How to use partial analysis increments in an LETKF data assimilation system

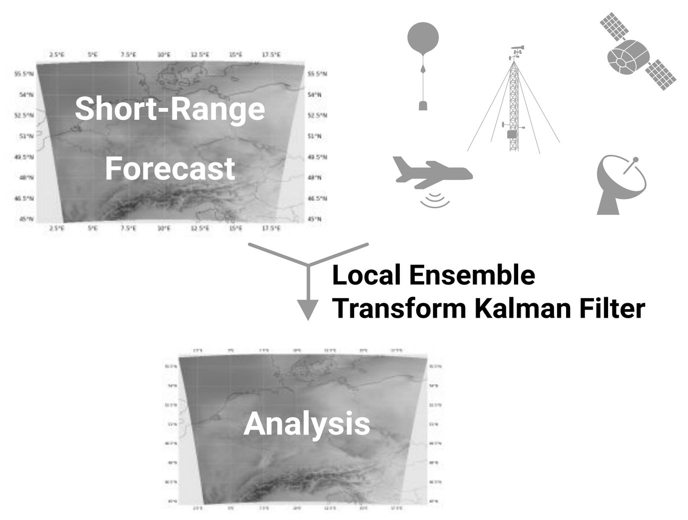
Theresa Diefenbach[1], George Craig[1], Christian Keil[1], Leonhard Scheck[2], Martin Weissmann[3]

1) Meteorologisches Institut, Ludwig-Maximilians-Universität München, Germany
2) Hans-Ertel-Centre for Weather Research, Data Assimilation Branch, LMU München, Germany
3) Institut für Meteorologie und Geophysik, Universität Wien, Germany

Motivation

- For convective-scale data assimilation there is potentially a vast amount of information available from ground-based remote-sensing instruments, various satellites and also human and economic activities e.g. smartphones, weather cameras, renewable energy production.
- The assimilation of such complex observations is non-trivial and requires better understanding of the processes and effects of the data assimilation system.

Goals of this project:
Development of a diagnostic tool to assess observation influence in 3D
Sensitivity studies of different observation types and assimilation settings



TL;DR

- Local Ensemble Transform Kalman Filters (LETKFs) allow us to explicitly calculate the Kalman Gain matrix and by this the contribution of every observation to the analysis field (partial analysis increment (PAI)).
- We propose their use to optimize LETKF systems in particular with respect to satellite data assimilation and vertical localization that constitute significant challenges.

Partial Analysis Increments

- The analysis is a statistical combination of the background state and the observations.
- The influence of the observations on the analysis is determined through the increment.

$$x_a = K\, y_o + (I - KH)x_b$$

$$x_a = x_b + \underbrace{K(y_o - Hx_b)}_{\text{increment}}$$

Variable	Description	Dimension
x_a	Analysis model state vector	n x 1
x_b	Background model state vector	n x 1
K	Kalman Gain	n x p
y_o	Observation vector	p x 1
H	Observation operator	p x n

- The Kalman Gain can be expressed using only available LETKF model output.
- It is not directly computed in the LETKF.

$$K = \frac{1}{k-1}\, X_a Y_a^T R_{loc}^{-1}$$

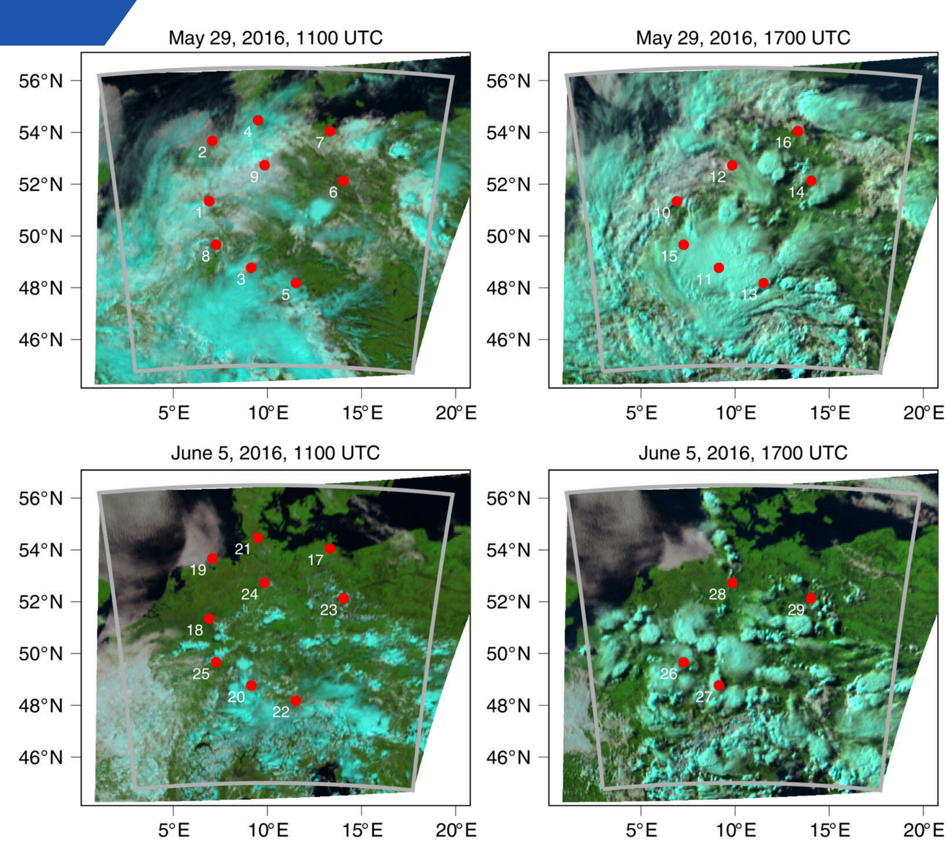
Variable	Description	Dimension
X_a	Analysis perturbation matrix	n x k
Y_a	Model equivalent of X_a	p x k
K	Kalman Gain	n x p
R_{loc}	Observation error covariance matrix localized with Gaspari-Cohn function	p x p
k	Ensemble size	1

- Using any subset of observations, i.e. only certain columns of K and rows of the innovation vector (y_o - Hx_b) allows for computing the partial analysis increments.
- Y_a is not available at every model grid point, however we demonstrate that this only introduces minor errors up to the localization length scale.

Data Assimilation Experiments

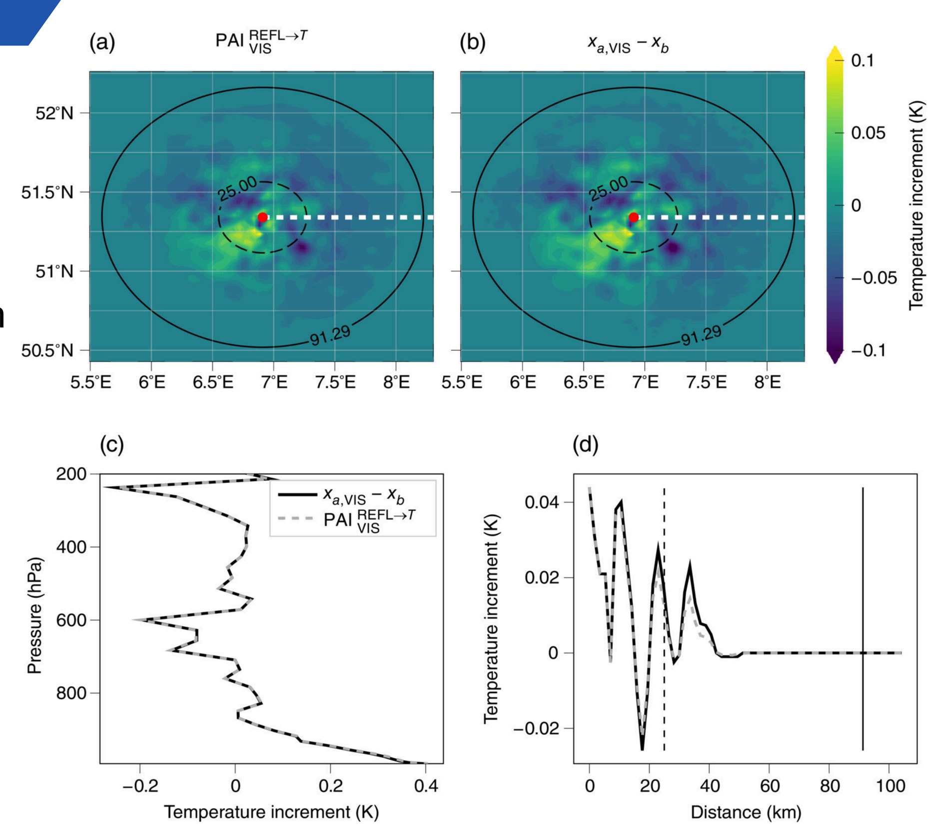
COSMO-KENDA setup:

- Assimilation of radiosonde (RASO) and colocated satellite observations (VIS, 0.6 µm wavelength)
- In total 29 colocated observation locations and ~ 773 radiosonde measurements per variable
- One analysis step, no inflation
- Horizontal localization radius such that every grid point is influenced by only one observation, i.e. multiple assimilation experiments in one model run



Validation of the Method

- Comparison between computed PAI and difference between background and analysis from single obs experiment with satellite observations only
- Horizontal localization length scale 25 km
- No vertical localization
- a) Temperature increment (ΔT), horizontal slice at ~ 500 hPa
  b) Same as a) but increment from LETKF output
  c) Vertical profile at obs location (red dot in a) and b))
  d) Horizontal cut through the domain (grey dotted line in a) and b))



Three Use Cases

Analyzing the influence of observations on different variables

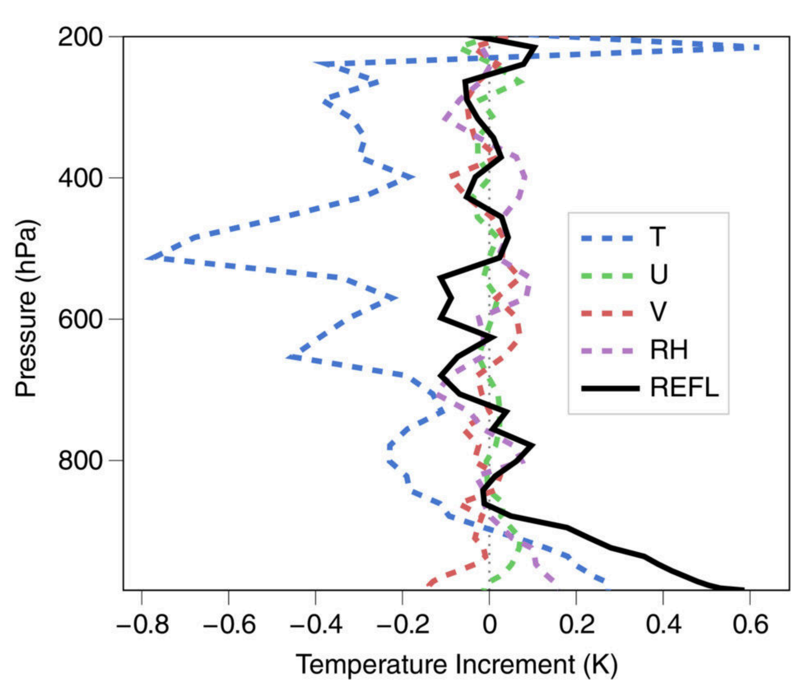


Figure: Vertical profile example of PAI contributions from all observations. The sum equals the total increment.

Model Variable \ Observation	T	U	V	RH	REFL
T	65.5	9.2	11.0	9.6	4.7
U	13.1	58.1	14.6	9.2	5.0
V	14.1	12.9	59.5	9.2	4.3
RH	12.5	9.4	11.5	59.5	7.1
W	28.3	19.6	28.7	15.8	7.6
Specific humidity	52.0	14.4	11.7	13.6	8.3
Specific cloud ice content	28.5	24.2	21.4	17.2	8.7
Cloud water mixing ratio	23.2	15.8	29.4	17.7	13.9

Table: Absolute PAI contributions in % summed over all profiles

Detecting detrimental observation influence

Δe defines the error reduction with respect to radiosonde observations. We consider only the contribution of assimilated satellite observations and compare the error reduction in a single obs experiment (VIS) and a combined experiment (RASO+VIS).

The error is defined as: $e_v = |H(x_v) - y_o|$ with $v \in a, b$

The error reduction is: $\Delta e = e_a - e_b$

$\Delta e < 0$ satellite pulls towards the radiosonde observation

$\Delta e > 0$ satellite pulls away from the radiosonde observation

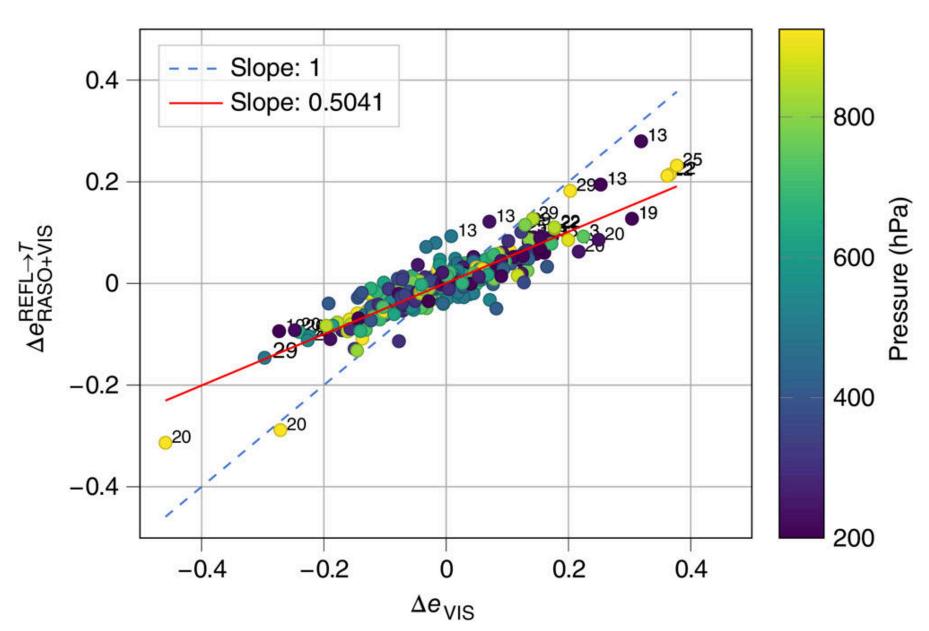
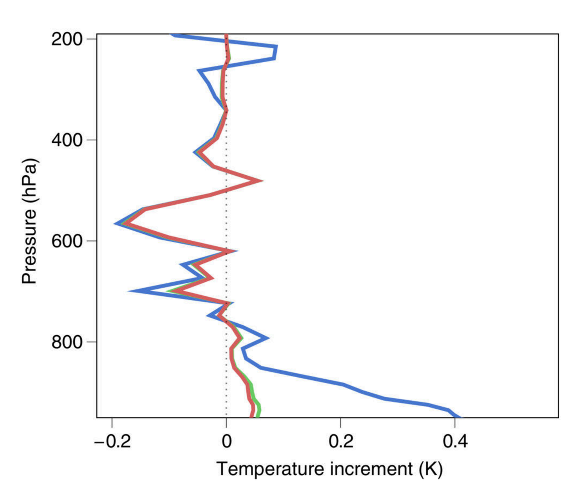


Figure: Δe for temperature, individual dots are associated with individual radiosonde measurements from 29 profiles.

Optimizing localization

The computation of PAI can be used to approximate the influence an observation would have when applying different localization length scales.

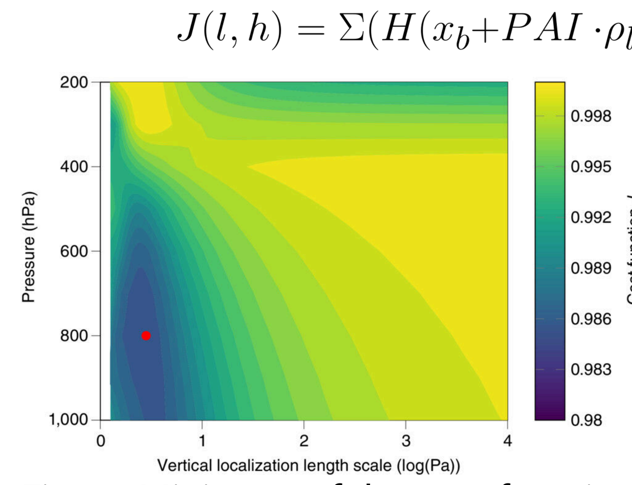


- non-localized
- localized
- retrospectively localized using PAI

Figure: Vertical profile of satellite PAI from three different experiments.

Determine the optimal parameters by iteratively minimizing a cost function of the form:

$$J(l,h) = \Sigma(H(x_b + PAI \cdot \rho_{loc}(l,h)) - y_o)^2$$



$\rho_{loc}(l,h)$ Gaspari-Cohn function

l localization length scale

h centering height

Figure: Minimum of the cost function indicated by red dot. J = 1 indicates cost in non-localized experiment.

Publication

Diefenbach, T., Craig, G., Keil, C., Scheck, L. & Weissmann, M.(2023) Partial analysis increments as diagnostic for LETKF data assimilation systems. Quarterly Journal of the Royal Meteorological Society, 149(752), 740–756. Available from: https://doi.org/10.1002/qj.4419

Contact:
Theresa Diefenbach
theresa.diefenbach@lmu.de