Downscaling, Data Fusion, and Data Assimilation in Hydro-meteorology:

A variational Framework based on Regularized Inverse Estimation in the Derivative Space

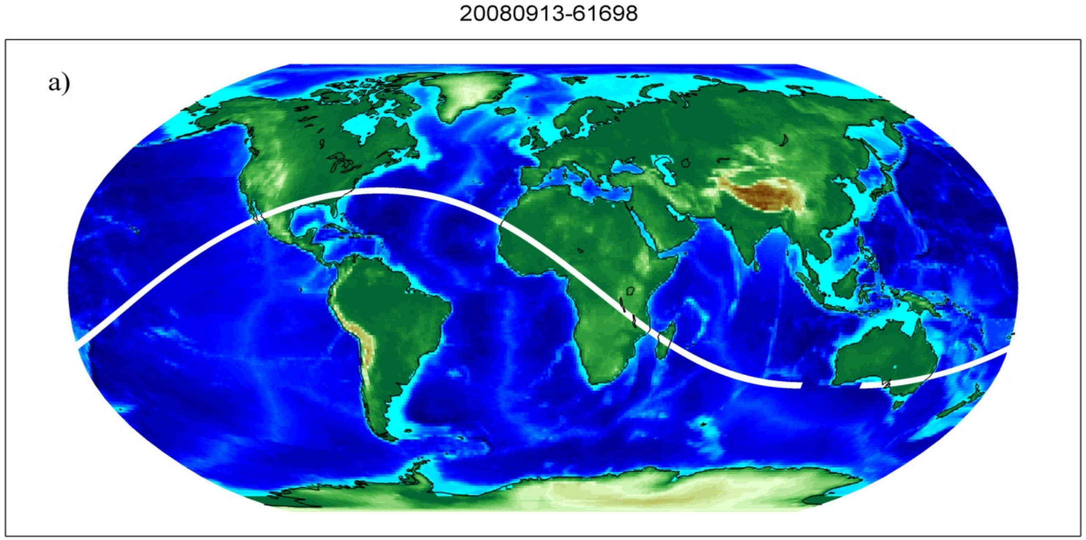
E. Foufoula-Georgiou and A.M. Ebtehaj

University of Minnesota

Princeton University

September, 2012

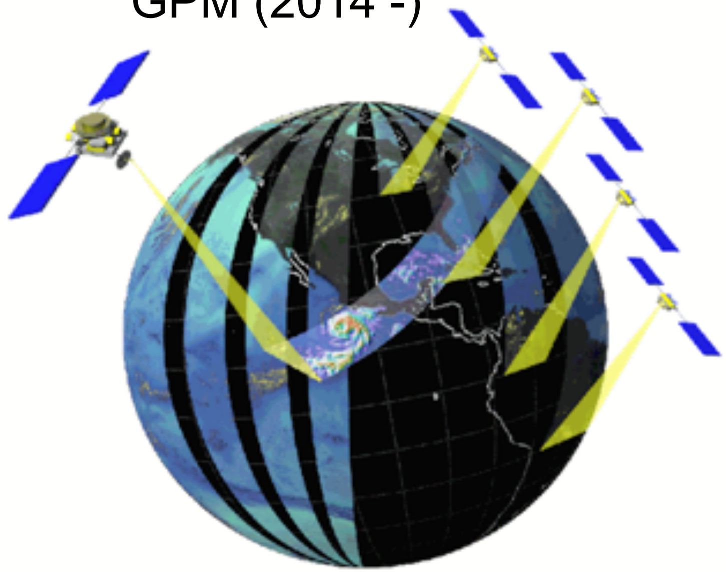
Precipitation from space: an important component of ES modeling

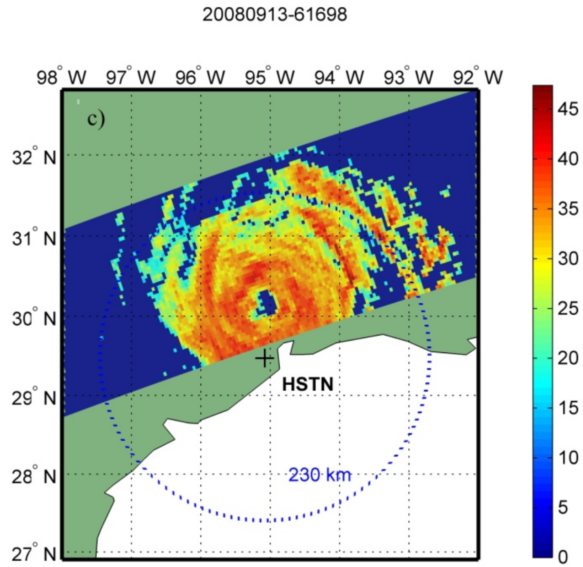


TRMM
(1997 -)

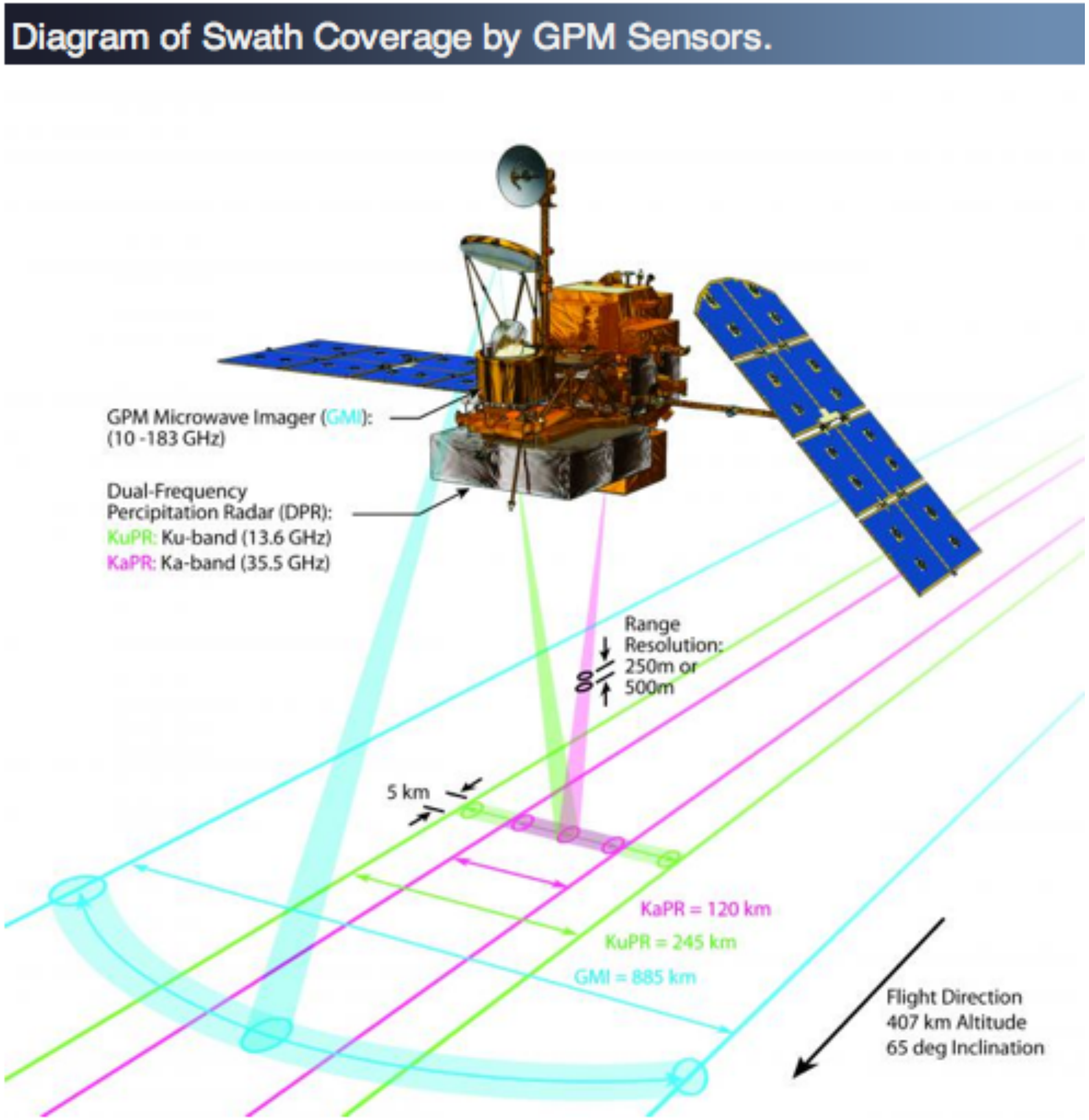
PR

GPM (2014 -)





GPM: a multi-satellite mission extending beyond the tropics



DPR:
125 and 245 Km swaths
Ka-band: 35.5 GHz
Ku-band: 13.6 GHz

GMI:
885 Km swath
13 channels 10 -183 GHz

From TRMM to GPM:
New opportunities &
new challenges
in retrieval, fusion,
and downscaling
of precipitation

Multi-sensor Data Fusion Problem

- Optimal merging of multi-sensor precipitation observations at different scales and with different observational errors

The challenge:

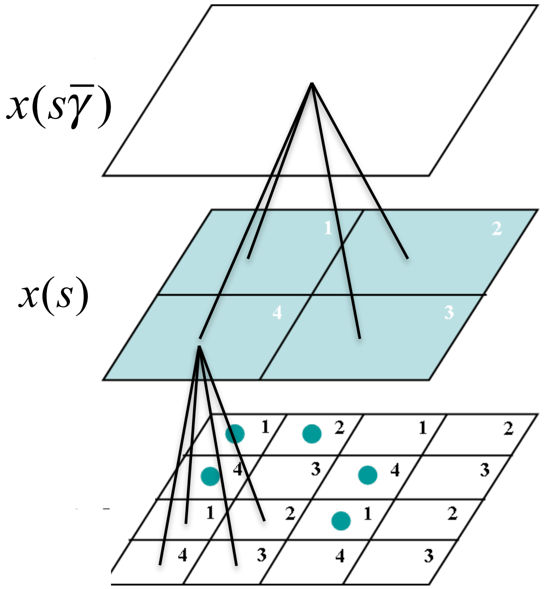
- Reproduce the non-Gaussian statistics of precipitation over multiple scales including the clustered structure and extreme intensities.

- Achieve efficient multi-scale optimal estimation for practical implementation over large space-time domains

Gauss-Markovian methods in real space

- A multiscale Kalman-Filtering methodology [SRE: scale-recursive estimation] by *Chou et. al (1994*) has been used for fusion of rainfall data (e.g. *Gorenburg, McLaughlin, and Entekhabi 2001; Tustison, Harris and Foufoula-Georgiou 2001; Gupta, Venugopal and Foufoula-Georgiou, 2003*).

- Assumptions: Linear Multiscale Gauss-Markovian representation of rainfall data on a quad-tree like structure in real spatial (or log-transformed space).



$$x(s) = A(s)x(s\bar{\gamma}) + w(s)$$
$$y(s) = C(s)x(s) + v(s)$$

Where:
$w(s) \sim N\left(0, R(s)\right);\ \ v(s) \sim N\left(0, Q(s)\right)$

Evolution of the Covariance (Lyapunov Equation):

$$\Sigma_{x(s)} = A(s)\Sigma_{x(s\bar{\gamma})}A(s)^T + R(s)$$

where

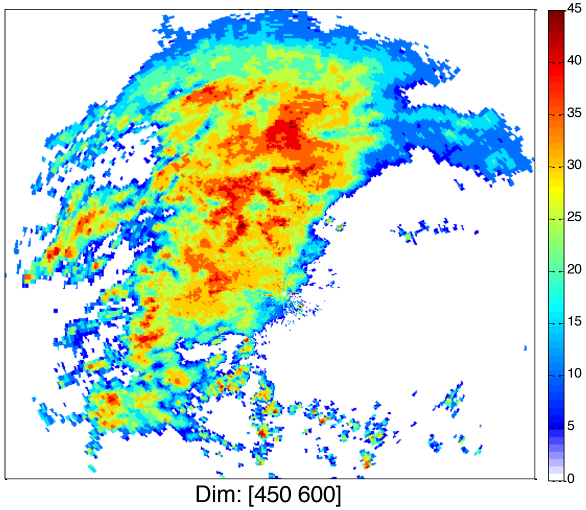
$$\Sigma_{x(s)} = \mathrm{E}\left[x(s)x(s)^T\right]$$

$\mathbf{x}$

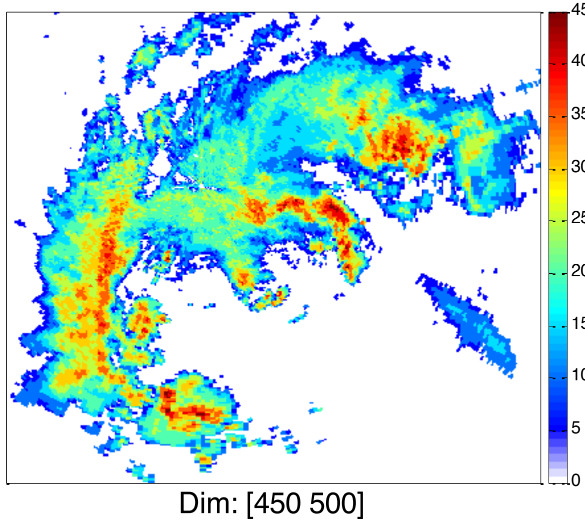
- Markovian structure assumption of the tree makes the linear least squares estimation optimal:

$$\hat{\mathbf{x}} = \Sigma_{\mathbf{x}}\mathbf{c}^T\left[\mathbf{c}\Sigma_{\mathbf{x}}\mathbf{c}^T + \mathbf{R}\right]^{-1}\mathbf{y} \qquad \mathbf{w} \sim N\left(0, \mathbf{R}\right)$$

Spatial Structure of Precipitation

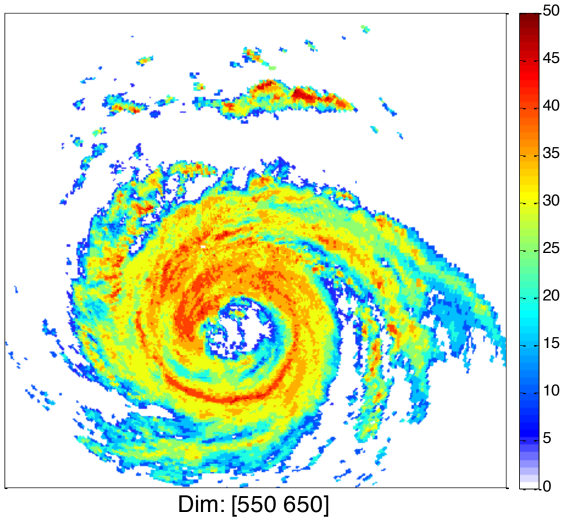


Houston, TX (HSTN) on
November 13, 1998 (02:00:00 UTC)



Houston, TX (HSTN) on
March 29, 1999 (20:13:00 UTC)

Observe in all cases the multi-cellular, "edgy", hierarchical structure

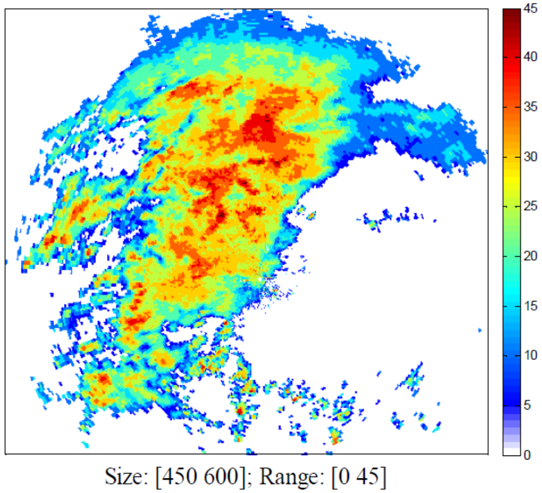


Melbourne, FL (MLB) on
September 26, 2004 (04:50:00 UTC)

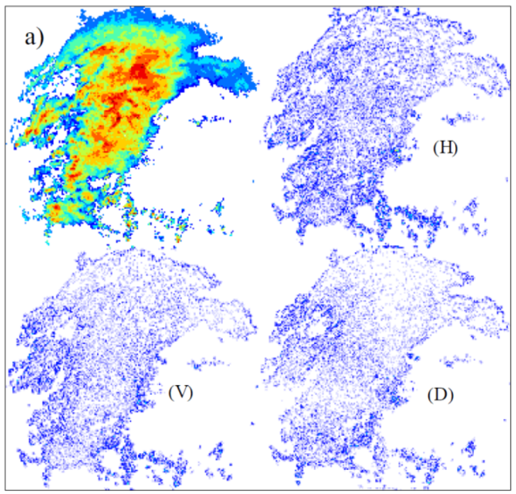
Rainfall is very Sparse in the Gradient Domain

- Storm of 1998/11/13 (02:00 UTC) over TRMM-GV site in Houston, TX.

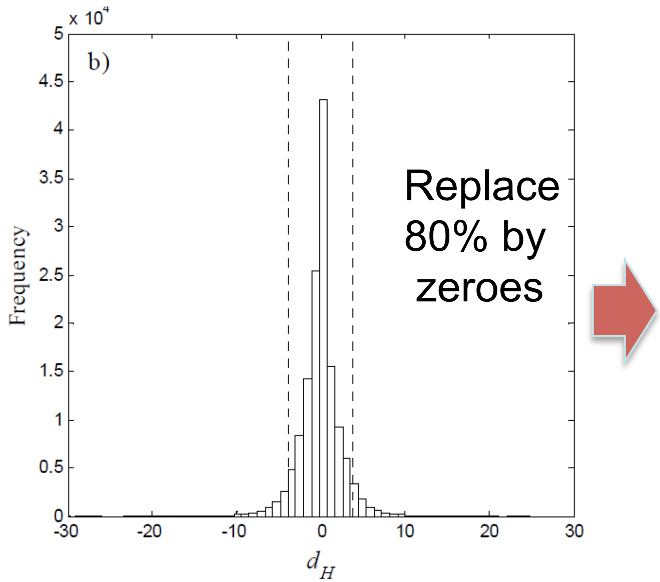
Sparsity: by projecting a signal into an appropriate basis most projection coefs are close to zero and the energy is contained in a few large coefs



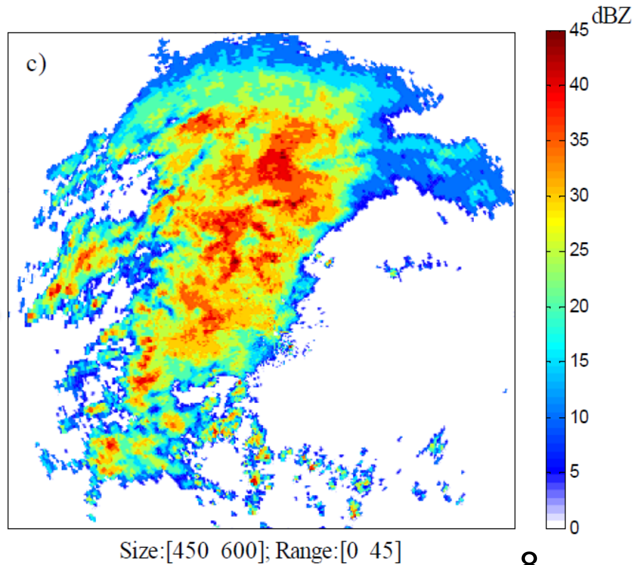
Rainfall is sparse or highly compressible in the wavelet domain!



Wavelet Decomposition



PDF of Coefficients



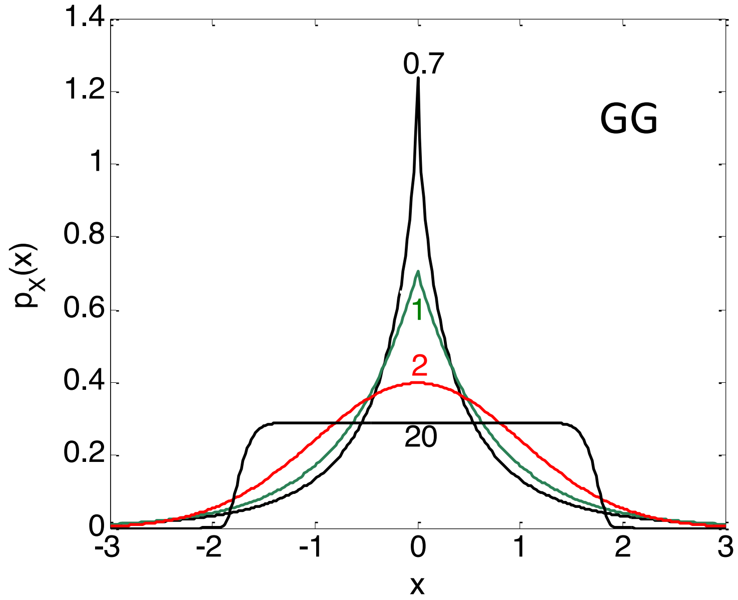
Reconstructed image

Probability Models

Generalized Gaussian (GG)

$$p_X(x) \propto \exp\left(-\left|\frac{x}{s}\right|^{\alpha}\right)$$

a=2 Gaussian, a=1 Laplace



Gaussian Scale Mixture (GSM)

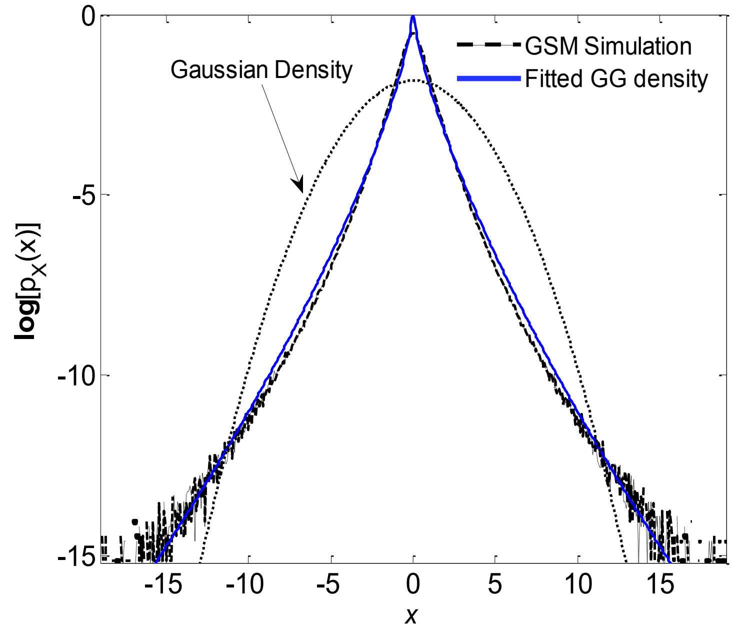
$$x \propto \sqrt{z}\, u$$

z: scalar multiplier
u: Gaussian random variable

$$p_X(x) = \int_0^{\infty} p_{x|z}(x\,|\,z)\, p_z(z)\, dz$$

$$= \int_0^{\infty} \frac{1}{\sqrt{2\pi z \sigma_u^2}} \exp\left(\frac{-x^2}{2z\sigma_u}\right) p_z(z) dz$$

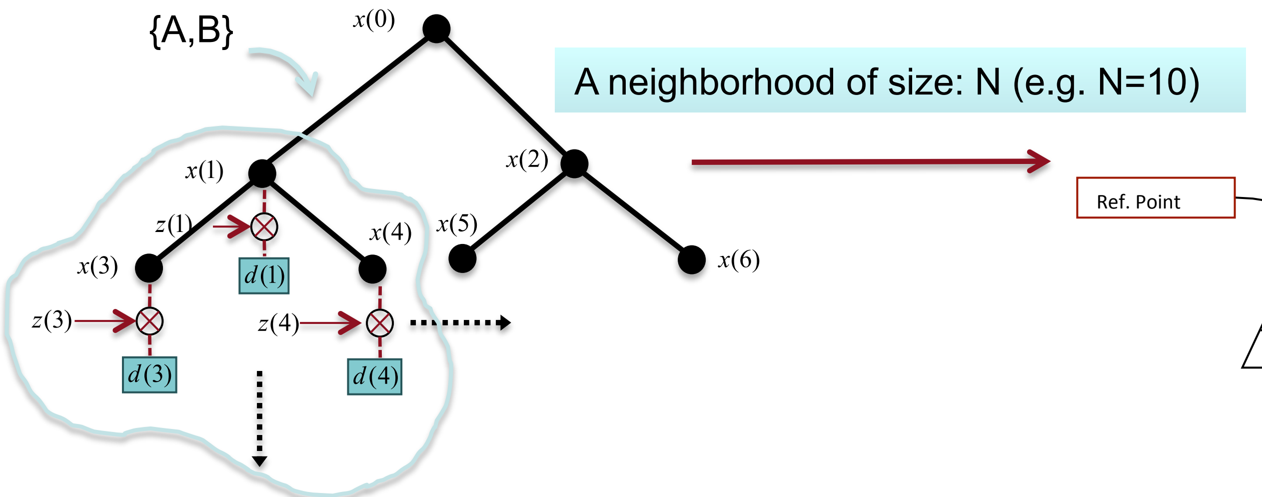
GSM well approximated by a GG

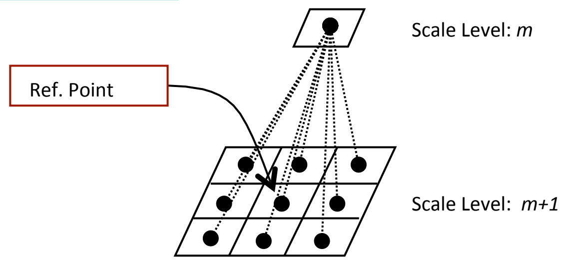


(Andrew, 1973; *West,* 1987)

Data Fusion: Random Cascade on Wavelet tree

- Because of the Decorrelation effect of the wavelet transform this estimation can be performed locally (cutting a local neighborhood) of the wavelet tree





- Local estimation and fusion of neighborhoods of wavelet coefficients with a prior GSM probability distribution

$\sqrt{z}$: is a scalar log-normal, controls the self-reinforcement and heavy tail structure of d(s)

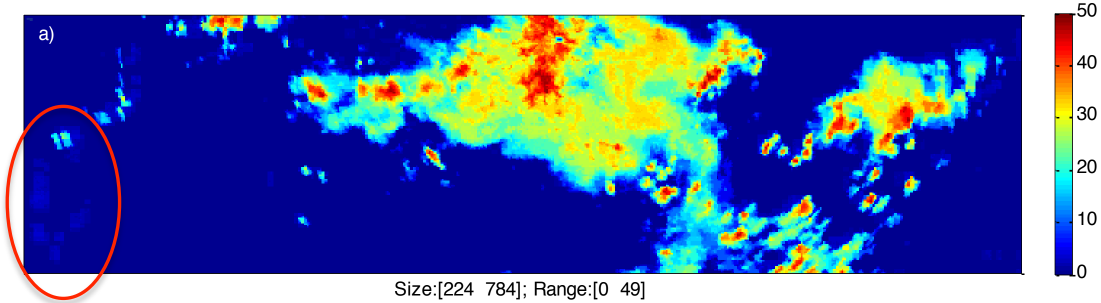
$$d(s) = \sqrt{z} \odot x(s) \qquad \Sigma_{d(s)} = \mathrm{E}[z]\Sigma_{x(s)} \qquad z \sim LN(\mu_Z, \sigma_Z)$$

x(s): MAR process on tree, controls the covariance and parent-to-child dynamics

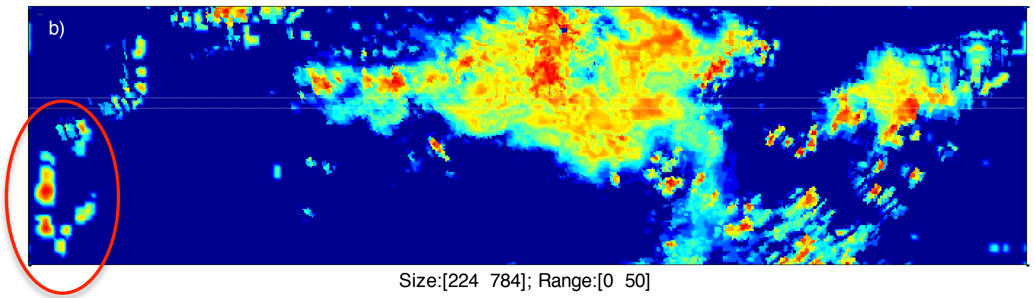
Portilla et al. 2001, Wainwright et al., 2001, Ebtehaj and Foufoula-Georgiou, 2011

Implementation for Multi-sensor Multiscale Rainfall Estimation & Fusion

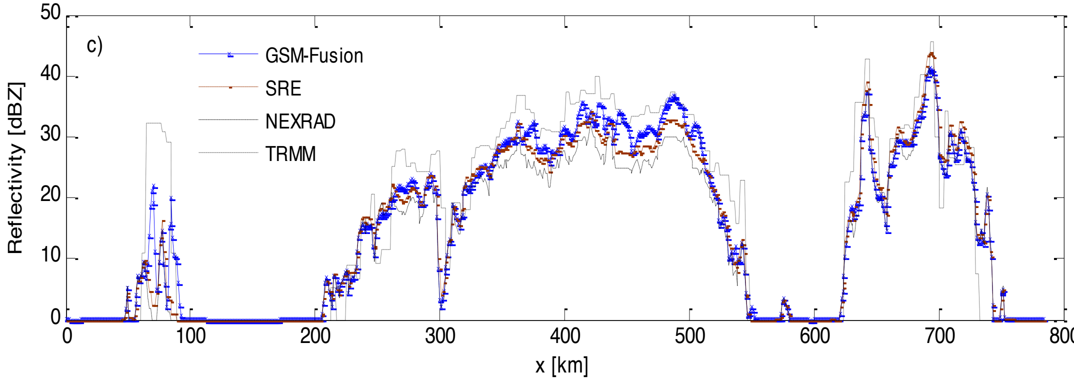
SRE-Fusion of the NEXRAD and TRMM-PR snapshot SNR=5.5 dB
(Overly smooth representation!)



GSM-Fusion with similar magnitude of measurement error, SNR=5.5 dB.
(More Detailed Structure of local rain cells)



A Transect which shows how GSM-fusion can better incorporate the detail structure of rain-cells and TRMM-PR data in the fusion process.

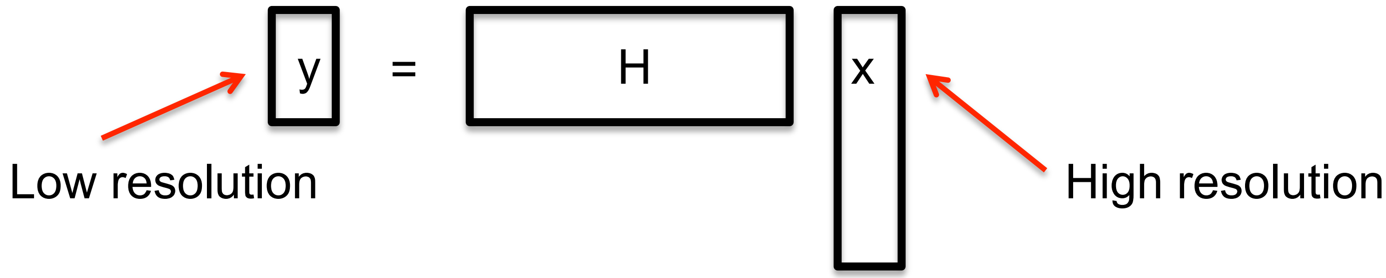


Ebtehaj, M., and E. Foufoula-Georgiou, **Statistics of Precipitation Images and Cascade of Gaussian Scale Mixtures in the Wavelet Domain**, *J. Geophys. Res.*, 2011.

Ebtehaj, M., and E.Foufoula-Georgiou, **Adaptive Fusion of Multi-sensor Precipitation using Gaussian Scale Mixture in the Wavelet Domain** , *J. Geophys.Res.*, 2011.

The Downscaling Problem: A New Perspective

- Sparse Inverse Estimation: Define the downscaling problem as an inverse ill-posed problem and solve it via non-linear constrained optimization



Candes and Tao, 2006; Donoho, 2006; Yang et al., 2010; Elad, 2010

Sparse Inverse Estimator

Recasting into an inverse problem

$$\mathbf{y} = H\mathbf{x} + \mathbf{e}$$

$\mathbf{x} \in \mathbb{R}^m$ (high-res. signal)

$\mathbf{y} \in \mathbb{R}^n$ (low-res. signal)

$H \in \mathbb{R}^{n\times m}$ (Downgrading Operator $m \geq n$)

$\mathbf{e}$ (Obs. Noise)

Sparse projection onto a suitable dictionary

$$\mathbf{x}_s = \mathbf{\Phi}\mathbf{c}$$

$\mathbf{\Phi} \in \mathbb{R}^{m\times M}$ (high-res. dictionary)

$\mathbf{c} \in \mathbb{R}^M$ (Representation Coefficients)

where,

$\|\mathbf{x} - \mathbf{x}_s\|_2 \rightarrow \epsilon$ (Small !)

$\|\mathbf{c}\|_0 \ll m$ (Sparsity !)

Results

$$\begin{aligned} y &= H\mathbf{\Phi}\mathbf{c} + \mathbf{e}' \\ &= \mathbf{\Psi}\mathbf{c} + \mathbf{e}' \end{aligned}$$

$\mathbf{e}' = H(\mathbf{x} - \mathbf{x}_s) + \mathbf{e}$

$\epsilon = \|e'\|_2$

$\mathbf{\Psi} \in \mathbb{R}^{n\times M}$ (low-res. dictionary)

Given y

Approximation of the Coefficients via Optimization

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \ \|\mathbf{c}\|_0 \quad \text{s.t.} \quad \|y - \mathbf{\Psi}\mathbf{c}\|_2 < \epsilon$$

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \ \lambda\|\mathbf{c}\|_1 + \frac{1}{2}\|y - \mathbf{\Psi}\mathbf{c}\|_2^2$$

Recovered Signal

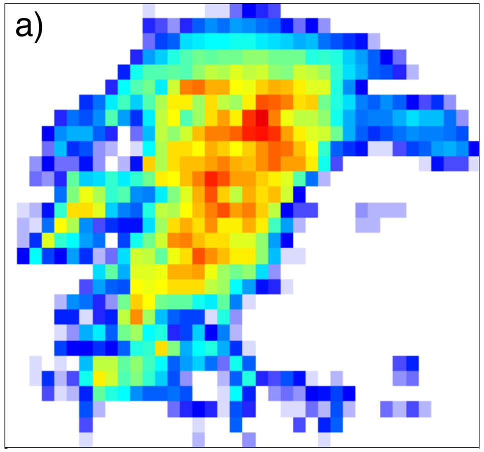
$$\hat{\mathbf{x}} = \mathbf{\Phi}\hat{\mathbf{c}}$$

See, Mallat 2010, Ebtehaj et al., 2011

Testing of the SParse Downscaling (SPaD) Methodology

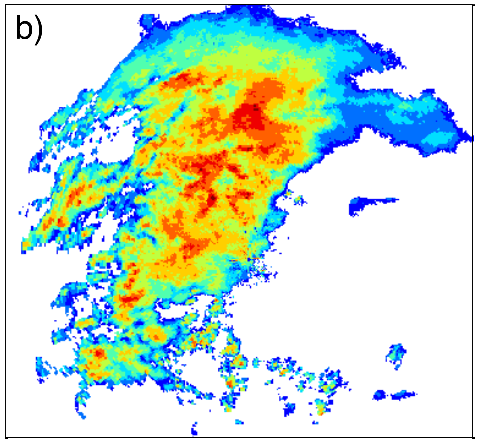
Smoothed and Downsampled by factor of 16x16

s=16



Dim: [29,38]; Range: [0,40]

SPaD Results (1x1 km)



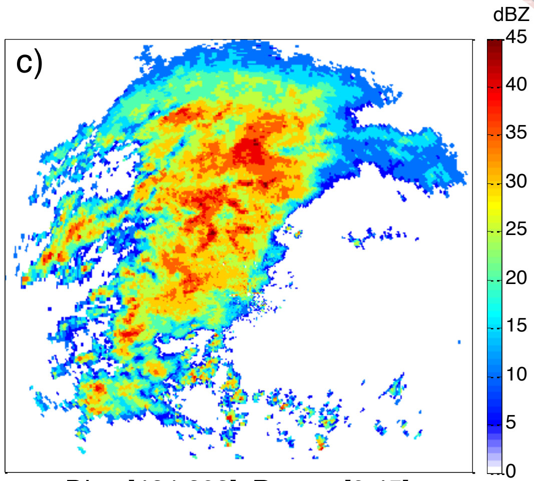
Dim: [464,608]; Range: [0,45]

Results

Structural similarity is almost **perfect** !

MSE is less than **1%** of image energy

Observation (1x1 km)



Dim: [464,608]; Range: [0,45]

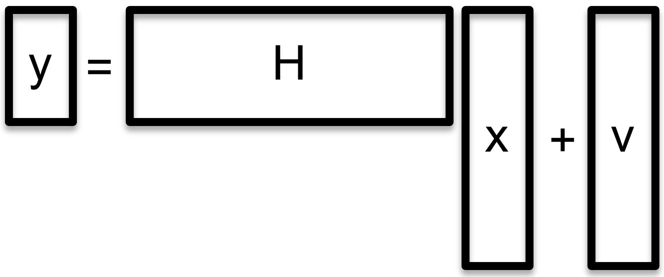
Ebtehaj, A.M., E.Foufoula-Georgiou and G. Lerman (2011), Sparse Regularization for Precipitation Downscaling, *J. Geophys. Res.*, 117, D08107,doi:10.1029/2011JD017057, 2012

A Unified Framework: Discrete Linear Inverse Problems

Downscaling

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$$

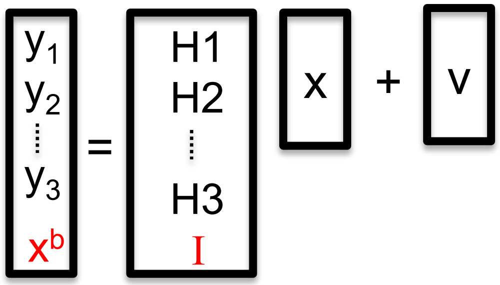
$$\mathbf{v} \sim \mathcal{N}(0, \mathbf{R})$$



Fusion and assimilation

$$\begin{cases} \mathbf{y}^i = \mathbf{H}^i\mathbf{x} + \mathbf{v} \\ \mathbf{x}^\mathbf{b} = \mathbf{x} + \mathbf{w} \end{cases}$$

$$\mathbf{w} \sim \mathcal{N}(0, \mathbf{B})$$



Ill-posed problems:

1. Existence
2. Uniqueness
3. Stability of the solution (inverted noise)

Downscaling as an Inverse Problem

Weighted Least Squares:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \left\| \mathbf{y} - \mathbf{H}\mathbf{x} \right\|_{\mathbf{R}^{-1}}^{2} \right\}$$

Derivative to zero → No unique solution

$$\left(\mathbf{H}^{T} \mathbf{R}^{-1} \mathbf{H} \right) \hat{\mathbf{x}} = \mathbf{H}^{T} \mathbf{R}^{-1} \mathbf{y}$$

Regularization: (Tikhonov)

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \left\| \mathbf{y} - \mathbf{H}\mathbf{x} \right\|_{\mathbf{R}^{-1}}^{2} + \lambda \left\| \mathbf{L}\mathbf{x} \right\|_{2}^{2} \right\}$$

Linear Problem → Unique solution

$$\hat{\mathbf{x}} = \left(\mathbf{H}^{T} \mathbf{R}^{-1} \mathbf{H} + \lambda \mathbf{L}^{T} \mathbf{L} \right)^{-1} \mathbf{H}^{T} \mathbf{R}^{-1} \mathbf{y}$$

L_1 regularization:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \left\| \mathbf{y} - \mathbf{H}\mathbf{x} \right\|_{\mathbf{R}^{-1}}^{2} + \lambda \left\| \mathbf{L}\mathbf{x} \right\|_{1} \right\}$$

Non-smooth → Not strictly convex

Harder to solve

(Constrained Quadric Programming, Proximal Gradient Method, Interior Point Methods…)

Huber regularization:

$$\left\| \mathbf{x} \right\|_{\mathrm{Hub}} = \sum_{i} \rho_{T}(x_i)$$

$$\rho_{T}(x) = \begin{cases} x^2 & |x| \le T \\ T\left(2|x| - T\right) & |x| > T \end{cases},$$

smooth → strictly convex close to the origin

Gradient descent methods

(large scale problems robustly)

Regularized Downscaling- statistical Interpretation

Frequentist Approach:

$$\hat{\mathbf{x}}_{ML} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p\left(\mathbf{y} \mid \mathbf{x}\right)$$

$$p\left(\mathbf{y} \mid \mathbf{x}\right) \propto \exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}\mathbf{x})\right)$$

$$\hat{\mathbf{x}}_{ML} = \underset{\mathbf{x}}{\operatorname{argmin}}\left\{\frac{1}{2}(\mathbf{y}-\mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}\mathbf{x})\right\} = \underset{\mathbf{x}}{\operatorname{argmin}}\left\{\frac{1}{2}\left\|\mathbf{y}-\mathbf{H}\mathbf{x}\right\|_{\mathbf{R}^{-1}}^2\right\}$$

Bayesian Approach:

$$\hat{\mathbf{x}}_{MAP} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p\left(\mathbf{x} \mid \mathbf{y}\right)$$

$$\hat{\mathbf{x}}_{MAP} = \underset{\mathbf{x}}{\operatorname{argmin}}\left\{-\log\left(\frac{p(\mathbf{y} \mid \mathbf{x})\, p(\mathbf{x})}{p(\mathbf{y})}\right)\right\} = \underset{\mathbf{x}}{\operatorname{argmin}}\left\{-\log p\left(\mathbf{y} \mid \mathbf{x}\right) - \underbrace{\log p(\mathbf{x})}_{\text{prior}}\right\}$$

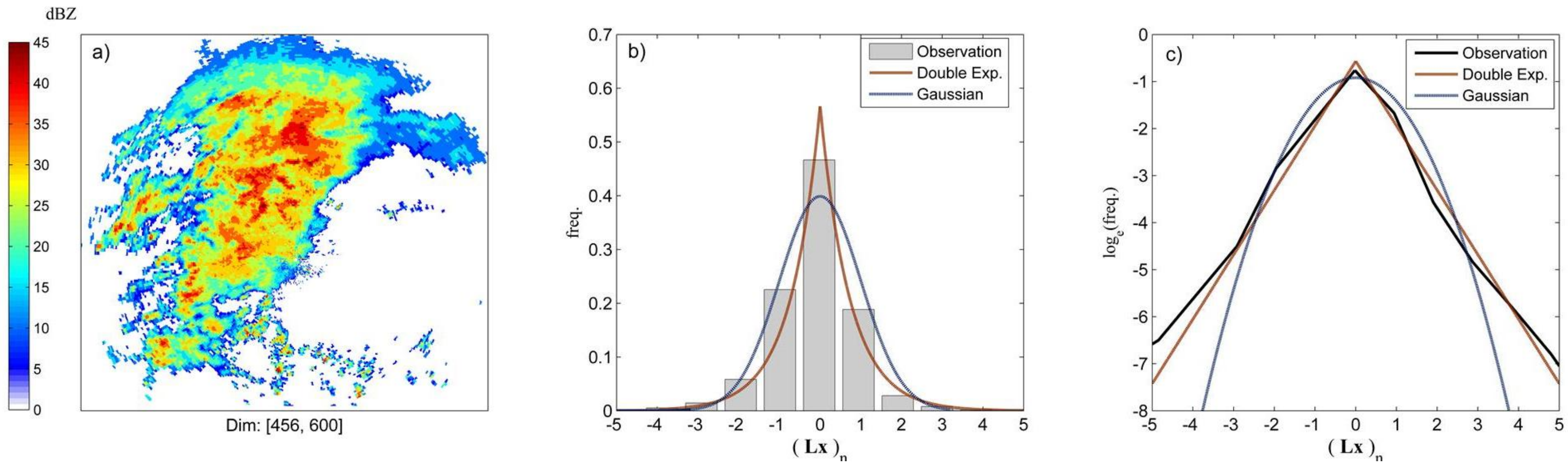
prior

$$\log p(\mathbf{x}) \propto \mathbf{x}^T \mathbf{Q}\mathbf{x} \qquad \mathbf{Q} = \mathbf{L}^T\mathbf{L}$$ ⟷ Tikhonov reg.

$$\log p(\mathbf{x}) \propto \left\|\mathbf{L}\mathbf{x}\right\|_1$$ ⟷ L_1 reg.

$$\log p(\mathbf{x}) \propto \sum_i \rho_T(x_i)$$ ⟷ Huber reg.

Statistics of rainfall under the Laplacian transform

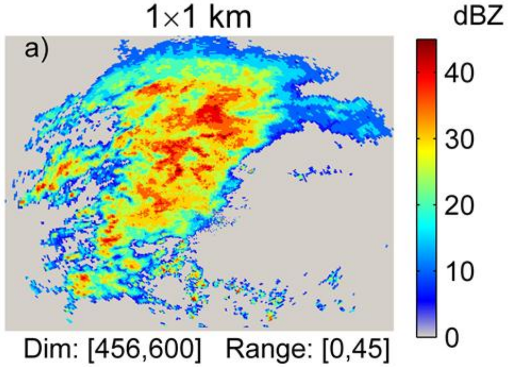


$$\log p(\mathbf{x}) \propto \left\| \mathbf{Lx} \right\|_1$$

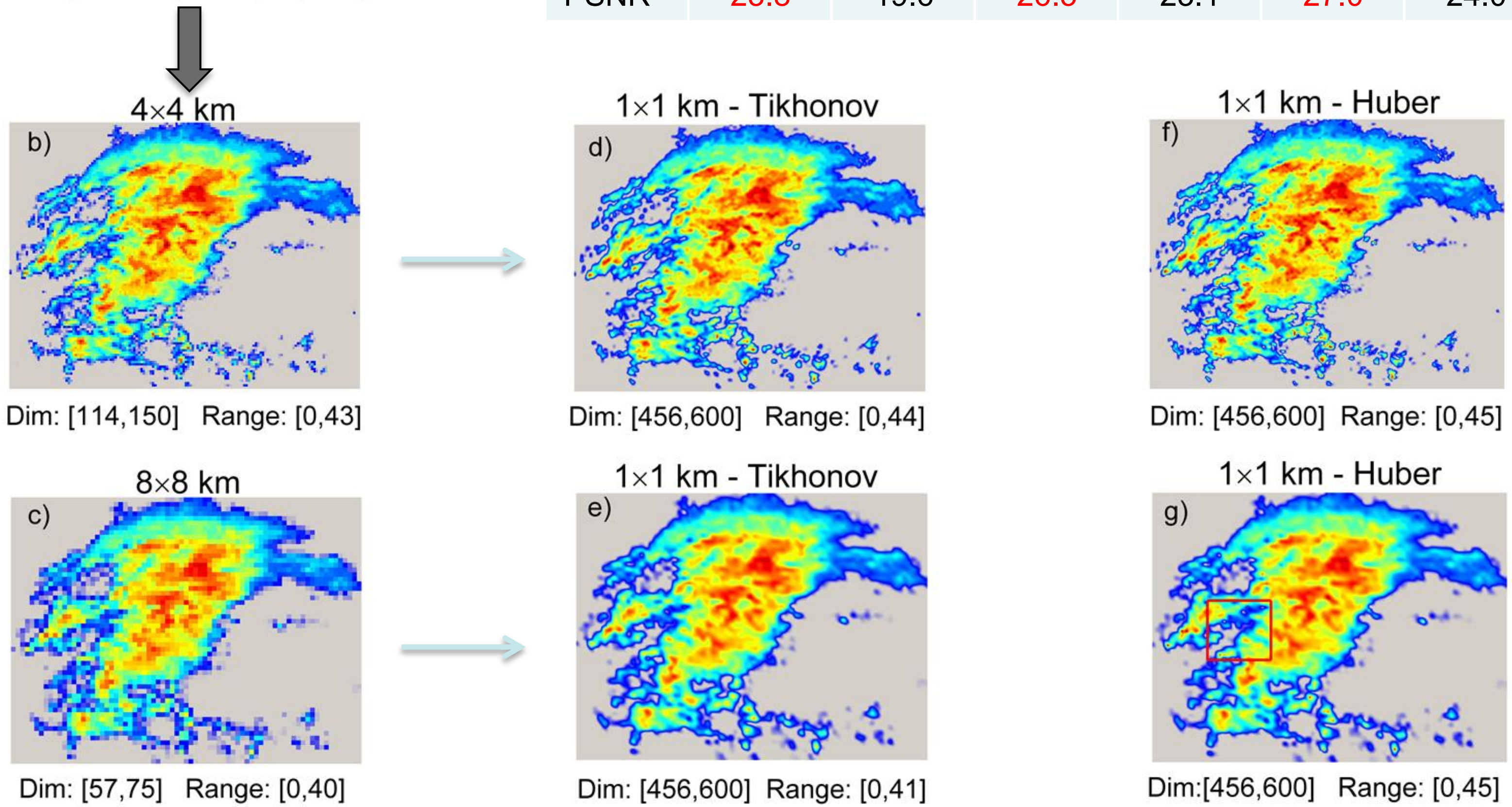
DS- Examples

High and low-res. counterparts

Downscaled



Metric	Observations		Tikhonov		Huber	
	4X4 km	8x8 km	4X4 km	8x8 km	4X4 km	8x8 km
RMSE	0.19	0.29	0.15	0.20	0.14	0.19
MAE	0.15	0.25	0.13	0.18	0.11	0.17
SSIM	0.71	0.56	0.78	0.66	0.80	0.66
PSNR	23.8	19.6	26.5	23.1	27.0	24.0

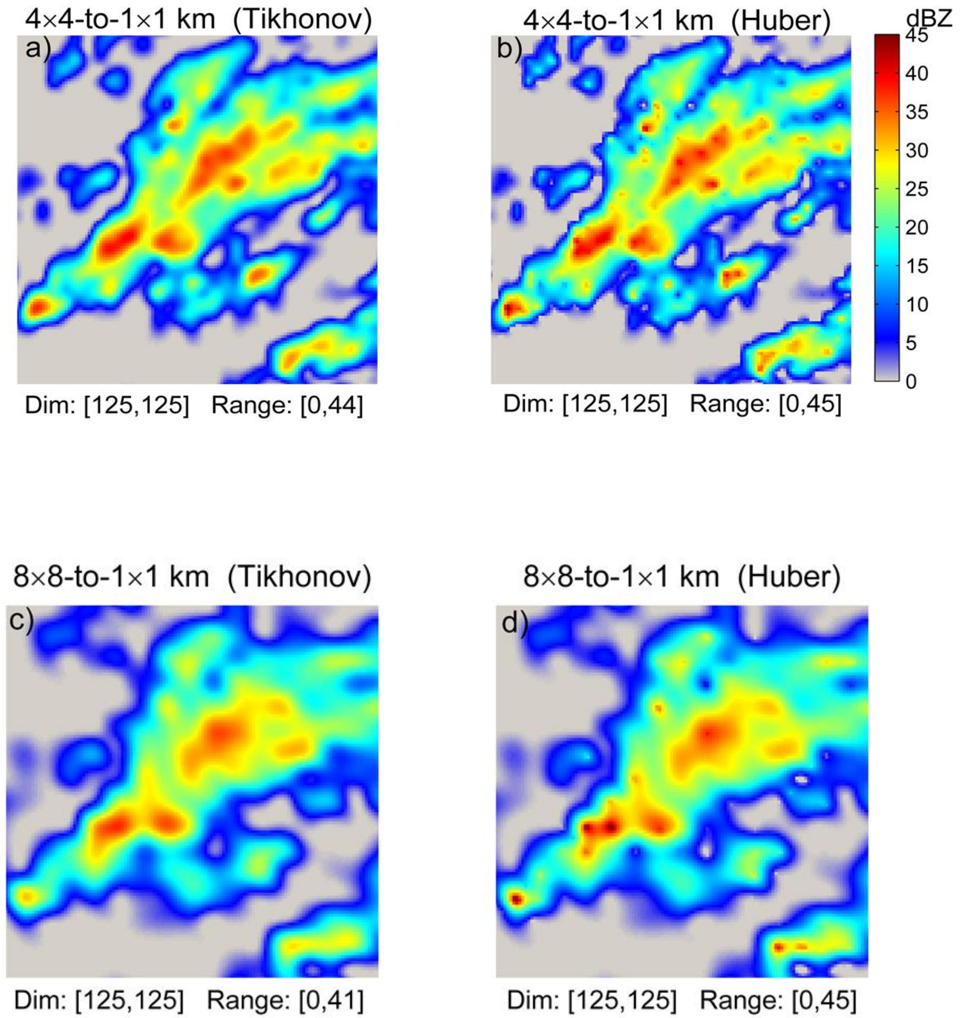


Downscaling: Example Results (zooming)

Tikhonov | Huber

4x4 to 1x1

8x8 to 1x1



Observe the suppressed range of the Tikhonov

Data Fusion

Weighted Least Squares:

$$\mathbf{y}^i = \mathbf{H}^i\mathbf{x} + \mathbf{v},$$
$$\mathbf{y}^i \in \mathbb{R}^{n_i}, i = 1,\ldots,N$$

Augmentation:

$$\begin{bmatrix} \mathbf{y}^1 \\ \vdots \\ \mathbf{y}^N \end{bmatrix} = \begin{bmatrix} \mathbf{H}^1 \\ \vdots \\ \mathbf{H}^N \end{bmatrix}\mathbf{x} + \begin{bmatrix} \mathbf{v}^1 \\ \vdots \\ \mathbf{v}^N \end{bmatrix} \Rightarrow \underline{\mathbf{y}} = \underline{\mathbf{H}}\mathbf{x} + \underline{\mathbf{v}} \qquad \underline{\mathbf{R}} = \mathbb{E}\left[\underline{\mathbf{v}}\,\underline{\mathbf{v}}^T\right] = \begin{bmatrix} \mathbf{R}^1 & & 0 \\ & \ddots & \\ 0 & & \mathbf{R}^N \end{bmatrix}.$$

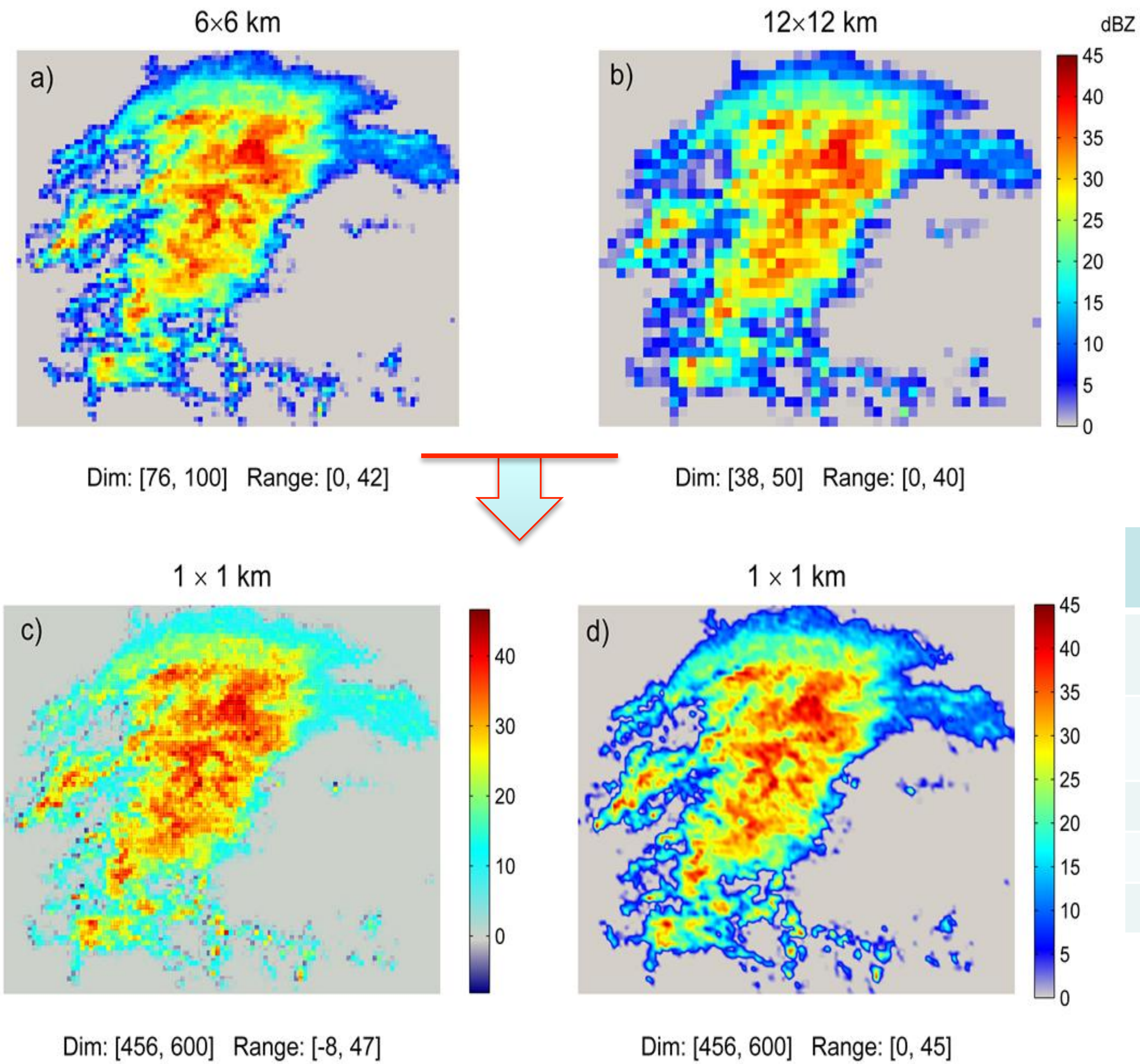
$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}}\left\{\frac{1}{2}\left\|\underline{\mathbf{y}} - \underline{\mathbf{H}}\mathbf{x}\right\|^2_{\underline{\mathbf{R}}^{-1}}\right\} \xrightarrow{\text{Unstable solution}} \left(\underline{\mathbf{H}}^T\underline{\mathbf{R}}^{-1}\underline{\mathbf{H}}\right)\hat{\mathbf{x}} = \underline{\mathbf{H}}^T\underline{\mathbf{R}}^{-1}\underline{\mathbf{y}}$$

Regularization:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}}\left\{\frac{1}{2}\left\|\underline{\mathbf{y}} - \underline{\mathbf{H}}\mathbf{x}\right\|^2_{\underline{\mathbf{R}}^{-1}} + \lambda\psi_{\mathbf{L}}(\mathbf{x})\right\}$$

$$\psi_{\mathbf{L}}(\mathbf{x}) = \begin{cases} \left\|\mathbf{L}\mathbf{x}\right\|_2 & \rightarrow \text{Tikhonov (L}_2\text{-norm)} \\ \left\|\mathbf{L}\mathbf{x}\right\|_1 & \rightarrow \text{L}_1\text{-norm} \\ \left\|\mathbf{L}\mathbf{x}\right\|_{\text{Hub}} & \rightarrow \text{Huber-norm} \end{cases}$$

Data Fusion - Examples



No regularization

Huber regularization

Metric	Observations		DF
	6X6 km	12x12 km	1x1 km
RMSE	0.25	0.35	0.17
MAE	0.21	0.32	0.15
SSIM	0.60	0.50	0.72
PSNR	21.3	18.1	25.0

Weighted Least Squares:

Classic 3D-VAR:

$$\mathcal{J}_{3D}(\mathbf{x}_k) = \frac{1}{2}\left\|\mathbf{x}_k^b - \mathbf{x}_k\right\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2}\left\|\mathbf{y}_k - \mathbf{H}\mathbf{x}_k\right\|_{\mathbf{R}^{-1}}^2 .$$

$$\mathbf{x}_k^a = \underset{\mathbf{x}_k}{\operatorname{argmin}}\left\{\mathcal{J}_{3D}(\mathbf{x}_k)\right\} \qquad \mathbf{x}_k^a = \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\left(\mathbf{B}^{-1}\mathbf{x}_k^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_k\right)$$

Regularized DA:

$$\mathbf{x}_k^a = \underset{\mathbf{x}_k}{\operatorname{argmin}}\left\{\mathcal{J}_{3D}(\mathbf{x}_k) + \lambda\psi_{\mathbf{L}}\left(\mathbf{x}_k\right)\right\} \qquad \psi_{\mathbf{L}}\left(\mathbf{x}\right) = \begin{cases} \left\|\mathbf{L}\mathbf{x}\right\|_2 & \rightarrow \text{Tikhonov } (\text{L}_2\text{-norm}) \\ \left\|\mathbf{L}\mathbf{x}\right\|_1 & \rightarrow \text{L}_1\text{-norm} \\ \left\|\mathbf{L}\mathbf{x}\right\|_{\text{Hub}} & \rightarrow \text{Huber-norm} \end{cases}$$

Weighted Least Squares:

3D-VAR:

$$\mathcal{J}_{3D}(\mathbf{x}_k) = \frac{1}{2}\left\|\mathbf{x}_k^b - \mathbf{x}_k\right\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2}\left\|\mathbf{y}_k - \mathbf{H}\mathbf{x}_k\right\|_{\mathbf{R}^{-1}}^2 .$$

$$\mathbf{x}_k^a = \underset{\mathbf{x}_k}{\operatorname{argmin}}\left\{\mathcal{J}_{3D}(\mathbf{x}_k)\right\} \qquad \mathbf{x}_k^a = \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\left(\mathbf{B}^{-1}\mathbf{x}_k^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_k\right)$$

4D-VAR:

$$\mathcal{J}_{4D}(\mathbf{x}_k) = \frac{1}{2}\left\|\mathbf{x}_k^b - \mathbf{x}_k\right\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2}\sum_{i=k}^{K+T}\left(\left\|\mathbf{y}_i - \mathbf{H}\mathbf{M}_{k,i}\mathbf{x}_k\right\|_{\mathbf{R}_i^{-1}}^2\right)$$

Regularized DA:

$$\mathbf{x}_k^a = \underset{\mathbf{x}_k}{\operatorname{argmin}}\left\{\mathcal{J}_{3D}(\mathbf{x}_k) + \lambda\psi_{\mathbf{L}}\left(\mathbf{x}_k\right)\right\}$$

$$\mathbf{x}_k^a = \underset{\mathbf{x}_k}{\operatorname{argmin}}\left\{\mathcal{J}_{4D}(\mathbf{x}_k) + \lambda\psi_{\mathbf{L}}\left(\mathbf{x}_k\right)\right\}$$

$$\psi_{\mathbf{L}}\left(\mathbf{x}\right) = \begin{cases} \left\|\mathbf{L}\mathbf{x}\right\|_2 & \rightarrow \text{Tikhonov } (\text{L}_2\text{-norm}) \\ \left\|\mathbf{L}\mathbf{x}\right\|_1 & \rightarrow \text{L}_1\text{-norm} \\ \left\|\mathbf{L}\mathbf{x}\right\|_{\text{Hub}} & \rightarrow \text{Huber-norm} \end{cases}$$

Statistical Interpretation (DA)-1

Frequentist Approach (3D-VAR):

$$\hat{\mathbf{x}}_{ML} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p\left(\underline{\mathbf{y}} \mid \mathbf{x}_k\right) \longrightarrow \text{Maximum Likelihood (ML)}$$

$$\begin{cases} \mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} & \mathbf{v} \sim \mathcal{N}\left(0, \mathbf{R}\right) \\ \mathbf{x}_k^b = \mathbf{x}_k + \mathbf{w} & \mathbf{w} \sim \mathcal{N}\left(0, \mathbf{B}\right) \end{cases}$$

$$\underline{\mathbf{y}} = \underline{\mathbf{H}}\mathbf{x}_k + \underline{\mathbf{v}}, \text{ where, } \quad \underline{\mathbf{y}} = \left[\left(\mathbf{x}_k^b\right)^T, \mathbf{y}_k^T\right]^T, \quad \underline{\mathbf{H}} = \left[\mathbf{I}, \mathbf{H}^T\right]^T, \text{and } \quad \underline{\mathbf{v}} \sim \mathcal{N}\left(0, \underline{\mathbf{R}}\right)$$

$$\underline{\mathbf{R}} = \begin{bmatrix} \mathbf{B} & 0 \\ 0 & \mathbf{R} \end{bmatrix}$$

Knowing that the log-likelihood is:

$$-\log p(\underline{\mathbf{y}} \mid \mathbf{x}_k) \propto \frac{1}{2}(\underline{\mathbf{y}} - \underline{\mathbf{H}}\mathbf{x}_k)^T \underline{\mathbf{R}}^{-1}(\underline{\mathbf{y}} - \underline{\mathbf{H}}\mathbf{x}_k)$$

$$\mathbf{x}_{ML} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2}\left\|\mathbf{x}_k^b - \mathbf{x}_k\right\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2}\left\|\mathbf{y}_k - \mathbf{H}\mathbf{x}_k\right\|_{\mathbf{R}^{-1}}^2 \right\}$$

Statistical Interpretation (DA)-2

Bayesian Approach (3D-VAR):

$$\hat{\mathbf{x}}_{MAP} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p\left(\mathbf{x} \mid \mathbf{y}\right)$$

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{v} \qquad & \mathbf{v} \sim \mathcal{N}\left(0, \mathbf{R}\right) \\ \mathbf{x}_k &= \mathbf{x}_k^b + \mathbf{w} \qquad & \mathbf{w} \sim \mathcal{N}\left(0, \mathbf{B}\right) \end{aligned}$$

$$p(\mathbf{x}_k) \sim \mathcal{N}(\mathbf{x}_k^b, \mathbf{B})$$

$$\begin{aligned} \hat{\mathbf{x}}_{MAP} &= \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ -\log\left(\frac{p(\mathbf{y} \mid \mathbf{x})\, p(\mathbf{x})}{p(\mathbf{y})} \right) \right\} \\ &= \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ -\log p\left(\mathbf{y} \mid \mathbf{x}\right) - \log p(\mathbf{x}) \right\} \end{aligned}$$

$$\hat{\mathbf{x}}_{\text{MAP}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \left\| \mathbf{x}_k^b - \mathbf{x}_k \right\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \left\| \mathbf{y}_k - \mathbf{H}\mathbf{x}_k \right\|_{\mathbf{R}^{-1}}^2 \right\}$$

Does not go beyond WLS estimator!

Statistical Interpretation (DA)-3

Bayesian Approach (Regularized 3D-VAR):

$$p(\mathbf{x}_k) \propto \lambda \psi_{\mathbf{L}}\left(\mathbf{x}_k\right)$$

$$\mathbf{x}_k^a = \underset{\mathbf{x}_k}{\operatorname{argmin}} \left\{ \mathcal{J}_{3D}(\mathbf{x}_k) + \lambda \psi_{\mathbf{L}}\left(\mathbf{x}_k\right) \right\}$$

The Regularized 3D-VAR might be interpreted as a MAP estimator which also accounts for an independent prior $p(\mathbf{x}_k) \propto \lambda \psi_{\mathbf{L}}\left(\mathbf{x}_k\right)$

MAP estimator in the derivative space
or in general in a transform domain!

Regularized Data Assimilation: Example 1-D Heat equation

>> Estimation of the initial condition from diffused and noisy observations:
an ill-posed deconvolution problem

>> Space-time representation of the 1-D scalar quantity x(s,t):

$$\frac{\partial x(s,t)}{\partial t} = \gamma \nabla^2 x(s,t)$$

$$x(s,0) = x_0(s)$$

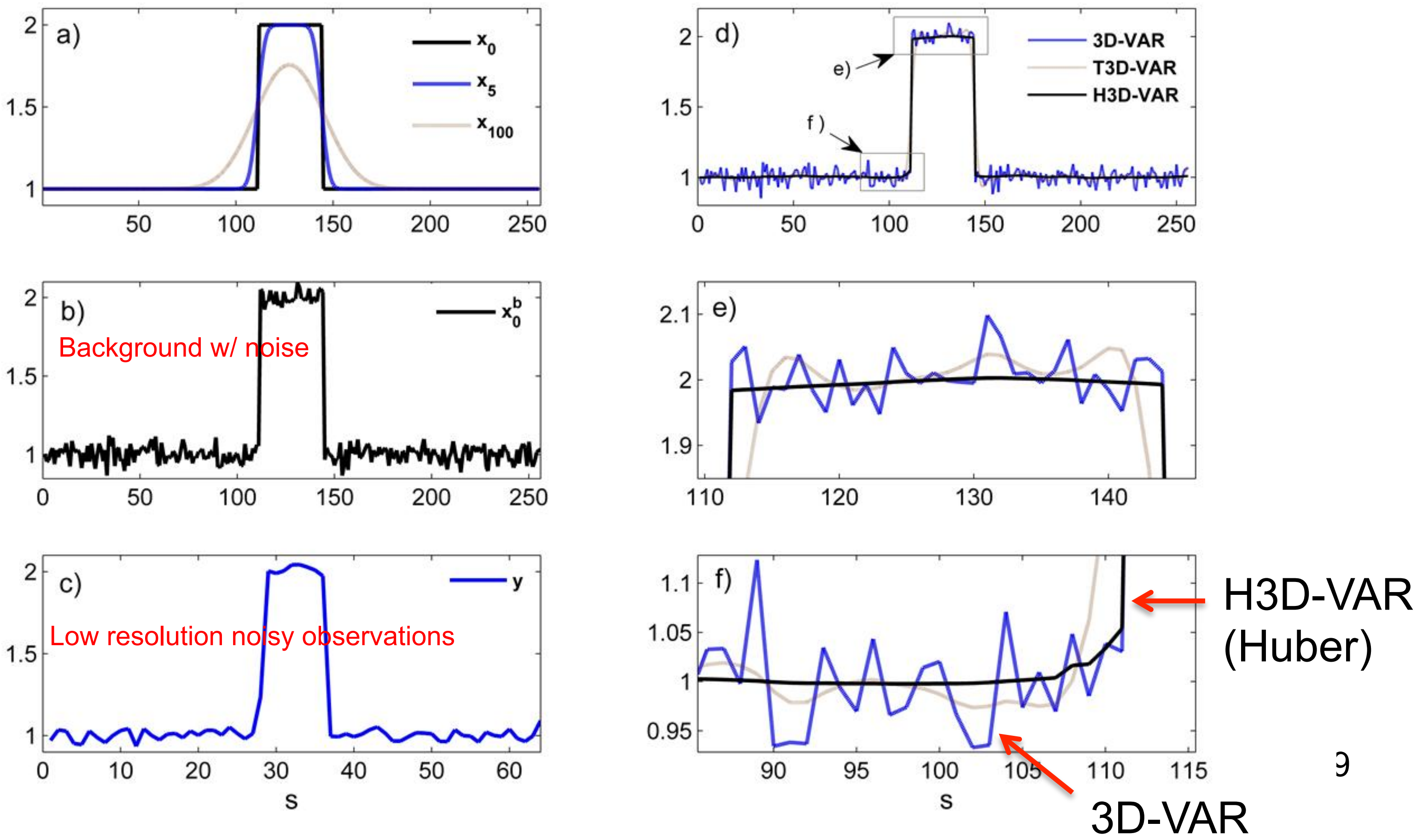
>> Solution:

$$x(s,t) = \int K\left(s-r,t\right) x_0(r)\,dr,$$

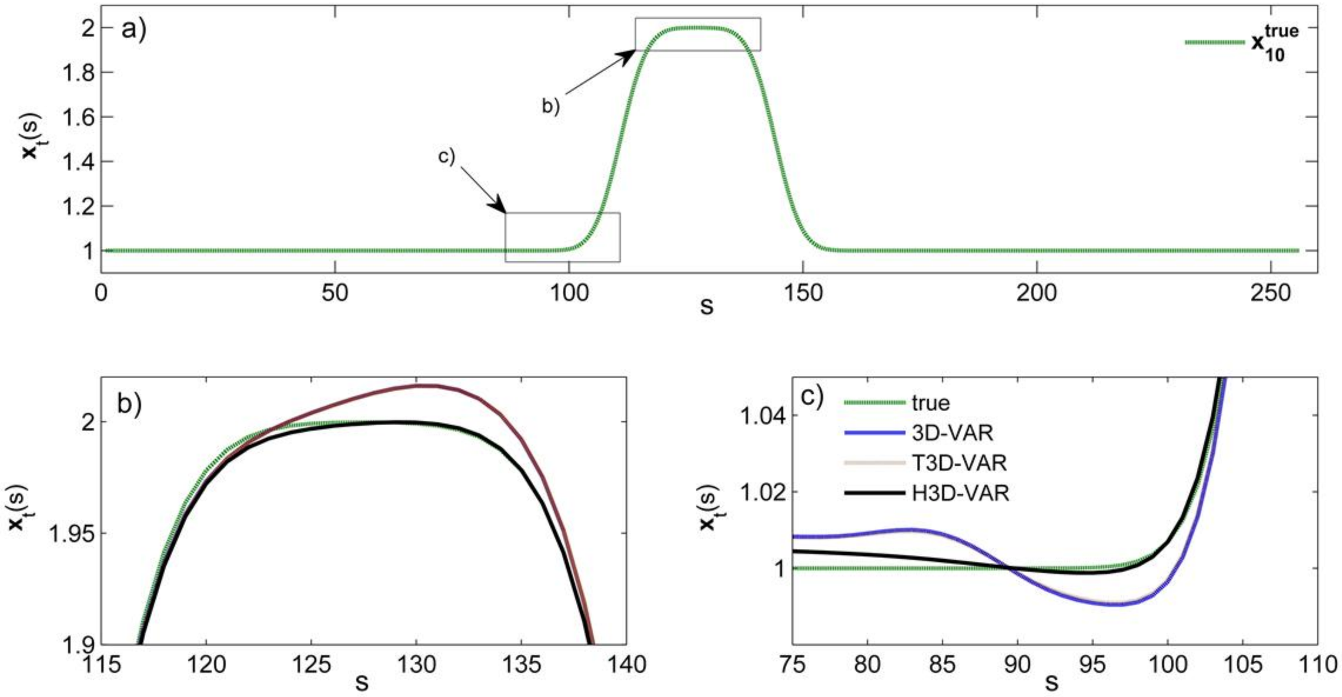
$$\text{where } K(s,t) = (4\pi t)^{-m/2} \exp\left(\frac{-\left|s\right|^2}{4t}\right)$$

Example: 1-D Heat equation

$$\mathbf{x}_0 = \begin{cases} 2 & 112 \le x_i \le 144 \\ 1 & \text{otherwise,} \end{cases} \qquad \mathbf{H} = \frac{1}{4}\begin{bmatrix} 1111 & 0000 & \cdots & 0000 \\ 0000 & 1111 & \cdots & 0000 \\ \vdots & \vdots & \vdots & \vdots \\ 0000 & 0000 & \cdots & 1111 \end{bmatrix} \in \mathbb{R}^{n\times m} \qquad \mathbf{L} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(m-1)\times m}$$



Example: 1-D Heat equation (Forecast at t=10)



Ebtehaj, A.M.,and E.Foufoula-Georgiou, Variational Downscaling, Fusion, and Assimilation of Hydrometeorological states via Regularized Estimation, reprint, 2012.

Fresh perspectives to old problems

1. Desire to preserve spatial coherency, abrupt gradients, and extremes
   (geometrically structured fields and non-Gaussian statistics)
2. Computationally efficient optimal estimation for large-scale applications

EXPLORE:

1. Sparsity in a transformed domain (gradient or wavelet domain)
2. Regularized Inverse Estimation (L2, L1, Huber norms of Lx)
3. Conditionally Gaussian form (GSM) of PDF: exploit linear estimation theory in an adaptive way and in locally defined operations