Data Assimilation for Reanalysis

Dick Dee

Massimo Bonavita, Mike Fisher, Paul Poli, David Tan

ECMWF

WCRP 4th International Conference on Reanalyses
7-11 May 2012
Silver Spring, Maryland

Reanalysis of the instrumental record

Special challenges for data assimilation:

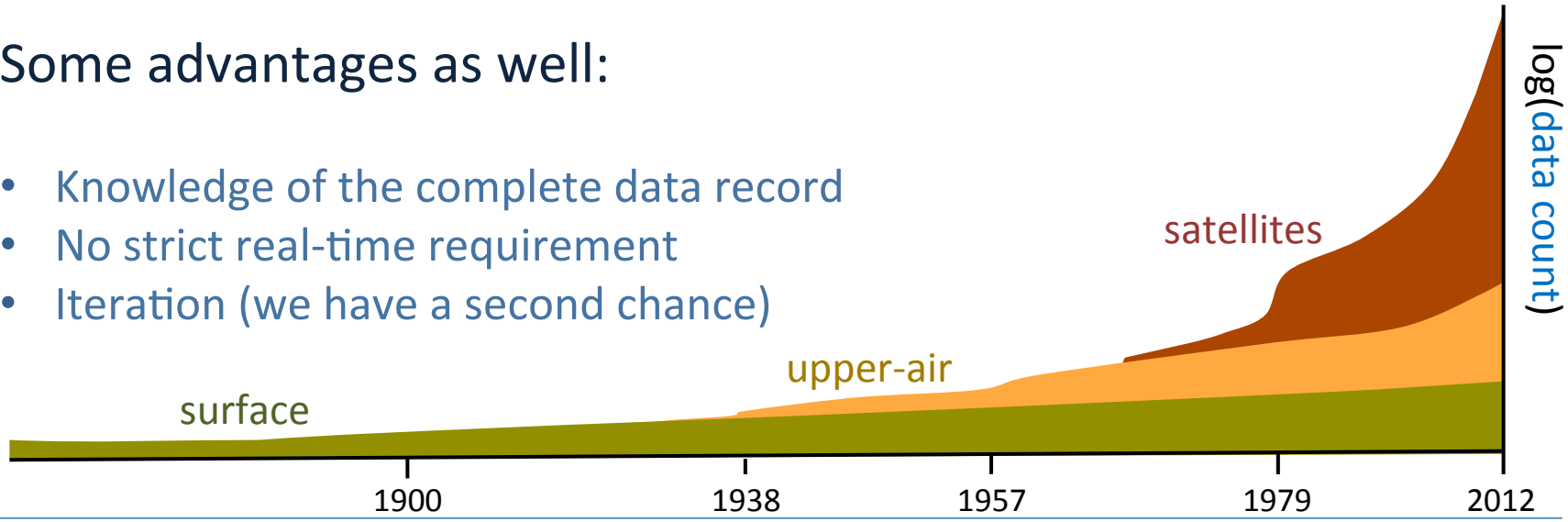
- Maintaining temporal consistency
- Time-varying background errors
- Using sparse observations
- Uncertainty estimates
- Computational cost

Topics for this talk:

- Ensemble Data Assimilation
- Long-window 4D-Var
- Use of weak-constraint 4D-Var to control model biases

Some advantages as well:

- Knowledge of the complete data record
- No strict real-time requirement
- Iteration (we have a second chance)



Approach taken in ERA-CLIM

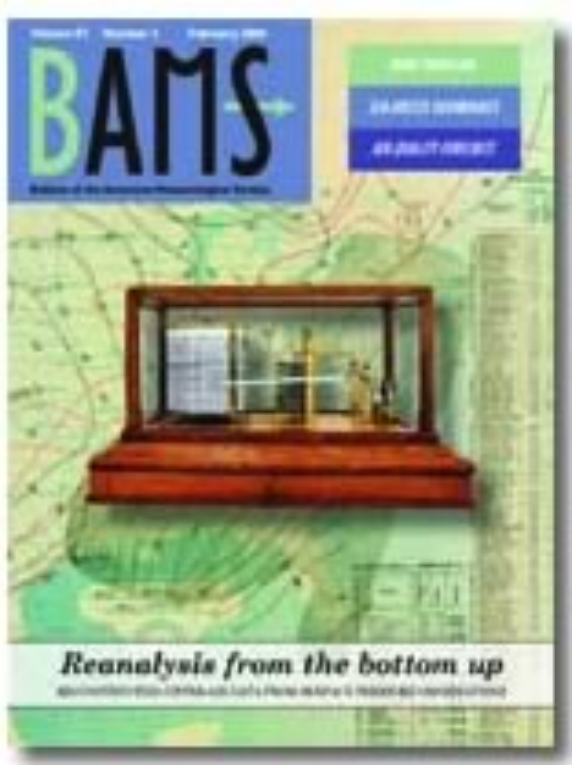
Develop a 20th-century climate reanalysis from the bottom up:

- Model + boundary conditions, atmospheric forcing data
- Reanalysis of surface observations only
- Reanalysis of early upper-air observations
- Reanalysis of reprocessed satellite data

Beyond ERA-CLIM:

- Comprehensive 20C atmospheric reanalysis
- Some form of coupling with the ocean
- Should begin production by end of 2014

This is a long-term project!



Choice of data assimilation method

ECMWF systems are based on 4D-Var analysis

- Has been successfully used in ERA-Interim
- Variational bias correction is important for reanalysis
- Options: longer analysis window, accounting for model errors

Major shortcoming: Background errors are not dynamic

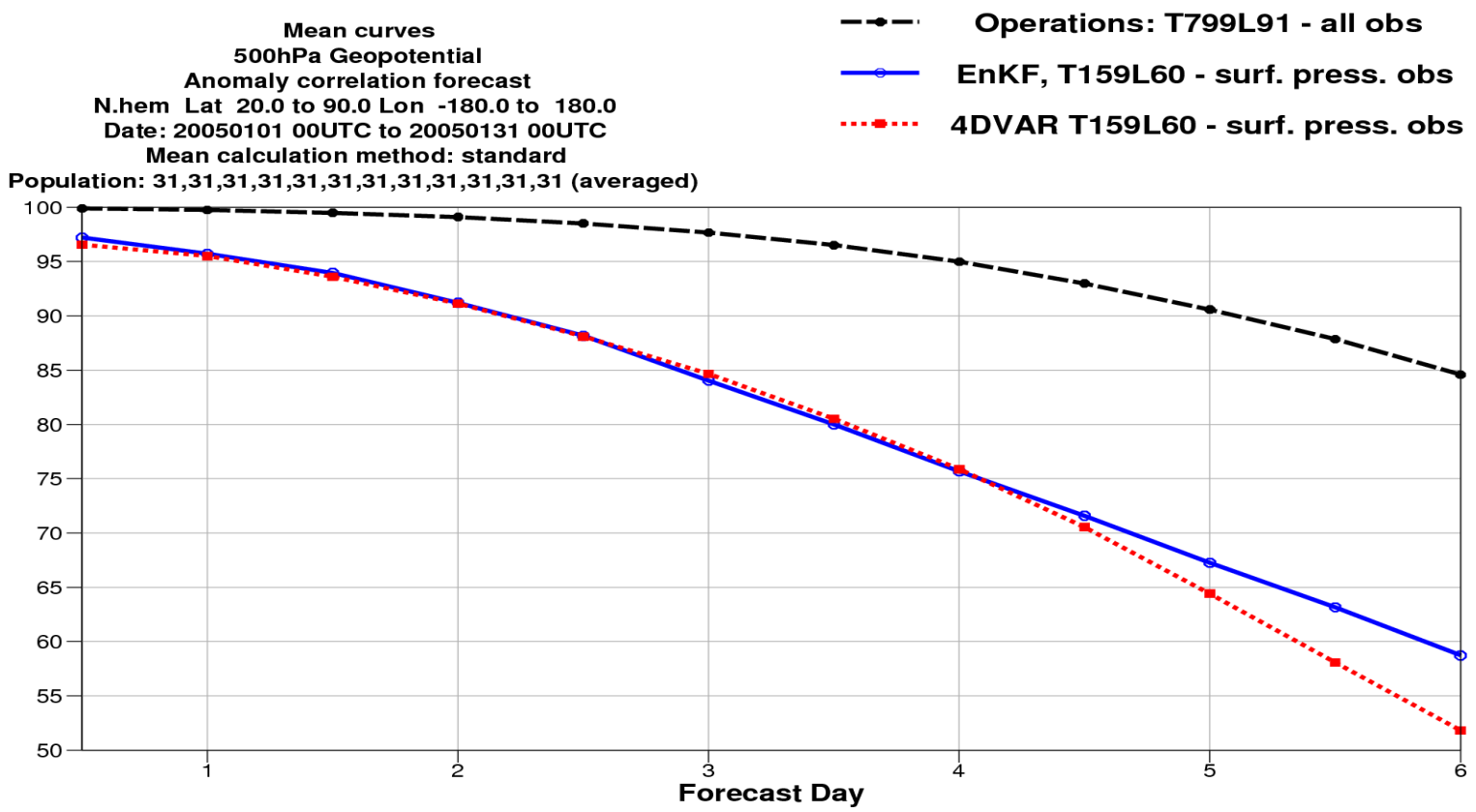
- Uses a stationary covariance model
- Flow-dependence induced by balance operators
- Manually tuned (model-dependent)

What about an ensemble Kalman filter (EnKF)?

- Implemented at ECMWF in collaboration with Jeff Whitaker

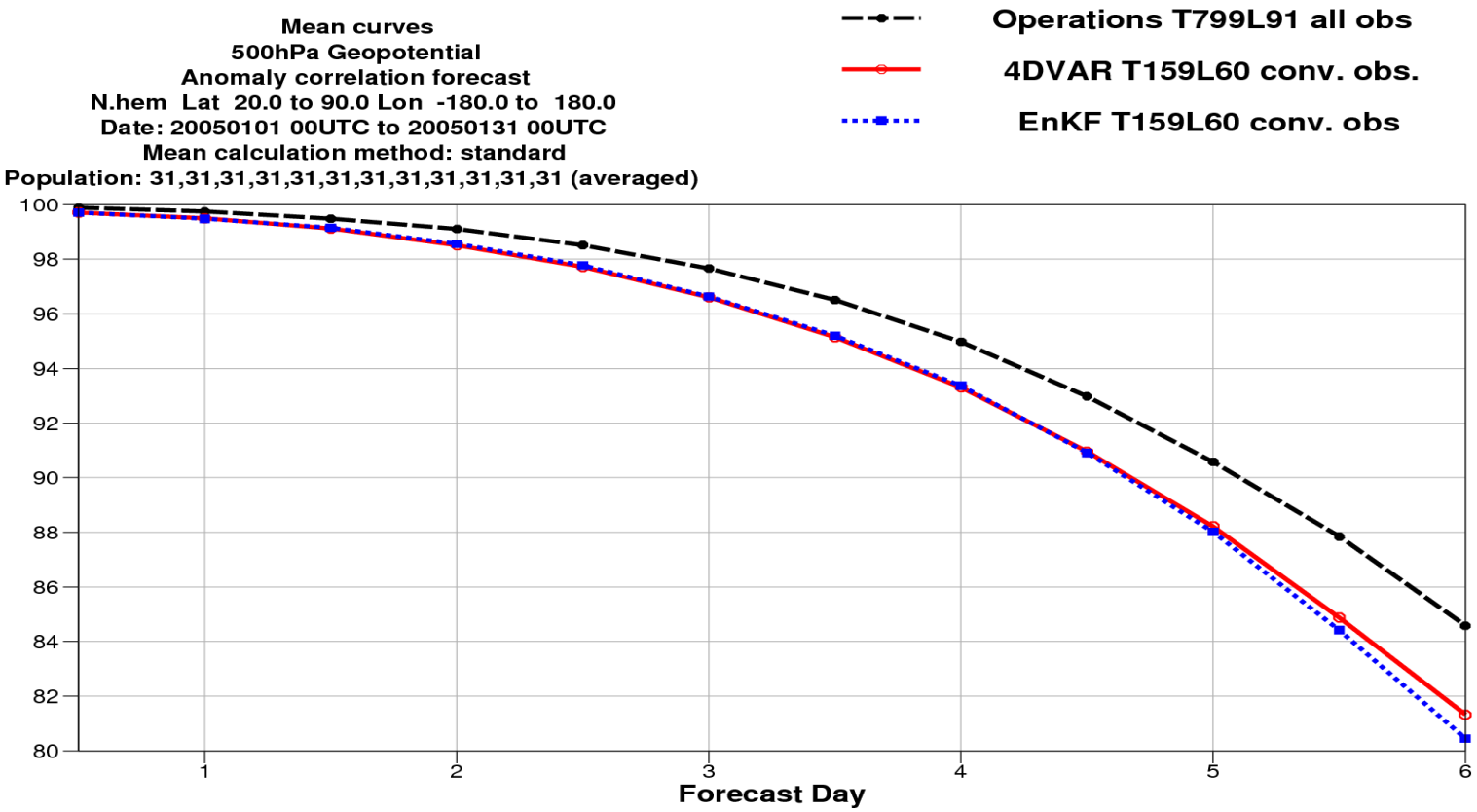
EnKF vs 4D-Var experiments

Assimilating surface pressure observations only



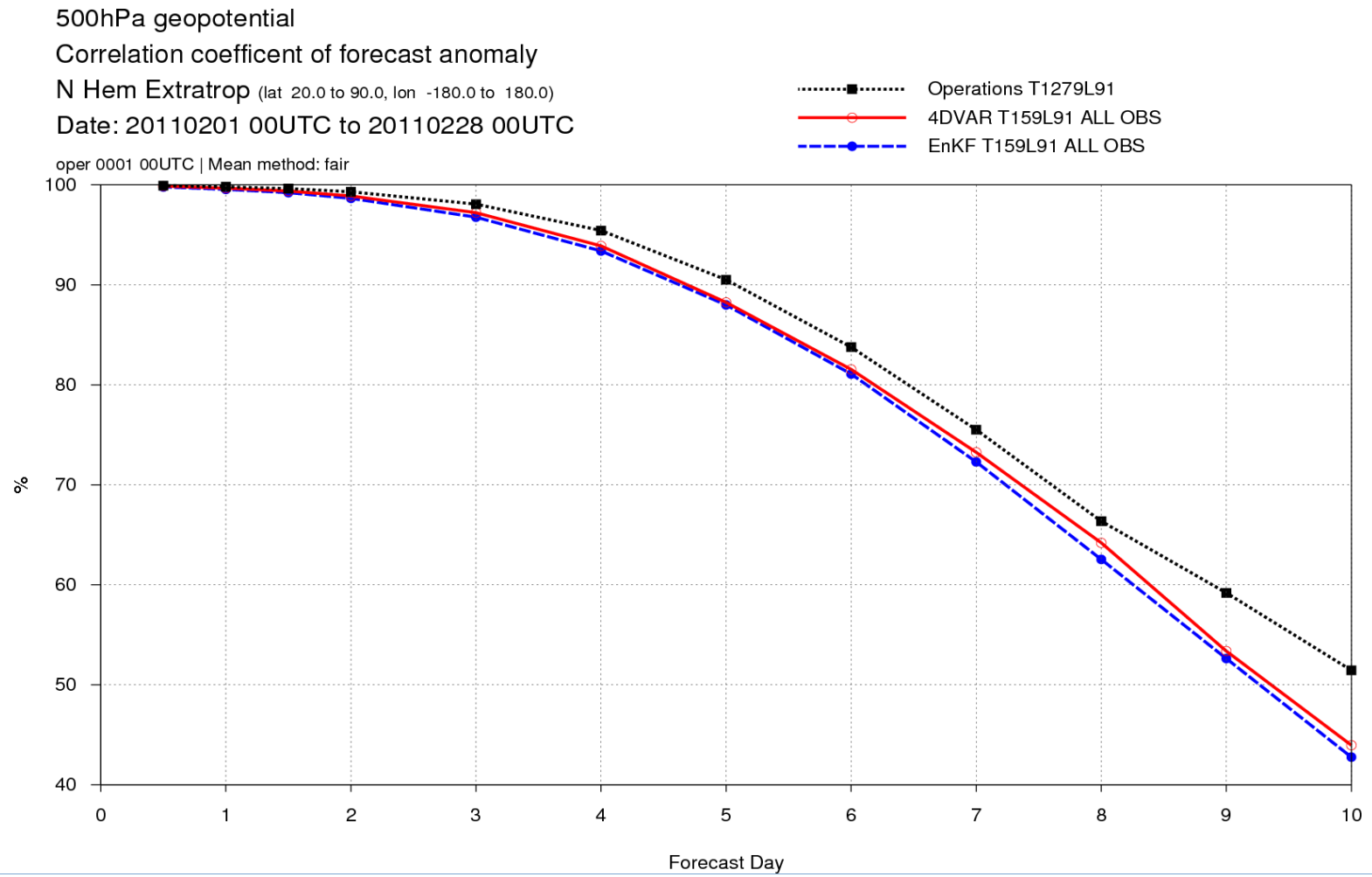
EnKF vs 4D-Var experiments

Assimilating all conventional observations



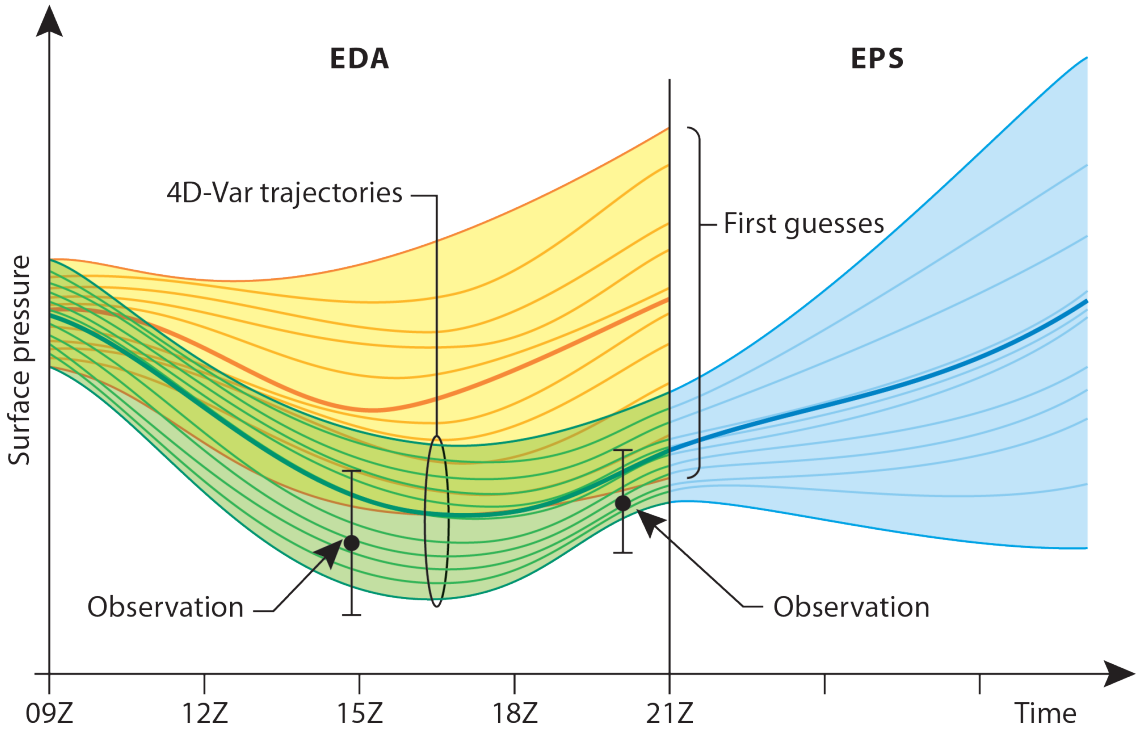
EnKF vs 4D-Var experiments

Assimilating all conventional and satellite observations



Development of a hybrid EDA/EnKF

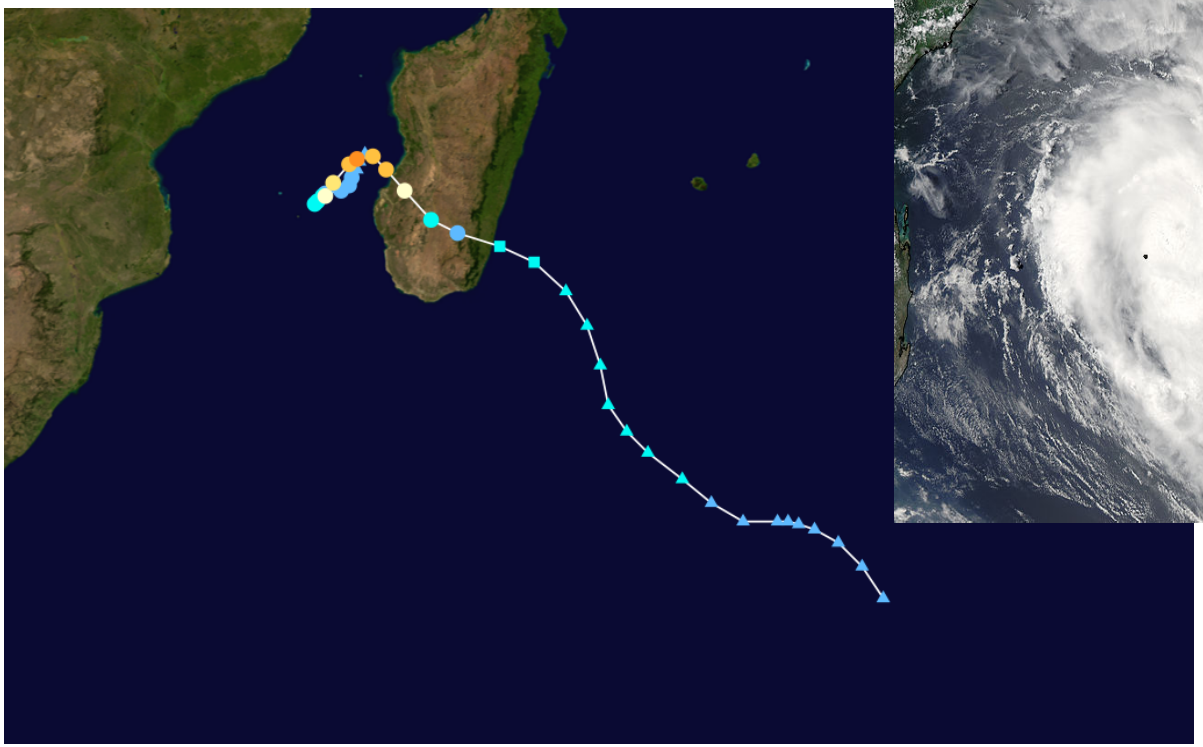
An ensemble of low-resolution (T399) 4D-Var data assimilations (EDA) is now used to estimate analysis and background errors for the operational forecasting system (T1279)

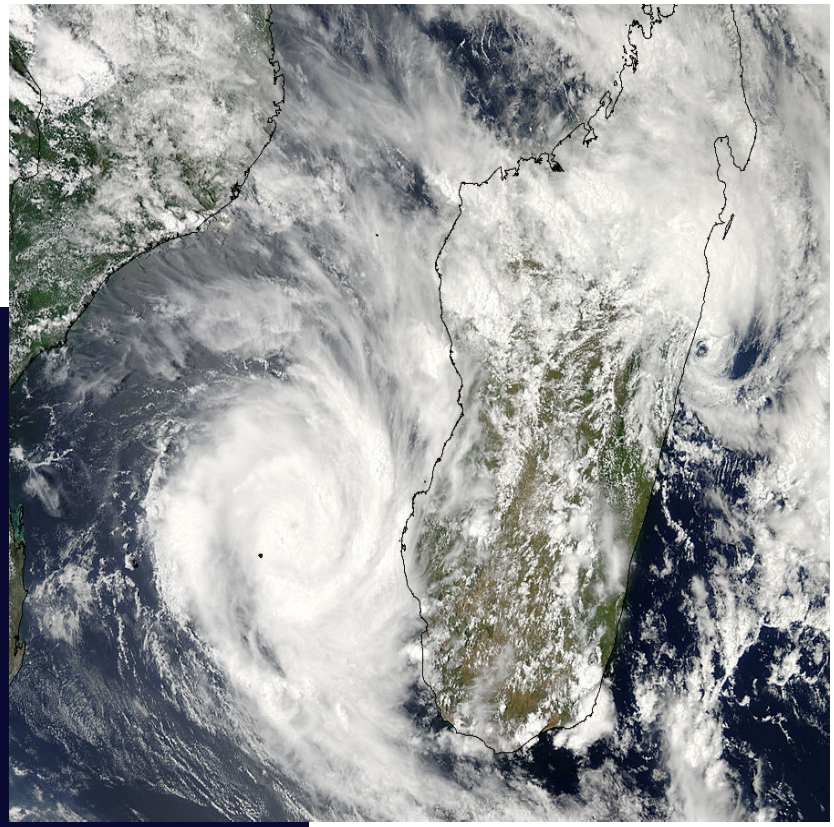


The EDA also creates perturbations for the ensemble prediction system (EPS)

Flow-dependent background errors

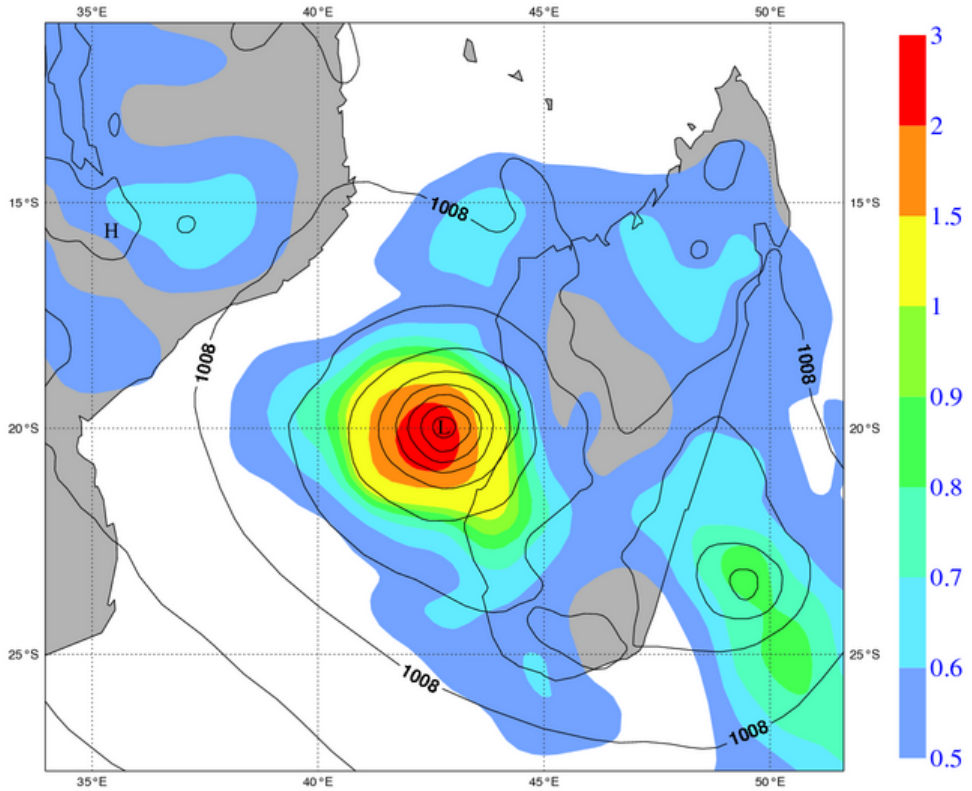
Hurricane Fanele
Indian Ocean, January 2009



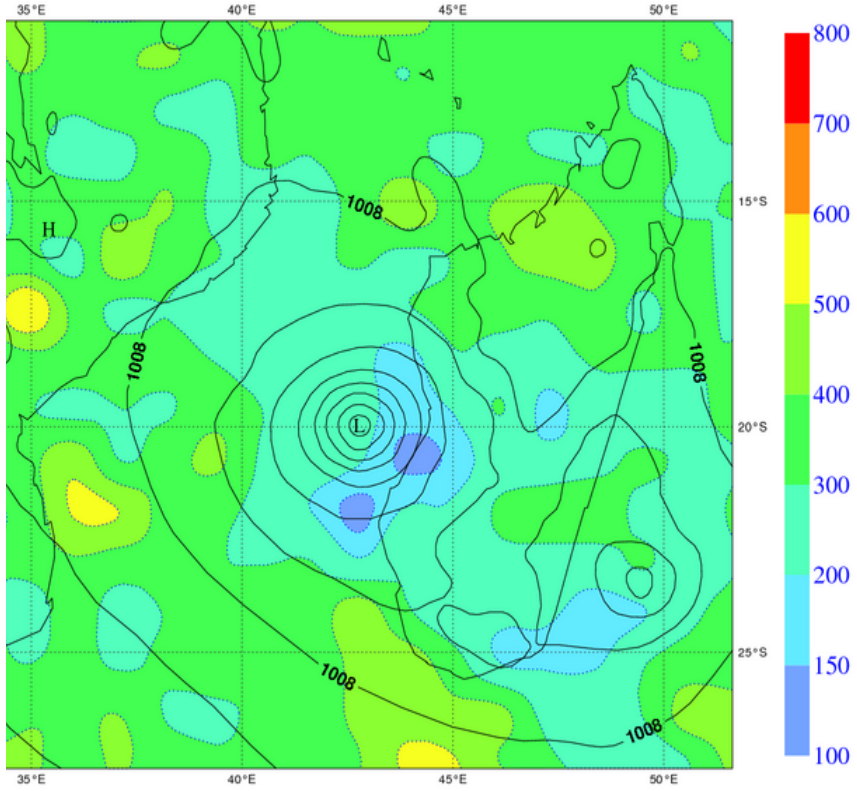


EDA with 20 members

Surface pressure error standard deviation [hPa]



Correlation length scales [km]



Variance estimates are fully flow-dependent; correlations not yet

EDA configuration for ERA-20C

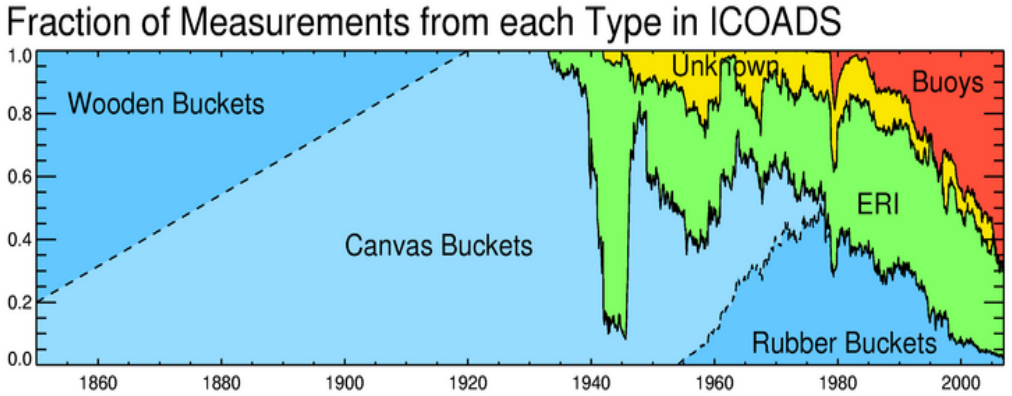
Ensemble of reanalyses from 1900, surface observations only, T159/L91/N10

10 members defined by:

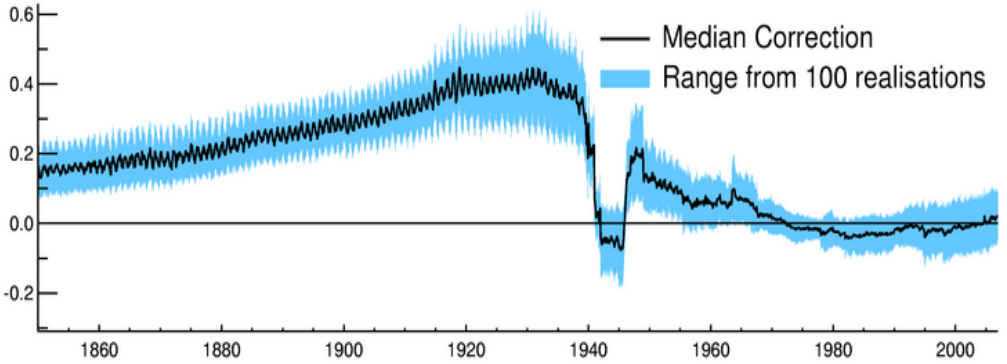
- Different HadISST2 realizations
- Stochastic physics in the forecast model
- Randomly perturbed observations

Background error variances estimated from ensemble

Still relying on stationary correlation structures

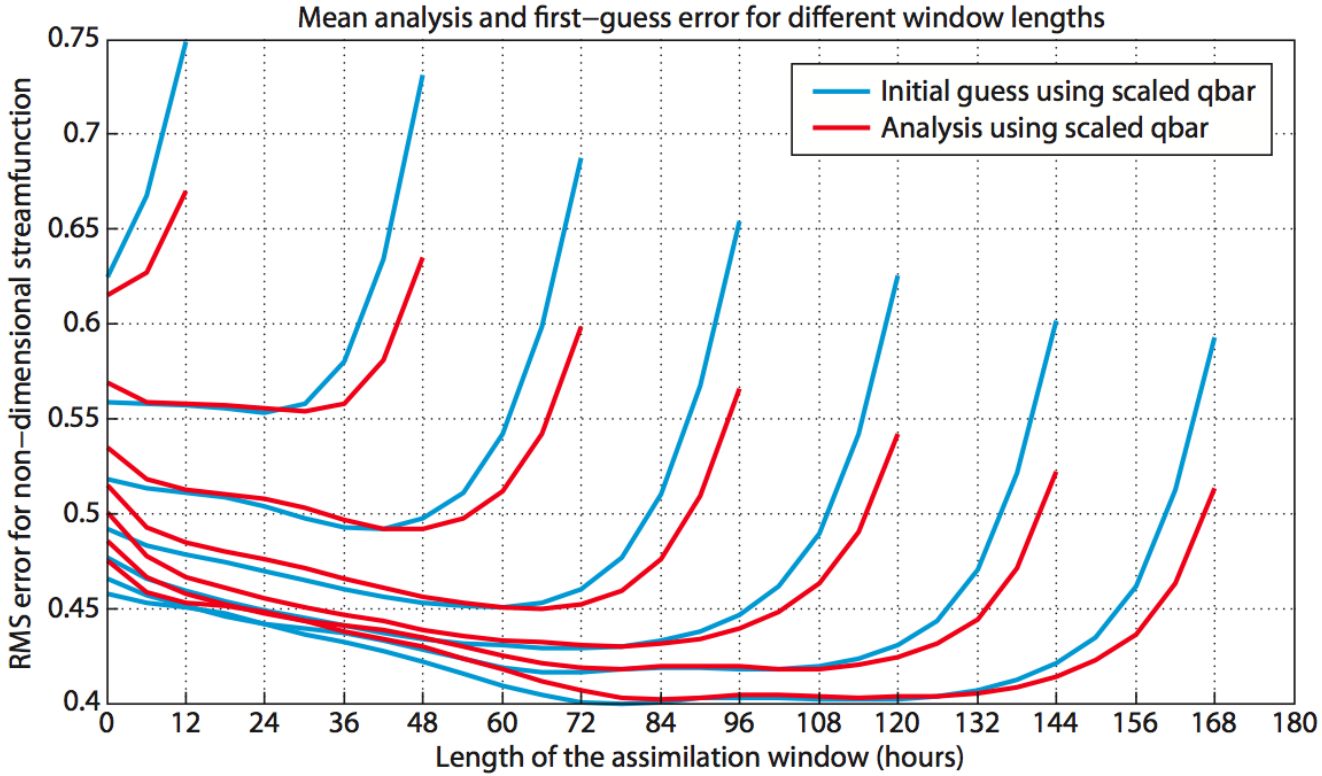


Bias corrections used in HadISST2



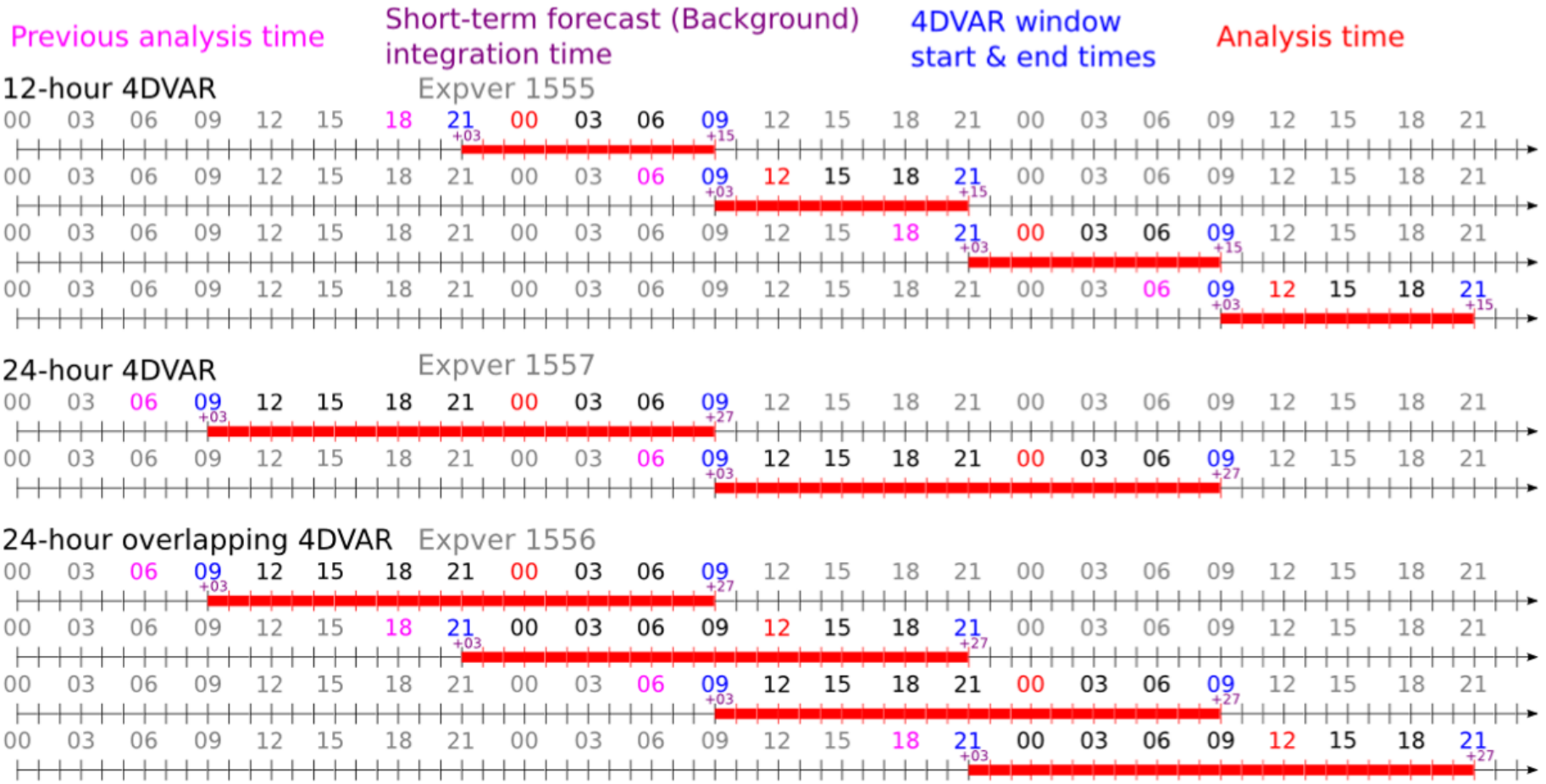
Long-window 4D-Var

Experiments with a 2-level QG model with realistic model errors



- Longer windows –> smaller background errors
- Analysis errors are smallest in the interior of the window
- This is easier to exploit in reanalysis than in forecasting

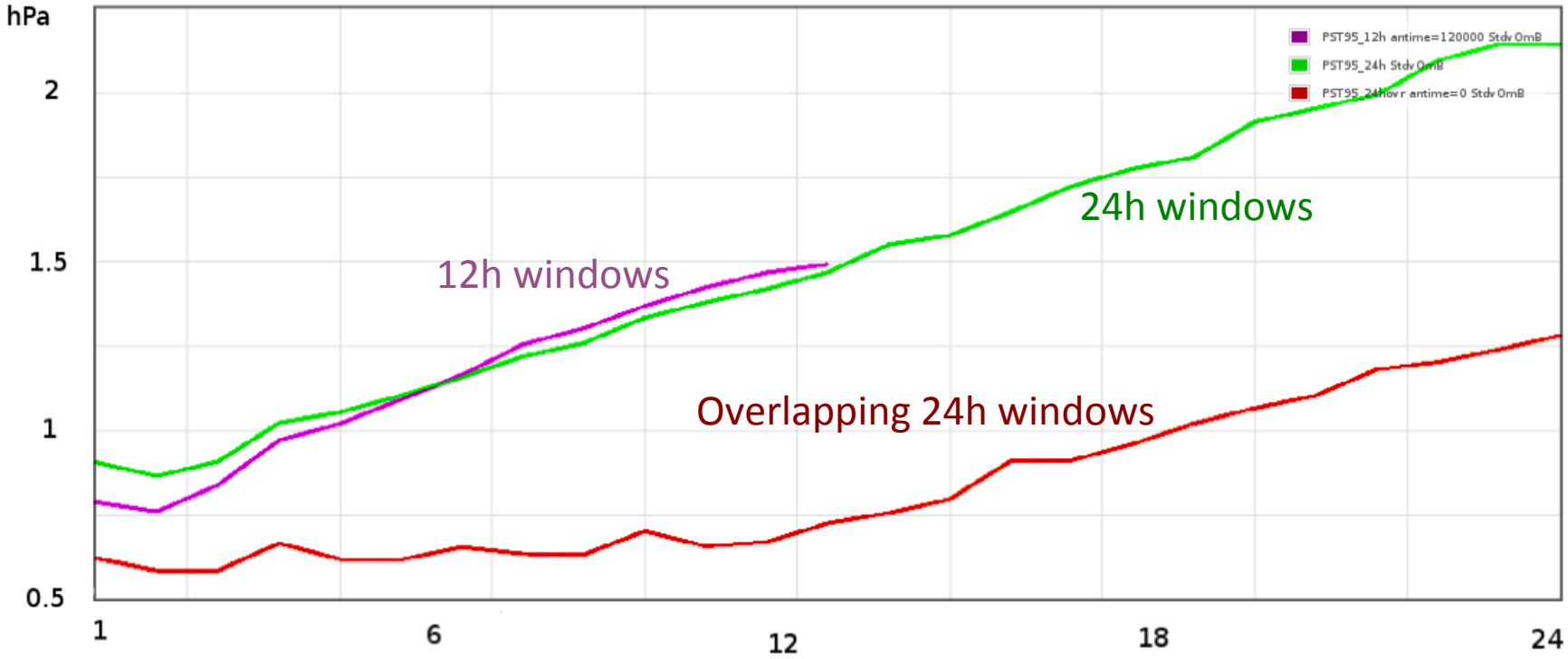
Cycling schemes



Overlap is necessary for even longer windows, to get an accurate first guess

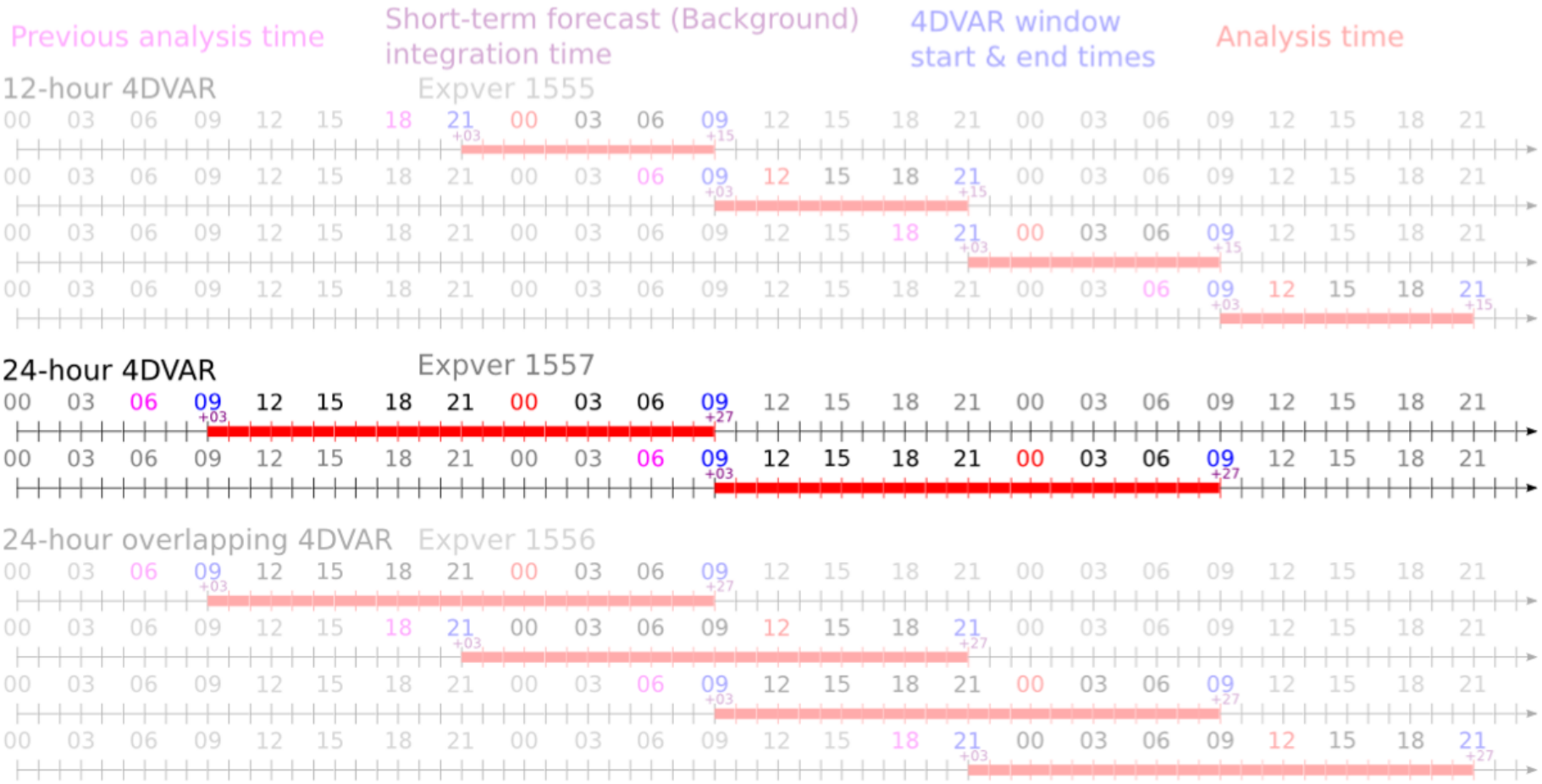
Assimilating surface pressure only

Background forecasts verified against surface pressure observations



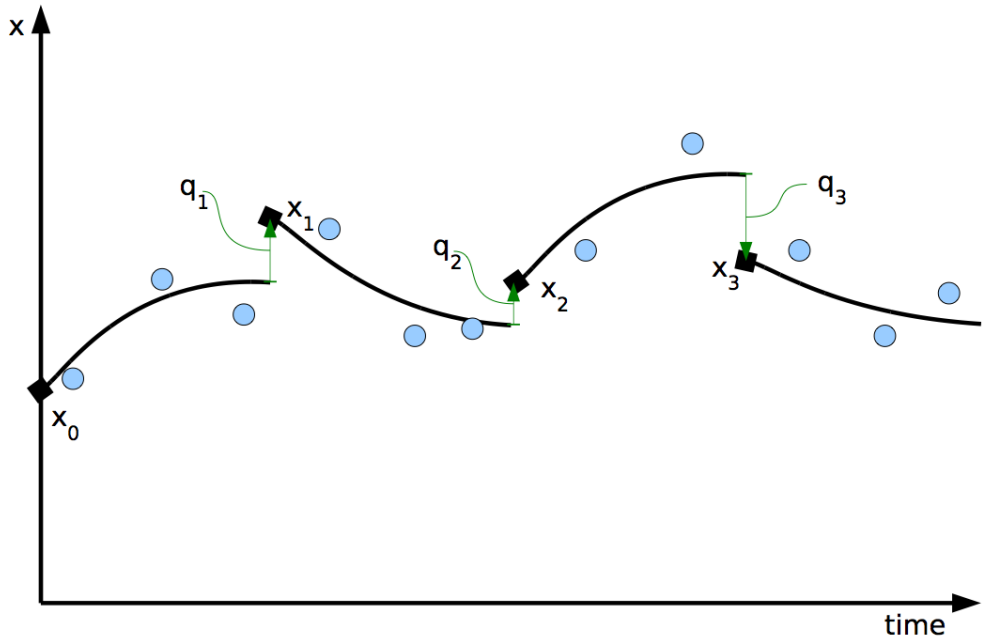
Overlapping 24h windows is much better – but too expensive for ERA-20C

ERA-20C will use 24h 4D-Var



Weak-constraint 4D-Var

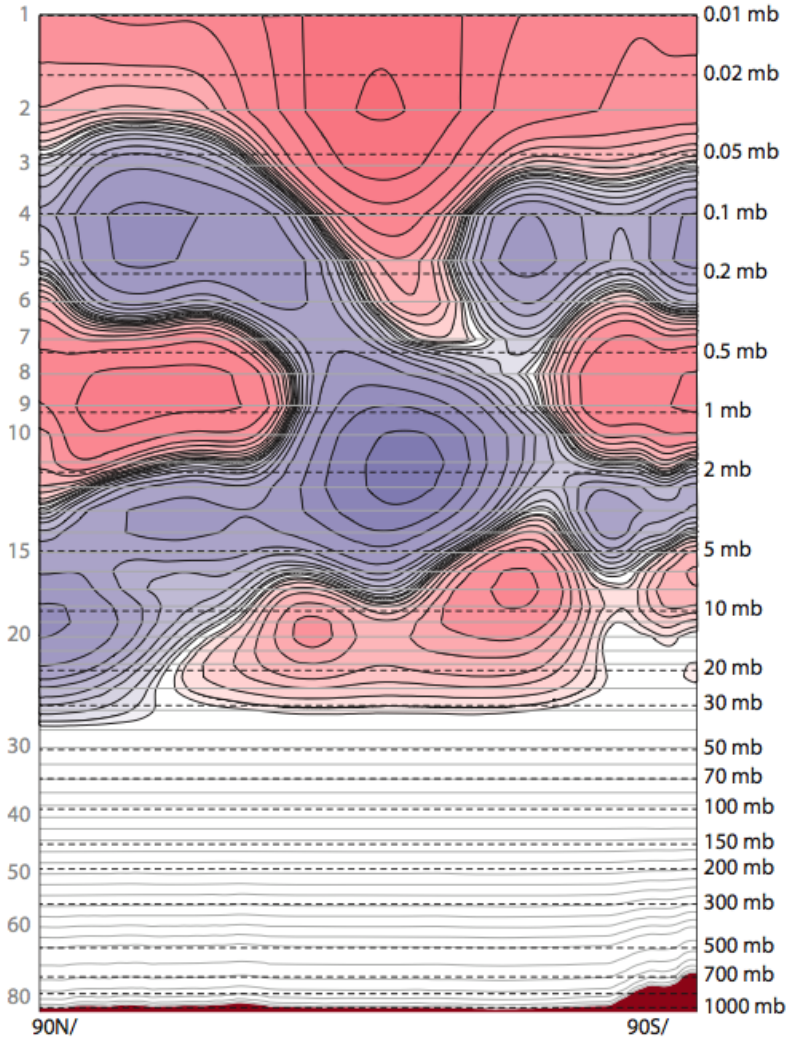
Implementation of long-window 4D-Var requires the addition of model error terms in the variational analysis equation:



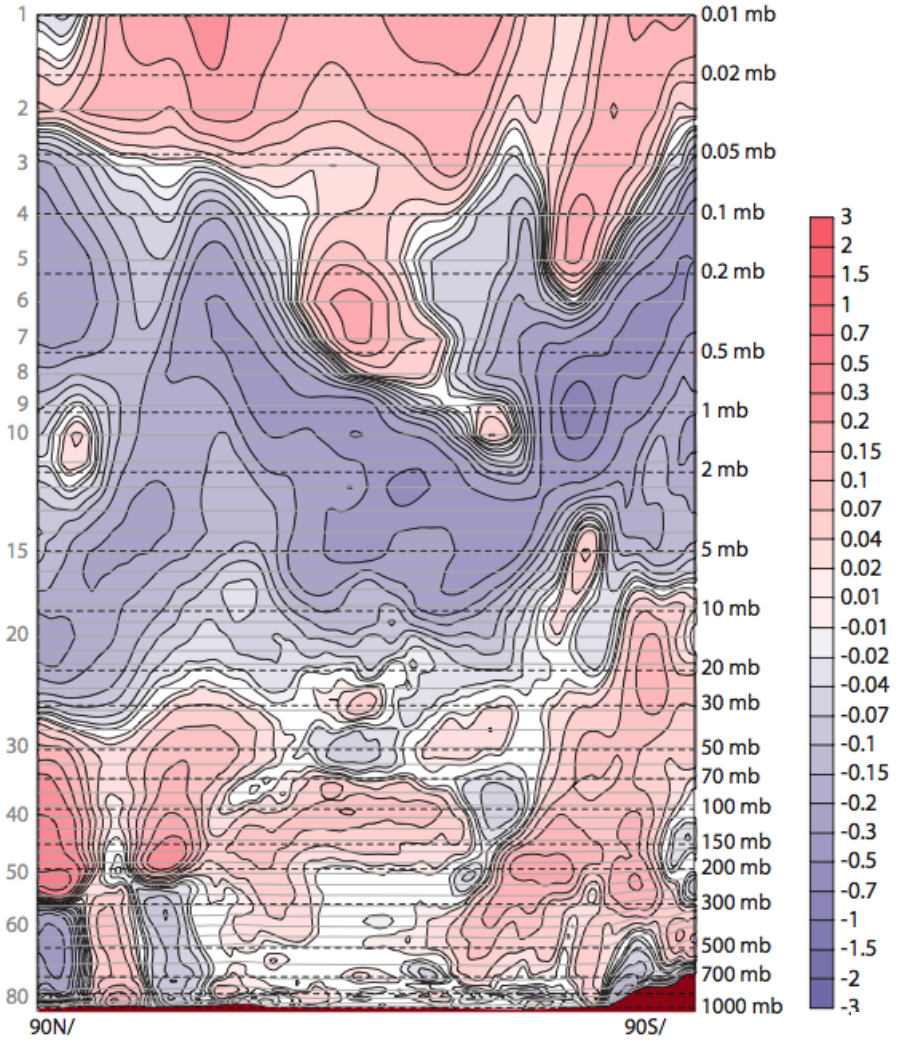
Weak-constraint 4D-Var can also be used to estimate persistent model errors.

Model bias estimates from weak-constraint 4D-Var

Estimated persistent model error [K/12h]



10-day minus 5-day forecast drift [K/day]



Controlling model bias in reanalysis

Can we use model error estimates from weak-constraint 4D-Var to control the effect of model bias in the absence of observations?

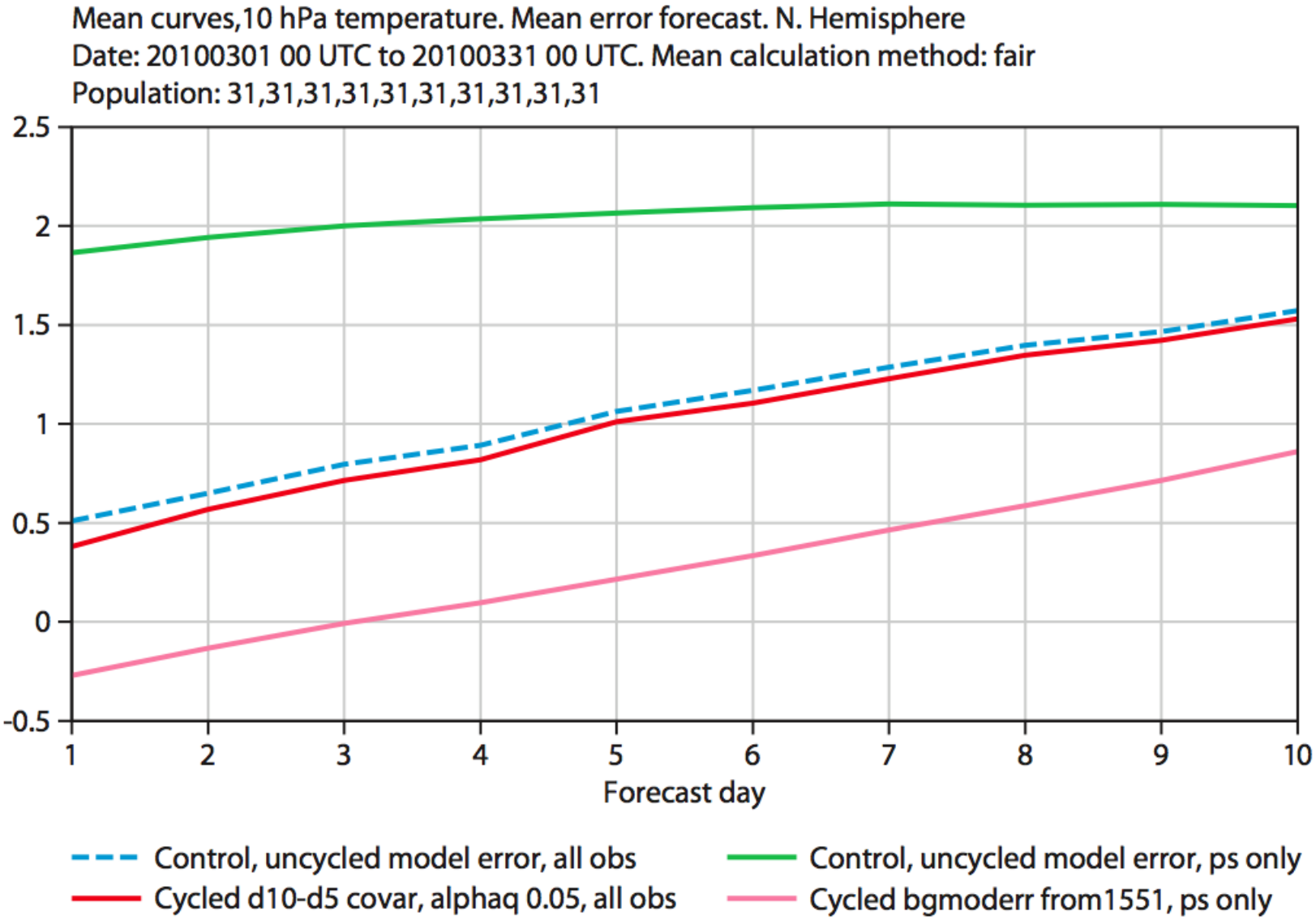
- For a given model, estimate persistent model errors in the recent (well-observed) period
- Apply these as a correction to the model in the past (poorly-observed) period

A similar approach has been successfully applied in ECMWF's ocean reanalysis *(Magdalena Balmaseda)*.

Experiment:

- Estimate persistent stratospheric model error in a fully observed system, using weak-constraint 4D-Var
- Apply the estimate to correct the model in an assimilation of surface pressure observations only

Mean forecast errors, T at 10hPa



Summary

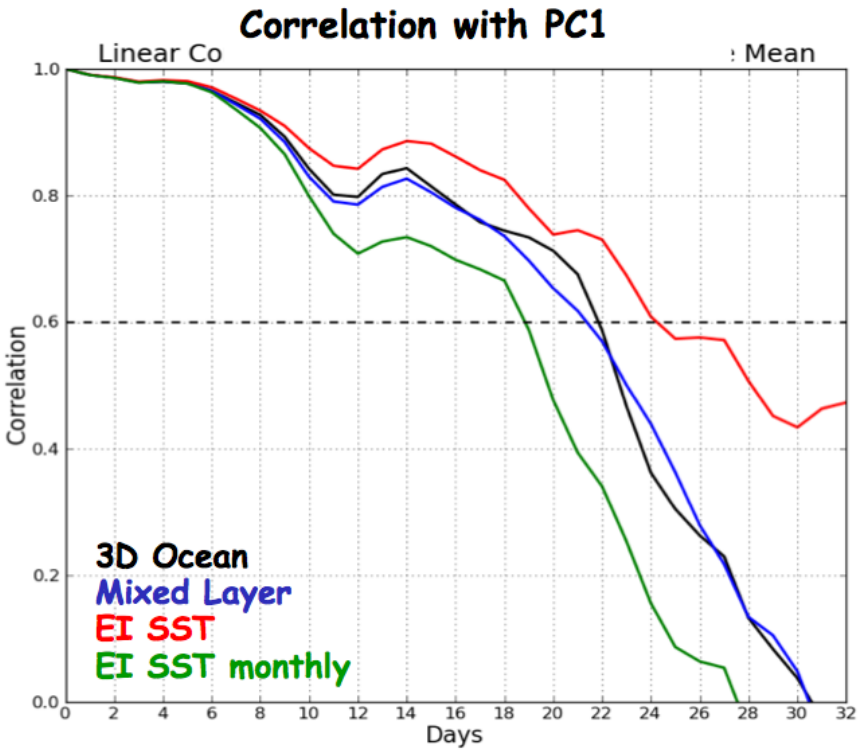
- ERA-CLIM uses the ECMWF Integrated Forecast System (IFS):
  - Ensemble of data assimilations (EDA)
  - Long-window weak-constraints 4D-Var
  - Variational bias correction of observations

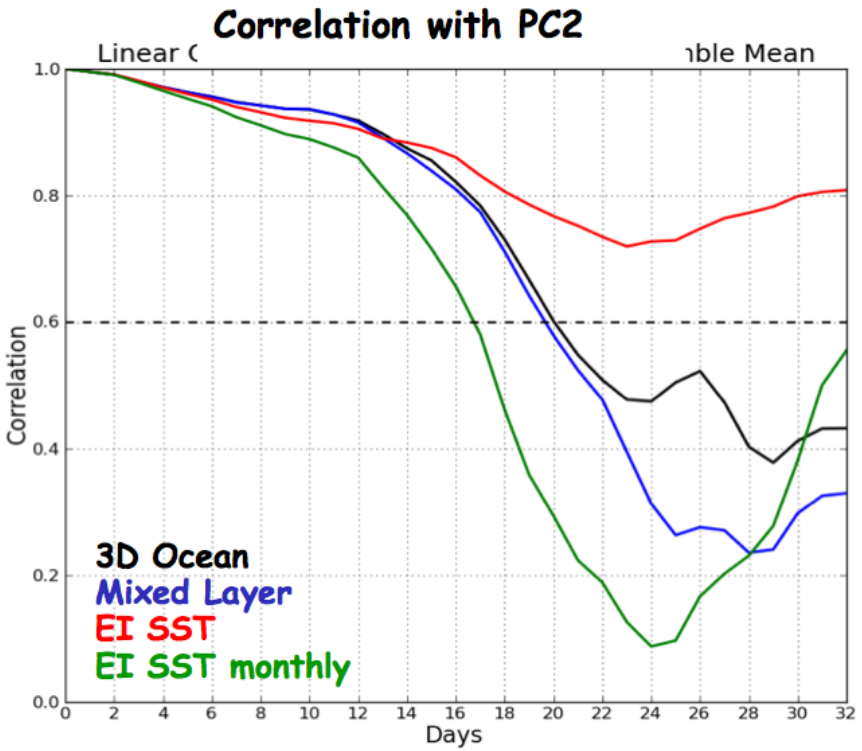
- For climate reanalysis:
  - Use prior knowledge of the full data record for QC and bias correction
  - Use the ensemble to represent key uncertainties (SST/sea-ice)
  - Configure the analysis window to make best use of sparse observations
  - Use weak-constraints 4D-Var to estimate and correct model biases

- Coupling the ocean
  - CFSR is the starting point
  - Focus on controlling model drift with SST observations

Need for a coupled system

Representation of the MJO in seasonal forecasting
(E. de Boisseson and M. Almaseda)





- Coupling does better than forcing monthly SST. Alternative before 1981 ?