A Data Assimilation Perspective on the Initialization of Climate Prediction

Robin Weber
Climate Forecasting Unit (CFU)
Catalan Institute of Climate Sciences (IC3),
Barcelona, Spain

2013 Summer School: Data Assimilation in Geosciences,
CSCAMM

Backdrop: s2d prediction

s2d prediction:

- ✓ External forcing
- ✓ Coupled models
- ✓ Signature of initial condition

Climate projection

NWP

Initialization of s2d prediction:

- Full Field Initialization (FFI)

Drift!

- Anomaly Initialization (AI)

Research objectives

- Propose a unified formalism from which the different initialization approaches can be derived
- Compare FFI and AI for a range of different observational and model error scenarios using an idealized coupled dynamics
- Propose two advanced formulations of FFI/AI to improve the assimilation of observations, and reduce the bias
- Illustrate the impact of initialization using a state-of-the-art Earth System Model

Posing the problem

Climate model:
$$\frac{d\vec{x}}{dt} = F(\vec{x},\vec{\lambda})$$

Assumptions

Nature:
$$\frac{d\vec{x}^{nat}}{dt} = F(\vec{x}^{nat},\vec{\lambda}^{nat}) + G(\vec{x}^{nat},\vec{\lambda}^{nat})$$

Observations:
$$\vec{y}_i^{\,o} = \vec{y}^{\,o}(t_i)$$

$$t_i = i\tau, \quad i = 0,1\ldots;$$

$$\vec{y}^{\,o} = \Phi(\vec{x}^{nat}) + \vec{\varepsilon}^{\,o}$$

$$\vec{x},\vec{x}^{nat} \in X$$

$$\left[\vec{x}^{nat}\right] = [\vec{x}] = I$$

$$\left[\vec{\lambda}^{nat}\right] = \left[\vec{\lambda}\right] = P$$

$$|G| << |F|$$

$$[\vec{y}] = M << I$$

DA formulation of standard initialization methods

FFI: Model state is replaced by the best-possible available estimate of the actual state

$$\vec{x}^a = \vec{x}^b + H^T\left[\vec{y}^o - H\vec{x}^b\right] \qquad \left[H\right] = M \times I$$

$\vec{x}^b$: Background state from long "control" run of the model

AI: Assimilation of observed climate anomalies on an estimate of the model climate

$$\vec{y}^{pso} = \vec{y}^o - \left(\bar{\vec{y}}^o - H\bar{\vec{x}}\right)$$

$$\vec{x}^a = \vec{x}^b + H^T\left[\vec{y}^{pso} - H\vec{x}^b\right]$$

Least Square Initialization (LSI)

Using a least-square framework, the background and observations are linearly combined in order to minimize the expected analysis error covariance. Errors are assumed to be Gaussian.

$$\vec{x}^a = \vec{x}^b + BH^T\left[HBH^T + R\right]^{-1}\left[\vec{y}^o - H\vec{x}^b\right]$$

How do we estimate **B** to initialize s2d predictions?

Proxy: $B^m = \overline{\alpha\left(\vec{x} - \bar{\vec{x}}\right)\left(\vec{x} - \bar{\vec{x}}\right)^T}$ Long time average

$$\alpha : \begin{cases} 1) & \text{Compensate estimation of the nature covariance} \\ 2) & \text{Optimize weighting} \end{cases}$$

Exploring Parameter Uncertainty (EPU)

Goal: Reduce model drift caused by parametric error

Approach: Online bias correction method based on a short-time, linear approximation of the bias evolution

$$\frac{d\delta\vec{x}}{dt} \approx \left.\frac{\partial F}{\partial \vec{x}}\right|_{\vec{x},\vec{\lambda}} \delta\vec{x} + \left.\frac{\partial F}{\partial \vec{\lambda}}\right|_{\vec{x},\vec{\lambda}} \delta\vec{\lambda} \quad \Rightarrow \quad \delta\vec{x} \approx \cancel{\mathbf{M}_{t,t_0}\delta\vec{x}_0} + \int_{t_0}^{t} \mathbf{M}_{t,t_0}\delta\mu(\tau)\,d\tau$$

Deterministic!

Bias estimation proposal:

$$\vec{b}(t) = \langle \delta\vec{x}(t) \rangle$$
$$\approx \langle \delta\mu_0 \rangle \left[t - t_0\right] = \langle \left.\frac{\partial F}{\partial \vec{\lambda}}\right|_{\vec{x},\vec{\lambda}} \delta\vec{\lambda} \rangle \left[t - t_0\right]$$

Unknown!

Exploring Parameter Uncertainty

Working Hypothesis:

Identification of a set of uncertain parameters
Identification of a range of possible parameter values $\left[\lambda_{\min}, \lambda_{\max}\right]$

$$\vec{x}^{un}(t_i) = \vec{x}(t_i) - \vec{b}(t_i)$$

$$= \vec{x}(t_i) - \left.\frac{\partial F}{\partial \lambda}\right|_{\vec{x}(t_{i-1}),\vec{\lambda}} \delta\lambda_i \Delta T_{Bias} \qquad t_i = i\Delta T_{Bias} \qquad i = 1,2,\ldots$$

$$\delta\vec{\lambda}_i \in \begin{cases} U(0, \vec{\lambda}^{\max} - \bar{\vec{\lambda}}) \\ U(\vec{\lambda}^{\min} - \bar{\vec{\lambda}}, 0) \end{cases} \quad \textbf{if} \quad \begin{matrix} \vec{\lambda} > \bar{\vec{\lambda}} \\ \vec{\lambda} < \bar{\vec{\lambda}} \end{matrix}$$

Idealized coupled model

Coupling

Extratropical Atmosphere

$$\dot{x}_e = \sigma(y_e - x_e) - c_e(Sx_t + k_1)$$
$$\dot{y}_e = rx_e - y_e - x_e z_e + c_e(Sy_t + k_1)$$
$$\dot{z}_e = x_e y_e - bz_e$$

c_e

Tropical Atmosphere

$$\dot{x}_t = \sigma(y_t - x_t) - c(SX + k_2) - c_e(Sx_e + k_1)$$
$$\dot{y}_t = rx_t - y_t - x_t z_t + c(SY + k_2) + c_e(Sy_e + k_1)$$
$$\dot{z}_t = x_t y_t - bz_t + c_z Z$$

c, c_z

Tropical Ocean

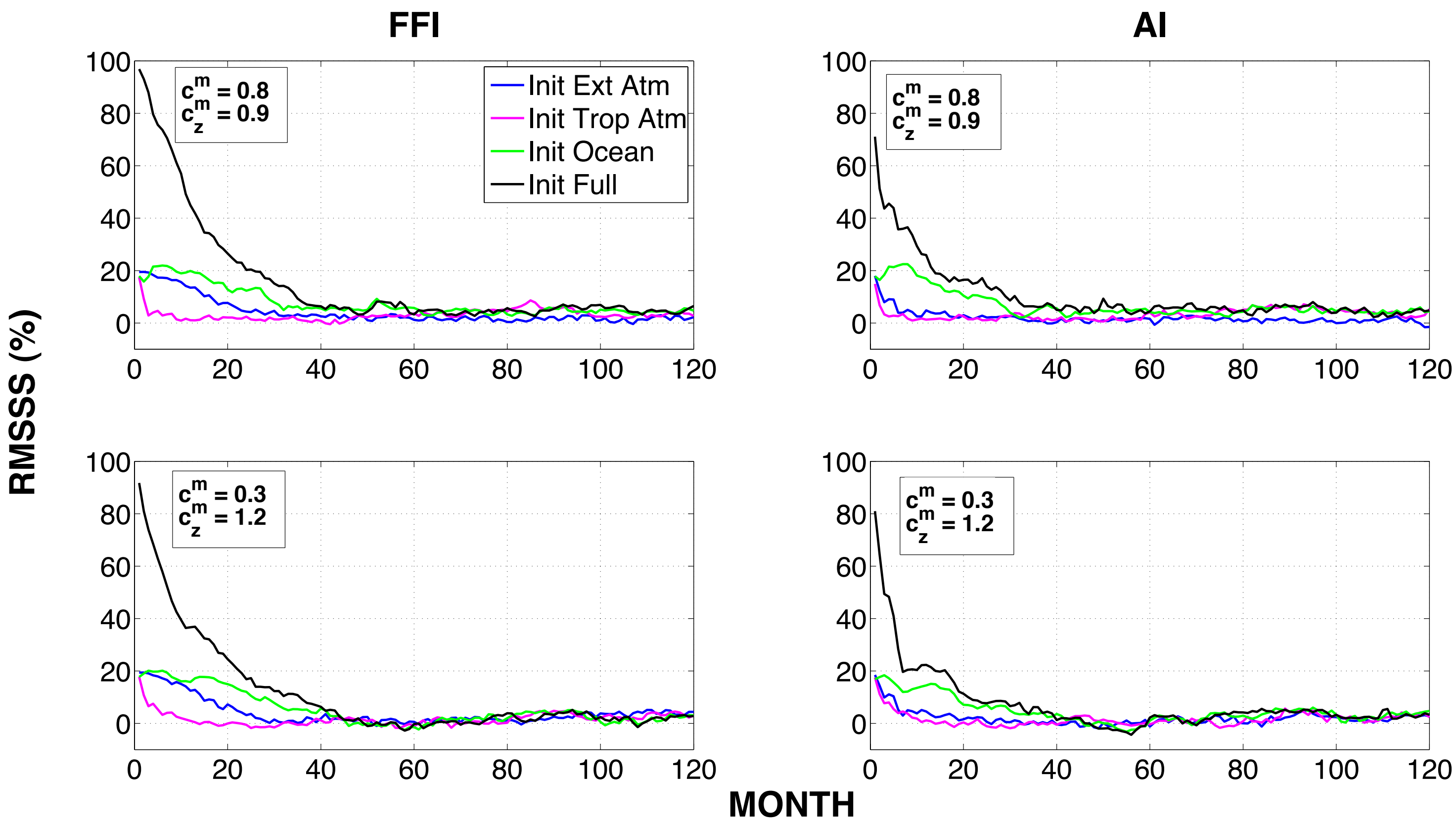
$$\dot{X} = \tau\sigma(Y - X) - c(x_t + k_2)$$
$$\dot{Y} = \tau rX - \tau Y - \tau SXZ + c(y_t + k_2)$$
$$\dot{Z} = \tau SXY - \tau bZ - c_z z_t$$

(Pena & Kalnay, 2004)

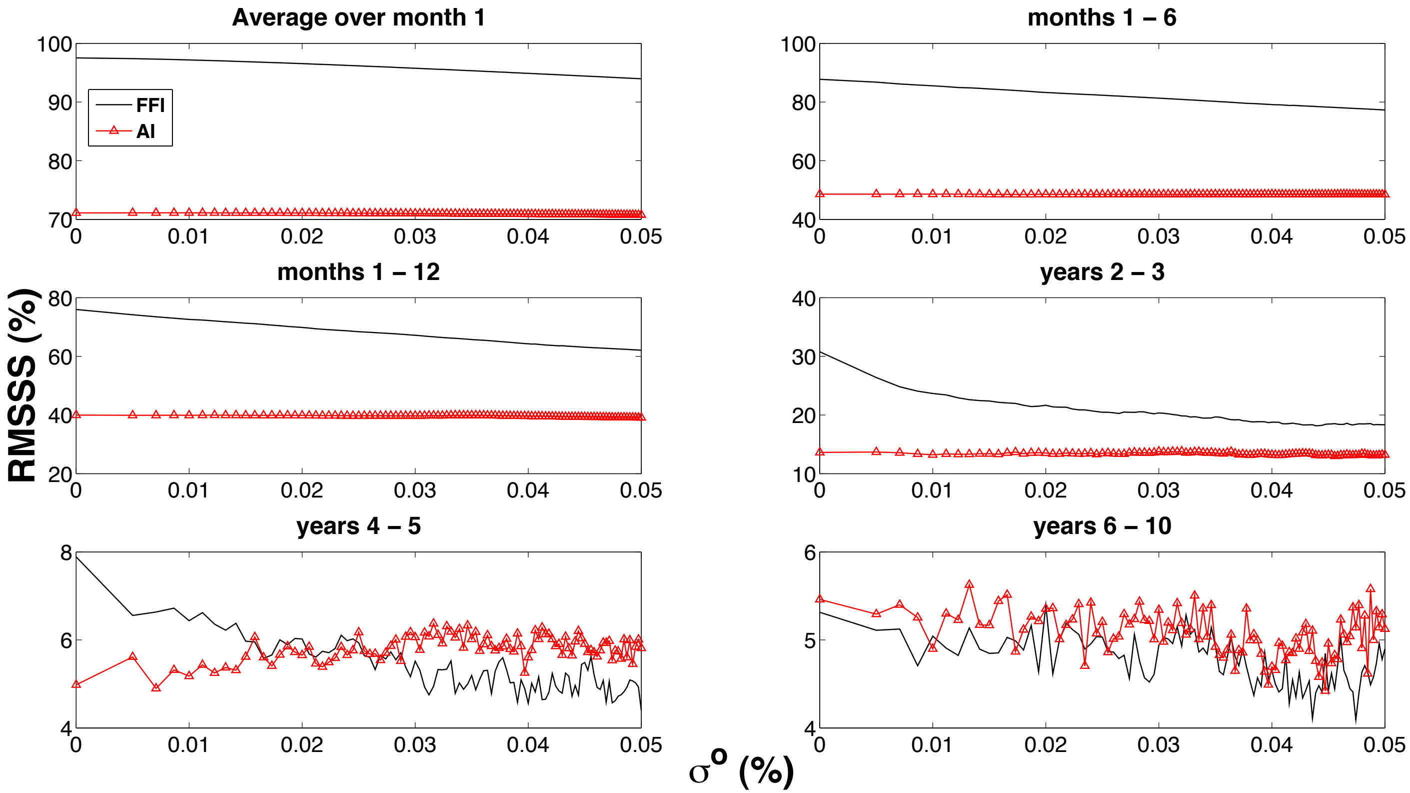
Experimental Setup

- Nature: Solution to model equations with coupling strength
  $c_e = 0.08, \quad c = c_z = 1$
- Introduction of parametric model error by modifying simultenously c^m, c_z^m in the range of $(0.1 - 1.5)$
- The observations are simulated by sampling the nature trajectory each month and adding a Gaussian white noise
- 10-yr long predictions are initialized each month for 30 yrs: Total of 360 startdates

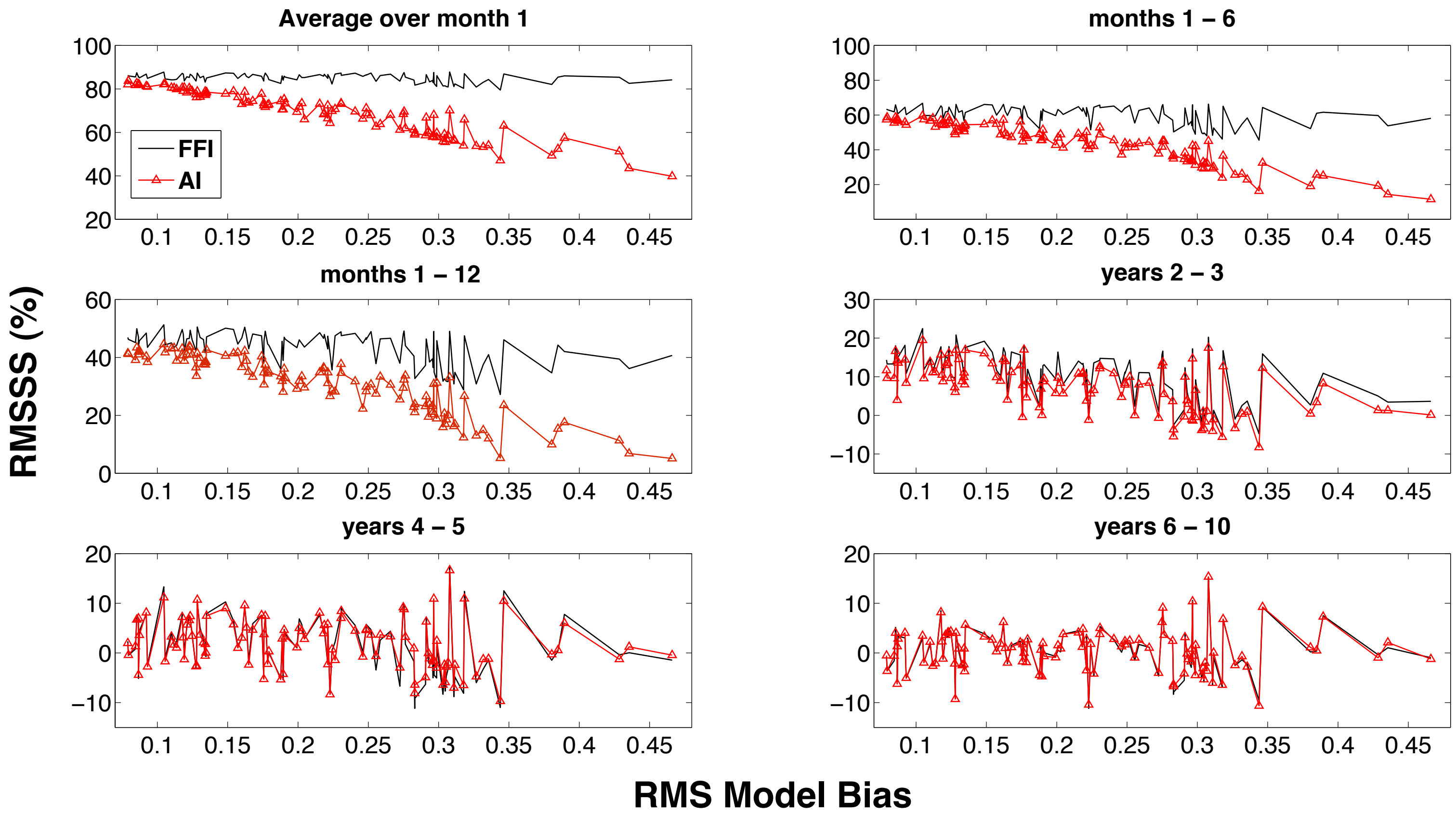
Comparison AI / FFI



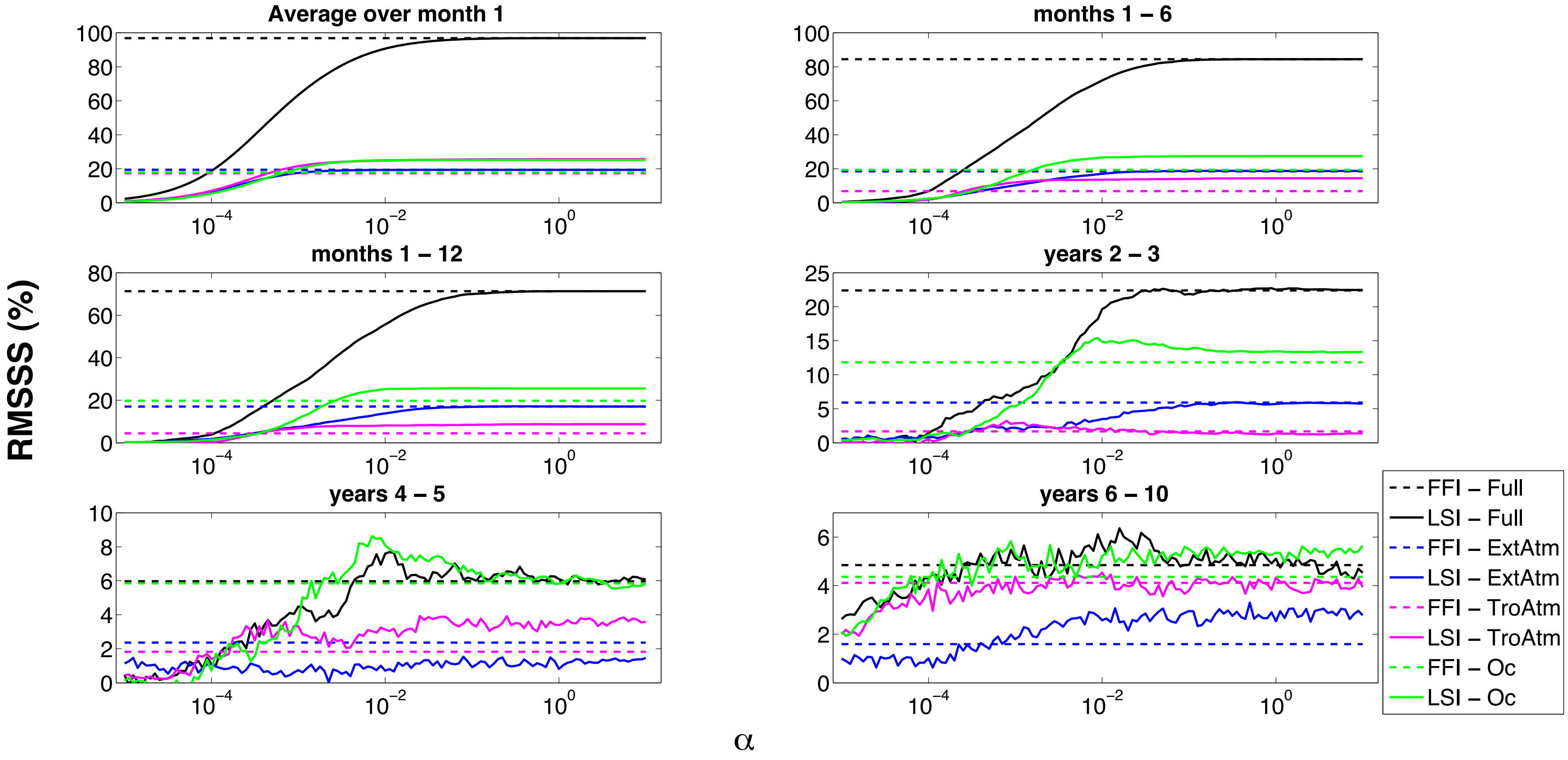
Influence of observational error



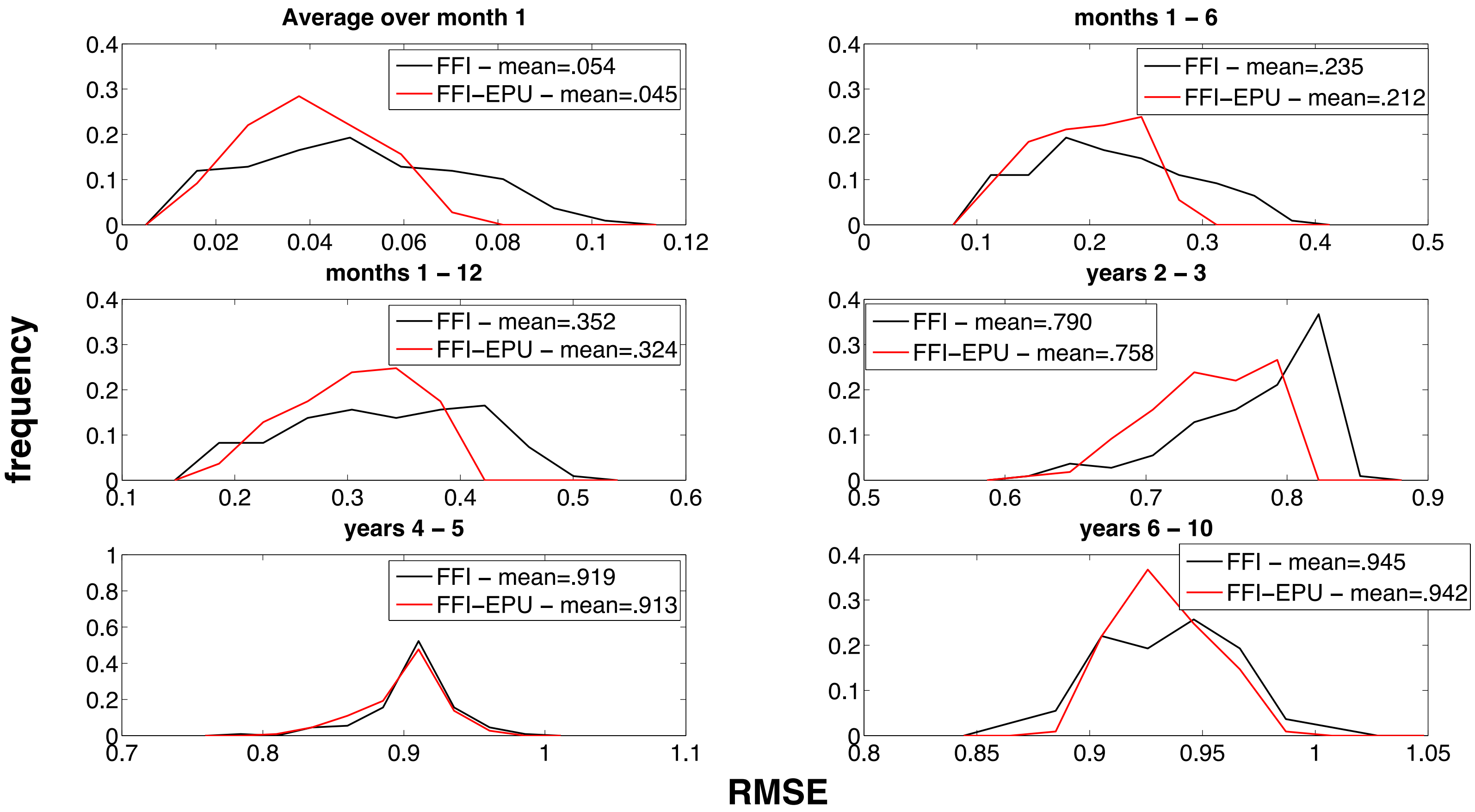
Influence of model bias



FF - LSI

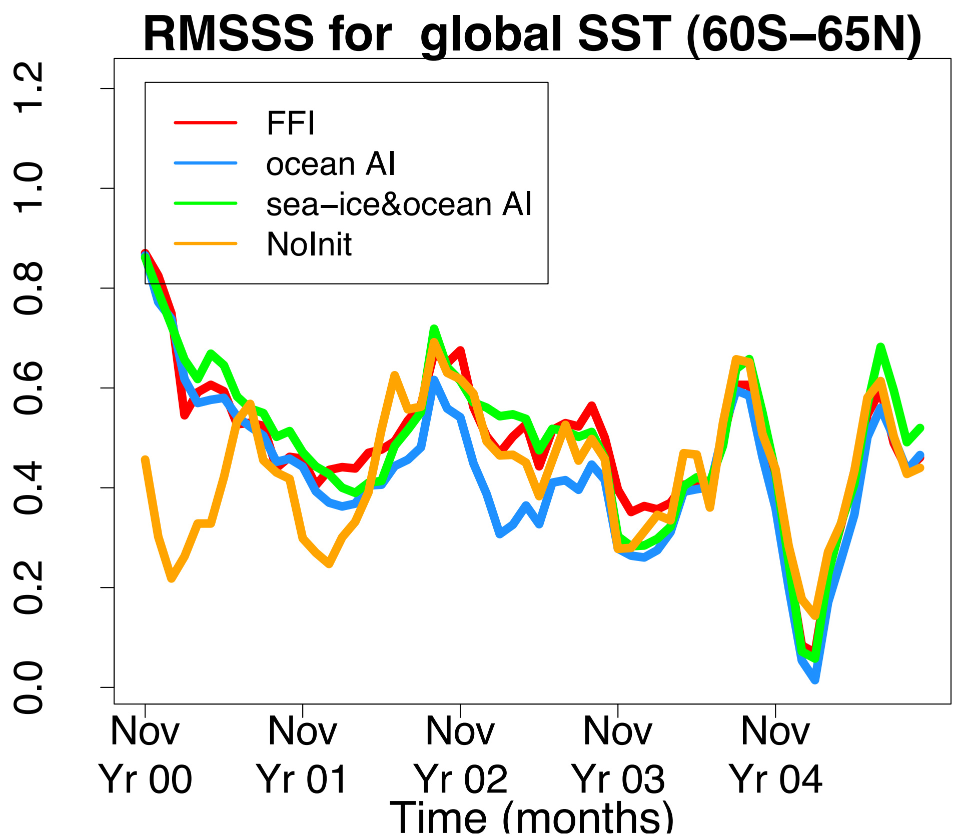

Average over month 1
months 1 – 6
months 1 – 12
years 2 – 3
years 4 – 5
years 6 – 10
RMSSS (%)
α
FFI – Full
LSI – Full
FFI – ExtAtm
LSI – ExtAtm
FFI – TroAtm
LSI – TroAtm
FFI – Oc
LSI – Oc

FFI - EPU



Average over month 1
FFI – mean=.054
FFI–EPU – mean=.045
months 1 – 6
FFI – mean=.235
FFI–EPU – mean=.212
months 1 – 12
FFI – mean=.352
FFI–EPU – mean=.324
years 2 – 3
FFI – mean=.790
FFI–EPU – mean=.758
years 4 – 5
FFI – mean=.919
FFI–EPU – mean=.913
years 6 – 10
FFI – mean=.945
FFI–EPU – mean=.942
frequency
RMSE

Coupled General Circulation Model: EC-Earth



Conclusions

- FFI slightly outperforms AI
- The ocean alone mainly determines the system's predictability horizon
- FFI sensitive to observational error, AI to model bias
- LSI helps improve prediction skill mainly by spreading out information, otherwise restricted to the observation subspace, to the entire model space
- Clear benefit of FFI-EPU up to 3 years in the forecast (4-17%)
- EC-Earth: added value from initialization for first forecast year, with only minor differences between initialization procedures

Reference: Carrassi A., R. Weber, V. Guemas, F. Doblas-Reyes, M. Asif, and D. Volpi 2013 (submitted to Nonlin. Processes Geophys.)

Thank you