Task 2.5: Research towards development of fully coupled data assimilation methods

Keith Haines Reading University

(1) **Analysis and inter-comparison of coupled error covariances (MetO, UREAD)**

- *Objective*: to assess strengths and weaknesses of weakly coupled data assimilation schemes, and to develop techniques for calculating coupled error covariances.
- Investigate and assess case studies of particular coupled phenomena, and develop methods and carry out the calculation of coupled covariances using the outputs of weakly coupled DA systems (eg. CERA).

(2) **Model bias correction in coupled data assimilation (UREAD)**

- *Objective*: development of bias correction methods explicitly designed to achieve more balanced atmosphere and ocean states as part of a coupled reanalysis system.
- Develop new advective approaches for correcting ocean biases in the tropical thermocline and mid-latitude boundary currents which are critical for atmospheric responses. New approaches will also be tested to reduce coupled model drifts in wind stress, buoyancy fluxes, and ocean thermocline.

(3) **Fully coupled data assimilation in simplified model systems (INRIA)**

- *Objective:* to explore advanced methods for fully coupled data assimilation, using the ECMWF OOPS framework.
- Develop the methodology by which some aspects of fully coupled data assimilation will be implemented. Investigate possibility of controlling the interfaces between the different components of the coupled system together with their initial conditions. Use a simplified coupled system in the OOPS framework, which would allow its extension to a realistic framework.

Popular approaches for Background error covariances: B

Method	Description and references
"Canadian quick" method	$\mathbf{x}_\mathrm{f} - \mathbf{x}_\mathrm{t} \sim (\mathbf{x}_\mathrm{f}(t+T) - \mathbf{x}_\mathrm{f}(T)) / \sqrt{2}$. Take population from one long time run. Polavarapu S., Ren S., Rochon Y., Sankey D., Ek N., Koshyk J., Tarasick D., Data assimilation with the Canadian middle atmosphere model. Atmos.-Ocean 43: 77–100 (2005).
Analysis of innovations $\mathbf{d} = \mathbf{y} - \mathbf{H}\mathbf{x}_\mathrm{f}$ "Hollingsworth and Lonnberg"	Choose a pair of direct and independent obs separated by r: $[y(r) - x_\mathrm{f}(r)]\,[y(r+\Delta r) - x_\mathrm{f}(r+\Delta r)] =$ $[\{y(r) - x_\mathrm{t}(r)\} - \{x_\mathrm{f}(r) - x_\mathrm{t}(r)\}]\,[\{y(r+\Delta r) - x_\mathrm{t}(r+\Delta r)\} - \{x_\mathrm{f}(r+\Delta r) - x_\mathrm{t}(r+\Delta r)\}]$ $\langle[\epsilon^y(r) - \epsilon^{x_\mathrm{f}}(r)]\,[\epsilon^y(r+\Delta r) - \epsilon^{x_\mathrm{f}}(r+\Delta r)]\rangle = \langle\epsilon^y(r)\epsilon^y(r+\Delta r)\rangle + \langle\epsilon^{x_\mathrm{f}}(r)\epsilon^{x_\mathrm{f}}(r+\Delta r)\rangle\,,$ (above assumes obs and bg errors are uncorrelated). Take population from many pairs with same Δr. Furthermore if $\Delta r > 0$: $\langle\epsilon^y(r)\epsilon^y(r+\Delta r)\rangle = 0$. Rutherford I.D. 1972. Data assimilation by statistical interpolation of forecast error fields. J. Atmos. Sci. 29: 809–815. Hollingsworth A., Lönnberg P., The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus 38A: 111–136 (1986). Järvinen H., Temporal evolution of innovation and residual statistics in the ECMWF variational data assimilation systems. Tellus 53A: 333–347 (2001).
NMC method	Choose pairs of lagged forecasts valid at the same time, e.g.: $\mathbf{x}_\mathrm{f} - \mathbf{x}_\mathrm{t} \sim \left(\mathbf{x}_\mathrm{f}^{48}(t) - \mathbf{x}_\mathrm{f}^{24}(t)\right) / \sqrt{2}$. Take population from difference at many times. Parrish D.F., Derber J.C., The National Meteorological Center's spectral statistical interpolation analysis system. Mon. Wea. Rev. 120 1747–1763 (1992). Berre L., Ştefănescu S.E., Pereira M.B., The representation of the analysis effect in three error simulation techniques. Tellus 58A 196–209 (2006).
Ensemble method	If you have an ensemble that is correctly spread: $\mathbf{x}_\mathrm{f} - \mathbf{x}_\mathrm{t} \sim \mathbf{x}_\mathrm{f}^{(i)} - \langle\mathbf{x}_\mathrm{f}\rangle$ or $\mathbf{x}_\mathrm{f} - \mathbf{x}_\mathrm{t} \sim \left(\mathbf{x}_\mathrm{f}^{(i)} - \mathbf{x}_\mathrm{f}^{(j)}\right) / \sqrt{2}$. Take population from ensemble members and over many times. Houtekamer P.L., Lefaivre L., Derome J., Ritchie H., Mitchell H.L., A system simulation approach to ensemble prediction. Mon. Wea. Rev. 124, 1225–1242 (1996). Buehner M., Ensemble derived stationary and flow dependent background error covariances: Evaluation in a quasi-operational NWP setting. Q.J.R. Meteorol. Soc. 131, 1013–1043 (2005).

Bannister

Available Coupled DA Model Runs

- ECMWF: CERA system
- Short 2 month CERA reanalyses: April-May 2008, Dec-Jan 2008-09, Apr-May 2010, Aug-Sept 2010, Dec-Jan 2010-11.
- 10-day forecasts run once per day (00Z) through all above periods. No ensemble
- 10-day forecasts run every 5 days (00Z) with uncoupled initialisation methods. No ensemble (Testing coupling shocks in Un/Coupled DA)
- MetO:
- 1-year Weakly coupled reanalysis: 2012
- 10-day forecasts run twice a day (00Z and 12Z) from 26/08/2012 to 15/09/2012, No ensemble
- 10-day forecasts run twice a day (00Z and 12Z) from 20/10/2012 to 31/10/2012. No ensemble

- No overlapping periods between systems yet

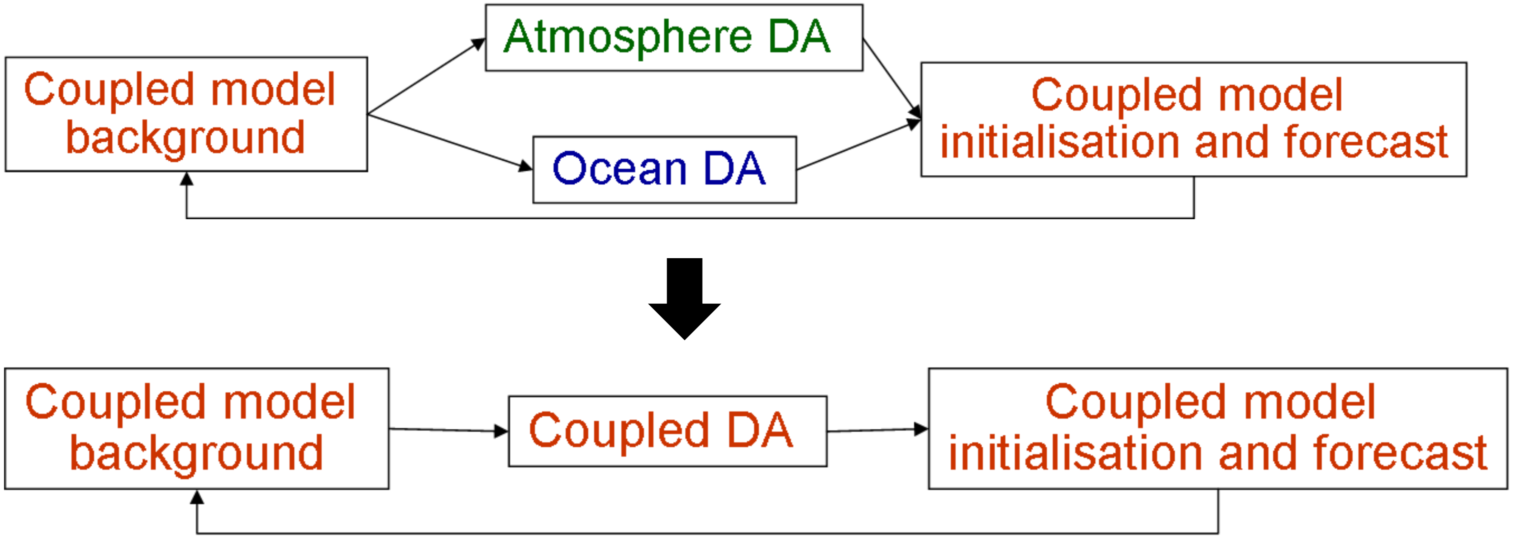
Met Office

Assess coupled covariances

Isabelle Mirouze

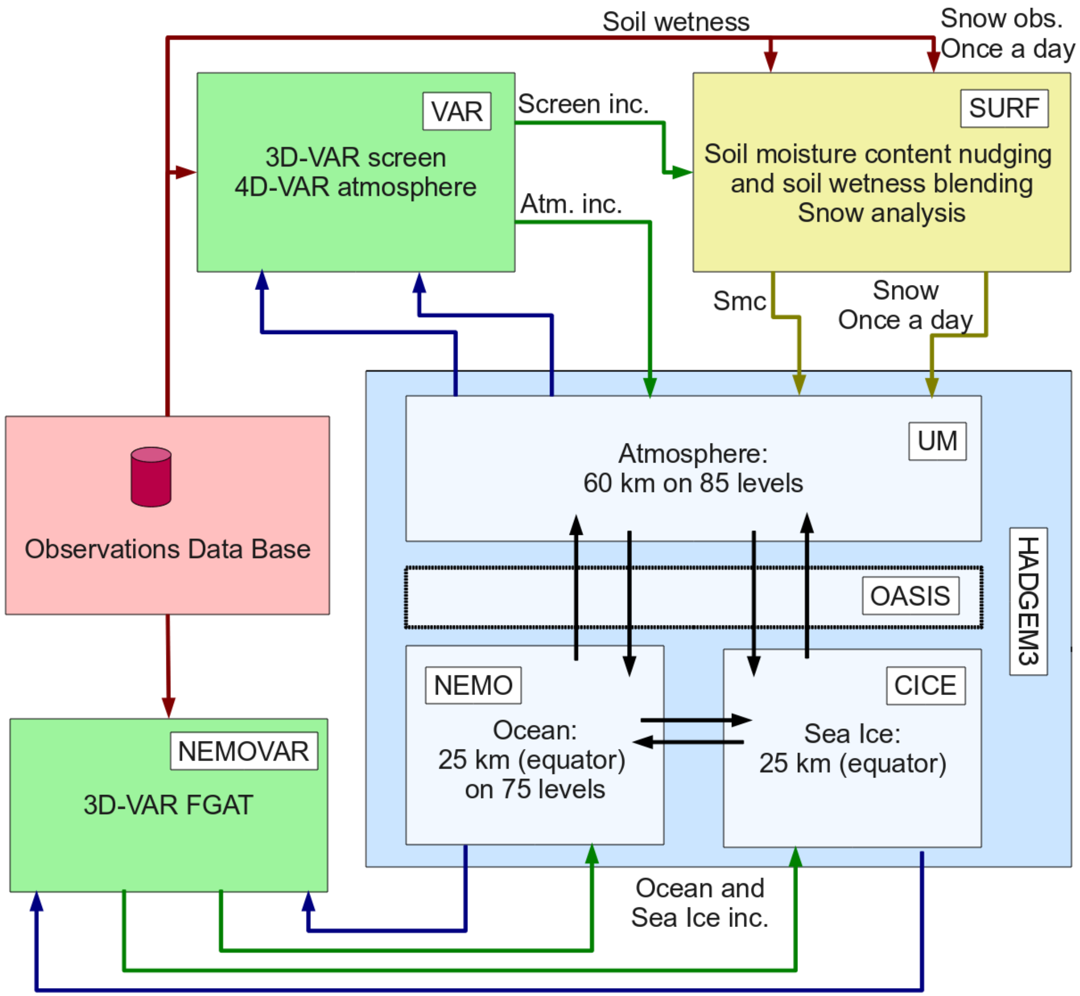
ERA_CLIM2 General Assembly November 2014

Weakly- to more fully-coupled data assimilation system



- We need a better understanding of the coupled forecast error distribution
  - What do the interfluid correlations look like?
  - How do we estimate them?

The Met Office coupled data assimilation system



Met Office / ECMWF system main differences

Atmosphere / land	Met Office	ECMWF
Model + DA system	UM + JULES VAR + SURF	IFS

Ocean / Ice / Waves	Met Office	ECMWF
Model	NEMO + CICE	NEMO + WAM
Ocean resolution	¼ degree 75 levels	1 degree 42 levels
DA system	NEMOVAR	
SST	Assimilated	Nudged

Coupled DA	Met Office	ECMWF
Window	6 h	24 h
Outer loop	1	2

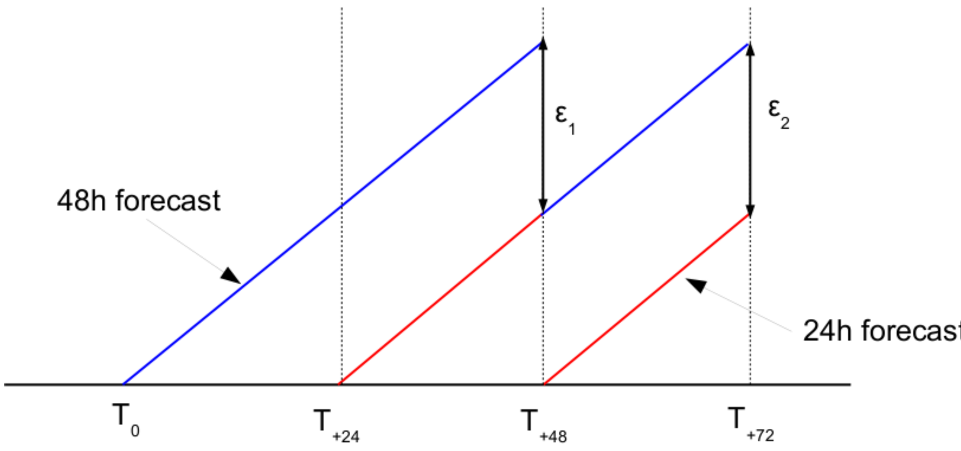
Coupled DA experiments

- 1-year reanalysis: 2012
- 10-day forecasts run twice a day (00Z and 12Z) from 26/08/2012 to 15/09/2012
  - Asian Monsoon: low pressure system between 3 and 10 September coincides with large peak rainfall
  - Tropical storms and hurricanes: Kirk, Leslie, **Michael**
- 10-day forecasts run twice a day (00Z and 12Z) from 20/10/2012 to 31/10/2012
  - Tropical storms and hurricanes: Tony, **Sandy**

➡ Use the forecast experiments to develop techniques for calculating coupled error covariances

The NMC method

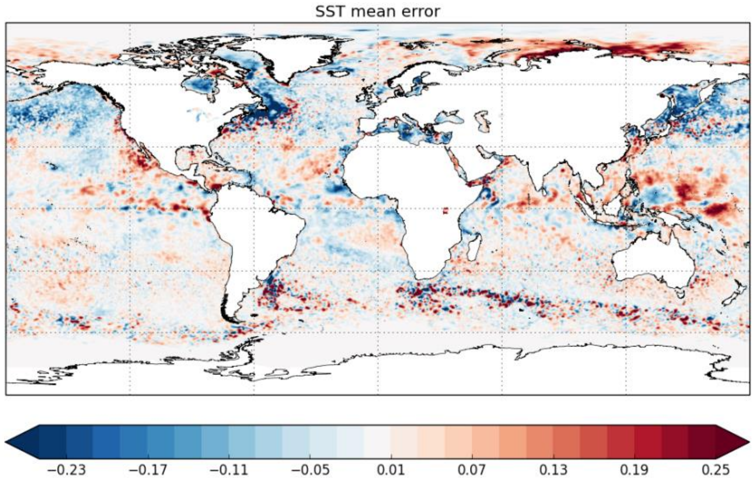
- Developed at the **N**ational **M**eteorological **C**enter
- Construct an ensemble of forecast differences as proxies for the background errors



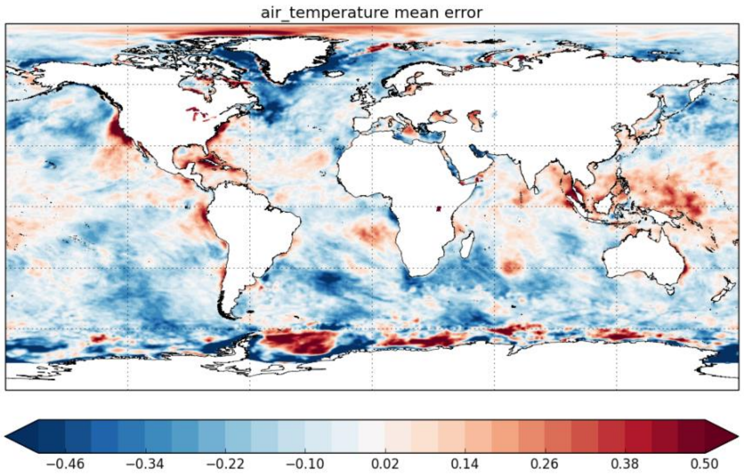
- Assuming the P(**ε**) is Gaussian, then P($\mathbf{x}^b$) is Gaussian and $$\mathbf{B} = \frac{1}{2}\left\langle \left(\boldsymbol{\varepsilon} - \bar{\boldsymbol{\varepsilon}}\right)\left(\boldsymbol{\varepsilon} - \bar{\boldsymbol{\varepsilon}}\right)^T \right\rangle$$

Ensemble of errors

- Shell scripts:
  - Retrieve atmosphere and/or ocean fields
  - Compute the forecast differences **ε** for each field
  - Interpolate the errors **ε** on the same grid
  - Output: NetCDF files with ocean mask



SST mean error



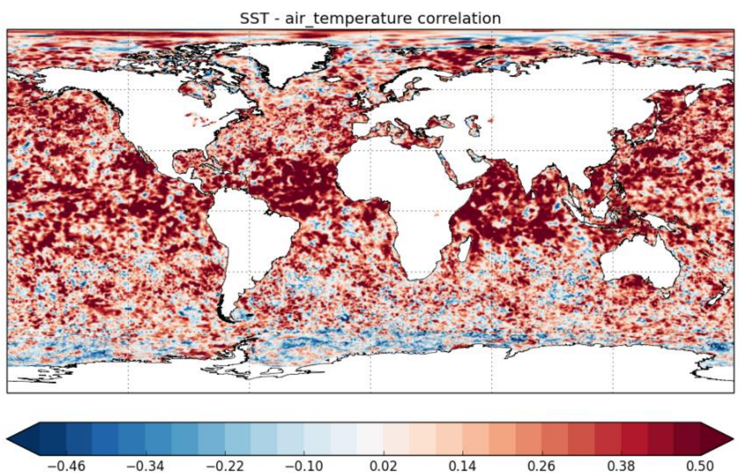
Air temperature 1.5 m mean error

26/08 to 15/09, 30h – 6h differences

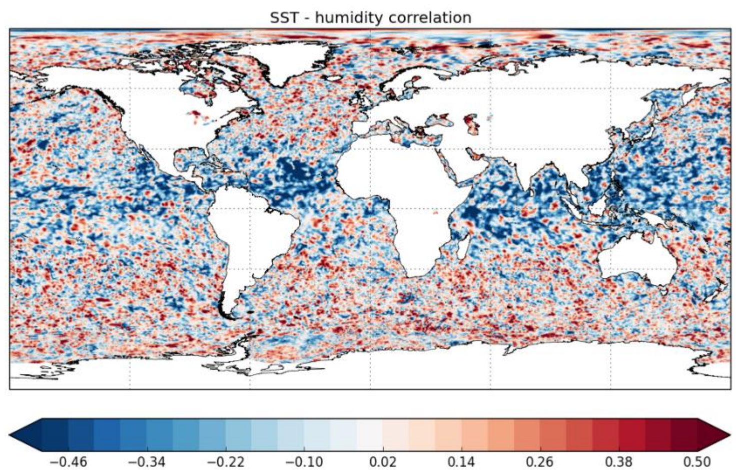
Point to point correlations between surface error fields

- Python scripts:
  - Compute point to point correlation between an ocean error field and an atmosphere error field
  - Plot the map of point to point correlation
  - Output: pickle file, png file

These are not yet results. Ensemble of 21 members only.



SST and air temperature 1.5 m

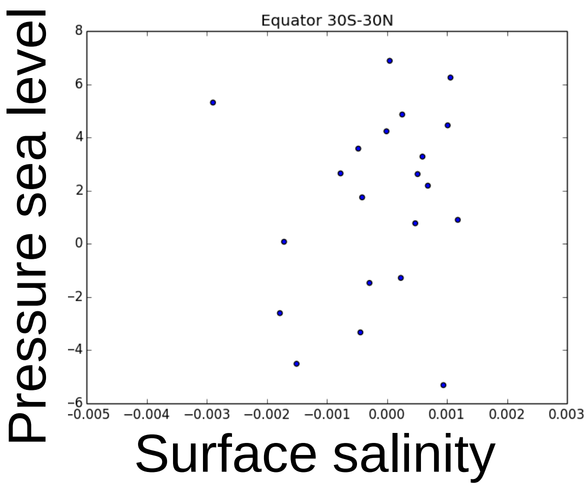


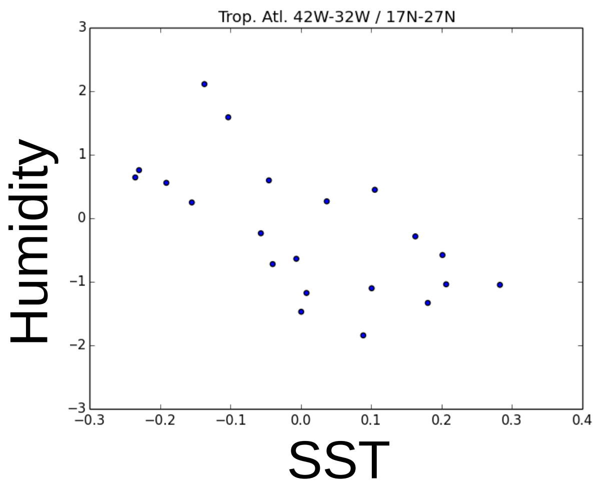
SST and humidity 1.5 m

Scatter plots

- Python scripts:
  - Compute regional average of ocean and atmosphere error fields
  - Scatterplot of the error fields
  - Output: pickle file, png file

These are not yet results. Ensemble of 21 members only.





Future plans

- To calculate ensemble of errors for more variables, using both periods
  - Use different forecast differences (e.g. 36h – 12h)
- To study further point to point correlations and scatter plots; to extend to vertical correlations, lagged correlations
  - Try and identify interesting pairs
- To compute the covariance between the interesting pairs
  - Balanced / unbalanced variables
- To focus on the particular phenomena that occurred during the available forecast periods
- Time permitting, run 2-day forecasts for 1 year and recalculate the covariance
  - Climatologic, seasonal covariance

Task 2.5 Develop the methodology by which some aspects of fully coupled data assimilation will be implemented. Investigate possibility of controlling the interfaces between the different components of the coupled system together with their initial conditions.

Objective: to explore advanced methods for fully coupled data assimilation, using the ECMWF OOPS framework.

We plan to explore the algorithmic part of coupled DA with a toy coupled system (at first 2 components and a bulk).

several leads :

- control of both initial conditions and coupling interfaces (weakly vs loosely coupled)
- implement global in time schwartz coupling method and study the interactions minimization / coupling iterations
- (robust) control of coupling parameter.

We just hired Rémi Pellerej (hidden somewhere in the audience) to work on these aspects.

Idealised Coupled DA studies @ Reading

PI: Amos Lawless, PDRA: Polly Smith

- 1D single column coupled atmosphere-ocean incremental 4D-Var assimilation system (Paper exists in draft)
- Current:
  - investigation of structure of cross-covariances between the atmosphere and ocean variables
  - Canadian Quick Covs method
- Future (from Jan 2015; 3-yr funding just secured):
  - development of new methods for incorporating state-dependent covariance information into strongly- and weakly-coupled variational systems
  - focus on estimation and use of atmosphere-ocean cross-covariances
  - comparison of climatological and flow dependent covariances plus ensemble-variational formulations

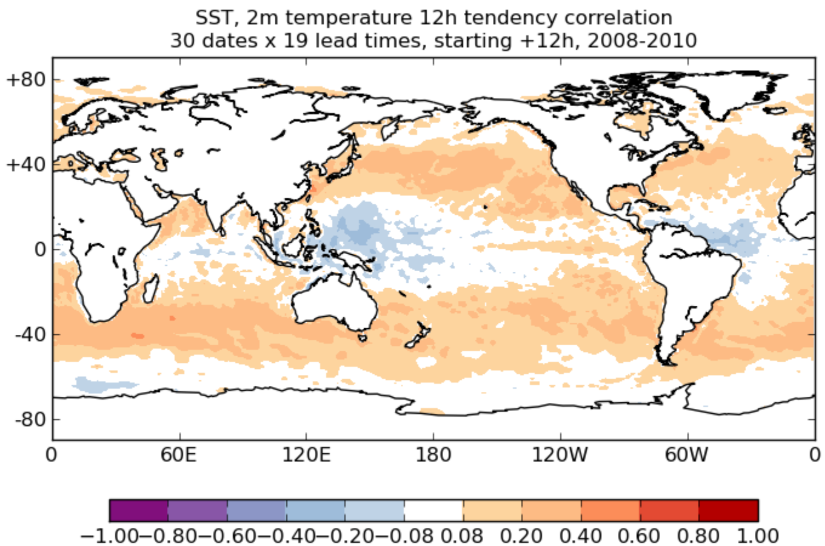
University of Reading

Canadian Quick applied to B for CERA hindcasts

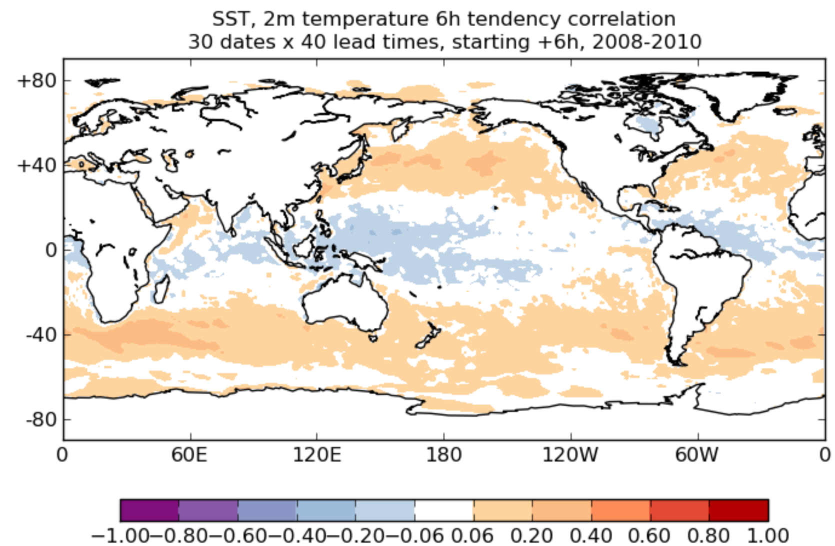
- Use Hindcasts instead of long climatology => Anchor to real world conditions.
- Don't need hindcasts everyday as NMC (Hindcasts every 5 days @ Reading for comparison with Uncoupled reanalyses)
- $B = ½ [(x_{t+\delta} - x_t) – <(x_{t+\delta} - x_t)>] [(x_{t+\delta} - x_t) – <(x_{t+\delta} - x_t)>]^T$ where δ=6hrs
- Used 4 different $<(x_{t+6} - x_t)>$ to account for diurnal cycle
- Used 3 x 2-month CERA hindcasts started every 5 days = 30x10 day hindcasts = 30x40 6hr tendencies

- Compared with:
  - Direct error covariances (Forecasts v Analyses) from CERA hindcasts
  - NMC method using daily CERA hindcasts

SST – SAT tendency covariances from CERA initialised hindcasts



CERA Weakly Coupled



Uncoupled

$$B = \tfrac{1}{2}\,[(\mathbf{x}_{t+\delta} - \mathbf{x}_t) - \langle(\mathbf{x}_{t+\delta} - \mathbf{x}_t)\rangle]\,[(\mathbf{x}_{t+\delta} - \mathbf{x}_t) - \langle(\mathbf{x}_{t+\delta} - \mathbf{x}_t)\rangle]^{\mathrm{T}}$$

Negative correlations => variations in strong coupling regions (Tropics)

Turbulent Air-Sea fluxes (Air-Sea Δt) => tendency anti-correlations

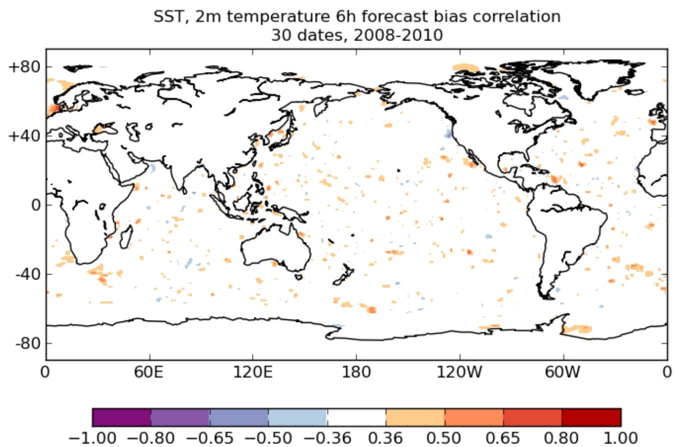
Is this realistic?

Test against other methods

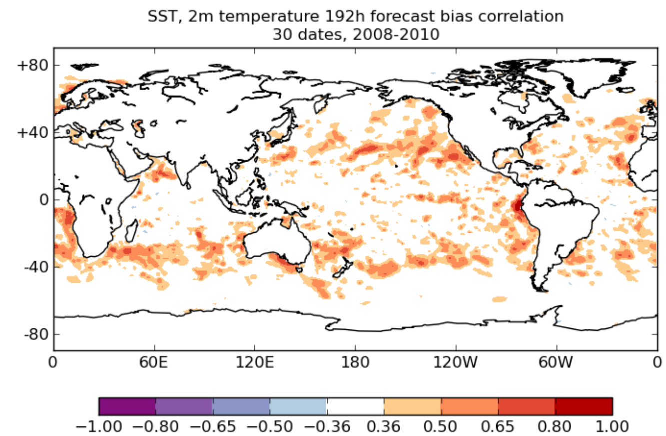
University of Reading

Direct CERA error covariances

6hr forecast SST-SAT error covariances:



8 day forecast SST-SAT error covariances



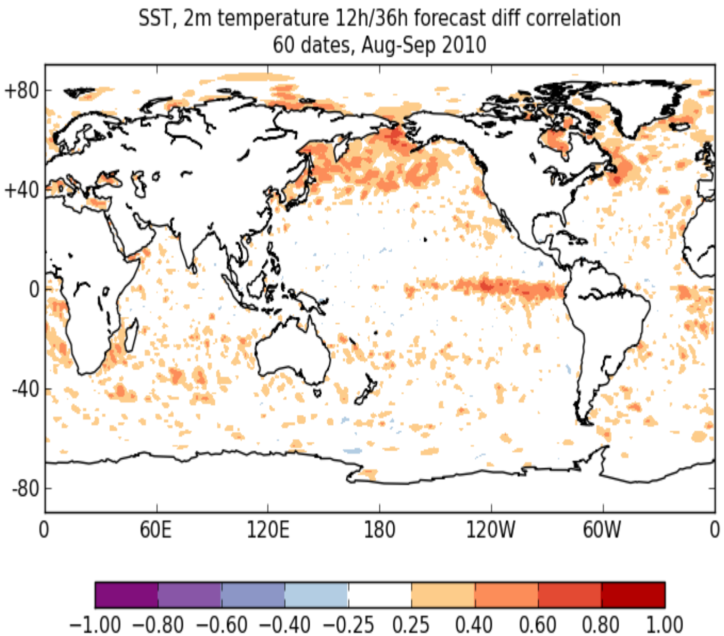
- CERA Single member hindcast runs initialised every 5 days during 3 2-month periods in 2008-2010 => 30 start times in all.
- Errors assessed against subsequent analyses (@6hrs and @8days)
- @8 days covariances starting to represent clim. bias?
- Not really enough data for robustness @6hrs

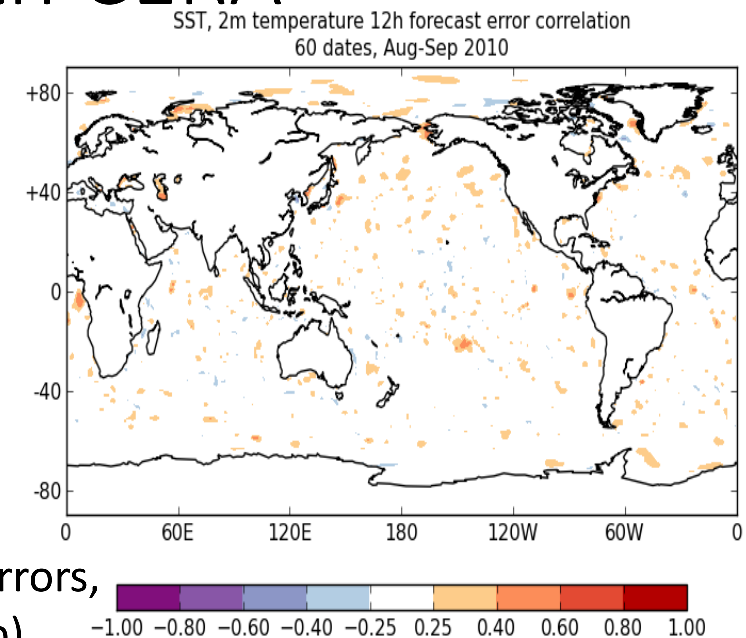
University of Reading

NMC method with CERA

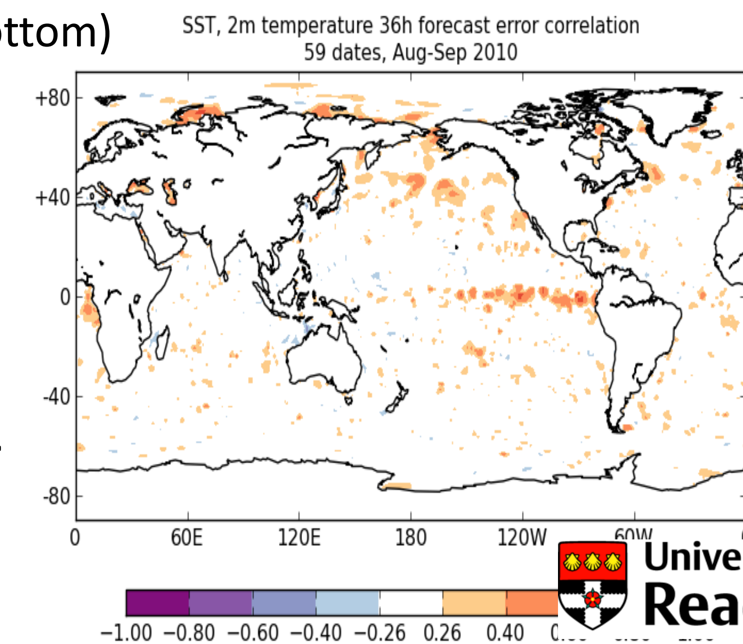
NMC, 36h-12h

Compare with MetO





Direct errors,
12h (top)
36h (bottom)

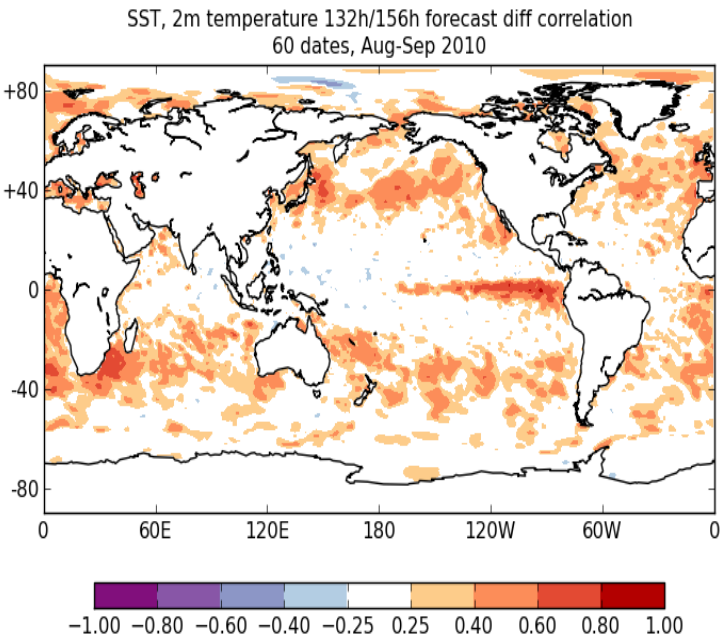


ECMWF CERA runs now initialised daily @ 00Z

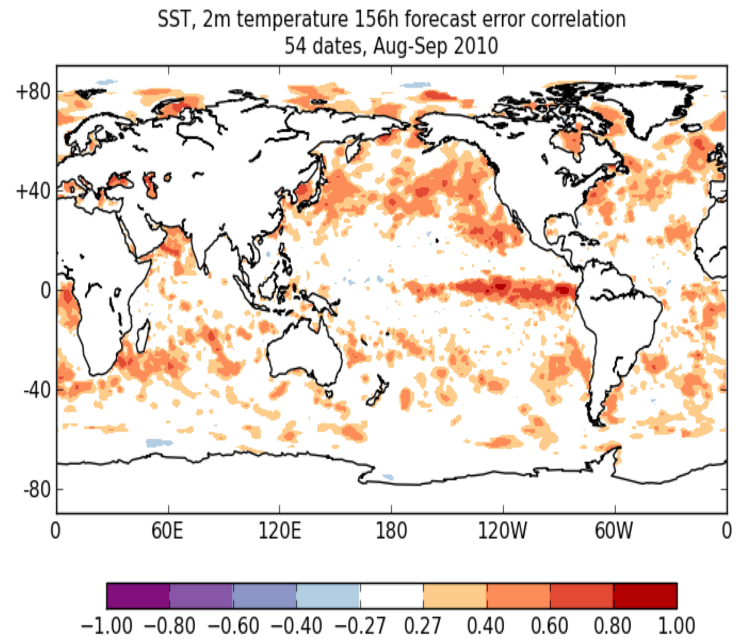
Comparison of NMC and direct error covariance methods

NMC with Longer lead times:

NMC, 6.5d-5.5d



Direct error, 6.5d

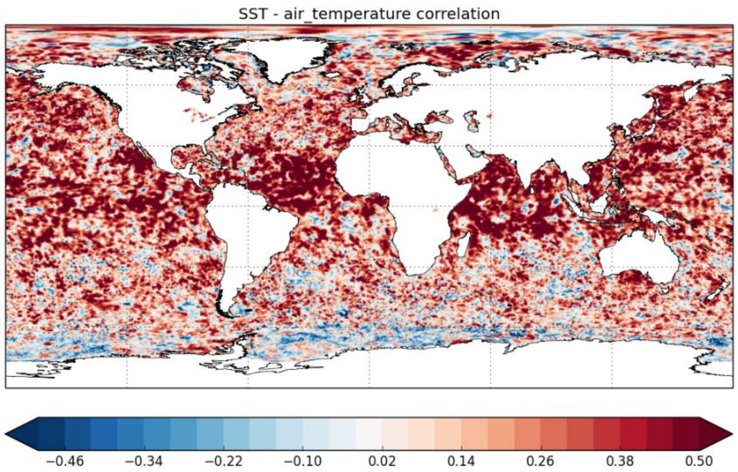


University of Reading

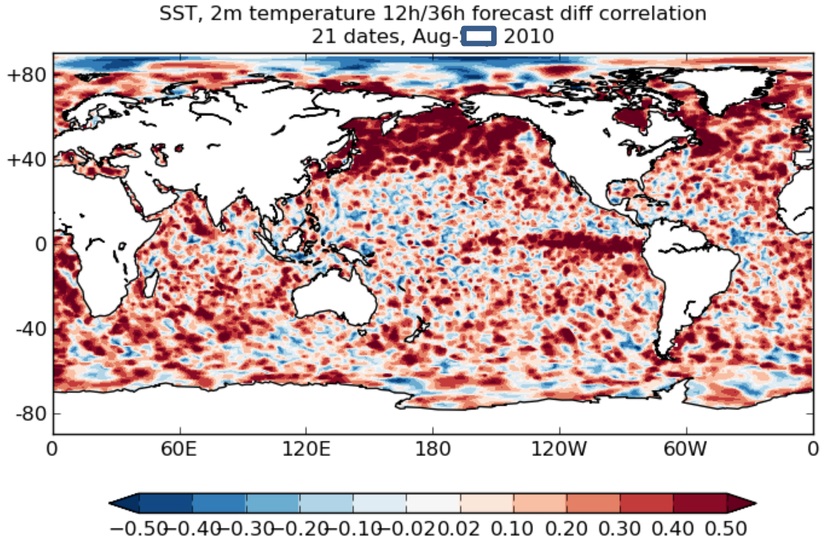
Point to point correlations between surface error fields

- Python scripts:
  - Compute point to point correlation between an ocean error field and an atmosphere error field
  - Plot the map of point to point correlation

These are not yet results. E



MetO NMC results

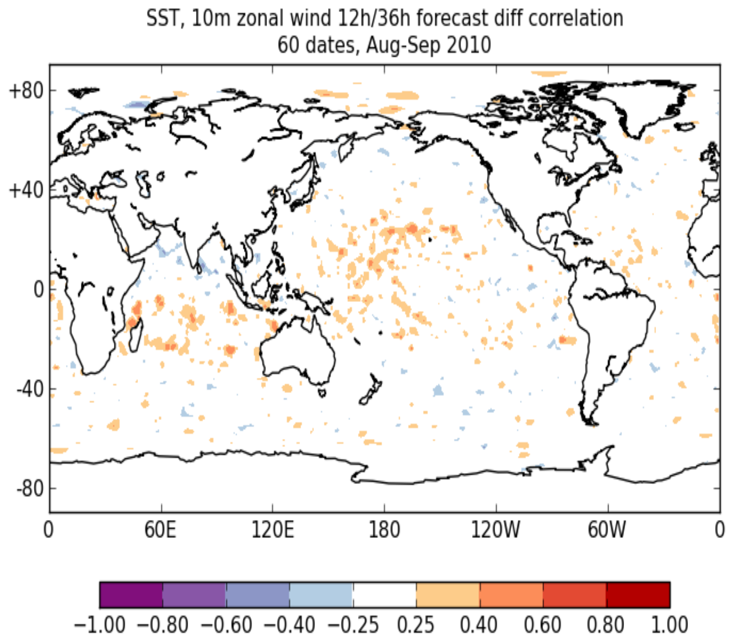


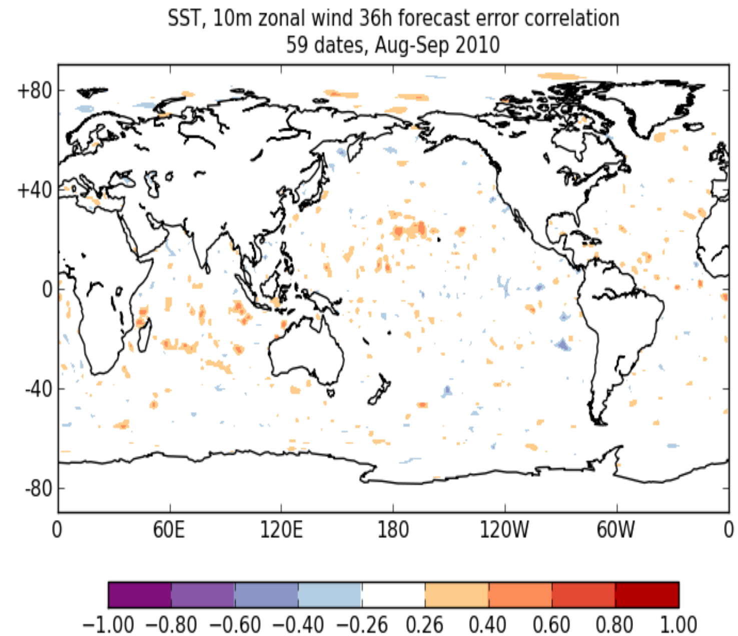
NMC from CERA

NMC with SST vs 10m zonal wind Covariances

NMC, 36h-12h

Direct error, 36h





University of Reading

Further studies

- Further weakly coupled covariance errors
  - SST v Humidity, Winds, Atmospheric boundary layer profiles, Precipitation,
- Compare and understand MetO and CERA differences: Ideally for common year comparison
  - Also understand which coupled covariances might add value to a reanalysis
- Some example ensemble covariance cases using different ensemble generation methods: Case studies eg. Typhoon or TIW periods

Deliverables

- D2.8 Report on strengths and weaknesses of weakly coupled data assimilation methods for Earth system reanalysis. UREAD 18
- D2.9 Report on techniques for calculating coupled error covariances from outputs of a weakly coupled data assimilation experiment. METO 18
- D2.10 Report on assessment of coupled-model drift and approaches for obtaining consistent ocean and atmospheric bias corrections. UREAD 34
- D2.11 Report on fully coupled data assimilation in simplified systems with implications for Earth system reanalysis. INRIA 34