



Climate  
Change

# High Resolution Variational Data Assimilation over the CARRA-2 Region

**Swapan Mallick**

*Swedish Meteorological and Hydrological Institute (SMHI)*

**Acknowledgements:** *CARRA-2 team members and colleagues.*

October 2, 2024: 46th EWGLAM and 31th SRNWP Meeting, PRAGUE

**SMHI**





Climate  
Change

## OUTLINES:

- Introduction to Copernicus Arctic Regional Reanalysis (CARRA-2).
- HarmonEPS: Configuration and Data Assimilation Techniques.
- High-Resolution Data Assimilation Across a Large Domain for B-Matrix Computation.
- Seasonal Variation Impact of the B-Matrix on the CARRA-2 Domain.
- Results and Future Plans with Uncertainty Estimation Using Deep Learning (AI).



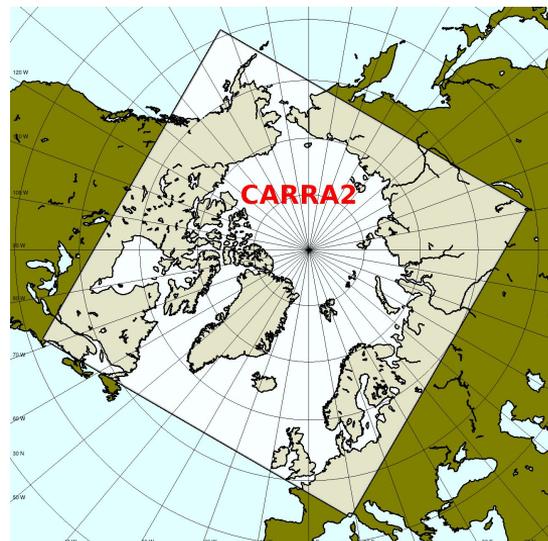
Climate  
Change

# Copernicus Arctic Regional Reanalysis (CARRA-2)

**Concept:** Reanalysis is a method of reconstructing past atmospheric states by using historical observations in conjunction with a weather forecasting model. CARRA-2 reanalysis is a high-quality climate data product by assimilating long time series of observations into Harmonie model and 3D-VAR data assimilation system to provide the best estimate of the atmospheric state.

HARMONIE-AROME	Cy46h1
Horizontal resolution	2.5 km
Number of grid points	2880 x 2880
Number of vertical levels	65
Model dynamics	Hydrostatic
Surface scheme	12-layer soil + 14-layer snow diffusive scheme
Upper-air data assimilation	3D-Var
Surface data assimilation	Simplified Extended Kalman Filter
Assimilation frequency	3-hourly, <b>6-hourly (for Ensemble)</b>
Output frequency	3-hourly analyses and hourly forecasts (3h after the 6h forecast range)
Forecast lengths	18h from 00 UTC and 12 UTC and 3h otherwise
Time coverage	September 1985 - December 2025 (timely updates are planned)
When data available	Full time series expected to be available spring 2026

## New generation Arctic reanalysis CARRA2: pan-Arctic extension

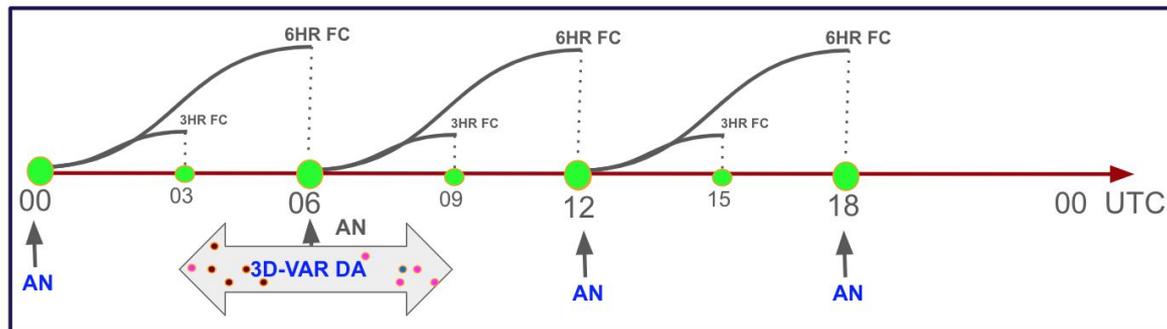


<https://climate.copernicus.eu/copernicus-arctic-regional-reanalysis-service>



Climate  
Change

# HarmonEPS: Configuration



**Model:** High resolution (2.5-km) convective perimete HARMONIE-AROME (CY46) regional numerical weather prediction model.

**Data Assimilation:** 3D-VAR for upper air observation assimilation from both in-situ and satellite sensors. Estimation of background errors from CARRA2-EDA with 10 ensemble members.

## Ensemble Prediction System (EPS):

- 9-member and a control.
- IC/BCs use members of an **ERA5 EDA** 10-ensemble members.



## B-Matrix:

### Cost Function for the Incremental Formulation:

$$\begin{aligned} J &= J_o + J_b \\ &= \frac{1}{2} \delta x'^T B'^{-1} \delta x' + \frac{1}{2} (Hx^b + H' \delta x' - y)^T R^T (Hx^b + H' \delta x' - y) \end{aligned}$$

$$B = \begin{pmatrix} \langle \varepsilon_1^2 \rangle & \langle \varepsilon_1 \varepsilon_2 \rangle & \dots \\ \langle \varepsilon_2 \varepsilon_3 \rangle & \langle \varepsilon_2^2 \rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

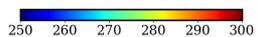
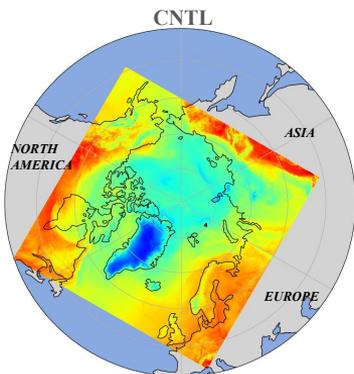
**STATIC**

The background error statistics (BGE) also referred as a Structure Functions, are produced through the standalone FESTAT (Forecast Error Statistics, MNC) method in control space. The control variables encompass vorticity, divergence, specific humidity, surface pressure, and temperature.

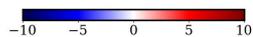
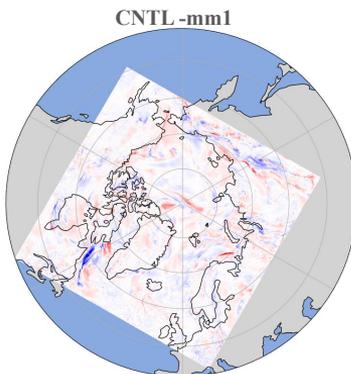
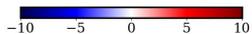
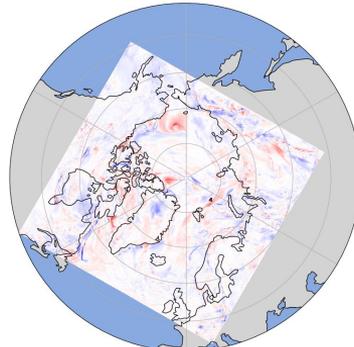


Climate  
Change

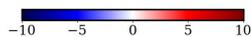
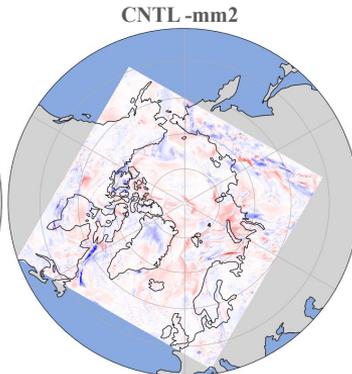
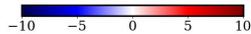
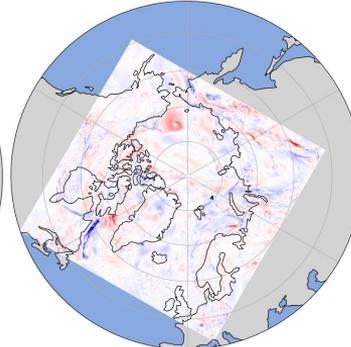
# Temperature (in K) at level 60 Valid on 20090902 at 00 UTC



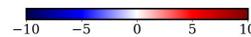
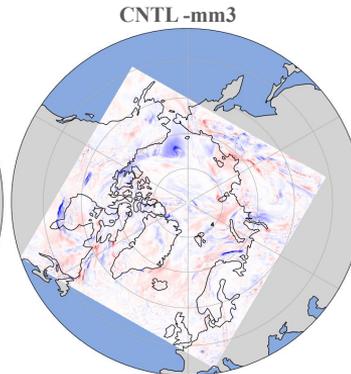
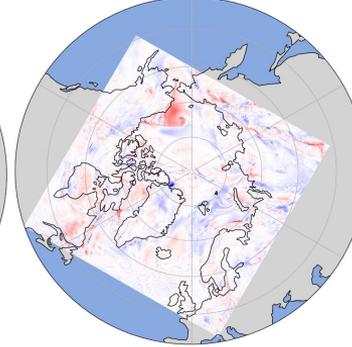
CNTL-mm4



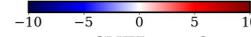
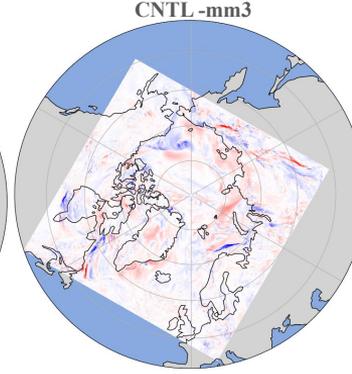
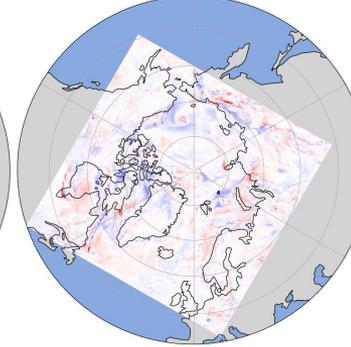
CNTL-mm5



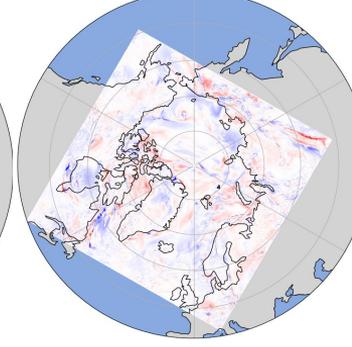
CNTL-mm6



CNTL-mm7



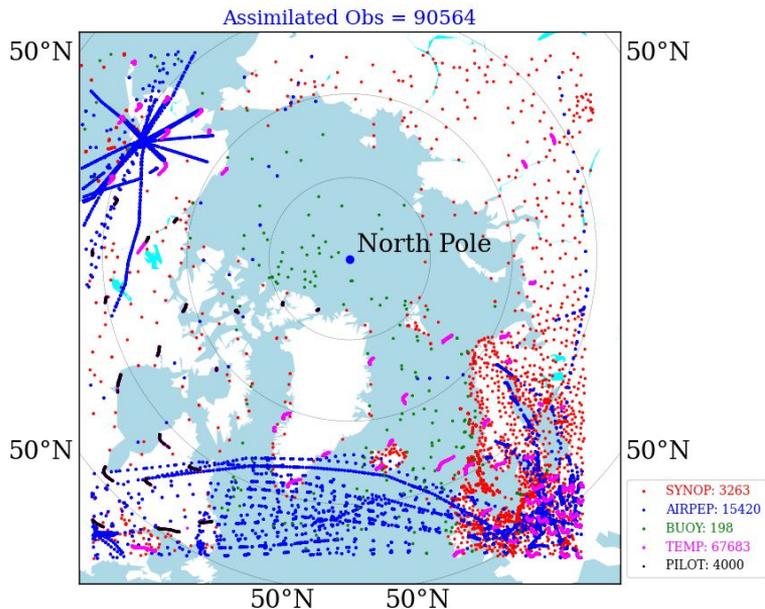
CNTL-mm8





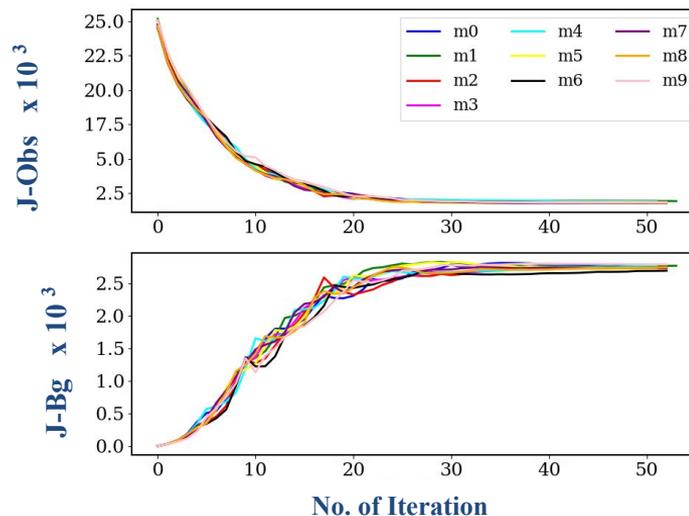
Climate  
Change

# Assimilated Observation and Cost Function:



Valid on 20220111 at 00 UTC

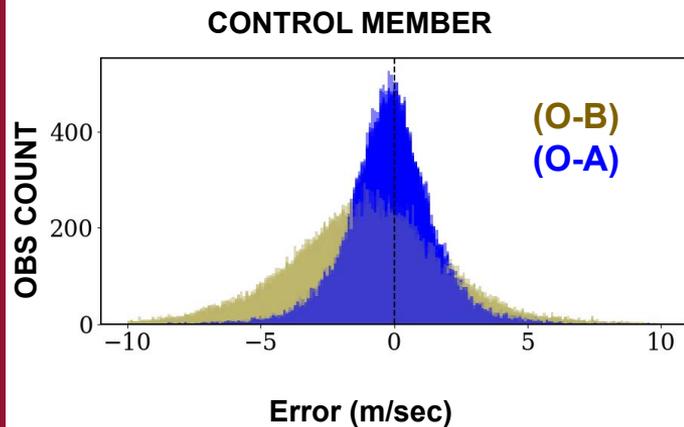
## Minimization of Cost Function ( $J$ ) for each Iteration:



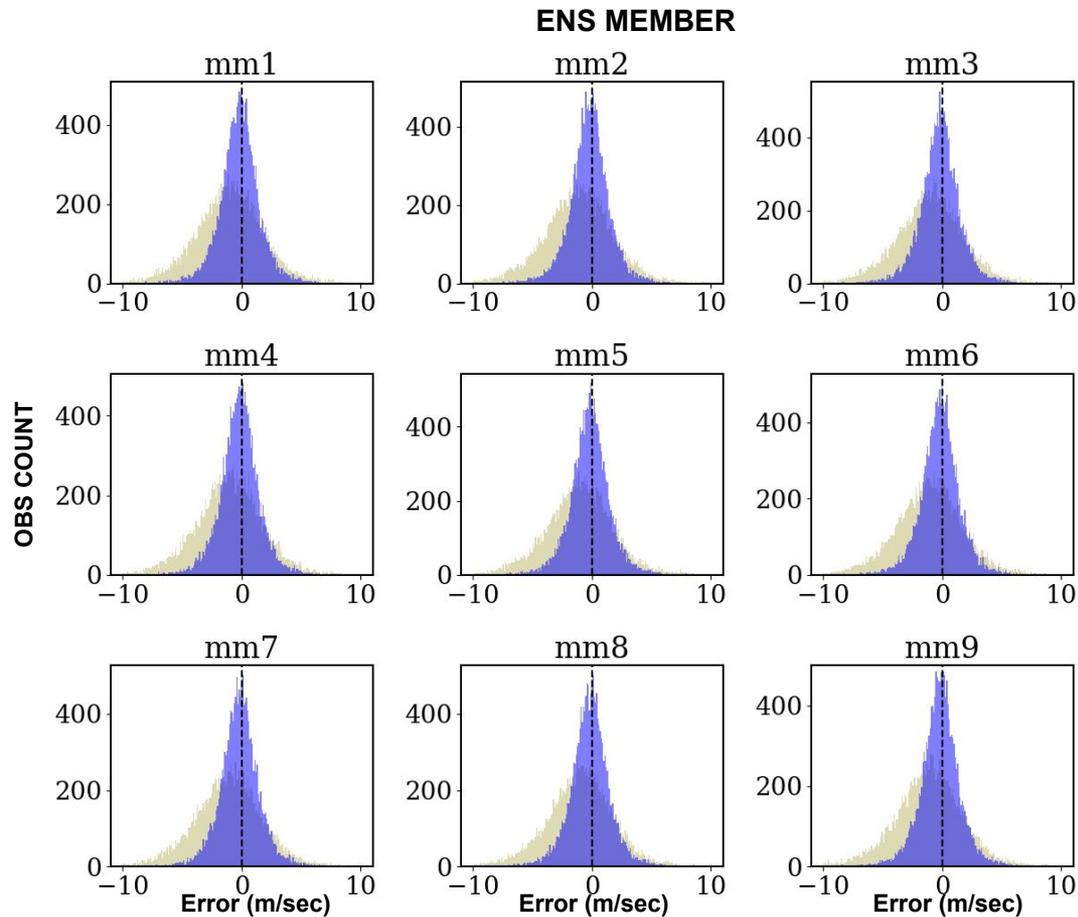


Climate  
Change

# Observation Space Statistics: ZONAL WIND (U)



Valid on 20220111 at 00 UTC



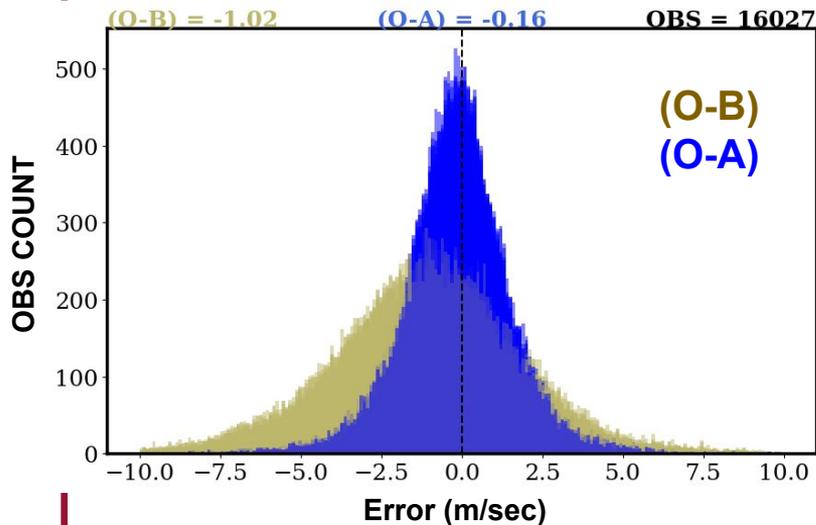


Climate  
Change

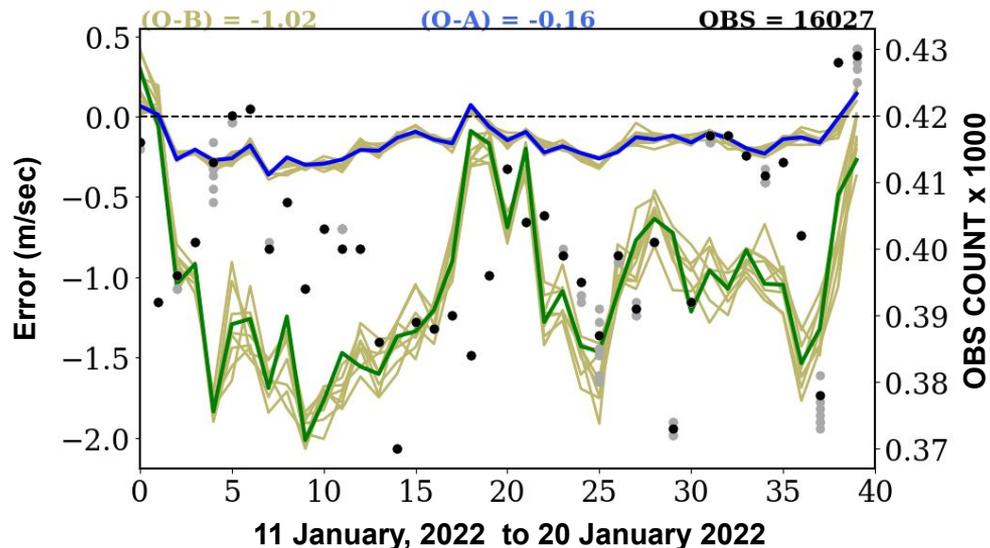
# Observation Space Statistics: ZONAL WIND (U)

## Systematic Error (or Bias in Zonal Wind)

DATA DISTRIBUTION



TIME SERIES ZONAL WIND (U)

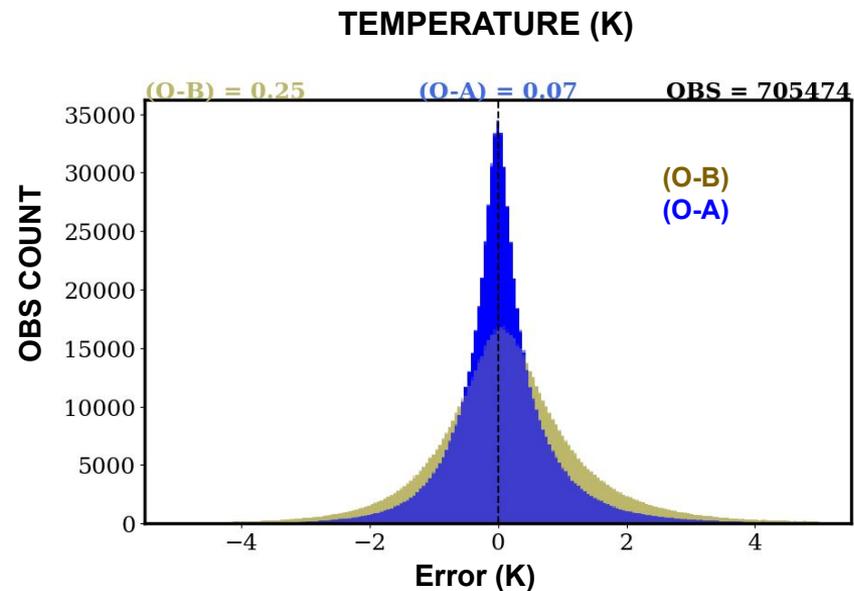
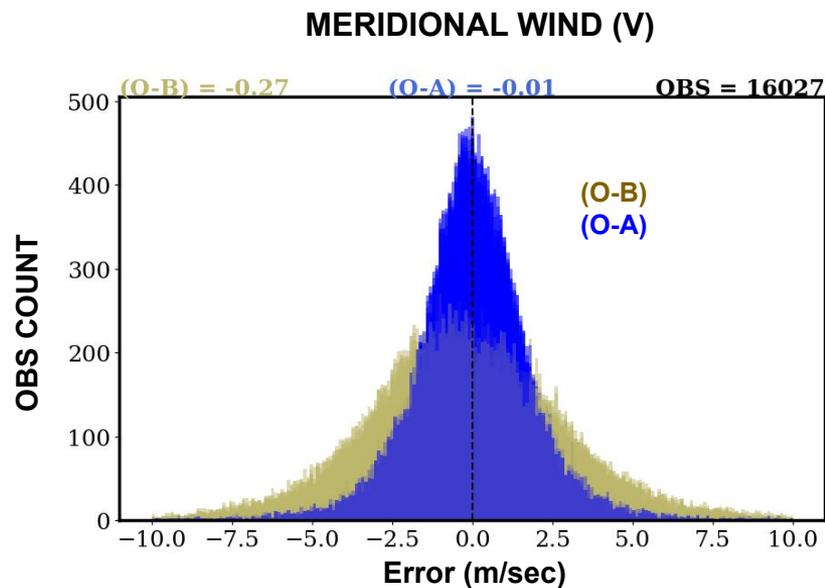


Valid: 11 January, 2022 to 20 January 2022



Climate  
Change

# Observation Space Statistics:





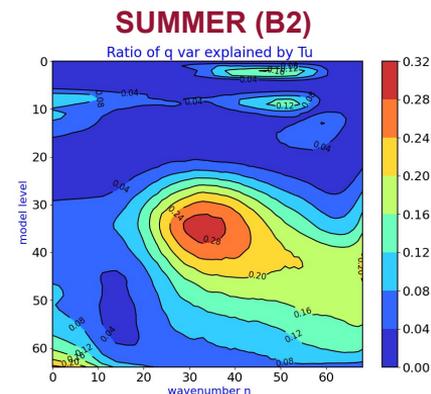
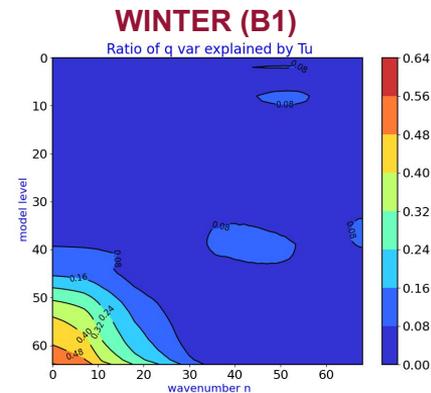
Climate  
Change

# Seasonal Variation of B-Matrix

## B-Matrix: Seasonal Variation

Cases	Valid	Season	Forecast hour	Total FC FILE (ENS mem x DA cyc)
B1	202201	Winter	6 HR	999 (9x111)
B2	202206	Summer	6 HR	999 (9x111)

*High-resolution analysis and a deterministic forecast were generated for December 1999 and September 2010 using the B1 and B2 matrices.*

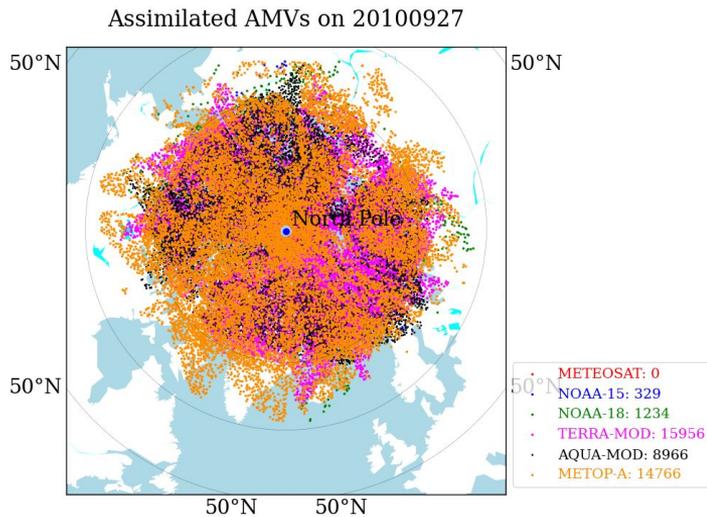


Percentage of the variance of specific humidity that is explained by unbalance temperature (Tu) as a function of height and wave number.

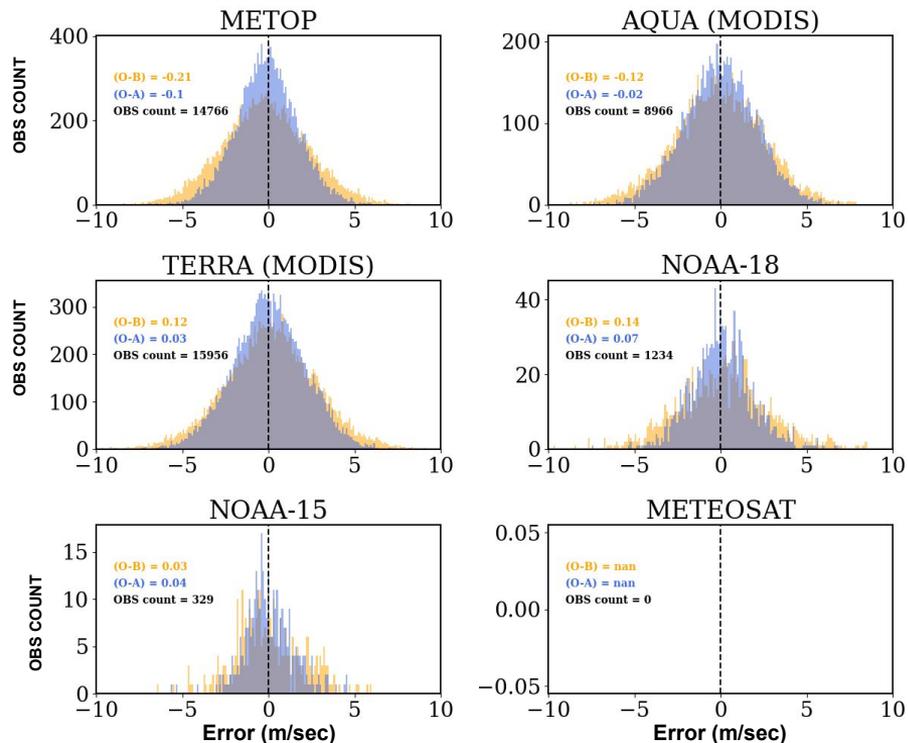


Climate  
Change

# Assimilated Atmospheric Motion Vectors (AMVs):



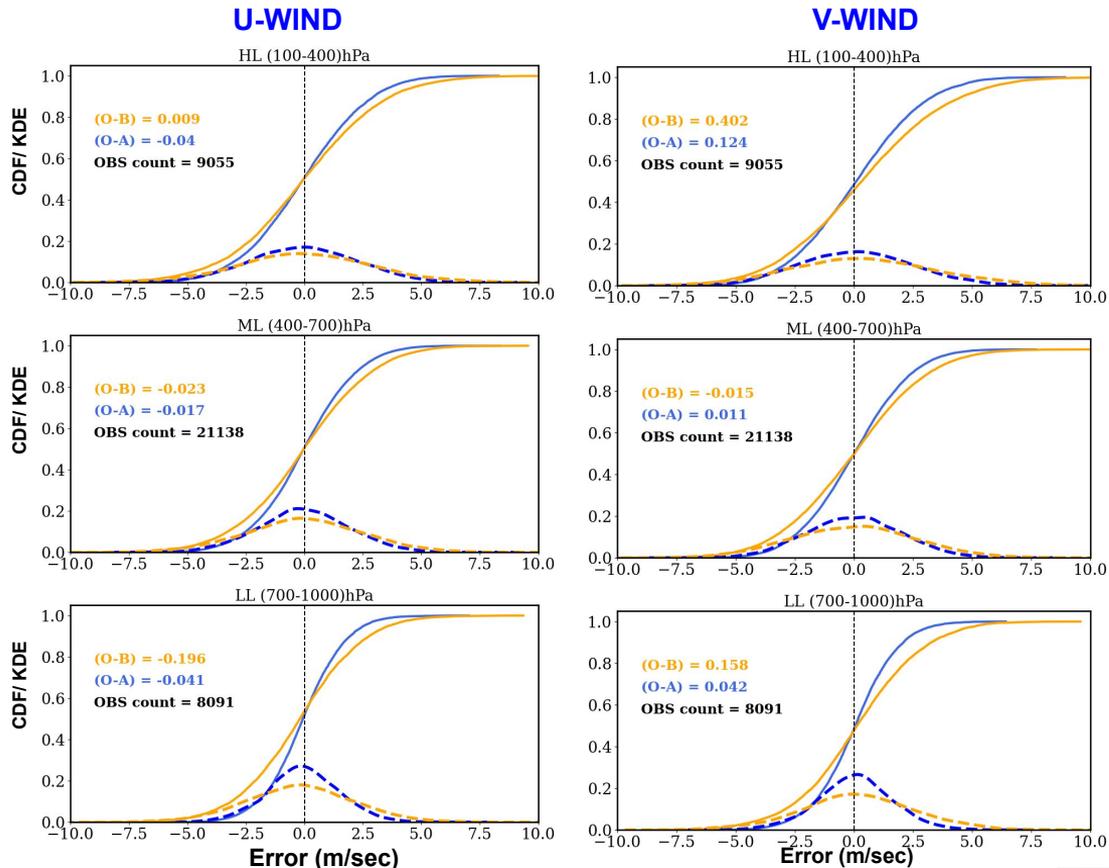
Valid on 20100927 and for all assimilation cycle





Climate  
Change

# Assimilated Atmospheric Motion Vectors (AMVs):



Statistics of the assimilated AMVs show significant improvement from the surface up to the 700 hPa level. Additionally, the spread of error is much lower compared to the higher levels.

CDF(Solid Lines): Cumulative Distribution Function.

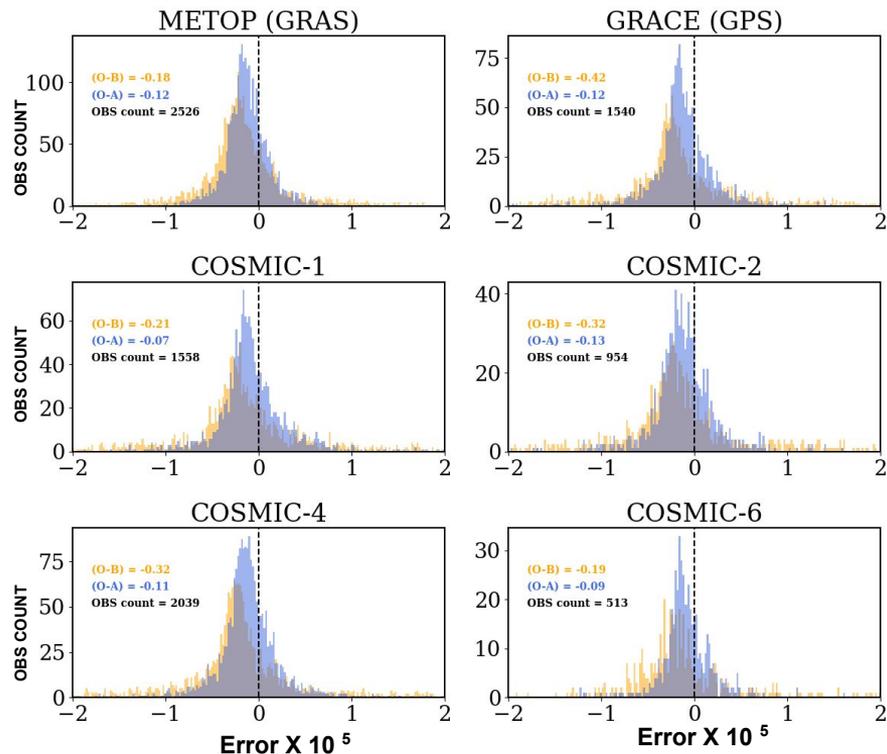
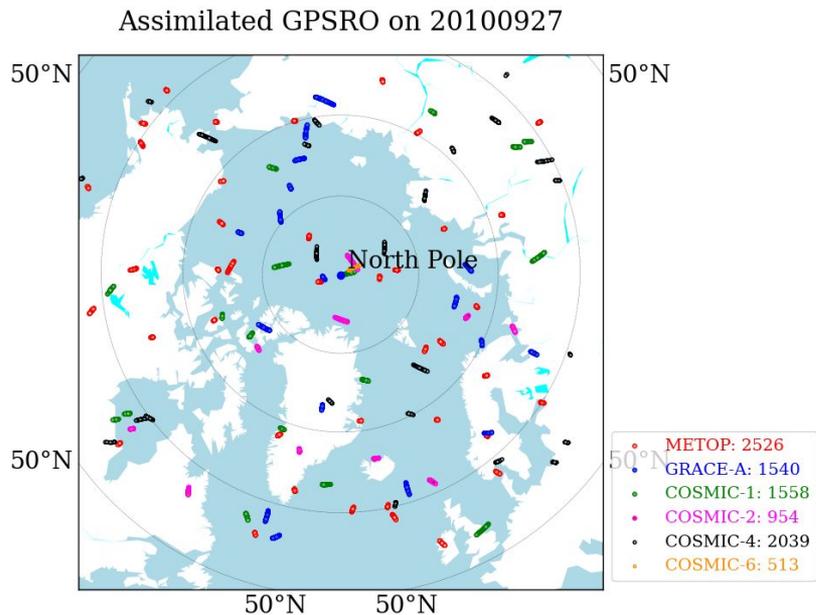
KDE (Desh lines) : Kernel Density Estimation.

Valid on 20100927 and including all assimilation cycle



Climate  
Change

# Assimilated GPSRO:



Statistics of the assimilated GPSRO (refractive index) show a systematic negative error in both the background (O-B) and analysis increments (O-A).

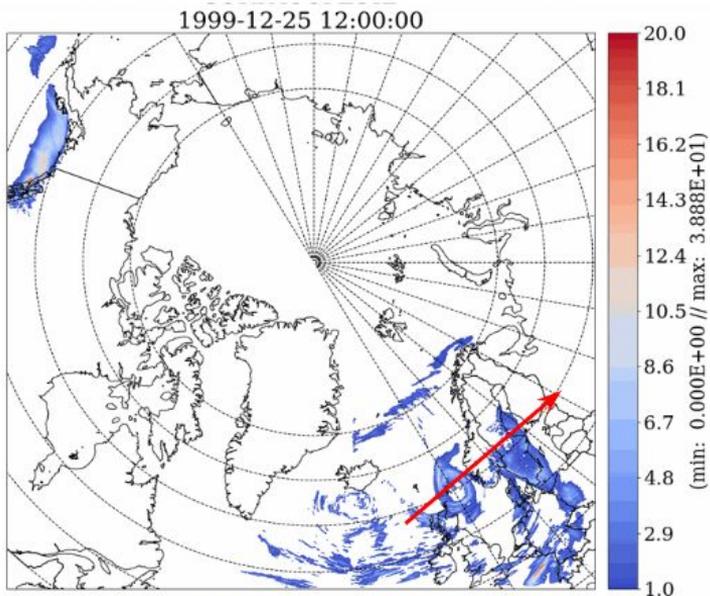


Climate  
Change

# Validation:

## Extreme Storms over Europe on 26 December 1999

### Accumulated Rain (kg/m<sup>2</sup>)

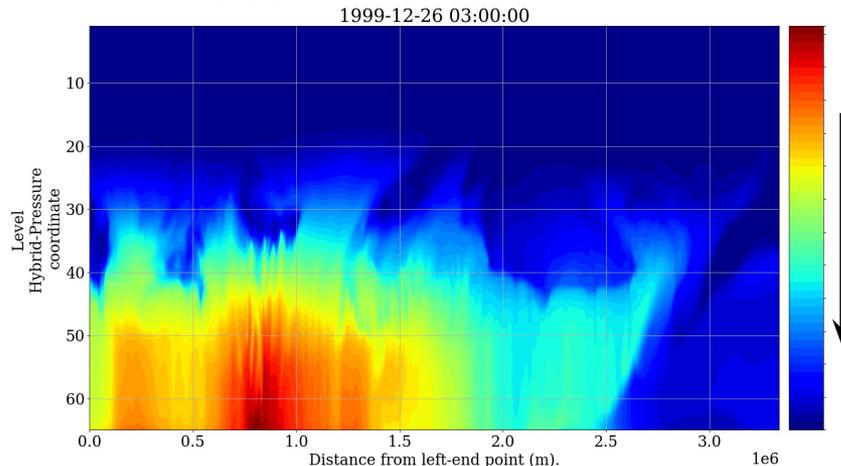


Analysis and Forecast generated using  
**WINTER B1-Matrix**

### Storm Lothar:(0000-1200 UTC)

It was caused by a low named Lothar, and left a trail of destruction from Denmark to southern Germany.

### Vertical Cross section of RH

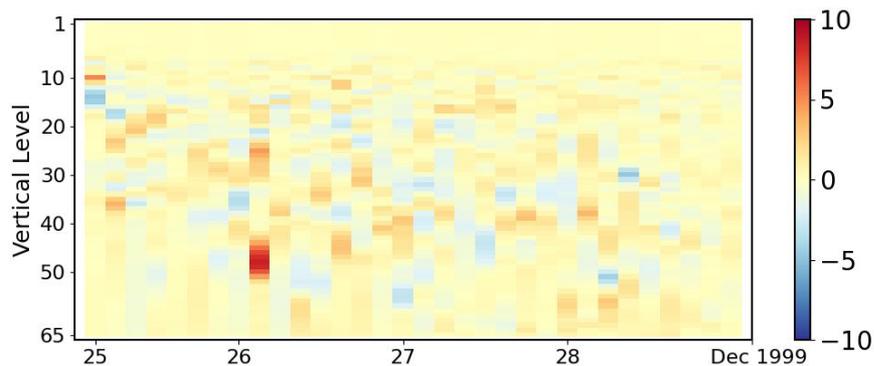




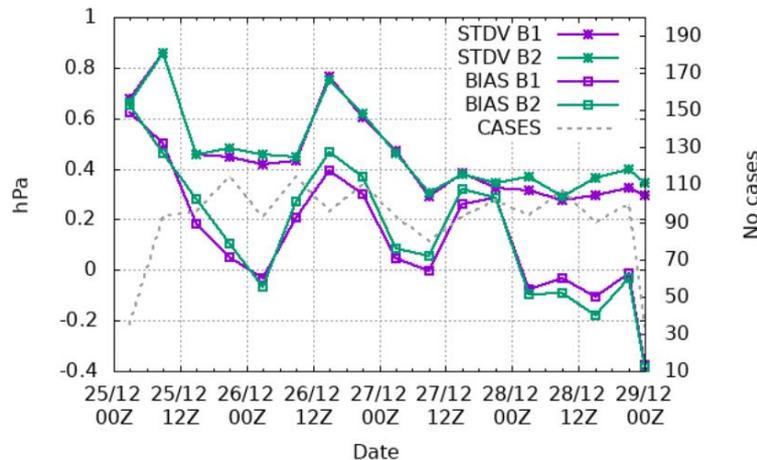
Climate  
Change

# Validation:

### 6-hr forecast difference (B1-B2) time series of Wind Speed (m/s)



Mslp  
Selection: Denmark 18 stations  
Used 00,03,...,21 + 00 03 06 09 12 15 18  
Averaging window: 6h



**B1- Analysis and Forecast generated using WINTER B-Matrix**

**B2- Analysis and Forecast generated using SUMMER B-Matrix**



Climate  
Change

## SUMMARY:

- Data assimilation in CARRA2 re-analysis plays a important role.
- High quality of analysis needed good quality of observation after quality control and data thinning.
- Assimilation of high-quality observations is important. It is always essential to incorporate good-quality observations into the CARRA2 domain.
- The seasonal variation of the B-matrix are significant in enhancing both analysis and forecast skill.





Climate  
Change

*Thanks for your Attention !*

[Swapan.Mallick@smhi.se](mailto:Swapan.Mallick@smhi.se)

**SMHI**