

CONTEXT

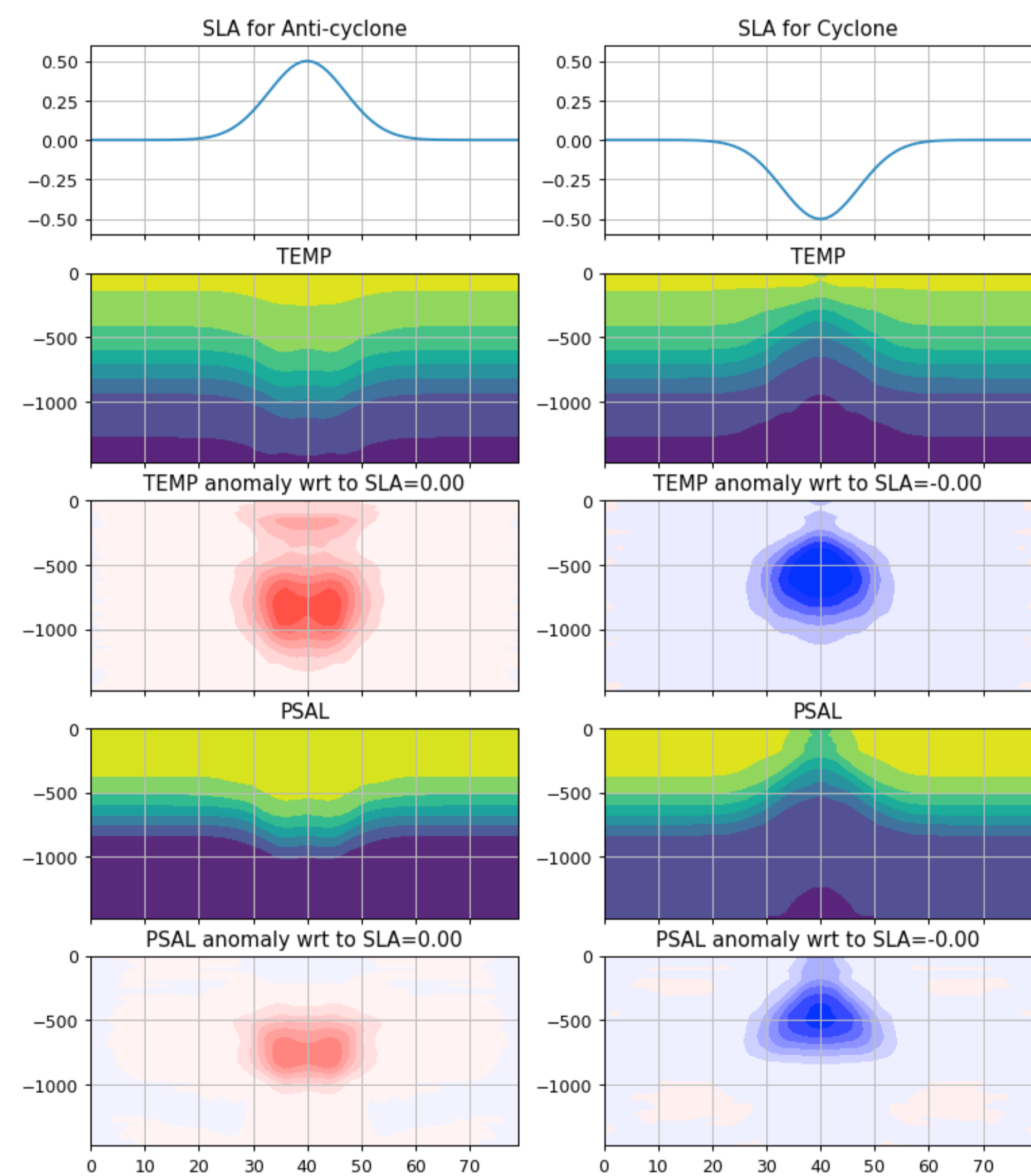
In order to determine the ocean's low frequency variability (sub-annual and beyond), it is necessary to define a reference state, i.e. a climatology. This requires information about the ocean's three-dimensional thermo-haline structure from in-situ observations. This has historically been synthesised using optimal approaches based on linear Gaussian statistics that have a significant drawback: a strong smoothing resulting in a loss of information. Developing ocean climatologies that can retain as much signal as possible from observations remains a statistical challenge.

Here, we propose a new approach we call OSnet (from "ocean state neural network"). We will present OSnet and its possible applications in physical oceanographic studies as well as in data management of in-situ observations (quality control).

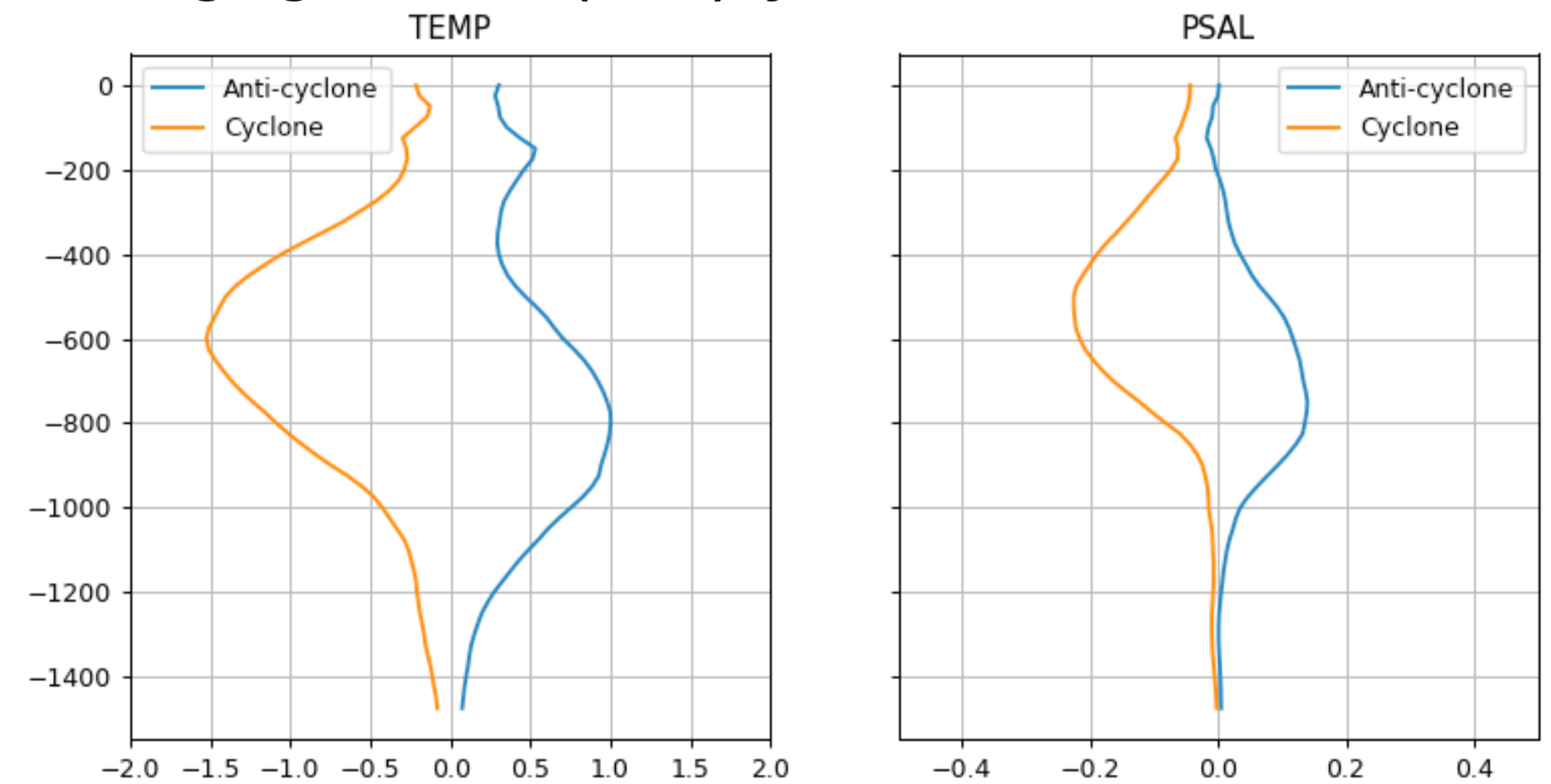
The specific and added value of OSnet compared to other models is that (i) it does not rely on an interpolation of observations onto a regular grid (ii) it can take into account dynamical information provided by altimetry and (iii) it automatically and implicitly determines the best covariance scales, limiting the amount of unnecessary smoothing. At this point, OSnet is able to predict sub-annual variations with similar accuracy to state of the art ocean state climatology (eg: ISAS15).

RESULTS

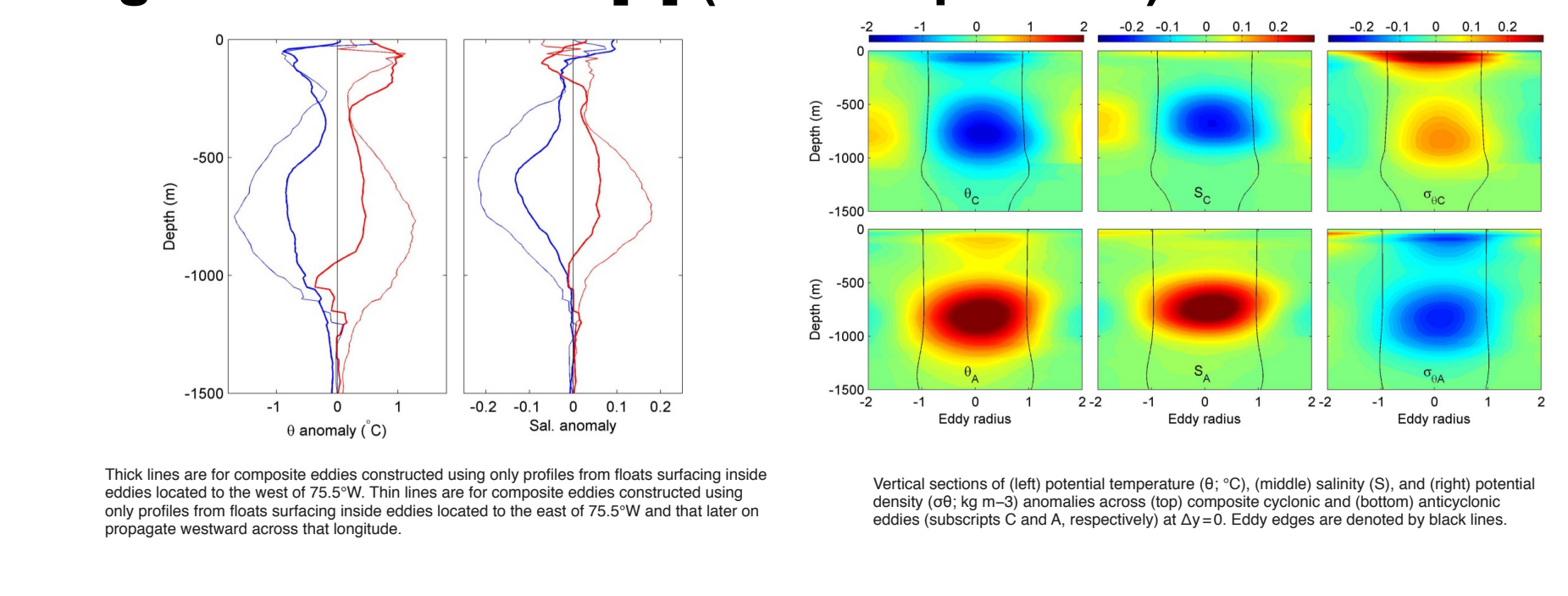
Role of altimetry: Since OSnet takes SSH as an input, we can determine the dependence of the vertical thermo-haline prediction to surface signature of eddies. Figures below show the temperature and salinity field and anomalies associated with realistic SLA from an anti-cyclone (left) and a cyclone (right) in the North Atlantic Bight. Figures at the bottom show the same diagnostics from observations [3]. We conducted the same successful comparison in the Kuroshio.



Averaging over the (anti-)cyclone:



Figures taken from [3] (for comparison):



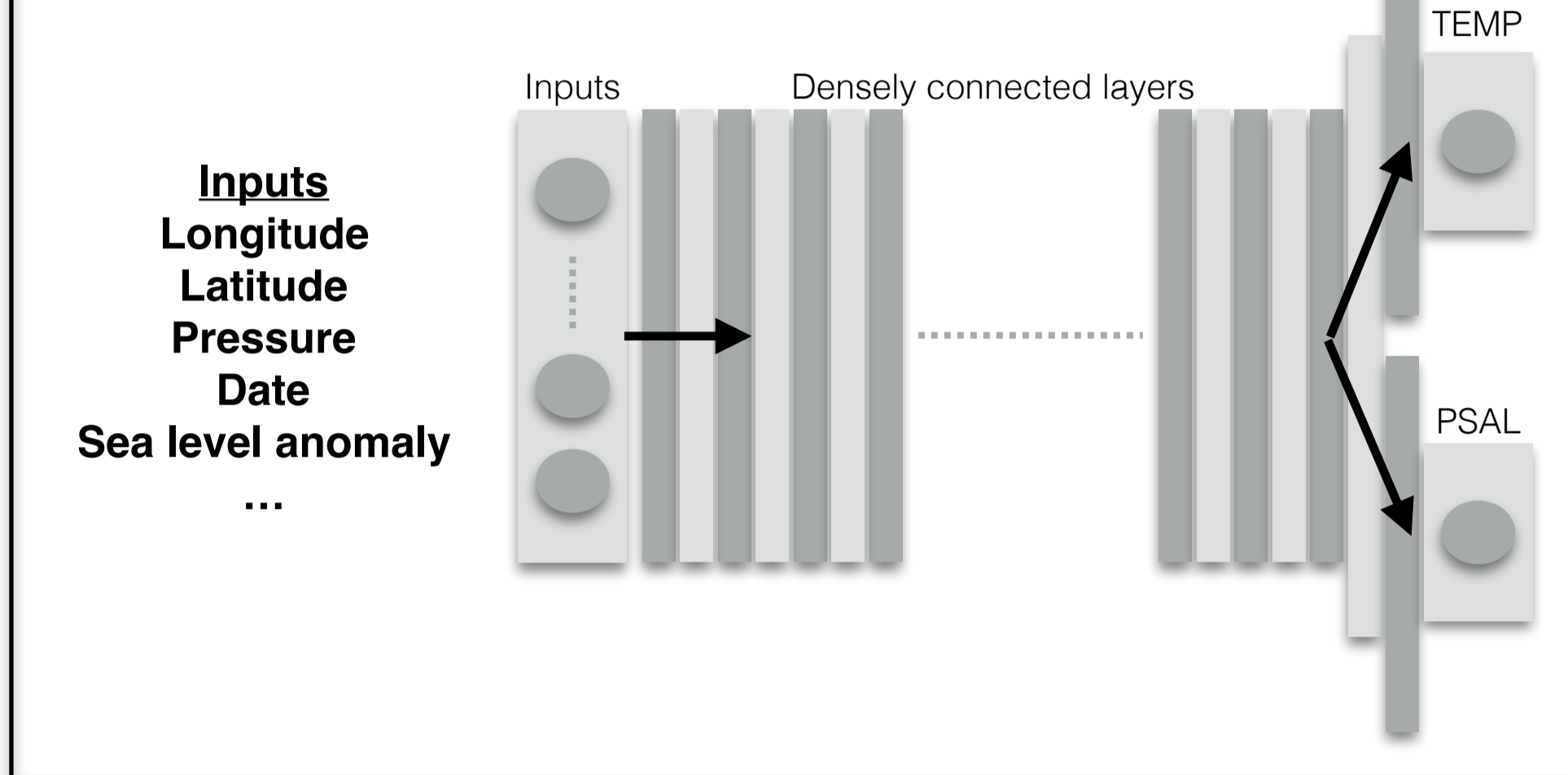
METHOD

The training phase:

We "fit" the neural network to the state of the ocean as measured by the ARGO & AVISO datasets. We do this using Keras [1] with a TensorFlow backend [2].

MLP-like structure:

~500k free parameters to adjust



The data:

- 15 years of ARGO data quality controlled for research (2001 - 2016).
- For each ARGO profile, the sea level anomaly is interpolated from AVISO.
- **255 million measurements**

Normalisations:

- lon/lat get converted into Cartesian coordinates
- Dates are either simply normalised (but kept as a linear axis), or rendered periodic, as desired

The prediction phase:

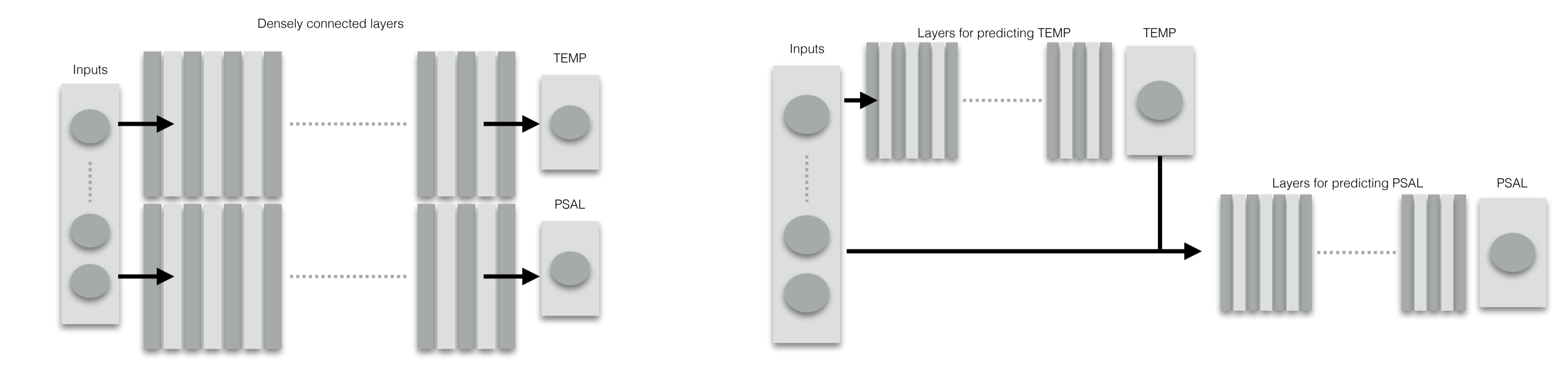
Once trained, the model makes predictions for ocean temperature and salinity anywhere in the world. These predictions can be compared to other models such as ISAS, as well as to other datasets (eg: moorings). See below.

Estimating the prediction error:

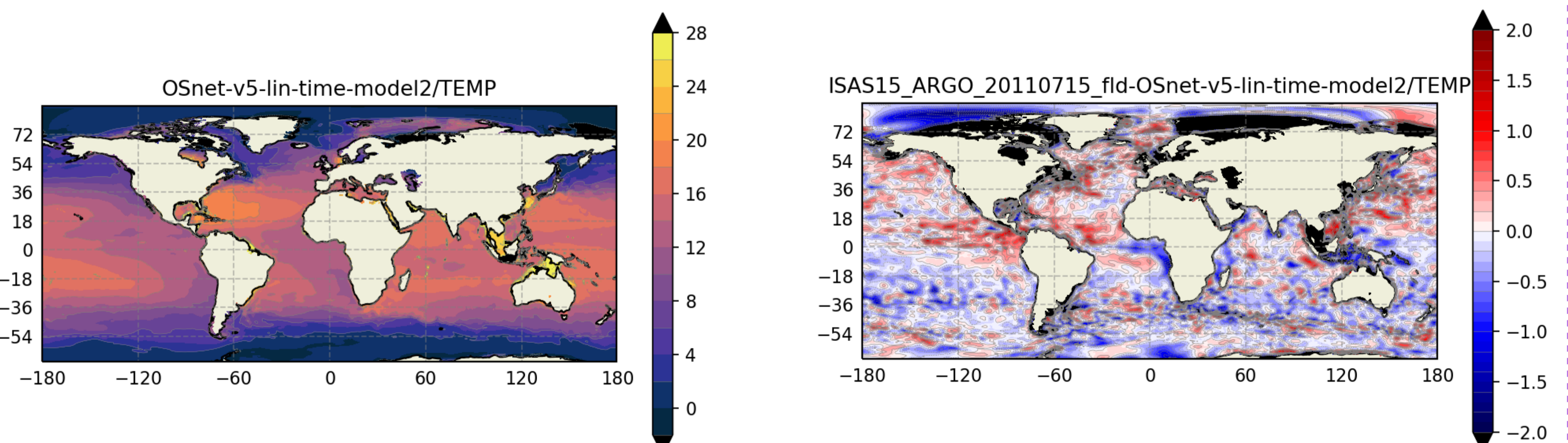
We do a 5-fold cross-validation, i.e. we actually train 5 models using subsets of our data, and consider the ensemble mean and variance of the predictions as the final output.

Other structures:

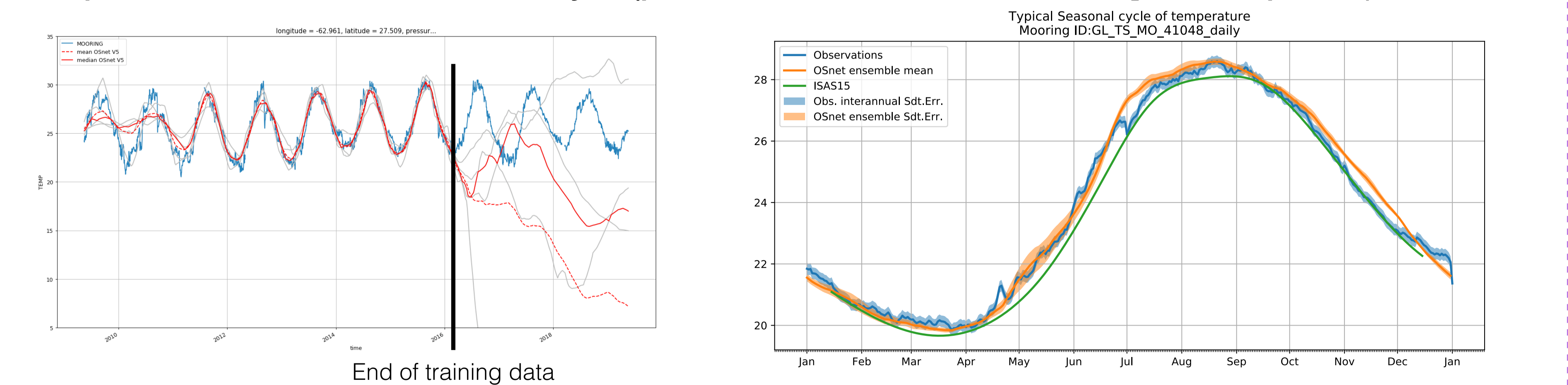
We're also trying different structures: training two separate networks (one for temperature and one for salinity), training temperature first, then salinity given a particular temperature... watch this space!



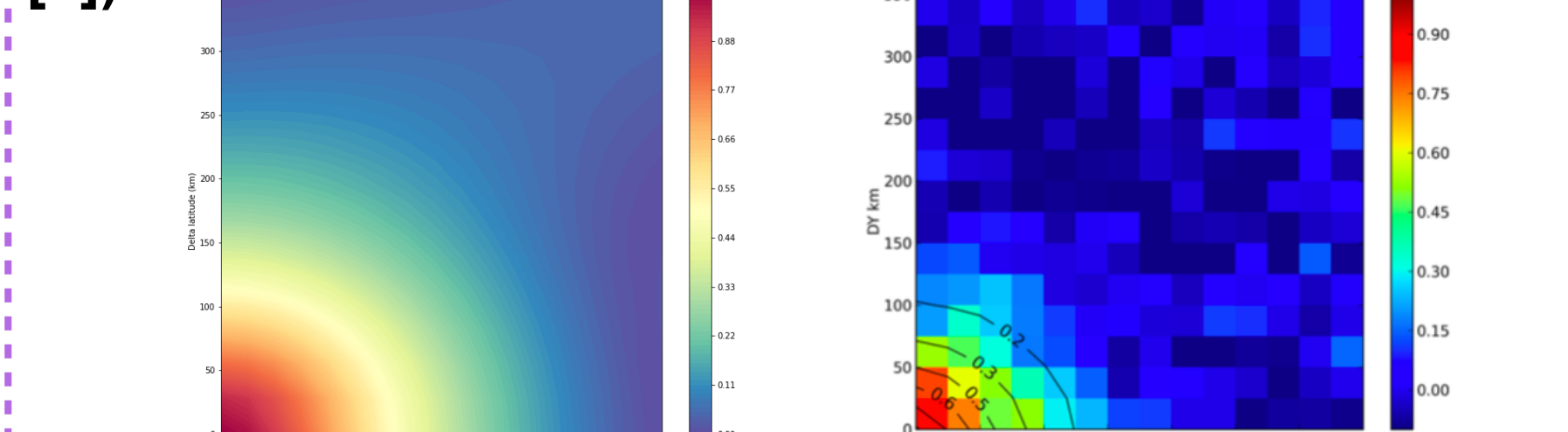
Map of ocean temperature at 100m depth for 2011/07: estimated global error: ~0.3K, mostly in poorly observed regions



Examples of a time series and a seasonal cycle (predicted in locations of fixed moorings for comparison):



Example of covariance matrix for temperature predictions at 200m in the Kuroshio Extension region (left) and comparison with observations (right [4])



REFERENCES

- [1] Chollet, F. et al., Keras www.keras.io (2015)
- [2] Abadi, M. et al., TensorFlow, tensorflow.org (2015)
- [3] Castelao, R. M., J. Geophys. Res. Oceans **119**, 3, 2048 (2014)
- [4] Ninove, et al, Oc. Sc, (2016)