Ensembles in the ocean: an overview

Magdalena Alonso Balmaseda

Contributions from
Hao Zuo, Eric de Boisseson, Patrick Laloyaux, Simon Lang, Kristian Mogensen, Tim Stockdale

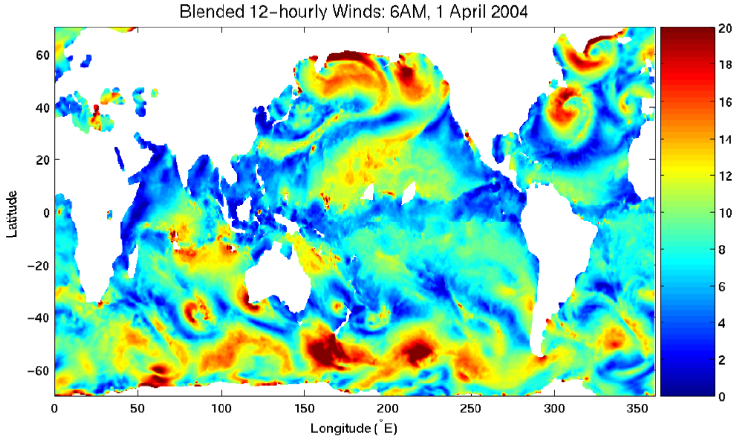
Outline

- **Setting the scene**
  - Ocean versus atmosphere
  - Marine versus weather and climate applications: different drivers and different practices
- **Ongoing developments using ensembles in the ocean**
  - **Ensemble of ocean reanalysis –Parameter Estimation –Quantifying ocean chaos**
- **Ocean ensembles for coupled forecasting: Focus on seasonal at ECMWF**
  - A history of developments
  - Ensembles in ORAS5
- **Present and Future**
  - Seamless Coupled Prediction – Hybrid Assimilation Methods -Coupled DA

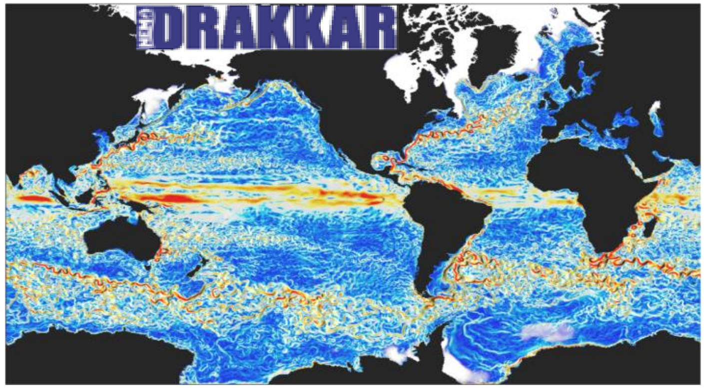
Ocean v Atmosphere: reminder

Spatial scales
Continental boundaries

Atmospheric wind speed (12h)



Ocean velocity 5-day mean



Temporal Scales,
Memory and
observations

- Ocean has **long memory** -thermal and dynamical inertia.
- Number of observations in atmosphere **~10^3-10^4 times** more than in the ocean.
- So the ocean (re-) analyses take long time to spin up.

Internal – forced –
coupled variability
and stability

- Ocean large scale variability is mostly **forced** by atmosphere.
- The ocean **internal chaotic** component is associated with eddy scale
- There are **unstable** ocean-atmospheric **coupled modes** at planetary scale, e.g. ENSO
- Ocean - atmosphere interaction also occurs at small eddy and frontal scale: stable or unstable?

Ensemble practices in the ocean: Marine and Weather/Climate

- **Data assimilation**
    - **Ensemble methods widely used for ocean DA**. EnKF started in the ocean (Evensen 1994).
    - Advantage to represent complex spatial structures in B.
    - Both in coupled and uncoupled mode data assimilation. Usually simplified (ensemble OI).
- **Forecasting**
    - **Marine (time scales 5-10 days): ensemble forecast hardly used.**
        - **The ocean is considered a deterministic system, forced by atmosphere fluxes (UNCOUPLED)**
        - **Chaotic nature, Uncertainty in atmospheric forcing** and **initial conditions** is also usually neglected.
    - **Coupled forecasting: seasonal, decadal and -more recently- medium and extended range.**
        - **Probabilistic** forecasts, therefore ensembles
        - **Uncertainty in initial conditions/forcing** acknowledged
- **Reanalysis and monitoring**
    - **Structural uncertainty acknowledged . Ensemble of Ocean Re-Analysis (ORAs)**

ORA-IP v1: Ocean ReAnalysis Intercomparison Project
Production Vintage 2010 - Analyzed 2013-2016

- 6 Observation only products (model independent)
- 13 Low resolution ocean reanalyses
- 8 High resolution reanalyses (1/3 or ¼ degree)
- 4 Coupled DA products
- 6 Long reanalyses, starting 1950's

Variable	Paper
Ocean Heat Content	Palmer et al
Steric Height	Storto et al
Heat Fluxes	Valdivieso et al
MLD	Toyoda et al a,b
Salinity	Shi et al
Sea Ice	Chevallier et al
AMOC	Karspec et al

Balmaseda et al 2015
+ Special Issue Climate Dynamic

Lessons Learnt: Signals and Sources of Uncertainty

Well constrained:

Temperature variability upper 300m
Sea Level
Mixed Layer Depth
Total Steric Height
Sea Ice Edge

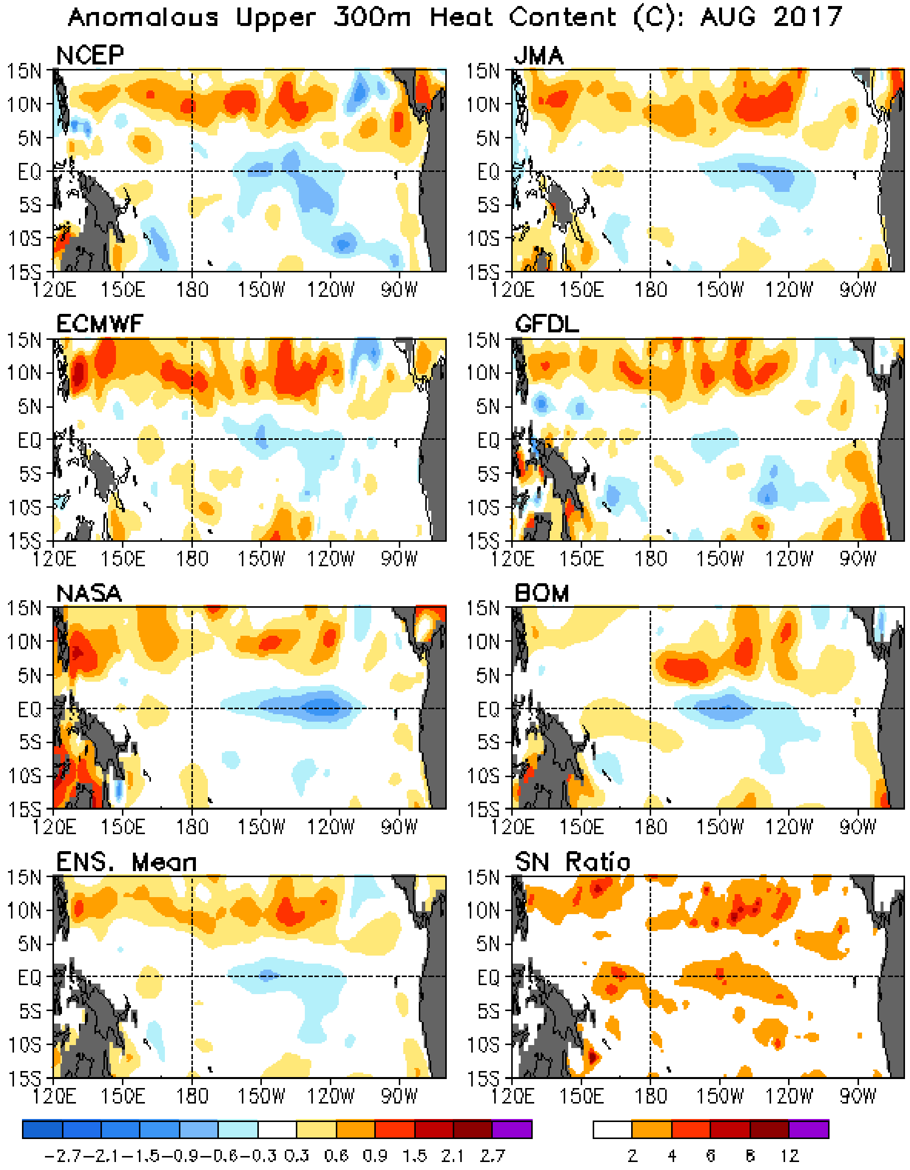
Poorly constrained:

Deep ocean (below 700m):
Steric Height Partition (Halo-Thermo) and depth range contribution
Atlantic Meridional Overturning Circulation
Salinity
Surface Heat Fluxes:
Sea Ice Thickness

Non-trivial result: Data Assimilation Method a main source of uncertainty.

ORA-IP Legacy

- **Lessons learnt (papers in peer reviewed literature)**
- **Multi-ORA archive:**
  - Climate Indicators and metrics with uncertainty estimates
  - Bench-mark to measure progress. Version control
- Multi-ORA real-time monitoring of climate:
  - Temperature (NCEP) and Salinity (BoM)
- European MULTI-ORA by Copernicus Marine Environmental Monitoring Services (CMEMS)
- PORA-IP and YOPP
- Further Scientific comparisons: assimilation increments and model error
- **Next: Multi-ORA for initialization of coupled modes?**

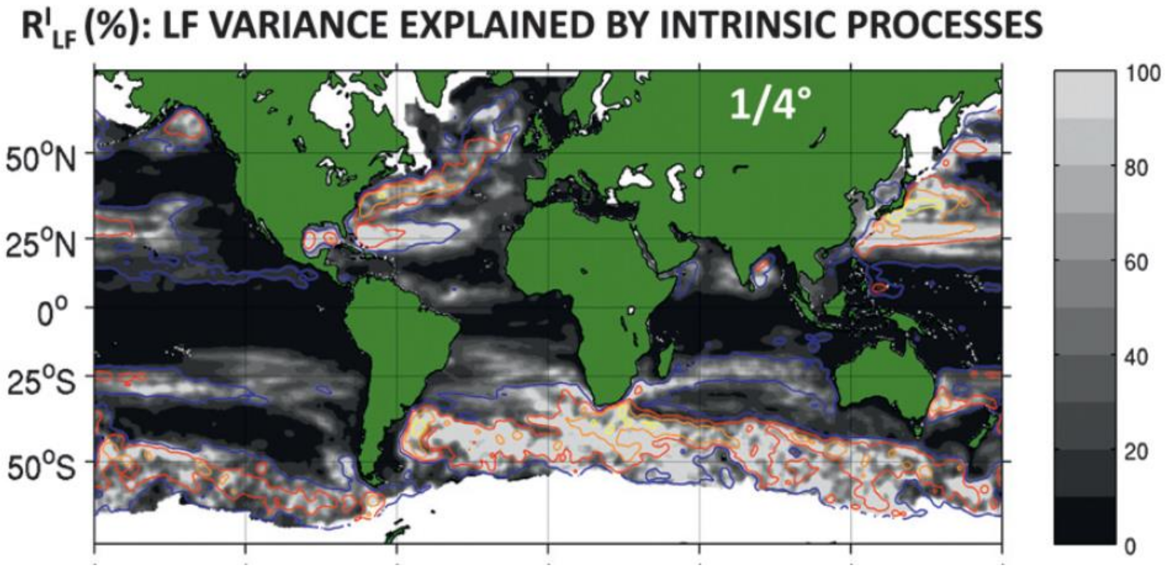


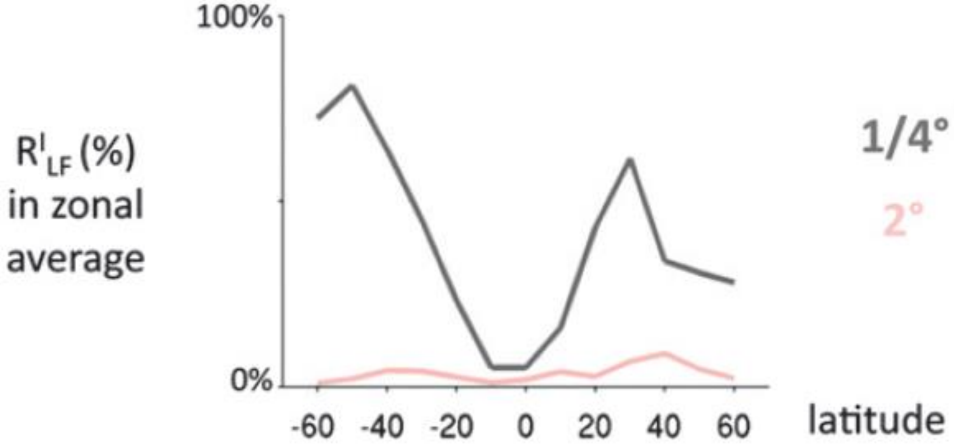
http://www.cpc.ncep.noaa.gov/products/GODAS/multiora_body.html

Other recent activities

- Model Error: Stochastic parameterizations (Laure Zanna's talk)

- Quantifying the Ocean Chaotic Component: Penduff et al 2011 and Project OCCIPUT

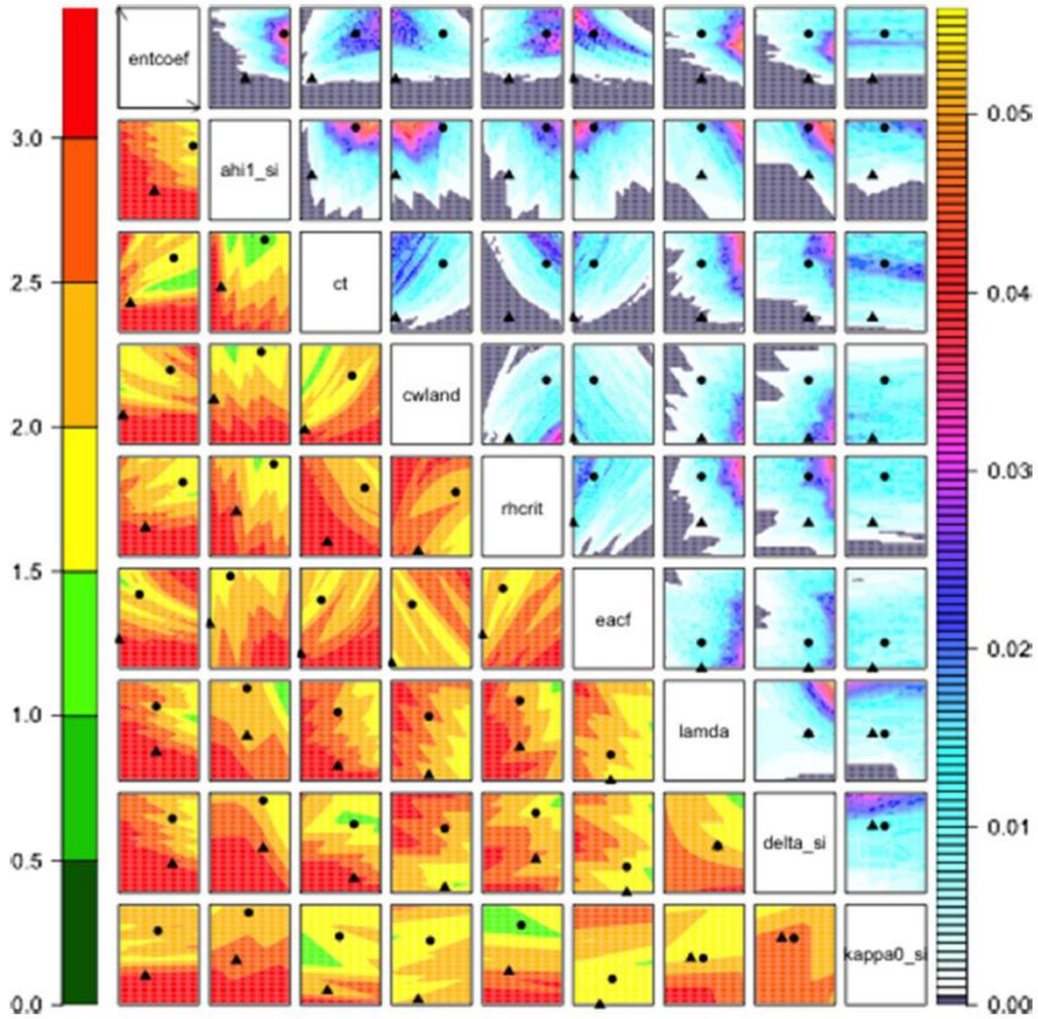




Internal variability high in frontal areas and ACC ~40% of total. Resolution dependent

Other recent activities wind ensembles

Using ocean ensembles and emulators for multi-variate parameter tuning and reducing parameter space. Williamson et al 2013,2014.



PPE ensembles with HadCM3
+
Emulators
+
"History matching Constraints"
+
Iterative "Not Ruled Out Yet" (NROY).

Less Implausible parameter combination

Ensembles in the Ocean: an ECMWF perspective

- A brief history of developments at ECMWF: 2002-2017

Main driver: seasonal forecasting –ENSO

Evolution and choices

- Widening the scope:
    - Seamless prediction
    - Hybrid DA in the ocean
    - Coupled DA

.

ECMWF

Ensembles for Seasonal / ENSO prediction: an ECMWF perspective

From the **atmospheric perspective**, seasonal forecast is a **boundary condition problem**

From the **ocean perspective**, seasonal forecast is an **initial value problem**

Need to sample uncertainty in ocean initial conditions

Difference with respects to Ensembles in NWP:

- **A-posteriori calibration of forecast PDF is needed in monthly seasonal**
  No major efforts in tuning the perturbations to achieve reliable forecasts.
  Assessment of the ensemble reliability is important

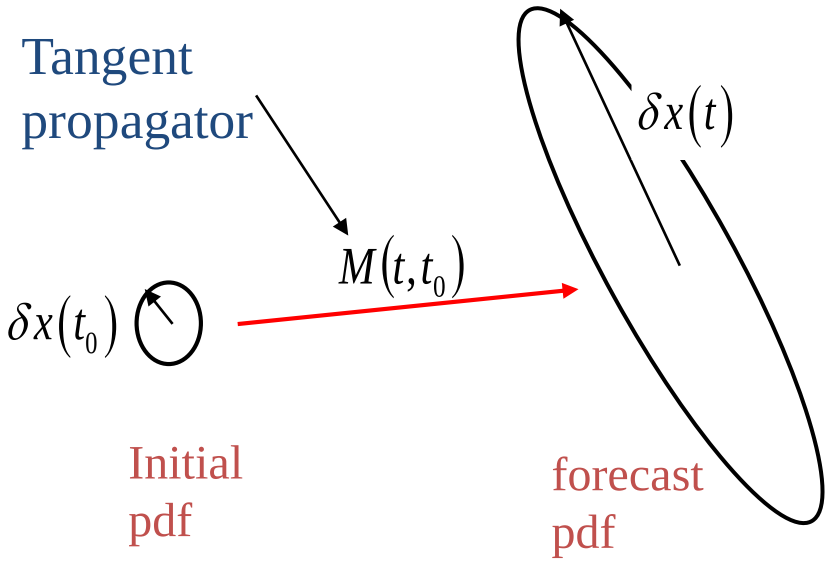
- **System with 2 time scales: days in the atmosphere and months in the ocean.**

How to generate the ensemble of ocean i.c for seasonal forecast?

Question ~ yr 2000

Are Singular Vectors a valid approach for operational seasonal forecasts?

Medium Range: Singular Vectors



$$\Rightarrow M^{*} M \,\delta x(t_0) = \lambda \,\delta x(t_0)$$

We need the TL& Adjoint of the full coupled model is required.

BUT…

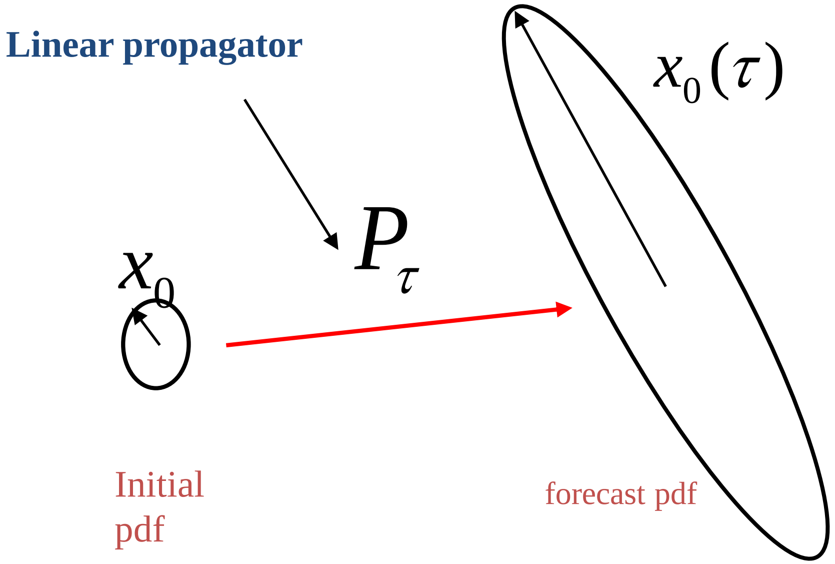
The linear assumption would fails for the atmosphere at lead times relevant for seasonal (~>1month).

Alternatives

1. Other approaches for **optimal** sampling of initial condition uncertainty:
   - Breeding Vectors (NASA, BoM. Not shown here)
   - SV using Generalized Linear Propagators
2. Sample **known** i.c. uncertainties, without considering optimality

Uncertainty in initial conditions may not be the dominant source of error

Generalized Singular Vector Problem (I)



Generalized Linearized Propagator (**not necessary tangent linear**)

$$x_0(\tau) = P_\tau x_0$$

Given a final **N** and initial norm **L**, the growth in **x** can be measured by

$$A(\tau) = \frac{\mathbf{x}(\tau)^{\mathrm{T}}\mathbf{N}\mathbf{x}(\tau)}{\mathbf{x}_0^{\mathrm{T}}\mathbf{L}\mathbf{x}_0} = \frac{\mathbf{x}_0^{\mathrm{T}}\mathbf{P}_\tau^{\mathrm{T}}\mathbf{N}\mathbf{P}_\tau\mathbf{x}_0}{\mathbf{x}_0^{\mathrm{T}}\mathbf{L}\mathbf{x}_0},$$

Optimal perturbations are those that maximize λ

$$\mathbf{P}_\tau^{\mathrm{T}}\mathbf{N}\mathbf{P}_\tau\mathbf{x}_0 = \lambda\mathbf{L}\mathbf{x}_0$$

Different ways of estimating the Linear Propagator P(τ)

I. Empirical (or Inverse modelling): basically a regression :
II. A simplified linear dynamical model (equilibrium atmosphere rather than tangent linear)
III. A hybrid system: Ocean GCM coupled to a simplified atmosphere

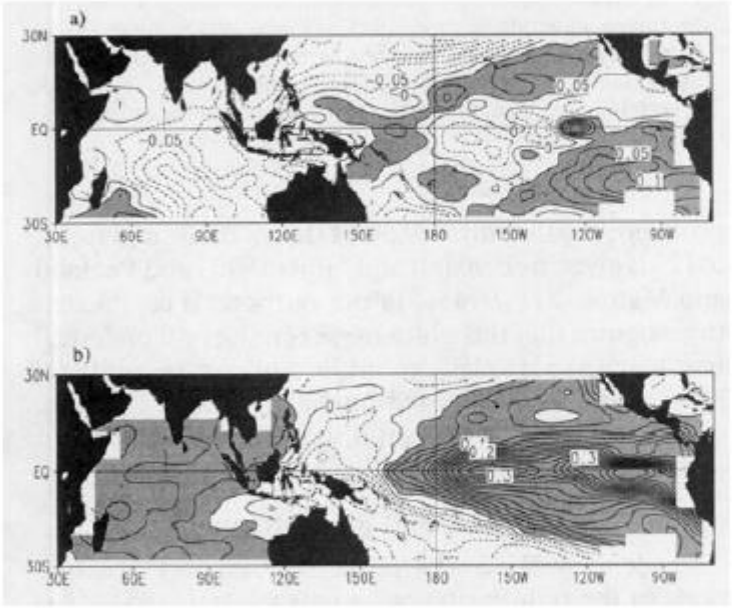
Generalized Singular Vector Problem (II)

Linear Propagator estimated empirically via regression model (Inverse modelling)

$$\frac{d\mathbf{x}}{dt} = \mathbf{B}\mathbf{x} + \xi,$$

- **From temporal records of observations**
  von Storch and Xu 1990 MJO   (POPs Principal Oscillation Patterns)
  Blumenthal   1991 ENSO
  Penland and Sadershmuck 1995, ENSO (inverse modelling)

- **From temporal records of model evolution**
  Xue et al 1997a,b; Fan et al 1999   ENSO
  Hawkins and Sutton 2009   Decadal Prediction AMOC



Initial SST

Final SST

Penland and Sadershmukh 1995

This approach is based on temporal sampling of existing timeseries: Difficult to capture flow dependence or errors of the day.
Judgement: not appropriate for ensemble generation.in operational systems.

These are **powerful tools for a-posteriori diagnostics of ensemble statistics for evaluation of forecasts**;. Ensemble Sensitivity.
Magnusson 2017 QJRMS

Generalized Singular Vector Problem (III)

Hybrid:

TL and Adjoint of Ocean model

+

simplified linear atmospheric model

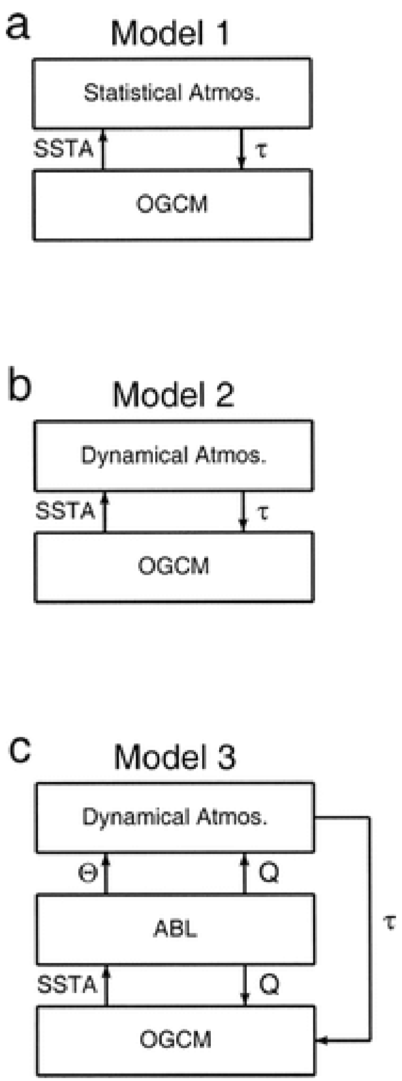
:

Moore et al 2003 used NEMO coupled to

a) A linear statistical atmosphere
b) A linearized dynamical linearized model
c) Linearized dynamical + ABL

Strong dependence on the details of the linearized atmosphere model.

Judgement: this approach not suitable for operational implementation.



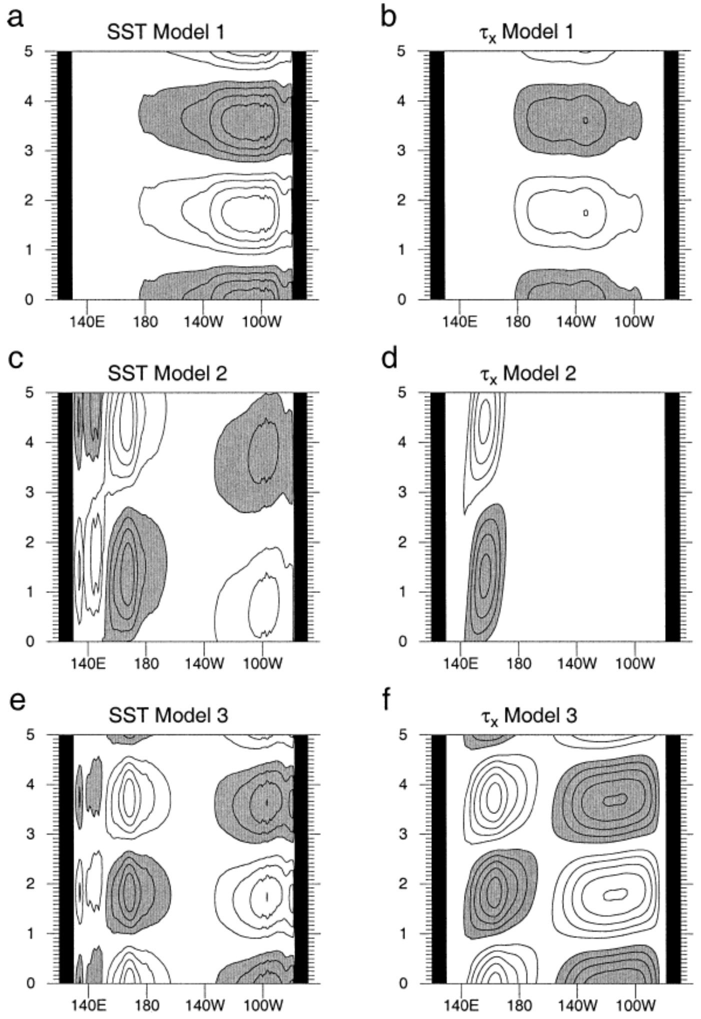
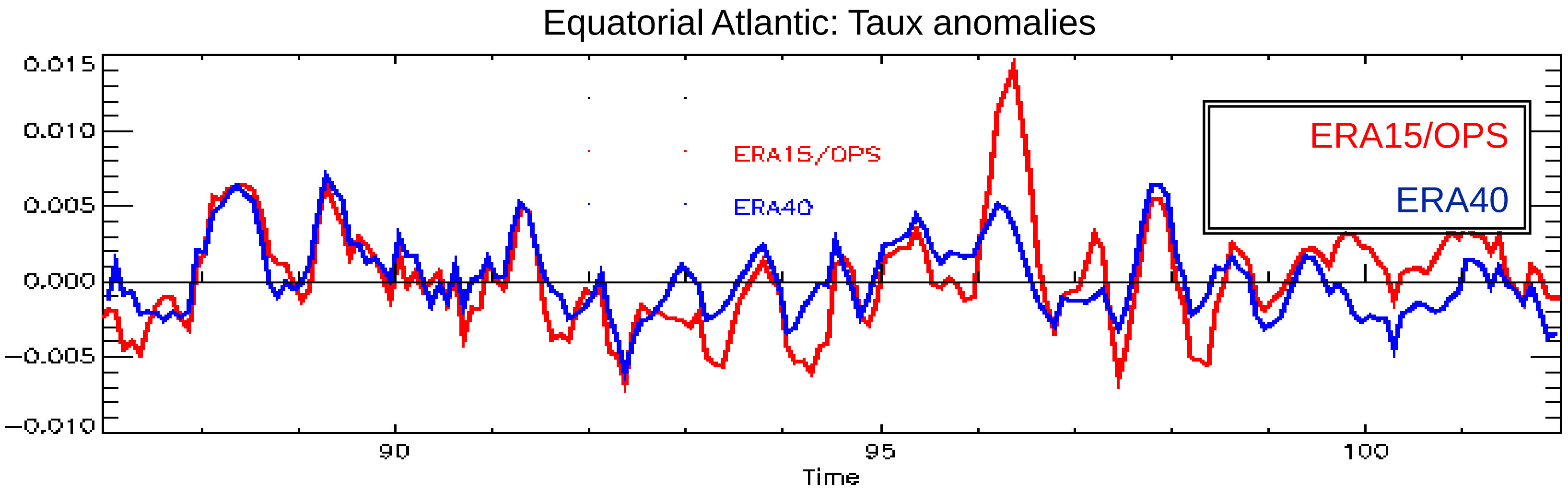


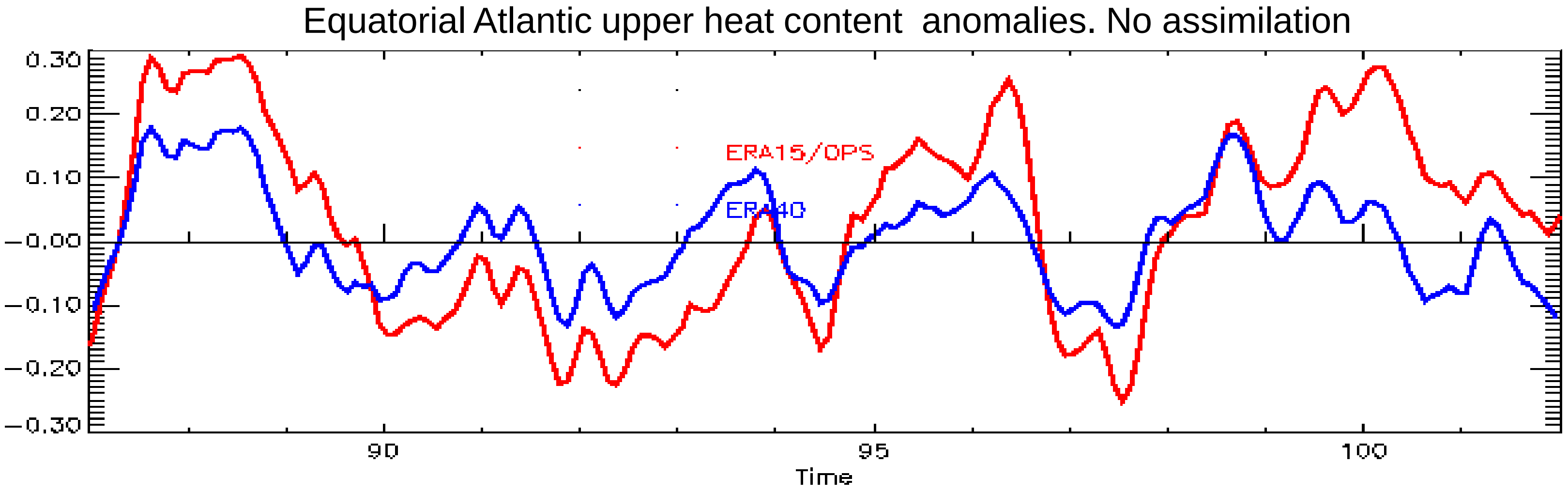
FIG. 4. Hovmöller diagrams of SST and zonal wind stress τ_x averaged between 2°N and 2°S for the most unstable eigenmode of (a), (b) model 1, (c), (d) model 2, and (e), (f) model 3. Shaded and unshaded regions indicate perturbations of opposite sign. The contour interval is arbitrary. In each case the exponential growth of the eigenmodes has been suppressed.

This approach was also used to estimate optimal forcing perturbations (or Stochastic Optimals, see later)

Representing Known Analysis Uncertainties

Uncertainty in surface wind stress a primary source of uncertainty in ocean initial conditions

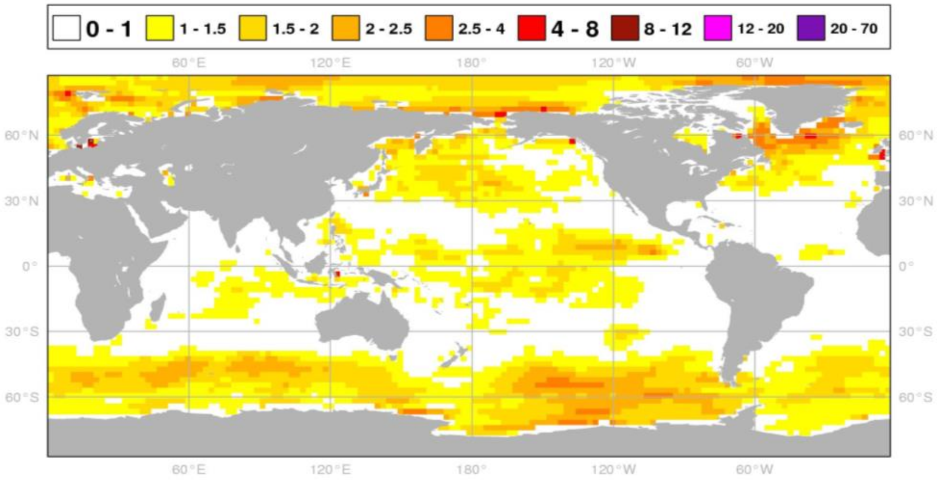




1.1 Wind stress perturbations

- Create data base with errors in the monthly anomalies of wind stress, arranged by calendar month:

SDV of Wind Stress Perturbation



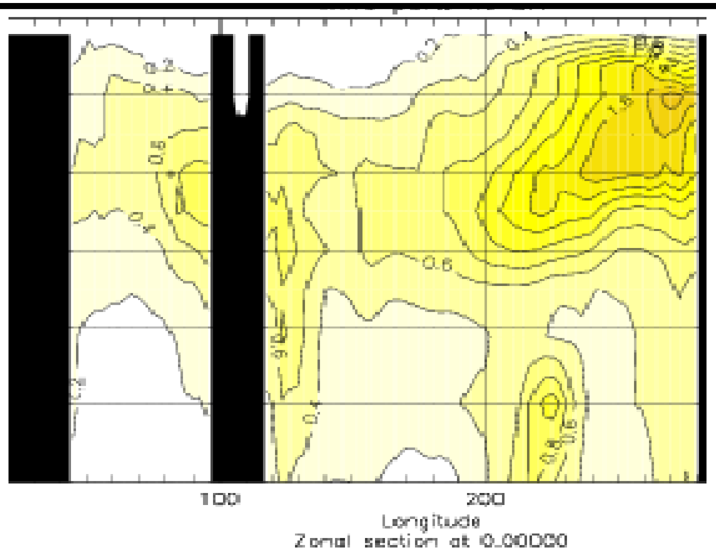
- Random draw of monthly perturbations, applied during the ocean analyses.

- **Create a centered ensemble of 5 reanalysis is constructed symmetric wind perturbations**

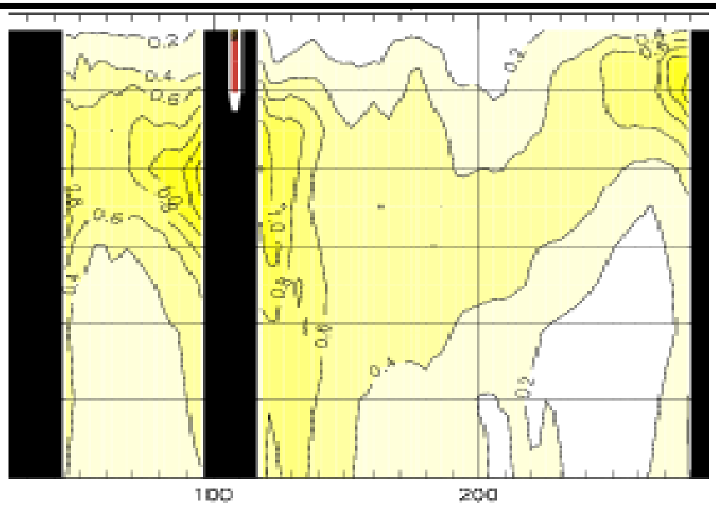
-P2 -P1 0 P1 P2

-

SDV Ocean Subsurface T: No Data assimilation



SDV Ocean Subsurface T: Data assimilation



C.I 0.2 deg C

1.2 SST Perturbations

-Create data base with errors of weekly SST anomalies, arranged by calendar week:

V1. Error in SST product: (differences between OIv2/OI2dvar)

+ Errors in time resolution: weekly versus daily SST

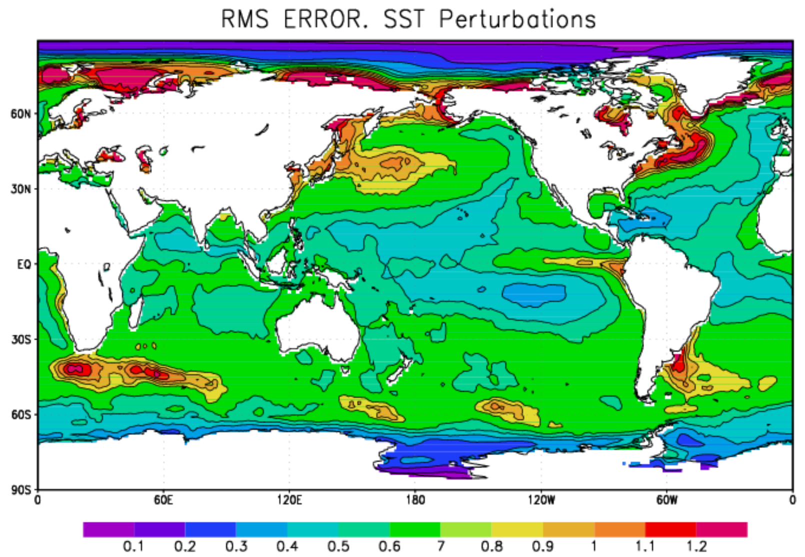
V2.

V3.

-Random draw of weekly perturbations, applied at the beginning of the coupled forecast. Over the mixed layer (~60m)

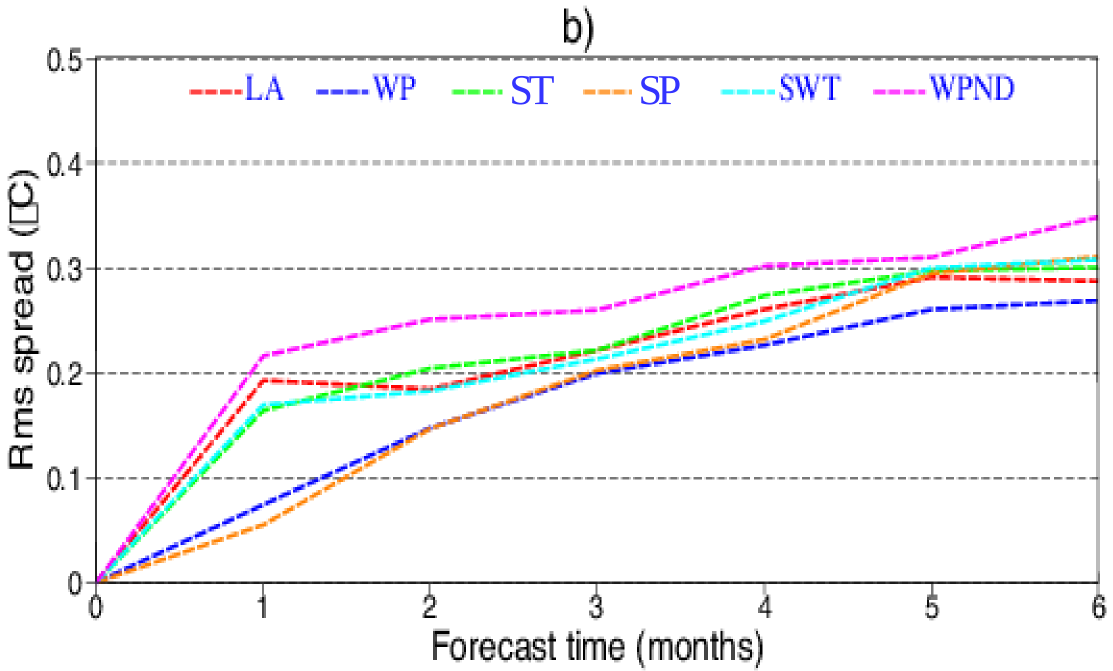
-A centred ensemble

SST Perturbations SDV



Comparison on contributions to the Ensemble Spread in seasonal forecasts

Wind Perturbations (WP) Wind Perturbations No DA (WPND)
SST Perturbations(ST) All(SWT)
Stochastic Physics (SP) Lag-averaged(LA)

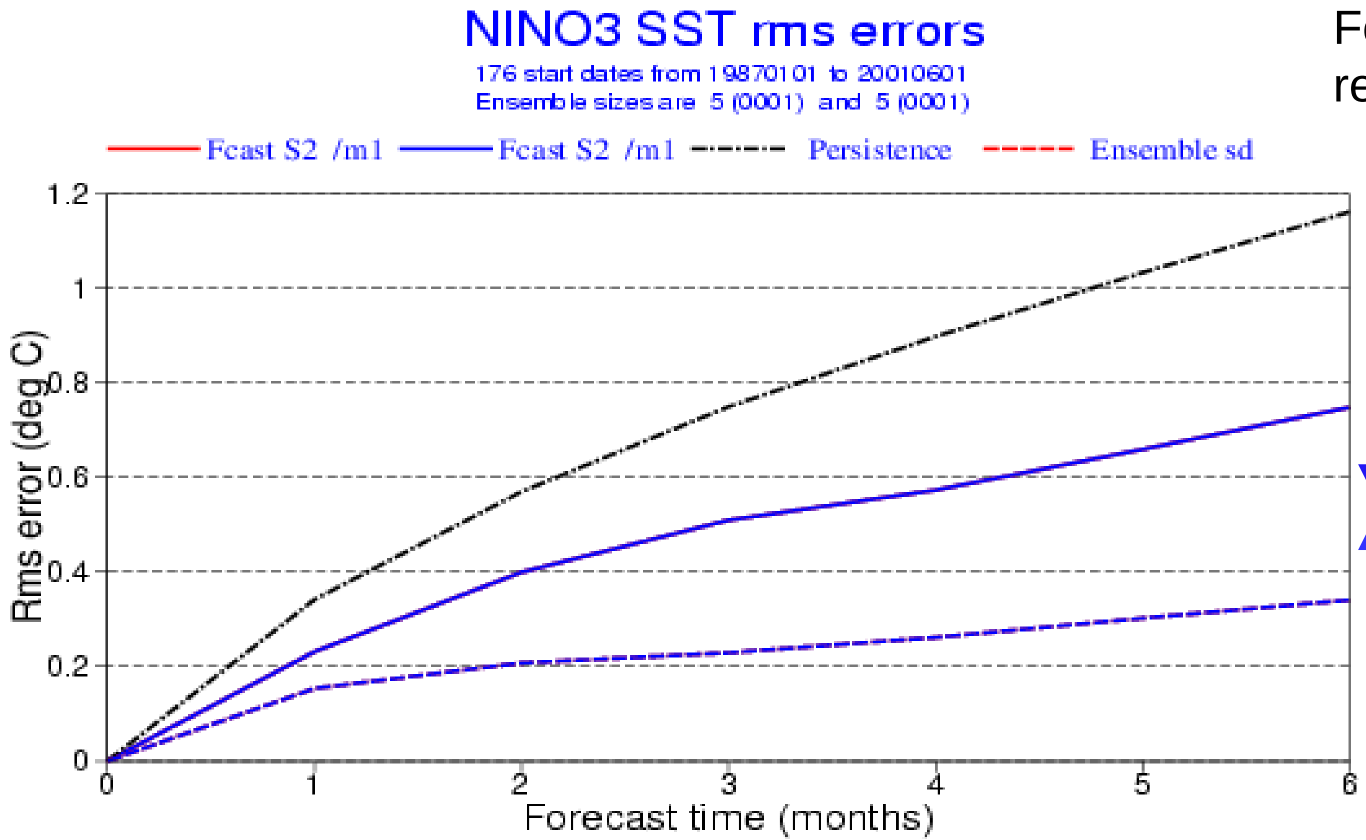


From Vialard et al, MWR 2005

- The spread by different methods converge to the same asymptotic value after after 5-6 months.
- SST and Lag-averaged perturbations dominate spread at ~1month lead time.
- With DA, the wind perturbations grow slowly, and notably influence the SST only after 3m. Similar growth as SP.
- Without DA, the initial spread (<3m) is larger. The asymptotic value is slightly larger

Is the level of spread sufficient?

Is the ensemble spread sufficient? Are the forecast reliable?



Forecast System is not reliable:

RMS > Spread

A. Can we reduce the error? How much? (Predictability limit)

B. Can we increase the spread by improving the ensemble generation?

To improve the ensemble generation we need to sample other sources of error:

a) Model error: multi-model

b) To design other optimal methods: Stochastic Optimals, Breeding Vectors, …

Or a-posteriori calibration of the ensemble.

Stochastic Optimals

Linear Theory:

Consider a stochastically forced linear ENSO

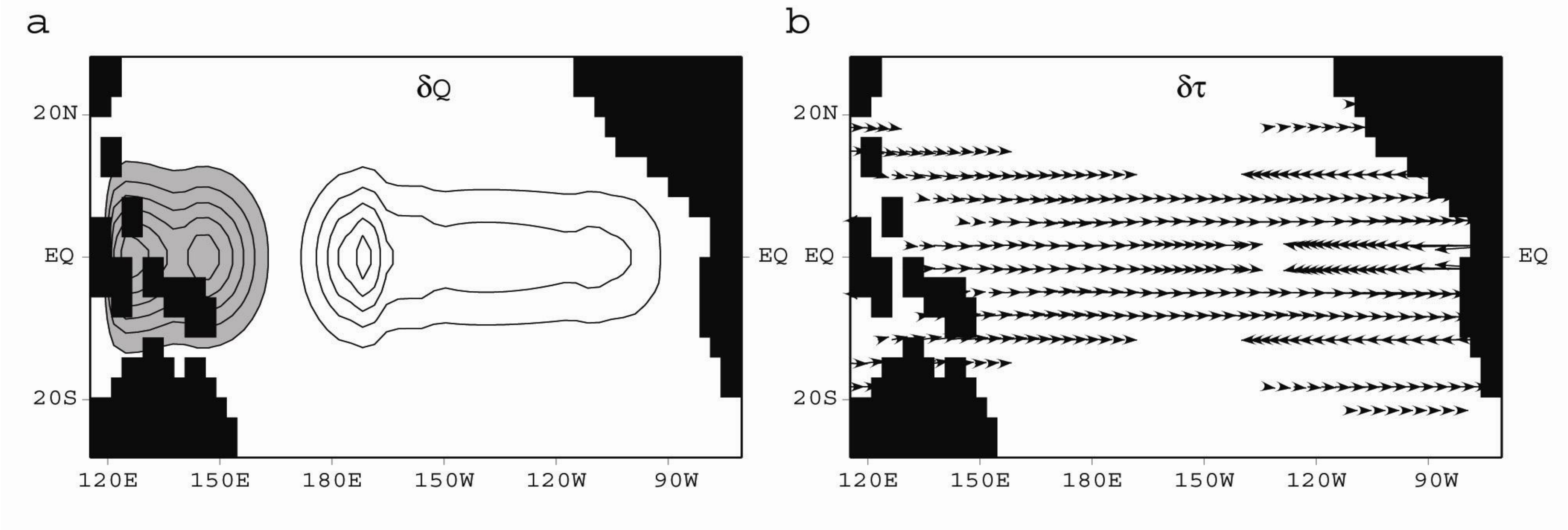
$$ds/dt = As + f(t)$$

- f(t) is coherent in space and white in time.
- Which are the patterns of f(t) that maximize the variance of s?

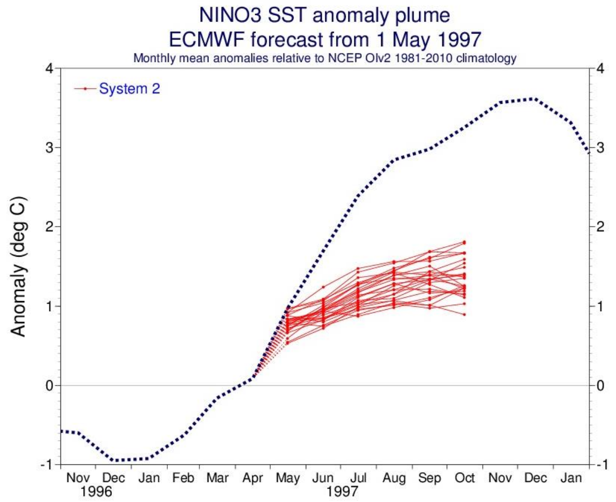
(Farrell and Ioannou, 1993, Moore et al 2003)

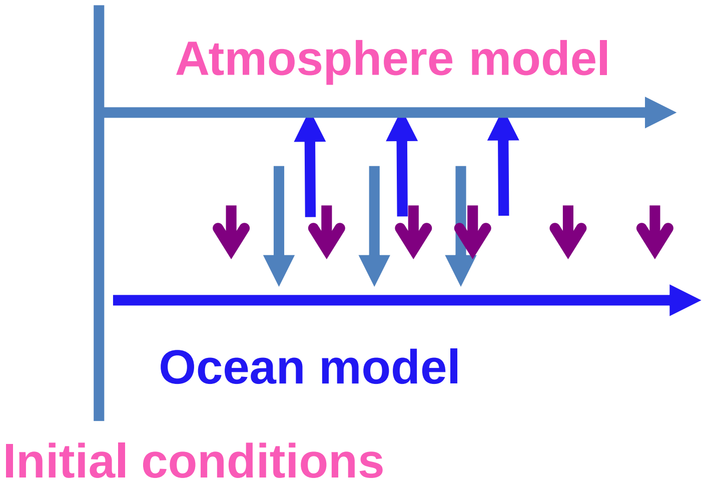
Example of Stochastic Optimals for and Intermediate Coupled Model for the Tropical Pacific

(Zabala-Garay et al,2003)



Adding Perturbations online



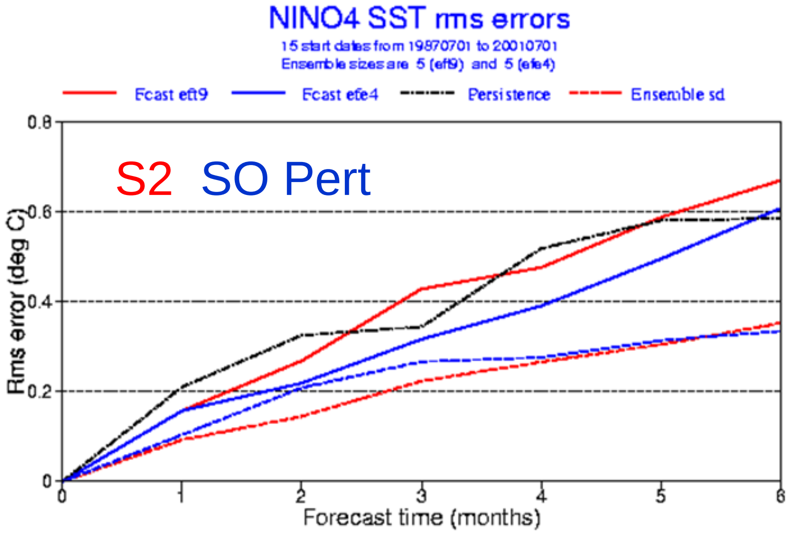


S2 failed to forecast the onset of the 1997-8 El Nino.

Vitart et al 2003 attributed failure to lack of Westerly Wind Bursts associated with the MJO.

Can we add perturbations to the coupled model to account for known deficiencies?.

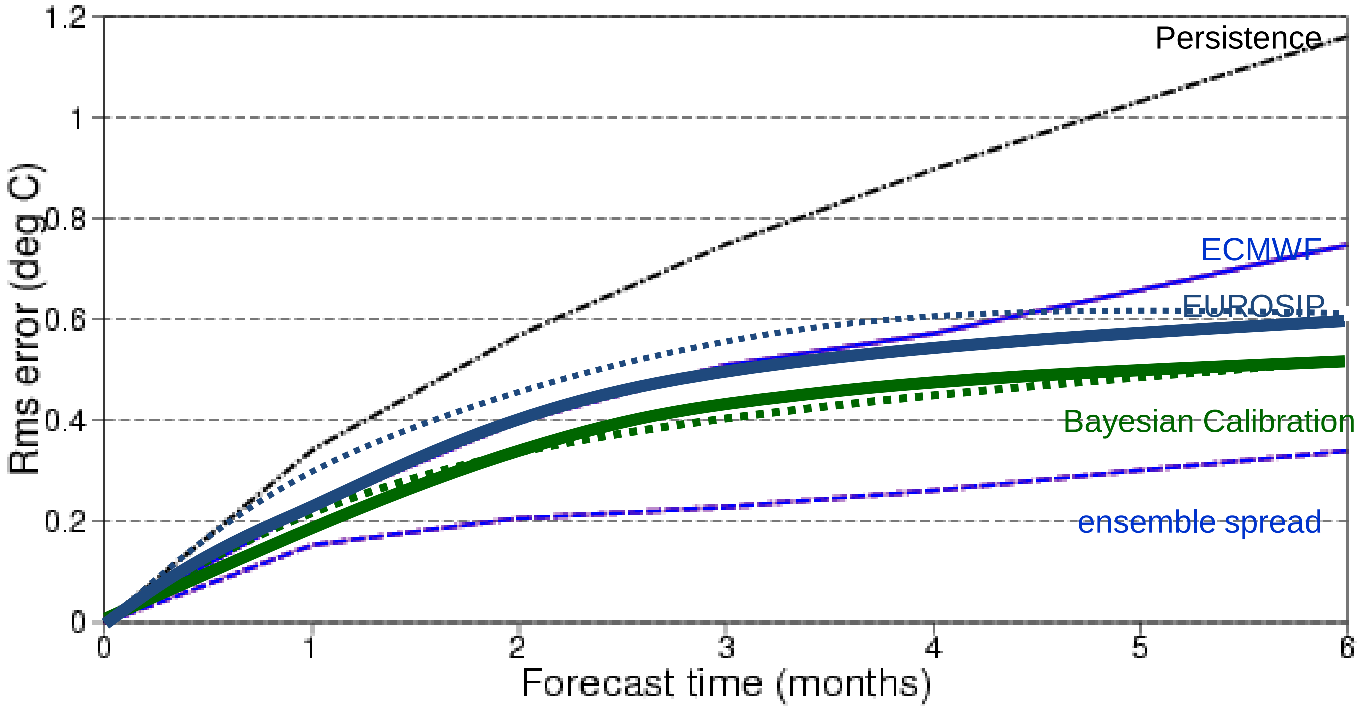
We tried Stochastic Optimals and Synthetic WWB



Improvement but still under dispersive ensemble. Other deficiencies related with too strong negative feedback : Model was not able to produce ENSO (Balmaseda et al 2003). **Judgement: do not implement. It should be sorted out by model improvement**

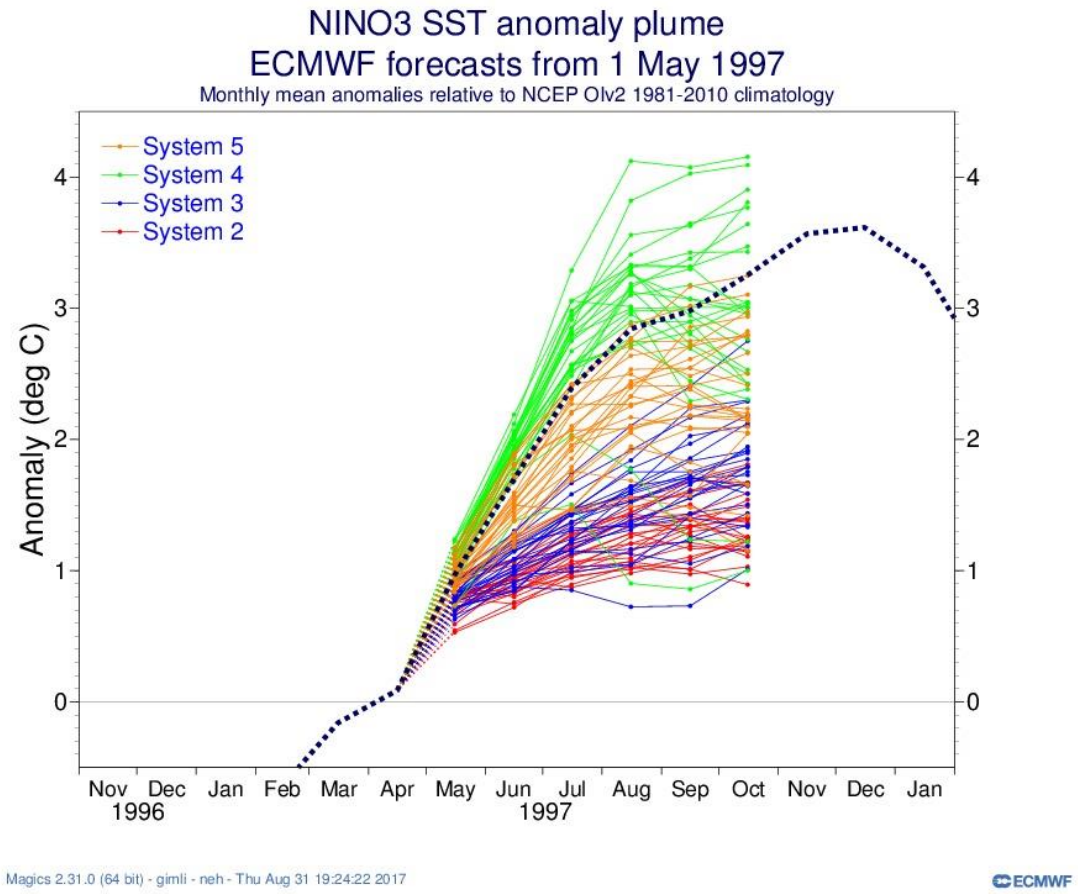
EUROSIP
ECMWF-UKMO-MeteoFrance

RMS error of Nino3 SST anomalies



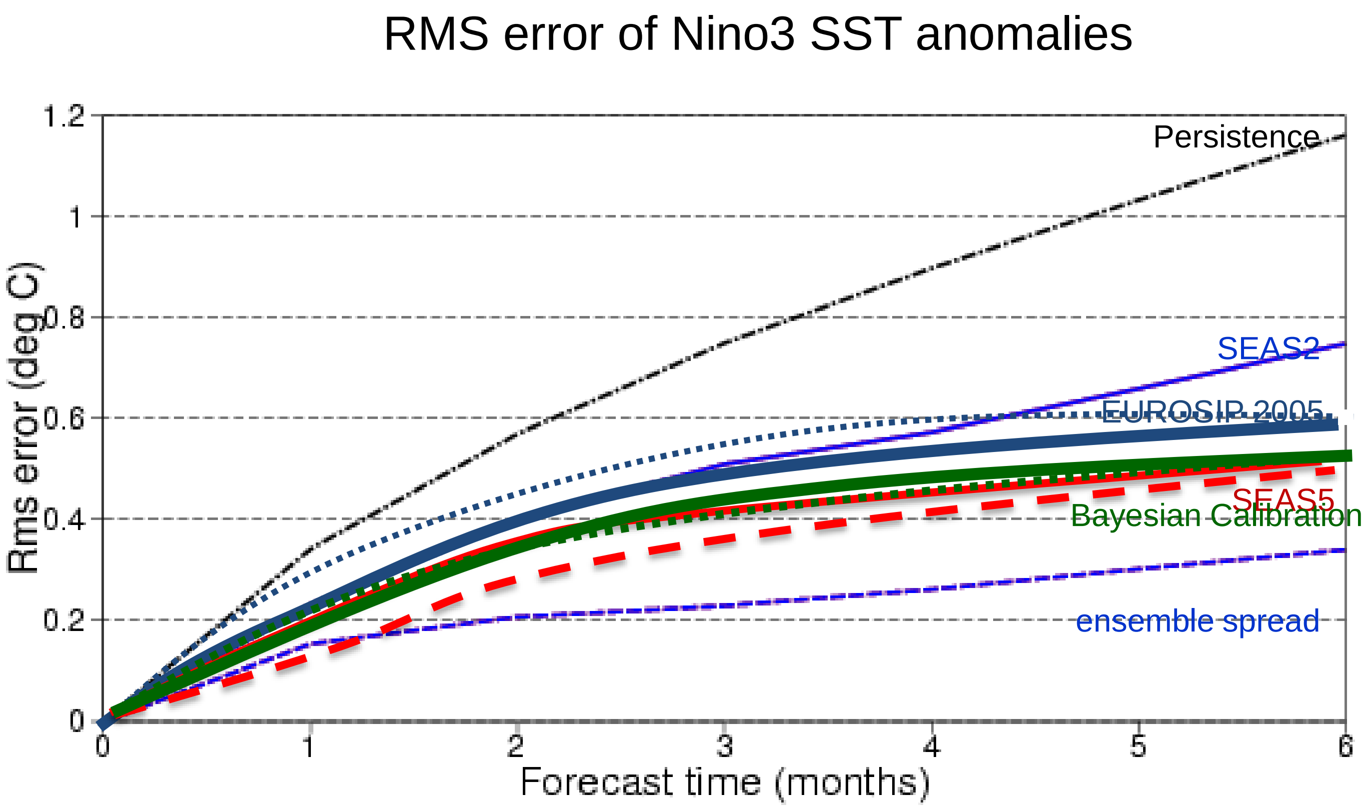
EUROSIP
~2005-2006

Over the years: SEAS2 – SEAS3 –SEAS4 – **SEAS5**



SEAS5 to become operational in Nov 2017

2002 - SEAS2
2017 – SEAS5



The ocean ensemble generation at ECMWF

A simple ensemble generation method with a legacy

1. Ensemble of 5 ocean reanalyses allows to sample uncertainty in the ocean subsurface.
   Re-analysis uncertainty

2. Perturbation package **distributed to the seasonal/decadal community** (EU projects DEMETER, ENSEMBLES). Still used.

3. It allowed going from lag-ensemble to **burst-ensemble**: 51-member ensemble forecast first of each month

4. SST perturbations were later used in the ensemble for the **atmospheric EDA**

5. Still use as a component of the ECMWF coupled ensemble
   Ocean ensemble + Atmospheric SV + Atmospheric EDA + Stochastic Physics

Ocean ensemble in the operational ECMWF coupled forecasting systems

Sampling uncertainty in ocean initial conditions via an ensemble of 5 ocean reanalyses.

2002: ORAS2-SEAS2

Forcing perturbations **v1**. Wind-SST

2006: ORAS3- SEAS3

Forcing perturbations **v2**. Wind-SST

2011: ORAS4-SEAS4

Forcing perturbations **v2:** Wind-SST

+ **Uncertainty in ocean reanalyses spin up**
+ **During forecast: sampling sea-ice recent climatology**

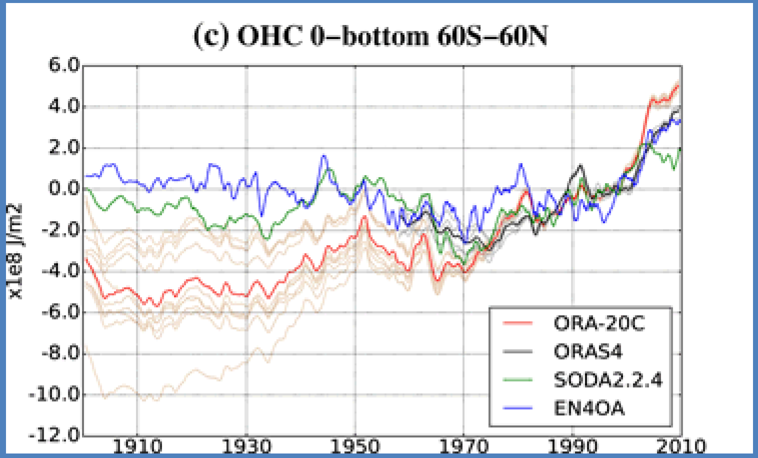
2016: ORAS5-SEAS5

Forcing perturbations **v3**:

Mutivariate Wind-SST-SeaIce-FreshWater-Solar Radiation

+ **Uncertainty in ocean reanalyses spin up updated**
+ **Observation perturbation: representativeness error**

Sampling Spin-Up uncertainty in ocean reanalyses



From de Boisseson et al 2017

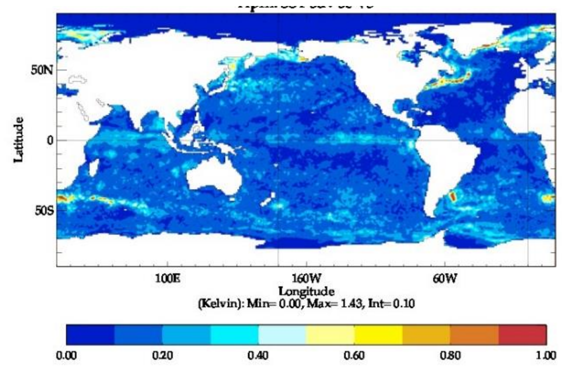
Also used in ORA-20C CERA-20C and CERA-SAT (see later)

V3 Forcing Perturbations in ORAS5

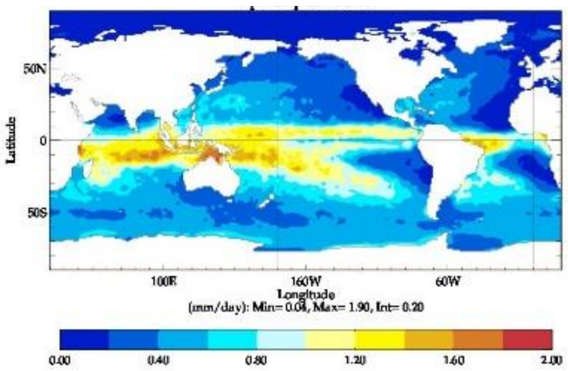
Multivariate - Updated data sets – 2 temporal scales – Multiple uncertainty sources

Still conservative: do not sample error in the mean.

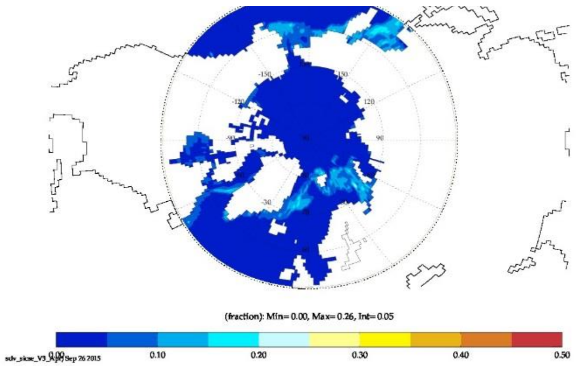
SDV SST SE New (V3)



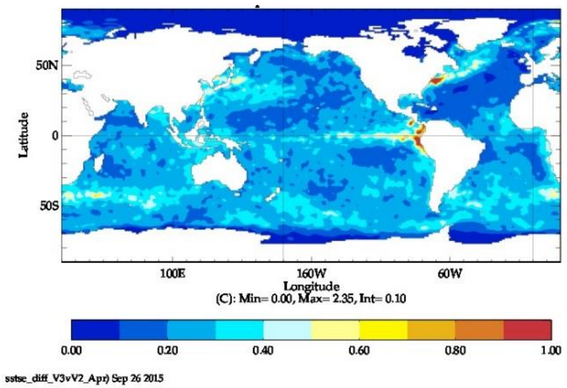
SDV PME



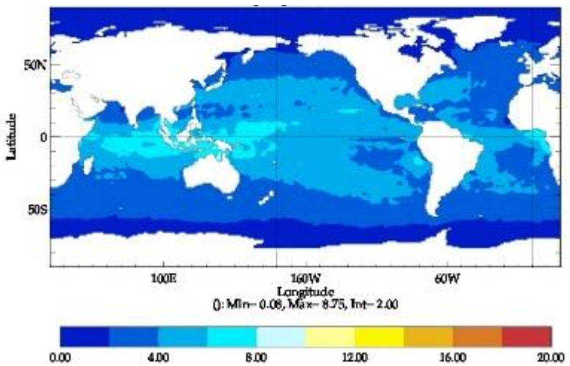
SDV SIC Apr



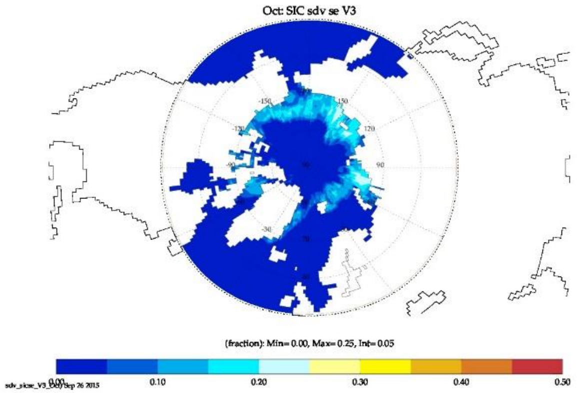
SDV SST SE OLD (V2)



SDV Abs Solar



SDV SIC Oct



Zuo et al 2017, Hirahara et al 2016

Perturbing the Observations

Representativeness error

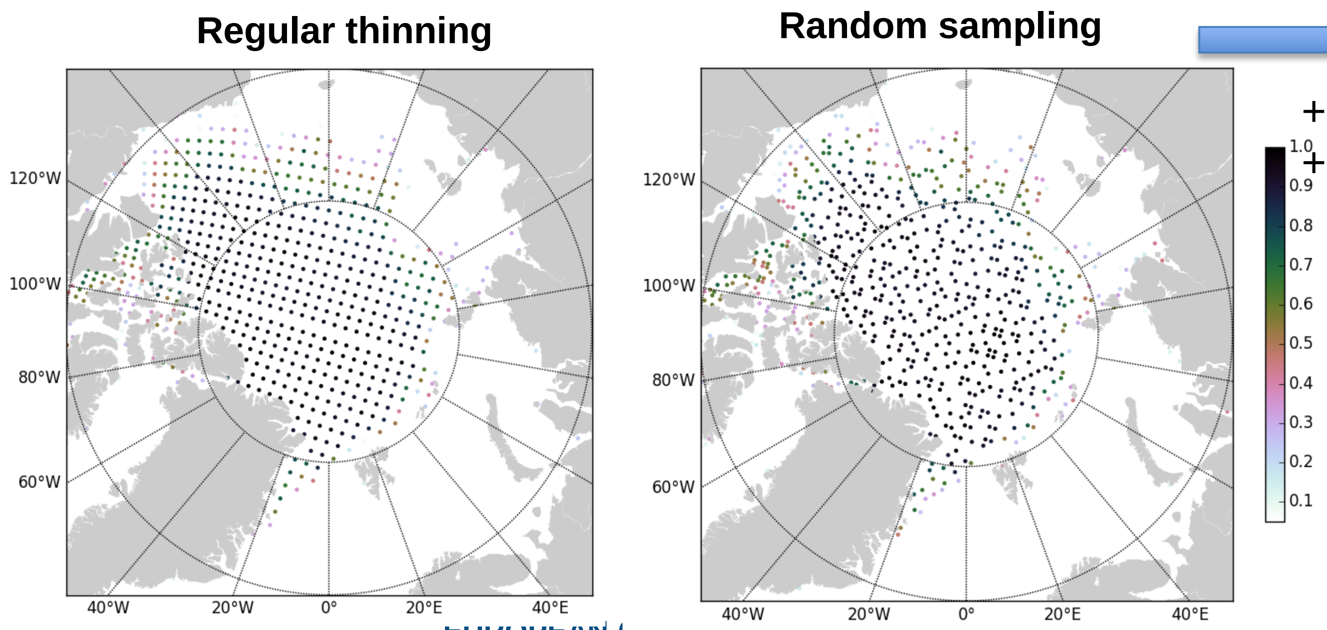
1) Profile displacement and stretching

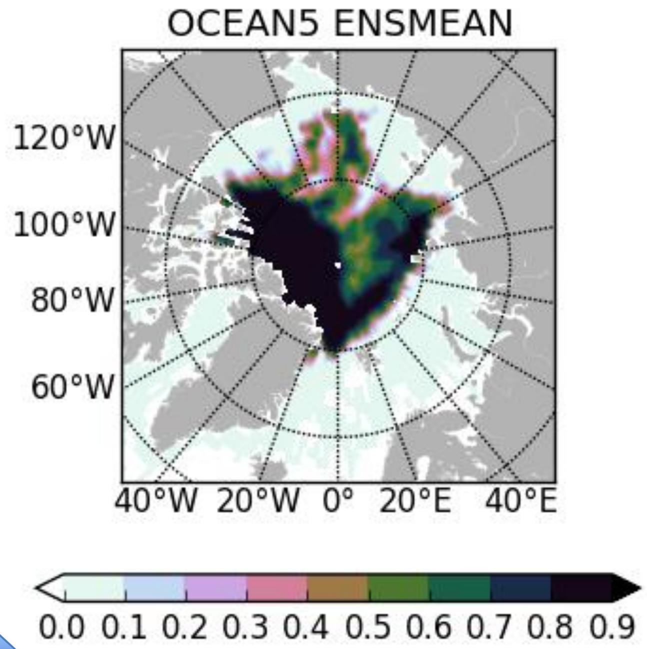
2) Thinning with random seed in different ensemble members:

More observations are used in the ensemble

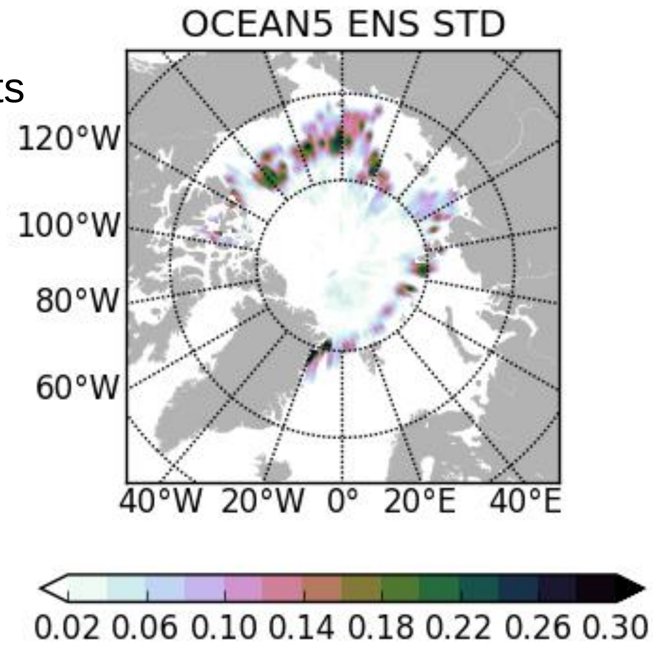
To be used in EDA for the ocean.

Thinning of Sea Ice Concentration Observations



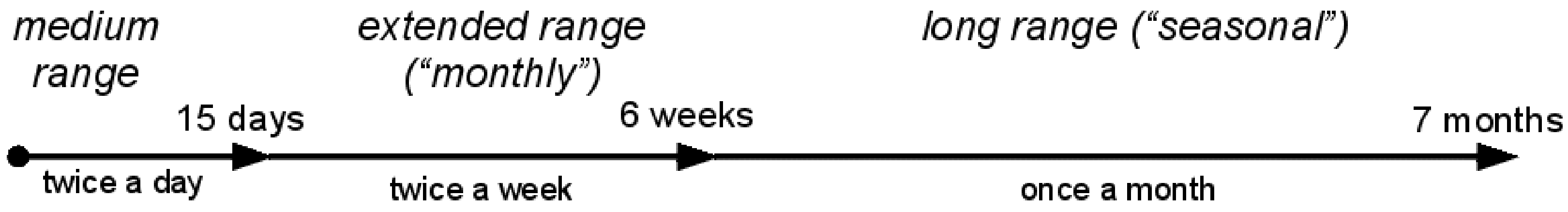


+ forcing perts
+ other obs perts

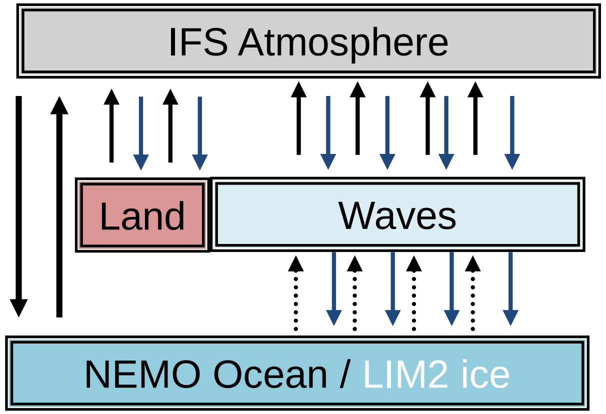


Zuo et al 2017, Tech. Memo 795

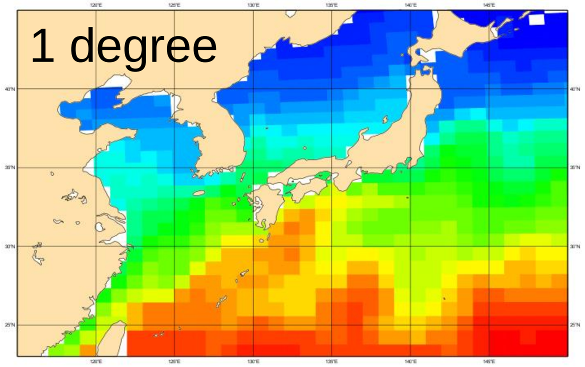
ECMWF Coupled Forecasting System



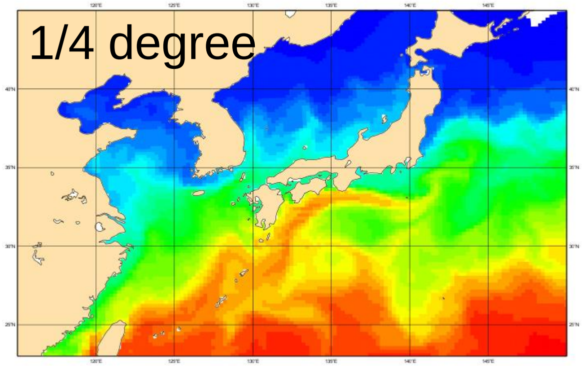
Real Time (re-)analysis of Atmosphere/ Ocean/SeaIce /Land



22 Nov 2016 (Cy43r1): increase in ocean resolution and inclusion of dynamic sea ice (LIM2).



NEMO ORCA1 Z42

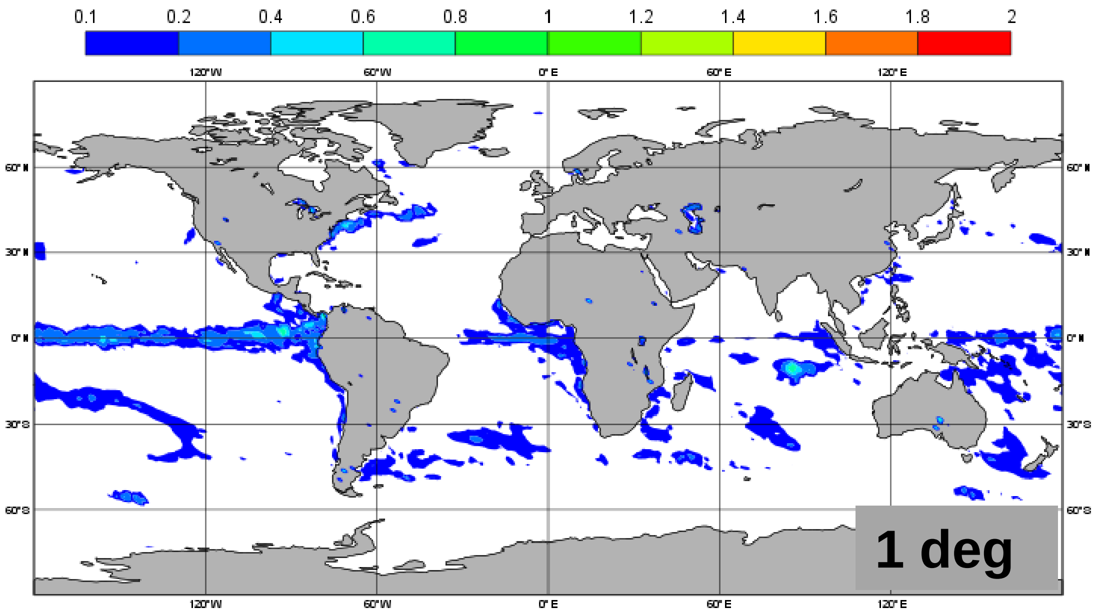


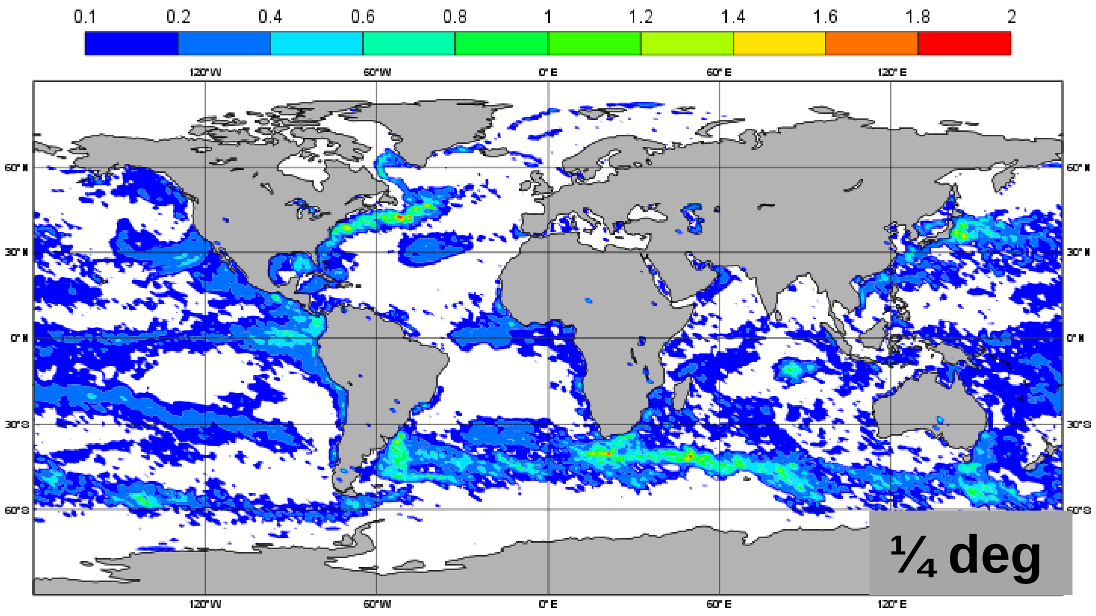
NEMO ORCA1 Z75

Hao Zuo

Ensemble spread in coupled medium-range forecasts
Impact of ocean resolution

10 member, TCO639 realtime system, 1 initial date, SST StDev, fc step 120h





From Simon Lang

Spread appears in areas where ocean intrinsic variability is large + convective regions

Question: does the local air-sea interaction favours/damps instability growth?

=> Can we quantify the chaotic behaviour of the ocean in uncoupled mode?

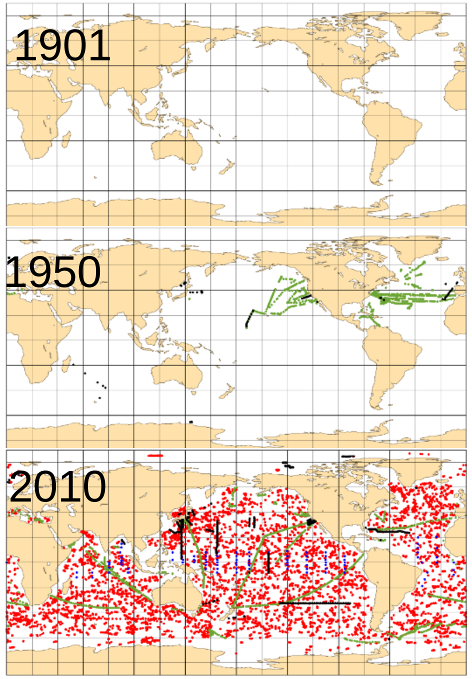
The ECMWF Coupled Reanalyses efforts

ensemble of earth system data assimilation

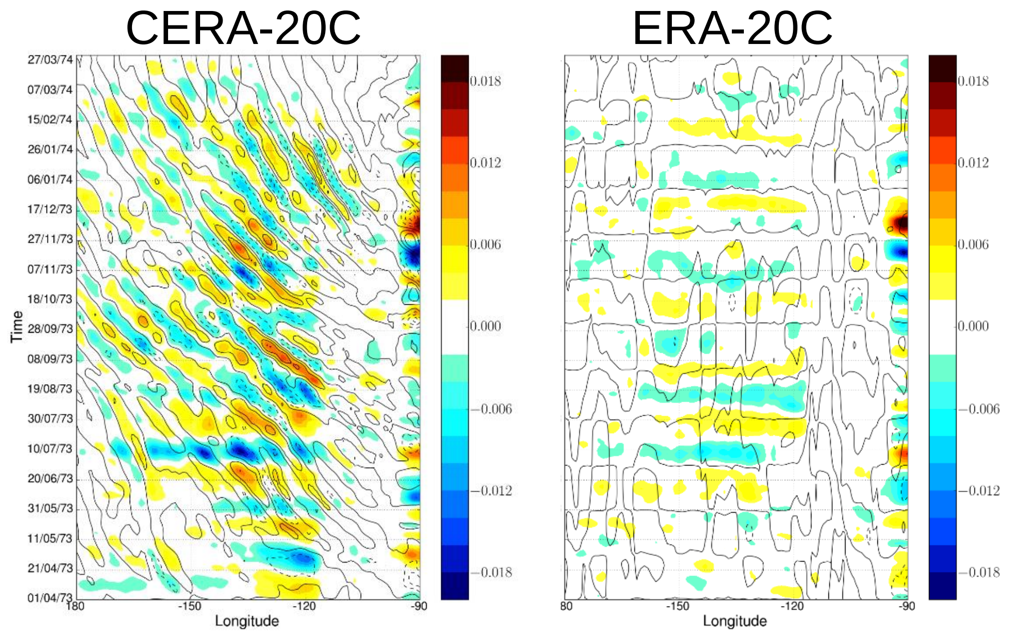
CERA-20C
CERA-SAT

Observation uncertainties + **Model uncertainties** + **Model forcing uncertainties** = **Reanalysis uncertainties**

Ocean temperature obs assimilated in CERA-20C



TIWs in Coupled/Uncoupled reanalysis of the 20th-Century



high-pass filtered SST (colour) and wind stress (contour)

Patrick Laloyaux

Eric de Boisseson

What about the ensemble spread in **coupled** data assimilation?

Compare ensemble spread of CERA-20C with equivalent uncoupled ocean reanalysis.

Uncoupled: Forcing and SST perturbations . By design, only capture only seasonal dependence
Coupled: Spread generated by coupling. SST from HadISST.

same observations, same data assimilation, same observation perturbations

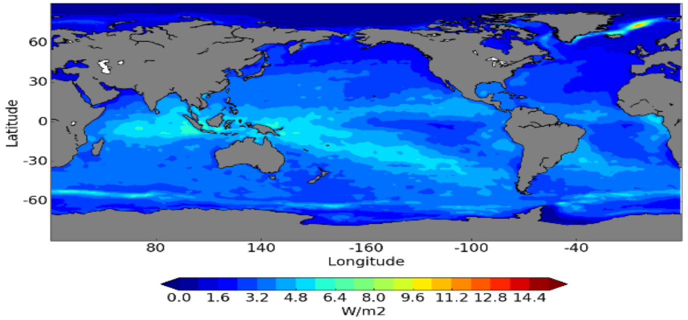
We diagnose the flow dependence of the spread: Decadal, interannual, intraseasonal

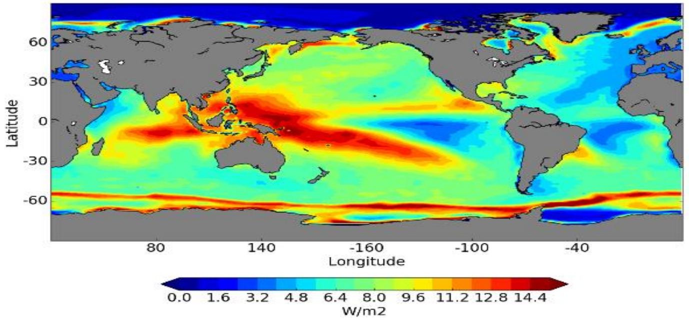
Work in progress

Decadal ..

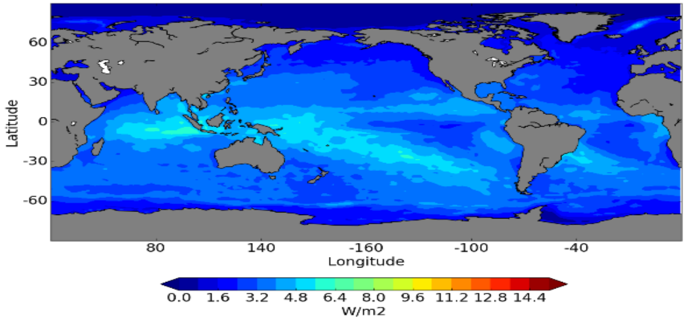
ORA-20C | Solar radiation | CERA-20C

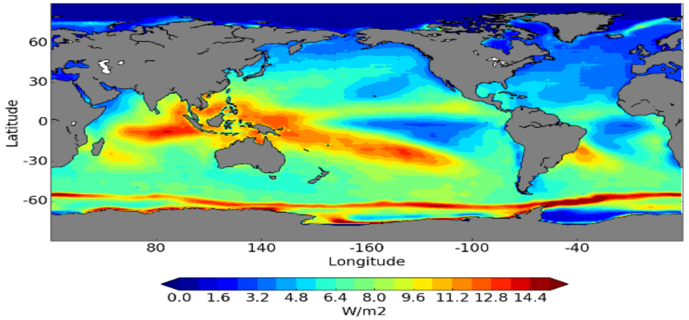
1900s



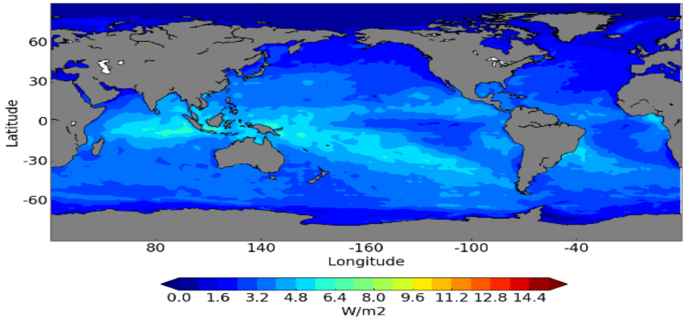


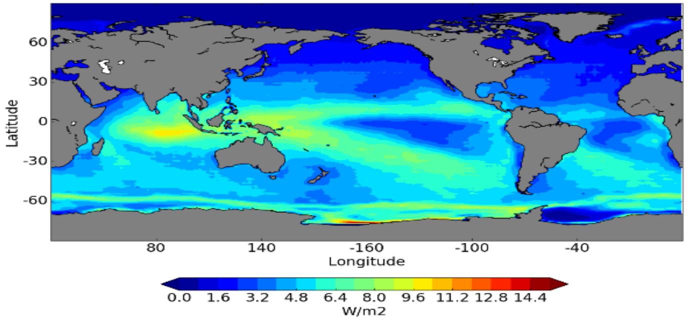
1940s



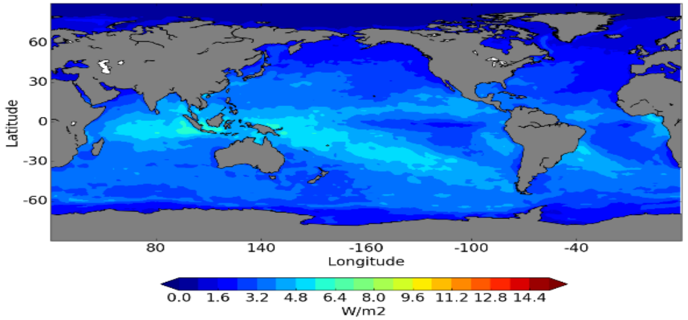


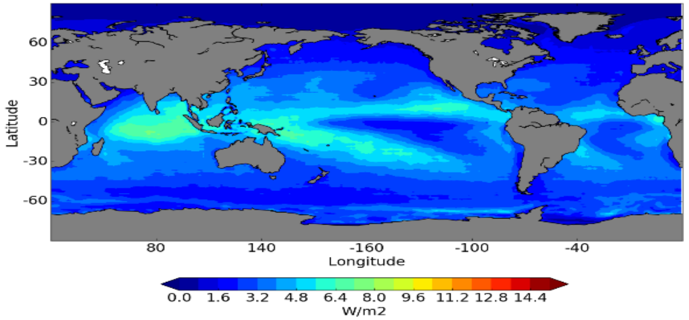
1970s





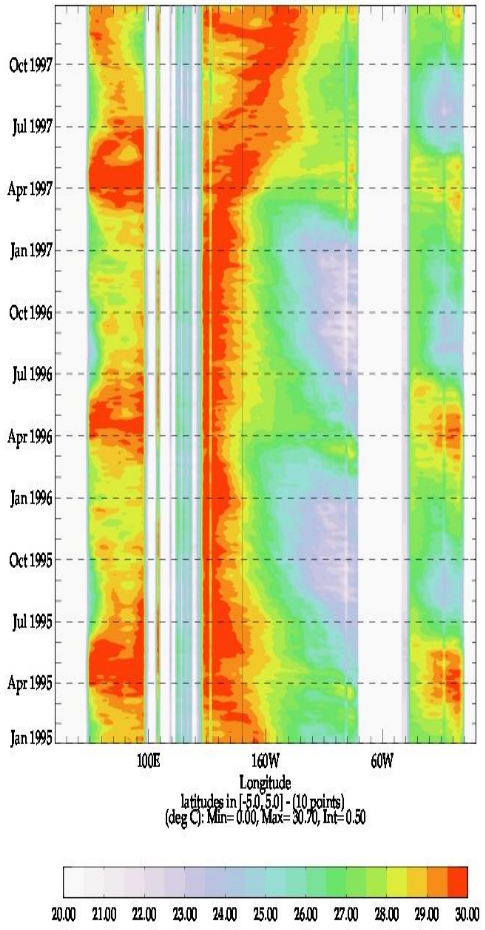
2000s



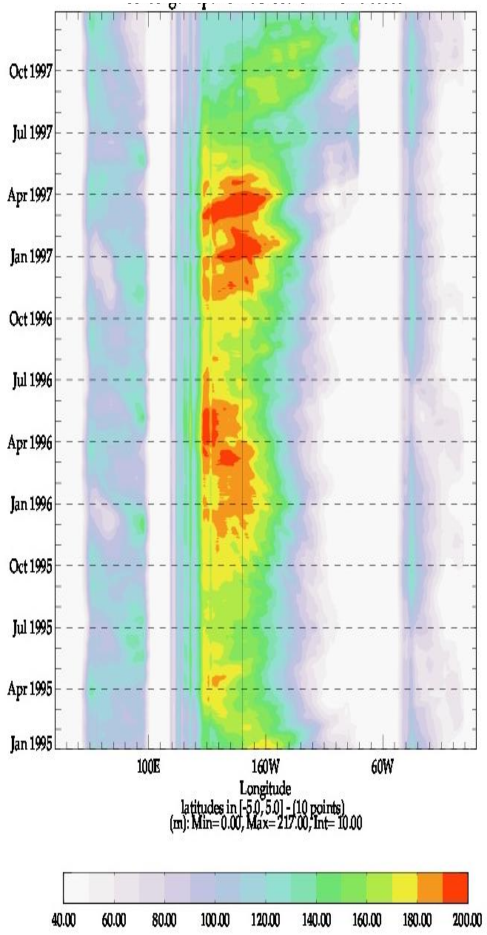


Zoom on 1996-1997: Onset of El Nino
Equatorial daily time series of actual reanalysis fields

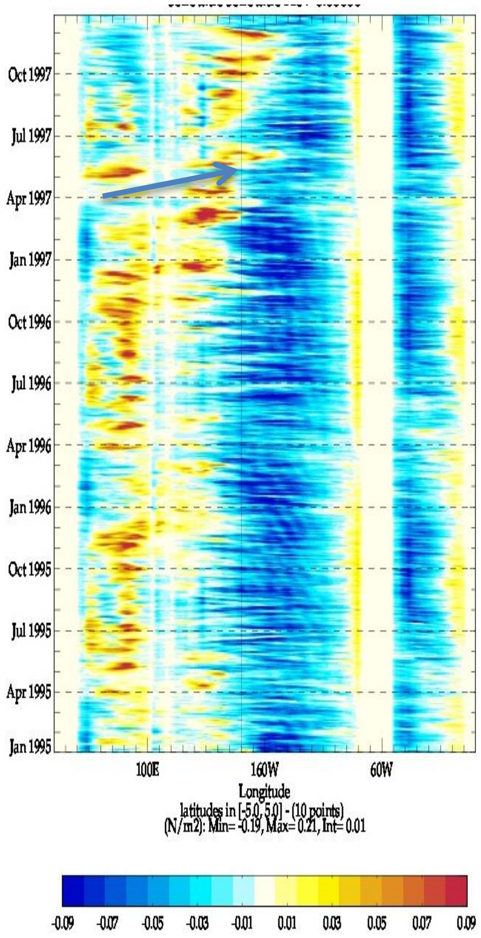
SST



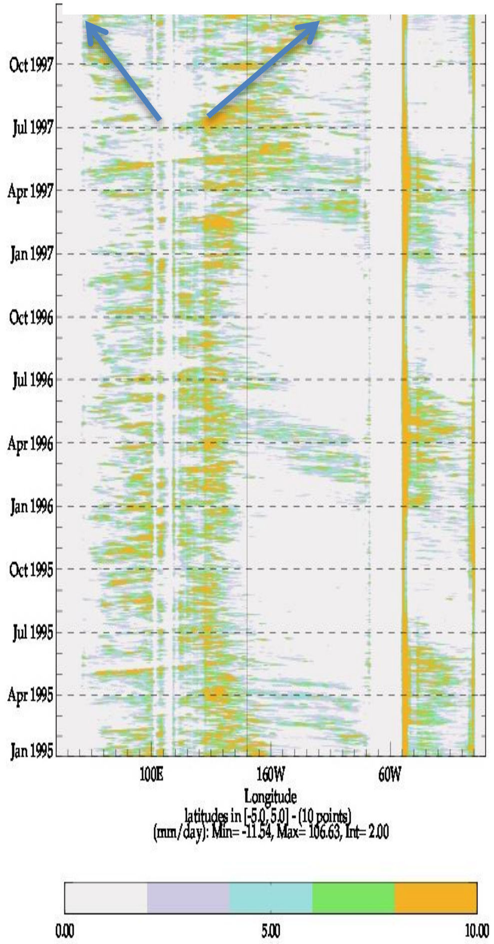
Thermocline depth



Zonal Wind Stress



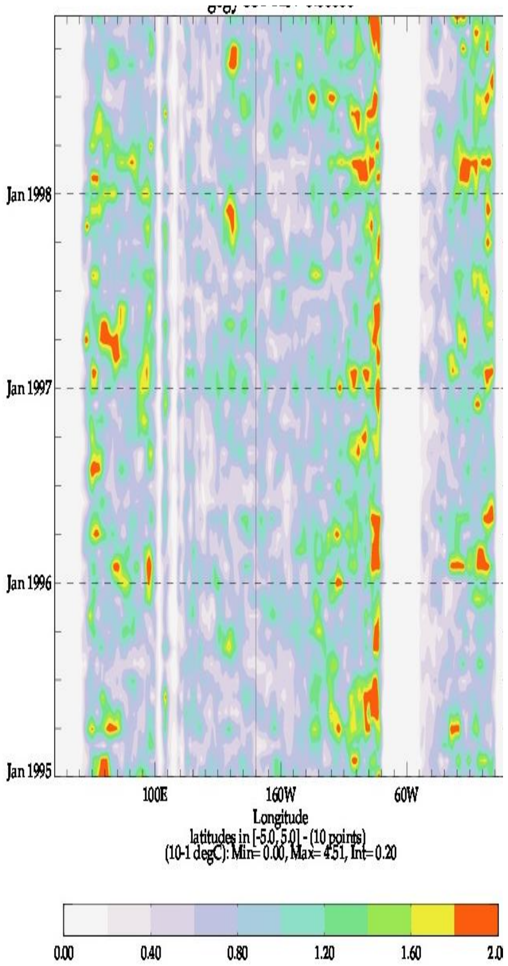
Fresh Water Flux



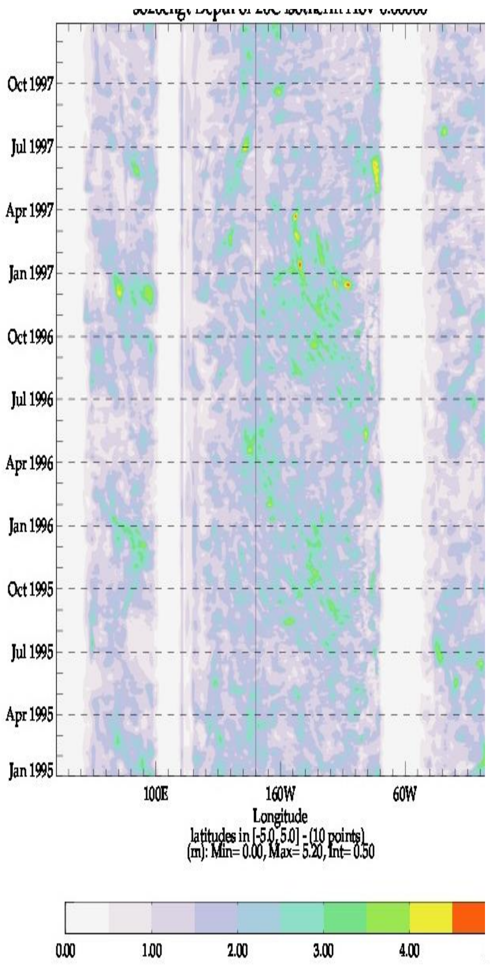
Coherent behaviour among variables SST-Precipation-Wind and thermocline response
Seasonal cycle, intraseasonal variability and onset of El Nino can be appreciated

Zoom on 1996-1997: Onset of El Nino
Equatorial daily time series of UNCOUPLED ensemble spread

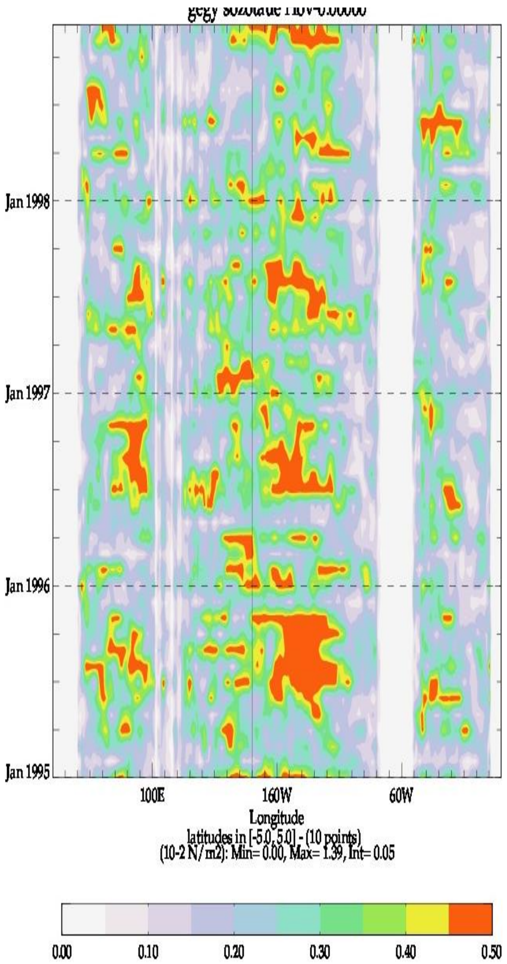
SST



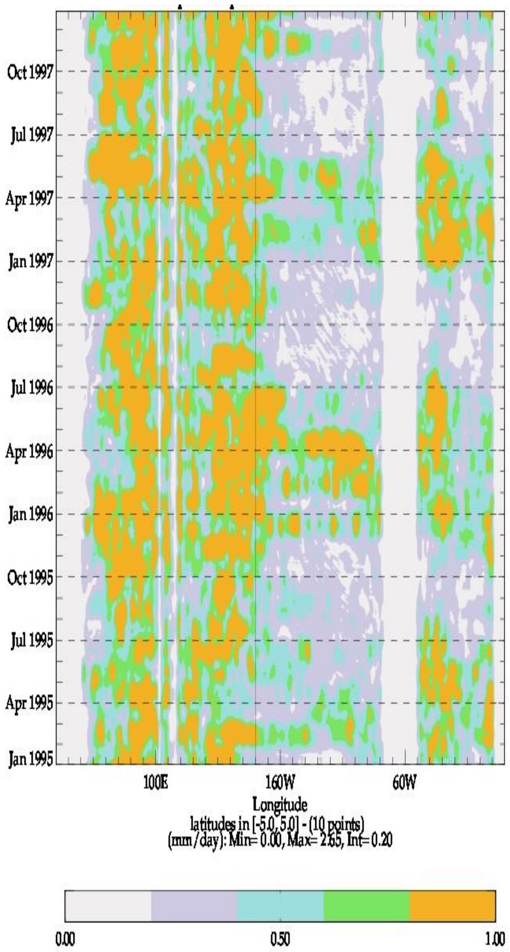
Thermocline depth



Zonal Wind Stress



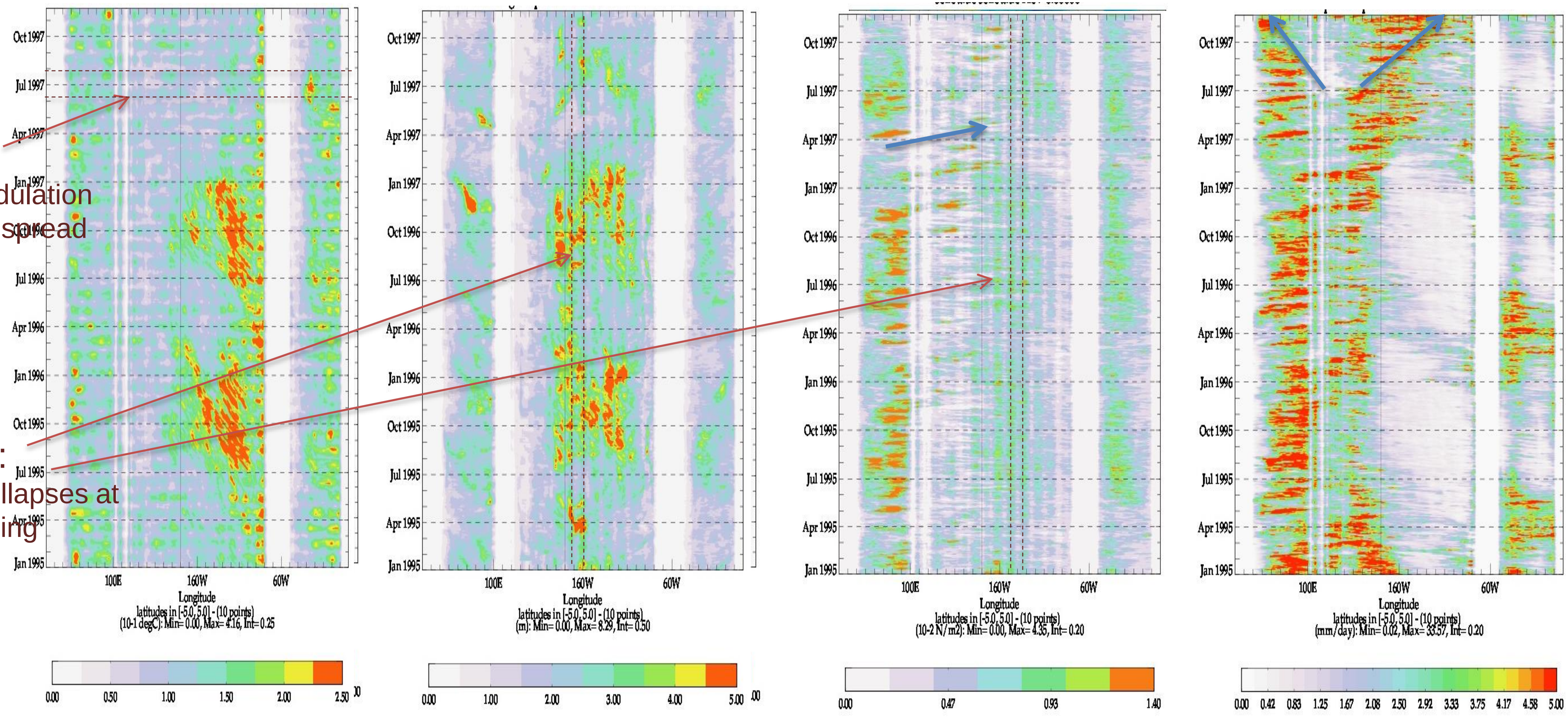
Fresh Water Flux



Coherent spread between ocean and atmopheric variables only at seasonal time scales (by design)
Ocean variables -SST and Thermocline depth- spread show intraseasonal –TIWs- and interannual modulation

Zoom on 1996-1997: Onset of El Nino
Equatorial daily time series of COUPLED ensemble spread

SST | Thermocline depth | Zonal Wind Stress | Fresh Water Flux



Artefact:
Monthly modulation of HadISST spread

Artefact:
Spread collapses at TAO mooring locations

Coherent behaviour among variables SST-Precipation-Wind and thermocline at seasonal-intraseasonal-interannual time scales

Summary and outlook

- **Differences between ocean-atmosphere and marine-weather/climate applications**

- **Ensemble of ocean reanalyses is now common practice in the ocean**.

- **A brief history of the ECWMF ensemble generation in the ocean**
  Evolution Design Opportunity

- **The new v3 ensemble generation used in ORAS5**
  - Random thinning allows using more observations in ensemble methods.
  - A first step towards the EDA in the ocean.

- **A comparison of ensemble in coupled and uncoupled data assimilation**:
  - The CERA system captures decadal – interannual – subseasonal dependence of spread
  - New opportunities for increased reliability of coupled forecasts, especially in tropics

- **Breeding Vectors: Parallel Evolution converging to the same point**
  Hybrid Methods for Coupled Data Assimilation
  Ensemble of coupled forecasts for seamless prediction