Exploring the power of machine learning to tune tracer diffusivities offline

Gustavo Marques
gmarques@ucar.edu

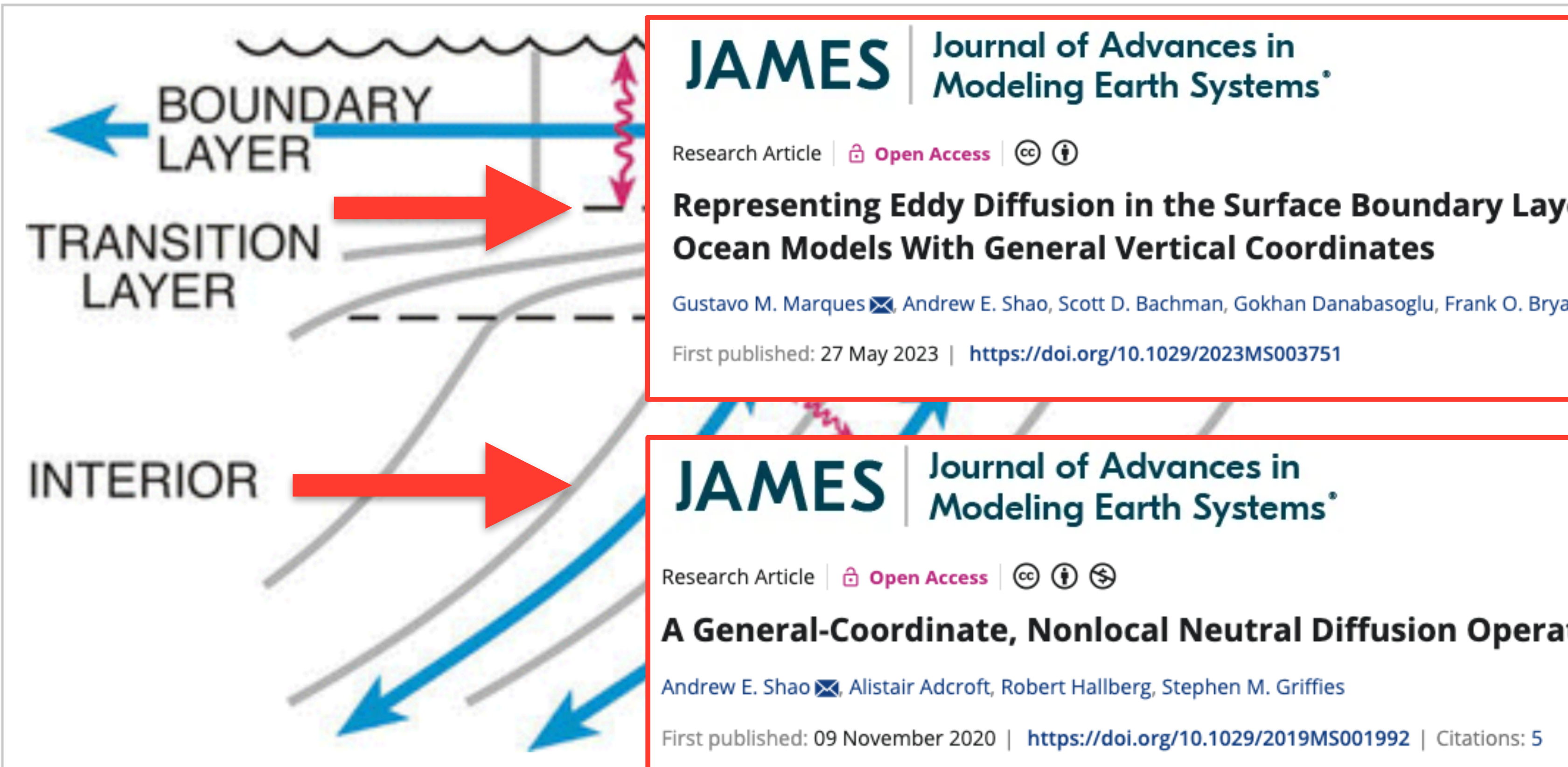
CESM OMWG, February 8, 2024

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Motivation: unification of mesoscale tracer mixing

The idea of (lateral) diabatic eddy fluxes near boundaries was initially proposed by Treguier et al., (1997).

Conceptual model of eddy fluxes in the upper ocean



JAMES | Journal of Advances in Modeling Earth Systems

Research Article | Open Access

Representing Eddy Diffusion in the Surface Boundary Layer of Ocean Models With General Vertical Coordinates

Gustavo M. Marques, Andrew E. Shao, Scott D. Bachman, Gokhan Danabasoglu, Frank O. Bryan

First published: 27 May 2023 | https://doi.org/10.1029/2023MS003751

JAMES | Journal of Advances in Modeling Earth Systems

Research Article | Open Access

A General-Coordinate, Nonlocal Neutral Diffusion Operator

Andrew E. Shao, Alistair Adcroft, Robert Hallberg, Stephen M. Griffies

First published: 09 November 2020 | https://doi.org/10.1029/2019MS001992 | Citations: 5

From: Ferrari et al., 2008

Defining eddy diffusivities (κ_{Redi})

2D Mesoscale Eddy Kinetic Energy (MEKE) (Jensen et al., 2015)**:**

MEKE

$$\partial_t E = \dot{E}_{bg} + \underbrace{\gamma_{GM}\dot{E}_{GM}}_{\text{SOURCE: Energy transferred from GM}} + \cancelto{0}{\gamma_\nu \dot{E}_\nu} - \underbrace{\left(2\frac{C_d}{H}|U_b|E_b\right)}_{\text{SINK: Energy lost to bottom drag}} + \underbrace{\mathcal{D}_E}_{\text{SMOOTHING: Diffusion}} \qquad (1)$$

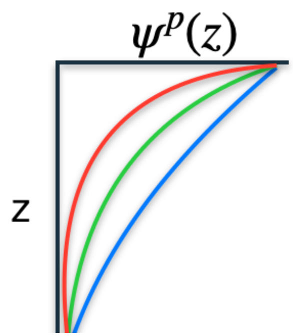
GEOMETRIC → (Marshall et al., 2012)

$$\kappa_{GM} = \alpha E \frac{N}{|\nabla b|} \qquad (2)$$

EBT

$$\kappa_{Redi}(z) = c\ \kappa_{GM}\ \psi^p(z) \qquad (3)$$

This is the default option in CESM/MOM6, with c = p = 1.



EBT is the equivalent barotropic vertical structure (non-dim).
Option to raise EBT to different powers (p = 1, 2, 3…).

Based on **QG PV** and neglecting the Beta term (Smith and Marshall, 2008 - JPO Eq. 4.3):

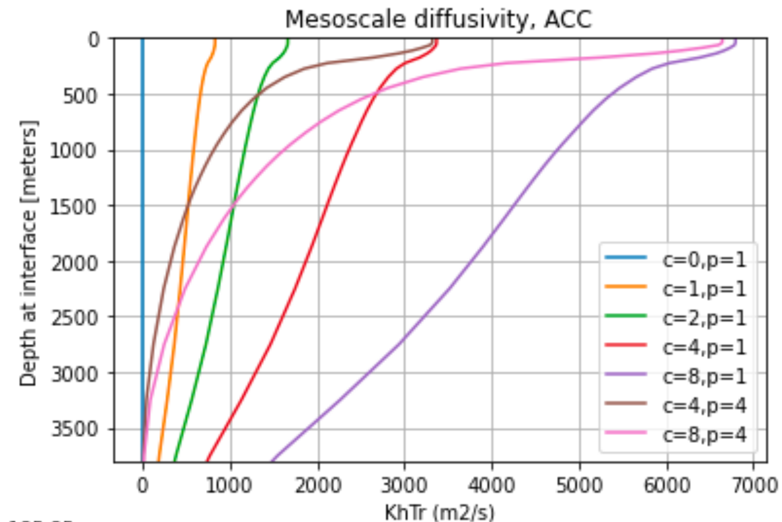
$$\kappa_{Redi}(z) = \kappa_{GM}(z) + \frac{\partial_z \kappa_{GM}(z)\ S}{\partial_z S} \qquad (4)$$

S = isopycnal slope.

Near-term plan to implement and test Eq (4).

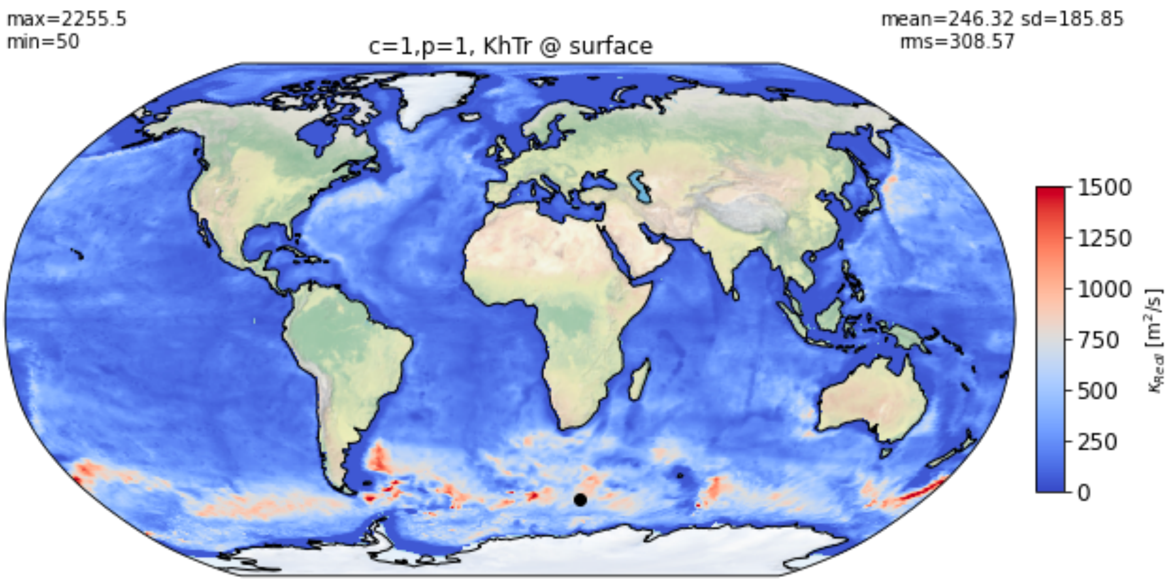
CESM workhorse forced simulations (1 cycle JRA-55)

Exploration of κ_{Redi} in forced experiments (7 cases) by varying **c** and **p** in equation (3).



$$\kappa_{Redi}(z) = c\ \kappa_{GM}\ \psi^{p}(z) \qquad (3)$$

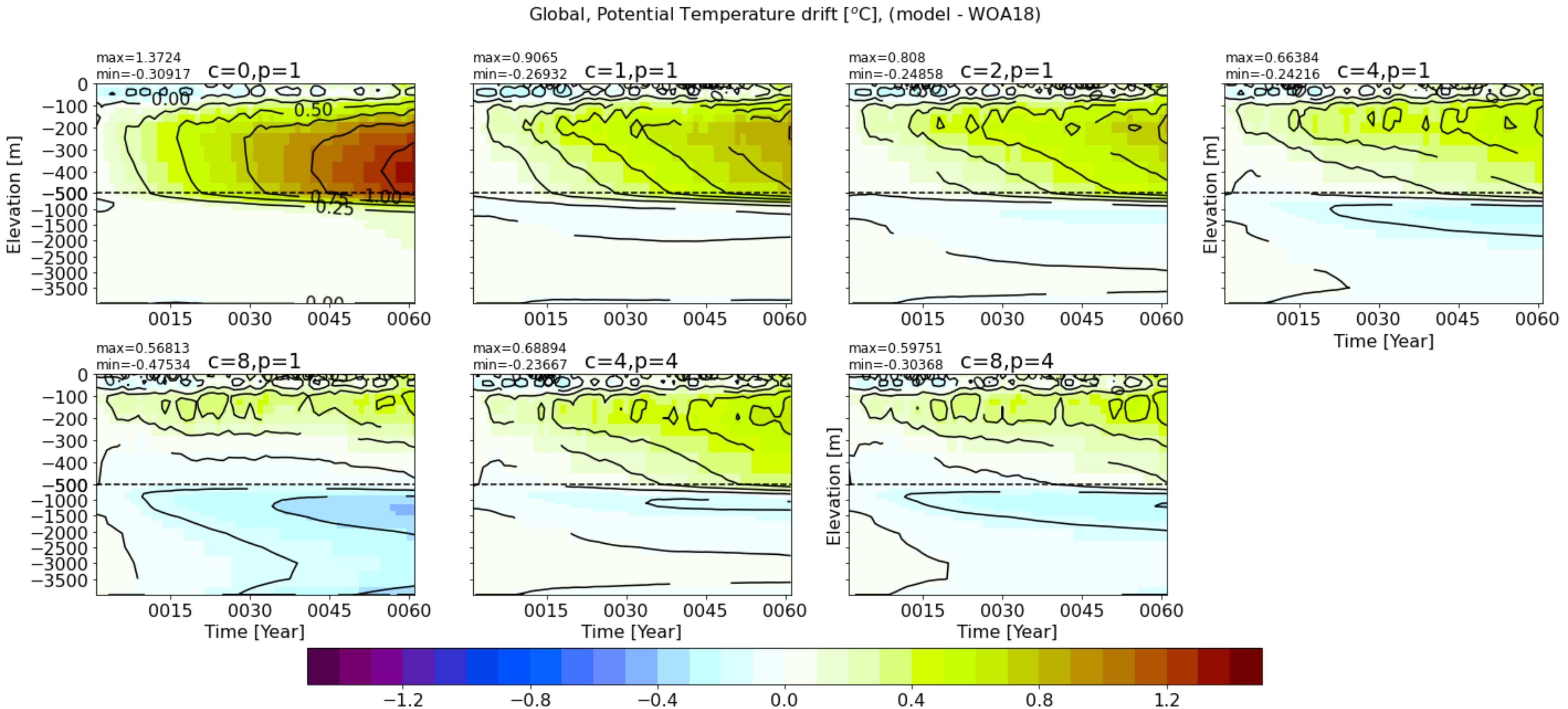
Tracer diffusivity (annual mean, year 61)



Overall structure does not change among the 7 cases;

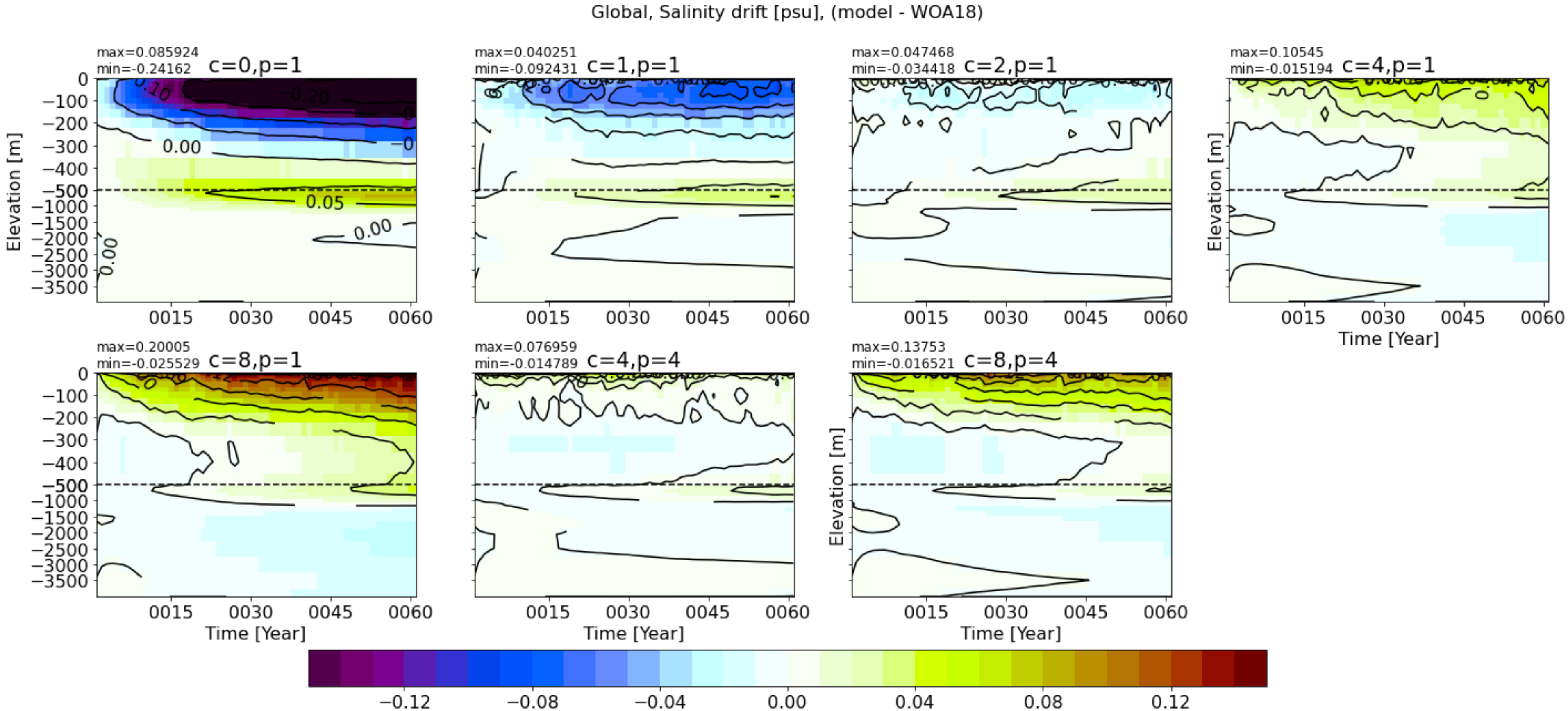
Lager values near western boundary currents and Southern Ocean.

Sensitivity of the global temperature drift to κ_{Redi}



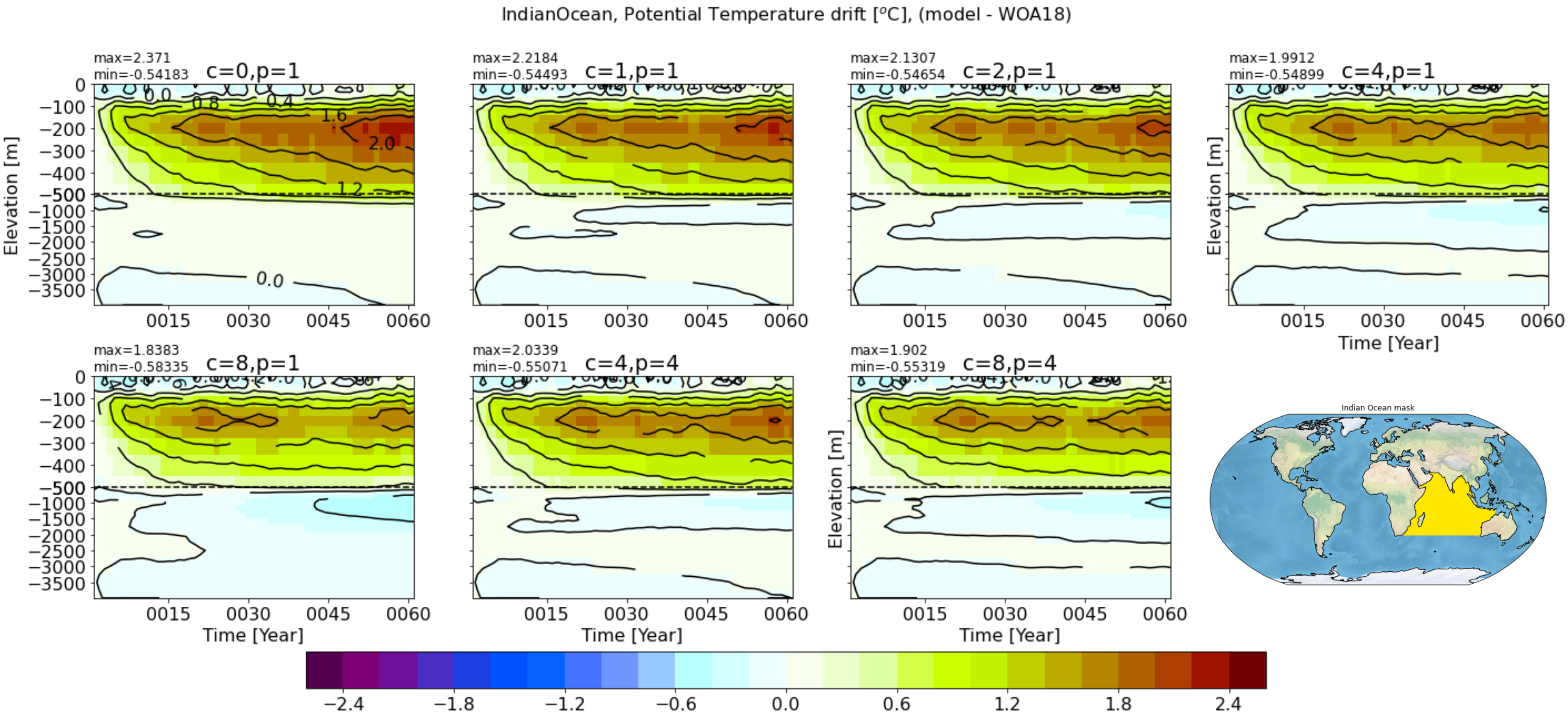
Changes in **c** and **p** (κ_{Redi}) affect global temperature drift.

Sensitivity of the global salinity drift to κ_{Redi}



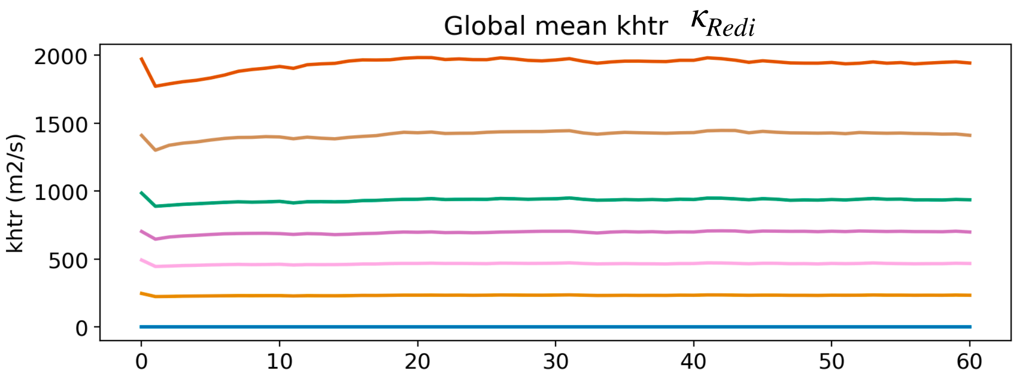
The effect of varying **c** and **p** (κ_{Redi}) on the global salinity bias is even more significant. It can flip the sign of the bias near the surface.

Sensitivity of the global temperature drift to κ_{Redi} (Indian Ocean)



Example of where κ_{Redi} has little effect on the temperature bias.

Mean T/S biases as a function of c and p (depth = 250 m)

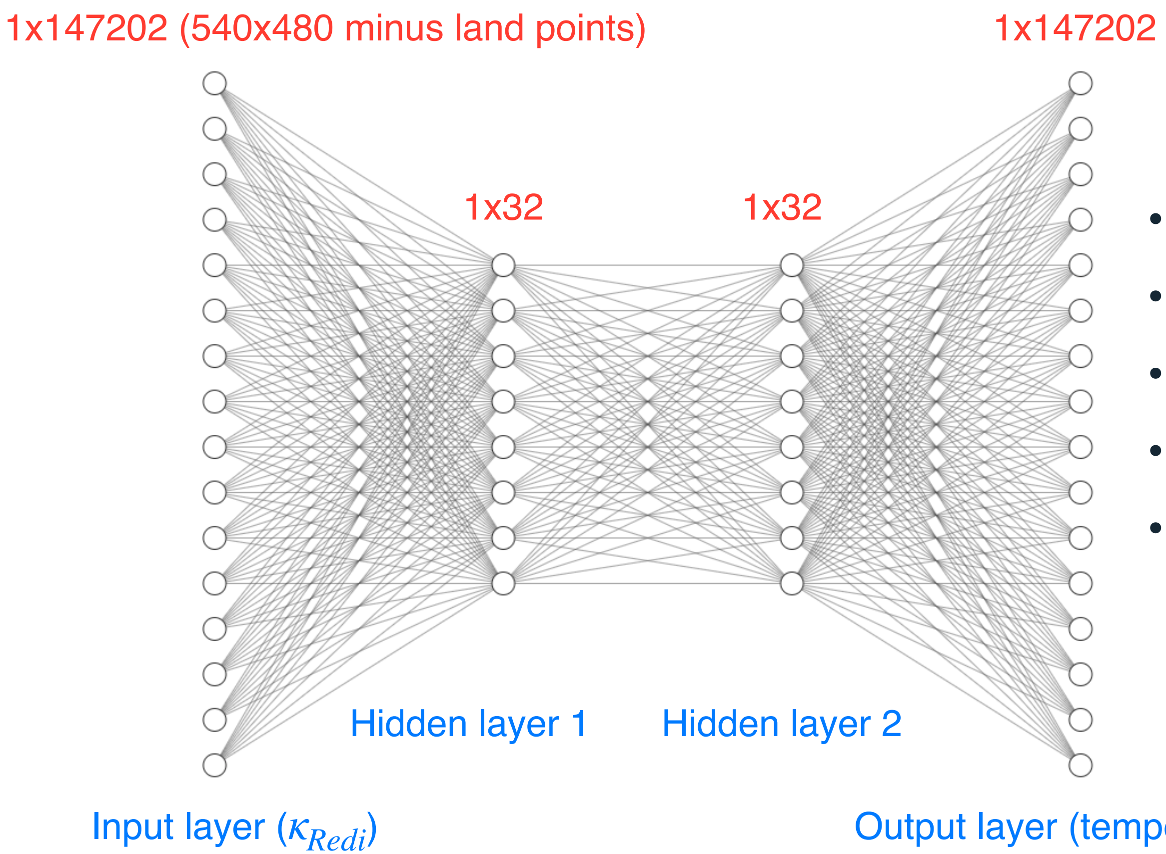


Global mean diffusivities equilibrate after a few years and then remain constant throughout the simulations;

Mean T&S biases grow with time and do not reach an "equilibrium";

Case c = 2, p = 1 **is not** used in the training/testing of the neural network (next slide).

A Neural Network model to predict temperature biases

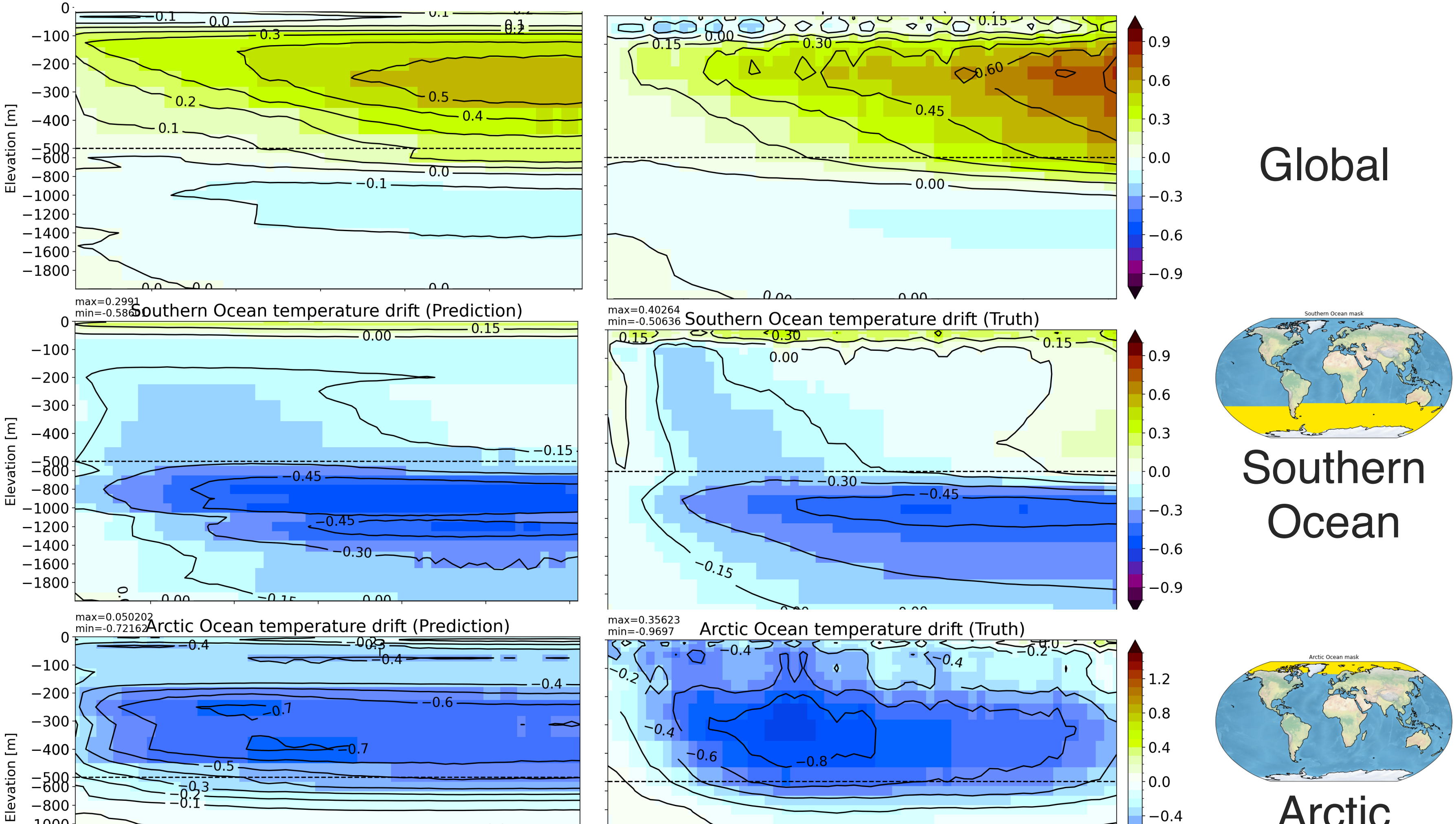


Hyperparameters:

- **# neurons** = 32;
- **# layers** = 2;
- **Activation function** = ReLu
- **Learning rate** = 1e-4;
- **Loss function** = mean absolute error.

- Hyperparameters optimized using KerasTuner's random search and manual exploration process;
- **For each vertical level**, a neural network was created using the “Sequential” model from Tensorflow Keras. It has ~ 8 million parameters trained using 4 Nvidia A100 GPUs.

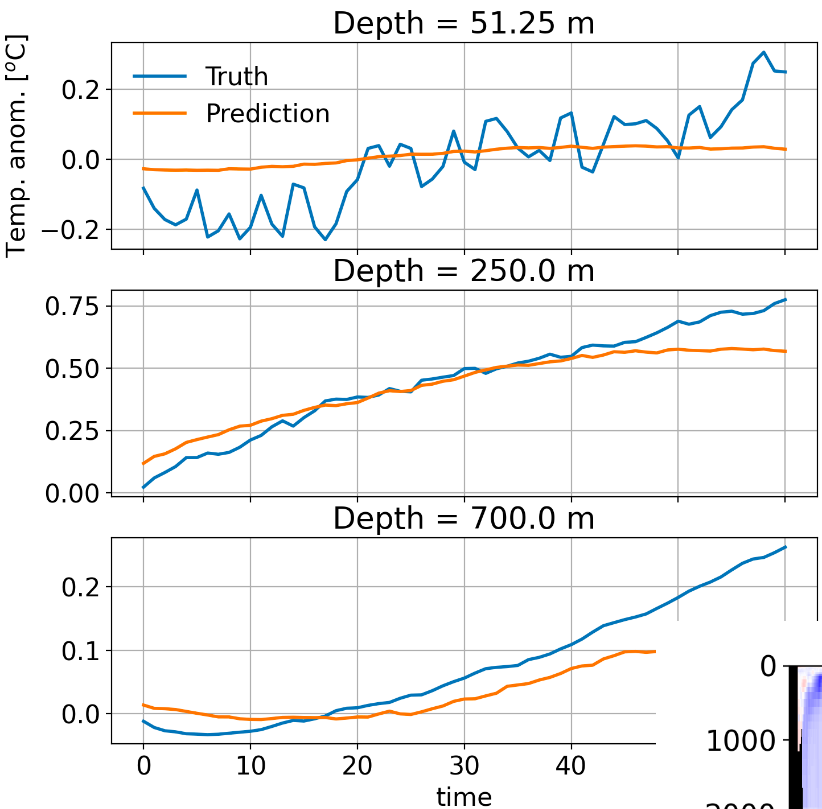
Neural network prediction vs truth: temperature bias (°C)



The NN model tends to mitigate (or “smooth out”) biases while generally capturing the overall pattern.

Neural network prediction vs truth: temperature bias (°C)

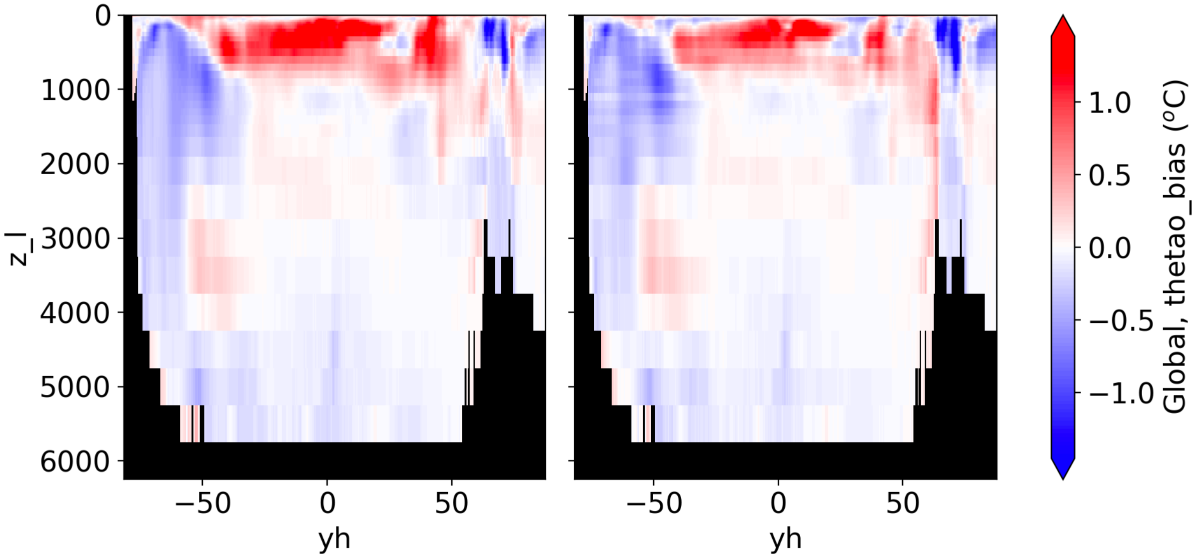
Global mean biases



NN model does not capture near-surface variability. κ_{Redi} has little influence at this depth;

Extreme values are attenuated.

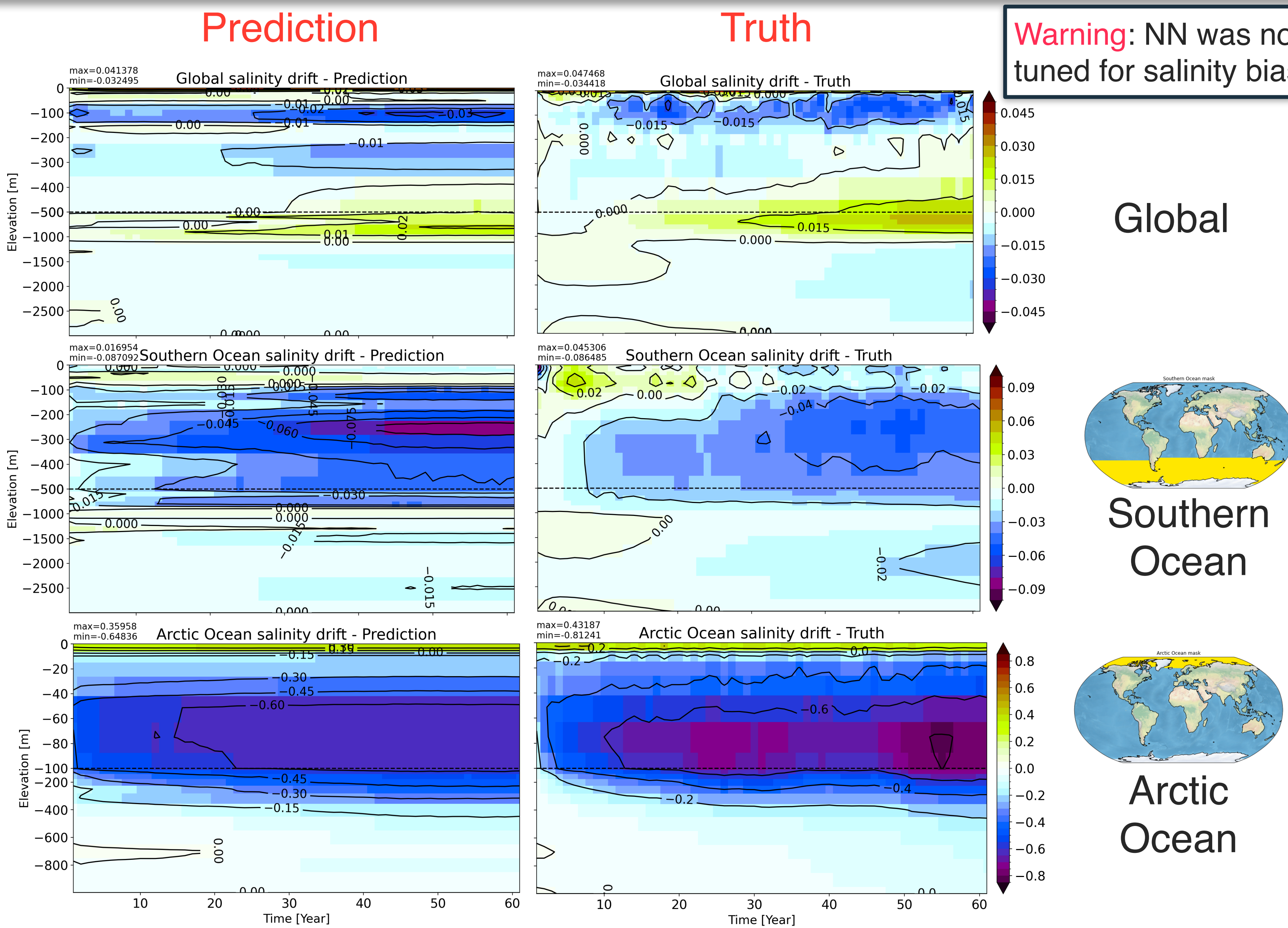
Zonal mean bias at year 61



Structure of zonal mean bias at year 61 is well represented.

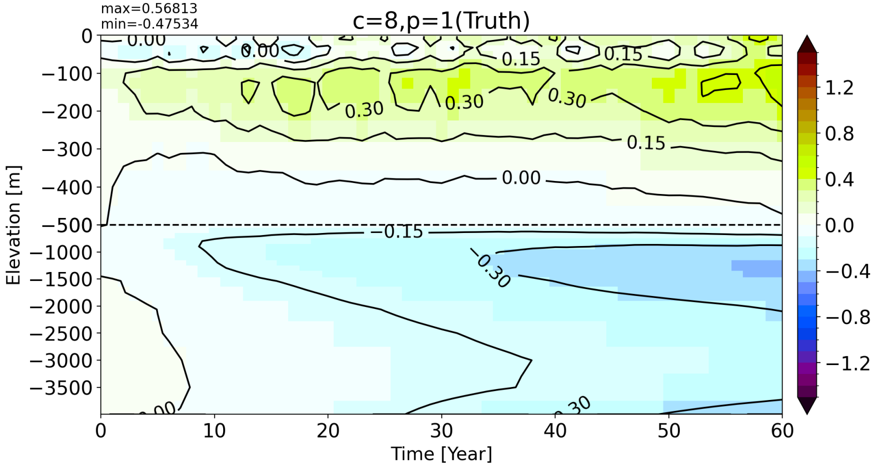
Neural network prediction vs truth: salinity bias (psu)

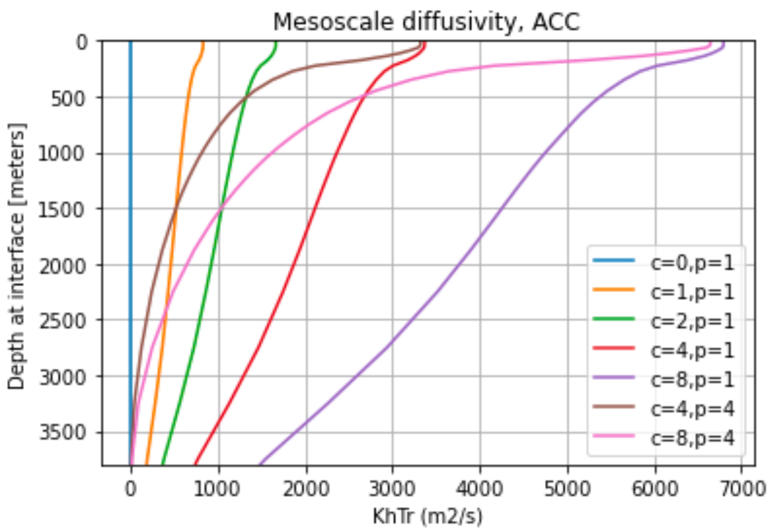
Warning: NN was not tuned for salinity bias



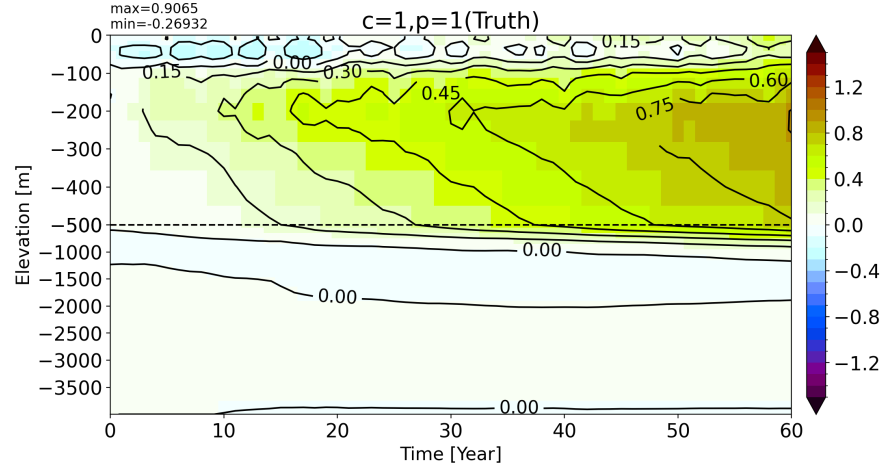
Case study: reduce near-surface global temperature bias

κ_{Redi} from case c=8,p=1 for depth = 0 to 800 m

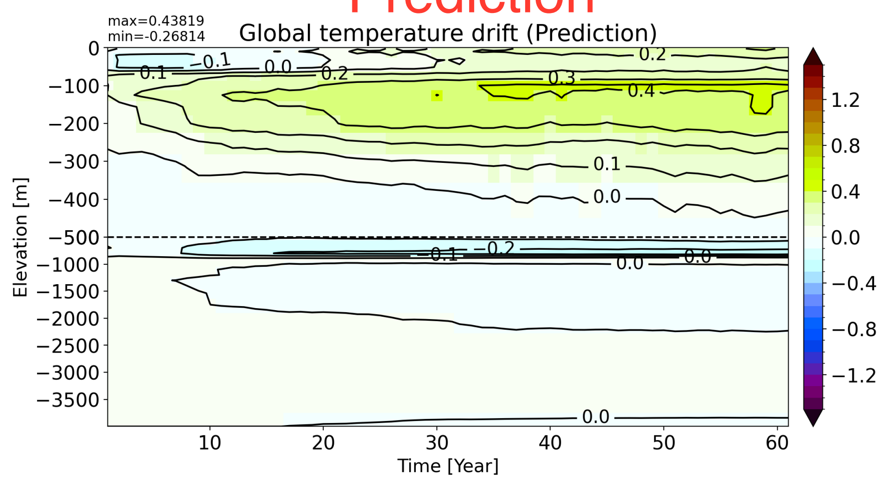




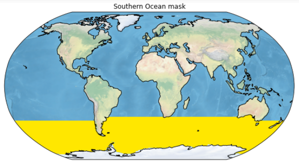
κ_{Redi} from case c=1,p=1 for depth > 800 m



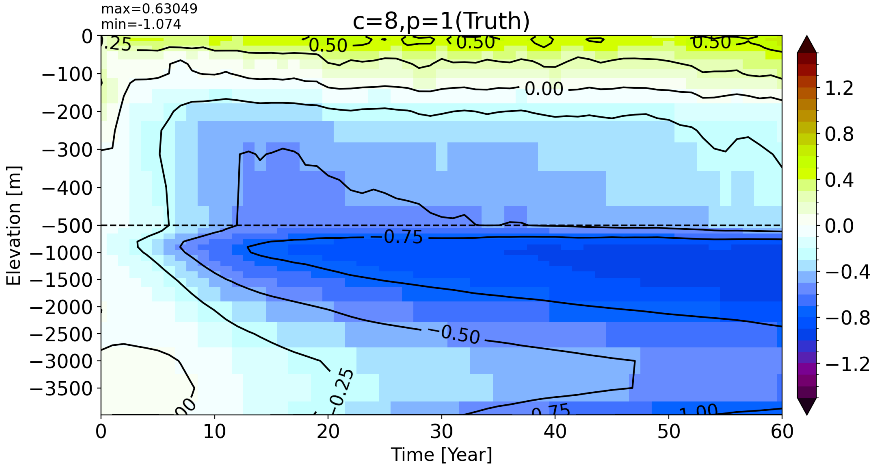
Prediction



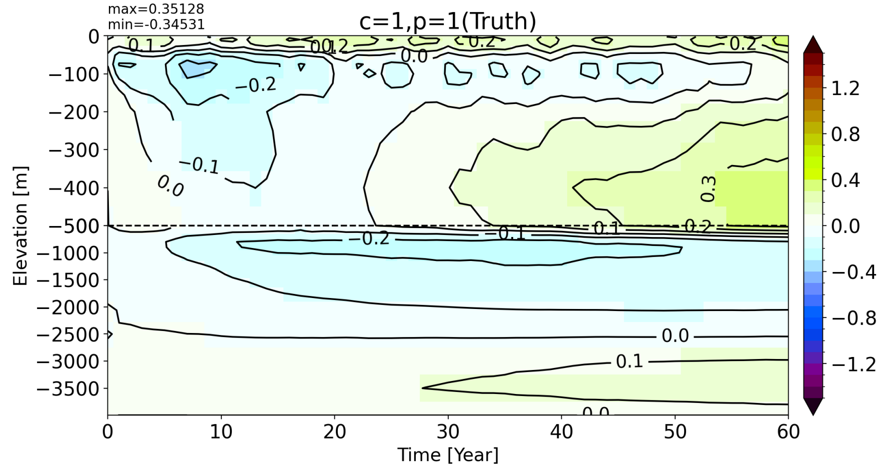
Case study: Southern Ocean temperature bias



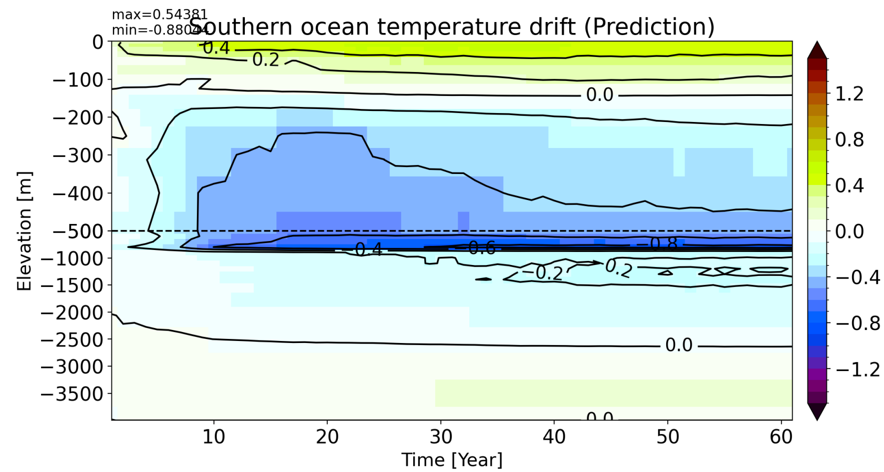
Case c=8,p=1



Case c=1,p=1



Prediction



Summary

Machine learning (ML) for offline model tuning (emulator):

- Neural network (NN) model to predict/emulate temperature biases using mesoscale diffusivities as input (work in progress);
- NN model tends to attenuate (or "smooth out") biases while generally capturing the overall pattern (good qualitatively);

Future work:

- Improve predictions by using more data and input "features";
- Try other ML techniques (convolutional NN, generative adversarial networks, random forest, gaussian process models);
- Apply this approach to emulate other metrics.

Thanks to LEAP NSF for the course on Machine Learning for Environmental Science by Pierre Gentine at Columbia University!

Thank you! gmarques@ucar.edu

Mean T/S biases as a function of c and p (depth = 51.25 m)