Mesoscale Modeling and Data Assimilation for Severe Weather and High Impact Snowstorms

Steven J. Greybush

Associate Professor of Meteorology

Center for Advanced Data Assimilation and Predictability Techniques
Institute for Computational and Data Sciences

Department of Meteorology & Atmospheric Science
The Pennsylvania State University

ADAPT

PennState

Cooperative Institute for Severe and High Impact Weather Research and Operations
Mesoscale and Storm Scale Modeling Workshop
December 6, 2021

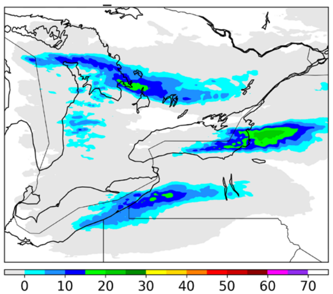
Collaborations with: Xingchao Chen, Rich Grumm, Matthew Kumjian, Jia Li, David Stensrud, George Young, Fuqing Zhang, Yunji Zhang

Grad Students: Da Fan, Glen Hanson, Robert Nystrom, Jon Seibert, Seth Saslo

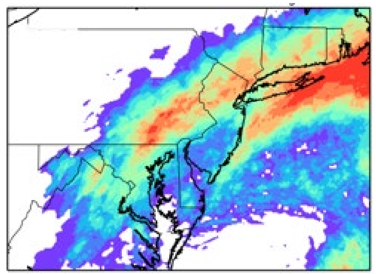
Applications of Data Assimilation

Classic NWP:

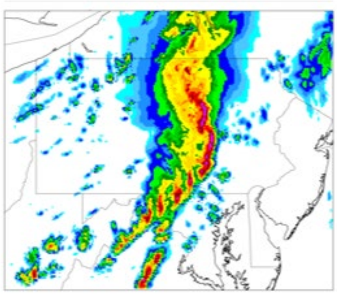
Lake-Effect Snow



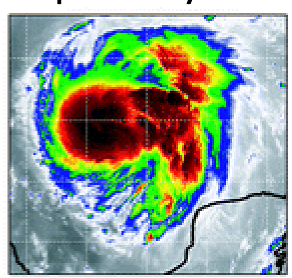
East Coast Winter Storms



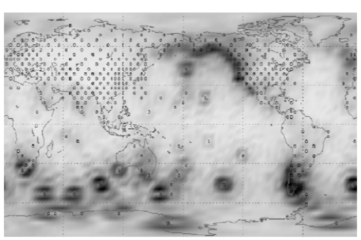
Severe Thunderstorms



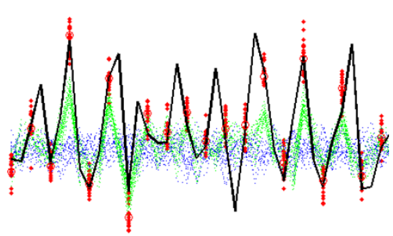
Tropical Cyclones



Global NWP

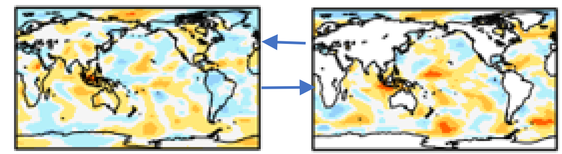


Idealized Models

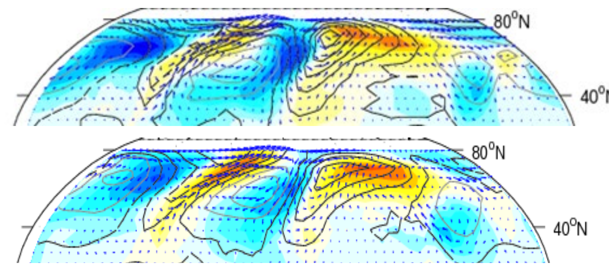


Beyond NWP:

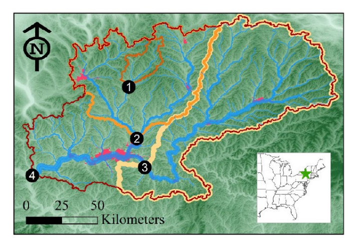
Coupled Earth System



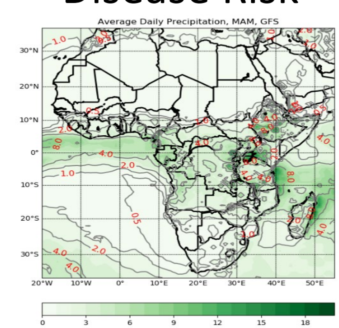
Mars Atmosphere



Hydrology



Disease Risk

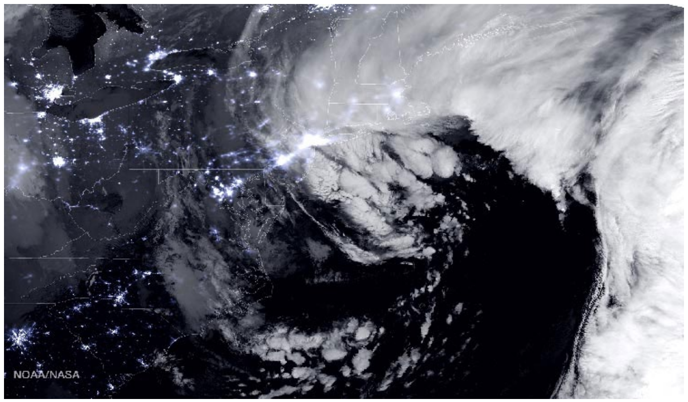


Collaborators/Image Courtesy: Seth Saslo, Rich Grumm, Glen Hanson, Robert Nystrom, Atsushi Okazaki, Xingchao Chen, Fuqing Zhang, Sanjib Sharma, Steve Schiff, Moussa Gueye

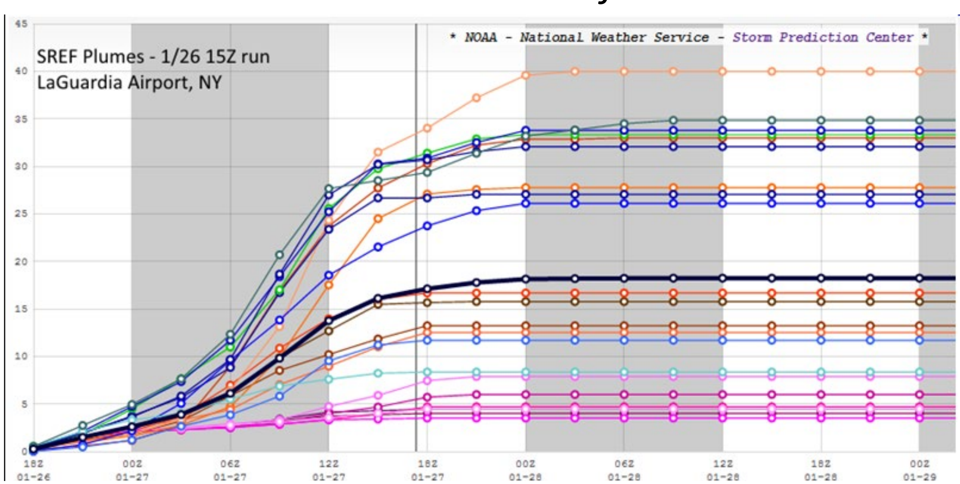
Outline

- East Coast Winter Storms Predictability and Sensitivity
- IMPACTS Modeling and Data Assimilation
- Tropical Cyclone Parameter Estimation
- Lake-effect Ensemble Design
- Geometry-Sensitive Ensemble Mean
- Smartphone Pressure Data Assimilation
- Convection Initiation Deep Learning

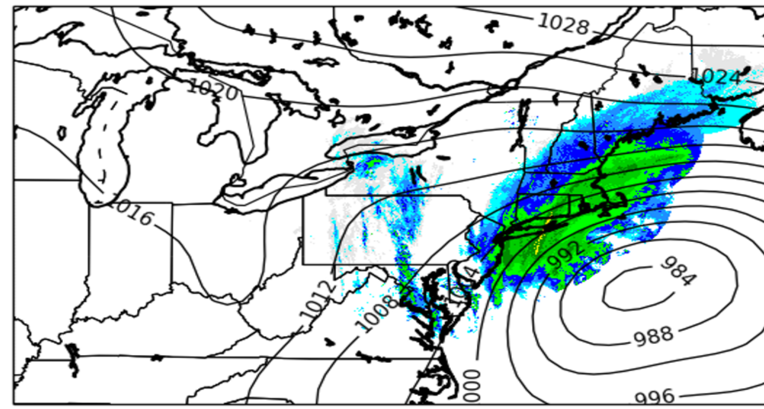
Assessing the Ensemble Predictability of East Coast Winter Storms



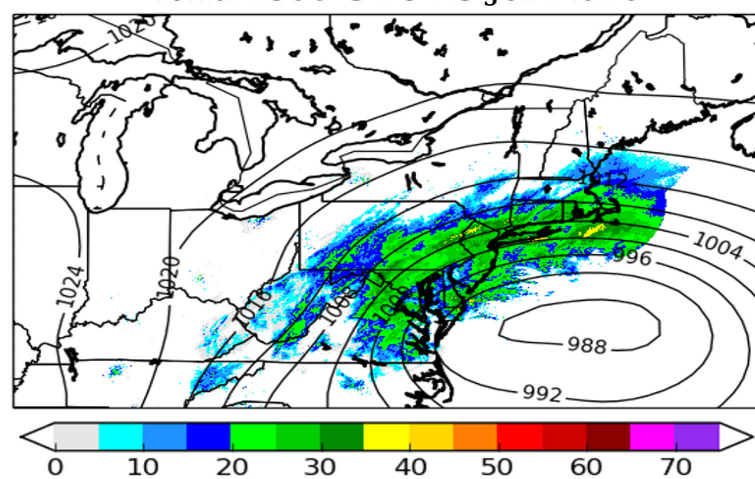
Deterministic Forecast: 24-30" for NYC
Ensemble Forecast: 5 to 40" of snow?



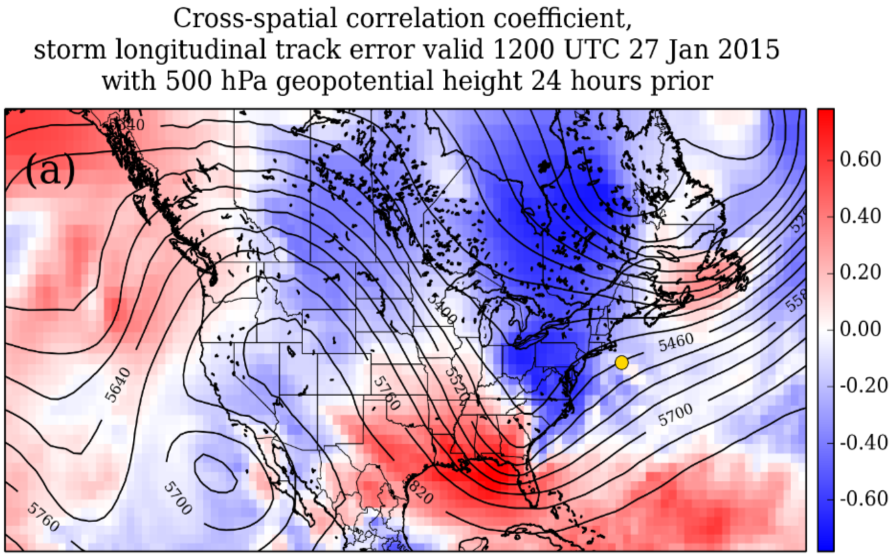
CFS and Composite Radar (dBZ)
Valid 0600 UTC 27 Jan 2015



CFS and composite radar (dBZ)
Valid 1800 UTC 23 Jan 2016



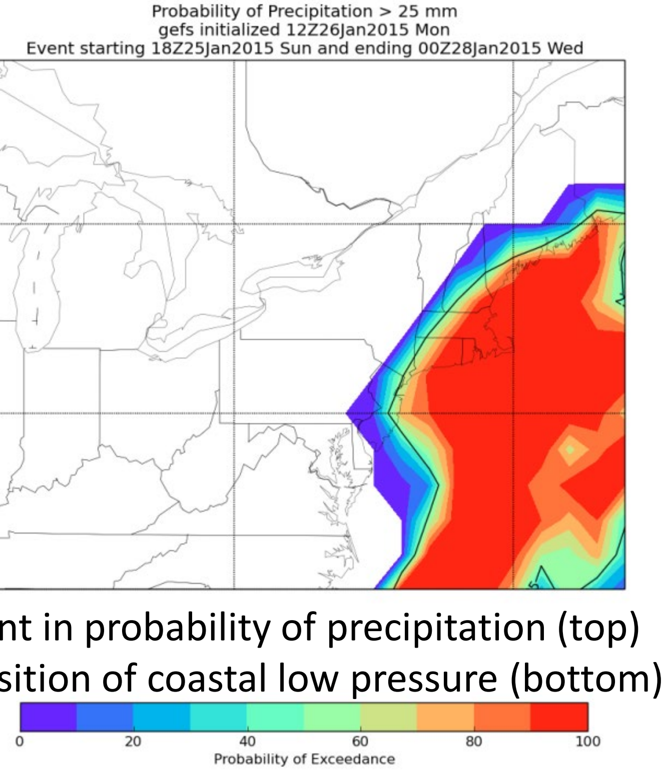
Ensemble Sensitivity: January 2015 Snowstorm



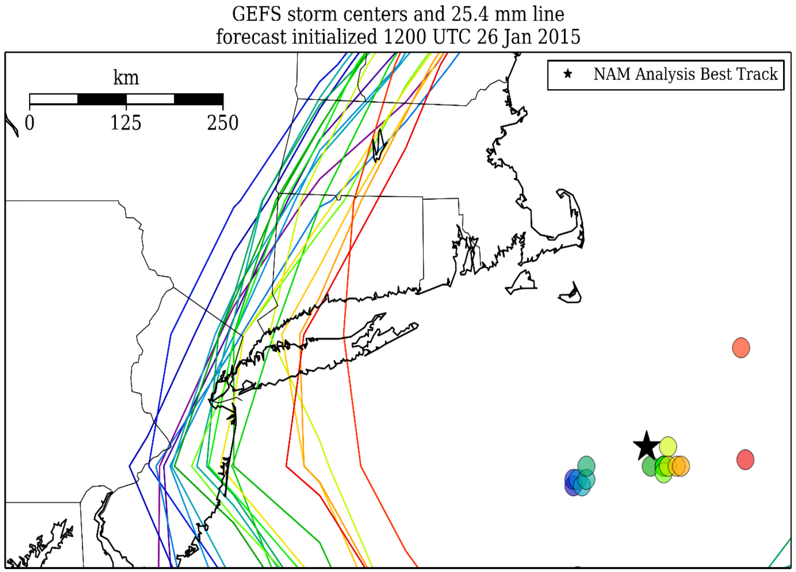
Position error in coastal low traced backwards in time to uncertainties in synoptic scale flow (contours) using ensemble sensitivity (shading).

Red: 500 mb height field is positively correlated with eastward track error.

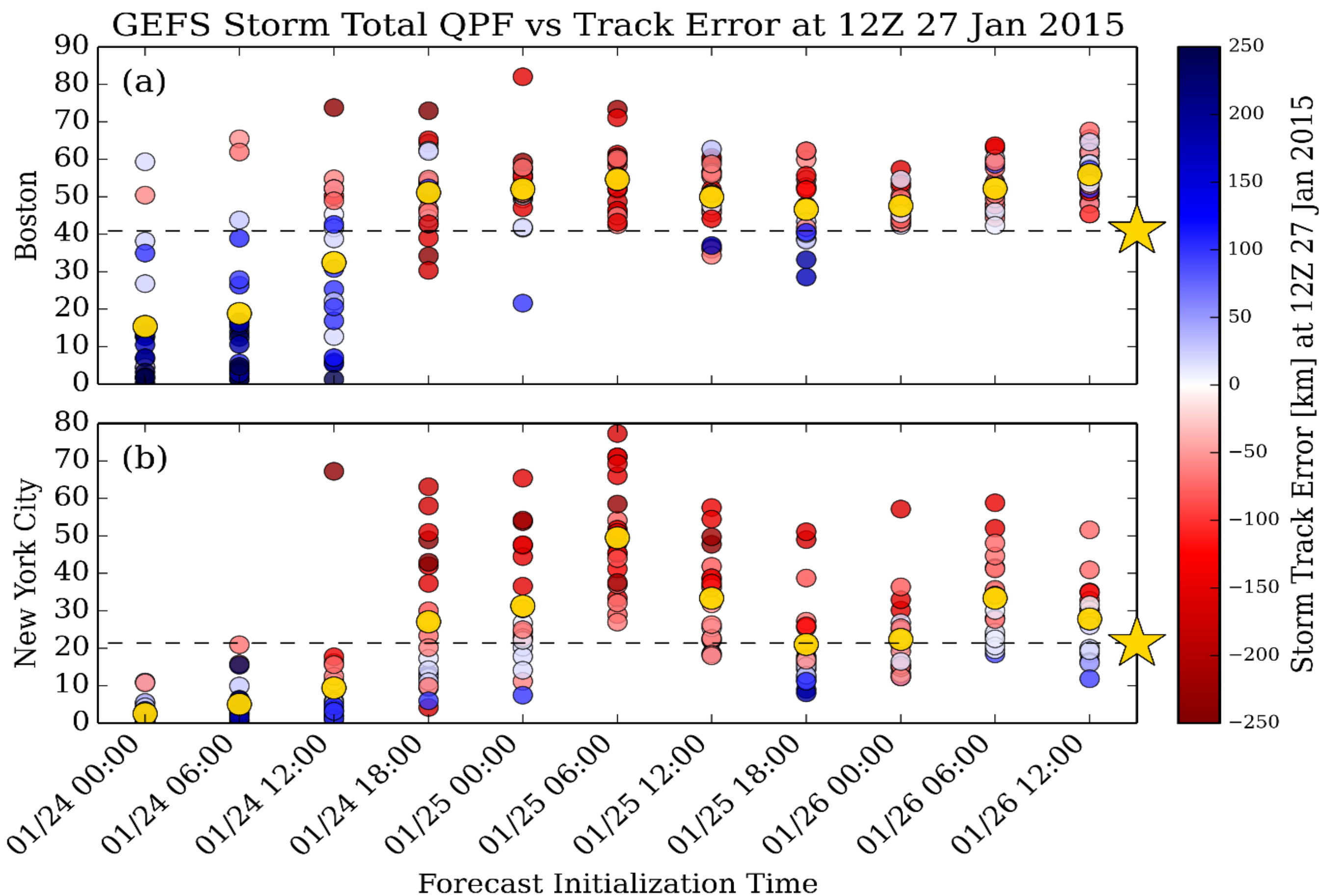
From Greybush et al. (2016)



Tight gradient in probability of precipitation (top) linked to position of coastal low pressure (bottom).

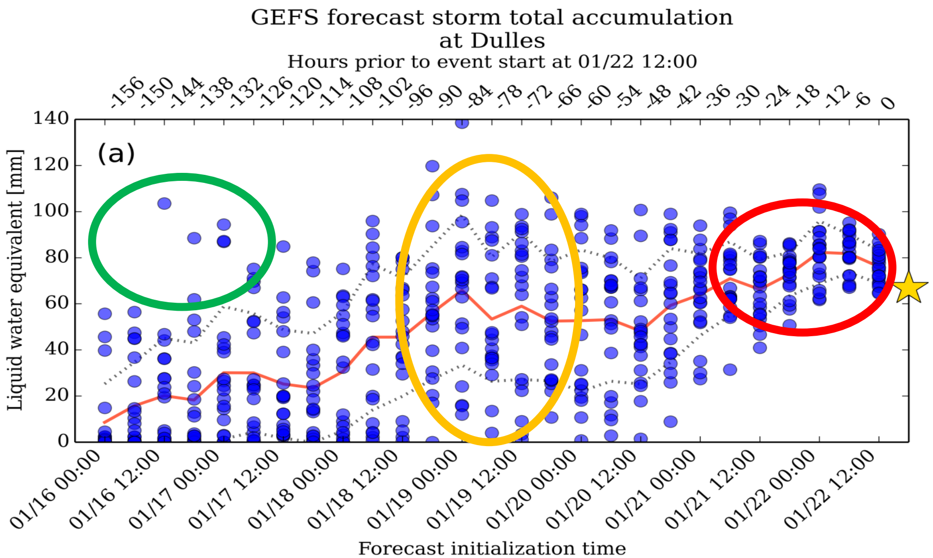


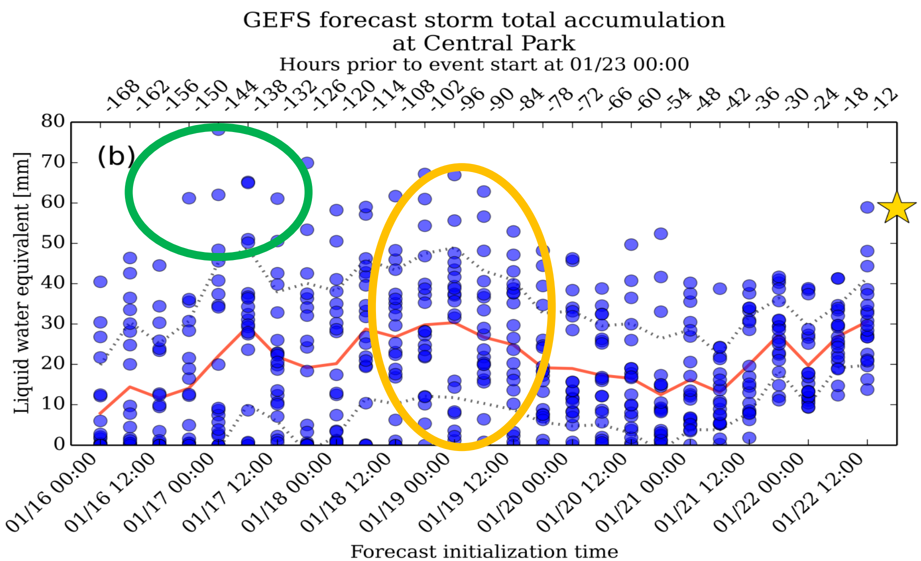
QPF and Track Error as function of Forecast Lead Time



From Greybush et al. (2016)

Predictability Horizons





To answer: How far in advance is a feature predictable?

First, identify an event (location, variable type, etc.).

Then characterize, using the ensemble, the:

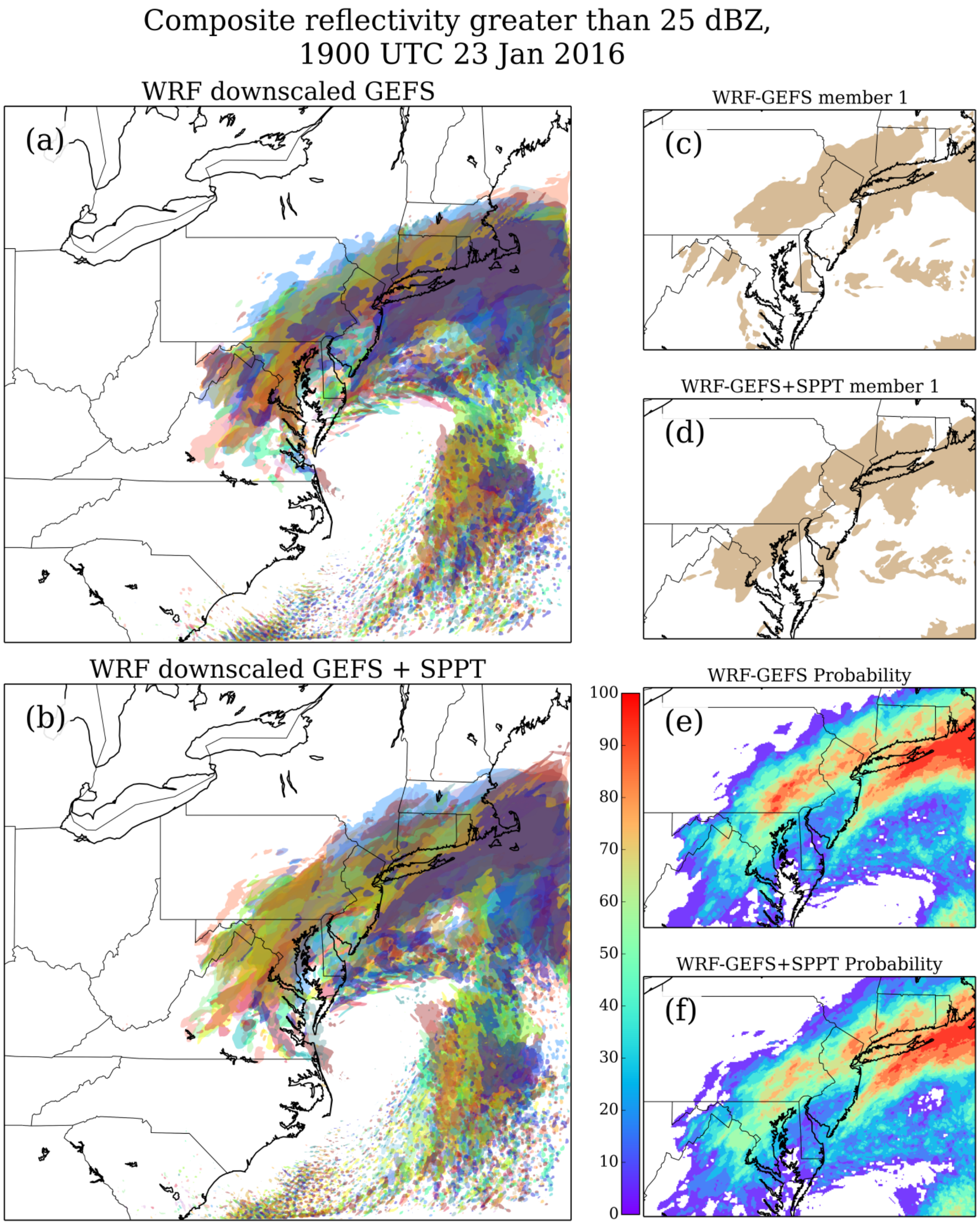
- initial detection
- emergence of a signal
- convergence of solutions

From Greybush et al. (2016)

Understanding Predictability with Convection-allowing Ensembles

Intrinsic Predictability:
Even if we have a perfect model, and nearly perfect initial conditions,
predictability is limited.
Estimate using ensemble spread of perfect model, as initial perturbations become smaller.

Practical Predictability:
Given our current (limited) observing system and (imperfect) models, how far ahead can we skillfully forecast a weather phenomenon.
Need to account for model error; e.g. include perturbations in forecast phase.

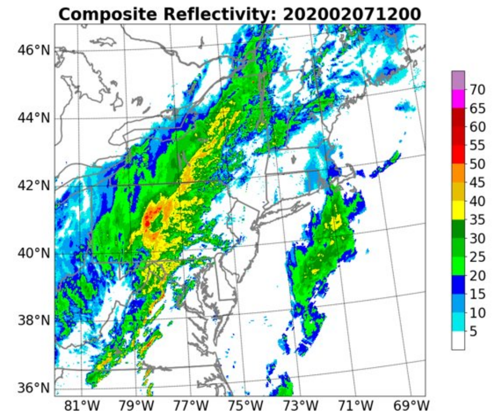


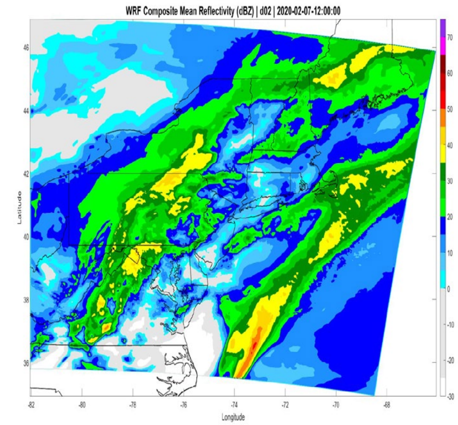
From Greybush et al. (2016)

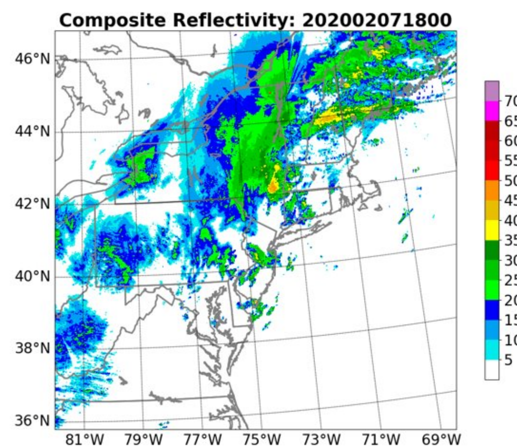
NASA IMPACTS Field Campaign: Data Assimilation and Parameter Estimation

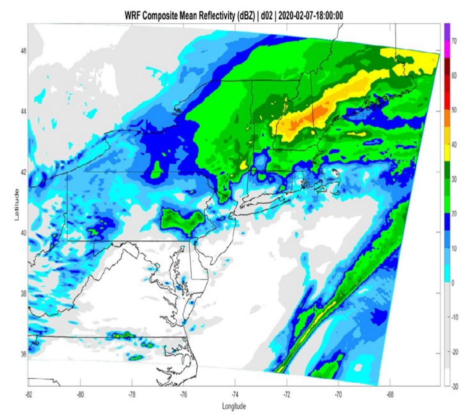
With Collaborators Matt Kumjian, Yunji Zhang

- Field Campaign includes aircraft in situ, radar observations
- PSU goals include:
- 4D regional reanalyses at convection-allowing resolution
- validation, then assimilation, with field campaign observations
- Finally, improve WRF microphysical modeling using ensemble parameter estimation









Potential for New Constraints on Tropical Cyclone Surface-Exchange Coefficients through Simultaneous Ensemble-Based State and Parameter Estimation

Robert Nystrom, Steven Greybush, Xingchao Chen, and Fuqing Zhang

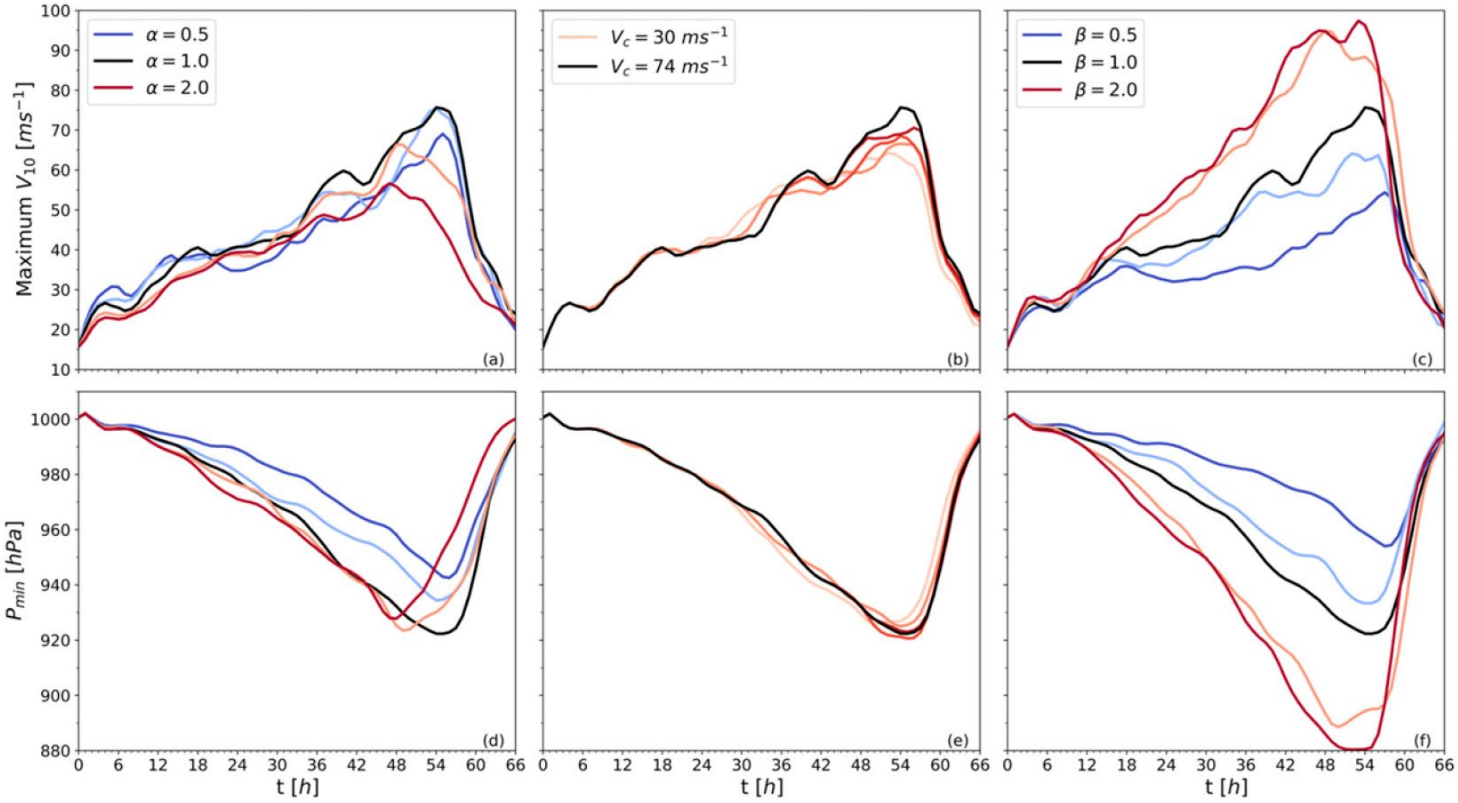


FIG. 2. Simulated (a)–(c) maximum 10 m wind speed and (d)–(f) minimum central pressure for single-parameter (a),(d) α; (b),(e) V_c; and (c),(f) β ensembles. The CNTL simulation is shown in black in all panels and warmer or cooler colors depict parameter values greater than or less than the CNTL parameter values, respectively.

From Nystrom et al. (2021)

Ensemble Parameter Estimation: Correlations

- Critical forecasting field needs to be sensitive to parameter within the model.
- Model parameter must also impact observable variables.

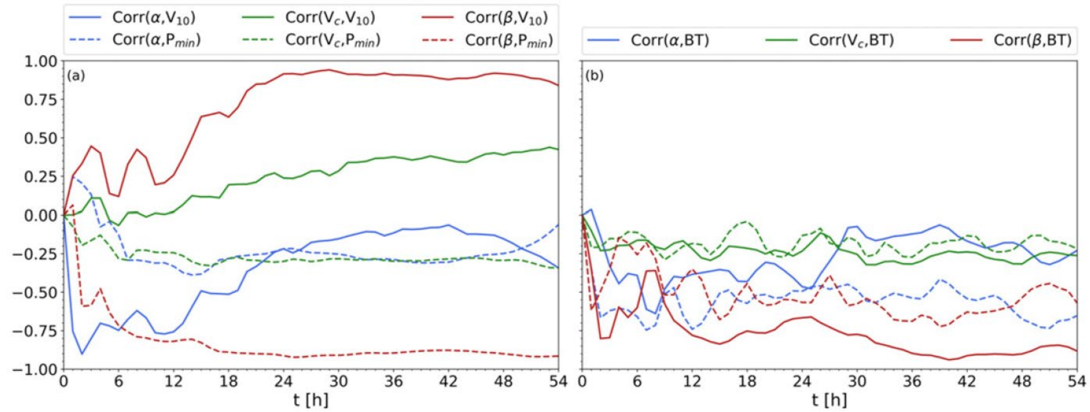


FIG. 5. Ensemble correlations between (blue) α, (green) V_c, or (red) β and (a) maximum 10 m wind speed (solid) or minimum central pressure (dashed) and (b) the average brightness temperature within 100 km (solid) or 50 km (dashed) from the TC center.

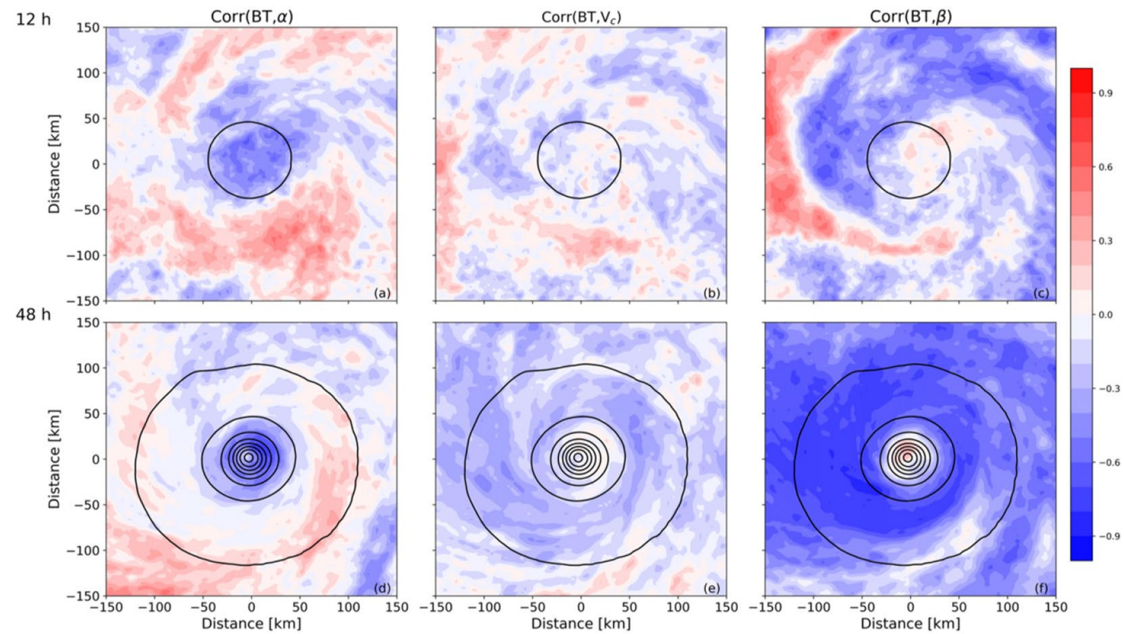


FIG. 8. Ensemble correlations between (a),(d) α; (b),(e) V_c; or (c),(f) β and simulated *GOES-13* channel-3 brightness temperature (BT) at (a)–(c) 12 and (d)–(f) 48 h. The ensemble-mean sea level pressure is contoured in black every 10 hPa.

From Nystrom et al. (2021)

Ensemble Parameter Estimation: OSSE Results

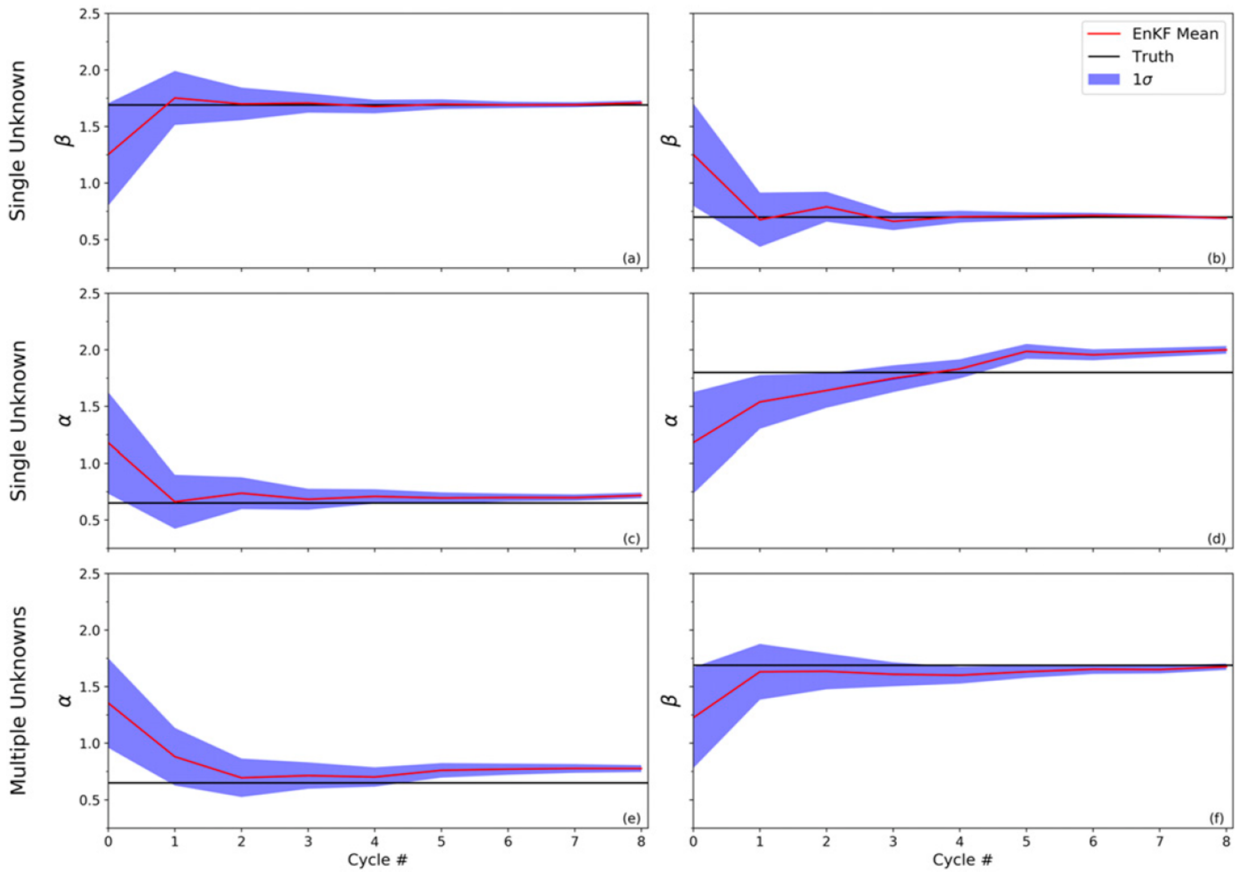
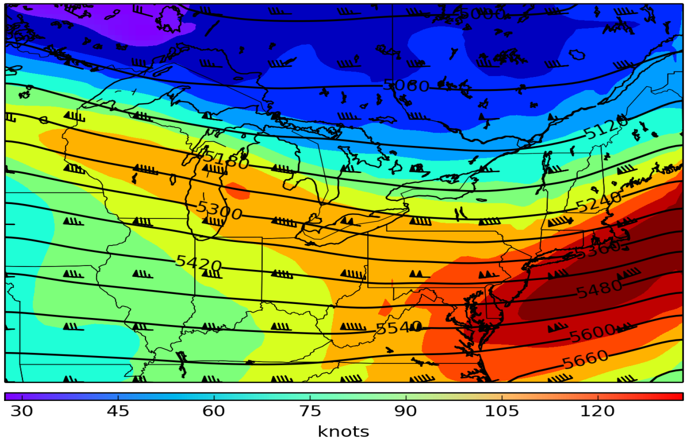


FIG. 9. Posterior EnKF updated ensemble estimation of (a) and (b) two cases with different "Truth" values and β as the only unknown parameter, (c) and (d) two cases with different "Truth" values and α as the only unknown parameter, and (e) and (f) a case with α and β as simultaneous unknown parameters. The "Truth" value for each parameter and experiment is shown in the black horizontal line.

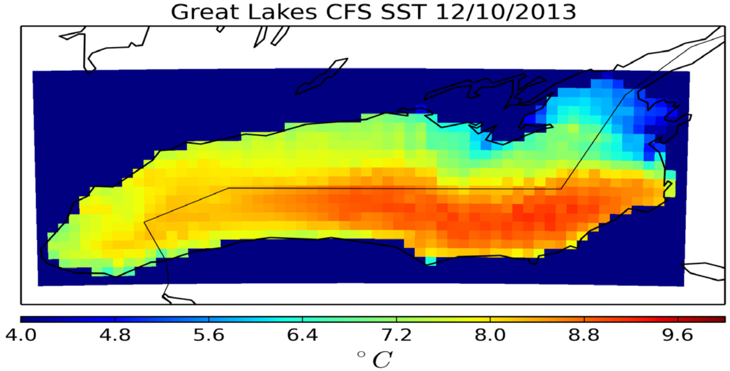
From Nystrom et al. (2021)

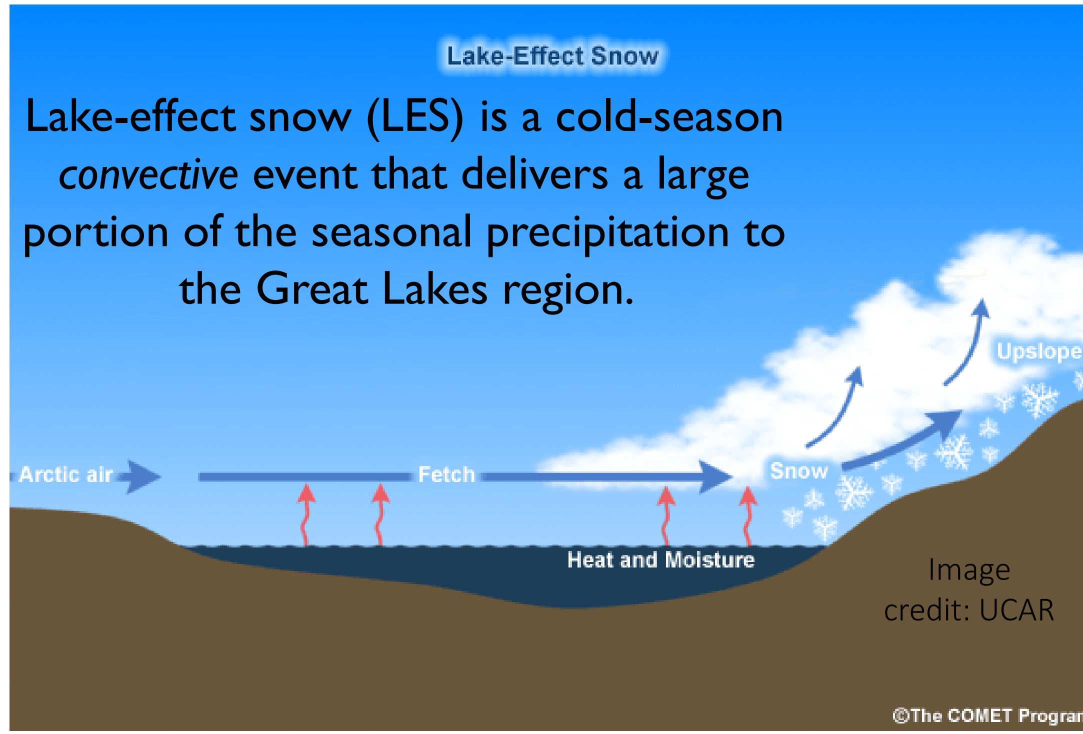
Convection-allowing ensemble forecasting and regional data assimilation for the prediction of lake-effect snow

domain d01, 12/11/2013 00:00



Great Lakes CFS SST 12/10/2013





Multi-Scale Phenomenon

Important meteorological variables:

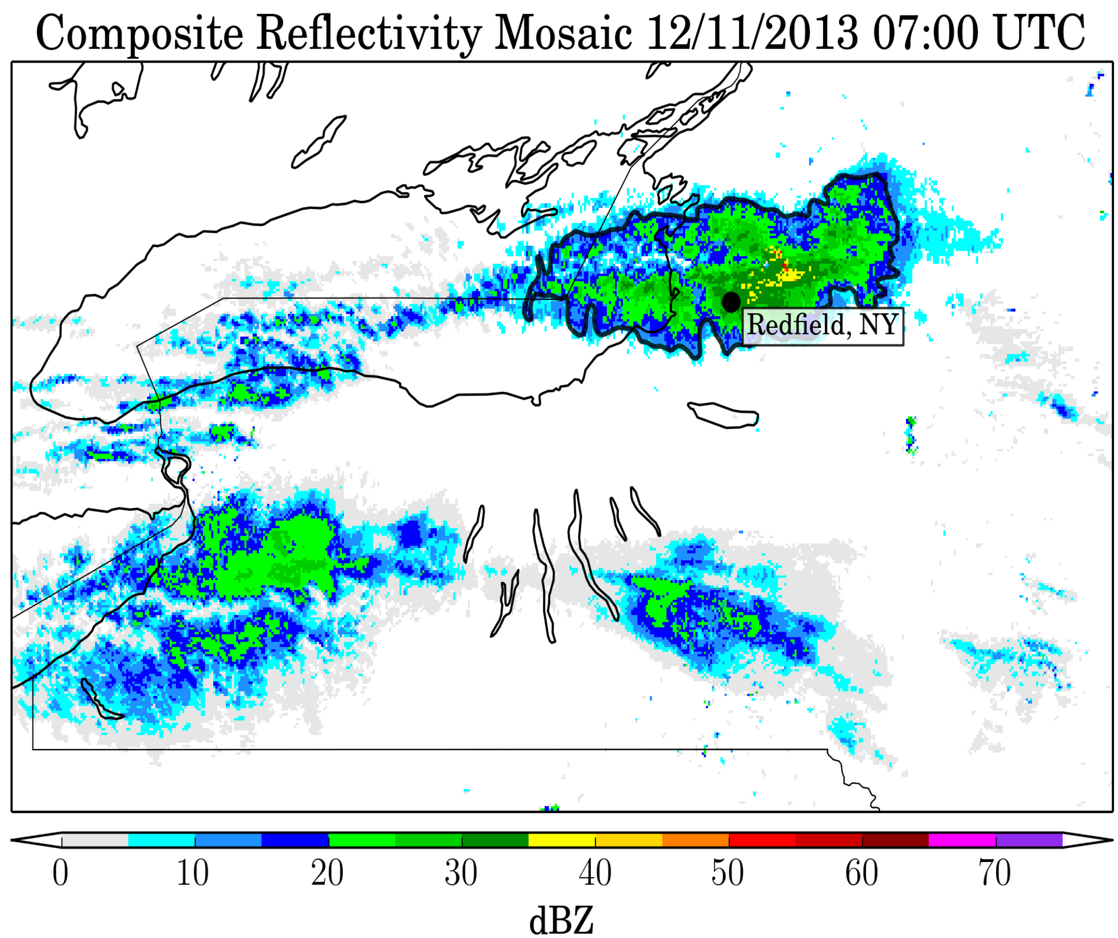
- lower-level wind speed and direction
- lower-level temperature and stability profiles (Niziol 1987, 1995)
- mid- and low-level temperature advection (Eipper et al., in preparation)

Important environmental variables:

- lake surface temperature
- lake surface ice coverage (Cordeira and Laird, 2008)
- lake shape and orientation (Laird et al., 2003)
- local topography at lake shore and inland (Onton and Steenburgh 2001; Alcott and Steenburgh 2013; Veals and Steenburgh 2015)

LES case study

- Long-lived long-lake-axis-parallel (LLAP) band
- Affected the Tug Hill Plateau region east of Lake Ontario, 10 December – 12 December 2013
- Significant snowfall, total accumulations exceeding 80 cm

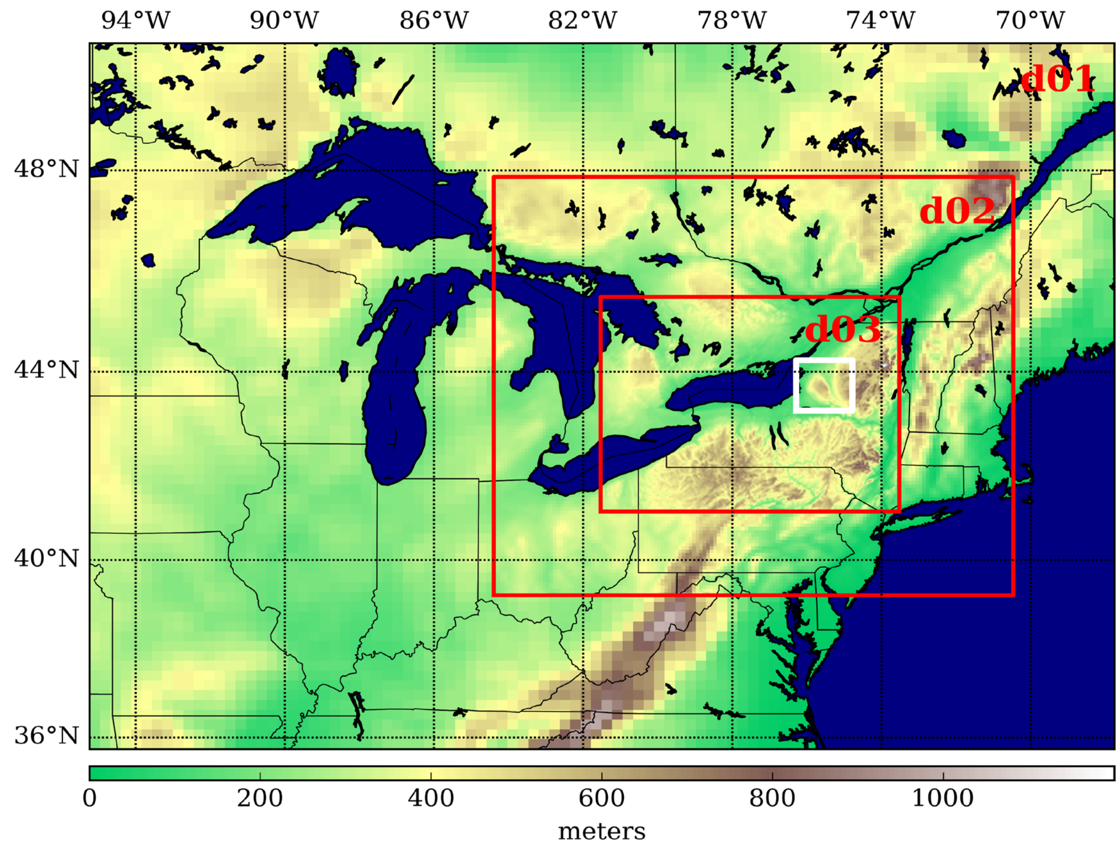


This event was particularly well observed as part of the Ontario Winter Lake-Effect Storms (OWLeS) project [*Kristovich et al.*, 2016], which provides some added in-situ observations.

Weather Research and Forecasting (WRF) V3.7.1

3 nested domains:
27, 9, 3 km
43 vertical levels

Parameterizations selected include:
Thompson MP;
MYJ PBL, Eta surface layer, NOAH LSM;
Dudhia & RRTMG short/longwave schemes;
Grell 3D cumulus domains 1&2 only



Ensemble Design

- Single physics vs. multi-physics
- GEFS vs. CV3 lateral boundary conditions

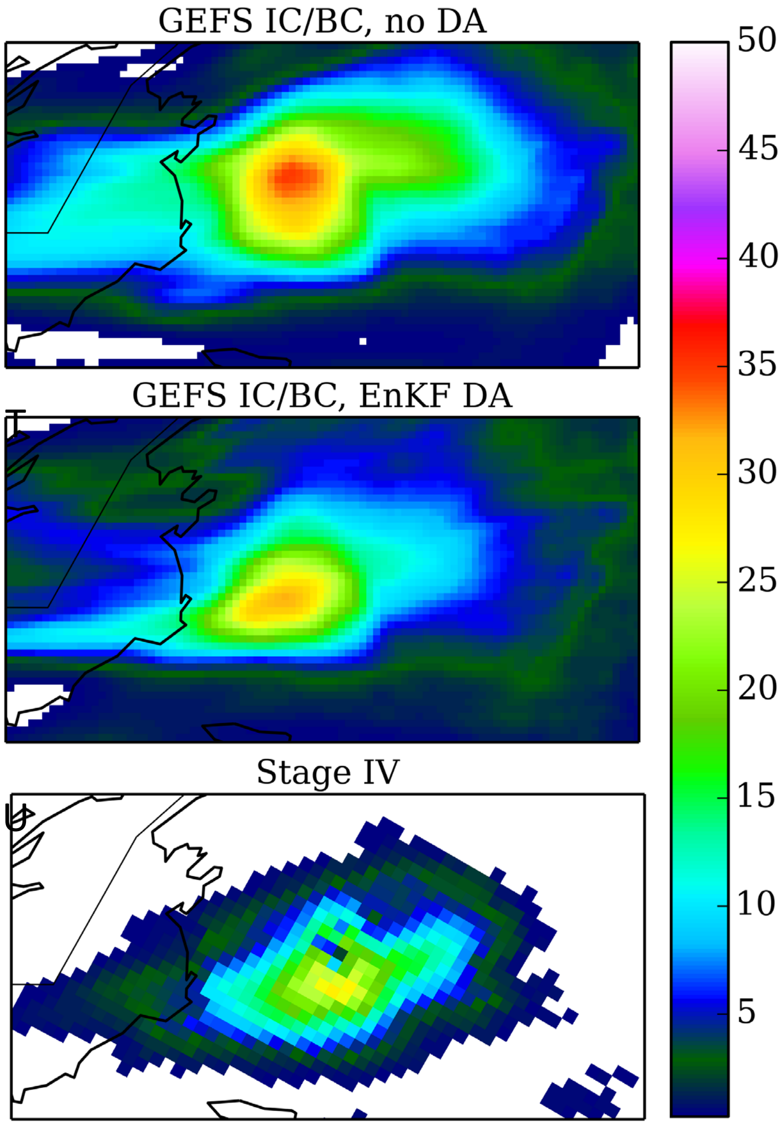
Characterize Uncertainty from:

Synoptic Scales
(from outside Great Lakes region; driven by GEFS)

Mesoscale
(simulated by convective-allowing WRF and regional EnKF)

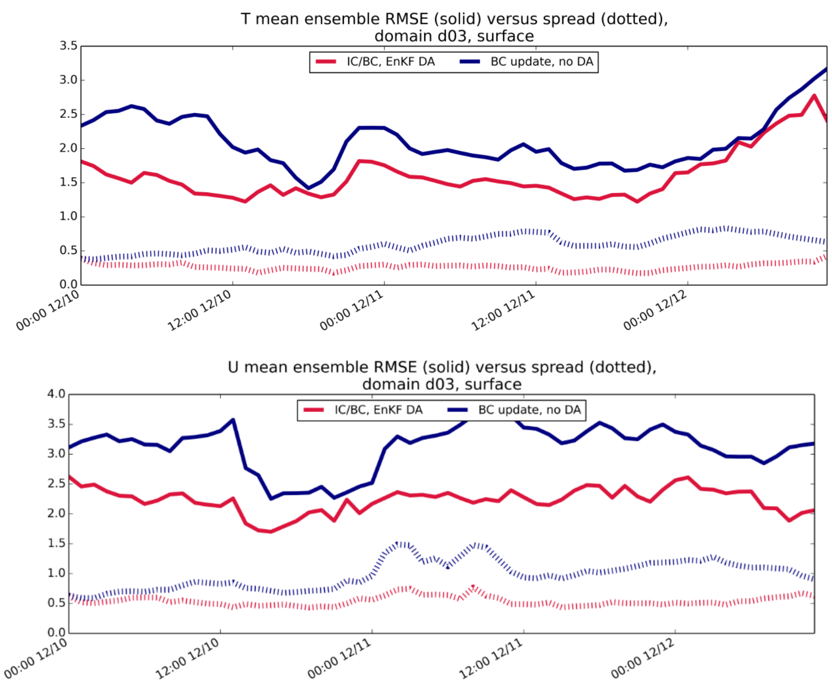
Parameterization Errors
(represent using multi-physics and stochastically perturbed ensembles)

EnKF data assimilation lead to significant **QPF** improvement



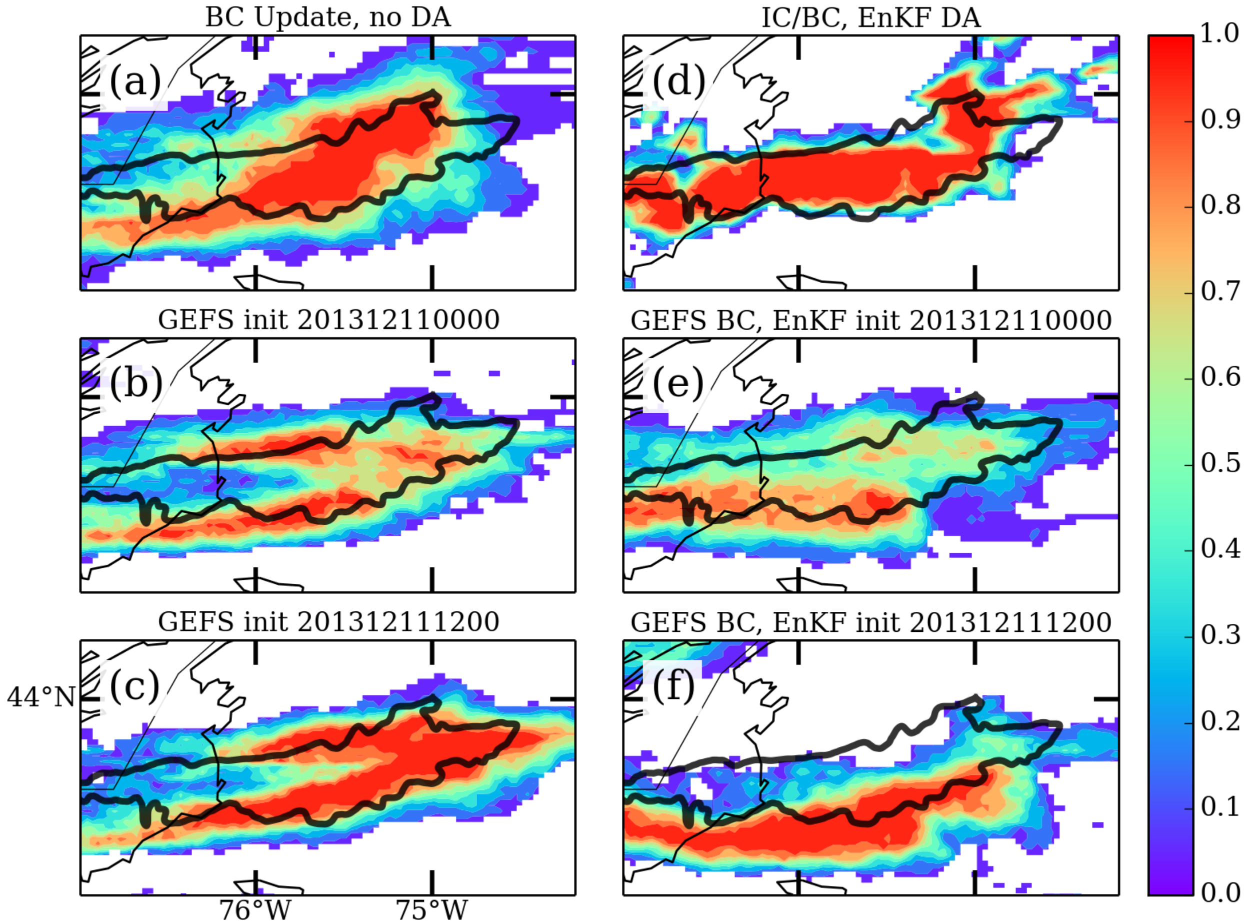
- Overforecast bias improved: maximum reduced, less overall precipitation produced within the domain
- Location bias reduced: precipitation maximum shifted south

Surface-based



Regional EnKF versus BC Update

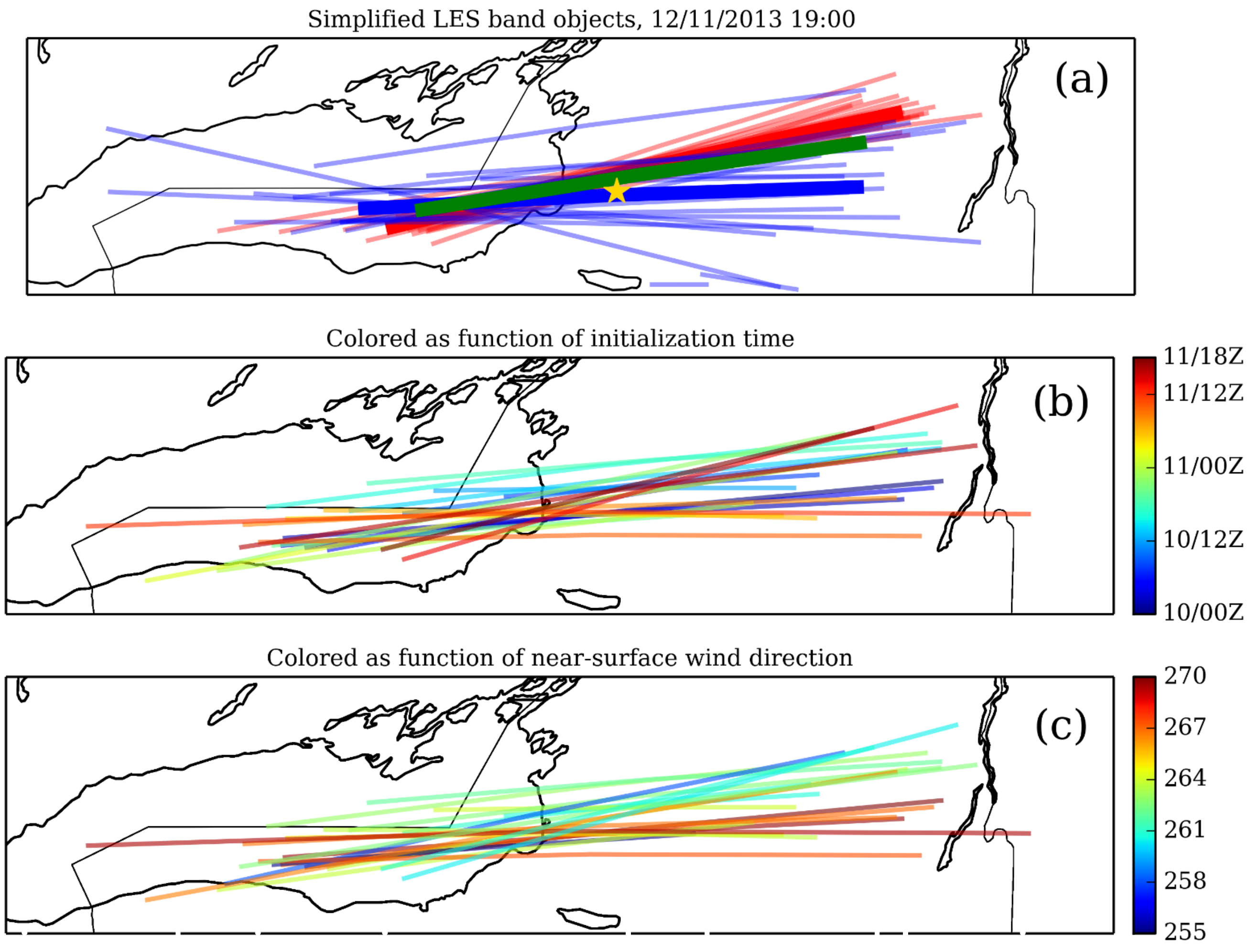
Position error?



Ensemble probability (colors) of exceeding composite reflectivity greater than 15 dBZ at 19 UTC 11 Dec; truth shown in black contour.

EnKF analysis recognizes single-banded structure in correct location; other forecasts do not.

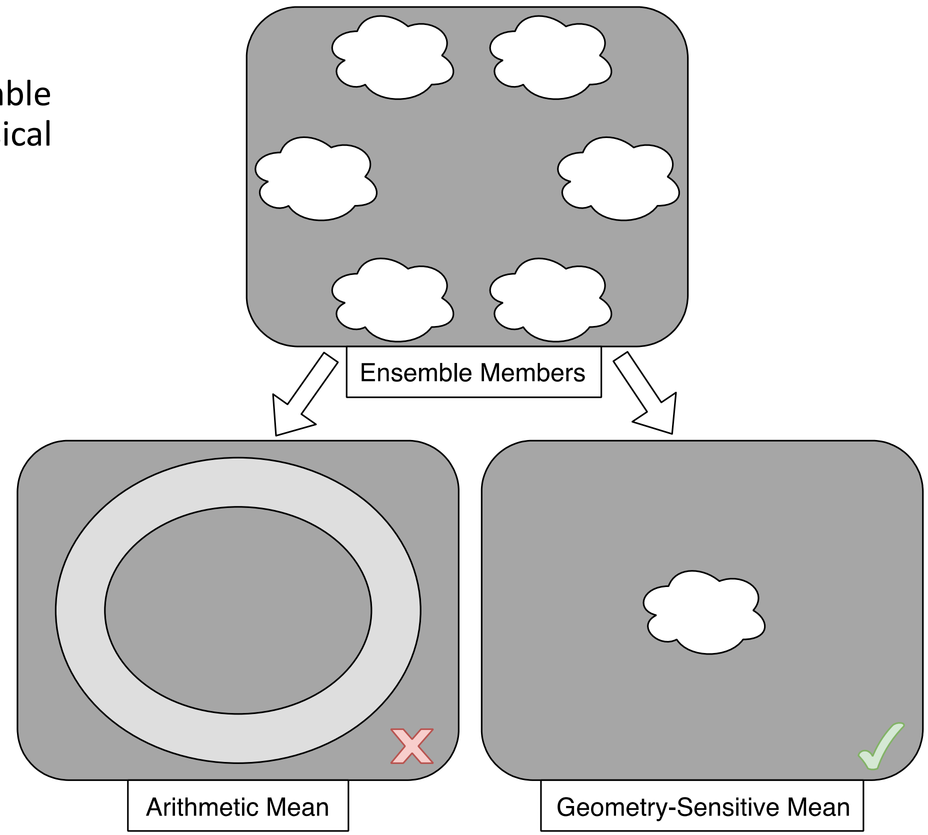
Object-Based Verification:
Regional ICs play an important role in LES band forecasting



Geometry-Sensitive Ensemble Mean

Jon Seibert, Steven Greybush, Jia Li, Zhoumin Zhang, and Fuqing Zhang

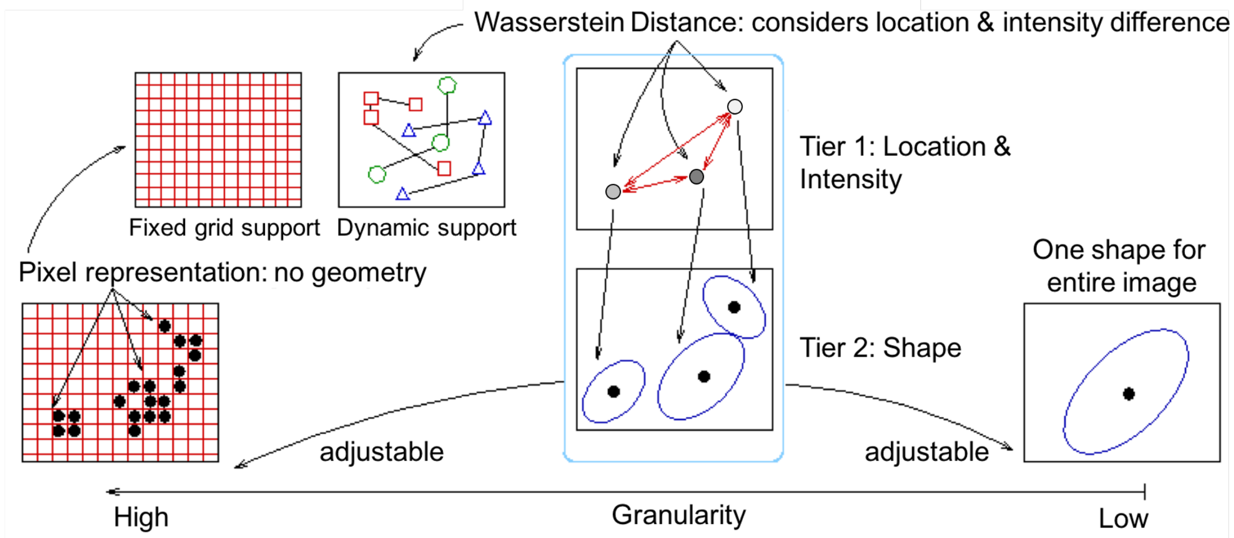
Motivation: Create an ensemble "mean" that retains the physical structures present in the individual members.



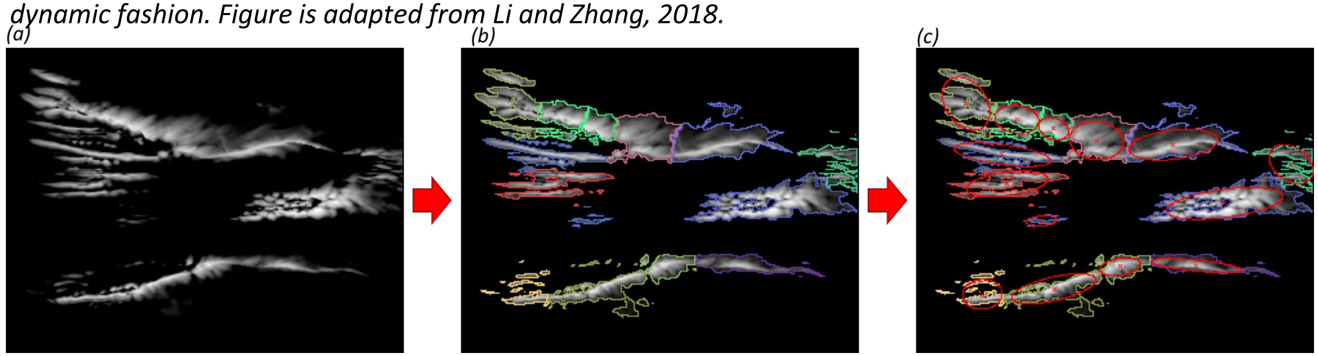
Idealized arithmetic (pixel-wise) mean (left) vs. GEM (right) for an ensemble of storm cells (top). While the arithmetic mean takes the strict average on a per-pixel basis, resulting in loss of intensity and structure information, the ideal GEM gives a more realistic mean that retains that information.

From Seibert et al. (2021)

Geometry-Sensitive Ensemble Mean: Overview of Methods



Schematic diagram illustrating the two-tiered signature (2TS), which captures the structural information by describing the location, intensity (shown here in grayscale), and shape of each cloud patch in a dynamic fashion. Figure is adapted from Li and Zhang, 2018.

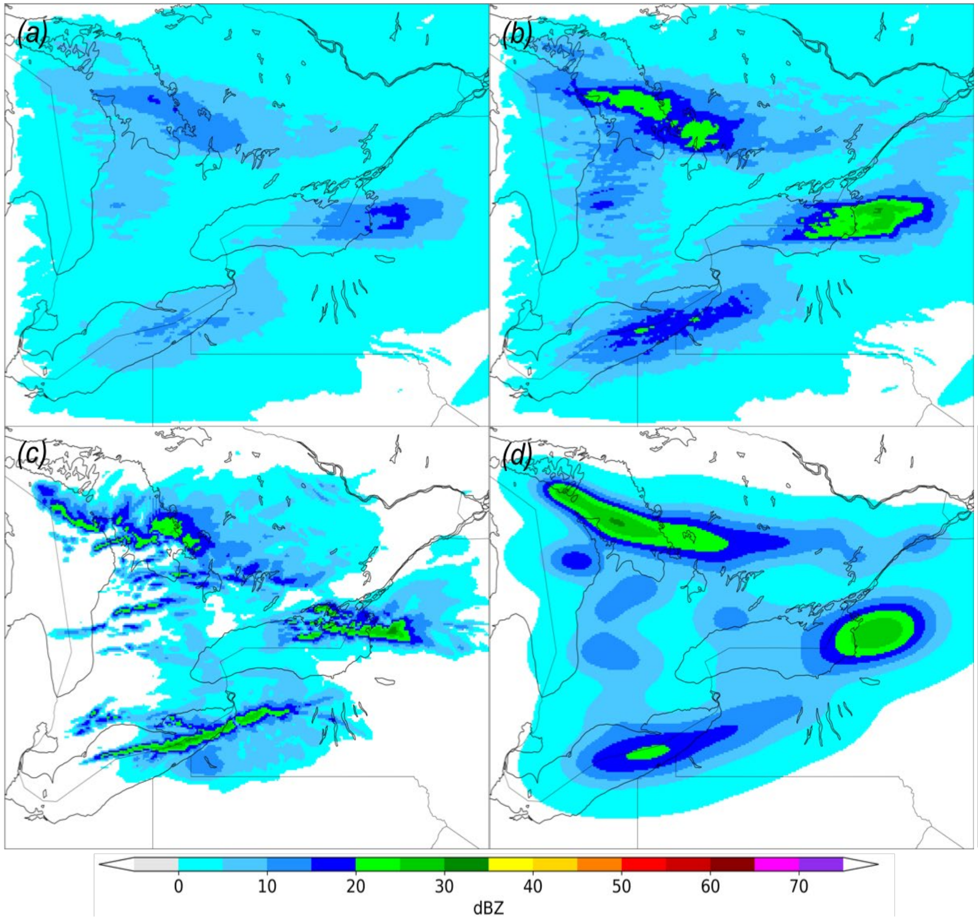


Schematic illustrating the process of clustering pixels into cloud patches and fitting the Gaussian of their 2TS: (a) The starting reflectivity source image, in grayscale; (b) image broken into patches, outlined in different colors; (c) intensity-weighted centroid (red star) and Gaussian distribution (red ellipse) calculated for each patch.

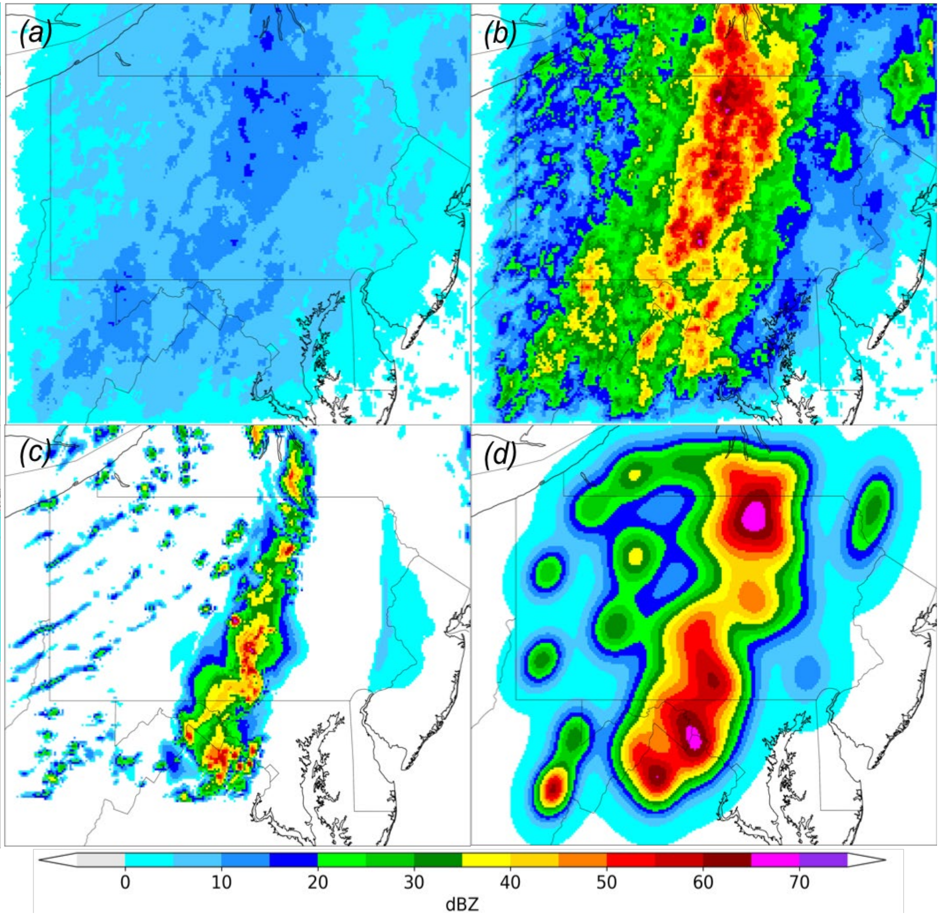
From Seibert et al. (2021)

GEM Applications Include: Lake-Effect Snow, Severe Thunderstorms

PWA = Pixel-Weighted Average
MDM = Mixture Density Mean
BPM = Bayesian Posterior Mean
ABM = Adjusted Best Member



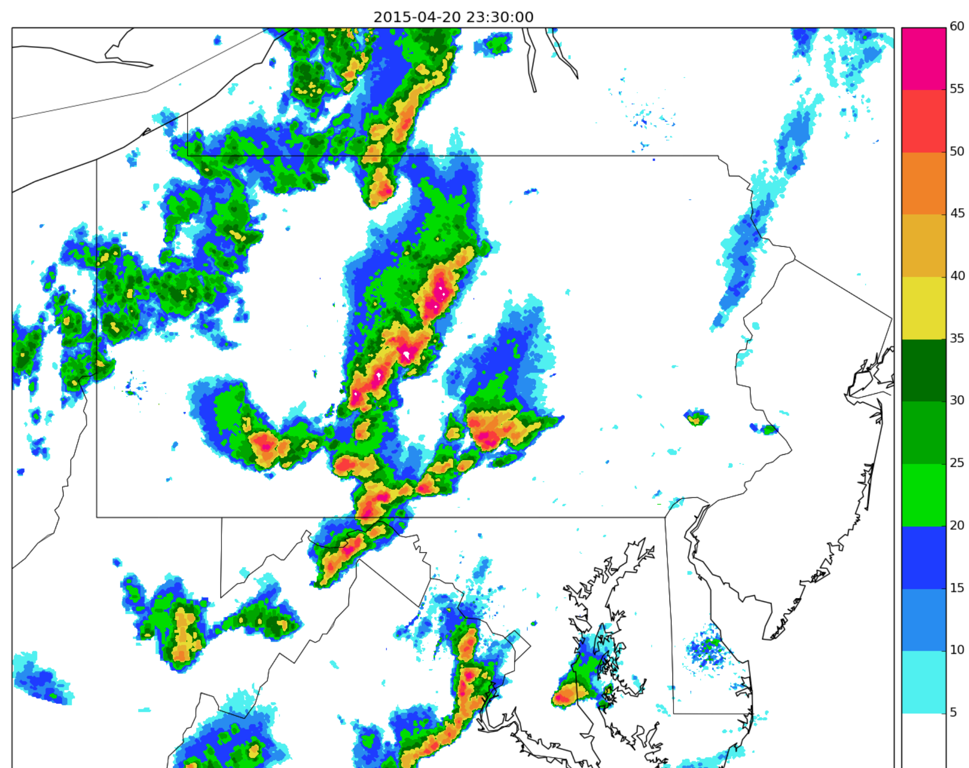
Thresholded Snowband ensemble results using GEM Bandwidth B, shown as composite reflectivity images: (a) PWA, (b) Scaled BPM, (c) ABM, and (d) Scaled MDM.



Thresholded Thunderstorm ensemble results using GEM Bandwidth B, shown as composite reflectivity images: (a) PWA, (b) Scaled BPM, (c) ABM, and (d) Scaled MDM.

Impact Of Assimilating Surface Pressure Observations From Smartphones On A Regional, High Resolution Ensemble Forecast: Observing System Simulation Experiments

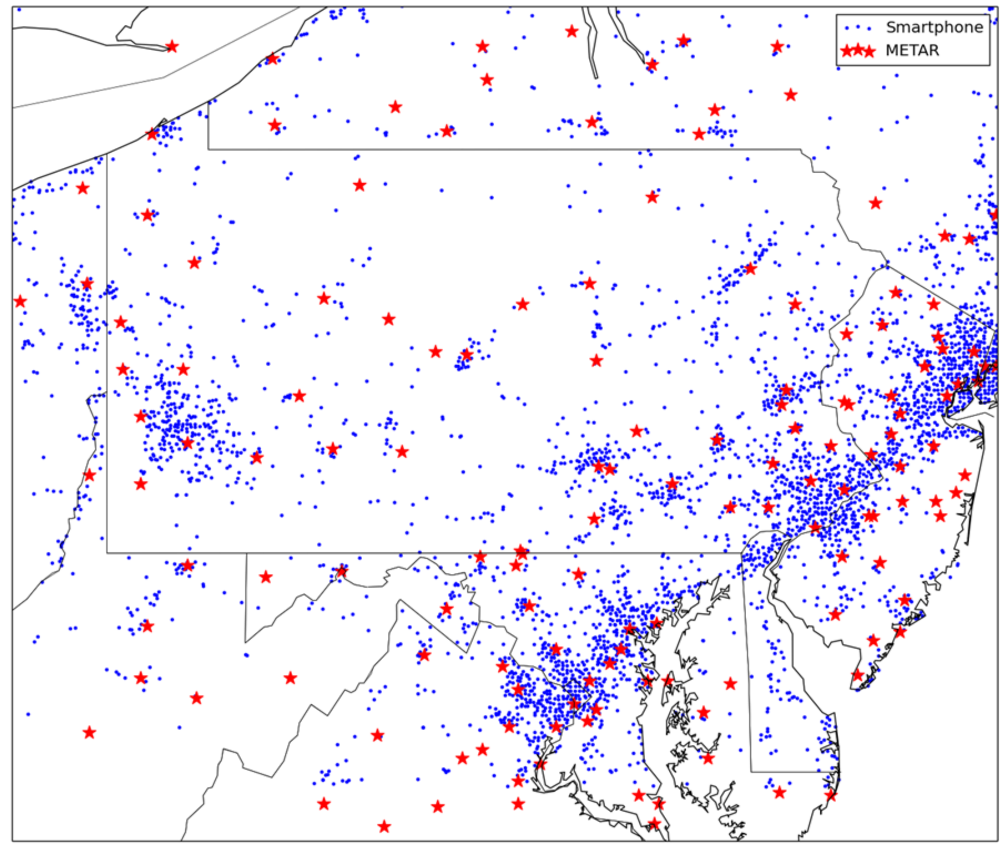
Case Study: Severe Thunderstorms in Pennsylvania



From Hanson (2016)

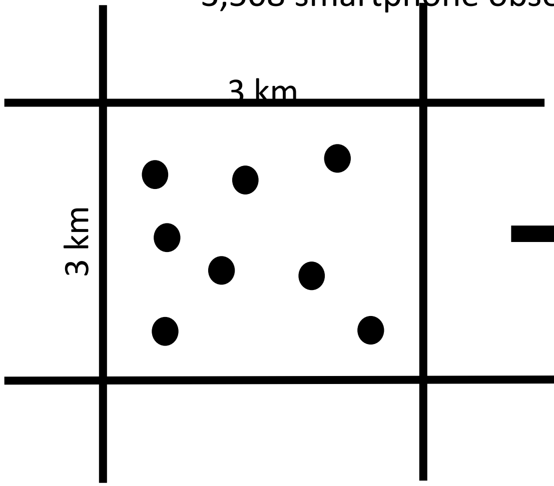
PressureNet Data

- Collected data from 27 February 2015 – 13 May 2015 (75 days)
- Hourly data sets contained an average of 15,000 observations on the domain shown



"Super-observations" created for smartphone observations

- Observations location identical for every experiment
    - 150 METAR observations
    - 3,508 smartphone observations



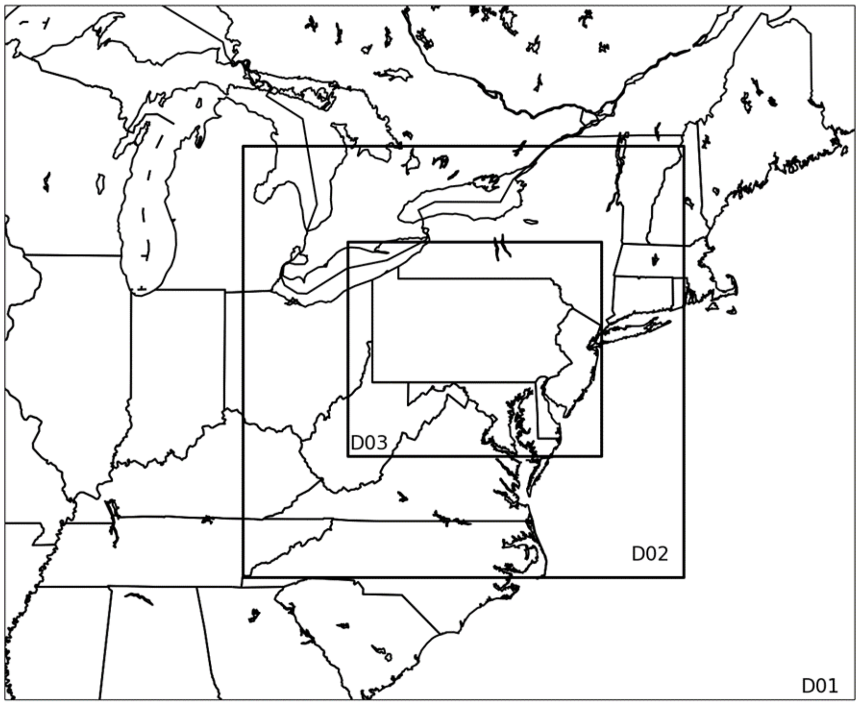
Standard deviation for each spatial grid averaged over entire domain... and then averaged for entire 75 days of data

Observation Error:

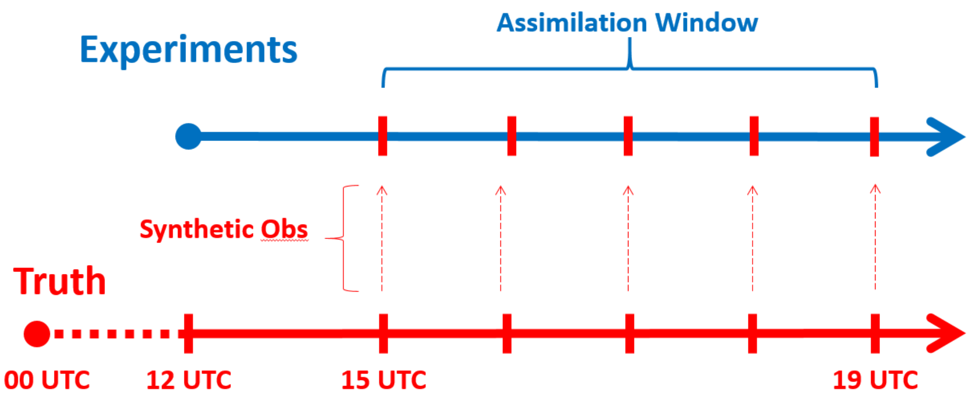
2.34 hPa

Challenge: representativeness errors and quality control.

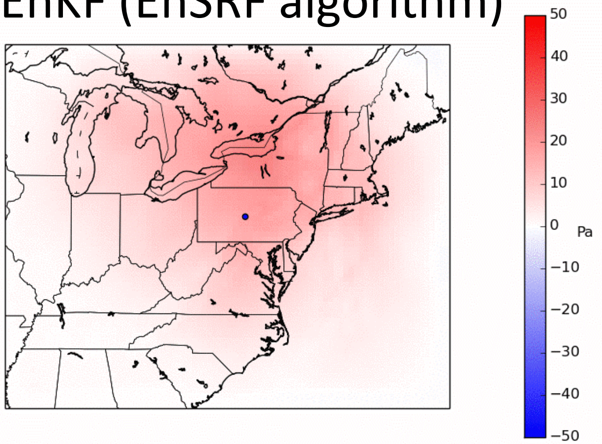
Observing System Simulation Experiment



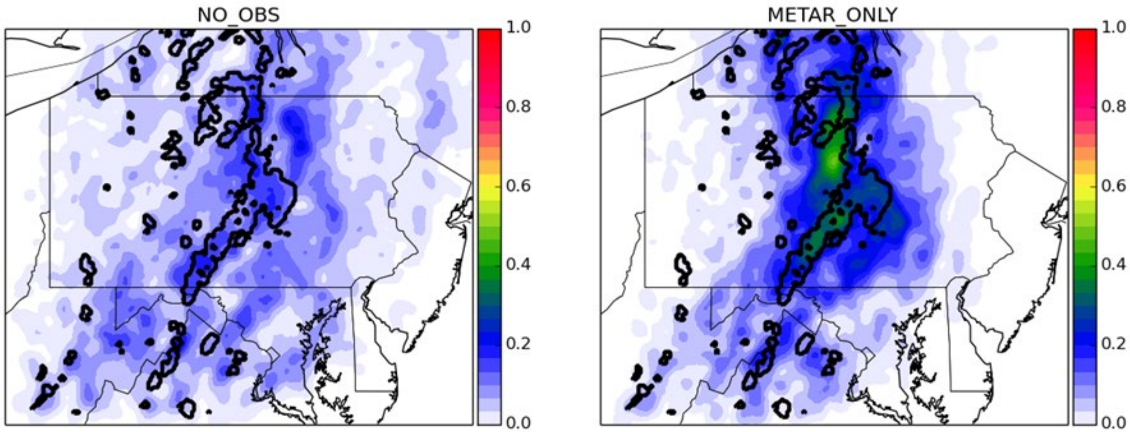
- WRF-ARW Version 3.7 and the PSU WRF-EnKF Data Assimilation System
- 27, 9, and 3 km grid spacing in domains
- No convective parameterization in D03
- Truth created from single deterministic WRF forecast initialized at 00 UTC 20 April 2015



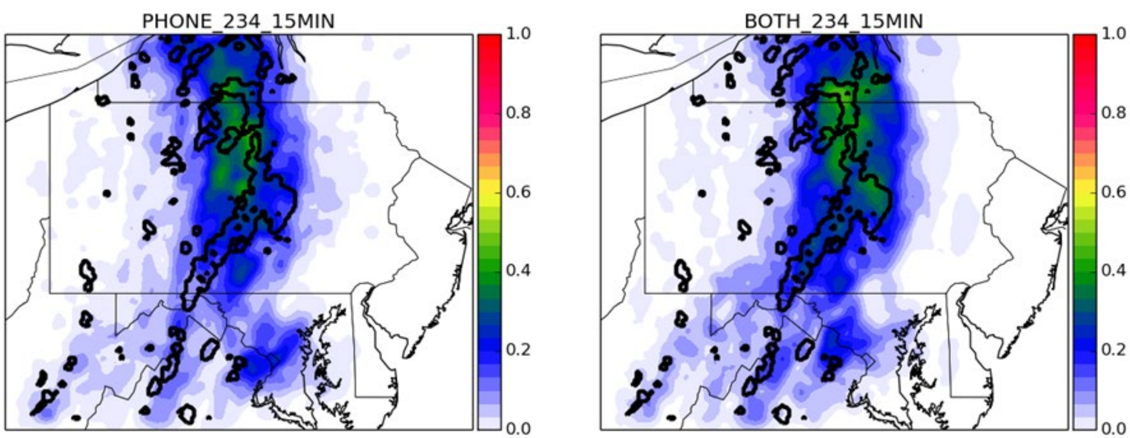
- Use PSU EnKF (EnSRF algorithm)



Results



Neighborhood Ensemble Probability for ≥ 35 dBZ: 23 UTC
Truth in black contour



All observations, regardless of type, led to higher probability regions that better match the truth than the NO_OBS case

Quantitative assessments included RMSD, FSS, ROC.

Smartphone observations can have a positive impact on the ensemble forecast of a convective event in a regional model using EnKF data assimilation.

Rapid assimilation of smartphone data (15 minute cycling) improved analysis results.

Assimilating only smartphone pressures successfully updated other state variables (e.g. temperature, winds), with skill approaching that of conventional networks.

Smartphone observations could be used in conjunction with conventional observations or possibly as the sole source of observations in a data-denied area.

Physically Interpretable Deep Learning for Convective Initiation Nowcasts

Da Fan, Steven J. Greybush

Short-term (0-1h) convective initiation (CI) nowcasting is challenging.

Goal: Evaluate the performance of deep learning on (radar-assessed) CI nowcasts using GOES-R and HRRR data as predictors, and use interpretable machine learning to assess the most important predictors

Definition of CI is adapted from Colbert et al. (2019). Grid points are defined as CI when the following conditions are met:

- **Composite reflectivity ≥ 35 dBZ**,
- No points within 15 km exhibit composite reflectivity ≥ 35 dBZ in the past 11 min, and
- At least three convective cells are grouped in the track that a CI belongs to.

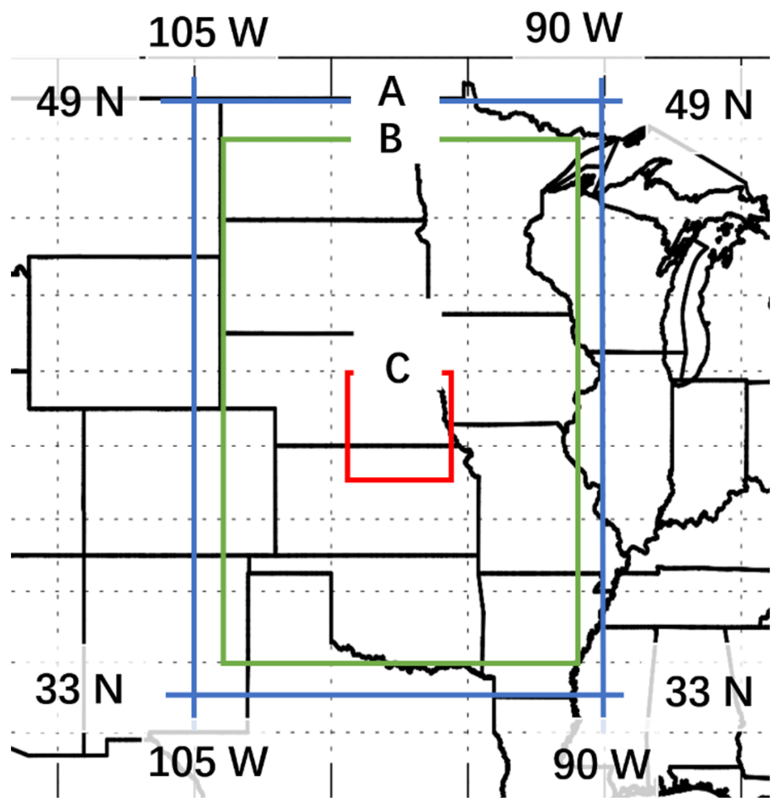
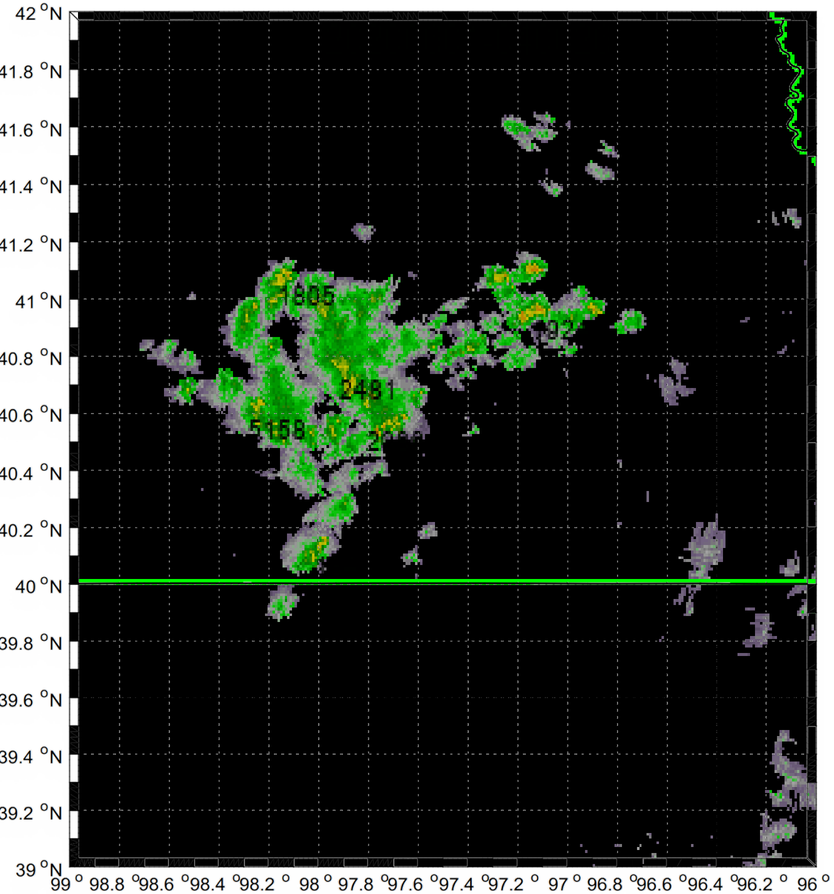


FIG. 1: Total area of study (blue outline labeled with A),

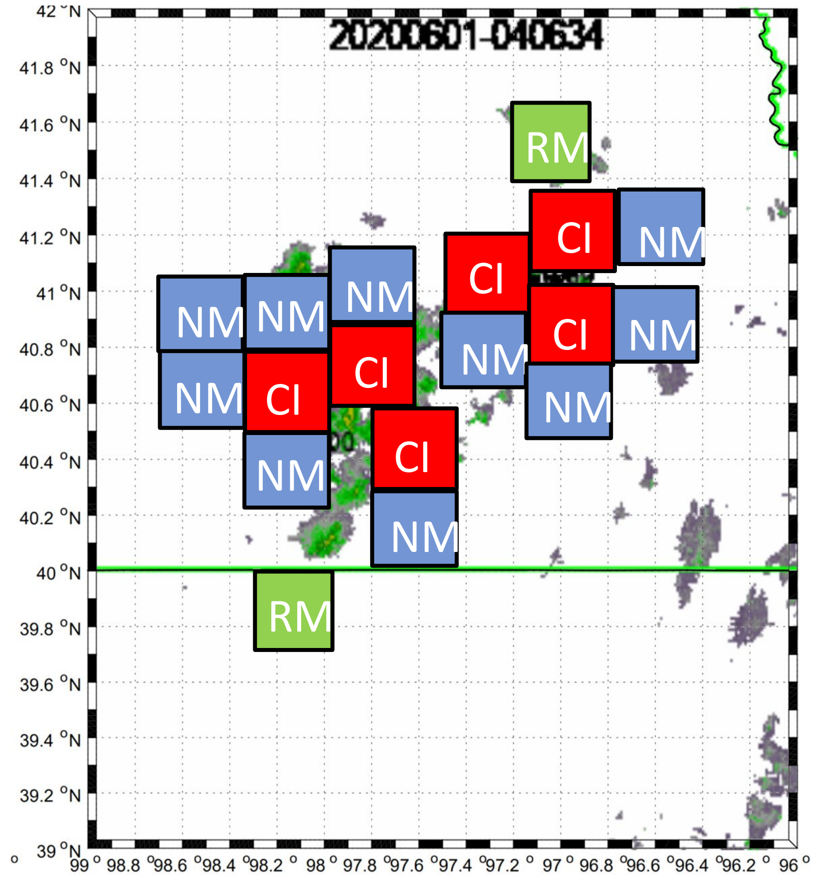
Data period: June 1-30, 2020

Assembling Training, Validation, and Testing Datasets



Total number of samples: 65849

- CI class (53.5%)
- NM class (35.7%)
- RM class (10.8%)



Data split:

- Training (60%)
- Validation (20%)
- Testing (20%): June 27-30

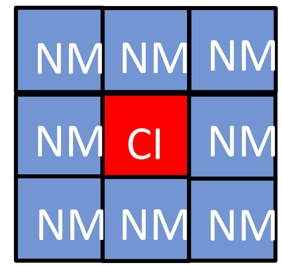
Predictor and Predictand

Predictors

Satellite and NWP predictors	Abbr
6.9-μm BT (K)	CH9 BT
7.3-μm BT (K)	**CH10 BT**
8.4-μm BT (K)	CH11 BT
9.6-μm BT (K)	CH12 BT
10.3-μm BT (K)	CH13 BT
11.2-μm BT (K)	**CH14 BT**
CAPE using highest average theta-e in six lowest 30 mb layers	**CAPE-6L**
Mixed-layer CAPE using average theta-e of the three lowest 30 mb layers	**MLCAPE**
Relative Humidity at 2 m above ground level	**RH-2m**
Relative humidity at highest tropospheric freezing level	**RH-FL**
Medium cloud cover	**MCDC**
Temperature at 2 m above ground level	**TMP**
Volumetric Soil Moisture Content at surface	**SOILW**
V-component of vector wind difference between wind in surface-500 m and 5.5-6.0 km layers above ground level	**VVCSH**

Predictands: 1/0 (whether CI occurred within 30×30 km patch over the next 15 min)

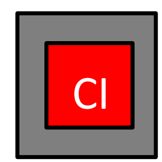
GOES-R
(2-km spacing,
5-min intervals)



NM: near-miss
RM: random

CH9-14: 30×30 km (15×15 points)
15 min prior to the timing of CI

HRRR
(3-km spacing,
hourly intervals)



Variables: 45×45 km (15×15 points)
the nearest hour before the timing of CI (e.g., Var at 3 am for CI at 3:30 am)

Machine learning models

Predictors

Satellite and NWP predictors	Abbr
6.9-μm BT (K)	CH9 BT
7.3-μm BT (K)	**CH10 BT**
8.4-μm BT (K)	CH11 BT
9.6-μm BT (K)	CH12 BT
10.3-μm BT (K)	CH13 BT
11.2-μm BT (K)	**CH14 BT**
CAPE using highest average theta-e in six lowest 30 mb layers	**CAPE-6L**
Mixed-layer CAPE using average theta-e of the three lowest 30 mb layers	**MLCAPE**
Relative Humidity at 2 m above ground level	**RH-2m**
Relative humidity at highest tropospheric freezing level	**RH-FL**
Medium cloud cover	**MCDC**
Temperature at 2 m above ground level	**TMP**
Volumetric Soil Moisture Content at surface	**SOILW**
V-component of vector wind difference between wind in surface-500 m and 5.5-6.0 km layers above ground level	**VVCSH**

- Logistic regression model (LR): representative values over patches (Mecikalski et al. 2015)
- Convolutional neural network (CNN):
  - CNN-Sat: All GOES-R predictors
  - CNN-NWP: All HRRR predictors
  - CNN-SatNWP: All HRRR + CH10 BT + CH14 BT

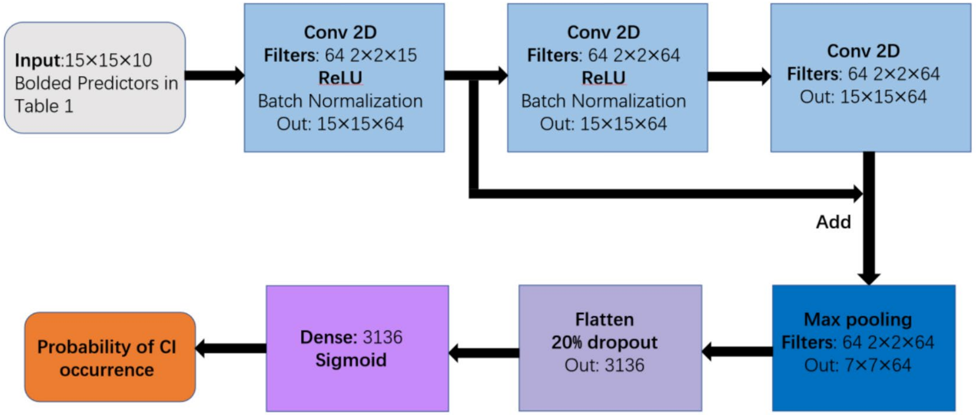


FIG. 2. Schematic of the convolutional neutral network for CI nowcasts.

Summary of Topics and Techniques:

- East Coast Winter Storms Ensemble Predictability and Sensitivity
- IMPACTS Modeling and Data Assimilation
- Tropical Cyclone Parameter Estimation
- Lake-effect Ensemble Design
- Object-based Approaches: Geometry-Sensitive Ensemble Mean
- Novel Datasets: Smartphone Pressure Data Assimilation
- Convection Initiation Deep Learning