Assimilation of satellite sea surface temperature in southern European seas using ML-based operators

Daniele Bigoni, Beniamino Tartufoli, Filipe Costa, Matteo Broccoli, Andrea Cipollone, Paolo Oddo, Emanuela Clementi, Ali Aydogdu

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

cmcc

Abstract – *Assimilating satellite sea surface temperature into ocean models is challenging due to the complex relationship between observed and modeled quantities. Observation operators are needed to project observations into model-compatible values. Deep learning offers flexible models capable of capturing these relationships while providing easy access to adjoint operators.*

The analysis and forecasting systems – The regional analysis forecasting systems for the Mediterranean (MedFS) and Black Sea (BSFS) provide regular analysis and forecasting of physical variables temperature (T), salinity (S), currents (V,U) and sea surface height (H) to the Copernicus Marine Service.

The forecasting systems are based on a NEMO configuration with quad horizontal discretizations (1/24 MedFS, 1/40 BSFS) and uneven vertical layers (141 MedFS, 121 BSFS). The systems are forced by ECMWF atmospheric forcing and EFAS rivers discharges, and implement open boundary conditions in the Atlantic and Dardanelles.

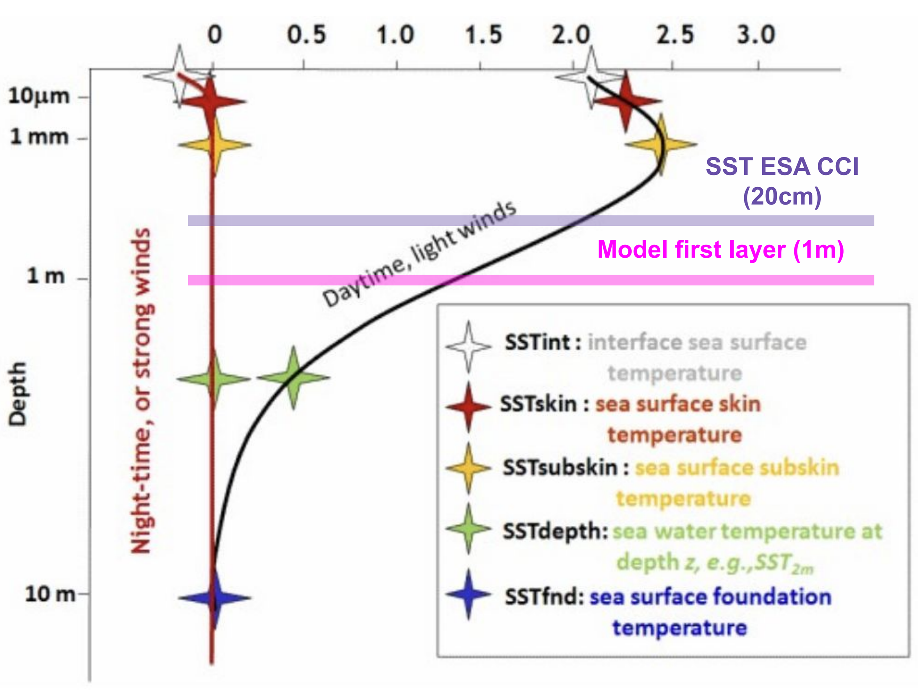
The data assimilation component is the 3D variational scheme OceanVar2.0 [1-2] where vertical covariances for T, S and H are described by Empirical Orthogonal Functions, while horizontal covariances are approximated by Gaussian functions and applied through a diffusion operator. A barotropic model is employed to assimilate sea level anomaly observations.

Assimilation of Sea Surface Temperature (SST)

Satellite observations of sea surface temperature need to be projected onto the model grid in order to be assimilated.

Near real time (NRT) L3 products provide foundation SST in non cloud-occluded areas that can be assimilated at nighttime (FGAT) and used to correct heat fluxes.

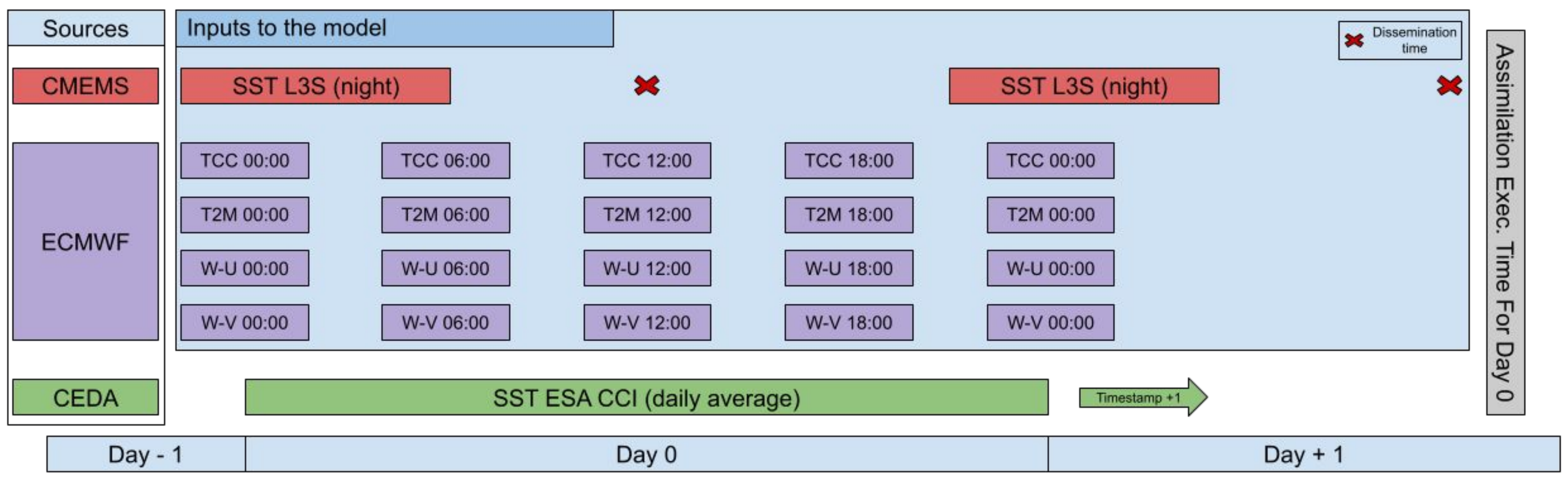
ESA CCI $SST_{0.2m}$ is an L4 analysis product representing the daily average temperature at 20cm.



SST shift due to diurnal cycle. Source: "The Recommended GHRSST Data Specification (GDS 2.1)"

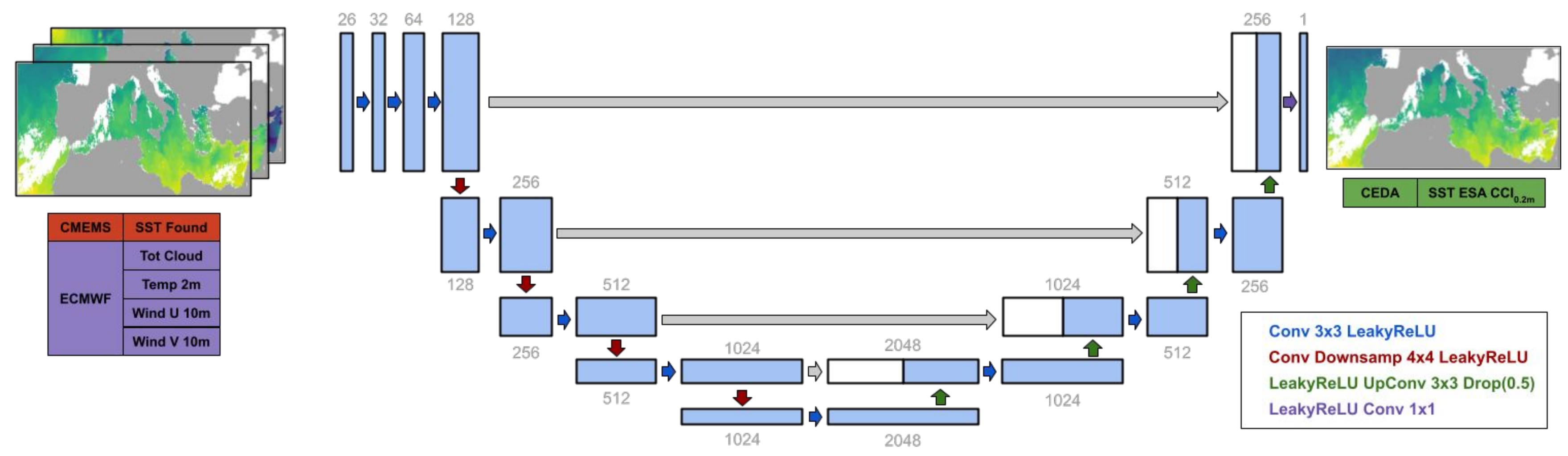
We follow [3] and try to assimilate higher quality SST information close to the model grid through a machine learning model capable of reproducing ESA CCI $SST_{0.2m}$ quality data from NRT L3 foundation SST and complementary variables.

Machine Learning for ESA CCI $SST_{0.2m}$ – Considering the dissemination schedule of the different products, we include the following spatial data layers as inputs and outputs to the ML model:



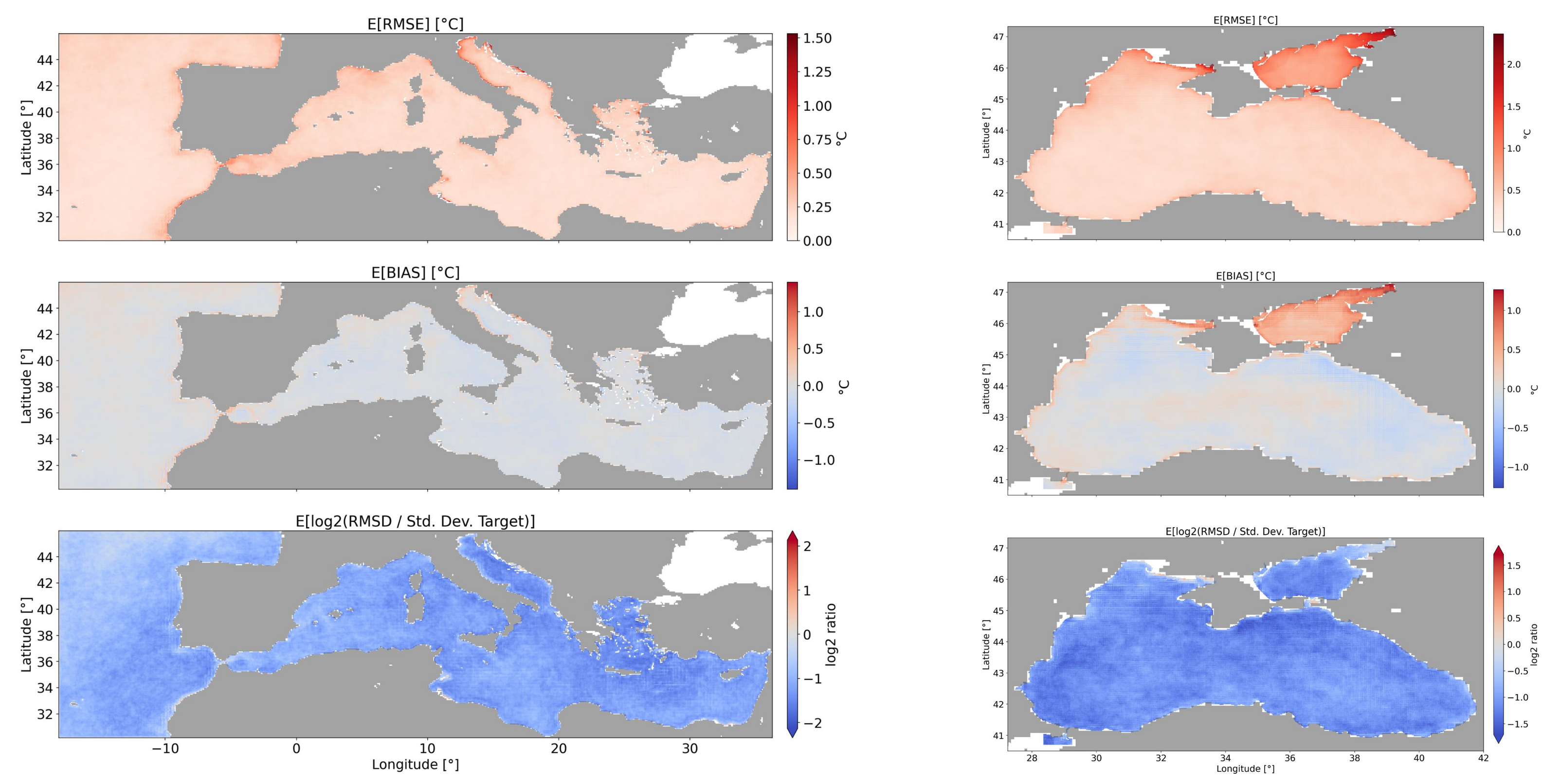
Data sources and layers used to construct the ML Model.
Not shown but included in the inputs are longitude, latitude and L3S masks layers for a total of 26 input channels.

Building on experience in [3] and after some tuning, we consider a UNet architecture for all our experiments with the following structure.



UNet architecture with 4 down/up blocks. Input features are standardized (L3S, T2M, U10M, V10M), min-max scaled (lon and lat), 0-100 scaled (TCC), and padded to fit the convolutional down/up scaling architecture.

The training loss is the RMSD over non cloud-occluded areas weighted by the inverse of the variance reported for ESA CCI $SST_{0.2m}$.
The network is trained on the data 2018-2022 and tested on 2023.



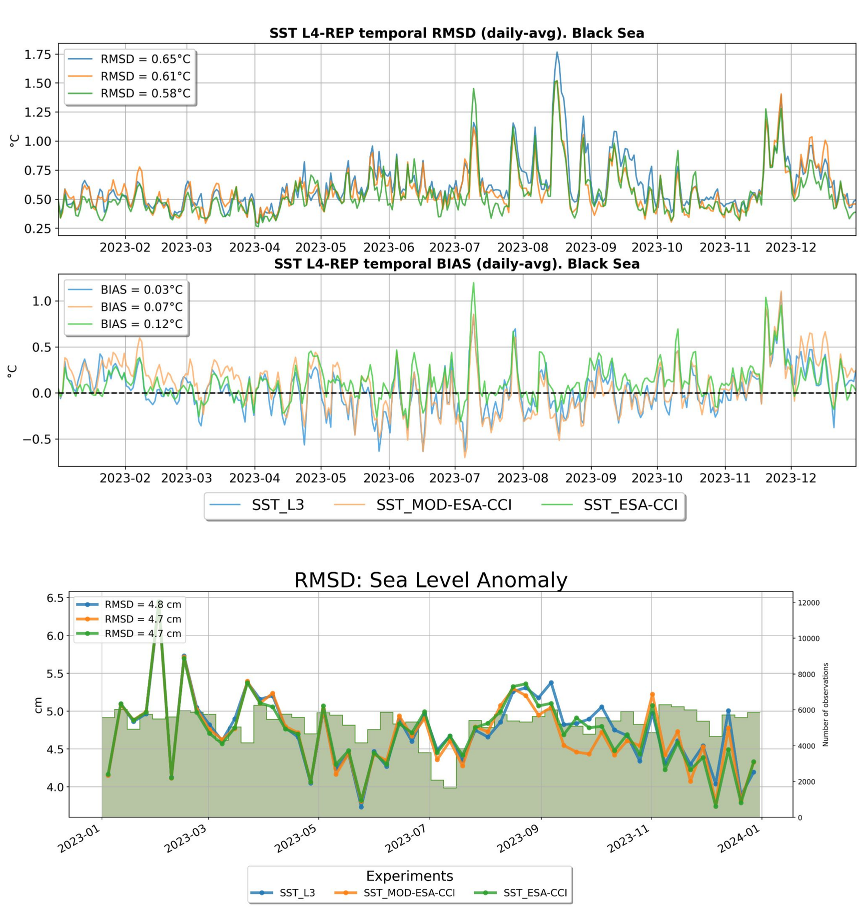
Averaged errors in SST ESA $CCI_{0.2m}$ reconstruction. The log2 ratio between RMSD and standard deviation of the data indicates that the model is already predicting below the estimated accuracy of the target data.

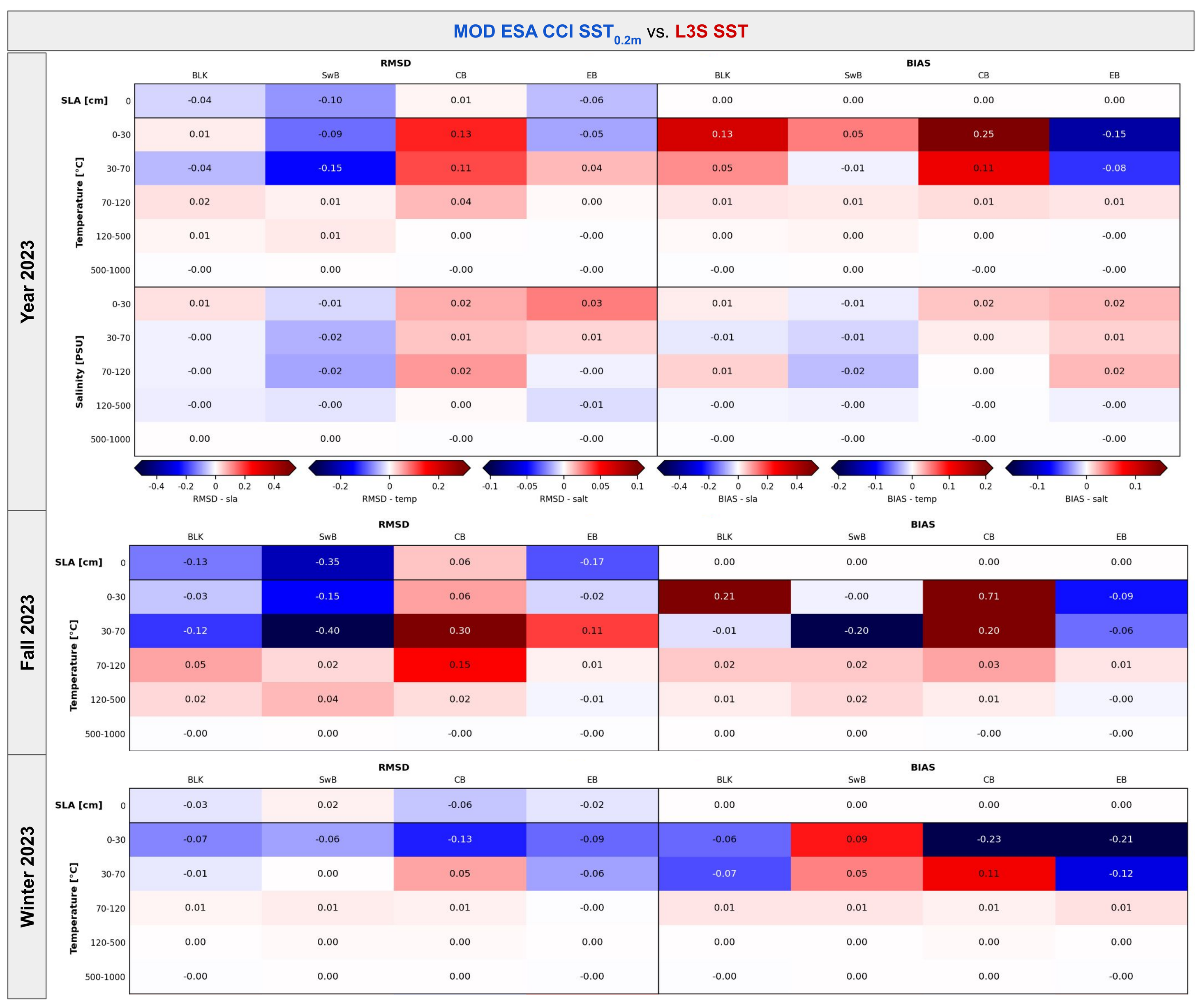
Experimental settings – Our initial testbed for the use of reconstructed ESA CCI $SST_{0.2m}$ is the Black Sea FS. We tested three scenarios:

- assimilation of **L3S SST**
- assimilation of **ESA CCI $SST_{0.2m}$**
- assimilation of ML **MOD ESA CCI $SST_{0.2m}$**.

In all the settings SST was the only observation assimilated. The test was run on year 2023, for which we have ESA CCI $SST_{0.2m}$ data and on which the ML model was tested.

We evaluate performance with respect to in-situ observations (ARGO), along track sea level anomalies and reprocessed L4 SST.





Results and next steps

Experiments show that ESA CCI assimilation

- improves with respect to the L4Rep SST,
- improves SLA across the basin,
- improves temperature in upper layers during fall and winter.

While results are promising, we are investigating several improvements

- Use of daily mean misfits rather than midnight FGAT (which introduces a positive bias due to the diurnal cycle)
- Investigation of seasonal biases which could account for degraded performance during spring and summer
- Extensive testing on the Mediterranean Sea

References

[1] Dobricic, S., & Pinardi, N. (2008). An oceanographic three-dimensional variational data assimilation scheme. *Ocean Modelling*, *22*(3–4), 89–105.

[2] Oddo, P., Adani, M., Carere, F., Cipollone, A., Goglio, A. C., Jansen, E., Aydogdu, A., Mele, F., Epicoco, I., Pistoia, J., Clementi, E., Pinardi, N., & Masina, S. (2025). *OceanVar2.0: An open-source variational ocean data assimilation scheme. Sensitivity to altimetry sea level anomaly assimilation.*

[3] Broccoli, M., & Cipollone, A. (2025). Assimilation of diurnal satellite retrieval of sea surface temperature with convolutional neural network. *Machine Learning: Earth*.