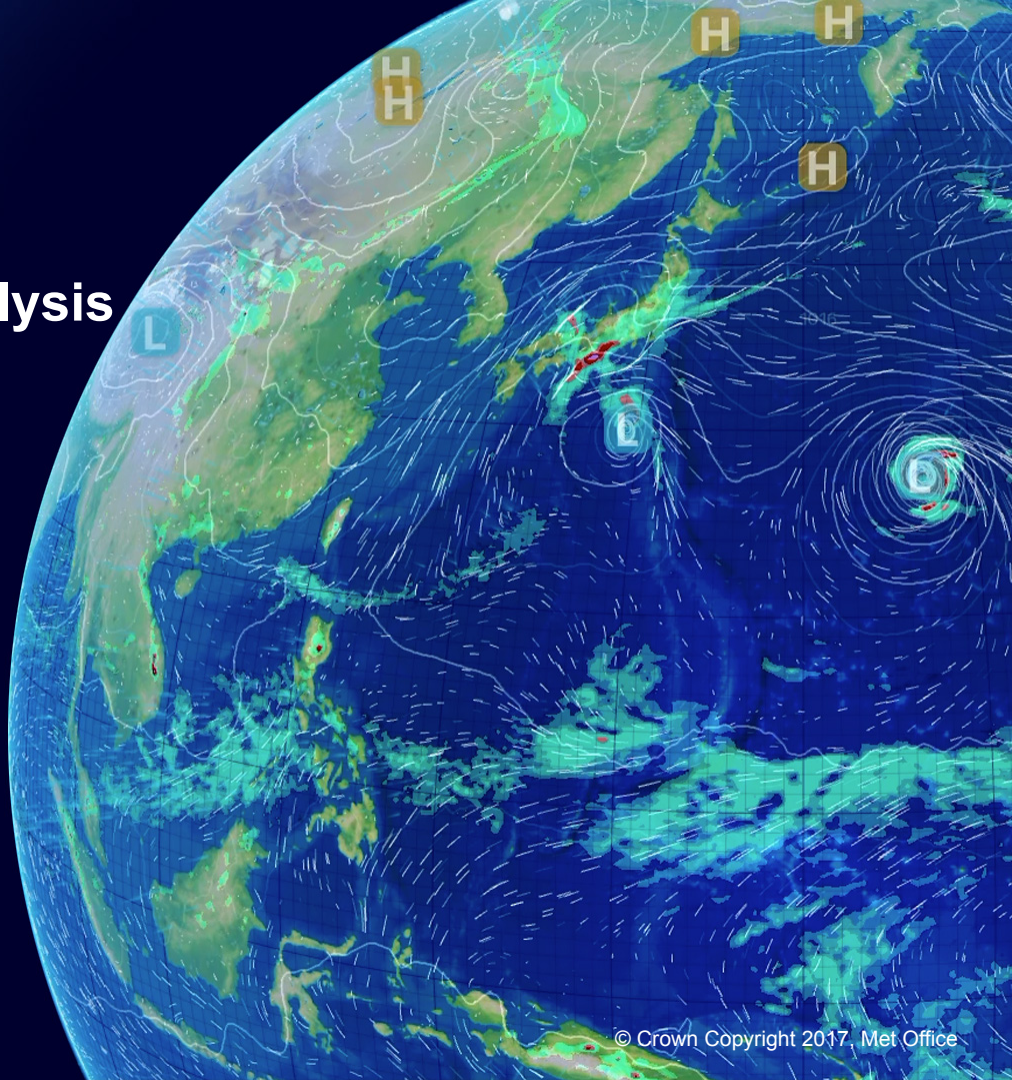
Met Office

Ocean data assimilation for reanalysis



Matt Martin.

ERA-CLIM2 Symposium, University of Bern,

14th December 2017.

Contents

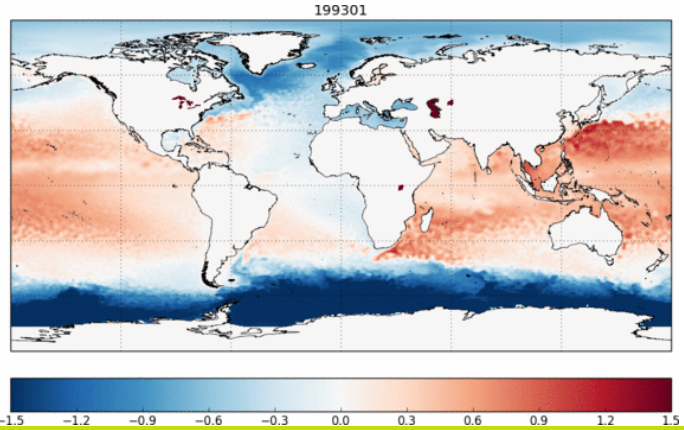
- Introduction.
- On-going developments to improve ocean data assimilation for reanalysis.
- Summary.

Ocean reanalysis

- A reanalysis combines observations with a model to produce a spatially complete representation of past climate. The aim is to produce a ***consistent*** time-series.
- It is possible to produce objective analyses of the available observations without a model, but assimilating data into a model should enable us to get a more complete picture of the climate.

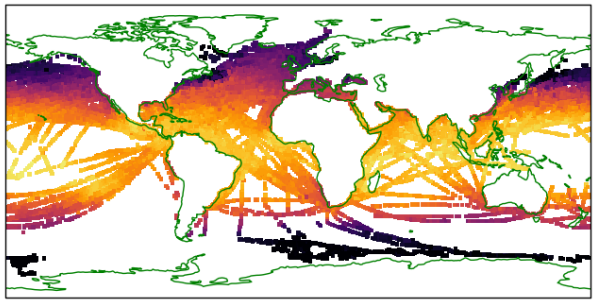
- Models contain variables which are often unobserved. They propagate information spatially, between variables, and in time in a dynamical way consistent with our understanding of the physics.
- Ocean data assimilation methods are constantly being improved for reanalysis applications. They generally are not as advanced as atmospheric systems, but the ERA-CLIM2 project has made good progress to develop various aspects.

Monthly sea surface height from GloSea5 reanalysis (1993-2012)



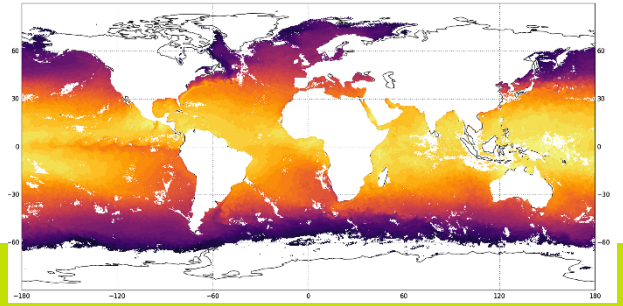
Challenges in producing reanalyses

- Models have biases and so do observations.
- The ocean observing system changes considerably over time, e.g. types of instruments change, sampling changes, variables measured change (SSH only well observed since 1993).
- Some of the challenges for data assimilation for reanalysis are therefore:
  - Deal with the changing observing system to make the most of sparse historical observations as well as the present day observing system.
  - Avoid model and observation biases introducing spurious trends/signals in the time-series.



Left: SST data for one month (Jan 1953)

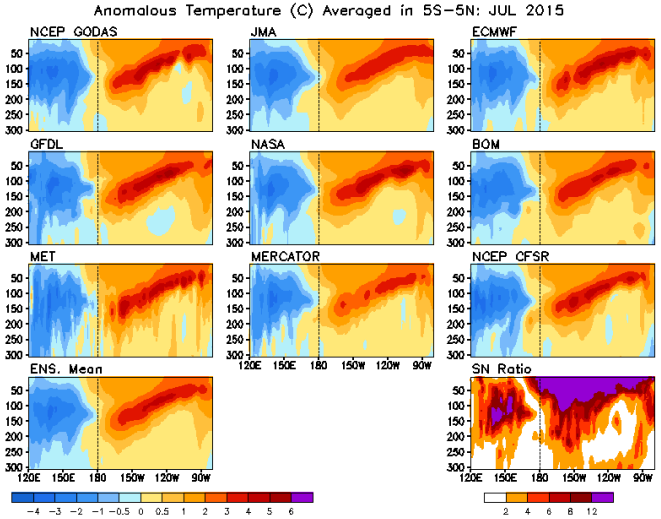
Right: SST data for one day (1st Dec 2017)



Uses of ocean reanalysis information

Some examples of the uses of ocean reanalysis.

1. Climate science – understanding past changes
2. Seasonal forecast calibration
3. Monitoring the current state of the ocean in the context of the climate of an area.
4. Engineering design of structures in the ocean.



Often, reanalyses are produced by groups who are also running operational near-real time forecasting systems, e.g. for seasonal forecasting. These reanalyses need to be made as consistent as possible with the real-time forecasting system.

Century scale reanalyses like CERA-20C take a different approach and assimilate only a consistent sub-set of the observations.

http://origin.cpc.ncep.noaa.gov/products/GODAS/ocean_briefing.shtml

Coordination of ocean reanalysis activities

- International coordination of the assessment of ocean reanalysis was carried out in the Ocean Reanalysis Inter-comparison Project (ORA-IP) and it's follow on. See Balmaseda et al. 2015.
- In Europe there are two main coordinated activities producing ocean reanalysis:

1. The Copernicus Marine Environment Monitoring Service (CMEMS) produces global and regional ***ocean*** reanalyses. A multi-system ensemble of global ocean reanalyses has been produced for the satellite era (1993-present):

Institute	Data assimilation scheme	Time-window
ECMWF	3DVar (NEMOVAR)	5-days
Met Office	3DVar (NEMOVAR)	1-day
CMCC	3DVar (OceanVar)	7-days
Mercator-Ocean	SEEK filter	7-days

All systems use ¼° global NEMO using ERA-Interim atmos forcing with similar observations assimilated.

2. ERA-CLIM2 develops ***coupled*** reanalyses for the 20th Century (CERA-20C) and the satellite era (CERA-SAT) using a similar scheme in the ocean to the one used to produce the ECMWF CMEMS reanalysis.

Data assimilation in ERA-CLIM2

- NEMOVAR scheme is a 3D-Var scheme which minimises:

$$J = \left(\mathbf{x} - \mathbf{x}^f\right)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^f) + (\mathbf{y} - h(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - h(\mathbf{x}))$$

- A short model forecast $\mathbf{x}^f$ is combined with the available observations **y** in a cost function which weights each component by its errors **(B** and **R).**
- The forecast (background) error covariances allow the scheme to spread information from the observation locations onto the model grid: horizontally, vertically and between variables.
- In CERA-20C, **y** includes only temperature and salinity profiles.
- In CERA-SAT, **y** includes T/S profiles, satellite altimeter SSH data, satellite SST data and sea-ice concentration data.
- The scheme produces increments (changes to the model) to the 3D temperature, salinity and velocity fields, the sea surface height, and the sea-ice concentration.

Data assimilation in ERA-CLIM2

- In ERA-CLIM2 various developments have been made to the NEMOVAR ocean DA system to improve future reanalysis (the next CERA-20C and CERA-SAT).
- Atmospheric component of CERA uses 4D-Var scheme which includes information from ensembles in a hybrid scheme. Aim to improve the ocean scheme to bring it in-line with the atmospheric one, and to improve aspects important for coupled reanalysis.
- Focus here on three main areas:

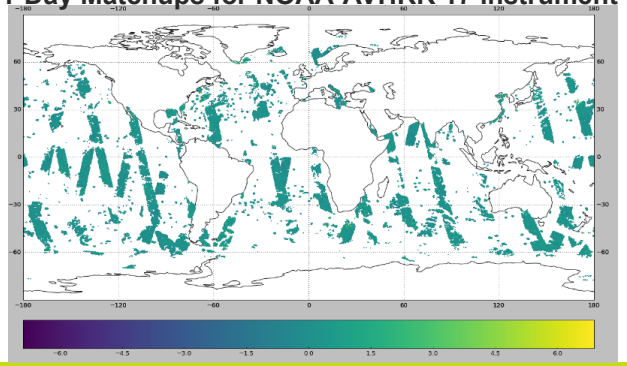
1. Improving the assimilation of SST data – particularly important in the context of coupled reanalysis.
2. Improving the representation of the background error covariances:
   - Making better use of sparse historical data through EOFs.
   - Using ensemble information in the data assimilation.
3. Testing potential of 4D-Var.

Assimilation of SST data

- SST important particularly in the context of coupled climate reanalysis because of the interactions with the atmosphere.
- In Met Office implementation of NEMOVAR, vertical propagation of SST data is based on the background mixed-layer depth, so SST assimilation has significant impact on the 3D ocean state.
- Satellite SST data contain biases due to various reasons, particularly because of errors in the atmospheric state used in the retrieval of the SST from the measured brightness temperatures.
- We want to avoid spurious trends in the reanalysis from the introduction of new instruments.

- We have developed an improved satellite SST bias correction which is based on a variational bias correction scheme.
- This includes the use of reference un-biased observations (some in situ data and ATSR-2/AATSR satellite) to improve the bias estimate and mitigate against possible model biases.

1-Day Matchups for NOAA-AVHRR 17 instrument



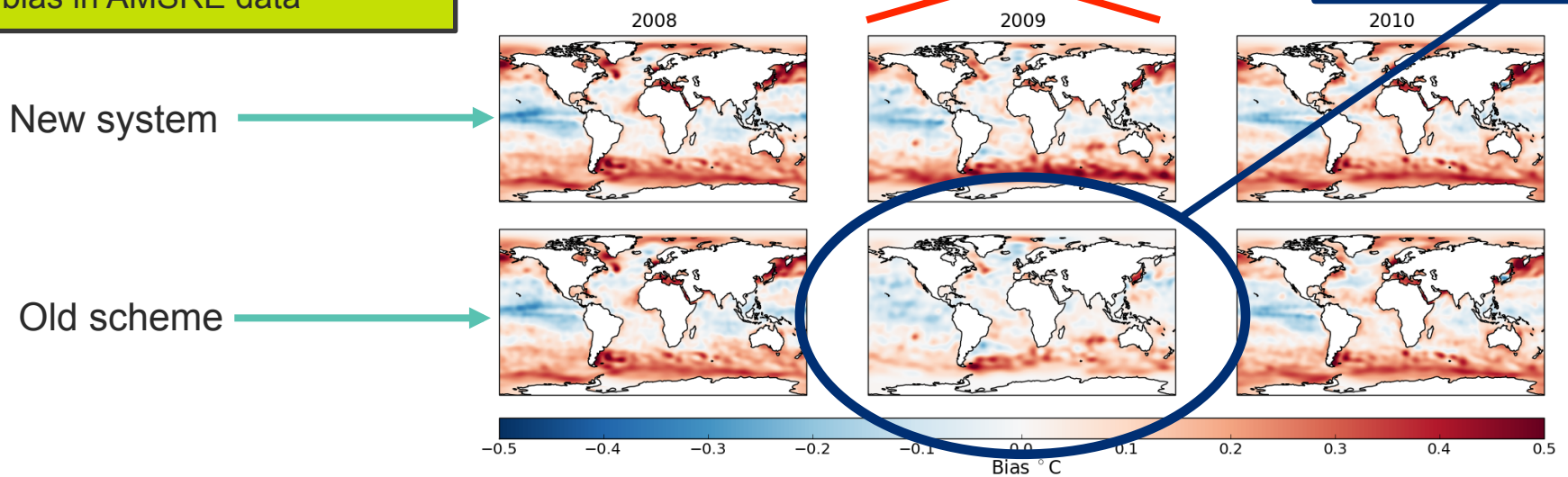
SST Bias correction

- A reanalysis experiment was run for 3 years, where the main reference satellite was removed in the middle year, to assess the sensitivity of various bias correction schemes to changes in the observing system. This aims to mimic what happens during periods when reference satellites are not available.
- Assimilates satellite and in situ SST data, as well as profiles of T/S, altimeter SSH data and sea-ice concentration data.

Average estimates of the bias in AMSRE data

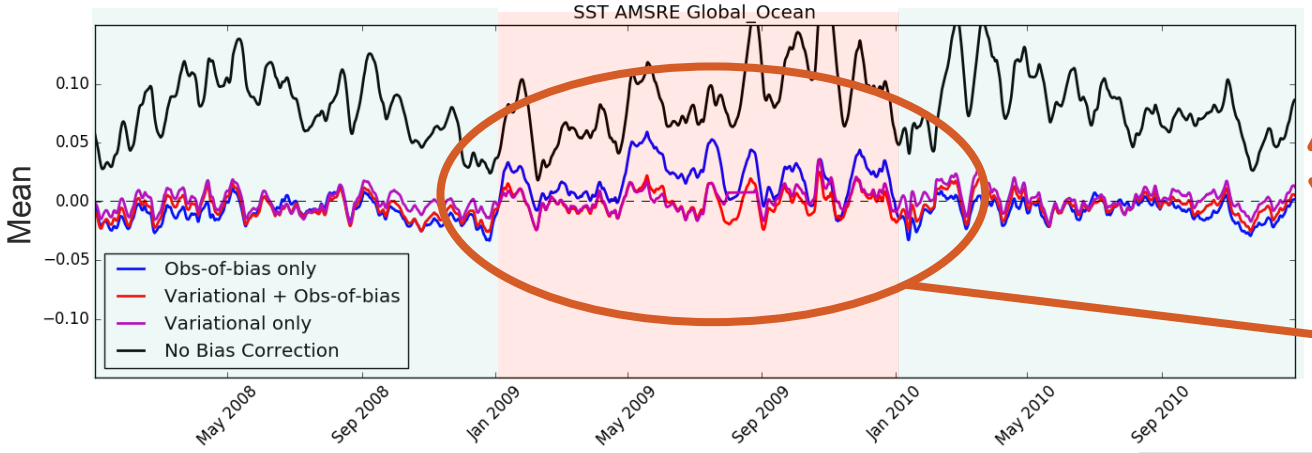
In 2009 we used many fewer reference observations

Bias estimate using the old system is inconsistent with the other years



SST Bias correction

- Plot shows mean 1-day forecast errors compared to one of the SST satellite data sources (AMSRE – a microwave instrument).
- No bias correction leads to significant biases ~ 0.1K.
- Various bias correction schemes tested.
- All reduce the bias, but the two which include a variational bias correction scheme reduce the impact of the lower number of reference data in the middle year.



The overall bias is much reduced

In the period with fewer reference observations, the old system (blue line) does not do as well

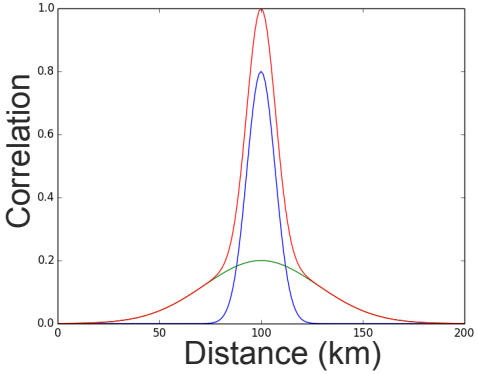
AATSR Data used as reference

AATSR Data not used as reference,

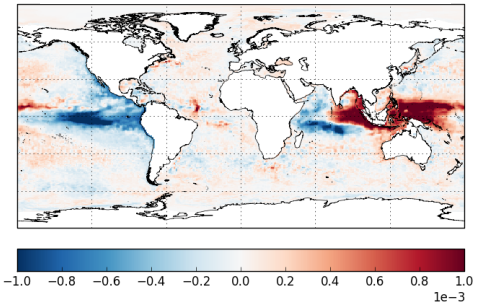
The plots show the difference between AMSRE data and a 1 day forecast of the model.

Improving the background error covariances
Spatial propagation

- Standard modern-day ocean data assimilation systems use local Gaussian functions to spread information spatially.
- NEMOVAR has the ability to combine multiple Gaussian functions with different length-scales. However, this only spreads information nearby the observation.
- Aim to make more use of sparse historical ocean observations by spreading information with Empirical Orthogonal Functions (EOFs).
- The EOF scheme has been implemented and combined with the standard Gaussian model so an observation can have a local impact as well as a global impact.

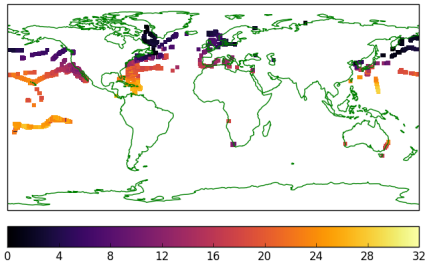


1st EOF of T at 97m depth

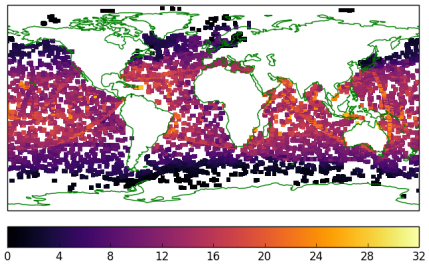


Sub-surface temperature profiles from HadIOD.

Jan 1953



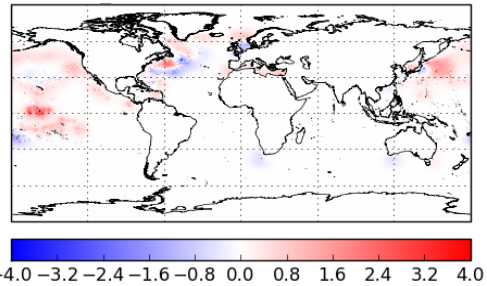
Jan 2010



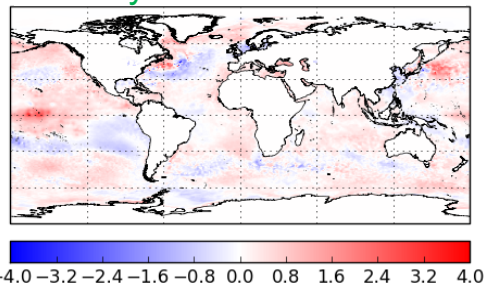
EOF data assimilation

SST increments /ºC from Jan 1953 profiles

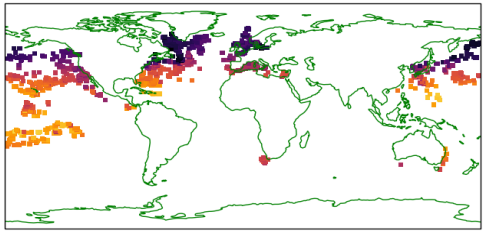
Standard DA



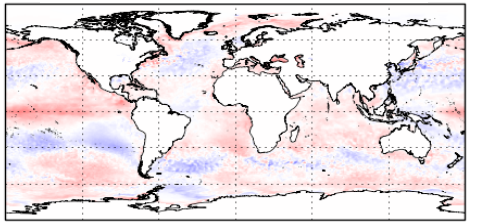
Hybrid 10% EOF



Observations



EOF DA

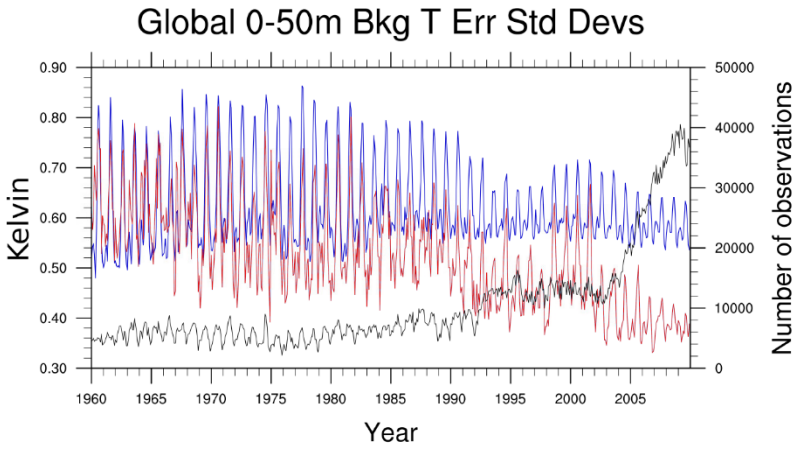


- EOFs generated from 100-year coupled climate model (HadGEM3).
- Hybrid EOF/standard DA spreads information globally, while retaining the structures from the standard DA near the observations

Improving the background error covariances Using ensembles

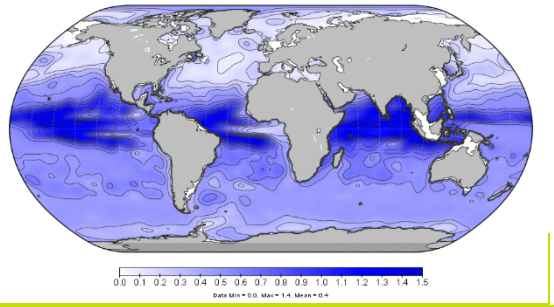
- Background errors change in time: as the observing system becomes more complete, the model forecast is gradually improved.
- Specified errors for temperature in current reanalysis vary based on parameterisations, but they do not properly represent the changes to the background errors. Either temporally or spatially.
- Example shows time-series of near surface ocean temperature forecast error standard deviations from ECMWF reanalysis system. Diagnosed errors gradually reduce over time.
- One way of representing these changes is by running an ensemble of reanalyses and using the ensemble spread as a measure of uncertainty.

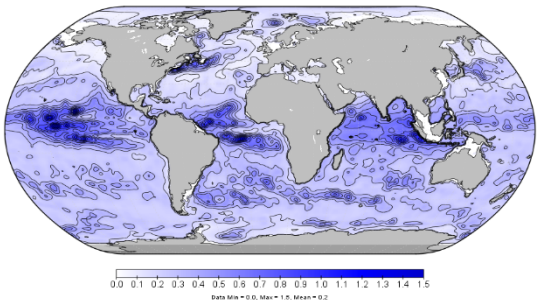
- Blue line shows the specified errors.
- Red line shows diagnosed errors a posteriori.
- Black line shows the number of observations.



Improving the background error covariances Using ensembles

- Running an ensemble of reanalyses provides uncertainty information to users, and the information can also be used to improve the data assimilation.
- However, it is very expensive computationally to run a large ensemble, so only limited numbers are used.
- The sampling errors associated with these small number of ensembles makes it difficult to use the information directly.
- Instead, hybrid ensemble-variational methods combine the ensemble information with existing parametrized error covariance information.
- Plots show parameterized and hybrid error estimates at 100m depth for temperature: ensemble improves error estimates in western boundary current regions particularly.



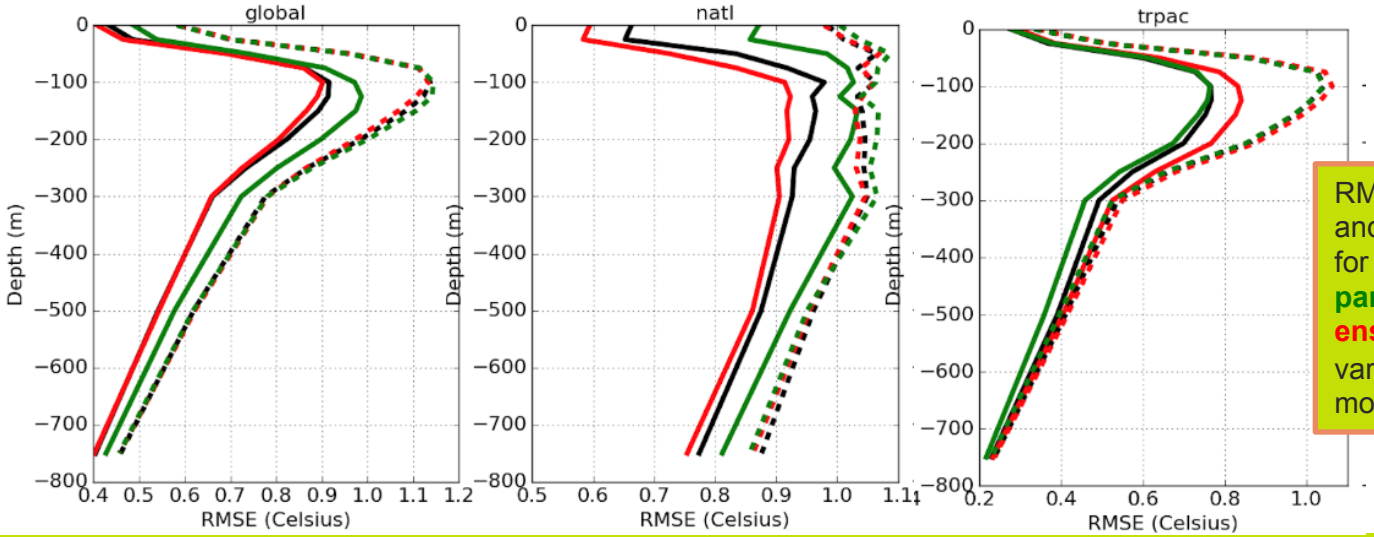


Example of parameterized and hybrid **temperature error standard deviations** at 100m, estimated from the ECMWF 11-member ensemble of ocean reanalyses.

Improving the background error covariances Using ensembles

Preliminary experiments testing ensemble and hybrid (parameterized + ensemble) variances show positive results compared to parameterized-alone variances.

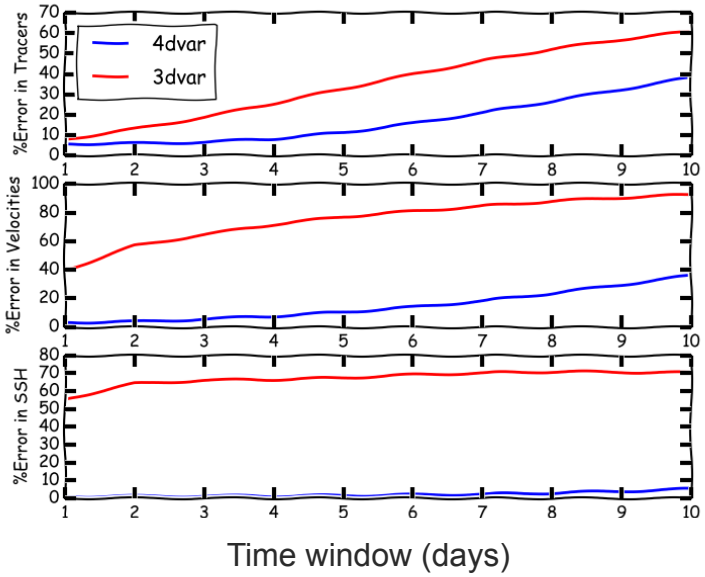
Plots show temperature errors compared to profile observations as a function of depth in different regions.



RMS of Obs-Ana (solid curves) and Obs-Bkg (dashed curves) for temperature with **parameterized-only, ensemble-only and hybrid** variances in a 1/4° global model.

Testing impact of 4DVar

- 4DVar uses a linearised version of the model to propagate observation information in order to produce increments which will allow the full non-linear model to fit the observations.
- With very short time-window, 4DVar is similar to 3DVar. There is the question of what time-window will allow 4D-Var to give improvements in the ocean.



- Experiments in ¼° NEMO.
- For short time-windows (1 day), 4DVar doesn't give much improvement for temperature/salinity, but does improve velocities and SSH.
- For longer time-windows, there is more impact for temperature and salinity.
- 4DVar expensive though, so developments have been made to enable it to be run more efficiently (resolution of inner loop, simplified linear model)

Summary

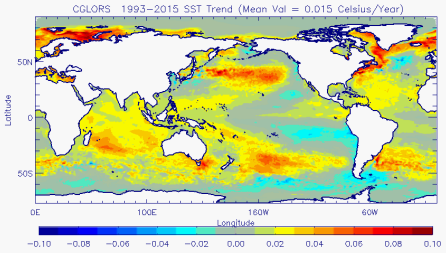
- CERA-20C and CERA-SAT reanalyses use the NEMO model and the NEMOVAR data assimilation in their ocean component.
- Many developments being made to NEMOVAR as part of ERA-CLIM2 to improve its use for reanalysis.
    - Assimilation of SST data is being improved by reducing spurious signals from satellite data-sets through observation bias correction.
    - EOF data assimilation allows us to spread information from sparse historical data-sets.
    - Hybrid variational/ensemble assimilation allows us to improve the representation of spatial and temporal changes in the error covariances which are crucial for making a good reanalysis.
    - 4DVar has the potential to improve the quality of the analysis.
- We aim to include these various developments in future versions of reanalyses produced using the CERA system.

Met Office

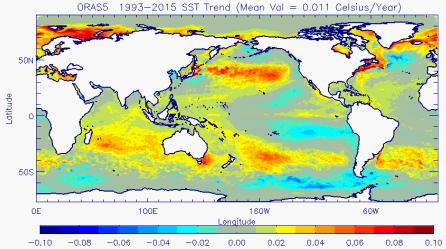
Thankyou for listening

SST Trends (1993-2015) in different reanalysis products

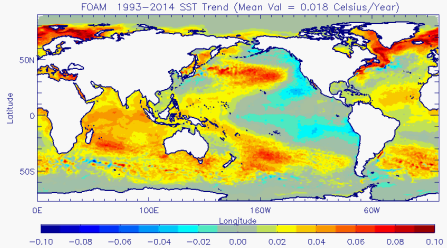
CMCC



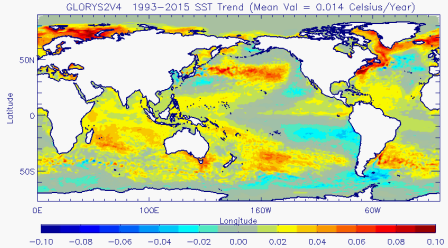
ECMWF

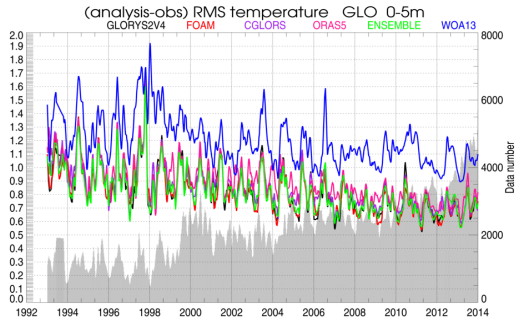


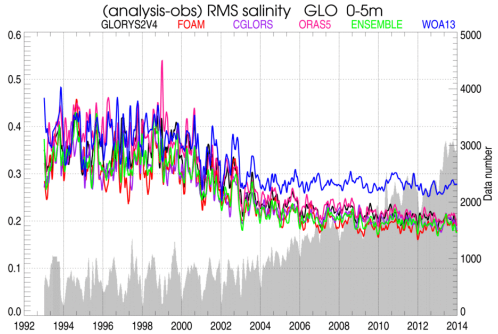
Met Office



Mercator

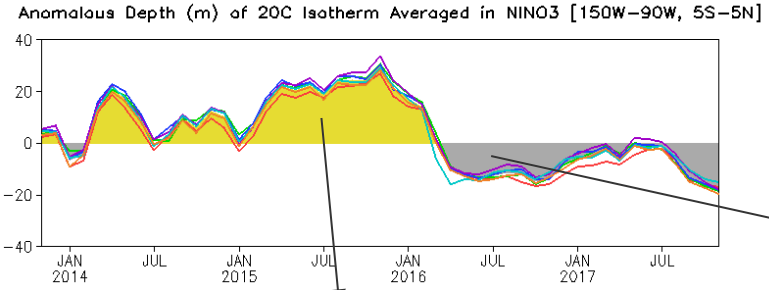


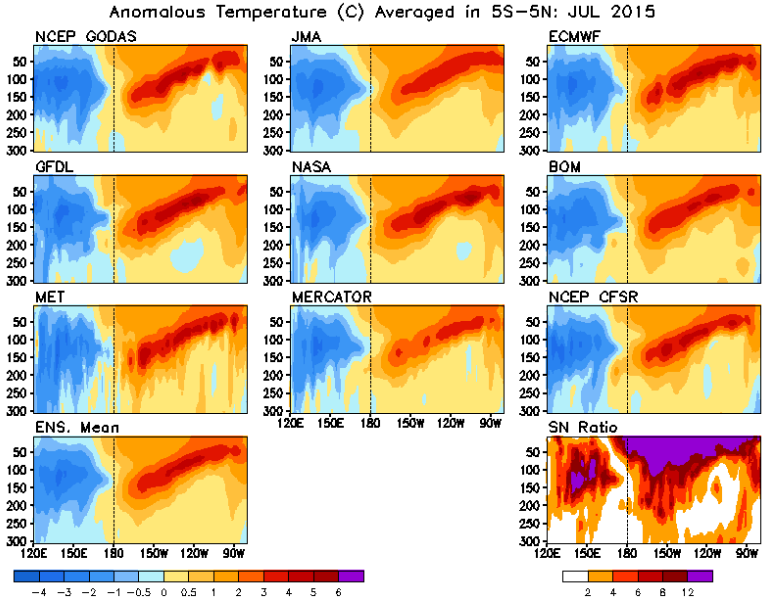


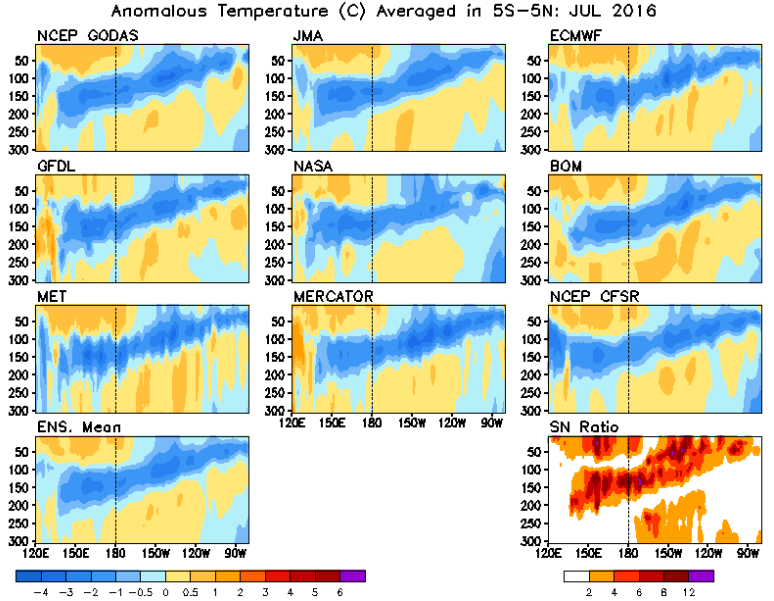


Introduction

Real-time climate monitoring







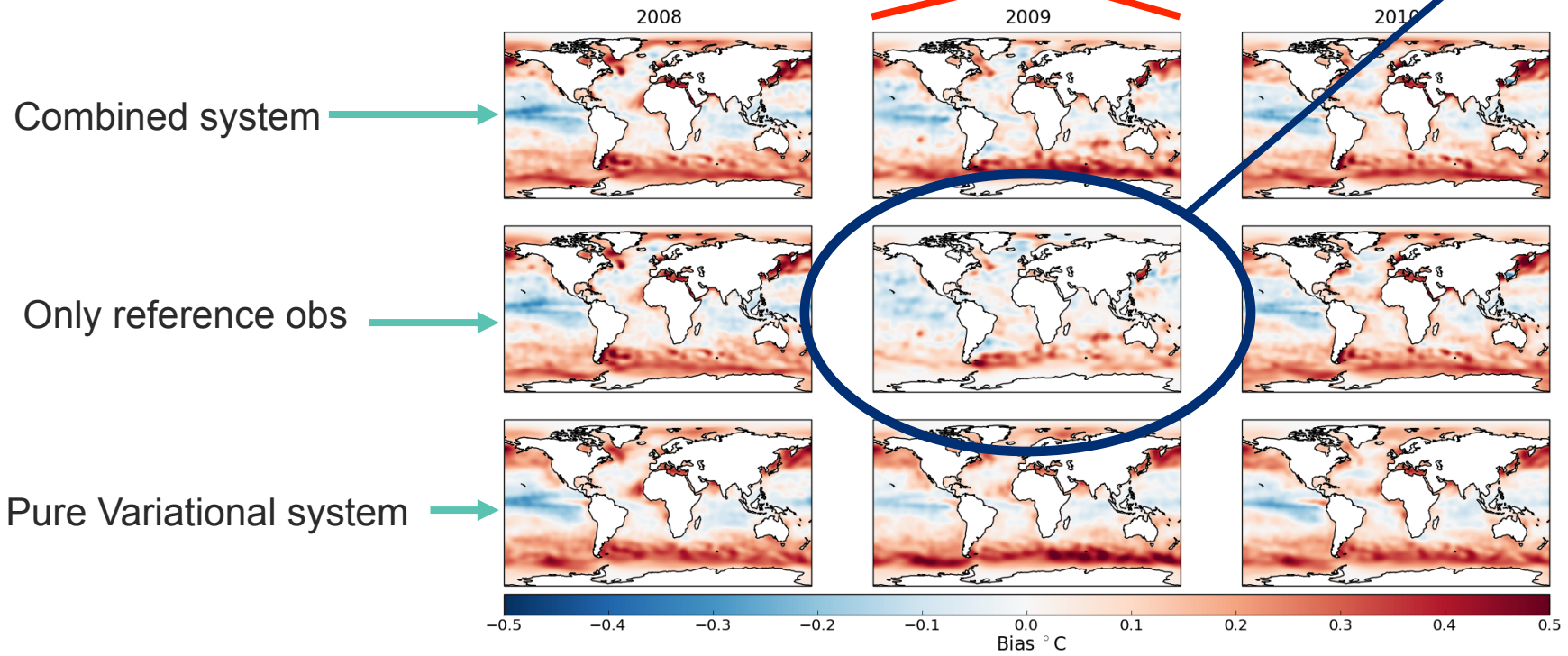
Real-time analysis needs to be put into the context of the reanalysis for each system so that anomalies can be compared.

http://origin.cpc.ncep.noaa.gov/products/GODAS/ocean_briefing.shtml

SST Bias correction

In 2009 we used many fewer reference observations

In 2009, bias
Using the old MO system is very patchy and inconsistent with the other years

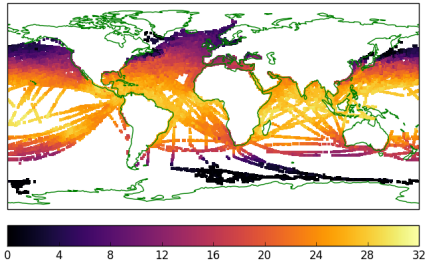


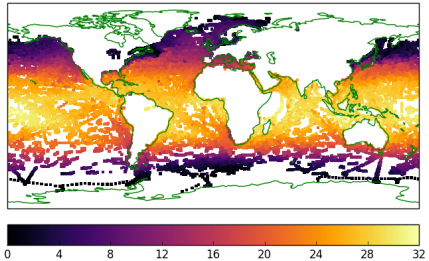
Average estimates of the bias in AMSRE data

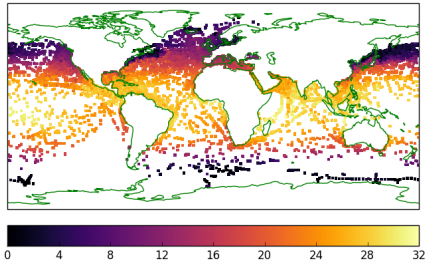
EOF data assimilation

The scheme is being tested by sub-sampling the modern day observing system to resemble an earlier period. The withheld observations can be used to assess the resulting analyses.

In-situ SST







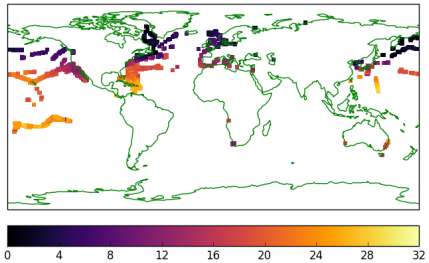
Data from HadIOD.

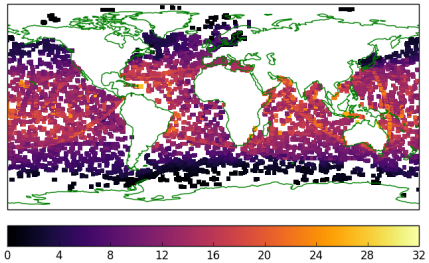
Jan 1953

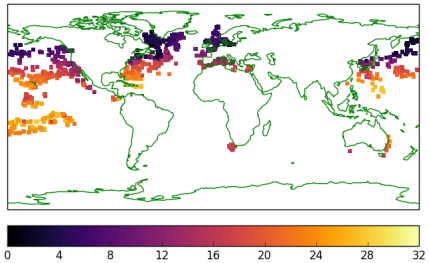
Jan 2010

2010 data sub-sampled

Profile T



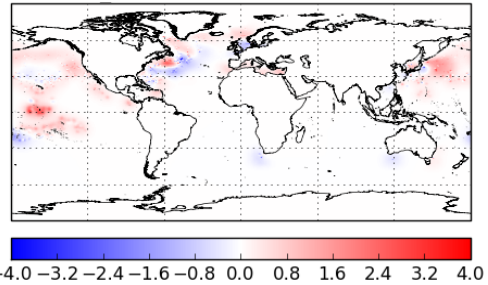




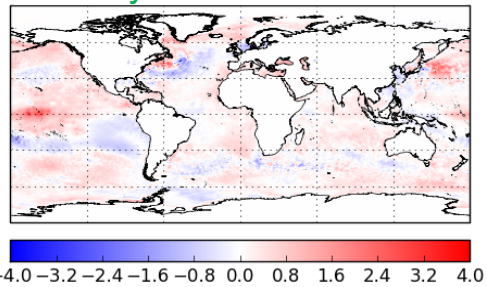
EOF data assimilation

SST increments /ºC from Jan 1953 profiles

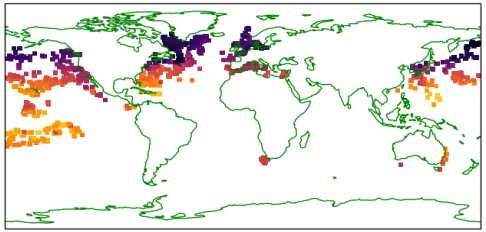
Standard DA



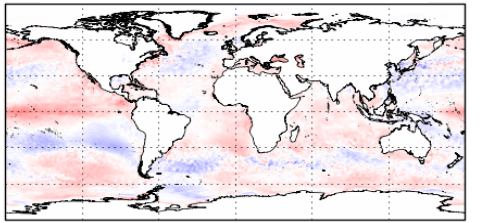
Hybrid 10% EOF



Observations



EOF DA



- EOFs generated from 100-year coupled climate model (HadGEM3).
- Hybrid EOF/standard DA spreads information globally, while retaining the structures from the standard DA near the observations

- The scheme is being tested by sub-sampling the modern day observing system to resemble an earlier period.
- The withheld observations can be used to assess the resulting analyses.
- Preliminary results show small improvements in SST statistics in some regions, e.g. in South Atlantic. Hybrid EOF assimilation in red.