EnKF

Look Ma, no ~~Hands~~!

Ensemble data Assimilation without Ensembles

Christian L. Keppenne[1,2], Guillaume Vernieres[1,2], Robin M. Kovach[1,2], Michele.M. Rienecker[1]

[1]NASA Global Modeling and Assimilation Office
[2]Science Systems and Applications, Inc. (SSAI)

4th World Climate Research Program International Conference on Reanalyses
Silver Spring, MD, May 07-11, 2012

We don't hate the EnKF!

Of course, it helps to have more than just one shot.

EnKF use at GMAO

- **Ocean data assimilation since 2001 (Keppenne et al.)**
- **Land assimilation and initialization since 2002 (Reichle et al.)**
- **Ensemble-based 4DVAR planned for next generation ADAS (Todling et al.)**

Objective

Maximize assimilation performance when only 1 model trajectory is available to estimate bckground covariances

3 EnKF alternatives

- Steady state ensemble (a.k.a. enOI, asymptotic EnKF, SE*K)
- Ensemble in time (LIFE)
- Ensemble in space (SAFE)

Other acronyms throughout:
RMSF: rms observation minus forecast errors
RMSA: rms observation minus analysis errors

Experimental context

- Model:
  - GFDL MOM4p1 (OGCM component of GEOS-5 modeling system)
  - LANL CICE IGCM
- Resolution: 720x410x40
- Tripolar grid – cartesian below 60°N (same as NCEP GODAS)

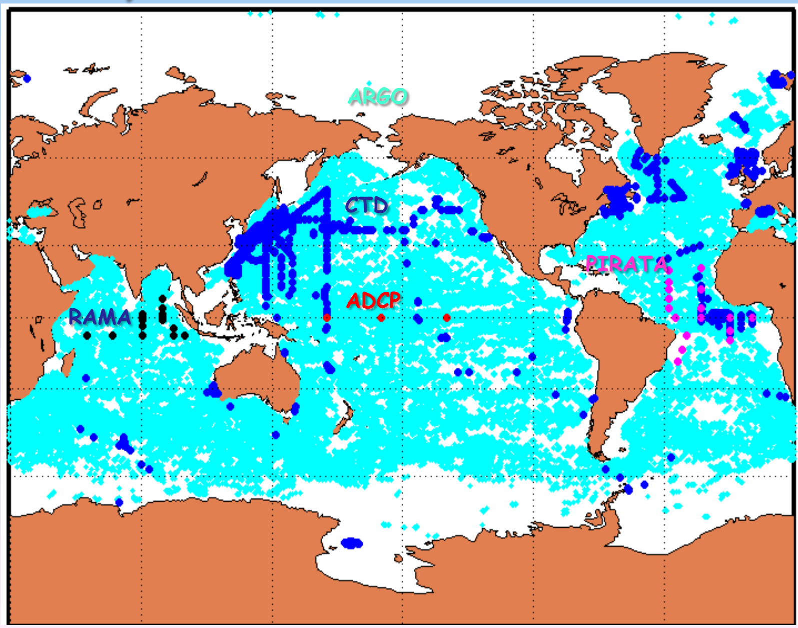
- Forcing:

  Surface fields from 2009-2011 coupled GEOS-5 run replaying MERRA

- Time span covered: 2010-2011

- Data assimilated:
  - Temperature profiles from ARGO, TAO, XBT, CTD & RAMA
  - Ice fractional coverage from NSIDC
  - Reynolds SSTs

- Data used for validation:
  - ARGO S, CTD S, RAMA S, Pirata S
  - ADCP u & v
  - PIOMAS ice thickness



Steady state ensemble: asymptotic ensemble filter (EnOI)

Ensemble of leading EOFs of forecast anomalies

Baseline ensemble	• Ensemble of 20 coupled GEOS5 runs without data assimilation
Signal EOFs	• "Signal EOFS" of baseline ensemble variability
1st stage analysis	• Ensemble OI using signal EOFs
Ensemble forecasts	• Ensemble forecasts initialized from perturbations of 1st stage analysis
Forecast-error EOFs	• EOFs of large ensemble of drift-removed 30-60 day forecasts
2nd stage analysis	• Ensemble OI with 20-leading forecast-error EOFs

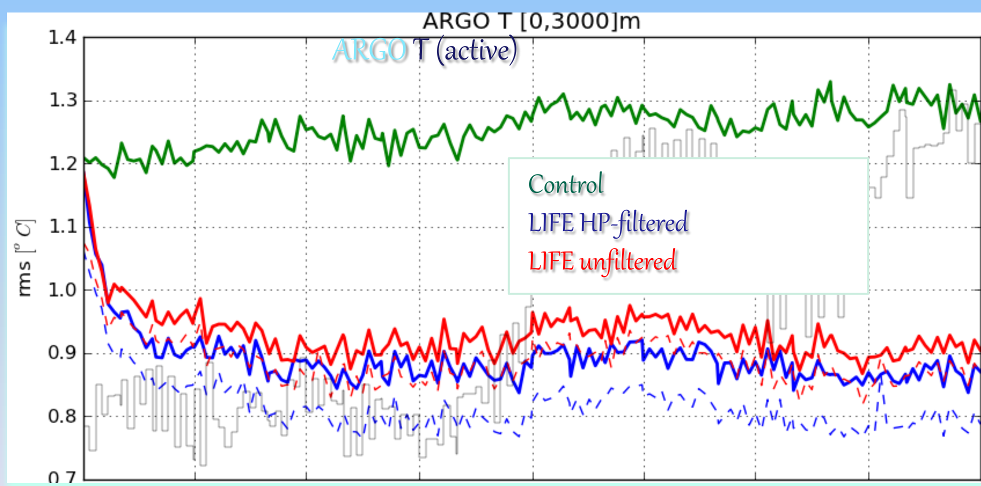
Ensemble in time:

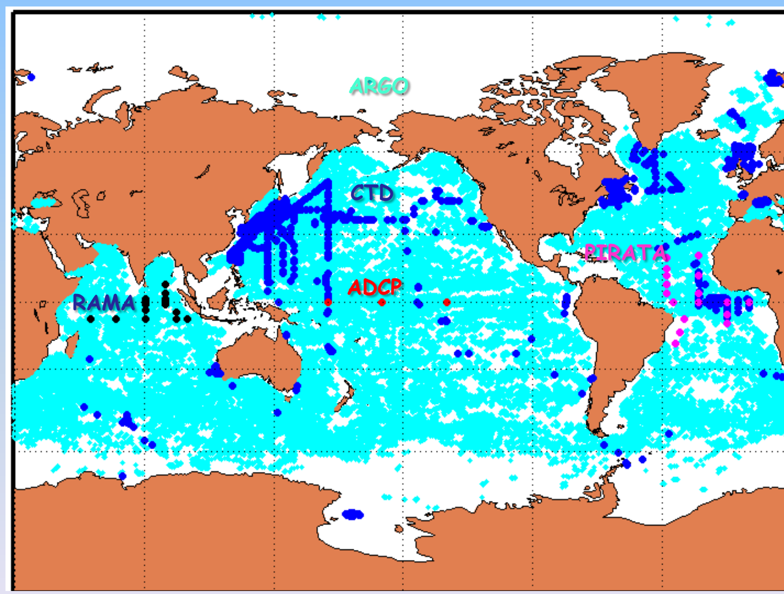
Lagged instantiation of forecast errors (LIFE)

Analysis increment Δx_k at time t_k computed from ensemble $(x'_k, x'_{k-1}, .., x'_{k-n})$ of high-pass filtered lagged background states of current model run

Window centering effect:

W/o high-pass filtering,
ensemble is dominated by lag = 0 and lag = n instances





Ensemble in space:

Space adaptive forecast errors (SAFE)

Covariances computed as if an ensemble were sampled near every gridpoint

Univariate update of observed variable

- Local variance estimated from spatial distribution of variables in high-pass filtered background

projection onto un-observed variables

- Local cross-field covariances computed from local spatial distribution of variables in high-pass filtered background serve to update un-observed variables

Assimilation procedure

. 5-day assimilation interval
- Incremental update
- SSTs processed first, T profiles second, ice last

- Flow adaptive error-covariance localization
  . SST and T profiles assimilation: x, y, z, t, T localization
  . Ice assimilation: x, y, z localization

- Multi-scale assimilation used with T profiles and ice

(unexplained part of innovation processed a 2nd time with shortened covariance localization scales)

- Logistic transformation of ice field to enhance covariances
- Adaptive error covariance inflation so as to maintain:
  $\mathrm{trace}(HPH^{t}R^{-1})/n_{obs}$ = pre-specified target value
  (1 when assimilating SSTs and T profiles, 1 for NH ice, 0.1 for SH ice)

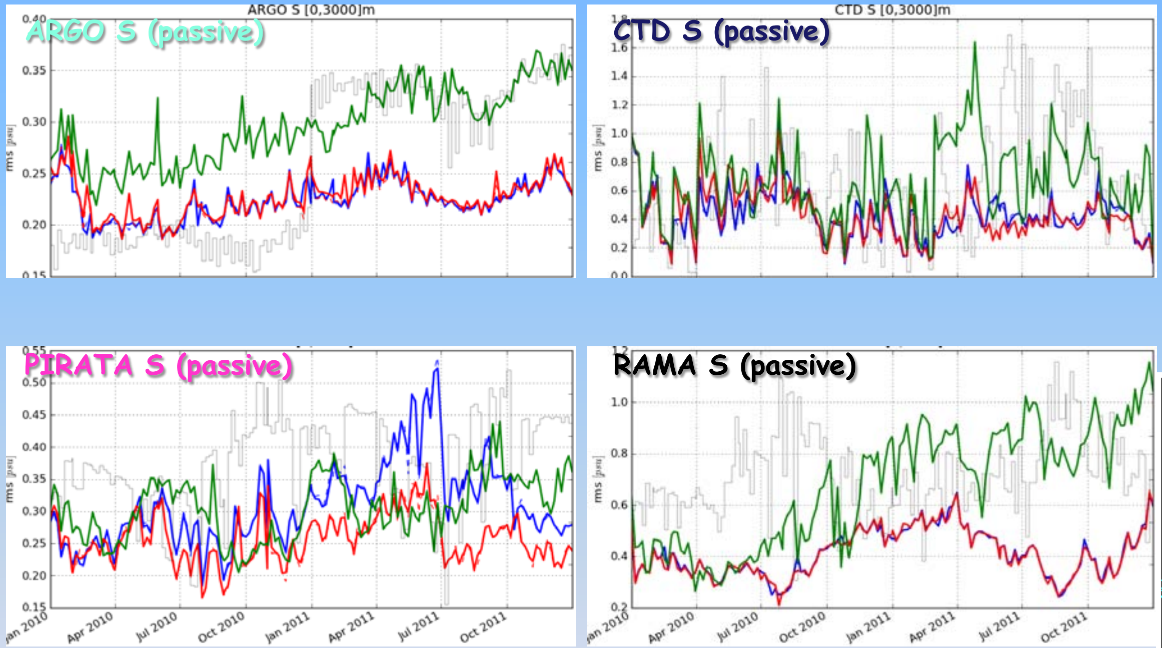
- Comparison benchmark:
  univariate OI enhanced with flow adaptive covariance localization (EUOI)

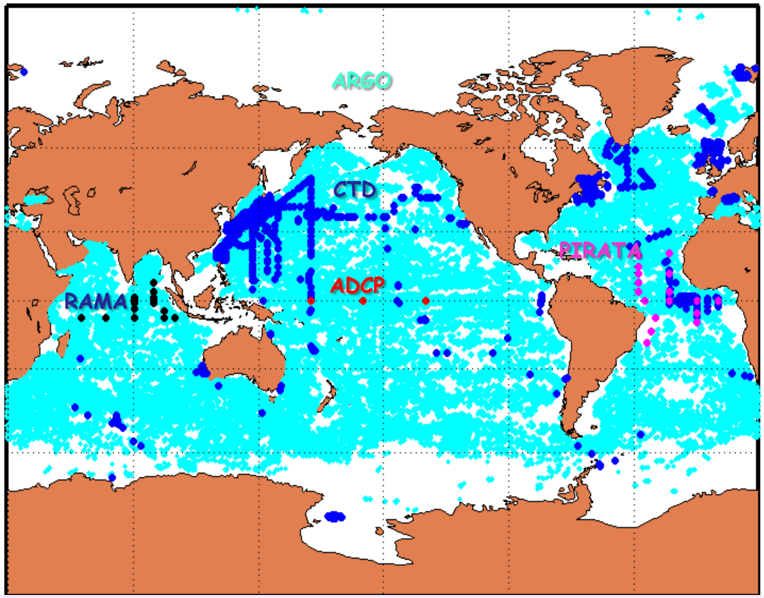
Multi-scale procedure

Covariance localization scales (x,y,z,t,T or ρ) repeatedly reduced

Each time

- background is updated
- error covariances are rescaled: $P_k = rms(oma_{k-1})/rms(omf_{k-1})P_{k-1}$
- unexplained portion of innovation is reprocessed





Control

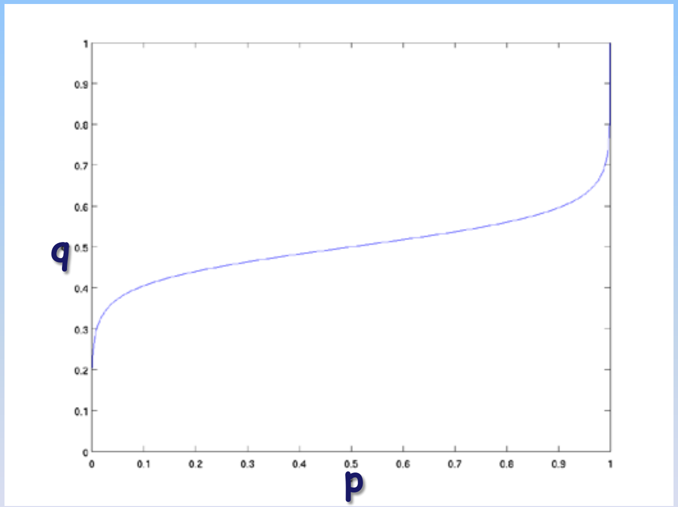
SAFE mono-scale

SAFE multi-scale 1st iter. (k=1)

Logistic ice transformation

Ice fraction p in [0, 1] replaced with $q = \log(p'/(1 - p'))$
where $p' = \min(1 - \varepsilon, \max(\varepsilon, p))$
q is rescaled to [0, 1]

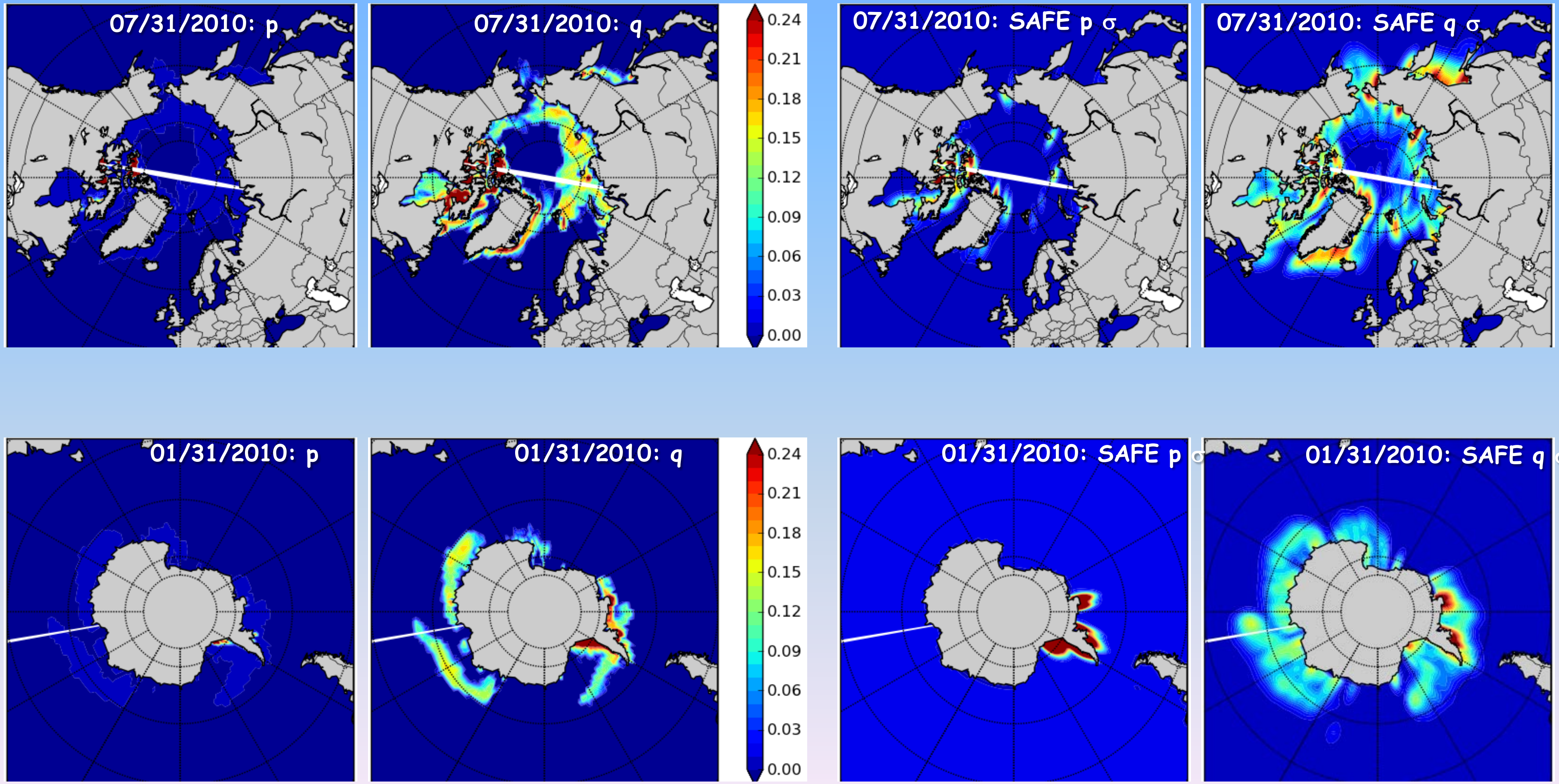
Purpose:
- **enhance spatial variance in SAFE**
- **enhance ensemble variance in LIFE/EnOI**



Logistic ice transformation vital in SAFE, LIFE

NB: In current un-coupled configuration,
forcing melts ice in Summer unless ice is assimilated!

P and Q and corresponding s from baseline run with no assimilation



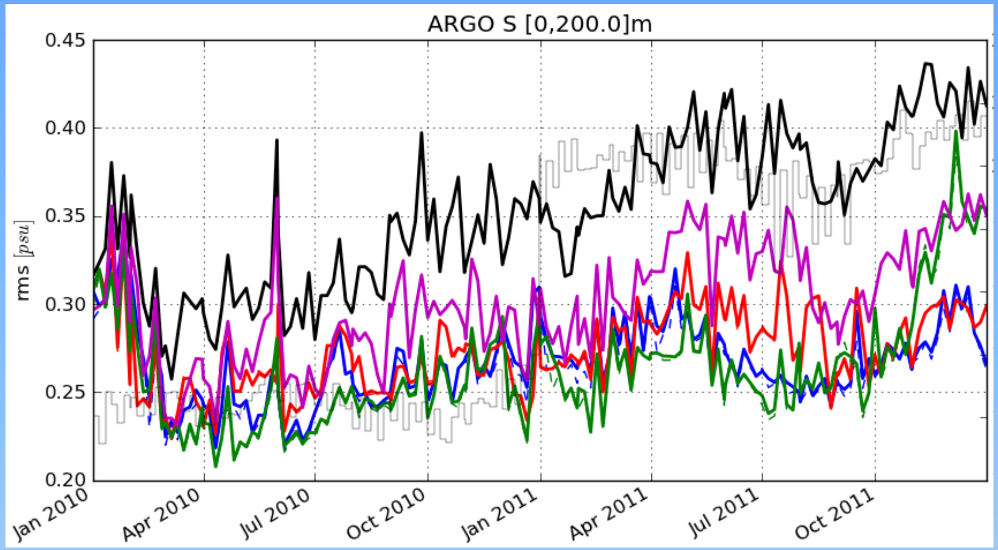
T-profile assimilation configuration summary

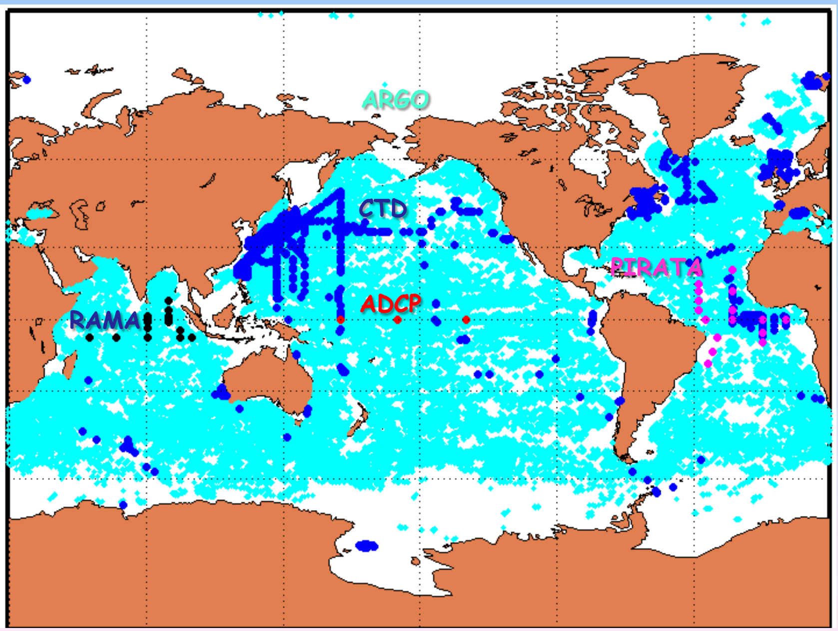
	EnOI	Ensemble in space (SAFE)	Ensemble in time (LIFE)	Enhanced Univariate OI (EUOI)
updated	T, S, u, v	T, S, u, v	T, S, u, v	T
multi-scale	Bi-scale	Bi-scale	Bi-scale	Mono-scale

T-Profiles assimilation results

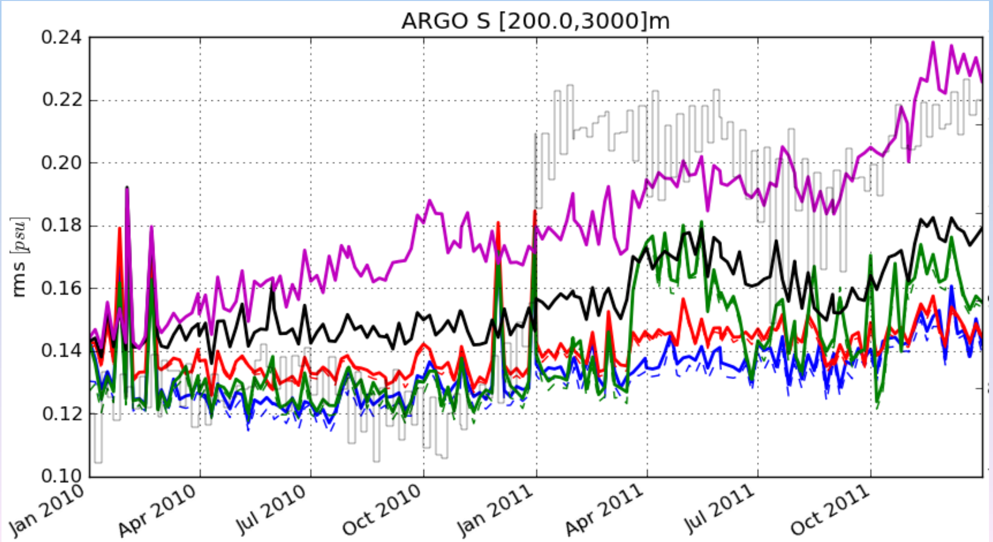
ARGO RMS passive S OMF/OMA
Control
EUOI
SAFE
LIFE
EnOI

Above 200m:
SAFE ≈ EnOI ≈ LIFE > EUOI > control





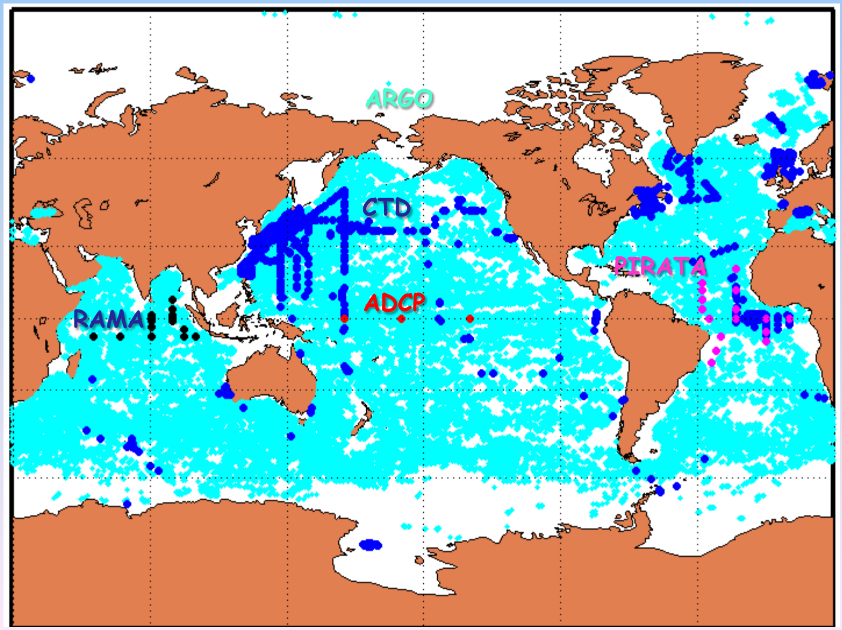
Below 200m:
SAFE > LIFE > EnOI > control > EUOI



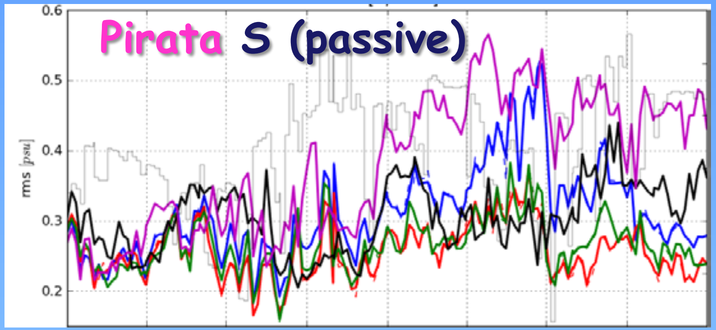
T-Profiles assimilation results

Pirata, RAMA, CTD RMS passive S
OMF/OM
Control
EUOI
SAFE
LIFE
EnOI

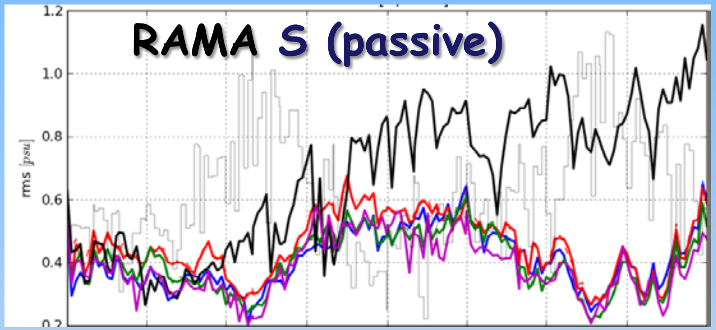
Multivariate update in SAFE, LIFE & EnOI preserves S when it is not assimilated



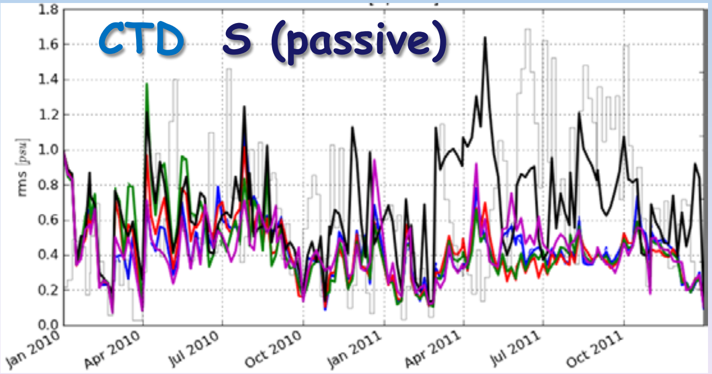
SAFE ≈ EnOI ≈ LIFE > control > EUOI



SAFE ≈ EnOI ≈ LIFE ≈ EUOI > control



SAFE ≈ EnOI ≈ LIFE ≈ EUOI > control

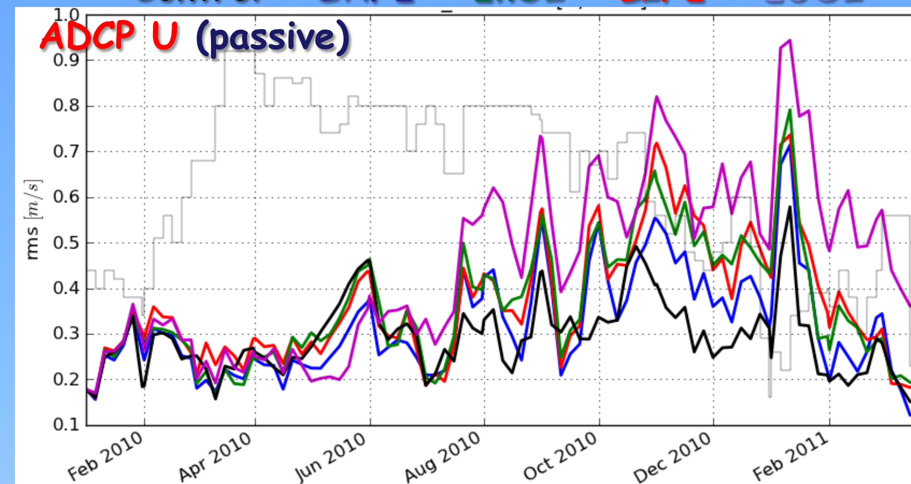


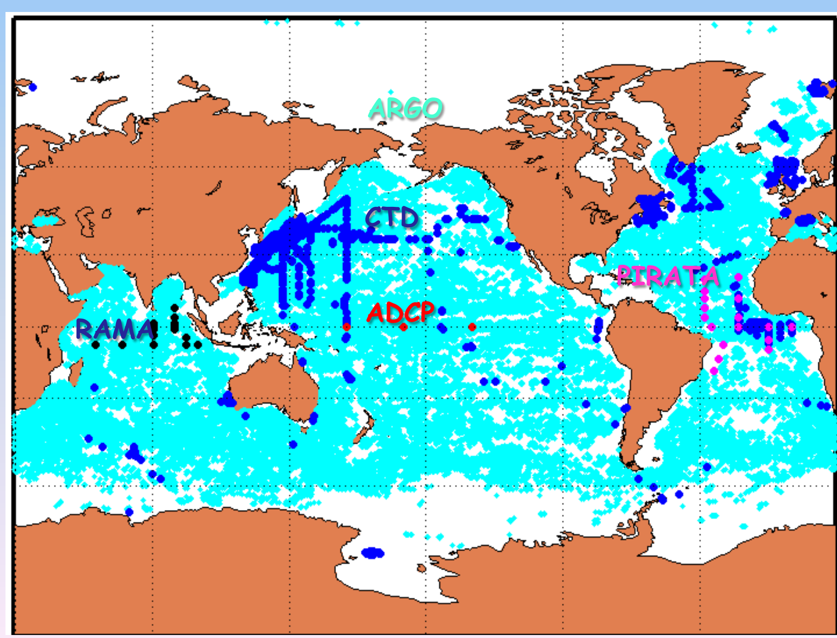
ADCP passive U, V RMS OMF/OMA

Control
EUOI
SAFE
LIFE
EnOI

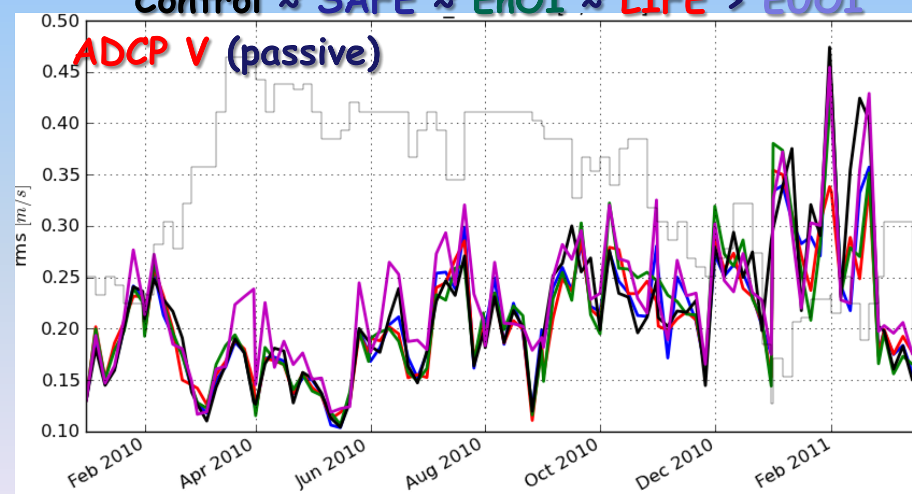
Not much to write home about in terms of current but SAFE, LIFE & EnOI all do better than univariate T + ice assimilation

Control > SAFE ≈ EnOI ≈ LIFE > EUOI





Control ≈ SAFE ≈ EnOI ≈ LIFE > EUOI



Multivariate ice update procedure (SAFE, LIFE, EnOI)

- Stage 1:

Compute Δice assimilation increment + use of <ice,T> and <ice,S> covariances to compute ΔT and ΔS increments

- Stage 2:

Calculate (T_{fr},S_{fr}) that minimize $[(T- T_{fr})/\sigma_T]^2+ [(S- S_{fr})/\sigma_S]^2$

- Stage 3:

Validation of stage 1 correction to ensure:
If Δice < > 0, then T analysis >= <= $(1\text{-ice})\, T_{ocean}$ + ice T_{fr}

NSIDC Ice fraction assimilation configuration summary

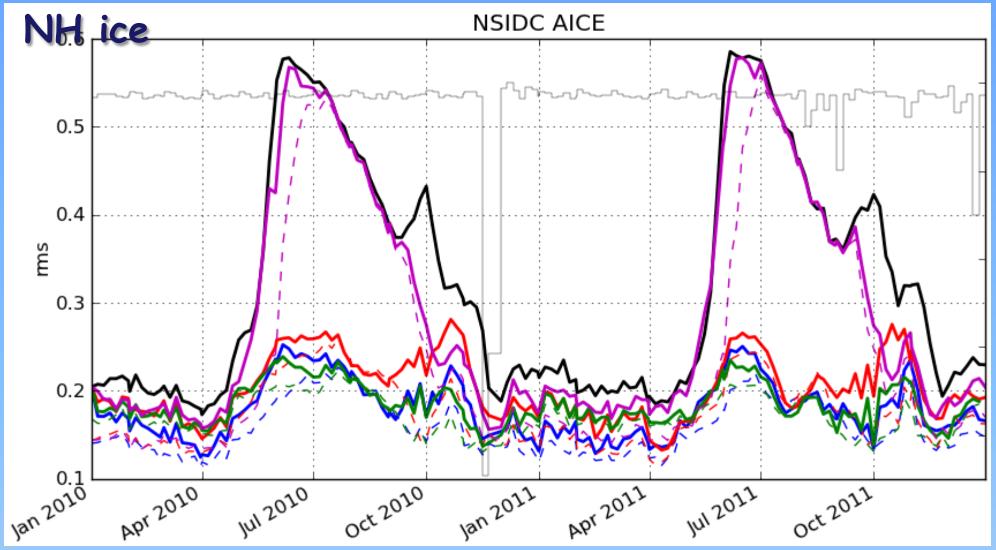
	EnOI	Ensemble in space (SAFE)	Ensemble in time (LIFE)	Enhanced Univariate OI (EUOI)
updated	Ice, T, S	Ice, T, S	Ice, T, S	Ice
multi-scale	Bi-scale	Bi-scale	Bi-scale	Mono-scale

Ice fraction assimilation results

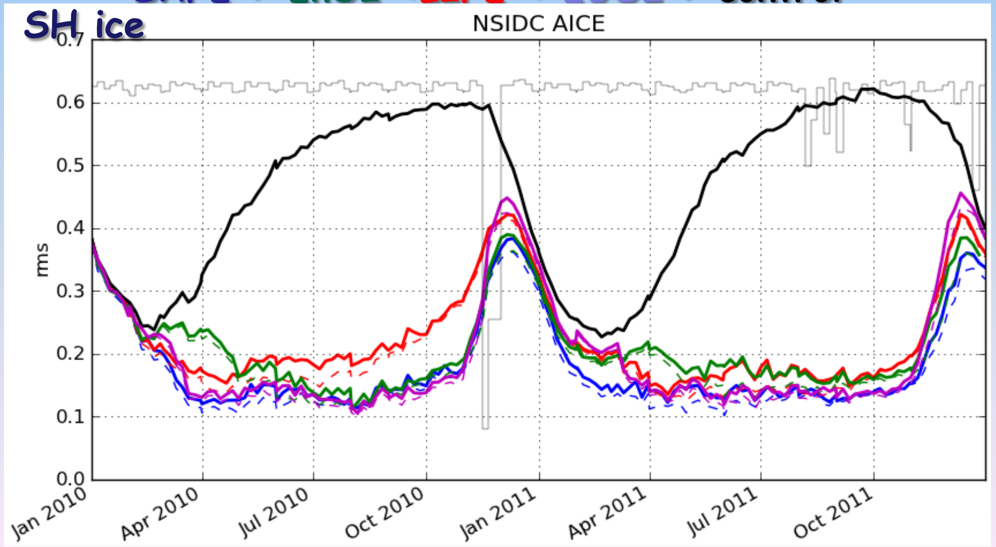
NSIDC RMS ice OMF/OMA
Control
EUOI
SAFE
LIFE
EnOI

Can't do without update of T, S when assimilating ice in NH!

Northern hemisphere ice:
SAFE > EnOI > LIFE > EUOI > control



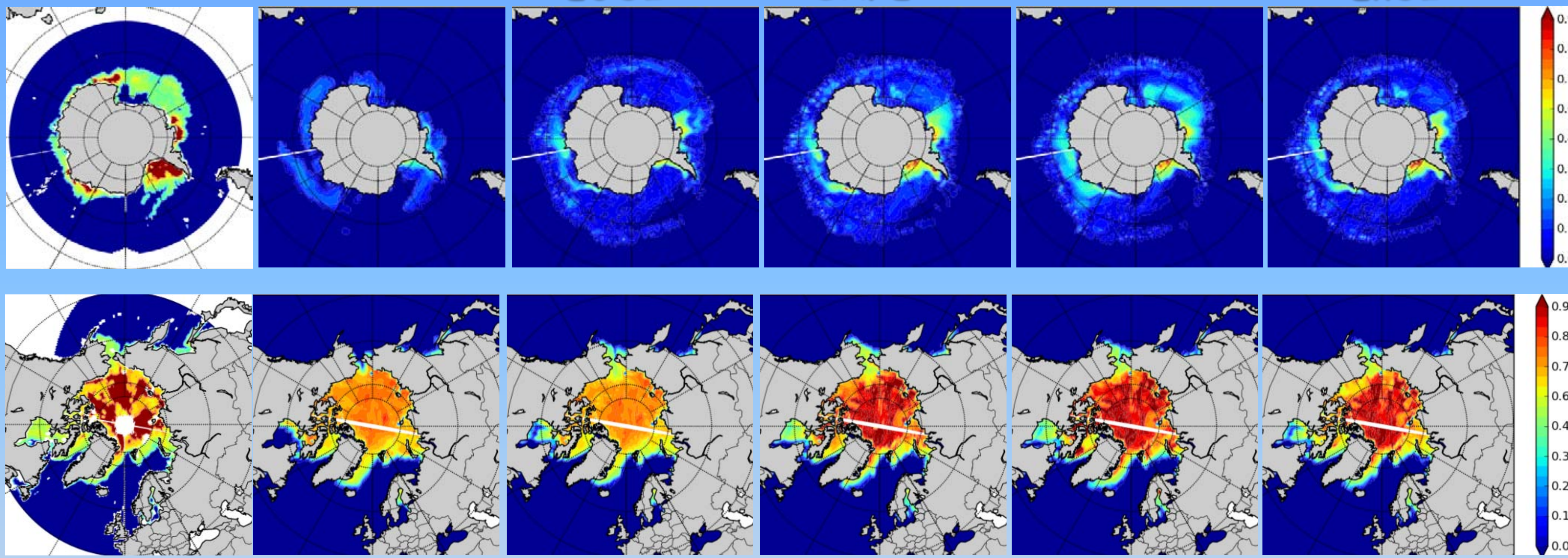
Southern hemisphere ice:
SAFE > EnOI ≈LIFE ≈ EUOI > control



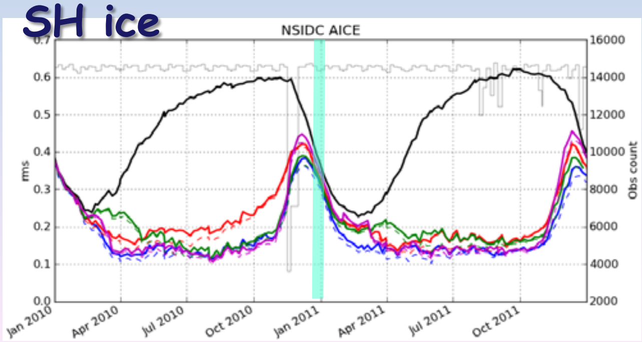
Ice assimilation results

01/01/2011

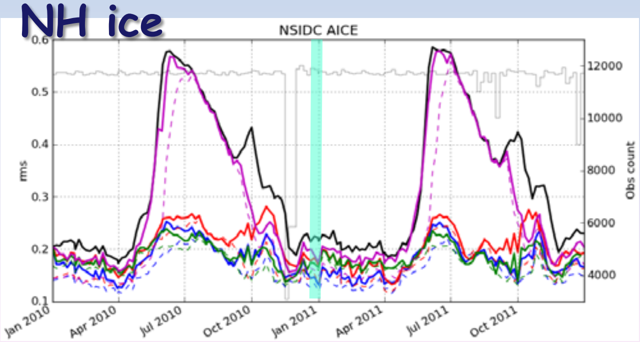
NSIDC control EUOI SAFE LIFE EnOI



SH ice



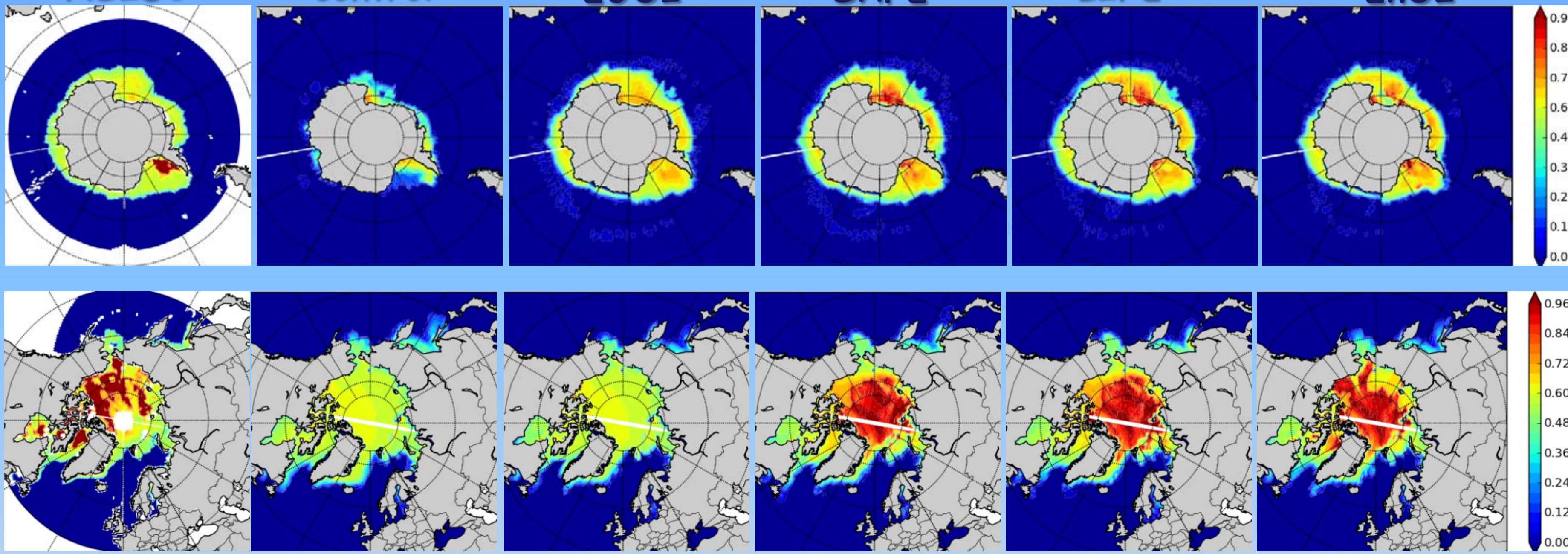
NH ice



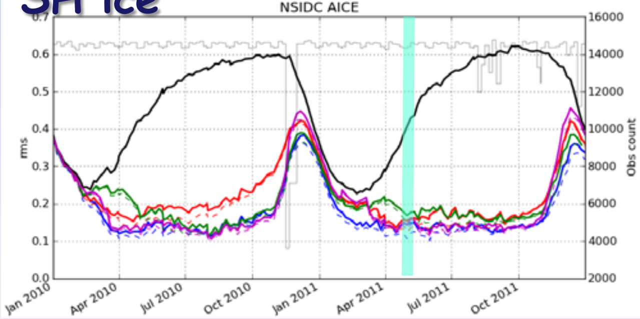
Ice assimilation results

05/01/2011

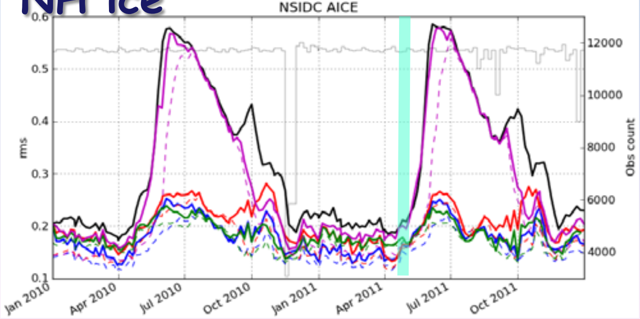
NSIDC | control | EUOI | SAFE | LIFE | EnOI



SH ice



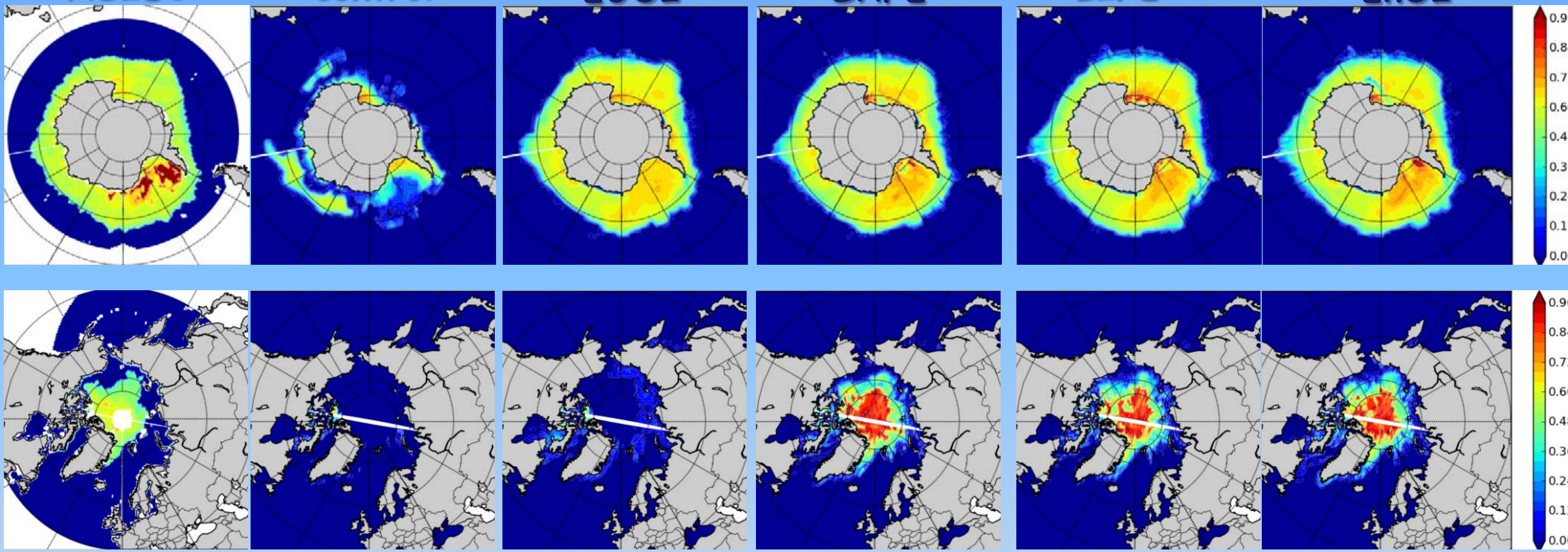
NH ice

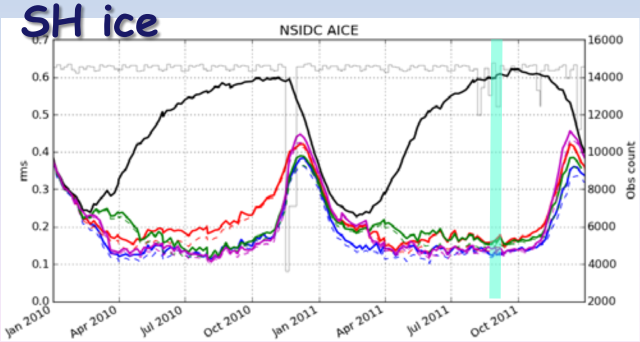


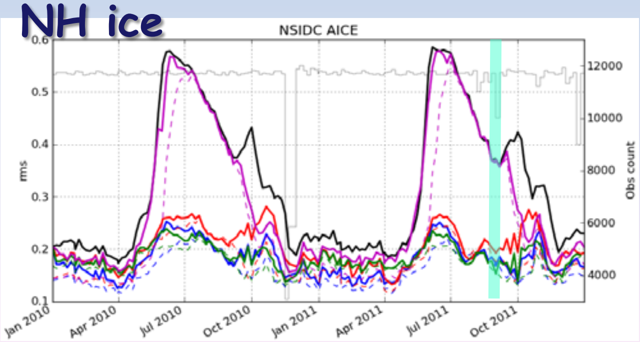
Ice assimilation results

09/01/2011

NSIDC	control	EUOI	SAFE	LIFE	EnOI

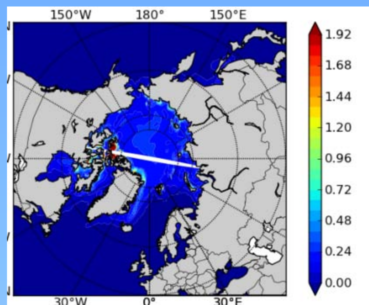






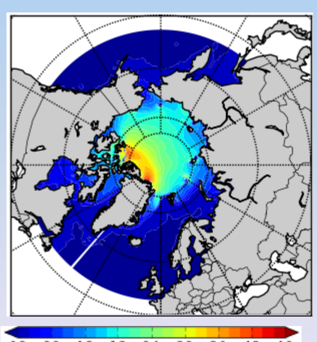
Ice assimilation results

Initial ice thickness 01/01/2010



Ice thickness at end of run (12/01/2011)

PIOMAS clim. (not assimilated)



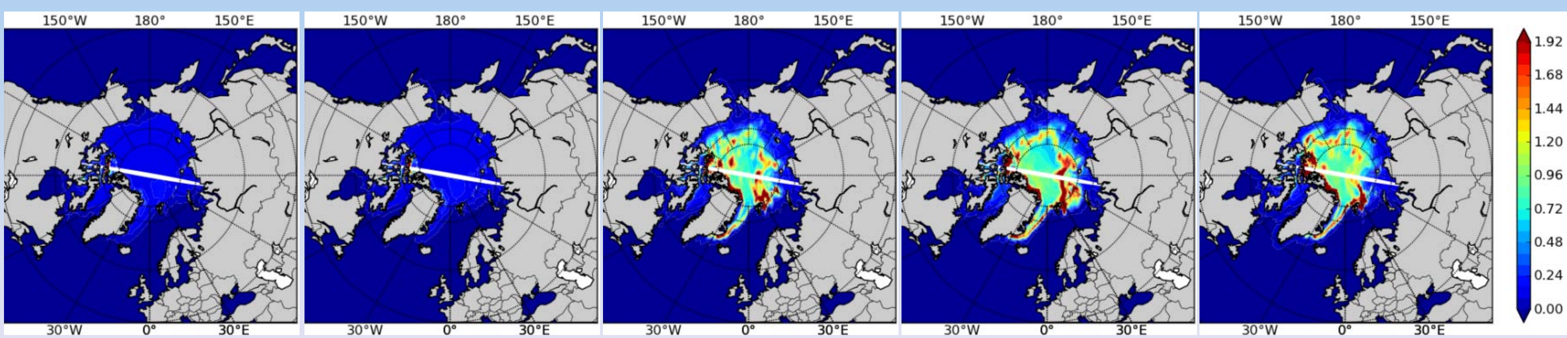
control

EUOI

SAFE

LIFE

EnOI



Conclusions

Some benefits of ensemble data assimilation schemes

- estimating background-error spatial distribution,
- multivariate update of unobserved variables,

are available with background-error covariances estimated from a single model run.