

FQI & Kantorovich Distance

Sukanta Basu

&

Efi Foufoula-Georgiou

(April 13, 2008)

Kantorovich (Transportation) Distance

“The general transportation problem is concerned with distributing any commodity from any group of supply centers, called sources, to any group of receiving centers, called destinations, in such a way as to minimize the total distribution costs”

Hillier and Lieberman (1990)

Source i ($i = 1, 2, \dots, m$) has a supply for s_i units to distribute to the destinations, and destination j ($j = 1, 2, \dots, n$) has a demand for d_j units to be received from the sources. The cost of distributing units from source i to destination j is directly proportional to the number distributed (x_{ij}). Lastly, c_{ij} denotes the cost per unit distributed.

Linear Programming Formulation

$$\text{minimize } \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}$$

$$\text{subject to } \sum_{j=1}^n x_{ij} = s_i, \quad \text{for } i = 1, 2, \dots, m$$

$$\sum_{i=1}^m x_{ij} = d_j, \quad \text{for } j = 1, 2, \dots, n$$

$$\text{and } x_{ij} \geq 0, \quad \text{for all } i \text{ and } j$$

Using Kantorovich Distance in QPF Verification

Pros:

- a single metric capturing amplitude and displacement errors
- can be used as an error index in predictability problems

Cons:

- supply = demand (bias corrections required)
- computationally expensive

Proof of Concept: Perturbed Case (Threshold = 20 mm)

Case #	d_i	λ	$d_i * \lambda$
1	5.55×10^5	1	5.55×10^5
2	5.63×10^5	1	5.63×10^5
3	6.10×10^5	1	6.10×10^5
4	7.81×10^5	1.01	7.89×10^5
5	12.15×10^5	1.04	12.64×10^5
6	5.23×10^5	1/0.31	16.87×10^5
7	5.13×10^5	1.25	6.41×10^5

NOTE:

Lower d_i : better forecast;

Lower λ (bias correction): better forecast;

Lower $d_i \lambda$: better forecast.

Forecast Quality Index (FQI)

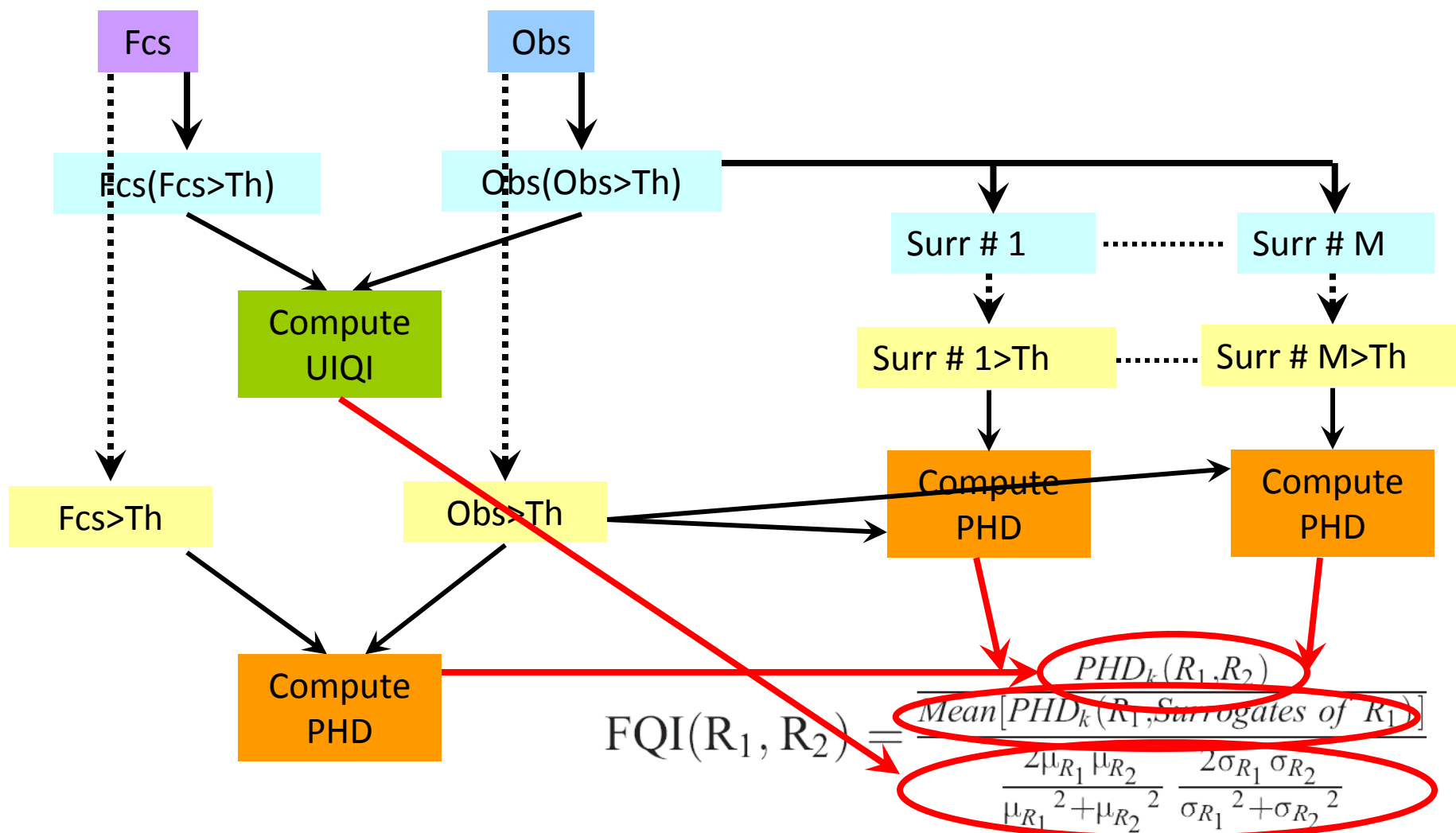
Ingredients:

- (a) Partial Hausdorff distance (PHD) → Displacement Error
- (b) (Modified) Universal Image Quality Index (UIQI) → Amplitude Error
- (c) Surrogates

References:

- (1) Basu, S., Dodov, B., and Foufoula-Georgiou, E. (2003). A novel measure for QPF verification and its usefulness in multimodel ensemble forecasting, *Geophys. Res. Abstracts*, 5, 04323, European Geophysical Society.
- (2) Venugopal, V., Basu, S., and Foufoula-Georgiou, E. (2005). A new metric for comparing precipitation patterns with an application to ensemble forecasts, *J. Geophys. Res.*, 110, D08111, doi:10.1029/2004JD005395.

Forecast Quality Index: Framework



Forecast Quality Index: Parameter Selection & Notations Used

Displacement Error:

75 percentile is used for Partial Hausdorff Distance calculation.

10 surrogates are created for observations (i.e., geom000_g240.txt and fake000_g240.txt)

Thresholds of 1,5,10, and 20 mm are used for 'perturbed' cases.

H_i denotes displacement error between observed and i-th forecast.

$\langle H_s \rangle$ denotes mean displacement errors between observation and its surrogates.

Amplitude Error:

In the original FQI approach, UIQI was calculated for 'only' non-zero pixels. This eliminated the influence of domain size on the FQI results. However, this approach fails for the 'geometric' cases 3 and 5 (same mean and variance for non-zero pixels).

To circumvent this issue, we have also reported UIQI based on 'all' the pixels.

Q_i denotes amplitude error between observed and i-th forecast (non-zero pixels).

Q_{ix} denotes amplitude error between observed and i-th forecast (all pixels).

Forecast Quality Index: Parameter Selection & Notations Used

Total Error:

FQI: forecast quality index based on Q_i

FQIx: forecast quality index based on Q_{ix}

Geometric Cases: Summary

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
1	41	1	1	0.19	0.19
2	191	1	1	0.87	0.87
3	145	1	0.38	0.66	1.72
4	141	1	1	0.64	0.64
5	186	1	0.16	0.85	5.13

$$\langle H_s \rangle = 220$$

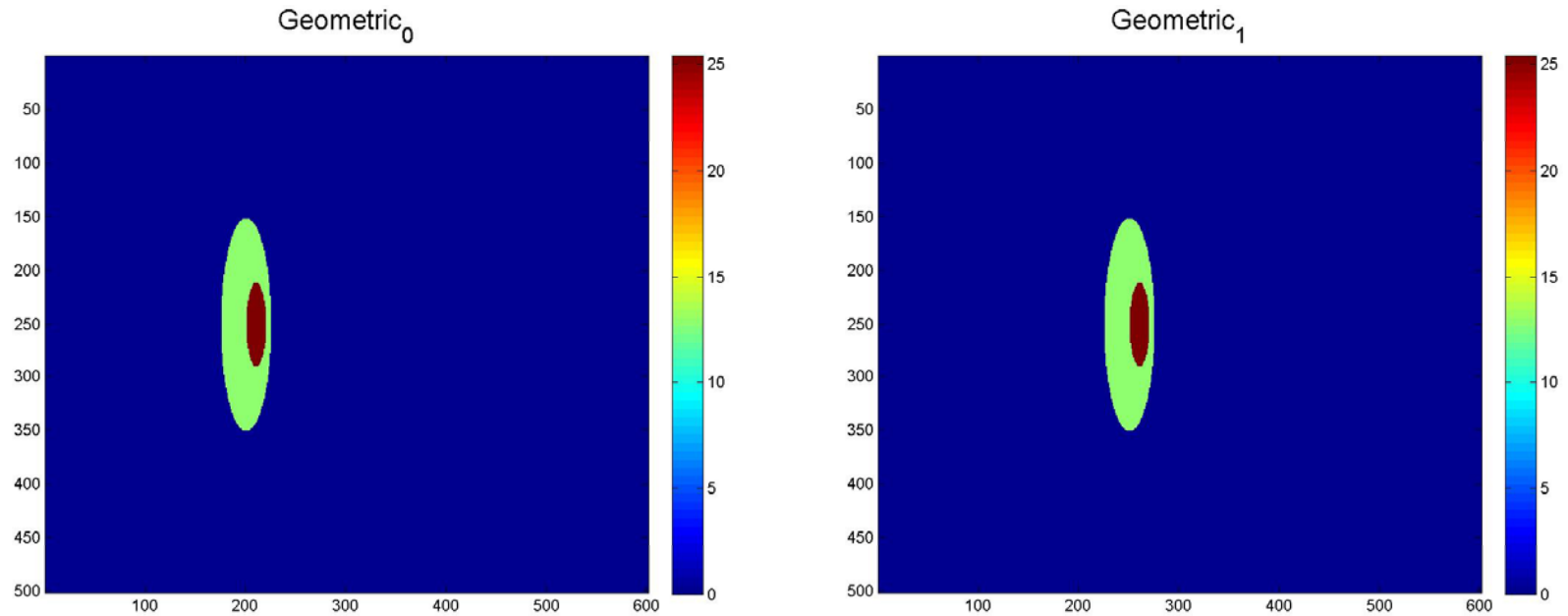
NOTE:

Lower H_i : better forecast;

Higher Q_i : better forecast;

Lower FQI (or FQIx): better forecast.

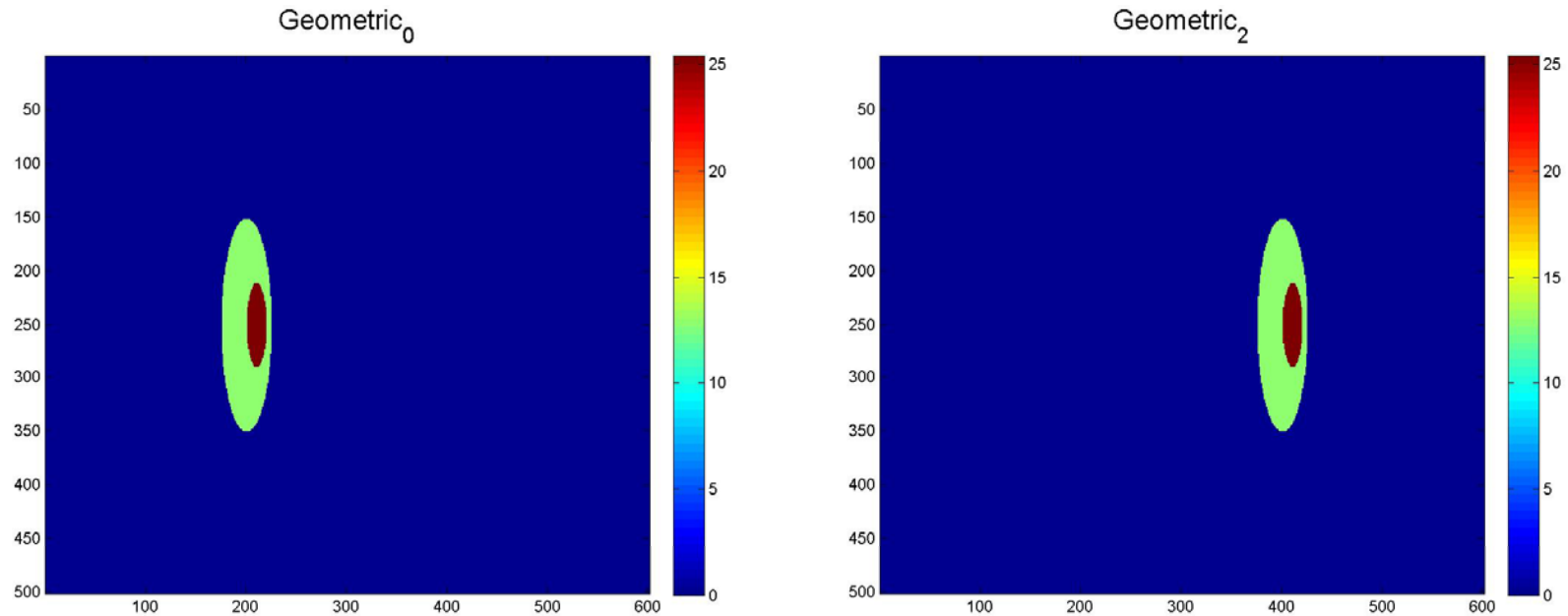
Geometric Cases



Case 1: 50 points to the right

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
1	41	1	1	0.19	0.19

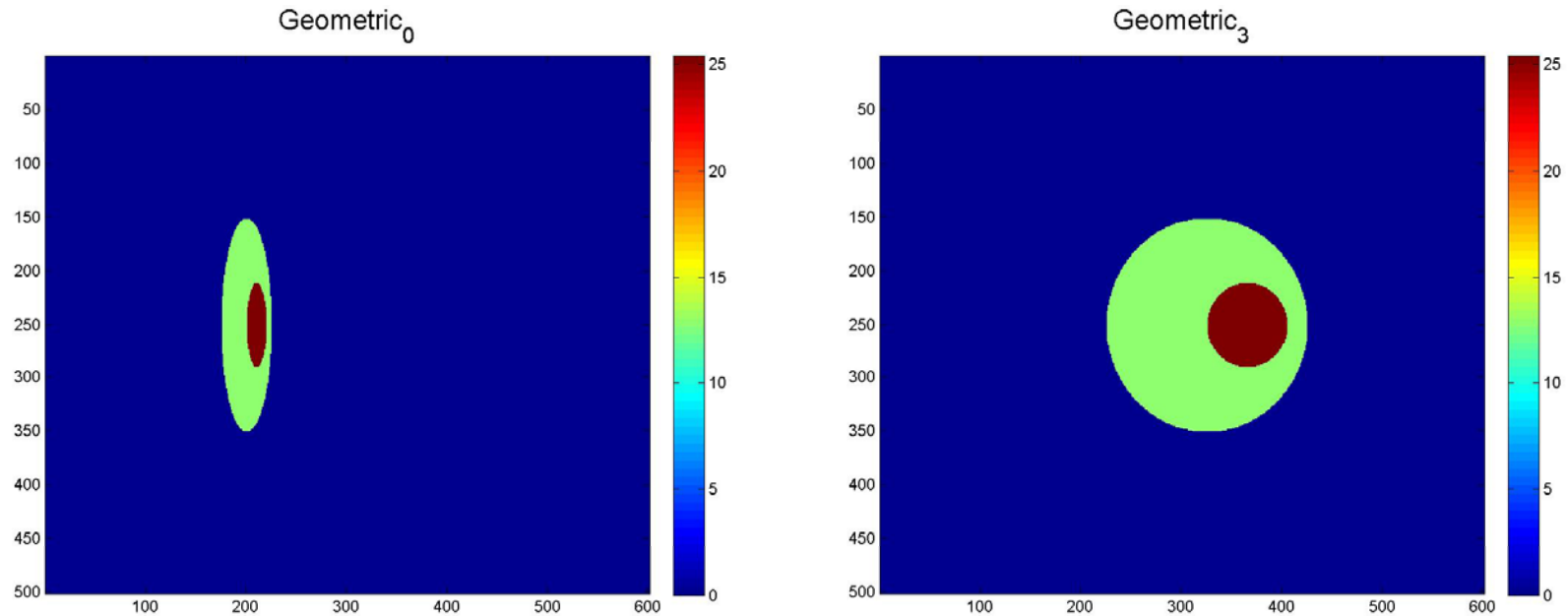
Geometric Cases (Cont...)



Case 2: 200 points to the right

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
2	191	1	1	0.87	0.87

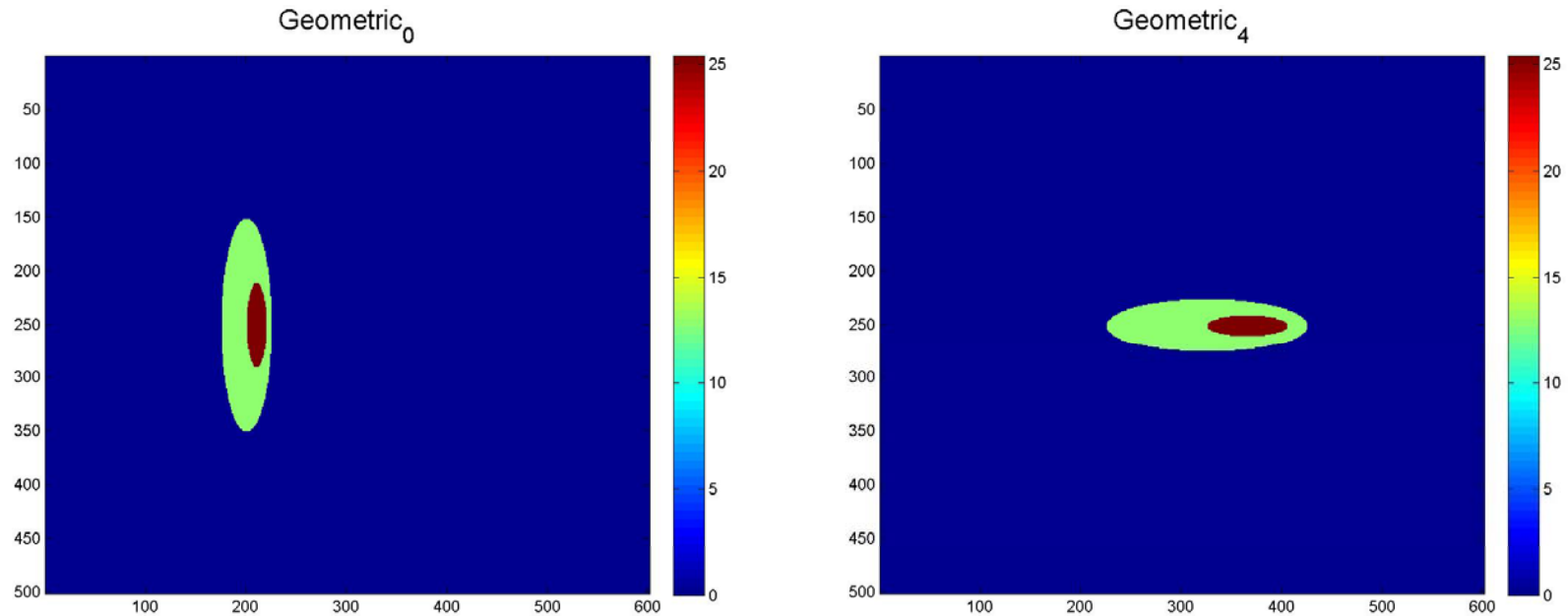
Geometric Cases (Cont...)



Case 3: 125 points to the right and big

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
3	145	1	0.38	0.66	1.72

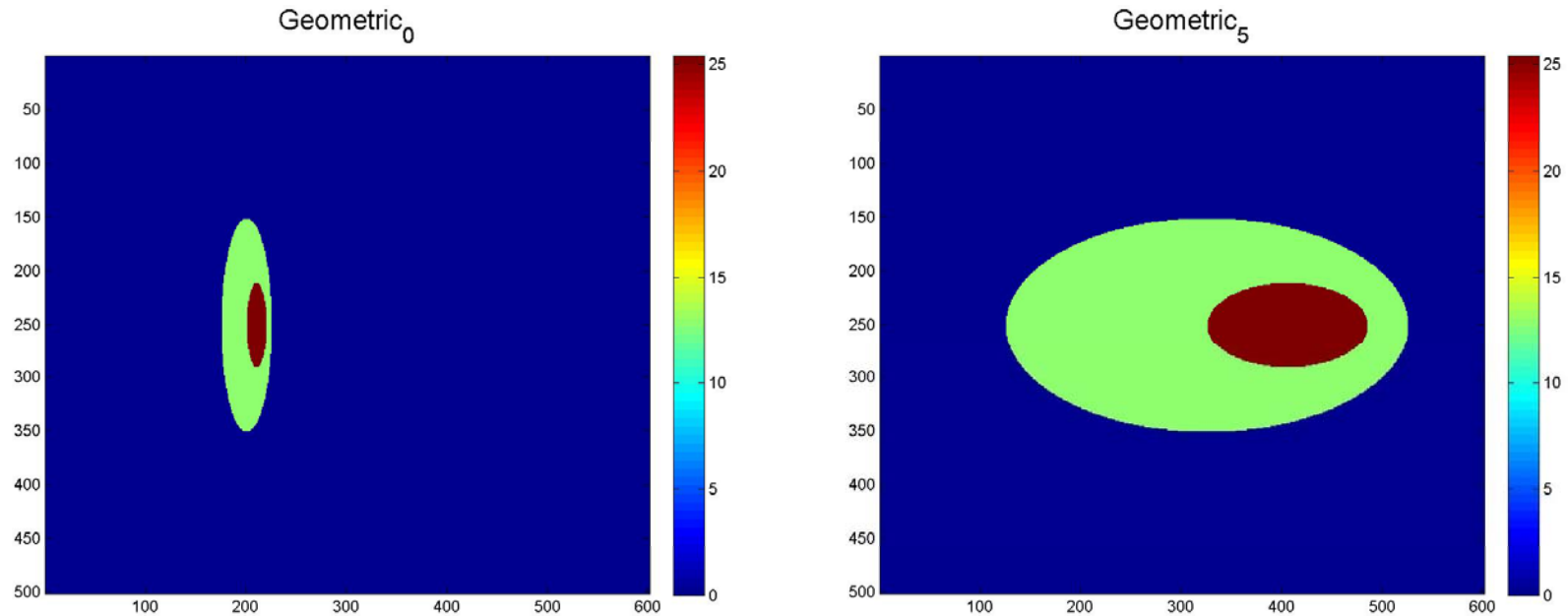
Geometric Cases (Cont...)



Case 4: 125 points to the right and rotated

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
4	141	1	1	0.64	0.64

Geometric Cases (Cont...)



Case 5: 125 points to the right and huge

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
5	186	1	0.16	0.85	5.13

Perturbed Cases: Summary (Threshold = 1 mm)

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
1	2	1	1	0.03	0.03
2	5	1	1	0.07	0.07
3	10	1	1	0.13	0.13
4	18	1	1	0.24	0.24
5	48	1	1	0.64	0.65
6	9	0.90	0.85	0.13	0.14
7	16	0.94	0.95	0.23	0.23

$\langle HS \rangle = 75$

NOTE:

Lower H_i : better forecast;

Higher Q_i : better forecast;

Lower FQI (or FQIx): better forecast.

Perturbed Cases: Summary (Threshold = 5 mm)

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
1	3	1	1	0.03	0.03
2	7	1	1	0.06	0.06
3	15	1	1	0.13	0.13
4	29	1	1	0.25	0.25
5	71	1	1	0.61	0.61
6	15	0.90	0.81	0.16	0.16
7	16	1	0.98	0.14	0.14

$\langle HS \rangle = 116$

NOTE:

Lower H_i : better forecast;

Higher Q_i : better forecast;

Lower FQI (or FQIx): better forecast.

Perturbed Cases: Summary (Threshold = 10 mm)

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
1	4	1	1	0.03	0.03
2	9	1	1	0.06	0.06
3	19	1	1	0.13	0.13
4	38	1	1	0.27	0.27
5	90	1	1	0.64	0.64
6	19	0.90	0.75	0.15	0.18
7	19	1	0.98	0.13	0.14

$\langle HS \rangle = 141$

NOTE:

Lower H_i : better forecast;

Higher Q_i : better forecast;

Lower FQI (or FQIx): better forecast.

Perturbed Cases: Summary (Threshold = 20 mm)

Case #	H_i	Q_i	Q_{ix}	FQI	FQIx
1	6	1	1	0.03	0.03
2	13	1	1	0.06	0.06
3	26	1	1	0.12	0.12
4	52	1	1	0.24	0.24
5	106	1	1	0.48	0.48
6	25	0.94	0.47	0.12	0.24
7	27	1	0.97	0.12	0.13

$\langle HS \rangle = 219$

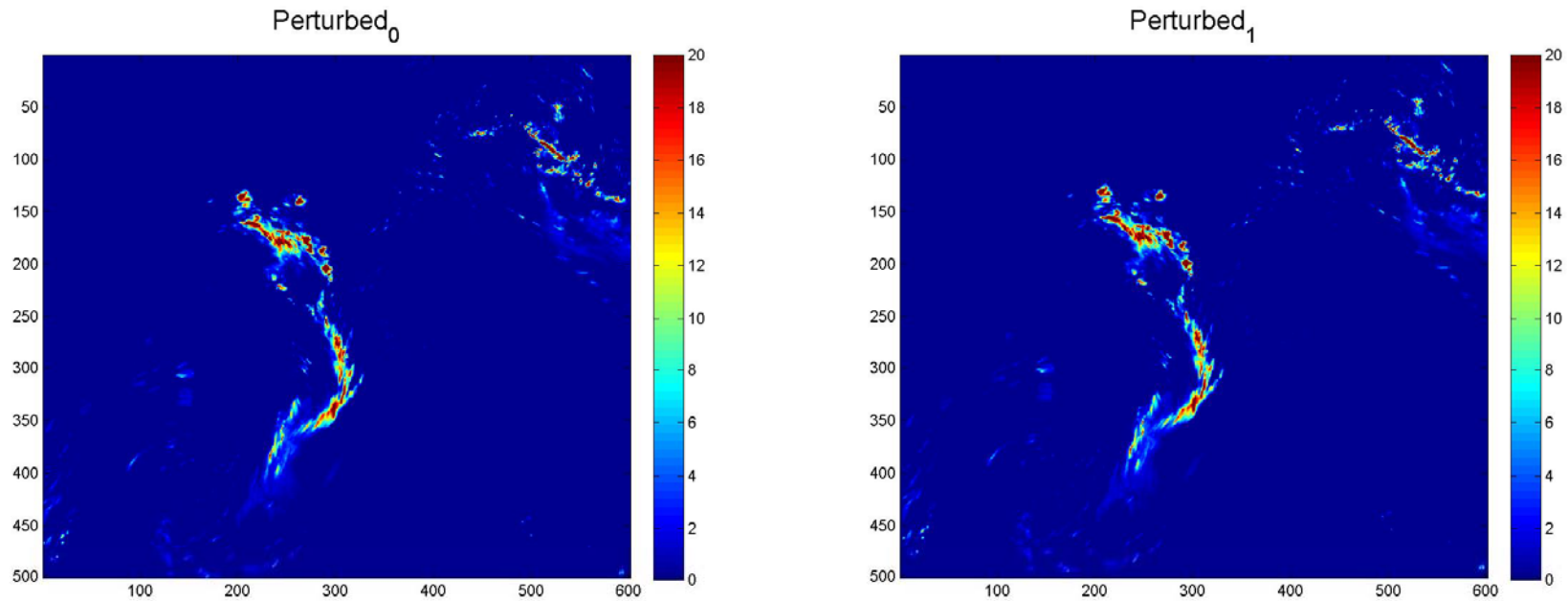
NOTE:

Lower H_i : better forecast;

Higher Q_i : better forecast;

Lower FQI (or FQIx): better forecast.

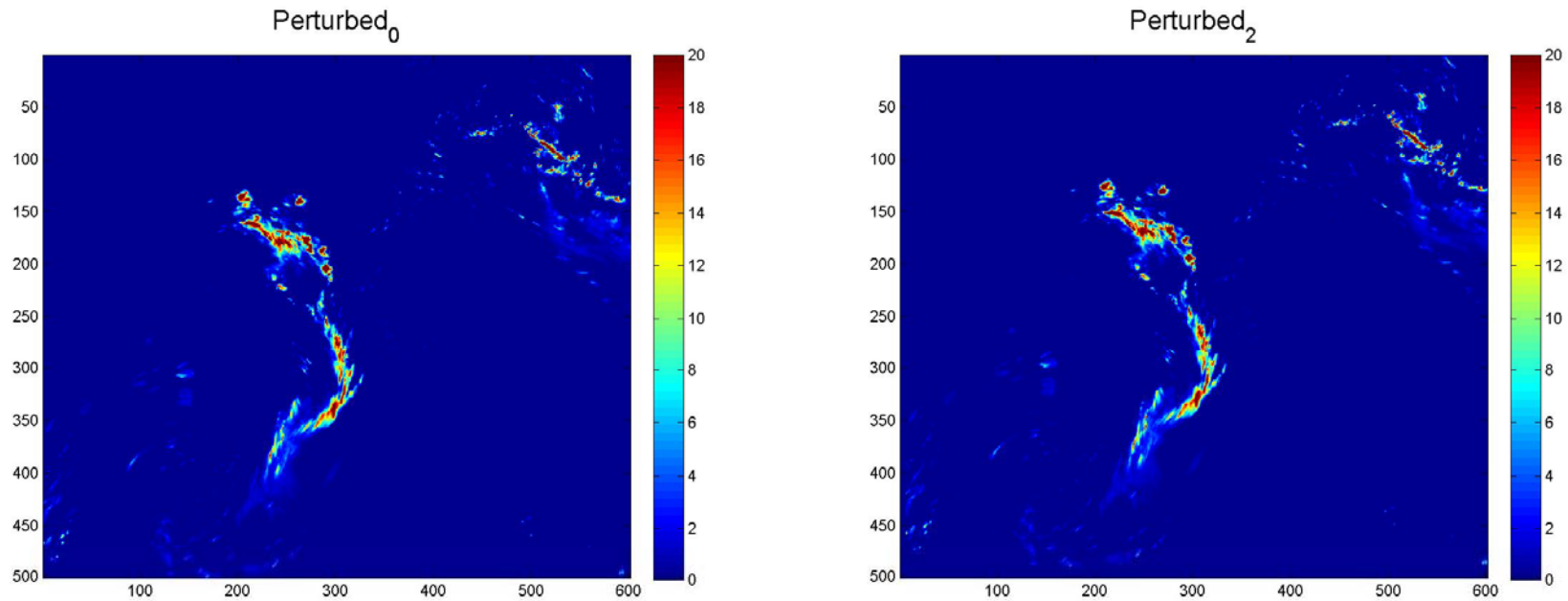
Perturbed Cases



Case 1: 3 points right, -5 pts up

Threshold	Hi	Qi	Qix	FQI	FQIx
1	2	1	1	0.03	0.03
5	3	1	1	0.03	0.03
10	4	1	1	0.03	0.03
20	6	1	1	0.03	0.03

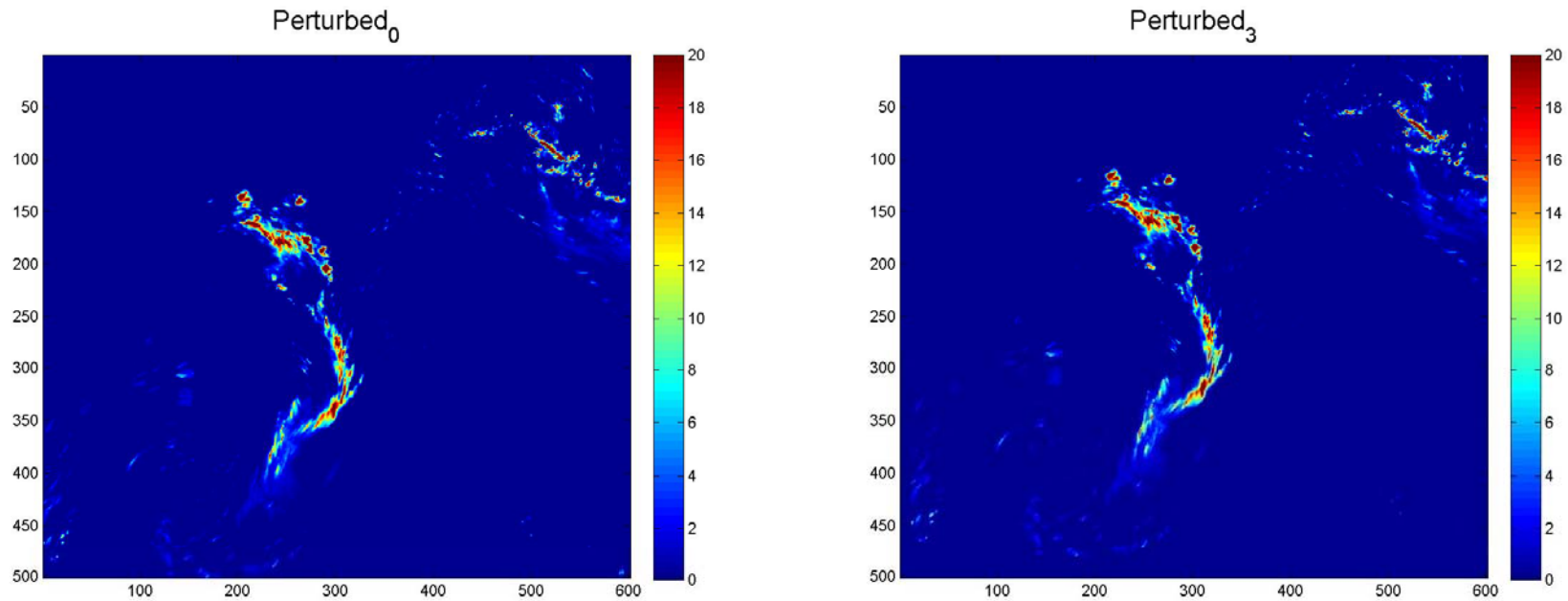
Perturbed Cases



Case 2: 6 points right, -10 pts up

Threshold	Hi	Qi	Qix	FQI	FQIx
1	5	1	1	0.07	0.07
5	7	1	1	0.06	0.06
10	9	1	1	0.06	0.06
20	13	1	1	0.06	0.06

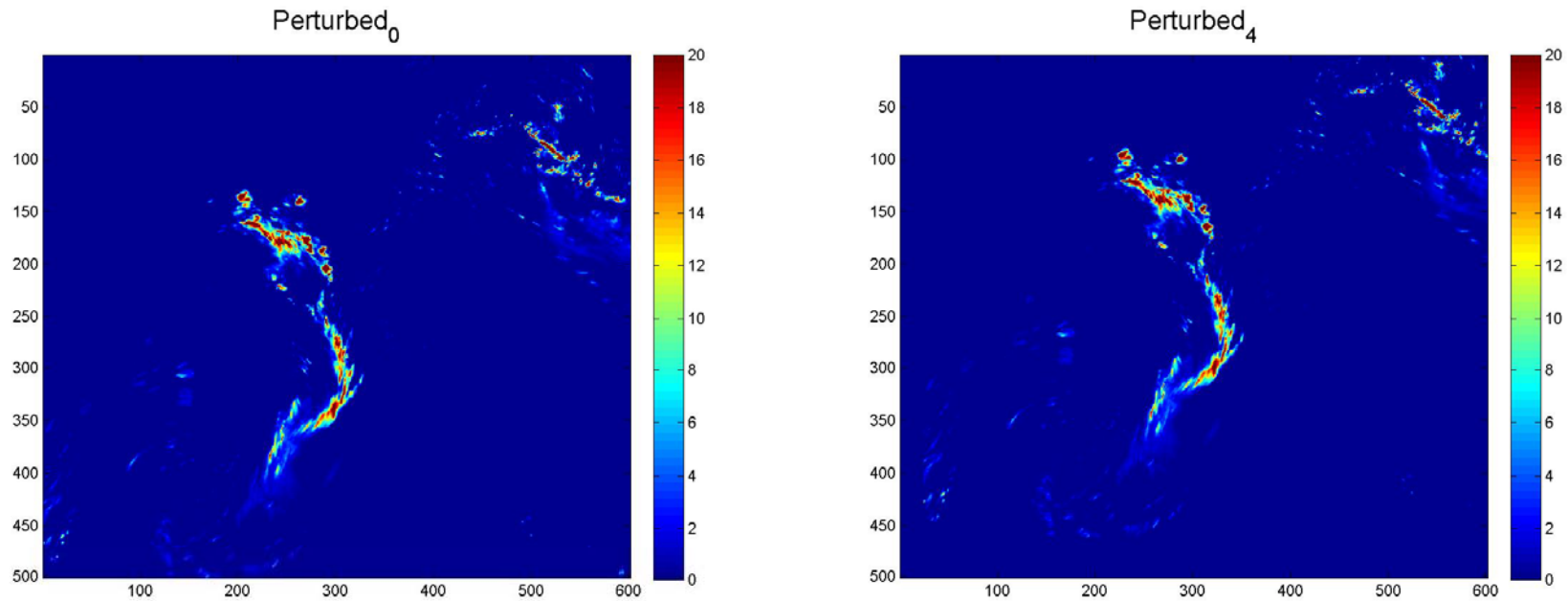
Perturbed Cases



Case 3: 12 points right, -20 pts up

Threshold	Hi	Qi	Qix	FQI	FQIx
1	10	1	1	0.13	0.13
5	15	1	1	0.13	0.13
10	19	1	1	0.13	0.13
20	26	1	1	0.12	0.12

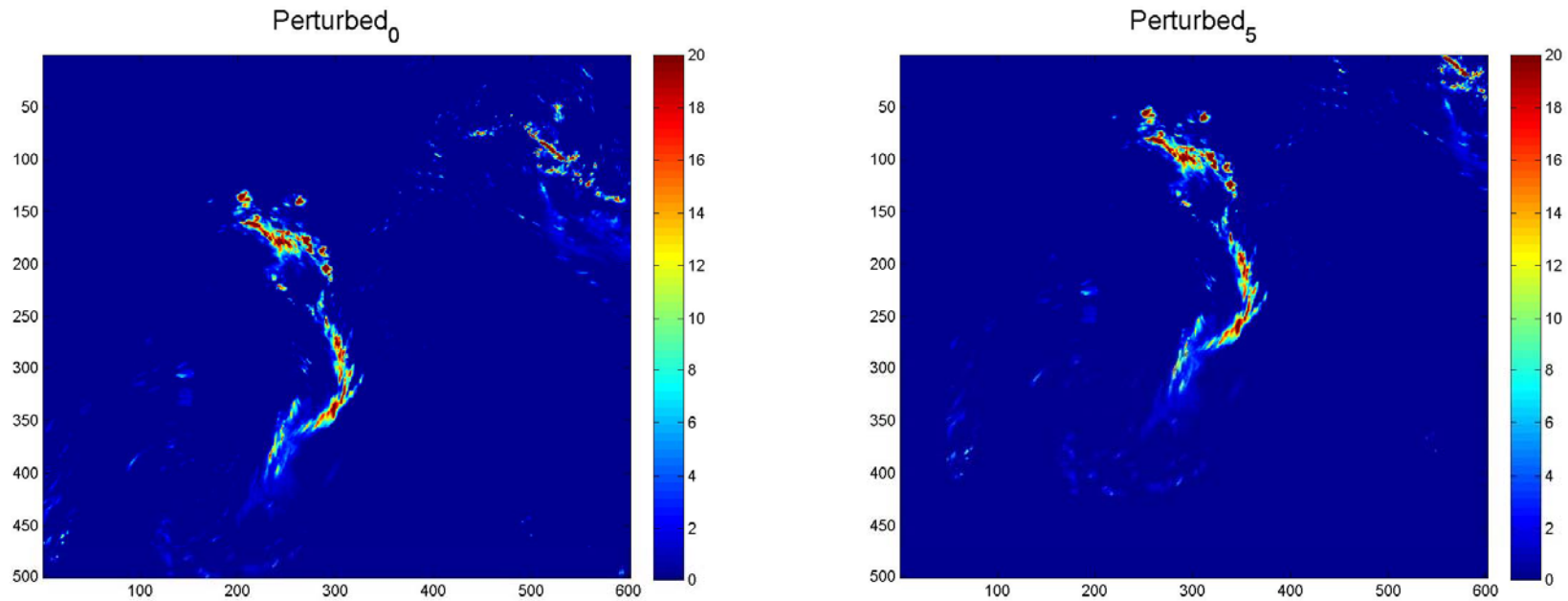
Perturbed Cases



Case 4: 24 points right, -40 pts up

Threshold	Hi	Qi	Qix	FQI	FQIx
1	18	1	1	0.24	0.24
5	29	1	1	0.25	0.25
10	38	1	1	0.27	0.27
20	52	1	1	0.24	0.24

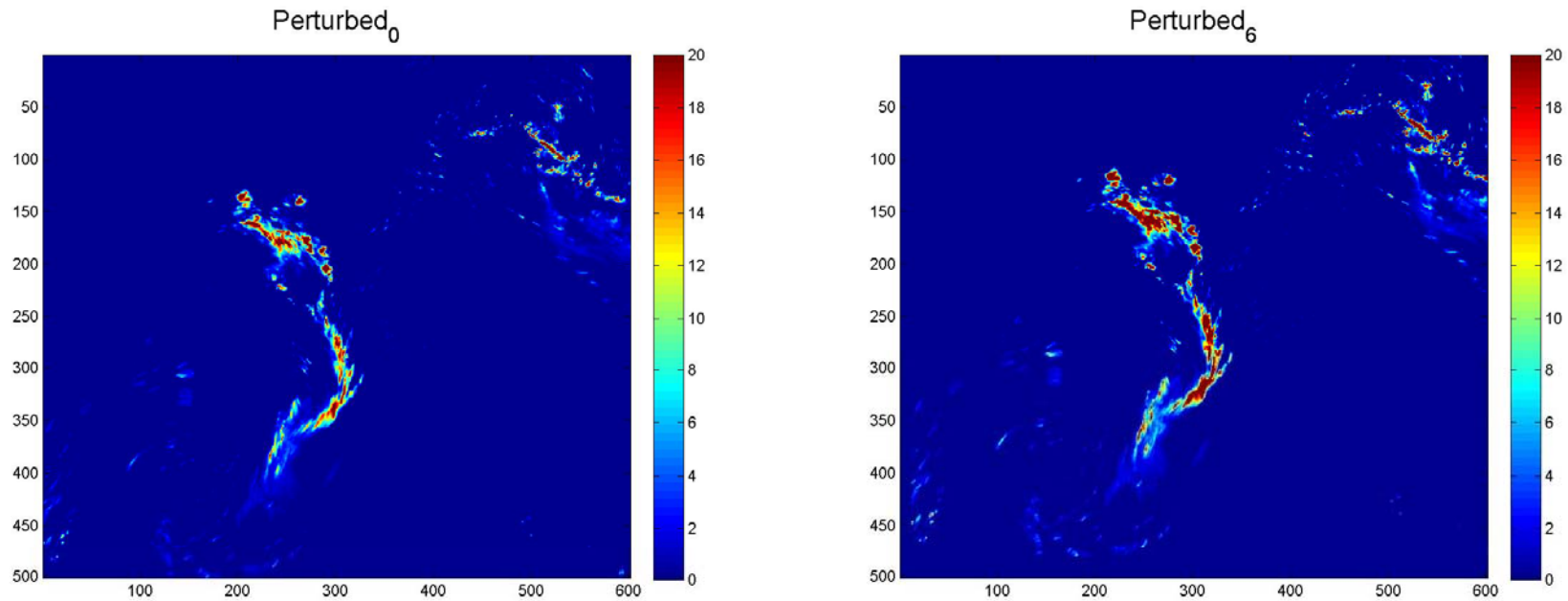
Perturbed Cases



Case 5: 48 points right, -80 pts up

Threshold	Hi	Qi	Qix	FQI	FQIx
1	48	1	1	0.64	0.65
5	71	1	1	0.61	0.61
10	90	1	1	0.64	0.64
20	106	1	1	0.48	0.48

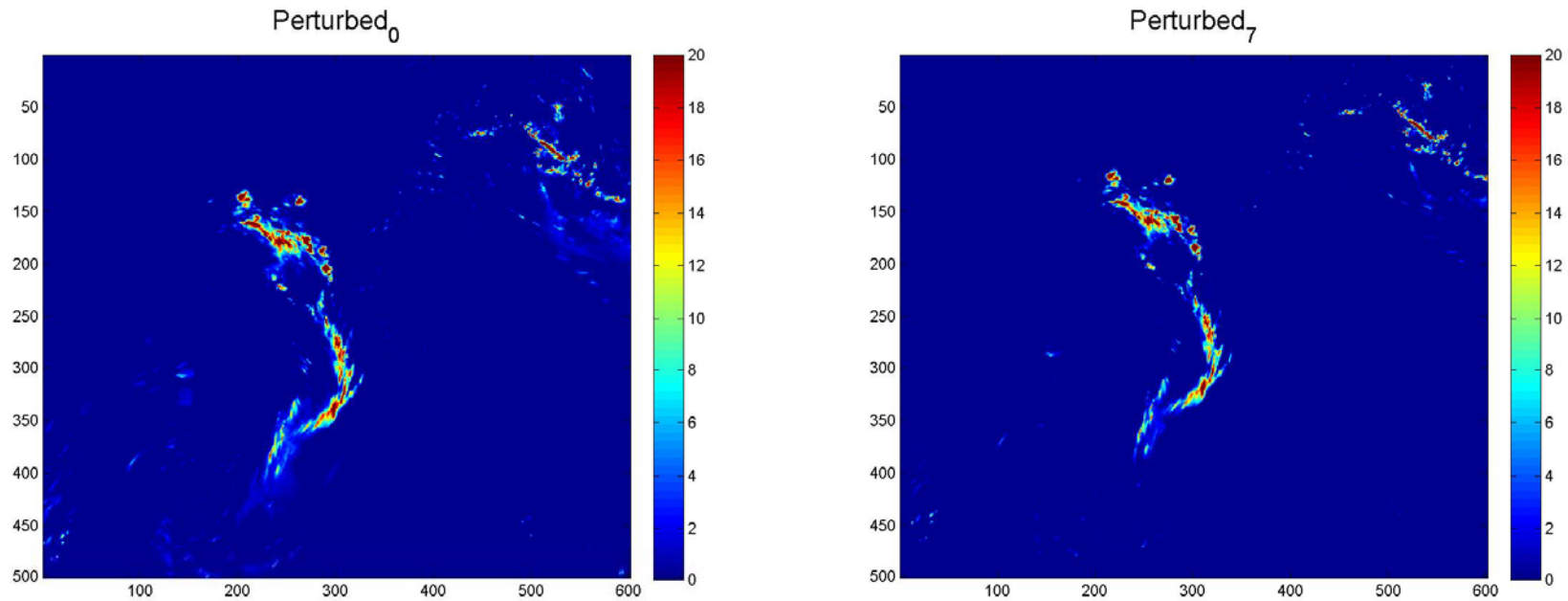
Perturbed Cases



Case 6: 12 points right, -20 pts up, times 1.5

Threshold	Hi	Qi	Qix	FQI	FQIx
1	9	0.90	0.85	0.13	0.14
5	15	0.90	0.81	0.16	0.16
10	19	0.90	0.75	0.15	0.18
20	25	0.94	0.47	0.12	0.24

Perturbed Cases



Case 7: 12 points right, -20 pts up, minus 0.05''

Threshold	Hi	Qi	Qix	FQI	FQIx
1	16	0.94	0.95	0.23	0.23
5	16	1	0.98	0.14	0.14
10	19	1	0.98	0.13	0.14
20	27	1	0.97	0.12	0.13