Spatial Forecast Verification methods overview

Eric Gilleland (EricG @ ucar.edu)
Research Application Laboratory
National Center for Atmospheric Research

SpatialVx: R package for performing spatial verification (in the works)

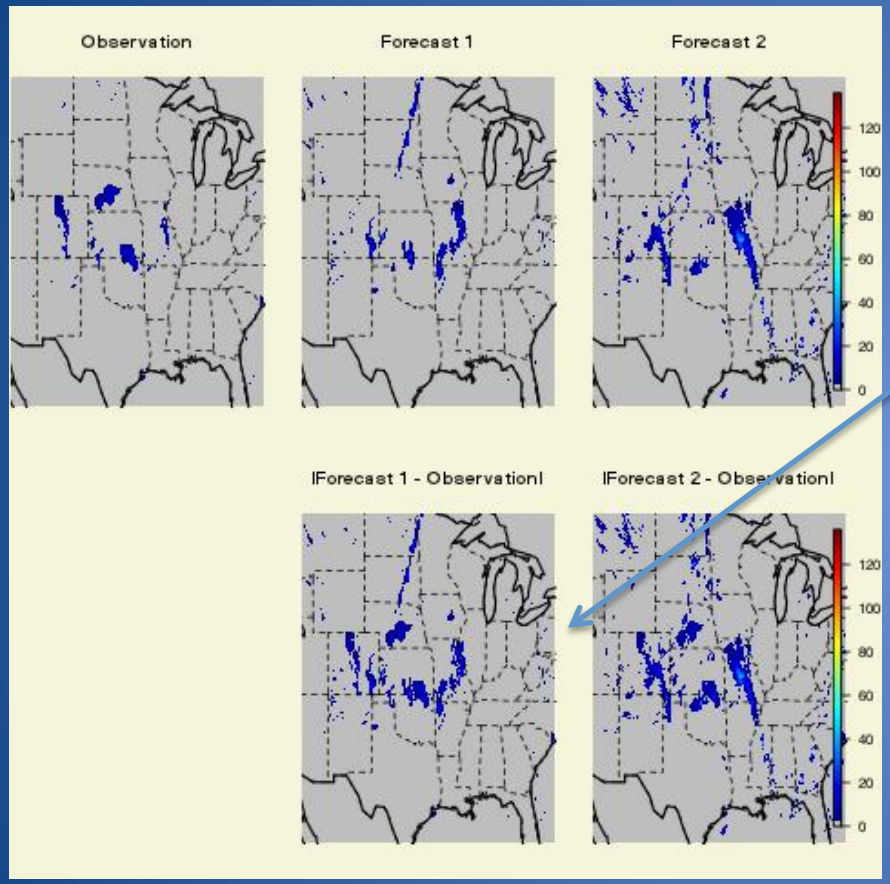
http://www.ral.ucar.edu/projects/icp/SpatialVx/

WRF/MET Tutorial, 23 – 24 January 2014
NCAR, Boulder, Colorado

Copyright 2014 NCAR

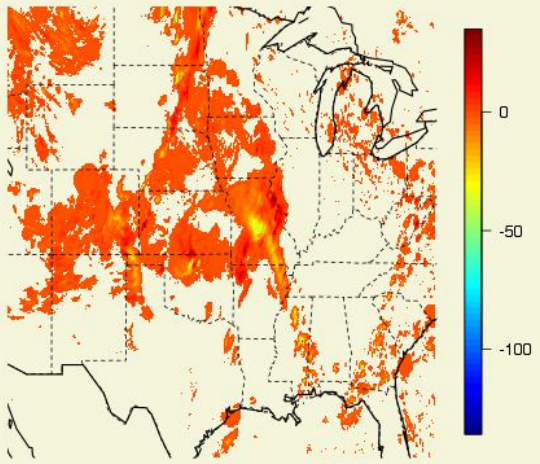
NCAR

Spatial Prediction Comparison Test

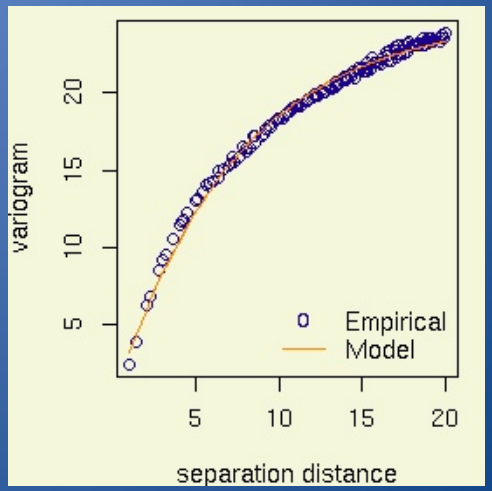


D1 D2

D = D1 – D2



Loss Differential Field

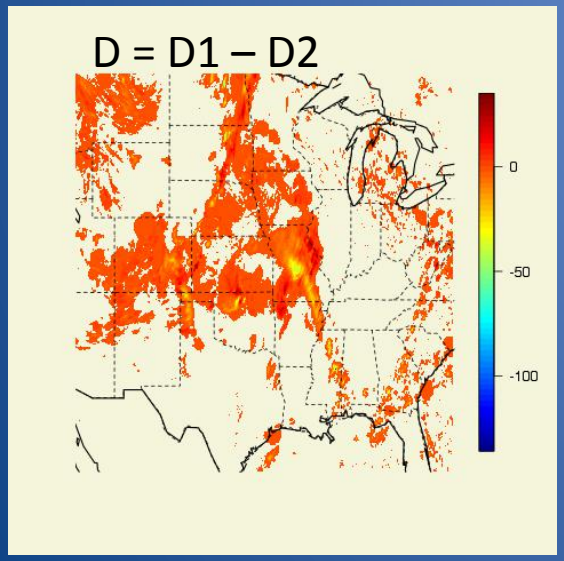


Copyright 2014 NCAR

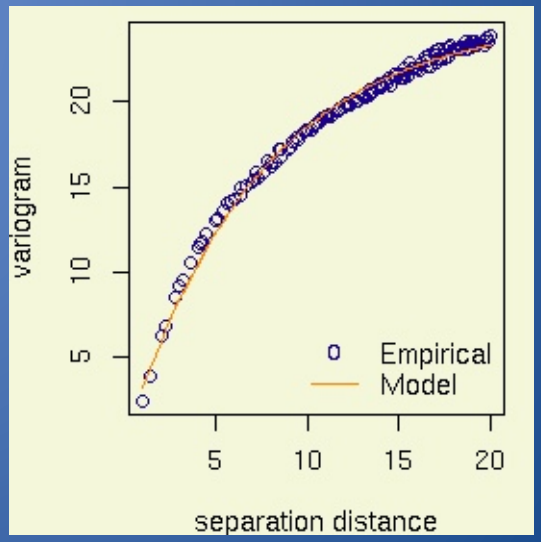
Spatial Prediction Comparison Test

Introduced by Hering and Genton (2011, *Technometrics*, **53**, 414 – 425)

Extension of the time series version introduced by Diebold and Mariano (1995, *J. Business and Economic Statistics*, **13**, 253 – 263).



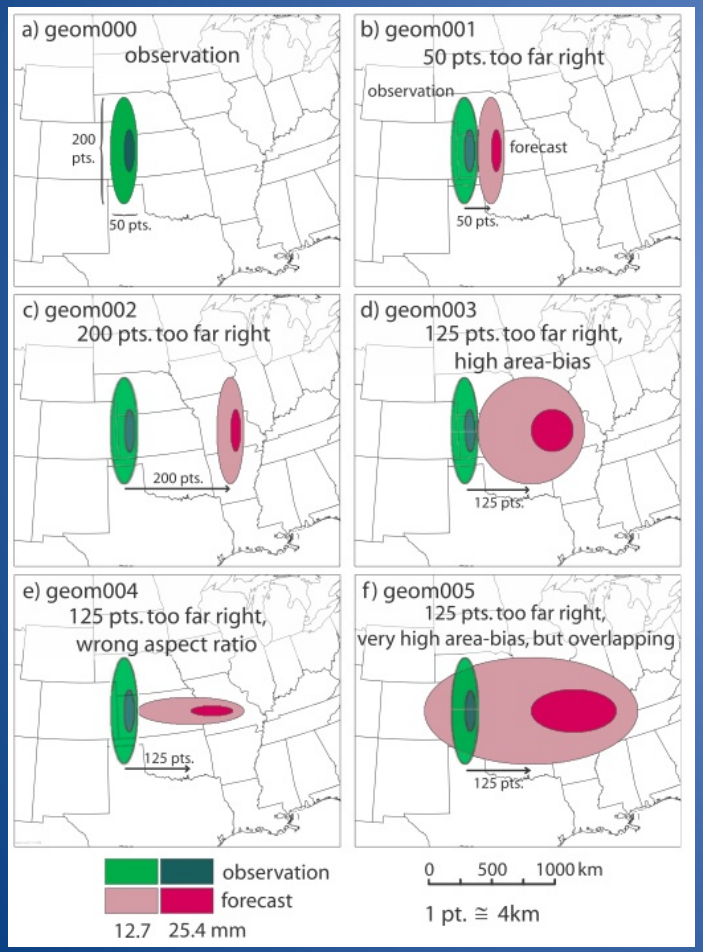
$$S = \frac{\bar{D}}{\sqrt{\mathrm{var}(\bar{D})}}$$

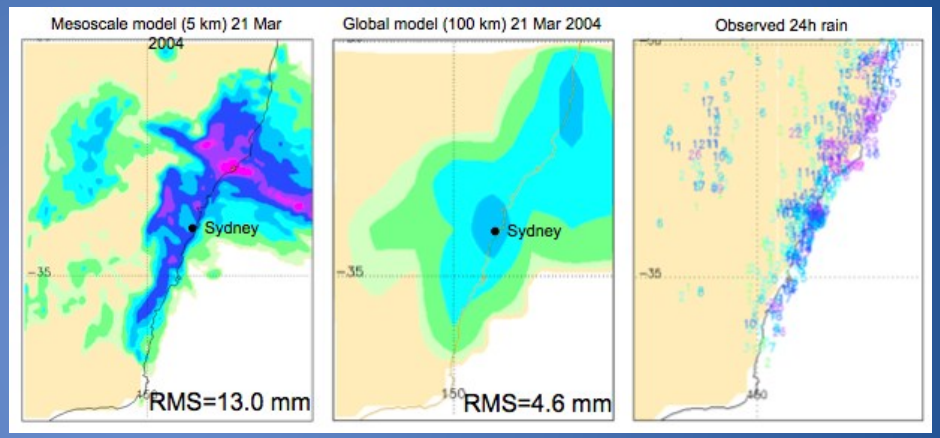


Copyright 2014 NCAR

Spatial Prediction Comparison Test

Accounting for Location Errors and Reducing Effects of Small Scale Errors





Traditional score	geom001/002/004	geom003	geom005
Accuracy	0.95	0.87	0.81
Frequency bias	1.00	4.02	8.03
Multiplicative intensity bias	1.00	4.02	8.04
RMSE (mm)	3.5	5.6	6.9
Bias-corrected RMSE (mm)	3.5	5.5	6.3
Correlation coefficient	−0.02	−0.05	0.20
Probability of detection	0.00	0.00	0.88
Probability of false detection	0.03	0.11	0.19
False alarm ratio	1.00	1.00	0.89
Hanssen–Kuipers discriminant (H-K)	−0.03	−0.11	0.69
Threat score or CSI	0.00	0.00	0.11
Equitable threat score or GSS	−0.01	−0.02	0.08
HSS	−0.03	−0.04	0.16

Above Figure from Beth Ebert

Fig. 1 and Table 2 from Ahijevych *et al*. (2009, *WAF*, **24**, 1485 – 1497)

Copyright 2014 NCAR

Spatial Prediction Comparison Test

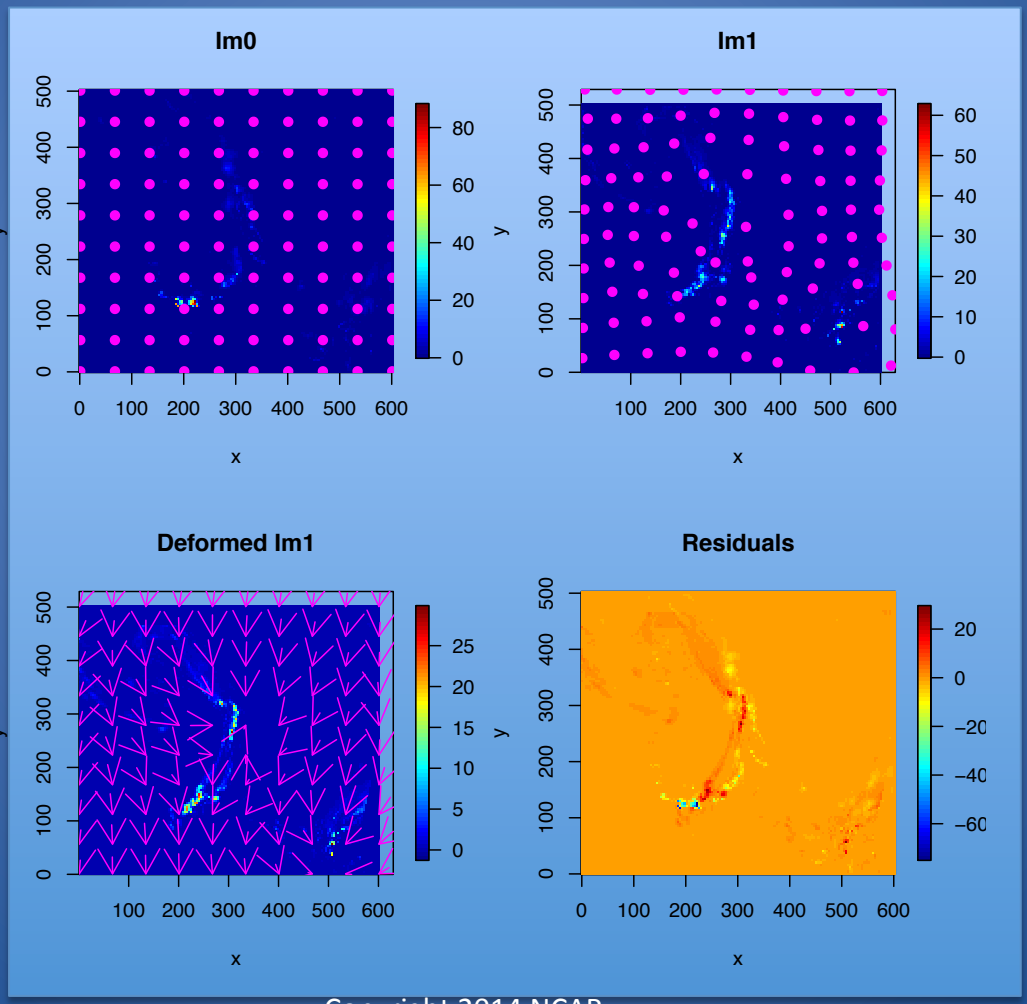
Accounting for Location Errors and Reducing Effects of Small Scale Errors

Above Figure from Johan Lindström

Copyright 2014 NCAR

Spatial Prediction Comparison Test

Accounting for Location Errors and Reducing Effects of Small-Scale Errors



Copyright 2014 NCAR

Spatial Prediction Comparison Test

Accounting for Location Errors and Reducing Effects of Small Scale Errors

Loss at each point =

Distance from original location of each point to warped location

Loss at each point between observation value and warped value

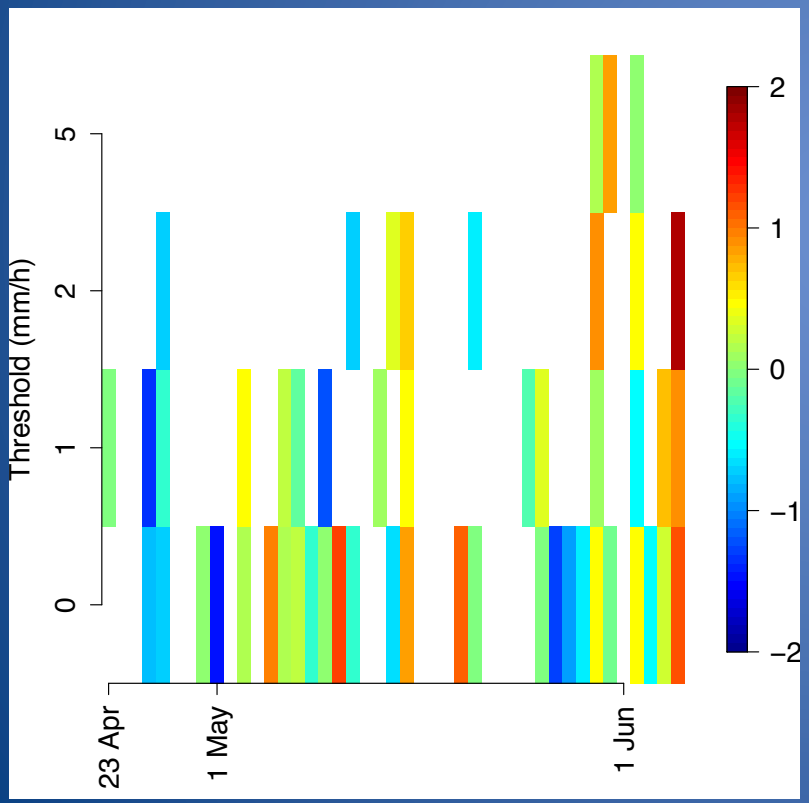
G. (2013, *MWR*, **141** (1), 340 – 355)

Copyright 2014 NCAR

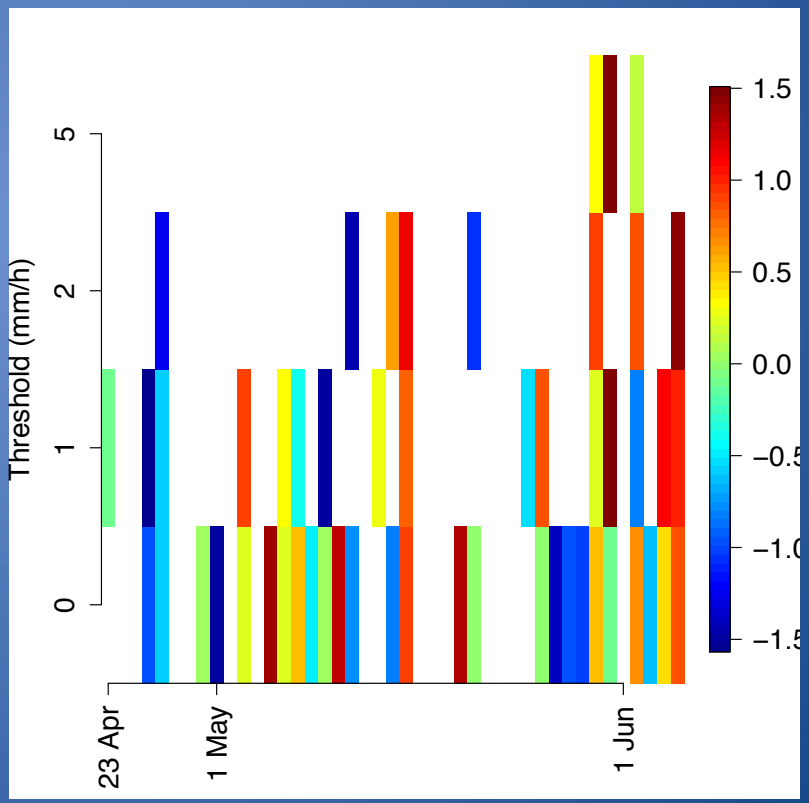
Spatial Prediction Comparison Test

Accounting for Location Errors and Reducing Effects of Small Scale Errors

$\bar{D}$



S



Copyright 2014 NCAR

Methods Overview

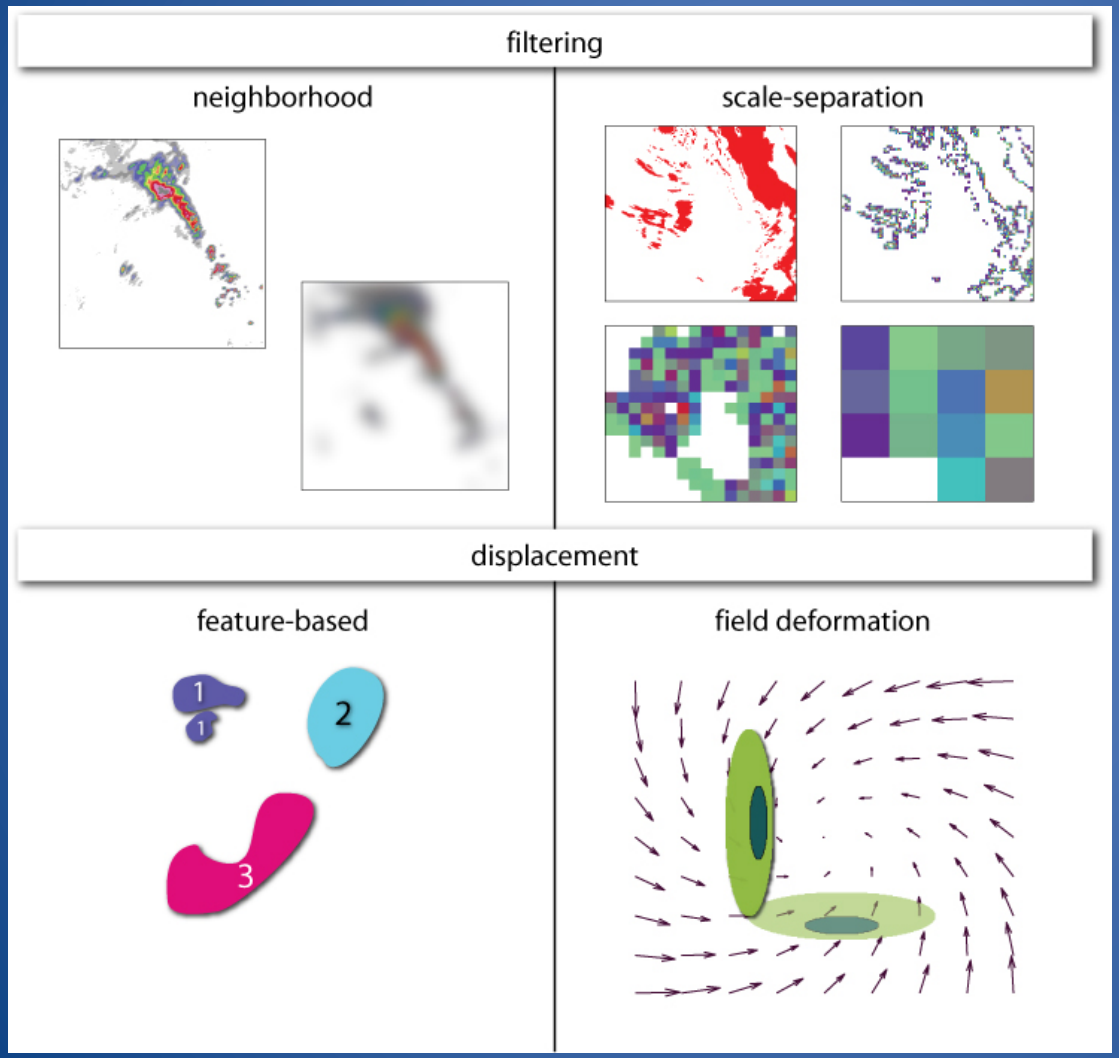
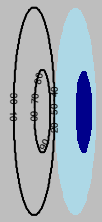


Fig. 2 from G. *et al*. (2010, *BAMS*, **91** (10), 1365 – 1373)

Copyright 2014 NCAR

Filter Methods: Smoothing

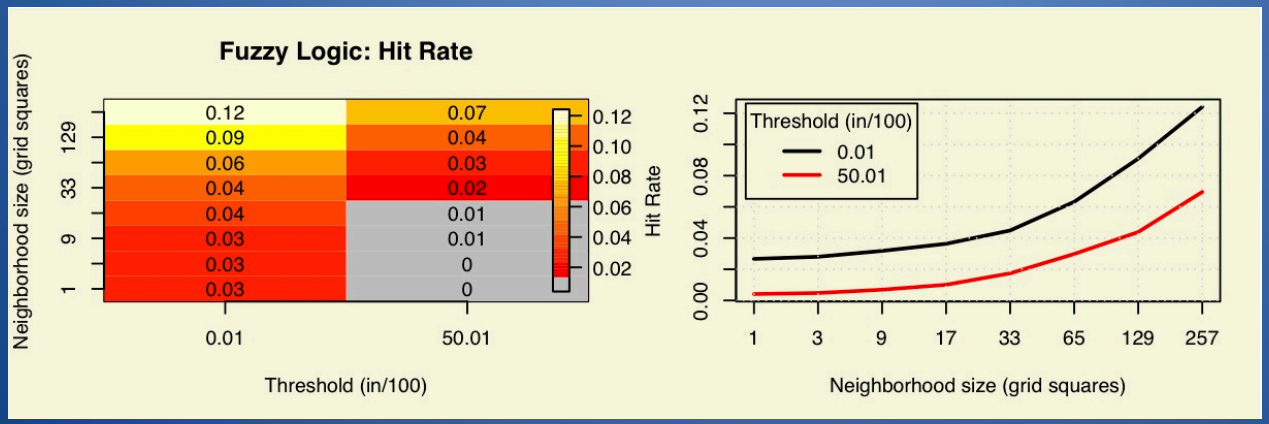
50 grid points displacement error



The fuzzy logic approach compares smoothed indicator fields:

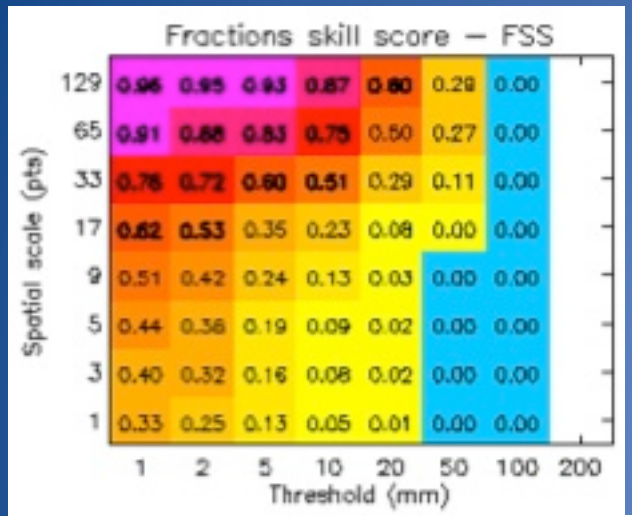
a new contingency table where
hits = sum of the minima between the two fields at each (smoothed) grid point.

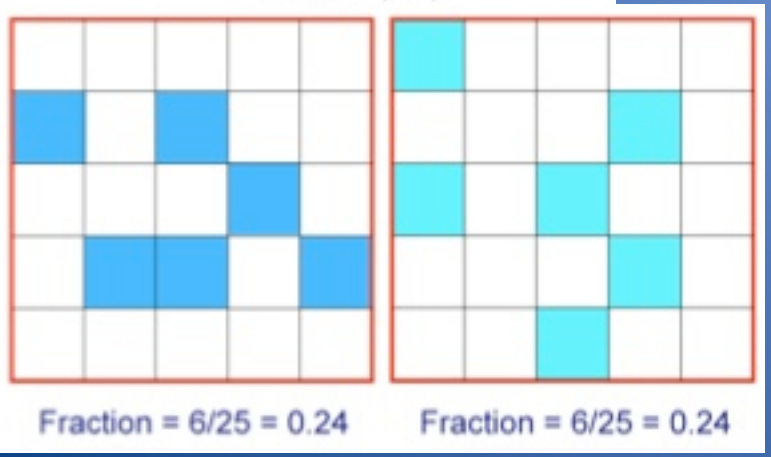
misses = sum of the minima between the smoothed indicator of the observed field and 1 – the smoothed indicator of the forecast field, and so on (cf. Ebert 2008, *Meteorol. Appl.*, **15**, 51 - 64).



Copyright 2014 NCAR

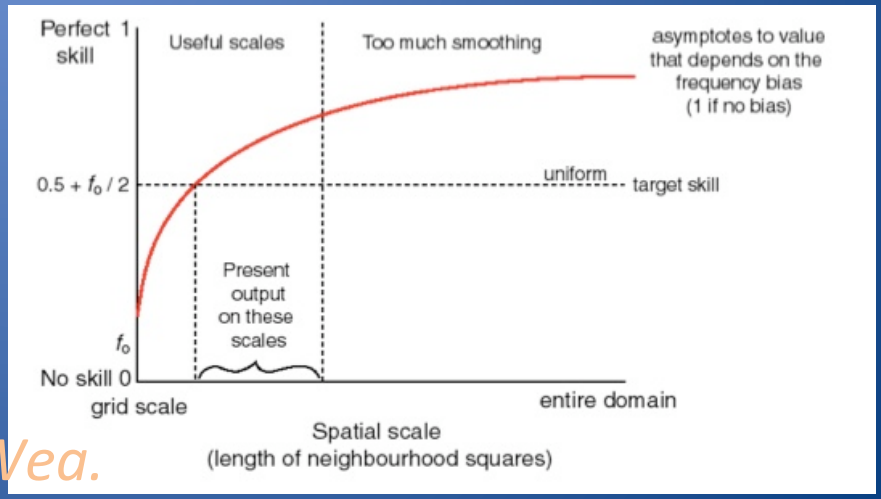
Filter Methods: Smoothing Fractions Skill Score (FSS)





$$\mathrm{FSS} = 1 - \frac{\sum_{s=1}^{n} \left(\hat{p}_s - p_s\right)^2}{\sum_{s=1}^{n} \hat{p}_s^2 + \sum_{s=1}^{n} p_s^2}$$

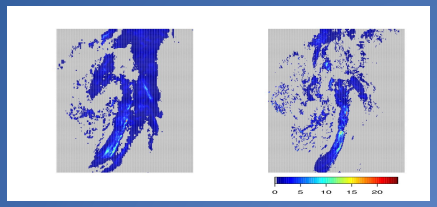
$p_s, \hat{p}_s$ are the fraction of events in a neighborhood centered on point s.

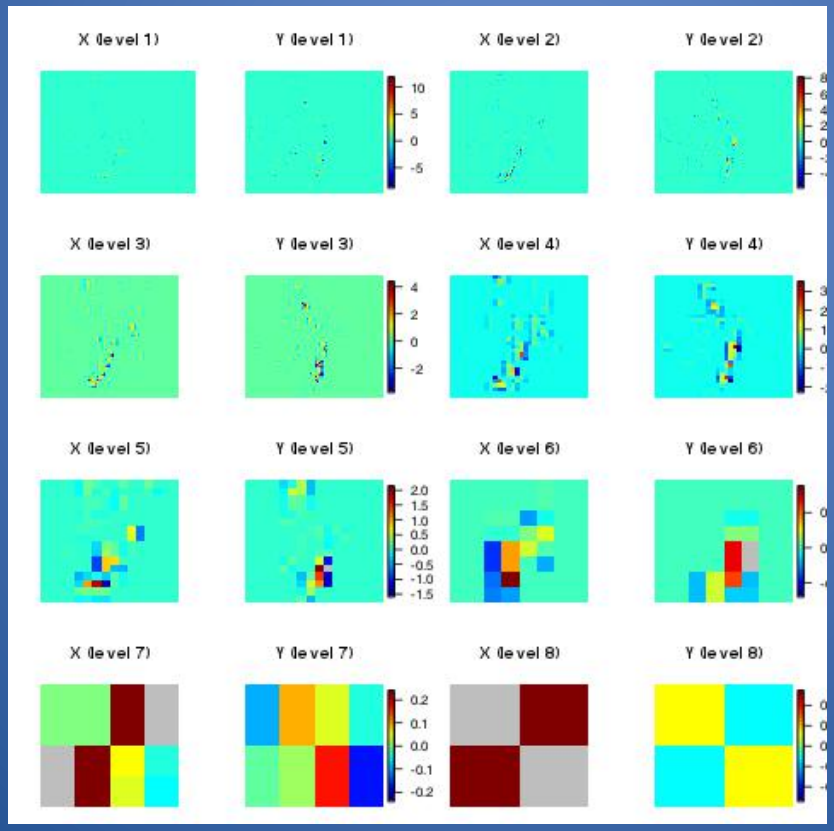


Roberts and Lean (2008, *Mon. Wea. Rev.* , **136**, 78 – 96)

Copyright 2014 NCAR

Filter Methods: Scale separation

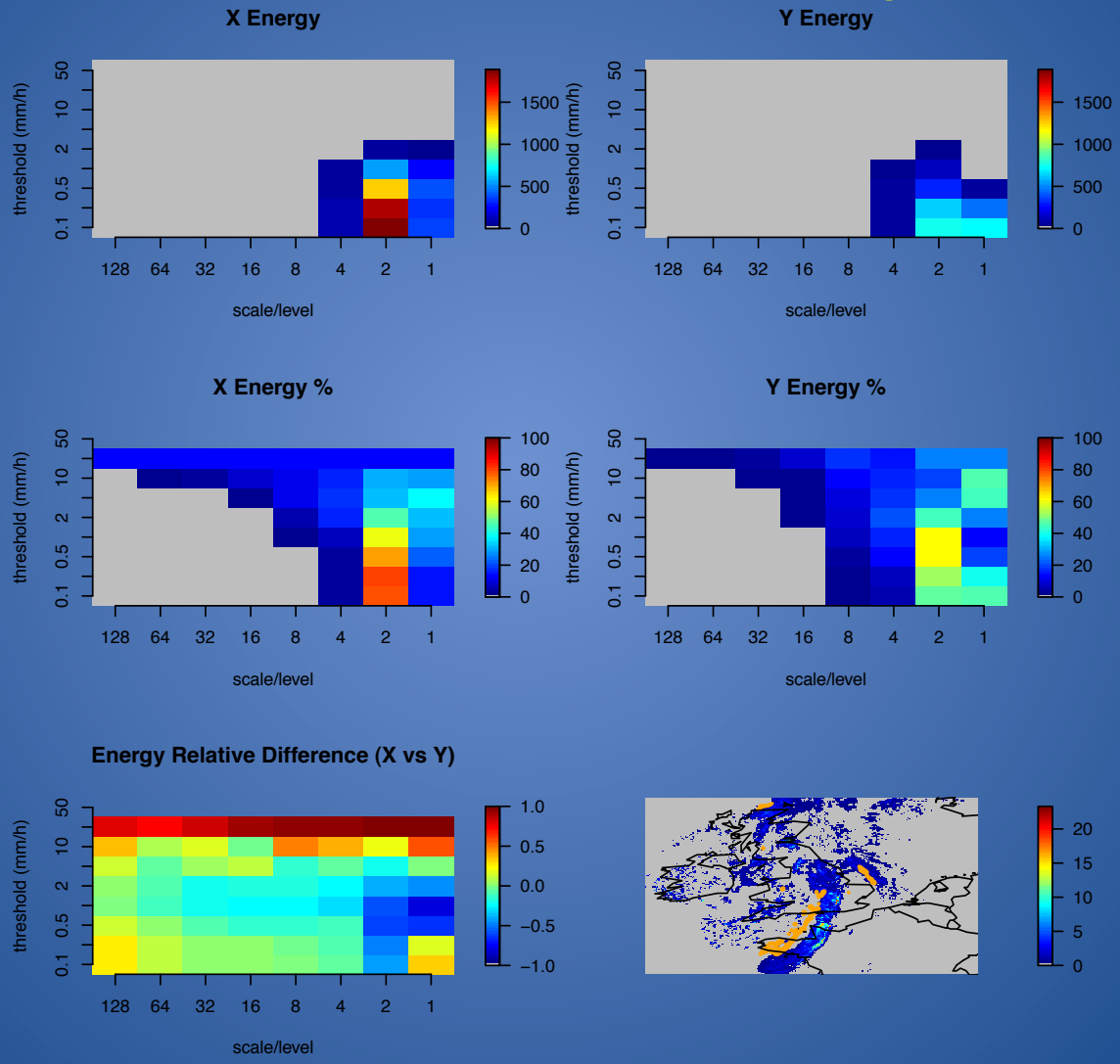




Copyright 2014 NCAR

Filter Methods: Scale separation

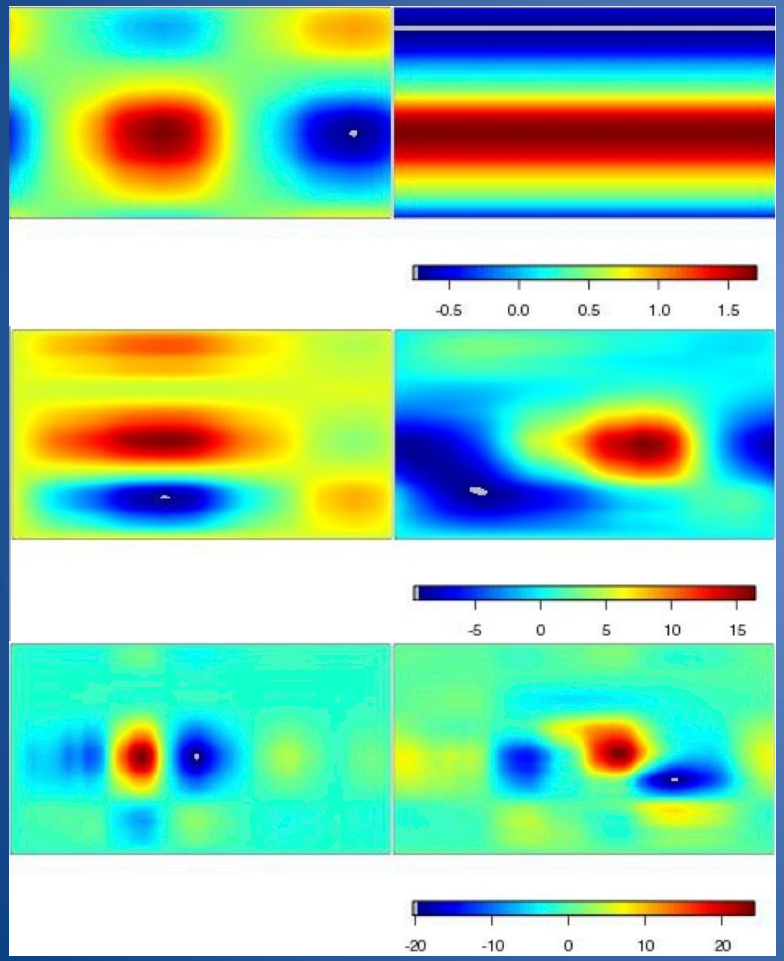
Intensity-Scale

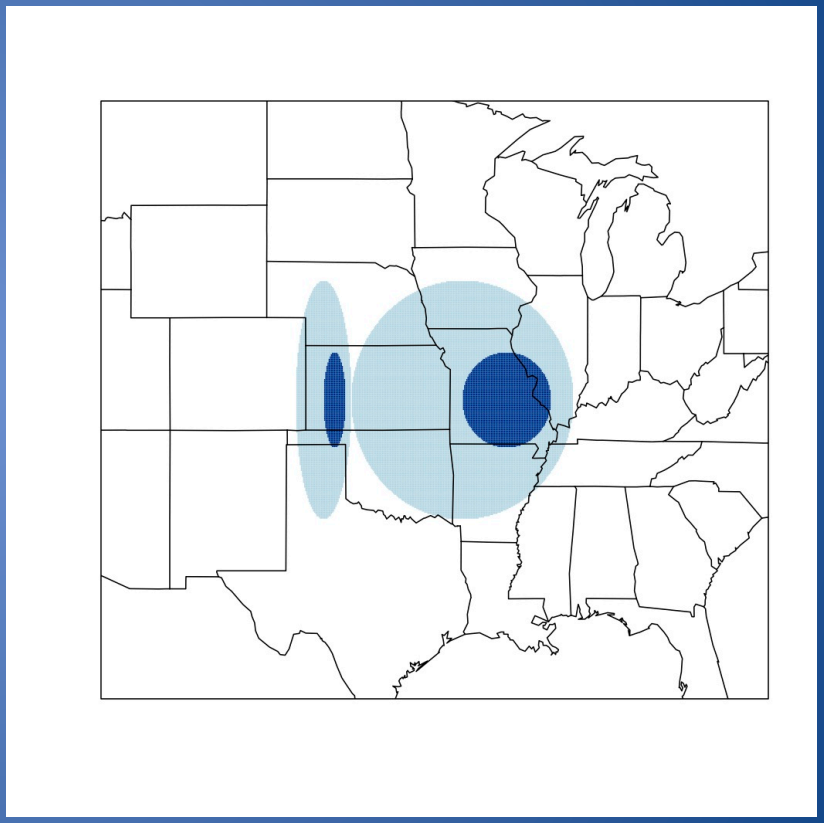


Copyright 2014 NCAR

Filter Methods: Scale separation

Wavelets





Copyright 2014 NCAR

Displacement Methods: Binary Image Metrics

- Perfect score (zero) only when verification and forecast are identical: M(A,B)=0 only if A=B.
- Symmetry ensures that answer does not depend on order of comparison: M(A,B)=M(B,A)
- Triangle inequality ensures that results are not overly sensitive: M(A,B) much lower than M(A,C), then M(B,C) is appropriately large.

Copyright 2014 NCAR

Displacement Methods: Binary Image Metrics

Examples of some binary image measures and **metrics**

- **Hausdorff Metric**
- **Baddeley's Δ Metric**
- Pratt's Figure of Merit (FOM)
- **Forecast Quality Index (FQI, also incorporates intensity information)**
- Minimum separation distance

Copyright 2014 NCAR

Displacement Methods: Baddeley's Δ Metric

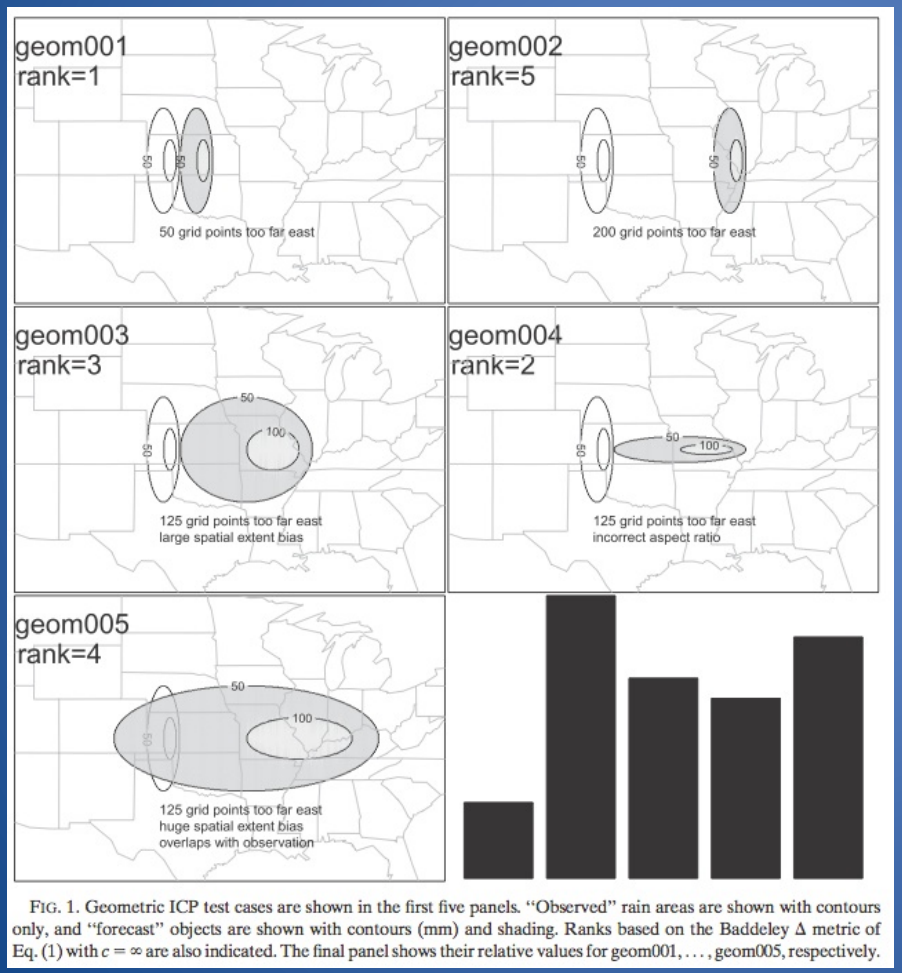


FIG. 1. Geometric ICP test cases are shown in the first five panels. "Observed" rain areas are shown with contours only, and "forecast" objects are shown with contours (mm) and shading. Ranks based on the Baddeley Δ metric of Eq. (1) with $c = \infty$ are also indicated. The final panel shows their relative values for geom001, . . . , geom005, respectively.

Copyright 2014 NCAR

From G. (2011, *WAF*, **26**, 409 - 415)

Geometric Indices

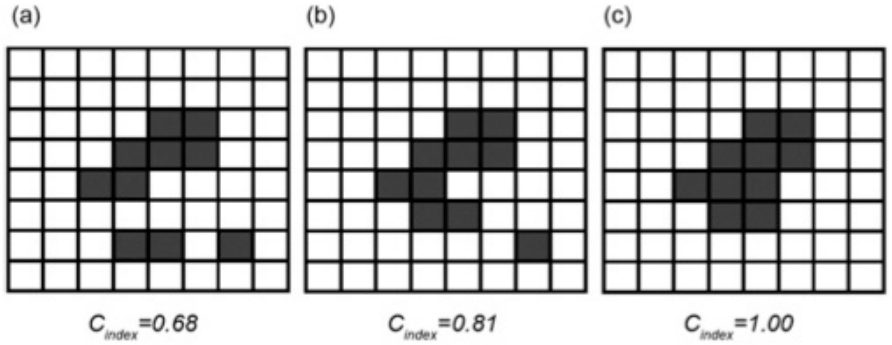


FIG. 3. The connectivity index (C_{index}) for three patterns.

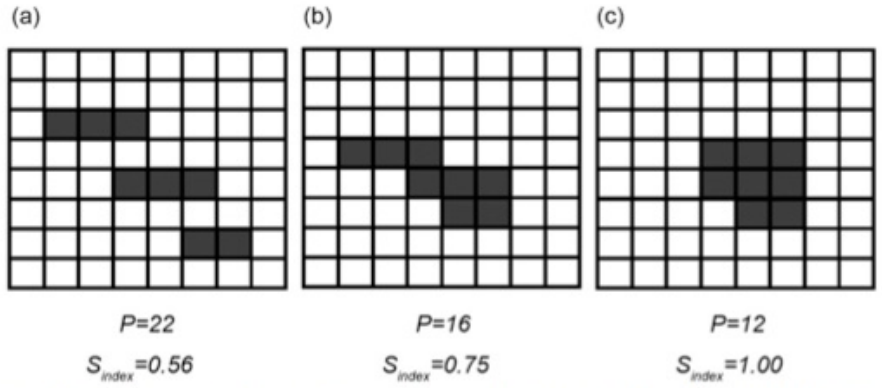


FIG. 7. The perimeter and S_{index} of 3 patterns that consist of 8 pixels with a P_{min} = 12: (a) P = 22, (b) P = 16, and (c) P = 12.

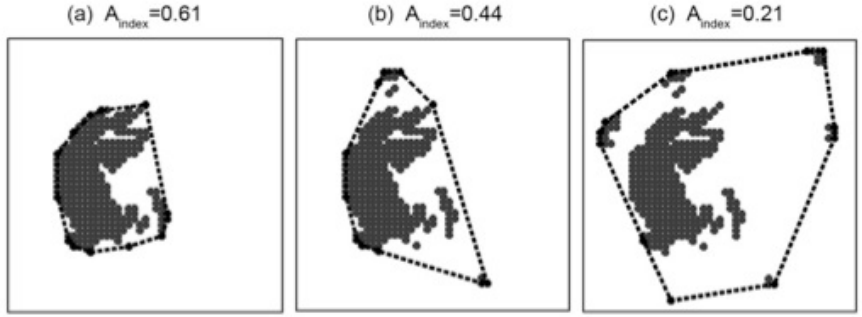
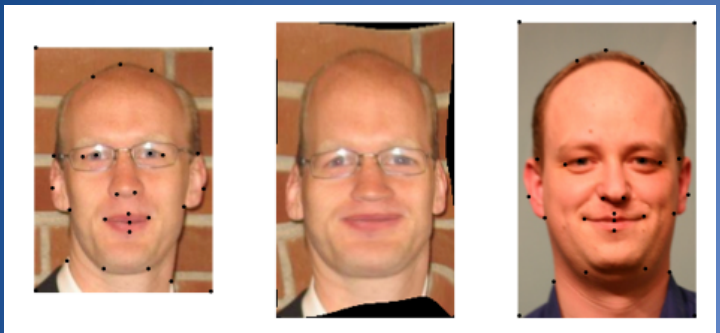


FIG. 9. The A_{index} for three example patterns: (a) A_{index} = 0.61, (b) A_{index} = 0.44, and (c) A_{index} = 0.21.

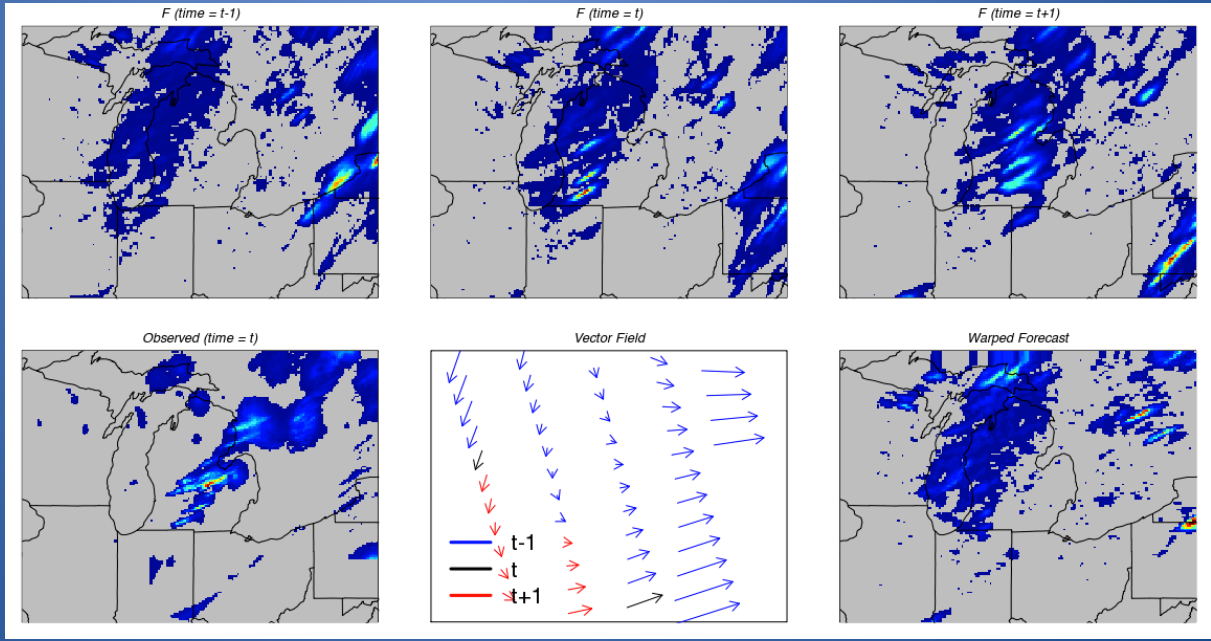
AghaKouchak *et al*. (2011, *J. Hyrdometeorology*, **12**, 274-285)

Copyright 2014 NCAR

Displacement Methods: Field deformation

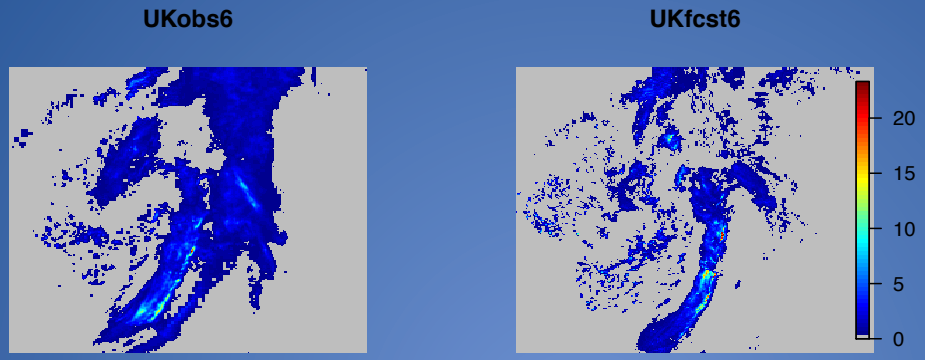


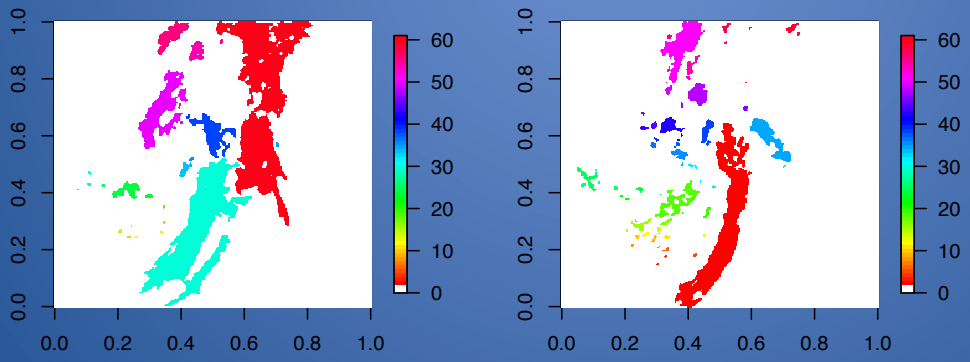
Reduction in RMSE is over 50% after applying the (space-time) warp.



Copyright 2014 NCAR

Displacement Methods: Feature-based





Copyright 2014 NCAR

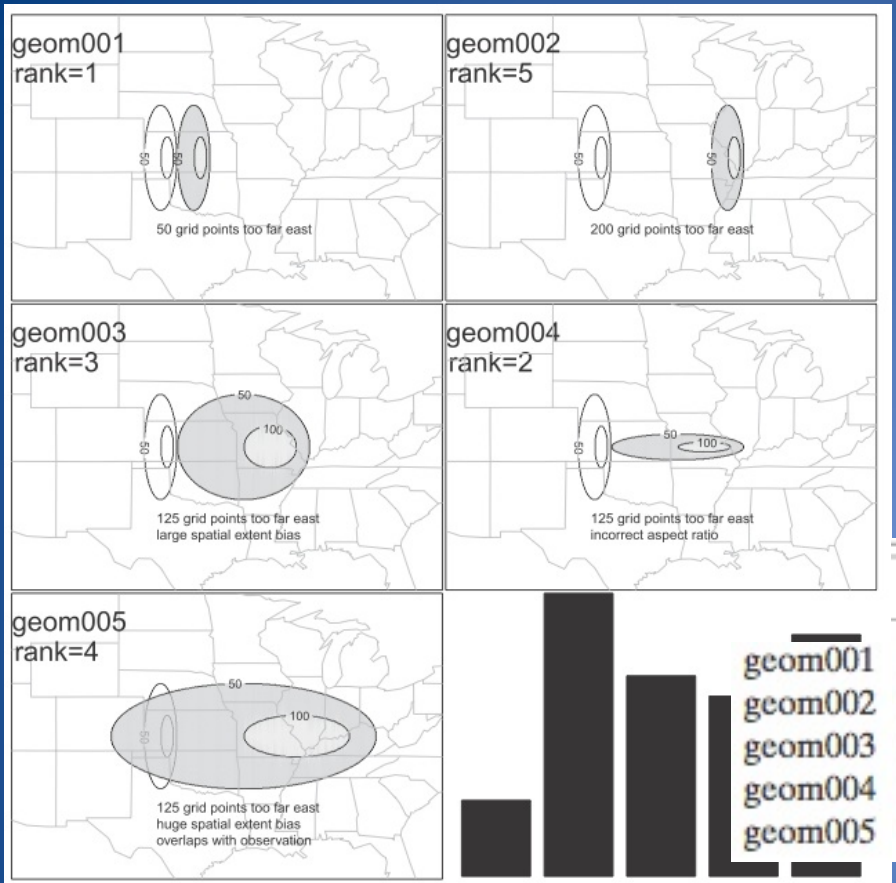


FIG. 1. Geometric ICP test cases are shown in the first five panels. "Observed" rain areas are shown with contours only, and "forecast" objects are shown with contours (mm) and shading. Ranks based on the Baddeley Δ metric of Eq. (1) with $c = \infty$ are also indicated. The final panel shows their relative values for geom001, ..., geom005, respectively.

Δ	Δ rank	DAS rank	MODE rank
44.98	1	1	1
167.34	5	2	5
118.52	3	5	3
106.5	2	4	4
142.86	4	3	2

From G. (2011, *WAF*, **26**, 409 - 415)

Copyright 2014 NCAR

Final Remarks

- R software package: `SpatialVx` (not yet ready for prime time)
- Spatial Forecast Verification Inter-Comparison Project (ICP)
  http://www.ral.ucar.edu/projects/icp
- ICP web page
  - List of references relevant to spatial forecast verification, as well as the ICP special collection in *WAF*.
    - See, e.g., review papers:
      - Ahijevych *et al*. (2009),
      - G. *et al*. (2010, 2010 *BAMS*),
      - Brown *et al*. (2012),
      - G. (2012): brief but more recent review.
  - ICP test cases available (geometric, perturbed and real)
  - Sign up to receive emails about the ICP
- ICP2 to begin soon, and to be called the Mesoscale Verification Intercomparison over Complex Terrain (MesoVICT) Project
  - New test cases
  - Includes: ensembles (forecast and observation), realistic meteorological cases over multiple time points, complex terrain, more variablesTech Note describing project plans:

Dorninger, M., M.P. Mittermaier, E. Gilleland (more) , 2013: MesoVICT: Mesoscale Verification Inter-Comparison over Complex Terrain. NCAR Technical Note NCAR/TN-505+STR, 23 pp, DOI: 10.5065/D6416V21.

Copyright 2014 NCAR