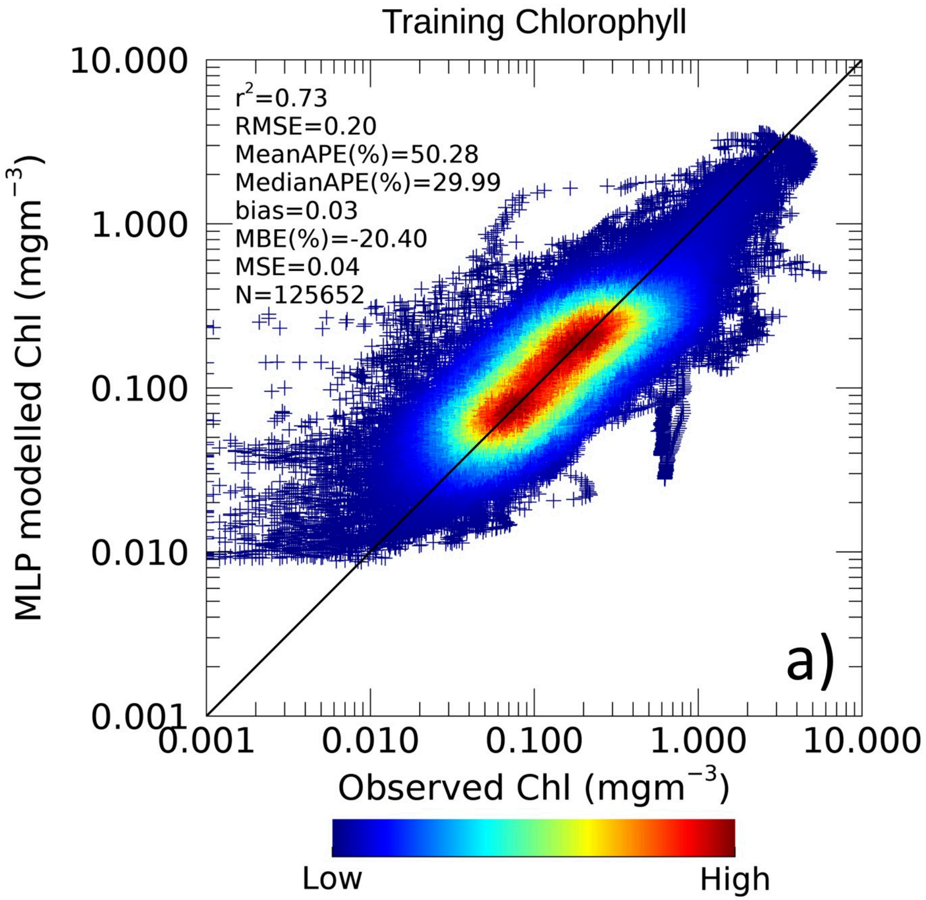
Vertical Profile

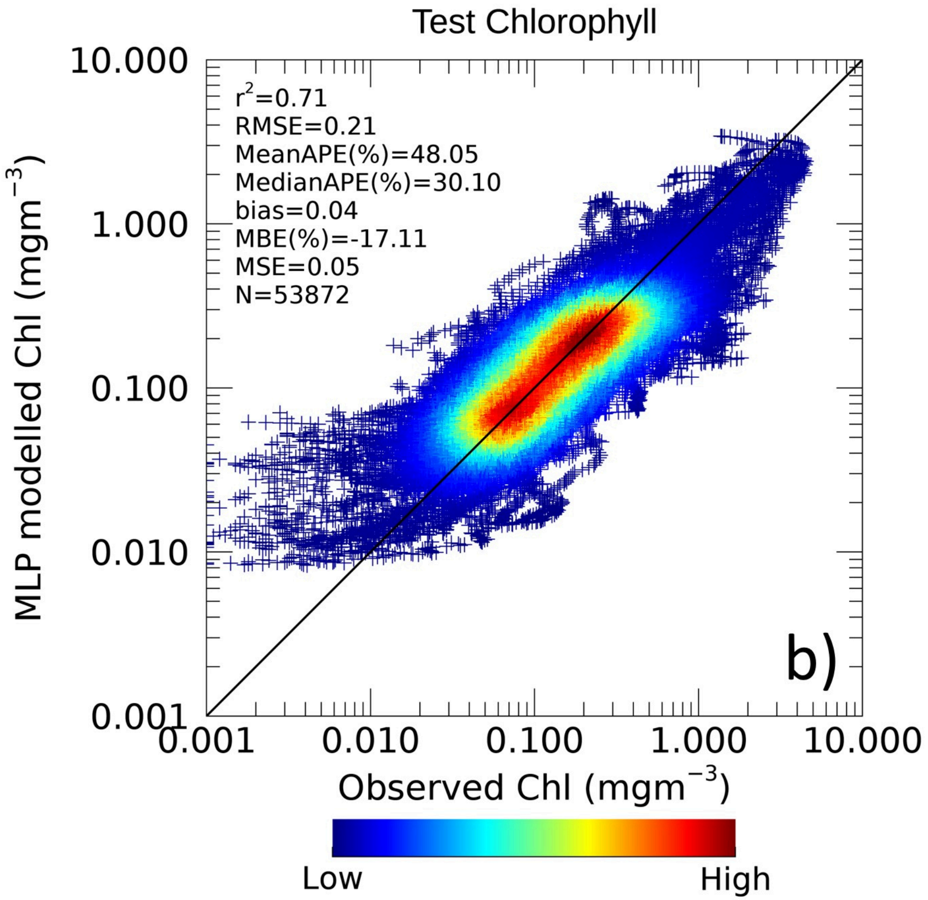
OUTPUT

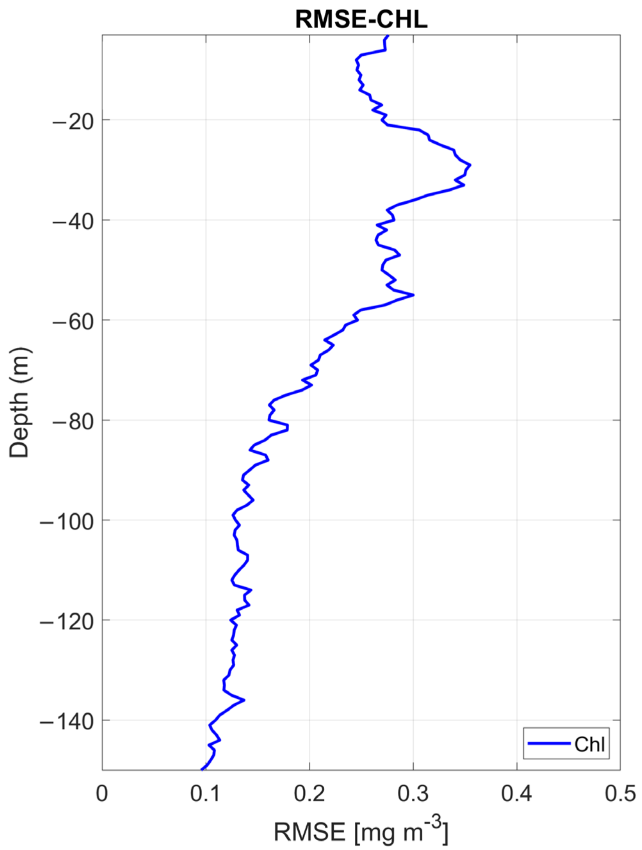


Chlorophyll results

- Network validation against *in situ* data showed very promising results, both on the training and test set;
- Comparable statistics both for training and test data set, suggesting the network is no-overfitting during the learning phase;
- The highest RMSE is observed from 20 m to 60 m of depth in agreement with the Deep Chlorophyll Maximum variations.

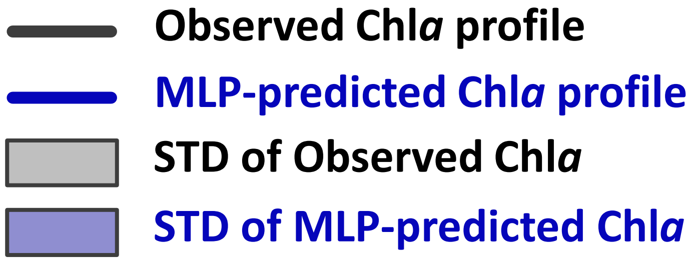


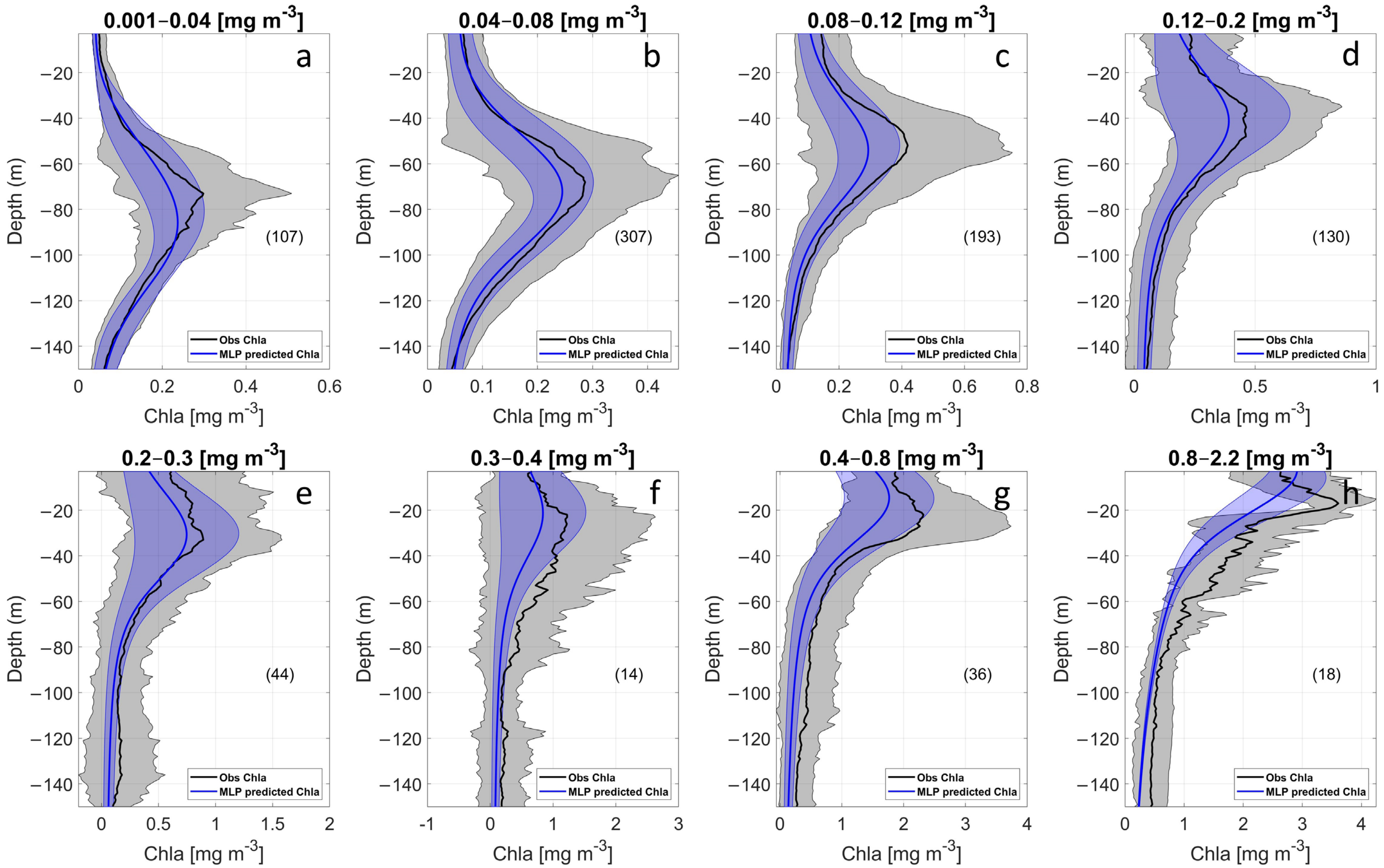




(from Sammartino et al., 2020)

Chlorophyll vertical profiles reconstructed from satellite



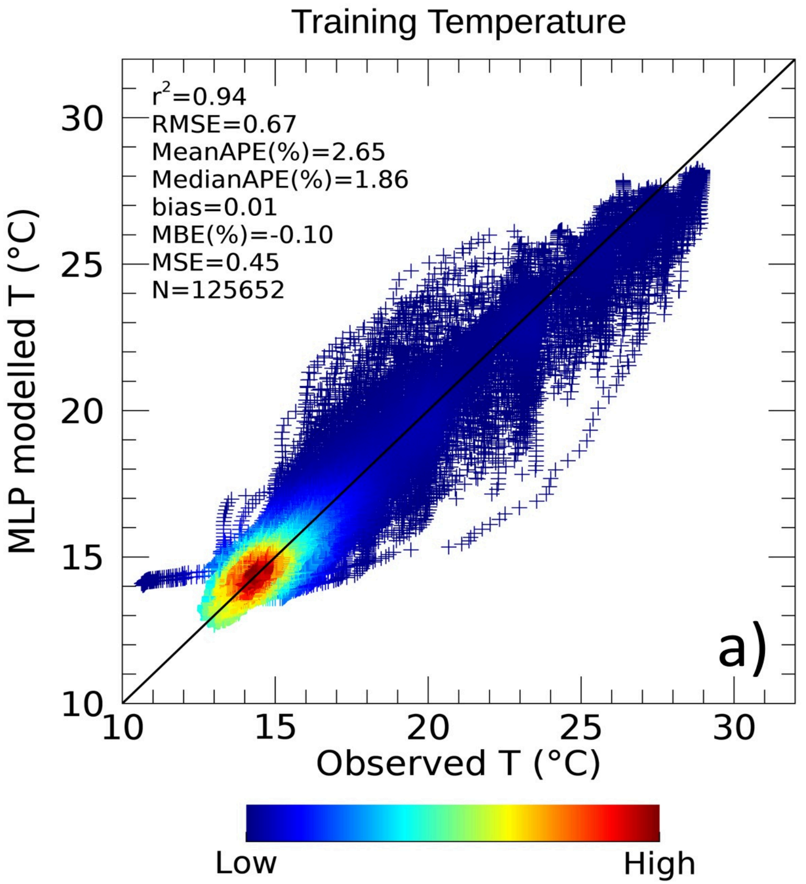


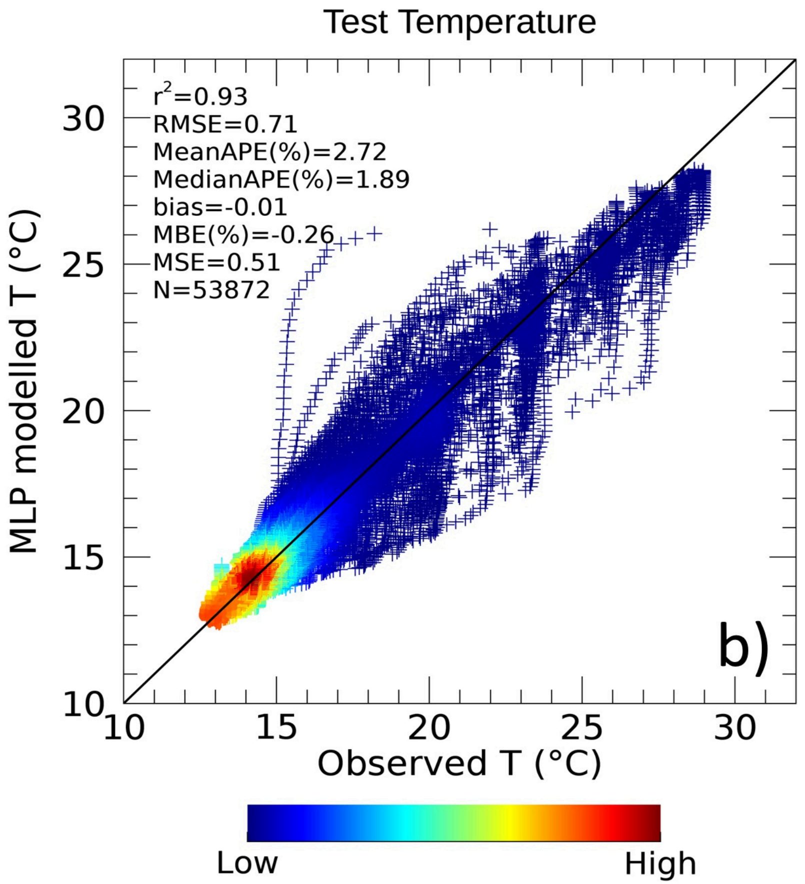
- Surface MLP-predicted Chl*a* very close to observed value for almost classes;
- Reconstruction of the DCM postion;
- Overlap of shaded areas (incorporation of high Chl*a* variability in the Med);
- Less accurate prediction for high [Chla_{surf}] ranges (maybe extreme bloom events less represented in the dataset).

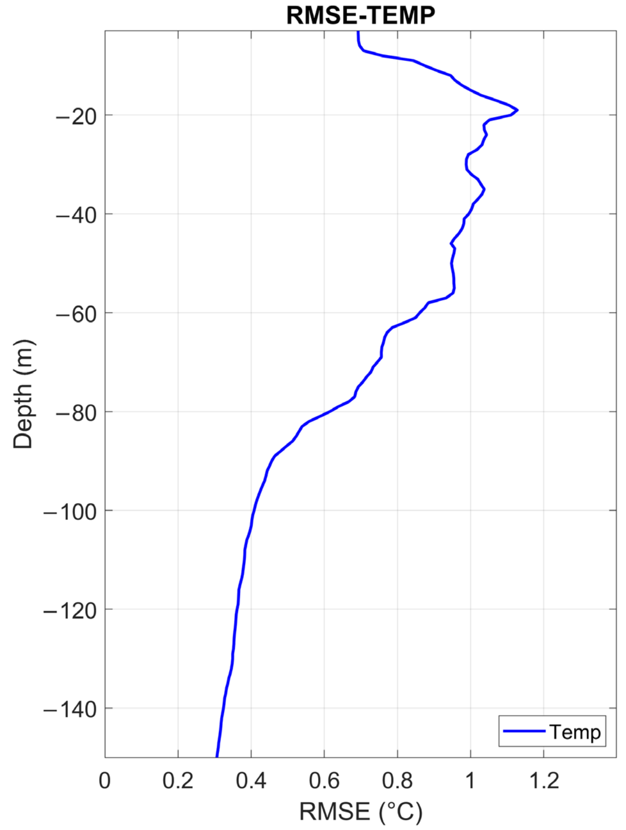
(from Sammartino et al., 2020)

Temperature results

- The network simulates the observed values with good accuracy;
- The statistics are better for temperature than those obtained for chlorophyll;
- No-overfitting during the learning phase;
- Maximum variability between 20 m and 80 m of RMSE curve, usually related to the mixed-layer depth (MLD) variations, impacting also the DCM position.

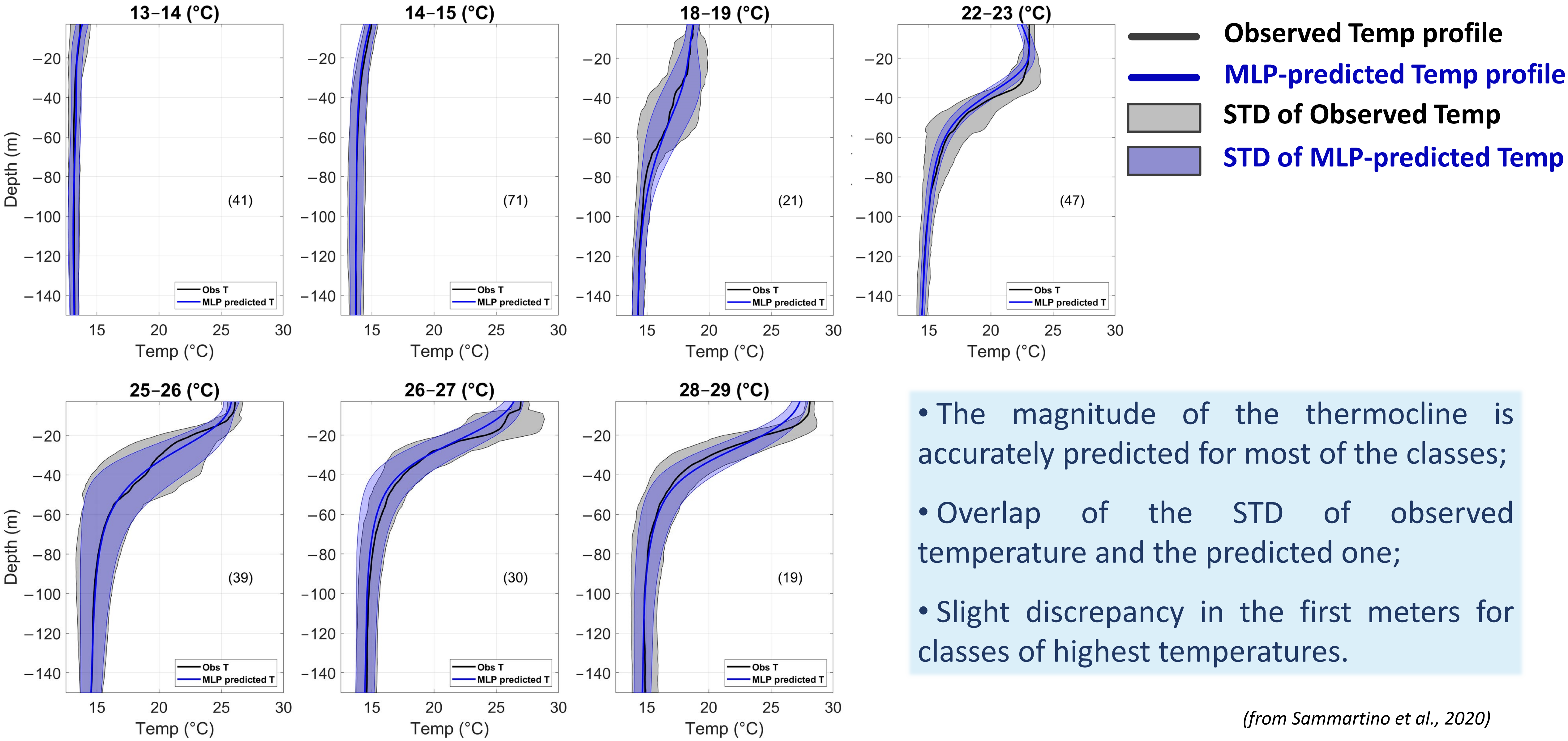






(from Sammartino et al., 2020)

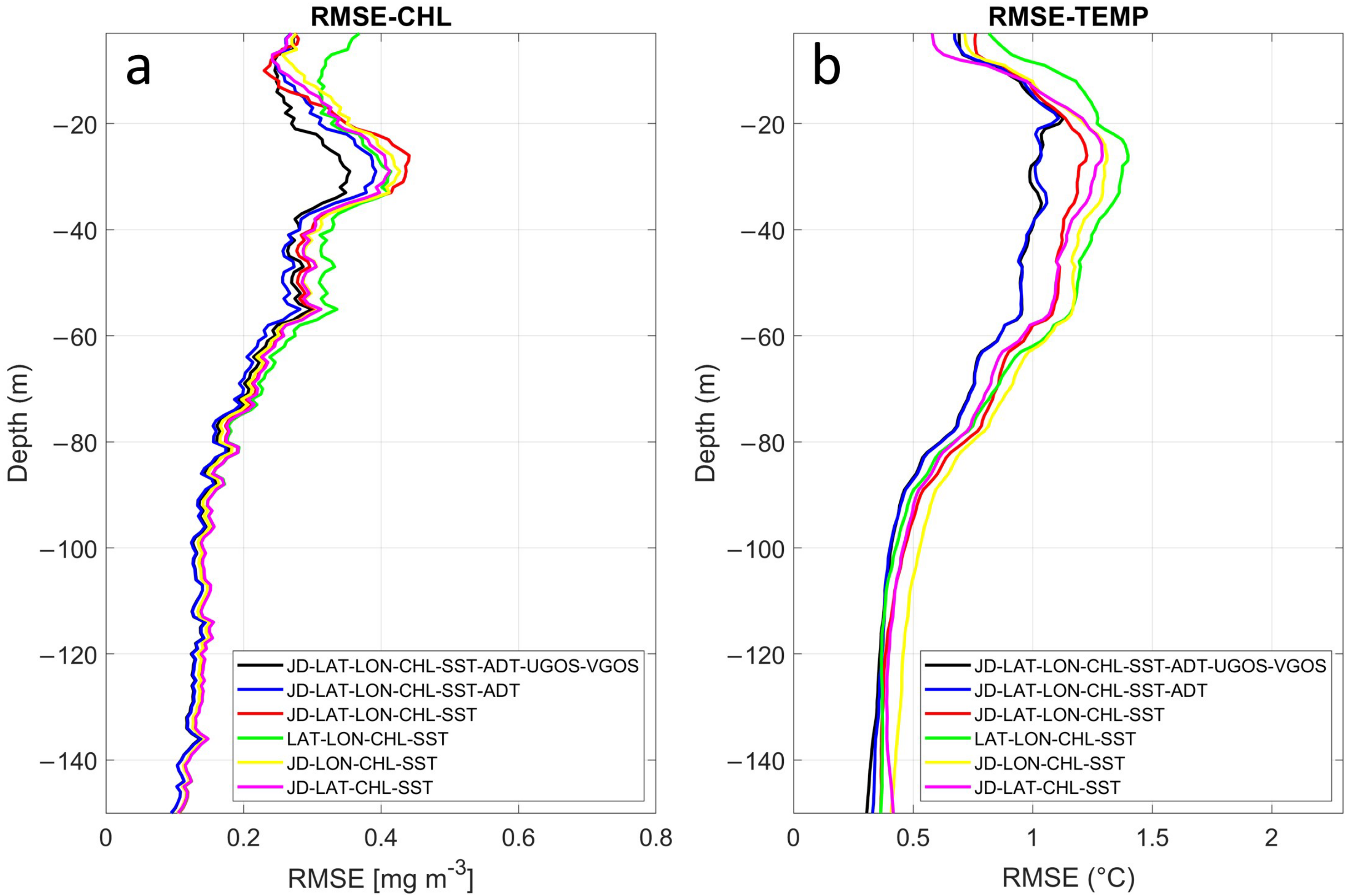
Temperature vertical profiles reconstructed from satellite



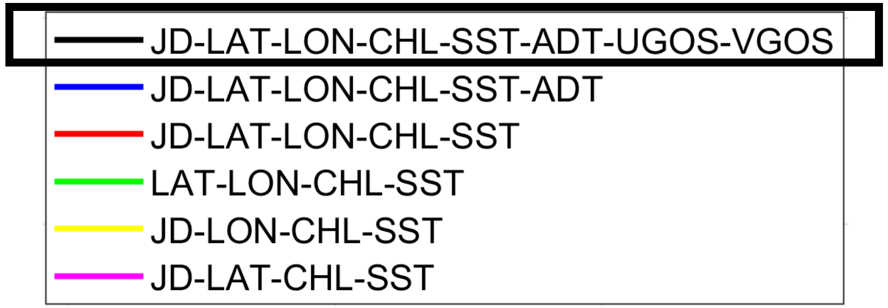
- The magnitude of the thermocline is accurately predicted for most of the classes;
- Overlap of the STD of observed temperature and the predicted one;
- Slight discrepancy in the first meters for classes of highest temperatures.

(from Sammartino et al., 2020)

A sensitivity analysis of the input influence on the prediction



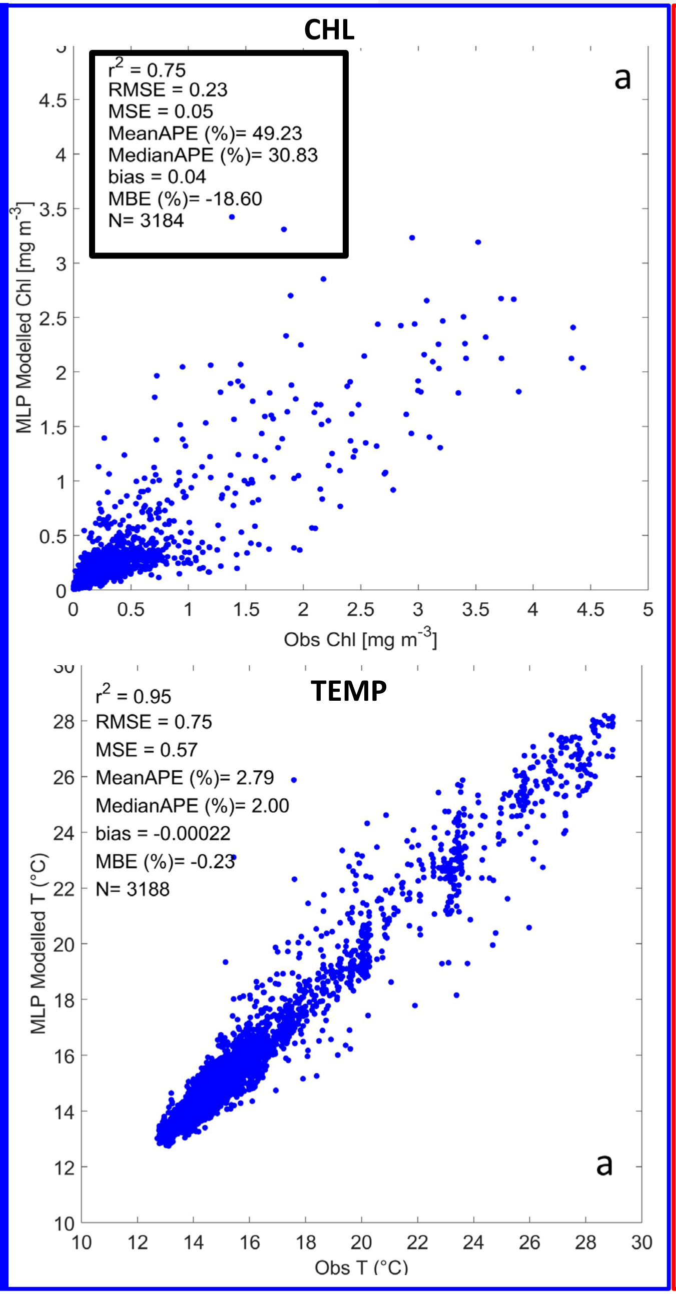
- Tests on different combinations of input surface data revealed that the best MLP setup included all the considered remote sensing variables;
- For both variables, the addition or subtraction of the ADT determines a strong impact and reduction of the errors

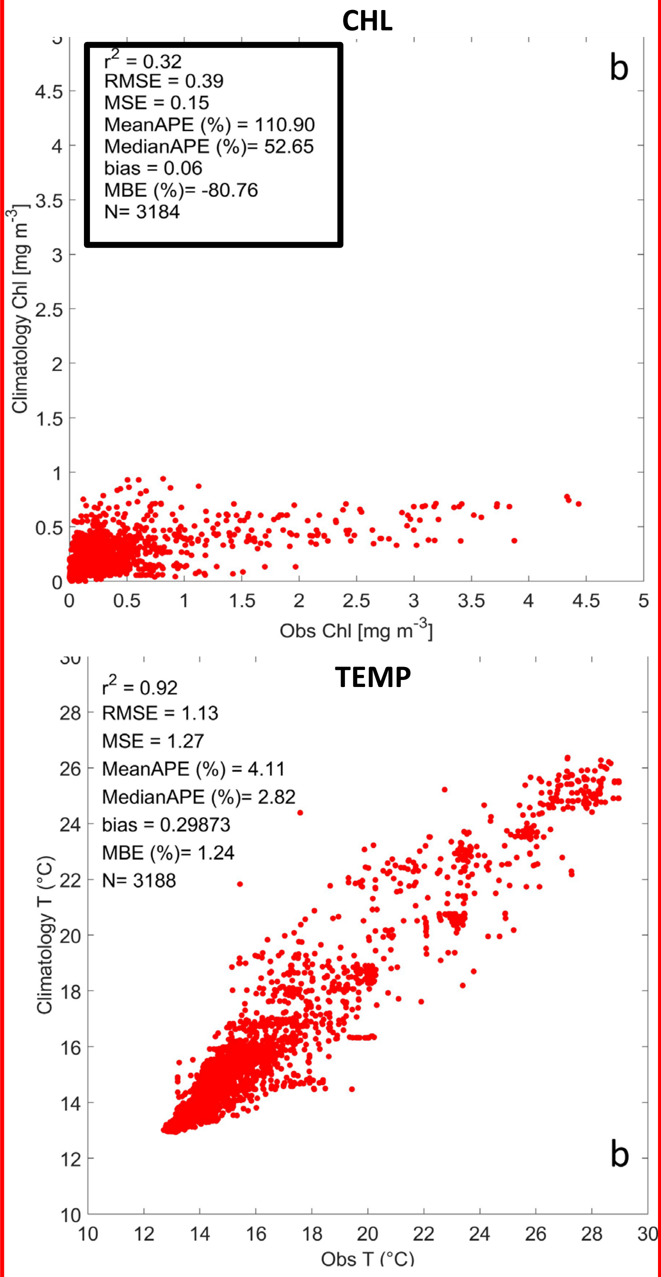


(from Sammartino et al., 2020)

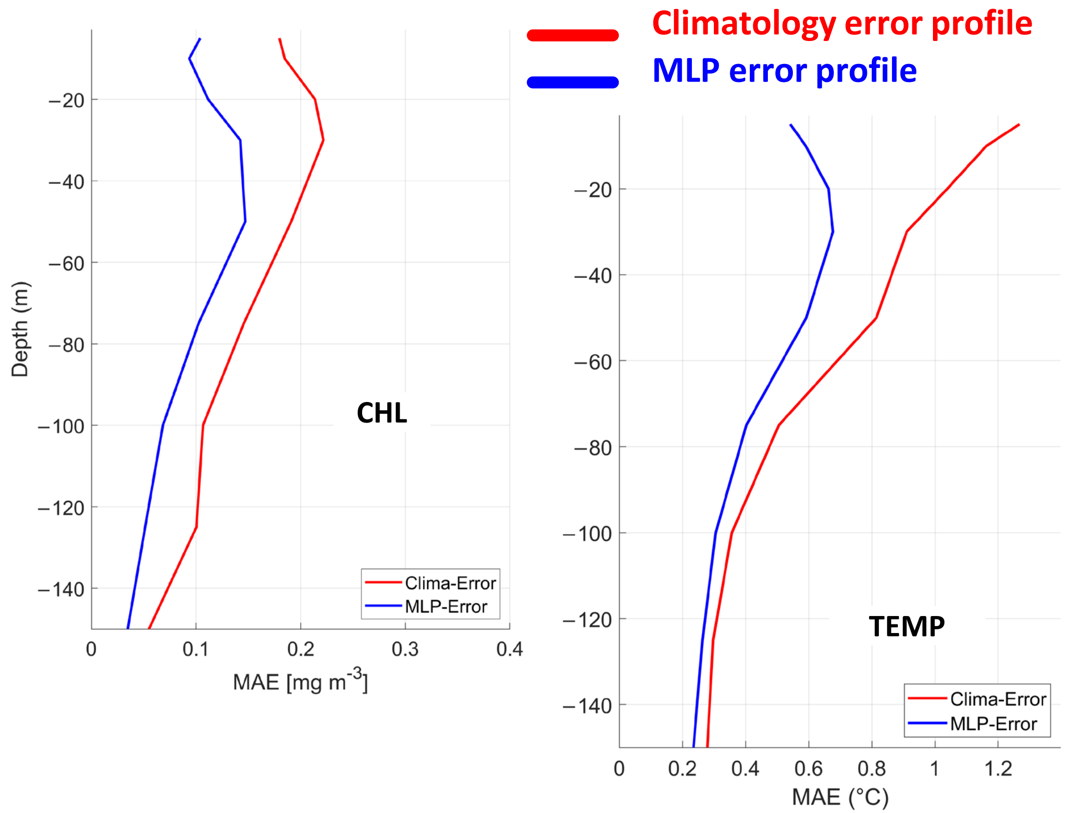
A comparison with reference Mediterranean Climatology (MEDATLAS)

NEURAL NETWORK COMPARISON





CLIMATOLOGY COMPARISON



- A significant improvement of the MLP Chl reconstruction with respect to the regional Mediterranean climatology
- MLP applied on satellite data can be an effective alternative to climatology to overcome the discontinuity and sparseness of the in-situ observations.

(from Sammartino et al., 2020)

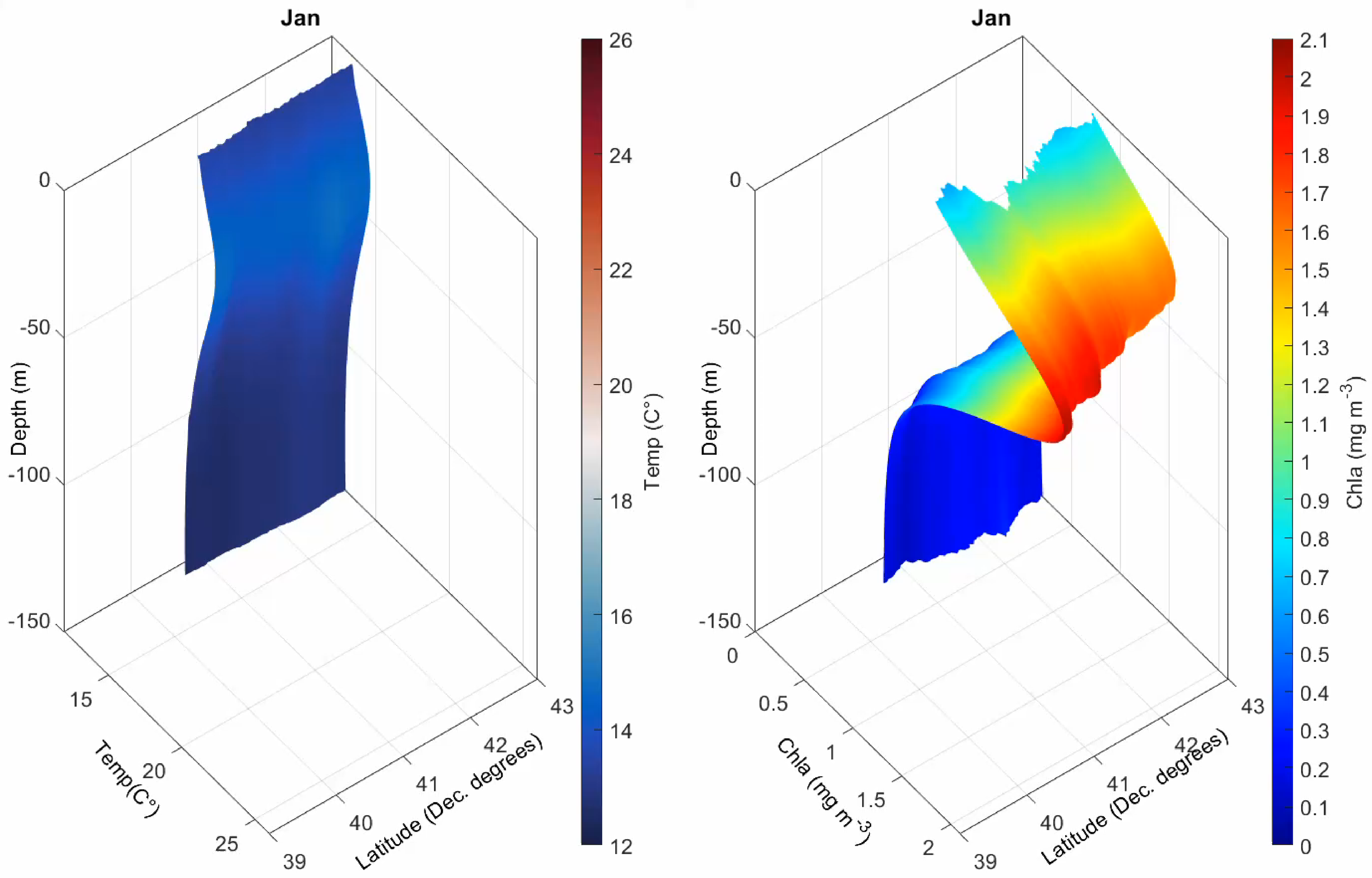
Conclusions and future works

- **Synergy** between data of different nature is possible, allowing the extrapolation of surface marine variables to deeper layer (from 2**D -> 3D)**;
- The prediction capability of such neural networks is strictly depending on **training dataset features** and the **choice of co-predictors**, that deeply influence the network's performance;
- The application of innovative techniques as those based on **artificial intelligence** to data acquired by **multiple and interdisciplinary observing systems** represents a useful approach to describe the ocean state evolution from surface to the deeper layers;
- Machine learning techniques applied to satellite estimates demonstrated huge and still only minimally exploited potentialities, both as predictive models and to better initialize numerical bio-geophysical models.

... works in progress

- **Test additional predictors**, for instance, directly ingesting radiances instead of satellite chlorophyll, or including input from new gap-free regional and multi-sensor sea surface salinity at high resolution (Sammartino et al. 2022);
- Implementing **different types of neural networks** (e.g. artificial recurrent neural network) for the reconstruction of the **3D structure of other variables** in the Mediterranean Sea (e.g. vertical salinity).


Thank you for your attention!



Jan
Depth (m)
Temp(C°)
Latitude (Dec. degrees)
Temp (C°)
Jan
Depth (m)
Chla (mg m -3)
Latitude (Dec. degrees)
Chla (mg m -3)

Principal references

Pubblication

- *Lek, S. and Giraudel J.L. and Guégan, J.-F. Neuronal networks: Algorithms and architectures for ecologists and evolutionary ecologists. In Artificial Neuronal Networks; Springer: Berlin, Germany, 2000; pp. 3–27.*

- *Sammartino, M.; Buongiorno Nardelli, B.; Marullo, S.; Santoleri, R. An Artificial Neural Network to Infer the Mediterranean 3D Chlorophyll-a and Temperature Fields from Remote Sensing Observations. Remote Sens. 2020, 12, 4123. https://doi.org/10.3390/rs12244123*

- *Sammartino, M.; Aronica, S.; Santoleri, R.; Buongiorno Nardelli, B. Retrieving Mediterranean Sea Surface Salinity Distribution and Interannual Trends from Multi-Sensor Satellite and In Situ Data. Remote Sens. 2022, 14, 2502. https://doi.org/10.3390/rs14102502*

Website

https://dficlub.org/future-of-machine-learning/
https://en.wikipedia.org/wiki/Earth_observation_satellite
https://www.gim-international.com/content/news/sentinel-2b-satellite-launched-into-orbit
http://gosweb.artov.isac.cnr.it/viewer/viewer.php