Formulating and Solving the Robust Data Assimilation Problem

Adrian Sandu[1]

with Vishwas Rao[1], Elias D. Nino[1], and Michael Ng[2]

[1]Computational Science Laboratory (CSL)
"Compute the Future!"
Department of Computer Science
Virginia Tech

[2]Hong Kong Baptist University

July 2, 2018

What happens when observations surprise the model?

- Presence of outliers in data is a common occurrence.
- Outliers can be bad data points that negatively impact the analysis quality.

  $\Rightarrow$ Data quality control is essential.

What happens when observations surprise the model?

- Presence of outliers in data is a common occurrence.
- Outliers can be bad data points that negatively impact the analysis quality.

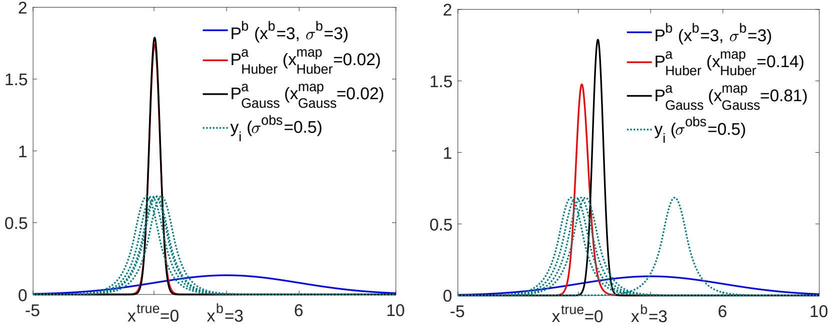
  $\Rightarrow$ Data quality control is essential.
- Outliers can also be data points containing new information that the model is unaware of.
- Data quality control by rejecting observations on the basis of background departure statistics leads to the inability to capture small scales in the analysis {Tavolato and Isaksen, QJRMS 2015}.

What happens when observations surprise the model?

- Presence of outliers in data is a common occurrence.
- Outliers can be bad data points that negatively impact the analysis quality.

  $\Rightarrow$ Data quality control is essential.
- Outliers can also be data points containing new information that the model is unaware of.
- Data quality control by rejecting observations on the basis of background departure statistics leads to the inability to capture small scales in the analysis {Tavolato and Isaksen, QJRMS 2015}.
- Robust data assimilation is needed to properly treat outliers:
  - make the analysis less sensitive to bad information, but
  - without throwing away (all) possibly good information.

Example: robust data assimilation increases resilience to observation outliers

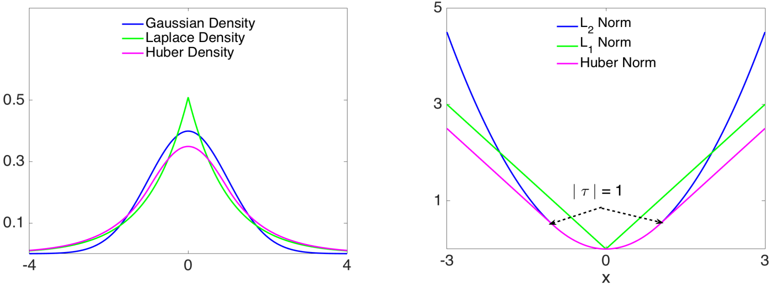


(a) Small observation errors. (b) One observation outlier.

Figure: Example: assimilation of a scalar variable using five measurements. Analytical PDFs are shown.

Robust data assimilation using the Huber norm

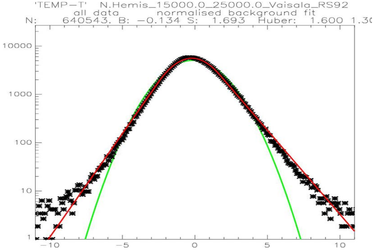
Assumption: $\mathcal{P}(\mathbf{y}|\mathbf{x}) = \begin{cases} \textit{Gaussian} & \text{for small obs. errors,} \\ \textit{Laplace} & \text{for large obs. errors.} \end{cases}$



(a) Gaussian, Laplace, Huber PDFs (b) L_2, L_1, Huber norms

$$\left\|\mathbf{x}\right\|_{\text{HUB}} = \sum_{\ell=1}^{\dim(\mathbf{x})} \mathrm{L}_\tau\left(\mathbf{x}_\ell\right), \quad \mathrm{L}_\tau\left(a\right) := \begin{cases} \frac{1}{2}\, a^2, & \text{for } |a| \le \tau \\ \tau\left(|a| - \frac{1}{2}\tau\right), & \text{otherwise}. \end{cases}$$

The need for a rigorous approach to solve Huber norm data assimilation



Departure statistics for radiosonde temperatures is well described by a Huber distribution {E. Holm, L. Isaksen, C. Tavolato, E. Andersson, ECMWF training course "Variational Quality Control", 2014}.

Current approaches to robust DA:

- *{Andersson and Jarvinen, QJRMS 1999}* 'Variational QC' assumes an observation likelihood convex combination of 'normal' and 'uniform'.
- *{Tavolato and Isaksen, ECMWF letter 22, 2009; QJRMS 2015}* 'Huber 4D-Var' by iterating over incremental 4D-Vars with rescaled observation cost functions.
- *{Roh, Szunyogh, et al MWR 2013}*: 'Huberized EnKF analysis' done by clipping EnKF innovations.

Robust 3D-Var using Huber norm I

- At time t_i traditional 3D-Var minimizes the cost function:

$$\mathcal{J}\left(\mathbf{x}_i\right) = \frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_i^{\mathrm{b}}\|^2_{\mathbf{B}_i^{-1}} + \frac{1}{2}\left\|\mathbf{R}_i^{-1/2}[\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i]\right\|_2^2.$$

- The 3D-Var cost function with Huber norm:

$$\mathcal{J}(\mathbf{x}_i) = \frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_i^{\mathrm{b}}\|^2_{\mathbf{B}_i^{-1}} + \frac{1}{2}\left\|\mathbf{R}_i^{-1/2}[\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i]\right\|_{\mathrm{HUB}}.$$

- The robust 3D-Var problem:

$$\min\ \mathcal{J}(\mathbf{x}_i, \mathbf{z}_i) = \frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_i^{\mathrm{b}}\|^2_{\mathbf{B}_i^{-1}} + \frac{1}{2}\left\|\mathbf{z}_i\right\|_{\mathrm{HUB}}$$

$$\text{subject to} \quad \mathbf{z}_i = \mathbf{R}_i^{-1/2}\left[\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i\right].$$

- The augmented Lagrangian:

$$\mathcal{L}(\mathbf{x}, \mathbf{z}, \boldsymbol{\lambda}, \mu) = \frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_i^{\mathrm{b}}\|^2_{\mathbf{B}_i^{-1}} + \frac{1}{2}\|\mathbf{z}_i\|_{\mathrm{HUB}} - \frac{1}{2\mu}\|\boldsymbol{\lambda}\|_2^2 + \frac{\mu}{2}\left\|\mathbf{R}_i^{-1/2}[\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i] - \mathbf{z}_i - \frac{\boldsymbol{\lambda}_i}{\mu}\right\|_2^2.$$

Solving robust 3D-Var by ADMM

Perform outer iterations for $k = 0, 1, \dots$:

1. Fix $\mathbf{z}_i^{\{k\}}$, $\boldsymbol{\lambda}^{\{k\}}$, $\mu^{\{k\}}$; solve L$_2$-3D-Var with updated observations:

$$\mathbf{x}_i^{\{k+1\}} \quad := \quad \arg\min_{\mathbf{x}} \; \frac{1}{2} \left\| \mathbf{x} - \mathbf{x}_i^{\mathrm{b}} \right\|_{\mathbf{B}_i^{-1}}^2 + \frac{\mu^{\{k\}}}{2} \left\| \mathbf{R}_i^{-1/2} [\mathcal{H}(\mathbf{x}) - \mathbf{y}_i] - \mathbf{z}_i^{\{k\}} - \frac{\boldsymbol{\lambda}_i^{\{k\}}}{\mu^{\{k\}}} \right\|_2^2 .$$

2. Fix $\mathbf{x}_i^{\{k+1\}}$, $\boldsymbol{\lambda}^{\{k\}}$, $\mu^{\{k\}}$, and solve via shrinkage procedure:

$$\mathbf{d}_i^{\{k+1\}} := \mathbf{R}_i^{-1/2} \left[\mathcal{H}(\mathbf{x}_i^{\{k+1\}}) - \mathbf{y}_i \right];$$

$$\mathbf{z}_i^{\{k+1\}} = \mathrm{HuberShrinkage}(\mu^{\{k\}}; \mathbf{d}_i^{\{k+1\}}; \boldsymbol{\lambda}^{\{k\}}).$$

3. Update $\boldsymbol{\lambda}_i$:

$$\boldsymbol{\lambda}_i^{\{k+1\}} := \boldsymbol{\lambda}_i^{\{k\}} - \mathbf{d}_i^{\{k+1\}} + \mathbf{z}_i^{\{k+1\}} .$$

4. Update μ:

$$\mu^{\{k+1\}} := \rho\, \mu^{\{k\}}, \qquad \rho > 1.$$

Repeat outer iteration.

Solving robust 3D-Var by half-quadratic minimization

Perform outer iterations for $k = 0, 1, \ldots$:

1. Compute $\mathbf{u}$ by component-wise regularizationn:

$$\mathbf{u}_i^{\{k+1\}} = \sigma\left(\mathbf{R}_i^{-1/2}\,\left[\mathcal{H}(\mathbf{x}^{\{k\}}) - \mathbf{y}_i\right]\right),$$

$$\sigma(a) = \begin{cases} 1, & |a| \le \tau; \\ \tau/|a|, & |a| > \tau. \end{cases}$$

2. Calculate $\mathbf{x}_i^{\{k+1\}}$ via L$_2$-3D-Var with scaled obs. covariance:

$$\begin{aligned} \mathbf{x}_i^{\{k+1\}} &= \arg\min_{\mathbf{x}_i}\ \frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_i^{\mathrm{b}}\|^2_{\mathbf{B}_i^{-1}} + \frac{1}{2}\left\|\mathbf{R}_i^{\{k\}-1/2}[\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i]\right\|_2^2 \\ \mathbf{R}_i^{\{k\}} &:= \mathbf{R}_i^{1/2}\,\mathrm{diag}\left(2/\mathbf{u}_i^{\{k+1\}}\right)\,\mathbf{R}_i^{1/2}. \end{aligned}$$

Repeat outer iteration

Comment: Another robust option is to replace $\|\cdot\|_{\mathrm{HUB}}$ by $\|\cdot\|_{\mathrm{L}_1}$. Different computational algorithm.

Robust strong-constraint 4D-Var data assimilation

- The robust strong-constraint 4D-Var problem:

$$\min_{\mathbf{x}_0} \mathcal{J}(\mathbf{x}_0, \mathbf{z}) := \frac{1}{2}\|\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}\|^2_{\mathbf{B}_0^{-1}} + \frac{1}{2}\sum_{i=1}^{N} \|\mathbf{z}_i\|_{\mathrm{HUB}}$$

$$\text{subject to: } \mathbf{z}_i = \mathbf{R}_i^{-1/2}[\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i], \quad i = 1, 2, \cdots, N,$$

$$\mathbf{x}_i = \mathcal{M}_{i-1,i}(\mathbf{x}_{i-1}), \quad i = 1, 2, \cdots, N.$$

- Can be solved by either approach:
  - ADMM, or
  - half-quadratic algorithms.
- Each inner iteration requires the solution of a (modified) new L_2-4D-Var problem.

Traditional EnKF data assimilation

- Ensemble space at time t_i:

$$\mathbf{x}_i = \overline{\mathbf{x}}_i^{\rm b} + \mathbf{X}_i^{\rm b}\, w_i, \quad \mathbf{X}_i^{\rm b} \in \mathbb{R}^{\rm N_{var} \times N_{ens}}, \quad w_i \in \mathbb{R}^{\rm N_{ens} \times 1}.$$

- Traditional L_2-EnKF: analysis mean weights {Hunt et al, 2007}:

$$\overline{w}_i^{\rm a} := \arg\min_{w} \quad ({\rm N_{ens}} - 1)\, \|w\|_2^2 + \left\| \mathcal{H}(\overline{\mathbf{x}}_i^{\rm b} + \mathbf{X}_i^{\rm b}\, w) - \mathbf{y}_i \right\|_{\mathbf{R}_i^{-1}}^2.$$

- EnSRF {LETKF, Hunt et al, 2007} analysis ensemble weights:

$$\mathbf{S}_i \quad := \quad \left(\mathbf{I} + \frac{1}{{\rm N_{ens}} - 1}\, \left(\mathbf{Y}_i^{\rm b}\right)^T\, \mathbf{R}_i^{-1}\, \mathbf{Y}_i^{\rm b} \right)^{-1} = \mathbf{W}_i\, \mathbf{W}_i^T,$$

$$w_i^{{\rm a}\langle \ell \rangle} \quad = \quad \overline{w}_i^{\rm a} + \mathbf{W}_i(:,\ell), \qquad \mathbf{x}_i^{{\rm a}\langle \ell \rangle} = \overline{\mathbf{x}}_i^{\rm b} + \mathbf{X}_i^{\rm b}\, w_i^{{\rm a}\langle \ell \rangle}.$$

- Perturbed observations EnKF analysis ensemble weights:

$$w_i^{{\rm a}\langle \ell \rangle} \approx \mathbf{S}_i \left(w_i^{{\rm b}\langle \ell \rangle} + \frac{1}{{\rm N_{ens}} - 1} (\mathbf{Y}_i^{\rm b})^T\, \mathbf{R}_i^{-1} \left(\mathbf{y}_i^{\langle \ell \rangle} - \mathcal{H}(\overline{\mathbf{x}}_i^{\rm b}) \right) \right).$$

Robust EnKF data assimilation

Huber EnKF optimization problem in ensemble space:

$$\min_{w,\mathbf{z}_i} \; (\mathrm{N_{ens}} - 1)\,\|w\|_2^2 + \|\mathbf{z}_i\|_{\mathrm{HUB}} \quad \text{s.t.} \quad \mathbf{z}_i = \mathbf{R}_i^{-1/2}\left[\mathcal{H}(\overline{\mathbf{x}}_i^{\mathrm{b}} + \mathbf{X}_i^{\mathrm{b}}\,w) - \mathbf{y}_i\right].$$

1. Calculate (component-wise):

$$\mathbf{u}_i^{\{k+1\}} = \sigma\left(\mathbf{R}_i^{-1/2}\left[\mathcal{H}(\overline{\mathbf{x}}_i^{\mathrm{b}} + \mathbf{X}_i^{\mathrm{b}}\,\overline{w}_i^{\{k\}}) - \mathbf{y}_i\right]\right).$$

2. Apply EnSRF with modified observation covariance:

$$\overline{w}_i^{\{k+1\}} = \arg\min_{w} \; (\mathrm{N_{ens}} - 1)\,\|w\|_2^2 + \|\mathcal{H}(\overline{\mathbf{x}}_i^{\mathrm{b}} + \mathbf{X}_i^{\mathrm{b}}\,w) - \mathbf{y}_i\|^2_{\mathbf{R}_i^{\{k+1\}\,-1}}$$

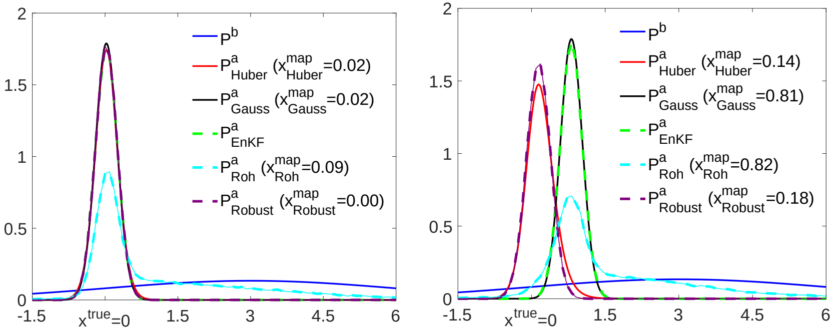
$$\mathbf{R}_i^{\{k+1\}\,-1} = \mathbf{R}_i^{-1/2}\,\mathrm{diag}\left(\mathbf{u}^{\{k\}}/2\right)\,\mathbf{R}_i^{-1/2}.$$

3. Satisfactory convergence after M iterations. The analysis:

$$\overline{w}_i^{\mathrm{a}} = \overline{w}_i^{\{M\}}, \qquad w_i^{\mathrm{a}\langle\ell\rangle} = \overline{w}_i^{\mathrm{a}} + \mathbf{W}_i^{\{M\}}(:,\ell),$$

$$\mathbf{S}_i^{\{M\}} := \left(\mathbf{I} + \frac{1}{\mathrm{N_{ens}} - 1}\,\left(\mathbf{Y}_i^{\mathrm{b}}\right)^T\,\mathbf{R}_i^{\{M\}\,-1}\,\mathbf{Y}_i^{\mathrm{b}}\right)^{-1} = \mathbf{W}_i^{\{M\}}\,\mathbf{W}_i^{\{M\}}.$$

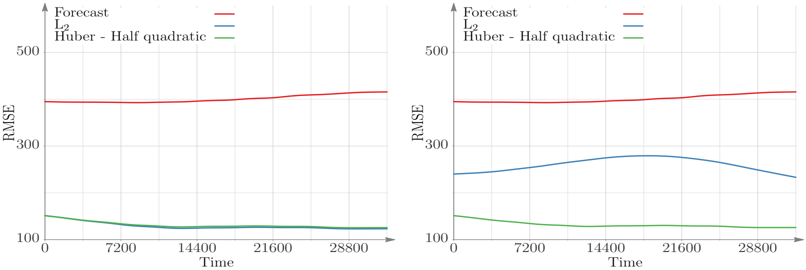
Robust EnKF results for the academic test problem



(c) Small observation errors.　　(d) One observation outlier.

Figure: Example: assimilation of a scalar variable using five measurements. Shown are analyses obtained by various EnKF algorithms.

Results with Shallow Water Equations on the sphere I

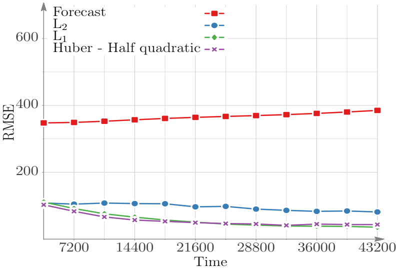


(a) Observations with small random errors.

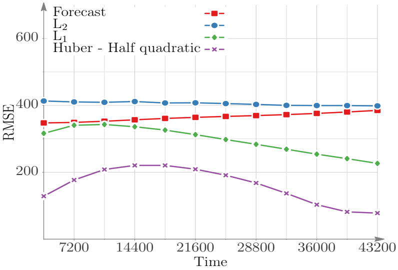
(b) Observations with outliers.

Figure: 4D-Var results for the shallow water model on the sphere. The Huber threshold is $\tau = 2$.

Results with Shallow Water Equations on the sphere II



(a) Observations with small random errors.



(b) Observations with outliers.

Figure: LETKF results for the shallow water model on the sphere. The Huber threshold is $\tau = 1$.

Summary

Robust data assimilation:

1. Rigorous framework for analyses using Huber, L_1 norms on observation errors
2. Applied to robustify: 3D-Var, 4D-Var, EnKF

References:

1. V. Rao, A. Sandu, M. Ng, and E. Nino-Ruiz: "Robust data assimilation using L_1 and Huber norms", SIAM Journal on Scientific Computing, Vol. 39, Issue 3, pp. B548–B570, 2017. https://doi.org/10.1137/15M1045910
2. Our CSL eprints library: http://csl.cs.vt.edu