Initialization Techniques in Seasonal Forecasting

Magdalena A. Balmaseda

Contributions from Linus Magnuson, Kristian Mogensen, Sarah Keeley

Outline

The importance of the ocean initial conditions in SF

Ocean Model initialization

The value of observational information: fluxes, SST, ocean observations

The difficulties

The traditional Full Initialization approach: pros and cons.

Other approaches. Assessment

Full Initialization, Anomaly Initialization

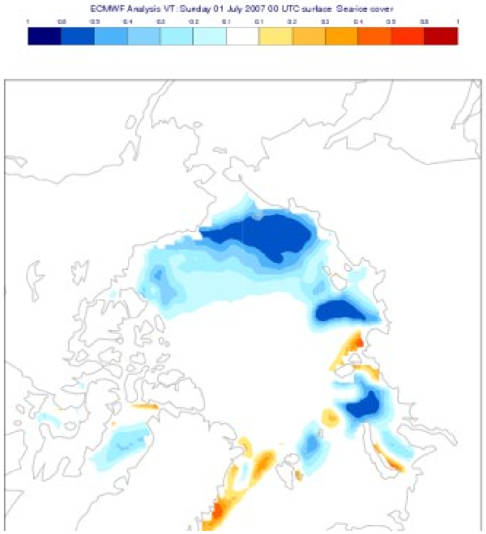
The basis for extended range forecasts

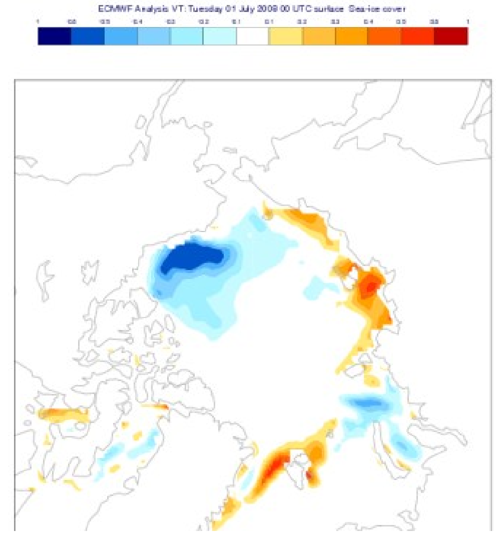
- Forcing by boundary conditions changes the atmospheric circulation, modifying the large scale patterns of temperature and rainfall, so that the probability of occurrence of certain events deviates significantly from climatology.
    - Important to bear in mind the probabilistic nature of SF

- The boundary conditions have longer memory, thus contributing to the predictability. Important boundary forcing:
    - **Tropical SST: ENSO, Indian Ocean Dipole, Atlantic SST**
    - Land: snow depth, soil moisture
    - Sea-Ice
    - Mid-Latitude SST
    - Atmospheric composition: green house gases, aerosols,…

Impact of Sea Ice:

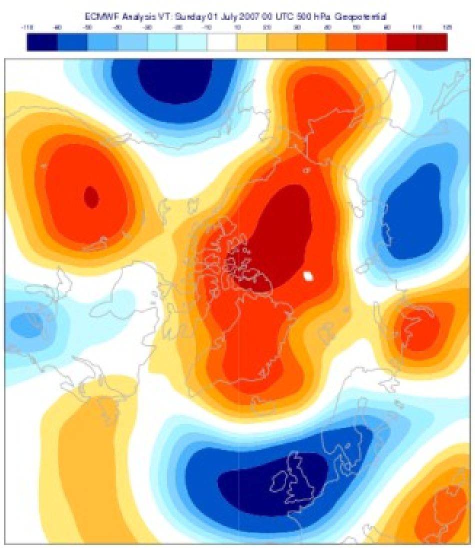
2007 **2008**

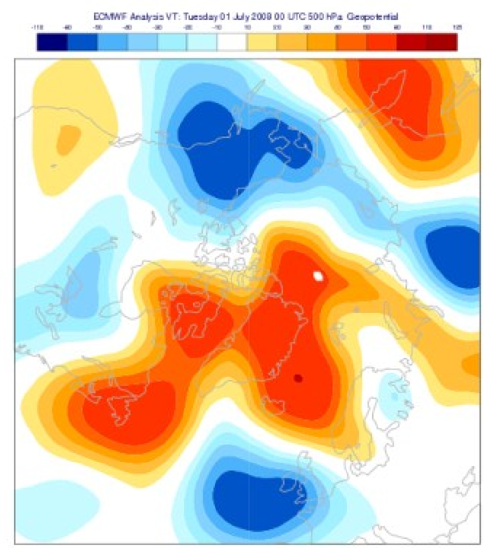




Ice Cover anomaly

CI .1





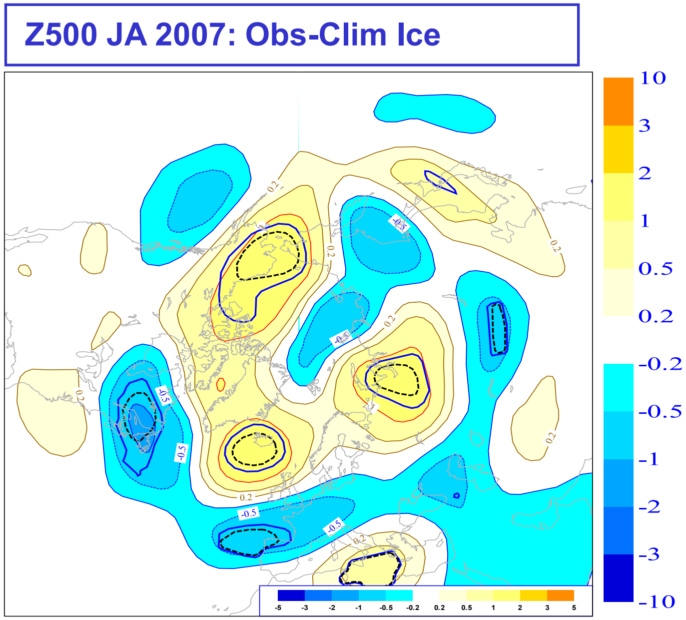
Observed Z500 anomaly

CI 1Dm

Courtesy of Sarah Keely

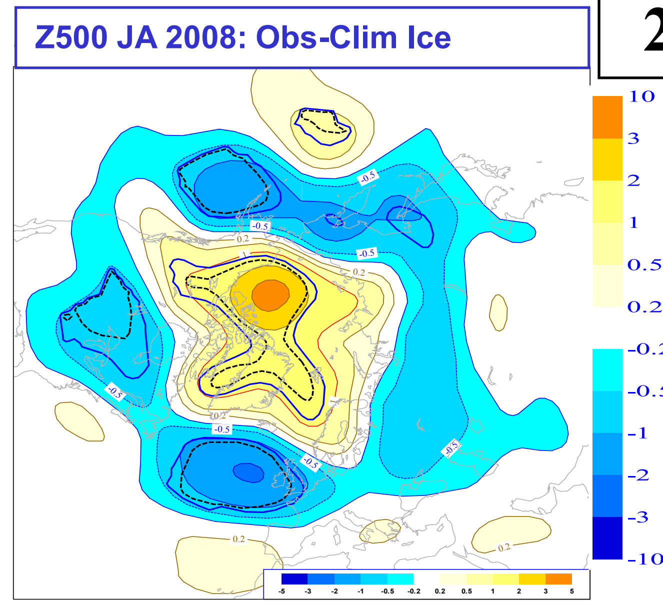
Impact on Z500:

2007



Atmos model

ObsIce-ClimIce



2008

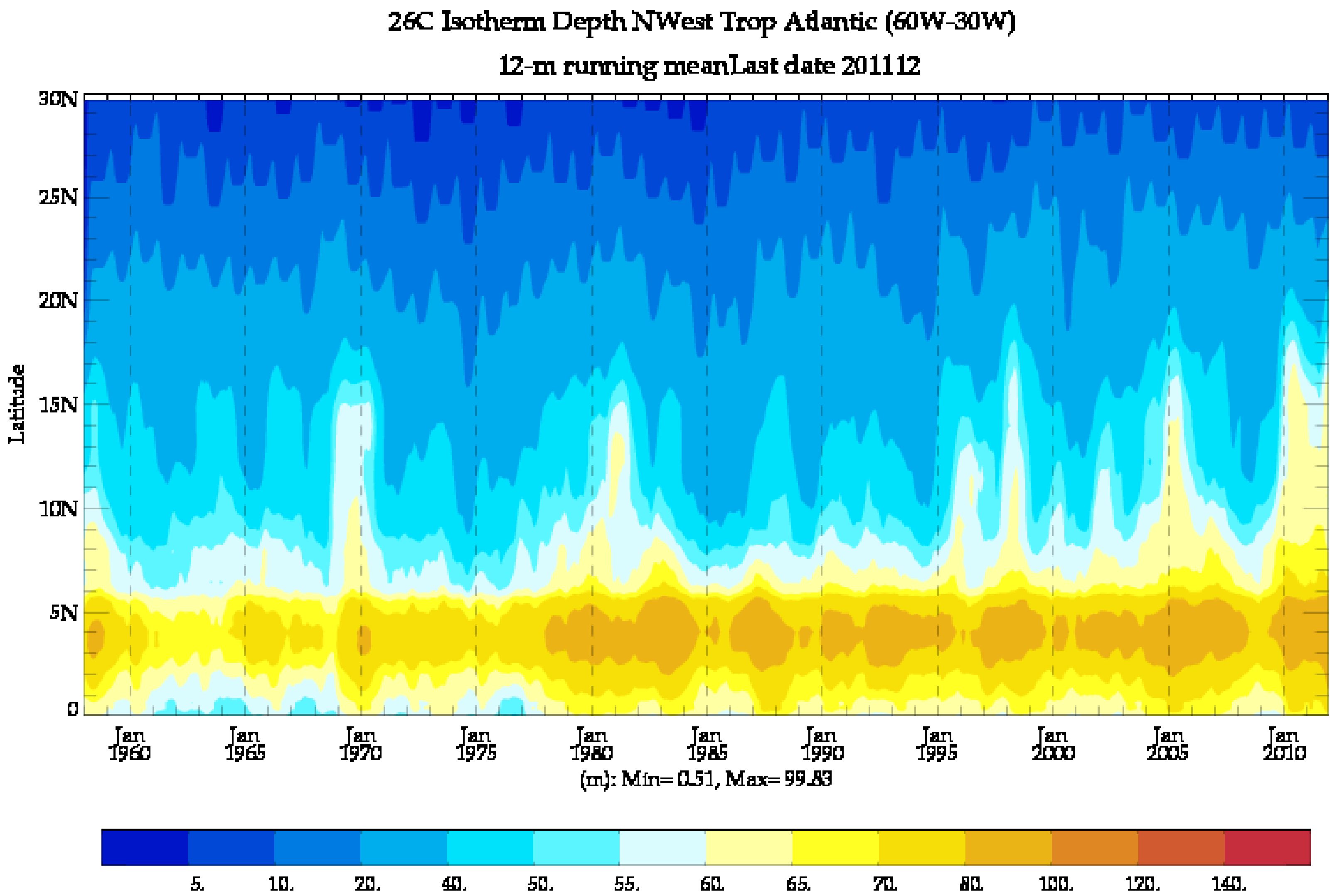
Low over Western Europe and Greenland high: similar response in both years. Consistent with observations

The response was conditioned by the SST, in particular the North Atlantic (Gulf Stream region), pointing towards the need of high resolution ocean models (or flux corrections).

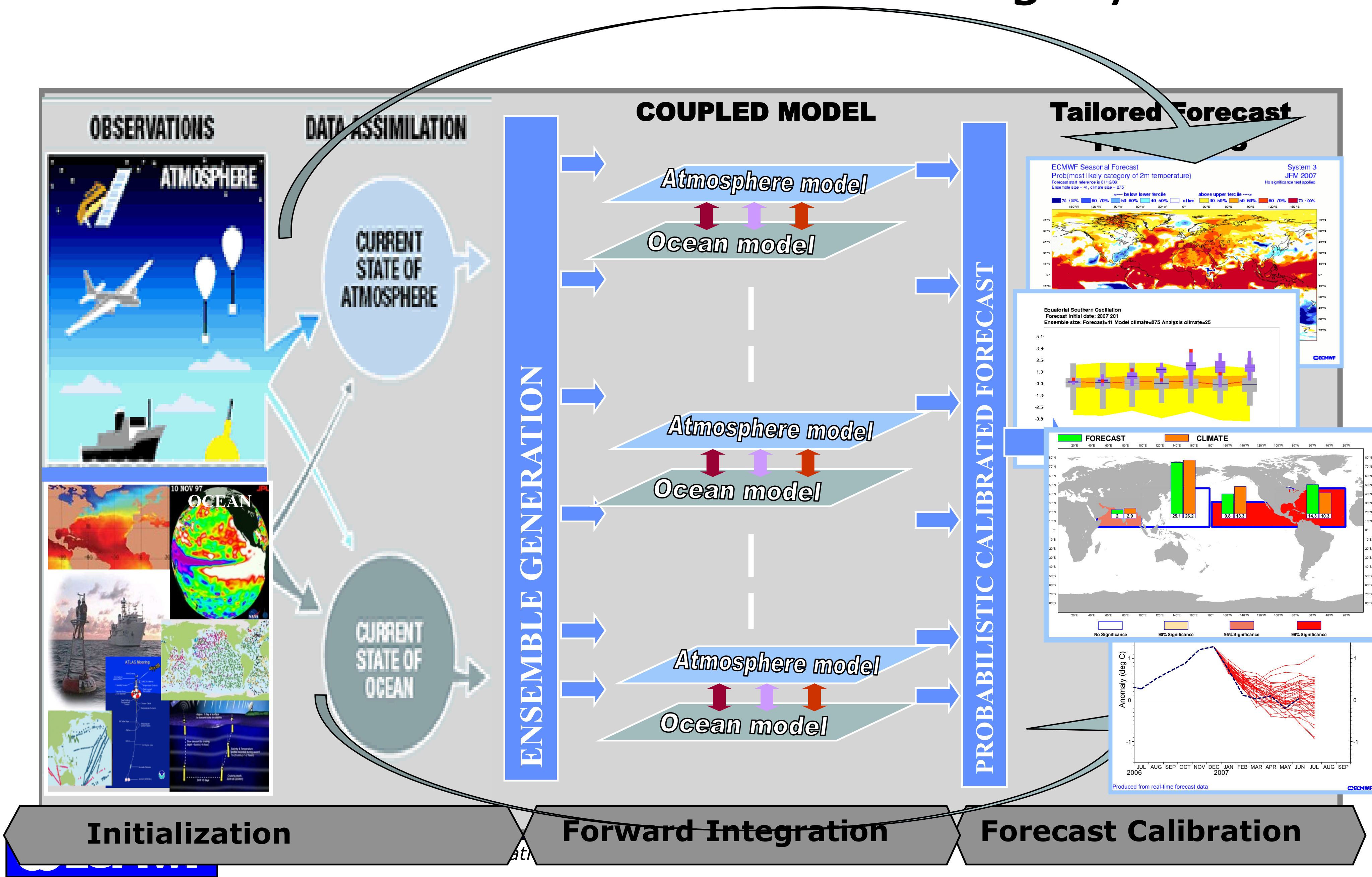
The question of the predictability of the sea-ice anomaly remains.

From Balmaseda et al 2010

Potential Energy for Tropical Cyclones

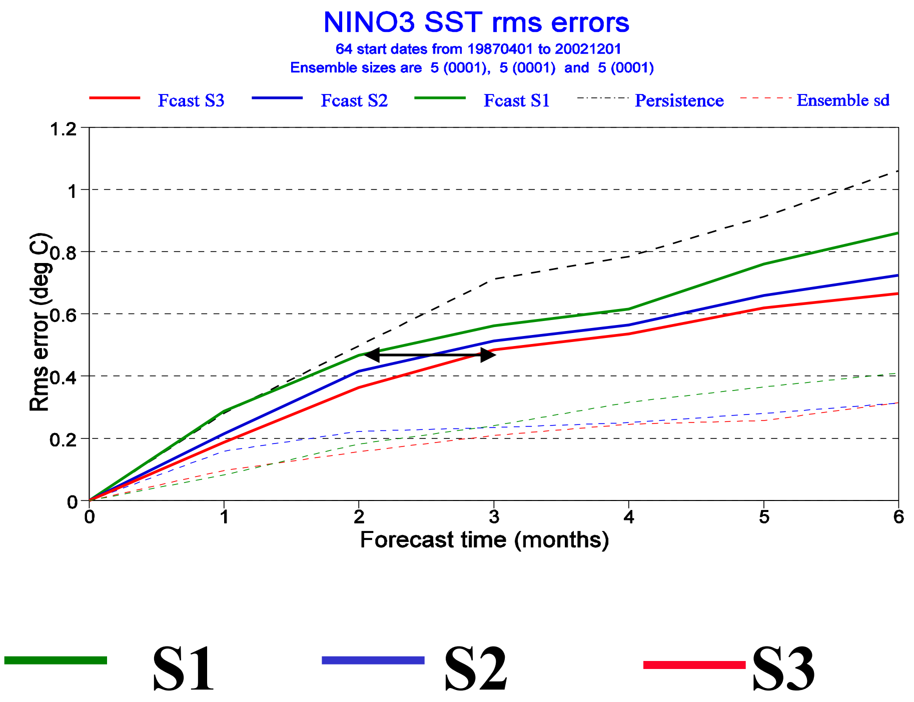


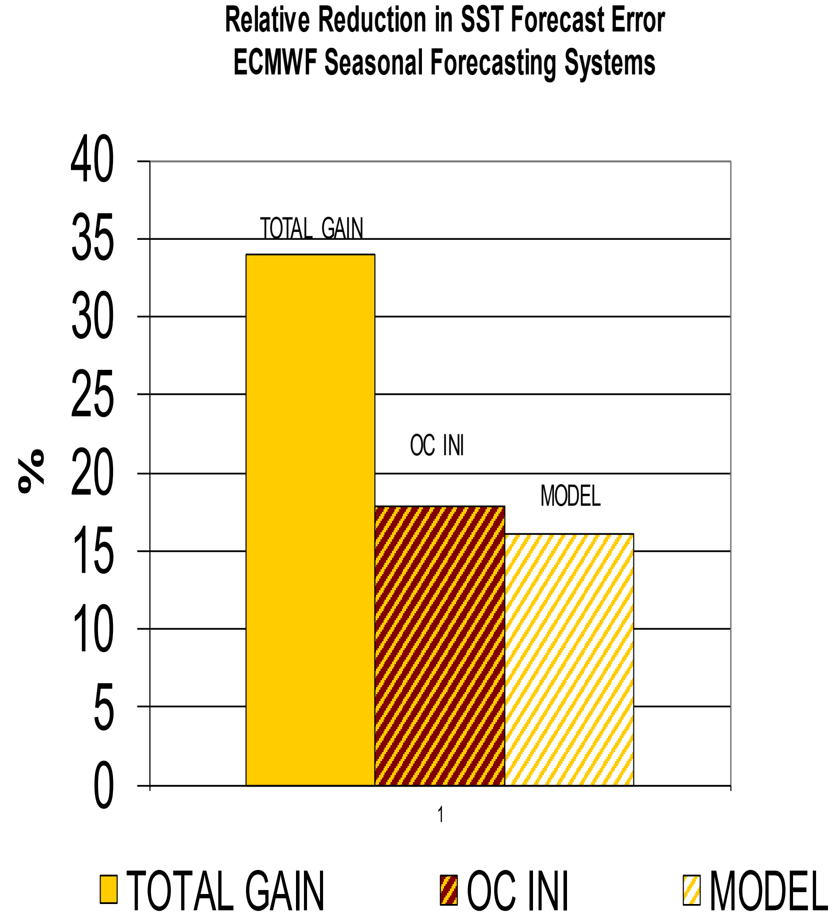
End-To-End Seasonal forecasting System



Initialization into Context

A decade of progress on ENSO prediction





- Steady progress: ~1 month/decade skill gain
- How much is due to the initialization, how much to model development?

Half of the gain on forecast skill is due to improved ocean initialization

Importance of Initialization

• *Atmospheric point of view: Boundary condition problem*

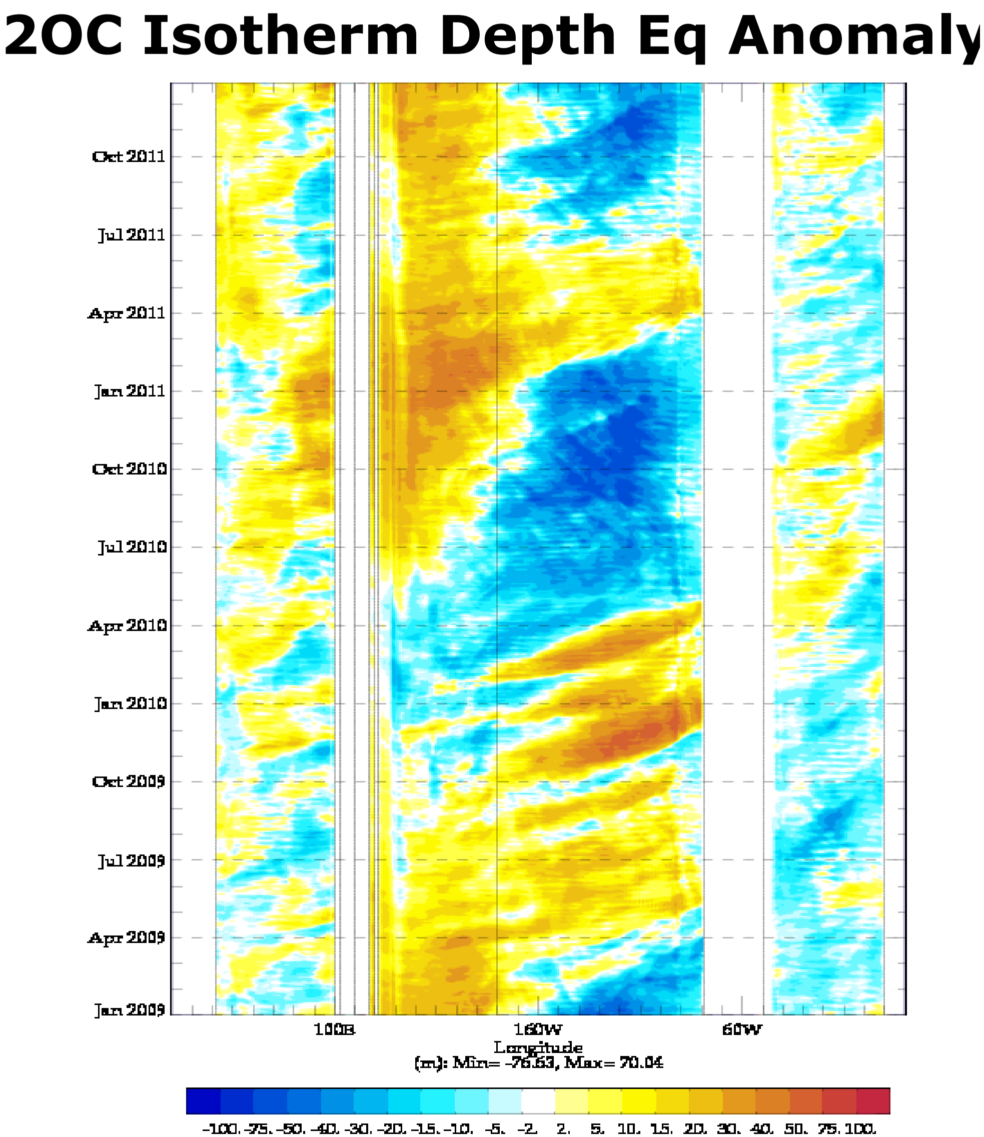
- Forcing by lower boundary conditions changes the PDF of the atmospheric attractor

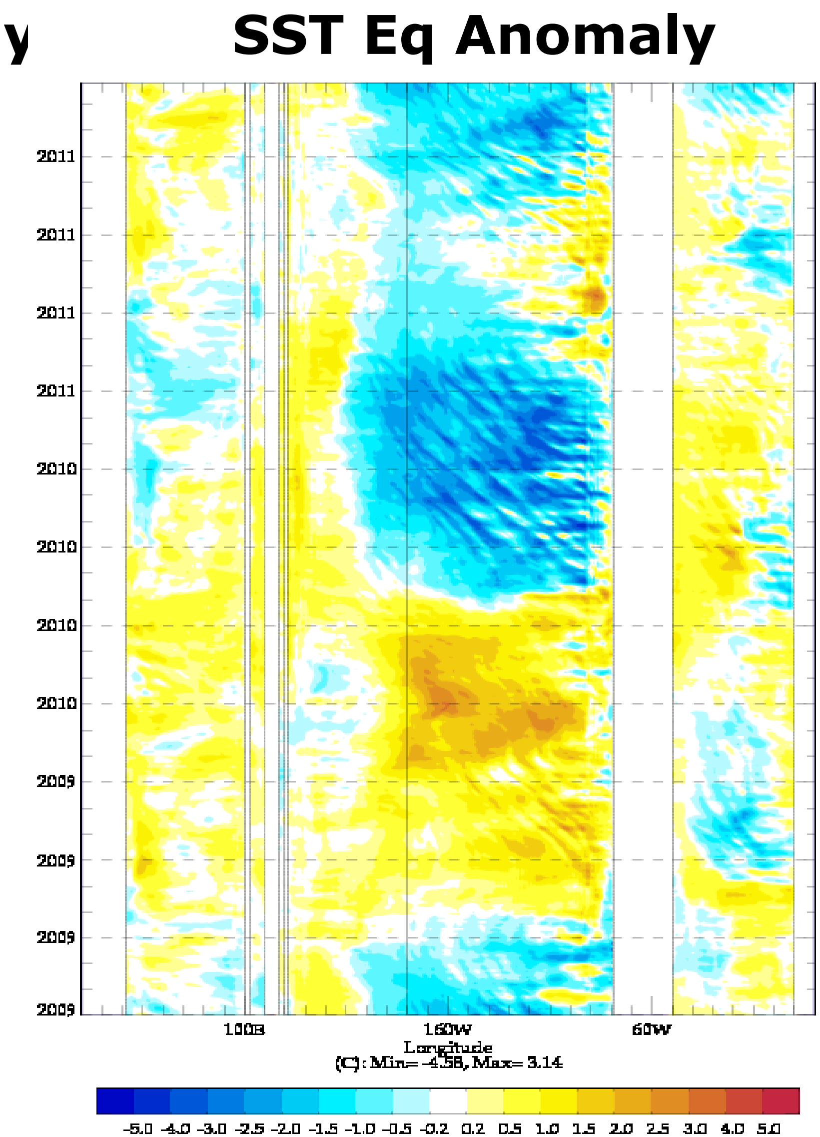
"Loaded dice"

• *Oceanic point of view: Initial value problem*

- *Prediction of tropical SST: need to initialize the ocean subsurface.*
  - *Emphasis on the thermal structure of the upper ocean*
  - *Predictability is due to higher heat capacity and predictable dynamics*

Need to Initialize the subsurface of the ocean





Initialization Problem: Production of Optimal I.C.

- **Optimal Initial Conditions: those that produce the best forecast.**

  Need of a metric: lead time, variable, region (i.e. subjective choice)
  Usually forecast of SST indices, lead time 1-6 months

- Theoretically, initial conditions should represent accurately the state of the real world and project into the model attractor, so the model is able to evolve them.

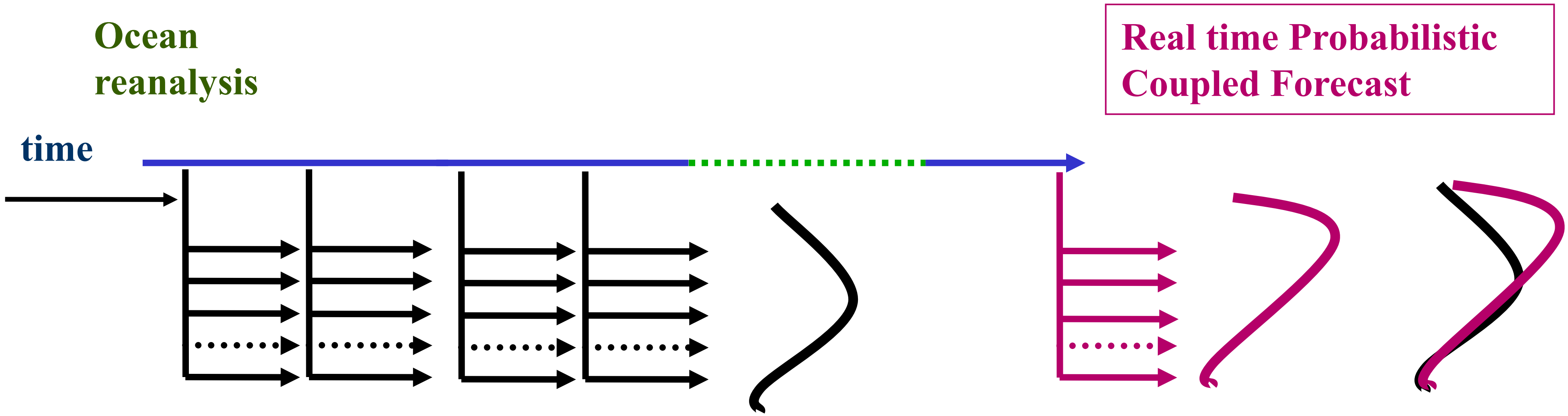
  Difficult in the presence of model error

- Practical requirements: Consistency between re-forecasts and real time fc

  Need for historical ocean reanalysis

- Current Priorities:
  - o Initialization of SST and ocean subsurface.
  - o Land/ice/snow

Dealing with model error: Hindcasts



Coupled Hindcasts, needed to estimate climatological PDF, require a historical ocean reanalysis

Consistency between historical and real-time initial conditions is required.

Hindcasts are also needed for skill estimation

Information to initialize the ocean

- Ocean model Plus:

  SST

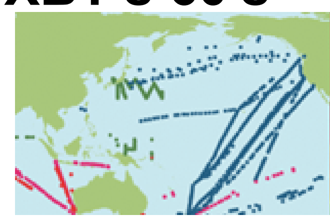
  Atmospheric fluxes from atmospheric reanalysis

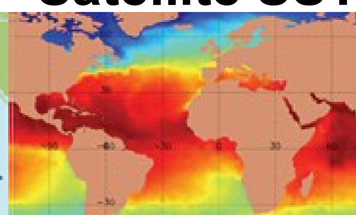
  Subsurface ocean information

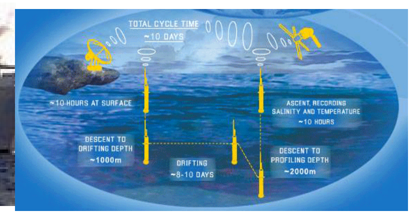
Time evolution of the Ocean Observing System

1982 **1993** **2001**

XBT's 60's **Satellite SST** **Moorings/Altimeter** **ARGO**







How do we initialize the ocean?

To a large extent, the large scale ocean variability is forced by the atmospheric surface fluxes.

Different ocean models forced by the same surface fluxes will produce similar tropical variability.
Daily fluxes of heat (short and long wave, latent, sensible heat), momentum and fresh water fluxes. **Wind stress is essential for the interannual variability**.

1. Constrained by SST: Fluxes from atmospheric models

   have large **systematic errors** and a **large unconstrained chaotic component**

2. Constrained by SST+ Atmos Observations: Surface fluxes from atmospheric reanalysis

   Reduced chaotic component. But still large errors/uncertainty

3. Constrained by SST+Atmos Observations+Ocean Observations: Ocean re-analysis

   Changing observing system and model error

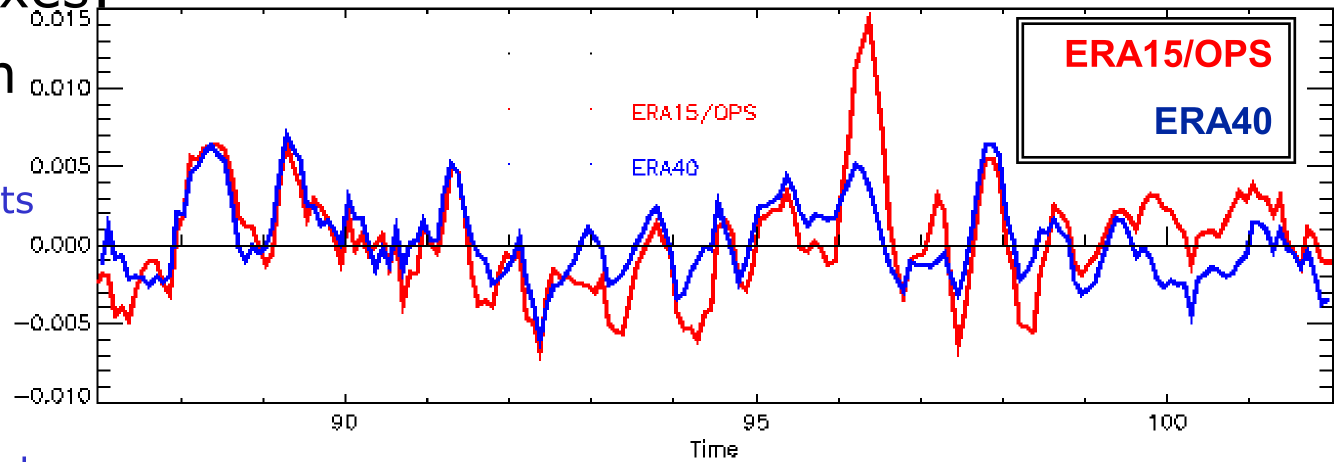
Uncertainty in Surface Fluxes: Need for Data Assimilation

- Large uncertainty in wind products lead to large uncertainty in the ocean subsurface
- The possibility is to use additional information from ocean data (temperature, others…)

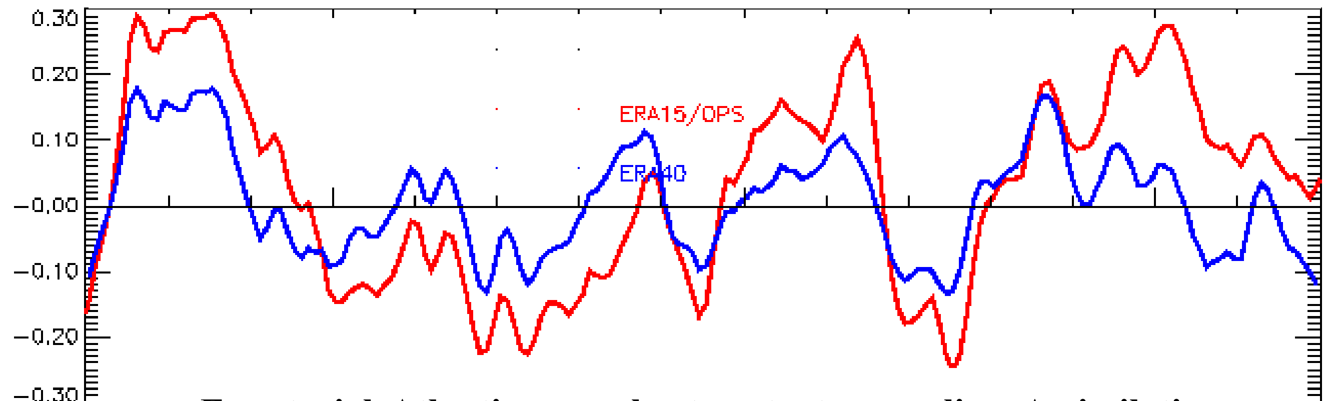
- Questions:

    1. Does assimilation of ocean data constrain the ocean state? **YES**
    2. Does the assimilation of ocean data improve the ocean estimate? **YES**
    3. Does the assimilation of ocean data improve the seasonal forecasts. **YES**

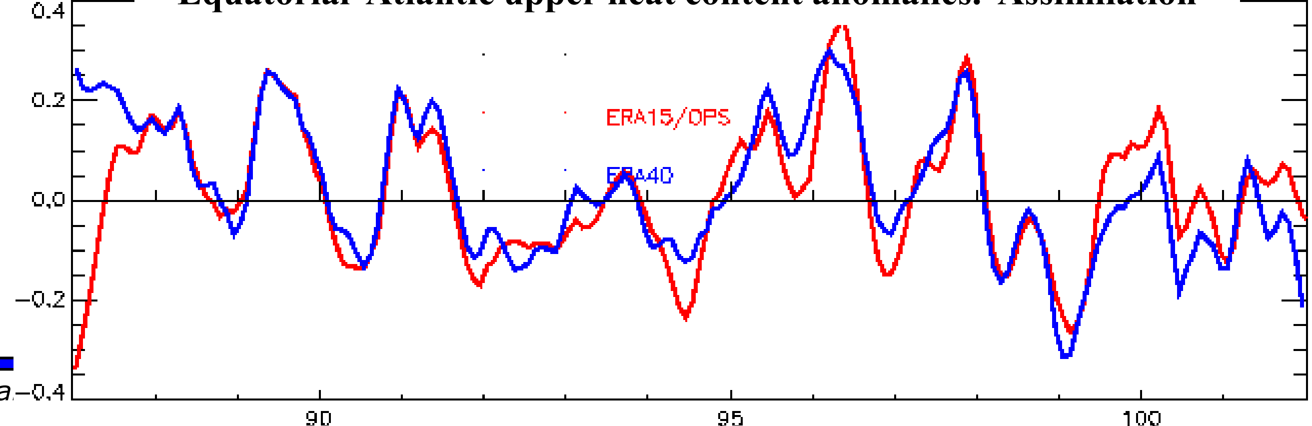
Equatorial Atlantic: Taux anomalies



Equatorial Atlantic upper heat content anomalies. No assimilation

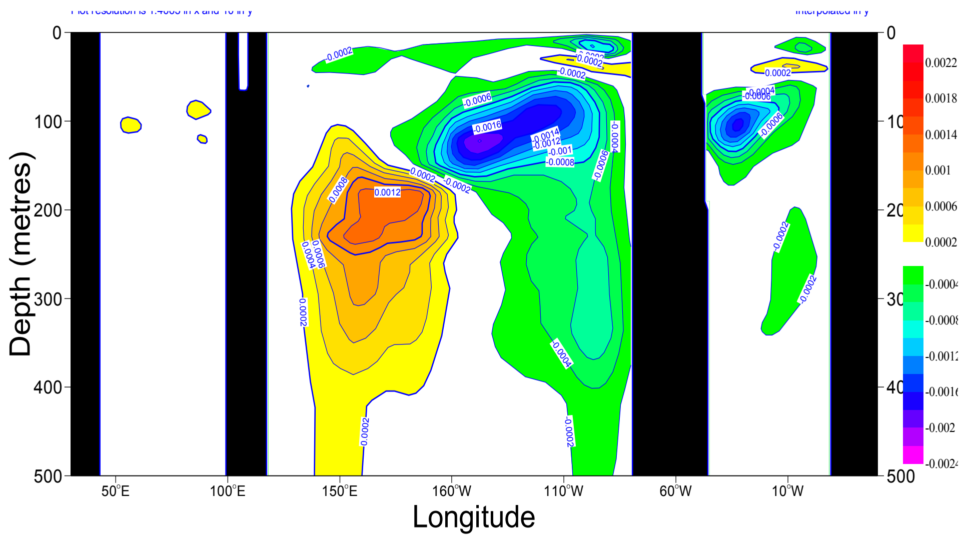


Equatorial Atlantic upper heat content anomalies. Assimilation

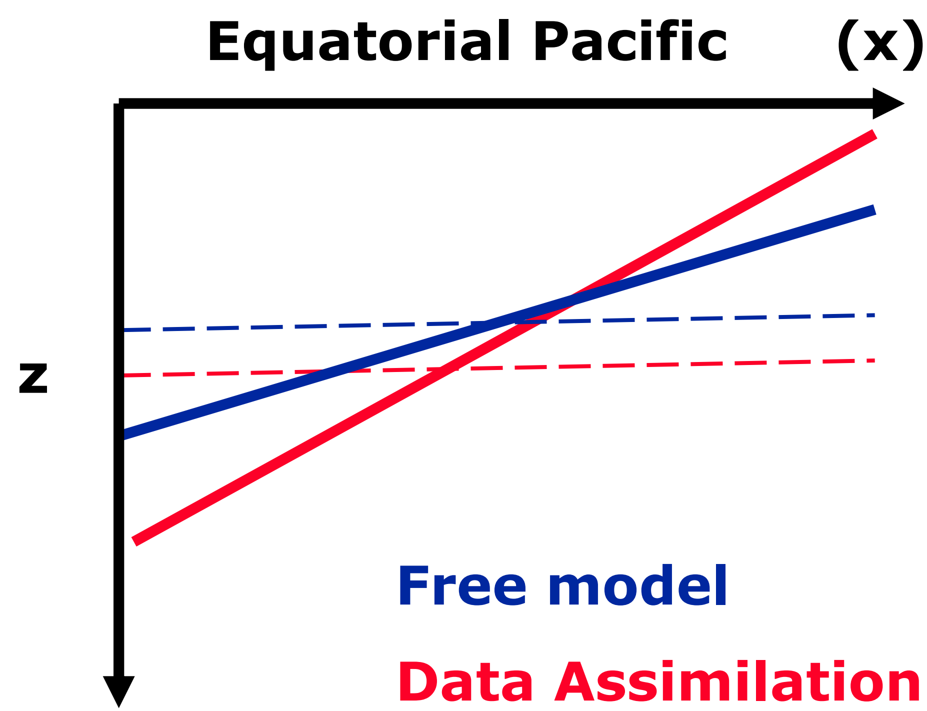


The Assimilation corrects the ocean mean state

Mean Assimation Temperature Increment

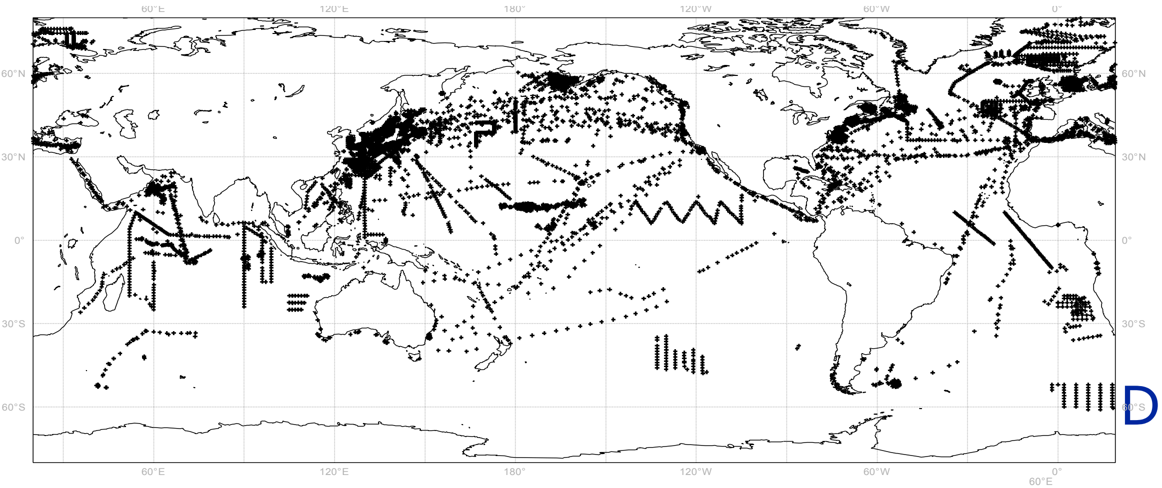


Data assimilation corrects the slope and mean depth of the equatorial thermocline



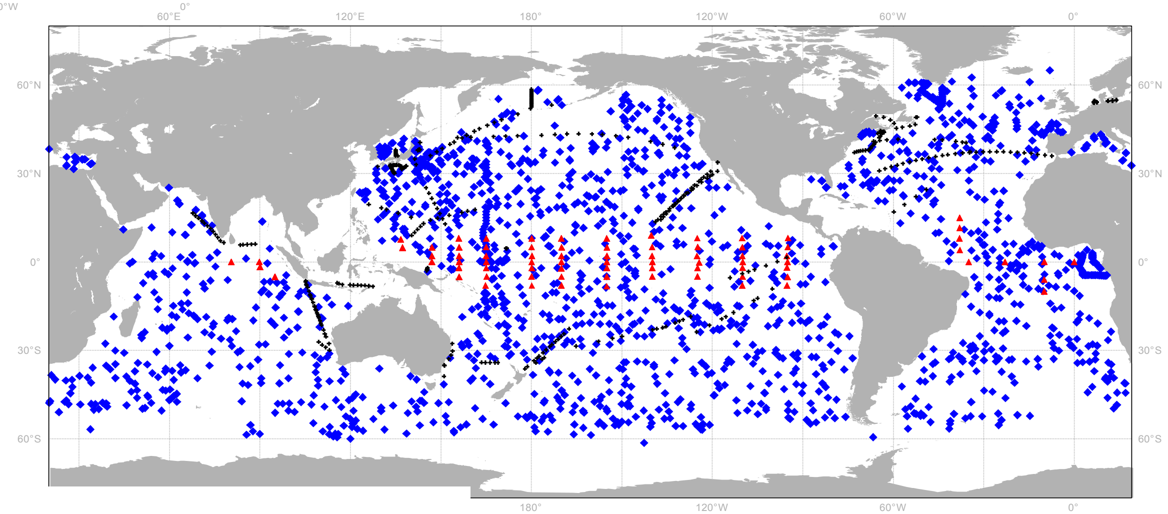
Ocean Observing System

Data coverage for June 1982



Changing observing system is a challenge for consistent reanalysis

Data coverage for Nov 2005



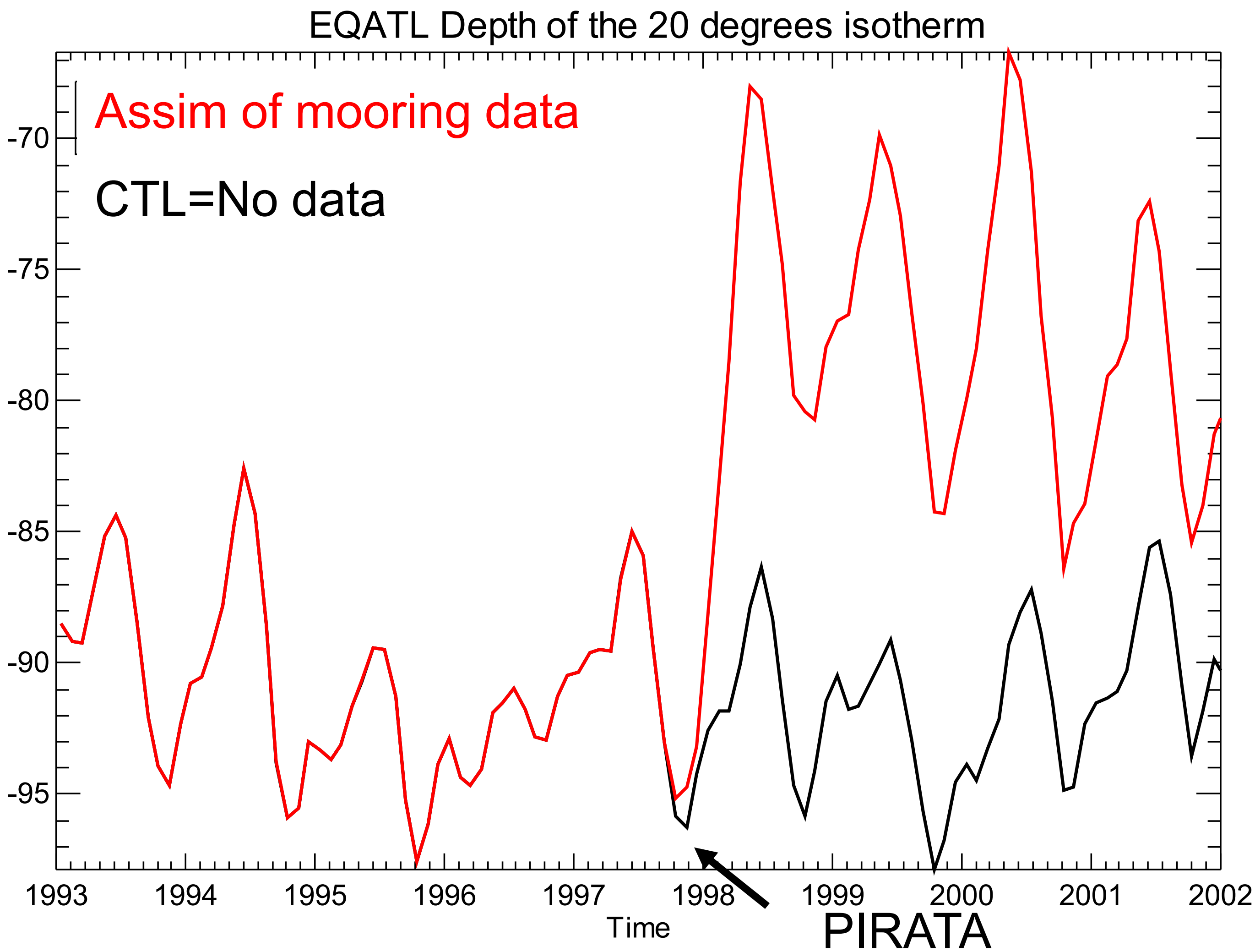
Today's Observations will be used in years to come

▲Moorings: SubsurfaceTemperature

◊ ARGO floats: Subsurface Temperature and Salinity

+ XBT : Subsurface Temperature

Impact of data assimilation on the mean



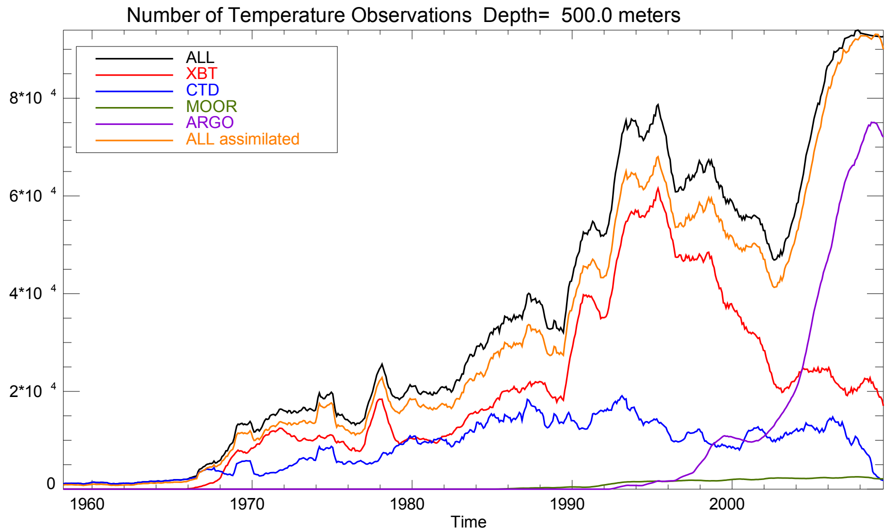
Large impact of data in the mean state leading to spurious variability

This is largely solved by the introduction of bias correction

Need to correct model bias during assimilation

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{b}^f + \mathbf{K}[\mathbf{y} - \mathbf{H}(\mathbf{x}^f + \mathbf{b}^f)]$$

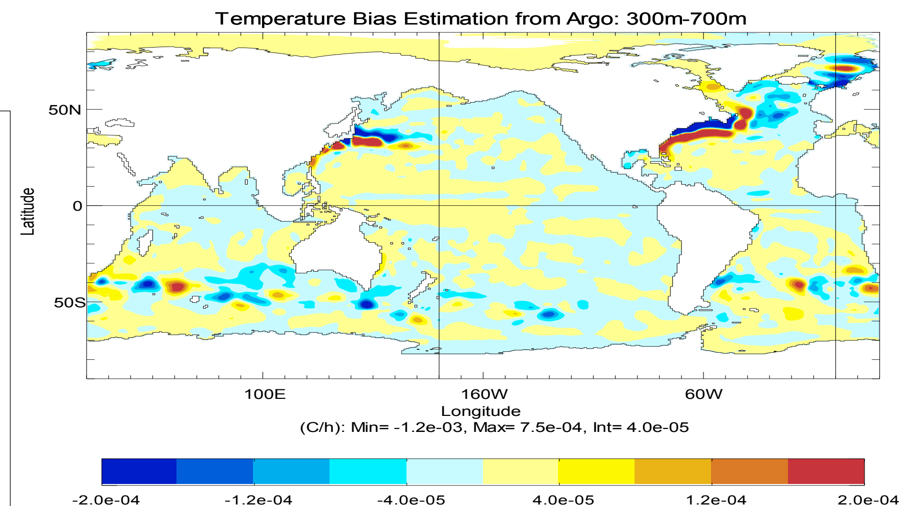
$$\mathbf{b}^a = \mathbf{b}^f + \mathbf{L}[\mathbf{y} - \mathbf{H}(\mathbf{x}^f + \mathbf{b}^f)]$$



There is a model for the time evolution of the bias

$$\mathbf{b}^f{}_k = \overline{\mathbf{b}}_k + \mathbf{b}'^f_k$$

This is an important difference with respect to the atmos data assimilation, where FG is assumed unbiased



Without explicit bias correction changes in the observing system can induce

- Spurious signals in the ocean reanalysis
- Non-stationarity of the forecast bias, leading to forecast errors.

Ideally, this information should be propagated during the forecast (for this the FG model and FC model should be the same, e.i. coupled model)

Ocean Initialization at the ECMWF

Ocean Reanalysis System 4 (ORAS4): 1958 to present. 5 ens members

Main Objective: Initialization of seasonal forecasts

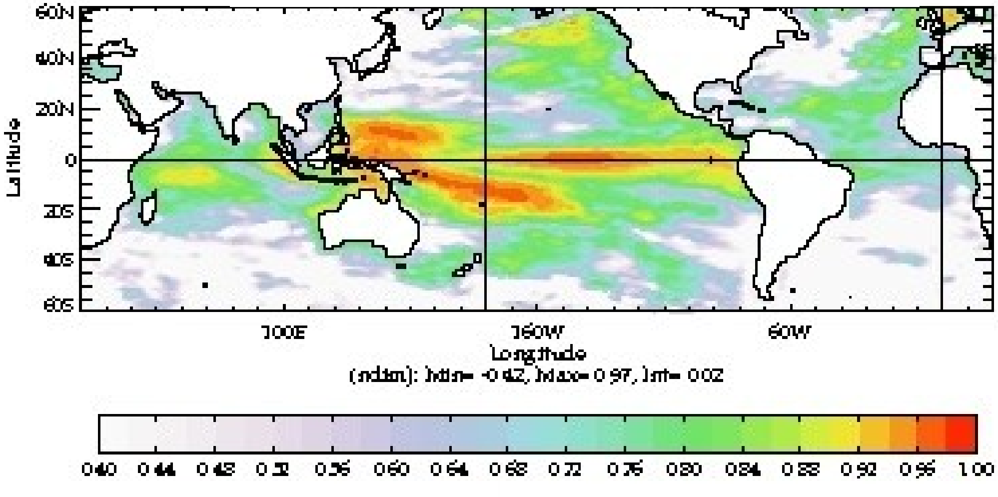
Operational ORA-S4 NEMO-NEMOVAR

- ERA-40 daily fluxes (1958-1989) and ERA-Interim thereafter
- Retrospective Ocean Reanalysis back to 1958, 5 ensemble members
- Multivariate offline+on-line Bias Correction (pressure gradient, Temp,Sal, offline from recent period)
- Assimilation of SST, temperature and salinity profiles, altimeter sea level anomalies an global sea level trends
- Balance constrains (T/S and geostrophy)
- Sequential, 10 days analysis cycle, 3D-Var FGAT. Incremental Analysis Update

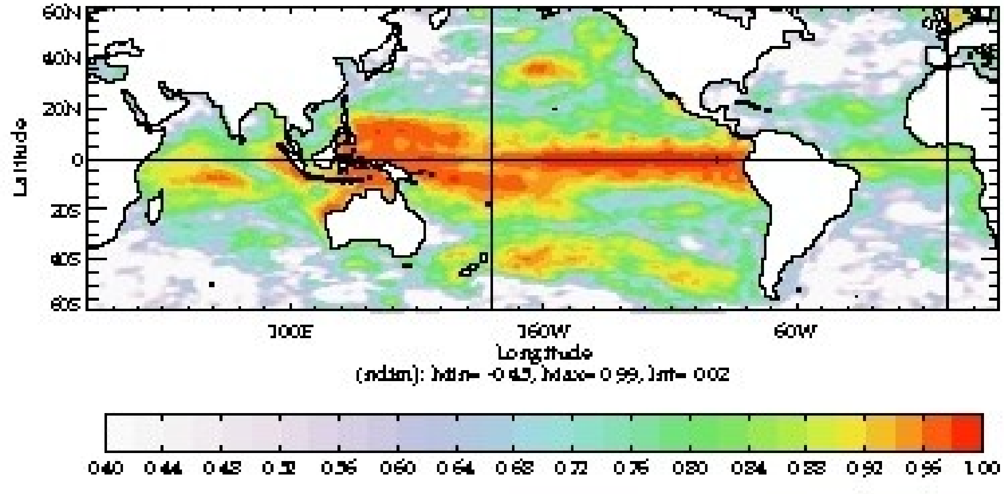
Mogensen et al 2012, Balmaseda et al 2012

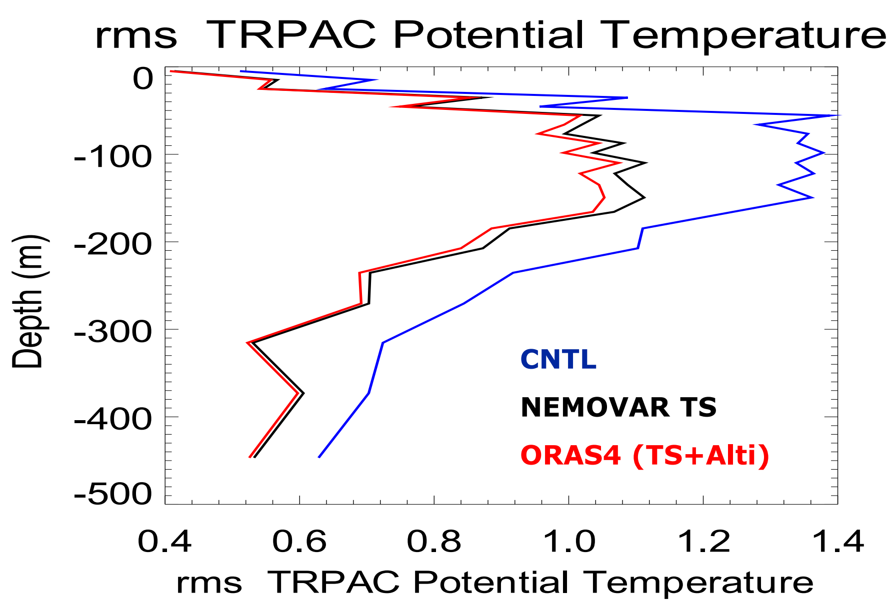
Time correlation with altimeter SL product

CNTL: NoObs

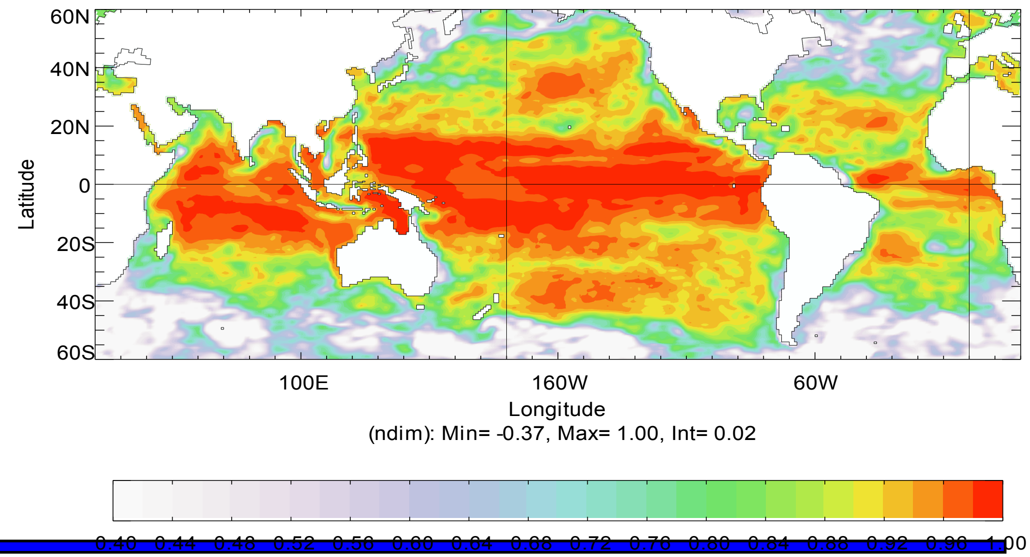


NEMOVAR T+S



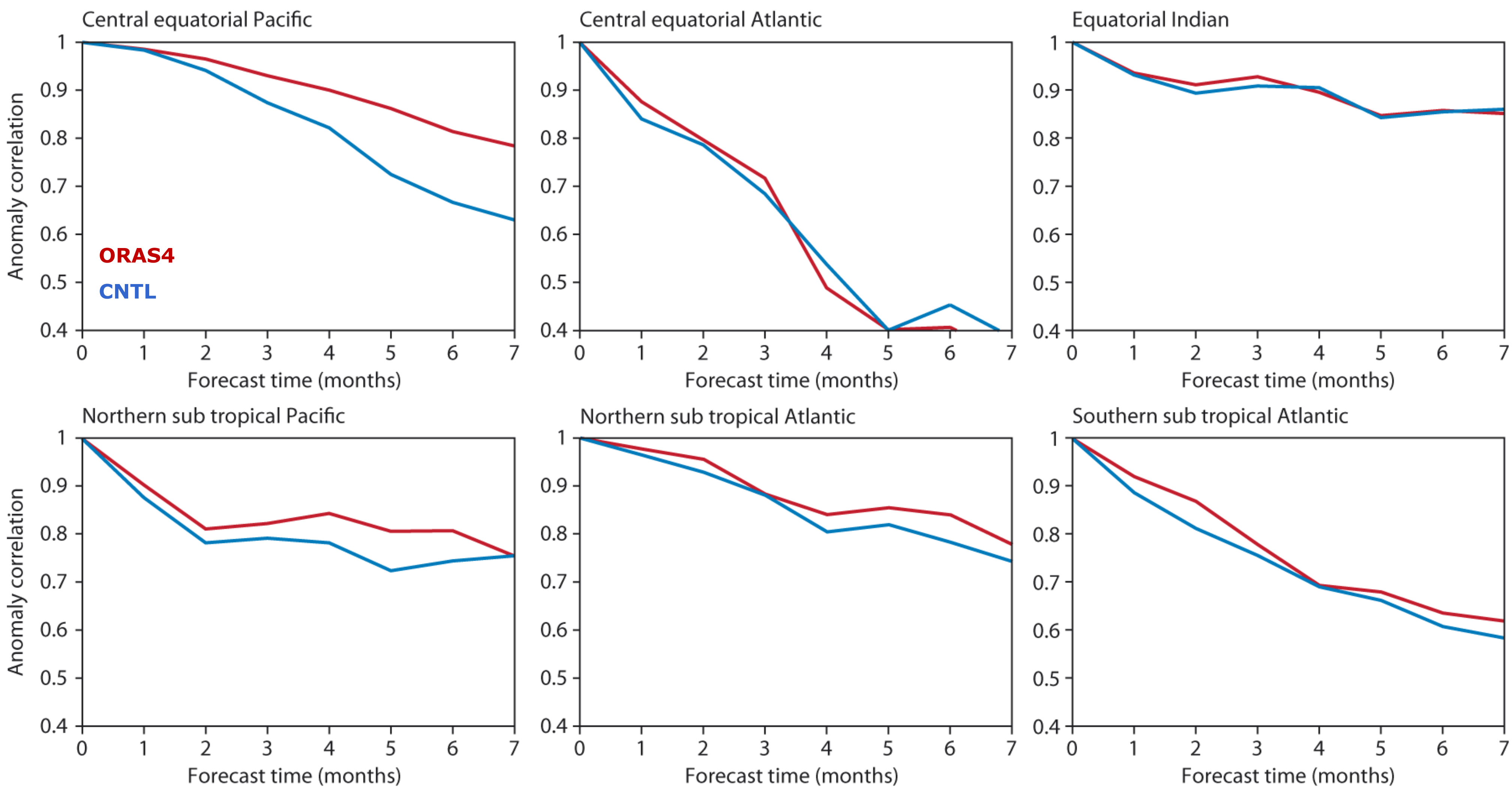


ORAS4 T+S+Alti



Impact on Seasonal Forecast Skill

Consistent Improvement everywhere. Even in the Atlantic, traditionally challenging area



Quantifying the value of observational information

- The outcome may depend on the coupled system
- In a good system information may be redundant, but not detrimental.

  If adding more information degrades the results, there is something wrong with the methodology (coupled/assim system)

- Experiments conducted with the ECMWF S3

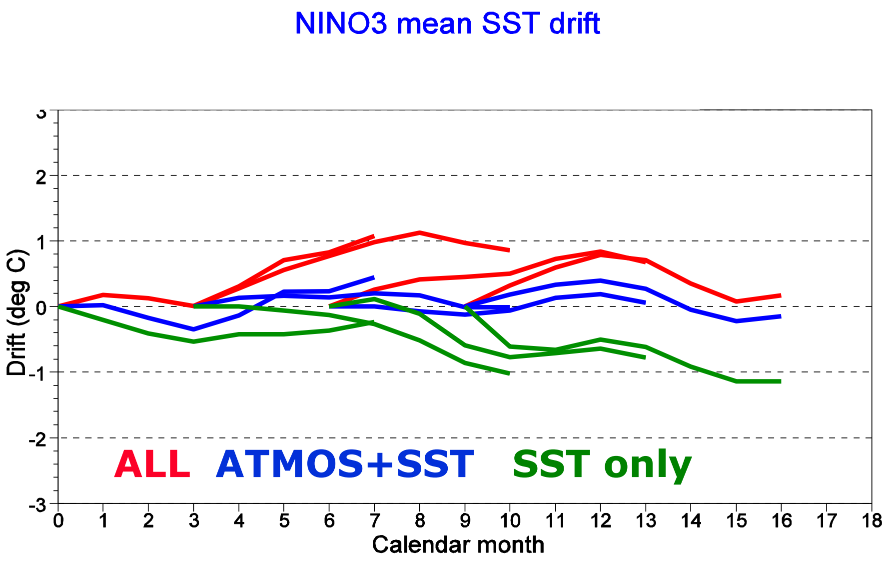
Balmaseda and Anderson 2009, GRL

SST (SYNTEX System Luo et al 2005, Decadal Forecasting Keenlyside et al, 2008)

SST+ Atmos observations (fluxes from atmos reanalysis)

SST+ Atmos observations+ Ocean Observations (ocean reanalysis)

Initialization and forecast drift



Different initializations produce different drift in the same coupled model.

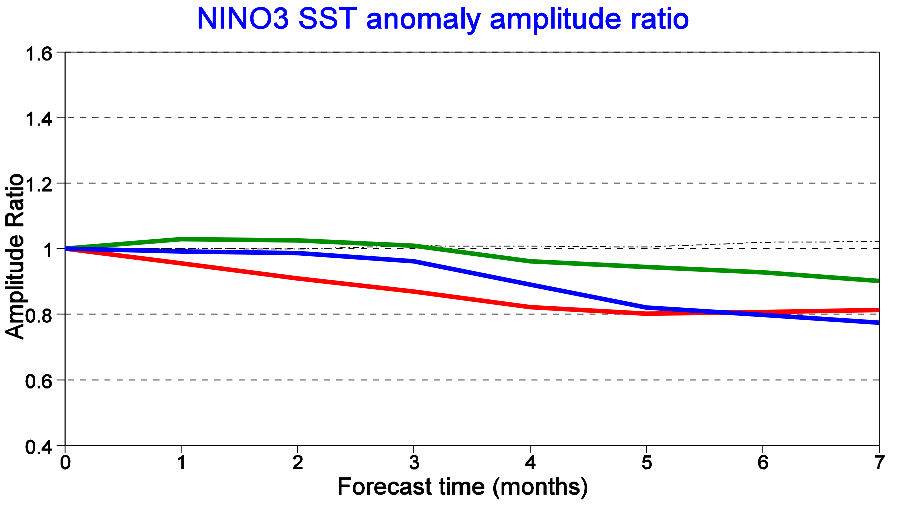
Warm drift in **ALL** caused by Kelvin Wave, triggered by the slackening of coupled model equatorial winds

SST only has very little equatorial heat content, and the SST cool s down very quickly.

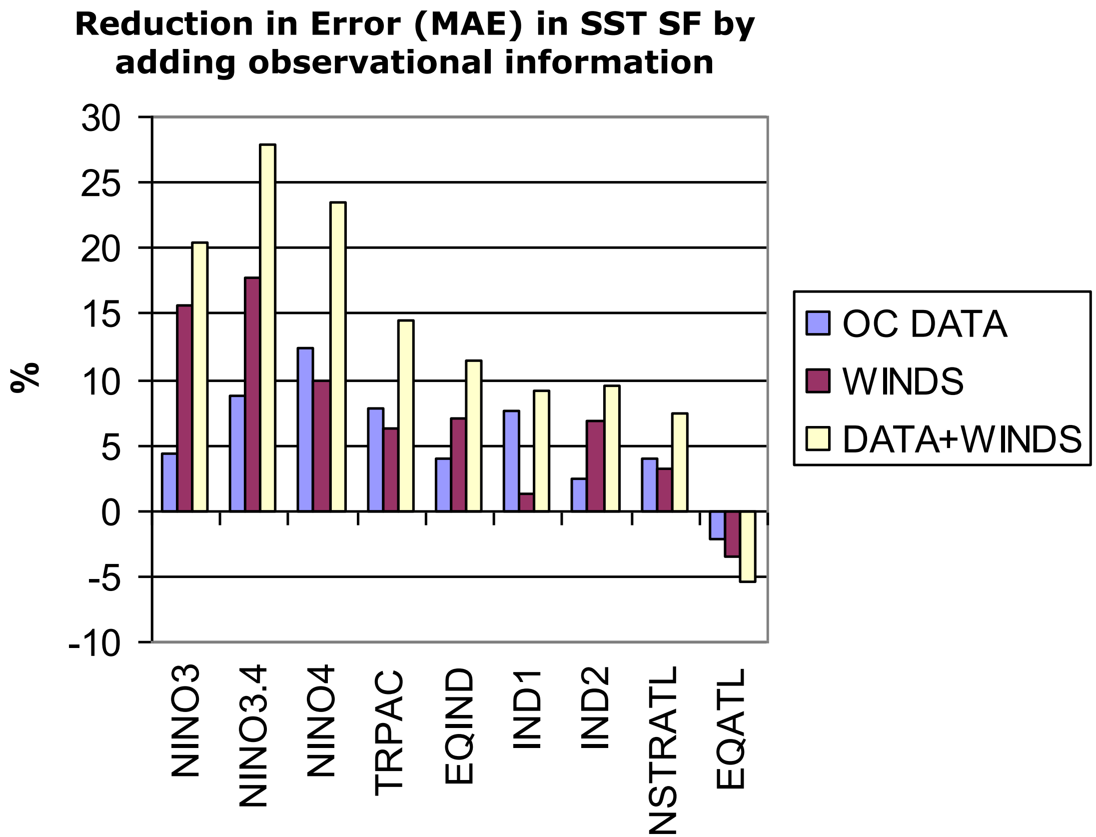
SST+ATMOS seems balanced in this region. Not in others

Sign of non linearity:

The drift in the mean affects the variability



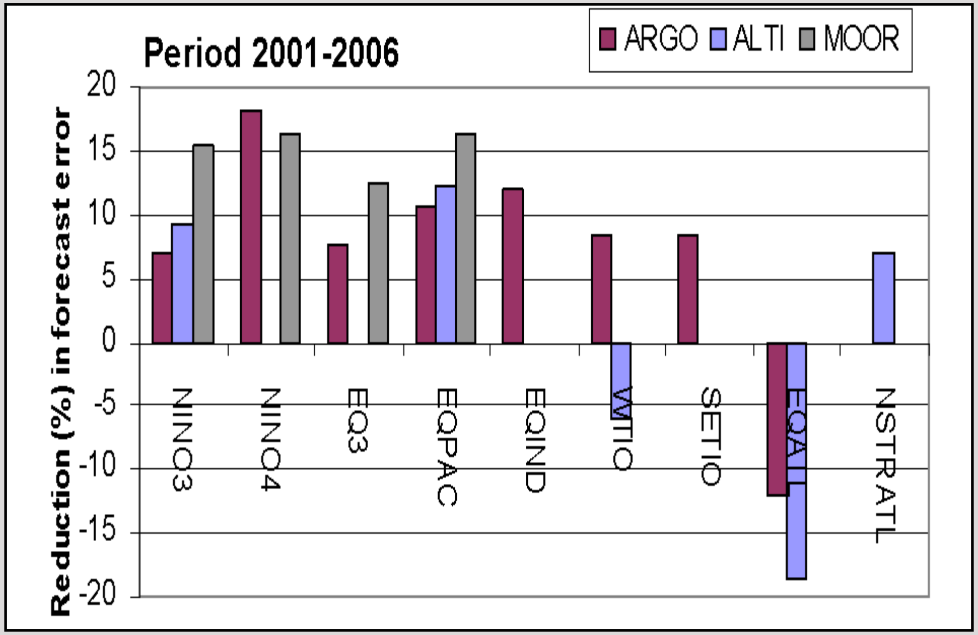
Impact of “real world” information on skill:



The additional information about the real world improved the forecast skill, except in the Equatorial Atlantic

Optimal use of the observations needs more sophisticated assimilation techniques and better models, to reduced initialization shock

Assessing the Ocean Observing System (S3)



From Balmaseda and Anderson 2009

See also Fujii et al 2008

1. **No observation system is redundant**

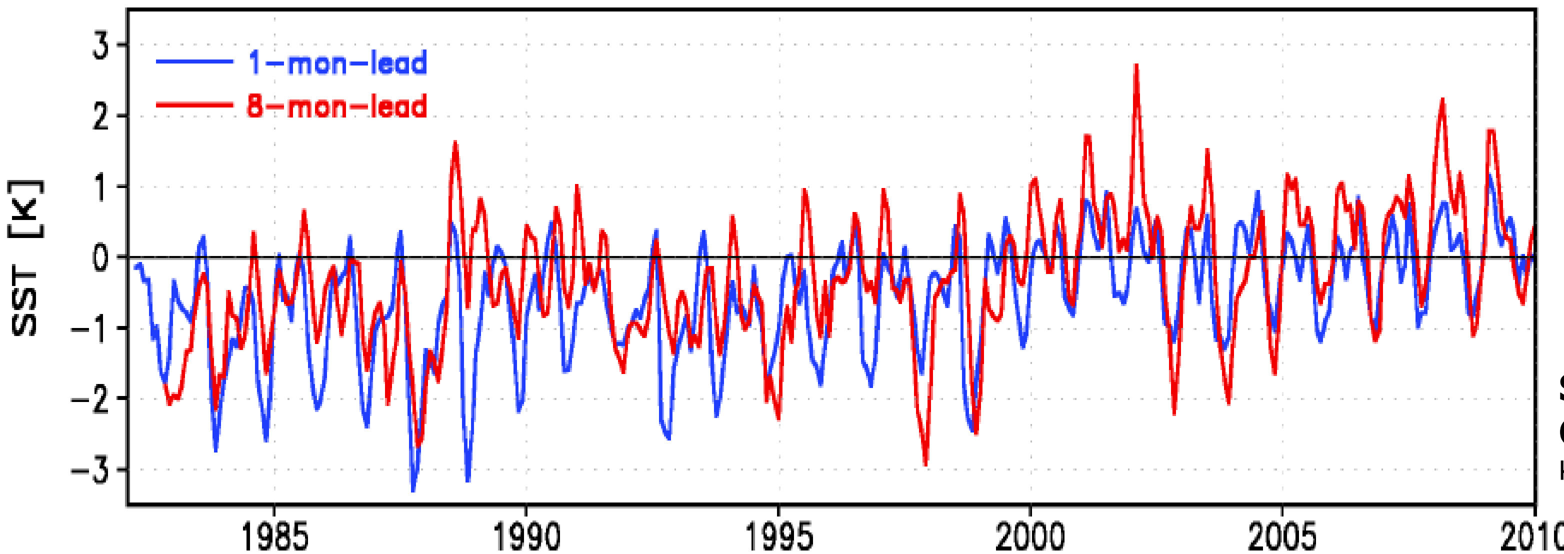
   Not even in the Pacific, where Argo, moorings and altimeter still complement. Lessons for other basins?

2. **There were obvious problems in the Eq Atlantic: model error, assimilation, and possibly insufficient observing system**

Important to bear in mind

1. The assessment depends on the quality of the coupled model
2. Need records long enough for results to be significant => any observing system needs to stay in place for a long time before any assessment is possible.

Seasonal Forecasts Approach
Some caveats



SST FC bias
CFS.v2
Kumar et al 2011 MWR

- Non-stationary model error. It depends on starting date. For example, seasonal cycle dependence, which is known. There are other unknown dependences
- Drift depends of lead time. A large number of hindcasts is needed. This is even more costly in decadal forecasts.
- Initialization shock can be larger than model bias

Non-linearities and non-stationarity can sometimes render the aposteriori calibration invalid

Perceived Paradigm for initialization of coupled forecasts

Real world **Model world**

Medium range

Being close to the real world is perceived as advantageous. Model retains information for these time scales.

Model attractor and real world are close?

Seasonal?

Decadal or longer

Need to initialize the model attractor on the relevant time and spatial scales.

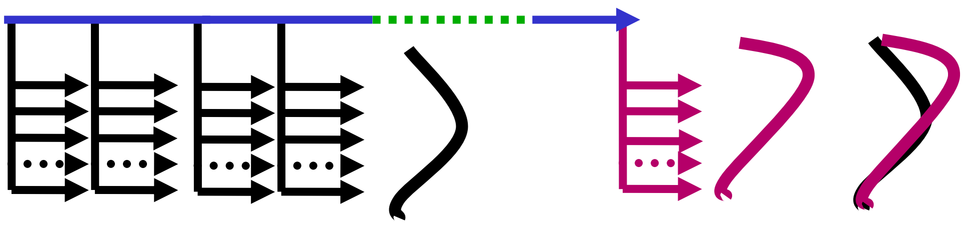
Model attractor different from real world.

At first sight, this paradigm would not allow a seamless prediction system.

- **Seasonal traditional approach as in Medium range BUT (see next slide)**

- **Not clear how to achieve initialization in model attractor**
  - Anomaly Initialization (decadal forecasts, Smith et al 2007)
  - Full initialization with coupled models of the slow component only
  - Other more sophisticated (EnKF, coupled DA, weakly coupled DA)

Seasonal Approach



As Medium range but:
Model bias taken into account during DA.

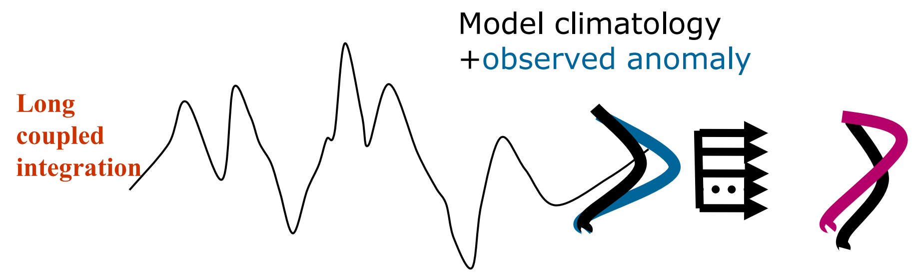
A posteriori calibration of forecast is needed. Calibration depends on lead time.

The model for first guess during the initialization is different from the forecast model. Bias correction estimated during initializaiton can not be applied during the forecasts

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{b}^f + \mathbf{K}[\mathbf{y} - \mathbf{H}(\mathbf{x}^f + \mathbf{b}^f)]$$

$$\mathbf{b}^a = \mathbf{b}^f + \mathbf{L}[\mathbf{y} - \mathbf{H}(\mathbf{x}^f + \mathbf{b}^f)]$$

Anomaly Initialization



The model climatology does not depend of forecast lead time. Cheaper in principle.

Hindcasts are still needed for skill estimation

Acknowledgment of existence of model error during initialization.

Model error is not corrected ("bias blind algorithm"):

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}[(\mathbf{y} - \overline{\mathbf{y}}) - \mathbf{H}(\mathbf{x}^f - \overline{\mathbf{x}})]$$

Anomaly Initialization (Cont)

Two flavours

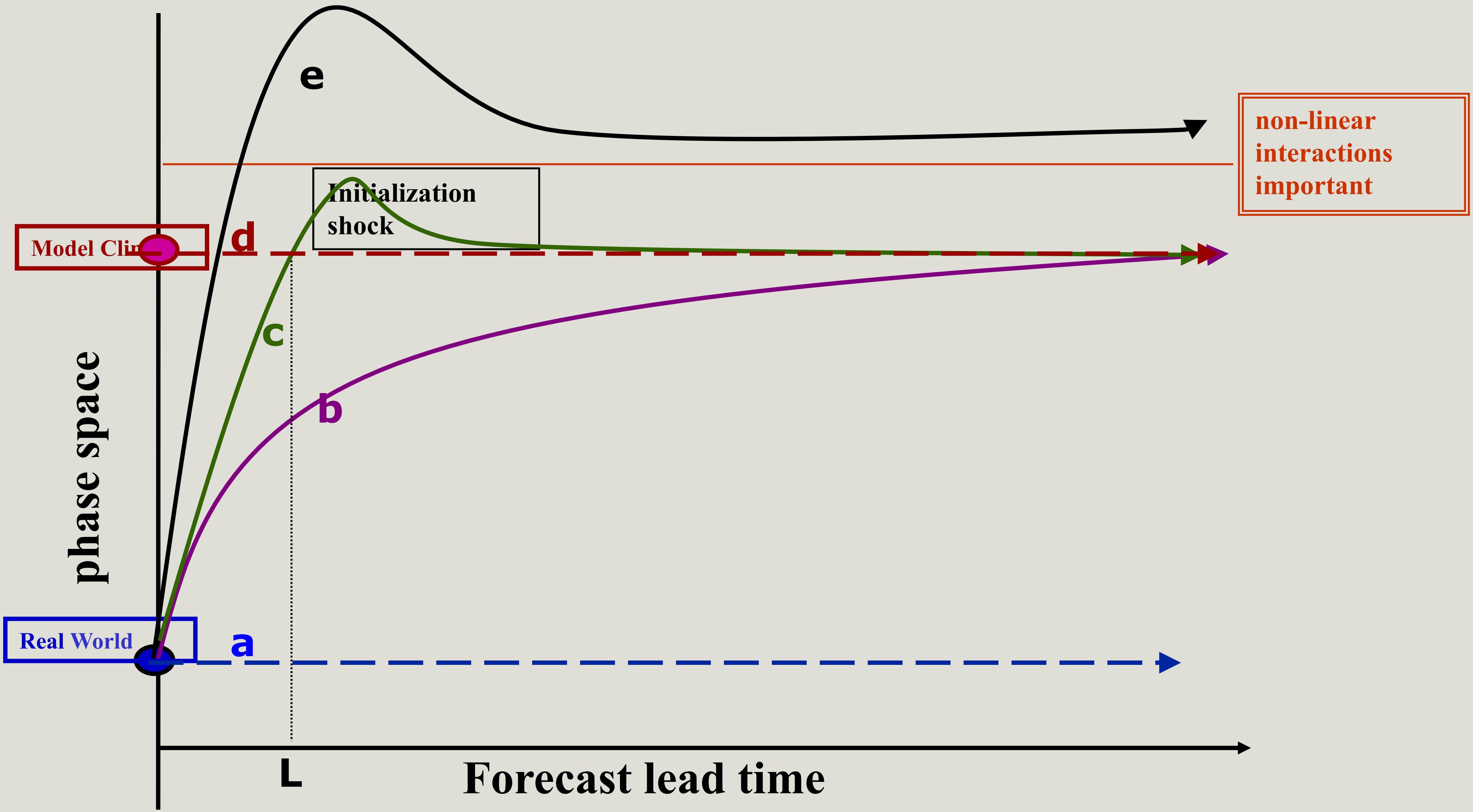
1. **One-Tier anomaly initialization (Smith et al 2007). Ocean observations are assimilated directly. Background error covariance formulation derived from coupled model (EOFs, EnOI, EnKF). Emphasis on large spatial scales**

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.**

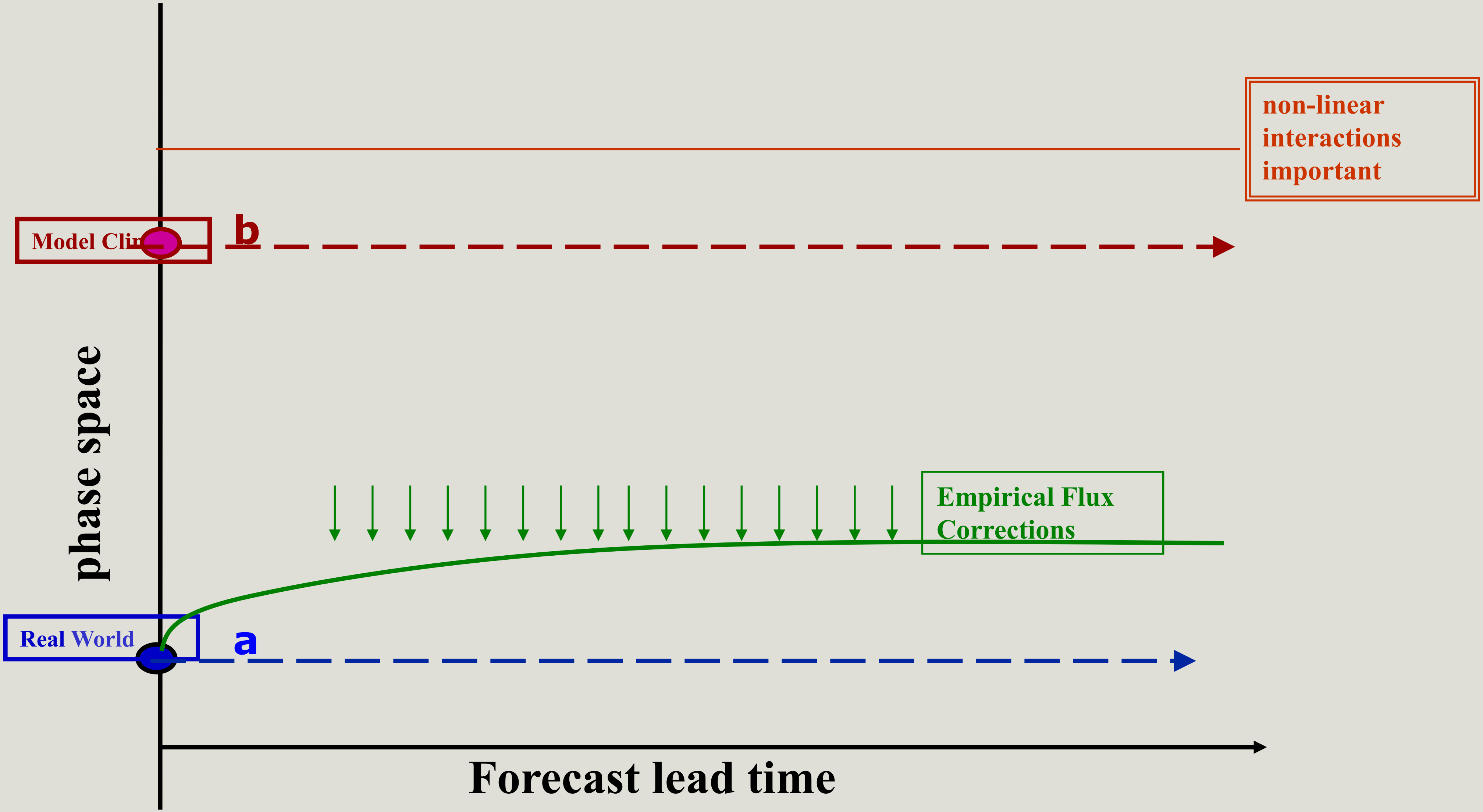
Limitations

- It assumes quasi-linear regime.
- Sampling: how to obtain an observed climatology equivalent to the model climate?

Initialization Shock and Skill

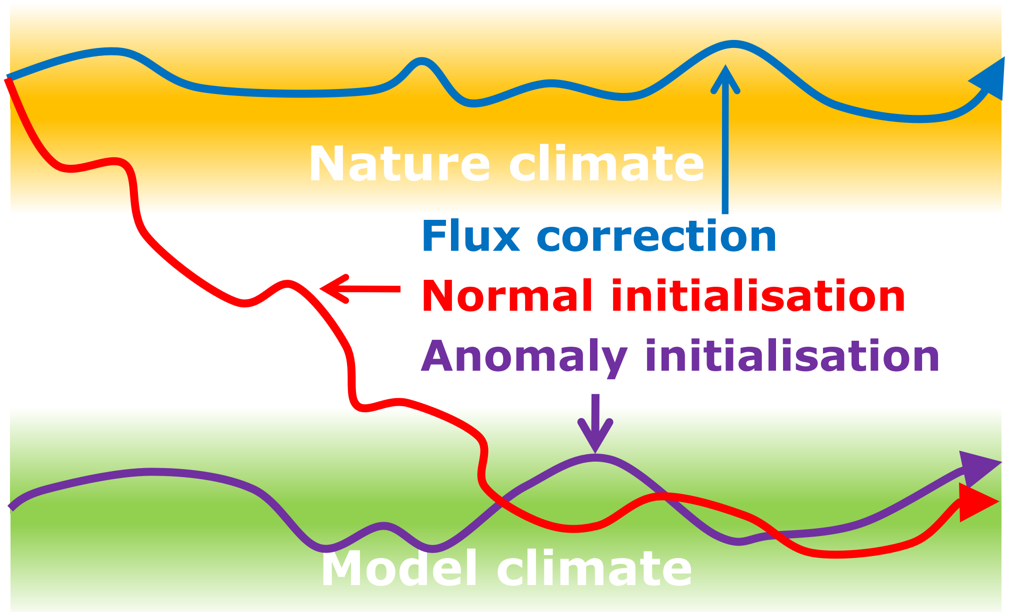


Initialization Shock and non linearities



Comparison of Strategies for dealing with systematic errors in a coupled ocean-atmosphere forecasting system

as part of the EU FP7 COMBINE project

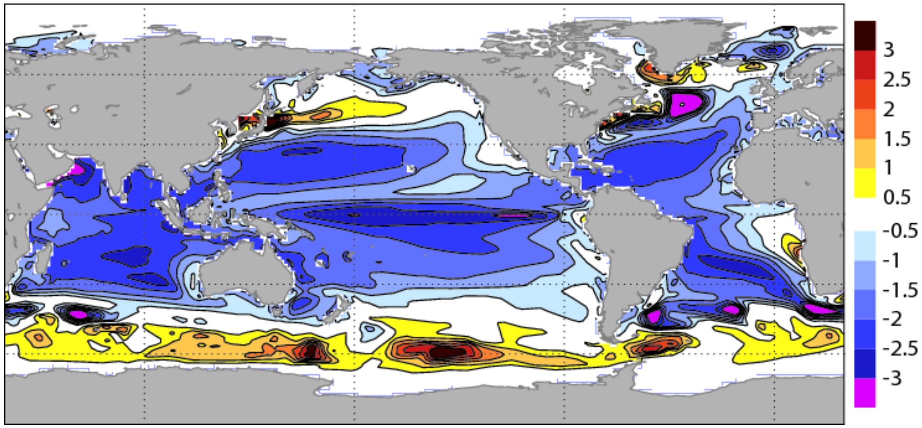


Magnusson et al. 2012, Clim Dyn Submitted. Also ECMWF Techmemo 658

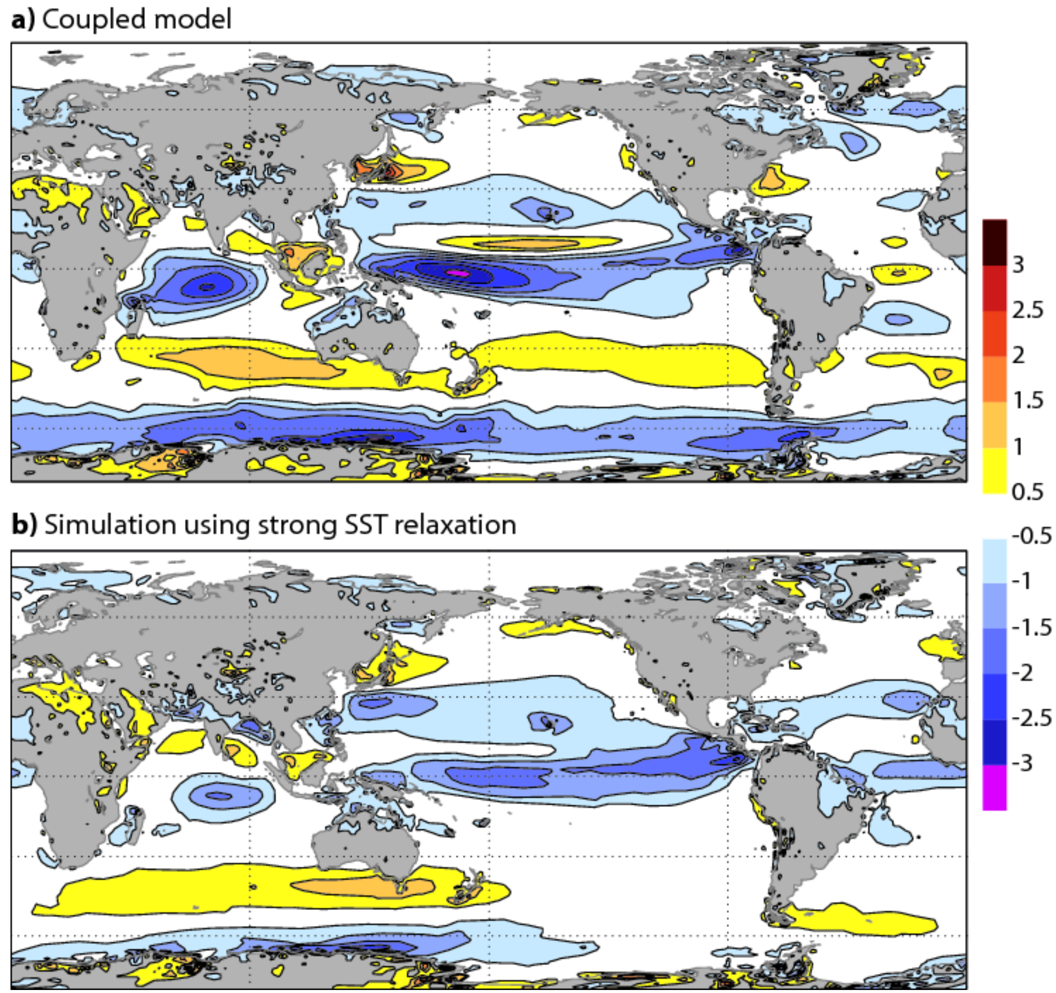
Magnusson et al. 2012, Clim Dyn Sumbitted. Also ECMWF Techmemo 676

Coupled model error

SST bias: model - analysis



10m winds: model - analysis



Part of the error comes from the atmospheric component (too strong easterlies at the equator)

The error amplifies in the couped model (positive Bjerkness feedback**)**

Possibility of flux correction

From Magnusson et al 2012 Clim Dyn. Submitted

Different mean states

Analysis

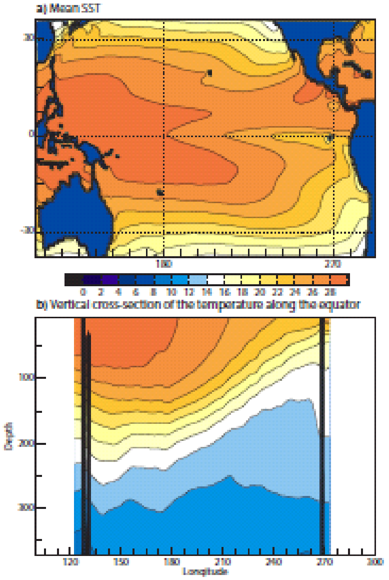
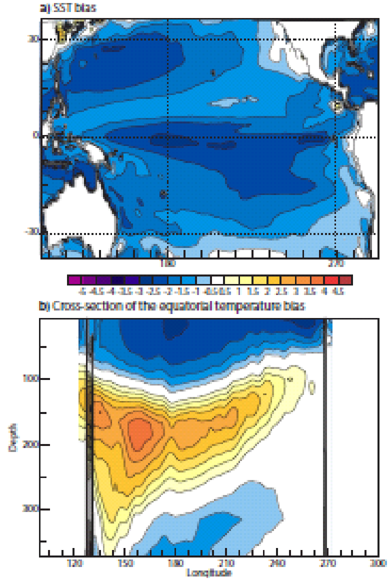
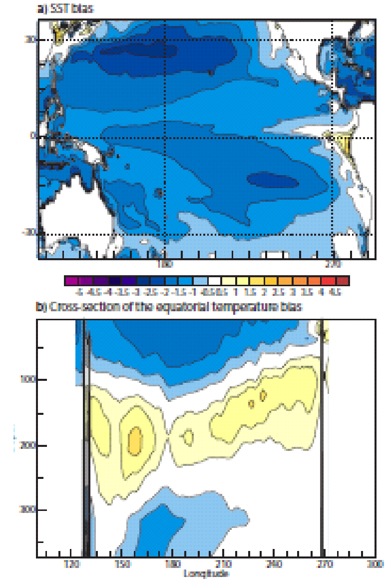


Figure 4: Mean of the reanalysis for 1983 to 1990.

Coupled Free

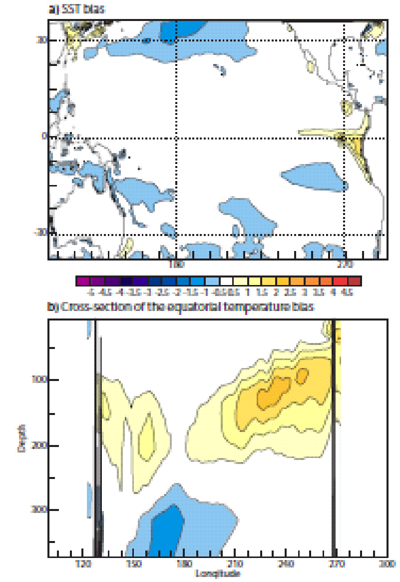


Coupled Ucor



Ucor: surface wind is corrected when passed to the ocean

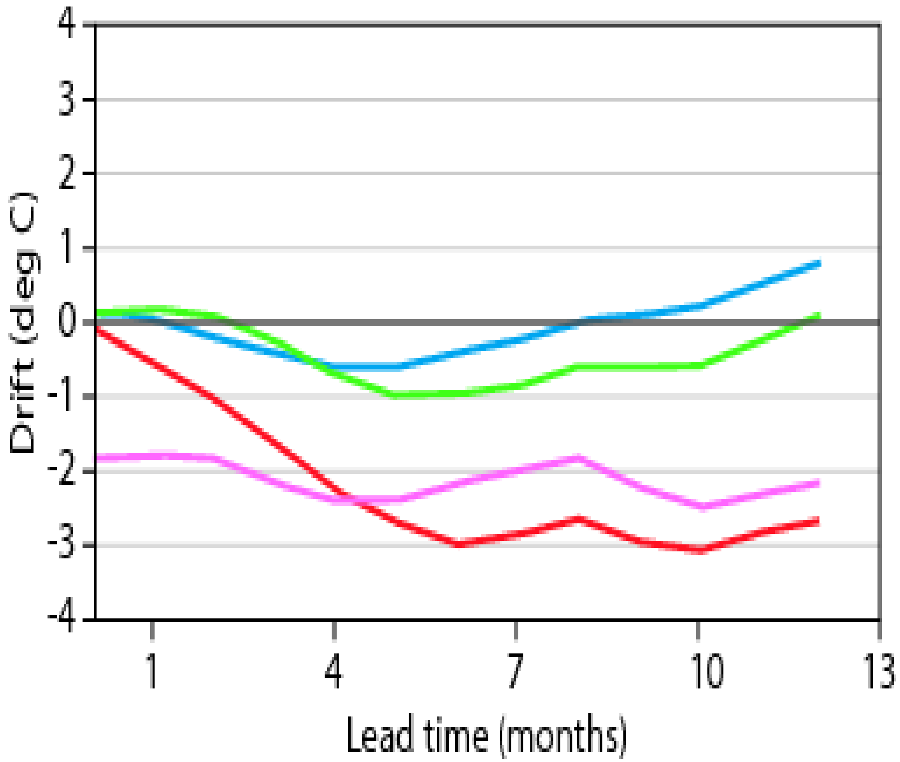
Coupled UHcor



UHCor: surface wind and heat flux are corrected when passed to the ocean

Comparison of Forecast Strategies: Drift

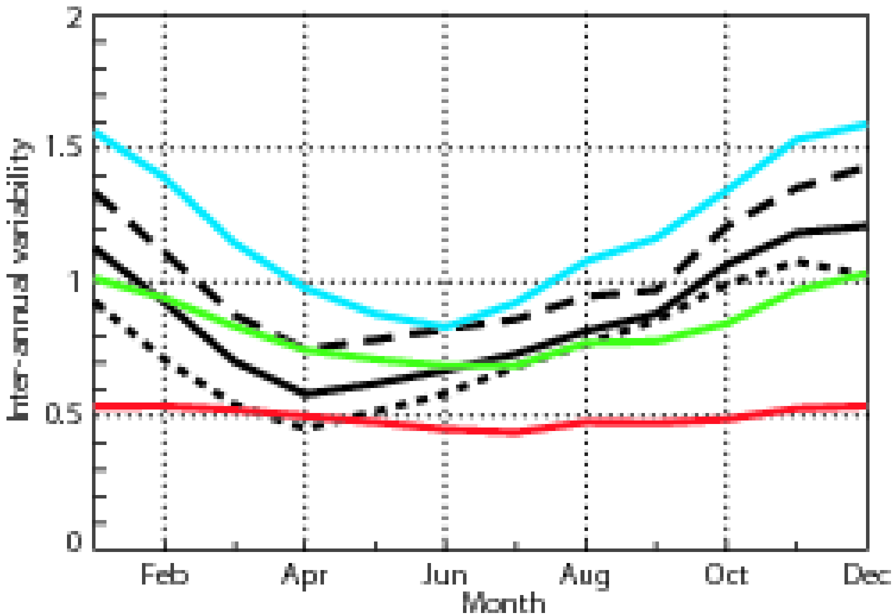
Nino 3 SST Drift 1-14 month forecast



Analysis Full Ini Anomaly Ini U Correction U+H correction

Comparison of Forecast Strategies: Variability

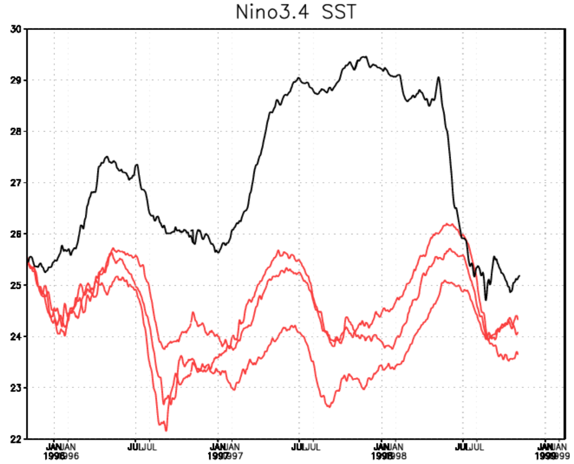
FC sdv / AN sdv



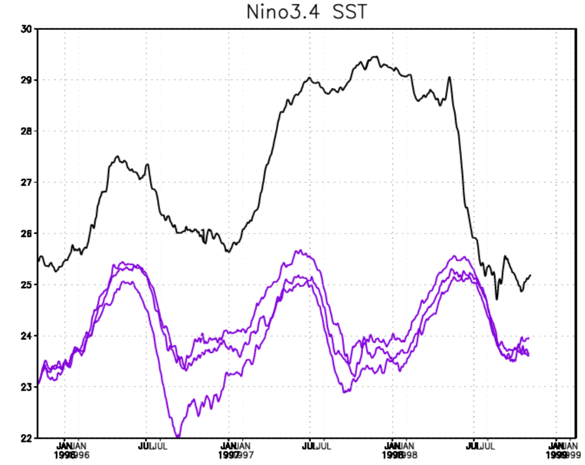
Analysis Full Ini Anomaly Ini U Correction U+H correction

Nino3.4 SST forecasts November 1995 – November 1998

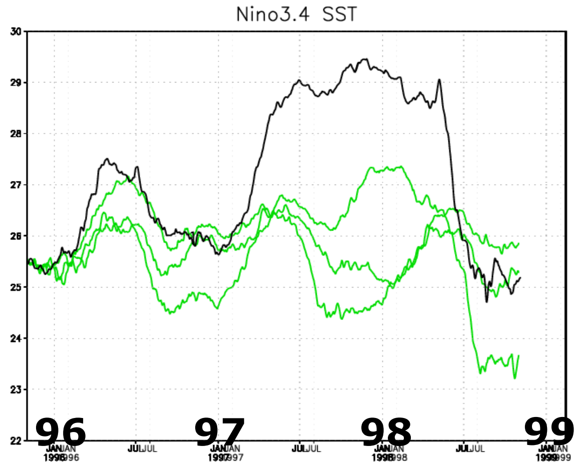
Full Initialization



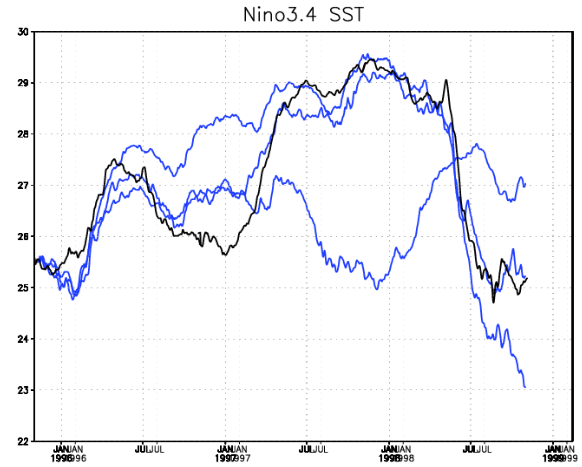
Anomaly Initialisation



U-flux correction

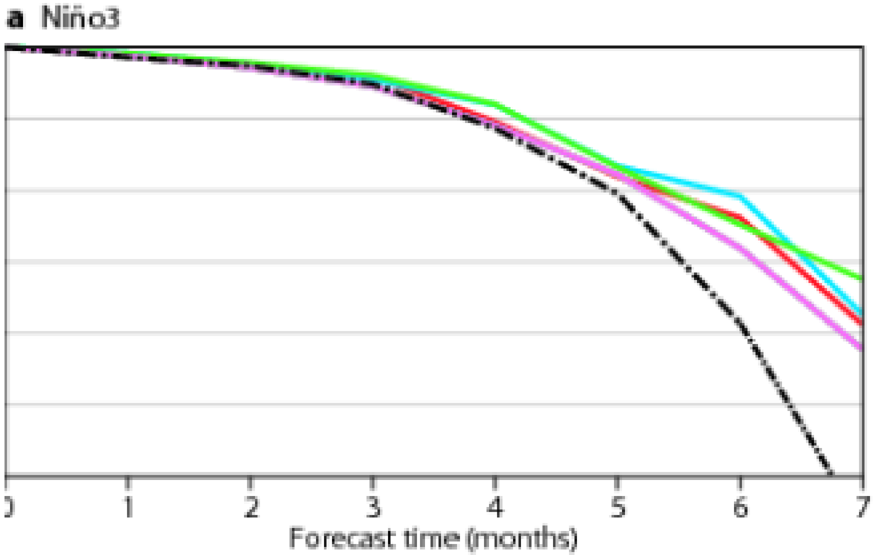


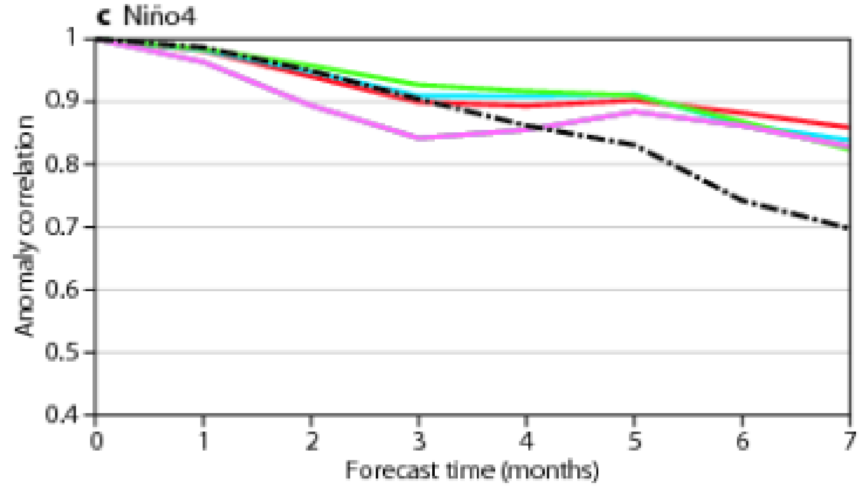
U- and H-flux correction



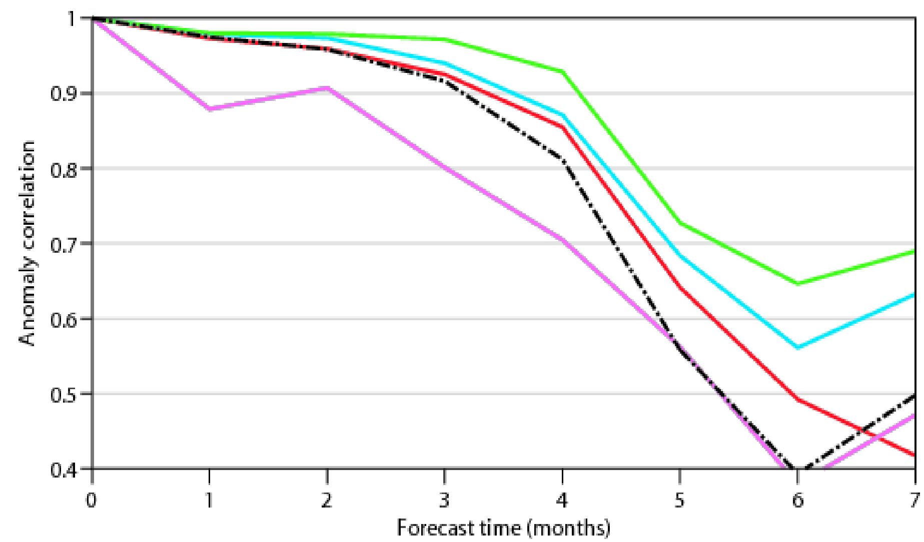
Linus Magnusson et al.

Impact on Forecast Skill (SST and Precip)





Persistence **Full Ini** **Anomaly Ini** **U Correction** **U+H correction**



The impact of initialization/forecast strategy depends on the region

When the mean state matters (convective precip), the anomaly Initialization underperforms

What about Coupled Initialization?

- Advantages:
  - Hopefully more balanced ocean-atmosphere i.c and perturbations. Important for tropical convection
  - Framework to treat model error during initialization and fc

    If the FG and FC models are the same, the (3D) bias correction estimated during the initialization can (should) be applied during the forecast.

  - Consistency across time scales (seamlessness):

    currently, weather forecasts up to 10 days use "extreme flux correction", since SST is prescribed. For longer lead times a free coupled model is used. More gradual transition?

- Current Approaches

  Weakly Coupled Data assimilation: FG with coupled model, separate DA of ocean and atmos. Example is NCEP with CFSR: coupled reanalysis to initialized and calibrate seasonal forecasts

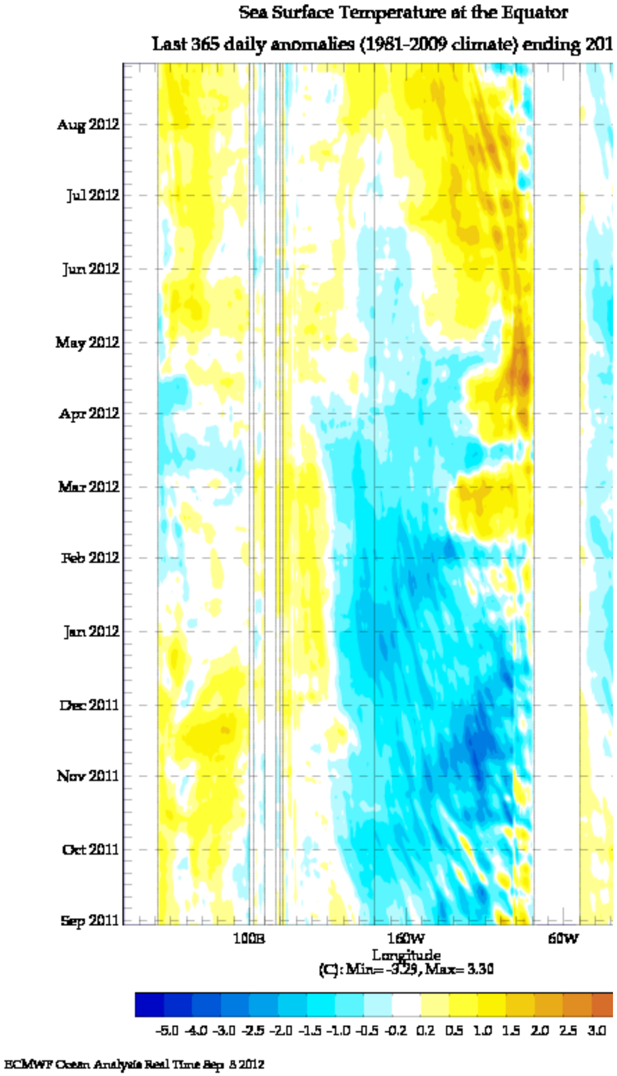
  Strongly Coupled Data assimilation: Coupled FG, Coupled Covariances**.** Usually EnKF

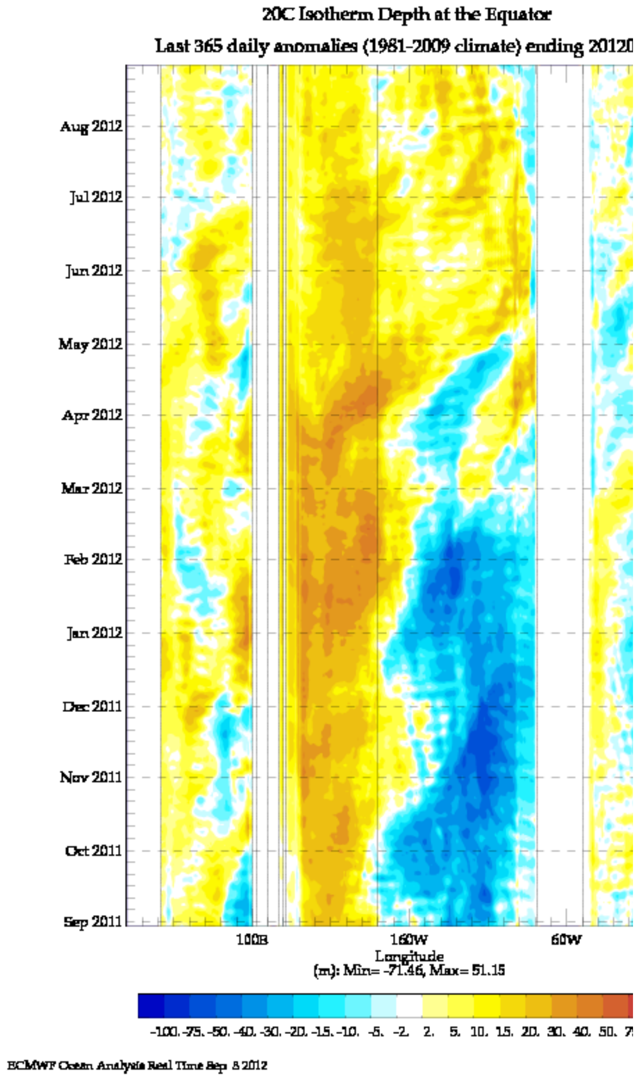
- Challenges:
  - Different time scales of ocean atmosphere . Long window weak constrain?
  - Cross-covariances. Ensemble methodology more natural?

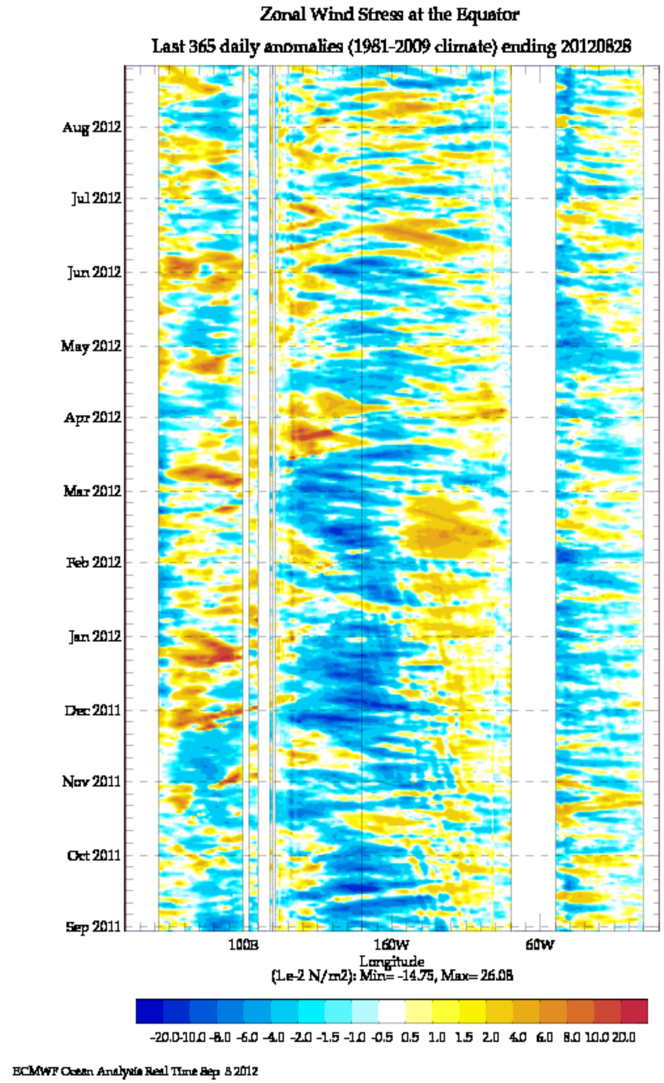
Summary

- Seasonal Forecasting (SF) of SST is an initial condition problem
- Assimilation of ocean observations reduces the large uncertainty (error) due to the forcing fluxes. Initialization of Seasonal Forecasts needs SST, subsurface temperature, salinity and altimeter derived sea level anomalies.
- Data assimilation improves forecast skill.
- Data assimilation changes the ocean mean state. Therefore, consistent ocean reanalysis requires an explicit treatment of the bias
- The separate initialization of the ocean and atmosphere systems can lead to initialization shock during the forecasts. A more balance “coupled” initialization is desirable, but it remains challenging.
- Initialization and forecast strategy go together. The best strategy may depend on the model. The anomaly initialization used in decadal forecasts can have problems in seasonal

Evolution over the last 365 days







Latest Conditions