A Rigorous Framework for Robust Data Assimilation

Adrian Sandu[1]

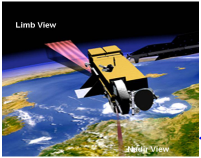
with Vishwas Rao[1], Elias D. Nino[1], and Michael Ng[2]

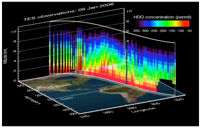
[1]Computational Science Laboratory (CSL)
Department of Computer Science
Virginia Tech
[2]Hong Kong Baptist University

May 25, 2016

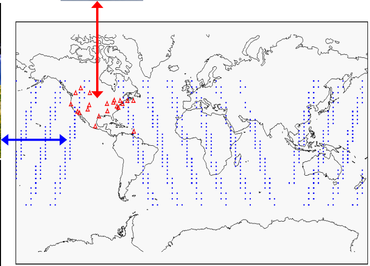
Data assimilation of satellite observations in global CTM improves global ozone predictions

TES ozone column observations (NASA EOS Aura)





IONS-6 ozonesonde data for validation



Improved global ozone prediction, 1-15 Aug. 2007

{Singh, Sandu, et al, GMD 2010}

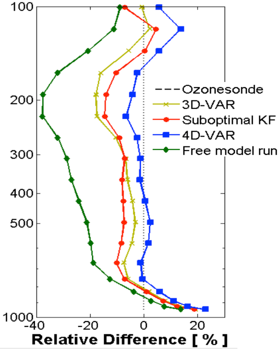
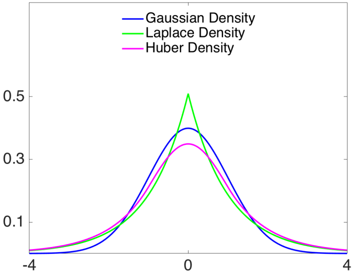


Figure: Past work: 3D/4D-Var systems for GEOS-Chem (NASA/Harvard), CMAQ (Environmental Protection Agency), STEM, partially WRF-Chem.

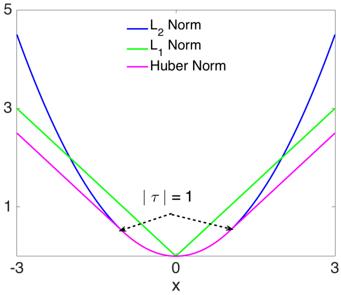
What happens when observations surprise the model?

- Presence of outliers in data is a common occurrence.
- These outliers negatively affect the quality of the solution. Quality control becomes an important step.
- 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 overcome this issue.

Robust data assimilation



(a) Gaussian, Laplace, and Huber densities



(b) L_2, L_1 and Huber norms

- Huber Norm :

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

Motivation for robust data assimilation: resilience to observation outliers

1. Consider the following case:
   - Background: $\mathbf{x}^{\rm b} = 10$, $\mathbf{B} = 1$.
   - Observations: $\mathbf{y}_1 = 11$, $\mathbf{y}_2 = 12$, $\mathbf{R} = 1$.
   - Assuming Gaussian observation error the analysis reads:

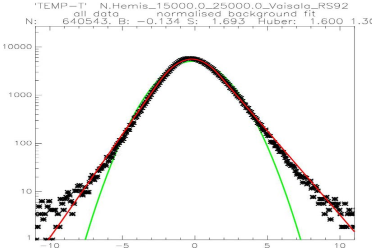
$$\mathbf{x}^{\rm a} = \arg\min \ \|\mathbf{x} - 10\|^2_{\mathbf{B}^{-1}} + \|\mathbf{x} - 11\|^2_{\mathbf{R}^{-1}} + \|\mathbf{x} - 12\|^2_{\mathbf{R}^{-1}}$$
$$\Rightarrow \quad \mathbf{x}^{\rm a} = 11.$$

2. With observation outlier: $\mathbf{y}_1 = 11$ and $\mathbf{y}_2 = 20$:
   Gaussian error assumption $\Rightarrow \quad \mathbf{x}^{\rm a} = 13.66$
3. Assuming :
   - small error distribution is Gaussian,
   - large error distribution is Laplace,

   leads to DA formulation in Huber norm (here $\tau = 2$):

$$\arg\min \ \|\mathbf{x} - 10\|^2_{\mathbf{B}_i^{-1}} + \left\|\mathbf{R}^{-1/2}(\mathbf{x} - 11)\right\|_{\rm HUB} + \left\|\mathbf{R}^{-1/2}(\mathbf{x} - 20)\right\|_{\rm HUB}$$
$$\Rightarrow \quad \mathbf{x}^{\rm a} = 11.37.$$

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



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

Current approaches *only approximately* solve the Huber norm DA:

- {Tavolato and Isaksen, ECMWF letter 22, 2009; QJRMS 2015} “Huberize norm” in 4D-Var: first adjust error covariances of observations based on their background departure statistics, then solve traditional 4D-Var problem.
- {Roh, Szunyogh, et al MWR 2013}: “Huberized analysis” done by clipping EnKF innovations.

Robust 3D-Var using Huber norm I

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

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

- 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:

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

Solving robust 3D-Var by ADMM I

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

1. Fix $\mathbf{z}_i^{\{k\}}$, $\boldsymbol{\lambda}^{\{k\}}$, $\mu^{\{k\}}$, and solve a new L_2-3D-Var problem:

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

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

$$\mathbf{z}_i^{\{k+1\}} := \arg\min_{\mathbf{z}} \quad \|\mathbf{z}\|_{\mathrm{HUB}} + \frac{\mu^{\{k\}}}{2} \left\|\mathbf{d}_i^{\{k+1\}} - \mathbf{z} - \frac{\boldsymbol{\lambda}_i^{\{k\}}}{\mu^{\{k\}}}\right\|_2^2,$$

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

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

Solving robust 3D-Var by ADMM II

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:

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

Algorithm 1 Huber_Shrinkage

procedure $\mathbf{z}$=HuberShrinkage(μ; $\mathbf{d}$; $\boldsymbol{\lambda}$)
    Input: $[\mu \in \mathbb{R};\ \mathbf{d} \in \mathbb{R}^m;\ \boldsymbol{\lambda} \in \mathbb{R}^m]$
    Output: $[\mathbf{z} \in \mathbb{R}^m]$
    for $\ell = 1, 2, \ldots, m$ **do**
        if $|\mathbf{d}_\ell| \geq \tau$ **then**
            $\mathbf{z}_\ell = \max\left\{\|\mathbf{d} - \boldsymbol{\lambda}/\mu\|_2 - 1/\mu, 0\right\} \cdot \dfrac{\mathbf{d}_\ell - \boldsymbol{\lambda}_\ell/\mu}{\|\mathbf{d} - \boldsymbol{\lambda}/\mu\|_2}.$
        else
            $\mathbf{z}_\ell = \dfrac{\mu}{1+\mu}(\mathbf{d}_\ell - \boldsymbol{\lambda}_\ell/\mu)$
        end if
    end for
end procedure

Solving robust 3D-Var by half-quadratic minimization

- Perform outer iterations for $k = 1, 2, \ldots$:
  1. Compute $\mathbf{u}$ by component-wise application of the regularization function:
  $$\mathbf{u}_i^{\{k+1\}} = \sigma\left(\mathbf{R}_i^{-1/2}\left[\mathcal{H}(\mathbf{x}^{\{k\}}) - \mathbf{y}_i\right]\right),$$
  $$\sigma(a) = \begin{cases} 1, & |a| \le \tau; \\ \tau/|a|, & |a| > \tau. \end{cases}$$
  2. Calculate $\mathbf{x}_i^{\{k+1\}}$ by solving a new L_2-3D-Var problem:
  $$\begin{aligned} \mathbf{x}_i^{\{k+1\}} &= \arg\min_{\mathbf{x}_i} \ \frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_i^{\mathrm{b}}\|^2_{\mathbf{B}_i^{-1}} \\ &+\frac{1}{2}\left(\mathbf{R}_i^{-1/2}\left[\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i\right]\right)^T \mathrm{diag}\left(\frac{\mathbf{u}_i^{\{k+1\}}}{2}\right)\left(\mathbf{R}_i^{-1/2}\left[\mathcal{H}(\mathbf{x}_i) - \right.\right. \end{aligned}$$
- Repeat outer iteration

Comment: Another robust option is to replace $\|\cdot\|_{\mathrm{HUB}}$ by $\|\cdot\|_{L_1}$.
Different computational algorithm, see paper.

Robust strong-constraint 4D-Var data assimilation

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

$$\begin{aligned}
\min_{\mathbf{x}_0} \ \mathcal{J}(\mathbf{x}_0, \mathbf{z}) \ &:= \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}\|^2_{\mathbf{B}_0^{-1}} + \frac{1}{2} \sum_{i=1}^{N} \|\mathbf{z}_i\|_{\mathrm{HUB}} \\
\text{subject to:} \ \mathbf{z}_i \ &= \mathbf{R}_i^{-1/2} [\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i], \quad i = 1, 2, \cdots, N, \\
\mathbf{x}_i \ &= \mathcal{M}_{i-1,i}(\mathbf{x}_{i-1}), \quad i = 1, 2, \cdots, N.
\end{aligned}$$

- 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:

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

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

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

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

$$\begin{aligned} \mathbf{S}_i &:= \left((\mathrm{N_{ens}} - 1)\, \mathbf{I} + \left(\mathbf{Y}_i^{\mathrm{b}}\right)^T \mathbf{R}_i^{-1}\, \mathbf{Y}_i^{\mathrm{b}} \right)^{-1} = (\mathrm{N_{ens}} - 1)^{-1}\, \mathbf{W}_i\, \mathbf{W}_i^T, \\ w_i^{\mathrm{a}\langle \ell \rangle} &= \overline{w}_i^{\mathrm{a}} + \mathbf{W}_i(:, \ell), \qquad \mathbf{x}_i^{\mathrm{a}\langle \ell \rangle} = \overline{\mathbf{x}}_i^{\mathrm{b}} + \mathbf{X}_i^{\mathrm{b}}\, w_i^{\mathrm{a}\langle \ell \rangle}. \end{aligned}$$

- Perturbed observations EnKF analysis ensemble weights:

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

Robust EnKF data assimilation I

1. The Huber EnKF optimization problem in ensemble space:

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

2. The problem is solved iteratively by half-quadratic minimization. Let the solution at iteration $\{k\}$ be $w_i^{\{k\}}, \mathbf{u}_i^{\{k\}}$. We proceed as follows.
3. Step 1: calculate (component-wise)

$$\mathbf{u}_i^{\{k+1\}} = \sigma\left(\mathbf{R}_i^{-1/2}\left[\mathcal{H}\left(\overline{\mathbf{x}}_i^{\mathrm{b}} + \mathbf{X}_i^{\mathrm{b}}\, w_i^{\{k\}}\right) - \mathbf{y}_i\right]\right),$$
$$\sigma(a) = \begin{cases} 1, & |a| \le \tau; \\ \tau/|a|, & |a| > \tau. \end{cases}$$

Robust EnKF data assimilation II

4. Step 2: update the state $\mathbf{x}_i^{\{k+1\}} = \overline{\mathbf{x}}_i^{\mathrm{b}} + \mathbf{X}_i^{\mathrm{b}}\, w_i^{\{k+1\}}$:

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

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

Apply again EnSRF with modified observation covariance matrix.

5. Satisfactory convergence after M iterations. The analysis:

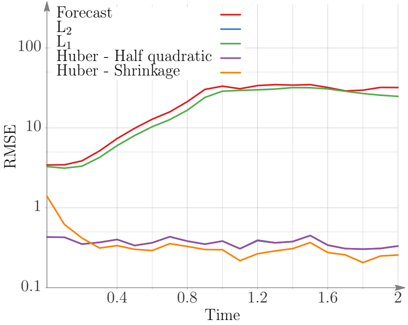
$$\overline{w}_i^{\mathrm{a}} = w_i^{\{M\}}$$

$$\mathbf{S}_i^{\{M\}} := \left((\mathrm{N_{ens}} - 1)\mathbf{I} + \left(\mathbf{Y}_i^{\mathrm{b}}\right)^T \mathbf{R}_i^{-1/2}\, \mathrm{diag}\left(\mathbf{u}^{\{M\}}/2\right)\, \mathbf{R}_i^{-1/2}\, \mathbf{Y}_i^{\mathrm{b}} \right)^{-1}$$

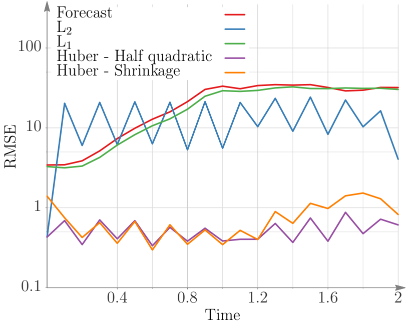
$$= (\mathrm{N_{ens}} - 1)^{-1}\, \mathbf{W}_i^{\{M\}}\, \mathbf{W}_i^{\{M\}\,T},$$

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

Numerical results with the Lorenz-96 model



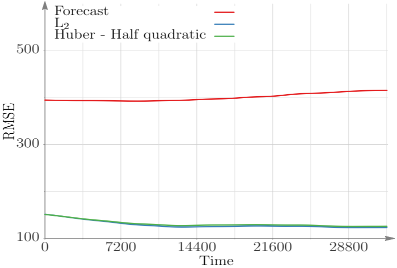
(a) Observations with small random errors



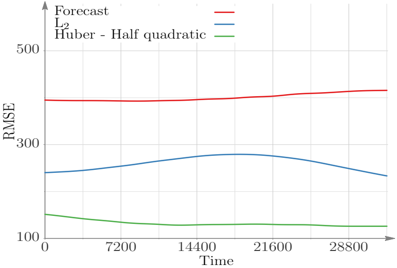
(b) Observations with outliers

Figure: 3D-Var results for Lorenz-96. The frequency of observations is 0.1 time units. Erroneous observations occur every 0.2 time units. The Huber norm uses $\tau = 1$.

Results with Shallow Water Equations on the sphere I



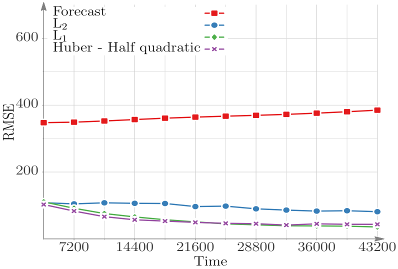
(a) Observations with small random errors.



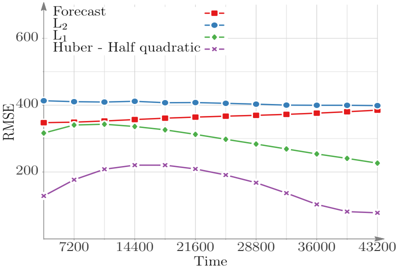
(b) Observations with outliers.

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 data errors
2. Applied to 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”. Under review.
2. Technical report: `https://arxiv.org/abs/1511.01593`
3. Our CSL eprints library: `http://csl.cs.vt.edu`