Increasing the sample size for testing decadal forecasts

Gregory J. Hakim

University of Washington

A. Perkins; L. Taylor; J. Emile-Geay; E. Steig; D. Noone; R. Tardif; D. Anderson; M. Newman; C. Snyder; S. Penny

1. **Forecasts need initialization with <u>observations</u>: coupled data assimilation**
2. **Forecasts need verification over a <u>large sample</u> of variability: pre-instrumental**

Sponsors: NOAA, NSF,
Heising-Simons Foundation

Main Challenges

- **Computational**
  - large ensembles of Earth System models
- **Observational**
  - Argo period is too short
  - satellite period is too short
  - instrumental era is too short
  - ∴ paleoclimate proxies essential to extending the sample

sample strongly
forced climate

- Capacity for **strongly coupled DA** (SCDA)
  - e.g. atmospheric observations to analyze the ocean
  - computational burden, again
  - weakly coupled approximations are common

Recent Progress

- **Computational**: emulators (linear inverse models)
  - e.g., Hawkins & Sutton (2009); Newman (2013); Perkins & Hakim (2019)
- **Observational**: instrumental & proxy assimilation
  - e.g., 20CR, ORA20C; Goose et al. (2010), Hakim et al. (2016), Franke et al. (2017)
- **SCDA**: ocean analysis from mainly terrestrial observations
  - Perkins & Hakim (2021): “LMR online”

I will show two examples:

1. Last Millennium Reanalysis (LMR)
2. S2S using ideas & results from LMR.

Last Millennium Reanalysis (LMR)

Hakim et al. 2016; Tardif et al. 2019; Perkins & Hakim 2021

- **Gridded, global, multivariate "reanalysis"**
  - combines proxies with physical constraints from climate models
  - aims to be like modern climate reanalyses
  - annual resolution, 2x2 degree (20CR grid)
  - 2m air T; pressure; 500 hPa height; wind; precipitation
  - SST; 0-700m ocean heat content; AMOC

- **Open data and code**
  - **https://www.atmos.uw.edu/~hakim/lmr/**
  - **https://www.ncei.noaa.gov/access/paleo-search/study/27850**

Coupled Atmos-Ocean Emulator

- LIM trained on CMIP5 last millennium simulations

Andre Perkins

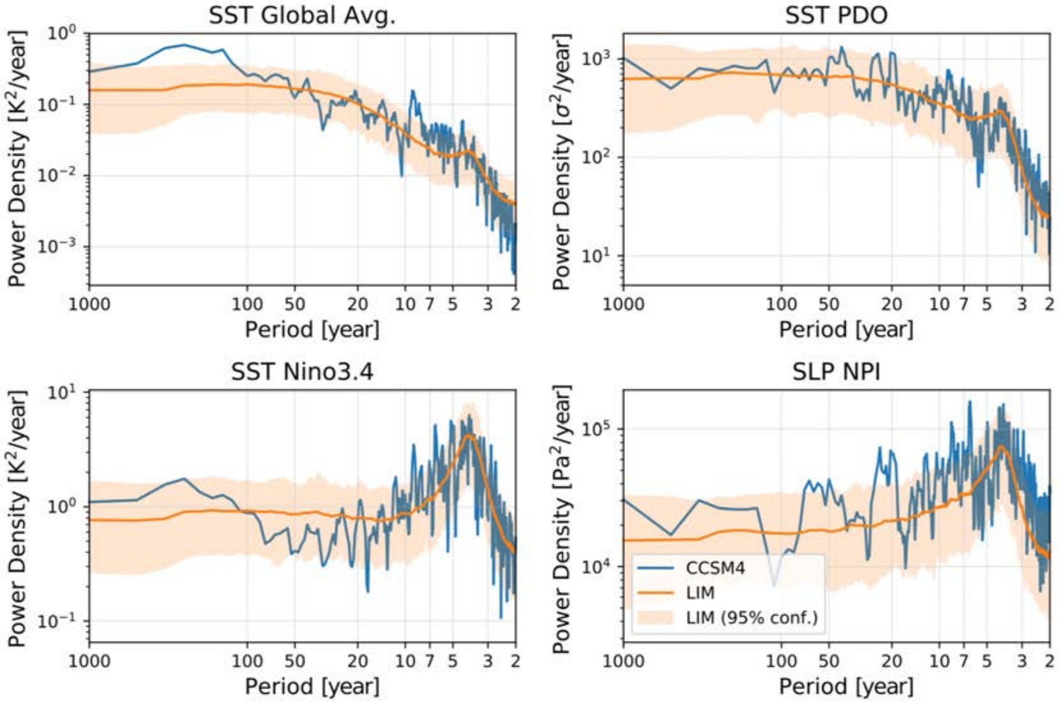


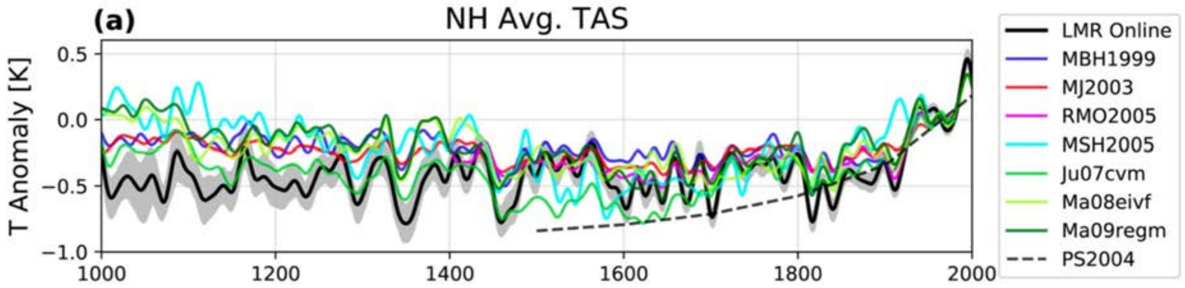
Figure 4. Power spectra for climate indices calculated from an ensemble of 1,000-year LIM integrations (orange) and from the CCSM4 past1000 data (blue). A 250-member ensemble of free-running integrations is used to determine the average LIM power spectra (solid line) and 95% confidence intervals (shaded region).

Perkins & Hakim (2019)

LMR Online: Coupled Atmos-Ocean DA

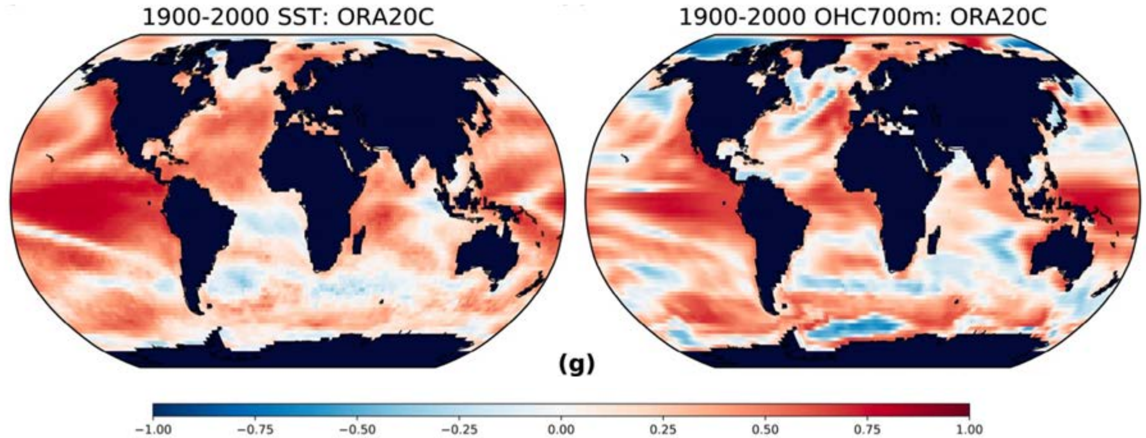
proxies only: trees, corals, & ice cores

2m air Temperature:



colder Medieval period, but decadal warm periods around Europe

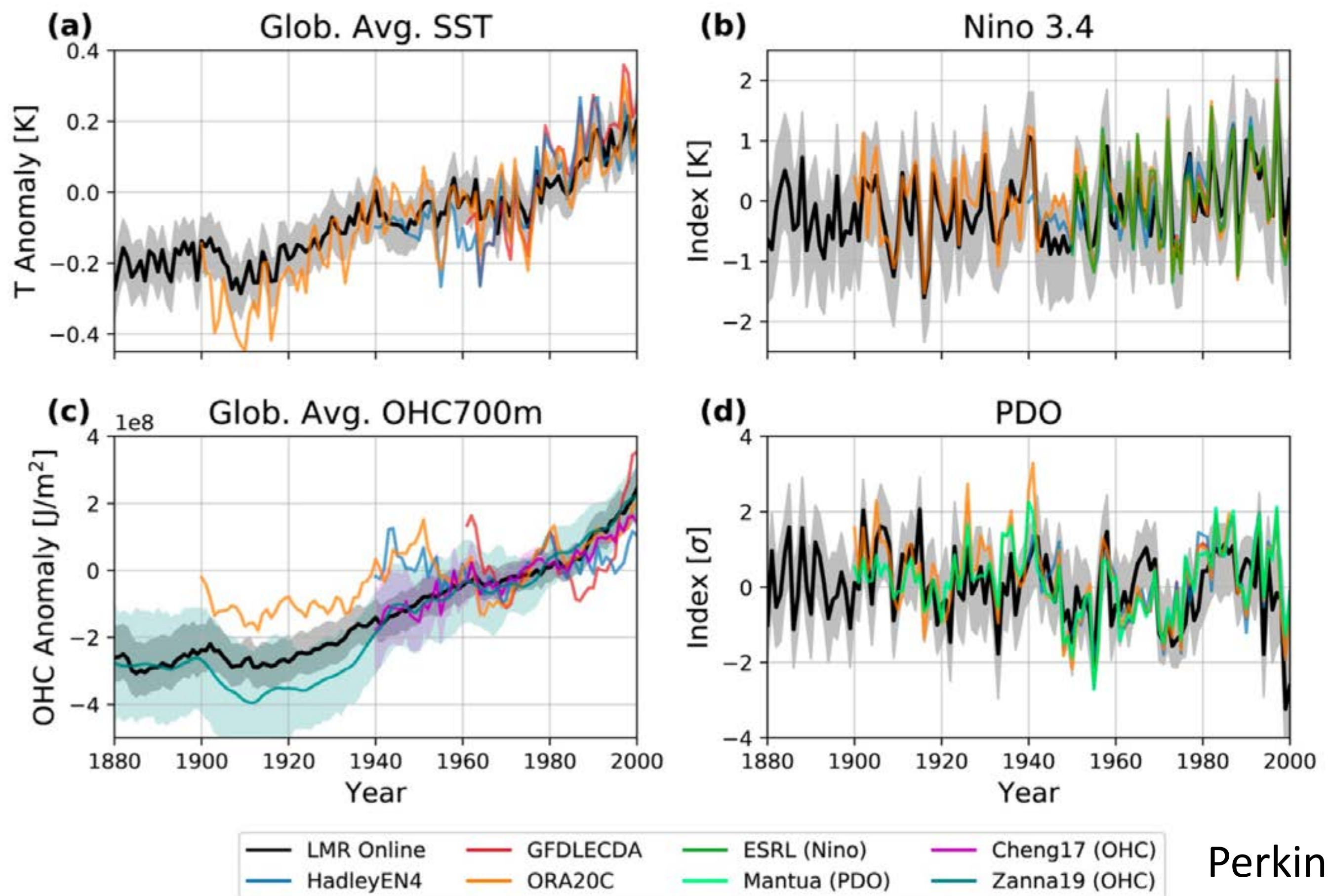
SST & ocean heat content validation:



good agreement with instrumental products

Perkins & Hakim (2021)

LMR Online: Instrumental Validation

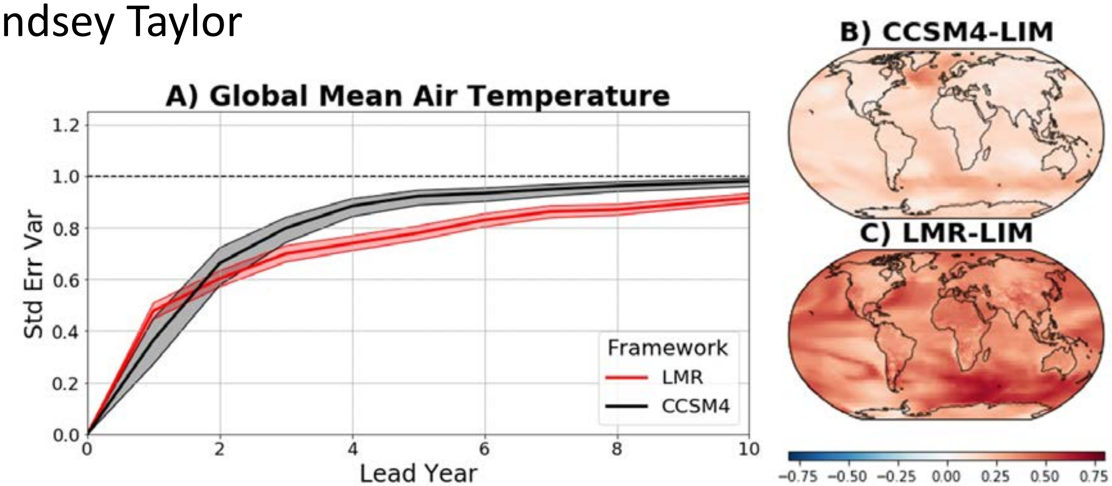


Perkins & Hakim (2021)

Figure 5. Scalar index comparison between the Last Millennium Reanalysis (LMR) Online (CCSM4-LIM) reconstruction (black with 95% confidence bounds in gray shading) and instrumental products for (a) SST, (b) Niño 3.4, (c) OHC700m, and (d) PDO. The HadleyEN4, GFDLECDA, and ORA-20C products are compared in all cases. Additionally, ESRL Niño 3.4 data, the Mantua et al. (1997) PDO index, and Cheng et al. (2017) and Zanna et al. (2019) OHC data are compared. Error bounds ($\pm 2\sigma$) are shown for the Cheng17 and Zanna19 OHC700m data.

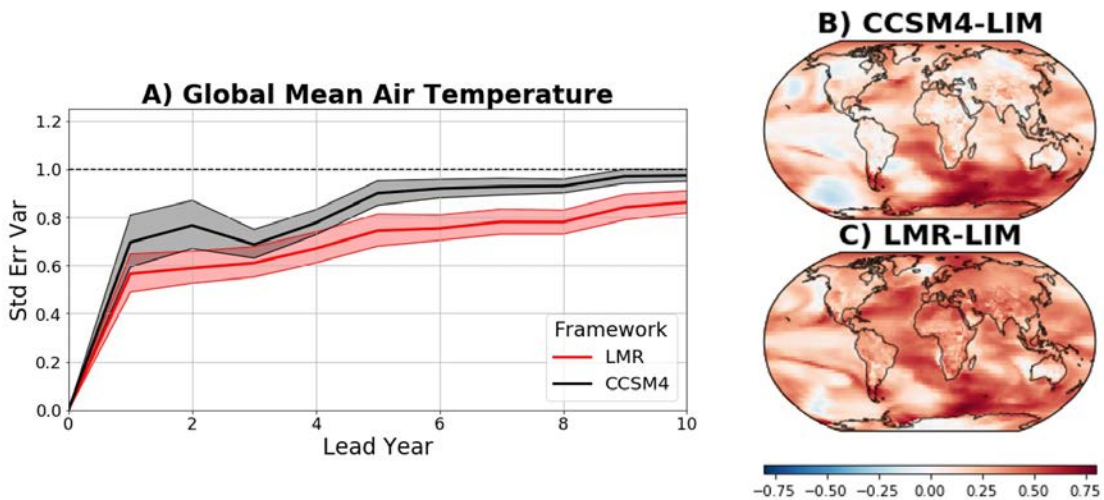
Lindsey Taylor

LMR-LIM Decadal Forecast Skill



In sample verification

- CCSM4-LIM: 850-1850 CE
- LMR-LIM: 1000-1850 CE



Out of sample verification

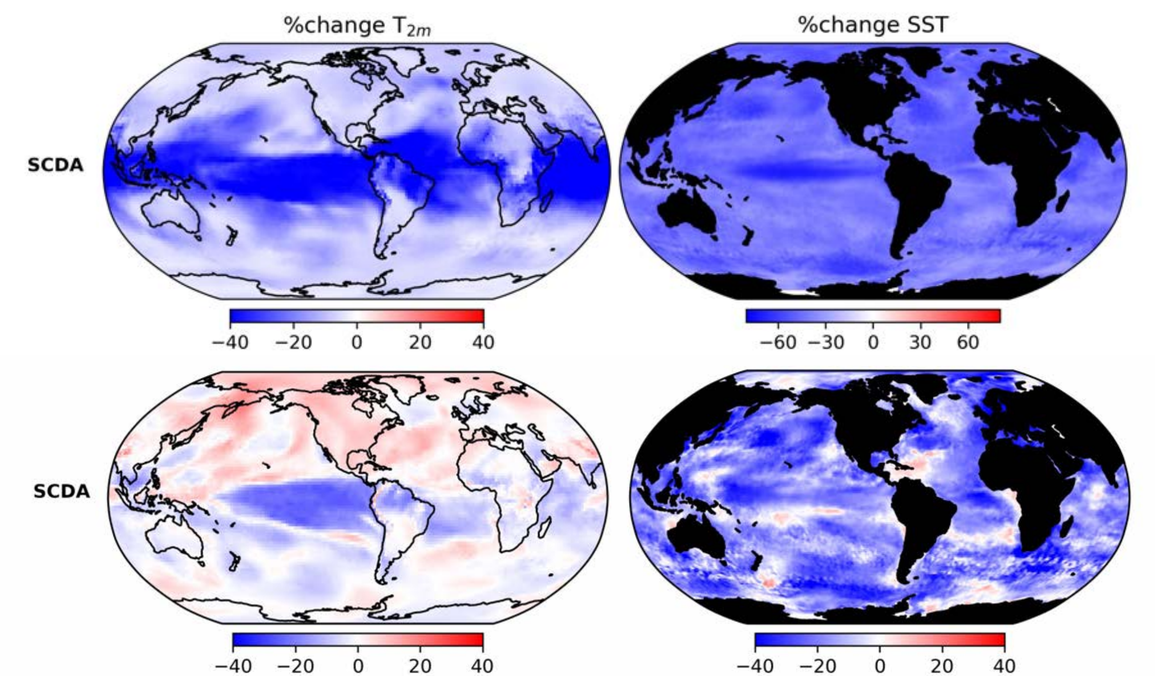
- 1850-2000 CE
- LMR initialization & verification

Taylor & Hakim (2022)

SCDA for S2S Analysis & Prediction

LIM trained on 5-day-mean atmosphere—ocean reanalysis (CFSR)
SCDA on reanalysis gridded observations using full Kalman filter

10-day forecast skill relative to controls (atmosphere-only, and ocean-only):



in-sample (1979-2003)

validation (2004-2010)

Hakim et al. (2022)

Summary

- Decadal forecasts need initialization with observations
  - strongly coupled DA
- Decadal forecasts need large-sample verification
  - utilize proxies & instrumental data
- LIMs provide a flexible tool for S2S--decadal DA & forecasts
  - skillful decadal forecasts
  - every Earth System model can have LIMs; multi-model ensembles
  - basis for evaluating nonlinear emulators ("machine learning")

Notes:

- Funding agencies (esp. NOAA) need input from modelers on paleo-DA!
- WCRP meeting 19-20 May 2022, Boulder: "DA needs for Climate Prediction"

Thank You!

Linear Inverse Models (LIMs)

e.g. Penland (1989); Newman et al. (2003)

$$\frac{d\mathbf{x}}{dt} = \mathcal{N}(\mathbf{x}) \approx \mathbf{L}\mathbf{x} + \xi \quad \rightarrow \quad \mathbf{x}(\tau) = \mathbf{G}_\tau \mathbf{x}_0 + \epsilon \qquad \mathbf{G}_\tau = e^{\mathbf{L}\tau}$$

Solve for G empirically from sample data:

$$\mathbf{G}_\tau = cov(\mathbf{x}_\tau, \mathbf{x}_0)(cov(\mathbf{x}_0, \mathbf{x}_0))^{-1}$$

and the noise error-covariance matrix

$$cov(\epsilon, \epsilon) = \mathbf{N}_\tau = \mathbf{C} - \mathbf{G}_\tau \mathbf{C} \mathbf{G}_\tau^{\mathrm{T}} \qquad \mathbf{C} = cov(\mathbf{x_0}, \mathbf{x}_0)$$

LIM Training

- Climate Forecast System Reanalysis (CFSR) (Saha et al. 2010)
- Global gridded fields of **2m air temperature** (T_{2m}), **SST**, 850 hPa **u** & **v**, and **OLR**
    - diurnal & 5-day running mean; seasonal cycle removed
- Truncate to leading 30 EOFs for each variable (150 degrees of freedom in total)
- LIM **training** period: 1979-2003 (9130 days)
- LIM **validation** period: 2004-2010 (2556 days)

Kalman Filter using the LIM

analysis

$$\mathbf{x}_a = \mathbf{x}_f + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_f)$$

$$\mathbf{P}_a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_f$$

$$\mathbf{K} = \mathbf{P}_f\mathbf{H}^{\mathrm{T}}\left[\mathbf{H}\mathbf{P}_f\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right]^{-1}$$

forecast

full matrices!

$$\mathbf{x}_f = \mathbf{G}_t\mathbf{x}_a$$

$$\mathbf{P}_f = \mathbf{G}_t\mathbf{P}_a\mathbf{G}_t^{\mathrm{T}} + \mathbf{N}_t$$

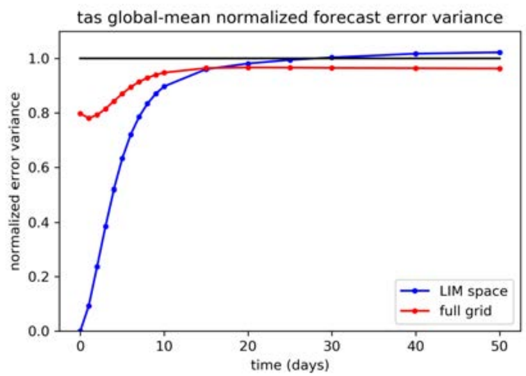
Cycling time *t* = 1 day

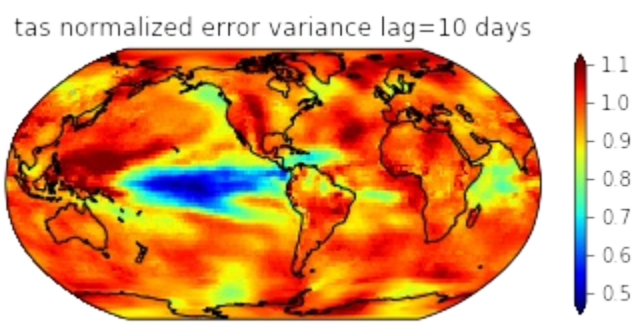
Strongly Coupled Data Assimilation (SCDA)

- Observation assimilation updates all modeled components of Earth system
- **SCDA**: Extremely computationally demanding (e.g. Liu et al. 2013; Penny & Hamill 2017)
  – Coupled forecast and "strongly coupled" DA: cross-medium updates
- "Weakly coupled" approximations (**WCDA**) (e.g. Saha et al. 2006; Zhang et al. 2007; Penny et al. 2019)
  – Separate DA in atmosphere & ocean with a coupled forecast step
  – No "cross covariance" influence from observations
  – Potentially "incompatible" states
  – Still very computationally demanding
- Here: break the computation bottleneck with a **linear emulator**
  – Low dimensionality promotes experimentation/prototyping
  – Unlike "toy" models, the emulator skillfully forecasts observed fields
  – Allows unapproximated Kalman filter, and WCDA-SCDA evaluation

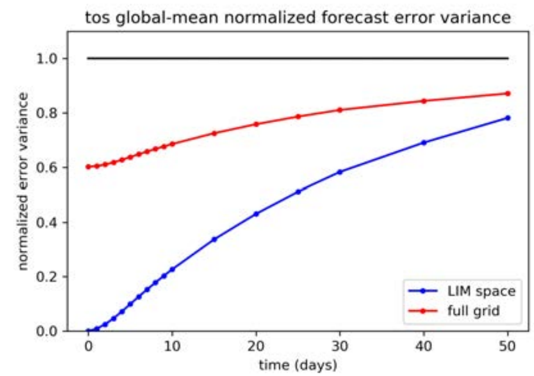
LIM forecast verification 2004-2010

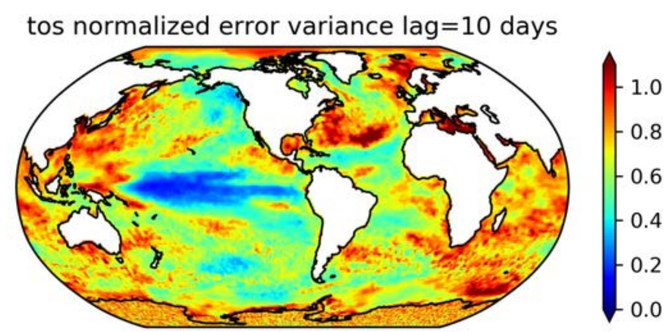
T_{2m}





SST





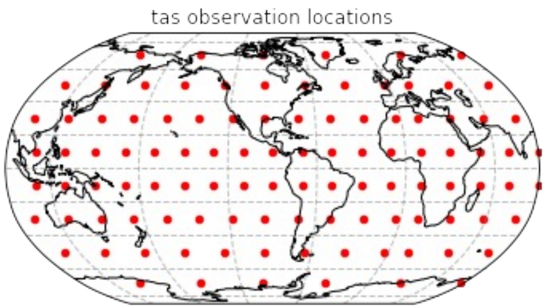
Pointwise observations from CFSR

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{e} = \underbrace{\mathbf{H}^*\mathbf{U}\mathbf{x}}_{\text{obs from LIM basis}} + \underbrace{\mathbf{H}^*\eta}_{\text{truncation error}}$$

$\mathbf{H}^*$: obs operator on CFSR grid

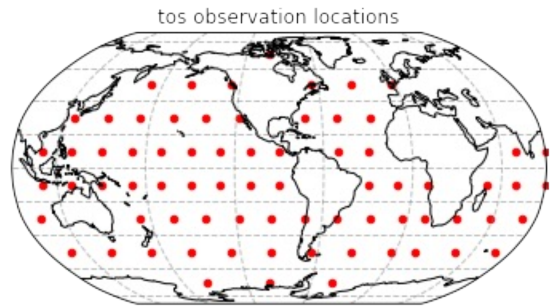
$\mathbf{R}$ derived from truncation error

T_{2m}



104 locations

SST

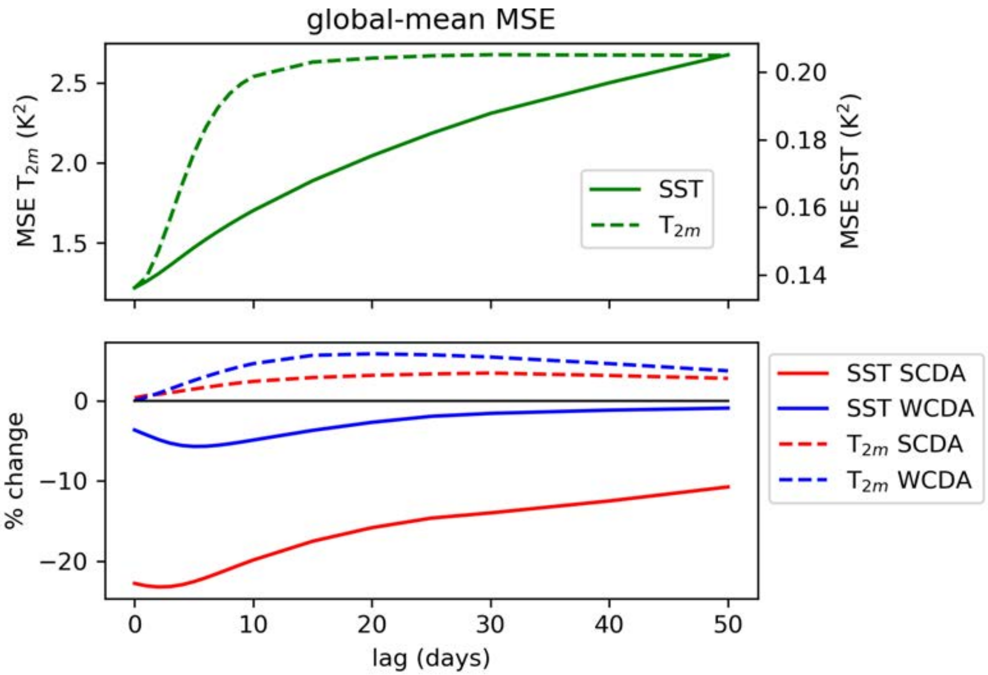


73 locations

DA Experiments

1. Uncoupled: **atmosphere**
   - Atmosphere components of the LIM & atmosphere observations
2. Uncoupled: **ocean**
   - Ocean components of the LIM & ocean observations
3. Weakly coupled (**WCDA**)
   - Full LIM; separate DA in atmosphere & ocean; no cross covariances
4. Strongly coupled (**SCDA**)
   - Full LIM; fully coupled DA

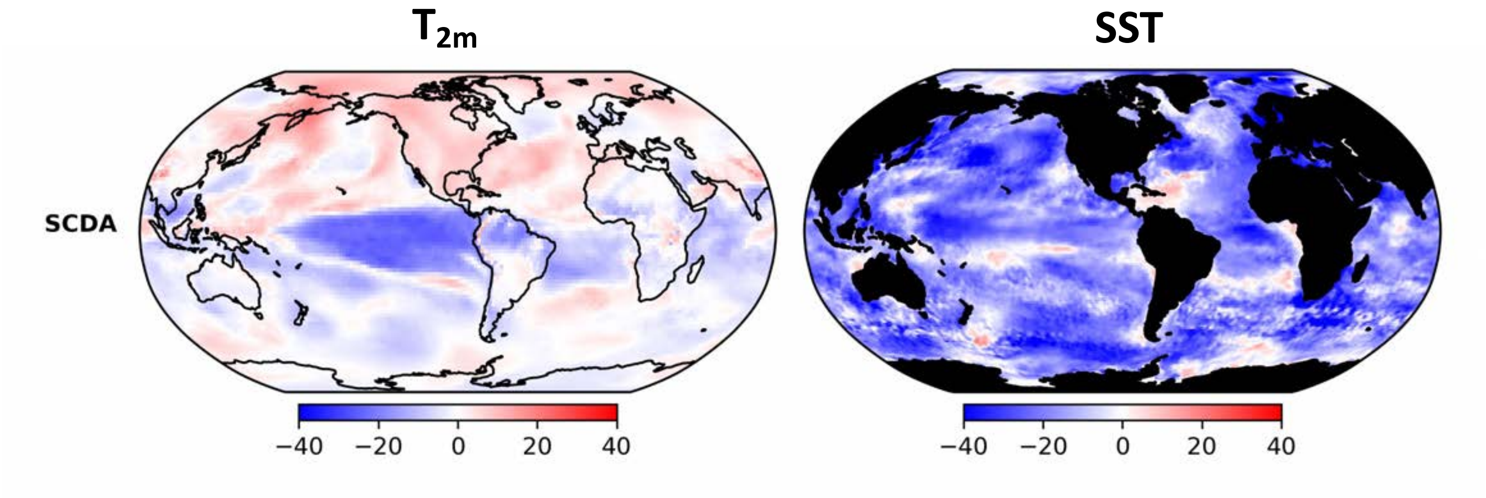
Global-mean Mean Squared Error



uncoupled (**control**) experiments

WCDA **&** **SCDA**
change from control experiments

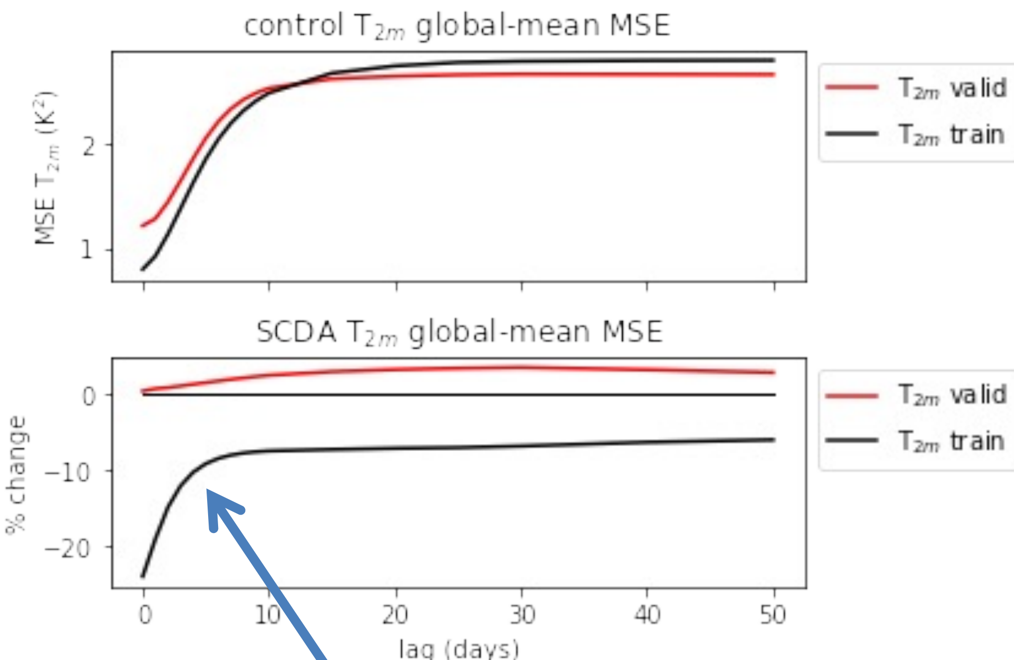
SCDA Change from Uncoupled 10-day Forecasts



- improved tropics
- degraded NH extratropics

- global improvement

T_{2m}: **Training** vs. **Validation** Periods



uncoupled (**control**) experiments

SCDA
change from control experiments

Large SCDA improvement
doesn't hold up in validation

Ensemble Kalman Filter (EnKF)

Kalman update equation:

Kalman gain

Observations

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}[\mathbf{y} - \mathcal{H}(\mathbf{x_b})]$$

Analysis Ensemble

Prior ensemble

Estimated Observations ($\mathbf{y}_e$)

Kalman gain:

$$\mathbf{K} = cov(\mathbf{x}_b, \mathbf{y}_e)[cov(\mathbf{y}_e, \mathbf{y}_e) + \mathbf{R}]^{-1}$$