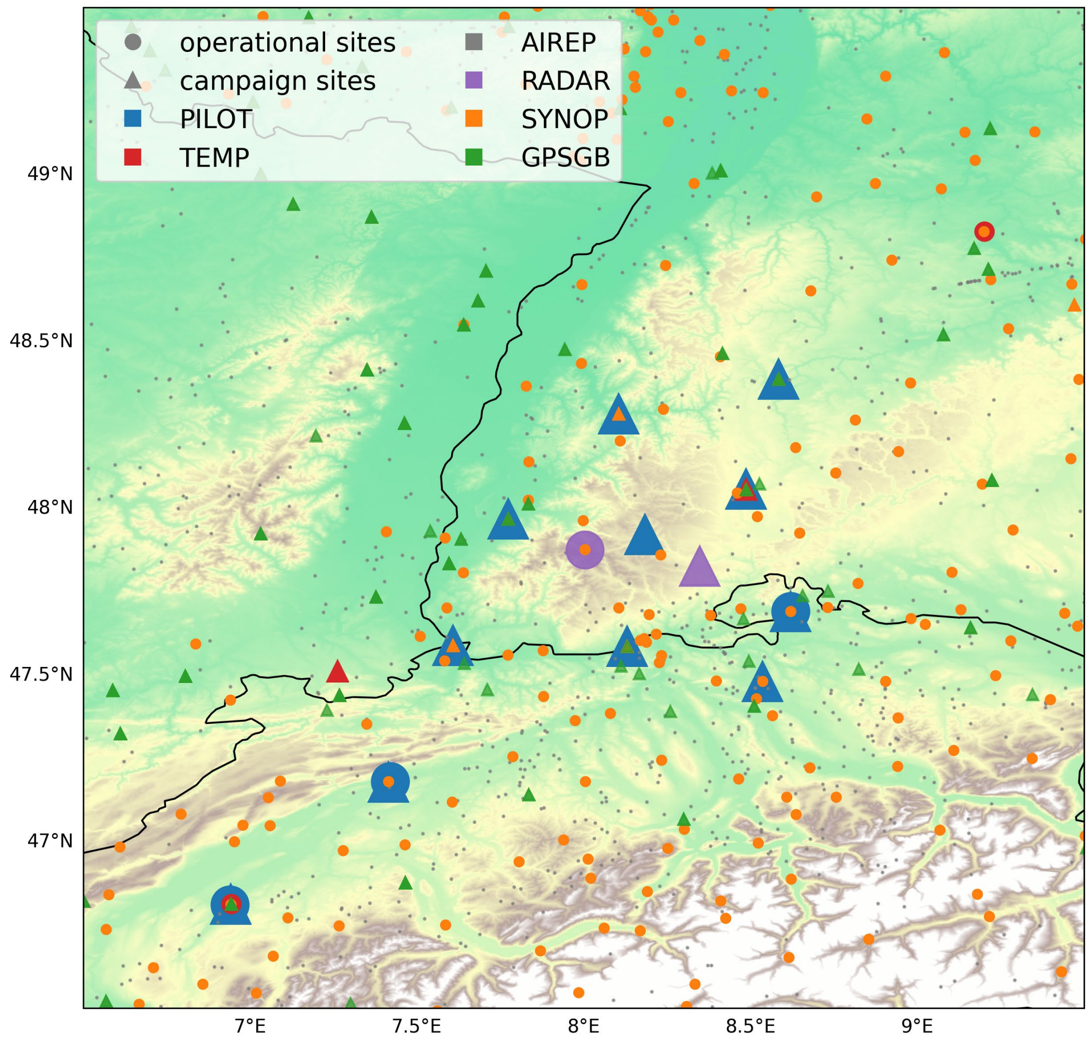
Observation influence in convective-scale data assimilation: Reanalysis and PAI diagnostics for the Swabian MOSES 2023 campaign in southwestern Germany

M. Borne, J.Thomas, H. Reich, J. Keller, and P. Knippertz, A. Oertel

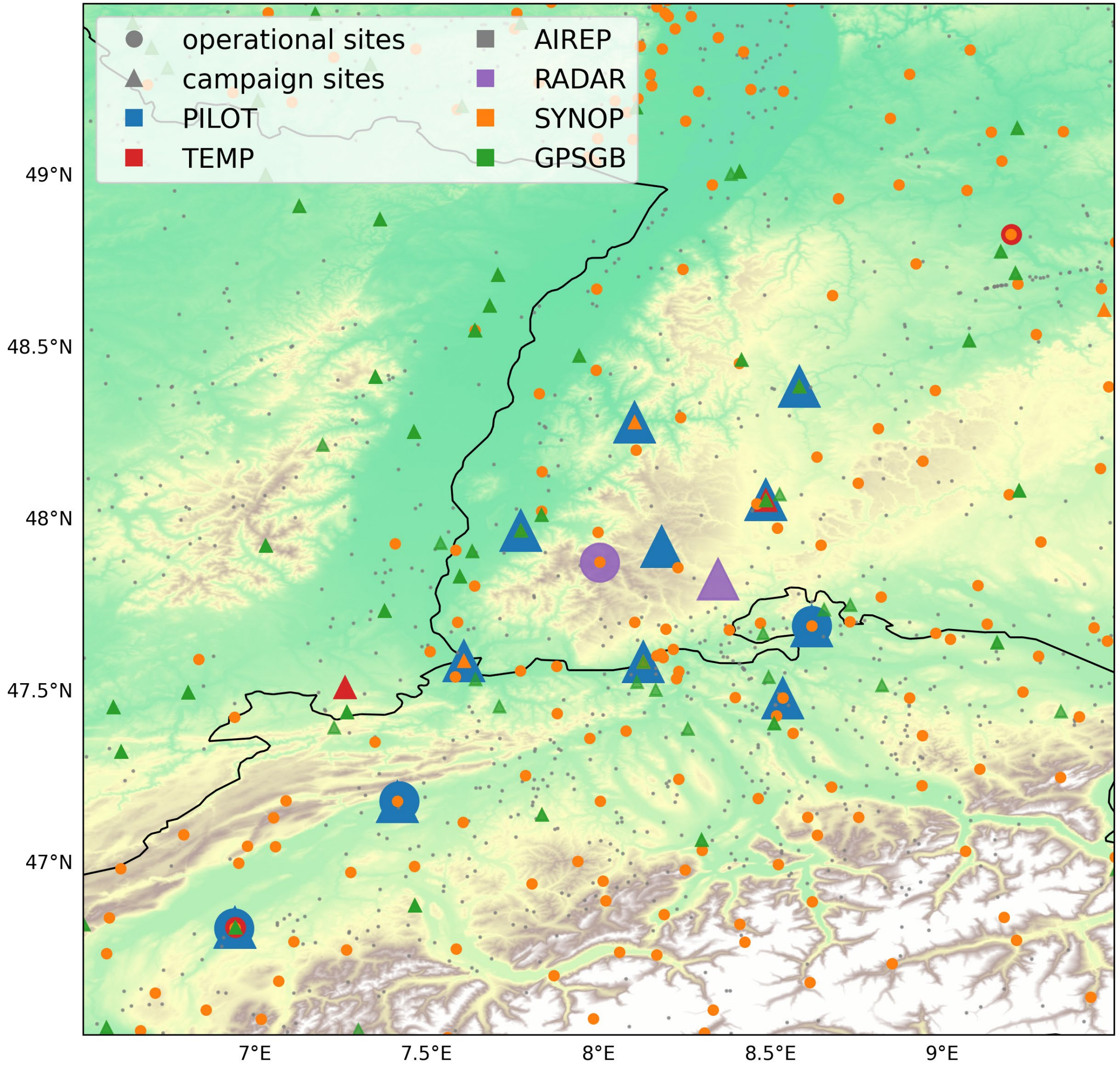
'Mesoscale Processes and Predictability', KIT

With contributions from DWD and MeteoSwiss



Swabian MOSES 2023 field campaign

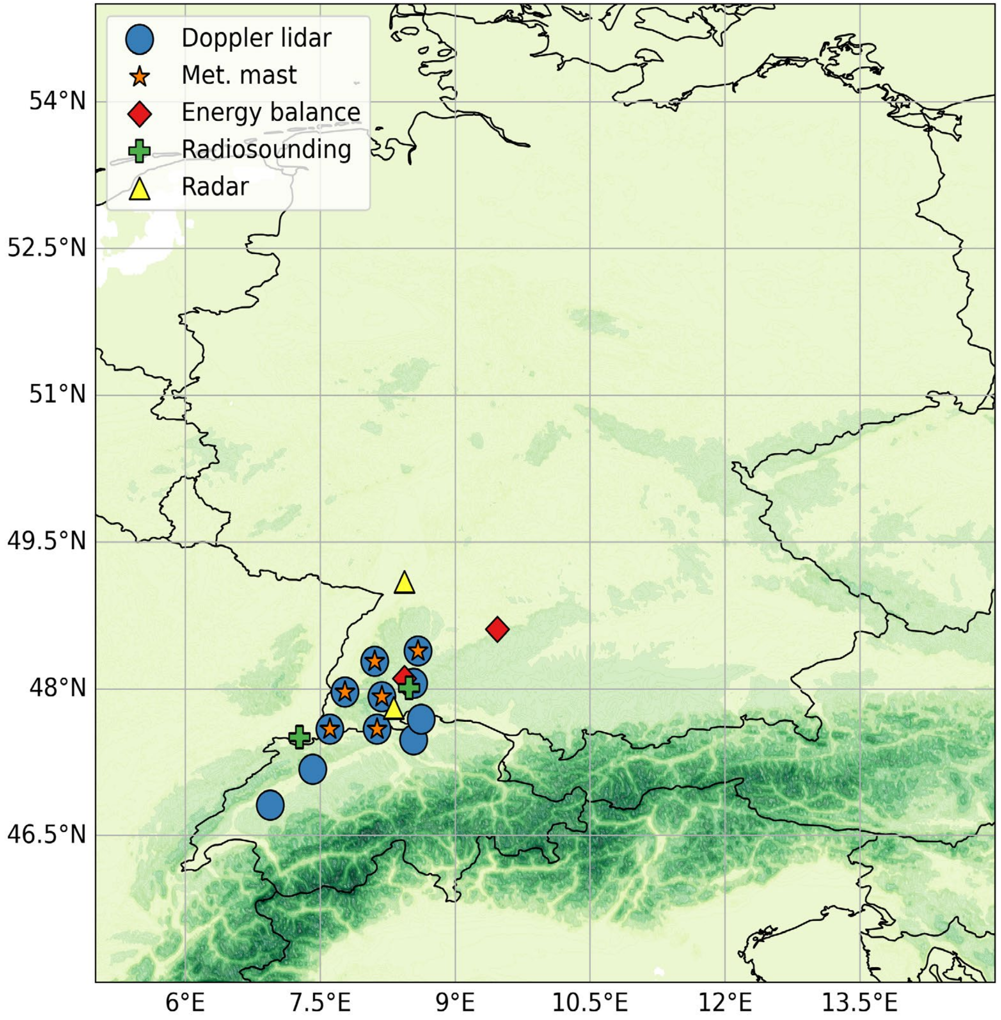
- Network of 12 Doppler wind lidars (DWLs) Handwerker et al., 2025
- June to August 2023
- Black Forest, southwestern Germany, moderately complex terrain
- Deployed network
  - 12 DWLs
  - 2 radiosounding sites
  - 1 X-band radar
  - 6 met. Masts
  - GNSS



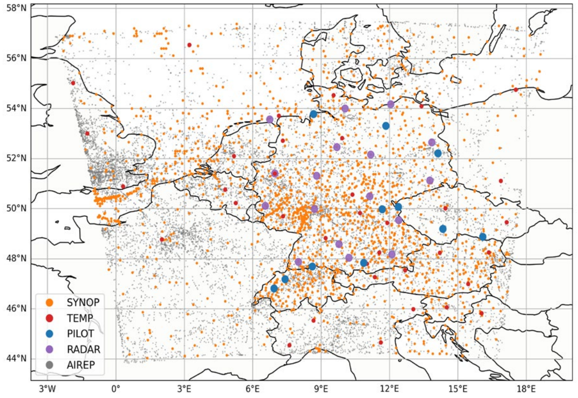
Data assimilation experiments

- ICON coupled to KENDA Zängl *et al.* 2015, Schraff *et al.* 2016, Hunt *et al.* 2007
  - Data Assimilation method: LETKF
  - Convective-scale resolution
  - 40 ensemble members
  - Continuous cycling June-August 2023

- ‘Campaign reanalysis’
  - Assimilation of
  - **operational observations**
  - **campaign observations**

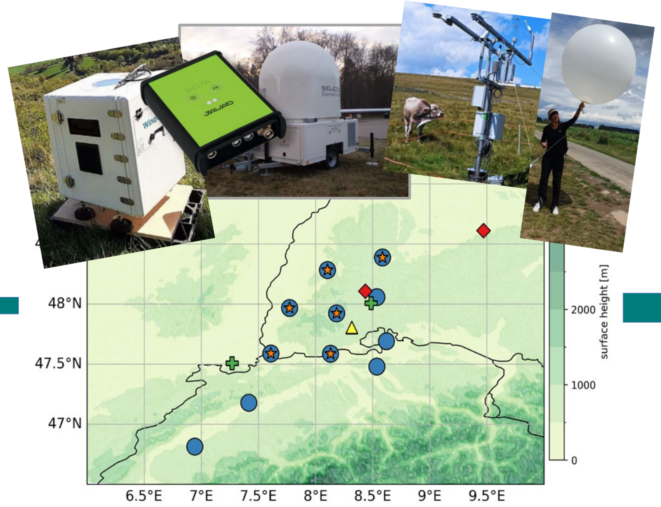


Campaign and Control reanalysis



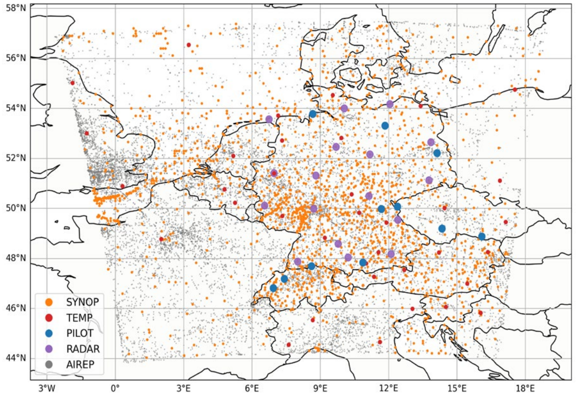
operational observations

+



campaign observations

→ 3-month KENDA cycling → **SWM** campaign reanalysis

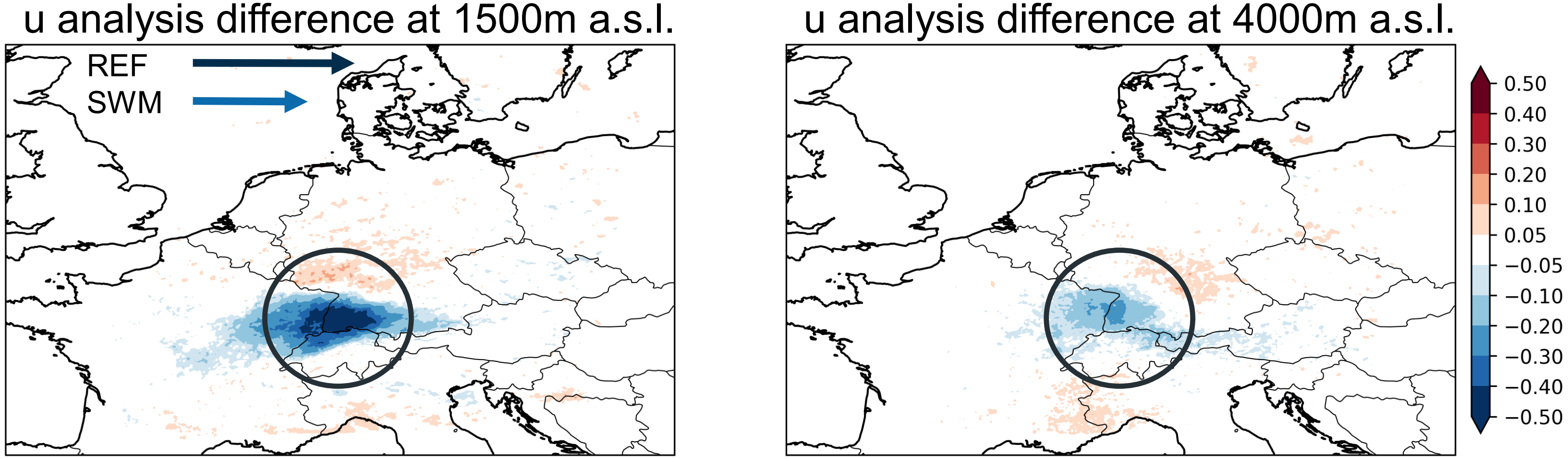


operational observations

→ 3-month KENDA cycling → **REF** control reanalysis

Systematic differences SWM - CTRL analyses

- In westerly flow conditions (85% of timesteps)



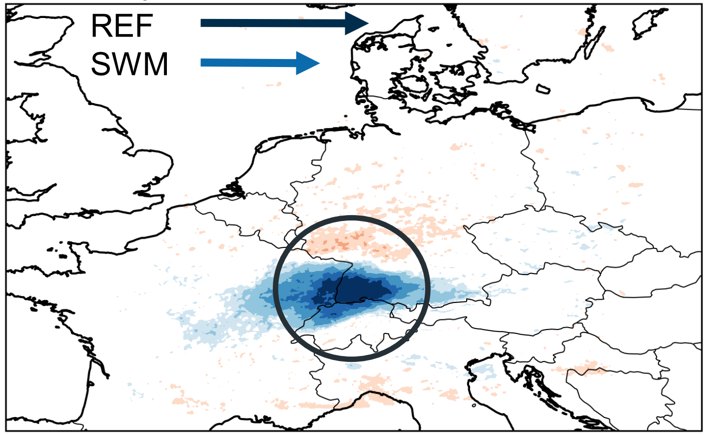
- The **DWL** network slows the zonal wind, most strongly in the lower troposphere.
- The signal spreads east and west, well beyond the campaign area.

Courtesy J. Thomas

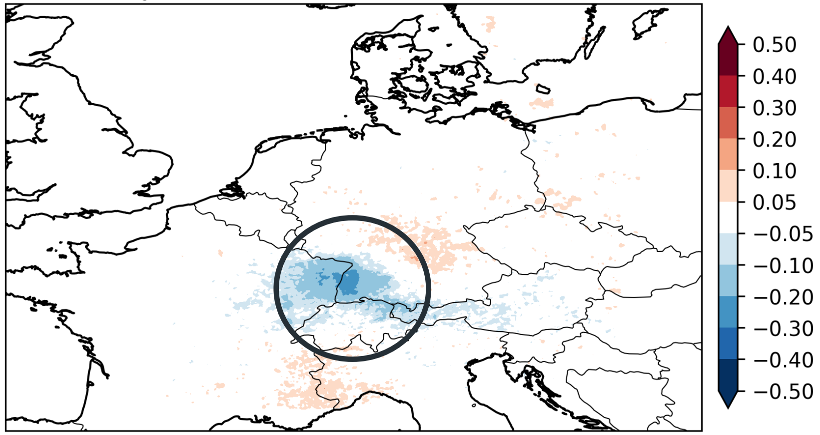
Systematic differences SWM - CTRL analyses

- In westerly flow conditions (85% of timesteps)

u analysis difference at 1500m a.s.l.

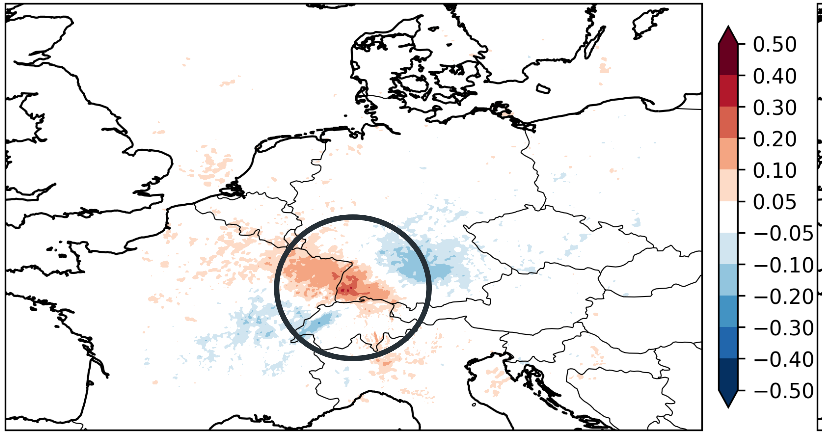


u analysis difference at 4000m a.s.l.

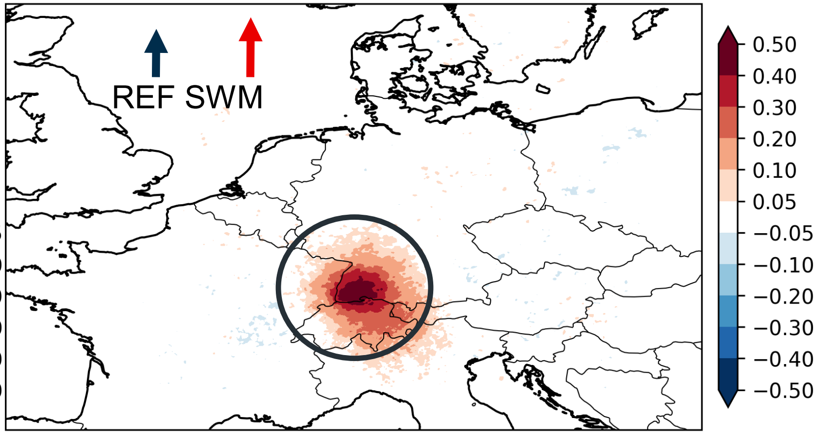


- The **DWL** network slows the zonal wind, most strongly in the lower troposphere.
- The signal spreads east and west, well beyond the campaign area.

v analysis difference at 1500m a.s.l.



v analysis difference at 4000m a.s.l.



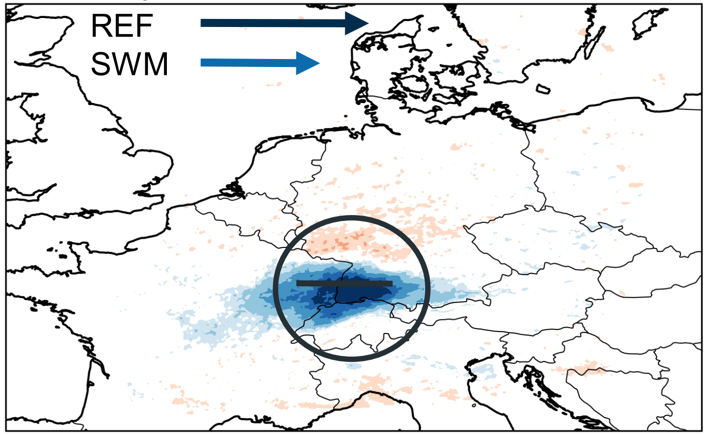
- It strengthens the meridional wind, peaking in the mid-troposphere.
- This signal spreads to the southeast.

Courtesy J. Thomas

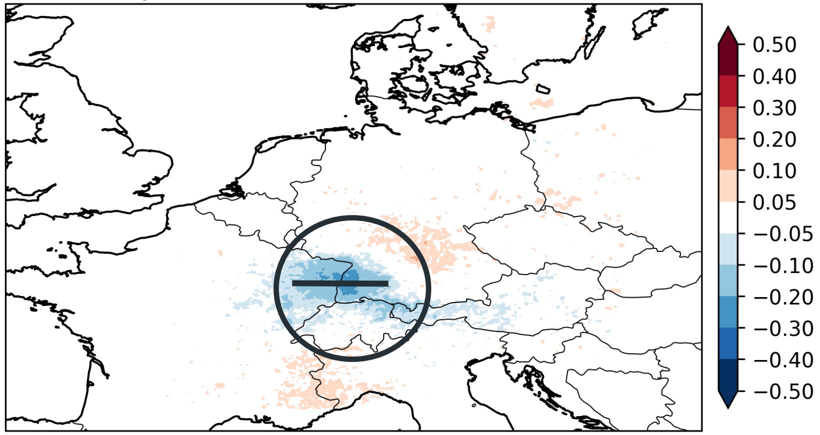
Systematic differences SWM - CTRL analyses

- In westerly flow conditions (85% of timesteps)

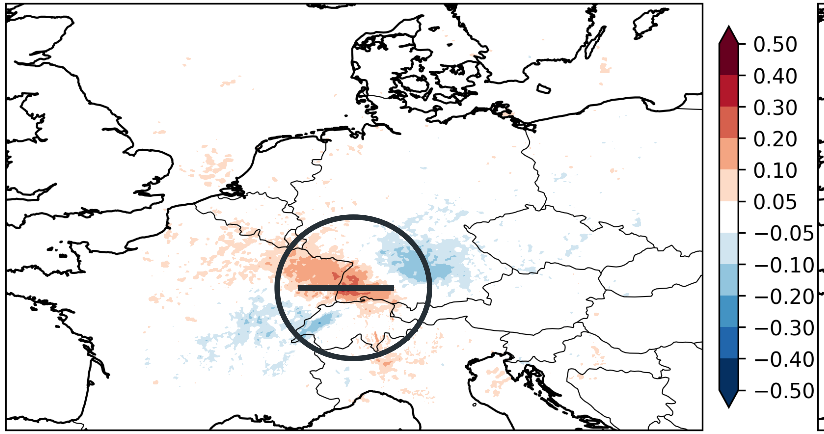
u analysis difference at 1500m a.s.l.



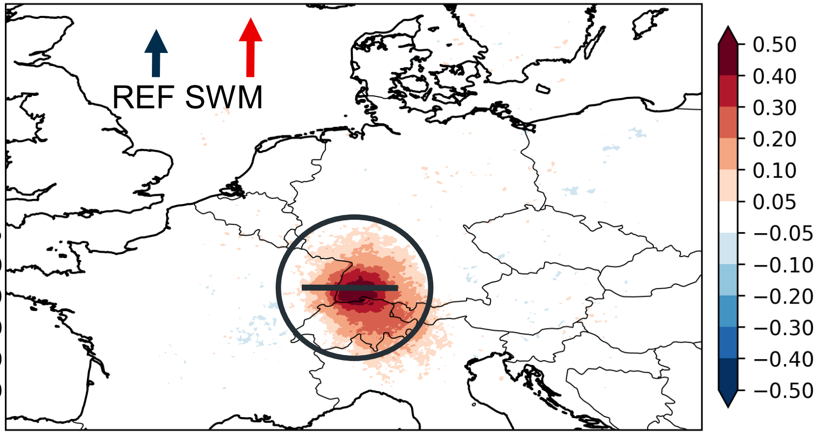
u analysis difference at 4000m a.s.l.

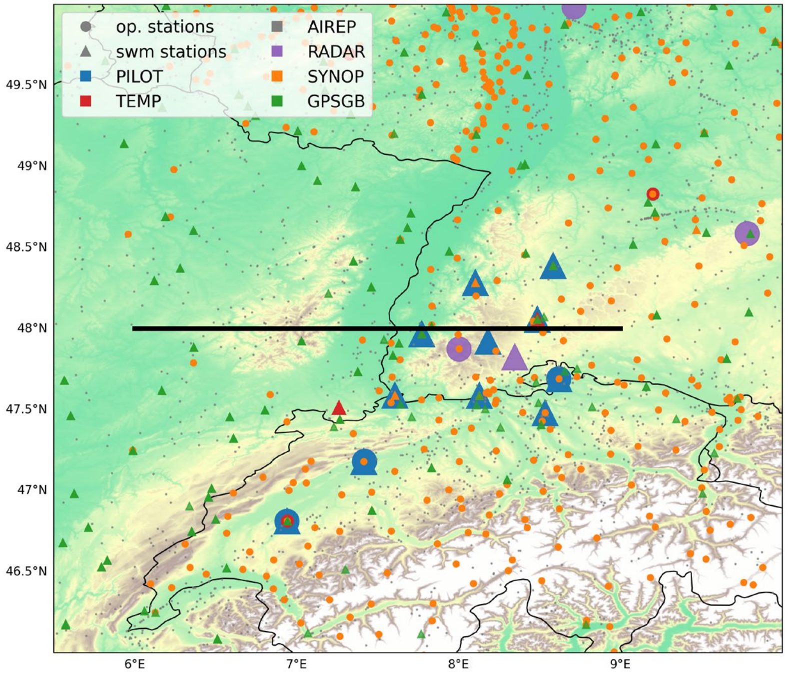


v analysis difference at 1500m a.s.l.



v analysis difference at 4000m a.s.l.

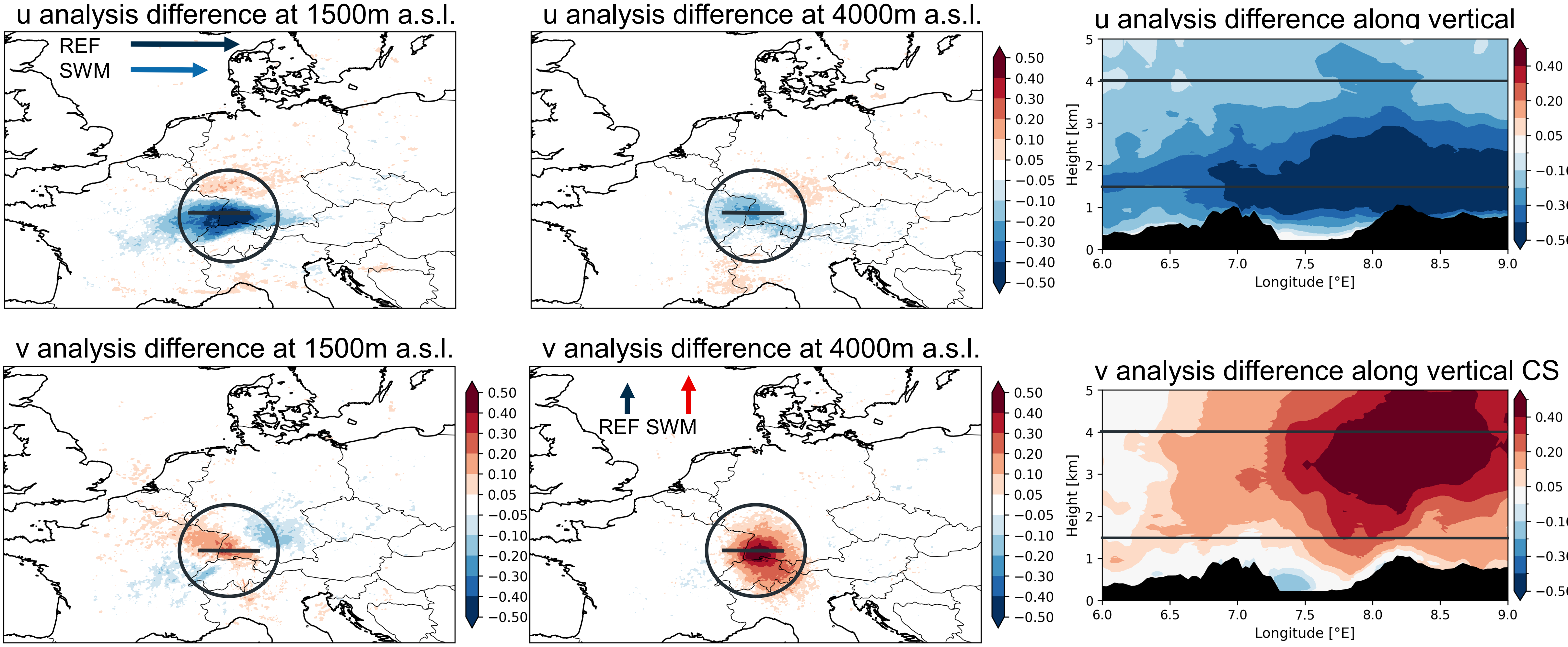




Courtesy J. Thomas

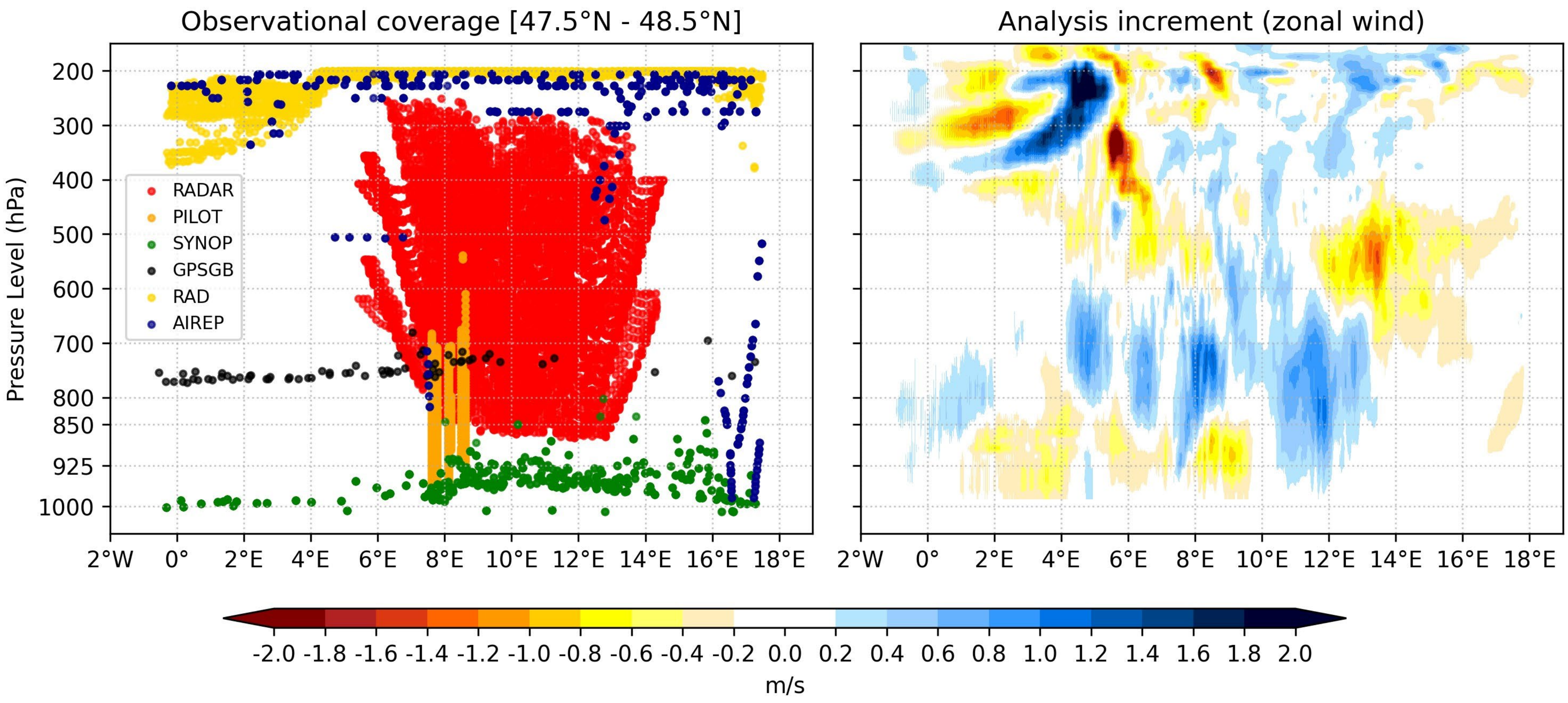
Systematic differences SWM - CTRL analyses

- In westerly flow conditions (85% of timesteps)



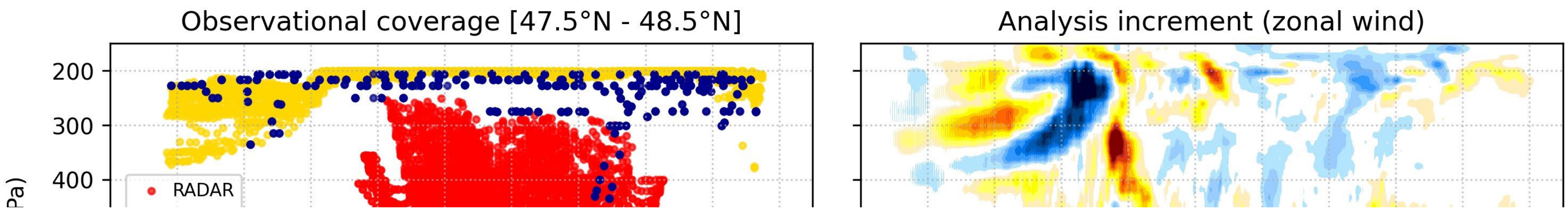
Courtesy J. Thomas

Observational coverage and increment

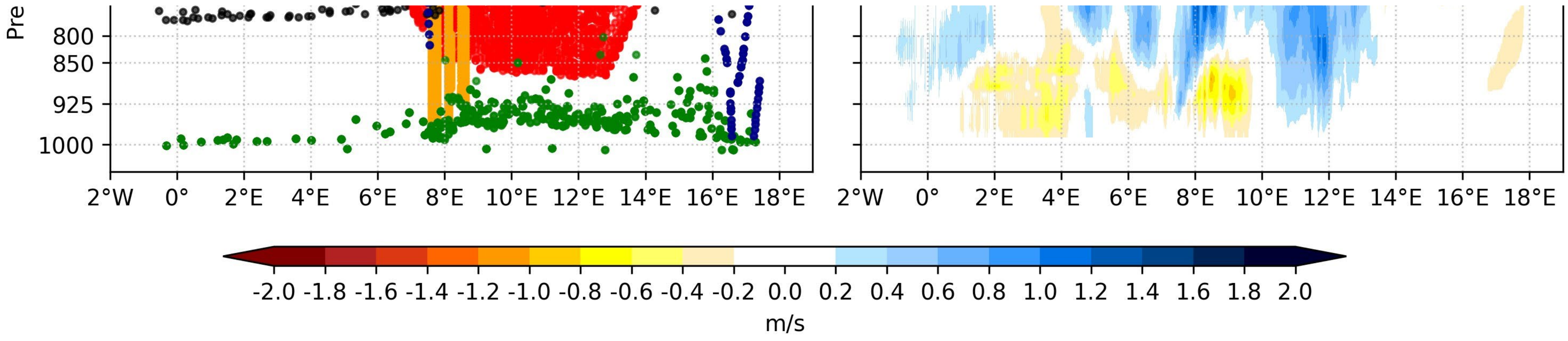


2023-08-01 01 UTC

Observational coverage and increment



What is the contribution of each observation to the total analysis increment?



2023-08-01 01 UTC

Partial analysis increments (PAI): formulation (1/2)

- **Partial analysis increments as diagnostic for LETKF data assimilation systems**, Diefenbach, 2023. *QJRMS.*
- **Investigate the 3D structure of the analysis contribution of one observation in model space**
- Analysis state $x_a = x_b + \boldsymbol{K}(y_o - H(x_b))$
- Analysis Increment $\delta x = x_a - x_b = \boldsymbol{K}(y_o - y_b)$
- Partial Analysis Increment $\delta x_j = K_j \ (y_o - y_b)_j$
- Total Analysis Increment $\sum_{j=1}^{p} \delta x_j = \delta x = \delta x_{\mathrm{PILOT}} + \delta x_{\mathrm{TEMP}} + \delta x_{\mathrm{SYNOP}} + \delta x_{\mathrm{GPSGB}} + \delta x_{\mathrm{RADAR}} + \delta x_{\mathrm{RAD}}$

Partial analysis increments (PAI): formulation (2/2)

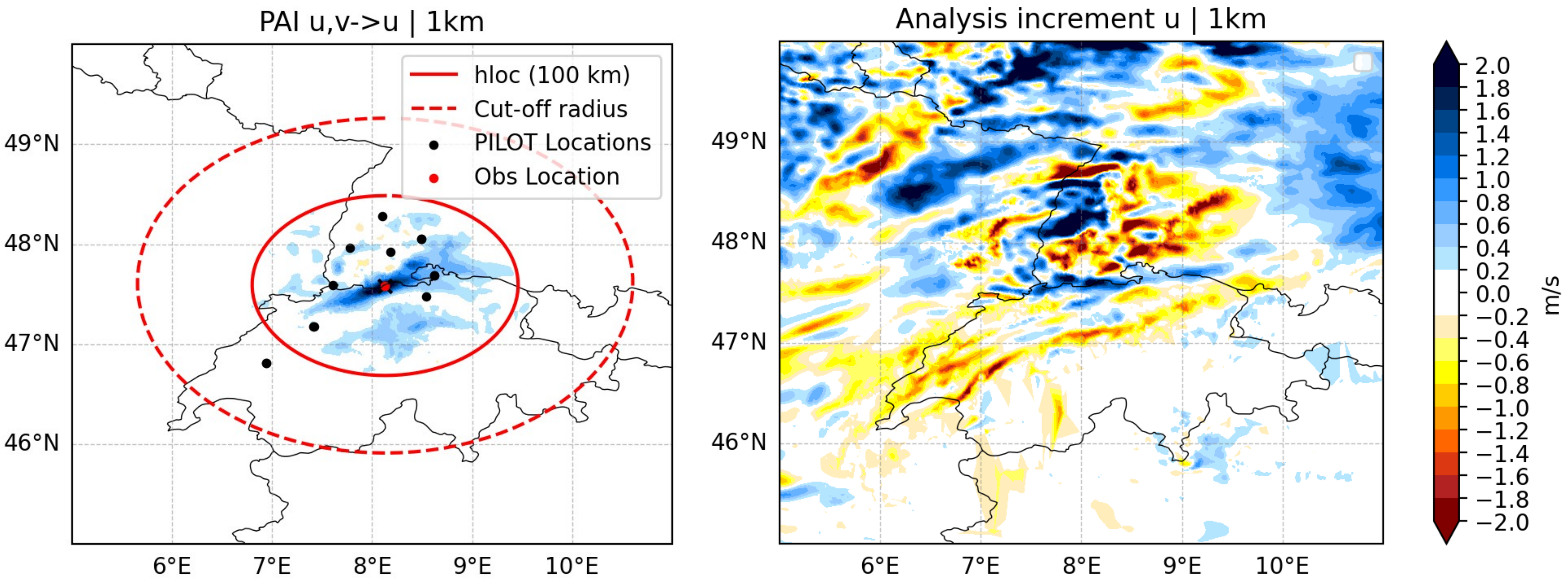
- It is possible to express K in terms of standard LETKF output data products as:

$$\mathbf{K}_{\mathrm{loc}} = \frac{1}{k-1}\,\mathbf{X}_a\,\mathbf{Y}_a^{\mathbf{T}}\,\mathbf{R}^{-1} \circ \boldsymbol{\rho}$$

- Ensemble analysis Perturbation Matrix in *model space* $\mathbf{X}_a$
- Ensemble analysis Perturbation Matrix in *observation space* $\mathbf{Y}_a$
- Observation Error Covariance $\mathbf{R}$ (diagonal)
- Localization matrix $\boldsymbol{\rho} = \boldsymbol{\rho}_{\mathrm{hor}} \circ \boldsymbol{\rho}_{\mathrm{ver}}$

PAI of single PILOT profile - single day (1/4)

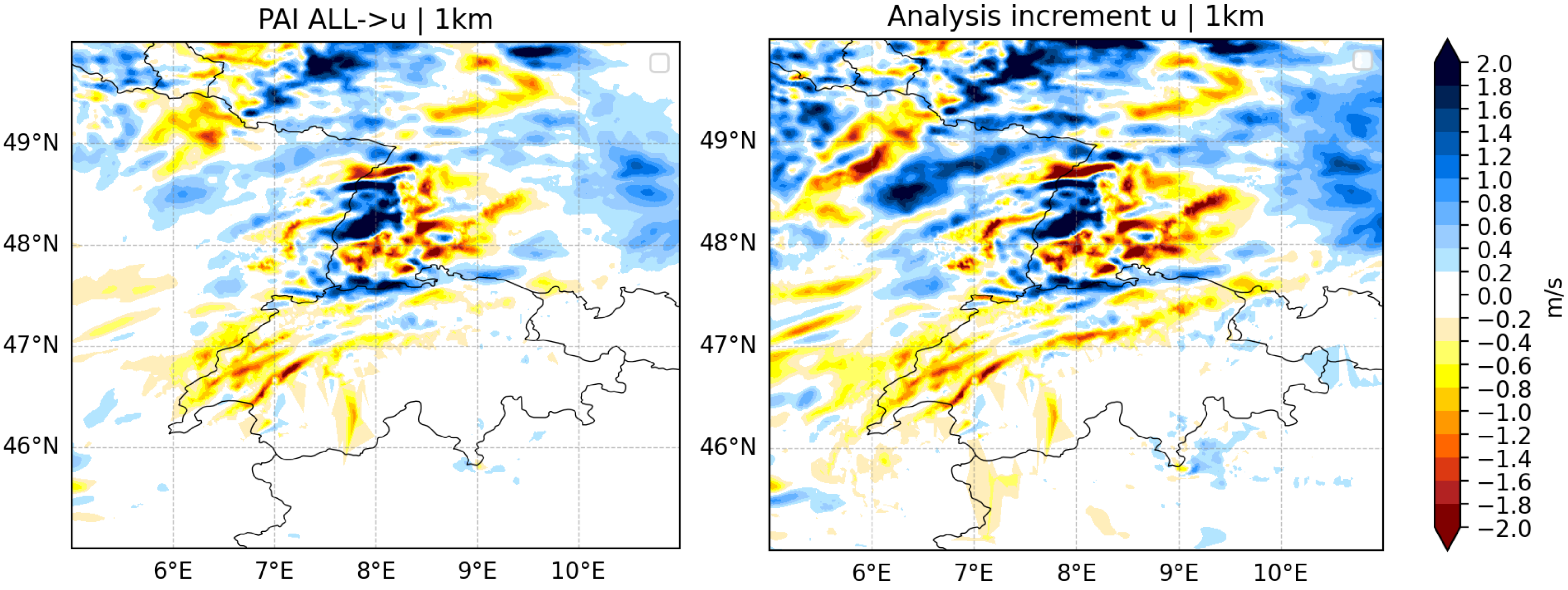
- PILOT at 47.59°N;8.12°E with 64 assimilated observations of u and v winds.



2023-08-01 01 UTC

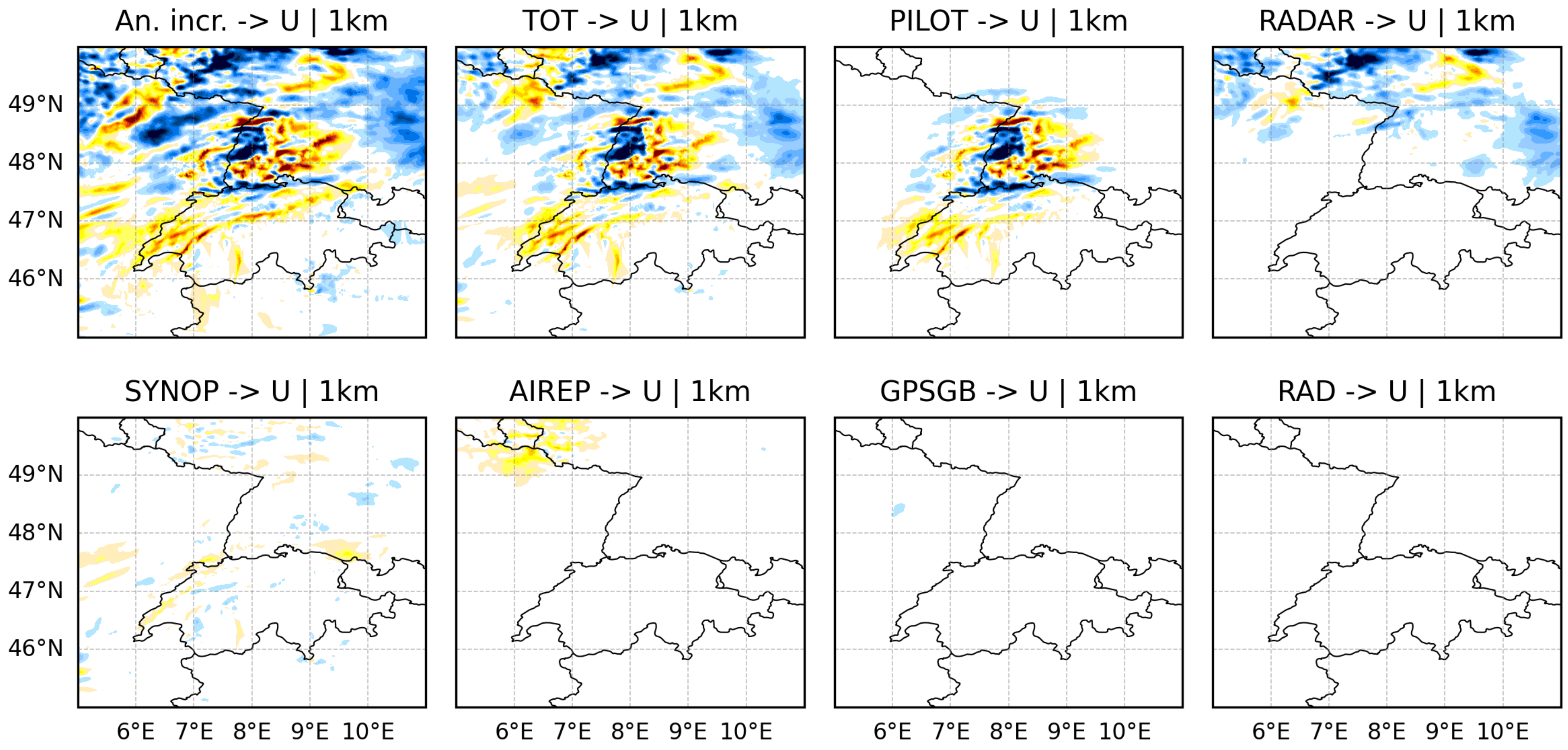
PAI of all observations - single day (2/4)

- ALL: PILOT, SYNOP, RADAR, RAD, AIREP, GPSGB



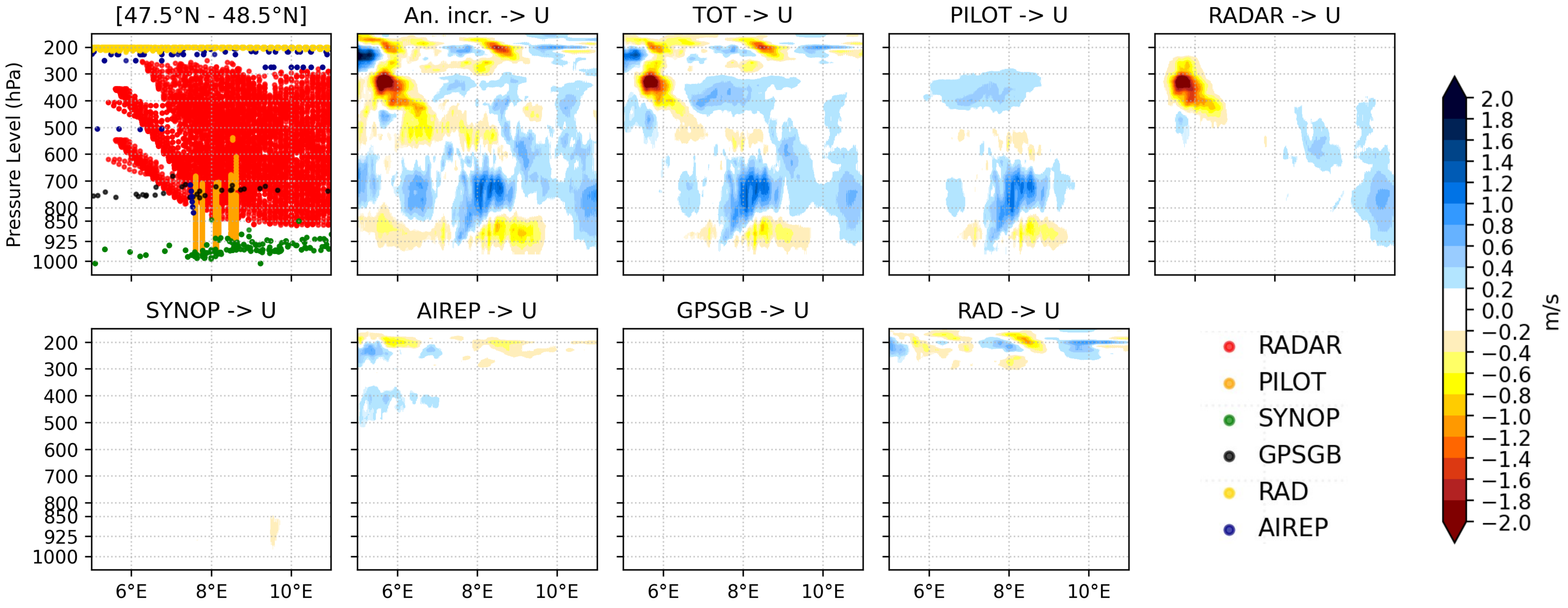
2023-08-01 01 UTC

PAI of all observations - single day (3/4)



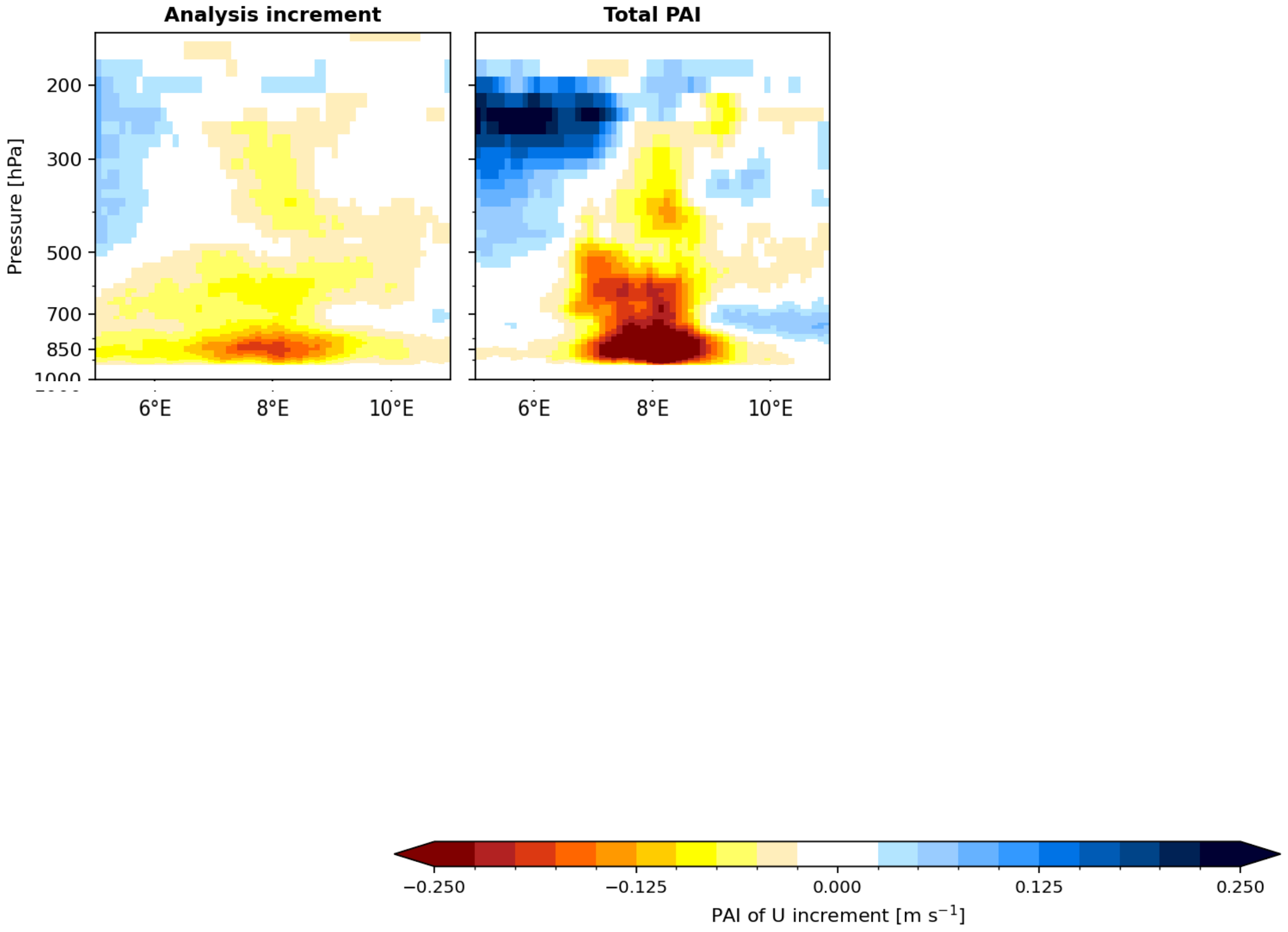
2023-08-01 01 UTC

PAI of all observations - single day (4/4)

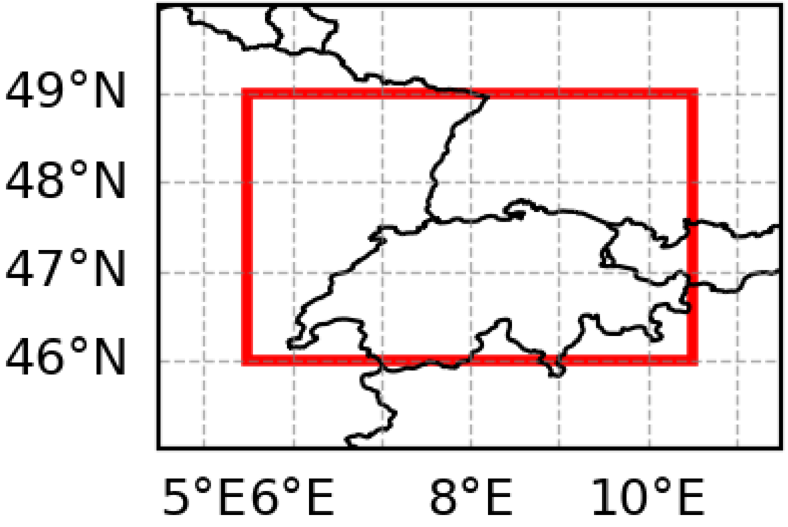


2023-08-01 01 UTC

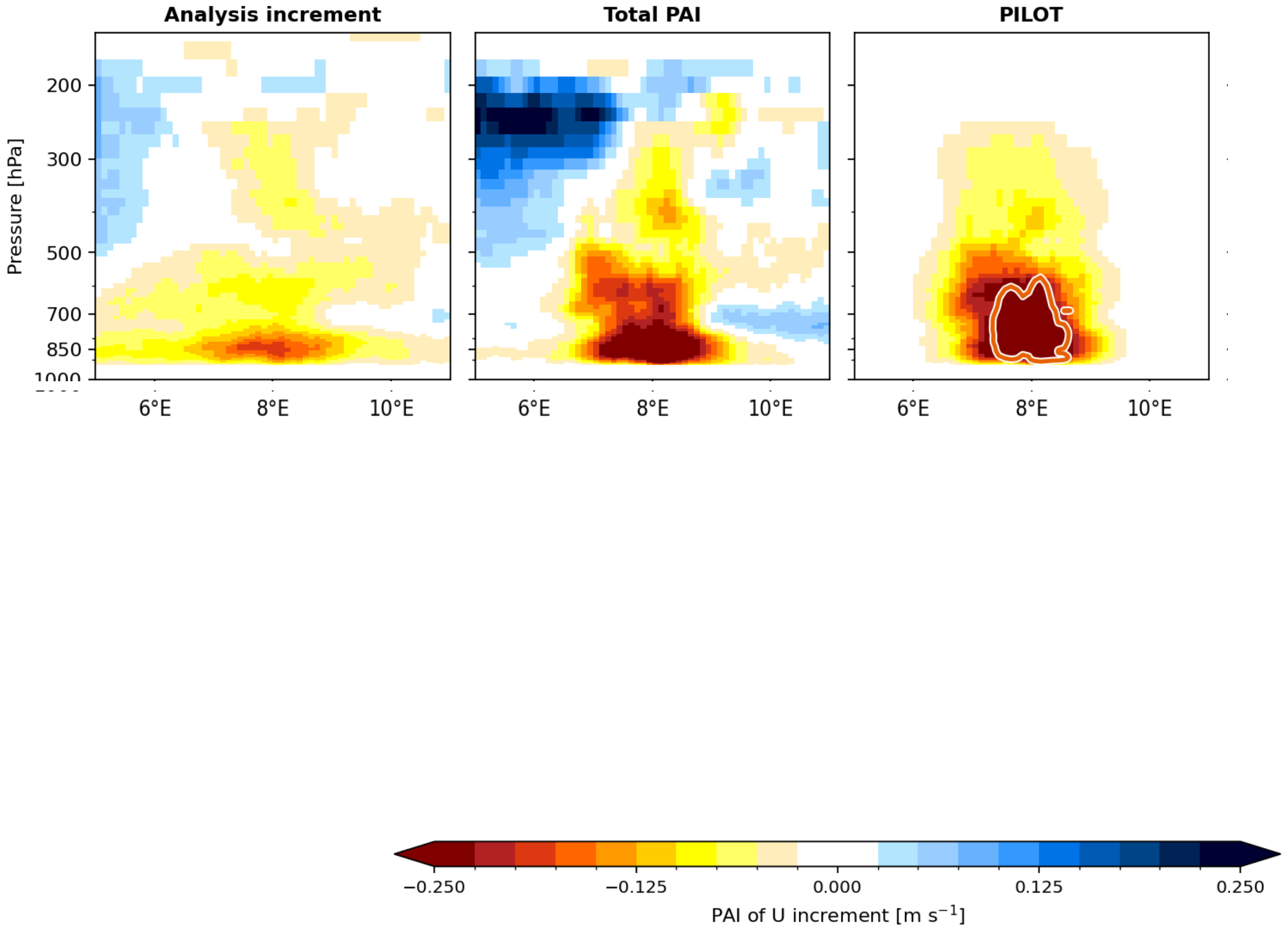
Mean PAI – June to August 2023 – Zonal Wind



- An. Incr. and PAI both fit the sign correctly; the difference is mainly due to **inflation effects** that PAI does not account for.



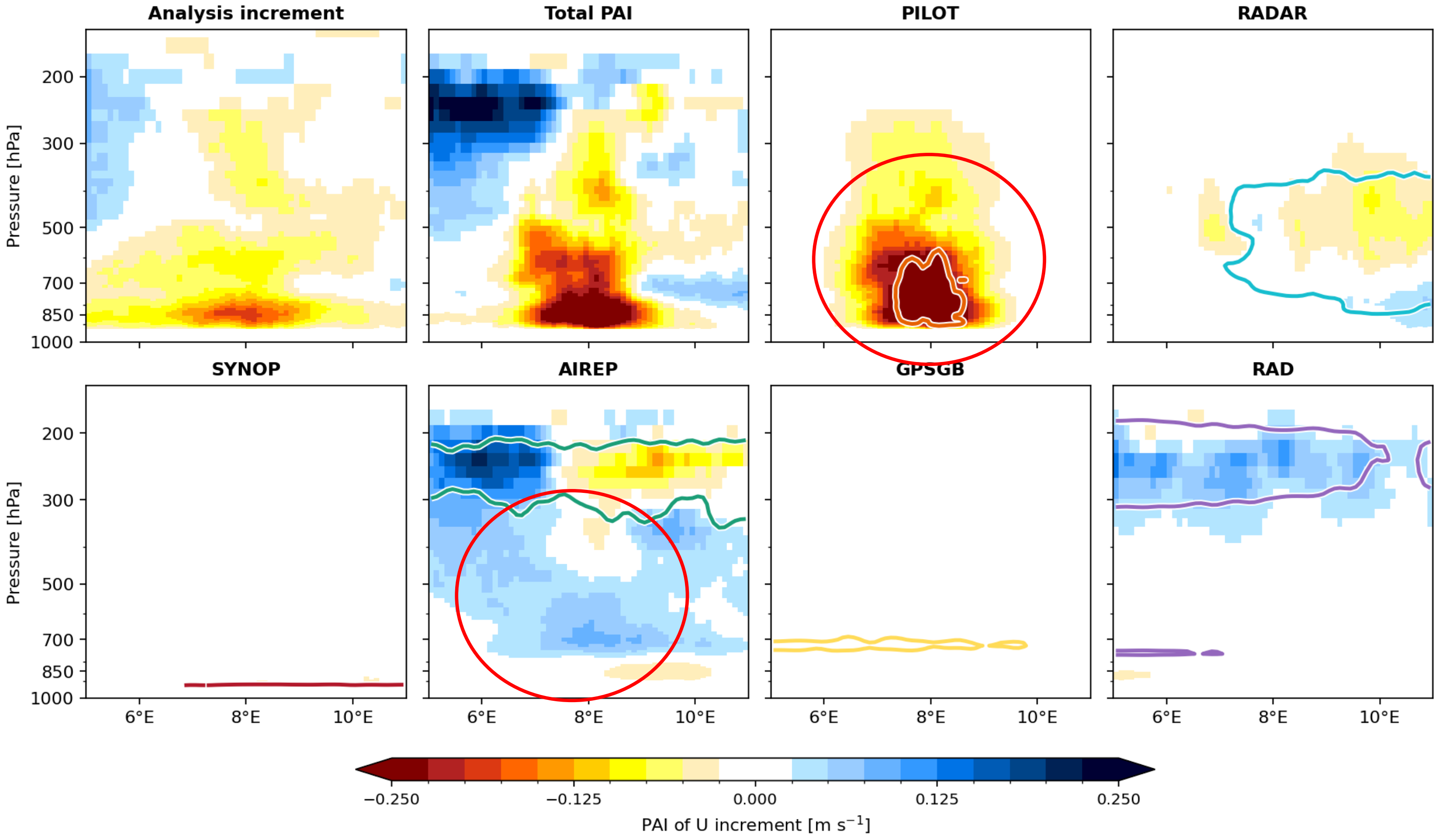
Mean PAI – June to August 2023 – Zonal Wind



U — Zonal wind

- PILOT drives the increment in the troposphere

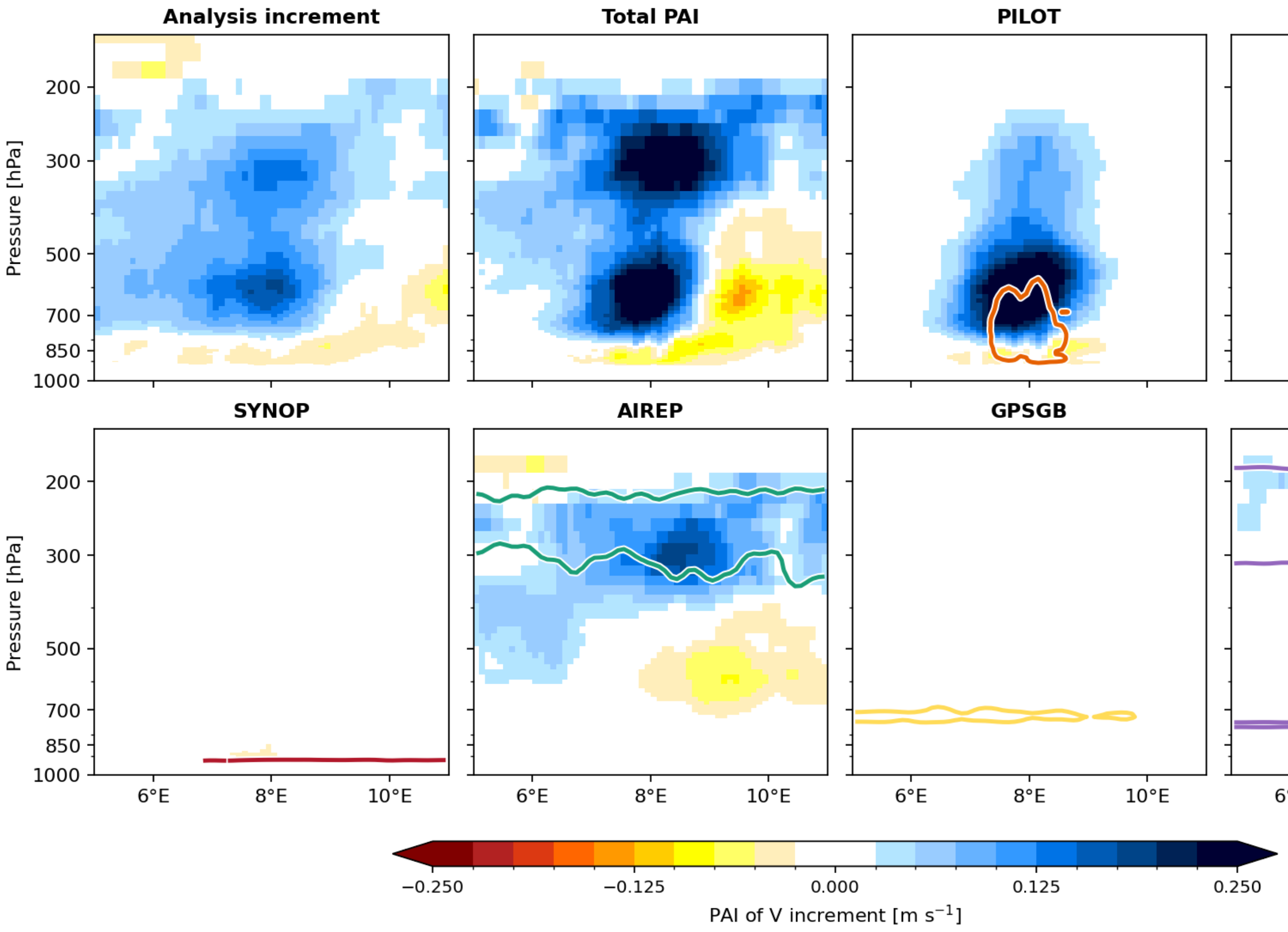
Mean PAI – June to August 2023 – Zonal Wind



U — Zonal wind

- PILOT drives the increment in the troposphere
- AIREP opposes it at 500–700 hPa: vertical localization effect

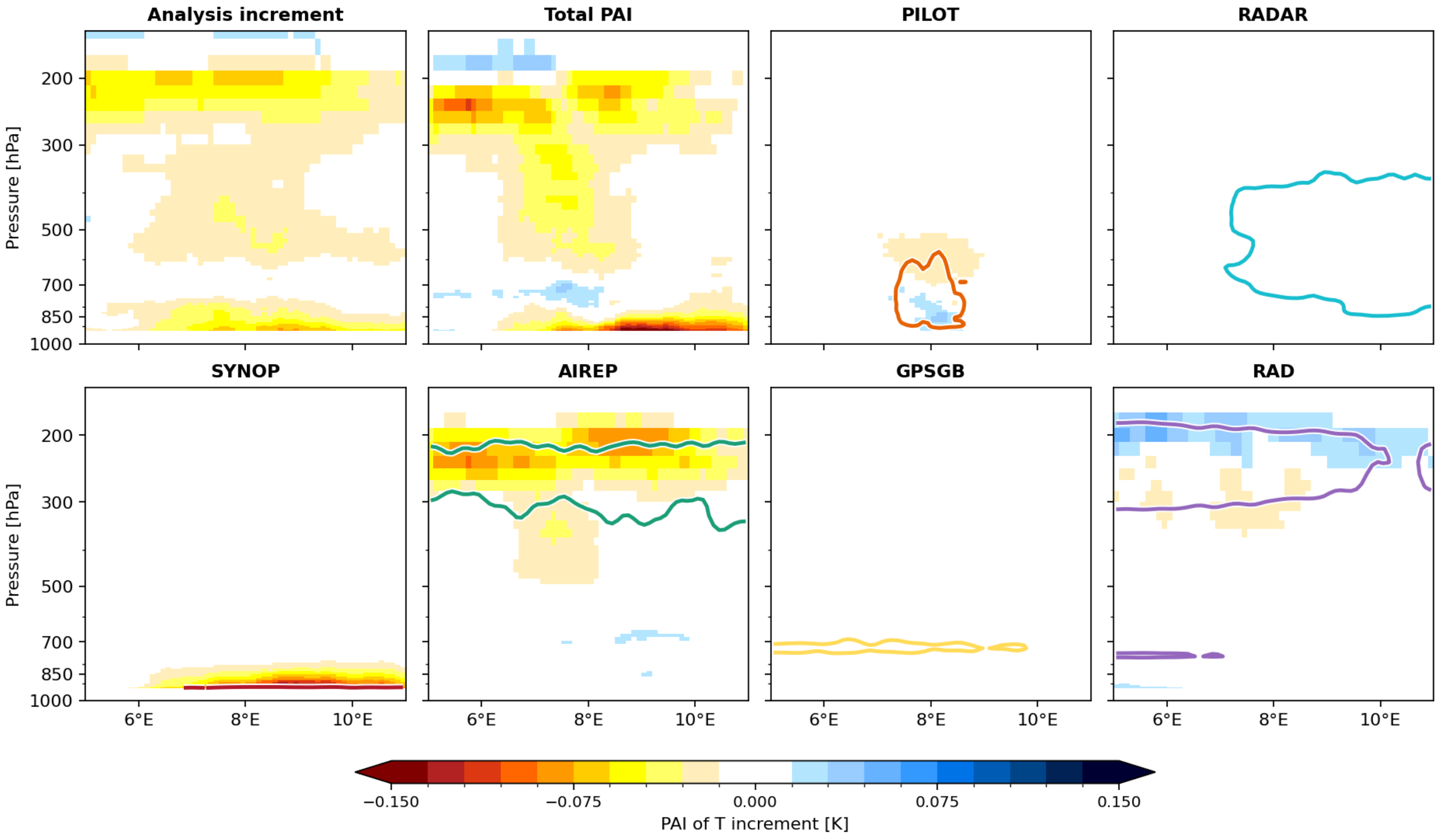
PAI – June to August 2023 – Meridional Wind



V — Meridional wind

- PILOT and AIREP both drive the increment
- Contributions are consistent (same sign)

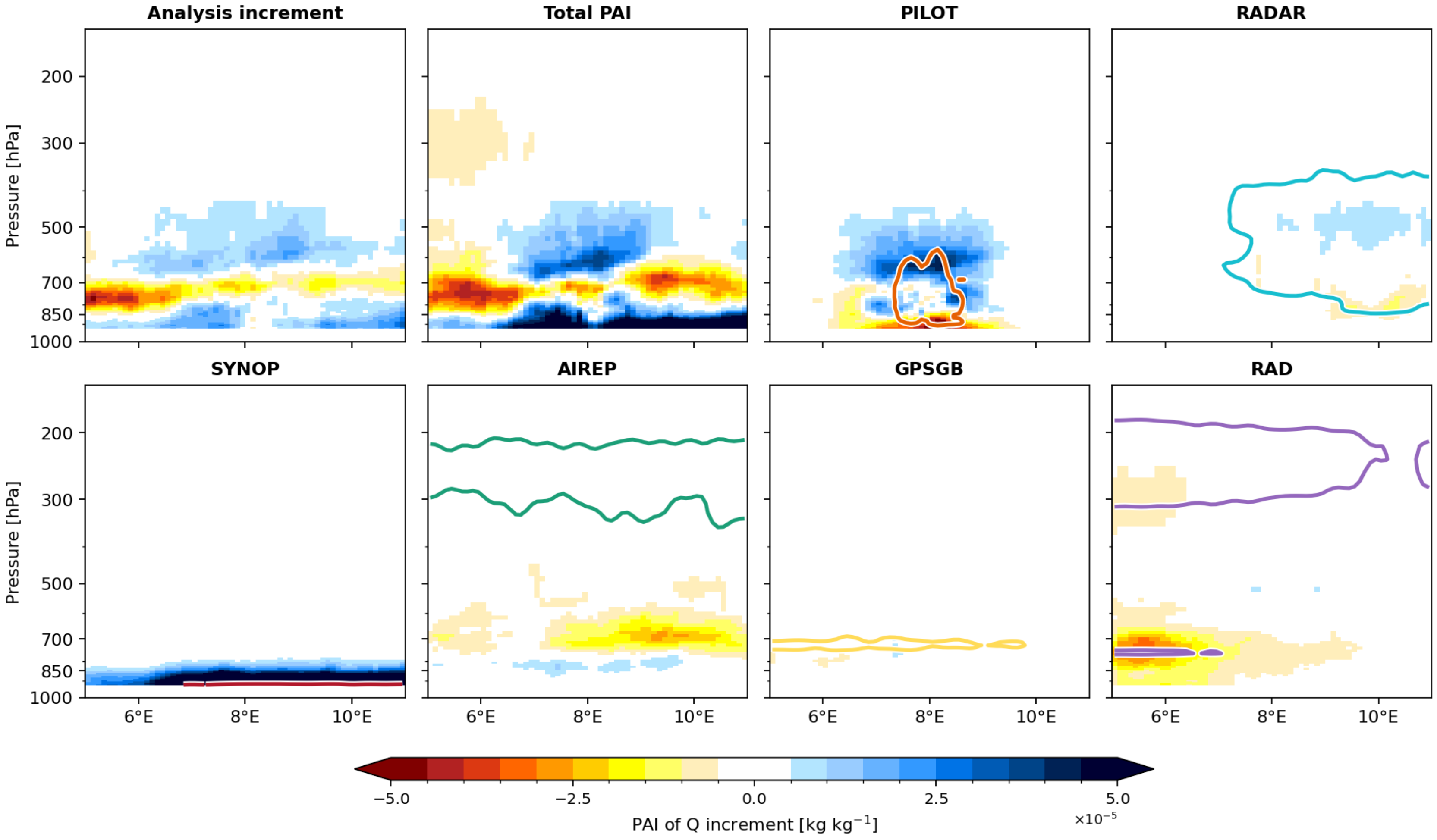
PAI – June to August 2023 – Temperature



T — Temperature

- SYNOP drives the surface; RAD and AIREP the UTLS
- RAD and AIREP oppose each other in the UTLS

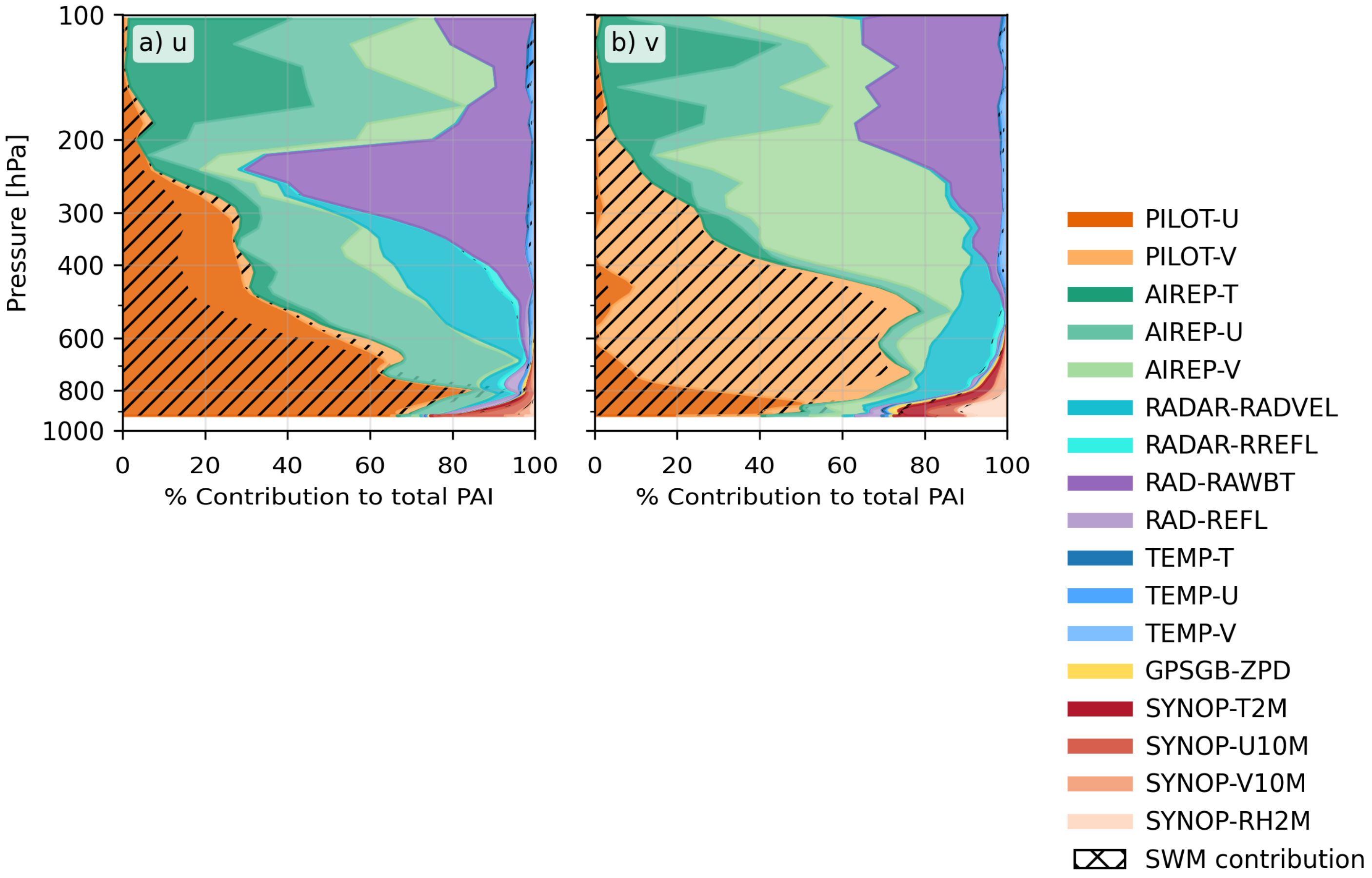
PAI – June to August 2023 – Humidity



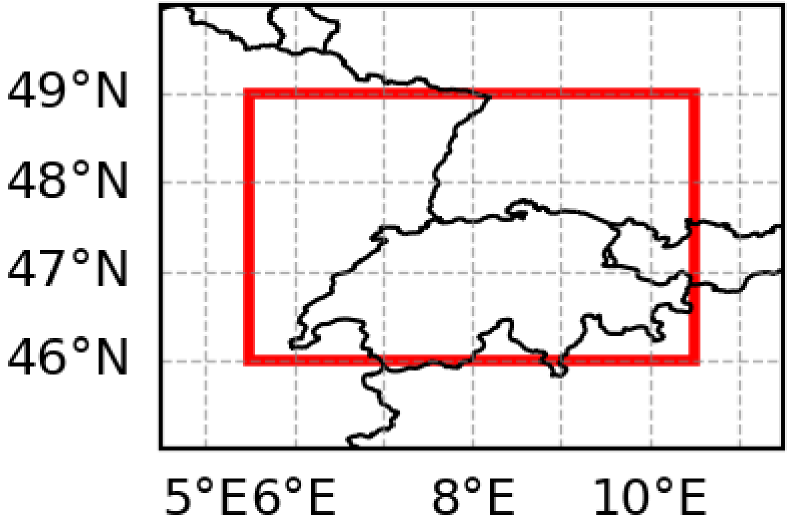
Q — Humidity

- SYNOP drives the surface increment
- RAD and AIREP contribute in the mid-troposphere (same sign)
- PILOT has a strong contribution in the mid-troposphere

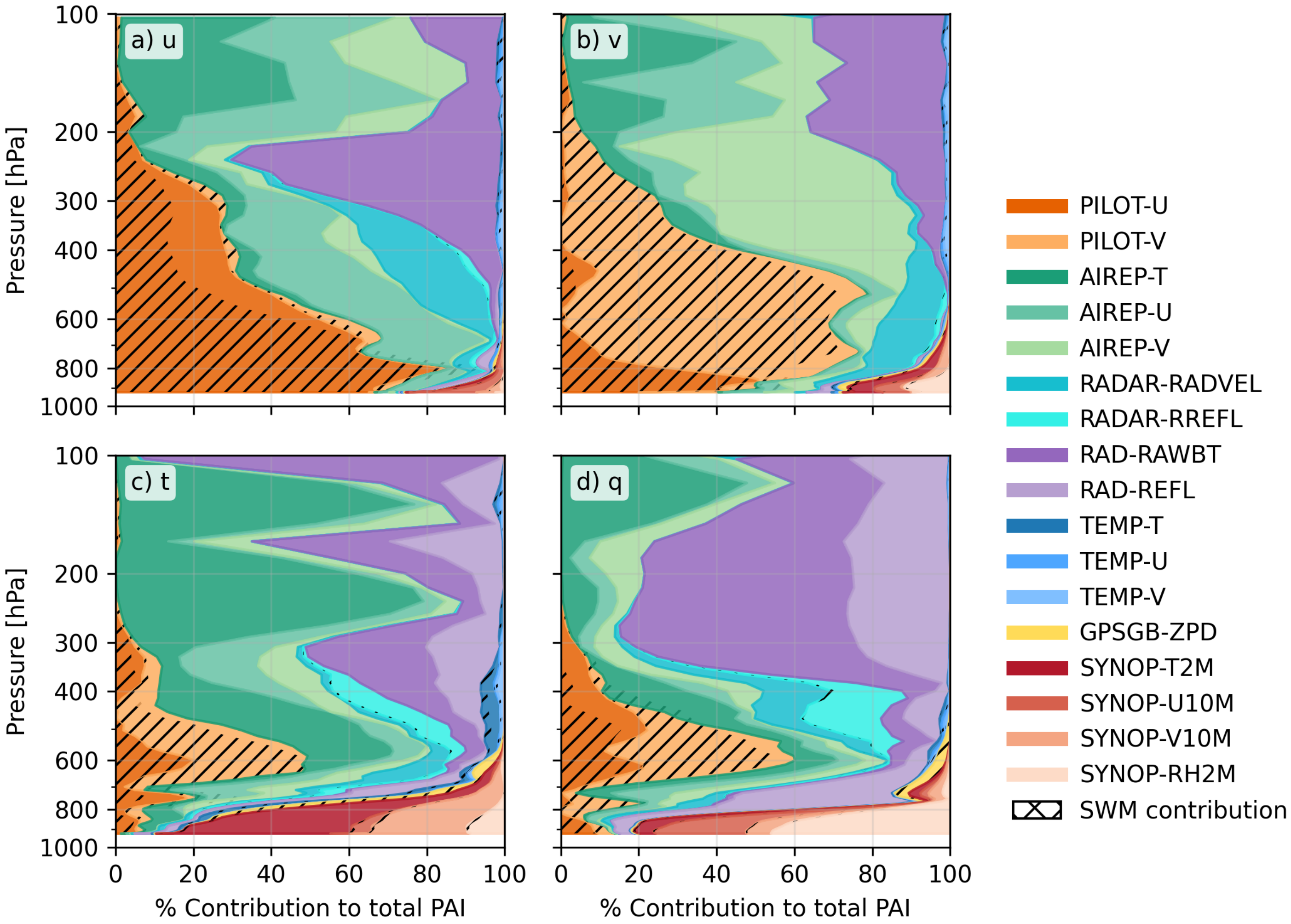
Relative Contributions of each observation type



- **PILOT contributes most to U and V in the lower and middle troposphere**, while the upper troposphere is dominated by **AIREP** and **RAD**.

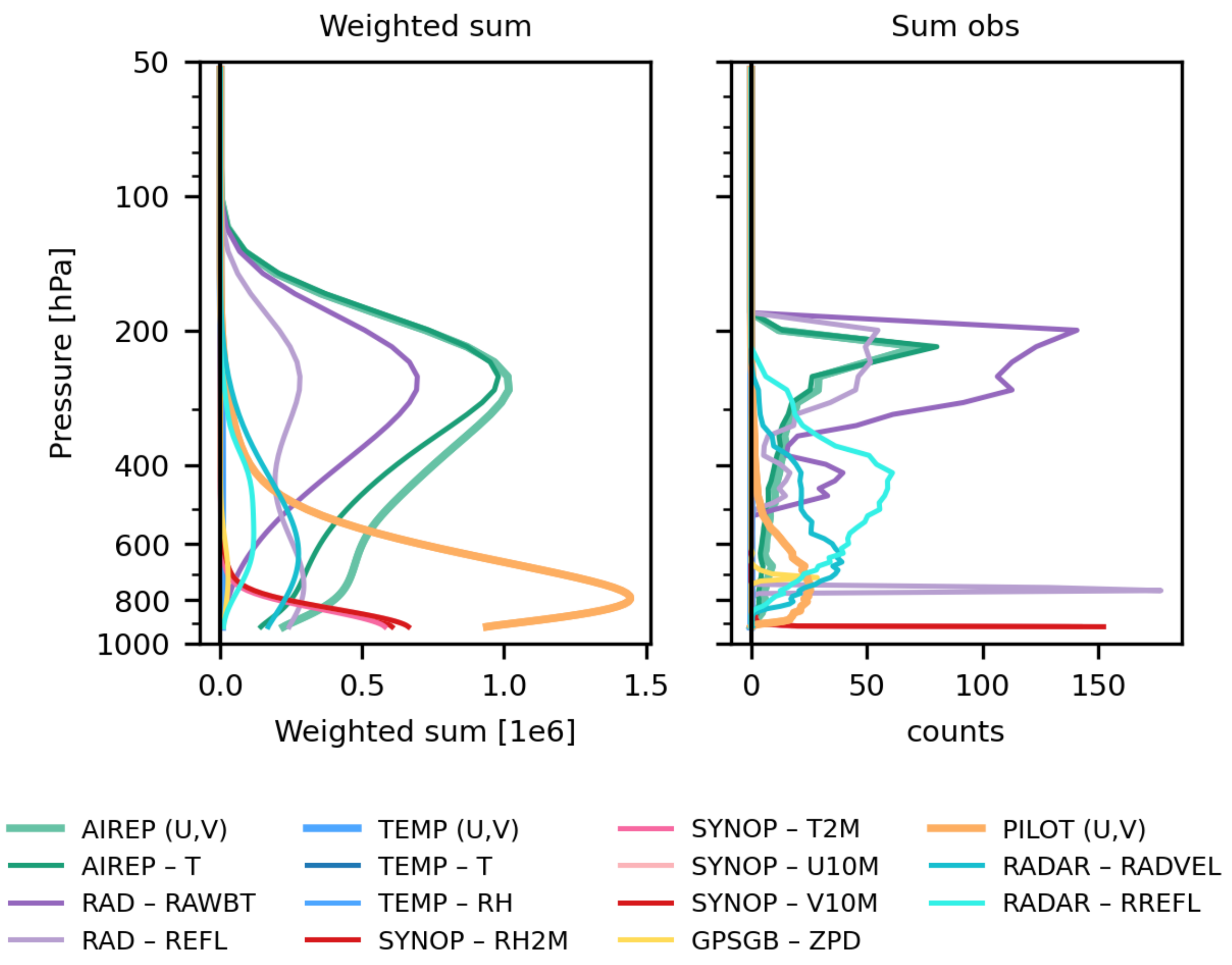


Relative Contributions of each observation type

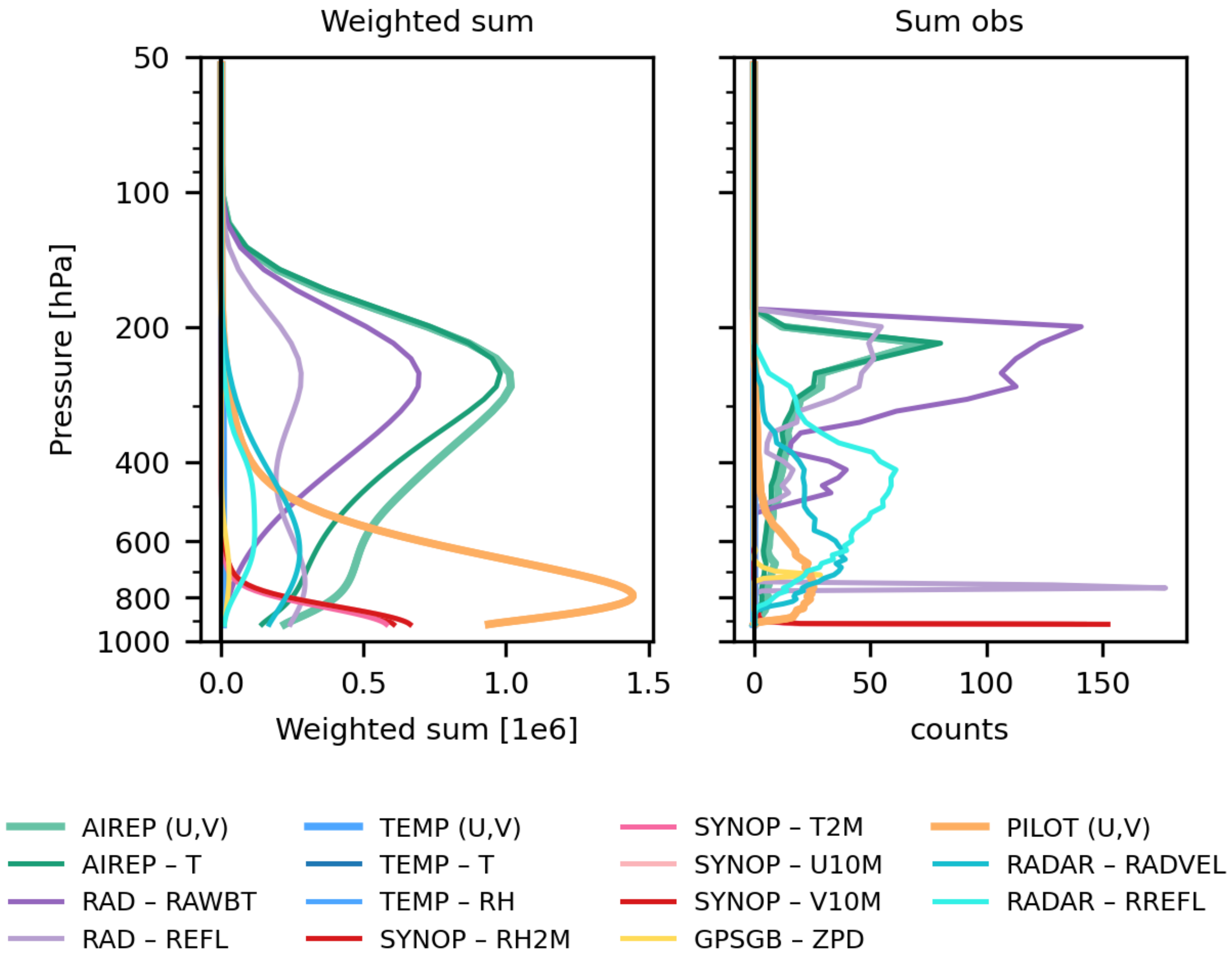


- **PILOT contributes most to U and V in the lower and middle troposphere**, while the upper troposphere is dominated by **AIREP** and **RAD**.
- **PILOT also has a strong contribution to q and T, particularly in the middle troposphere around 600 hPa**.
- **Lower troposphere q and T** is mostly influenced by **SYNOP**, while **RAD** and **AIREP** play a larger role in the **upper troposphere.**

Observation distribution (1/2)



Observation distribution (2/2)

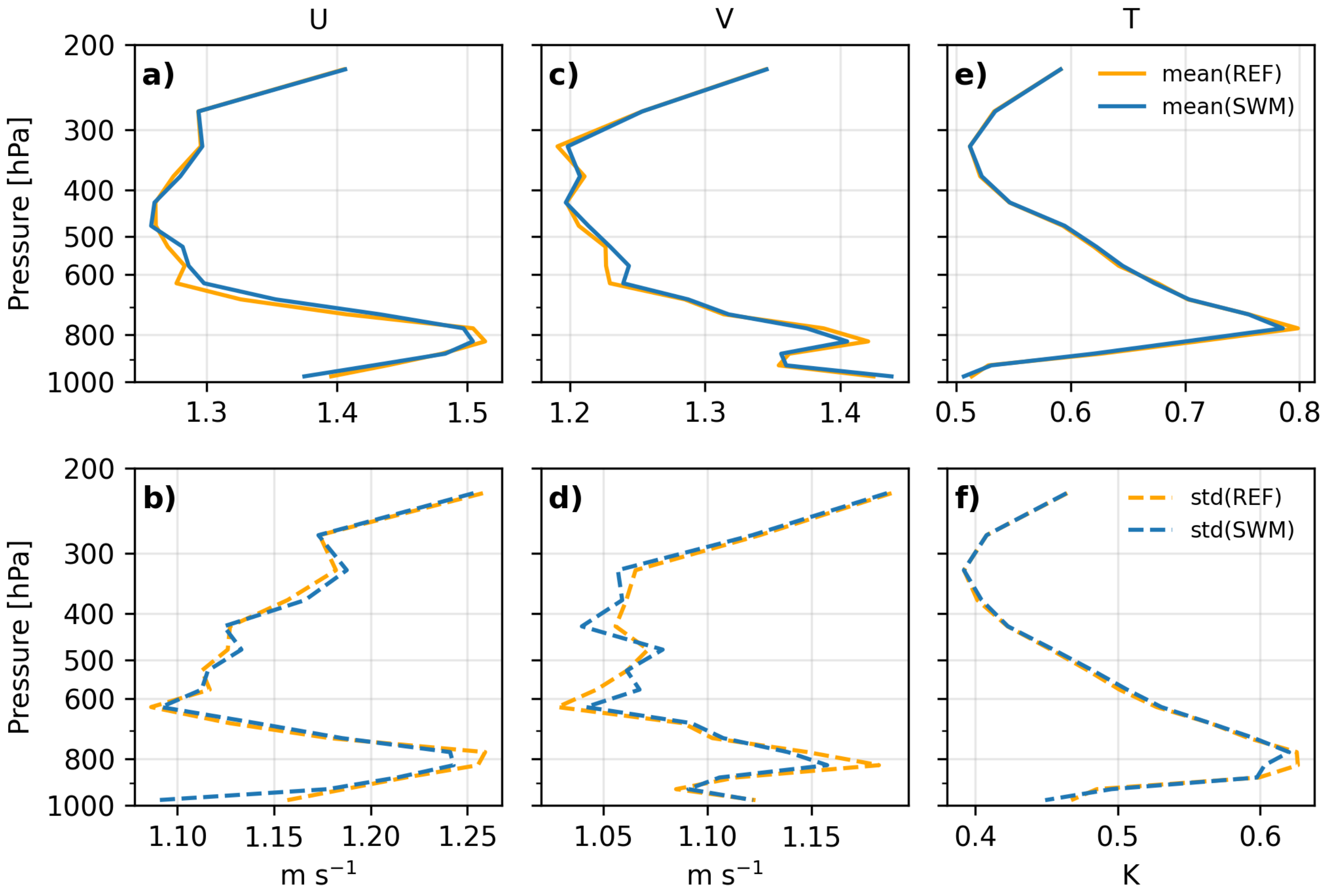


Localization length by observation type

Observation	Horizontal (km)	Vertical (log-p)
PILOT, TEMP, AIREP, SYNOP	50–100*	0.075–0.5
RAD	25	0.3
GPSGB	20	0.17
RADAR	16	0.07–0.3†

** adaptive horizontal localisation*
† RADAR vertical: 0.07 (reflectivity) – 0.3 (radial velocity)
Vertical localisation: Gaspari–Cohn function on log-pressure distance

Departure statistics (O-B) of AIREP observations

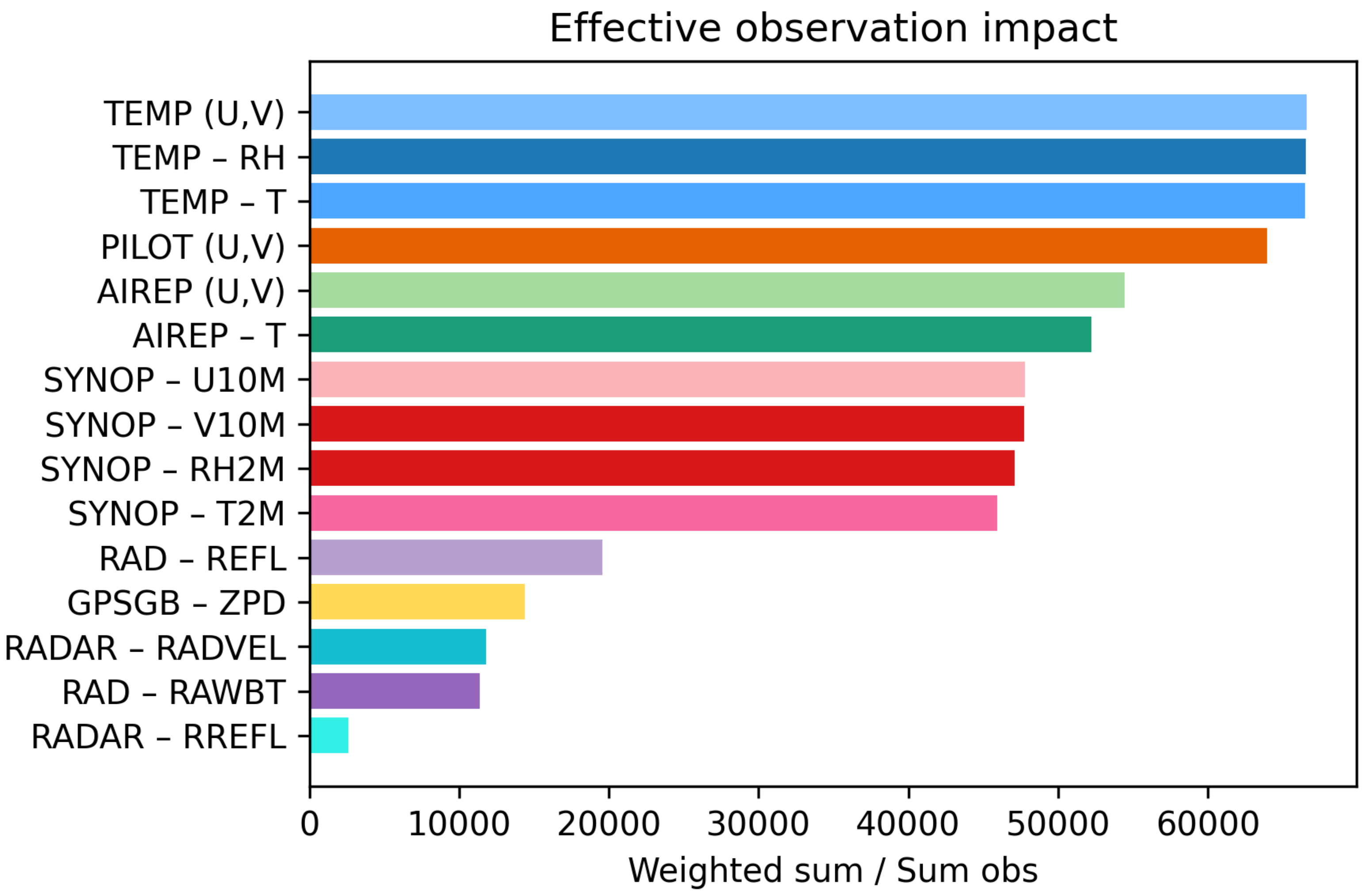


- Departures peak in the lower troposphere (~800 hPa), smallest at mid-levels.
- Departure shows improvement near 800 hPa, a first sign of positive impact

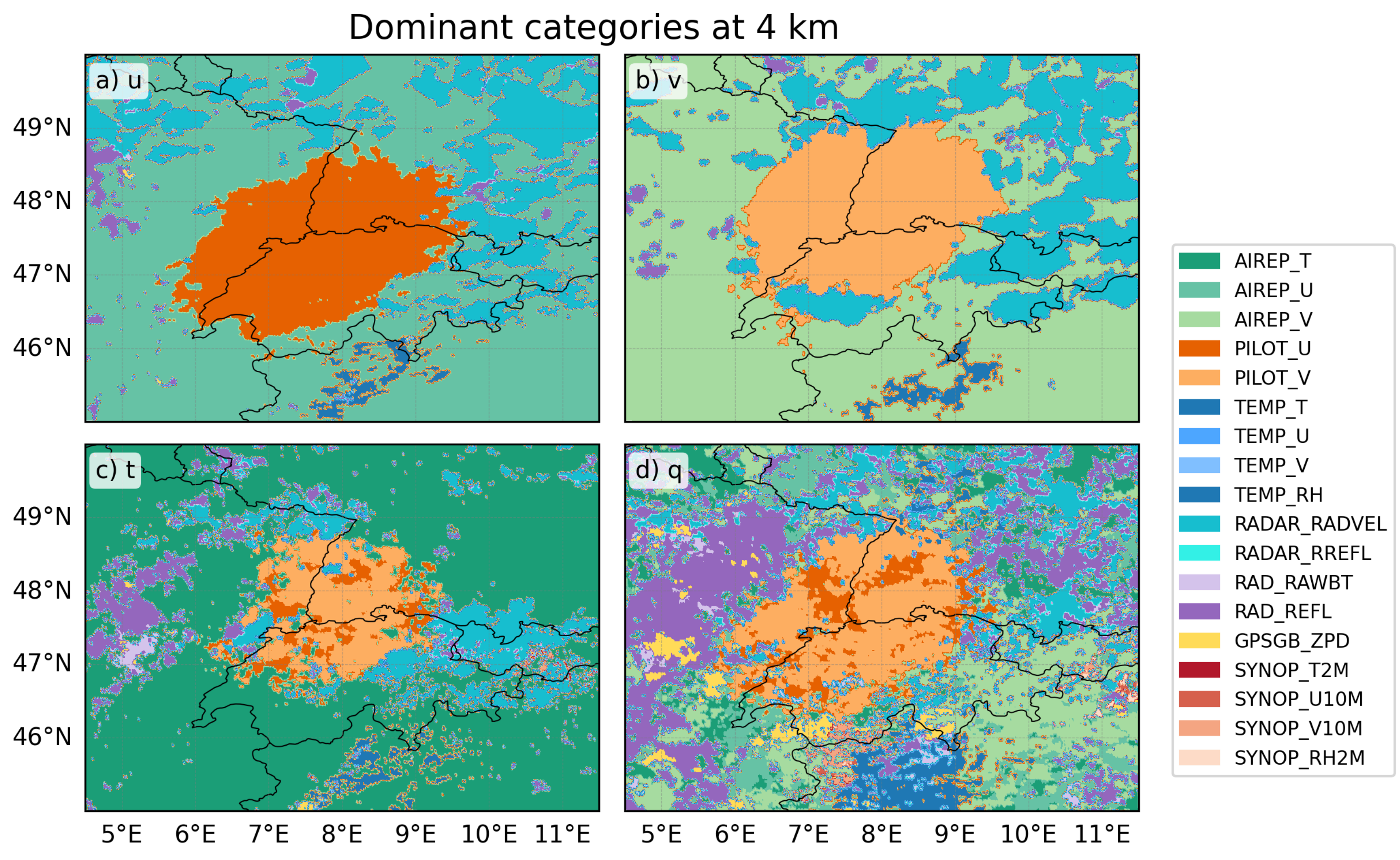
Conclusions

- The SWM wind-lidar network leaves a coherent, systematic fingerprint on the mesoscale analysis: In westerly flow conditions, it slows the zonal wind, strengthens the meridional wind, and spreads well beyond the campaign area
- PILOT/DWL dominates the wind increments in the lower and middle troposphere, thanks to both the dense network and its large localization length; it also reaches T and q near 600 hPa.
- Some observation types contribute with opposite signs in the same layer (AIREP vs PILOT wind in the mid-troposphere). This is suboptimal and points to a vertical-localization issue: observations still influence the analysis a few km above or below where they measure.
- The departure statistics of AIREP shows improvement in the lower troposphere (~800 hPa): a first indication of positive impact.
- Residuals mainly arise from inflation effect and observations outside the localization radius.

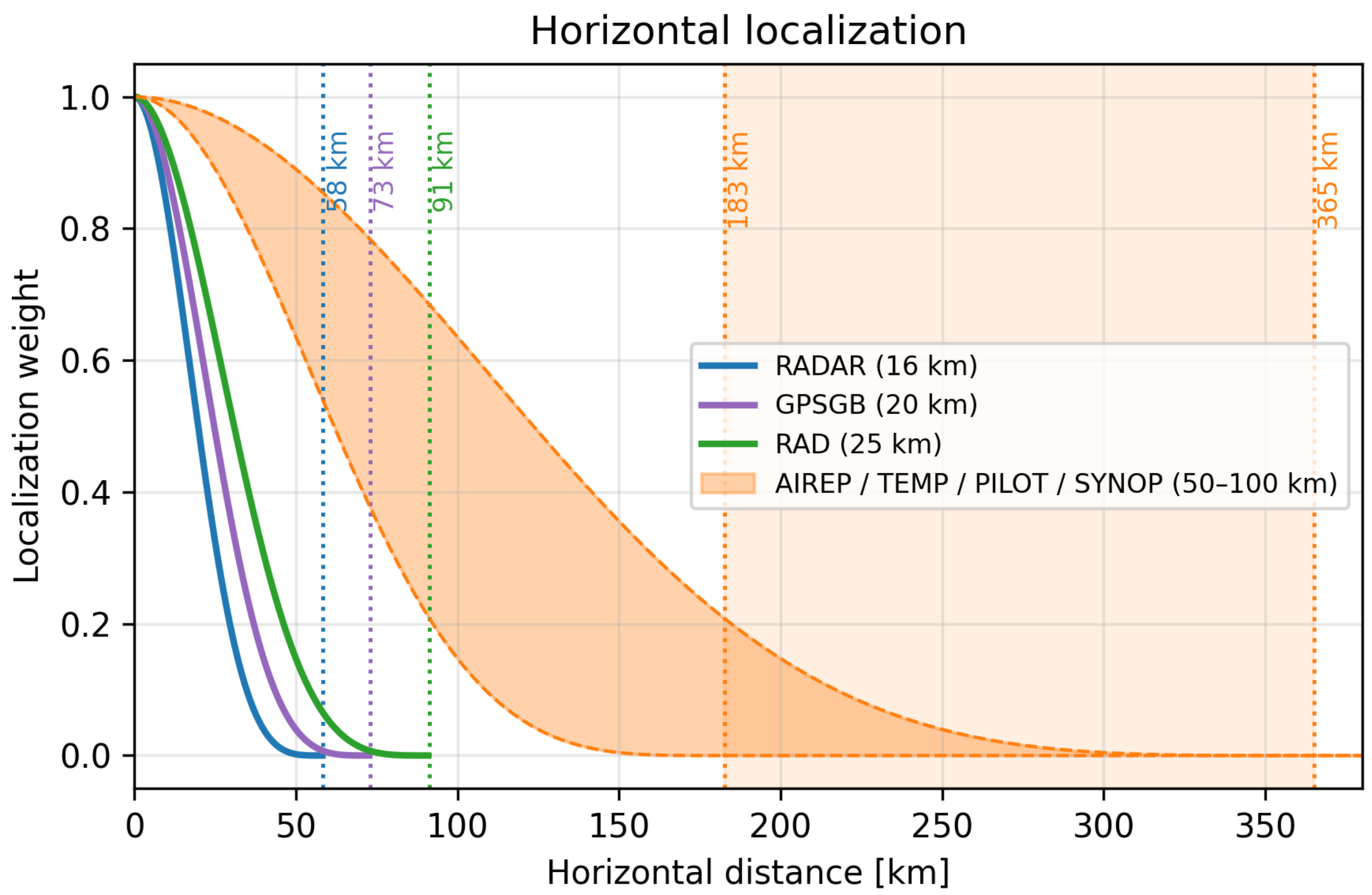
Effective observation impact

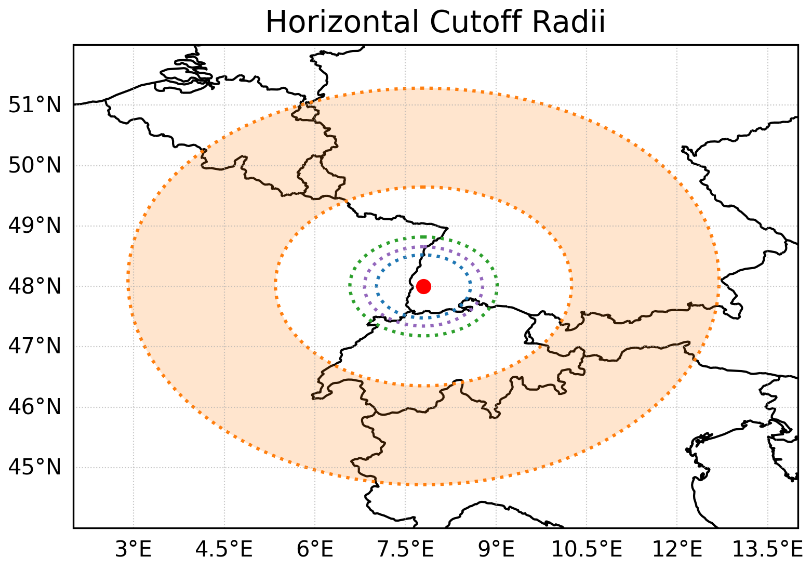


Relative Contributions of each observation type



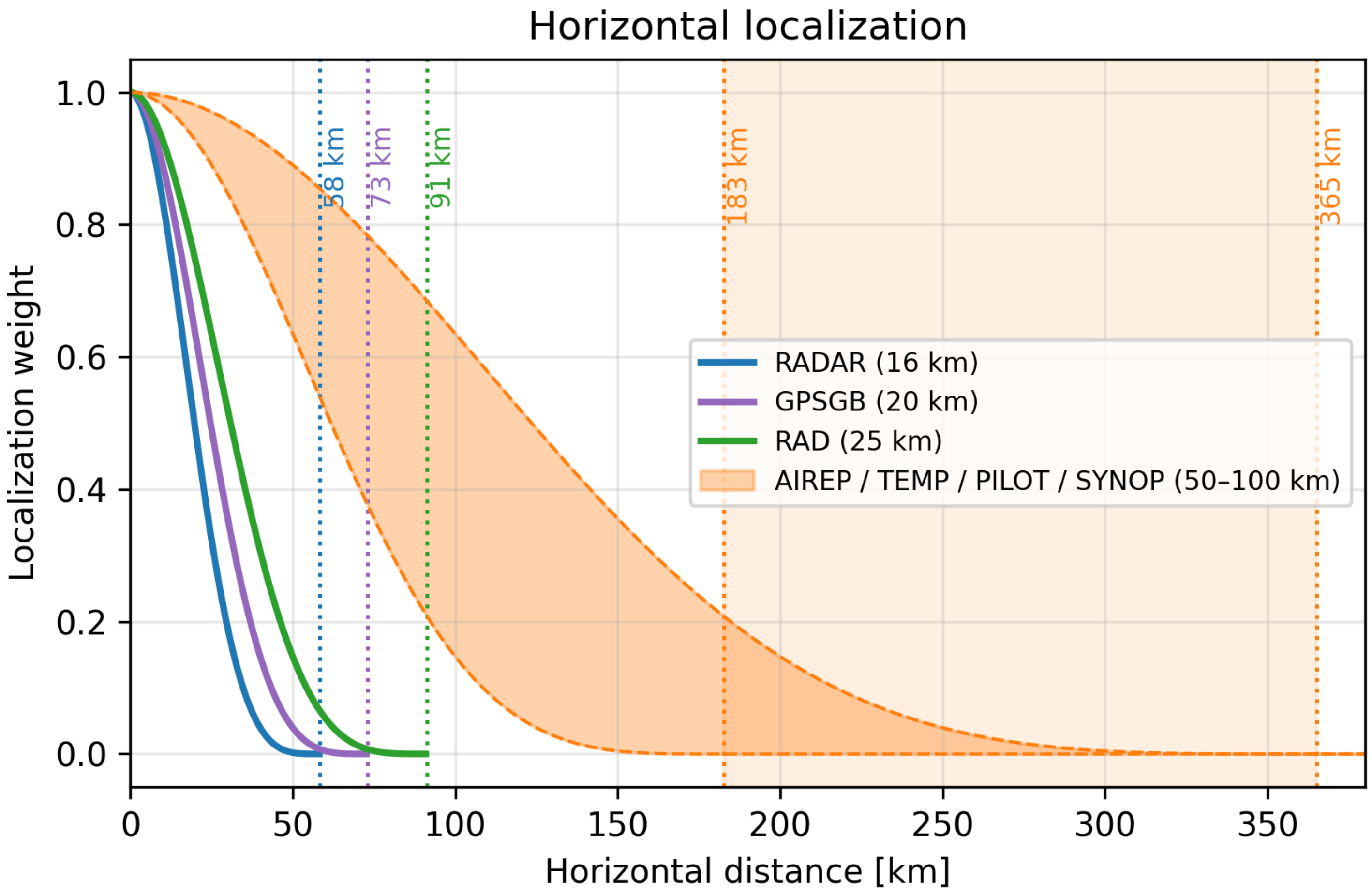
How are the observation localized?

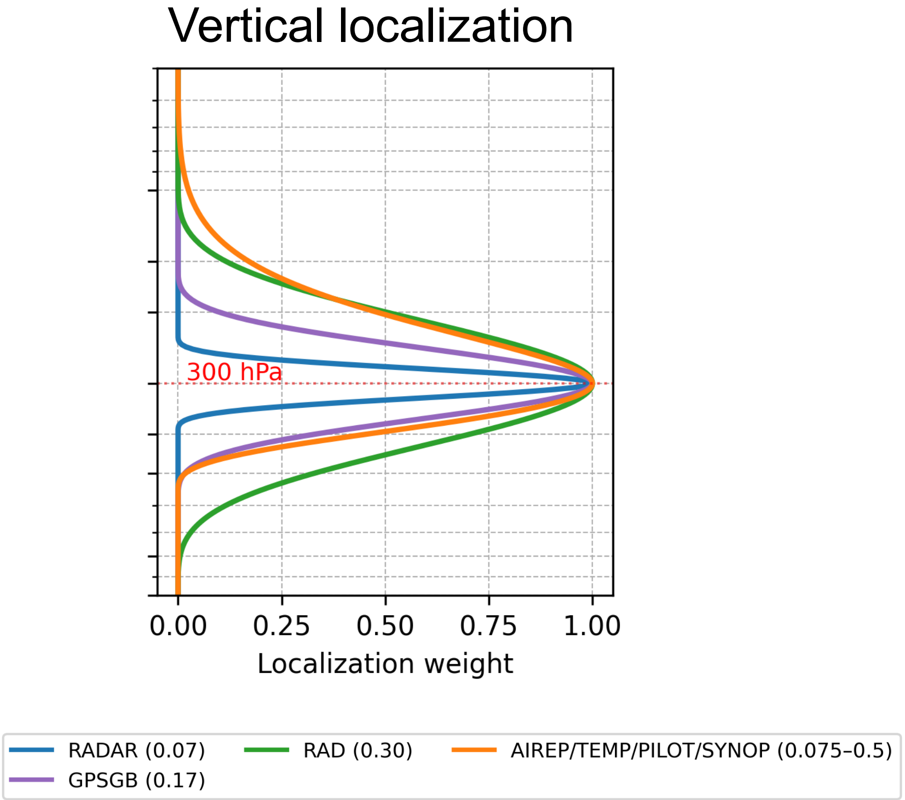




Gaspari–Cohn function: Gaussian-shaped curve that decays to zero after the cutoff radius

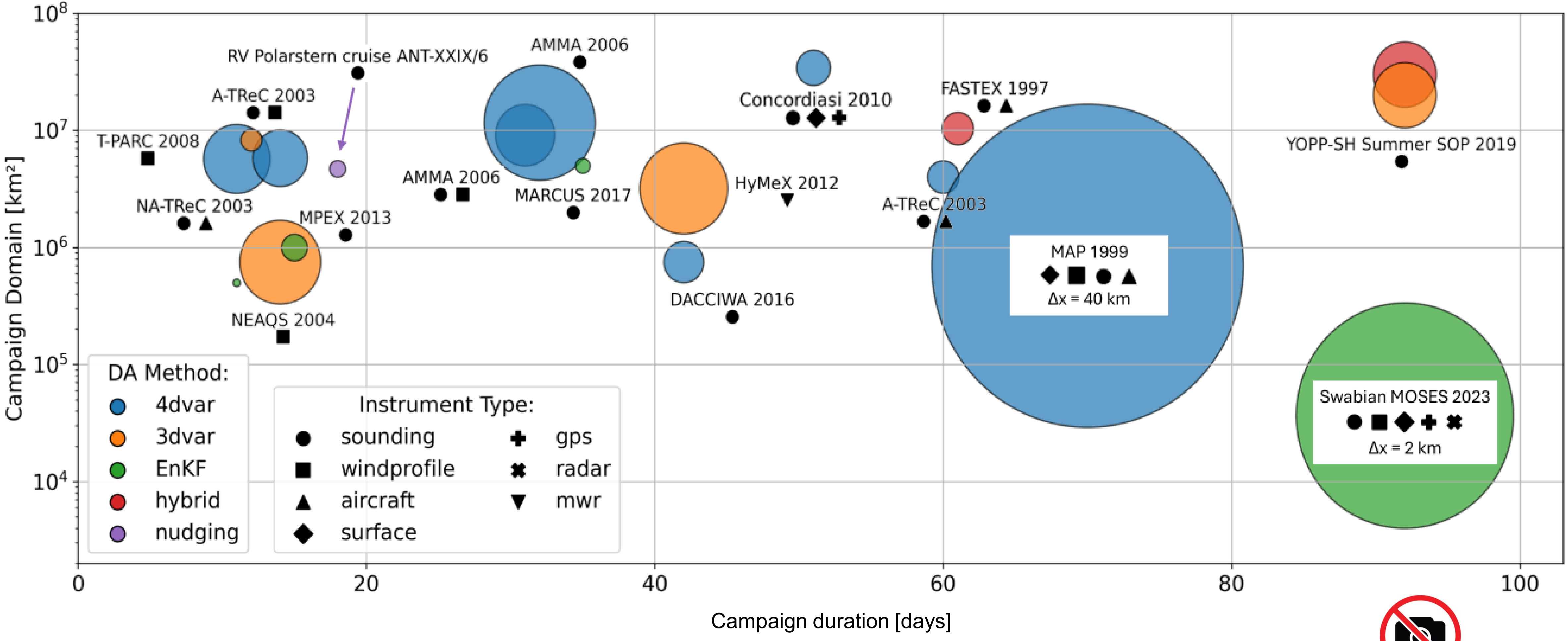
How are the observation localized?





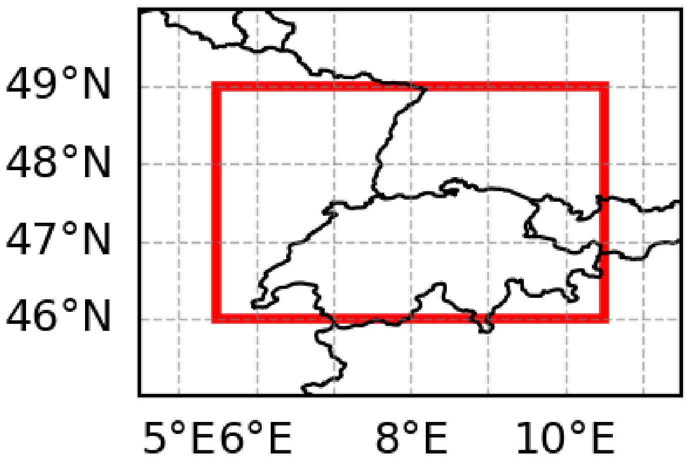
Gaspari–Cohn function: Gaussian-shaped curve that decays to zero after the cutoff radius

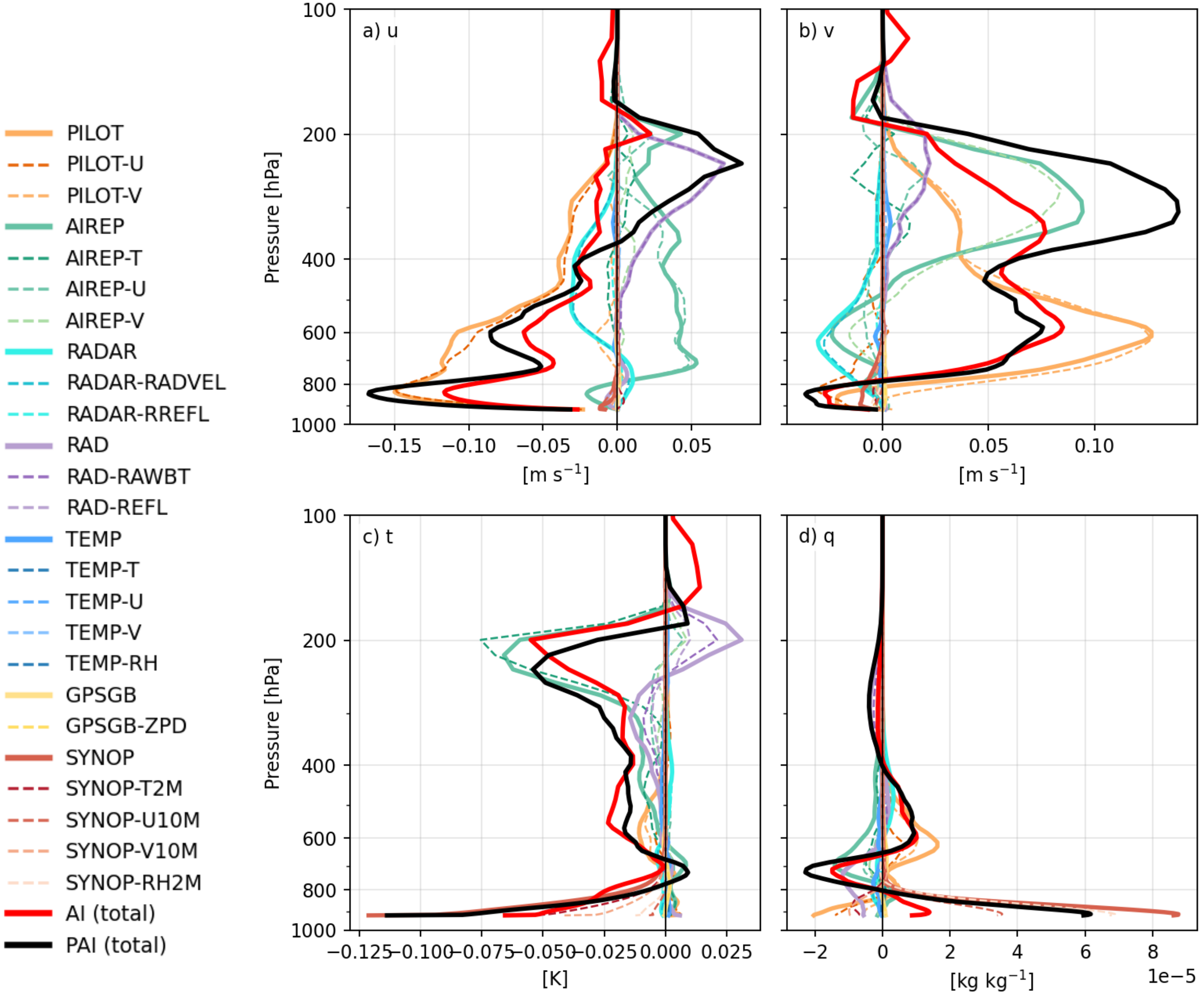
Overview of ground-based campaigns



Courtesy J. Thomas

TOTAL PAI


49°N
48°N
47°N
46°N
5°E6°E
8°E
10°E

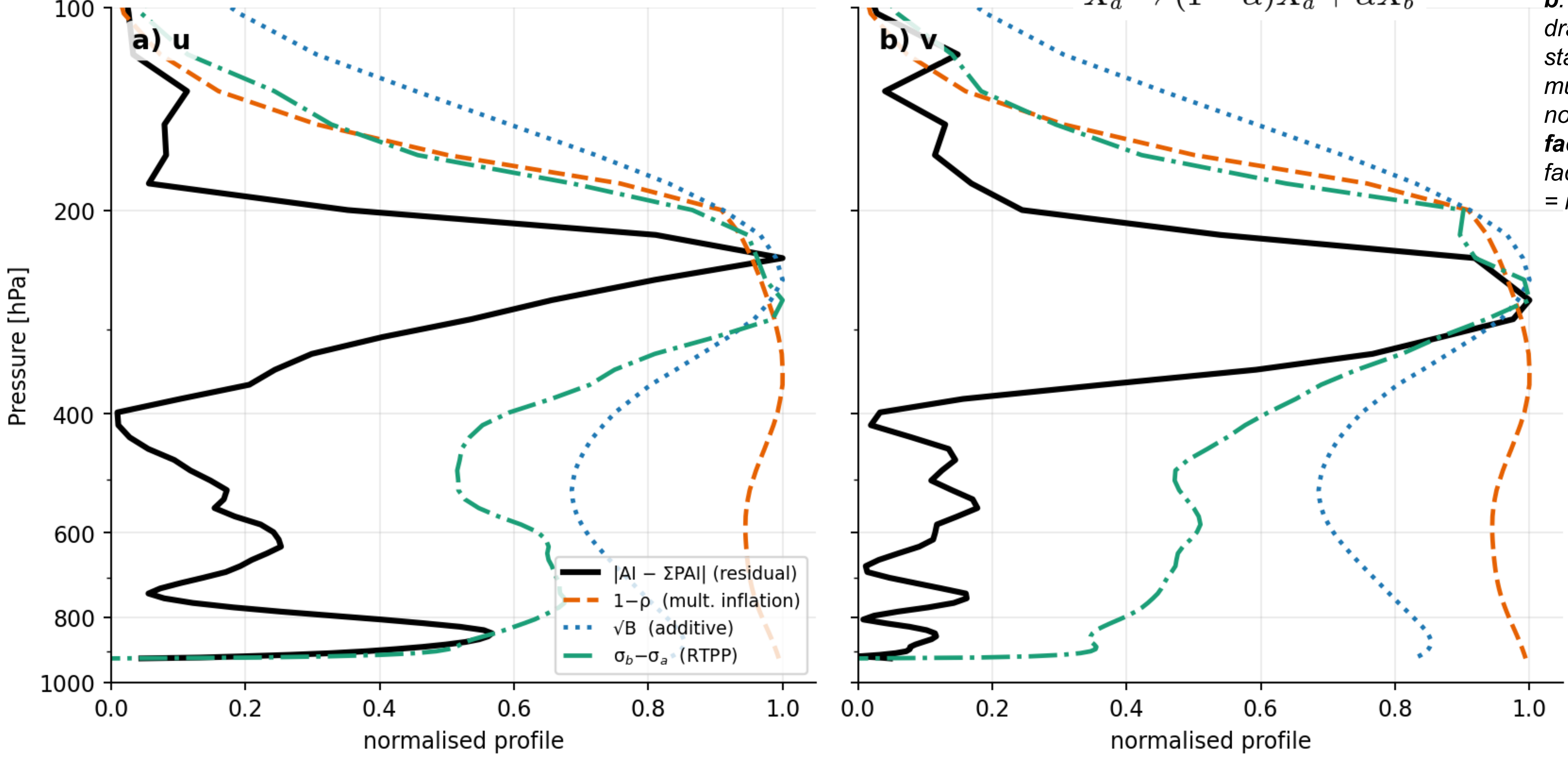

PILOT
PILOT-U
PILOT-V
AIREP
AIREP-T
AIREP-U
AIREP-V
RADAR
RADAR-RADVEL
RADAR-RREFL
RAD
RAD-RAWBT
RAD-REFL
TEMP
TEMP-T
TEMP-U
TEMP-V
TEMP-RH
GPSGB
GPSGB-ZPD
SYNOP
SYNOP-T2M
SYNOP-U10M
SYNOP-V10M
SYNOP-RH2M
AI (total)
PAI (total)
a) u
b) v
c) t
d) q
Pressure [hPa]
100
200
400
600
800
1000
−0.15 −0.10 −0.05 0.00 0.05
[m s^{-1}]
0.00 0.05 0.10
[m s^{-1}]
−0.125 −0.100 −0.075 −0.050 −0.025 0.000 0.025
[K]
−2 0 2 4 6 8
[kg kg^{-1}]
1e−5

Inflation

$$X_a \rightarrow \sqrt{\rho} X_a$$

$$X_a \rightarrow X_a + f_{\mathrm{add}} B^{1/2} b$$

$$X_a \rightarrow (1 - \alpha) X_a + \alpha X_b$$



B: *Background covariance matrix*
b: *Random vector drawn from the standard multivariate normal distribution*
fadd: *Scaling factor (0-1) with 0 = no inflation*