Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

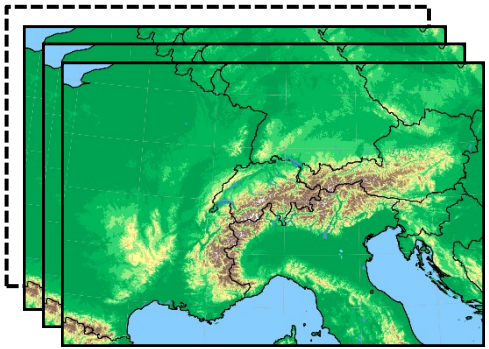
Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

Diagnostics on 2m Temperature assimilation in the KENDA-CH1 system at MeteoSwiss

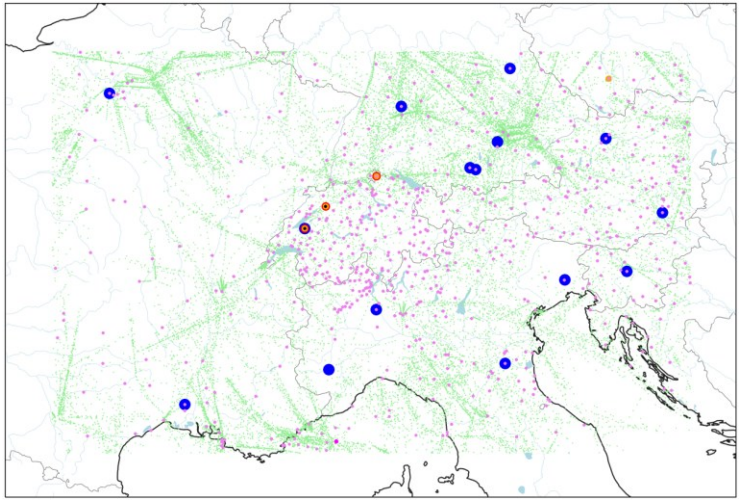
Krishnamoorthy Chandramouli, Daniel Leuenberger, Claire Merker
MeteoSwiss, Zürich, Switzerland

Background

- Operational KENDA-CH1 assimilates variety of observations including the near surface variables like 2m Temperature, 2m Relative Humidity and surface pressure



ICON with lateral boundary conditions from IFS (IFS ENS 18km, IFS HRES 9km)



- Aircraft
- Radiosondes
- Wind Lidar
- Wind Radar
- Land Surface Stations
- Ship Surface Stations
- Buoys

+ Radar QPE with Latent Heat Nudging

KENDA-CH1 ensemble data assimilation system (Schraff et al. 2016): LETKF (Hunt et al. 2007), 40+1 members at 1.1km, hourly cycling

Background

- Assimilation of near surface variables has resulted in a positive impact on the forecast of fog. Still a visible degradation of the analysis was observe in some cases which results in a poor fit to radiosondes.

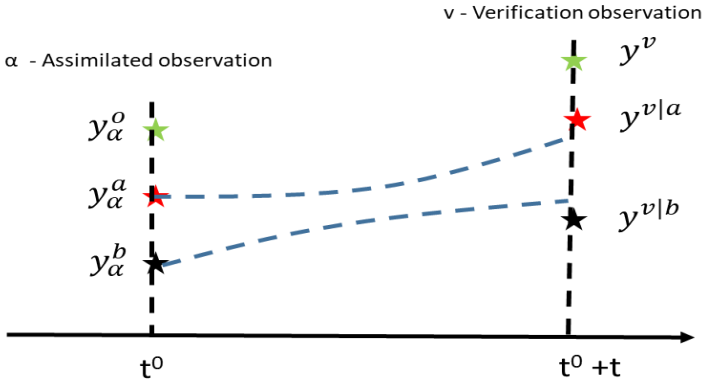
- Given the model deficiencies, dense surface observation network and complex topography encountered within the NWP model domain, we want to answer the following questions

  - ✓ Can we optimize the KENDA system (tuning the localization radius etc.) to reduce the negative impacts of 2m Temperature assimilation?

  - ✓ Can we employ any selection strategy to flag stations which could not be optimally assimilated (possibly due to model biases / representativity issue) ?

Cross-validation diagnostics (Stiller, 2022)

- CV tool : Observation space diagnostic to assess the impact of a given set of observation against a verifying truth
- For an observation to create positive impact, the J^b term should be positive (assimilation of a given observations pulls the forecast towards the verifying truth)
- Cross-validation refers to the comparison of J^b against its reference value
- Comparing the observation vs the ensemble estimated covariances
- Single observation equivalent of the CV with focus on the analysis



$$J = 0.5\left[\|y^v - y^{v|a}\|^2 - \|y^v - y^{v|b}\|^2\right]$$

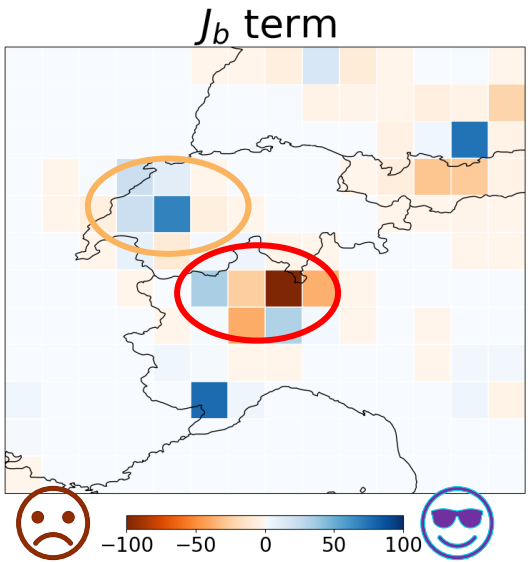
$$J = -0.5(2J^b_\alpha - J^{ab}_\alpha)$$

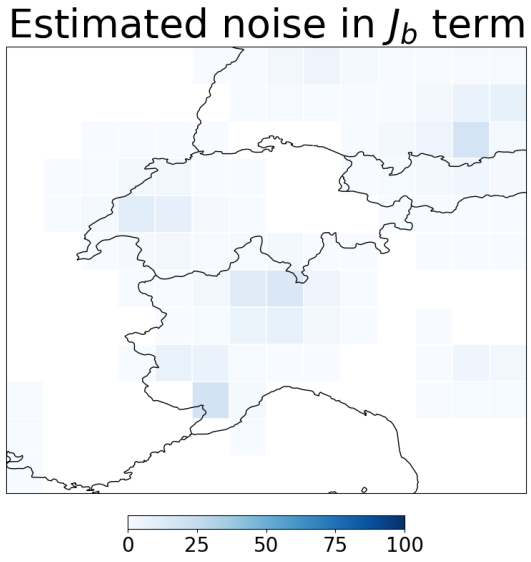
$$J^b_\alpha = \sum_v \hat{P}^a_{\mathrm{en}[v,\alpha]} \frac{(y^{\mathbf{v}}_v - y^{\mathbf{v}|b}_v)(y^o_\alpha - y^b_\alpha)}{R_{vv}R_{\alpha\alpha}}$$

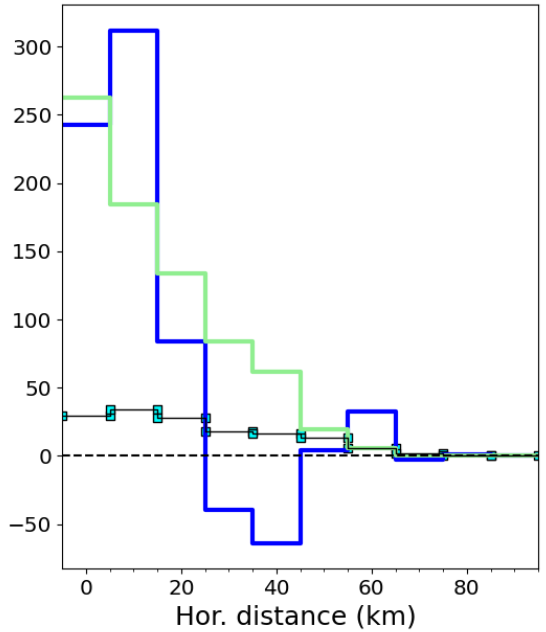
$$\langle J^b_\alpha \rangle_{\mathrm{estim}} = \sum_v \hat{P}^a_{\mathrm{en}[v,\alpha]} \frac{\hat{P}^b_{\mathrm{en}[v,\alpha]}}{R_{vv}R_{\alpha\alpha}}.$$

Impact of T2m assimilation

- Known issue with the T2m is revisited with CV tool (period: OND2023)
- Visible degradation due to the 2m Temperature assimilation
- Detailed diagnostics over Po valley and Payerne





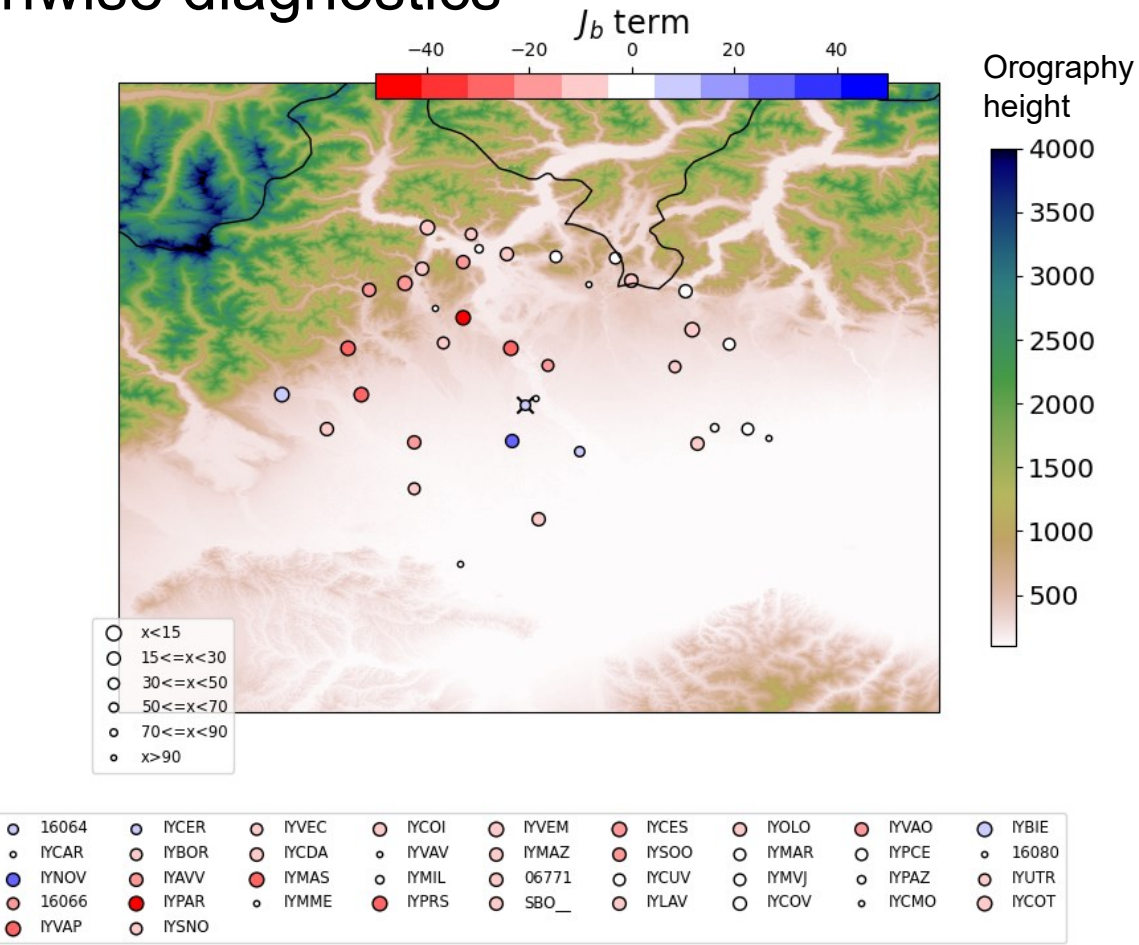


Mismatch between observation and ensemble estimated covariances

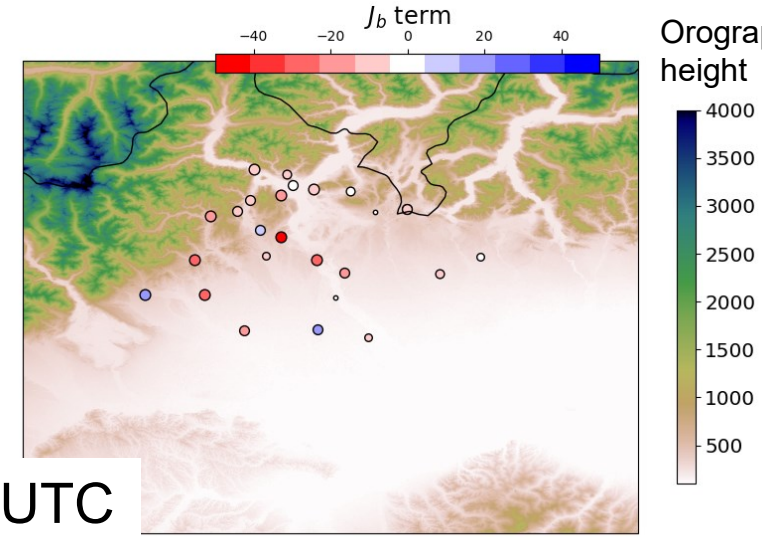
Possibility to reduce the horizontal localization

Po valley: Stationwise diagnostics

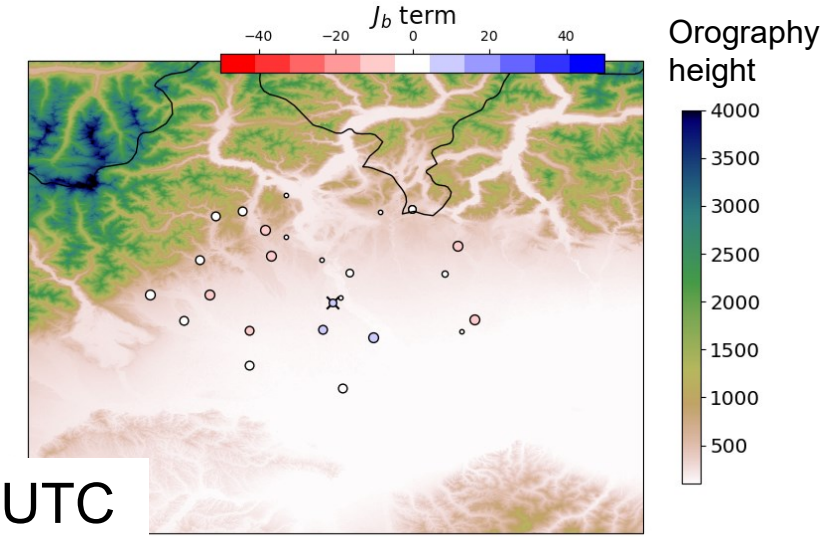
- Only stations which are located below 950hPa are considered here
- Most of the stations result in a negative impact
- Some of the stations are located close to the orography/ in valleys
- Any diurnal variability in the impact?



Po valley: Stationwise diagnostics



00UTC



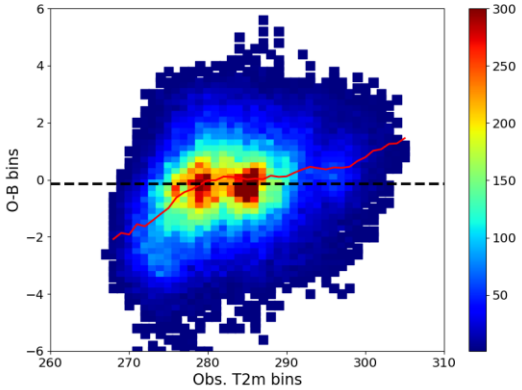
12UTC

More severe degradation during the nighttime compared to daytime.

Further diagnostics to understand systematic dependence of the impact on the observed surface temperature.

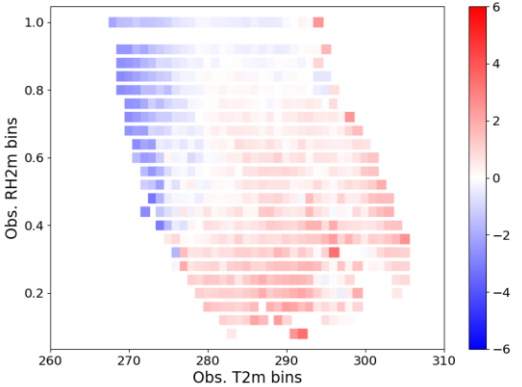
Po valley: Observation space diagnostics

Statistics from stations which measure both 2m T and RH (OND 2023)



Binned obs. T2m vs O-B with colors representing the number of data points in each bin.

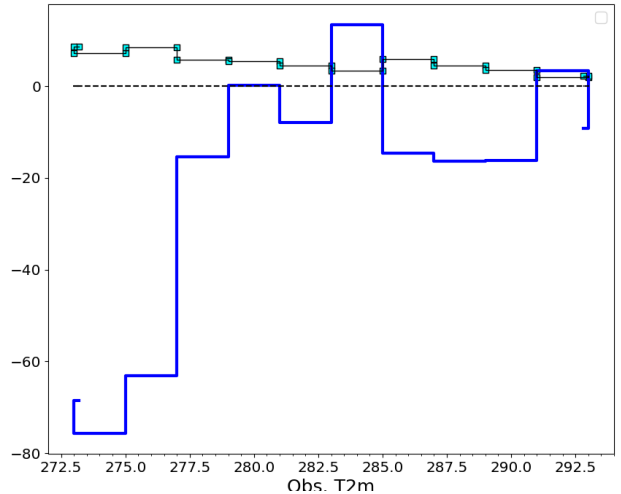
Red line represents the bias in O-B at each T2m bins,



Binned obs. T2m vs Obs. RH2m with colors representing the T2m O-B in each bin

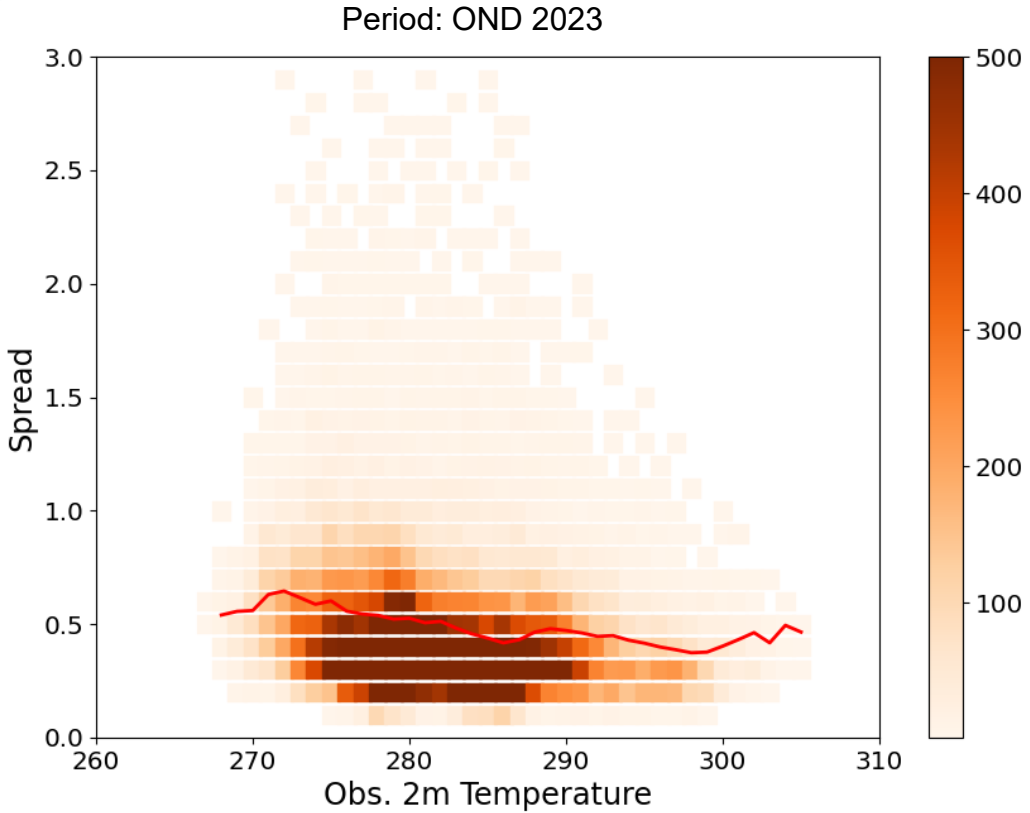
Negative O-B bias in T2m not necessarily related to the RH saturation conditions

- Systematic bias at colder near surface temperatures
- Consistent with this, CV tool also points to a large degradation of the analysis from cold surface observations



J^b term (blue) as a function of observed 2m temperature along with the noise in the estimate (black)

Po valley: Observation space diagnostics



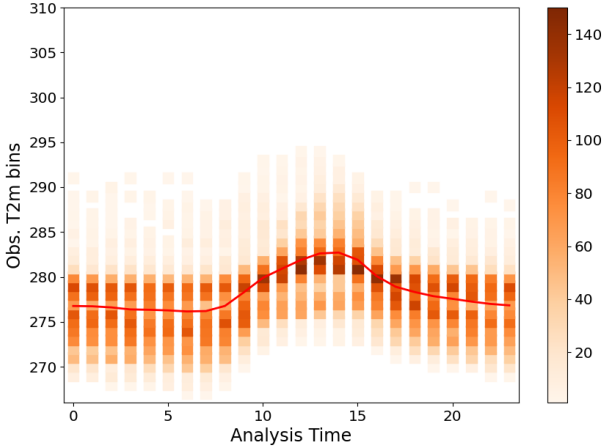
Binned background ensemble spread at station locations (38 station) vs observed 2m temperature with color representing the number of data points in each bin.

Red line represents the average spread in each obs. T2m bin.

Larger ensemble spread at observed cold 2m temperatures.

Po valley: Observation space diagnostics

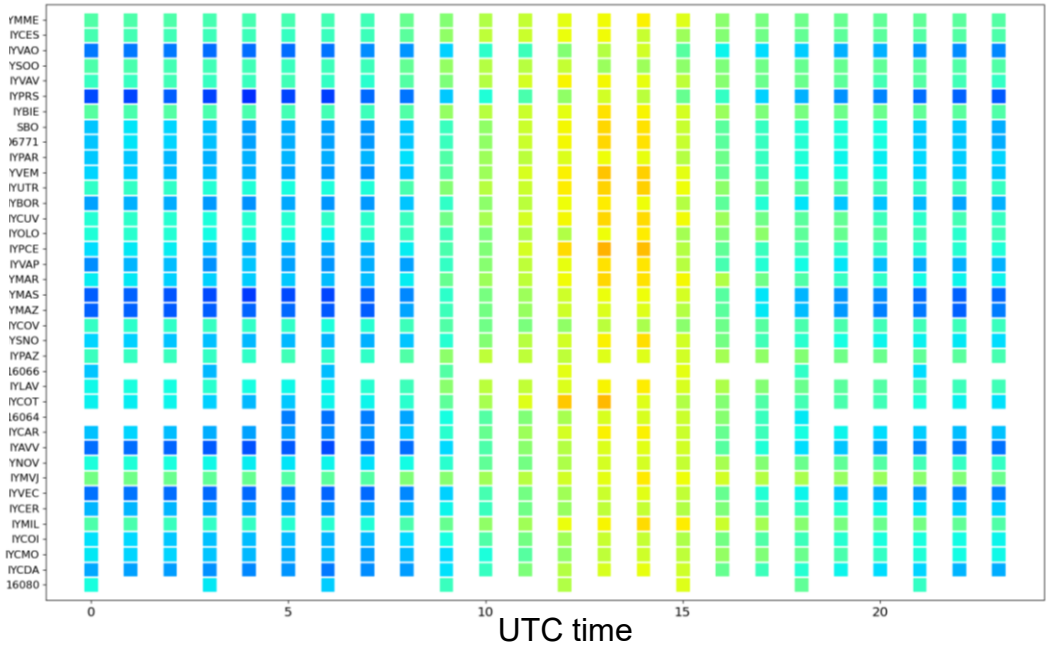
Period: Dec 2023



Binned obs. 2m temperature from the 38 stations at each analysis cycle with colors represent the data density

Two distinct clusters of observed T2m

Station-wise average observed T2m at each analysis cycle



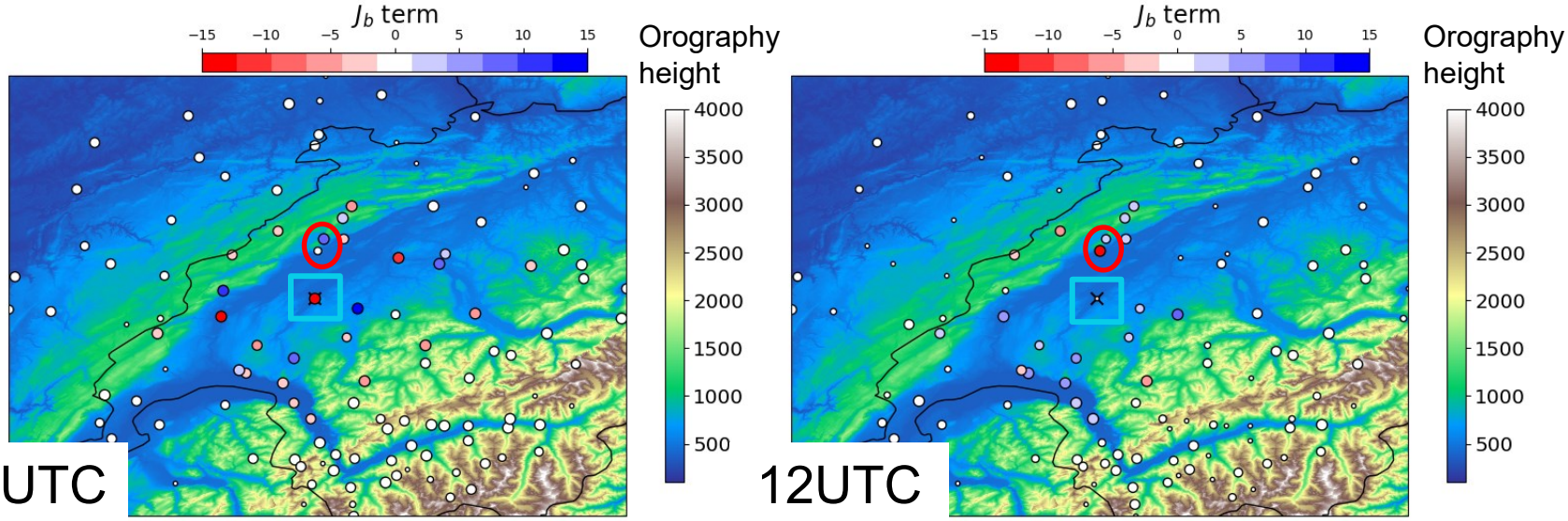
Clusters of cold observed 2m temperatures points to a possible issue of representativity. More detailed study underway on station selection strategy

Payerne: Stationwise diagnostics

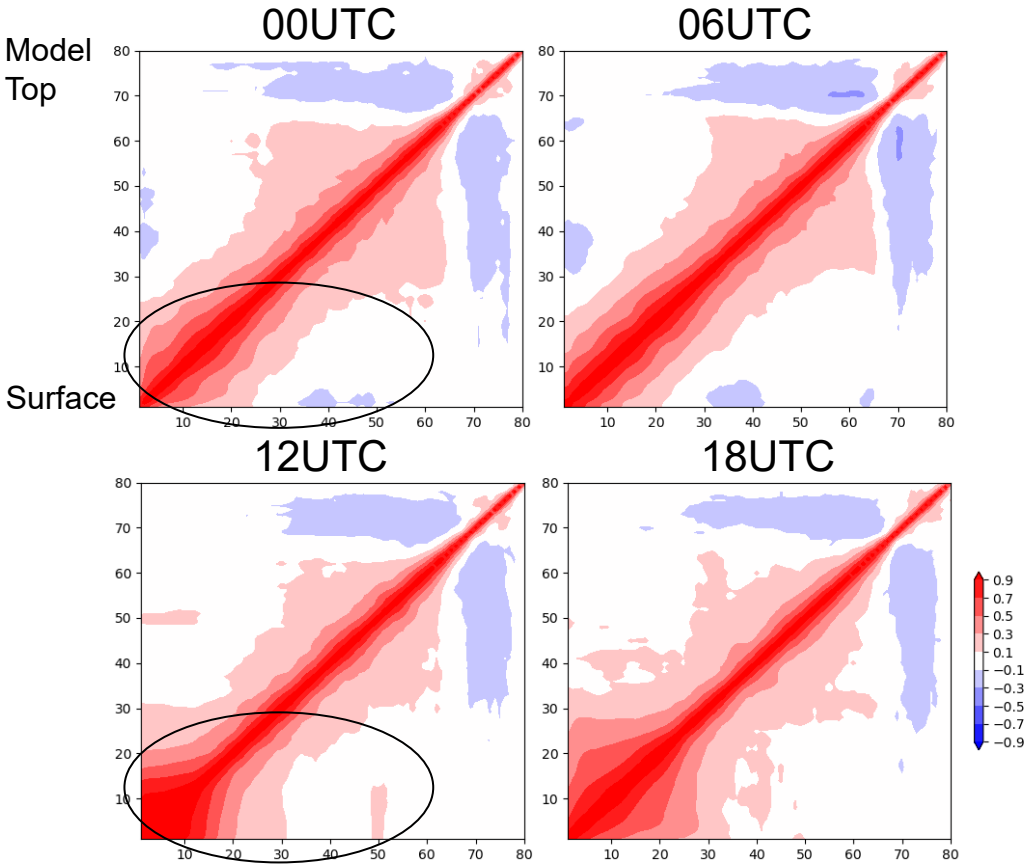
Effect of nighttime low-level inversion: Large negative impact during the night

Illustrates the difference in impact of high-altitude stations

Period: JJA 2023



Payerne: Model space diagnostics



Ensemble vertical temperature cross-correlation at grid point close to Payerne (temporally averaged over June 2023)

Difference in the vertical correlation structure in the lower model levels.

More detailed study underway to understand the effect of the vertical localization scale on low level inversion cases.

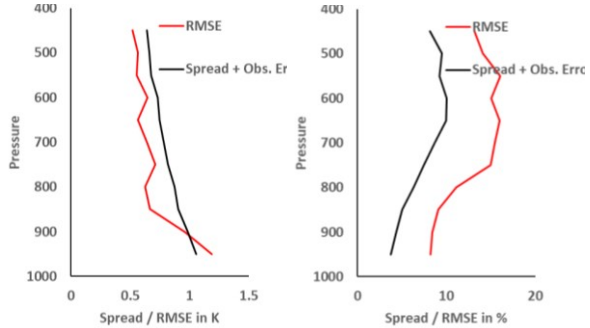
Outstanding issues with the assimilation of surface variables in KENDA-CH1

- Spin-up issue in 30min surface pressure
- Double counting of 2m variables
- Loss of 30min observations due to redundancy check

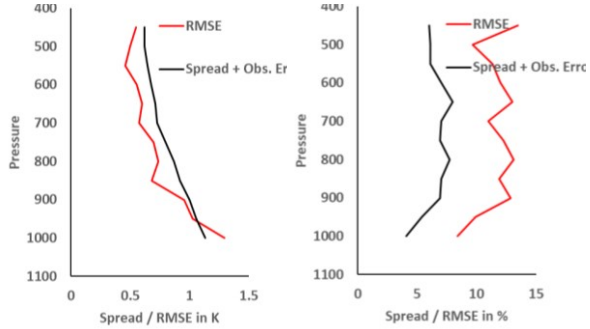
Spin-up issue in 30min surface pressure

Temperature **RH**

Payerne RS



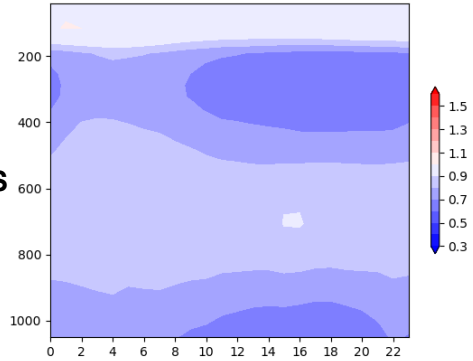
Milan RS



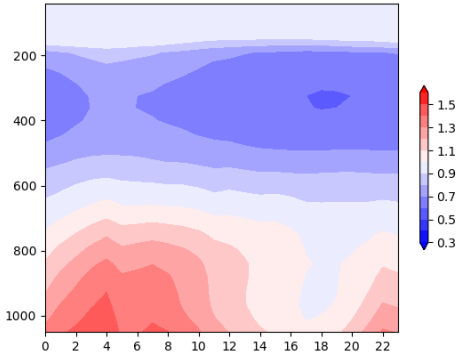
Averaged over July 2023

Adaptive inflation factor (rho) at different analysis times

Gridpoint close to Payerne RS



Gridpoint close to Milan RS



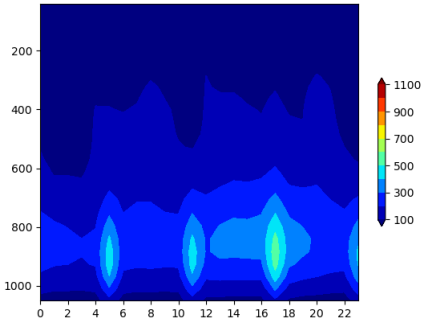
Based on the diagnostics from RS, under-dispersion can be seen clearly in the T and RH. The winds are slightly over dispersive.

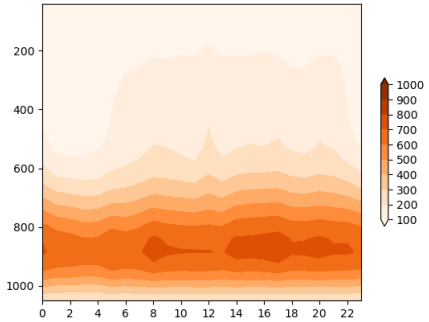
- The time average adaptive inflation over Payerne at different analysis times shows deflation.
- More diagnostics based on the diag files…

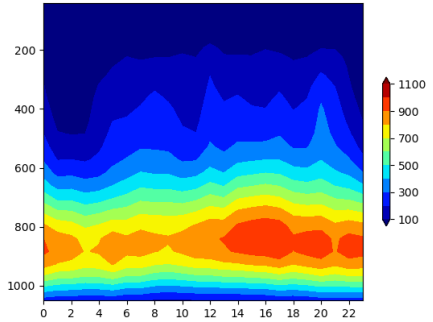
Spin-up issue in 30min surface pressure

tr(HPH^T o R^{-1}) | **Sum of obs. weights** | **tr(dd^T o R^{-1})**

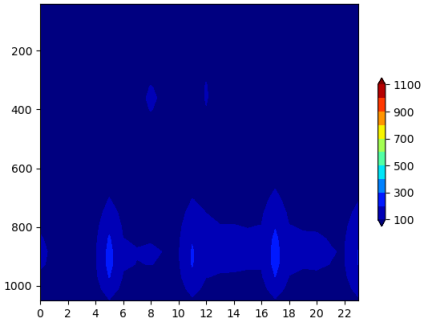
Gridpoint close to Payerne RS

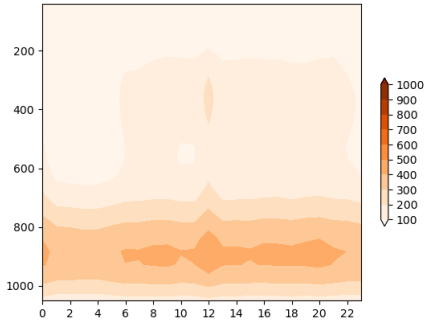


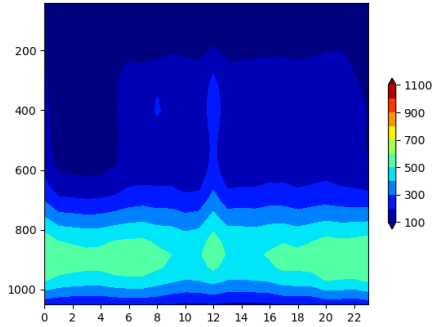




Gridpoint close to Milan RS







Averaged over July 2023

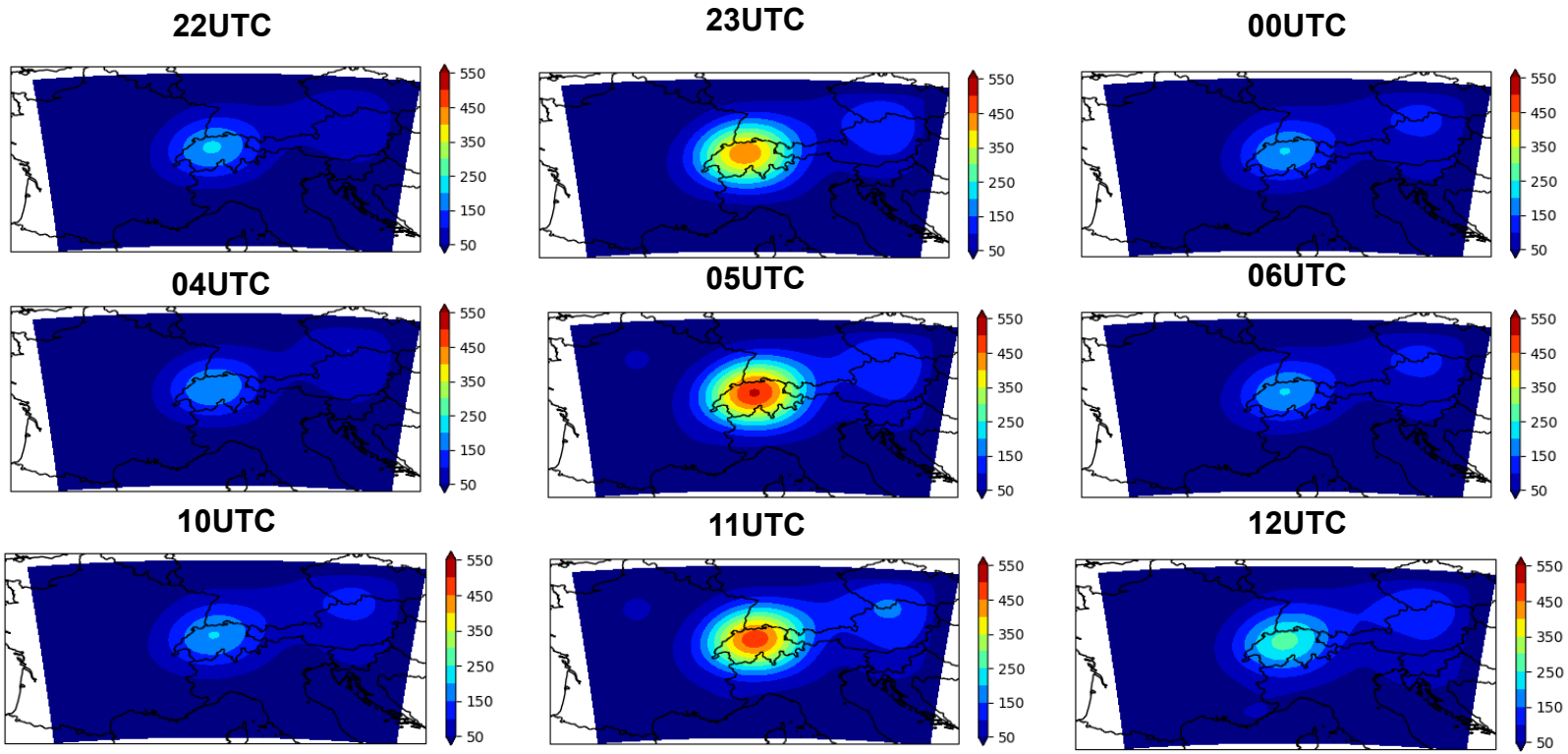
Quantities used in the calculation of adaptive inflation factor.

Plotted from the diag files from KENDA-CH1 pre-operational runs

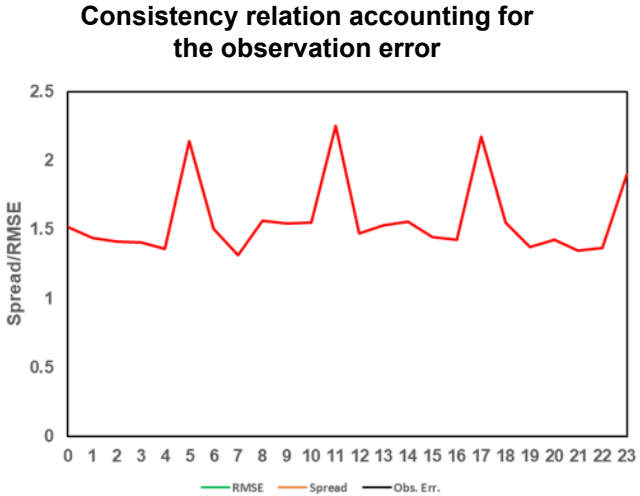
Spikes in the values of **tr(HPH^T o R^{-1})** term at regular intervals

Spin-up issue in 30min surface pressure

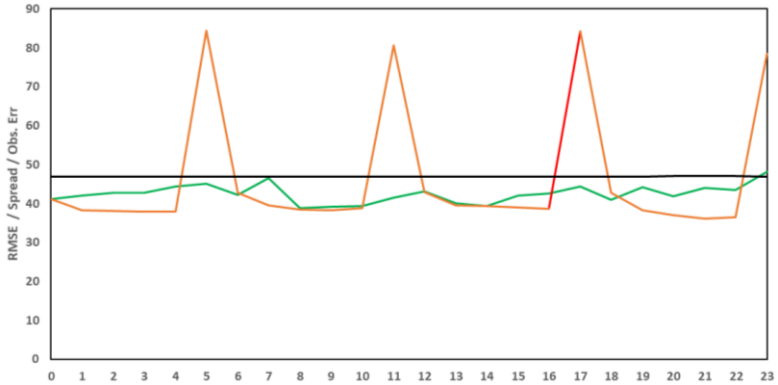
Spatial plot of tr(HPH^T o R^{-1}) at ~960hPa level averaged over July 2023

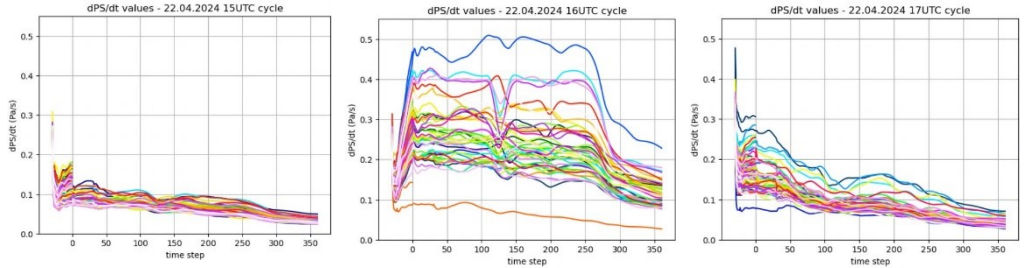


Spin-up issue in 30min surface pressure



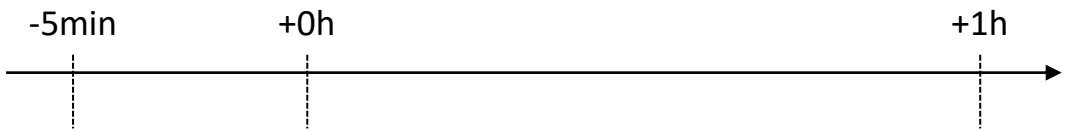
- Observations over the specific region of interest was analyzed to understand the spikes in **tr(HPH^T o R^{-1})**
- Based on this, it was seen that the winds (U and V) were generally over-dispersive.
- But the periodical spike is explained only from the Ps observations.
- Problem with the spin up when new boundary conditions are used -> this translates as larger ensemble spread in Ps especially at T-0.5 hrs



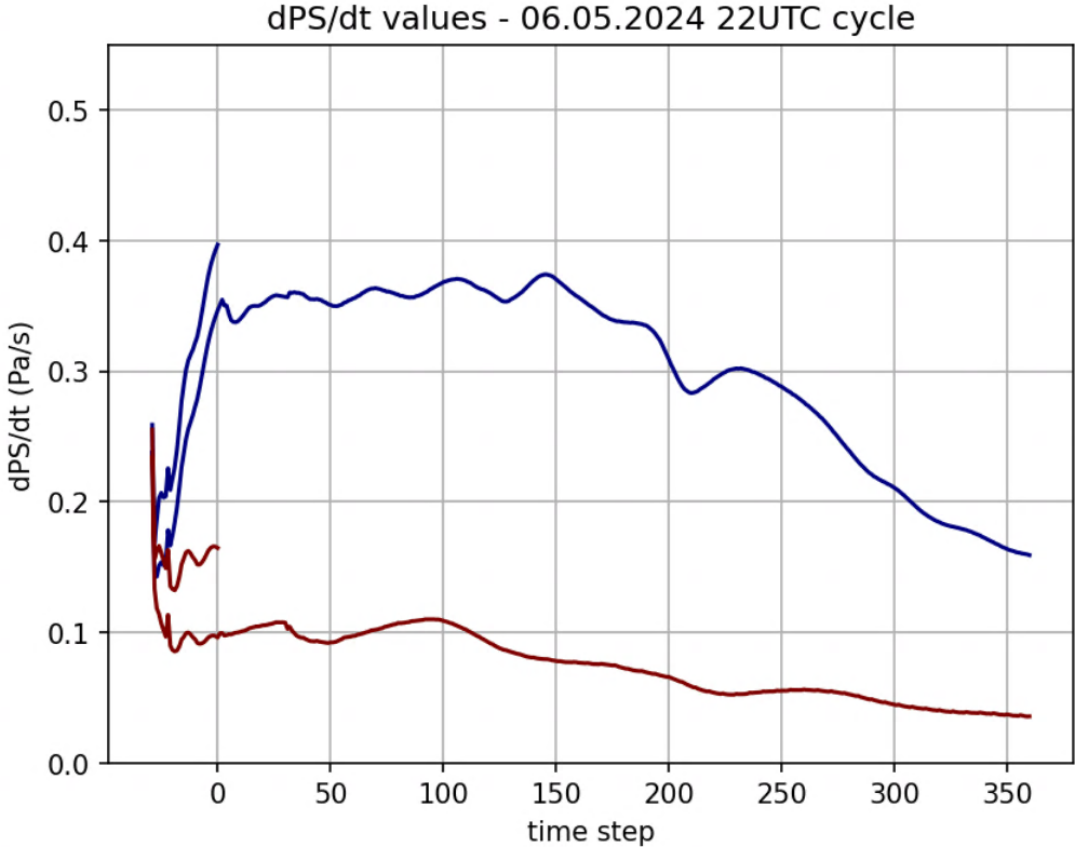


Hypothesis: change in LBCs (happening in 04, 10, 16 and 22 UTC cycles)

- new IFS LBCs
- det: no effect because HRES is quite stable?
- members: IFS ENS perturbations are added to the HRES, and perturbations probably don't have memory after new init → a KENDA member can get very different perturbations after LBC update
- Due to IAU, the interpolation between old and new LBCs is only 5min



KENDA-CH1	ICs	new IFS	new IFS
KENDA-1	/	ICs	new IFS
test	ICs	old IFS	new IFS

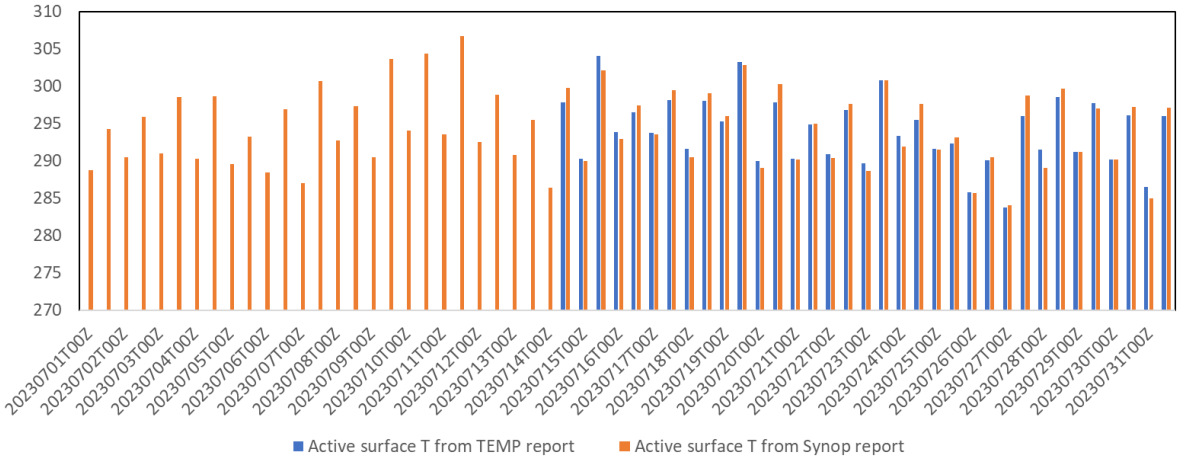


member 10

- preopr setup
- using old LBCs

Double counting of 2m variables

- From 14 July 2023 onwards, the surface obs in both TEMP and SYNOP are active over Payerne
- After dace_icon code change in July 14,2023. The surface reports in fofTEMP are set to active



fofTEMP_20230701110000 Header

STATID	TIME	STATUS	FLAG	LEVELS
6610	60	1	0	57
6610	60	7	524288	1

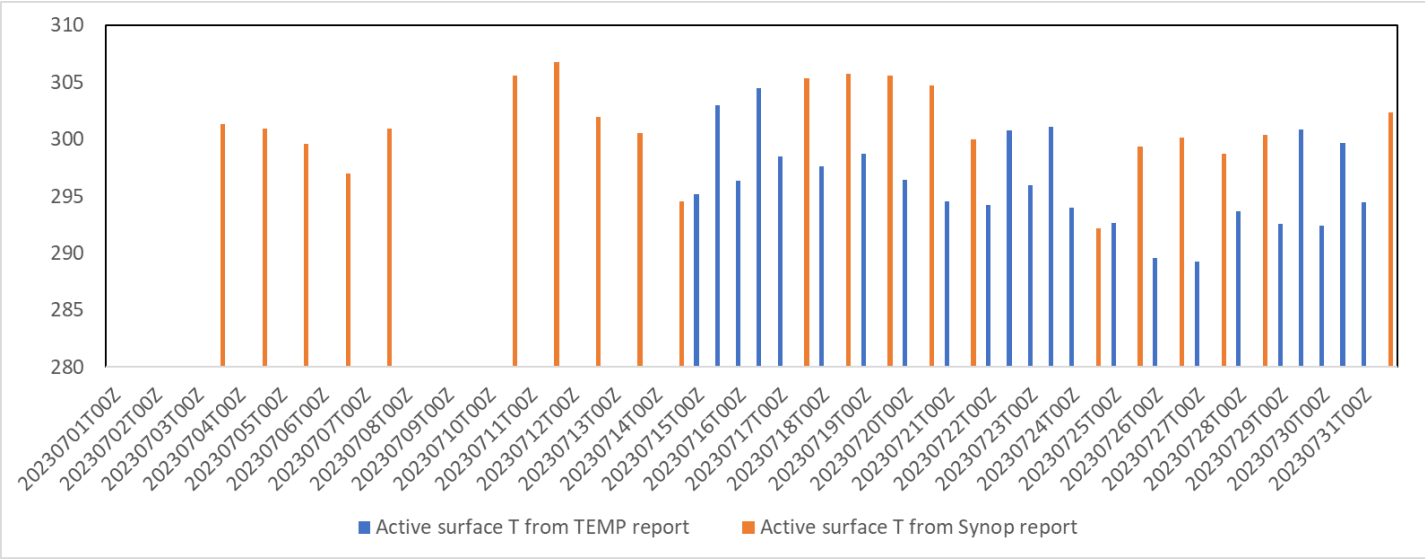
fofTEMP_20230715110000 Header

STATID	TIME	STATUS	FLAG	LEVELS
6610	60	1	0	42
6610	60	1	0	1

Active observations over Payerne in ekfTEMP and ekfSYNOP files from July 2023

Double counting of 2m variables

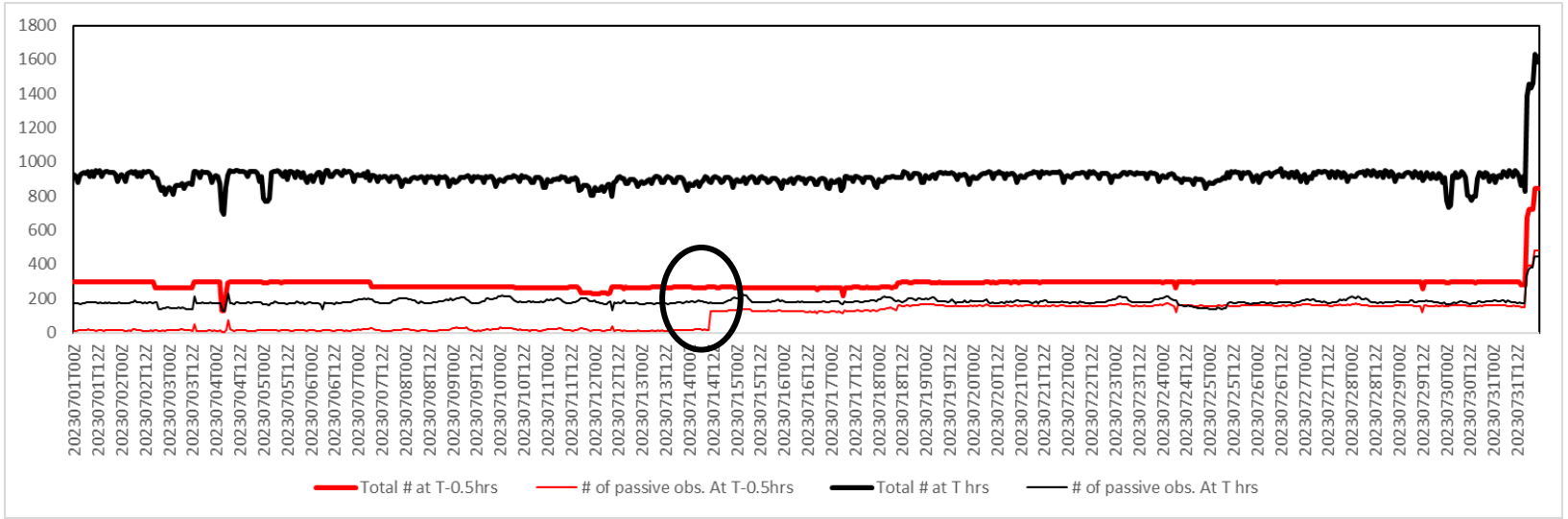
- The assimilation of surface obs from TEMP as well as SYNOP does not happen for all stations.
- Over Milan, after July 14. 2023, the surface variables from TEMP are assimilated but only when the variables are missing from SYNOP files. There is no double counting here.



Active observations over Milan in ekfTEMP and ekfSYNOP files from July 2023

Loss of 30min observations due to redundancy check

- After July 14, icon-dace update, the number of 30 min observations of 2m T/RH set to passive (due to redundancy flag) has increased.
- Possible due to the inclusion of a new redundancy check criteria based on the analysis time ?
- The redundancy check seems to flag mostly the partner stations which report obs every 10 min.



Summary

- Diagnostics on the pre-operational KENDA-CH1 system shows some degradation of the analysis due to the assimilation of 2m Temperature in some regions of the model domain.
- With guidance from the Cross-Validation tool, we are able to locate different regions where this impact is more significant.
- With additional diagnostics over the Po valley region, we could illustrate the presence of a strong negative systematic bias in O-B with respect to observed 2m temperature which increases with decreasing surface temperature.
- Further, we could see some hints of the representativity issue where certain stations (which are located in valleys or closer to the orography) tend to have persistently colder temperature compared to others.
- Diagnostics over the Payerne region shows a negative impact from surface stations during nighttime inversion scenarios. This also points to a possible optimization of localization scales to optimize the use of these stations.
- Further investigations into KENDA system revealed some outstanding issues with the assimilation of surface variables.

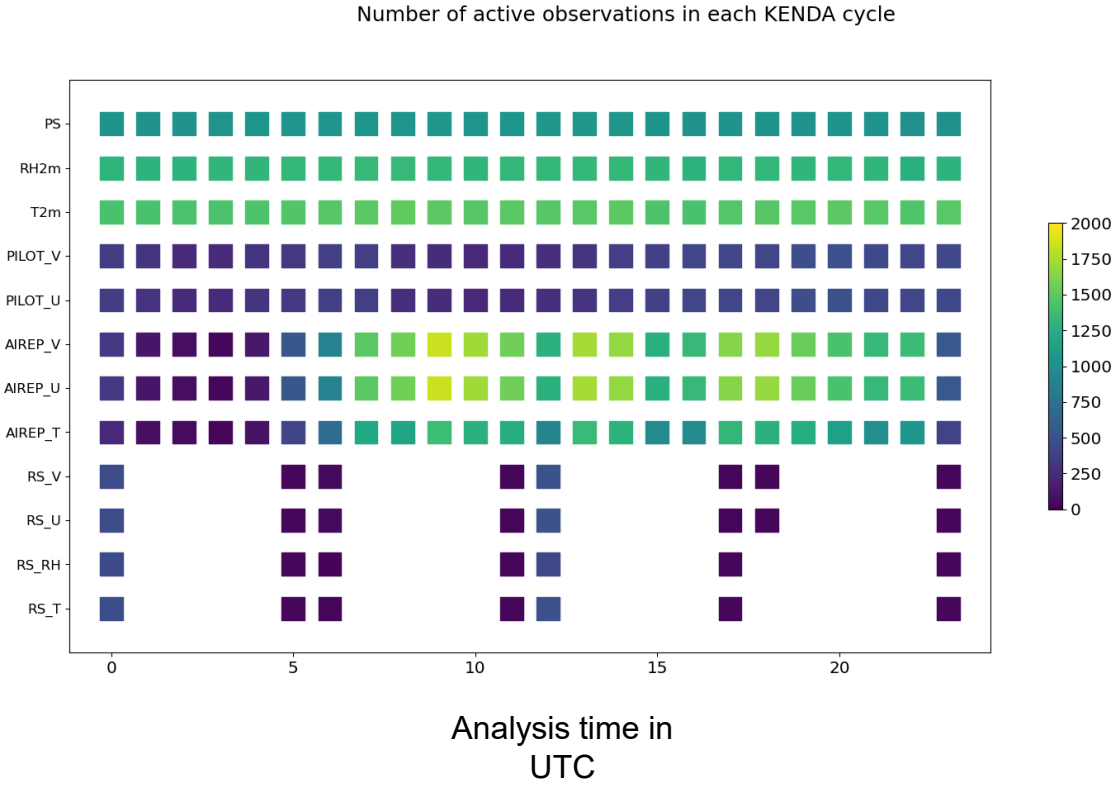
Work in progress …

- Given the spin-up issue with the sub-hourly surface pressure observations, we are currently testing the hourly assimilation of surface level quantities.

- With the guidance from CV and other diagnostics, we are now testing with reduced horizontal localization scale for SYNOP observations.

- Resolving other outstanding issues with the assimilation of surface level observations.

Thank you!



Surface observations are the major contributors from 23 to 06UTC

There are only two sources of information on humidity: RH from Sondes and 2m RH