PARTIAL ANALYSIS INCREMENTS AS DIAGNOSTIC FOR LETKF DATA ASSIMILATION SYSTEMS

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

1) Hans Ertel Center for Weather Research / LMU Munich, Germany
2) Deutscher Wetterdienst, Offenbach, Germany
3) Institut für Meteorologie und Geophysik, Universität Wien, Germany

MOTIVATION

For convective-scale data assimilation, a vast amount of information is available from ground-based remote sensing instruments, various satellites, and also from human activities such as smartphones and weather cameras.

The assimilation of such complex observations is non-trivial as in many cases algorithms rely on assumptions that do not reflect reality, for example linear forward operators.

Monitoring the system and understanding which observations are effectively used is crucial for improving data assimilation.

The local ensemble transform Kalman filter (LETKF) enables direct computation of the Kalman Gain, allowing assessment of observation influence.

A diagnostic tool is developed to evaluate 3D observation influence and analyze sensitivity to observation types and assimilation settings.

KEY POINTS

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 diagnose and optimize LETKF systems in particular with respect to satellite data assimilation and vertical localization.

THEORY

The influence of the observations on the analysis mean is determined through the increment

$$\bar{\mathbf{x}}_a = \bar{\mathbf{x}}_b + \mathbf{K}\underbrace{\left[\mathbf{y}_o - \overline{\boldsymbol{H}(\mathbf{x}_b)}\right]}_{\text{increment}}.$$

The Kalman Gain can be expressed using only available LETKF model output

$$\mathbf{K} = \frac{1}{k-1}\mathbf{X}_a\mathbf{Y}_a^T\mathbf{R}_{loc}^{-1}.$$

K is not directly computed in the LETKF.

Using any subset of observations indicated by index j, i.e. only certain columns of **K**, and rows of the innovation vector $[(\mathbf{y}_o - \boldsymbol{H}(\mathbf{x}_b))]$ allows for computing the **partial analysis increments** (PAI)

$$PAI_j = \mathbf{K}_j\left[\mathbf{y}_o - \overline{\boldsymbol{H}(\mathbf{x}_b)}\right]_j.$$

	Description	Dim.
$\bar{\mathbf{x}}_a$	Analysis ensemble mean state vector	n x 1
$\bar{\mathbf{x}}_b$	Background ensemble mean state vector	n x 1
$\mathbf{y}_o$	Observation vector	p x 1
H	Observation operator	p x n
$\mathbf{K}$	Kalman Gain	n x p
k	Number of ensemble members	1
X_a	Analysis perturbation matrix	n x k
Y_a	Model equivalent of $\mathbf{X}_a$	p x k
R_{loc}	Observation error covariance, localized	p x p

EXPERIMENT

Data assimilation set-up

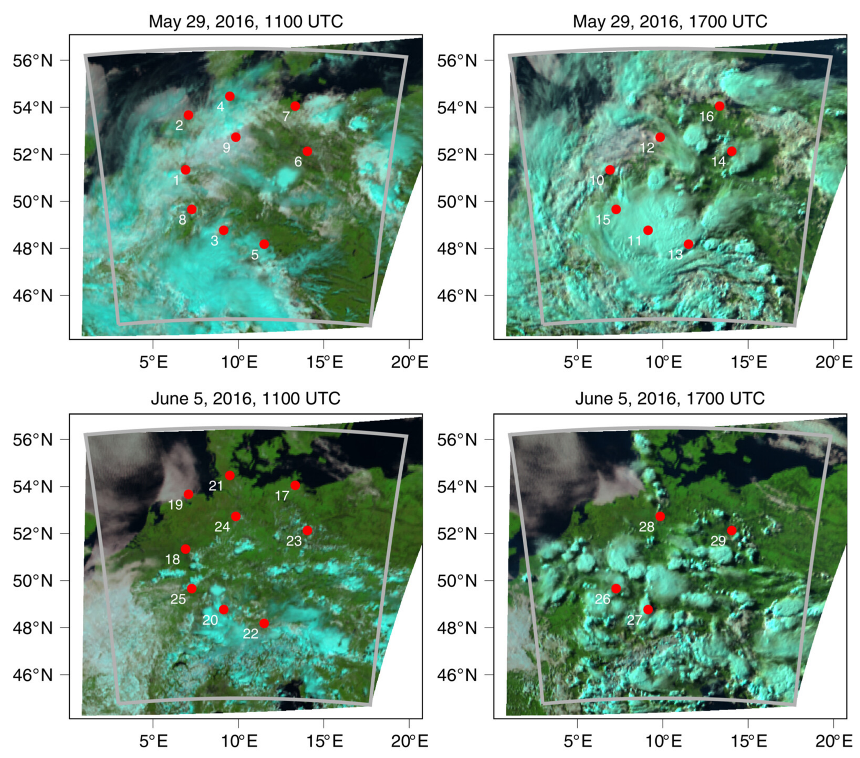
Assimilation of radiosonde (RASO) and colocated satellite observations (VIS, 0.6 µm) using the KENDA system of the German Weather Service (DWD).

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

Different settings for vertical localization



THREE USE CASES

Partial analysis increments in the experiment

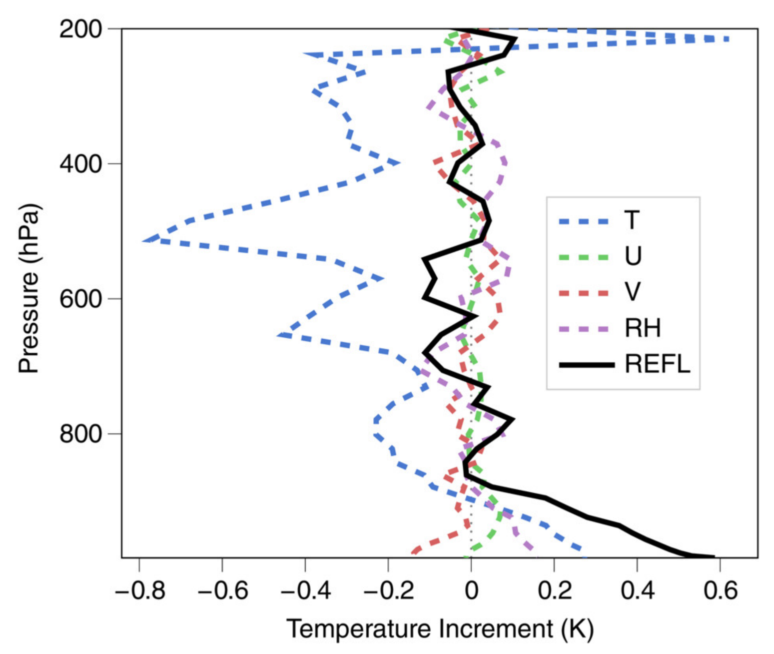


Figure: Vertical profile example of PAI contributions from all assimilated observations. The sum equals the total increment.

Observation / *Model Variable*	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.

The method assesses the 3D influence of specific observations in presence of other assimilated observations.

Beneficial and detrimental impact

$\Delta\mathbf{e}$ represents the error reduction relative to radiosonde observations, considering only the contribution of assimilated satellite data. The error reduction is analyzed for a single-observation experiment (VIS) and a combined experiment (RASO + VIS).

The error is defined as $\mathbf{e}_v = |\mathbf{H}(\bar{\mathbf{x}}_v) - \mathbf{y}_o|$ with $v \in a, b$.

The error reduction is $\Delta\mathbf{e} = \mathbf{e}_a - \mathbf{e}_b$.

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

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

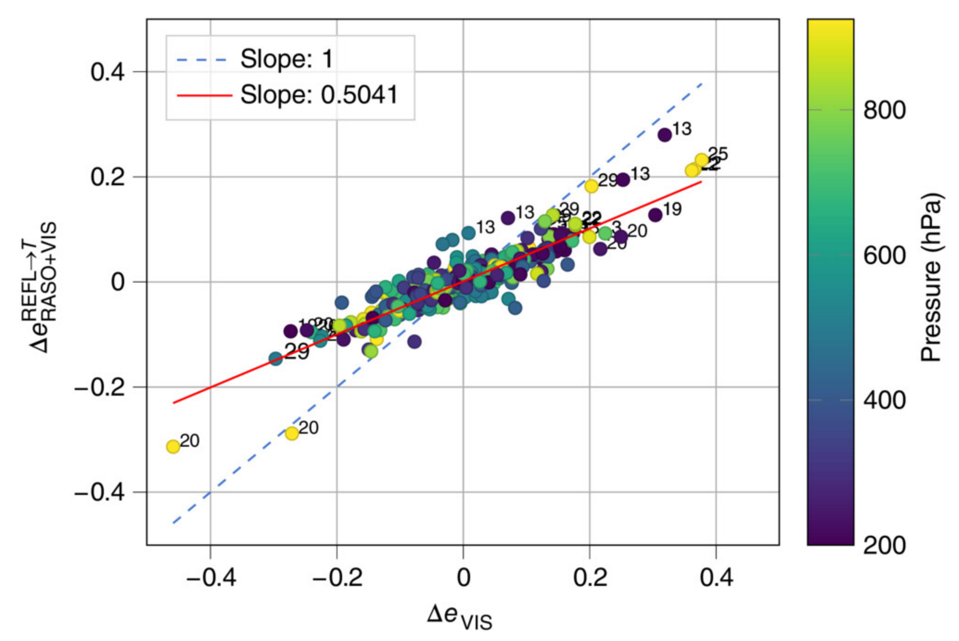


Figure: $\Delta\mathbf{e}$ for temperature, individual dots are associated with individual radiosonde measurements from 29 profiles.

When VIS is assimilated with RASO, its qualitative influence remains unchanged, but its magnitude decreases, with the strongest beneficial influence near the surface.

Optimizing localization

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

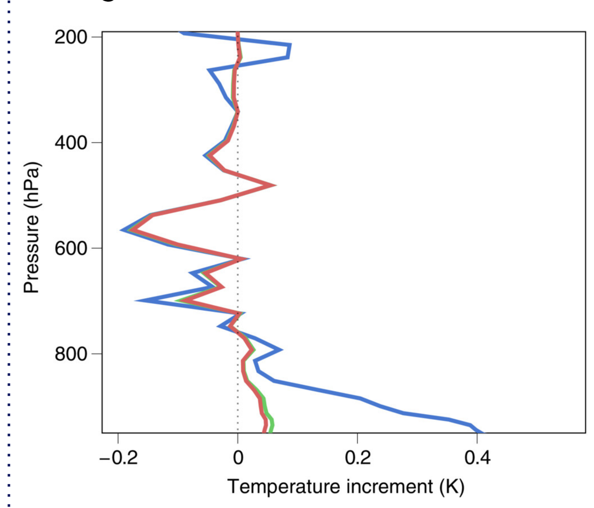
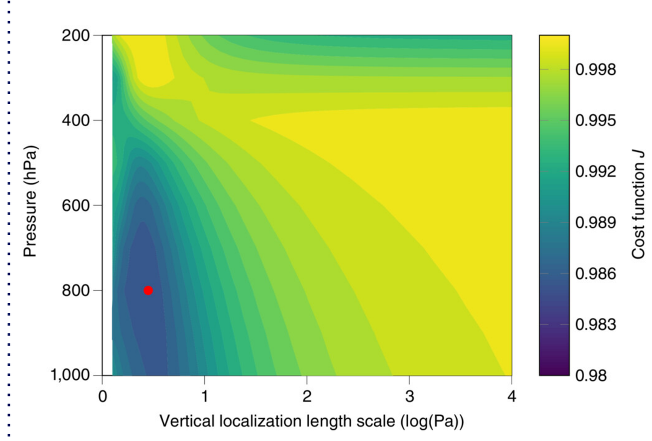


Figure: Vertical profile of satellite PAI from

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

$$J(l,h) = \Sigma\left[\mathbf{H}(\bar{\mathbf{x}}_b + PAI \cdot \boldsymbol{\rho}(l,h)) - \mathbf{y}_o\right]^2.$$

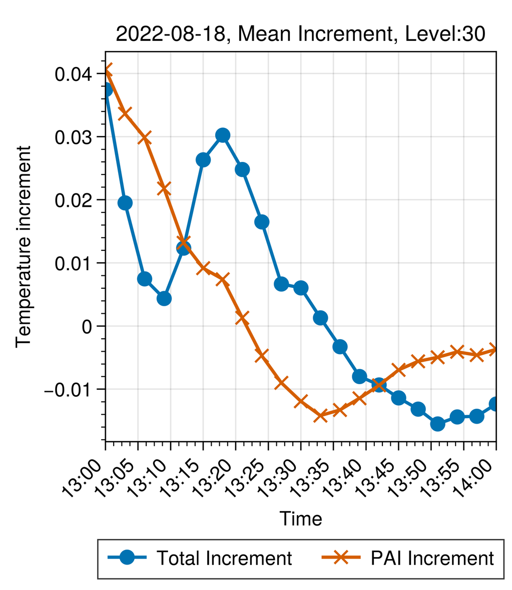


	Description	Dim.
ρ	Gaspari-Cohn function	n x 1
l	localization length scale	1
h	centering height	1

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

Verification with radiosondes shows optimal VIS assimilation when localized around 800 hPa in this metric.

EXTENDING PAI INTO THE FORECAST



Measure the forecast impact of observations using PAI.

$$M(\mathbf{x}_a) - M(\mathbf{x}_b) \approx \mathbf{MK}[\mathbf{y}_o - \boldsymbol{H}(\mathbf{x}_b)]$$

Assume linear model operator **M**.

$$\mathbf{MK} = \frac{1}{k-1}\mathbf{MX}_a\mathbf{Y}_a^T\mathbf{R}_{loc}^{-1}$$

Replace $\mathbf{X}_a$ with $\mathbf{X}_f$ in the computation of **K**, similar as in EFSOI.

$$\approx \frac{1}{k-1}\mathbf{X}_f\mathbf{Y}_a^T\mathbf{R}_{loc}^{-1}$$

Figure: Temperature increment and PAI, absolute mean over horizontal model layer from an experiment with only assimilated visible satellite obs, analysis time 13:00 UTC.

PAI SOFTWARE

A python package is currently under development.

The code is available here:

Features:

Supports reading model data from GRIB or NetCDF files.

Works on the native triangular ICON grid.

Computes the Kalman gain matrix, localization with Gaspari-Cohn function and PAI.

Utilizes localization to make matrix operations computationally feasible.

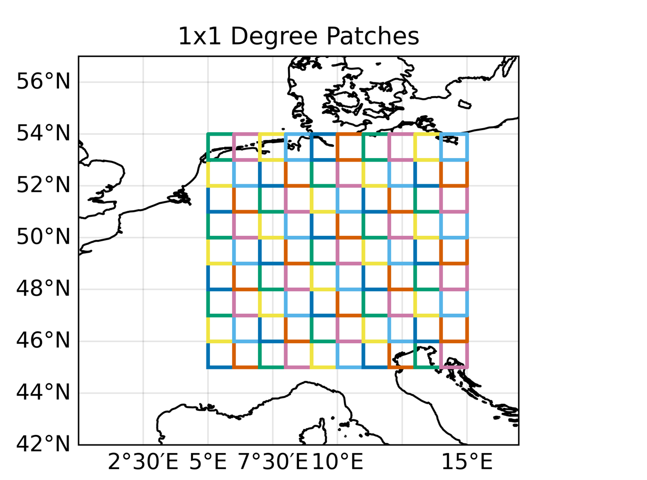


Figure: Parallel computation of PAI is enabled by splitting the model domain in patches.

Diefenbach, T., Craig, G., Keil, C., Scheck, L. & Weissmann, M.(2023): Partial analysis increments as diagnostic for LETKF data assimilation systems.

Contact: theresa.diefenbach@lmu.de