Reconstructing the Tropical Pacific Upper Ocean using Online Data Assimilation with a Deep Learning model

Zilu Meng (zilumeng@uw.edu), Gregory J. Hakim
Department of Atmospheric Sciences, University of Washington

QR code for **paper**

QR code for **code**

Grate Project

University of Washington Atmospheric Sciences

Key Point

- A deep learning (DL) model exhibits superior prediction skill in the tropical Pacific compared to a linear inverse model.
- Implement a novel inflation technique by incorporating noise from hindcast experiments to address DL signal damping.
- Data assimilation (DA) on a sparse network of observations accurately reconstructs the monthly upper ocean spatial fields.
- The superior reconstruction skill of the deep learning model stems from its enhanced prediction skill.

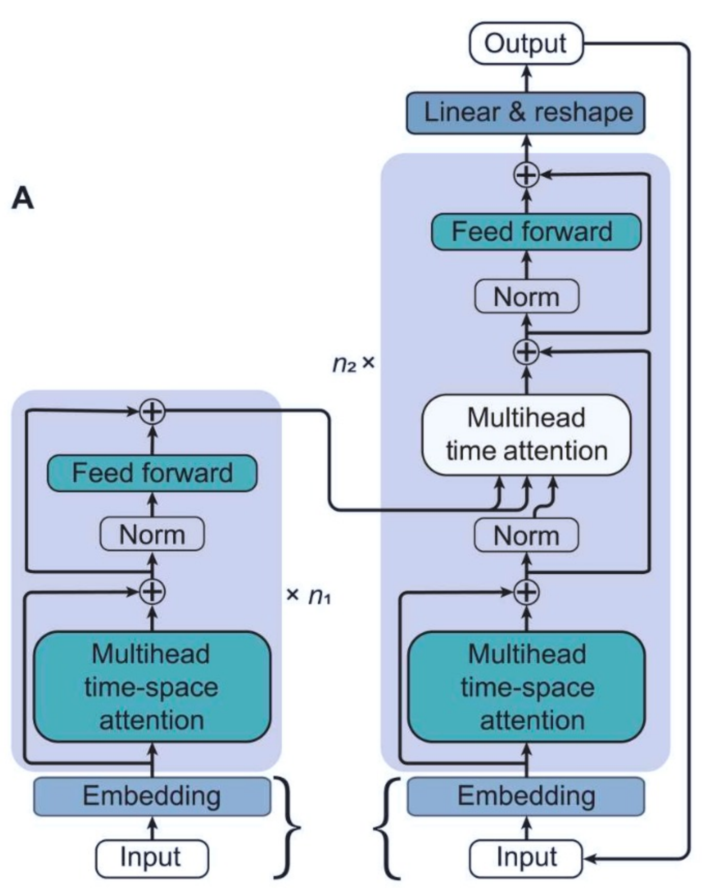
Baseline: Linear Inverse Model

$$\frac{d\boldsymbol{x}}{dt} = \mathbf{L}\boldsymbol{x} + \boldsymbol{\xi}$$

$\mathbf{L}$: deterministic Matrix; $\boldsymbol{\xi}$: random noise; $\boldsymbol{x}$: state vector (PCs from EOF).

$$\mathbf{G}_\tau = \mathbf{C}(\tau)\mathbf{C}(0)^{-1} \qquad \mathbf{C}(\tau) = <\boldsymbol{x}^T(\tau)\boldsymbol{x}(0)> \qquad \mathbf{G}_\tau = exp(\mathbf{L}\tau)$$

Deep Learning Model (Zhou&Zhang, 2023)



$$\boldsymbol{X}^{\text{out}}_{t+1:t+12} = \boldsymbol{DL}(\boldsymbol{X}^{\text{in}}_{t-12:t})$$

$\boldsymbol{X}$: Surface wind stress, upper ocean temperature.

Model structure: self-attention and encoder-decoder structure.

Training: trained on CMIP6, verified on Simple Ocean Data Assimilation (SODA) products, tested on Global Ocean Data Assimilation System (GODAS) reanalysis (Independent verification).

Loss: all fields RMSE.

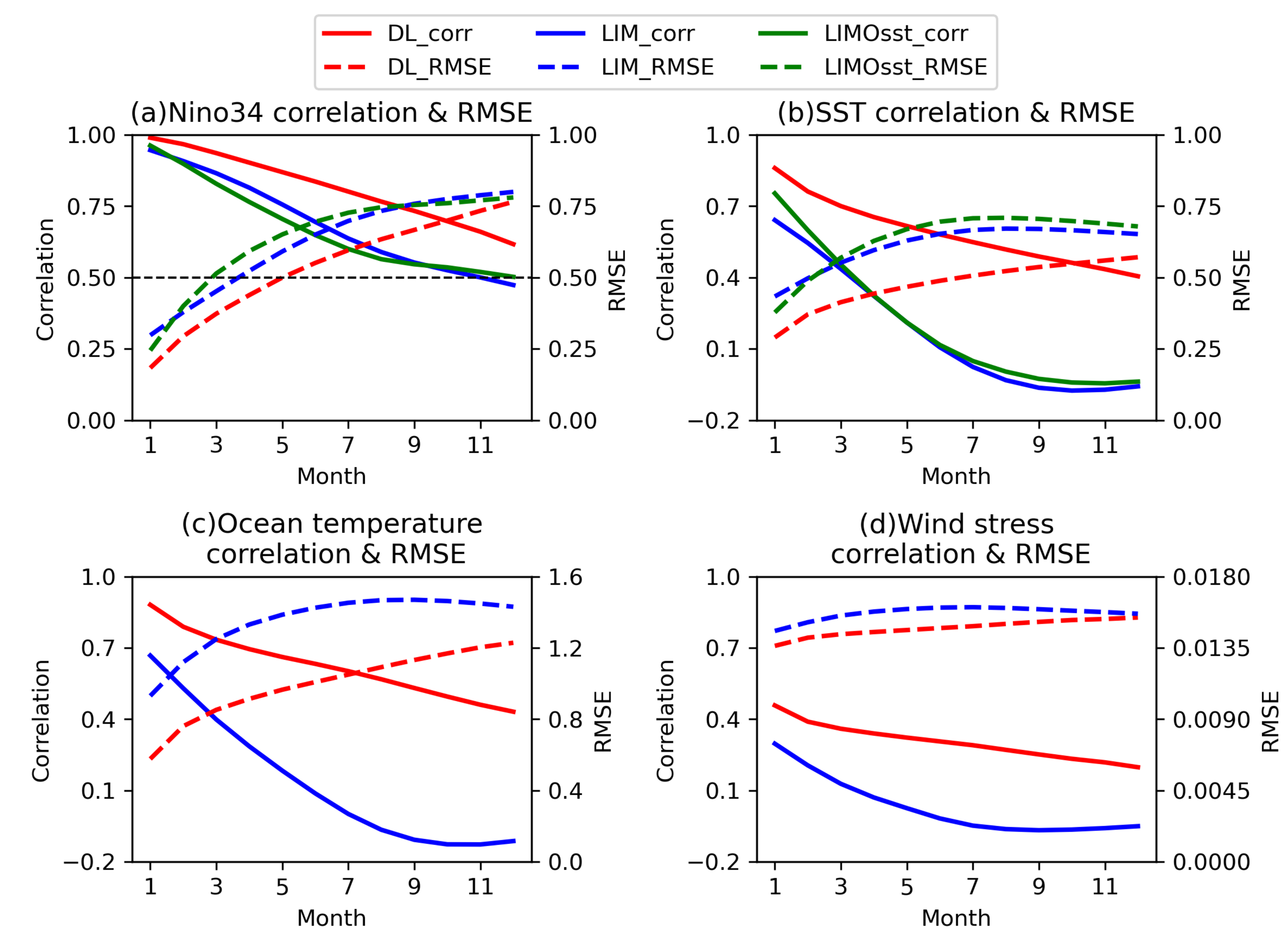
Advantages: Capture the nonlinear dynamics; No need to truncate data; Trained on larger ensemble data.

DA Method: Ensemble Kalman Filter

$$\boldsymbol{x}_a = \boldsymbol{x}_p + \mathbf{K}\left[\boldsymbol{y} - \mathcal{H}(\boldsymbol{x}_p)\right] \qquad \boldsymbol{x}_{p,t+1} = \mathcal{M}(\boldsymbol{x}_{a,t})$$

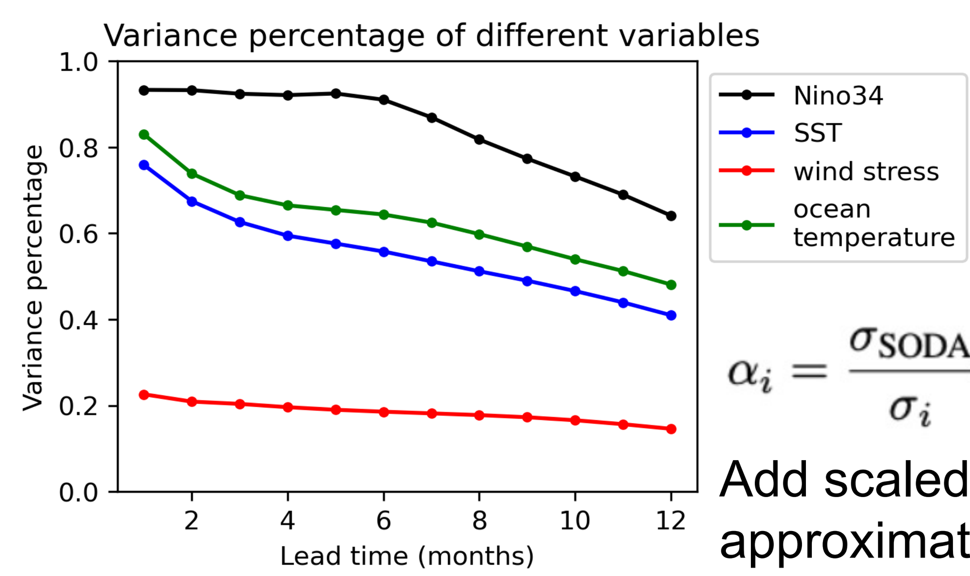
$\boldsymbol{x_a}$: posterior; $\boldsymbol{x_p}$: prior; $\mathbf{K}$: Kalman gain; $\boldsymbol{y}$: Observation; $\mathcal{H}$: observation operator. $\mathcal{M}$: Model operator.

Prediction Skill



DL forecasts consistently outperform both the LIM and LIMOsst (SST only LIM) forecasts across all variables and at all lead times.

DL Model Ensemble Inflation



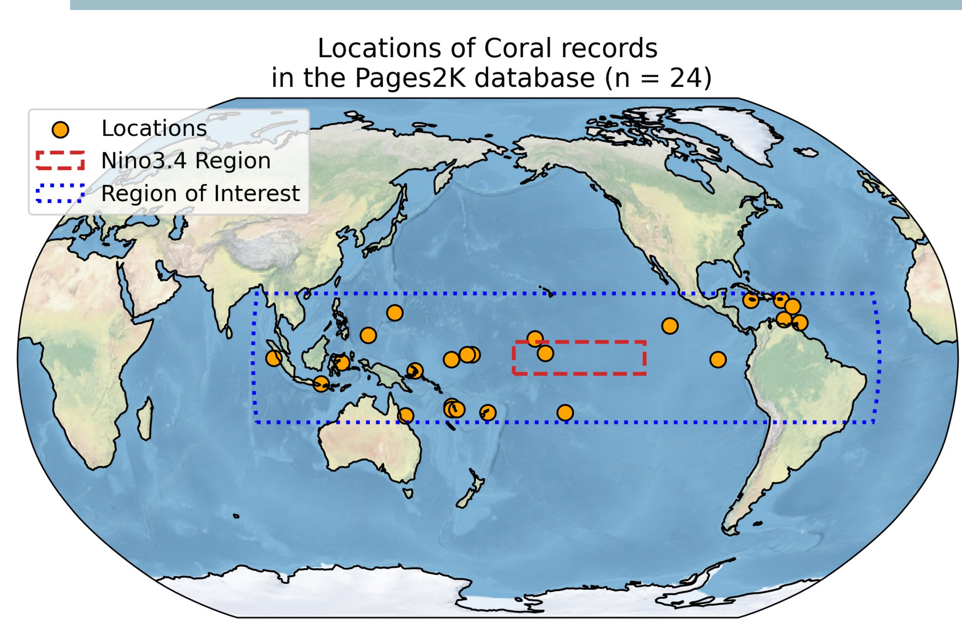
DL models tend to lose error associated with the unpredictable signal.

$$\boldsymbol{\eta}_{m,l,i} = \alpha_i(\boldsymbol{x}^{\text{true}}_{m,l,i} - \boldsymbol{x}^{\text{out}}_{m,l,i})$$

$$\alpha_i = \frac{\sigma_{\text{SODA}}}{\sigma_i} \qquad \boldsymbol{x}_{t+1} = \boldsymbol{x}^{\text{out}}_{m,l} + \boldsymbol{\eta}_{m,l}$$

Add scaled noise from hindcasting to better approximate errors.

Pseudo Proxy Experiment

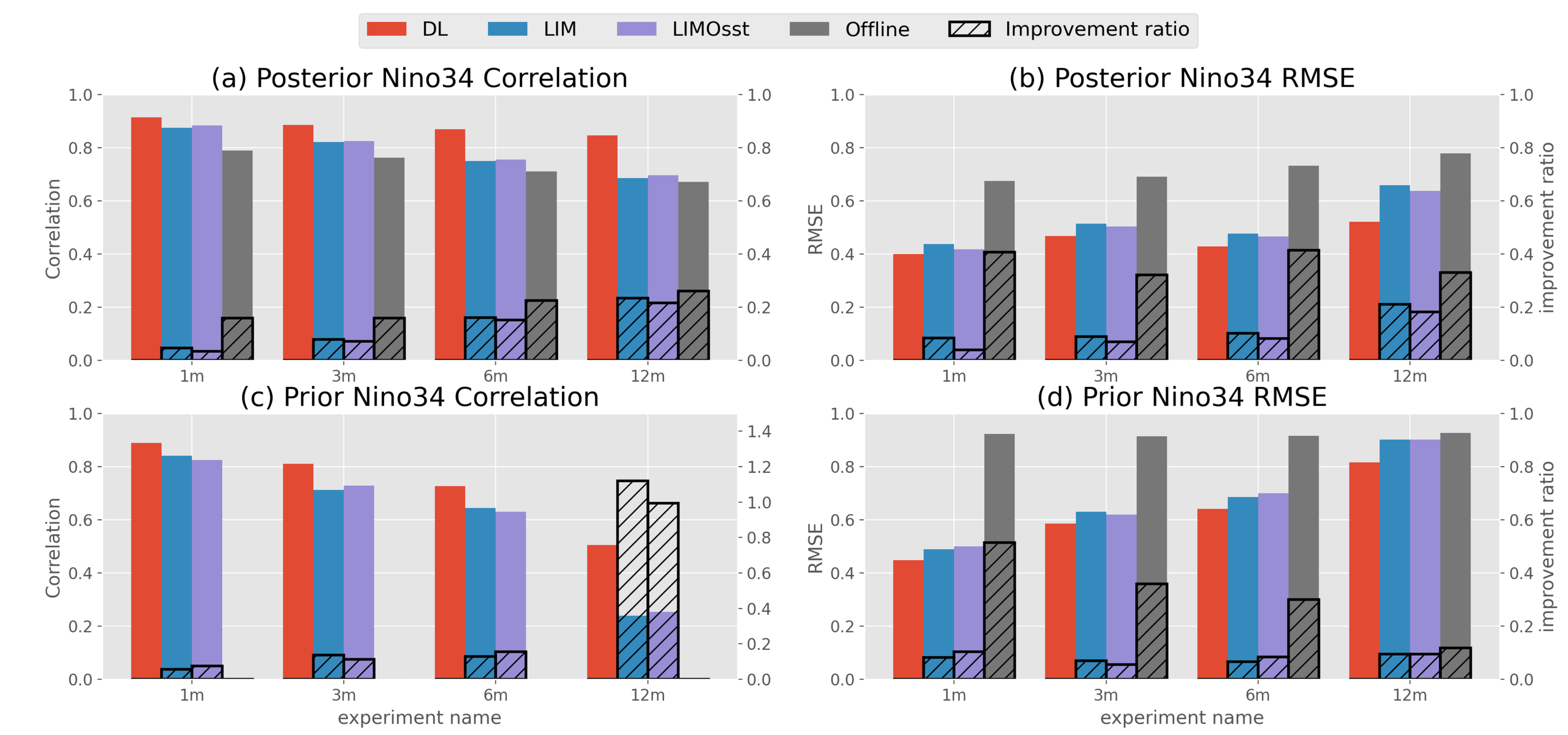


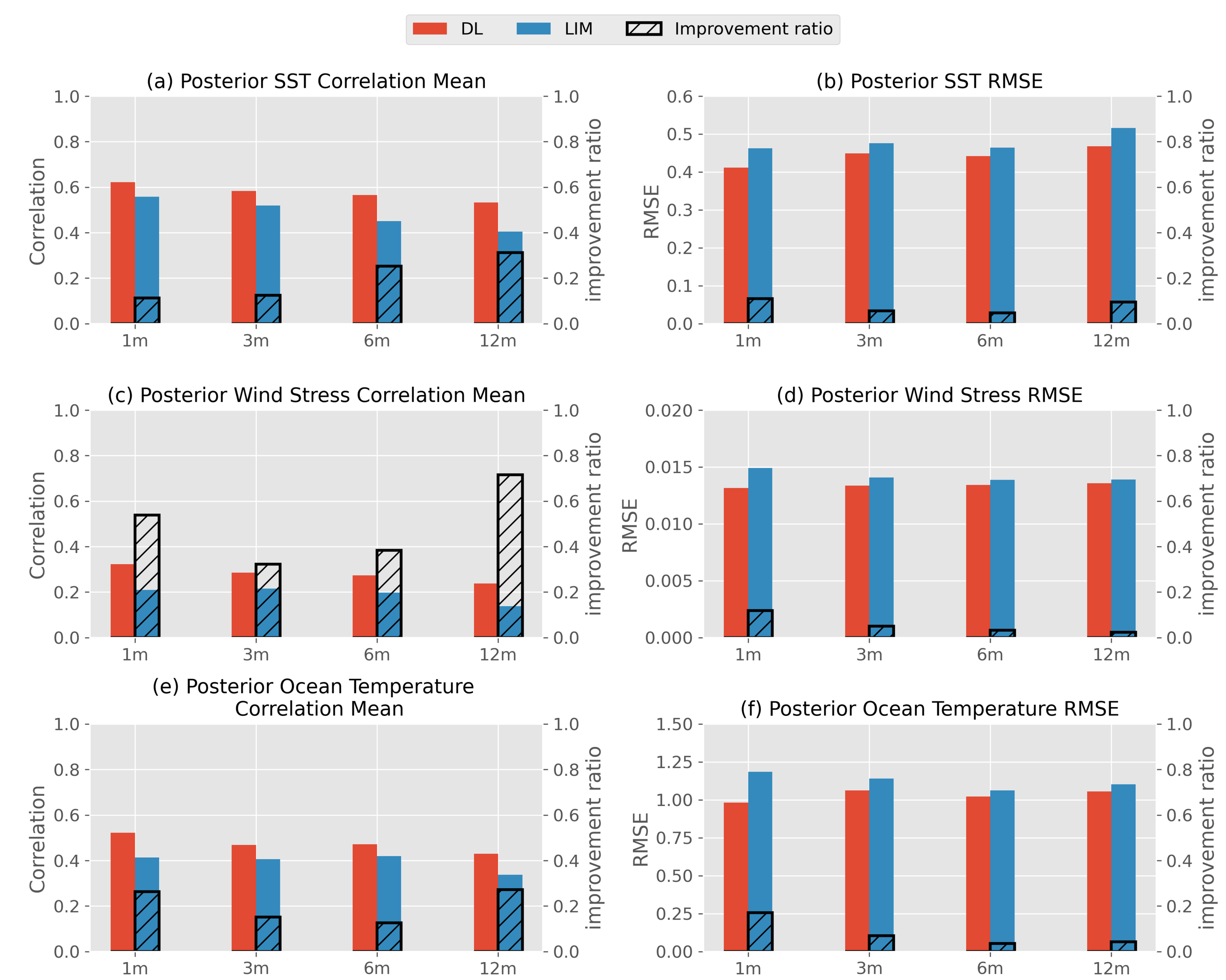
$$y_{\text{avg},N} = \frac{1}{N}\sum_{i=k+1}^{k+N} y_i$$

$$y'_{\text{avg},N} = y_{\text{avg},N} + \zeta$$

Average SST at coral locations over periods of 1, 3, 6 and 12 months and add white noise to mimic coral situation.

Data Assimilation Result





- DL model outperforms the others by around from 10 % to 30 % in reconstructing the Nino3.4 index.
- Improvement using the DL model increases with observation averaging time.
- In terms of skill across the entire domain, the DL model outperforms the LIMs for all observation averaging times and variables, most notably for correlation, and less so for RMSE.
- These improvements reflect a combination of the forecast-skill improvements, which better retain the memory of past observations, and improved spatial covariance, which spread information from the sparse network of observations.