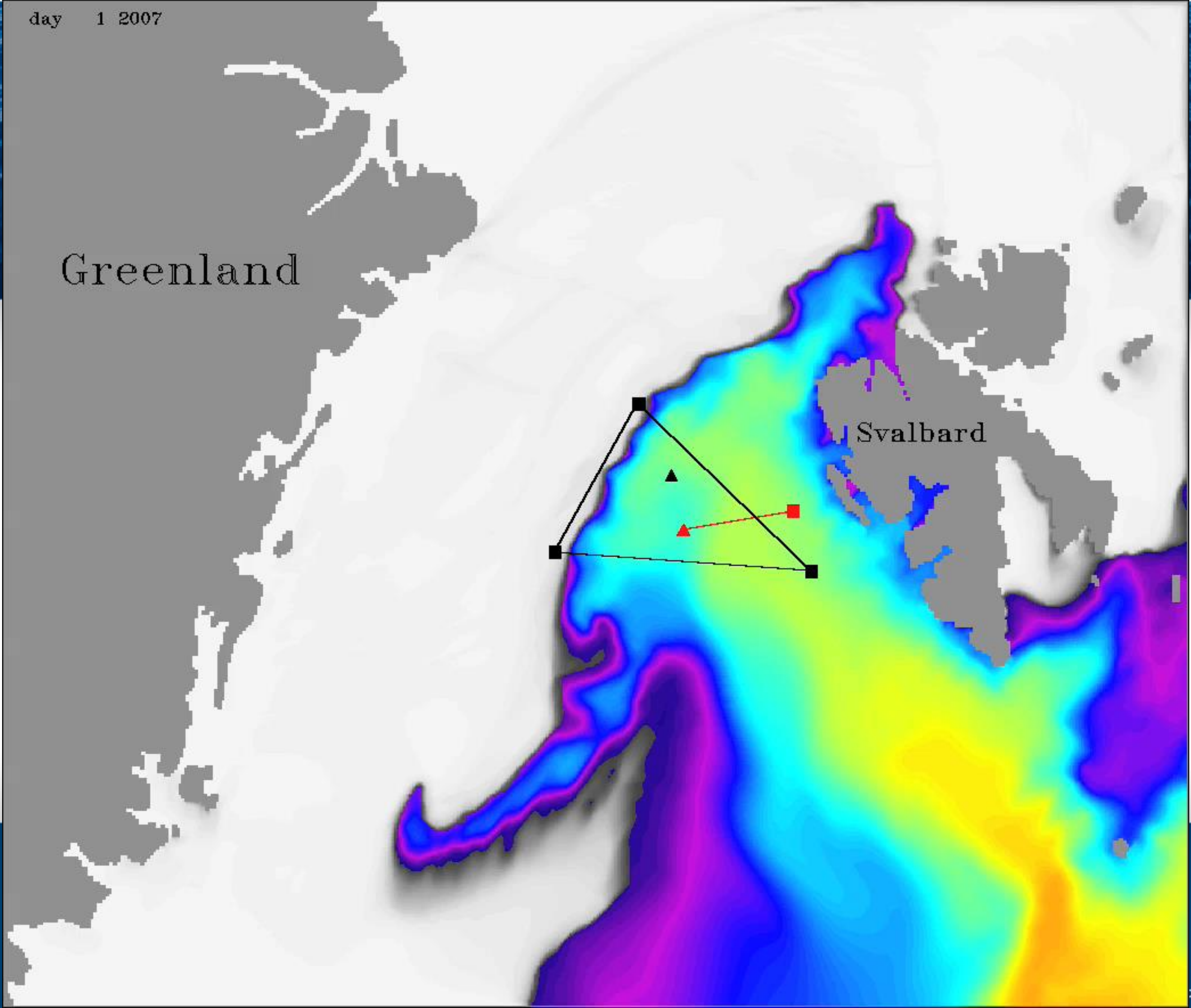
Ocean data assimilation

L. Bertino, NERSC

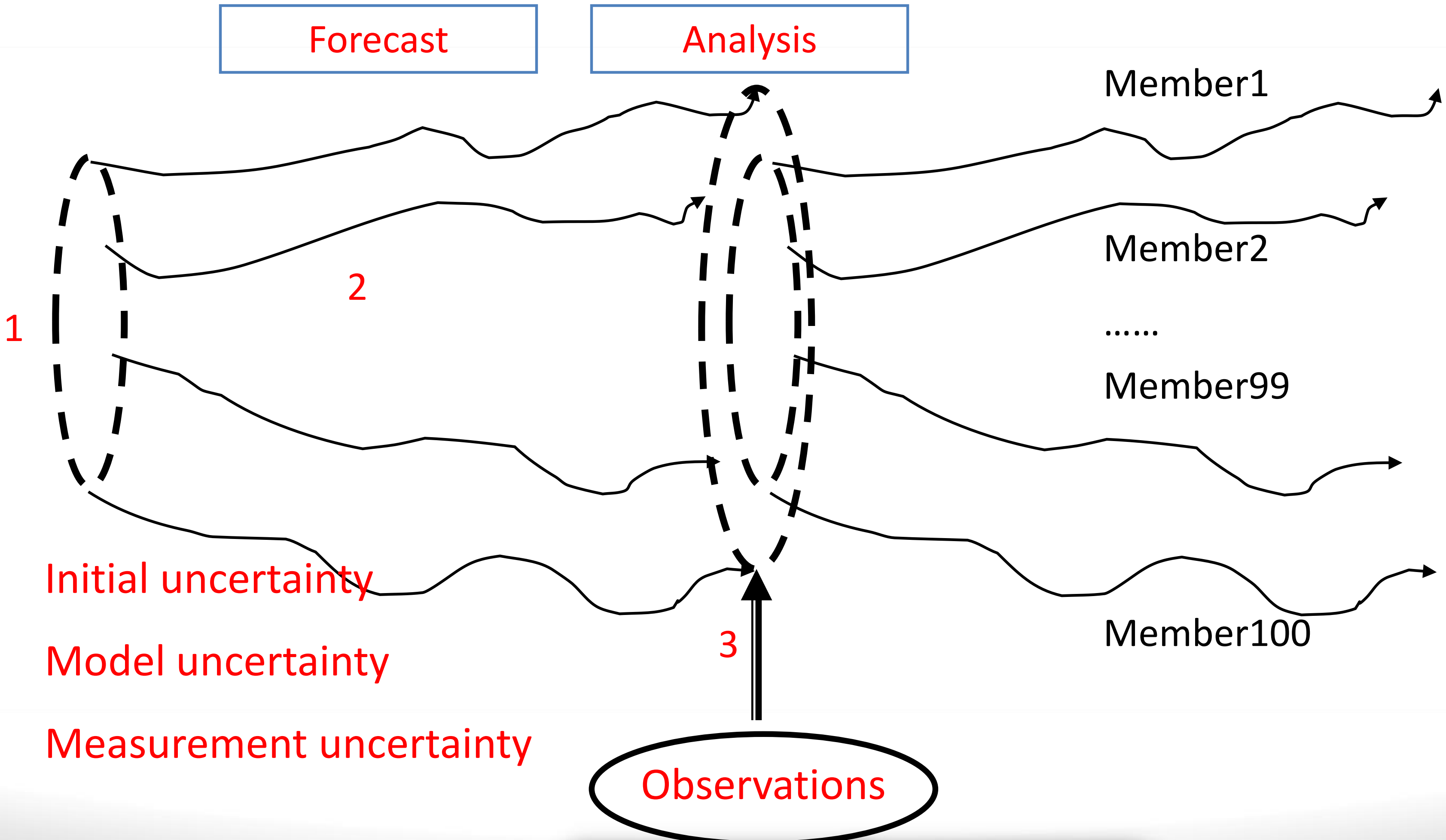
With inputs from Johnny Johannessen, François Counillon

ECMWF seminar 8th Sept 2014



day 1 2007
Greenland
Svalbard

Ensemble Kalman filtering



1. Initial uncertainty
2. Model uncertainty
3. Measurement uncertainty

Vocabulary and Analogies with 4D-VAR

- 4D-VAR
  - Cost function (quadratic)
  - Adjoint sensitivities
  - 4D assimilation
  - Optimal solution in linear cases
- Not analogous
  - Powerful iterative gradient descent
  - Strong constraint

- EnKF
  - Posterior variance (min)
  - Cross-covariances
  - Asynchronous EnKF
  - Optimal solution in linear cases
    - but sampling errors
- Not analogous
  - Monte-Carlo framework
  - Explicit model errors

MyOcean

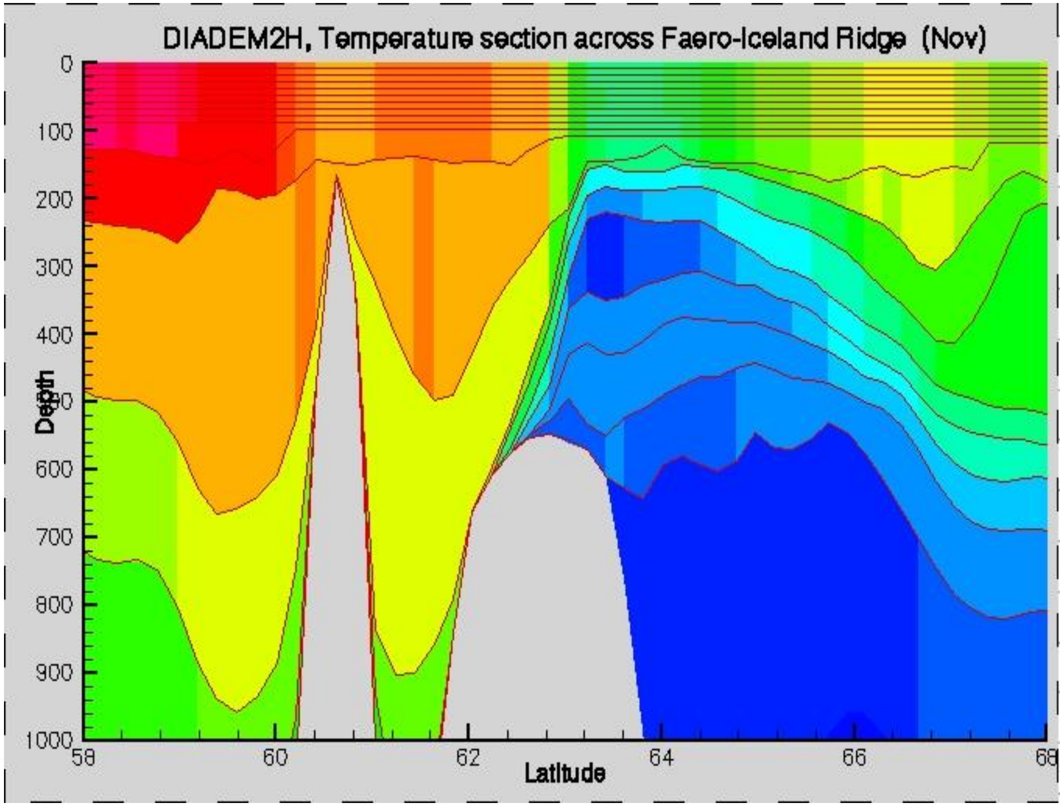
GMES Marine Service

The MyOcean “Tordesillas“

1. Global Modeling and Forecasting Center
   - Lead Mercator
     - NEMO + fixed based SEEK filter
2. Arctic Modeling and Forecasting Center
   - Lead developments NERSC
   - Exploited operationally at MET Norway
   - Based on the TOPAZ system
     - HYCOM + EnKF
3. until 7, see http://myocean.eu

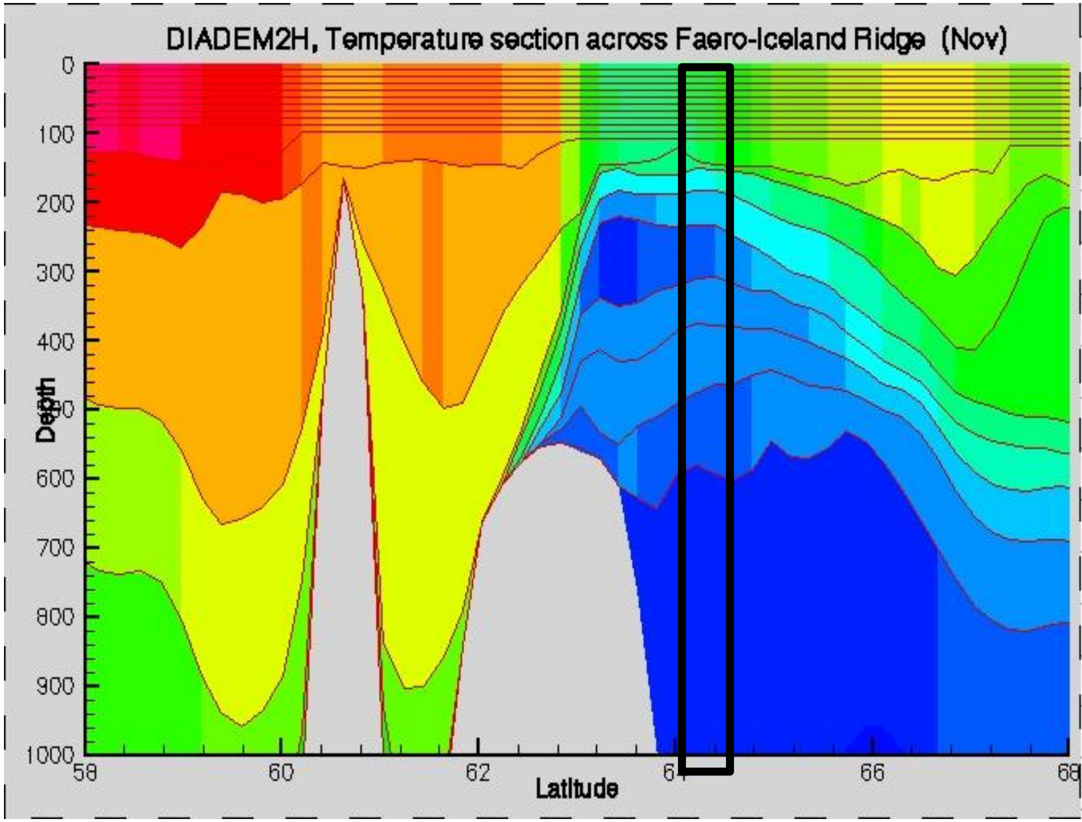
The HYCOM model at NERSC

- 3D numerical ocean model
  - Hybrid Coordinate Ocean model, HYCOM (U. Miami)
- Hybrid vertical coordinate
  - Isopycnal in the interior
  - Z-coordinate at the surface
  - TOPAZ4 uses 28 layers
- Coupling to sea ice model
  - EVP dynamics …
  - Semtner Thermodynamics
- Data assimilation:
  - EnKF (probabilistic) …



The state vector *X*

- 3D variables
  - Temperature
  - Salinity
  - Layer thickness (can be zero)
  - X-current
  - Y-current
- 2D variables
  - Sea ice area
  - Sea ice thickness
  - Snow depths
  - Barotropic currents + pressure
- Typical grid size
  - Horizontal: 800x880
  - Vertical: 28
  - **Total unknowns: ~10^8**
    - **Need to perform *local* analyses**



Evensen 2002

Computations

DEnKF 100 members

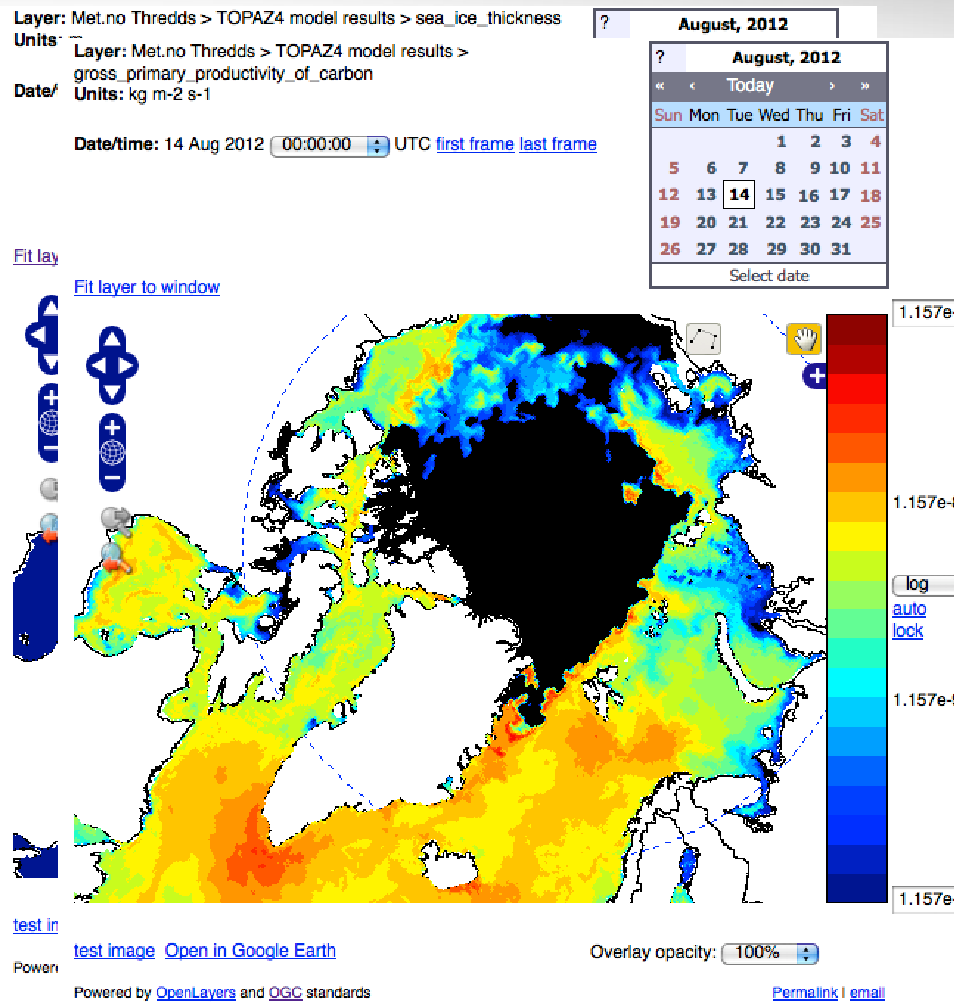
- **Ensemble Forecast**
- 2500 CPU hours / cycle
- Embarrassingly parallel
- 100x **133 CPU 11 min** jobs
- Each job requires **400 Mb**
  - MPI parallelization

- **Analysis**
- 20 CPU hours / update
- 6 datasets simultaneously
- One **20 CPU 1h job**
- Memory required **1 Gb**
  - MPI parallelization

- HPC Machine:
- Cray XE6m, updated 2012
  - 22272 cores, 205 Tflop/s
  - 676 nodes (32-cores)
  - 1-4 Gb per node

The TOPAZ system

- Exploited operationally at met.no
  - Since 2008
  - Ecosystem added in Jan. 2012
- 20 years reanalysis at NERSC
  - Took 2 years to produce
  - 3-years ecosystem reanalysis
- MyOcean (Arctic MFC)
  - Free distribution of data
  - Dynamical viewing (Godiva2)
- Data used by ECMWF wave model (J. Bidlot)
  - Sea ice edge forecast
  - Surface currents



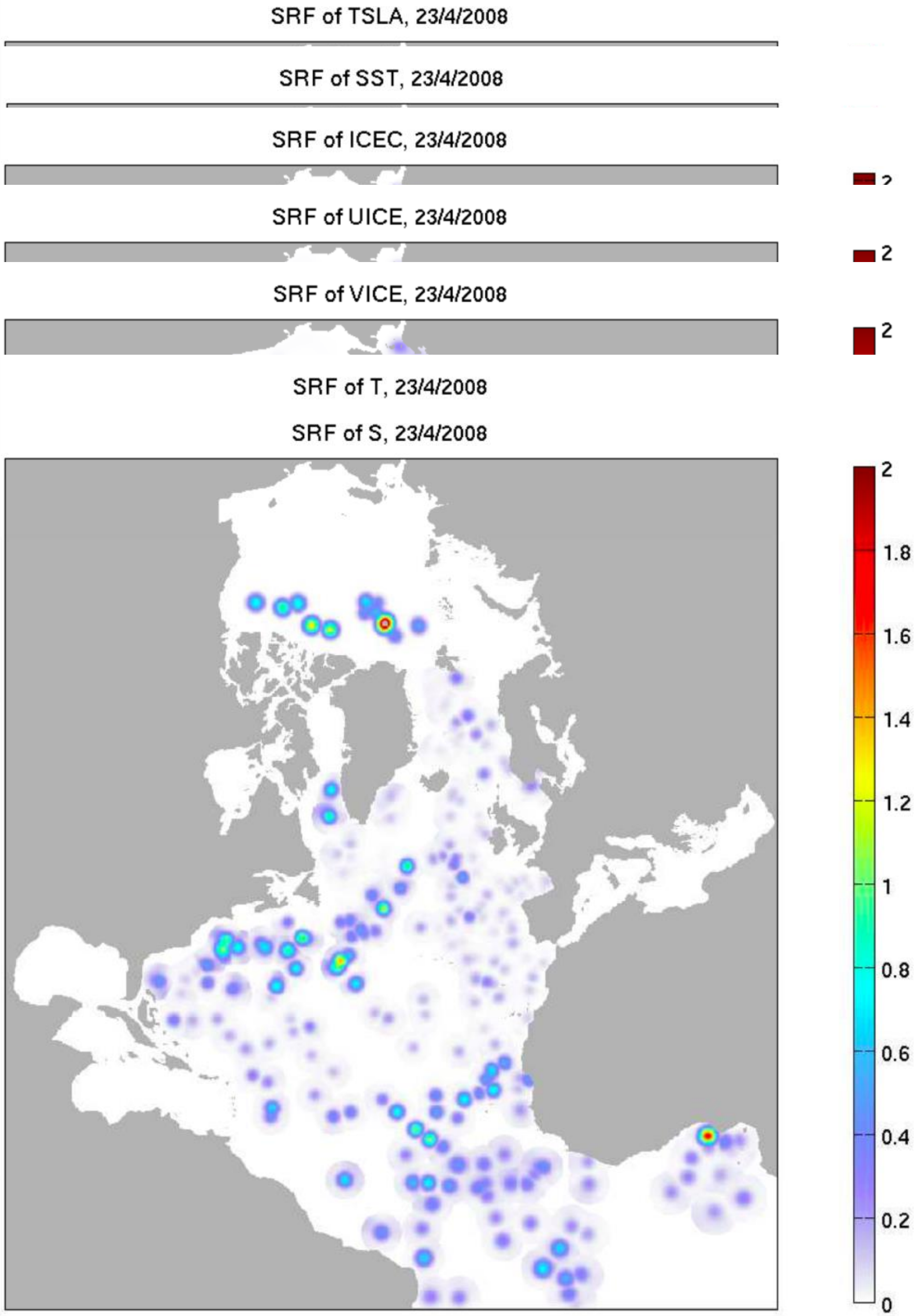
Ice thickness forecast for 14th Aug. 2012

TOPAZ Assimilation

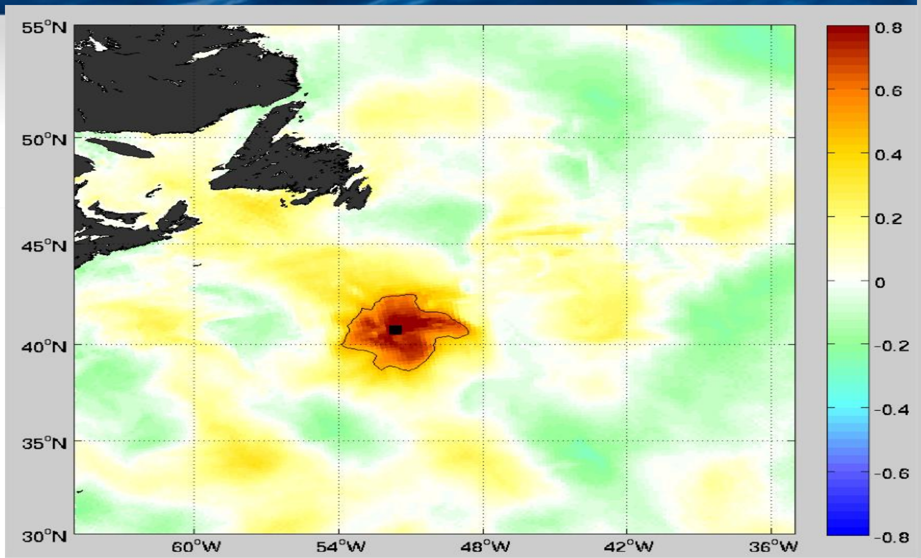
- DEnKF, asynchronous
  - 100 members
  - Local analysis (~90 km radius)
  - Ensemble inflation by 1%
- Observations:
  - Sea Level Anomalies (CLS)
  - SST (NOAA, then UK Met)
  - Sea Ice Concentr. (OSI-SAF)
  - Sea ice drift (CERSAT)
  - T/S profiles (Coriolis)
  - **400.000 observations** per week
  - ~100 in each local radius

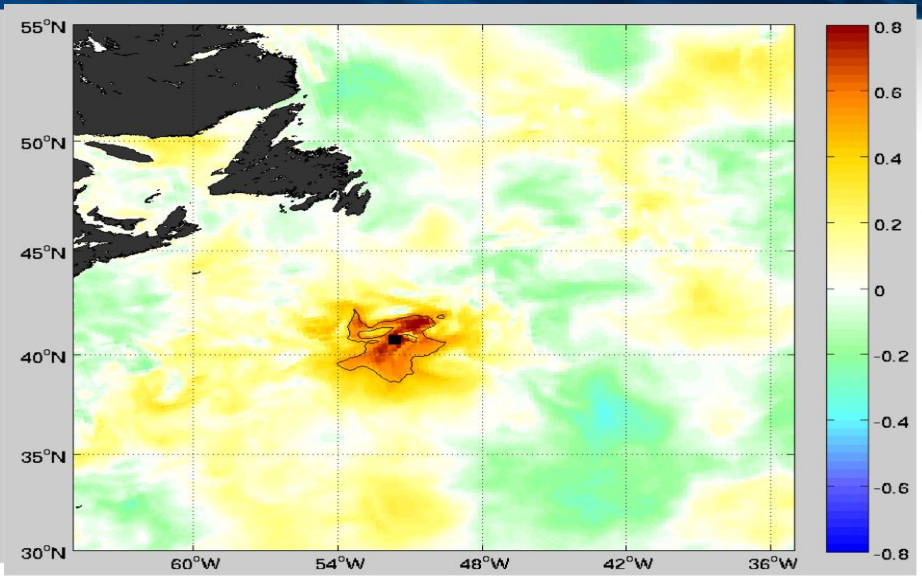
SRF: local spread reduction factor

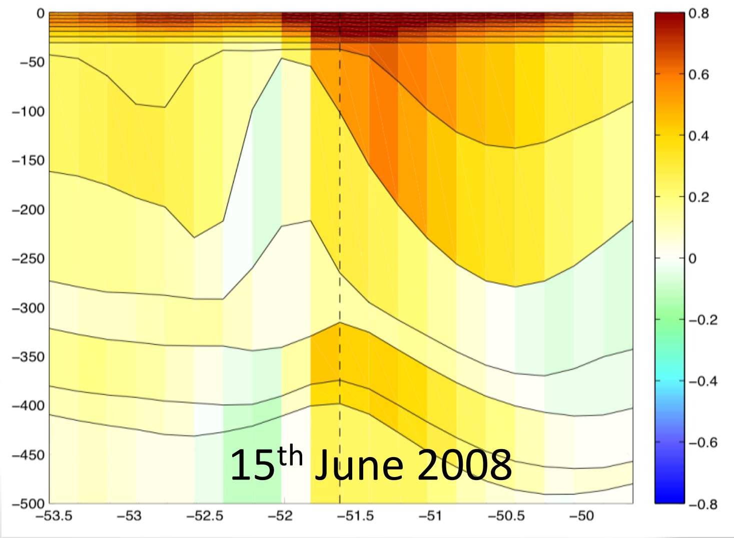
$$\mathrm{SRF} = \sqrt{\frac{\mathrm{tr}(\mathbf{H}\mathbf{P}^f\mathbf{H}^\mathrm{T}\mathbf{R}^{-1})}{\mathrm{tr}(\mathbf{H}\mathbf{P}^a\mathbf{H}^\mathrm{T}\mathbf{R}^{-1})}} - 1$$

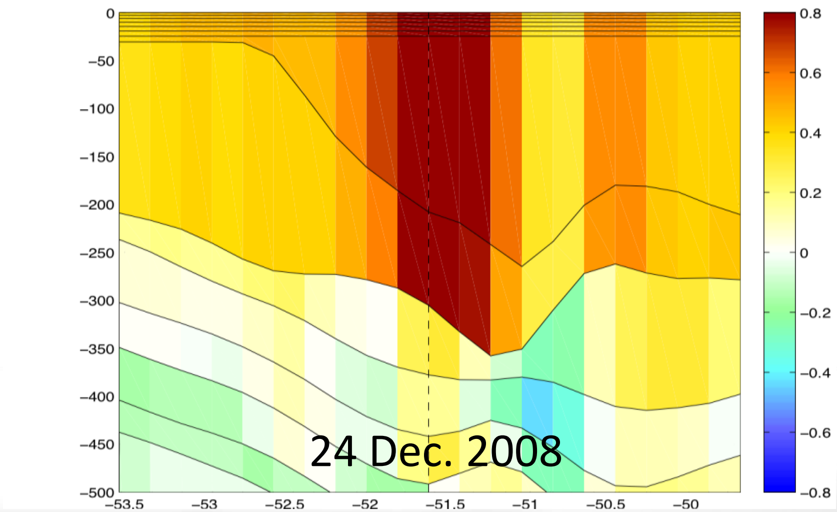


EnKF Correlations, SST



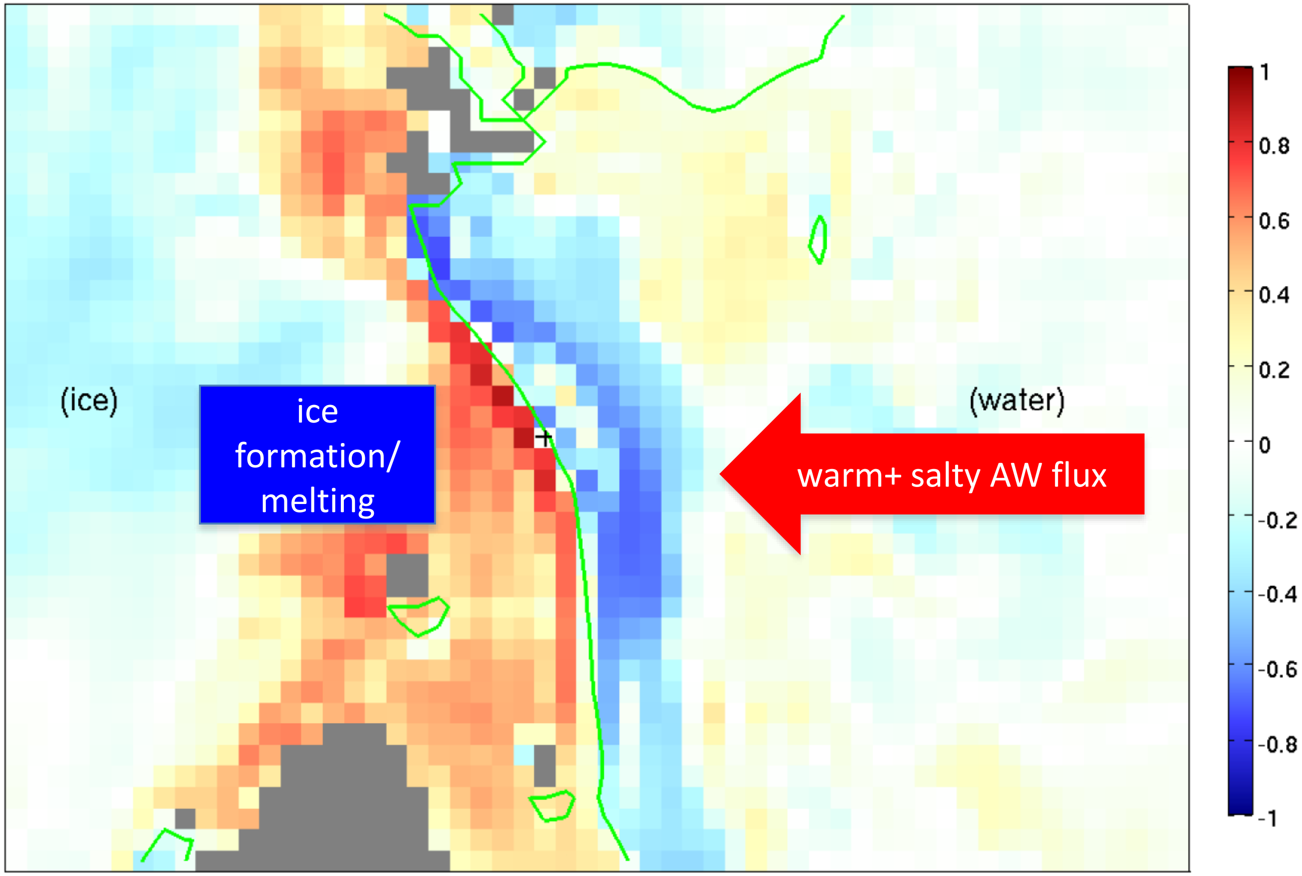






Why dynamic Data Assimilation in the Arctic?

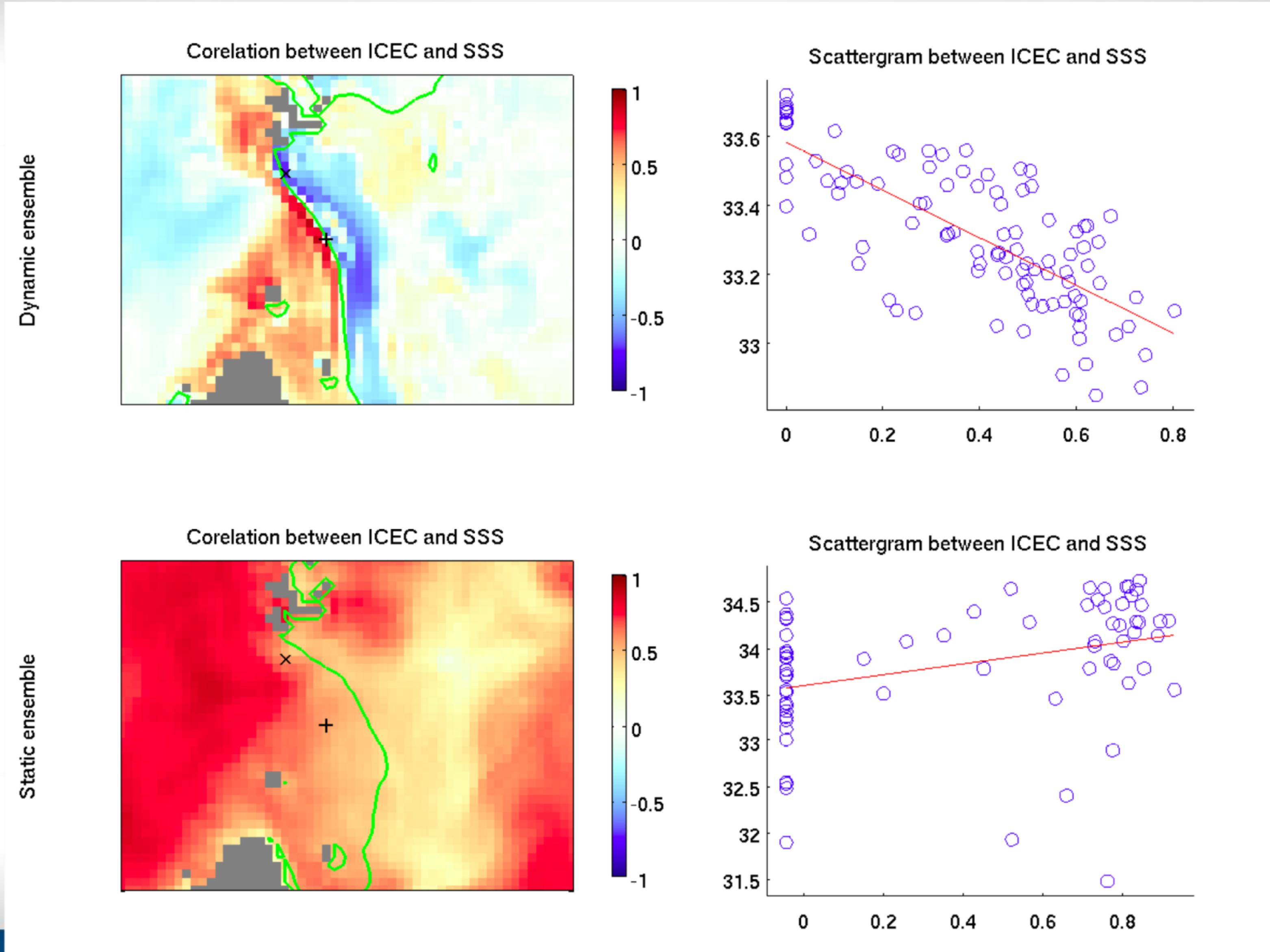
Example of ice-salinity correlations in the Barents Sea



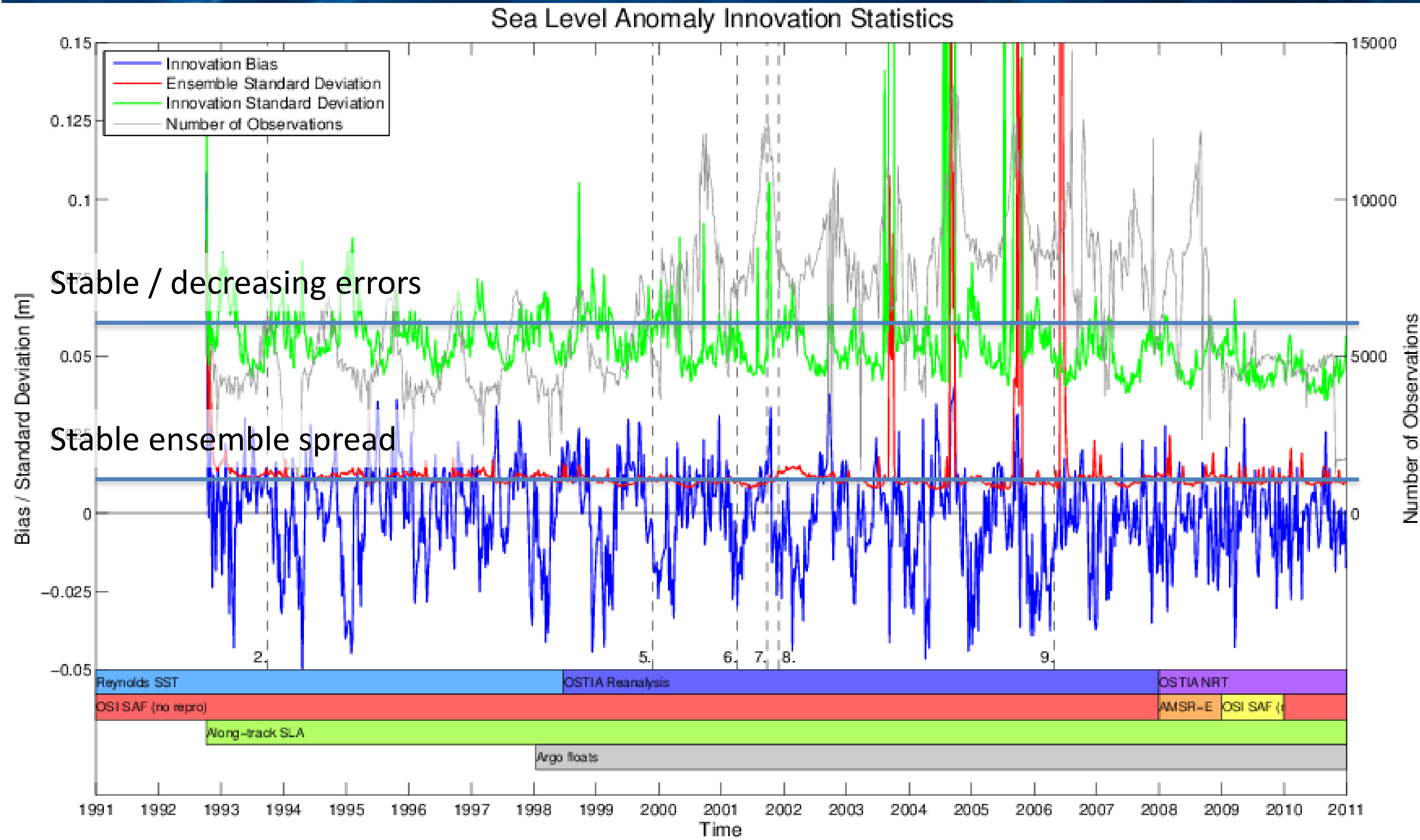
Sakov et al., the TOPAZ4 system, OS 2012

Also see *Lisæter et al.* Oc. Dyn. 2003

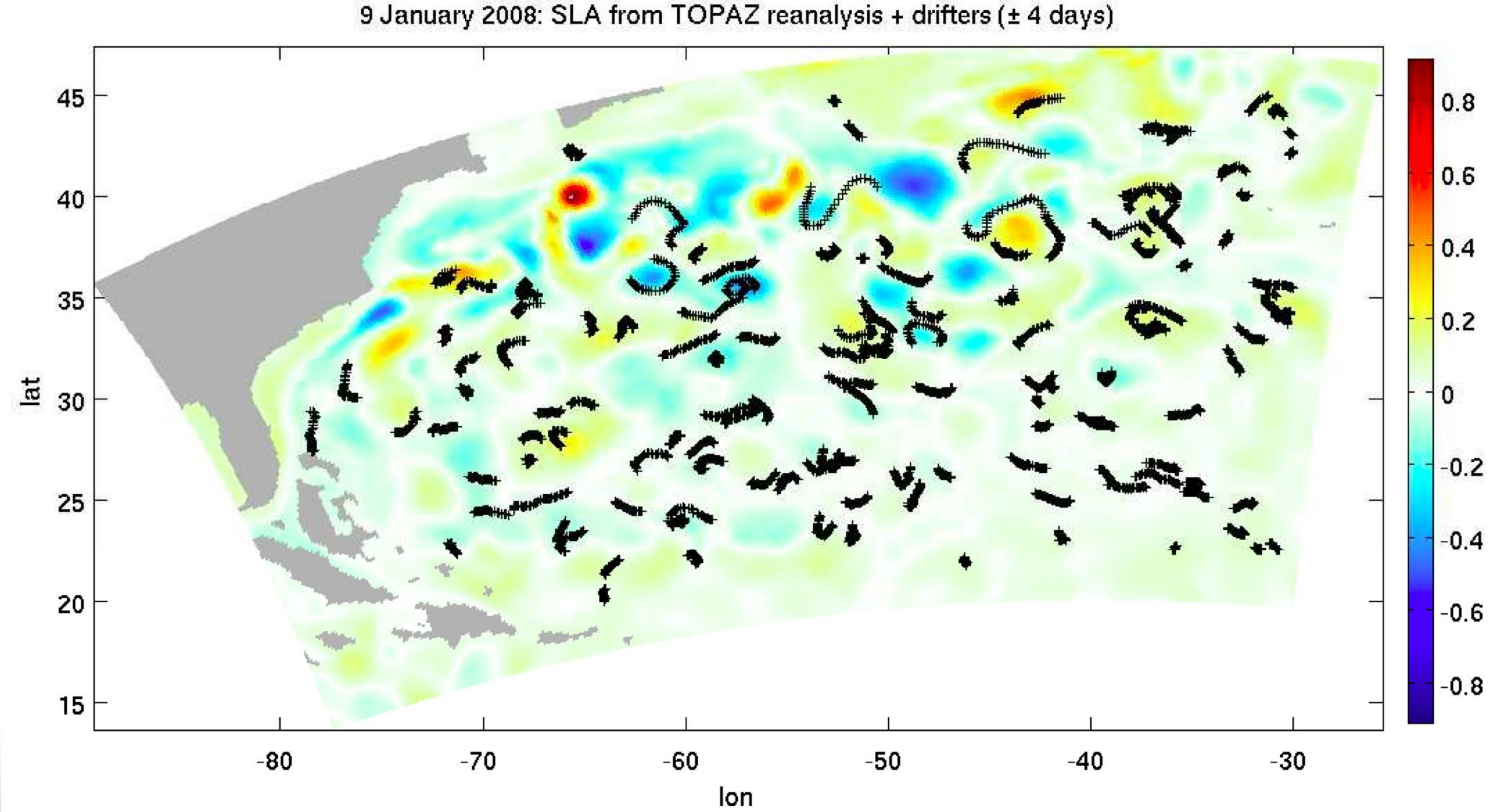
Comparison to static / climatological covariances



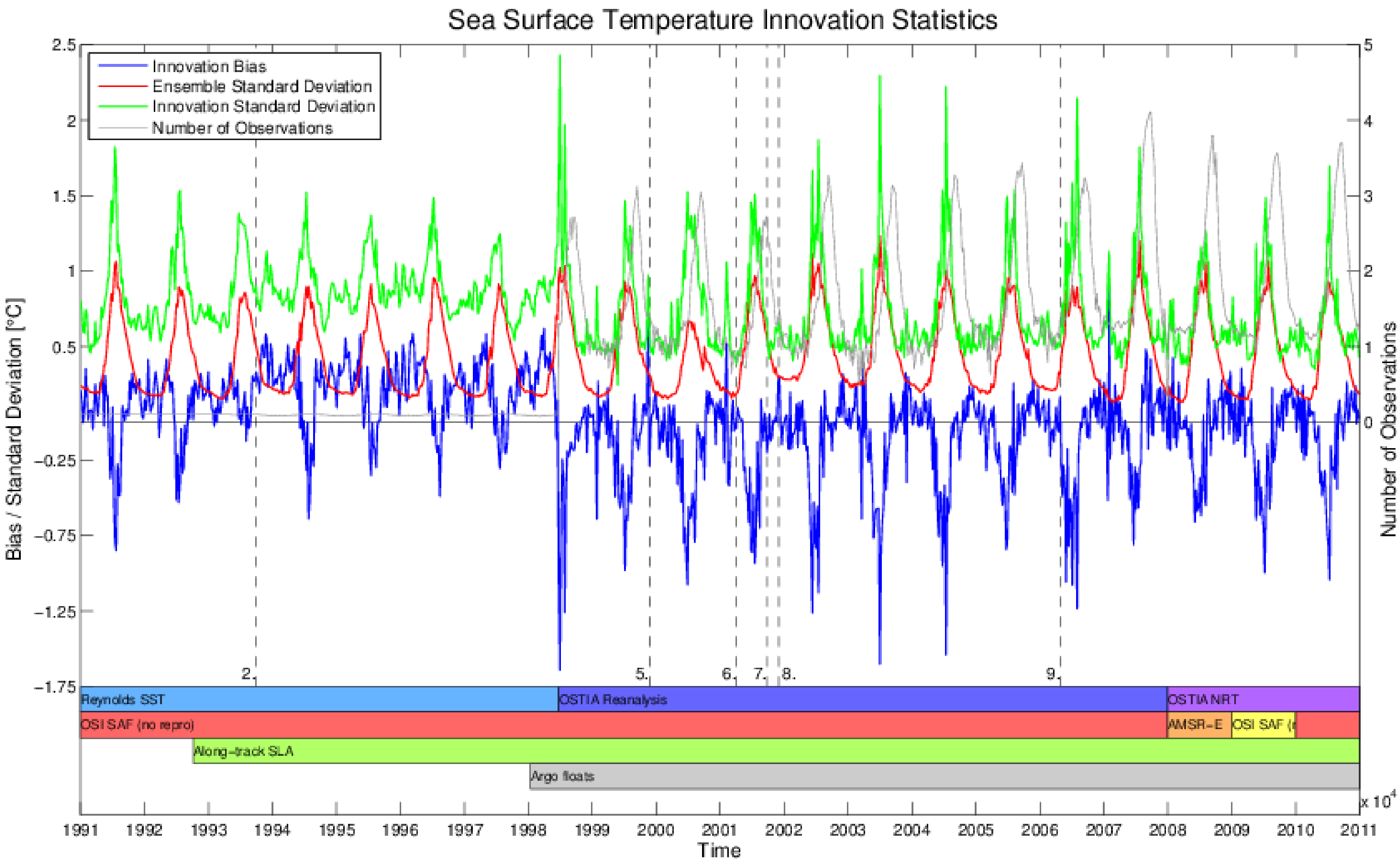
Data assimilation statistics SLA



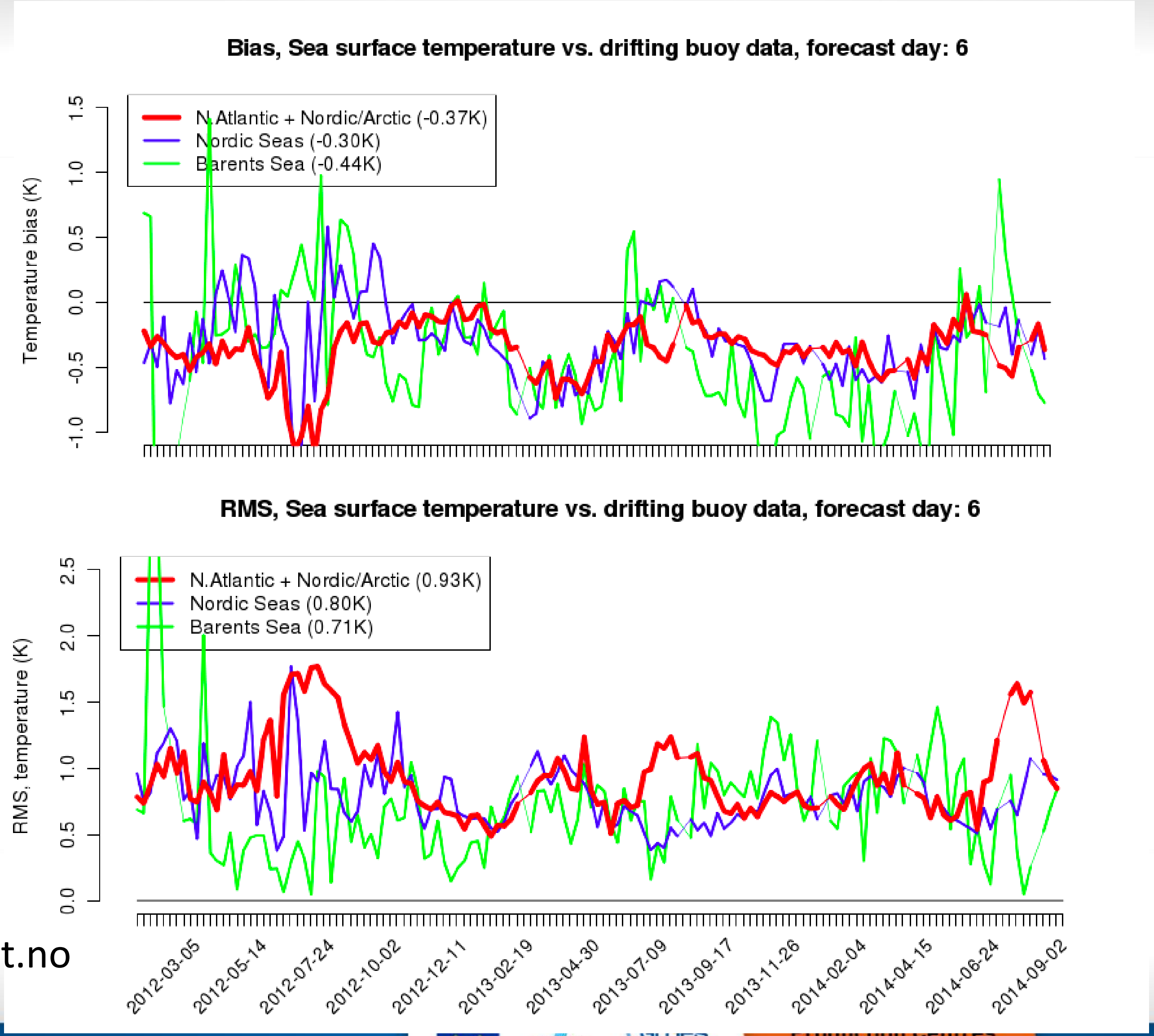
Independent data: surface drifters



Data assimilation statistics SST



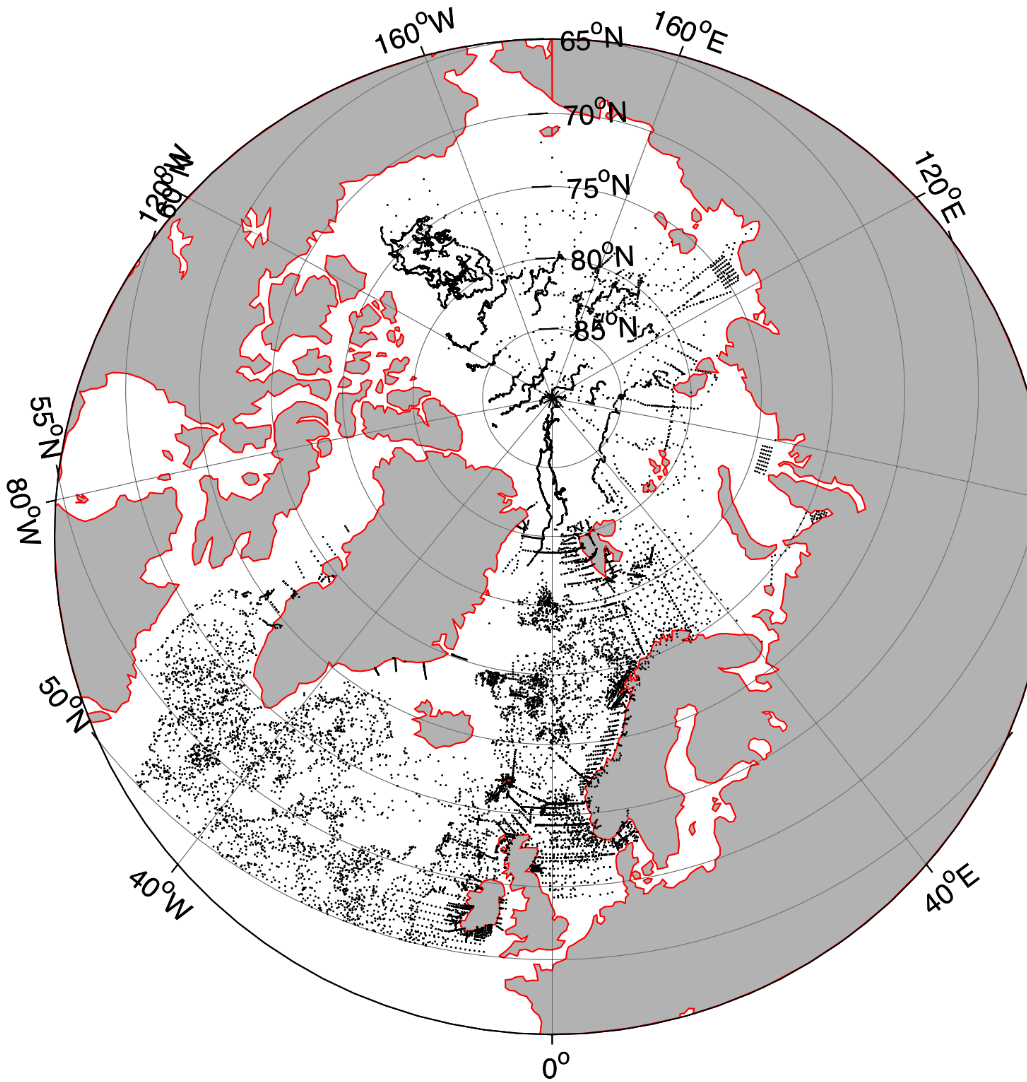
SST forecasts in real-time



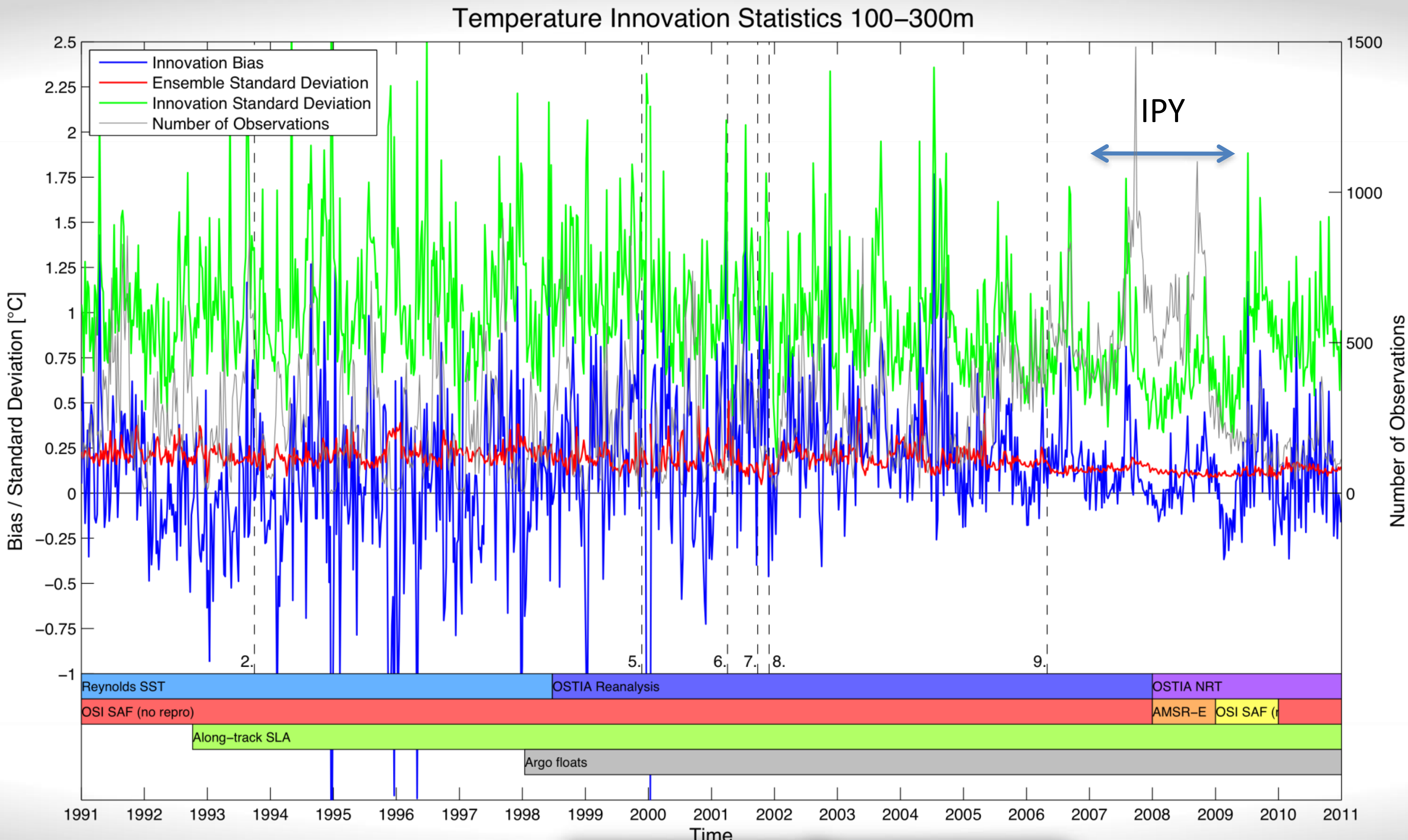
Myocean.met.no

In situ profiles assimilated

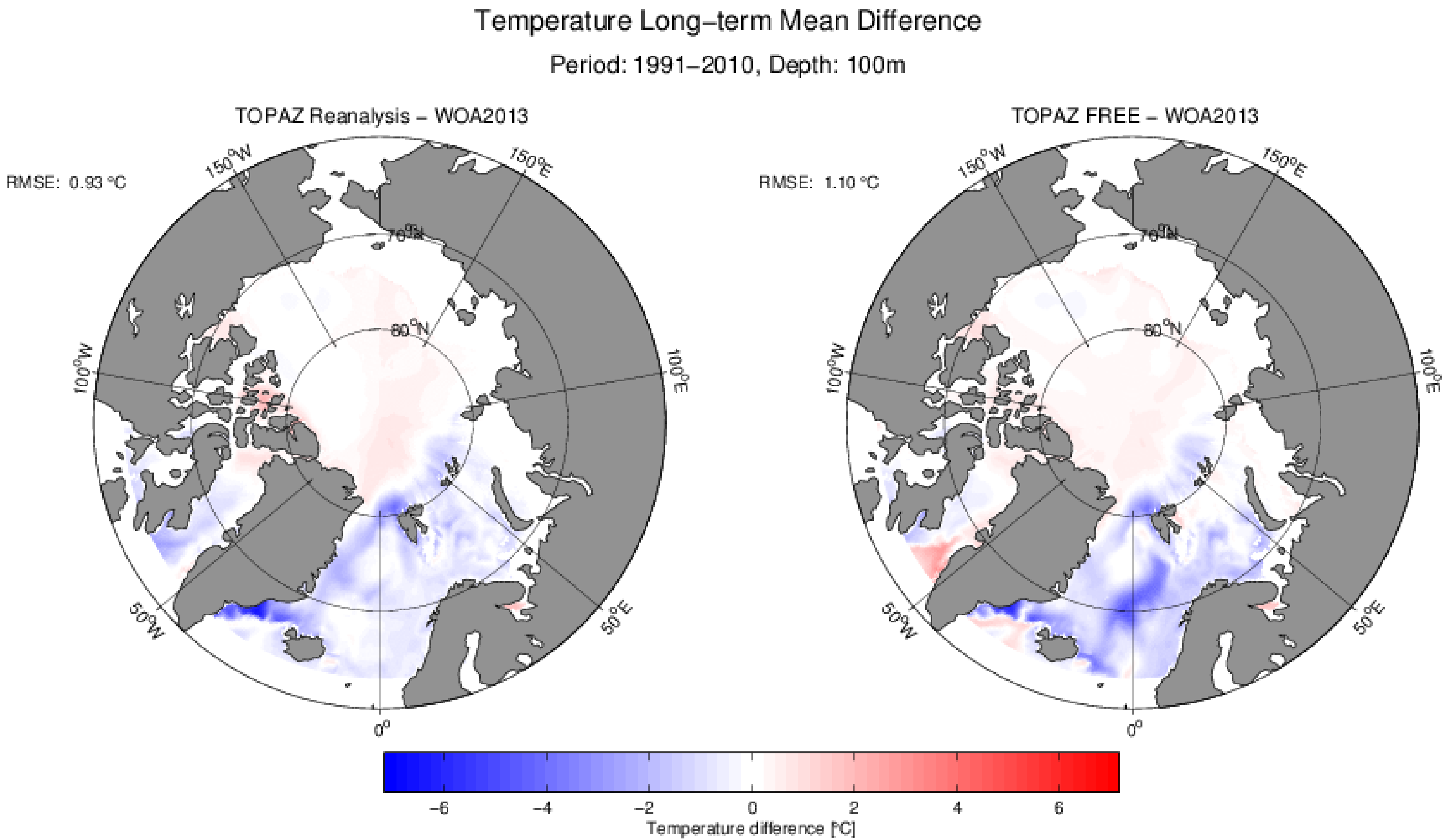
- A "Good period" 2003-2008
  - Argo floats
  - Sections
  - Ice-Tethered Profilers from Damocles IPY
  - All reprocessed quality controlled data
  - Not all profiles contain salinity
- Still very poor coverage compared to atmosphere



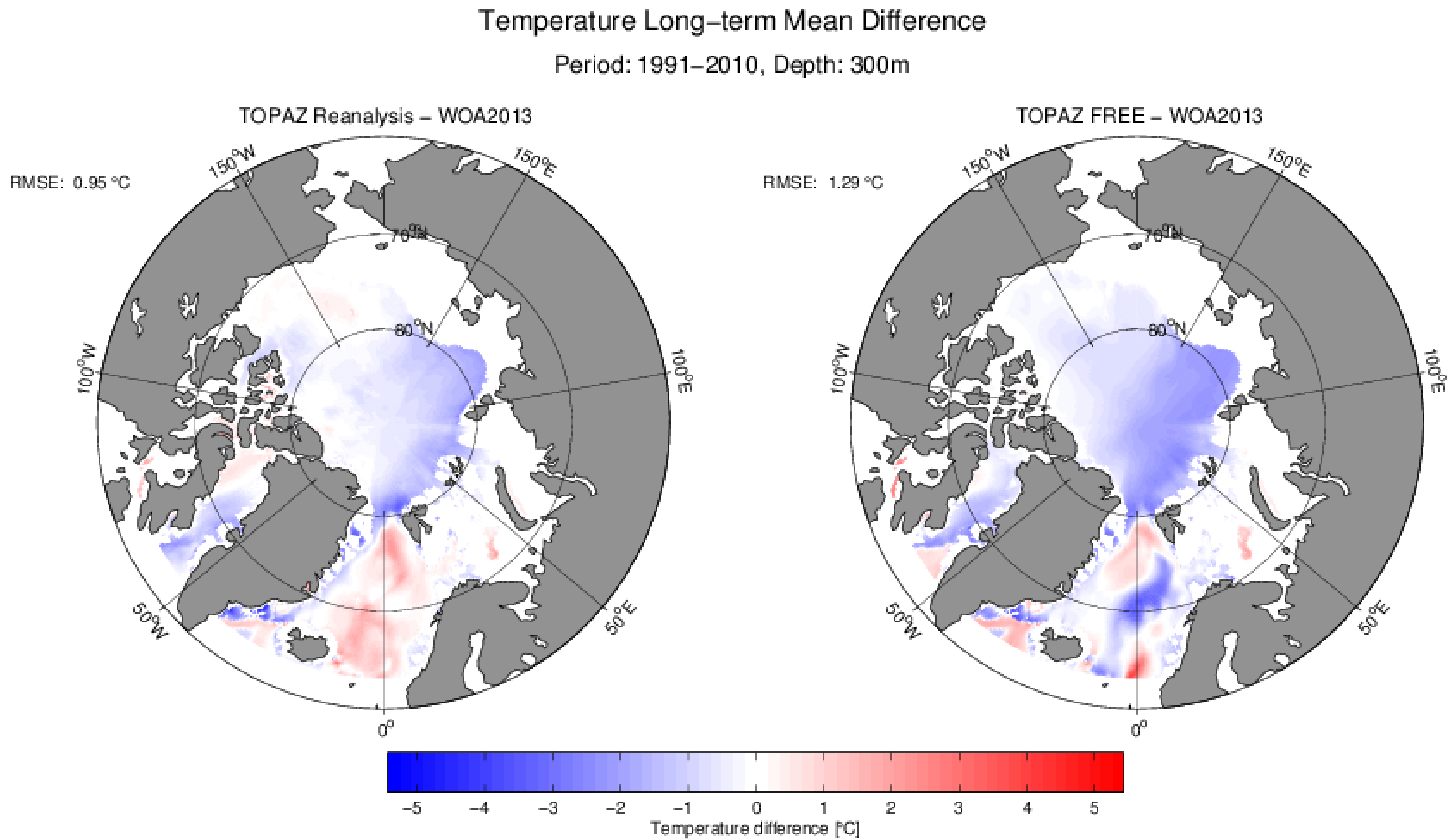
Data assimilation stats T100-300



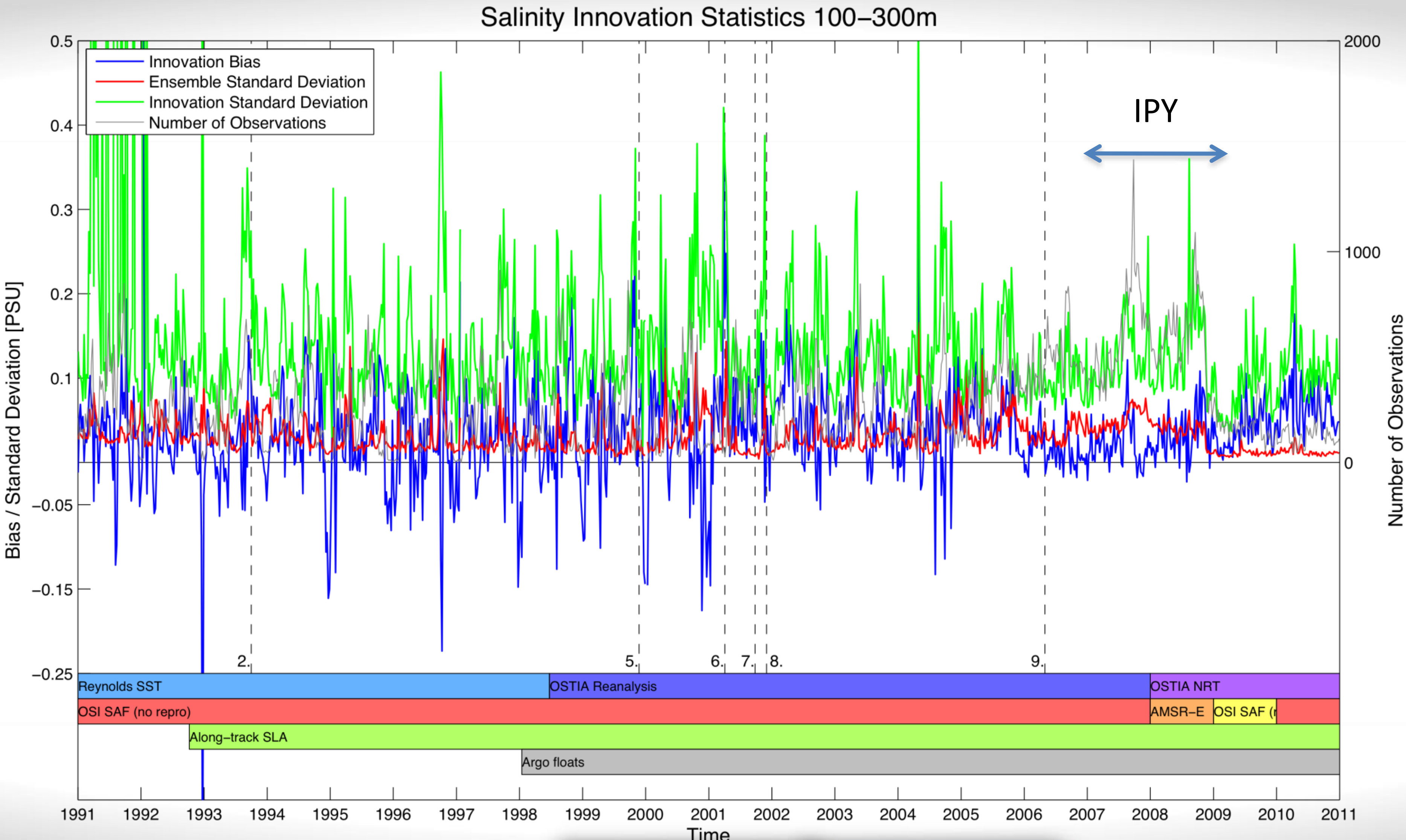
TEM biases at depths



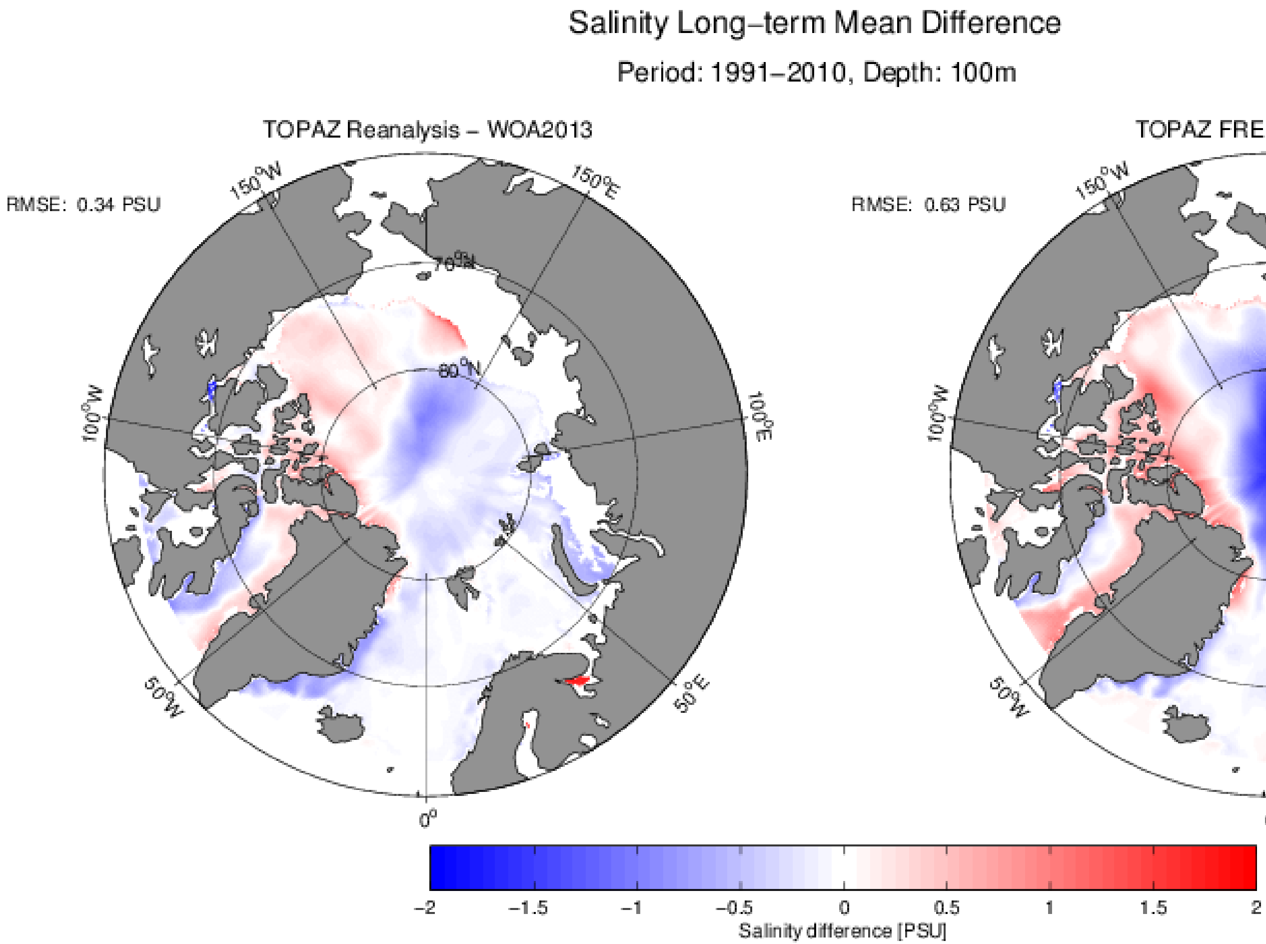
TEM biases at depths



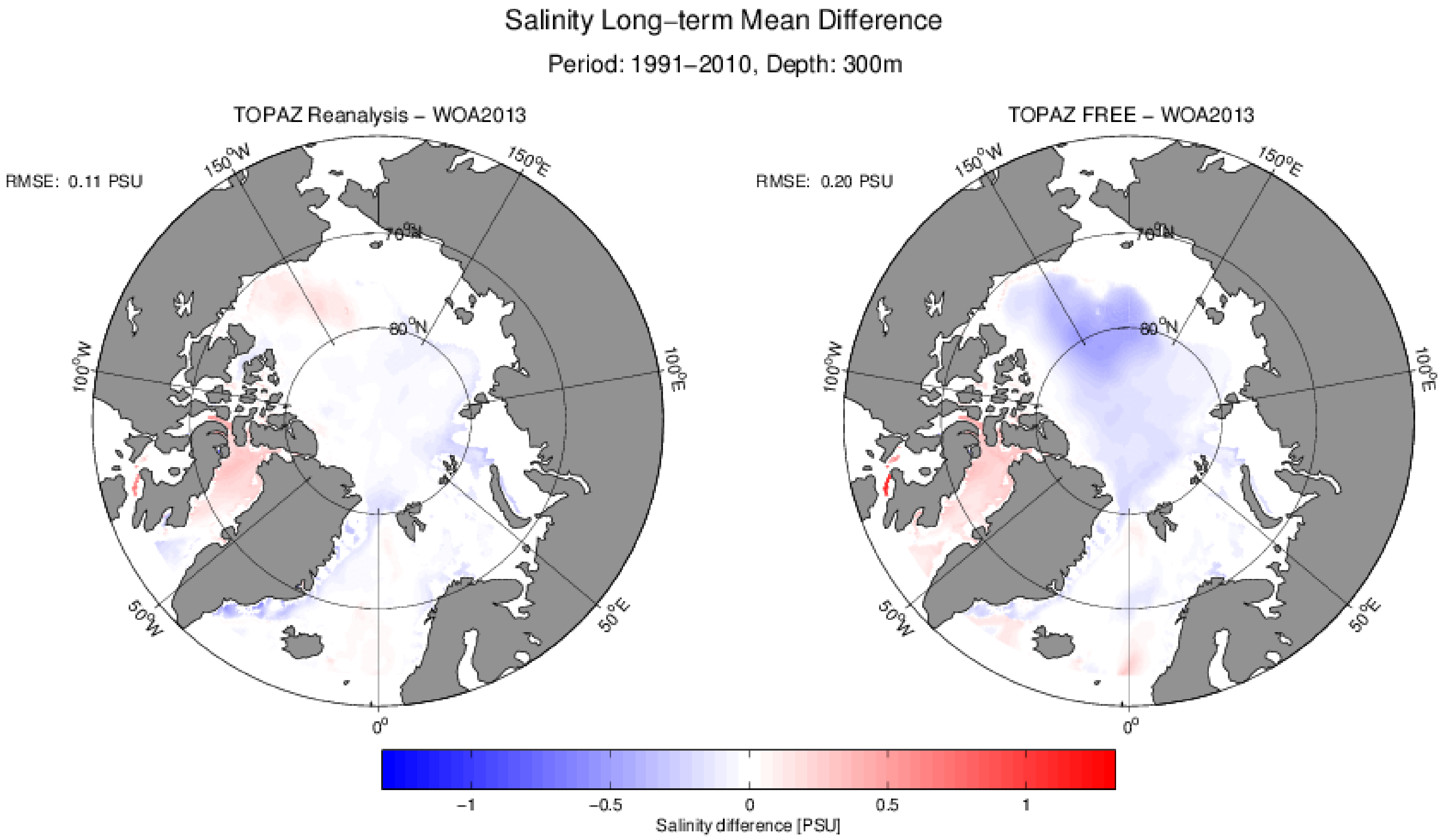
Data assimilation stats S100-300



Salinity bias at 100m depths



Salinity bias at 300m depths

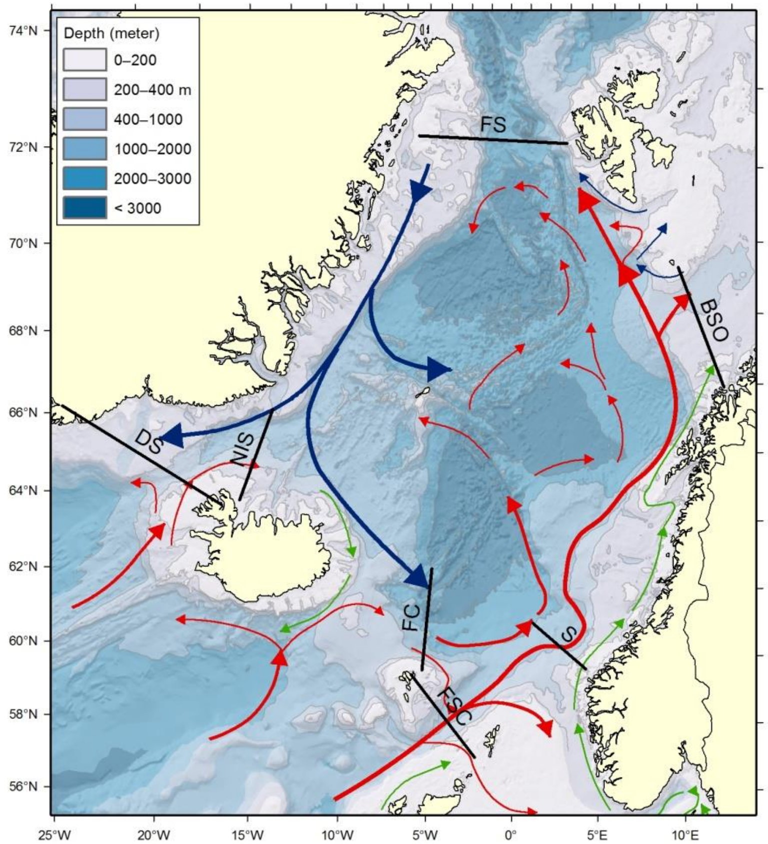


Validation of 1993-2009 reanalyses

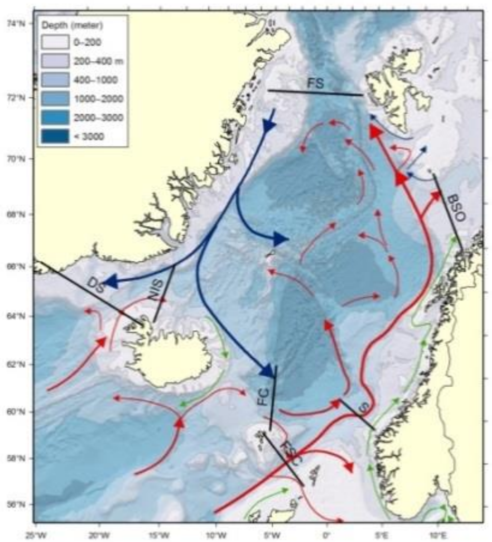
Solfrid Hjøllo, Vidar Lien, Morten, Henning, Einar, Gilles, Francois

TASK

- Validation of 1993-2009 reanalyses, focus on vol & heat fluxes, hydrography
- Global / Arctic MFC / (ROMS)
- Monthly means ,both free and assimilated runs
- Mean, std, seasonal cycle and trends



Færøy Shetland



	NE	SW	Net (Atl. Inflow)
Surface-to-bottom (Sv)	3.5	−3.1	0.4
Warmer than 5 °C (Sv)	3.5	−0.8	2.7
Colder than 5 °C (Sv)	0.0	−2.3	−2.3
Relative Heat Transport (TW)	131	−24	107
Salt Transport (10^6 kg s^{-1})	125	−27	98

Atlantic water T>5°C, S>35.0

Berx et al 2013

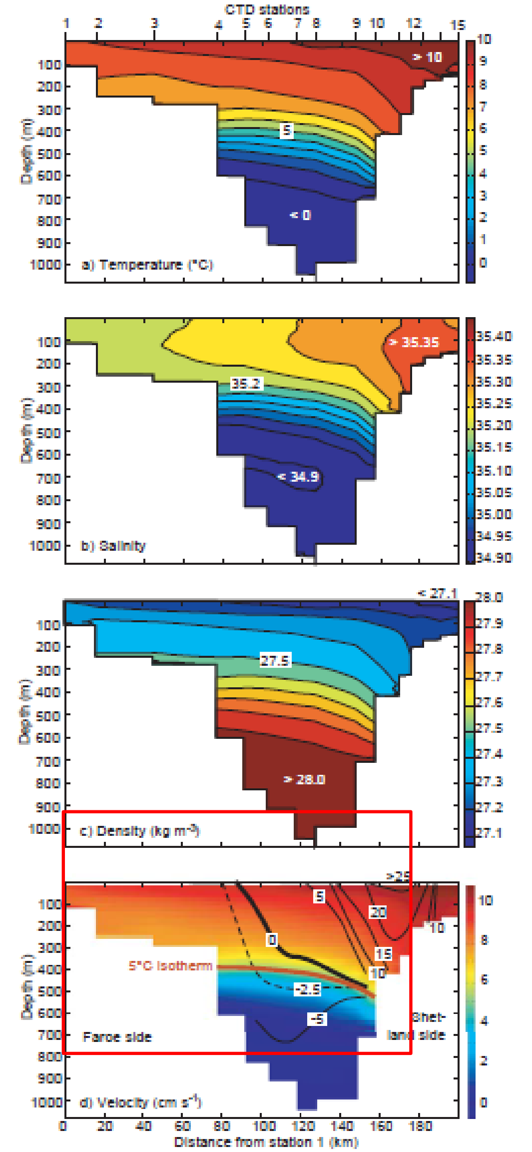
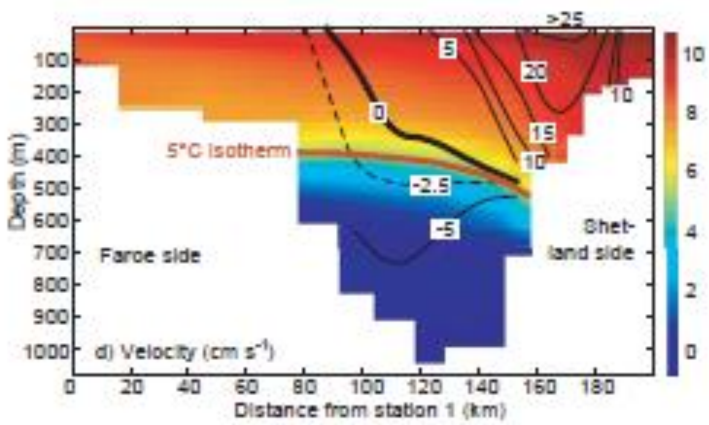
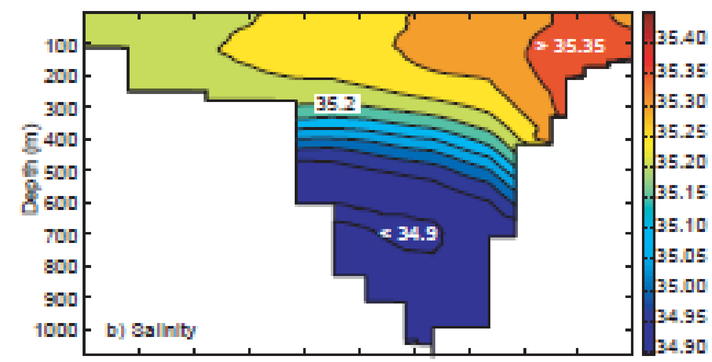
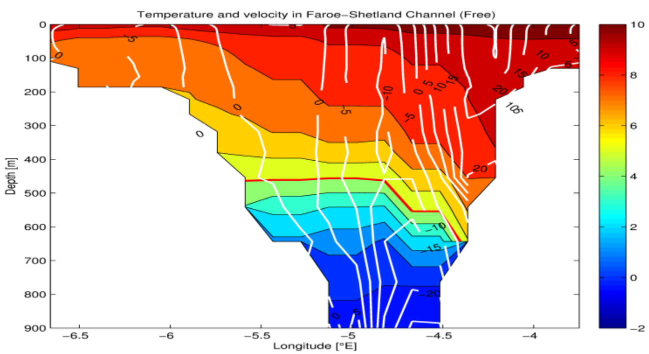


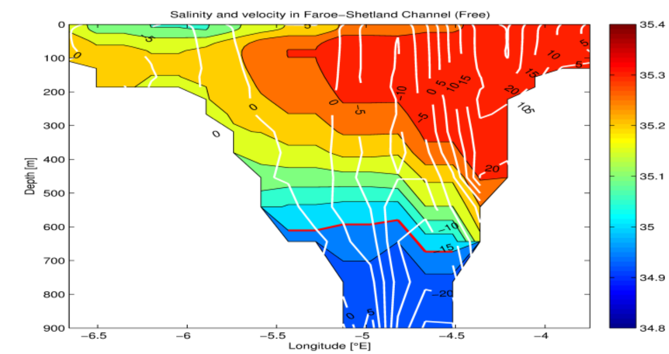
Fig. 4. Averages for the period 1995–2009 of temperature (a), salinity (b), density (σ_t) (c), and along-channel velocity superposed on the average temperature distribution (d). The velocities are based on geostrophy from CTD data corrected with altimetry-adjusted velocity data using the core interpolation. The average depth of the 5 °C isotherm along the section is indicated by the brown line on (d).

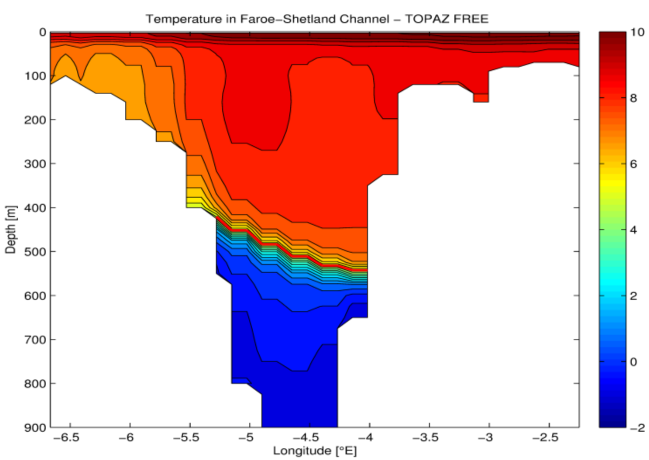




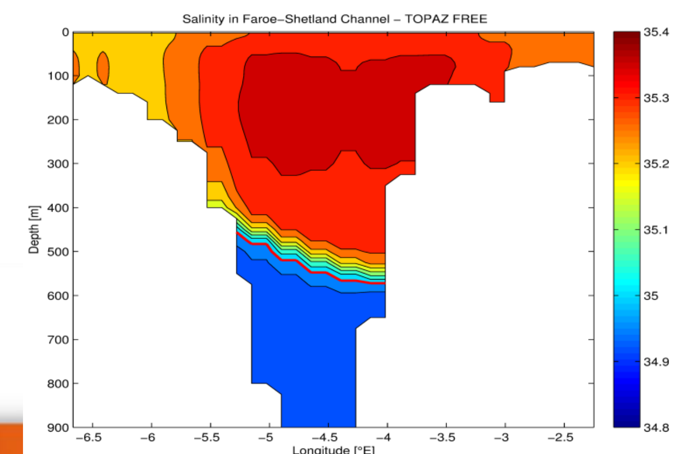


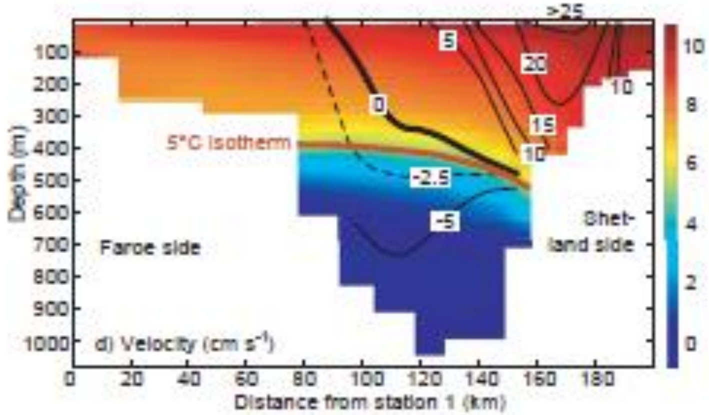
Nemo free:
Slightly higher salinity, temperature and speed than in assimilated run

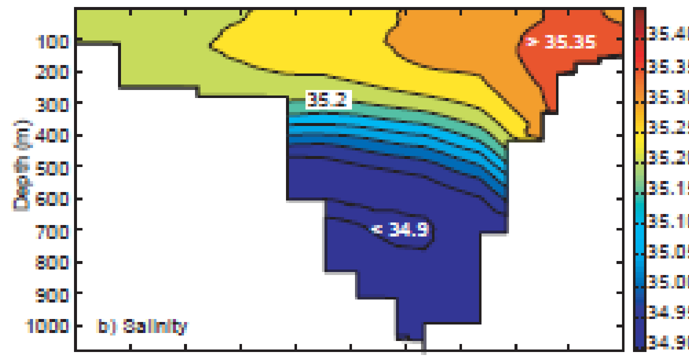


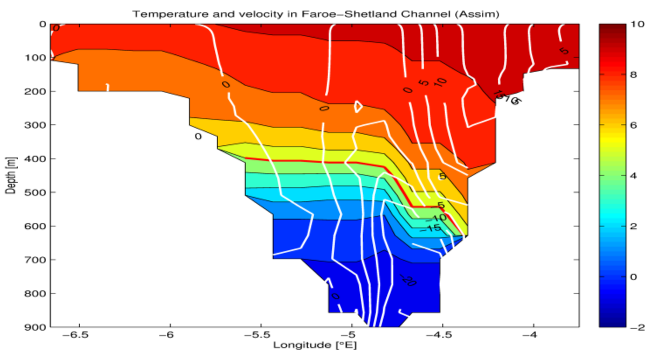


TOPAZ free:
More saline AW core than in assimilated run, but AW depth similar

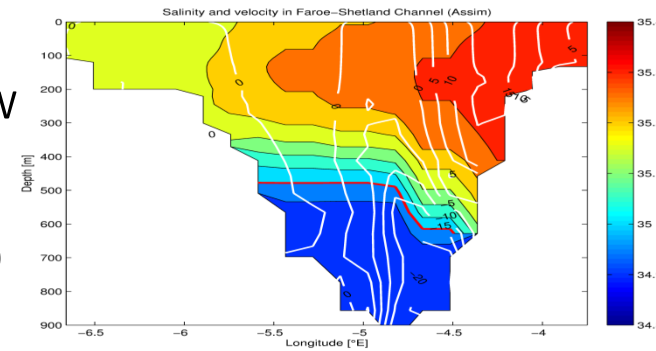


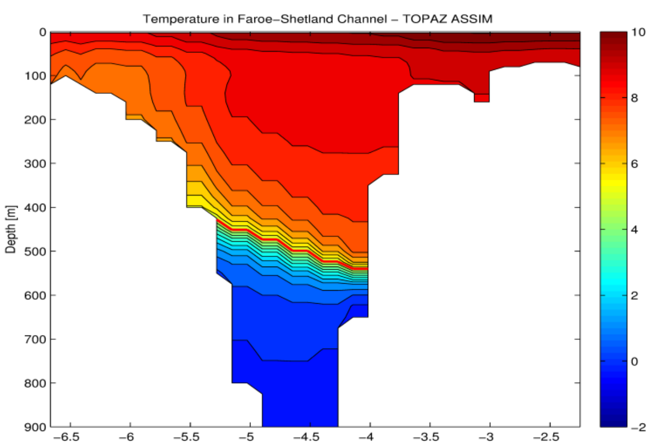




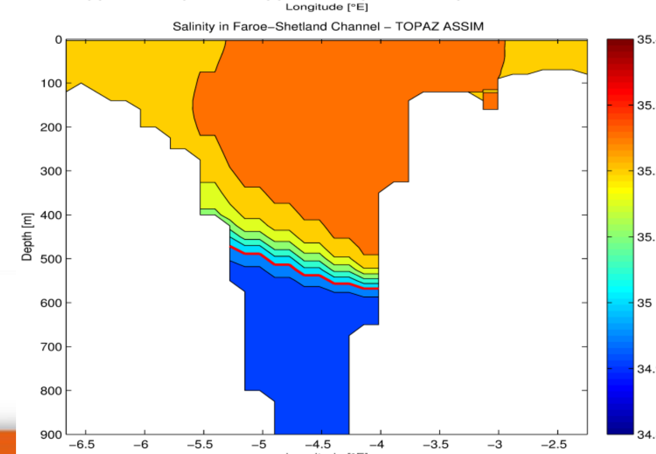


Nemo assim:
Realistic hydrography: AW core at Shetland shelf slope; sloping T and S surfaces; AW above ~500 m. Too weak currents

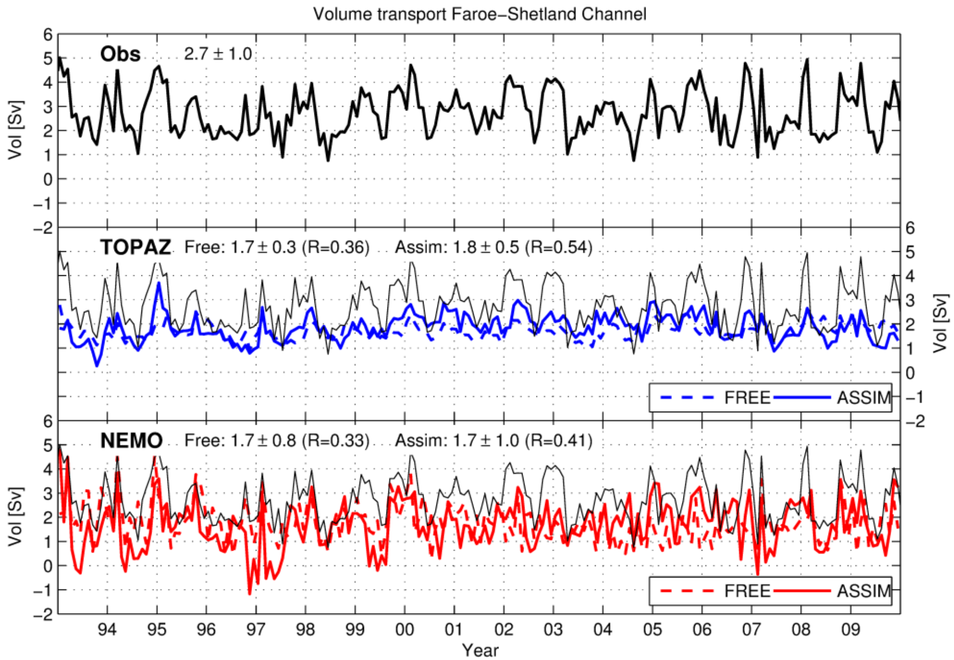


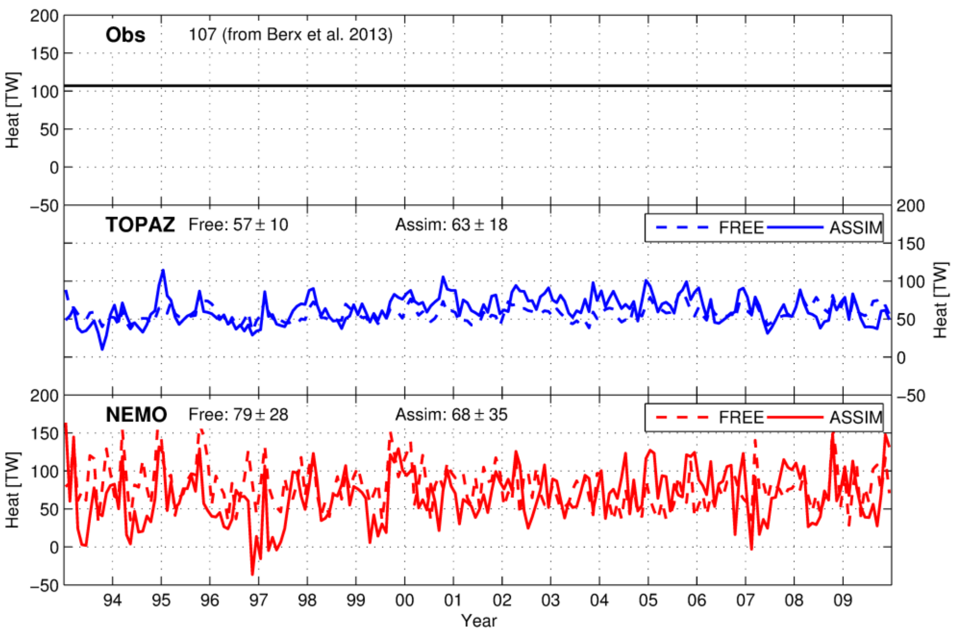


TOPAZ assim:
Realistic hydrography: AW core at Shetland shelf slope; sloping T and S surfaces; AW above ~500 m.



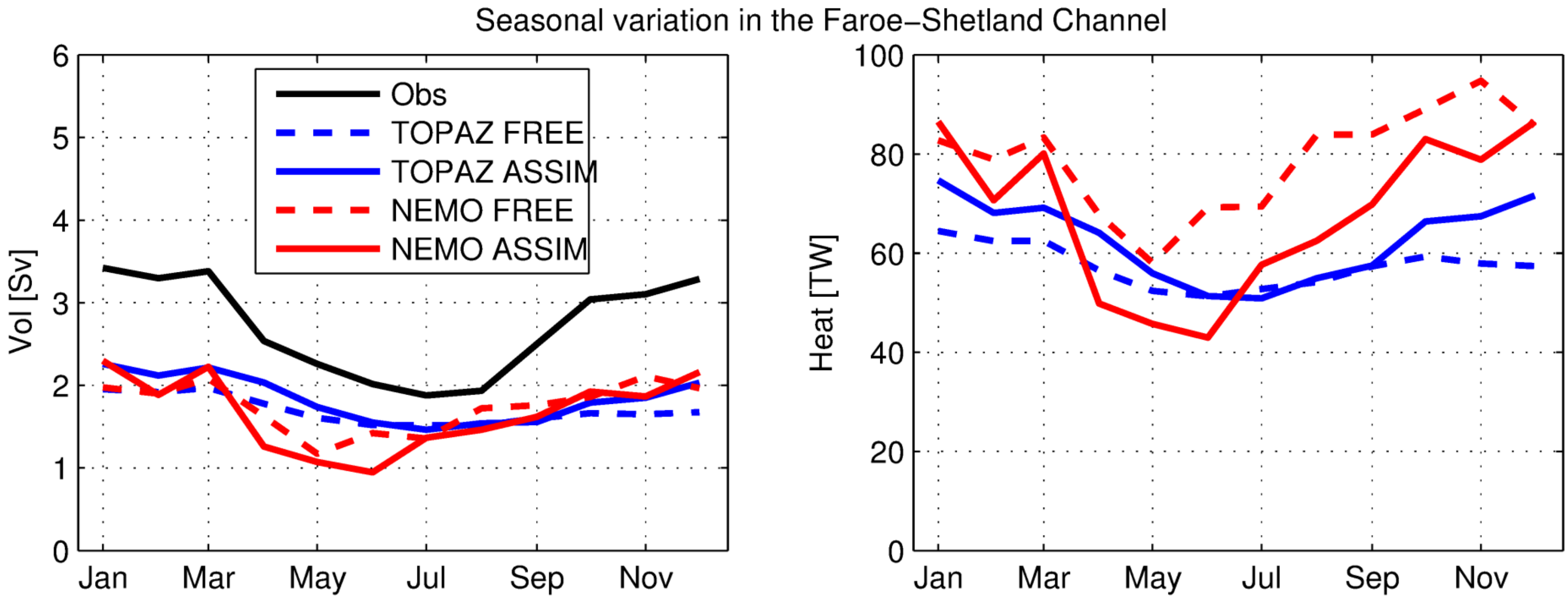
Færøy Shetland



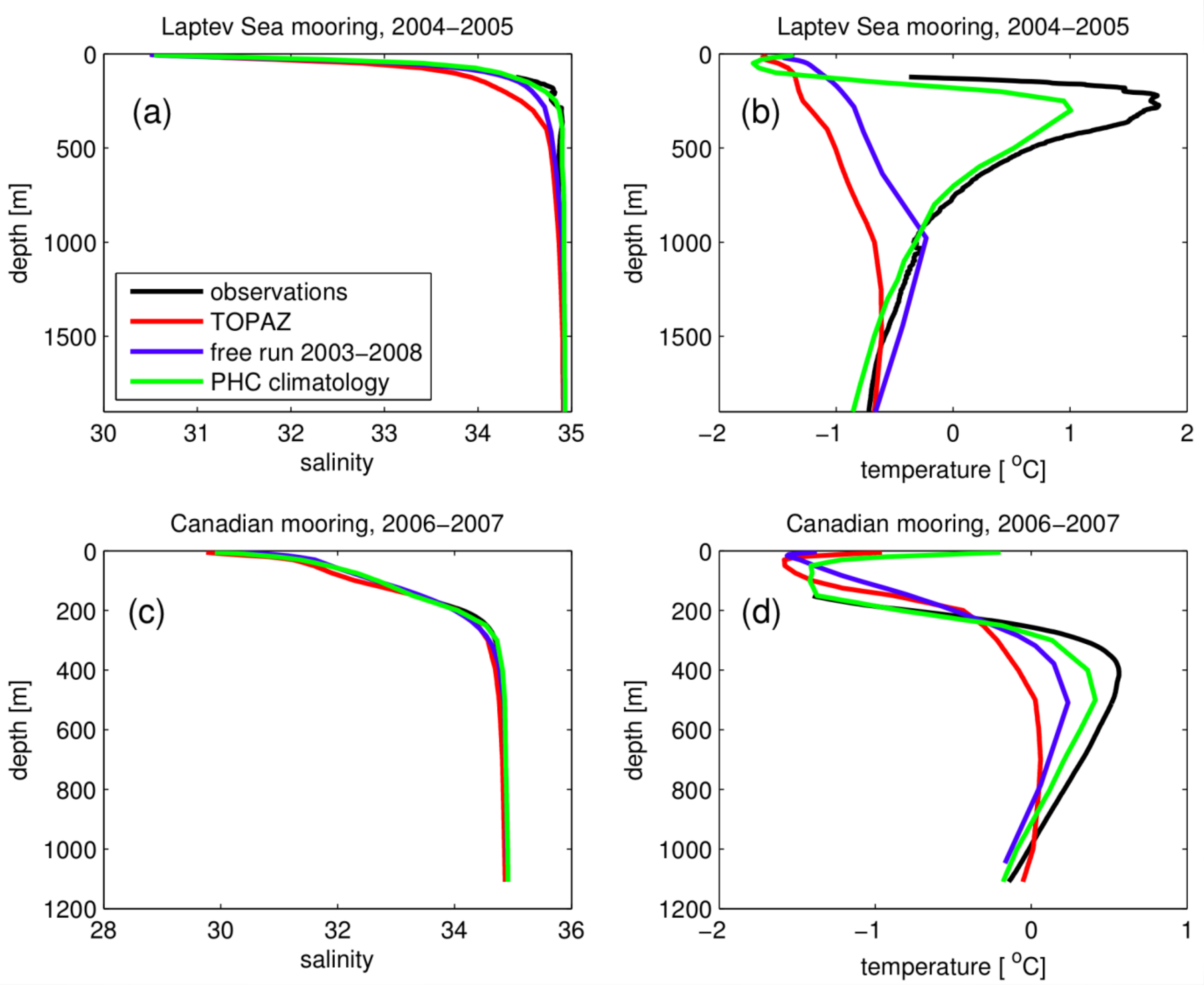


- All model simulations show too low AW volume and heat transports
- Assimilation improves correlation slightly

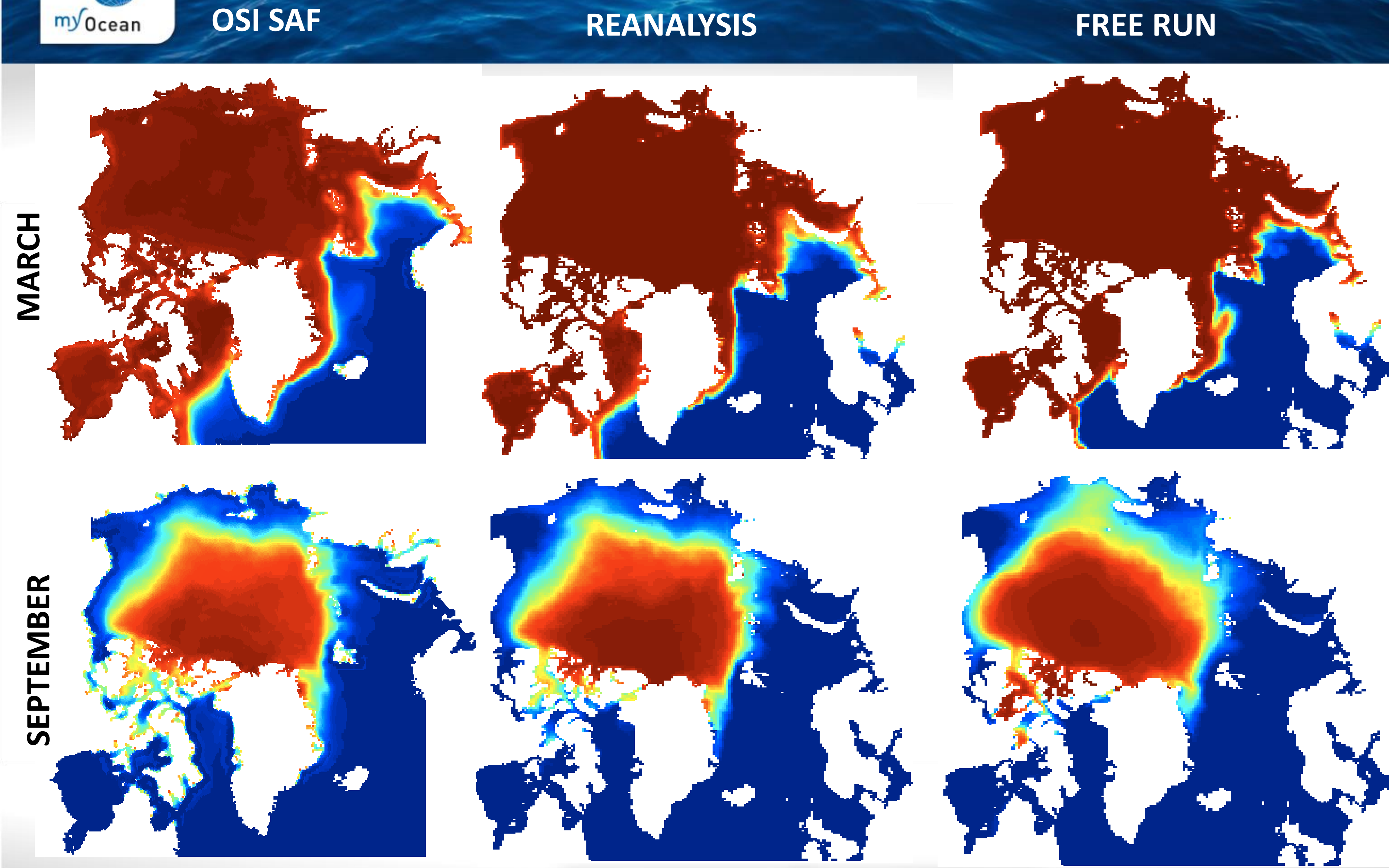
Færøy Shetland



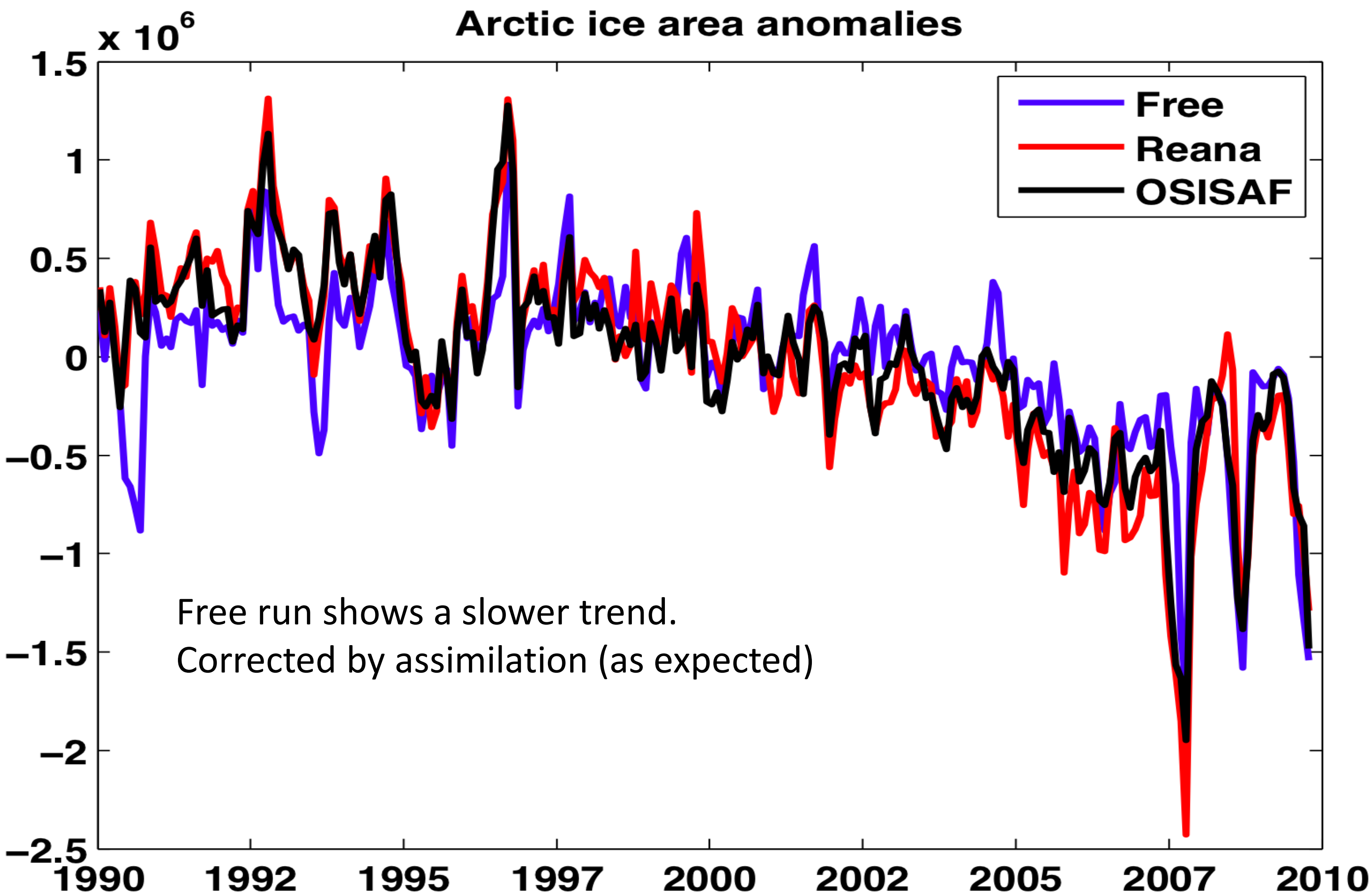
Problem of AW representation



Ice concentrations climatology 1991-2010



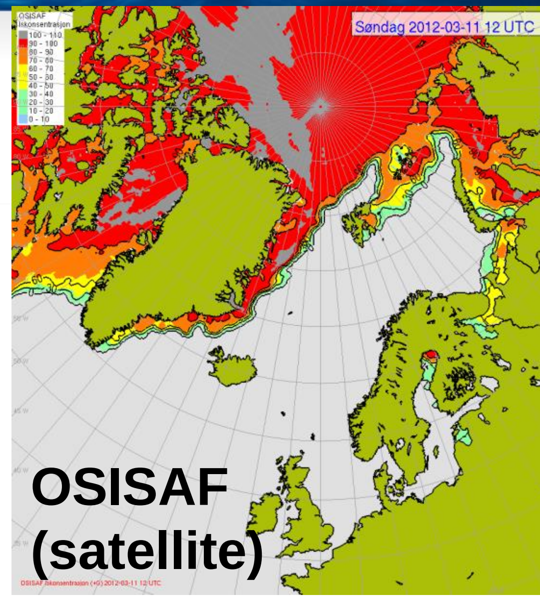
Icea area anomalies

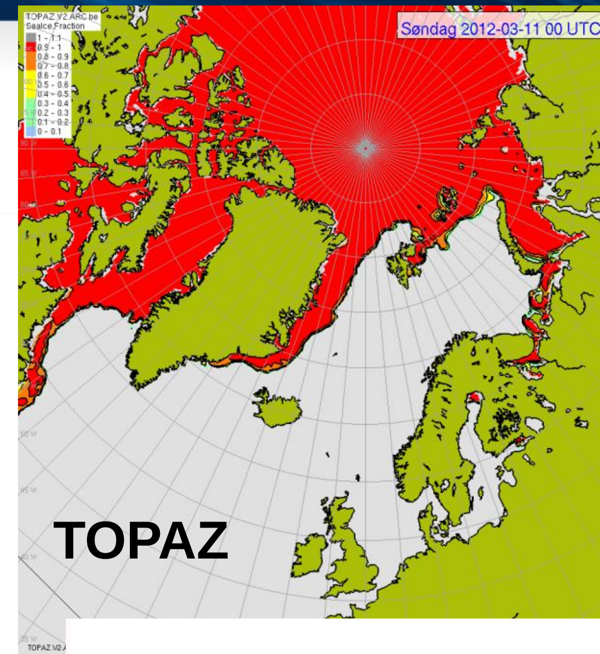


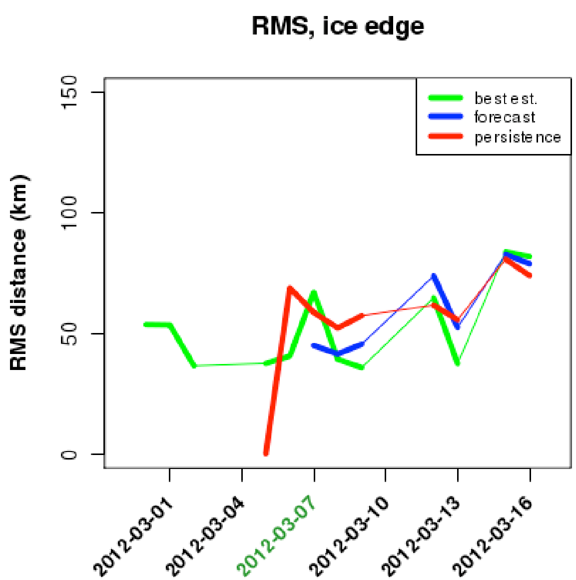
Validation of operational forecasts

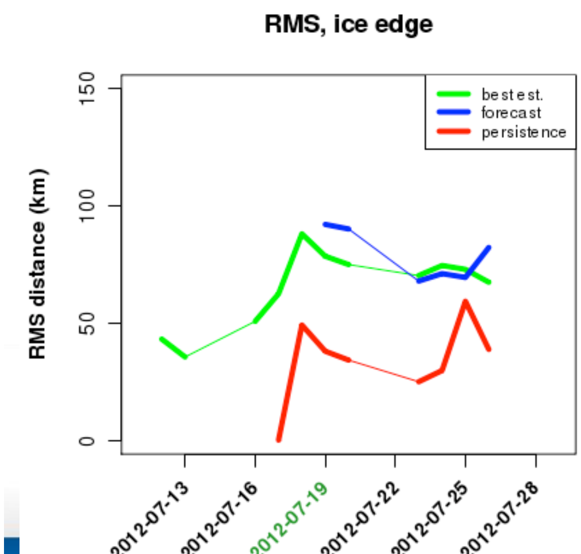
Ice edge *(A. Melsom met.no)*

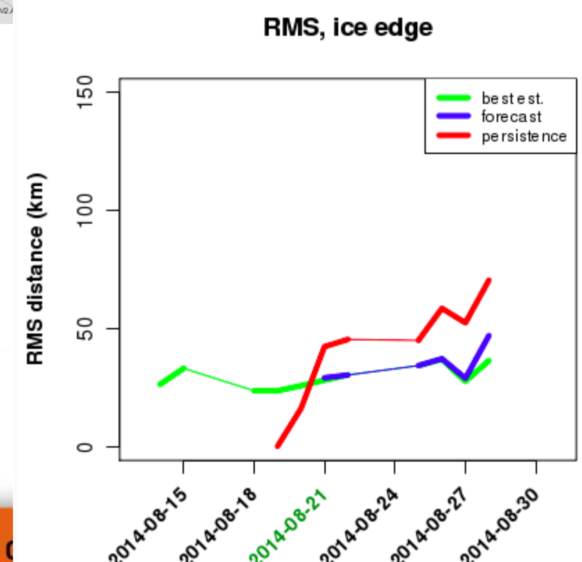
- http://myocean.met.no/ARC-MFC/V2Validation/index.html
- Weekly monitoring of forecast skills
- Error on ice edge: 50 km on average in European areas
- Larger errors in Summer
  - Expected from reanalysis









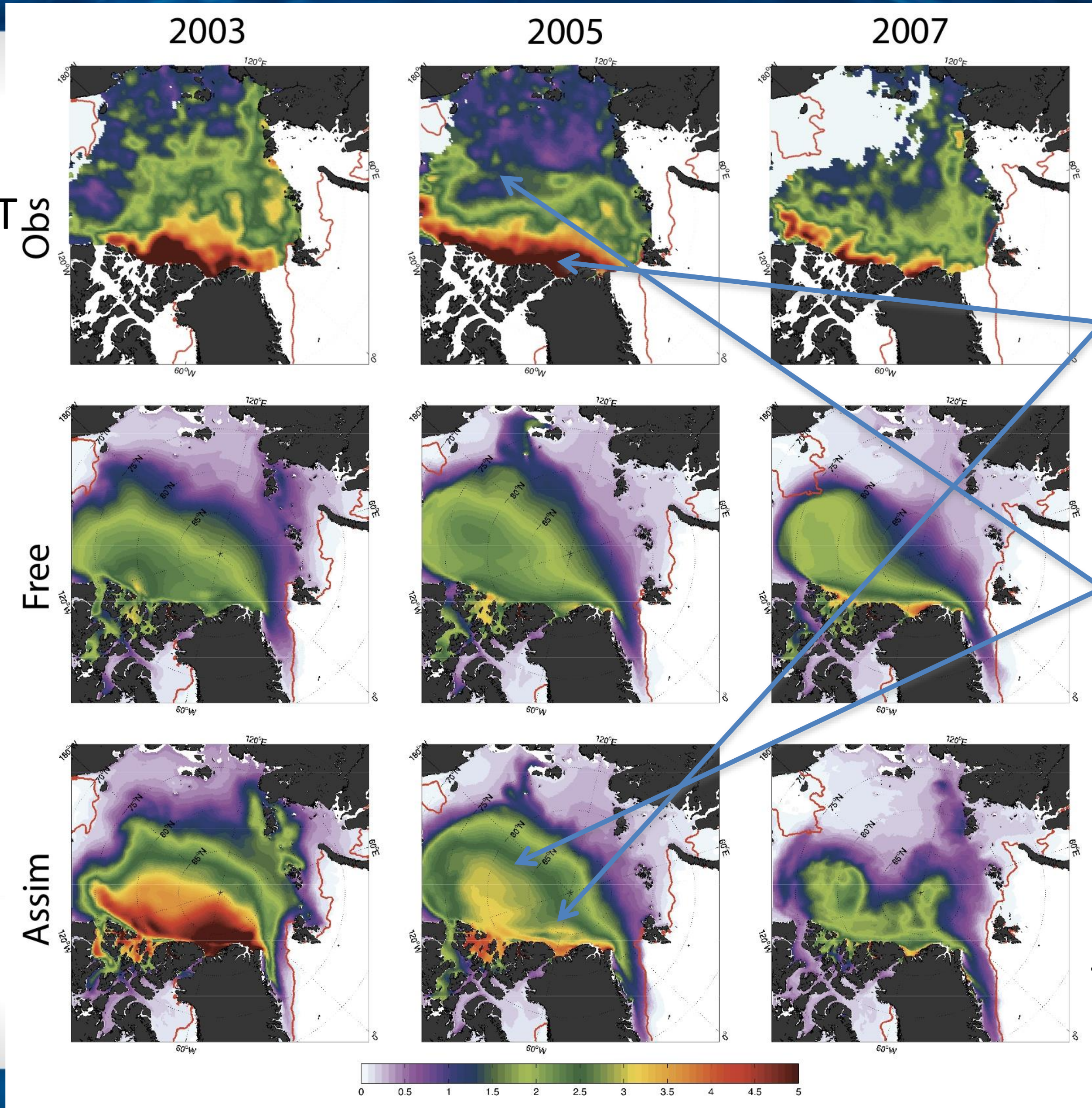


Ice thickness validation

Independent satellite IceSAT (Kwok, JPL)

TOPAZ free run

TOPAZ pilot reanalysis

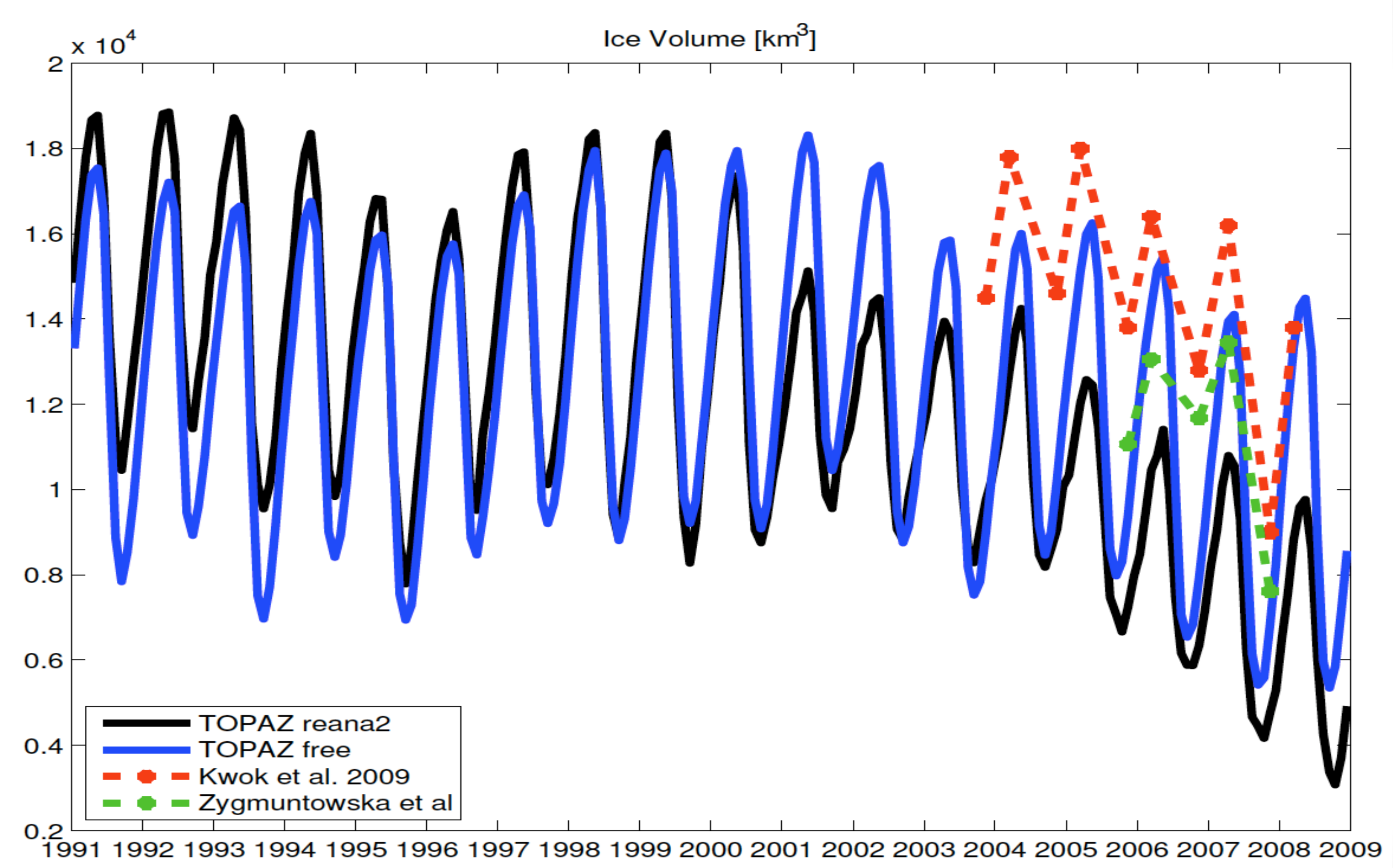


Underestimates thick ice

Overestimates thin ice

Common feature of AOMIP models (Johnson et al. JGR 2012)

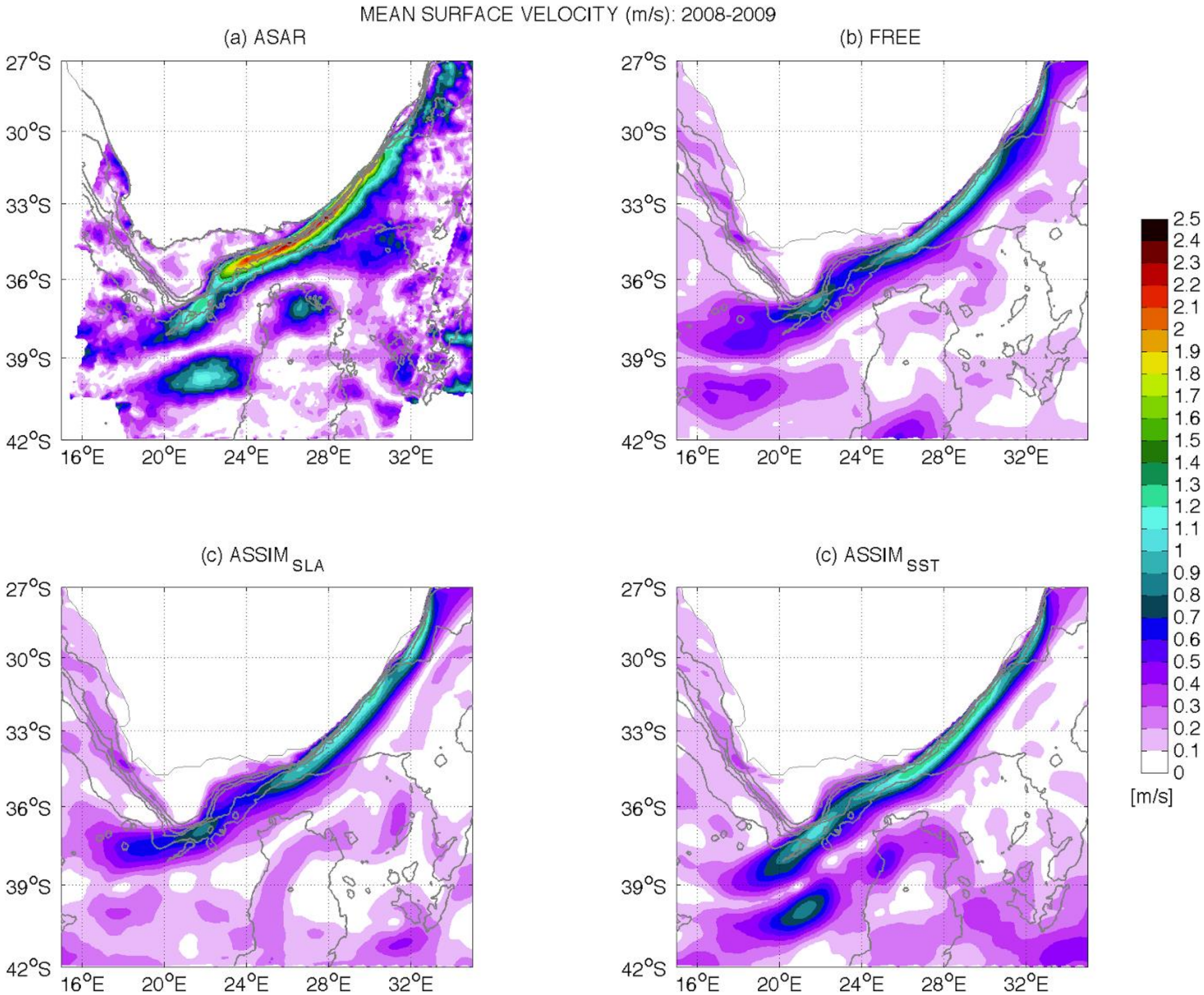
Ice thickness validation



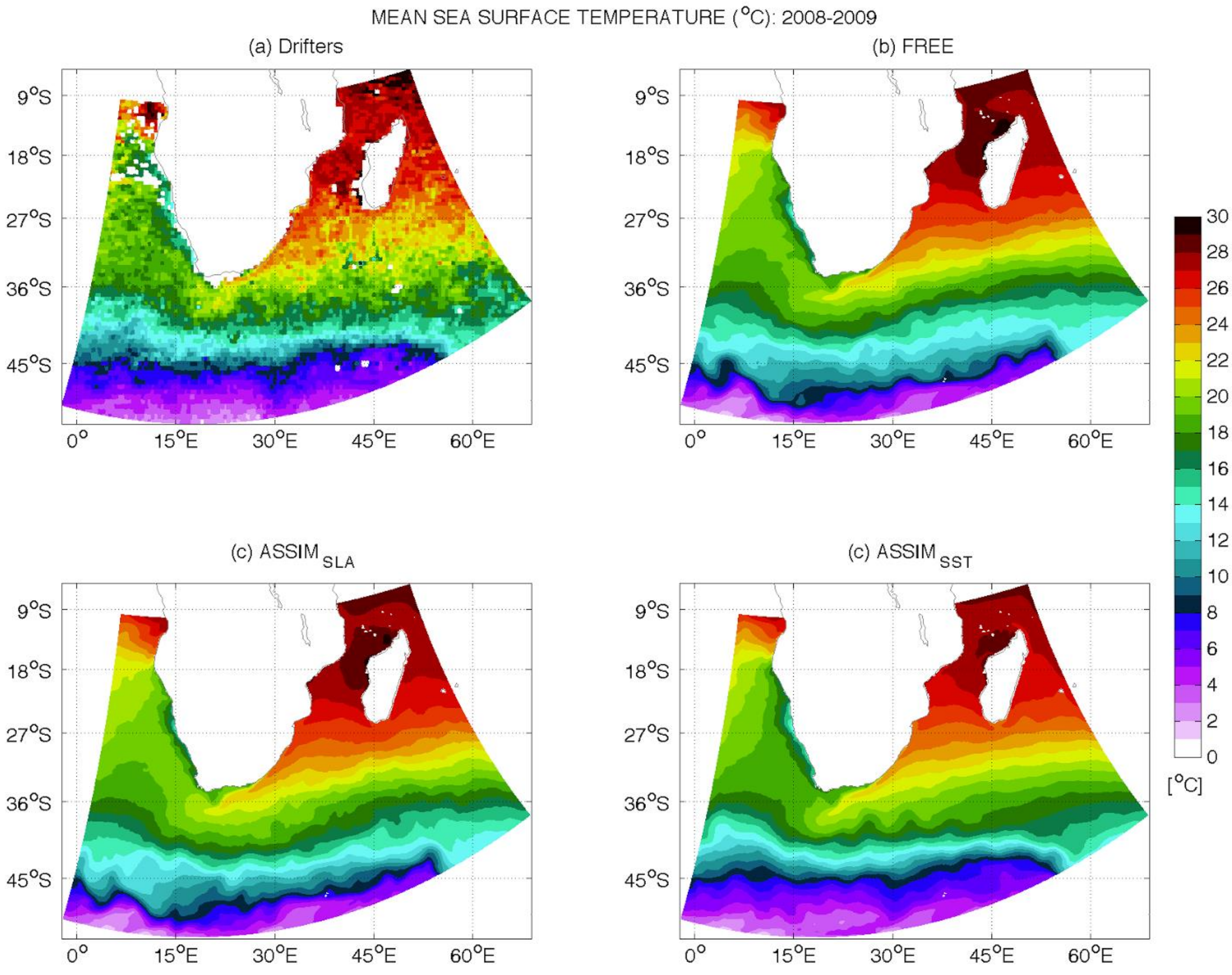
Assimilation in regional models

- We do not afford an EnKF for nested high-res models
  - Resort to a cheaper, locally-tuned EnOI
    - “Static ensemble” instead of dynamic ensemble
    - Relies on a model climatology
  - Most operational ocean data assimilation methods today are similar to an EnOI
    - Srinivasan et al. OM, 2011
  - Our experience: Gulf of Mexico, South China Sea, Agulhas currents
    - Able to constrain identifiable mesoscale features
    - Also able to handle tides while assimilating Altimeter data

Agulhas current



Agulhas current

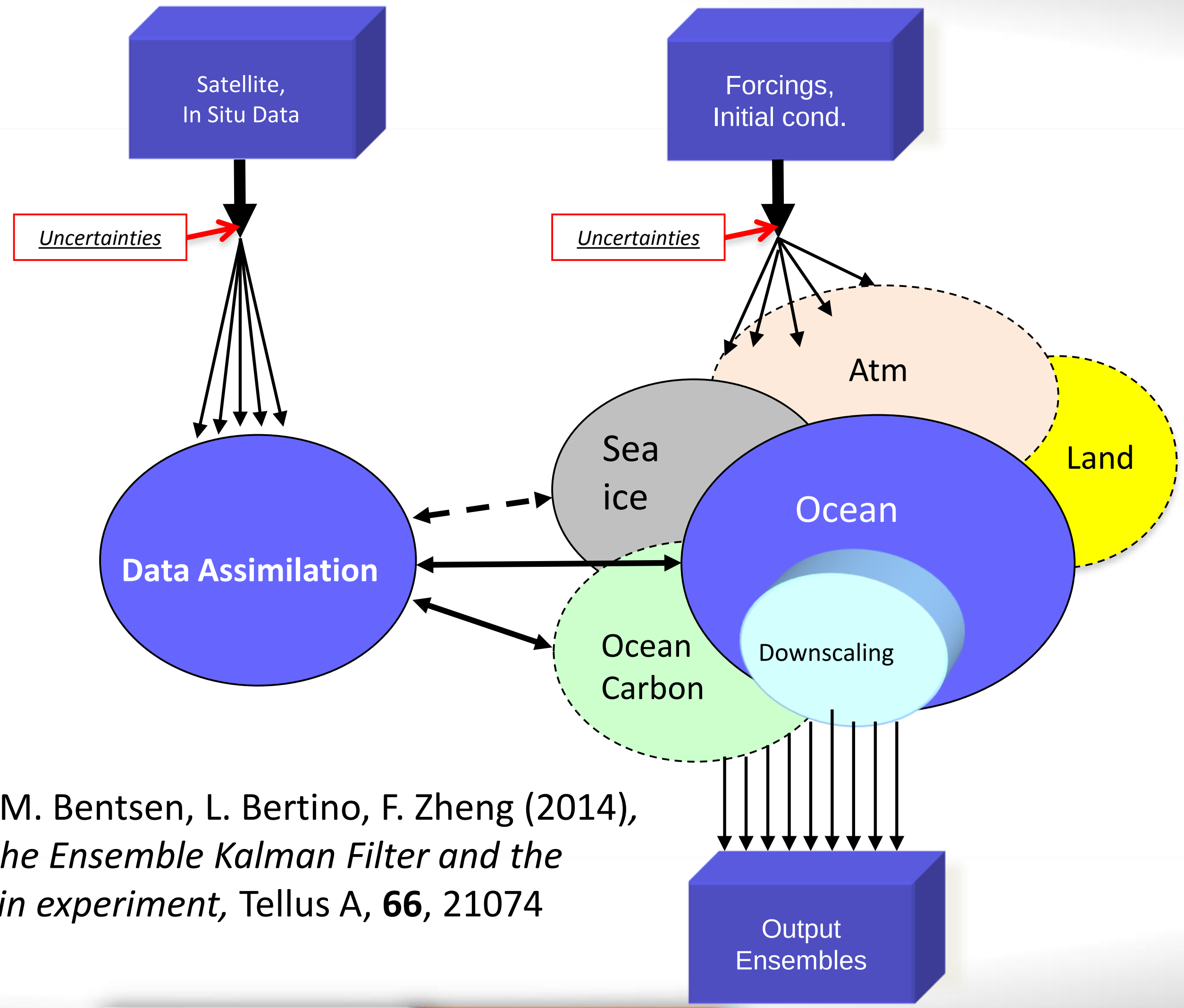


Seasonal-to-decadal prediction with the Norwegian Climate Prediction Model

Counillon F., Bethke I., Keenlyside N., Wang Y. , Bentsen M., Bertino L., Zheng F.

Norwegian Climate prediction system

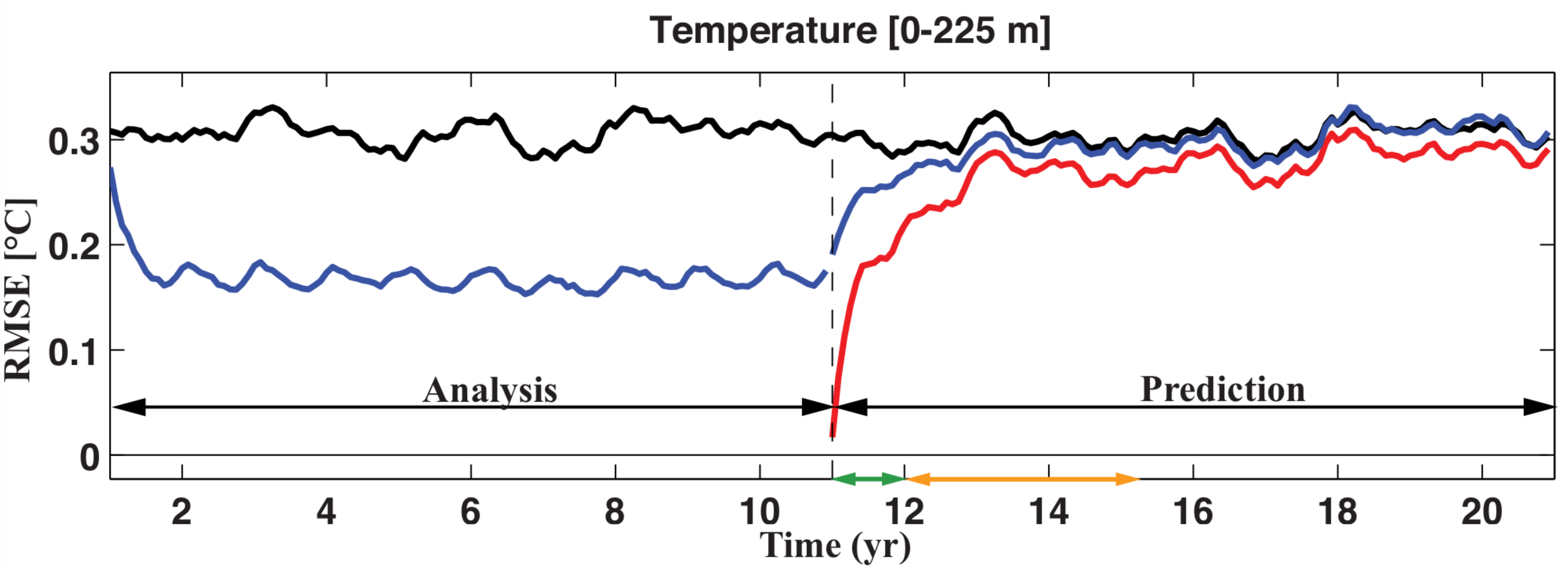
- Model: NorESM
  - Ocean: UniRe Klima
  - Carbon: UiB/UniRe
- Assimilation: EnKF
  - NERSC



F. Counillon, I. Bethke, N. Keenlyside, M. Bentsen, L. Bertino, F. Zheng (2014), *Seasonal-to-decadal prediction with the Ensemble Kalman Filter and the Norwegian Earth System Model: a twin experiment,* Tellus A, **66**, 21074

Global skill assessment: Upper ocean temperature

RMSE calculated over the full model domain (averaged over the 10 prediction cycles)



For all model variables at 1-year lead average; 2-5 lead year average

- Analyze reduction of RMSE in **EnKF-SST** relative to **Free**
- Compare the improvements relative to **Perfect**

Conclusions

- Ocean data assimilation is worth the hassle
  - EnKF framework makes probabilistic forecasts seamless.
  - Also useful in a coupled climate model (…)
- Possible to correct both (poorly) observed and non-observed variables in the ocean
  - Still some regressions but not catastrophic
  - Ice edge accuracy within 50km, SST less than 1 deg C
- Other features are difficult to reproduce
  - Acceleration of ice drift
    - Model drift is too fast
    - Even the drift seasonality is not respected
  - Thinning of the sea ice
- R'n D to do
  - Ocean models Arctic water mass properties: better numerics or resolution
  - Sea ice validation argues for a change of the model EVP rheology