

A Consortium for COnvection-scale modelling
Research and Development

Algorithmic developments in ACCORD

Magnus Lindskog on behalf of ACCORD data assimilation team

Contents

- B-matrix
- HARMONIE-AROME 4D-Var
- Application and improvements of ENVAR
- Exploring some complementary algorithmic avenues
- Background errors in observation space
- Summary and conclusions

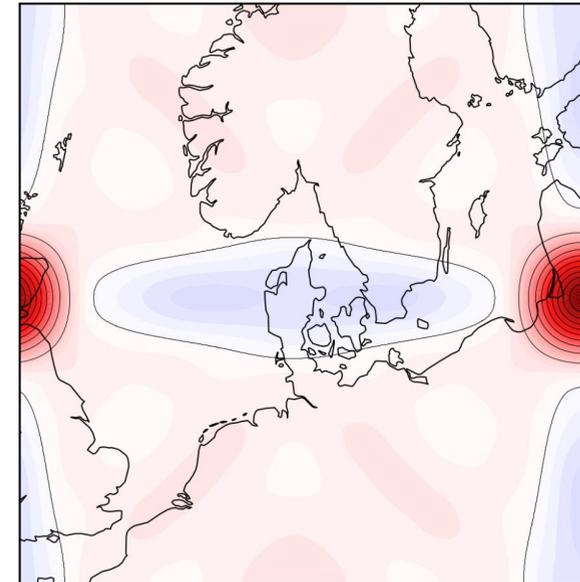
Recent Focus Areas

- Large recent efforts into porting of our data assimilation algorithms too OOPS (Object Oriented Prediction System).
- Exploit/adapt/tune ENVAR to be used with various types of ensemble systems.

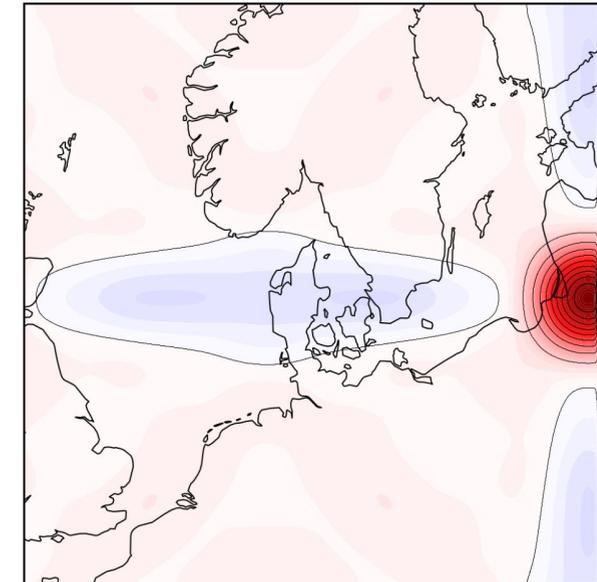
B matrix refactoring for OOPS

- **The current B matrix code is rusty:**
 - Obsolete coding and development standards, no testing
 - Hard-coded variables and operators
 - Almost impossible to compile/run on local machines
 - Inconsistencies with the B matrix calibration tool FESTAT, which scales poorly
 - Painful development with multiple IAL cycles
- **New generic implementation of the spectral B within the SABER framework, interfaced with OOPS:**
 - Cleaner object-oriented C++ code
 - Easier maintenance and development, complete testing
 - Improved calibration process
 - Backward compatibility with the current file format
 - Possible combination with other SABER blocks
 - Enhanced collaboration with other centers (UKMO)

Old code with an 11-point biperiodization zone



New code with an 81-point biperiodization zone



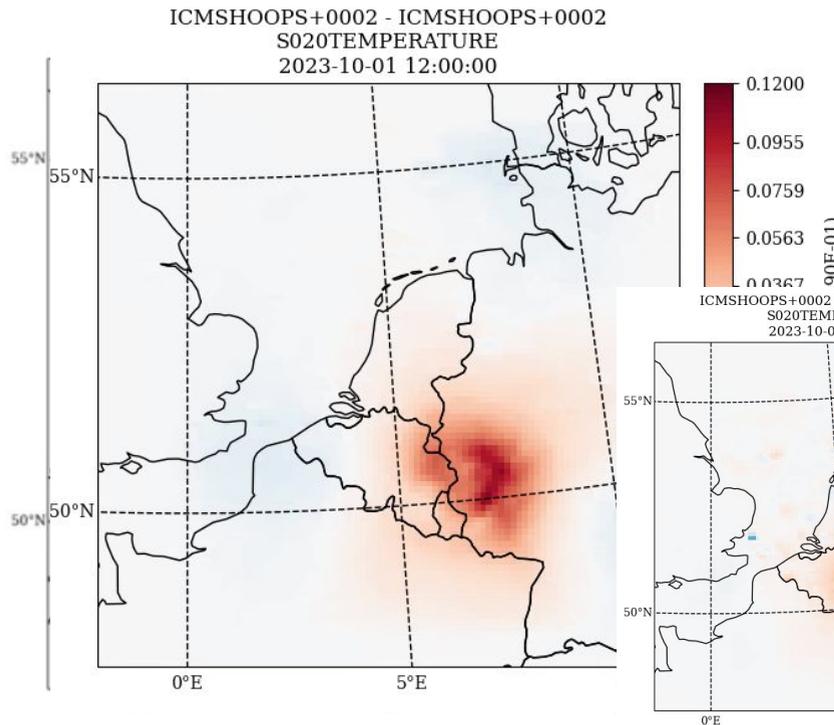
Example of possible enhancement: using an additional SABER block to extend the periodization zone and avoid increment wrapping around the domain, without changing the periodization in the minimization/forecast

(See Benjamins presentation)

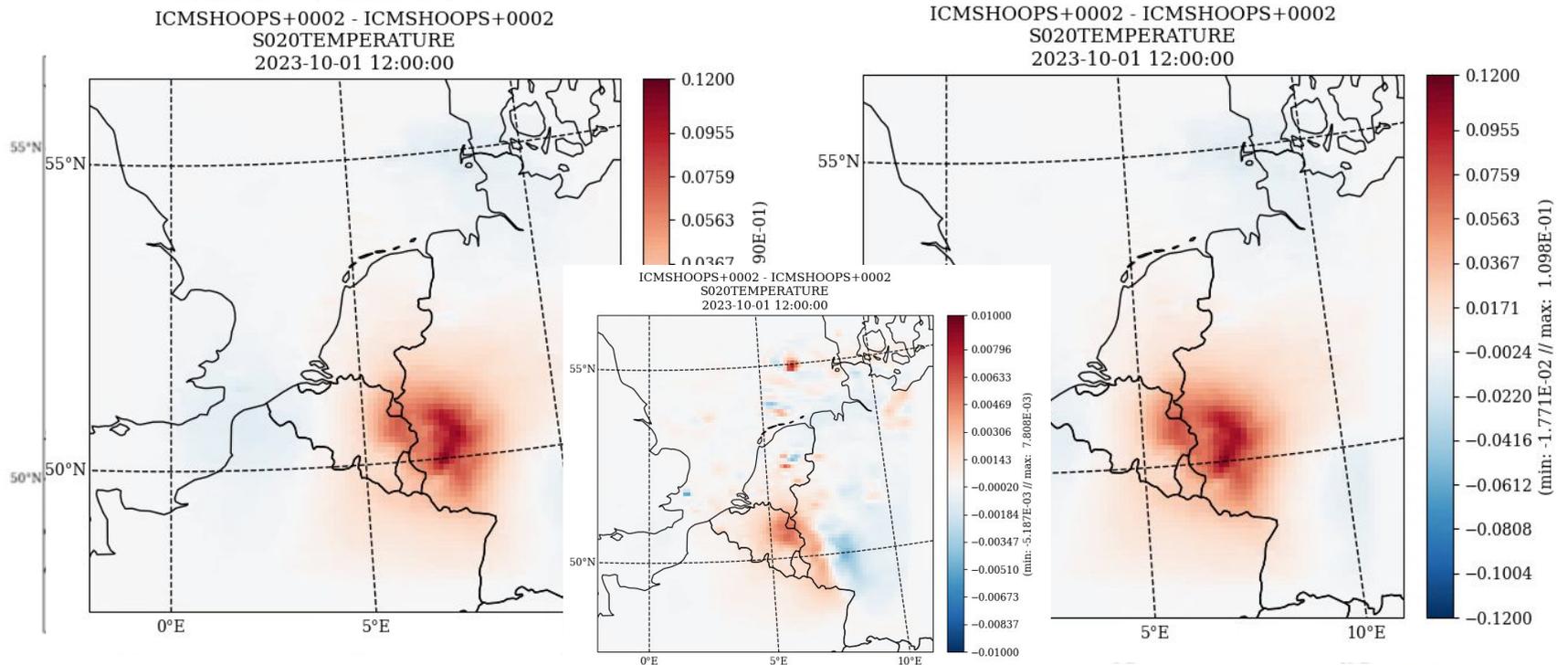
HARMONIE-AROME OOPS 4D-VAR

- Change of resolution of the increment in spectral space implemented in OOPS (as MASTERODB version)
- Equivalence of OOPS -vs- MASTER for a 2 outer loop (2x,1x resolution respect to back), demonstrated for a single obs experiment.
- Working for the equivalence of the convergence of VARBC coefficients, when assimilating satellite obs.
- A next step will be to perform a high-resolution (2.5 km) long run comparison of OOPS -vs- MASTER 4D-Var with an all-obs assimilated operational-like environment.

OOPS 2x Fullpos

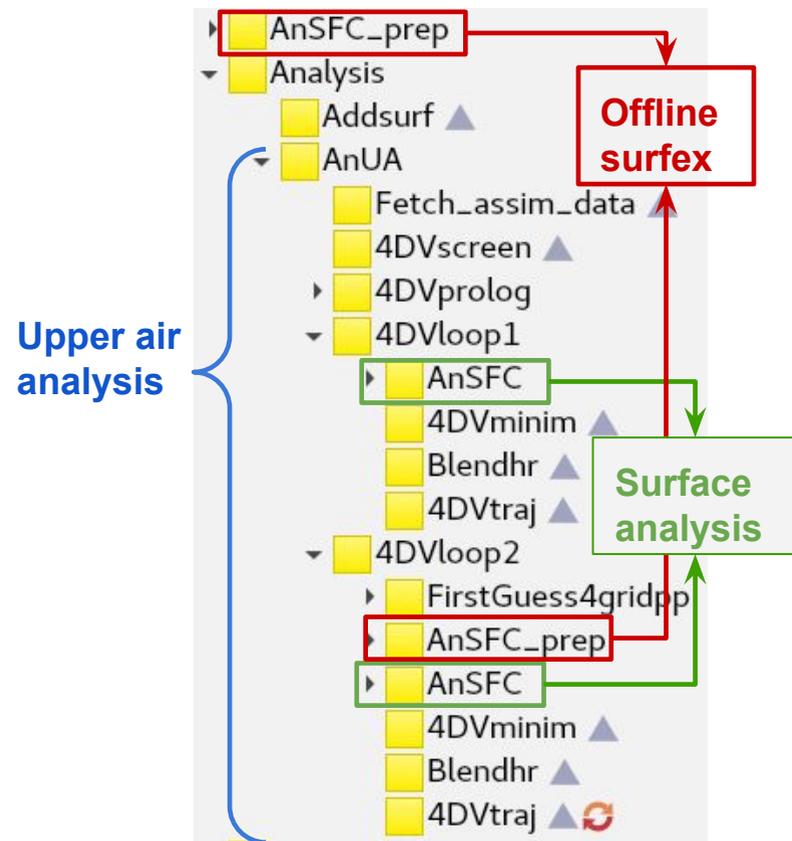


OOPS 2x SpecChg DX



— *Temperature analysis increments at model level 20 (~420 hPa) for one single observation (AMDAR T and 2 h into the window) assimilation. Grid point change resolution of increment (left) -vs- spectral change (right)*

Towards a weak surface/upper-air coupling in HARMONIE-AROME 4D-Var



- Performing surface assimilation (**SEKF**) inside each outer loop of upper-air assimilation (**4DVar**)
- This allows the sequential update of soil and atmospheric fields, and therefore their influence into each other's analysis within the same assimilation window
- The **offline-surfex** block computes the sensitivities from 2m measurements to soil. The recomputation of the atmospheric forcing in the second loop with the updated upper-air fields makes these sensitivities flow-dependent

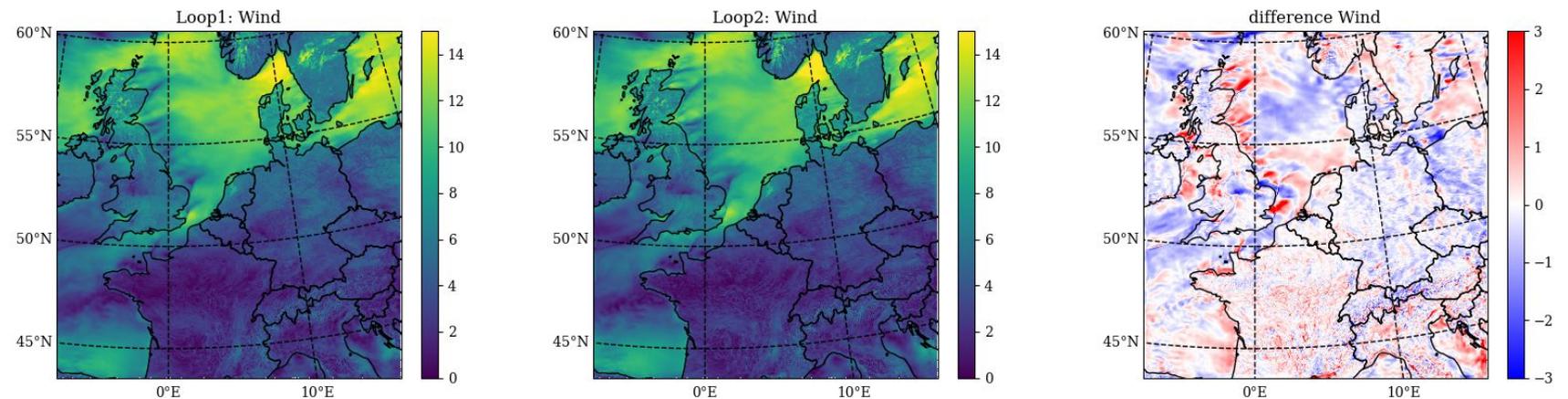


Figure: Wind forcing used in the first (left) and second (middle) loop, and their difference (right).



Meteo France LAM : Generalisation of the EnVar scheme under OOPS

Operational implementation of OOPS (all DA) & 3DEnVar (control run) in cy48 since 15/10/2024, with major positive impacts on forecast scores and on consistency of successive runs.

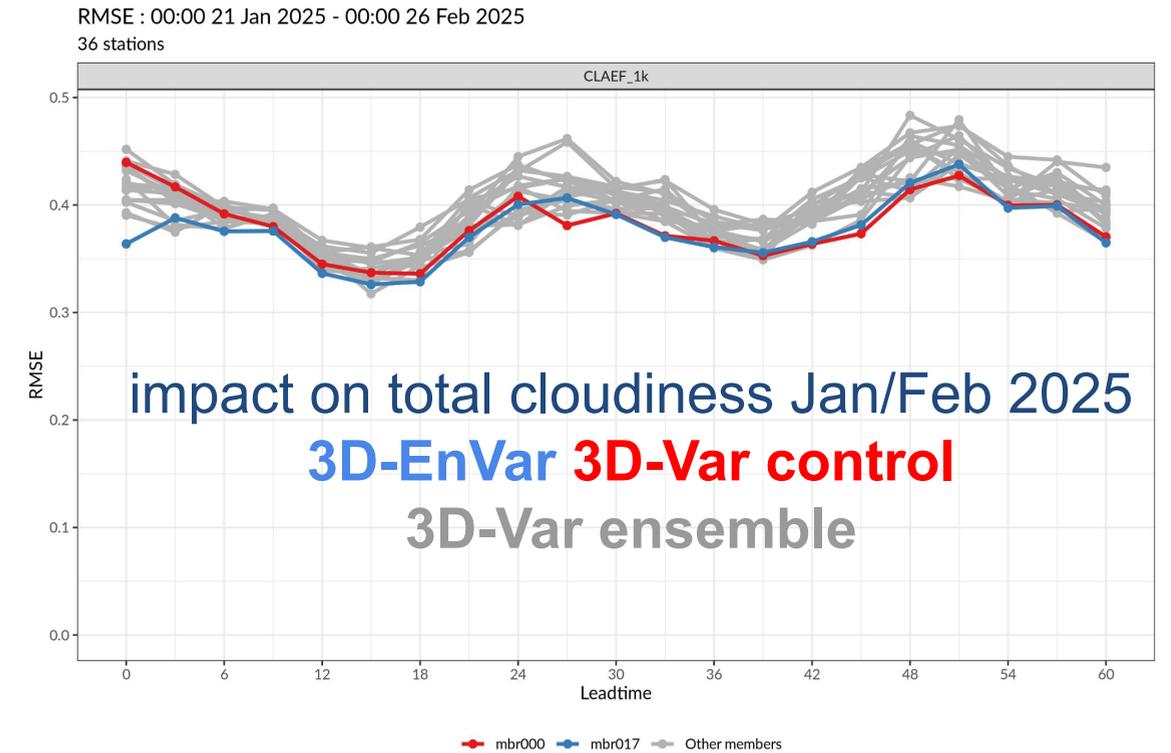
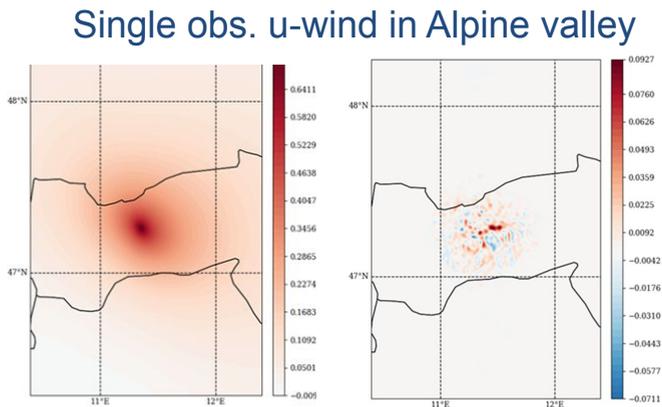
	o-suite (cy48)	e-suite (cy49)
AROME-France (deterministic)	3DEnVar (OOPS) 1.3km, L90 Hourly cycle	4DEnVar + SDL + direct assimilation of reflectivities
AROME EDA	3DVar (OOPS) 3.2km, L90 3hr cycle 50 members	3DEnVar + SDL
AROME NWC (nowcasting)	3DVar (OOPS) 1.3km, L90 Hourly short cut-off analysis + 6hr forecasts assimilation window [-10mn ; 10mn]	3DEnVar + SDL + direct assimilation of reflectivities

**The e-suite in cy49 has been in service since July 2025.
All DA configurations have switched to EnVar in cy49.**

(See Valérie's presentation)

3D-EnVar as a control member of Claef1k

- Claef1k: regional km-scale ensemble with 16 members, used in 3D-EnVar (32 members in lagged mode)
- Positive impact on several variables wrt to 3D-Var control in 2025 winter period.
- Experience over summer 2025: reduction of inflation needed for stable performance.
- SDL under evaluation to replace vertically-dependent horizontal localization.
- Candidate for operationalization in January 2026



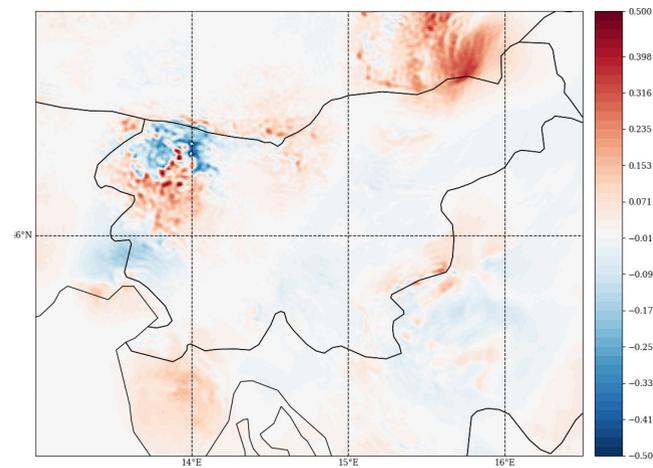
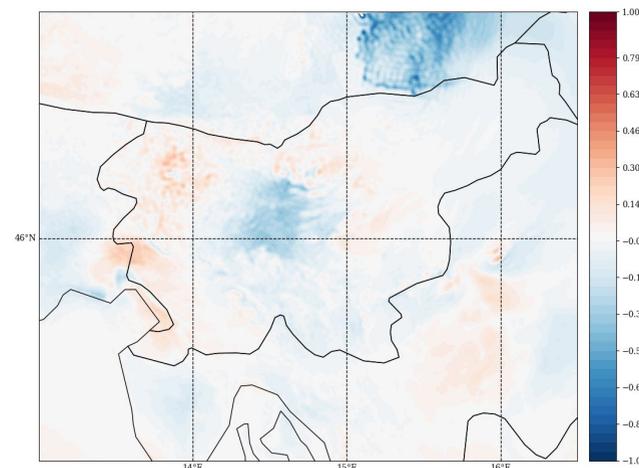
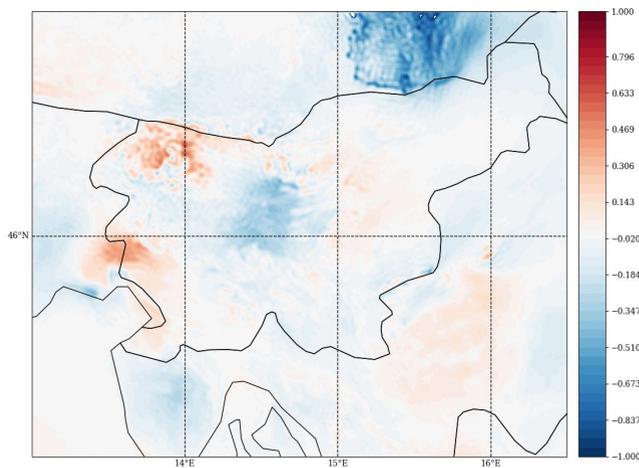
Verification for CCtot

Sub-ensembles for 3DEnVar perturbations

Sub-ensembles in 3DEnVar

Perturbations that belong to different sub-ensembles scaled separately, resulting in more realistic increments

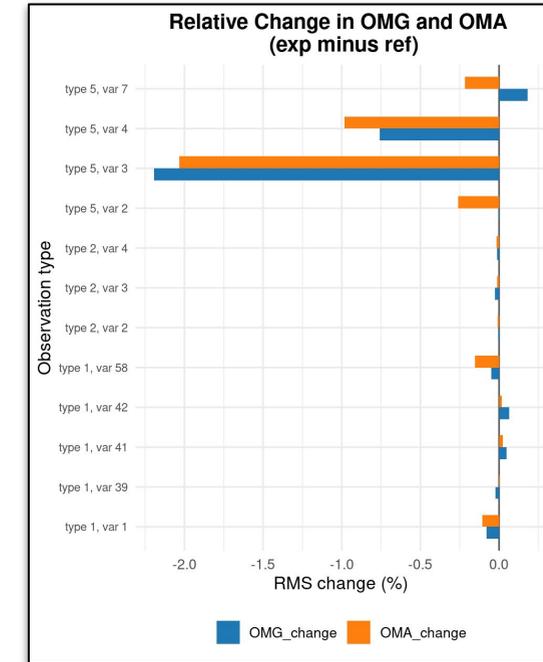
$$\epsilon_B^{i,l} = \frac{1}{\sqrt{N - N_s}} (\mathbf{X}_B^{i,l} - \langle \mathbf{X}_B \rangle^l)$$



one ensemble (16 + 16) two sub-ensembles (each 16)

difference

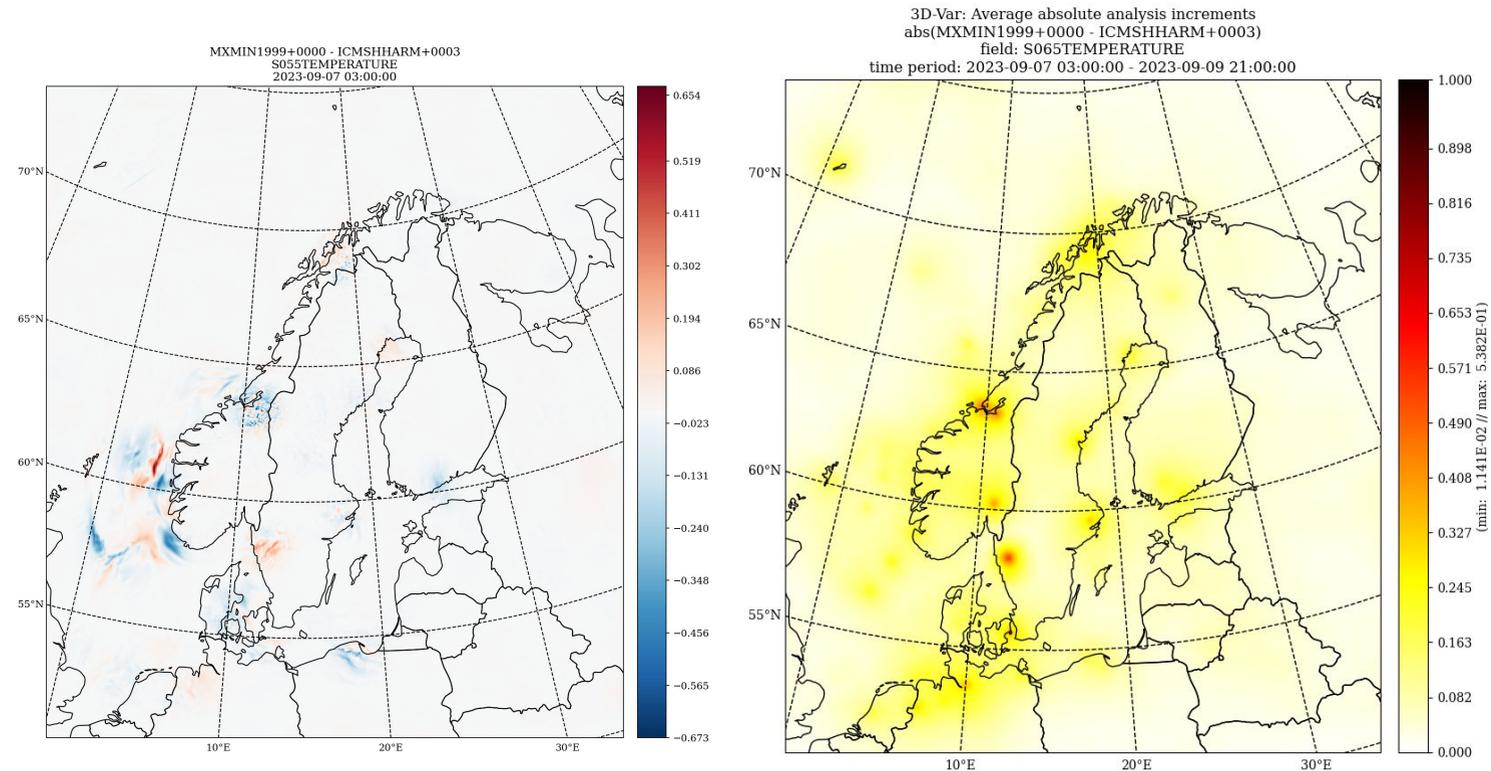
Analysis increment of low-level (L88) temperature



impact on
OMG/OMA
statistics
(reduction for
radiosondes)

HARMONIE-AROME ENVAR

- Ongoing exploitation of ENVAR approaches in HIRLAM.
- Quality of the ENVAR heavily dependent on ensemble.
- Ensemble data under investigation
 - ECMWF global ensemble forecasts
 - MetCoOp limited area ensemble and various flavours of this.
 - UWC-W
- First assimilation experiments, but further extended studies planned.
 - Some work on localisation tuning in cy49



ENVAR Temperature assimilation increments (K) at lowest model level (65) for (one particular assimilation cycle (left) and absolute increment value accumulated over three days of assimilation (right).

Exploring some complementary algorithmic avenues

- Cloud data assimilation by additional cost function penalty term
- Gaussian integrals for modelling of background error covariances
- Machine learning for various components in data assimilation
 - ML-based observation operators for satellite radiances.
 - Introduction of hybrid ML/DA approach for improving the MW observation operator over sea ice.
 - Enhancing quality of products to be used for assimilation.
 - Bias correction of satellite radiances.
 - Observation quality control procedures (crowdsourced observations).
 - Ensemble generation techniques of relevance for ensemble-based data assimilation.
 - Constraining the to a "balanced" manifold in latent-space.
 - First prototype for an EnKF using ensemble from AIFS/Bris.

Cloud data assimilation

Observation operator defined by the Penalty Function

Observations:

- 1) Ceilometer observation of cloud base altitude (p_b)
- 2) Satellite observation of cloud top altitude (p_t)

Penalty Function definition:

$$P(vl) = \begin{cases} \frac{RH(vl) - RH_{crit}(vl)}{\beta}, & \text{for } vl \in [h_b, h_t] \\ \frac{RH(vl) - \alpha(vl)}{\beta}, & \text{for } vl \notin [h_b, h_t] \text{ and } RH(vl) > \alpha \\ 0, & \text{otherwise} \end{cases}$$

where vl – vertical model level; RH – model relative humidity;
 RH_{crit} – critical relative humidity threshold above which cloud starts to form; α - threshold for the humidity level below saturation, β -uncertainty in RH_{crit}

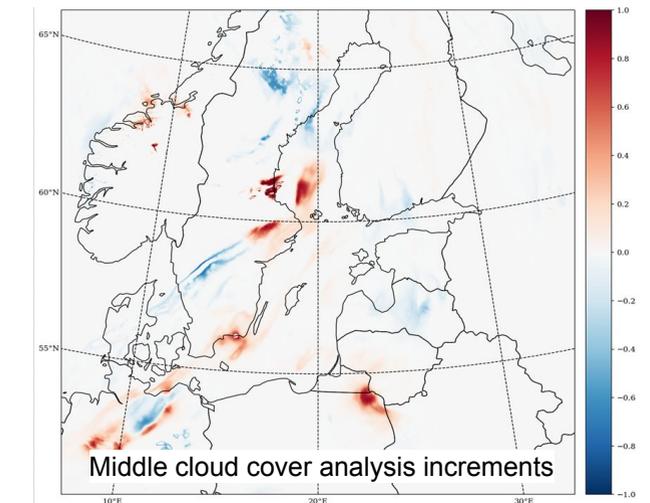
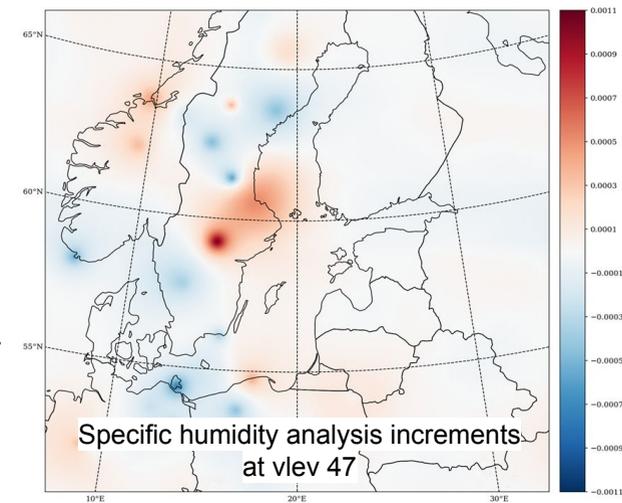
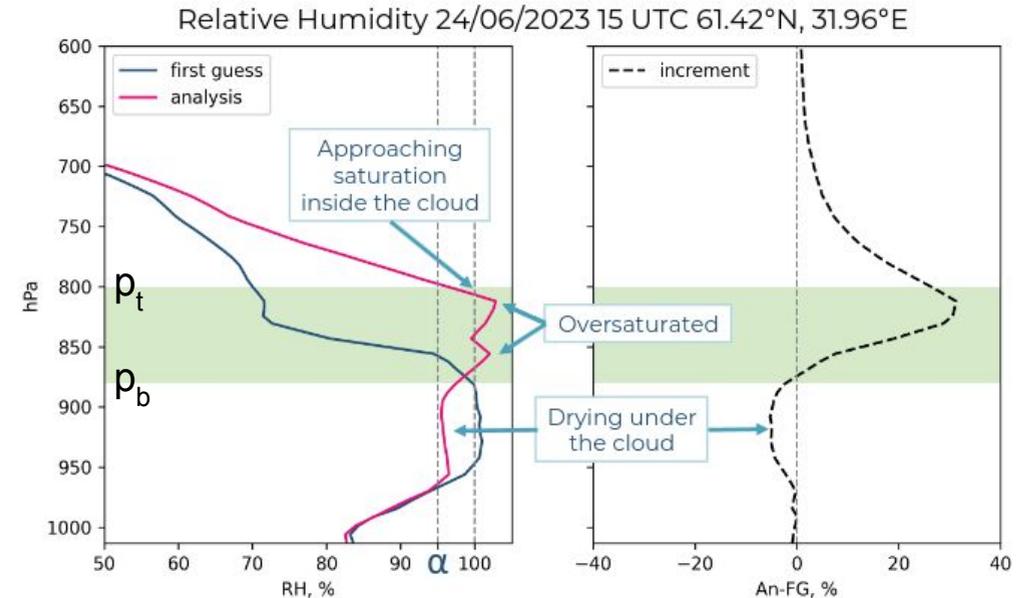
$$J = J_b + J_o + J_c$$

$$J_c = \frac{1}{2} (\mathbf{y}_c - (\mathbf{y}_c + \mathbf{P}))^T (\mathbf{y}_c - (\mathbf{y}_c + \mathbf{P})) = \frac{1}{2} \mathbf{P}^T \mathbf{P}$$

Real observations experiment 20240816_09
 SYNOP (ceilometers) + NWCSAF (CTTH)



Single pseudo-observation illustration

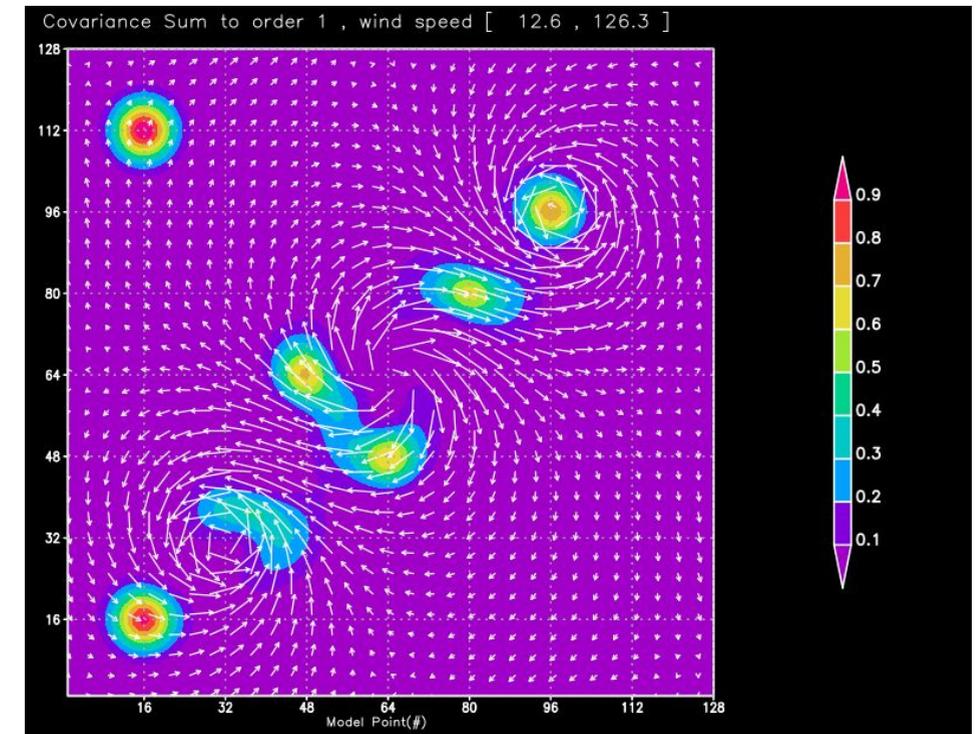


Modelling Background Error Covariances with Gaussian Integrals

- It is possible to give flow-dependency to the \mathbf{B} by introducing a positive quadratic form that depends on the \mathbf{V} field:

$$Z(S) = \int d\Delta_1 \dots d\Delta_N e^{-\frac{1}{2} \left(\Delta^T B^{-1} \Delta + \mu \operatorname{tr} \left[(\mathbf{V} \cdot \vec{\nabla} \Delta) (\mathbf{V} \cdot \vec{\nabla} \Delta)^T \right] \right) + \Delta^T S}$$

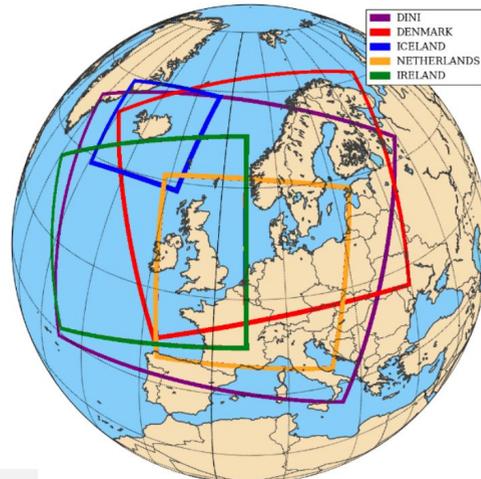
- Implementation of prototype and “Proof of Concept” successful.
- Computing performance analysis on CPUs and standard MPI and OpenMP parallelization APIs carried out.
- The idea is appealing, in the first place, for its potential utility in ENVAR, $\mathbf{B} \rightarrow \mathbf{B}[\mathbf{V}]$.



Bias-correction of microwave radiance observations using ML

- VarBC updates the bias coefficients β 's at every cycle:
- Update β 's with a NN instead of VarBC
 - Training: DINI domain
 - Testing: Dutch domain
- Use the NN predictions of bias coefficients in a Harmonie-Arome run, and compare:
 - Traditional H-A
 - ML bias correction H-A
 - Denial H-A

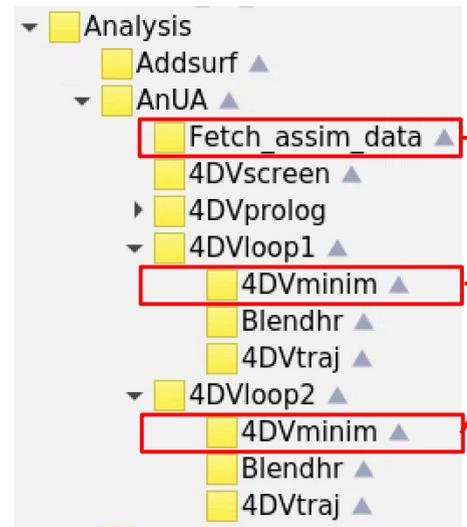
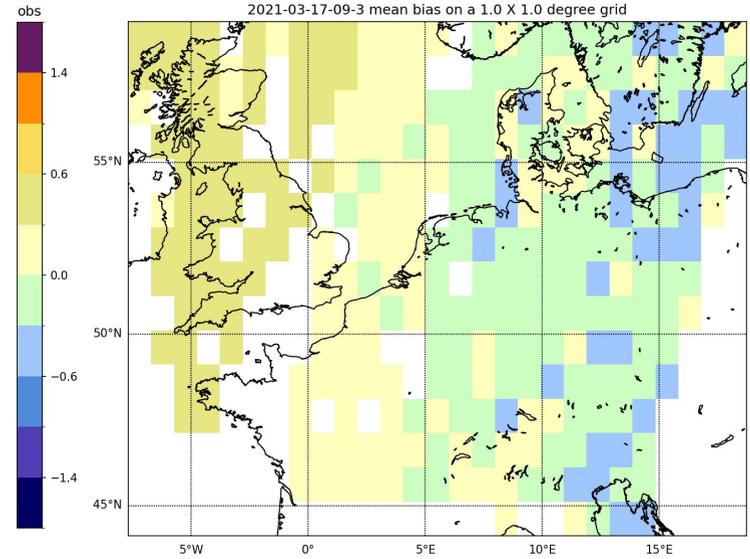
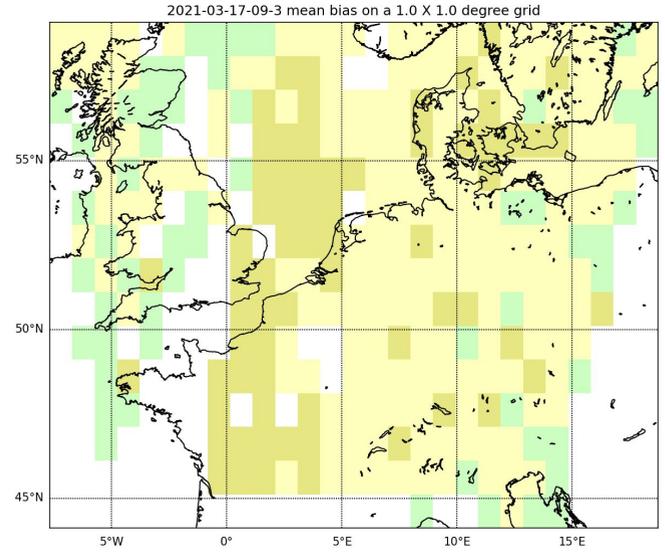
$$b(x, \beta) = \beta_0 + \beta_1 p_1(x) + \beta_2 p_2(x) + \dots + \beta_n p_n(x)$$



Mean bias from AMSU-A 09:00 UTC

Default exp

ML exp



Fetch my own ML VarBC cycles

Prevent update of ML VarBC cycles

BackGround error statistics in Observation Space (BGOS)

Application developed in the OOPS framework to compute background error standard deviation in observation space:

$$\sigma_b^o = \sqrt{\frac{1}{N} \sum_{i=1}^N (HU\chi_i)^2}$$

where

σ_b^o bg error std in ob space

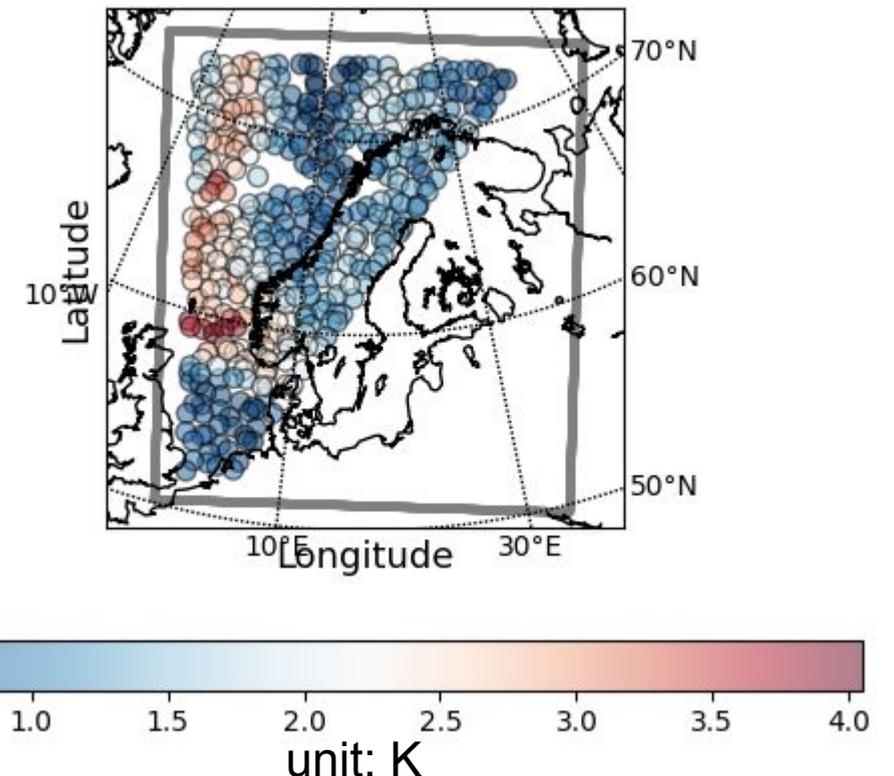
N sample size

U is the series of transform applied to get a unit B matrix in minimization, H is the observation operator (tl or nl) and χ_i is the control vector (containing Gaussian errors) for the individual member i.

$$\left(\begin{array}{l} \delta x = U\chi \\ J_b = \frac{1}{2} \delta x^T B^{-1} \delta x = \frac{1}{2} \chi^T \chi \end{array} \right)$$

Example

MHS channel 3 background error standard deviations in observation space 15 July 2022, 00 UTC. MetCoOp domain.



(corresponding observation error standard deviation is 1.8 K)

Summary and conclusions

- Ongoing porting of 4D-Var algorithm into OOPS and evaluation.
- Weakly coupled surface/upper-air data assimilation is being explored.
- Extended operational implementations of EnVar with OOPS and further research.
- Machine-learning provides an opportunity in various aspects of data assimilation.
- Important to have diagnostic tools, verification systems and metrics relevant for high-resolution data assimilation.

Thank you for your attention!