Coupled reanalyses methodologies

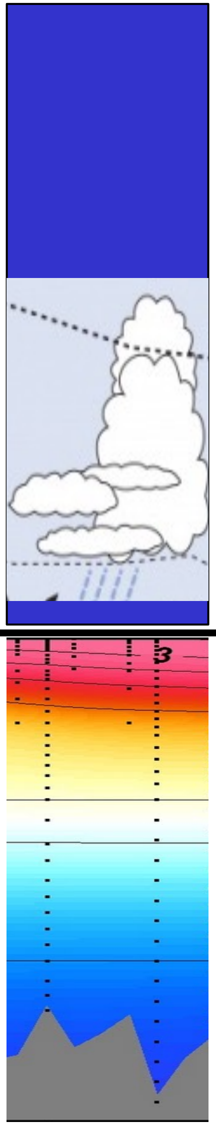
Sergey Frolov

Presented at:
CLIVAR Reanalysis Workshop
May 2022

The coupled reanalysis opportunity: effective use of observations

Atmosphere

Ocean



Well-established DA system for each fluid

Cross-fluid correlations that may benefit DA

Highly valuable satellite observations that are sensitive to both fluids

How do we exploit these opportunities to best use sparse historic observations?

Soup of acronyms

Quasi coupled

Weakly coupled

Intermediate coupling

Strongly coupled

Interface solver

Outerloop coupling

Inter component coupling

Multicomponent coupling

Outline

Uncoupled reanalysis

Weakly coupled replay

Weakly coupled reanalysis

…

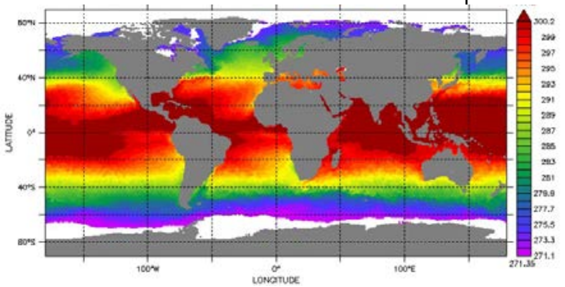
Stronger coupled reanalysis

…

Strongly coupled reanalysis

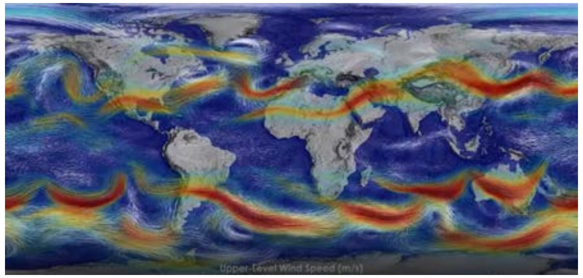
Uncoupled reanalysis

SST analysis



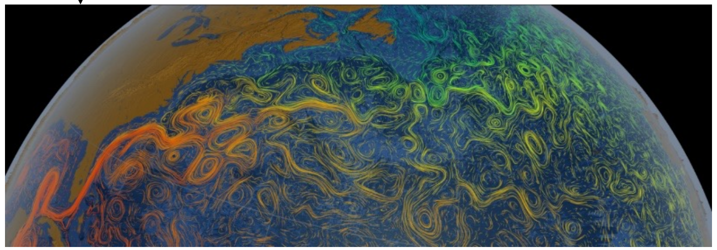
SST observations

Atmospheric analysis



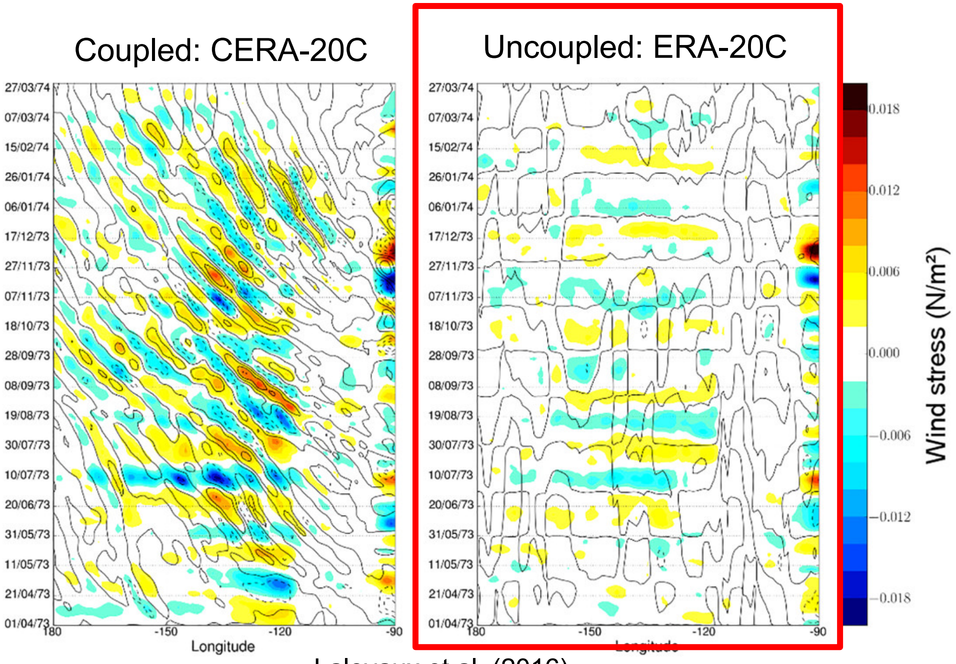
Atmospheric observations

Ocean analysis



Ocean observations

- Most of the existing reanalysis are produced sequentially.
- These reanalysis do not capture the full range of ocean-atmosphere interactions.



Laloyaux et al. (2016)

Examples: ERA5, MERA2, ORAS5, GLORYS

Weakly coupled replay

Replay integration

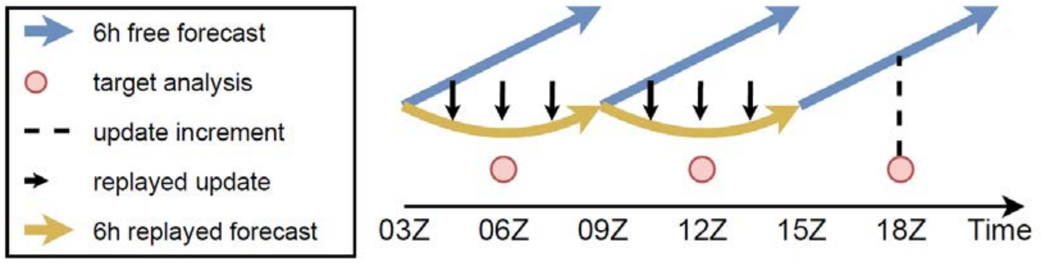
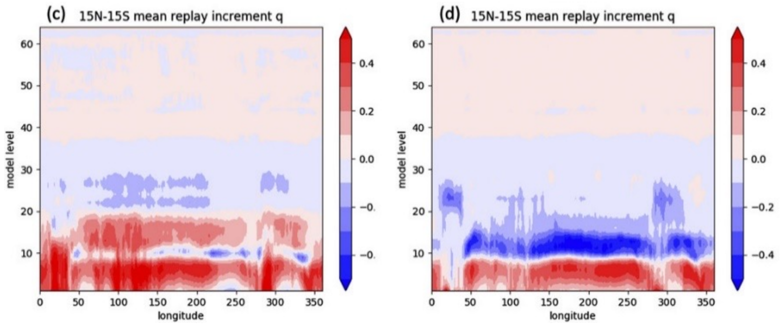


Figure from Chen et.al in prep

Mean replay increments

GFS physics replayed to ERA5

IFS physics replayed to ERA5



Bengtsson et.al 2019

- (right) Replay nudges forecast integration to an external analysis using additional tendency term.
- (left) Differences in the choice of moist physics are apparent in replay increments even if replayed to the same reference analysis.
- We plan to replay coupled UFS to ERA5 and ORAS5 to initialize coupled re-forecast for the next GFS and GEFS system.

Examples: UFS replay/reforecast for GFSv17/GEFSv13

Outline

Uncoupled reanalysis

Uncoupled replay

Weakly coupled reanalysis

…

Stronger coupled reanalysis

…

Strongly coupled reanalysis

Framework for formal notation

Kalman gain:
maps observation innovations to model space

$$\underbrace{x_k^a}_{\text{analysis}} = \underbrace{\mathscr{M}(x_{k-1}^a)}_{\text{forecast}} + \mathbf{K}\underbrace{\left[y - \mathscr{H}(\mathscr{M}(x_{k-1}^a))\right]}_{\text{innovation}}$$

innovation:
difference between observations and forecasts

For didactic purposes, lets start with something simple:
Observational space estimator with one outerloop

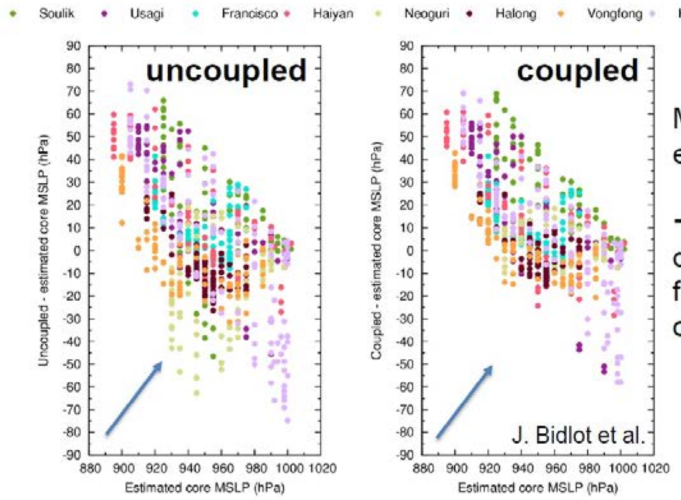
$$x_k^a = \mathscr{M}(x_{k-1}^a) + \mathbf{P}_0\mathbf{M}^T\mathbf{H}^T\left(\mathbf{H}\mathbf{M}\mathbf{P}_0\mathbf{M}^T\mathbf{H}^T + \mathbf{R}\right)^{-1}\left[y - \mathscr{H}(\mathscr{M}(x_{k-1}^a))\right]$$

Definitions: weakly coupled DA

Weakly coupled data assimilation

$$x_k^a = \mathcal{M}(x_{k-1}^a) + \mathbf{P}_0\mathbf{M}^T\mathbf{H}^T\left(\mathbf{H}\mathbf{M}\mathbf{P}_0\mathbf{M}^T\mathbf{H}^T + \mathbf{R}\right)^{-1}\left[y - \mathcal{H}\left(\mathcal{M}(x_{k-1}^a)\right)\right]$$

Coupled forecast model: $x_{k+1}^{coupled} = \begin{bmatrix} x_{k+1}^{atm} \\ x_{k+1}^{oce} \end{bmatrix} = \mathcal{M}^{coupled}\left(\begin{bmatrix} x_k^{atm} \\ x_k^{oce} \end{bmatrix}\right)$



MSLP forecast error versus the estimated core MSLP

→ In the uncoupled case, too deep TCs (no cold wake feedback effect, infinite source of energy)

- Impact of coupled forecast models have been widely documented:
  - TC strength (ECMWF left)
  - Tropical wind-SST coupling (ECMWF, NRL)
  - Ice extent prediction (NRL)

Examples: CFSR

Definitions: coupling through an outerloop

Data assimilation coupled through 4DVAR outerloop

$$x_k^{a[i]} = \mathcal{M}\left(x_{k-1}^a + \sum_i \delta x_{k-1}^{[i]}\right) + \mathbf{P}_0\mathbf{M}^T\mathbf{H}^T\left(\mathbf{HMP}_0\mathbf{M}^T\mathbf{H}^T + \mathbf{R}\right)^{-1}\left[y - \mathcal{H}\left(\mathcal{M}\left(x_{k-1}^a + \sum_i \delta x_{k-1}^{[i]}\right)\right) - \mathbf{H}\sum_i \delta x_{k-1}^{[i]}\right]$$

Coupled forecast model:

$$x_{k+1}^{coupled} = \begin{bmatrix} x_{k+1}^{atm} \\ x_{k+1}^{oce} \end{bmatrix} = \mathcal{M}^{coupled}\left(\begin{bmatrix} x_k^{atm} \\ x_k^{oce} \end{bmatrix}\right)$$

TLM/ADJ of the forecast model:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}^{AA} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

TLM/ADJ of the observation operator:

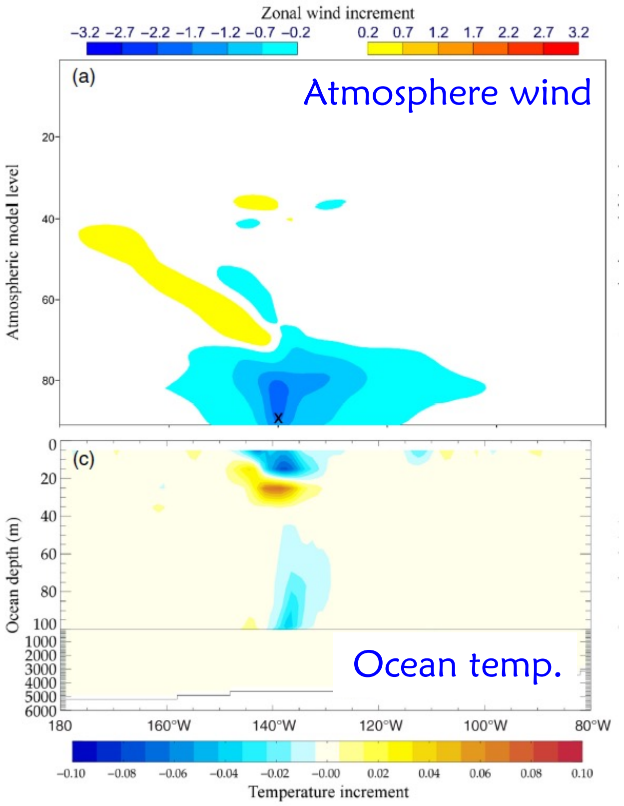
$$y^{radiance} = \mathbf{H}^{rtm} x^{atm} = \begin{bmatrix} \mathbf{J}^{atm} \\ 0 \end{bmatrix}\begin{bmatrix} x^{atm} \\ 0 \end{bmatrix}$$

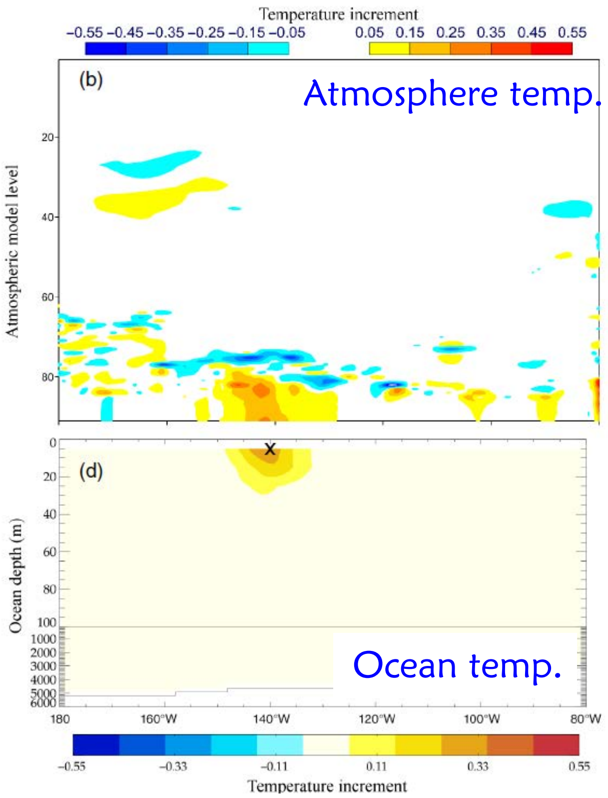
Initial-time covariance:

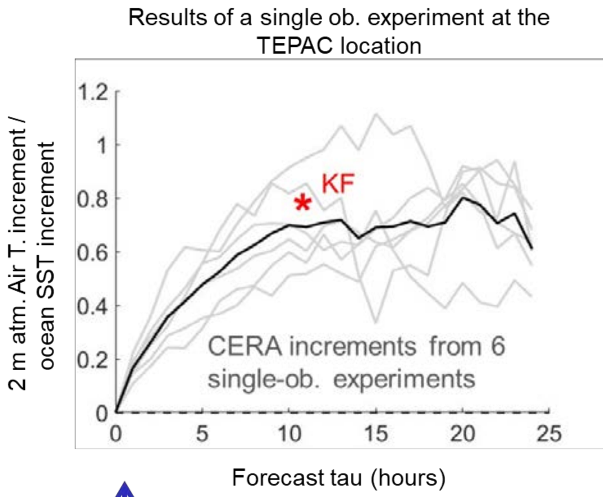
$$\mathbf{P}_0 = \begin{bmatrix} \mathbf{P}^{AA} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{OO} \end{bmatrix}$$

Examples: CERA20C

An example of DA coupled through an outerloop







Outerloop approximates strong coupling poorly at the beginning of the window

Outerloop approximates strong coupling well after 12 hours

- Laloyaux et al. (2016) showed that outerloop coupling is effective at propagating information between assimilated fluids: E.g.
  - (left) Impact of wind observation on the mixed layer depth
  - (center) Impact of SST assimilation on the boundary layer depth
- Laloyaux et al. (2018) also showed (right) that outerloop is most effective with long assimilation windows (>12 hours)

Definitions: strongly coupled DA

Strongly coupled data assimilation

$$x_k^a = \mathcal{M}(x_{k-1}^a) + \mathbf{P}_0\mathbf{M}^T\mathbf{H}^T\left(\mathbf{H}\mathbf{M}\mathbf{P}_0\mathbf{M}^T\mathbf{H}^T + \mathbf{R}\right)^{-1}\left[y - \mathcal{H}(\mathcal{M}(x_{k-1}^a))\right]$$

Coupled forecast model:

$$x_{k+1}^{coupled} = \begin{bmatrix} x_{k+1}^{atm} \\ x_{k+1}^{oce} \end{bmatrix} = \mathcal{M}^{coupled}\left(\begin{bmatrix} x_k^{atm} \\ x_k^{oce} \end{bmatrix}\right)$$

Coupled TLM/ADJ of the forecast model (if used):

$$\mathbf{M}^{coupled} = \begin{bmatrix} \mathbf{M}^{AA} & \mathbf{M}^{AO} \\ \mathbf{M}^{OA} & \mathbf{M}^{OO} \end{bmatrix}$$

Coupled TLM/ADJ of the observation operator:

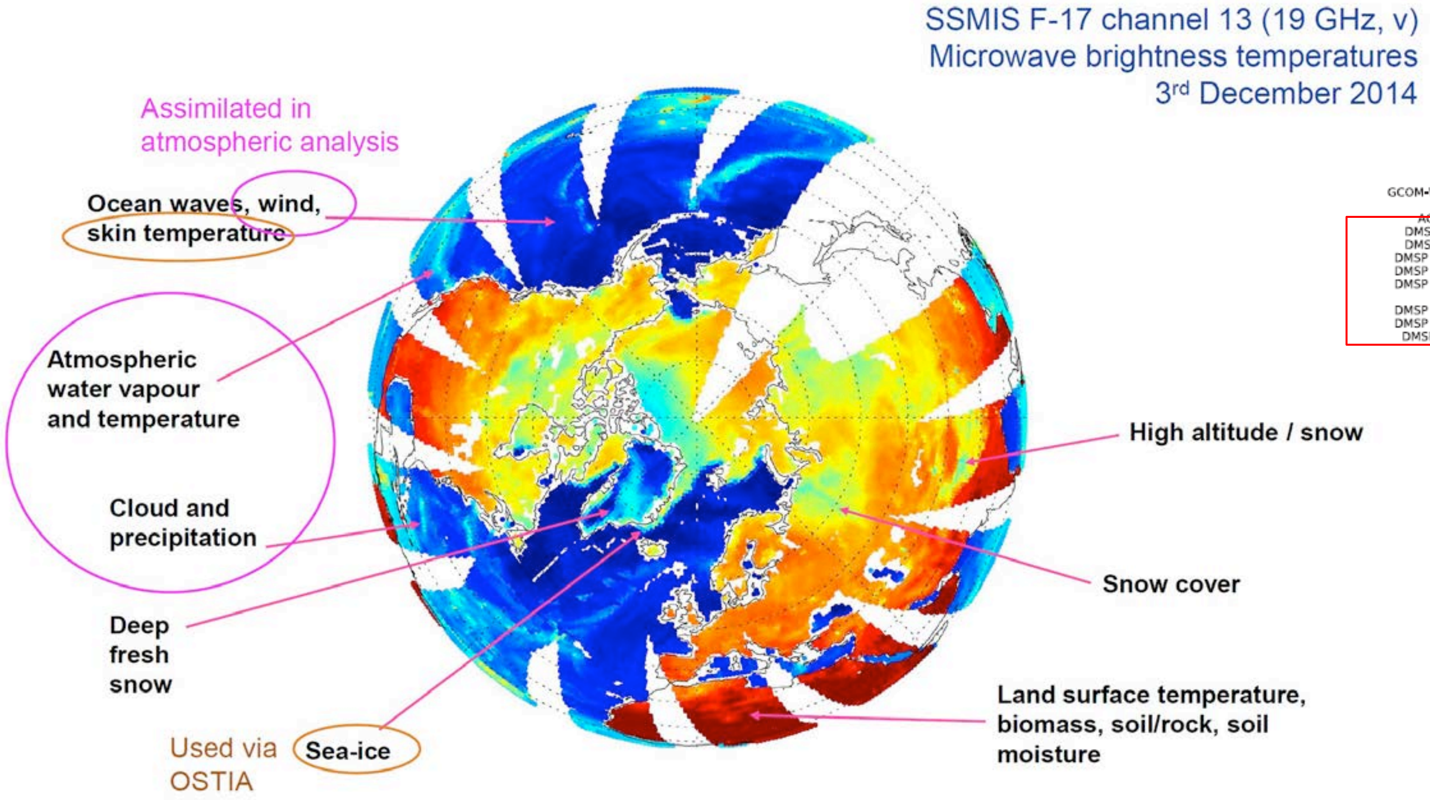
$$y^{radiance} = \mathbf{H}^{rtm-coupled} x^{coupled} = \begin{bmatrix} \mathbf{J}^{atm} \\ \mathbf{J}^{ocean} \end{bmatrix}\begin{bmatrix} x^{atm} \\ x^{ocean} \end{bmatrix}$$

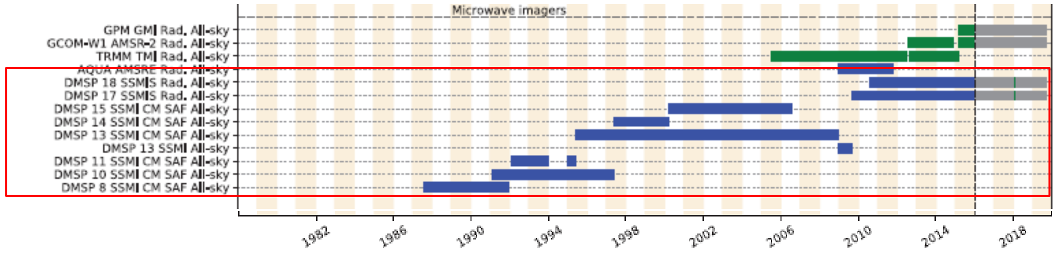
Coupled initial-time covariance:

$$\mathbf{P}_0^{coupled} = \begin{bmatrix} \mathbf{P}^{AA} & \mathbf{P}^{AO} \\ \mathbf{P}^{OA} & \mathbf{P}^{OO} \end{bmatrix}$$

No implementations so far in models of operational complexity

Better use of all-sky / all-surface information from historic radiance observations



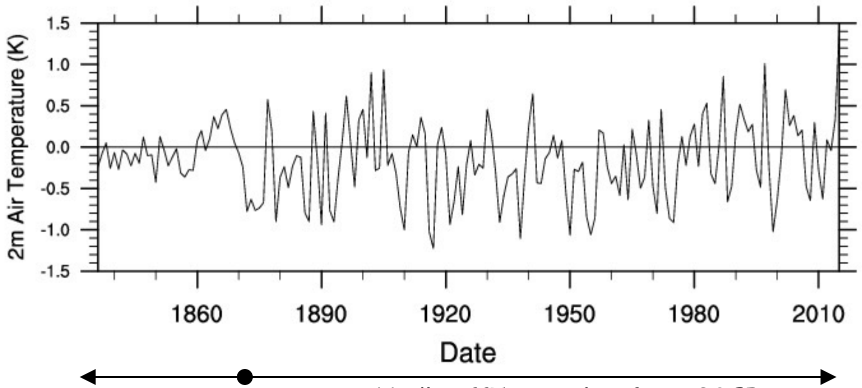


- Many of the historic observations are underexploited (e.g. no assimilated over all surfaces and in all conditions).

Figure courtesy: Allan Geer and Hersbach et.al 2020.

Coupled DA for sparse input

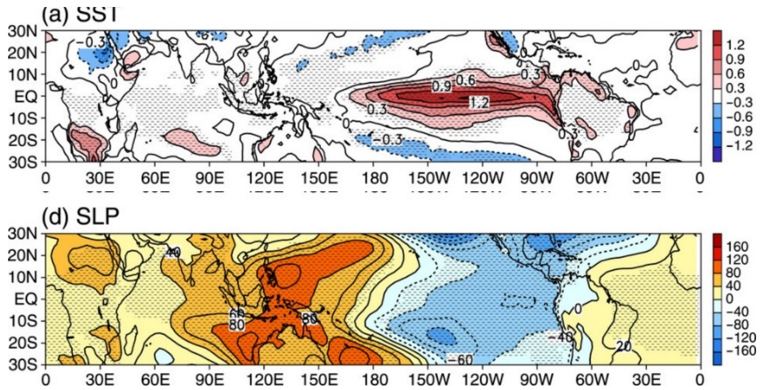
ENSO-like signal in 20CR



Strong relaxation to climatology

Hadley SST is used to force 20CR

ENSO anomaly in SST and SLP



- (right) 20CR has poorly constrained before 1880.
- Direct SST observations in East Tropical Pacific a required to directly detect ENSO (possibly one ship crossing in a month in 1860)
- Sea level pressure anomaly for ENSO is much large-scale. Reliable SLP records in Australia and US West coast date back to mid-19th centaury.
- **Hypothesis**: coupled data assimilation can invert atmospheric observations to constrain large-scale ocean signals.

Figure courtesy: Laura Slivinski and 10.1175/JCLI-D-14-00716.1

Conclusions

- Next versions of major reanalyses will be weakly coupled for atmosphere, ocean, ice, and land:
  - ERA6, MERA3, NOAA’s replacement for CFSR.
- Unclear what is the best strategy for other components of the system:
  - Composition, biomass, biogeochemistry, .. Etc.
- Stronger coupling will require sustained effort:
  - Translating advances in all-sky / all surface assimilation in the NWP system to historic reanalysis.
  - Modeling and exploiting coupled covariances.
  - …

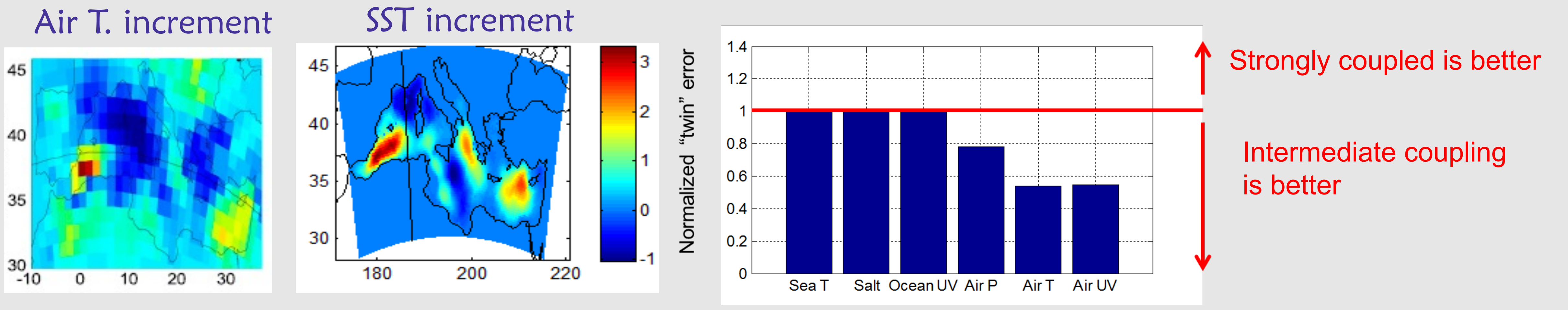
Questions

- What do you see are the most significant advances for the field of reanalysis in the next 5-10 years?
  - Advancement of the methodology that can extract more information from historic/sparse observations.
  - Production of fully coupled reanalysis.
- What do you see are the most significant barriers to progress in the field of reanalysis?
  - Computational cost.
  - Meaningful ways to share cost of development and production across centers.
- Which collaborations need to be fostered?
  - Private-public partnerships.
  - Collaboration between reanalysis centers.

End

Examples of strongly coupled DA

- None so far in the models of operational complexity
- Early indications of promise in simplified models
  - Lu et al. (2015), Sluka (2016), Smith et al. (2015, 2017)
- Early indications of caution against strong coupling
  - Lu et al. (2015), Frolov et.al. (2016)



Frolov et al. (2016) MWR

Definitions: coupling through initial time error covariance

Data assimilation coupled through initial time covariance

$$x_k^a = \mathscr{M}(x_{k-1}^a) + \mathbf{P}_0 \mathbf{M}^T \mathbf{H}^T \left(\mathbf{H}\mathbf{M}\mathbf{P}_0 \mathbf{M}^T \mathbf{H}^T + \mathbf{R} \right)^{-1} \left[y - \mathscr{H}(\mathscr{M}(x_{k-1}^a)) \right]$$

Coupled forecast model:

$$x_{k+1}^{coupled} = \begin{bmatrix} x_{k+1}^{atm} \\ x_{k+1}^{EST} \end{bmatrix} = \mathscr{M}^{coupled}\left(\begin{bmatrix} x_k^{atm} \\ x_k^{EST} \end{bmatrix} \right)$$

TLM/ADJ of the forecast model:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}^{AA} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

Coupled TLM/ADJ of the observation operator:

$$y^{radiance} = \mathbf{H}^{rtm-coupled} x^{coupled} = \begin{bmatrix} \mathbf{J}^{atm} \\ \mathbf{J}^{SST} \end{bmatrix} \begin{bmatrix} x^{atm} \\ x^{SST} \end{bmatrix}$$

Coupled initial-time covariance:

$$\mathbf{P}_0^{coupled} = \begin{bmatrix} \mathbf{P}^{AA} & \mathbf{P}^{A,EST} \\ \mathbf{P}^{EST,A} & \mathbf{P}^{EST,EST} \end{bmatrix}$$

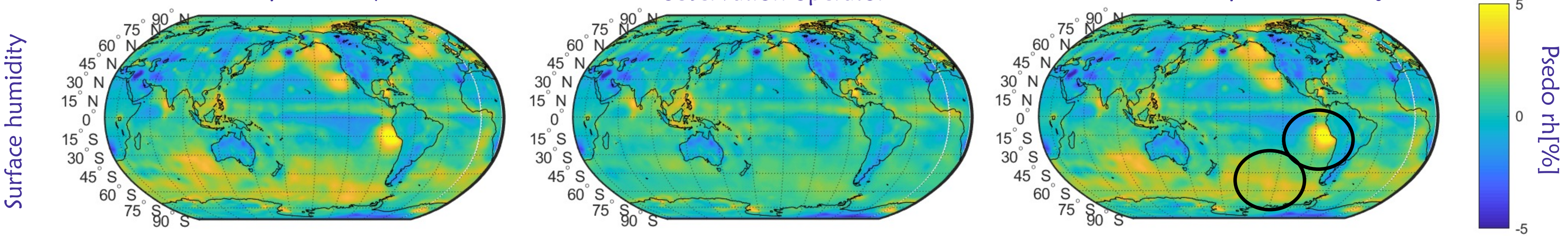
Example of coupling through observation operator

Average increment (July-August 2016)

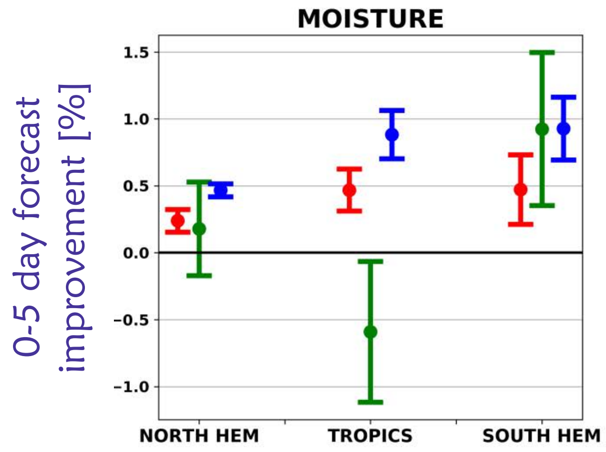
Atmosphere-only DA

Coupling through observation operator

Coupling through observation operator and P_0



- Coupling through observation operator alone might alias atmospheric signal into the ocean.



Atmosphere + diurnal SST

Atmosphere + diurnal SST + coupled H

Atmosphere + diurnal SST + coupled H + coupled P_0

Frolov et.al. (2020)

$$x_k^{replay} = \int_{k-1}^{k} \left[M x_{k-1}^{replay} + \frac{x_k^{era5} - \mathscr{M}_{atm}(x_{k-1}^{replay})}{(dk)/k-(k-1)} \right] dk$$