Estimating observation impact in a convective-scale localized ensemble transform Kalman filter

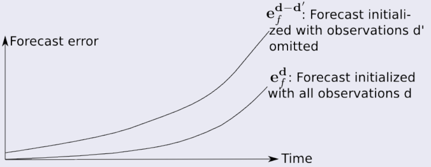
Martin Weissmann and Matthias Sommer

Hans–Ertel–Centre for Weather Research
Data Assimilation Branch
Ludwig–Maximilians–Universität München

November 6, 2014

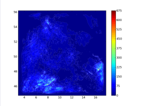
Observation impact: Definition

Data denial impact of observations $\mathbf{d}'$ with respect to complete set of observations $\mathbf{d}$

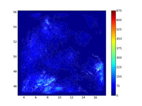


$$J'(\mathbf{d}') = |\mathbf{e}_f^{\mathbf{d}}|^2 - |\mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}|^2$$

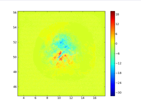
Example



(a) $|\mathbf{e}_f^{\mathbf{d}}|^2$

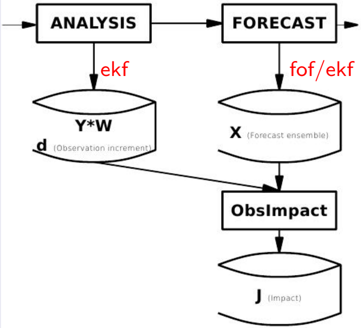


(b) $|\mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}|^2$



(c) $J = \frac{1}{2}\left(|\mathbf{e}_f^{\mathbf{d}}|^2 - |\mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}|^2\right)$

Algorithm



Approximation by ensemble perturbations

LETKF update equation

$$\overline{x}_{aj} = \mathsf{X}_{bj}\widetilde{\mathsf{P}}_a(j)\mathsf{Y}_b{}^\intercal \mathsf{R}^{-1}(j)(\mathbf{y}_o - \overline{\mathbf{y}}_b) + \overline{x}_{bj}$$

Data denial observation impact

$$J'(\mathbf{d}') = |\mathbf{e}_f^{\mathbf{d}}|^2 - |\mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}|^2 = \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}\right)\left(\mathbf{e}_f^{\mathbf{d}} - \mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}\right)$$

Direct approach [Kalnay et al., 2012]

$$\mathbf{e}_f^{\mathbf{d}} - \mathbf{e}_f^{\mathbf{0}} \approx \frac{1}{K-1}\mathsf{Y}_f^{\mathbf{d}}(\mathsf{X}_b\mathsf{W}^{\mathbf{d}})^\intercal \mathsf{R}^{-1}\mathbf{d}$$

$$\Rightarrow J'(\mathbf{d}')$$

$$= \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}\right)\left(\mathbf{e}_f^{\mathbf{d}} - \mathbf{e}_f^{\mathbf{0}} - \left(\mathbf{e}_f^{\mathbf{d}-\mathbf{d}'} - \mathbf{e}_f^{\mathbf{0}}\right)\right)$$

$$\approx \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}\right)\left(\frac{1}{K-1}\mathsf{X}_f^{\mathbf{d}}(\mathsf{X}_b\mathsf{W}^{\mathbf{d}})^\intercal \mathsf{R}^{-1}\mathbf{d}'\right)$$

$$\approx \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{0}}\right)\left(\frac{1}{K-1}\mathsf{X}_f^{\mathbf{d}}(\mathsf{X}_b\mathsf{W}^{\mathbf{d}})^\intercal \mathsf{R}^{-1}\mathbf{d}'\right)$$

Approximation by ensemble perturbations

LETKF update equation

$$\overline{x}_{aj} = \mathsf{X}_{bj}\widetilde{\mathsf{P}}_a(j)\mathsf{Y}_b{}^\intercal\mathsf{R}^{-1}(j)(\mathbf{y}_o - \overline{\mathbf{y}}_b) + \overline{x}_{bj}$$

Data denial observation impact

$$J'(\mathbf{d}') = |\mathbf{e}_f^{\mathbf{d}}|^2 - |\mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}|^2 = \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}\right)\left(\mathbf{e}_f^{\mathbf{d}} - \mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}\right)$$

Direct approach [Kalnay et al., 2012]

$$\mathbf{e}_f^{\mathbf{d}} - \mathbf{e}_f^{\mathbf{0}} \approx \frac{1}{K-1}\mathsf{Y}_f^{\mathbf{d}}(\mathsf{X}_b\mathsf{W}^{\mathbf{d}})^\intercal\mathsf{R}^{-1}\mathbf{d}$$

$$\Rightarrow J'(\mathbf{d}')$$

$$= \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}\right)\left(\mathbf{e}_f^{\mathbf{d}} - \mathbf{e}_f^{\mathbf{0}} - \left(\mathbf{e}_f^{\mathbf{d}-\mathbf{d}'} - \mathbf{e}_f^{\mathbf{0}}\right)\right)$$

$$\approx \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{d}-\mathbf{d}'}\right)\left(\frac{1}{K-1}\mathsf{X}_f^{\mathbf{d}}(\mathsf{X}_b\mathsf{W}^{\mathbf{d}})^\intercal\mathsf{R}^{-1}\mathbf{d}'\right)$$

$$\approx \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{0}}\right)\left(\frac{1}{K-1}\mathsf{X}_f^{\mathbf{d}}(\mathsf{X}_b\mathsf{W}^{\mathbf{d}})^\intercal\mathsf{R}^{-1}\mathbf{d}'\right)$$

Taylor expansion [Sommer and Weissmann, 2015]

$$J'(\mathbf{d}') = J'(\mathbf{0}) + \left.\frac{d}{d\mathbf{d}'}\right|_{\mathbf{d}'=\mathbf{0}} J'(\mathbf{d}') + \mathcal{O}\left(|\mathbf{d}|^2\right)$$

$$= 2\mathbf{e}_f^{\mathbf{d}}\left(\left.\frac{d}{d\mathbf{d}'}\right|_{\mathbf{d}'=\mathbf{d}} \overline{x_f^{\mathbf{d}'}}\right)\mathbf{d}' + \mathcal{O}\left(|\mathbf{d}|^2\right)$$

$$\approx 2\mathbf{e}_f^{\mathbf{d}}\left(\frac{1}{K-1}\mathsf{X}_f^{\mathbf{d}}(\mathsf{X}_b\mathsf{W}^{\mathbf{d}})^\intercal\mathsf{R}^{-1}\mathbf{d}'\right)$$

Verification in observation space

Verification with analysis [Kalnay et al., 2012]

$$\mathbf{e}_f = \overline{\mathbf{x}_f} - \mathbf{x}_a$$

$$J'(\mathbf{d}') \approx 2\mathbf{e}_f^{\mathbf{d}} \left(\frac{1}{K-1} \mathsf{X}_f^{\mathbf{d}} (\mathsf{X}_b \mathsf{W}^{\mathbf{d}})^{\intercal} \mathsf{R}^{-1} \mathbf{d}' \right)$$

$$|\mathbf{e}_f|^2 = \sum_{gridpoints} \frac{1}{2} (\overline{\mathbf{u}_f} - \mathbf{u}_a)^2 + \frac{1}{2} (\overline{\mathbf{v}_f} - \mathbf{v}_a)^2$$

+ Homogeneous in space and time

Verification with observations [Sommer and Weissmann, 2015]

$$\mathbf{e}_f = H(\overline{\mathbf{x}_f}) - \mathbf{y}_o$$

$$J'(\mathbf{d}') \approx 2\mathbf{e}_f^{\mathbf{d}} \left(\frac{1}{K-1} \mathsf{Y}_f^{\mathbf{d}} (\mathsf{X}_b \mathsf{W}^{\mathbf{d}})^{\intercal} \mathsf{R}^{-1} \mathbf{d}' \right)$$

$$|\mathbf{e}_f|^2 = \sum_{observations} \frac{H(\overline{\mathbf{x}_f}) - \mathbf{y}_o}{\sigma}$$

+ Independent of forecast
+ Computationally cheaper

DWD Convective-scale assimilation and forecasting systems

Kilometer-scale Ensemble Data Assimilation (KENDA)

- **L**ocalized **E**nsemble **T**ransform **K**alman **F**ilter for use with COSMO-DE (in development)

Consortium for Small-scale Modelling (COSMO)

- Operational limited-area model of Deutscher Wetterdienst
- Grid point model of non-hydrostatic equations
- Horizontal resolution: 2.8 km; 50 vertical levels

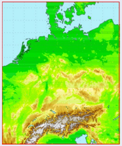
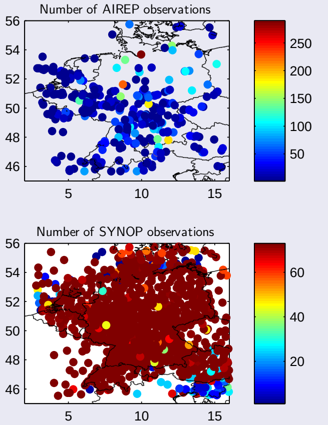


Figure : COSMO-DE domain ($\approx$ 1300 km $\times$ 1200 km)

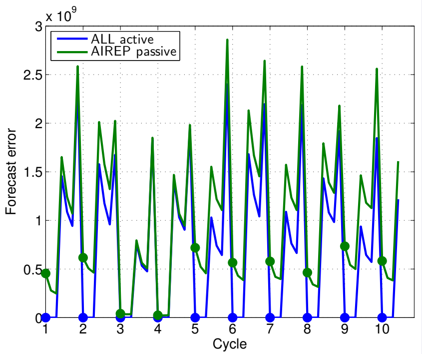
Experimental settings

- Test period: 10 June 2012 12:00 UTC – 13 June 2012 15:00 UTC
- Initialization every 3 h
- Forecast length 6 h
- 40-members ensemble
- Observations used:
  - AIREP (Aircrafts): U, V, T
  - PROF (Wind profiler): U, V
  - SYNOP (Ground stations): U, V, T, RH
  - TEMP (Weather Balloons): U, V, T, RH

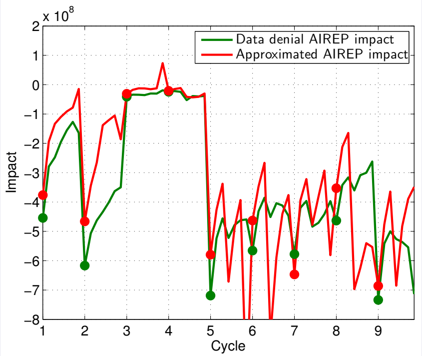
Number of assimilated observations per station



AIREP observations



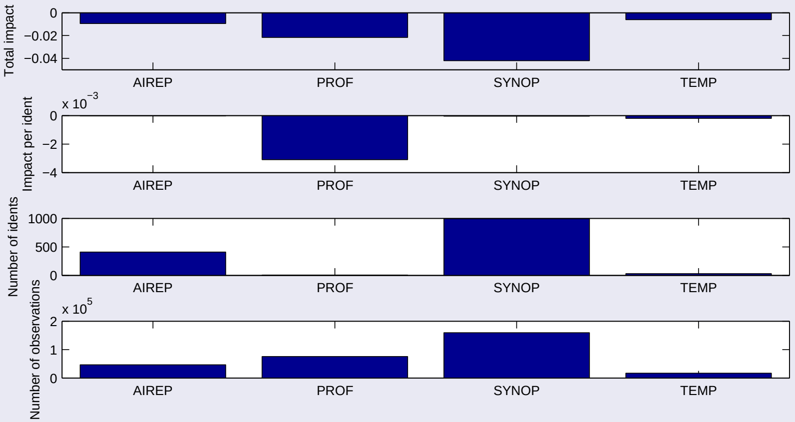
(a) Data denial forecast errors



(b) AIREP impact

- Observation impact = Forecast error difference
- Data denial result well reflected by approximation

Impact per observation type



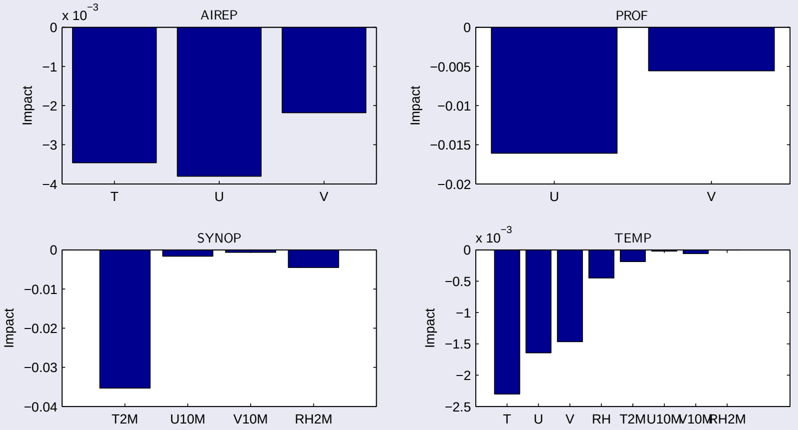
One wind profiler equivals. . .

134 aircrafts

73 ground stations

16 weather balloons

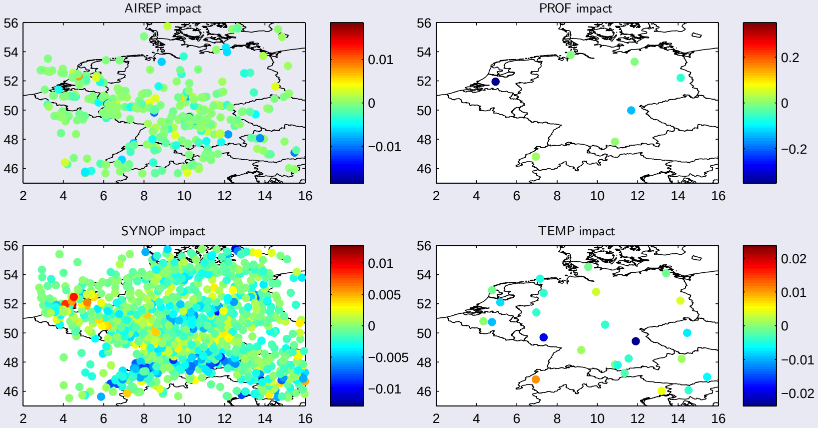
Impact per observed variable



- Large temperature impacts
- Small SYNOP wind impacts
- Anisotropy of wind component impacts

Spatial impact distribution

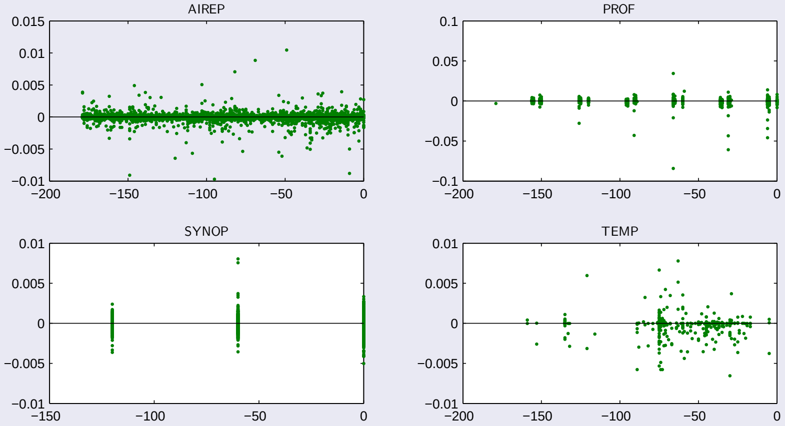
Impact per ident



- Weak specificity of regions with positive / negative impact.

Temporal impact distribution

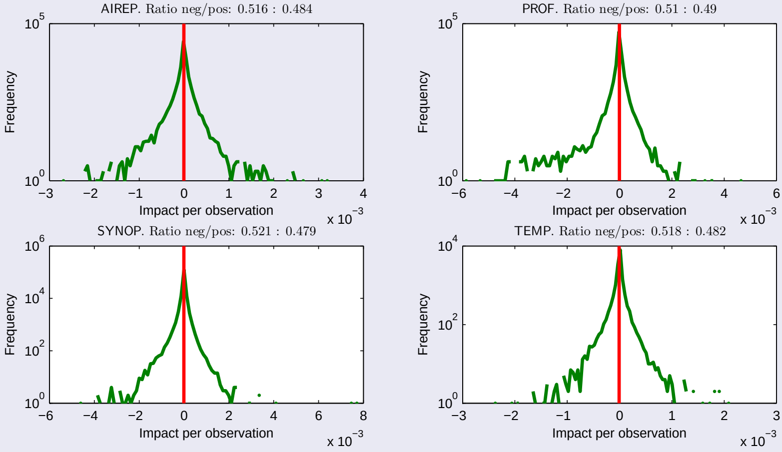
Observation time vs impact



- Temporally uniform distribution

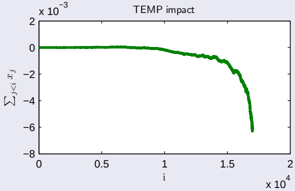
Distribution of impact values

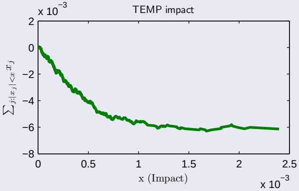
Histogram of individual impact values



- Impact distribution dominated by extreme values
- Ratio of negative to positive values approximately 52:48

Verification with TEMP observations

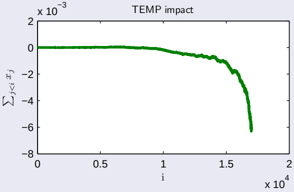


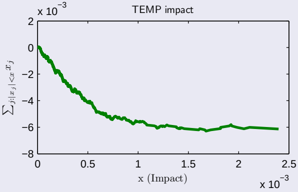


"Sum of i smallest (absolute value) impacts"

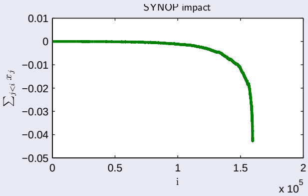
"Sum of impacts whos absolute value is smaller than x"

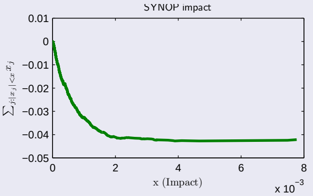
Verification with TEMP observations





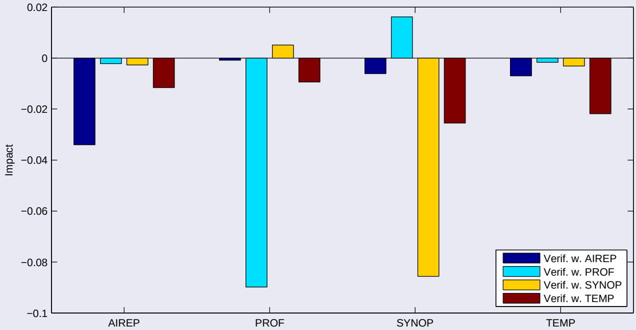
Verification with SYNOP observations





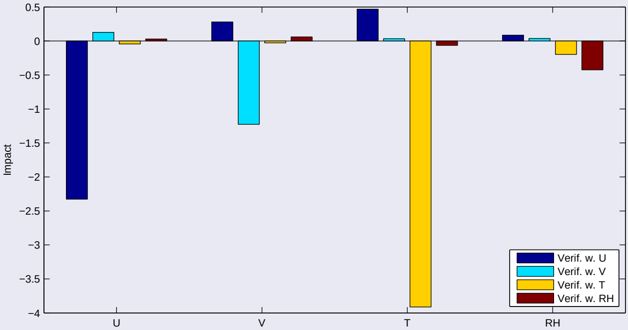
"Sum of *i* smallest (absolute value) impacts" "Sum of impacts whos absolute value is smaller than *x*"

Verification with conventional obstypes



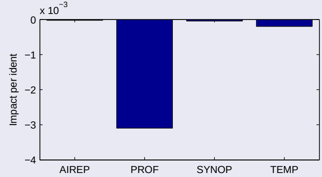
Dependence on verification

Verification with individual variables

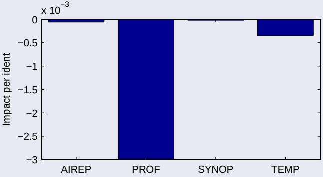


Dependence on verification

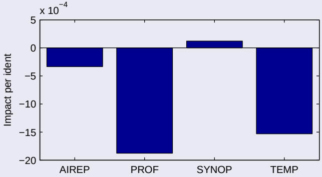
Verification with ALL ACTIVE (no PS)



Verification with ALL ACTIVE AND PS



Verification with ONLY PS



Verification with ONLY RH

Summary

Main progress steps

- Verification in observation space.
- More accurate formula for data denial approximation.
- 3-days experimental period give more robust results.
- Optionally use `ekf`-files for verification.
- Application to VIS/NIR radiances and data from MeteoSwiss.

Outlook

- From mid-2015: PhD working on observation impact.
  - Relative impact of observation at different stages.
  - Impact on unobserved variables.
  - Impact time.
  - Ensemble sensitivity analysis.

Literature

Eugenia Kalnay, Yoichiro Ota, Takemasa Miyoshi, and Junjie Liu. A simpler formulation of forecast sensitivity to observations: application to ensemble Kalman filters. *Tellus A*, 64, 2012. ISSN 1600-0870. URL http://www.tellusa.net/index.php/tellusa/article/view/18462.

Matthias Sommer and Martin Weissmann. Estimating obeservation impact using an observation-based verification metric. *Quarterly Journal of the Royal Meteorological Society*, 2015.