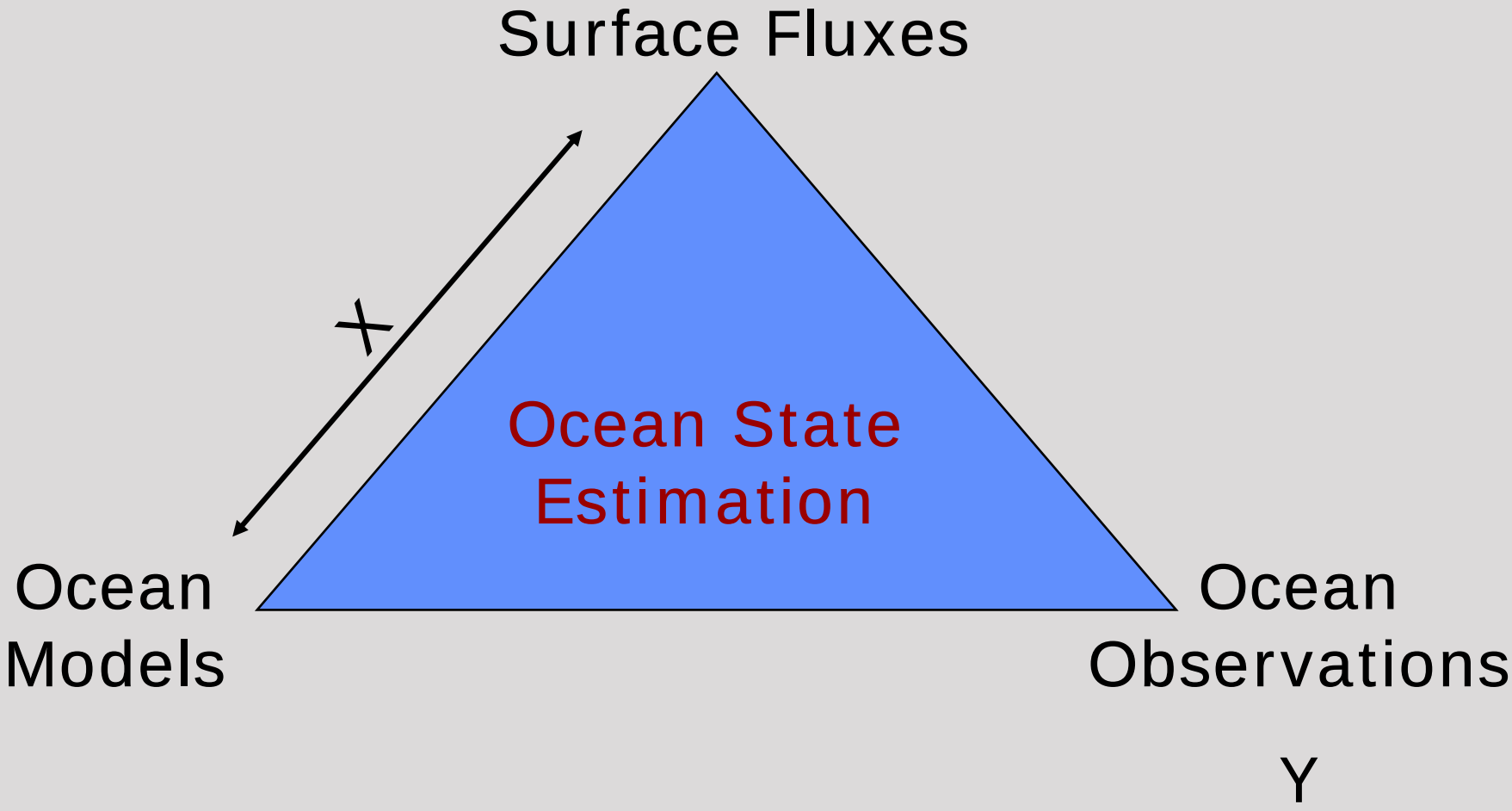
A Critical Perspective of Ocean Reanalysis

Magdalena A. Balmaseda ECMWF (UK)

Overview

- **Estimation of climate variability**
  - Uncertainties and Approximations
  - Some outstanding questions
- **Initialization of coupled forecasts**
  - Real world versus model attractor: Initialization shock and forecast skill
  - Overview of approaches
- **Statistical corrections & Dynamical models**

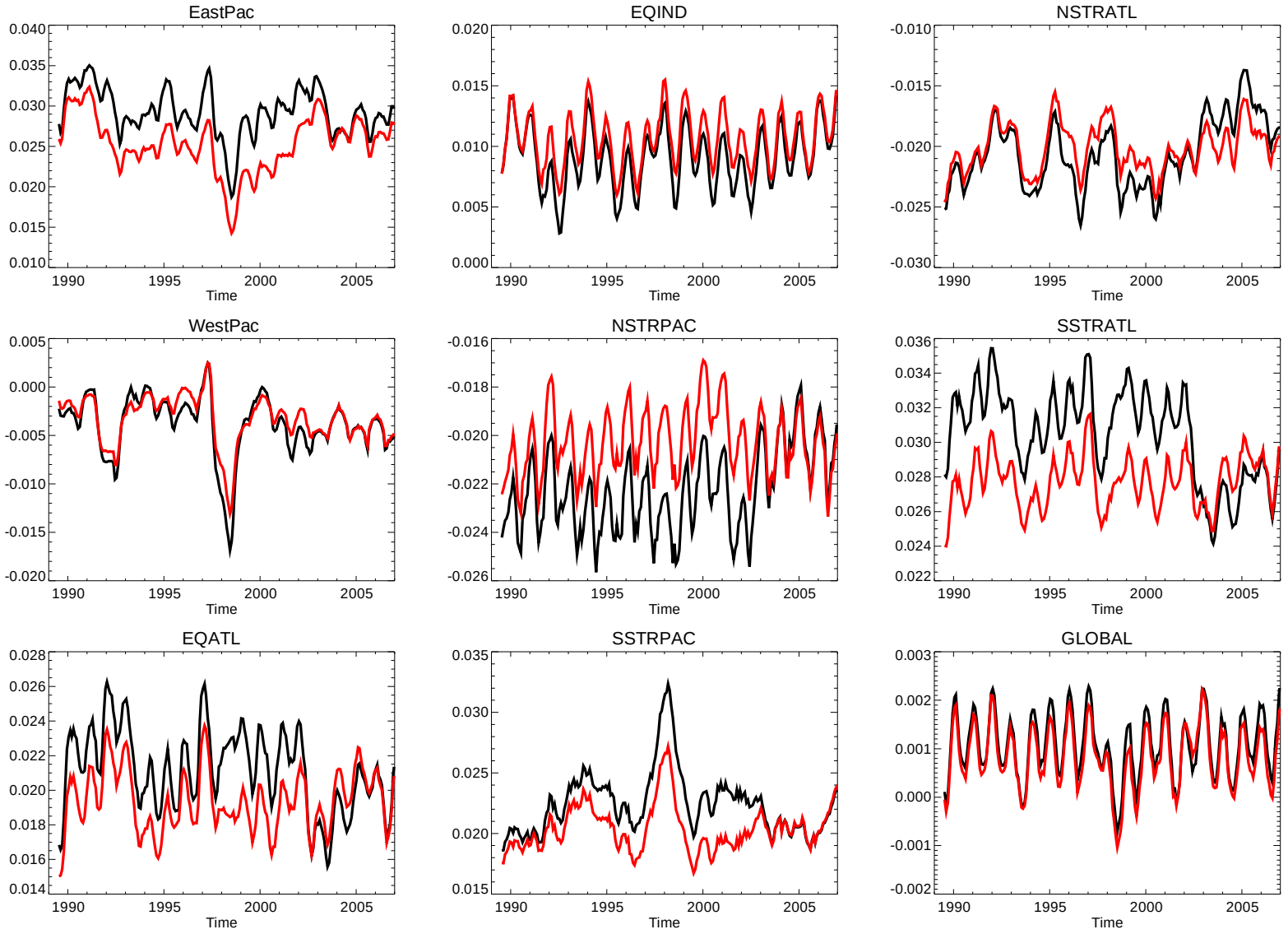


$$\mathbf{J}(\delta\mathbf{x}) = (\mathbf{x}+\delta\mathbf{x})^{\mathrm{T}}\mathbf{B}(\mathbf{x}+\delta\mathbf{x}) + (H(\mathbf{x}+\delta\mathbf{x})-\mathbf{y})^{\mathrm{T}}\mathbf{R}(H(\mathbf{x}+\delta\mathbf{x})-\mathbf{y})$$

Uncertainties and Approximations

- Surface fluxes:
  - From atmospheric reanalysis. Large uncertainties and errors in mean and time variability.
  - There have been some efforts on quantifying uncertainty.
- Ocean models
  - Resolution (~ low resolution (1x1 deg), although SODA ¼ deg. Ongoing efforts at ¼ of deg).
  - Physical parameterizations.
  - Uncertainty far from being quantified.
- Ocean observations:
  - Measurement ad representativeness error
  - Limited sample of variables (T,S, SSH). Non stationary in time and space
  - Still large source of uncertainty. Especially in the historical record
- Cost function
  - Control vector: which variables should be modified? Forcing, initial conditions?
  - Analysis cycle: time discretization
  - Treatment of bias
  - **B** and **H**: how to propagate the information from control vector to other variables?
  - **R**: observation error (usually assumed diagonal, observation error not correlated)

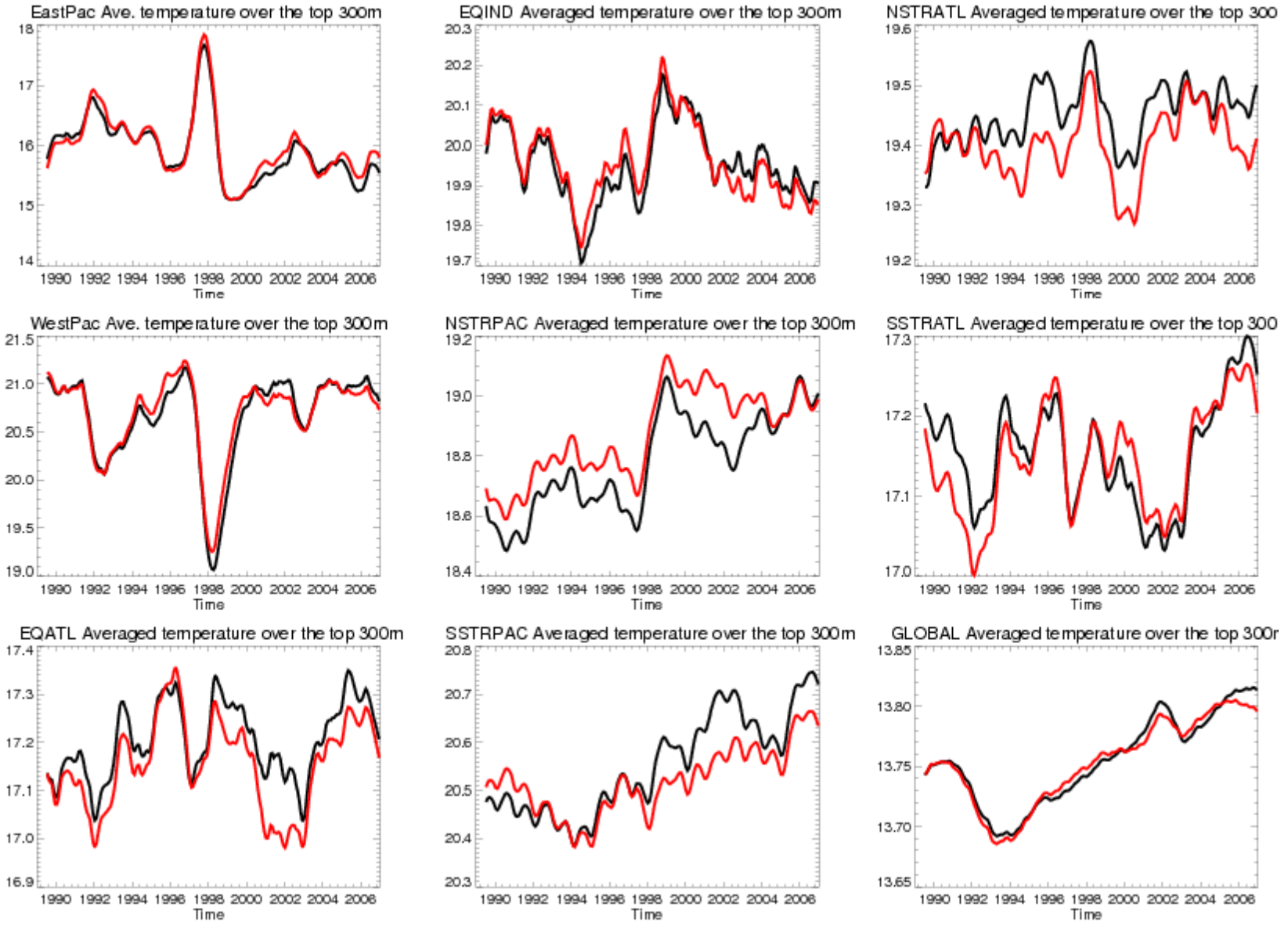
Uncertainty in forcing fluxes: Meridional wind stress



ERA-40/OPS

ERA-INTERIM

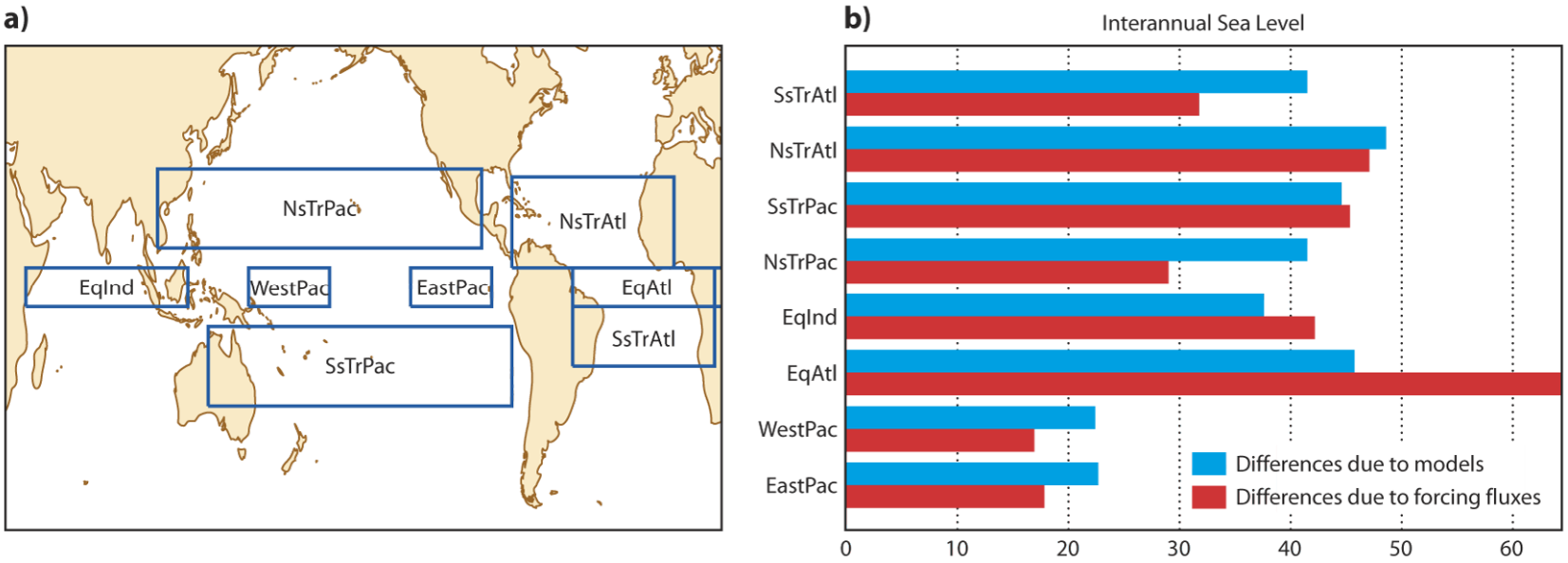
Upper Ocean HEAT CONTENT



ERA-40/OPS

ERA-INTERIM

Uncertainty: Fluxes and Ocean Model



Ocean models: HOPE v NEMO :(1x1 deg with eq. refinement)

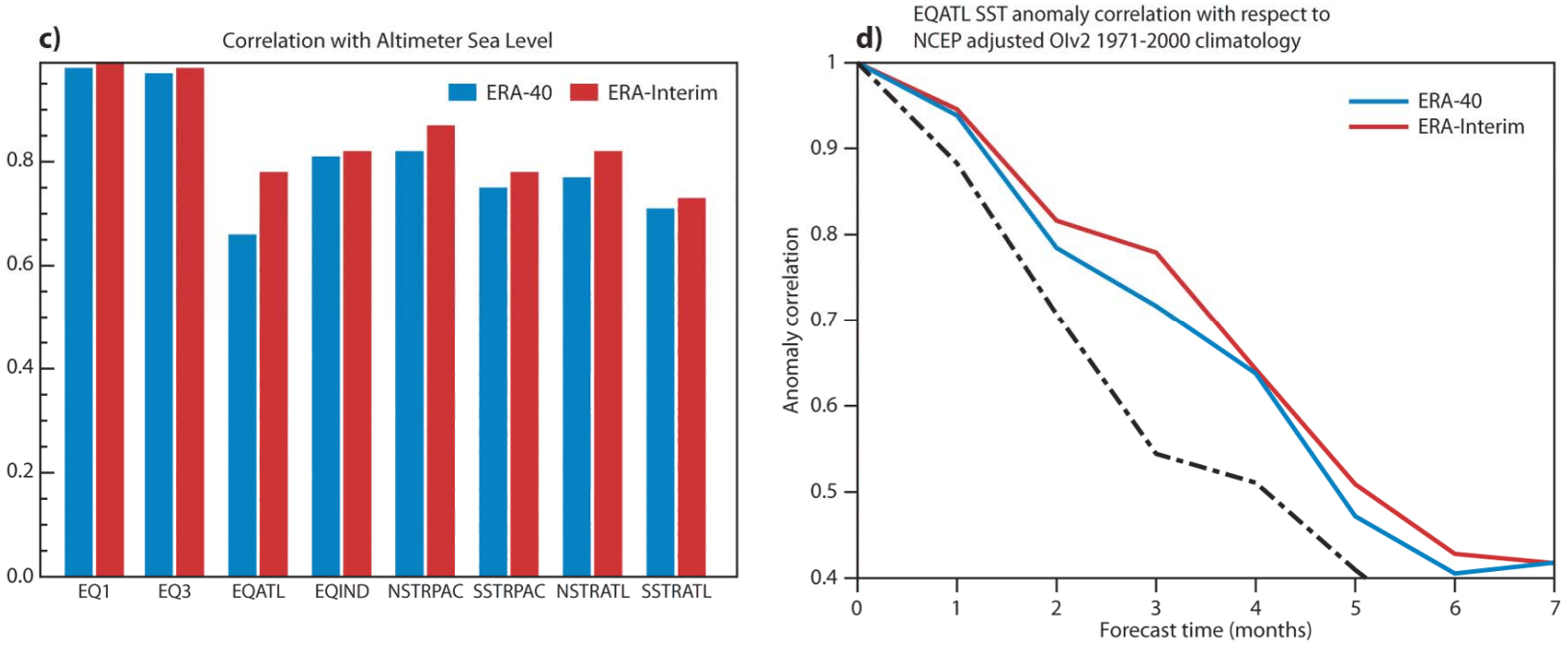
Fluxes: ERA40 v ERA-INTERIM

Integration 1989-2007, strongly relaxed to SST.

Uncertainty in Interannual Variability~ 40%-60%

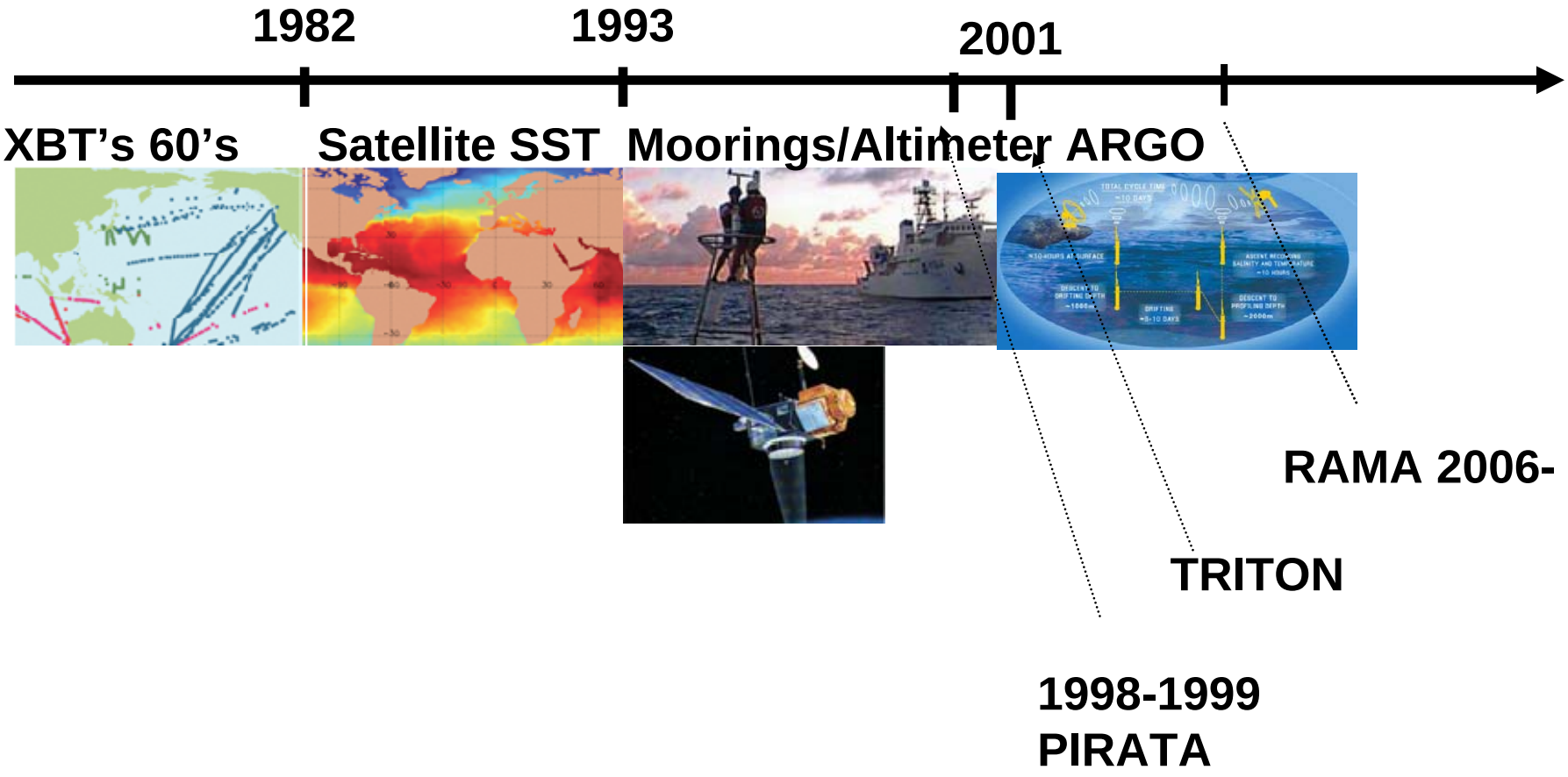
Best constrained: Eq. Pacific. Worst constrain: Eq. Atlantic

Continuous Improvements

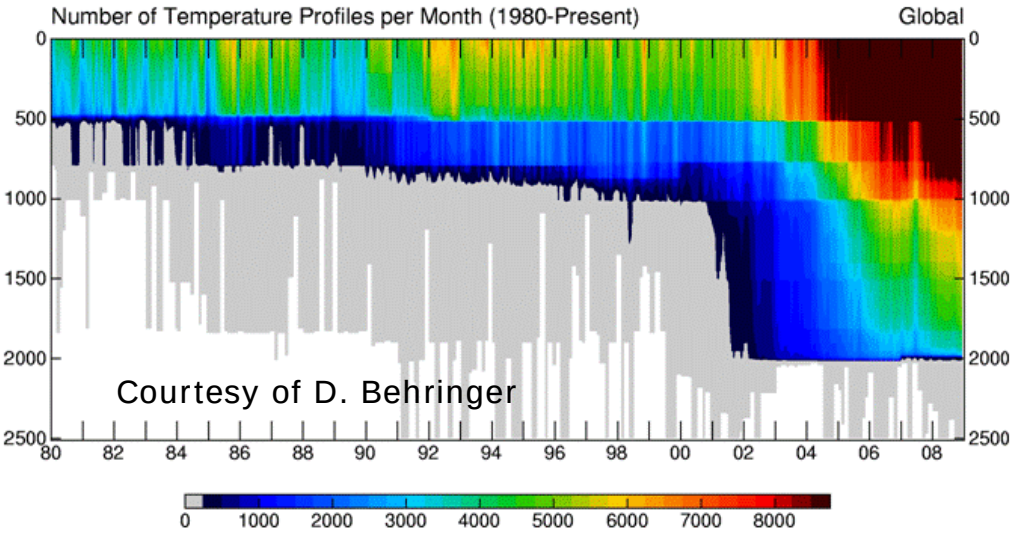


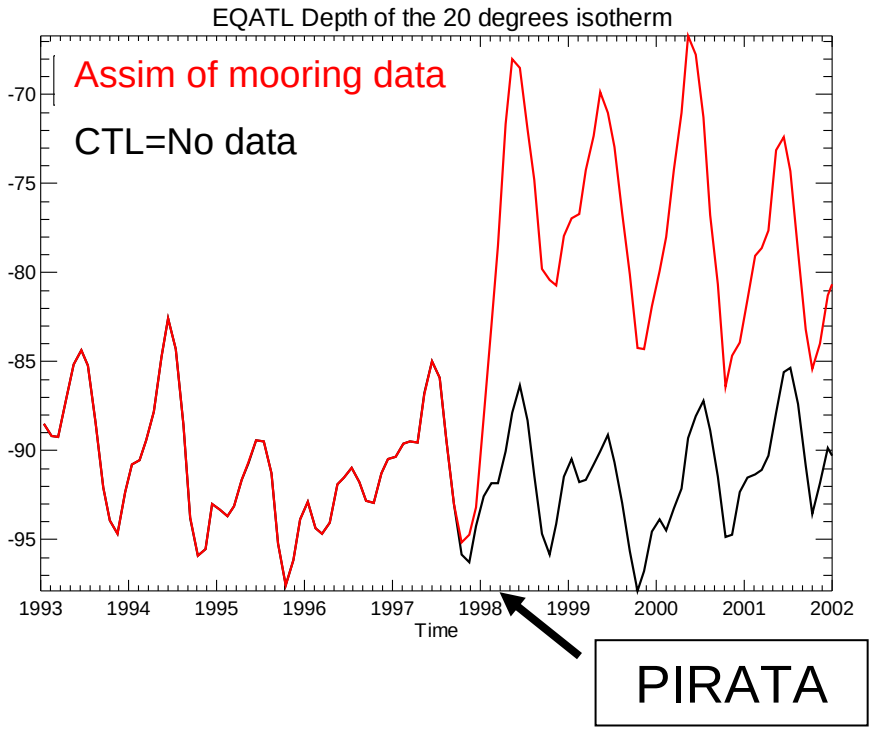
- ERA-Interim seems superior to ERA-40/Ops. But it only goes back to 1989.
- Still large uncertainties in heat fluxes and P-E
- There is room for a-posteriori correction of direct fluxes from the atmospheric reanalysis (CORE, DRAKKAR, OAFlux)
- It is important to represent the daily variability: monthly means are not enough

Time evolution of the Ocean Observing System



Bias+Changing GOOS = Spurious Variability





Large impact of data in the mean state:
Shallower thermocline

Impact of ocean observations in the mean state

Model estimates are largely biased.

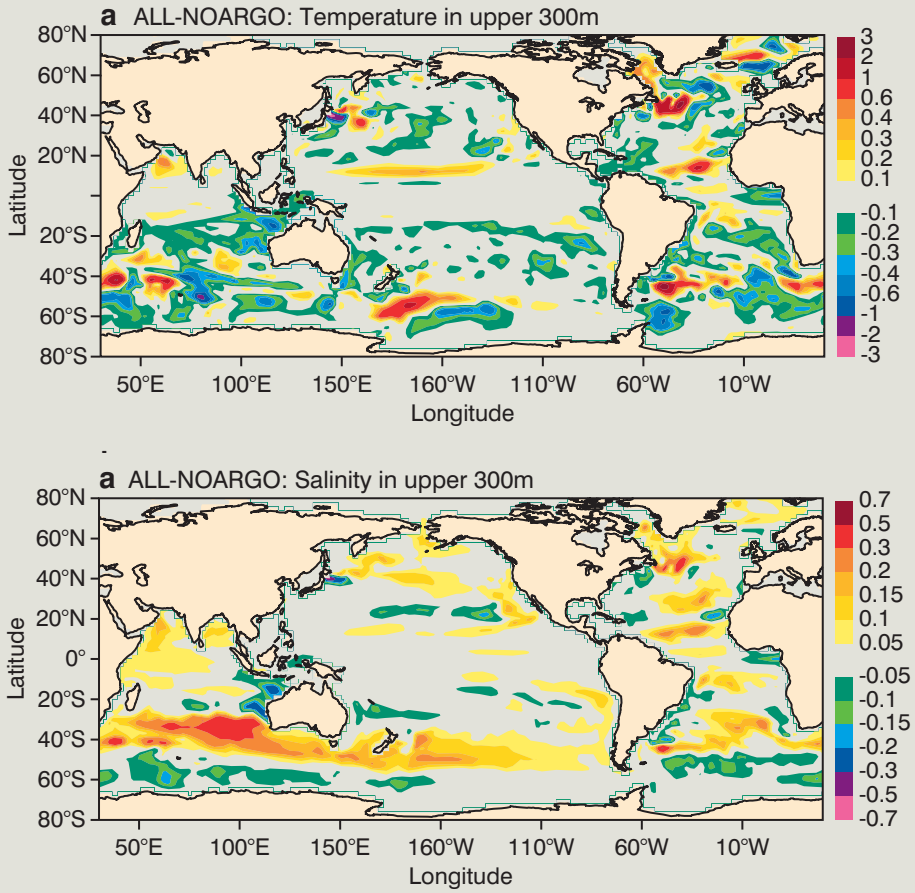
Usually, the assimilation systems assume un-biased states.

A common approach is the relaxation to climatology (1-3 yr time scale): this is not acceptable for decadal variability

Alternative: additive and adaptive bias corrections schemes.

A better alternative is to reduce the bias by improving model/fluxes

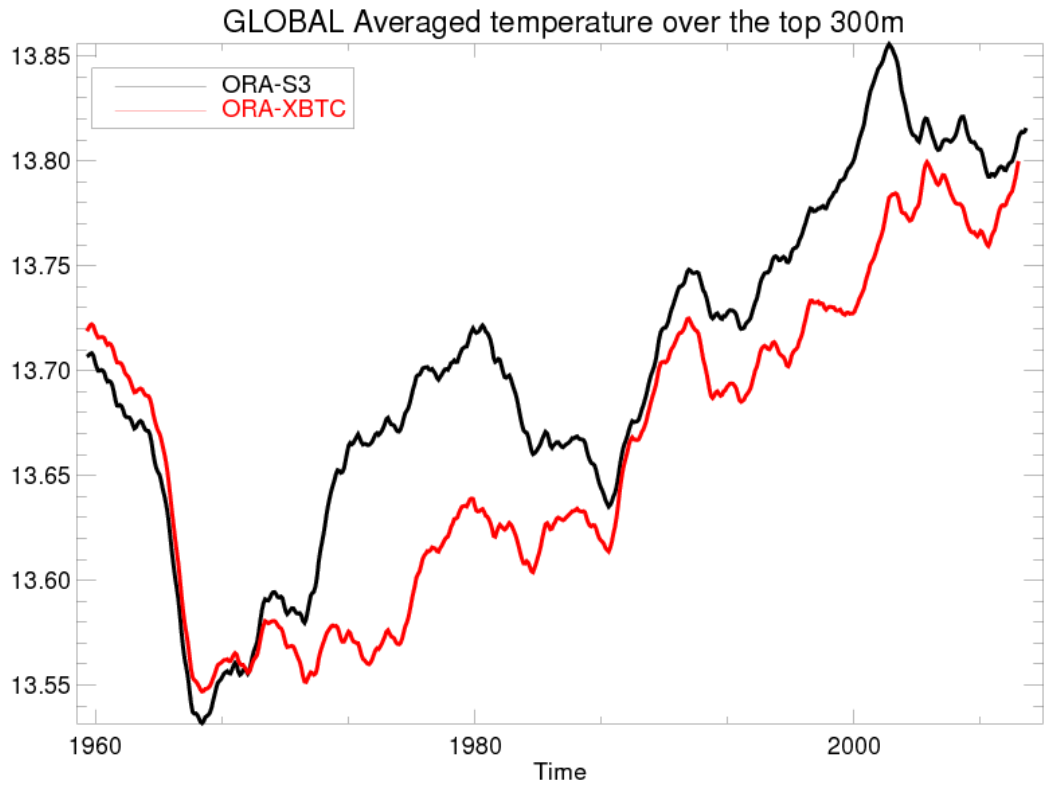
Effect of Argo in the upper ocean



Balmaseda et al 2007

Observation error: XBT example

Upper 300m Global Ocean heat content



Averaged temperature in upper 300m

ORA-S3: operational re-analysis, used in ENSEMBLES. XBT's were not corrected

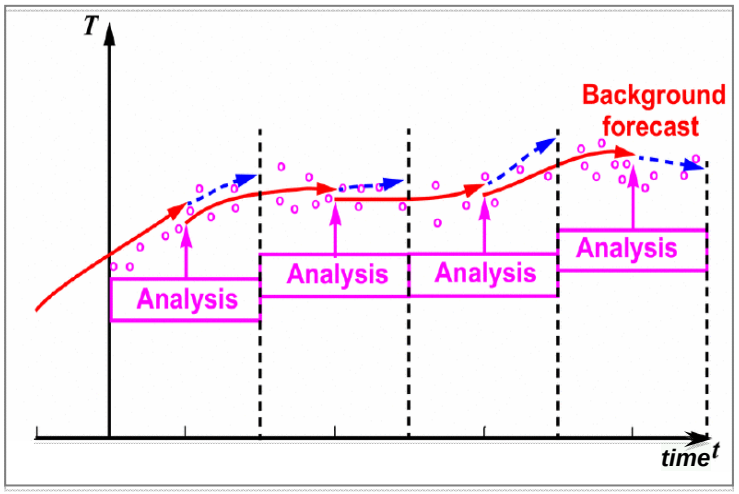
ORA-XBTC: as ORA-S3, but with corrected XBT's.

70's: uncorrected XBT's induce spurious warming

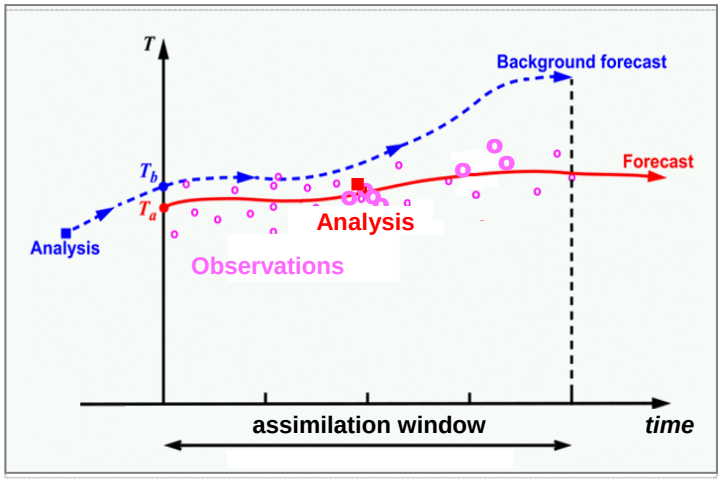
Quite some uncertainty in XXI century evolution.

Assimilation Methods

- Ocean DA is a hugely under-determined problem
- A variety of methods is currently used: OI, EnKF, En-OI, 3D-var, 4D-var (or adjoint methods)
- Sometimes methods allow/require different implementation.
  - Example: Length of assimilation cycle.



Sequential Methods: 10days



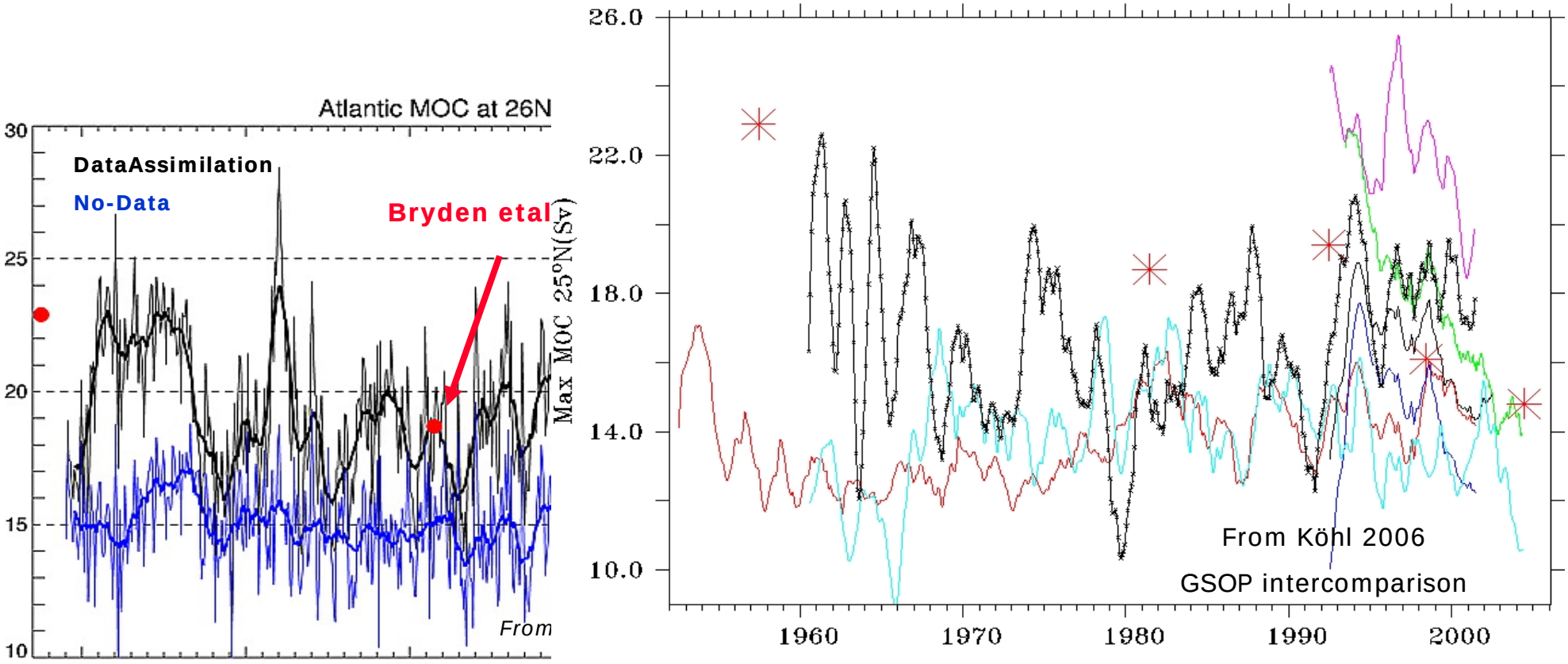
Adjoint methods: 2 months or 40 years?

But often, the solution depends on much simpler parameters

- Relative weight to observations and background
  - Simple yet highly sensitive parameter !!
  - Should depend on time and space.
  - Current approximations are quite ad-hoc.
- De-correlation scales
  - What is the area of influence of an observation?
  - Representation of complex spatial structures
  - What to do below the lowest observation?
- Multivariate Constraints
  - What is the relation between errors in different variables?
    - Temperature and salinity
    - Density/velocity
  - Need to preserve dynamical balances, water mass characteristics…

More work needed to improve the error covariances and understand their impacts

Estimation of the Atlantic MOC



Reasons for the large dispersion among different ocean reanalysis are not understood. Possibilities:

Correction of the pressure gradients

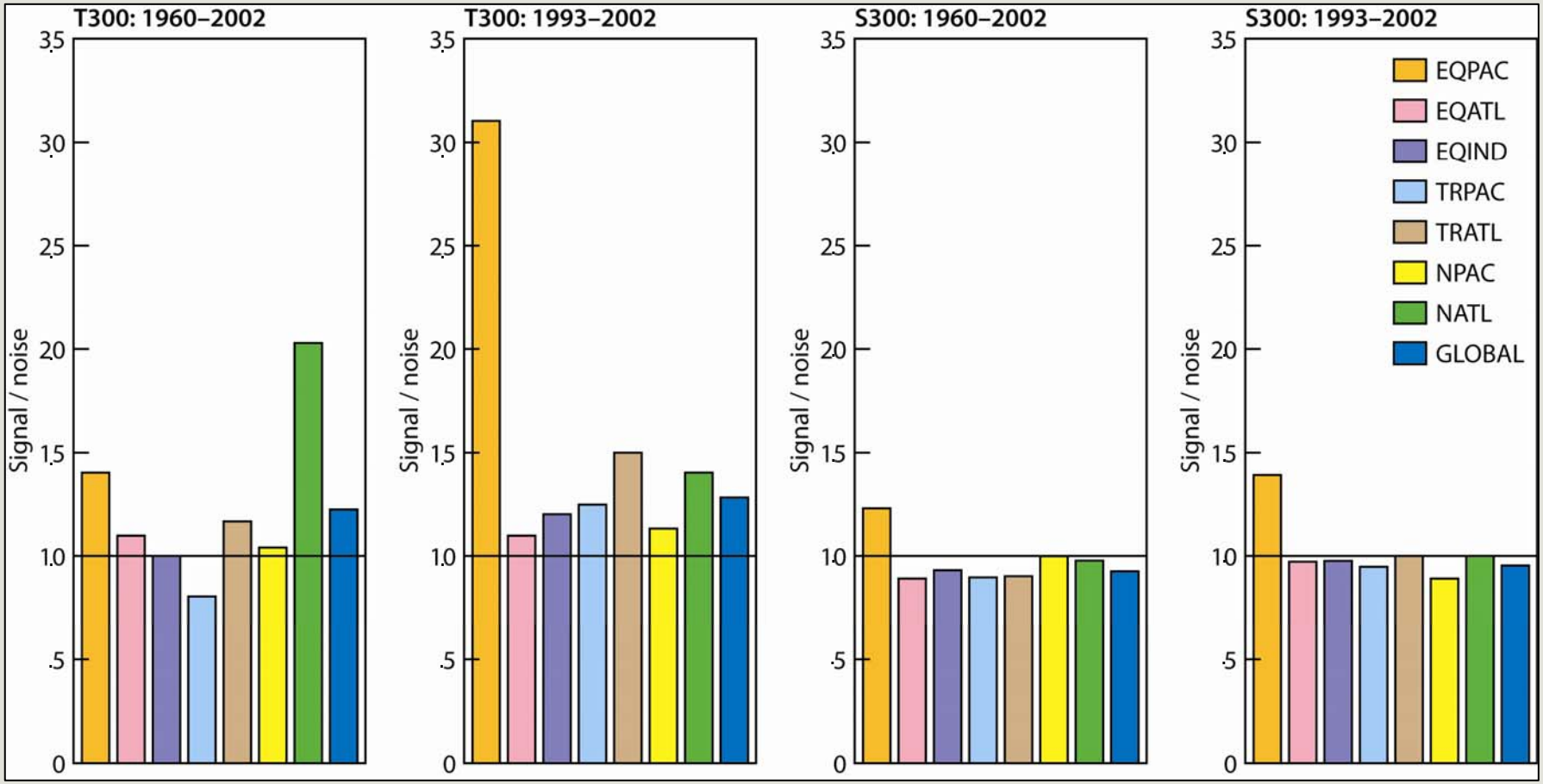
Propagation of information below the lowest observational level

Observations close to the coast: representativeness error?

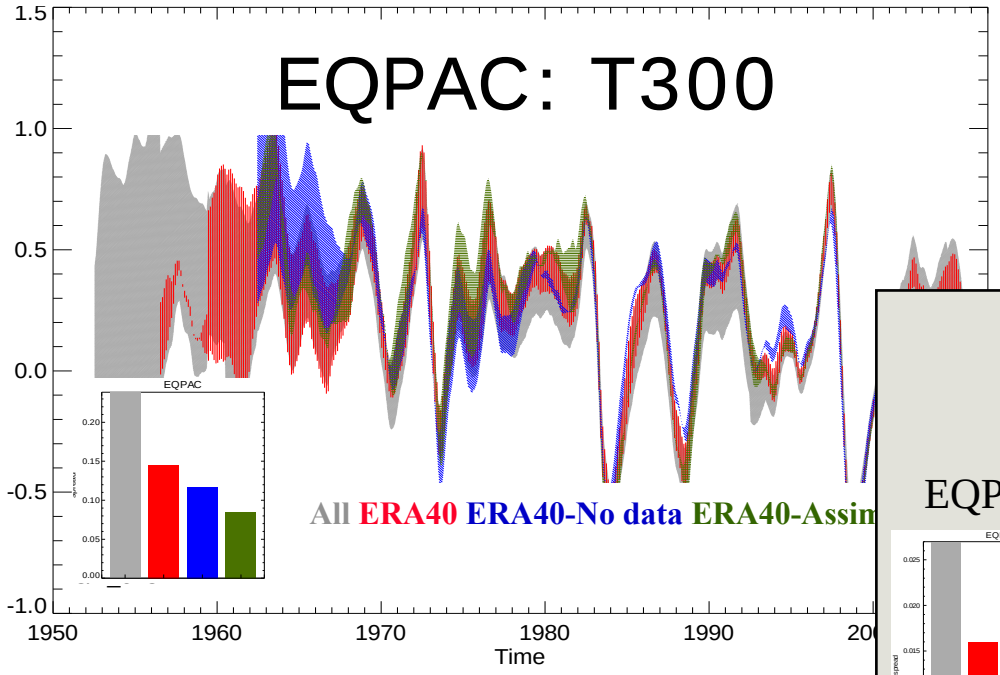
It is imperative to understand the AMOC sensitivity to the error covariances

How well do ocean reanalysis detect climate variability? --- Signal/Noise Ratio

Signal/Noise ratio of upper 300m averaged T & S in various regions for 20 products (Balmaseda & Weaver 2006)



Sources of Uncertainty

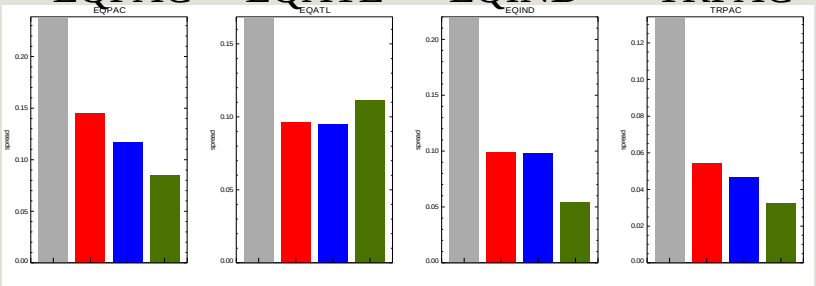


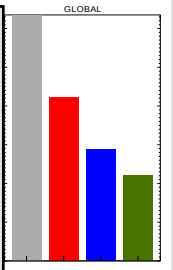
Upper ocean temperature: Assimilation reduces uncertainty. Except for Tropical Atlantic

Upper ocean salinity: Assimilation increases uncertainty!!

Temperature

EQPAC EQATL EQIND TRPAC

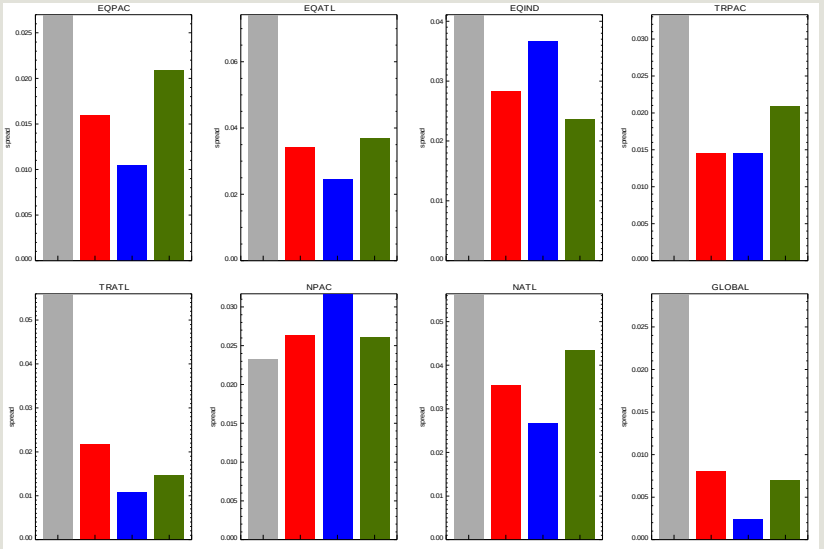




GLOBAL

Salinity

EQPAC EQATL EQIND TRPAC



TRATL NPAC NATL GLOBAL

Validation of Ocean Reanalysis

- How good are the products?
- Is there any product better than other? Any methodology better than other?
  - ➢ Does data assimilation provide a better estimate than ocean forced by surface forcing?
    - o Yes for interannual variability, at least from the impact on forecast skill
    - o For decadal variability and trends?

Assessment is difficult, due to lack of independent data

Improving the Estimation of Ocean Variability

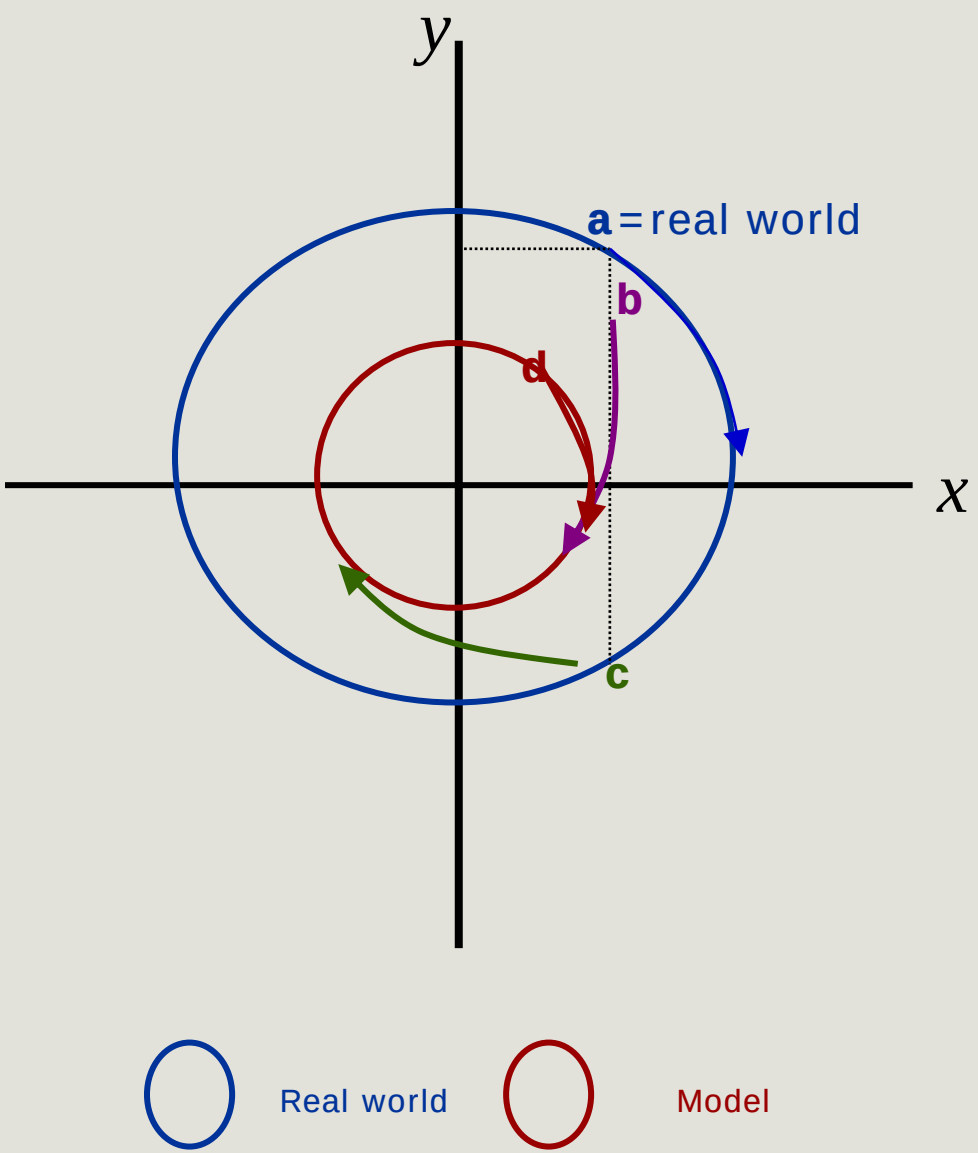
- **Improvement the forcing fluxes**
  - Improved atmospheric reanalysis
  - Better treatment of air/sea interaction in the production of atmospheric reanalysis (e.g. SST as a control variable?)
  - The reanalysis should extend as far back as possible, and be continued in real time.
  - A-posteriori correction to the direct fluxes from atmospheric re-analysis. (CORE/DRAKKAR/OAFlux)
- **Ocean model improvement**
  - What do we gain with increased resolution?
  - Other?
- **Improved data assimilation methods**
  - Error Covariance Modelling (making the most of observations)
  - Revisit standard practices
  - Careful choice of control vectors.
  - Bias corrections. Extrapolation to the past of recent observations
  - Systematic development and less “production”.
- **Continuous efforts on archives of ocean observations**
  - Sustained observing system. New observational campaigns
  - Observation retrieval & quality control
  - Data for direct assimilation as well as for validation (transports and currents are very valuable independent data)
- **Coupled data assimilation (of ocean observations) is not the panacea for estate estimation**

Ocean Reanalysis and Initialization of Forecasts

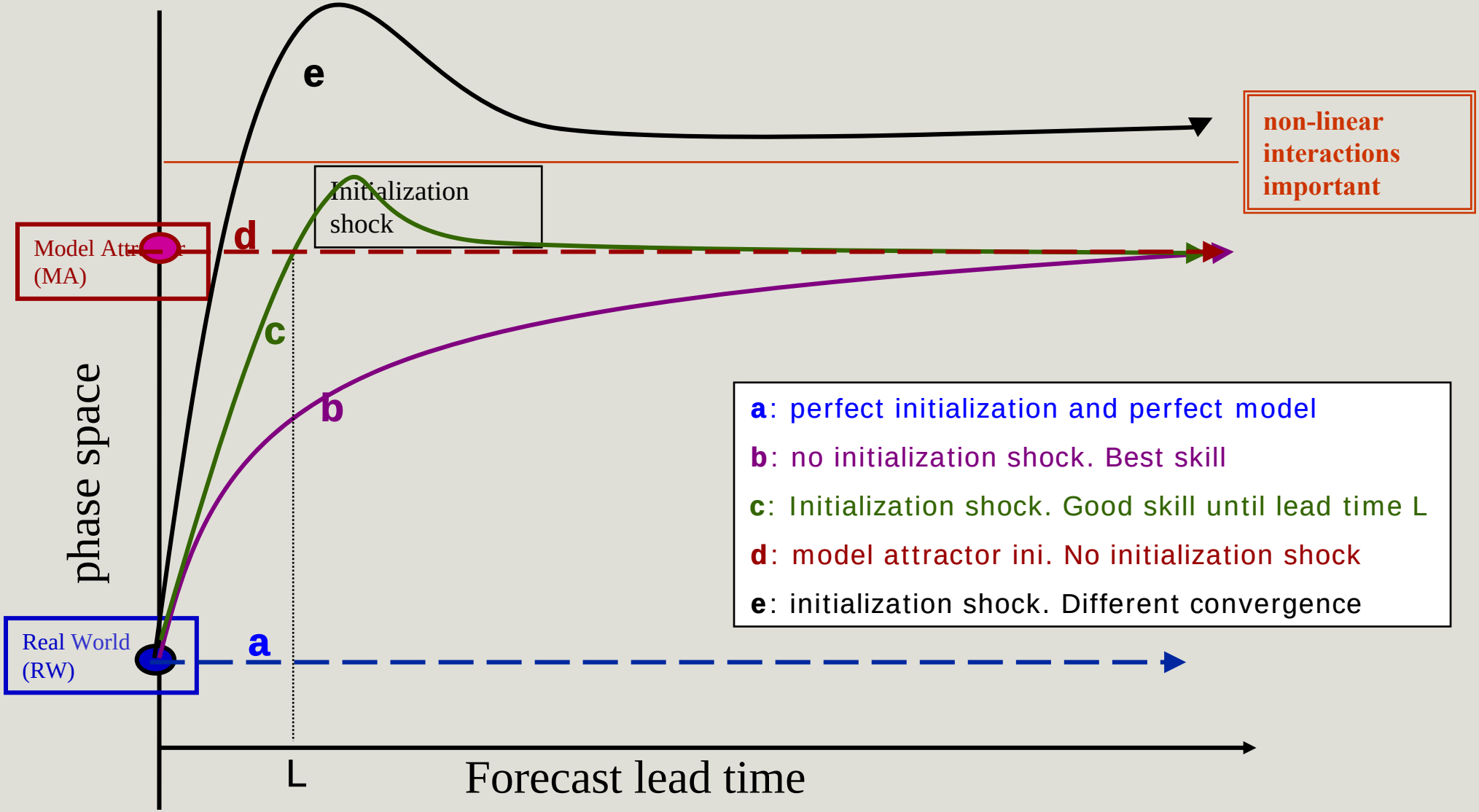
Even if the ocean estate estimation is perfect (x(t)=obs(t)).

It would not guarantee good initialization. For the forecast problem the time derivative is also important.

Besides, models are imperfect, and model attractor is not the same as the real world attractor.



Initialization Shock and Skill



Initialization Shock not opposite to Forecast Skill

- It is possible to have real world initialization without initialization shock and "best skill"
- The impact of initialization shock on skill depends on the lead time.
  - At seasonal time scales , the benefits of real-world initialization in forecasts skill are clear, in spite of initialization shock. The Equatorial Atlantic is an exception (Balmaseda and Anderson 2009).
  - At decadal time scales we don't know yet, since it is difficult to quantify skill.
- Coupled Data Assimilation and Model Attractor Initialization are not synonyms.
- Anomaly Initialization does not imply Model Attractor Initialization.
- Coupled Data Assimilation does not imply Anomaly Initialization

1) Coupled Data Assimilation: MEAN+ ANOMALY

•**The assimilation method apply mainly to the ocean observations**.

•**Not intention to initialize the fast time scales of the atmospheric component**. The atmospheric observations are either neglected or binned in long windows. They can also be used indirectly via nudging to existing atmospheric reanalysis.

•The aim is to produce better (more balanced) initial conditions rather than an accurate estimation of the ocean variability.

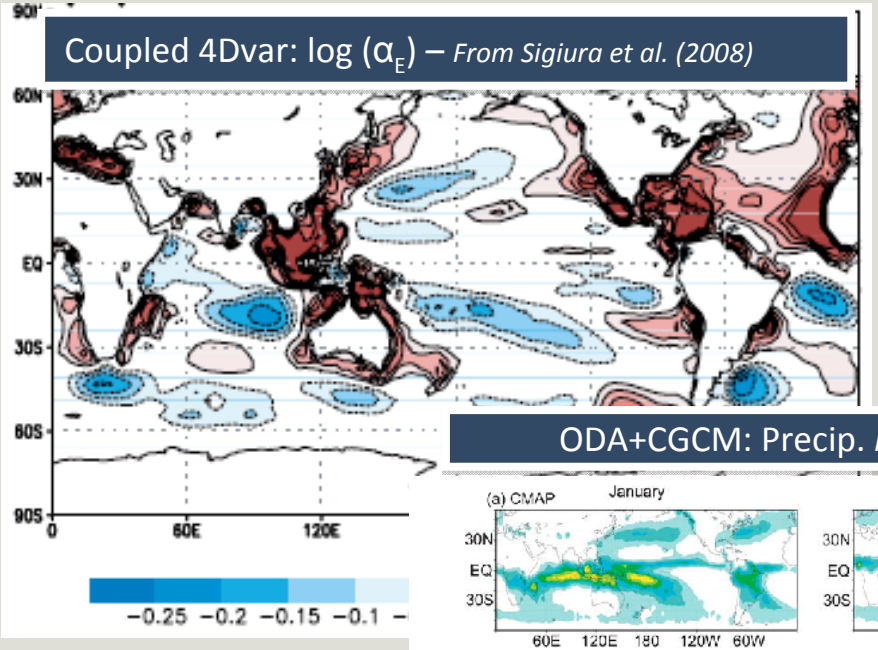
•**Observations are used to correct both the mean and the variability.**

Existing efforts

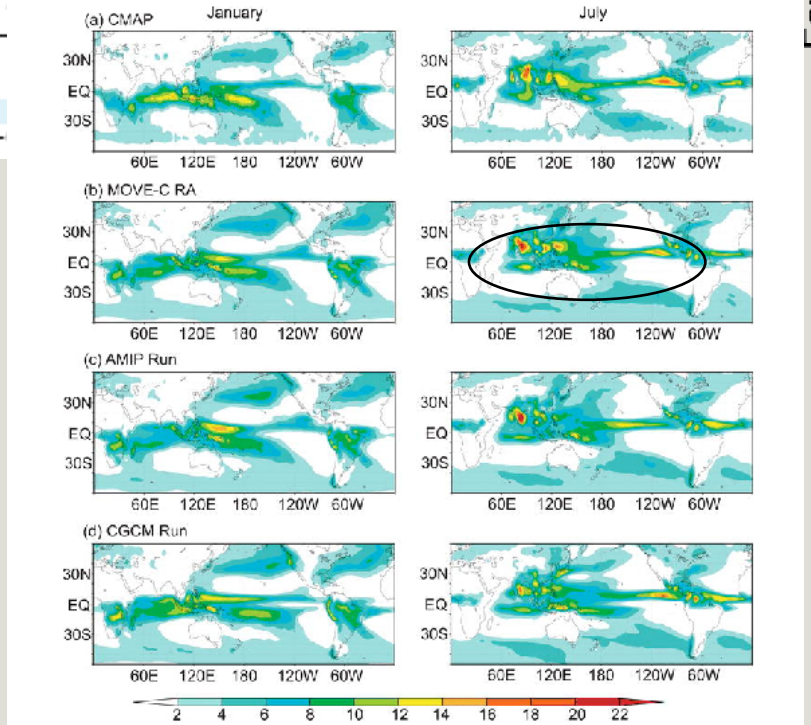
1. **Coupled 4D-var (Suguira et al 2008).** Both ocean and atmospheric observations (binned). 9 months assimilation cycle. Control vector: coupling coefficients and ocean initial conditions.
2. **Coupled EnKF (Zhang et al 2007).** Only ocean observations are used directly. Atmospheric information is nudged during the integration.
3. **Ocean Data Assimilation with a coupled model. (Fujii et al 2009)** Atmospheric model is free (AMIP). Spectral control of SST variability. Free at time scales < 1month.
4. **Coupled model+SST constraint** : Luo et al 2005, Kynleeside et al 2005

Examples of Coupled Data Assimilation: MEAN+ ANOMALY

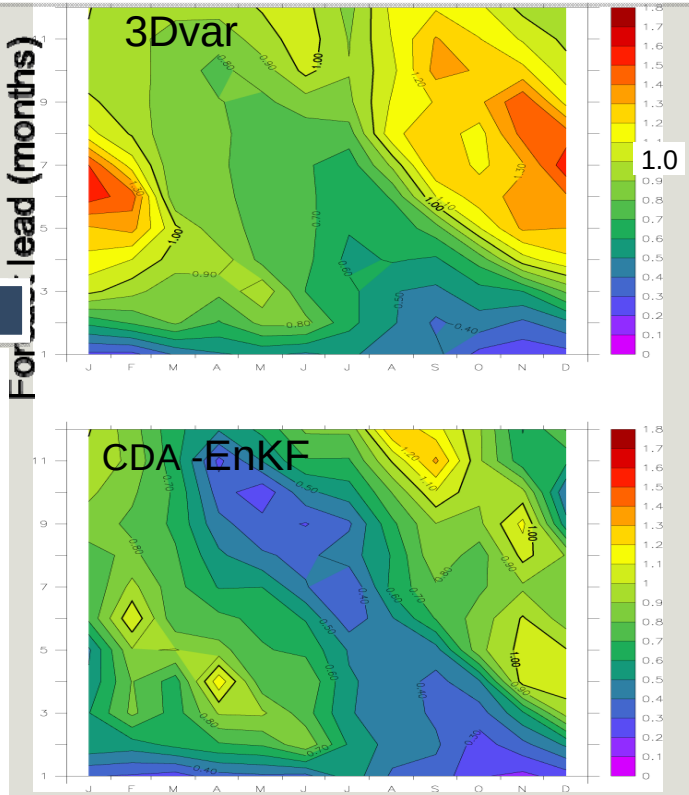
Coupled 4Dvar: log (α_E) – *From Sigiura et al. (2008)*



ODA+CGCM: Precip. *From Fujii etal 2009*



Coupled- EnKF: Forecast RMSE of Niño-3 SST
From Tony Rosati



Initial month

2) ANOMALY INITIALIZATION

•**Observational information is used to initialize only the anomalies, which are superimposed into the model climate.** It assumes quasi-linear regime.

•**Observational information is used either directly or from existing reanalysis.** Usually only the ocean component is initialized.

•**Background given by the coupled model**

•**To obtain observational anomalies an observational climatology is assumed.** In poor observed areas the time sampling for the climatology may be limited

Two flavours

1. **One-Tier anomaly initialization (Smith et al 2007).**

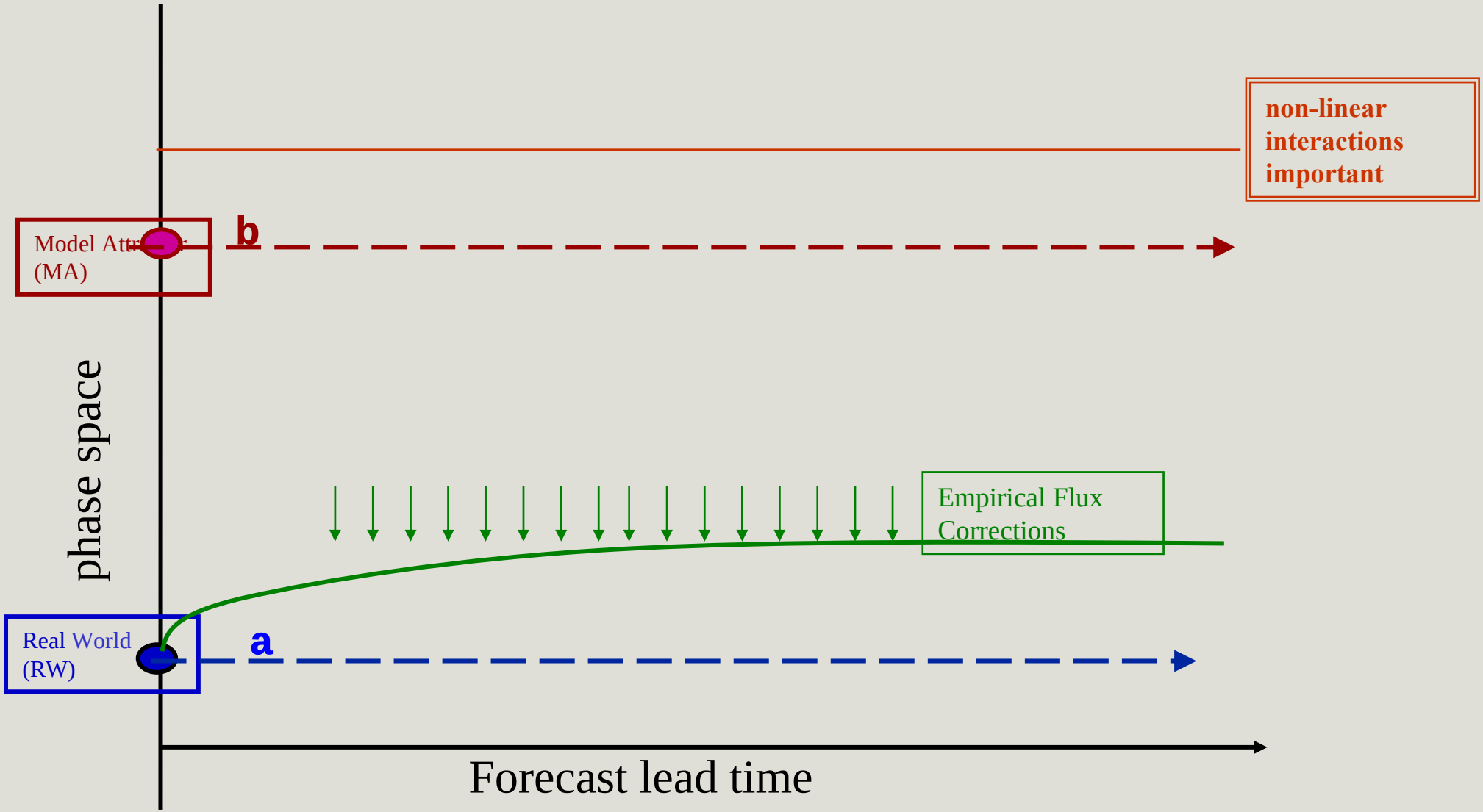
   Ocean observations are assimilatated directly. Background error covariance formulation derived from coupled model. Emphasis on large spatial scales

2. **Two-Tier anomaly initialization (Pohlmann et al 2009).**

   Nudging of anomalies from existing ocean re-analysis. The spatial structures are those provided by the source re-analysis.

   - Special case when only SST anomalies are used. (Keenlyside et al 2008)

Initialization Shock and non linearities



Statistical Corrections & Dynamical Models

- **Statistical corrections are at the heart of any data assimilation system.**
- **Statistical corrections are also routinely used in the a-posteriori calibration of forecasts**
- **There seem to be reluctance to use these corrections in the forward integration of coupled models.**
- **Statistical corrections are not always robust.**

Answer depends on Objective

1. **Delivering information**:
   - Historical reconstruction of climate
   - Forecast

Use all possible information and best available tools.

1. **Scientific Progress:**
   - Model Improvement
   - Process studies

Statistical corrections can be an obstacle

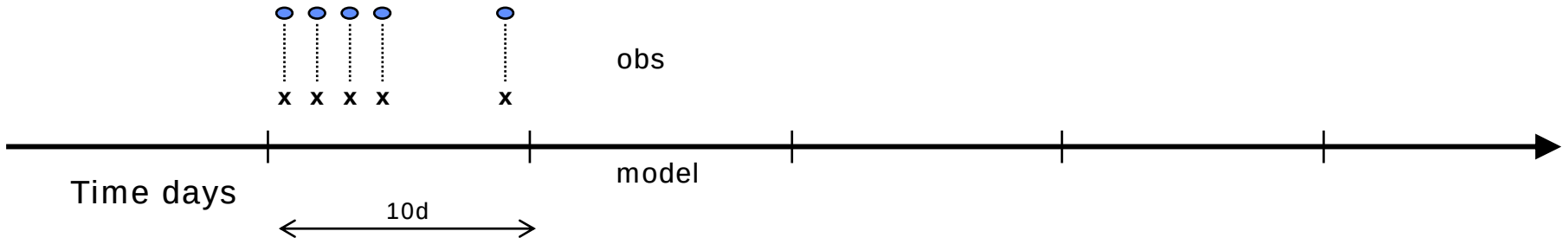
Summary

- Ocean reanalysis for historical reconstruction of the climate:
  - Uncertainties in important climate variables remain large
  - Difficult to validate.
  - More effort needed on:
    - Consolidation of assimilation methodology.
    - Improved atmospheric fluxes
    - Observation retrieval and QC.
    - Improved ocean models
- There are many myths and ambiguous terminology around the initialization problem.
  - Initialization shock does not imply loss of skill
  - Balanced initialization does not imply anomaly initialization
  - Anomaly initialization has problems when strong non-linearities
  - One-Tier and Two-Tier anomaly initialization can be quite different
- Statistical corrections to Dynamical models:
  - The answer is not universal and depends on the objective

Thanks for your attention

Assimilation Methods

- A variety of methods is currently used: OI, EnKF, En-OI, 3D-var, 4D-var (or adjoint methods)
- Different methods imply different assumptions and make different approximations.
- Example: length of assimilation cycle. 10 d or 40yr?



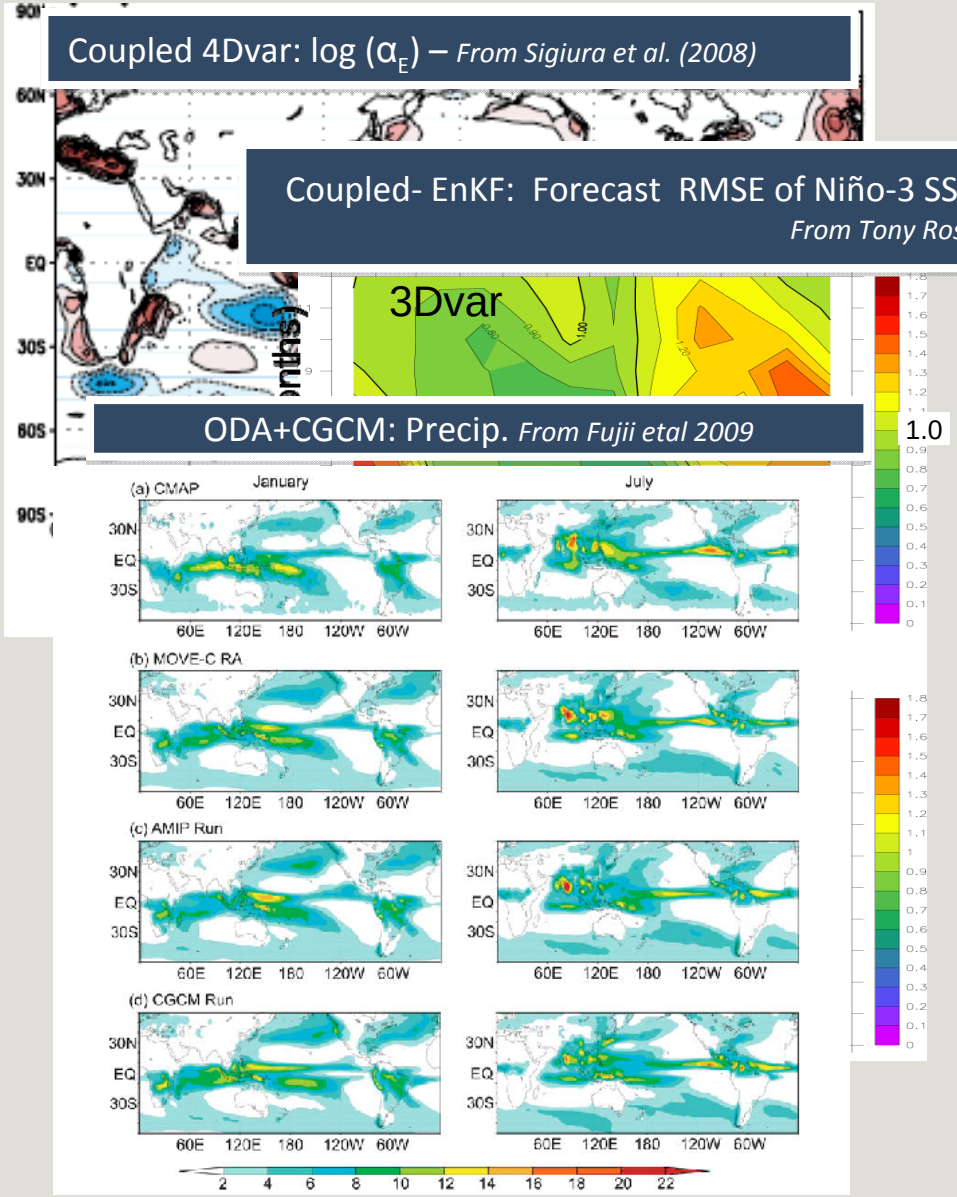
Ocean Reanalysis and Initialization of Forecasts

Even if ocean estate estimation is perfect (x(t)=y(t)).

It would not guarantee good initialization. For the forecast problem the time derivative $\dot{x}$ is also important.

Besides, models are imperfect, and model attractor is not the same as the real world attractor.

1) Coupled Data Assimilation: MEAN+ ANOMALY



Existing efforts

1. **Coupled 4D-var (Sigiura et al 2008).** Both ocean and atmospheric observations (binned). 9 months assimilation cycle. Control vector: coupling coefficients and ocean initial conditions.

2. **Coupled EnKF (Zhang et al 2007).** Only ocean observations are used directly. Atmospheric information is nudged during the integration.

3. **Ocean Data Assimilation with a coupled model. (Fujii et al 2009)** Atmospheric model is free (AMIP). Spectral control of SST variability. Free at time scales < 1month.

4. **Coupled model+SST constraint** : Luo et al 2005, Kynleeside et al 2005