Data Mining of High-Resolution Storm-Scale Datasets

PIs: Travis Smith, Chris Karstens,
Jimmy Correia, Kiel Ortega
U. Of Oklahoma / CIMMS (+ NOAA/NSSL&SPC)

CIMMS
Cooperative Institute for Mesoscale Meteorological Studies

NOAA Collaborators

- Lans Rothfusz, Alan Gerard (National Severe Storms Laboratory / Hazardous Weather Testbed)
- NWS Storm Prediction Center staff
- many NWSFO staff
- NCEI

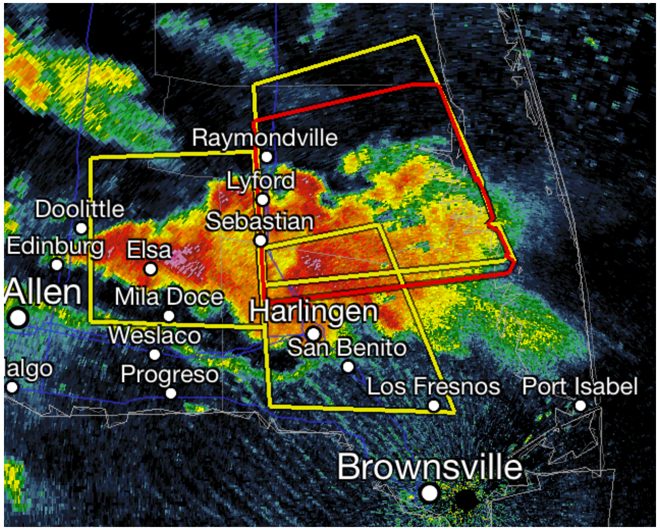
Big Picture

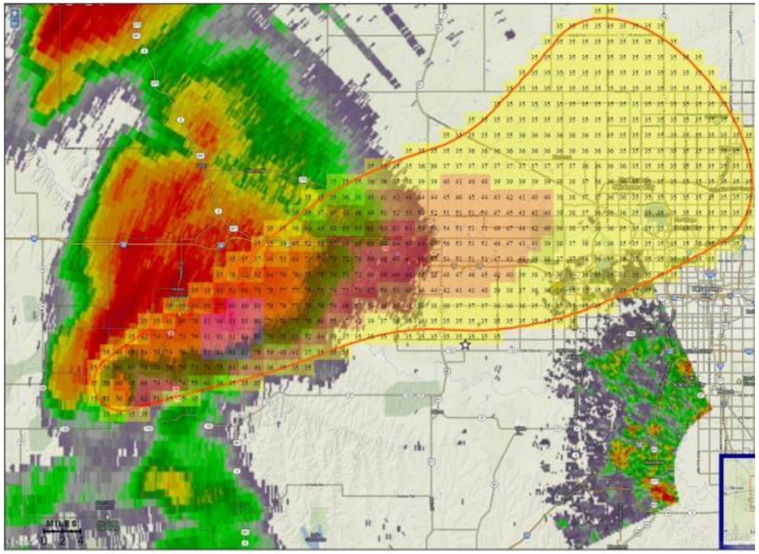
Current warning paradigm:

- Make extrapolative prediction based on radar and storm spotter observations
- “Warn-on-Detection”

Forecasting A Continuum of Environmental Threats (FACETs):

- Continuously updating flow of information
- Storm-scale ensembles (“Warn on Forecast”)
- Probabilistic Hazard Information





Goals

Measure model improvements at the scale of individual thunderstorms (short term predictability).

Understand the strengths and limitation in the models' simulation of storms and storm evolution over a diverse spectrum of convective modes.

Demonstrate in Hazardous Weather Testbed.

Project Plan

1. Acquire and prepare data sets
2. Develop and refine software
3. Objectively classify convective storms
   - Observational data
   - Model output
4. Compare observed and model storm-typing and severity
5. Real-time testing and evaluation with forecasters

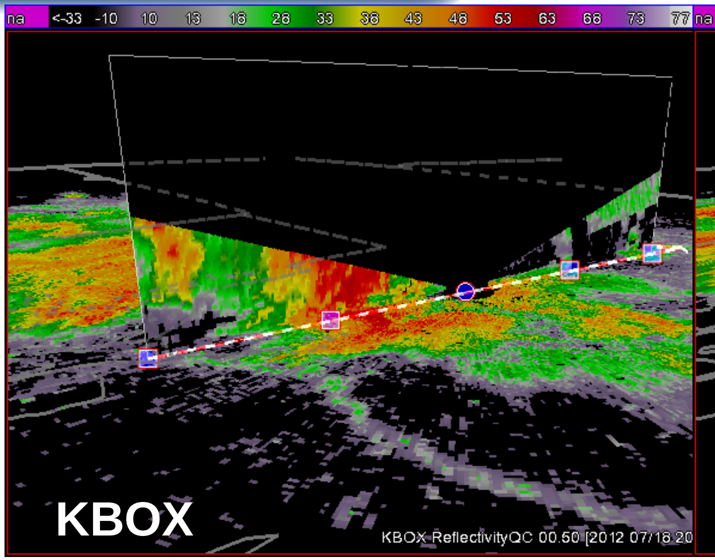
Acquire/Prepare Data Sets

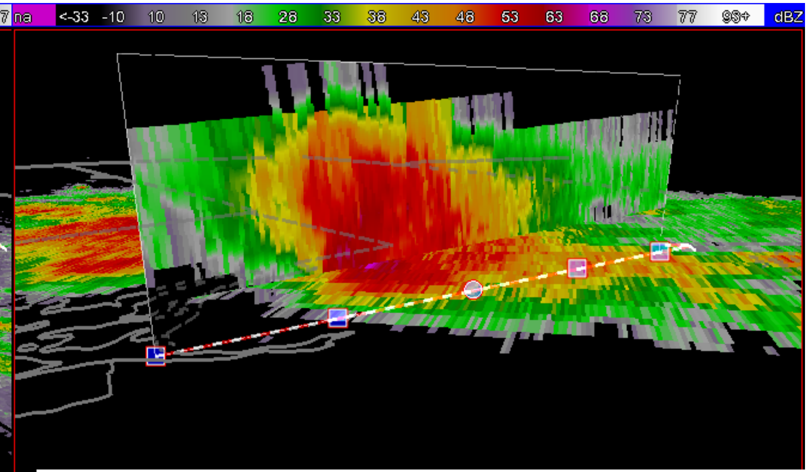
Multi-Year Reanalysis of Remotely Sensed Storms (MYRORSS; pronounced “mirrors”)

Storm Prediction Center database of convective modes

NSSL WRF storm-object data set

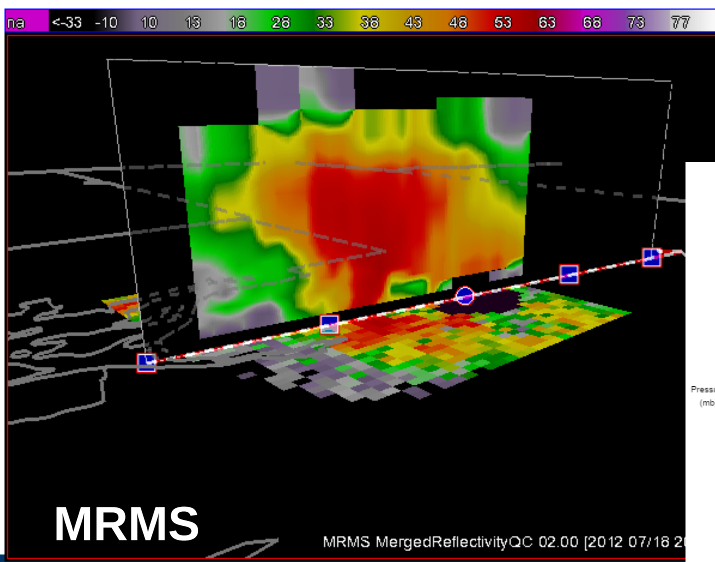
Multi-Radar Multi-Sensor System

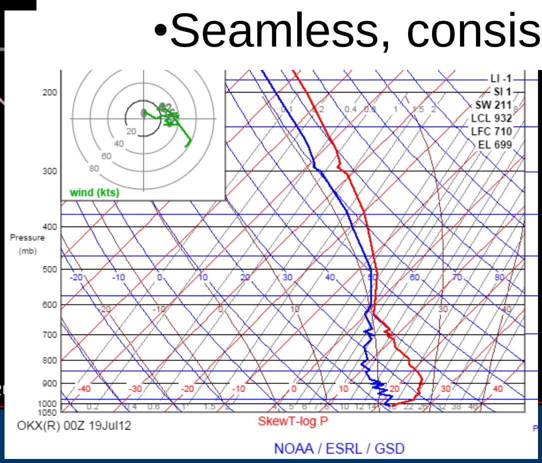


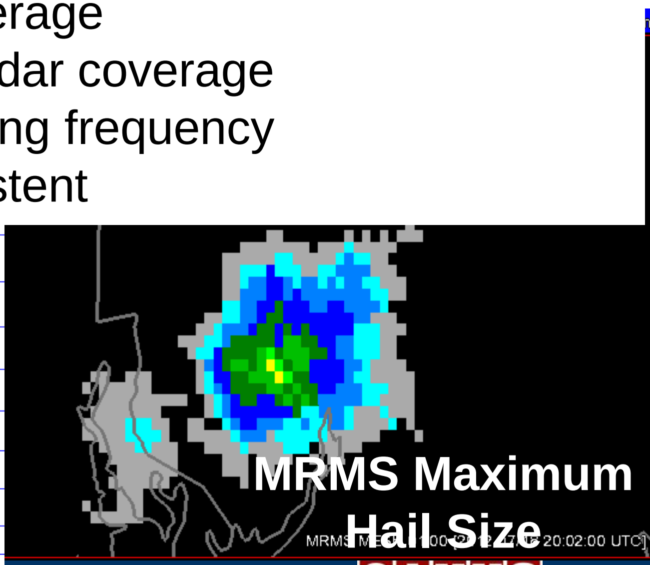


+ other neighboring radars

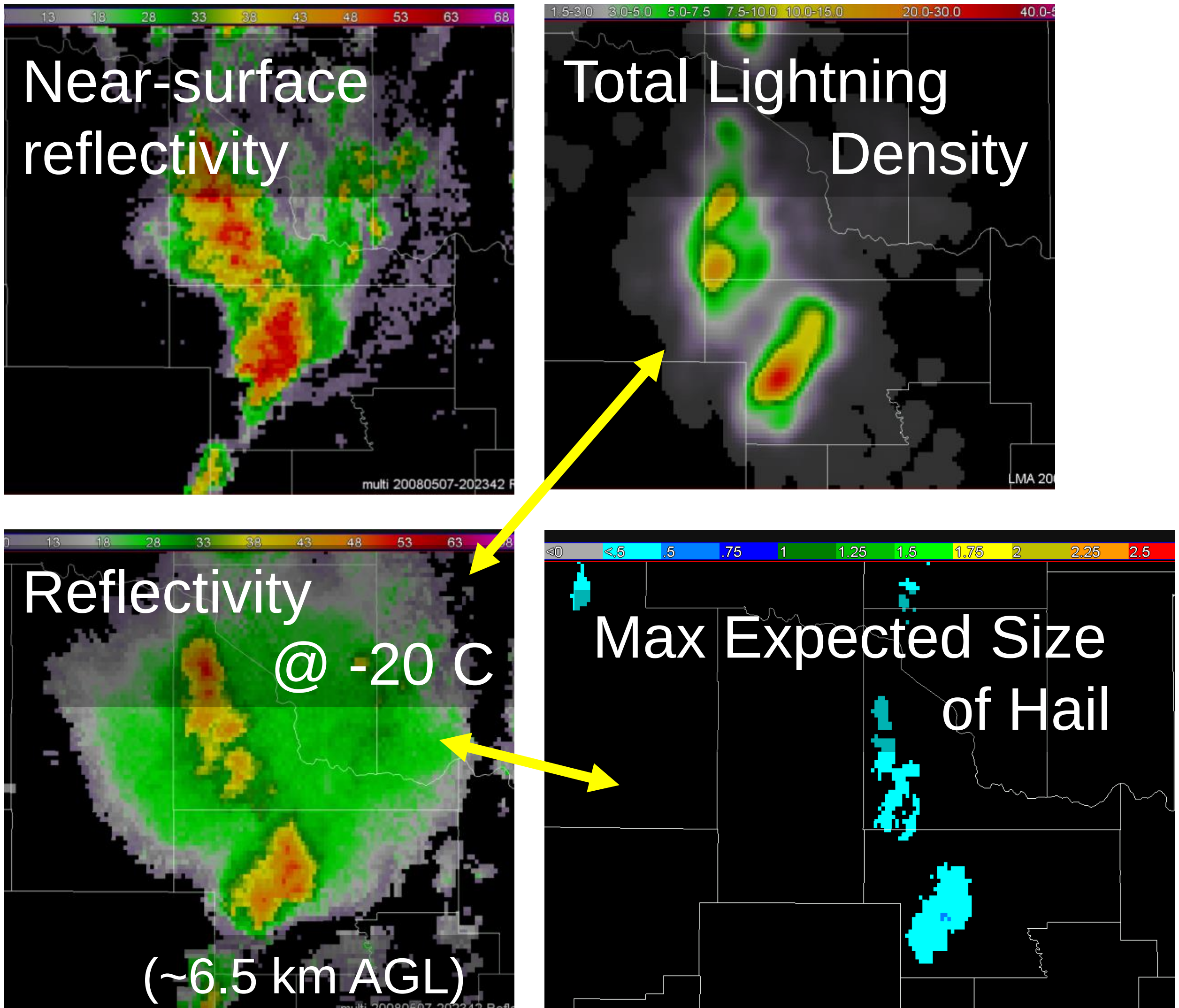
- Radars in network supplement each other:
  - Overlapping coverage
  - Fills in gaps in radar coverage
  - Increased sampling frequency
  - Seamless, consistent





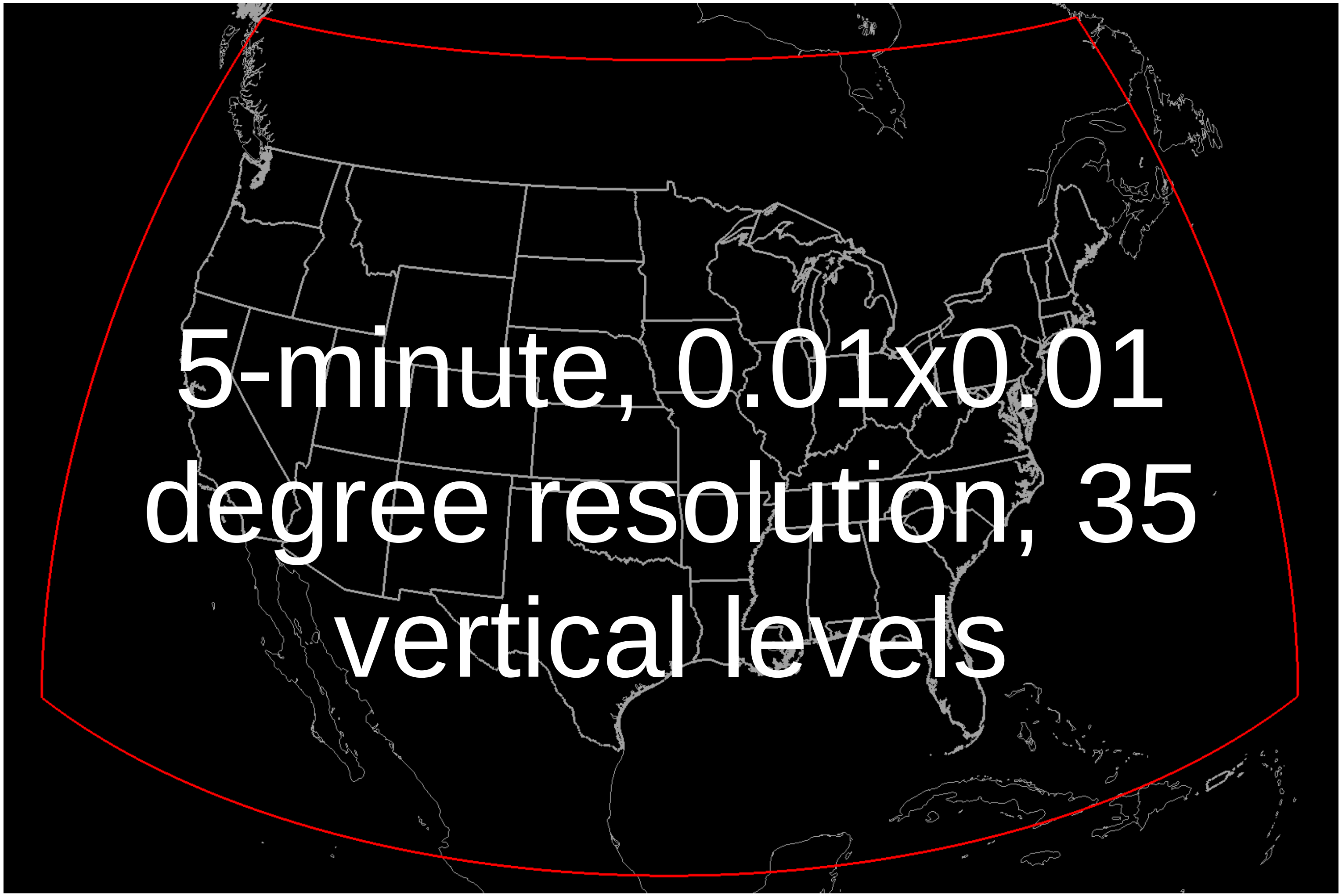


Examples: Multi-sensor data fields



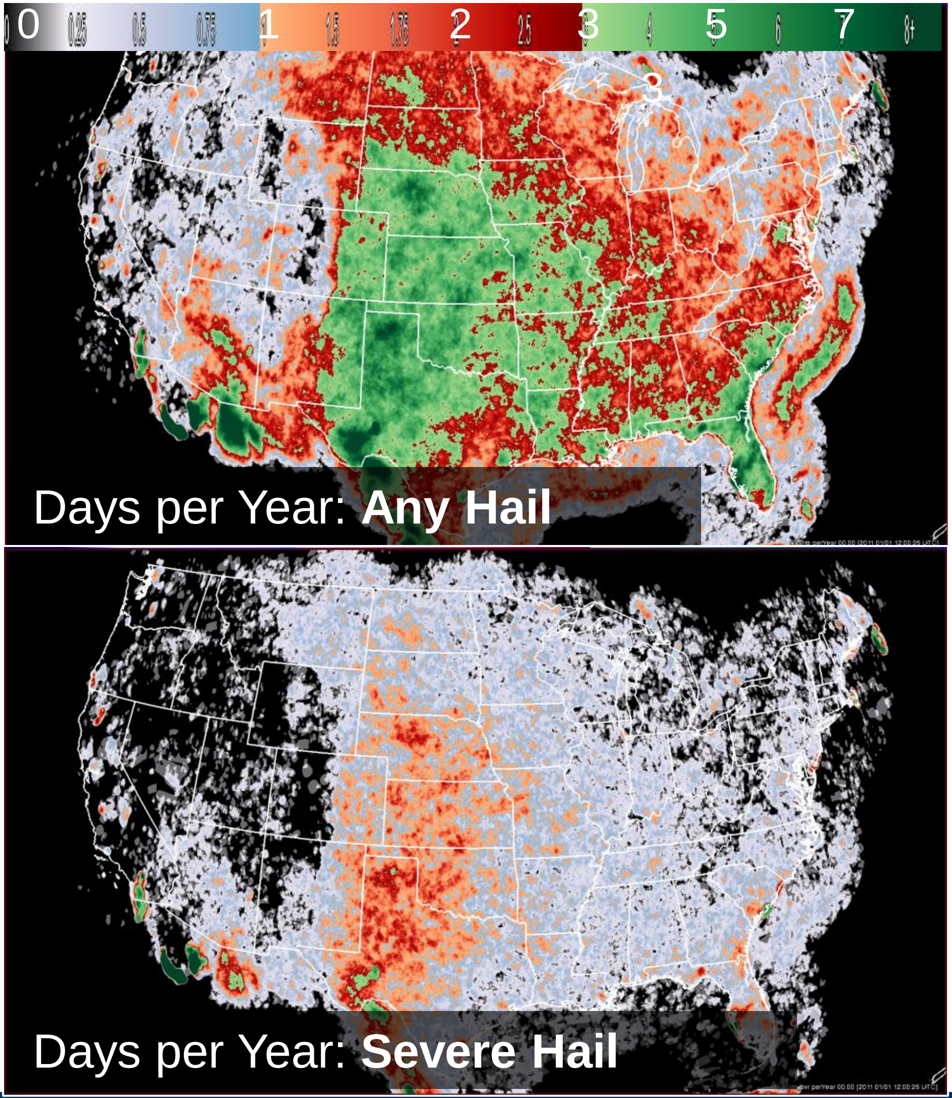
- Show physical relationships between data fields from multiple sensors
- Storm tracks and trends can be generated at any spatial scale, for any data fields

MRMS Domain



5-minute, 0.01x0.01 degree resolution, 35 vertical levels

MYRORSS



Multi-Year Reanalysis of Remotely Sensed Storms (MYRORSS)

- 15+ years of storm statistics
- Data Mining
- MRMS & MYRORSS are foundational to the effort to improve NWS warning services

Storm classification

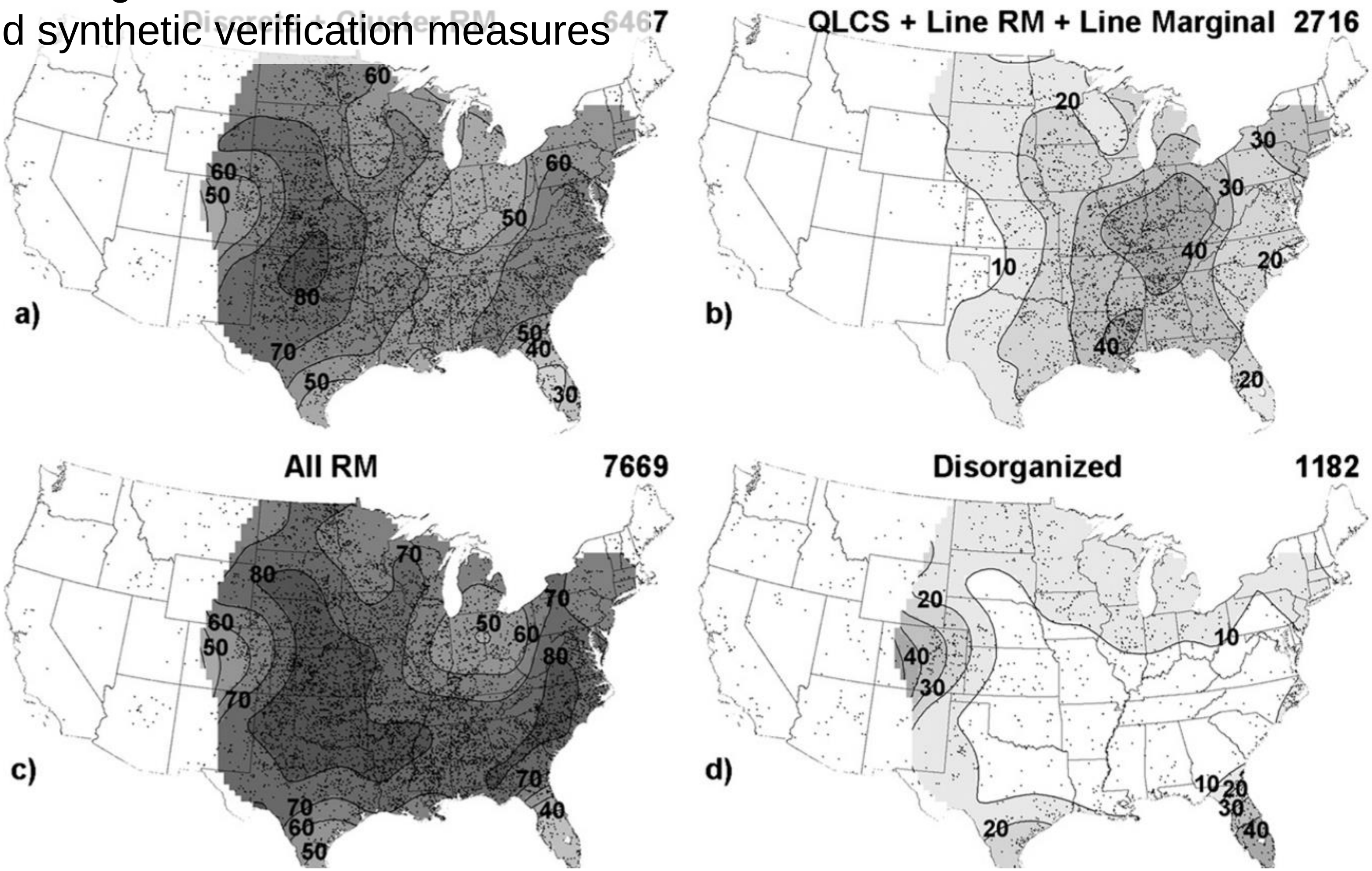
- Disorganized
  - Discrete
  - In Cluster
  - In Line
- QLCS
  - Bow Echo
- Supercell Right Mover
  - Discrete
  - In Cluster
  - In Line
- Supercell Left Mover
  - Discrete
  - In Cluster
  - In Line

Based on:

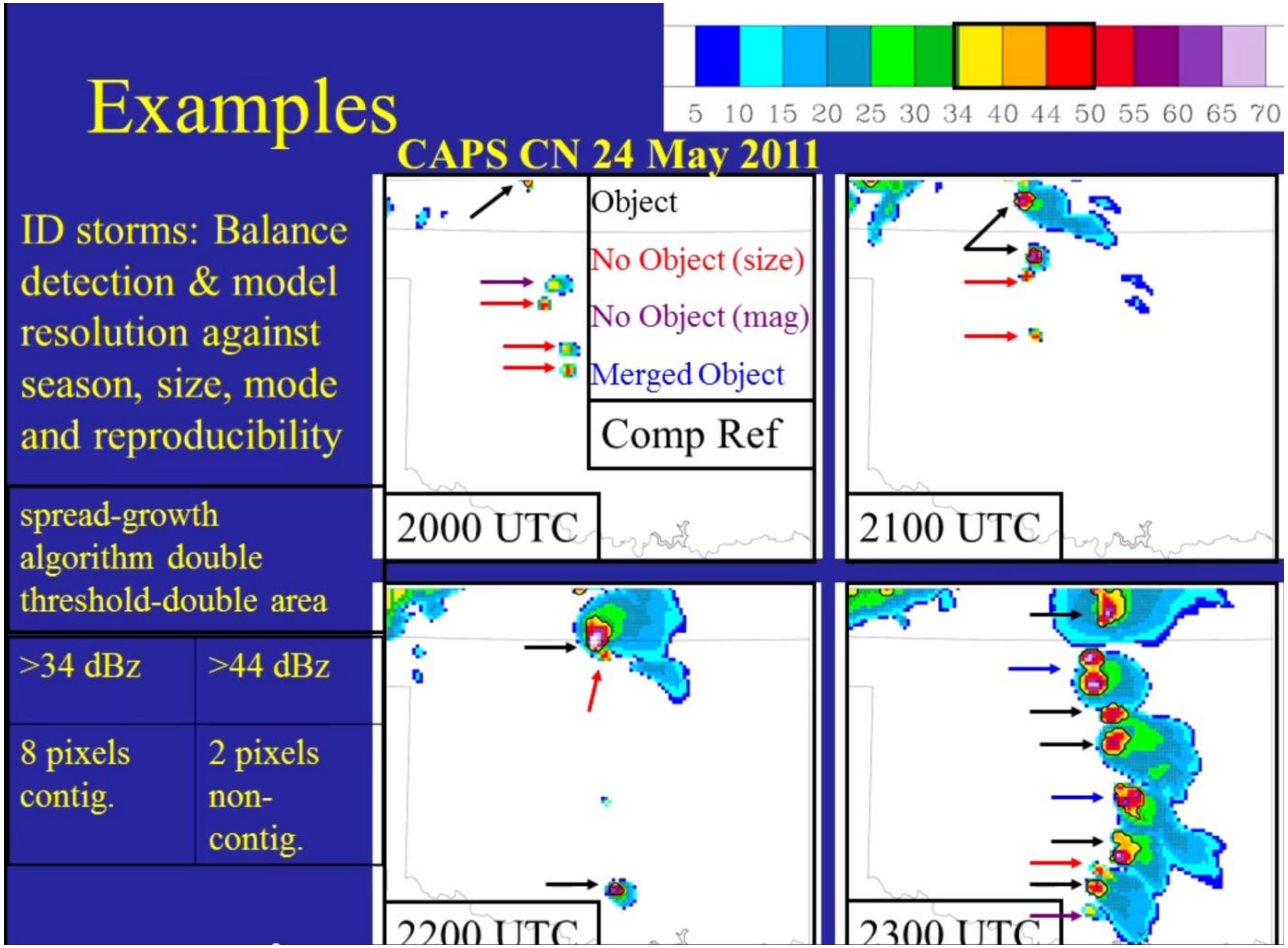
Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135.

Storm Convective Modes Dataset

We supplement Smith et al. 2012, including weak severe, non-severe, and synthetic verification measures



NSSL-WRF Storm Features



Correia, Jr, J., J. Kain, A. J. Clark, 2014: A four year climatology of simulated convective storms from NSSL WRF. 94th AMS Annual Meeting, Atlanta, GA, J11.2.

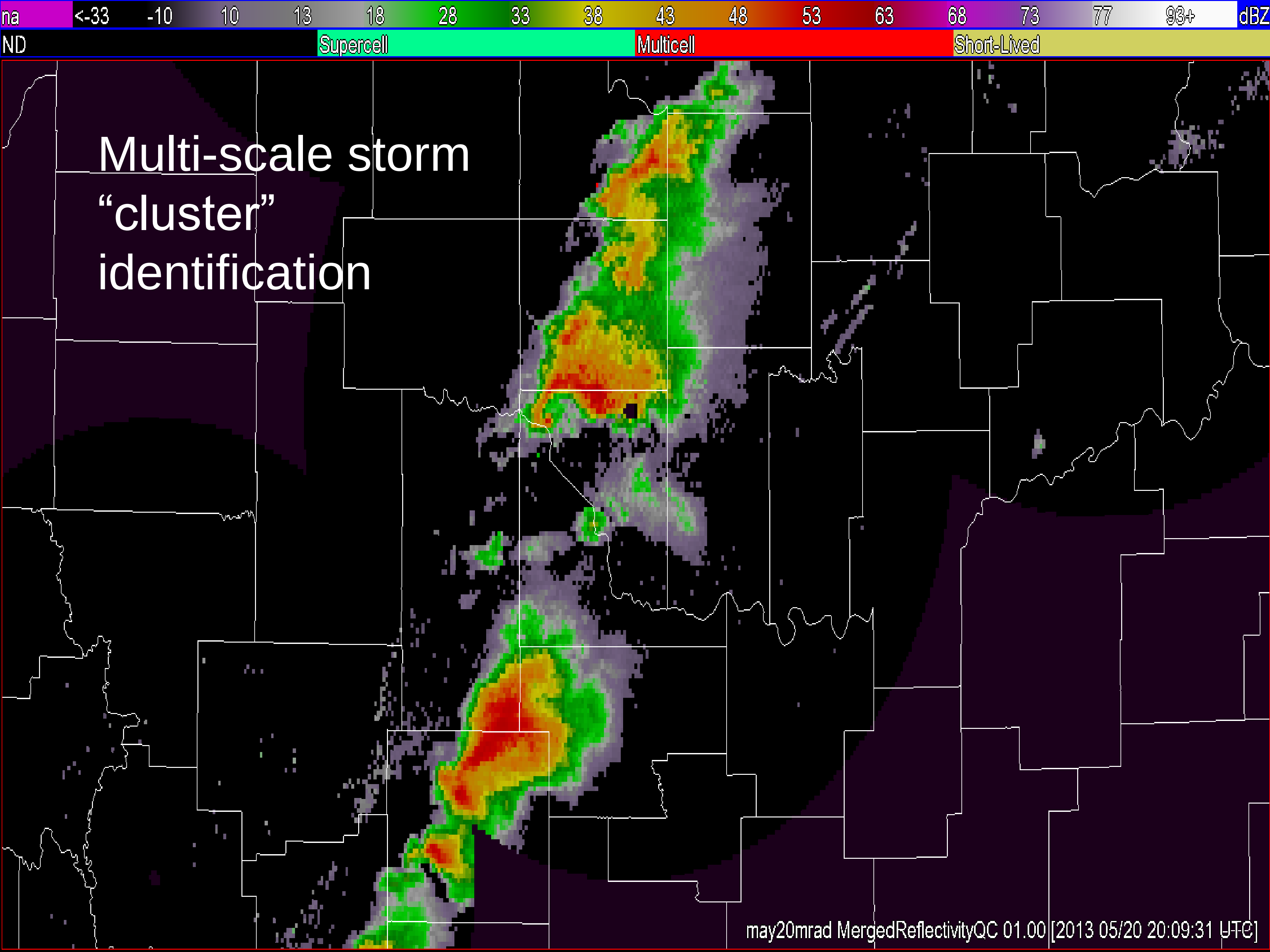
Develop / Refine Software

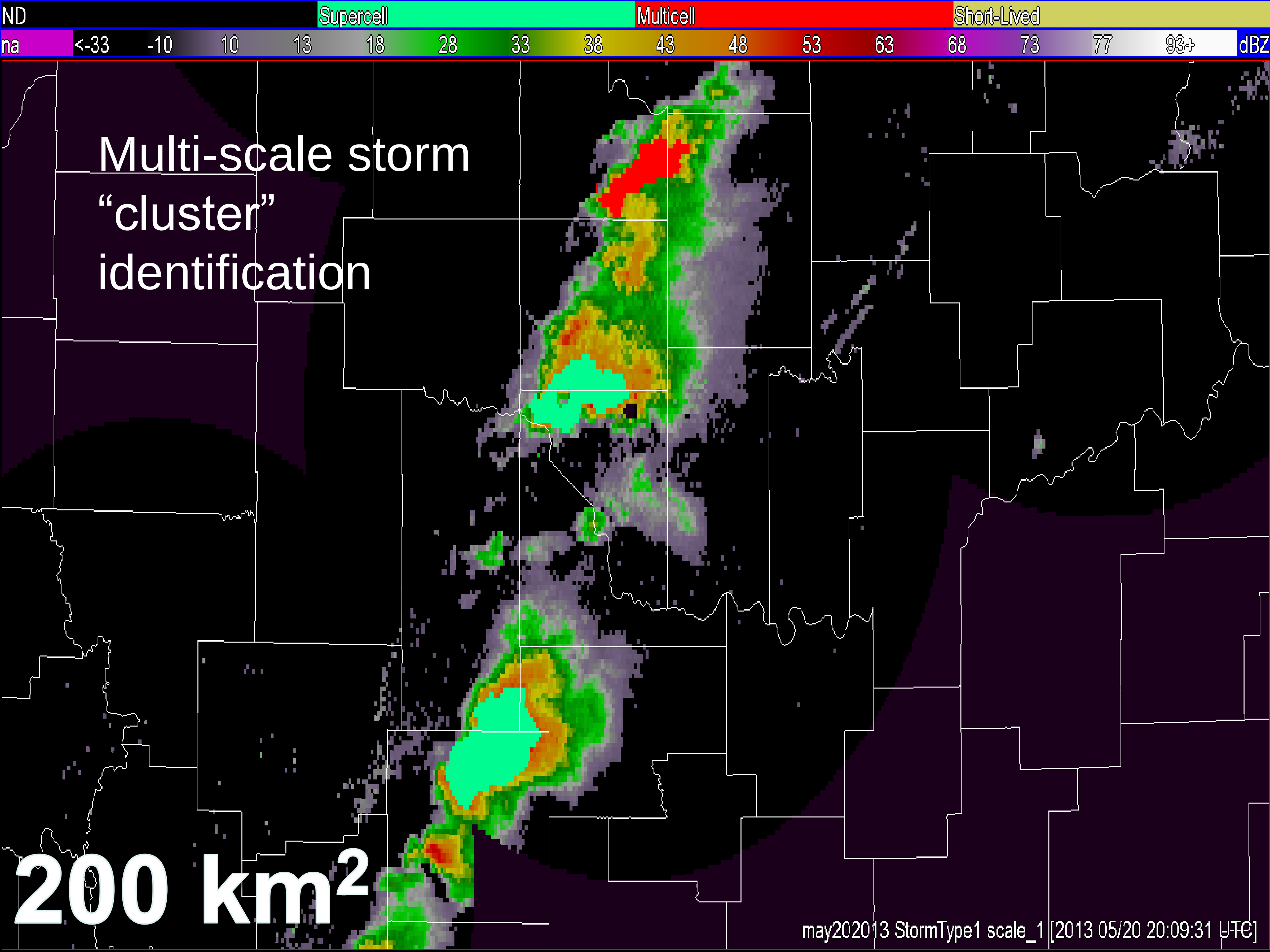
Storm object identification / tracking

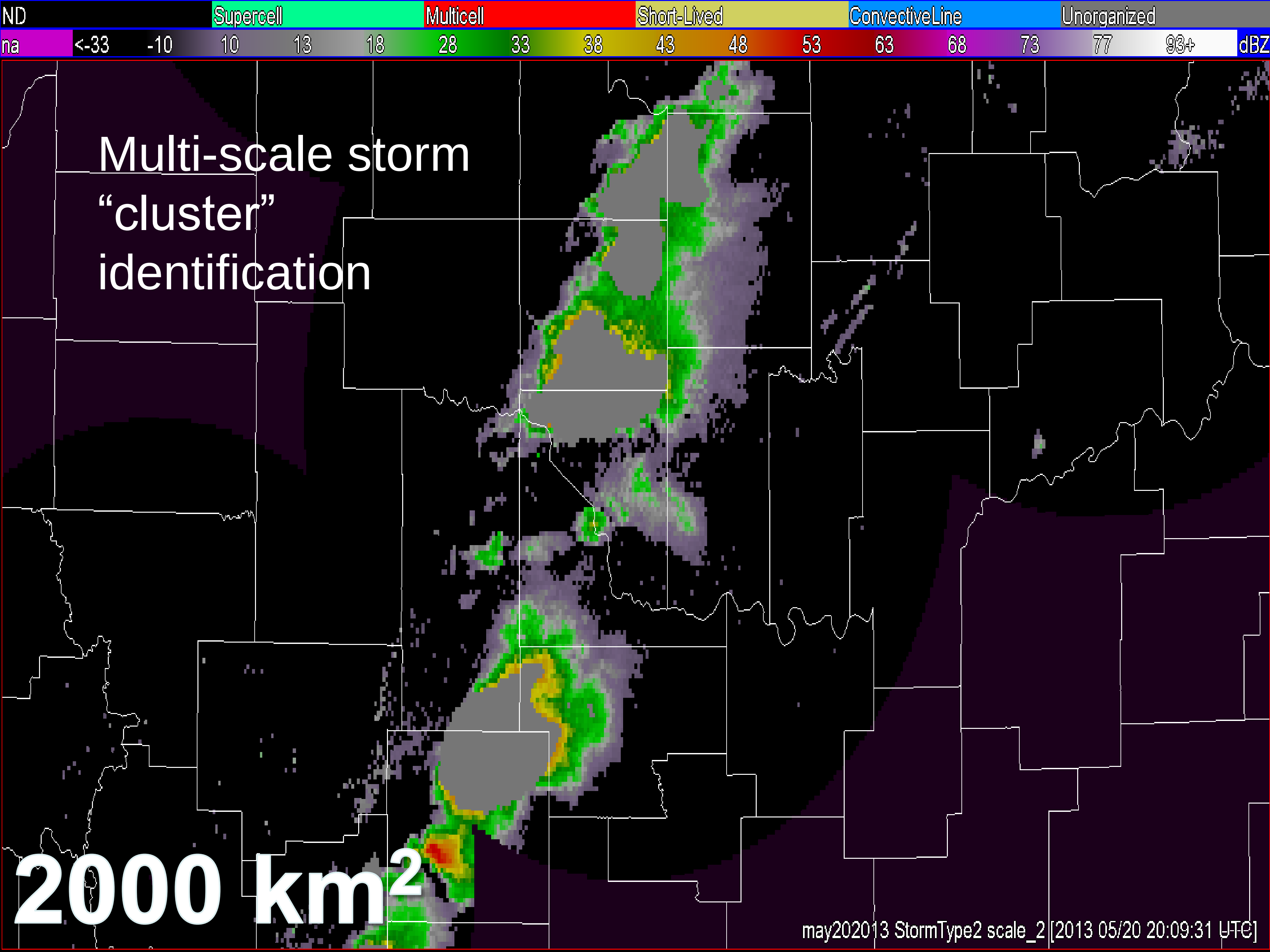
Machine learning

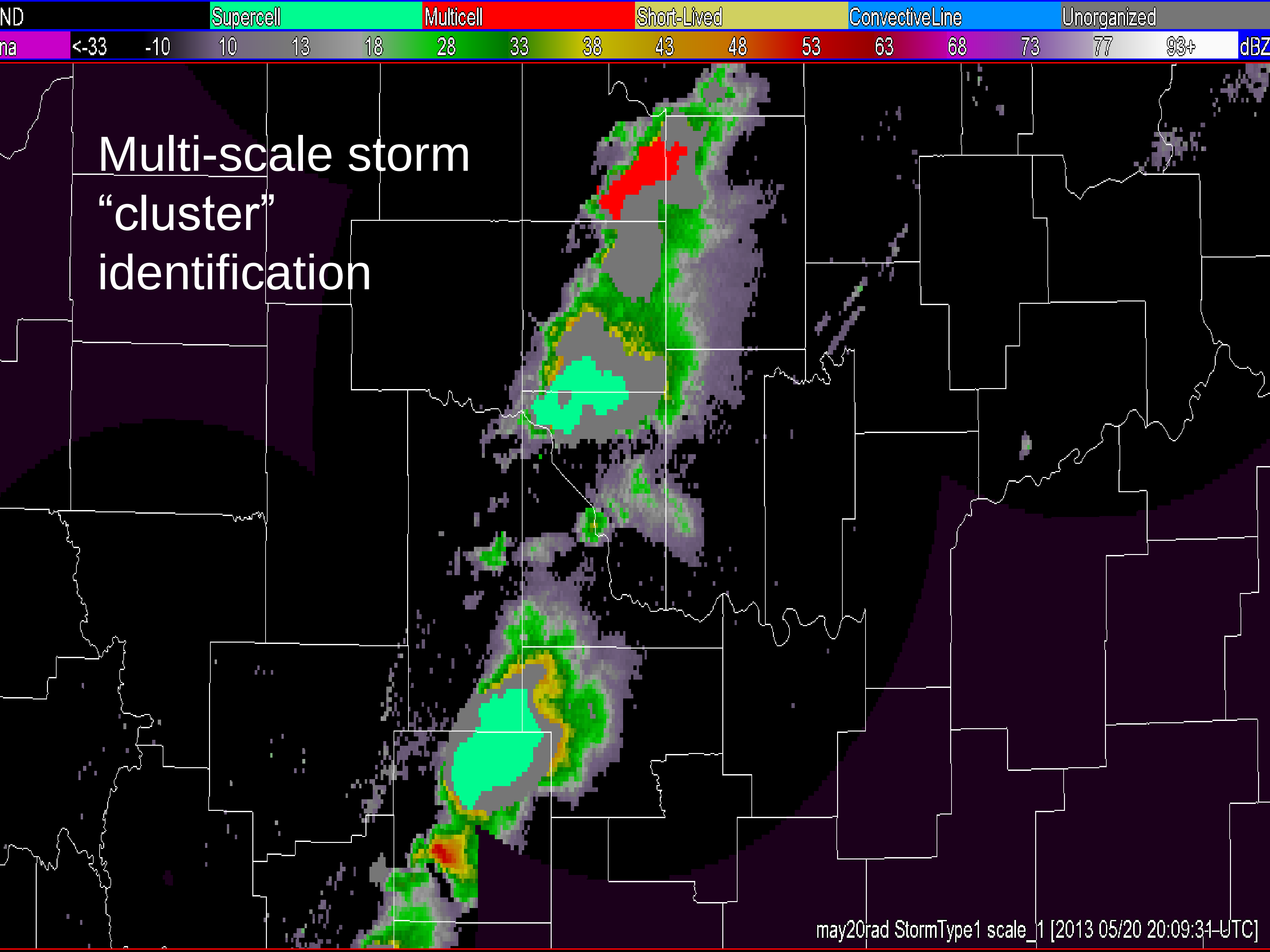
Visualization / blending of MRMS + model

Multi-scale storm "cluster" identification




ND
Supercell
Multicell
Short-Lived
na
<-33
-10
10
13
18
28
33
38
43
48
53
63
68
73
77
93+
dBZ
Multi-scale storm “cluster” identification
200 km²
may202013 StormType1 scale_1 [2013 05/20 20:09:31 UTC]

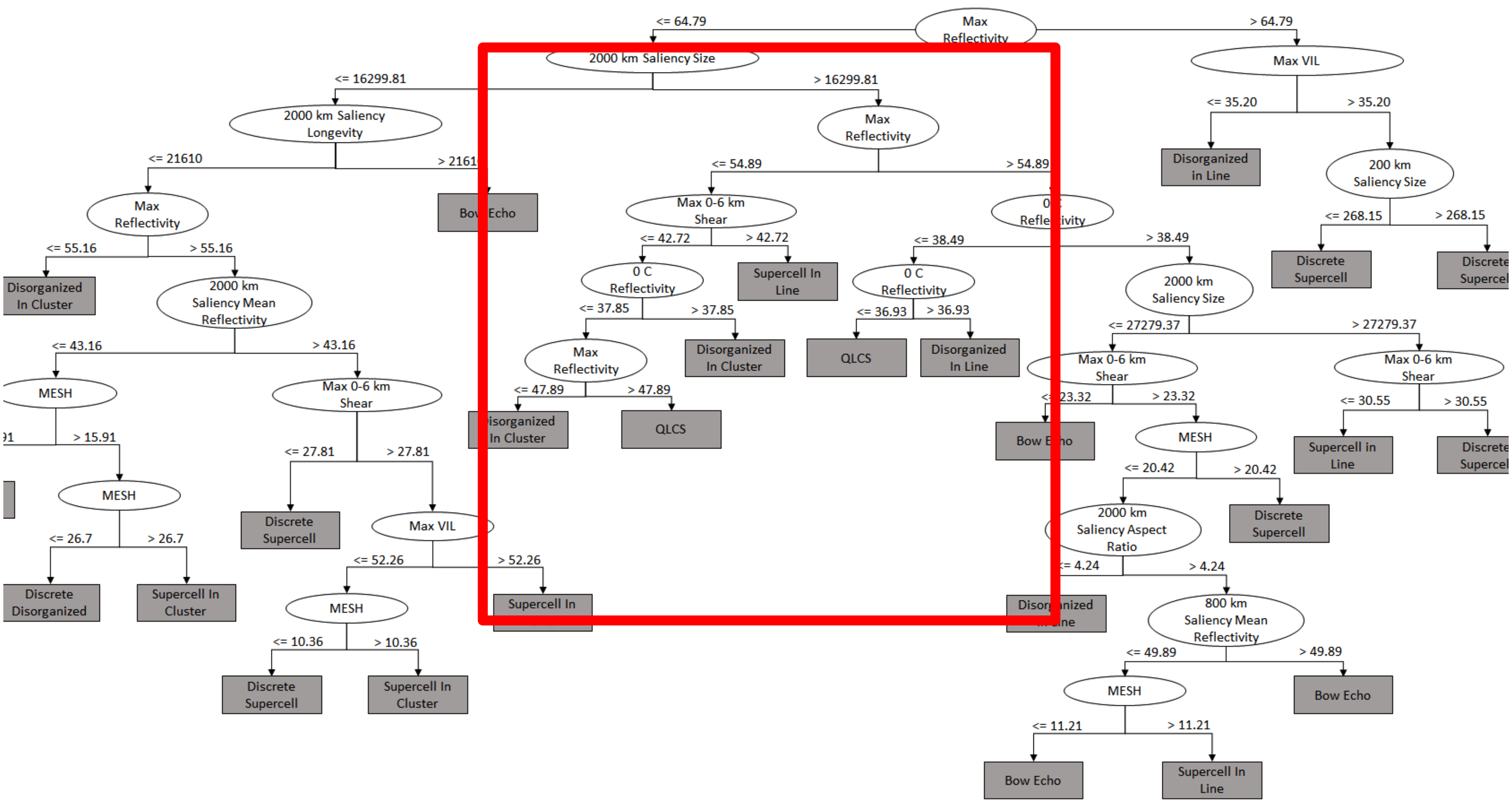

ND
Supercell
Multicell
Short-Lived
ConvectiveLine
Unorganized
na
<-33
-10
10
13
18
28
33
38
43
48
53
63
68
73
77
93+
dBZ
Multi-scale storm “cluster” identification
2000 km2
may202013 StormType2 scale_2 [2013 05/20 20:09:31 UTC]

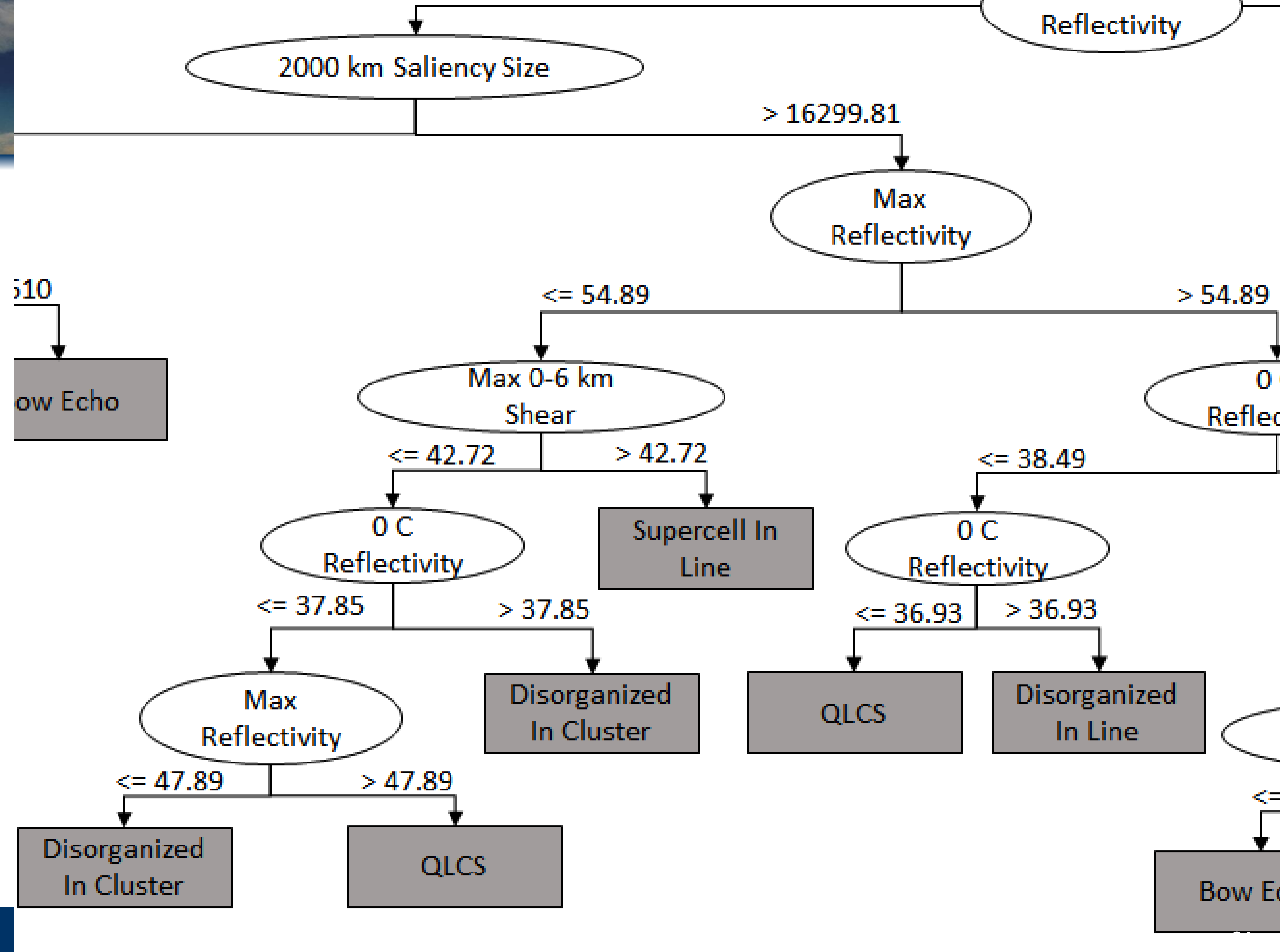

ND
Supercell
Multicell
Short-Lived
ConvectiveLine
Unorganized
na
<-33
-10
10
13
18
28
33
38
43
48
53
63
68
73
77
93+
dBZ
Multi-scale storm
“cluster”
identification
may20rad StormType1 scale_1 [2013 05/20 20:09:31 UTC]

Storm classification inputs from MYRORSS / MRMS

- Storm Attribute
- -20 C Merged Reflectivity
- 0 C Merged Reflectivity
- Aspect Ratio
- 0-2 km Merged Azimuthal Shear
- 3-6 km Merged Azimuthal Shear
- 0-6 km Shear Magnitude
- 0-1 km Storm Relative Helicity
- 0-3 km Storm Relative Helicity
- Longevity
- Maximum Expected Size of Hail (MESH)
- Max 30 Minute MESH
- Most Unstable CAPE
- Most Unstable LCL Height
- Probability of Severe Hail (POSH)
- Quality Controled Merged Reflectivity Composite
- Severe Hail Index (SHI)
- Storm Size
- Surface CAPE
- Surface Dewpoint
- Surface Temperature
- Vertically Integrated Liquid (VIL)

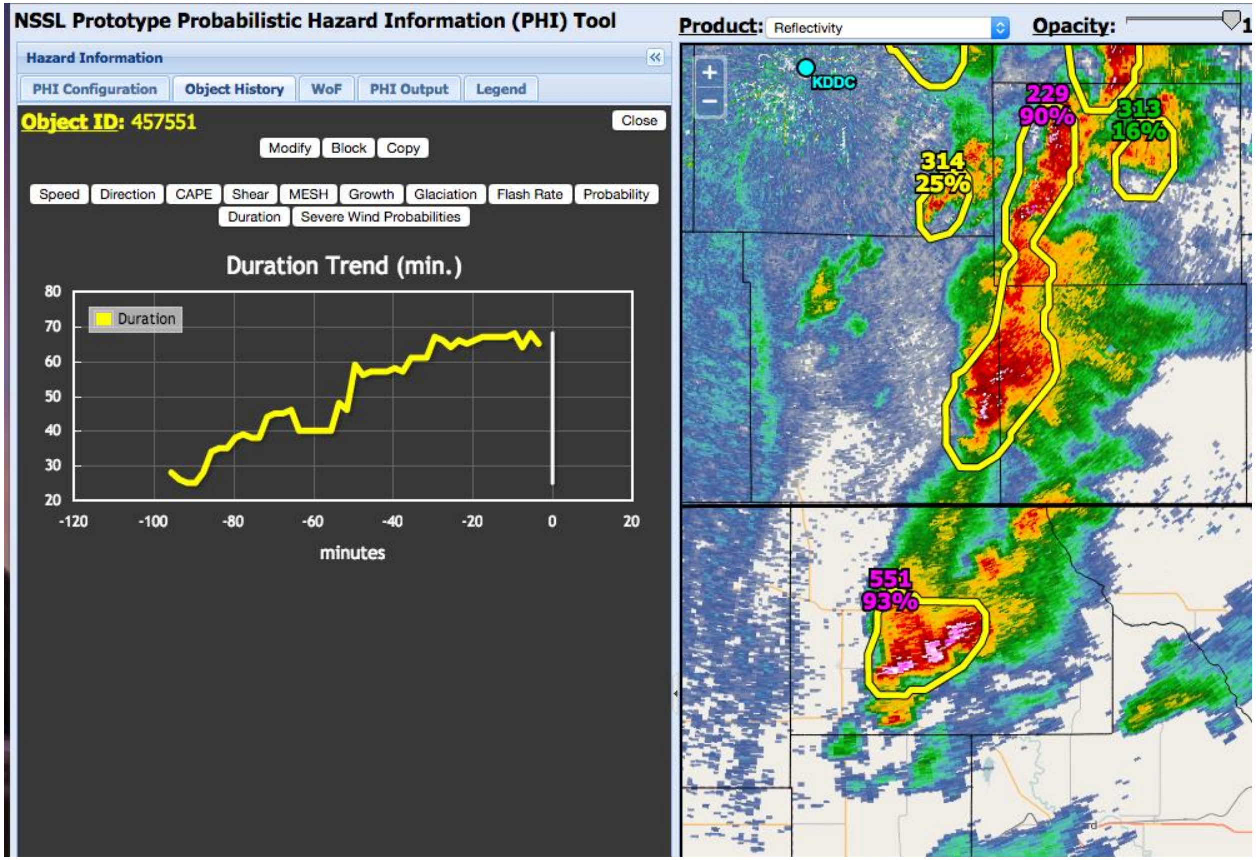
Storm classification: Example Decision Tree





Reflectivity
2000 km Saliency Size
> 16299.81
Max Reflectivity
<= 54.89
> 54.89
Max 0-6 km Shear
<= 42.72
> 42.72
0 C Reflectivity
Supercell In Line
<= 37.85
> 37.85
Max Reflectivity
Disorganized In Cluster
<= 47.89
> 47.89
Disorganized In Cluster
QLCS
<= 38.49
0 C Reflectivity
<= 36.93
> 36.93
QLCS
Disorganized In Line

Machine Learning

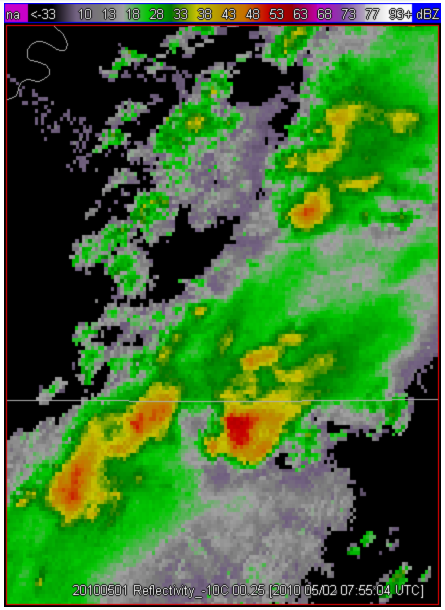


Trend of the expected storm lifetime.

Tested on storm lifetime forecast

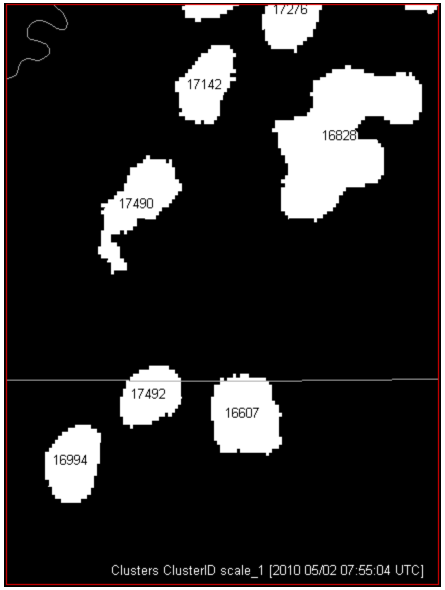
- gradient boosted regression trees
- random forests
- logistical regression
- AdaBoosting

MYRORSS Data Mining



1) Reflectivity (or other image that can be clustered)

2) Storm Clustering and Tracking



3) Then extract storm properties:

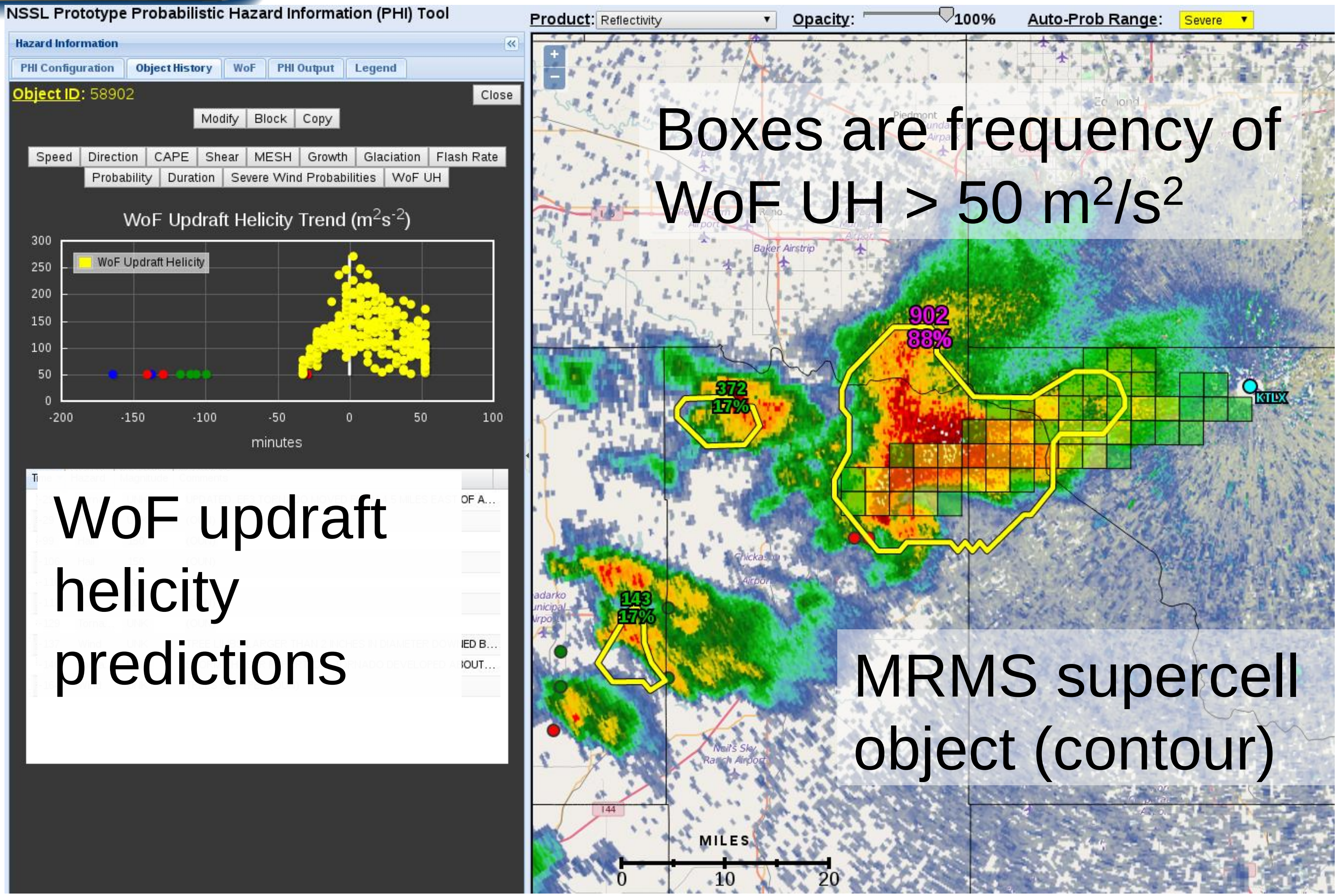
- Other MRMS data for each cluster (radar, satellite, lightning)
- Background environment (from NWP model analysis)

12 analyses / hour
X 24 hours / day
X 365 days / year

105,120 analyses / year

Millions of storms / year

HWT Real-Time Visualization



Accomplishments

1. Acquire/prepare data sets
    - Multi-Year Reanalysis of Remotely Sensed Storms (MYRORSS; pronounced “mirrors”)
    - Storm Prediction Center database of convective modes (+ supplemental data)
    - NSSL WRF storm-object data set
2. Develop / refine software
    - Storm object identification / tracking
    - Machine learning
    - Visualization / blending of MRMS + model

Ongoing work / future plans

Objectively classify storms

- Machine learning algorithms applied to storm object data
- Distributions of storm type and lifetime based on environment

Compare observed and model storm-typing and severity

Real-time testing and evaluation with forecasters (Spring 2017)