Learning Ocean Model Errors from Data Assimilation Increments

Tarun Verma, Feiyu Lu, Alistair Adcroft

02/09/2023

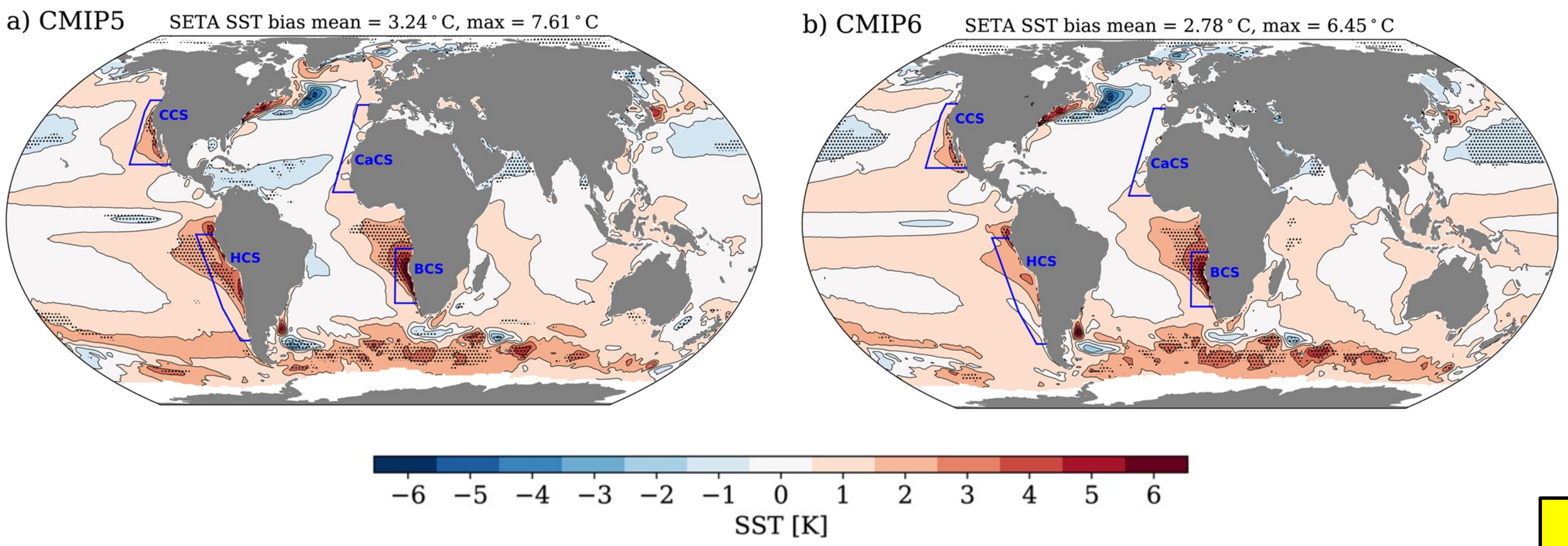
CESM Ocean Working Group Meeting

Bias: a persistent problem in climate modeling

Multi-model Mean SST Bias
(Farneti et al. 2022)

High-res MIP (1.78º C)

CMIP5 (3.25º C) **CMIP6 (2.78º C)**



Ocean MIP2 (1.14º C)

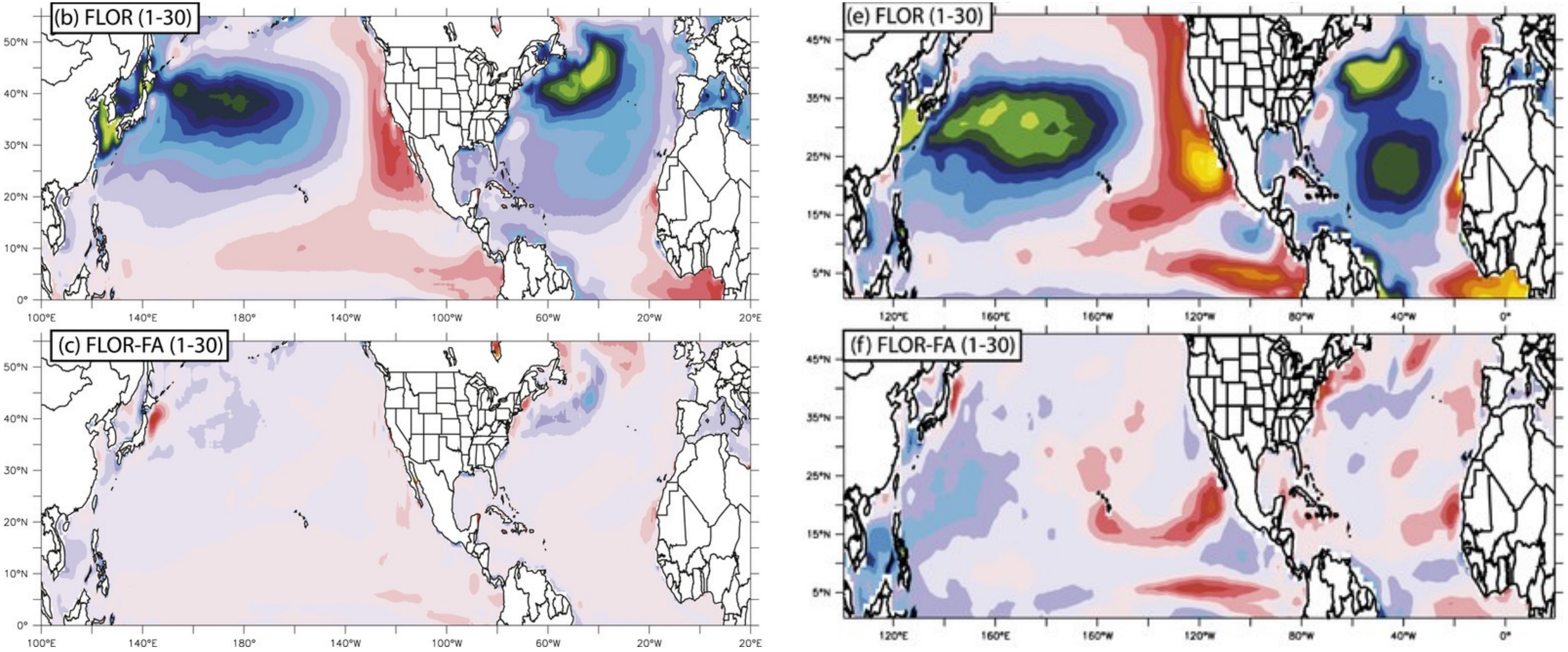
SST bias affects many other important phenomena, e.g., Tropical Cyclones (TC)

Reduction in SST bias improves TC activity (Vecchi et al. 2014)

SST bias **TC Potential Intensity**

Control

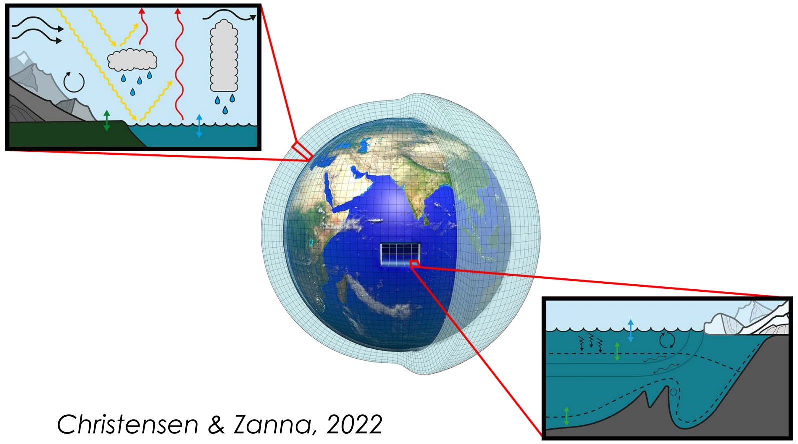
Flux-adjusted



Part of the bias stems from fast physics errors

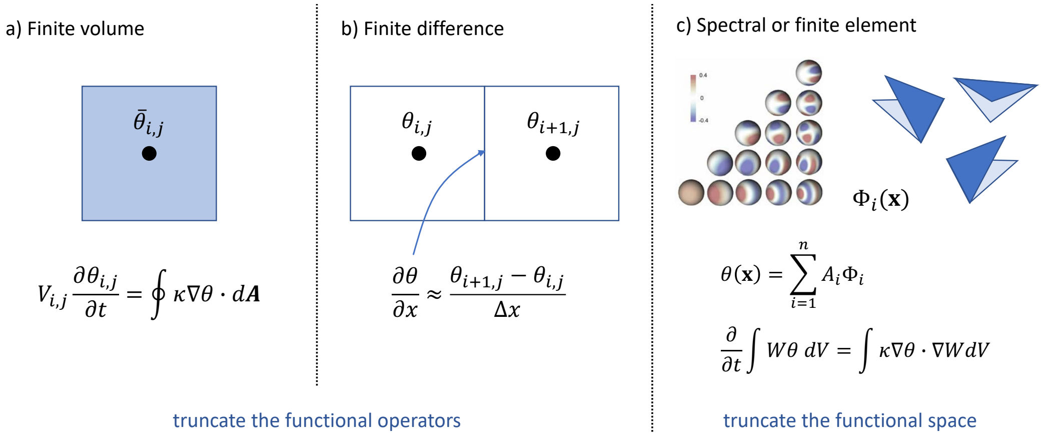
Model error

➡ Poor or lacking representation of key processes



Christensen & Zanna, 2022

➡ Numerics

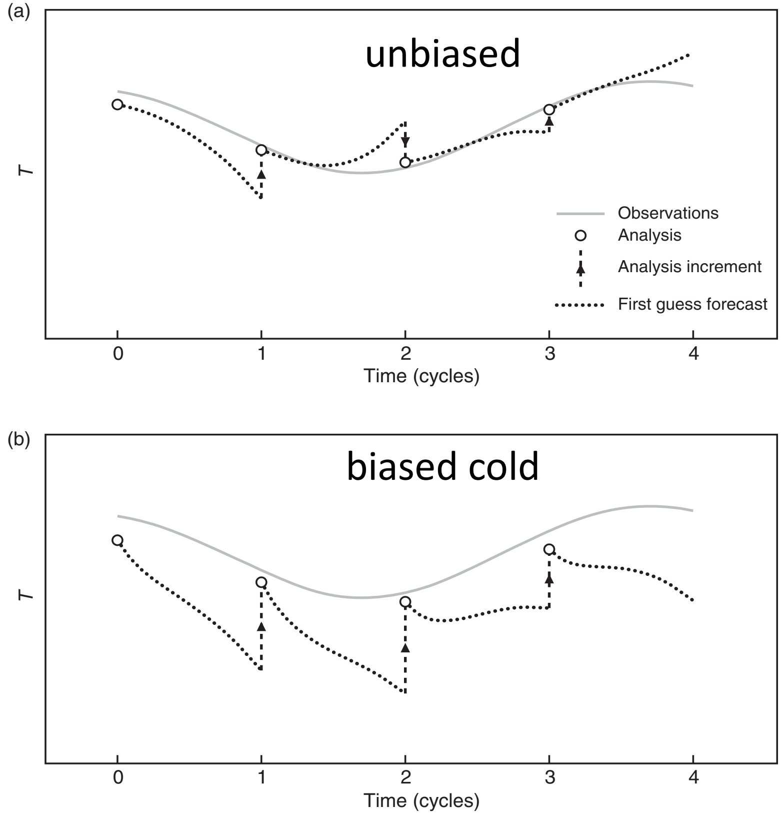


- Missing or inaccurate physics,
  - vertical mixing, sub-meso-mesoscale eddies etc.
- Numerics.

Here we propose to (machine) learn ocean model corrections at these fine-scales using data assimilation increments.

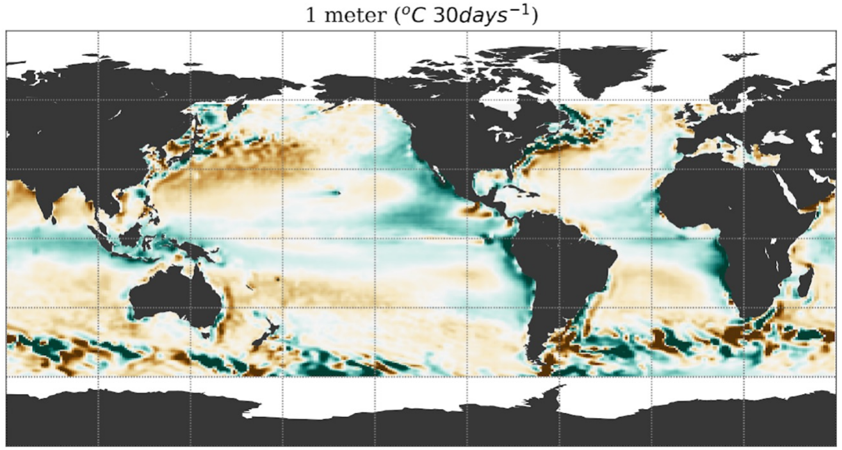
Data Assimilation (DA) Increments and Model Bias

Palmer and Weisheimer 2011

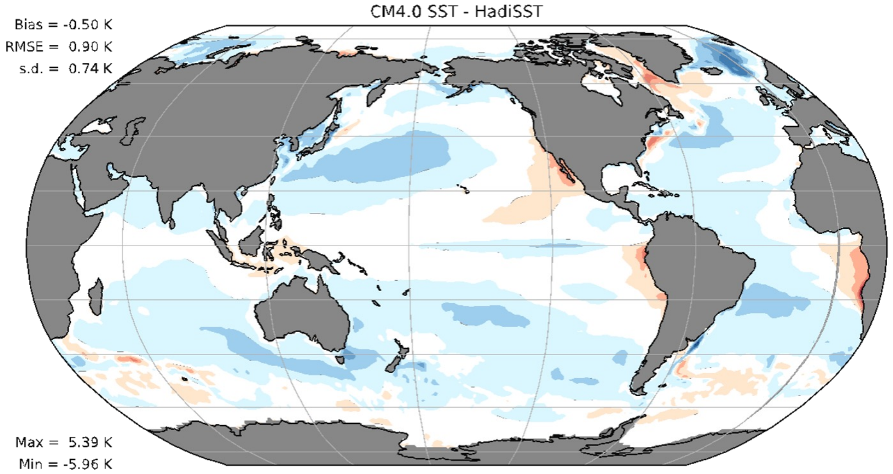


a) DA corrects random errors.
b) DA corrects systematic errors.
Learning corrections = Learning Increments

Mean Increments in GFDL SPEAR

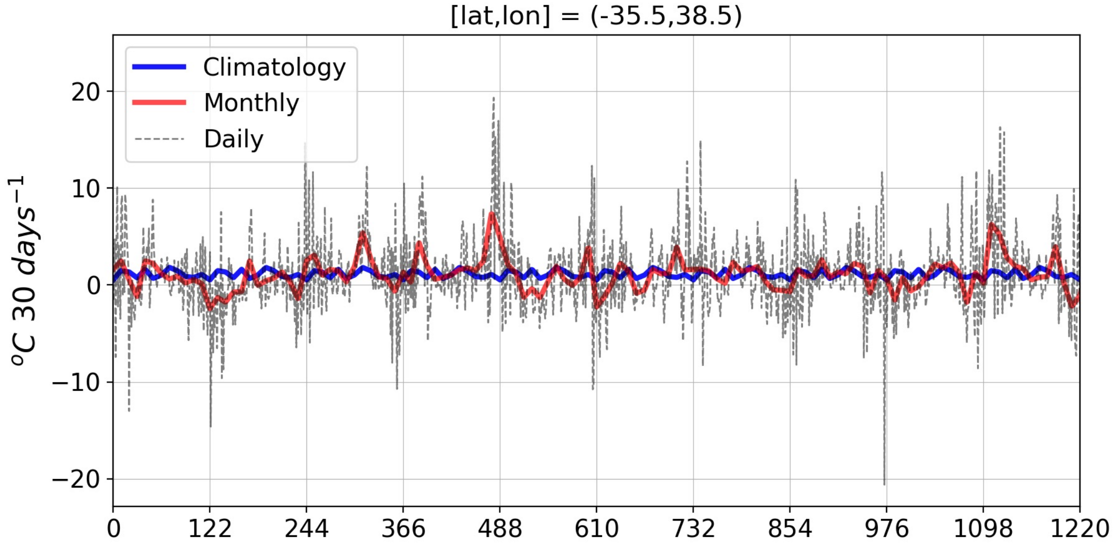


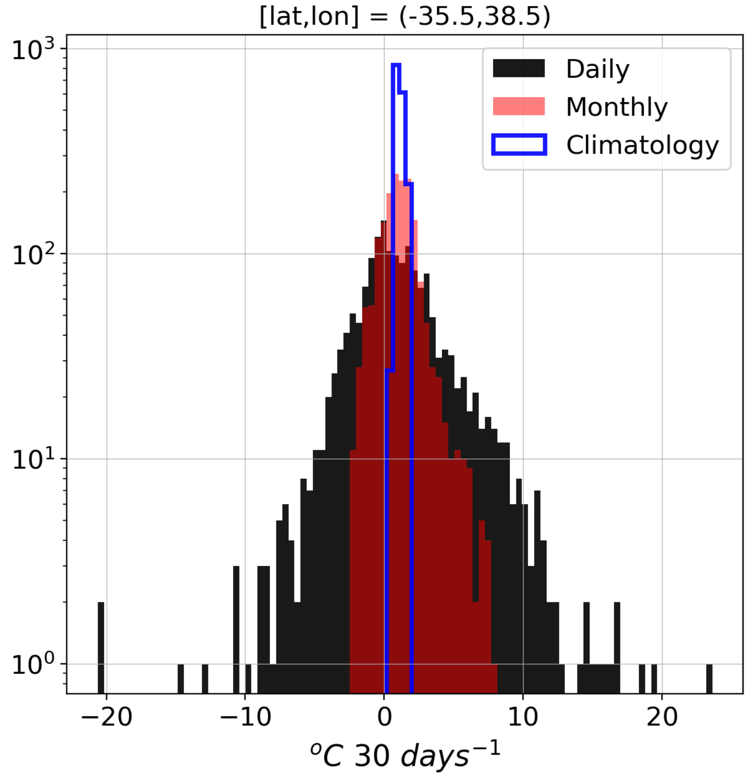
Mean SST Bias in CM4.0



Held et al. 2019

Learn high frequency or low frequency increments?



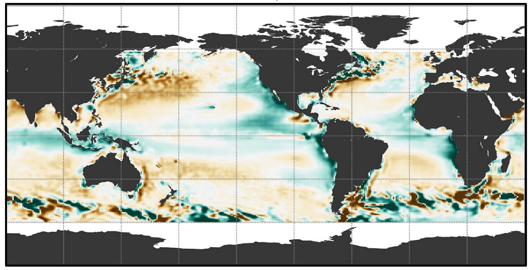


- $X_{t,s}$ is a daily DA increment with t and s indices representing time and space.
- $\langle X_{t,s} \rangle^{d,m,c}$ is a smoothing (or high frequency filtering) operation with daily, monthly or climatological timescales
- Loss function for the neural network can then be written as $\boldsymbol{E}\left[\left(\hat{X}_{t,s} - \langle X_{t,s} \rangle^{d,m,c}\right)^2\right]$

Use Neural Nets to learn nonlinear mapping from model state to DA increments

- ✓ **Goal:** local, state-dependent (~parameterization)
- ✓ **Inputs:** daily T, T_z, U_z,V_z, hfds, taux, tauy etc.
- ✓ **Output:**
  1) daily climatology of temperature increments
  2) raw daily temperature increments
- ✓ **Architecture:** Fully connected neural network (2h32 to 5h256)

60S-59N, 0-360E



- ✓ near-global domain
- ✓ sub-sampled to ~2^o horizontally
- ✓ coarsened to 19 levels in vertical
- ✓ sub-sampled every 3^{rd} day

- ✓ Using **GFDL SPEAR DA system** which assimilates SST, ARGO observations
- ✓ 2004-2021

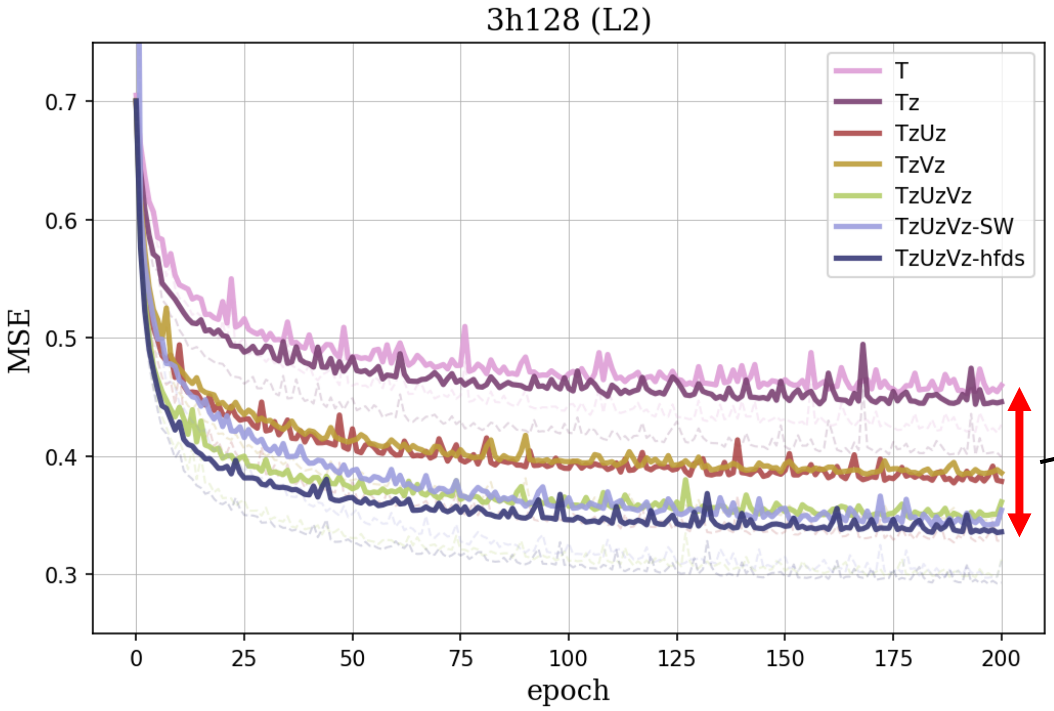
Train | Test

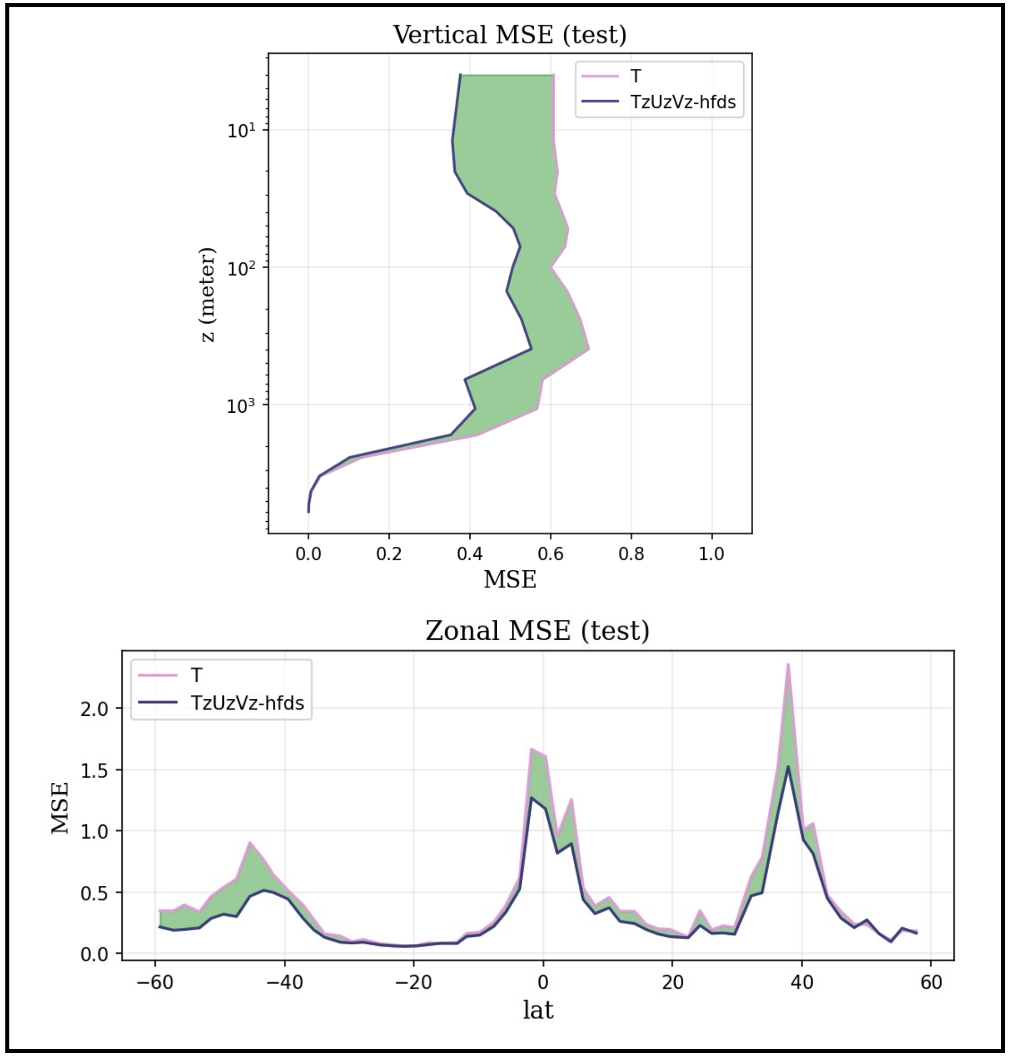
2008-2018 | 2019-21

1) Learning daily climatology of temperature increments

Sensitivity to different input predictors

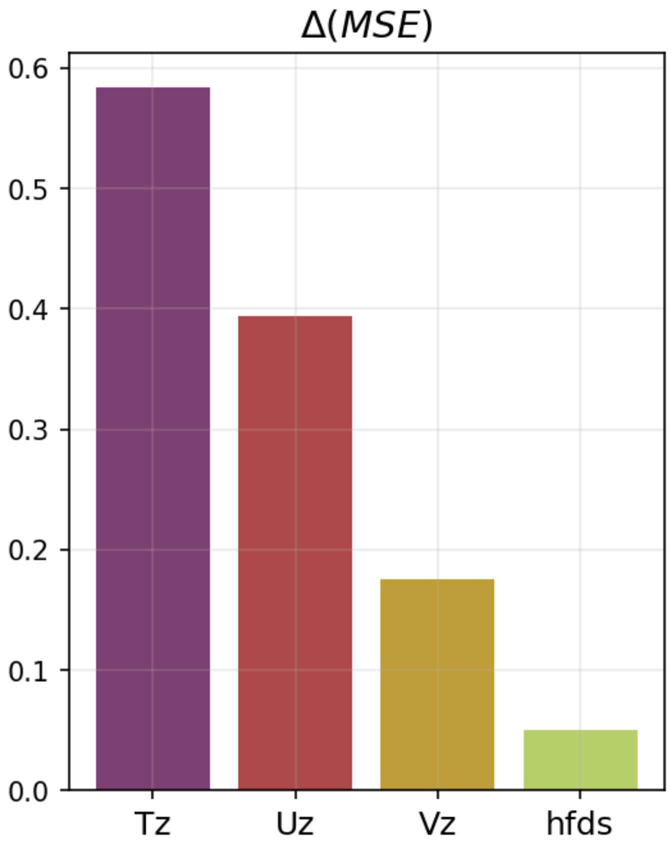
$[T]$, $[T_z]$, $[T_zU_z]$, $[T_zV_z]$, $[T_zU_zV_z]$, $[T_zU_zV_z\,sw]$, $[T_zU_zV_z\,hfds]$

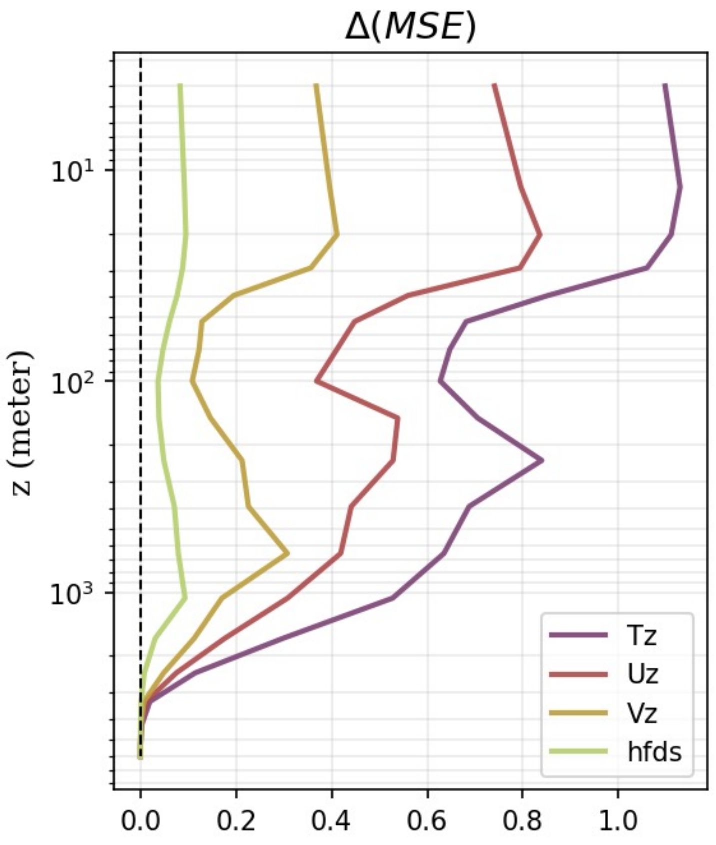




Change in MSE due to randomizing input predictors

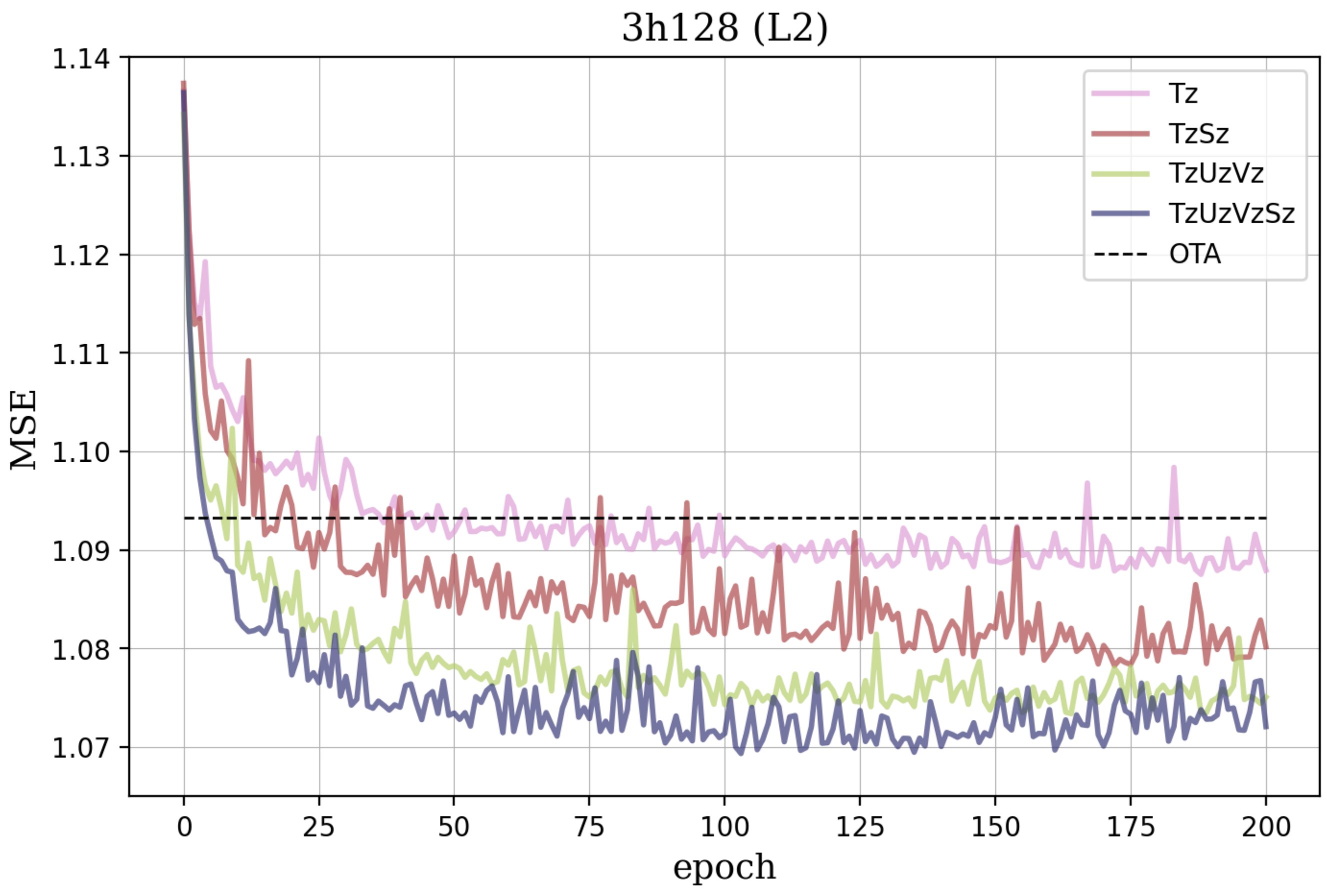
$$\overline{\delta T} = \hat{\mathcal{F}}[T_z, Uz, Vz, hfds; \theta] \quad ; \text{NN} = 3\text{h}128$$





2) Learning raw daily temperature increments

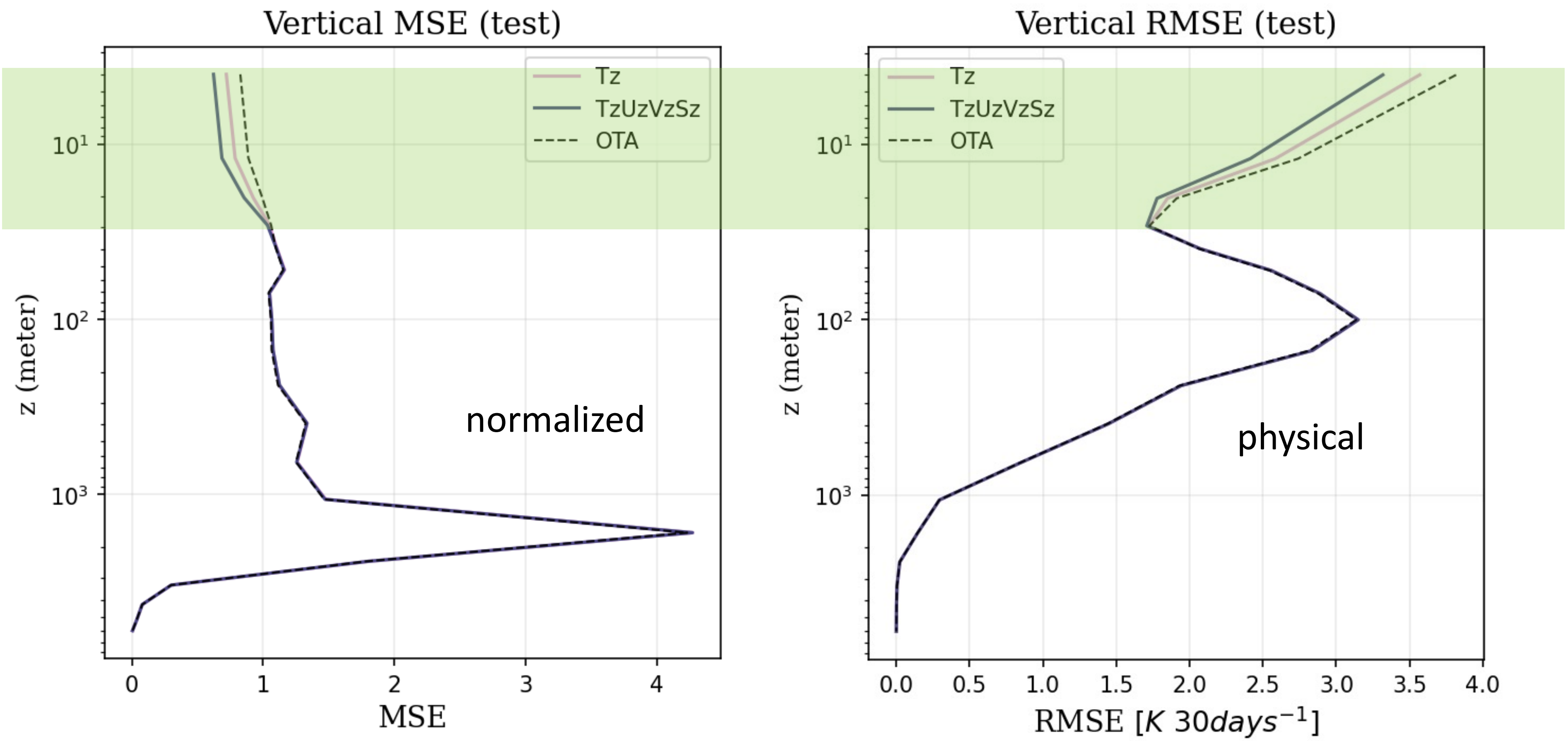
Sensitivity to different input predictors



OTA = benchmark based on daily climatology of temperature increments

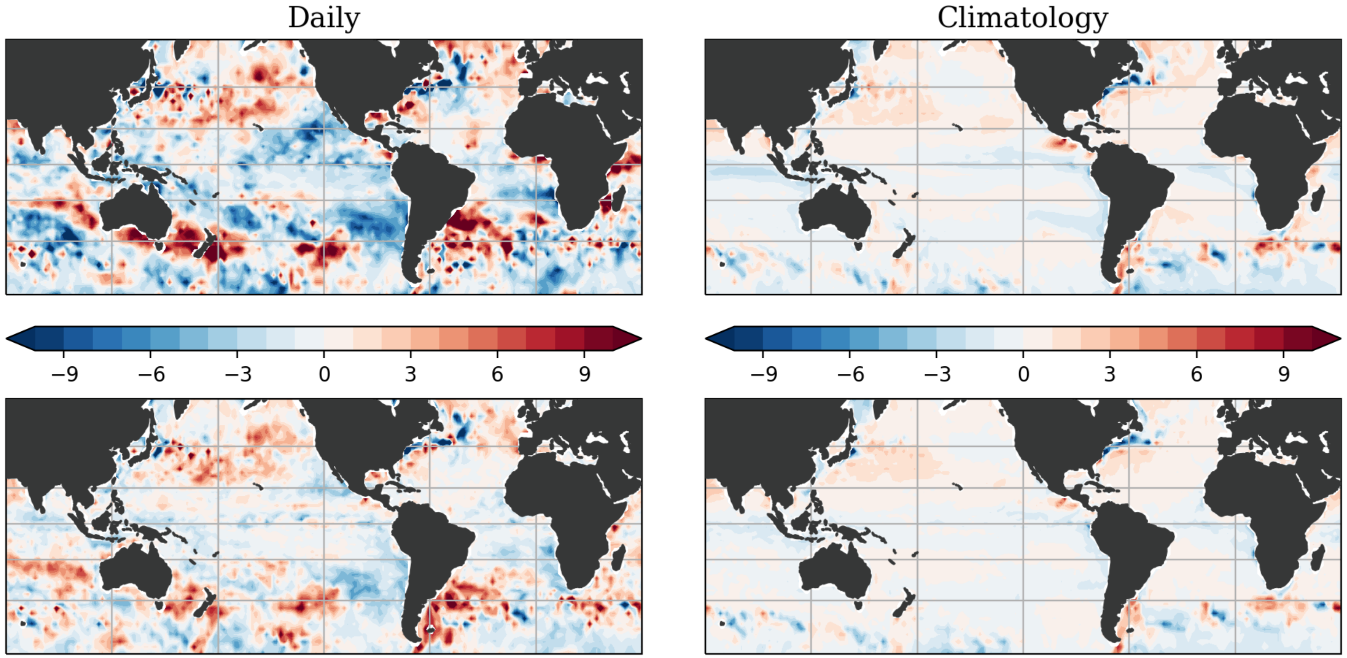
Vertical Structure of MSE/RMSE

OTA = benchmark based on daily climatology of temperature increments



	OTA	Tz	TzUzVzSz
R^2 (0-30m) (%)	12.1	18.4	25.0

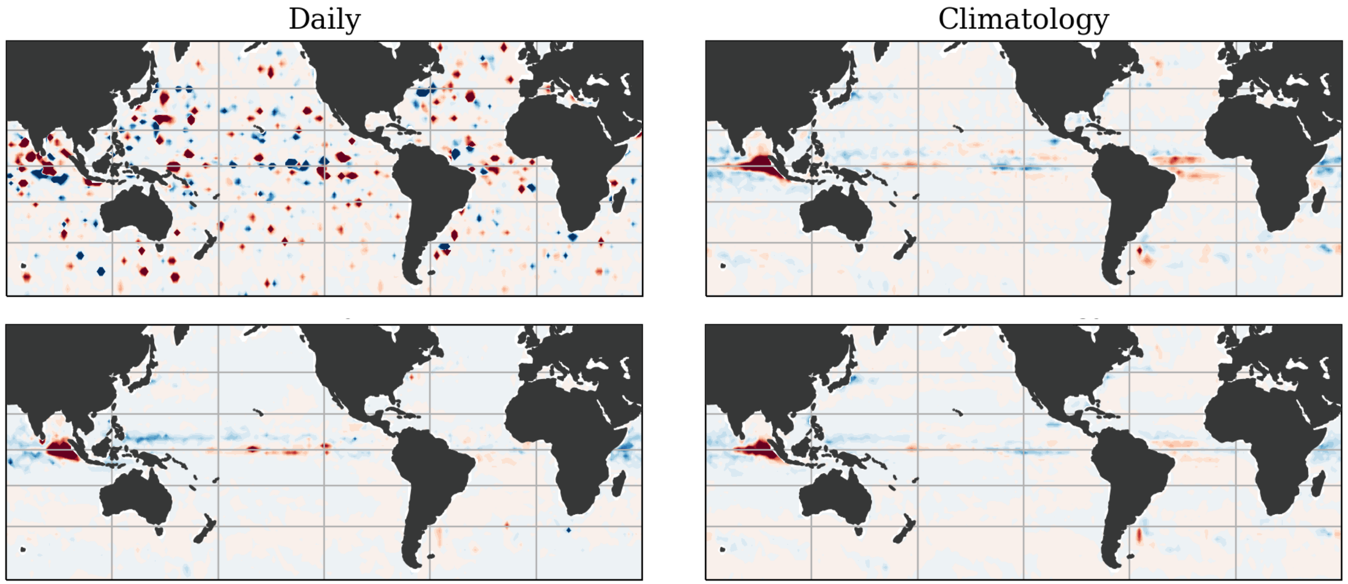
2019-01-01 Truth and Predictions (learning daily vs climatology)



Truth

4 meter (K 30-days^{-1})

NN

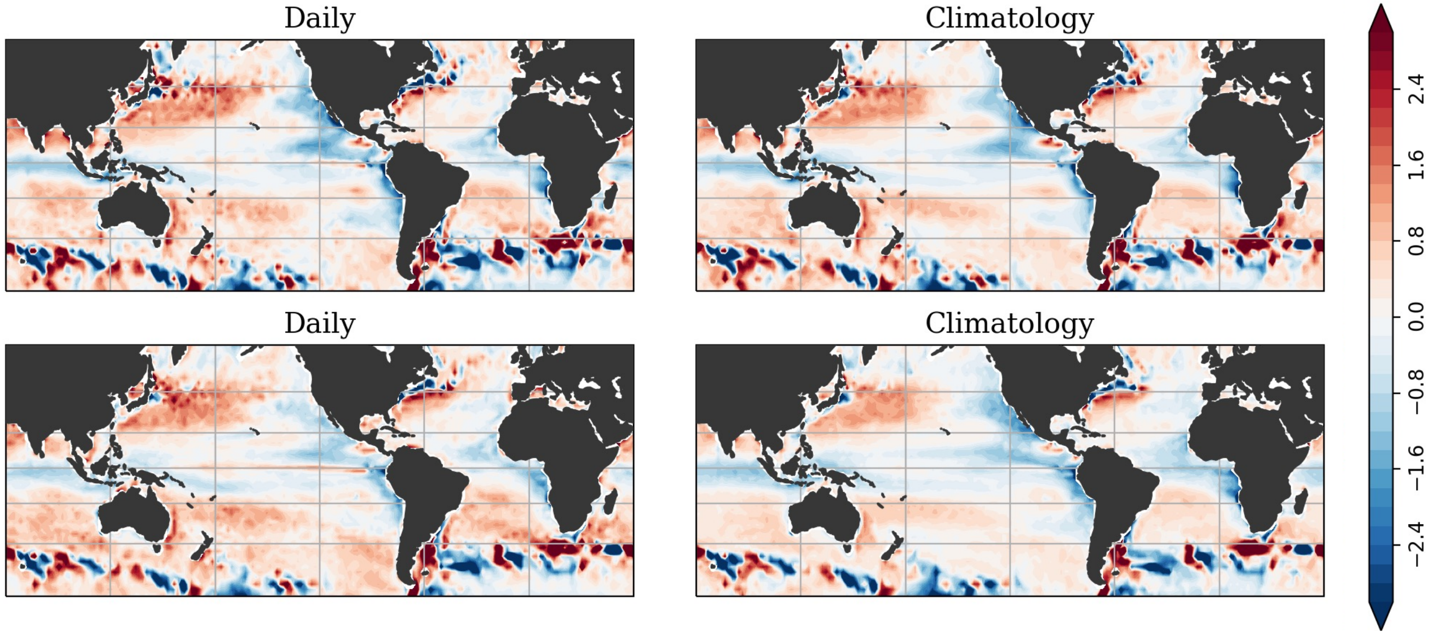


Truth

100 meter (K 90-days^{-1})

NN

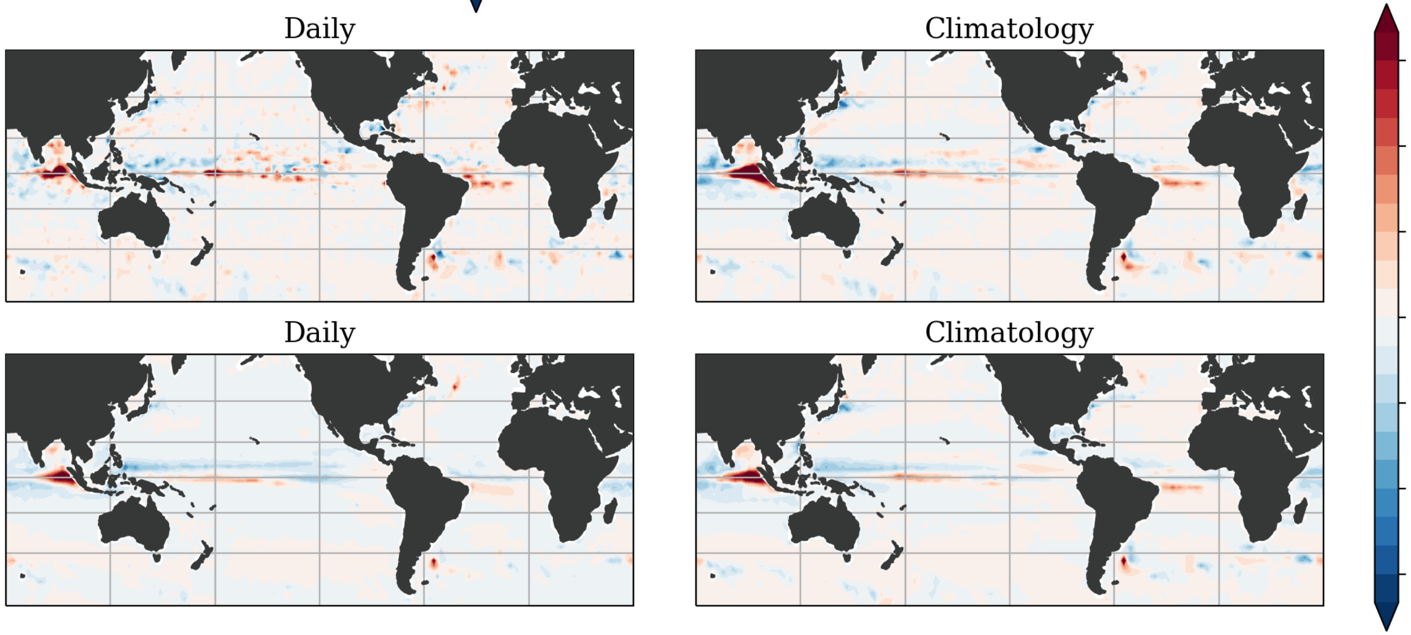
2019-2021 mean Truth and Predictions (learning daily vs climatology; K 30-days^{-1})



Truth

4 meter

NN



Truth

100 meter

NN

Conclusions

- ✓ DA increments can be used to learn systematic fast ocean model errors.
- ✓ Neural network shows significant offline skill in learning nonlinear relationships between model state and DA increments.
- ✓ The offline skill improves with the addition of dynamically relevant quantities.
- ✓ A network trained on raw daily temperature increments has better skill in the upper 30 meters which relaxes to the one trained on daily climatology of temperature increments below that.

Next…

- ✓ Expand to global domain, remove subsampling, coarsening
- ✓ Joint prediction of temperature and salinity increments.
- ✓ Interpretation and learn missing/inaccurate physics
- ✓ Online inference and evaluation