Comparison of Object-Based and Grid-Based Verification of Warn-on-Forecast System HAILCAST Forecasts

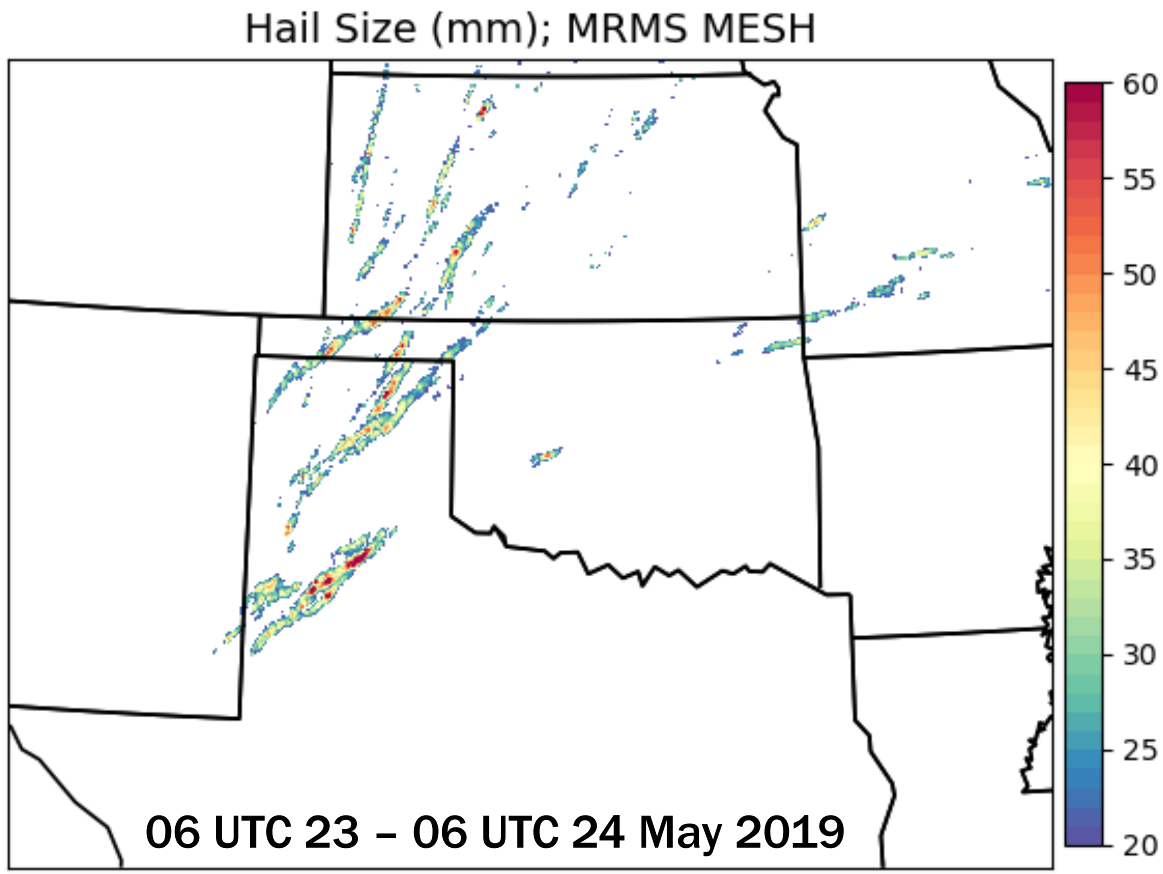
Rebecca Adams-Selin[1], Christina P. Kalb[2], Patrick S. Skinner[3], Tara Jensen[2]

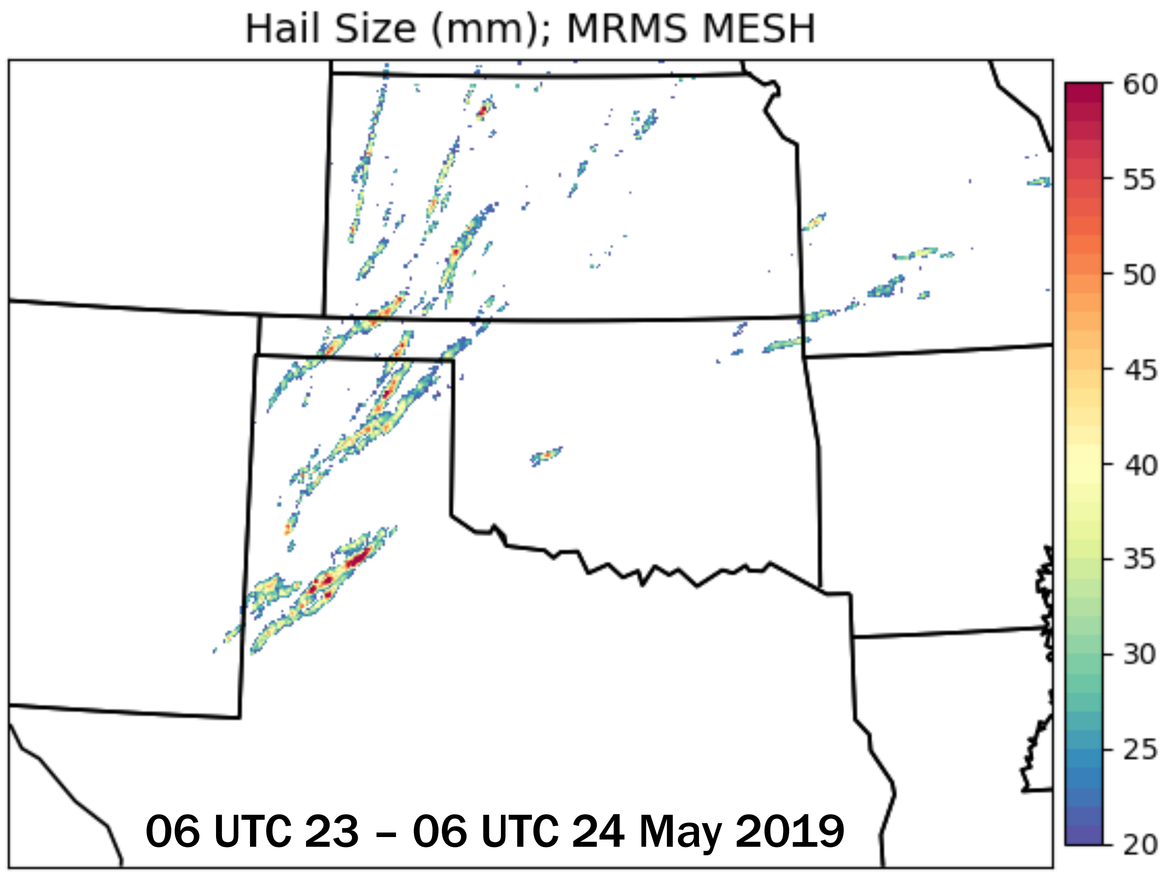
[1]Atmospheric and Environmental Research (AER)
[2]National Center for Atmospheric Research Developmental Testbed Center (NCAR/DTC)
[3]OU/CIMMS and NOAA/NSSL

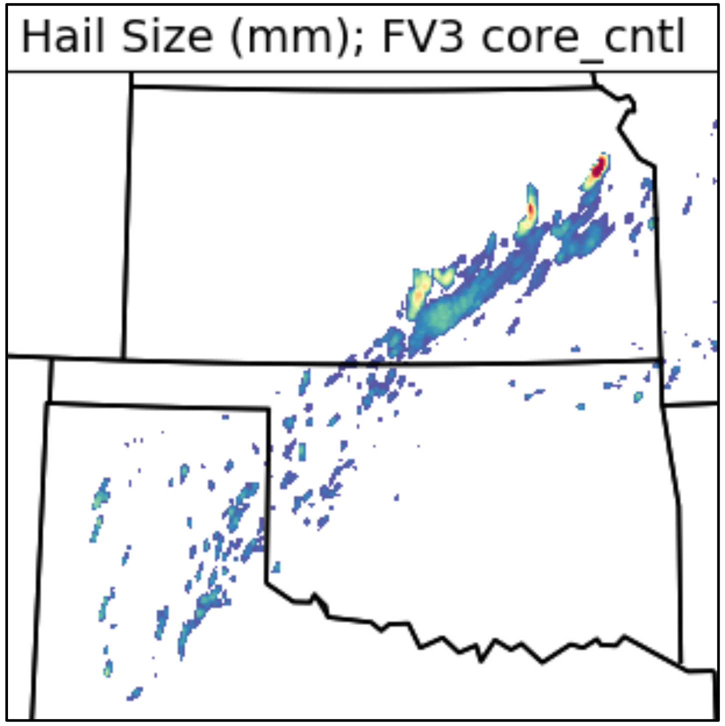
Funding provided by NOAA JTTI award no. NA18OAR4590388

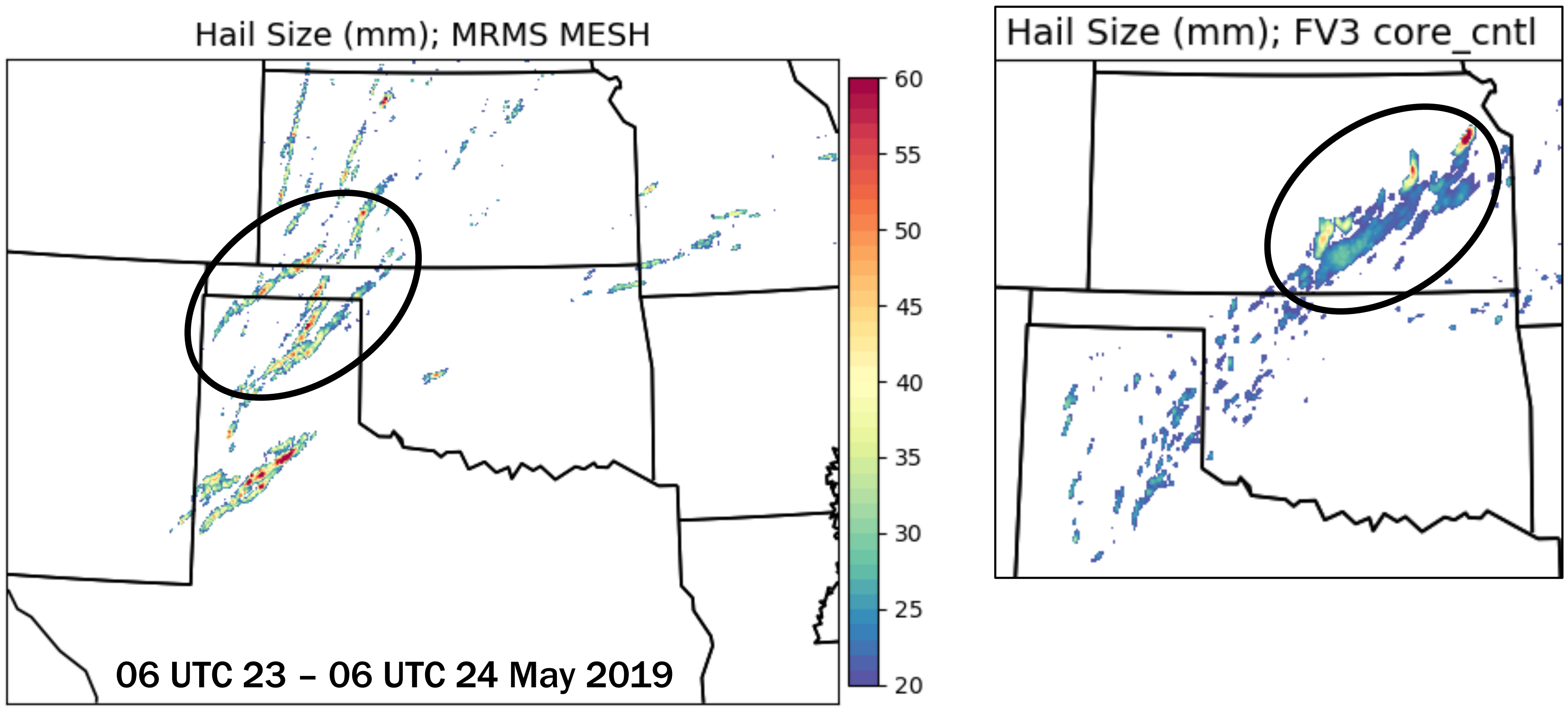
Motivation

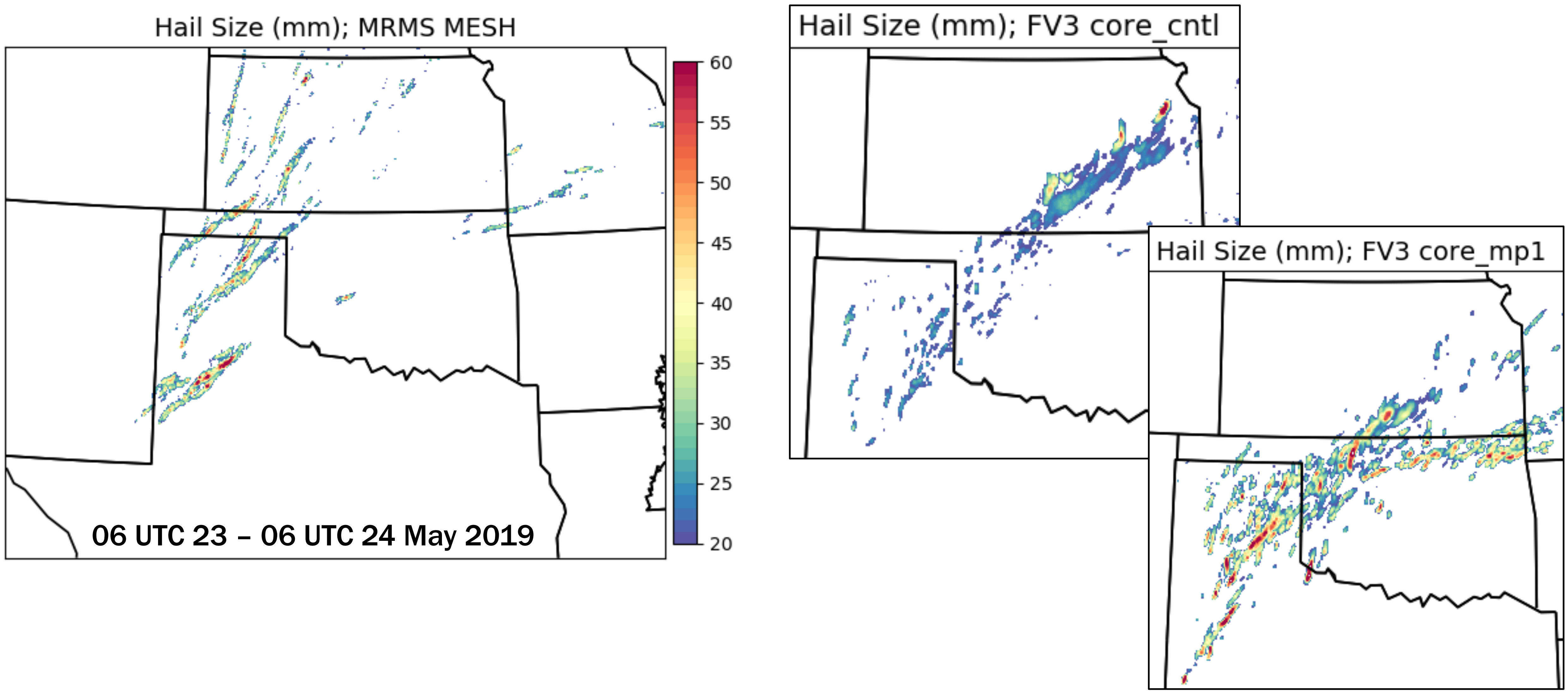
- What makes a “good” hail size forecast?

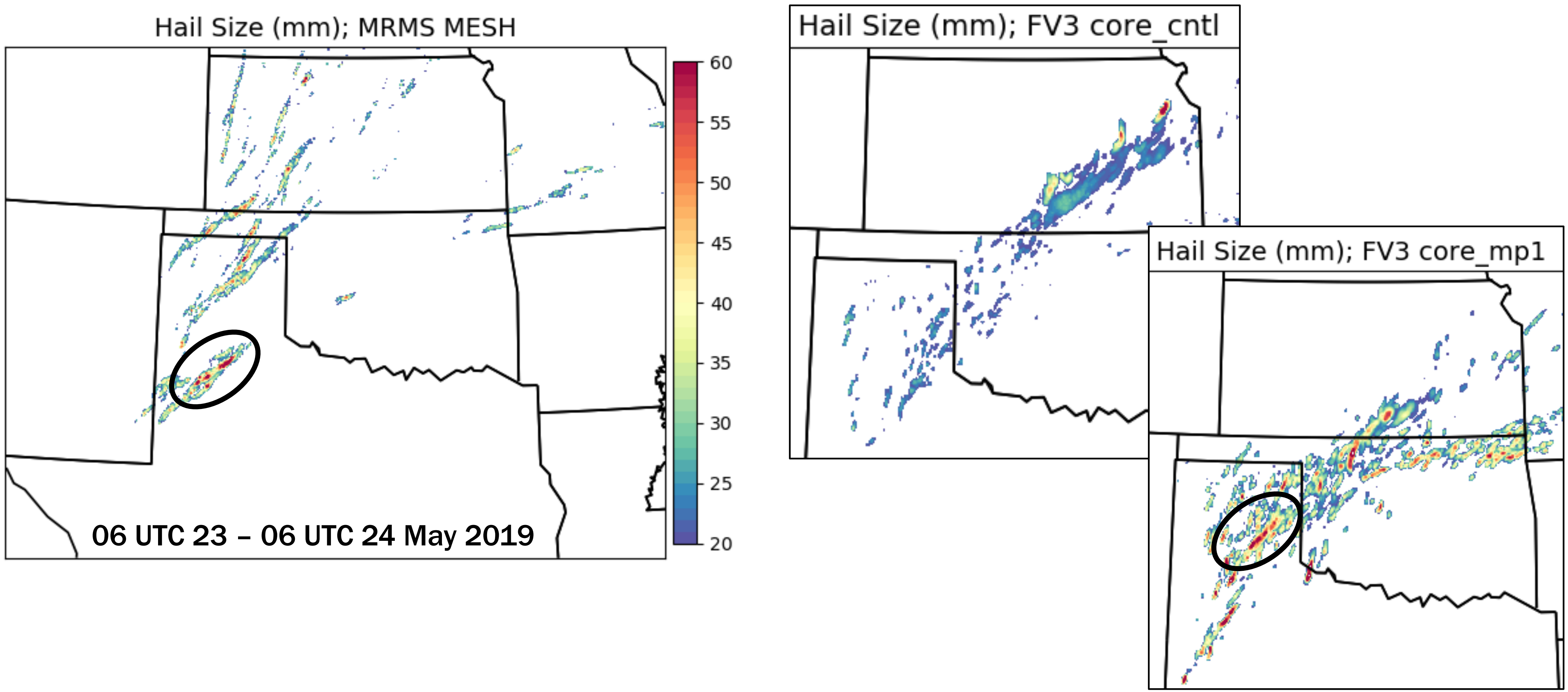

Hail Size (mm); MRMS MESH
06 UTC 23 – 06 UTC 24 May 2019
60
55
50
45
40
35
30
25
20


Hail Size (mm); MRMS MESH
60
55
50
45
40
35
30
25
20
06 UTC 23 – 06 UTC 24 May 2019

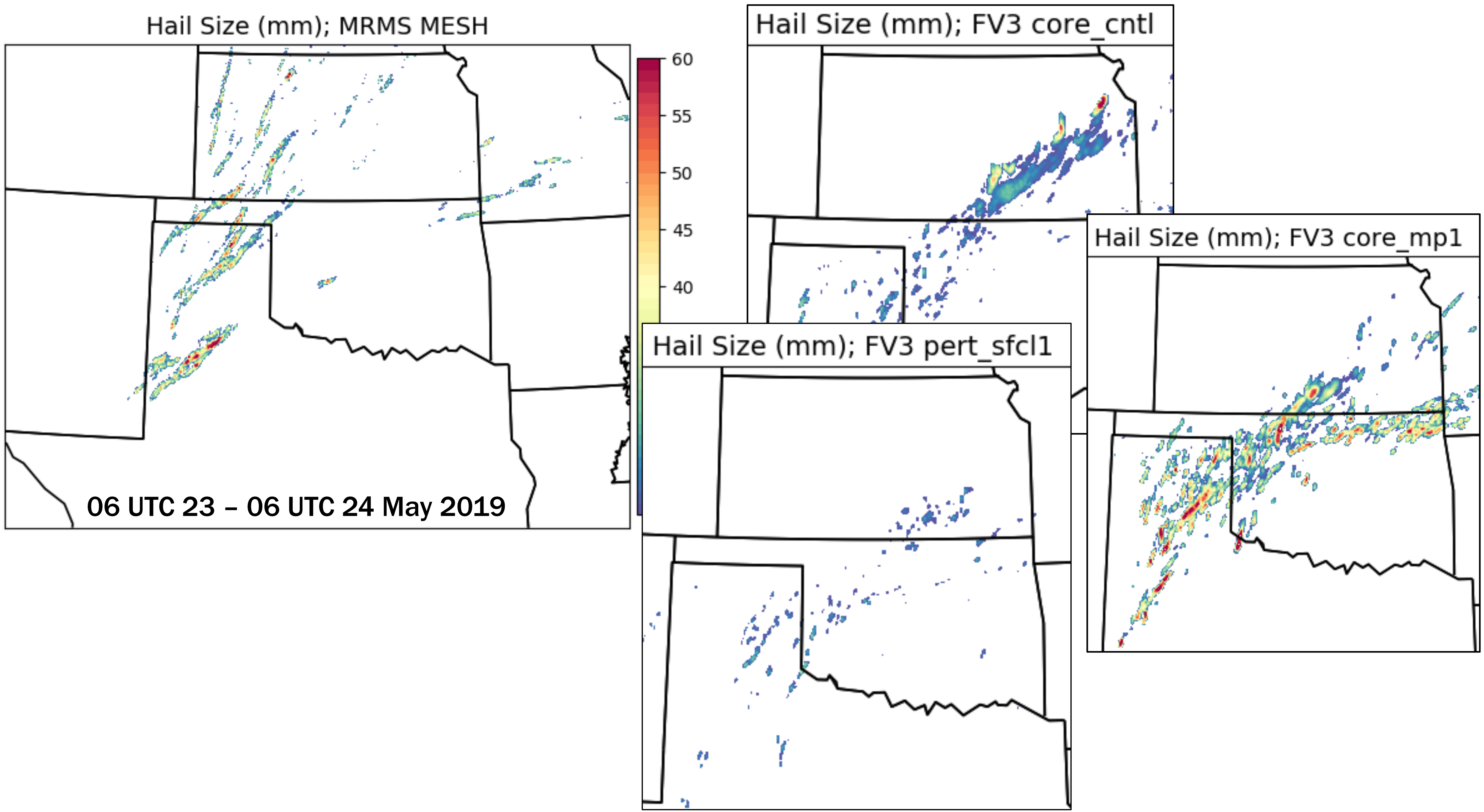

Hail Size (mm); FV3 core_cntl

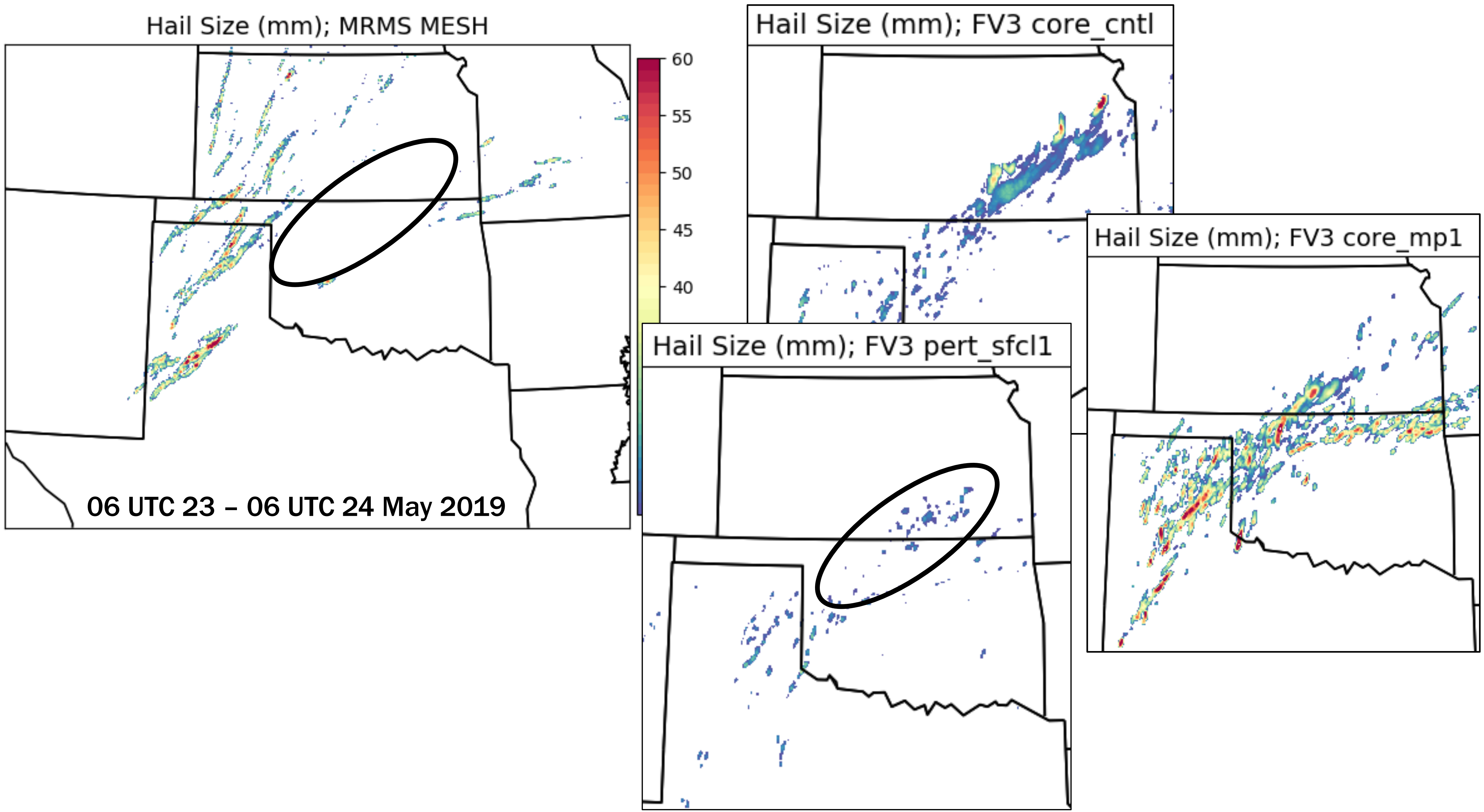

Hail Size (mm); MRMS MESH
60
55
50
45
40
35
30
25
20
06 UTC 23 – 06 UTC 24 May 2019
Hail Size (mm); FV3 core_cntl


Hail Size (mm); MRMS MESH
06 UTC 23 – 06 UTC 24 May 2019
60
55
50
45
40
35
30
25
20
Hail Size (mm); FV3 core_cntl
Hail Size (mm); FV3 core_mp1



Hail Size (mm); MRMS MESH
60
55
50
45
40
35
30
25
20
06 UTC 23 – 06 UTC 24 May 2019
Hail Size (mm); FV3 core_cntl
Hail Size (mm); FV3 core_mp1


Hail Size (mm); MRMS MESH
60
55
50
45
40
06 UTC 23 – 06 UTC 24 May 2019
Hail Size (mm); FV3 core_cntl
Hail Size (mm); FV3 core_mp1
Hail Size (mm); FV3 pert_sfcl1


Hail Size (mm); MRMS MESH
60
55
50
45
40
06 UTC 23 – 06 UTC 24 May 2019
Hail Size (mm); FV3 core_cntl
Hail Size (mm); FV3 core_mp1
Hail Size (mm); FV3 pert_sfcl1

Motivation

- Different end users have different definitions of "good"
  - forecaster: issuing a watch, or a warning?
  - model developer
- What verification methods can we use to understand if our forecast is "good"?

Models

- CAM-HAILCAST hail size forecasts from:
  - CAPS SAR FV3 ensemble, run during the 2019 NOAA HWT for May 14-17,20-24, 27-31
  - NSSL-WRF, from the 2019 NOAA HWT for May 14-17,20-24, 27-31
  - NSSL Warn-on-Forecast System ensemble, 18 member WRF ensemble from the 2019 NOAA HWT
- Validation: Multi-Radar Multi-Sensor Maximum Estimated Size of Hail (MRMS MESH)

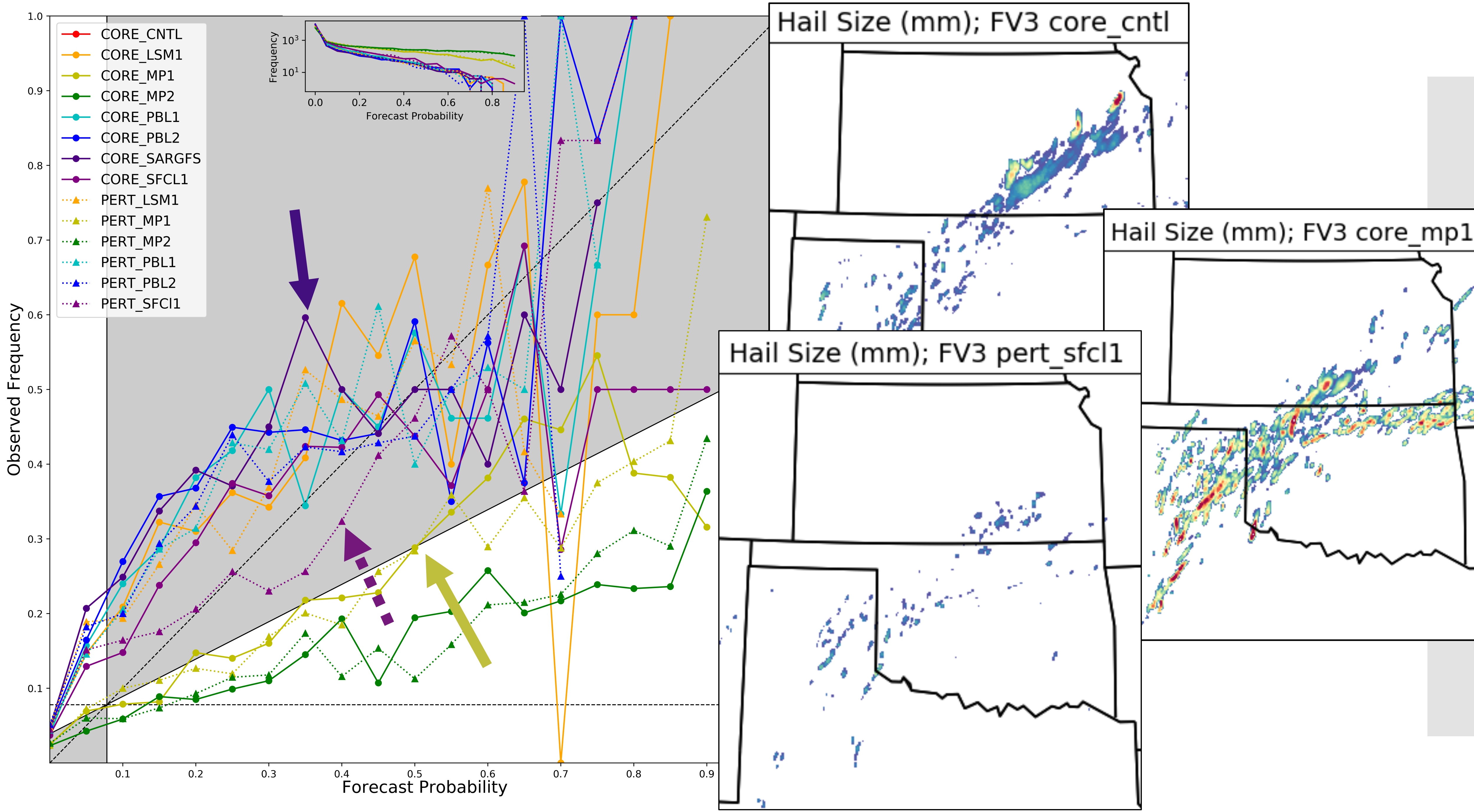
Models

- HAILCAST hail size forecasts from:
  - CAPS SAR FV3 ensemble, run during the 2019 NOAA HWT for May 14-17,20-24, 27-31
  - NSSL-WRF, from the 2019 NOAA HWT for May 14-17,20-24, 27-31
  - NSSL Warn-on-Forecast System ensemble, 18 member WRF ensemble from the 2019 NOAA HWT
- Validation: Multi-Radar Multi-Sensor Maximum Estimated Size of Hail (MRMS MESH)

Interested in knowing more about HAILCAST's implementation in the FV3? Visit poster #191: Henderson, J., C. Calkins, T. Supinie, L. M. Harris, Y. Wang, and R. Adams-Selin: *Implementation of CAM-HAILCAST in the Stand-Alone Regional FV3*

Methods

- Neighborhood grid-based verification
  - Neighborhood Maximum Ensemble Probability of Schwartz and Sobash (2017, *Mon. Wea. Rev.*)
  - Convert hail size forecast to a yes/no forecast on a coarse grid, smooth, and calculate its reliability
  - Coarse grid tries to ensure forecasts close in space are still viewed positively by the verification method
- Can be useful for end users interested in:
  - mesoscale (not storm-scale) forecasts
  - the performance of the model as a whole (convection and convective hazard forecasts together)

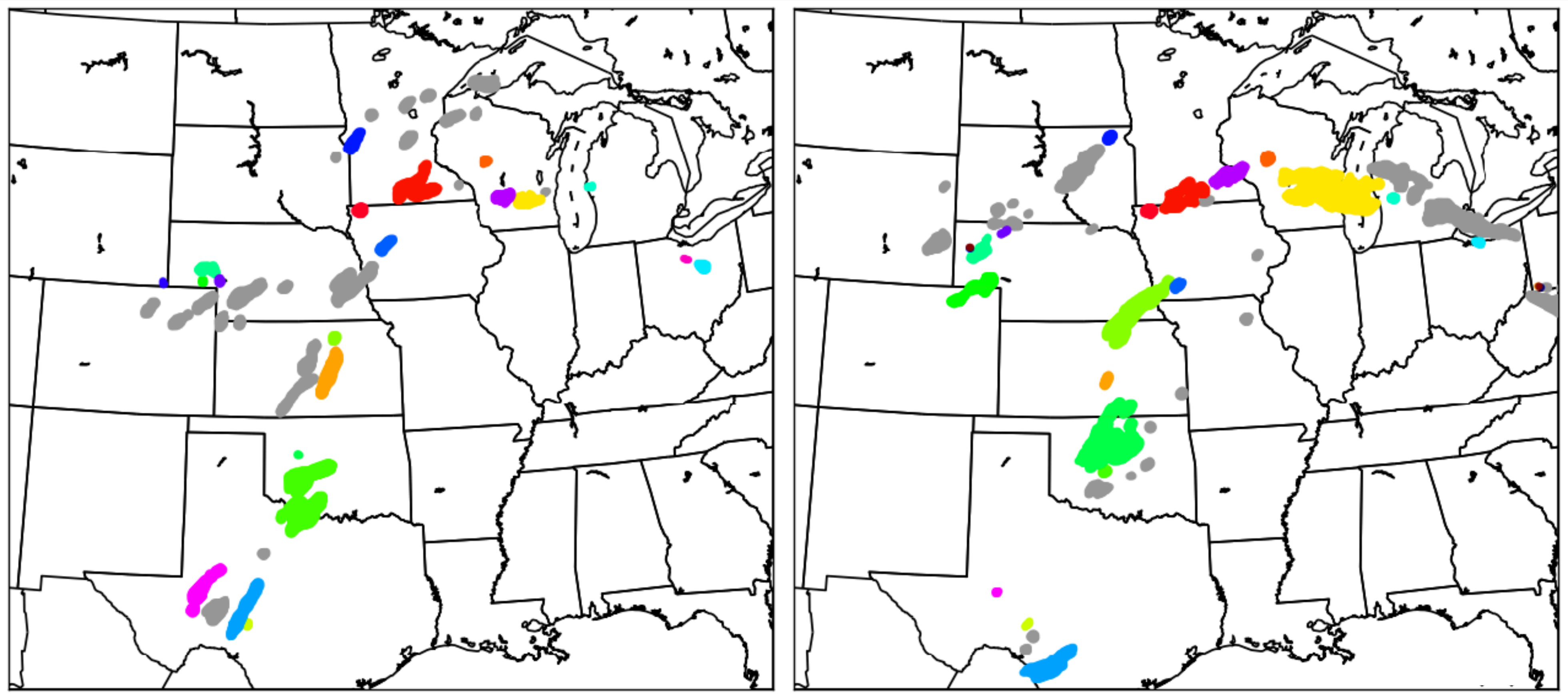


CORE_CNTL
CORE_LSM1
CORE_MP1
CORE_MP2
CORE_PBL1
CORE_PBL2
CORE_SARGFS
CORE_SFCL1
PERT_LSM1
PERT_MP1
PERT_MP2
PERT_PBL1
PERT_PBL2
PERT_SFCl1
Frequency
Forecast Probability
Observed Frequency
Forecast Probability
Hail Size (mm); FV3 core_cntl
Hail Size (mm); FV3 core_mp1
Hail Size (mm); FV3 pert_sfcl1

Methods

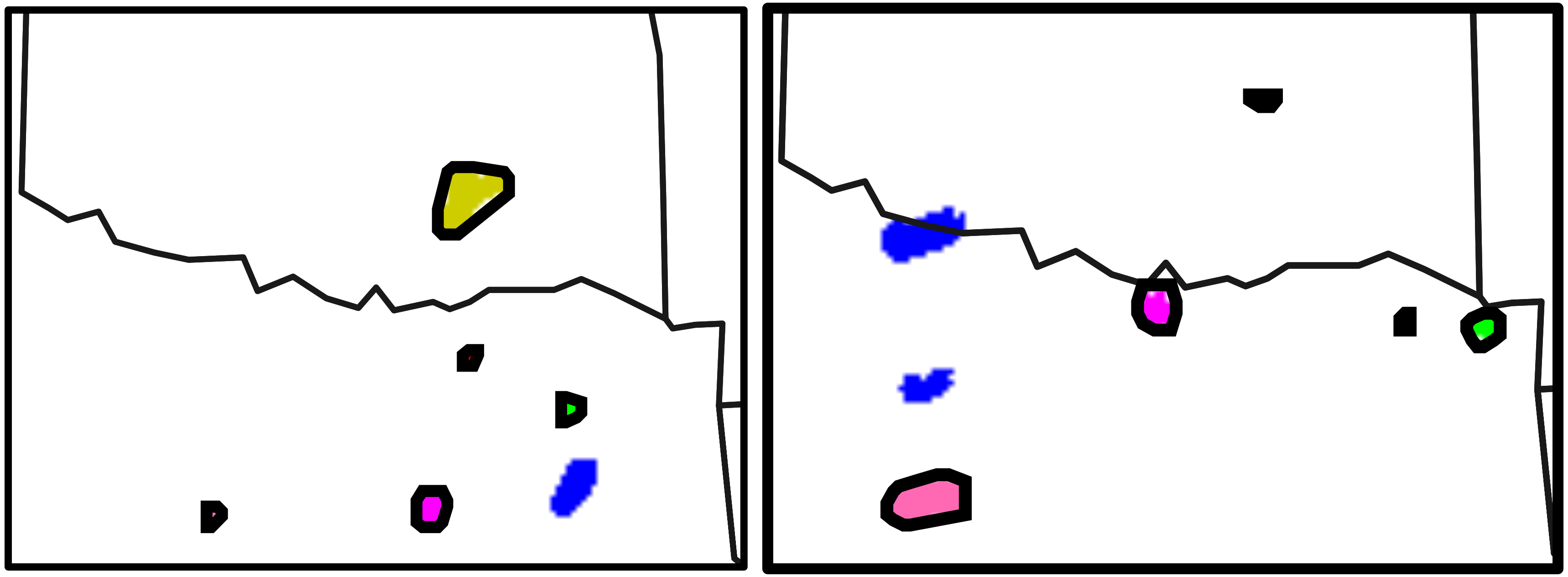
- Object-based forecasts Model Evaluation Toolkit (MET) Method for Object-Based Diagnostic Evaluation (MODE; Davis et al. 2006a,b, *Mon. Wea. Rev.*)
  - Avoid the "double penalty" by matching hail swath to hail swath
  - Mesoscale configuration: 24-h hail swaths, scale of supercell family or single MCS

06 UTC 07 May – 06 UTC 08 May 2014



Methods

- Object-based forecasts Model Evaluation Toolkit (MET) Method for Object-Based Diagnostic Evaluation (MODE; Davis et al. 2006a,b, *Mon. Wea. Rev.*)
  - Avoid the “double penalty” by matching hail swath to hail swath
  - Mesoscale configuration: 24-h hail swaths, scale of supercell family or single MCS
  - Storm-scale configuration: 1-h hail swaths, individual storm scale



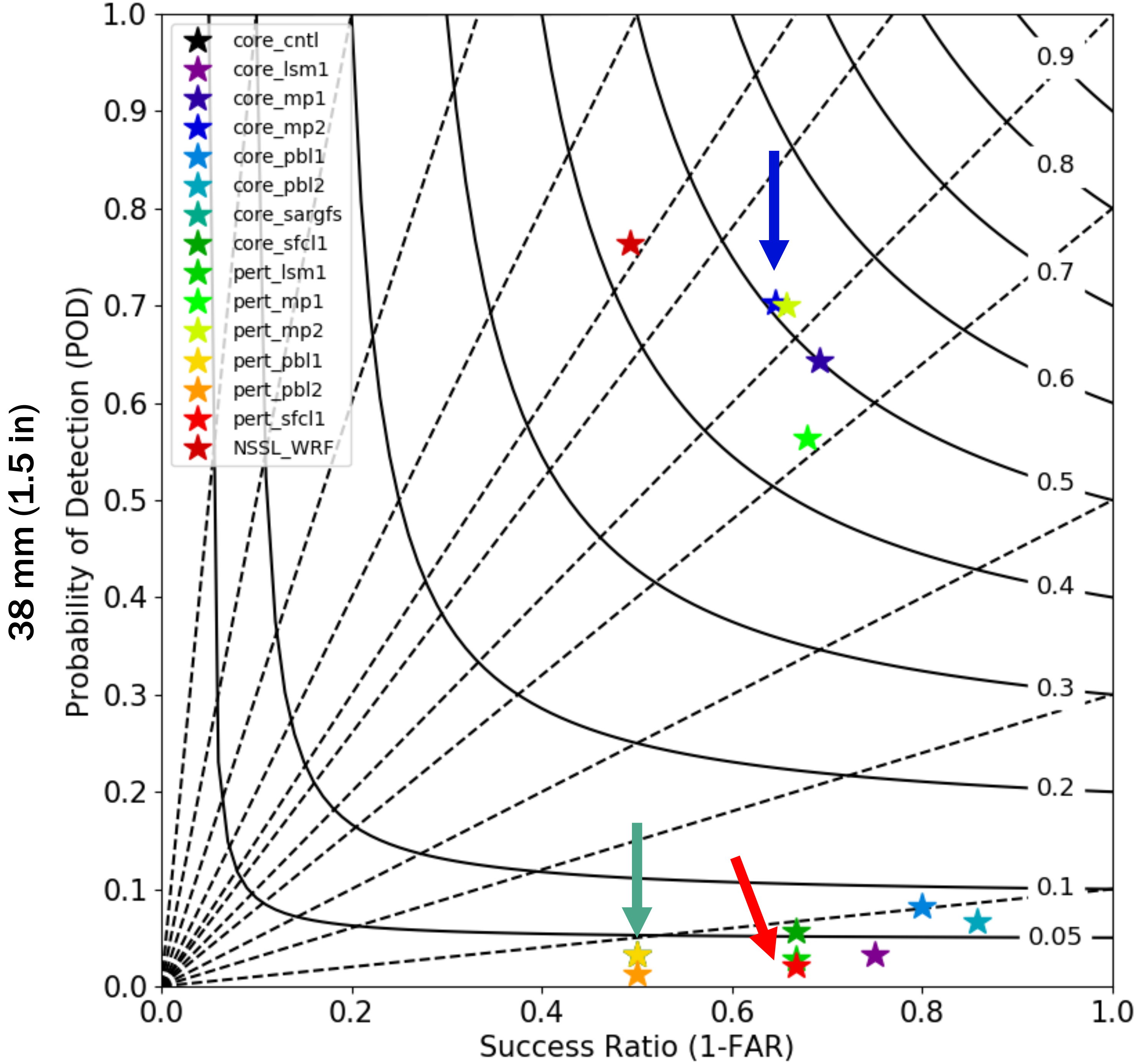
HAILCAST MRMS MESH

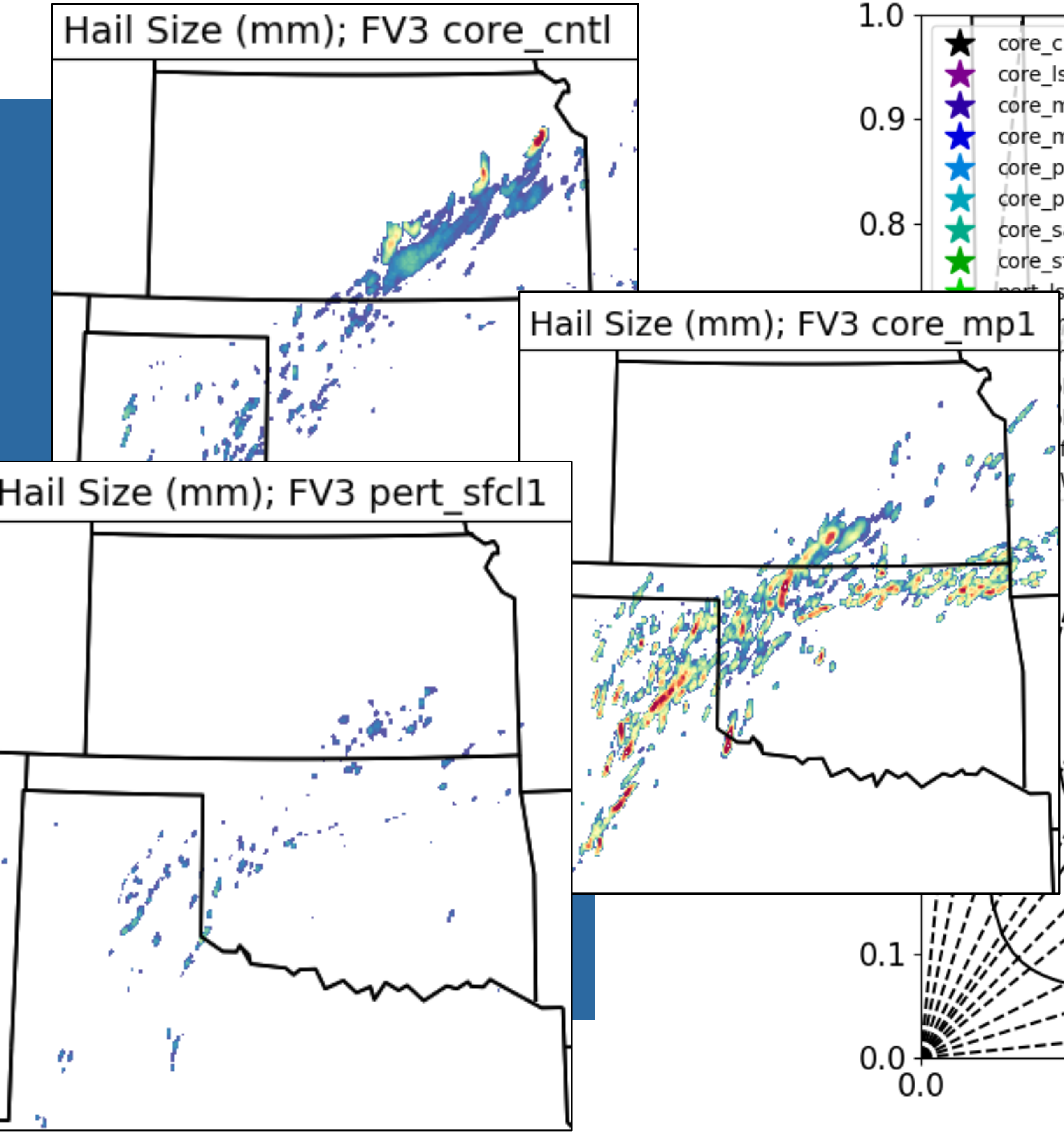
Methods

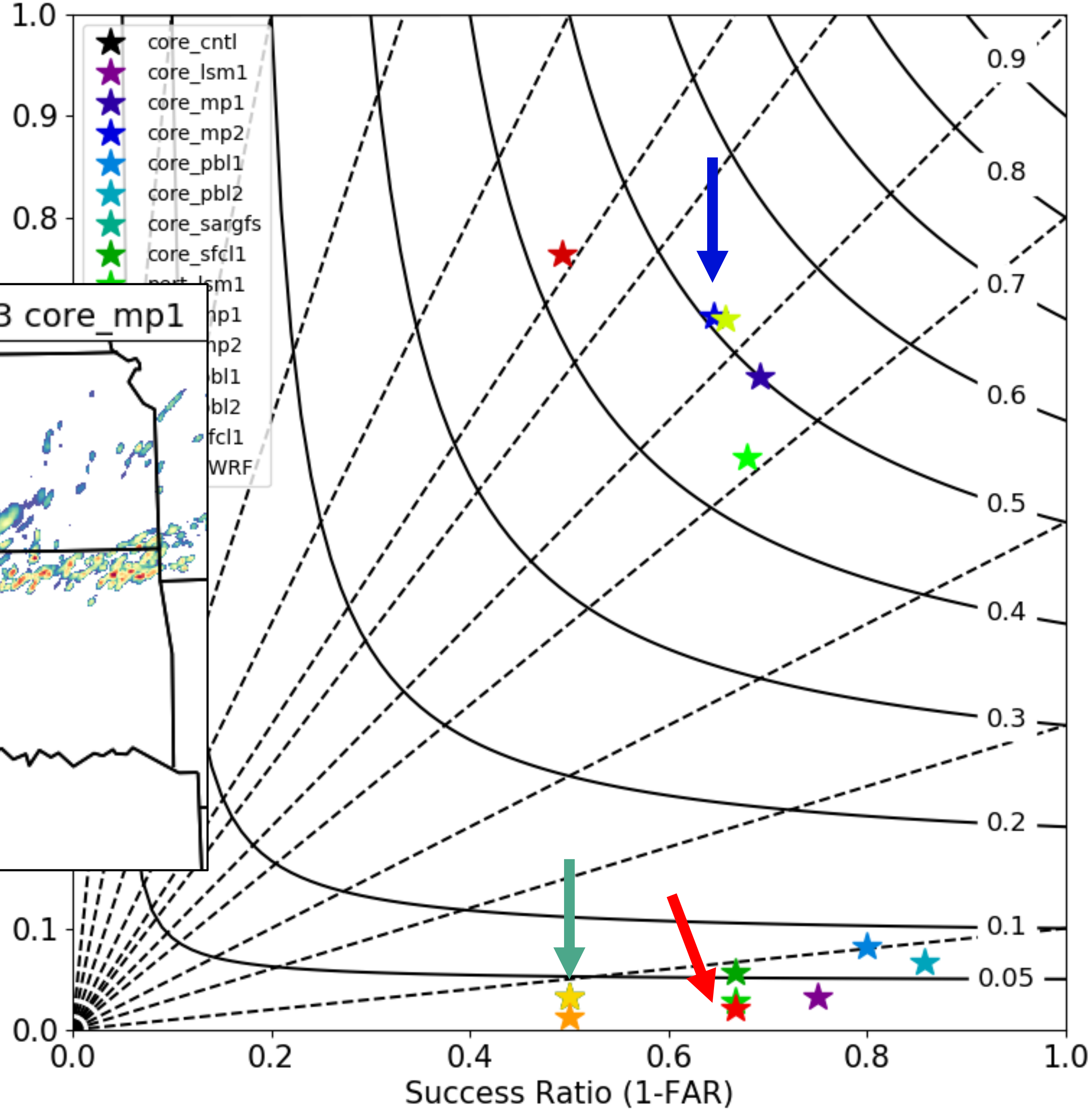
- Object-based forecasts Model Evaluation Toolkit (MET) Method for Object-Based Diagnostic Evaluation (MODE; Davis et al. 2006a,b, *Mon. Wea. Rev.*)
  - Avoid the "double penalty" by matching hail swath to hail swath
  - Mesoscale configuration: 24-h hail swaths, scale of supercell family or single MCS
  - Storm-scale configuration: 1-h hail swaths, individual storm scale
- **Matched object requirement: only include forecast hail swaths with a matching observed object**
  - Avoids "unfair penalty": what if the model fails to produce any convection?
- Useful for end users interested in:
  - performance of their convective hazard model!

Mesoscale configuration

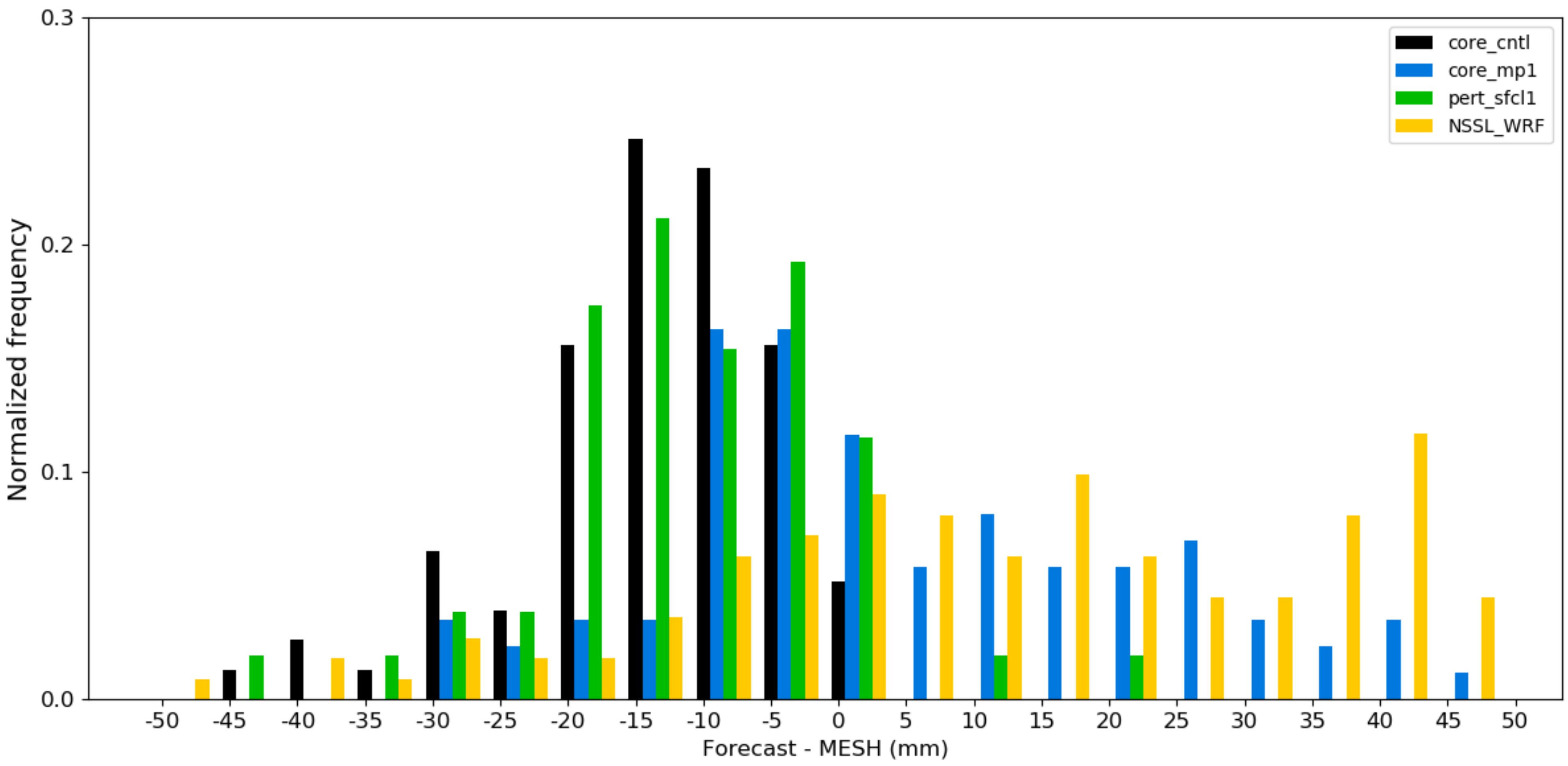
CAPS-FV3 SAR; NSSL-WRF




Hail Size (mm); FV3 core_cntl
Hail Size (mm); FV3 core_mp1
Hail Size (mm); FV3 pert_sfcl1


core_cntl
core_lsm1
core_mp1
core_mp2
core_pbl1
core_pbl2
core_sargfs
core_sfcl1
pert_lsm1
0.0
0.1
0.8
0.9
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Success Ratio (1-FAR)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.05

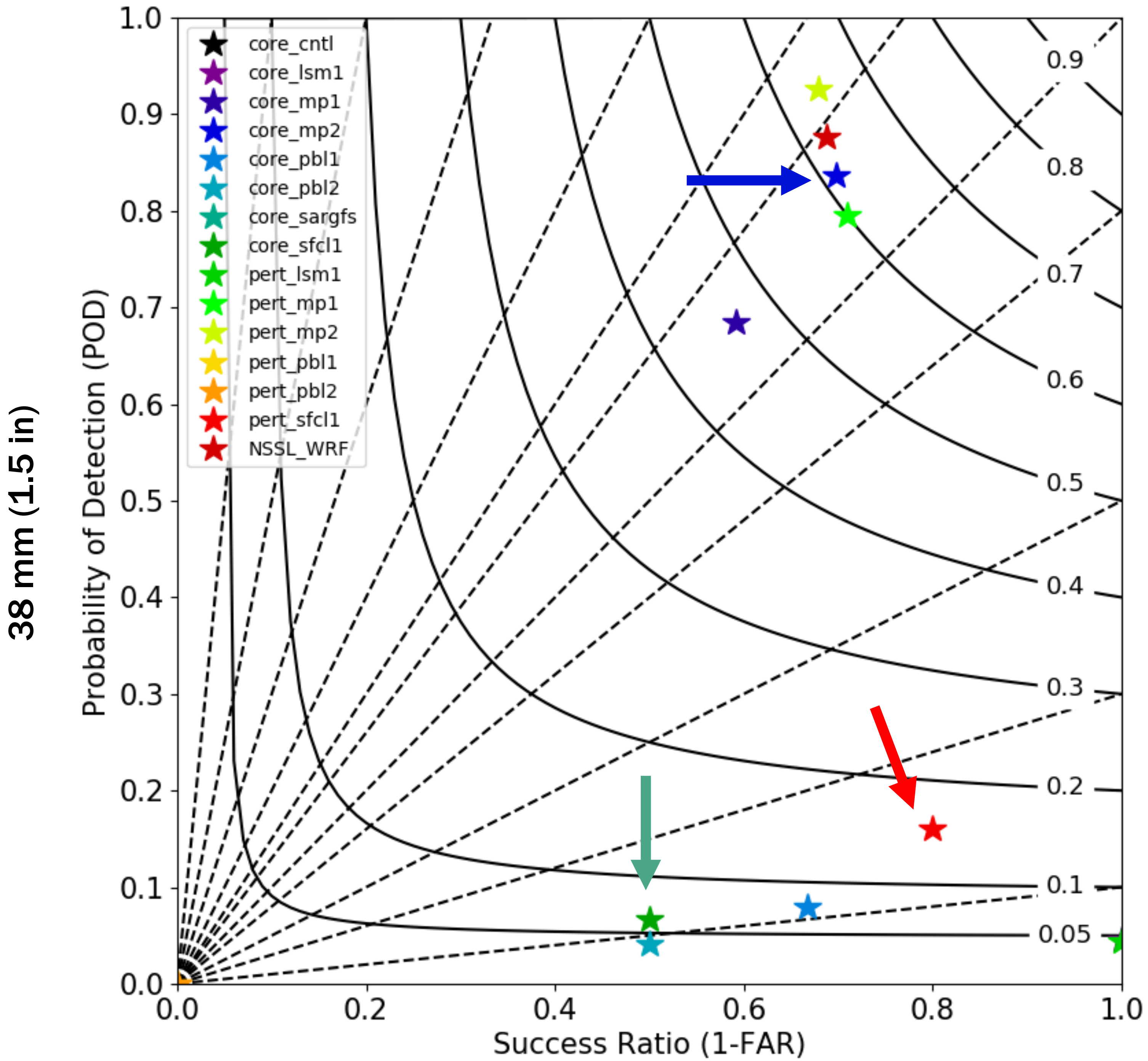
Mesoscale configuration



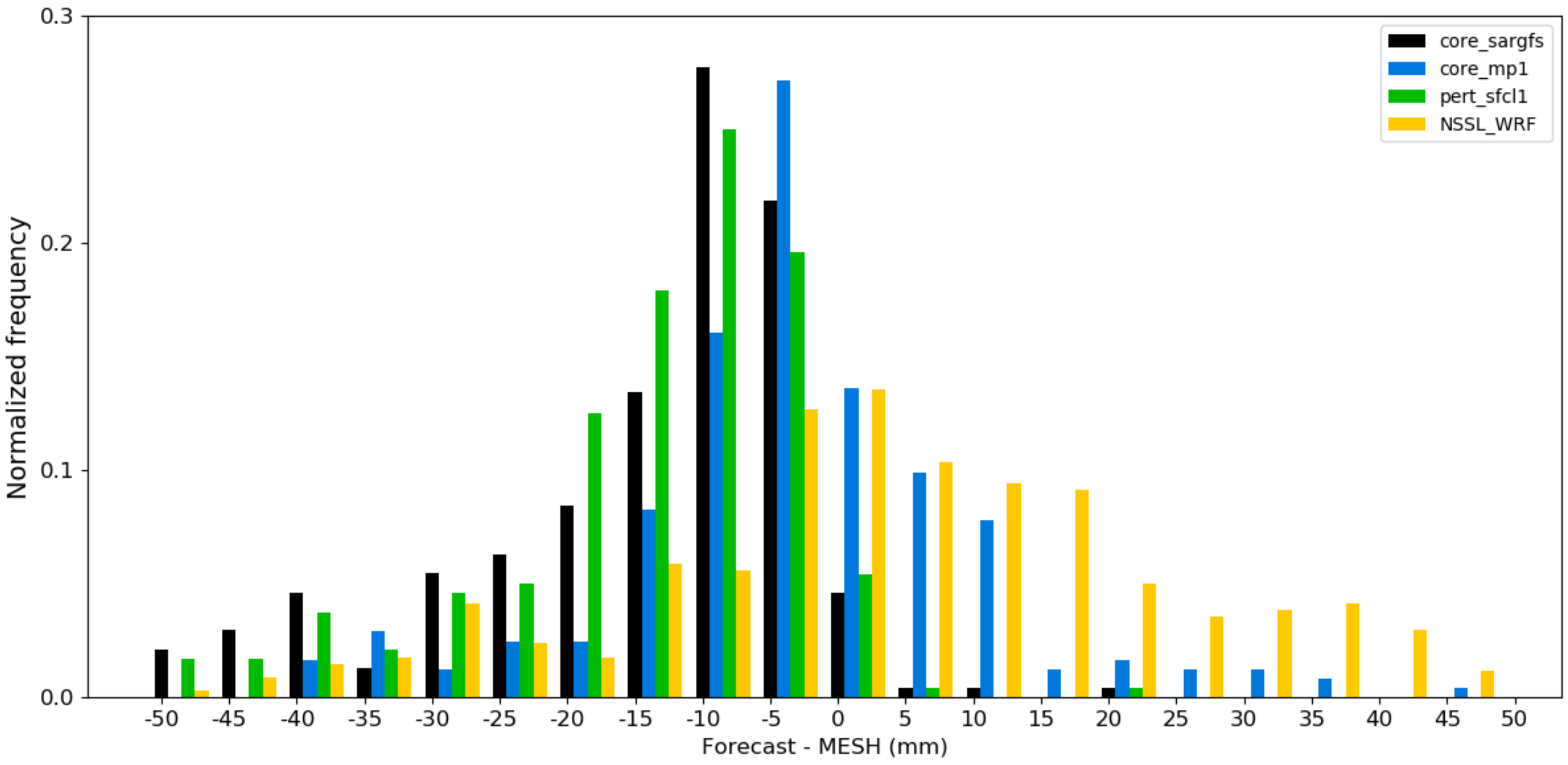
CAPS-FV3 SAR; NSSL-WRF

Storm-scale configuration

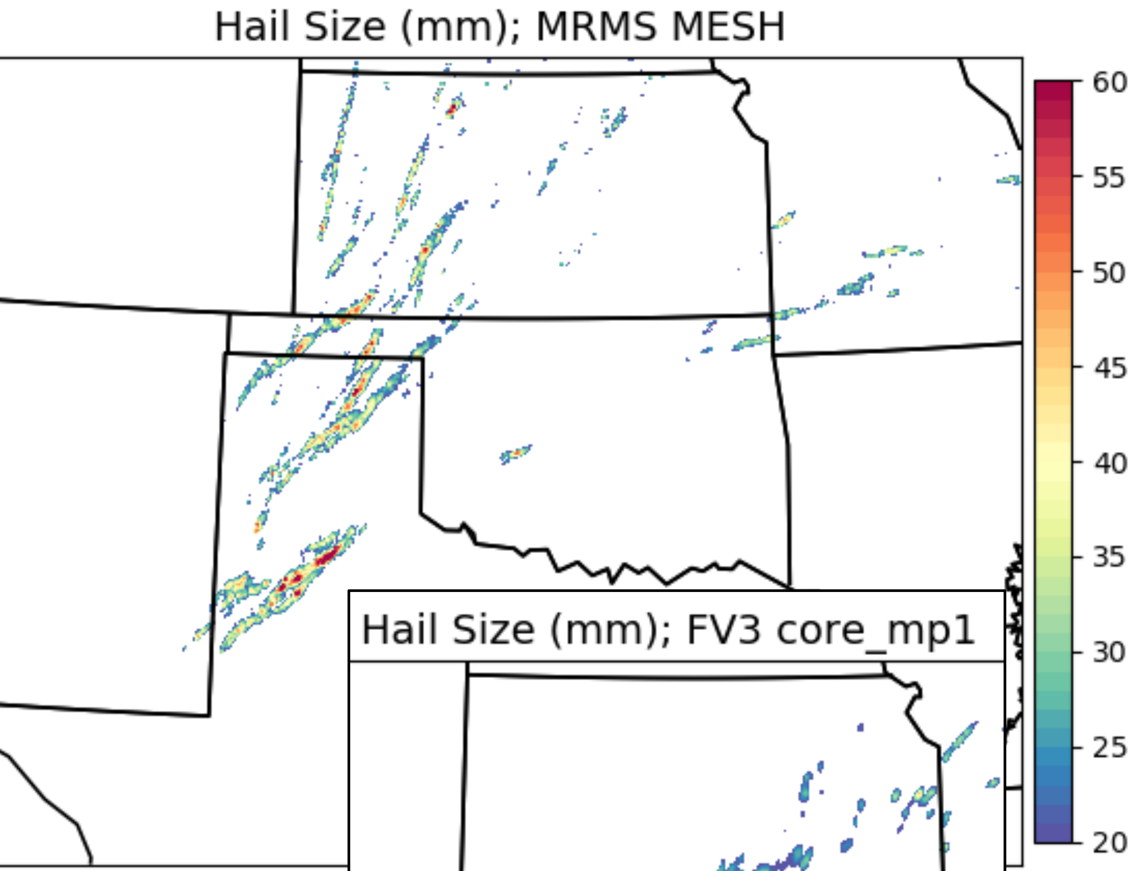
CAPS-FV3 SAR; NSSL-WRF

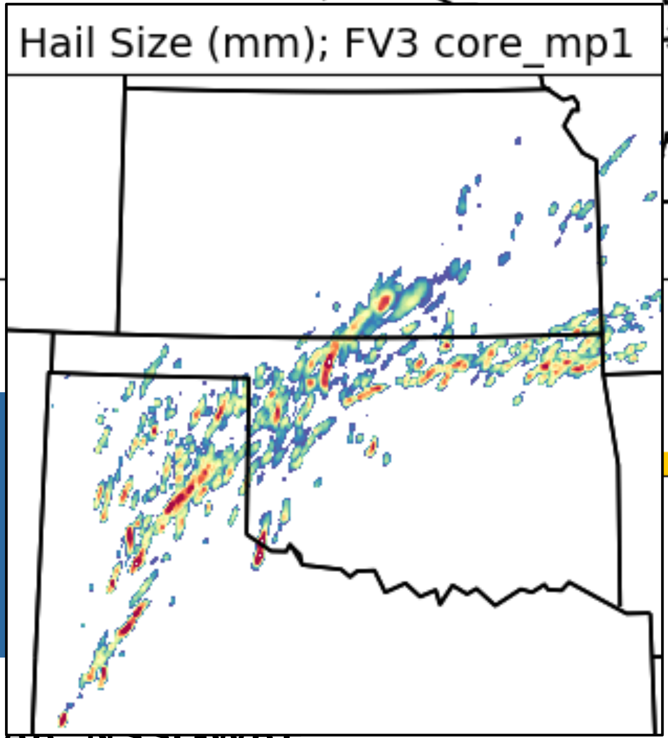


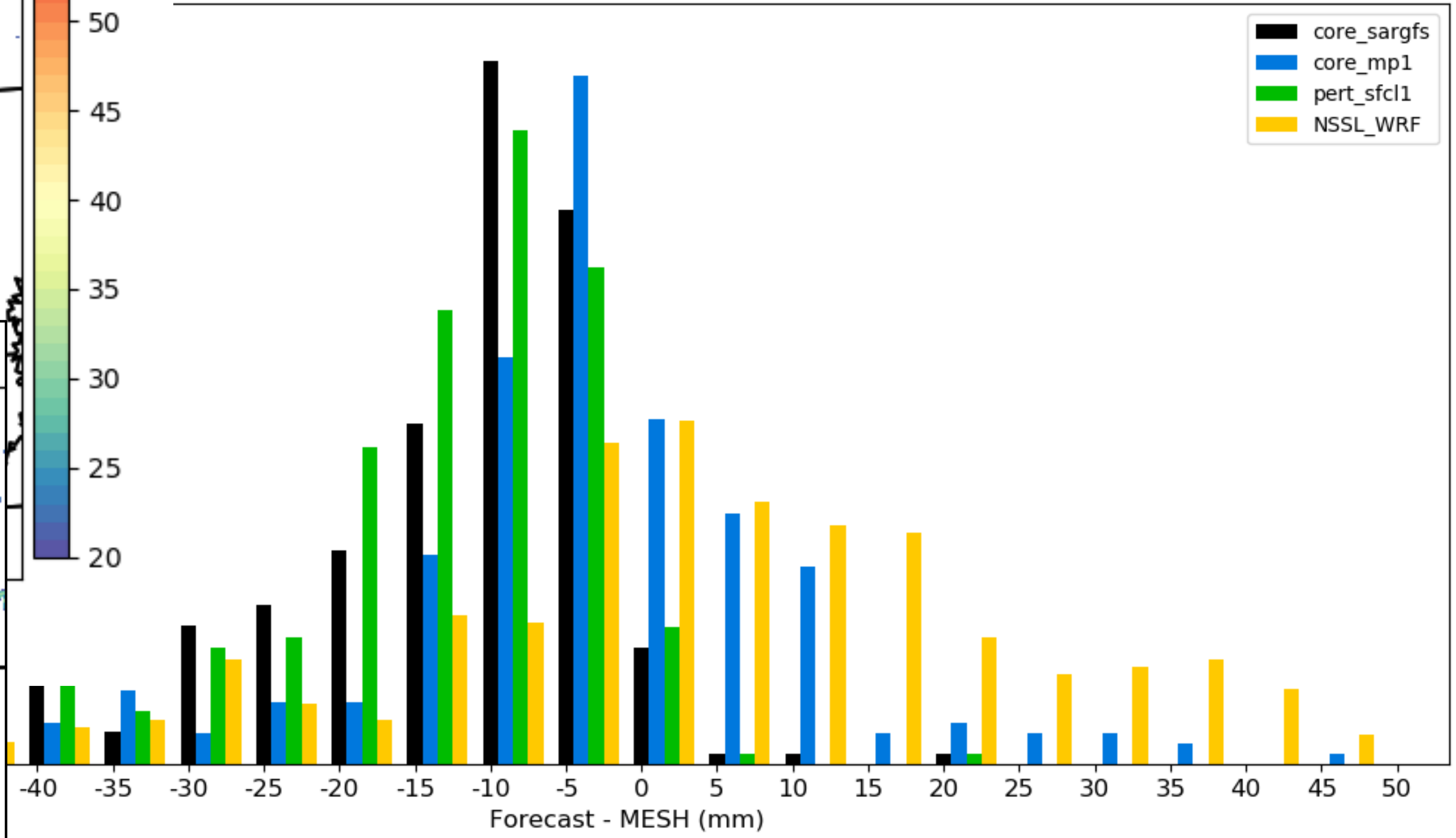
Storm-scale configuration



CAPS-FV3 SAR; NSSL-WRF

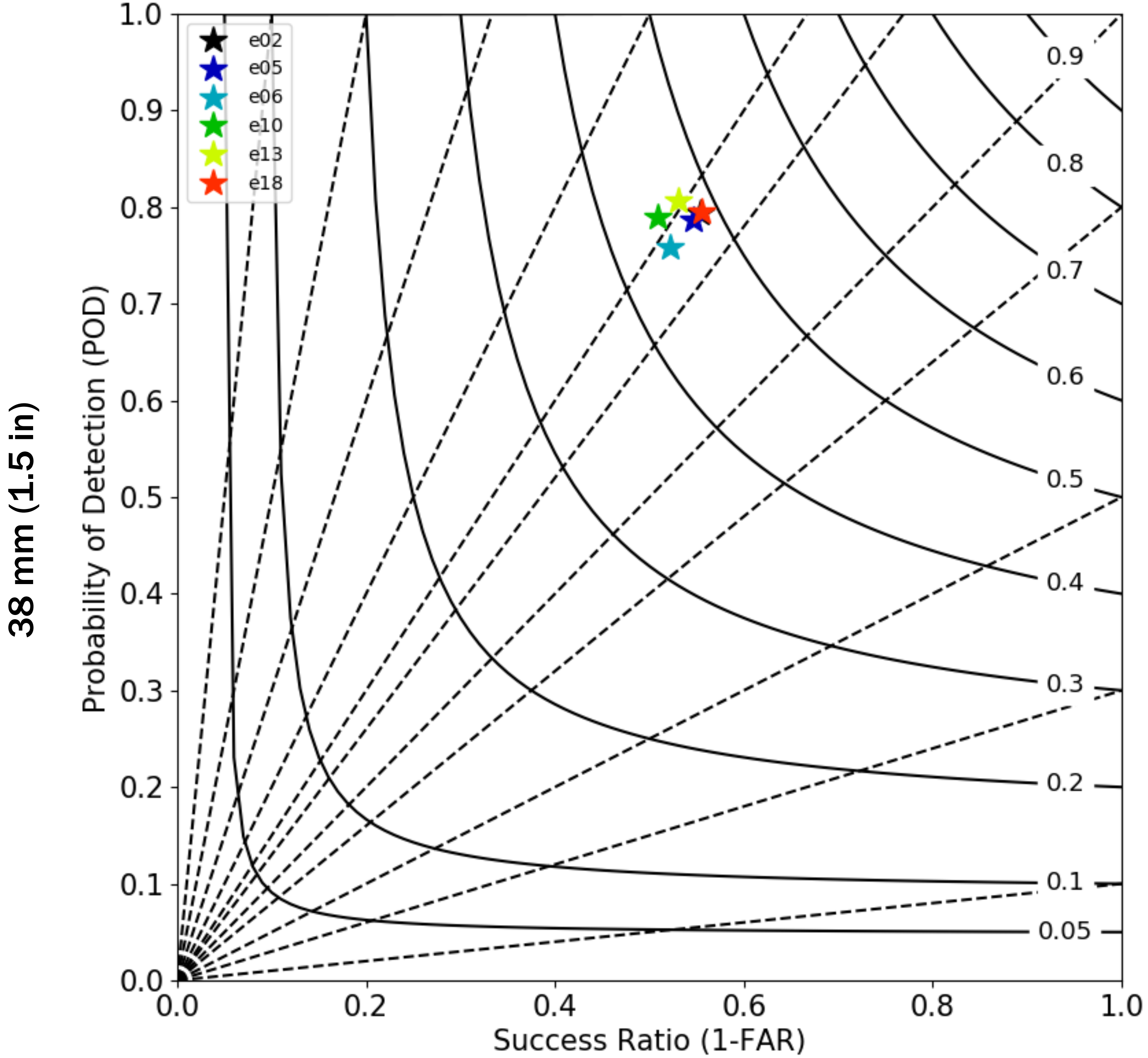






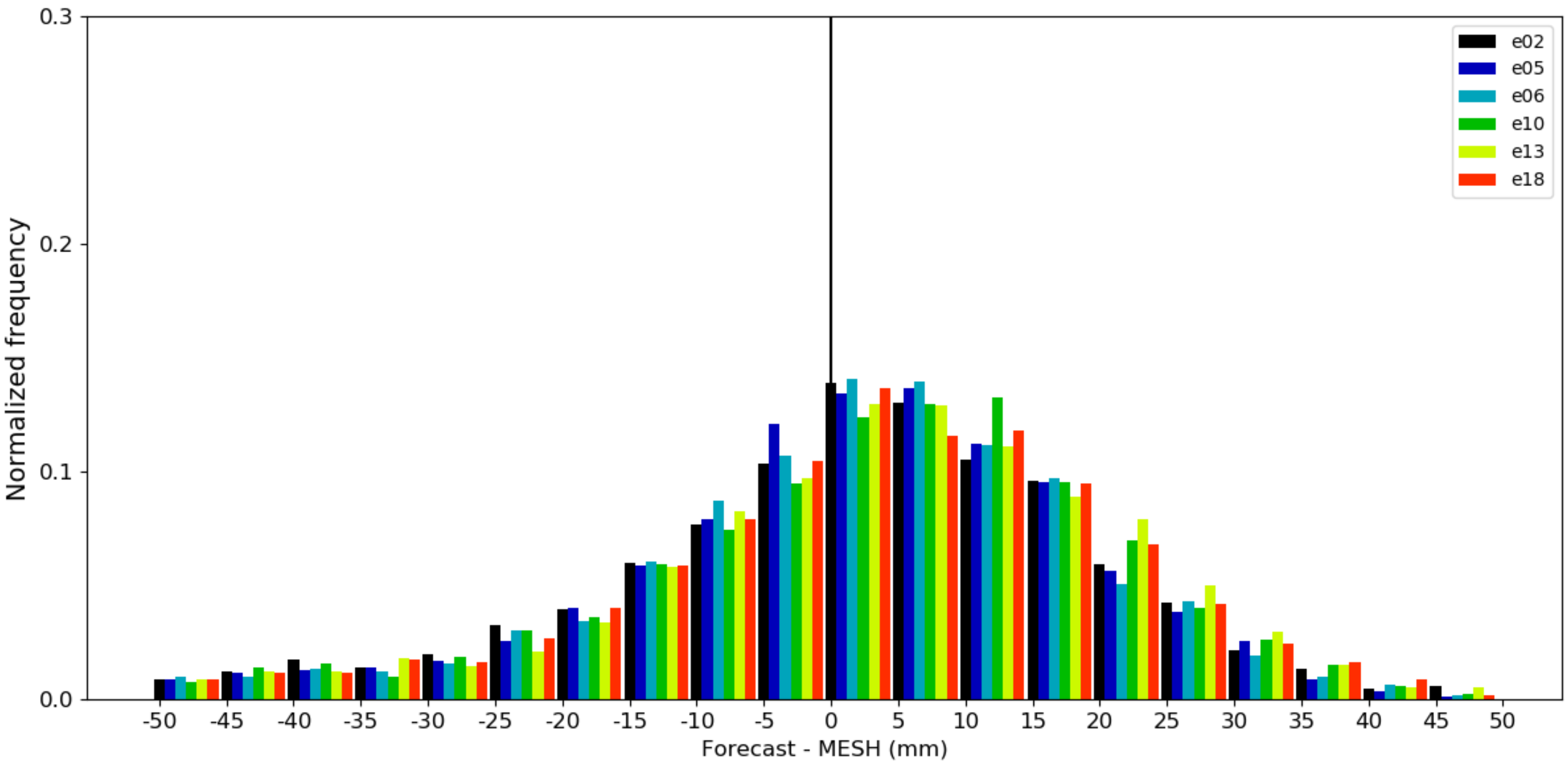
CAPS-FV3 SAR, NSSL-WRF

Storm-scale configuration



Warn-on-Forecast System ensemble (WRF)

Storm-scale configuration



Warn-on-Forecast System ensemble (WRF)

Skinner configuration (Skinner et al. 2018)

Object Identification:

- Find regions of HAILCAST or MESH exceeding a prescribed threshold of 1 inch
- Regions must be larger than 60 km^2 to be retained as an object
- Second maximum intensity threshold optionally applied to HAILCAST objects where objects are only retained if the max intensity exceeds a prescribed value

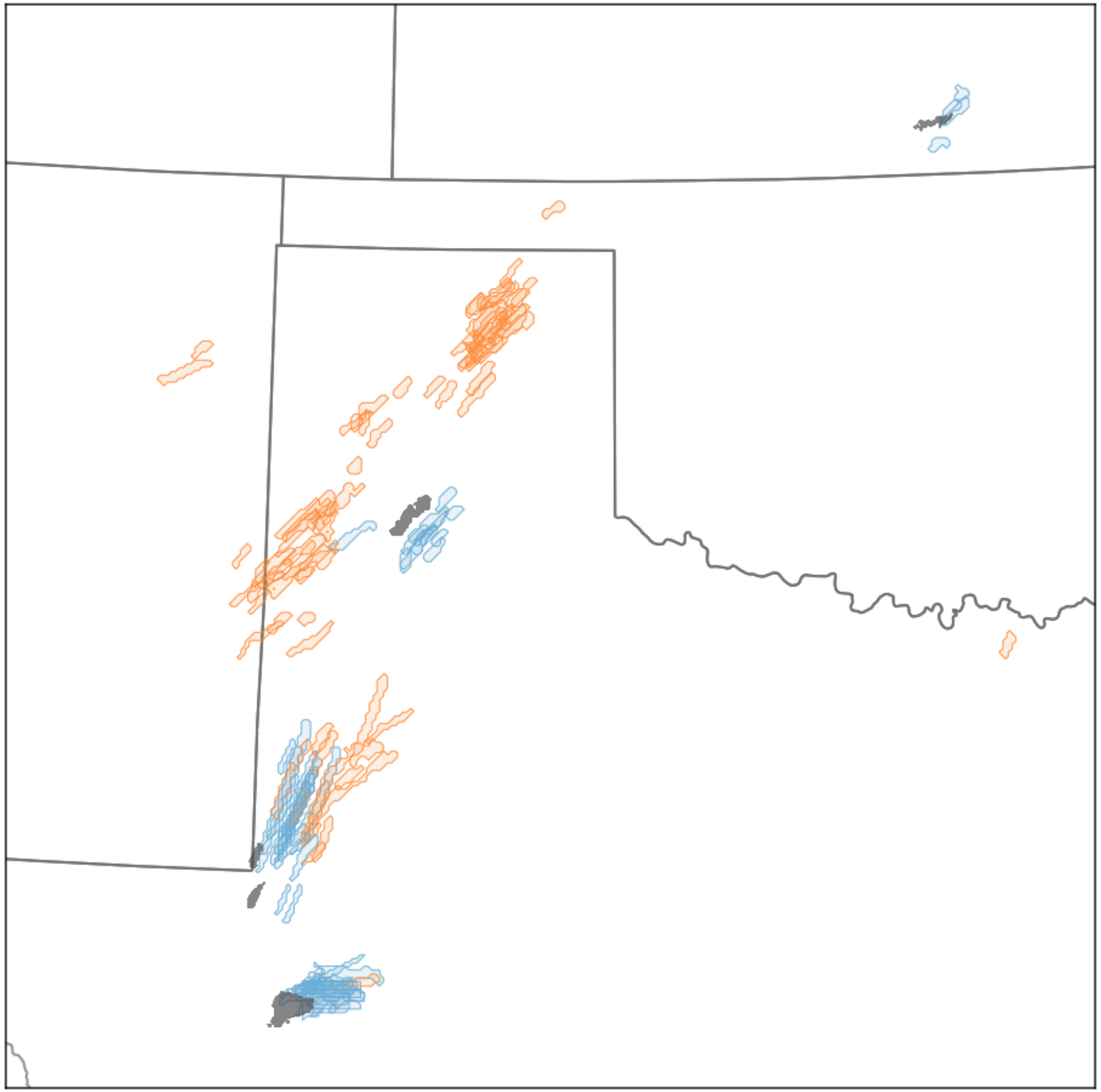
Object Matching:

- Objects matched using a total interest score defined as the average between centroid and boundary (minimum) displacement
- Both displacement measures use a maximum allowable displacement of 40 km
- Total interest must exceed 0.2 for objects to be matched
- Only a single forecast object is allowed to be matched to a single observed object

Unmatched objects included in statistical calculations!

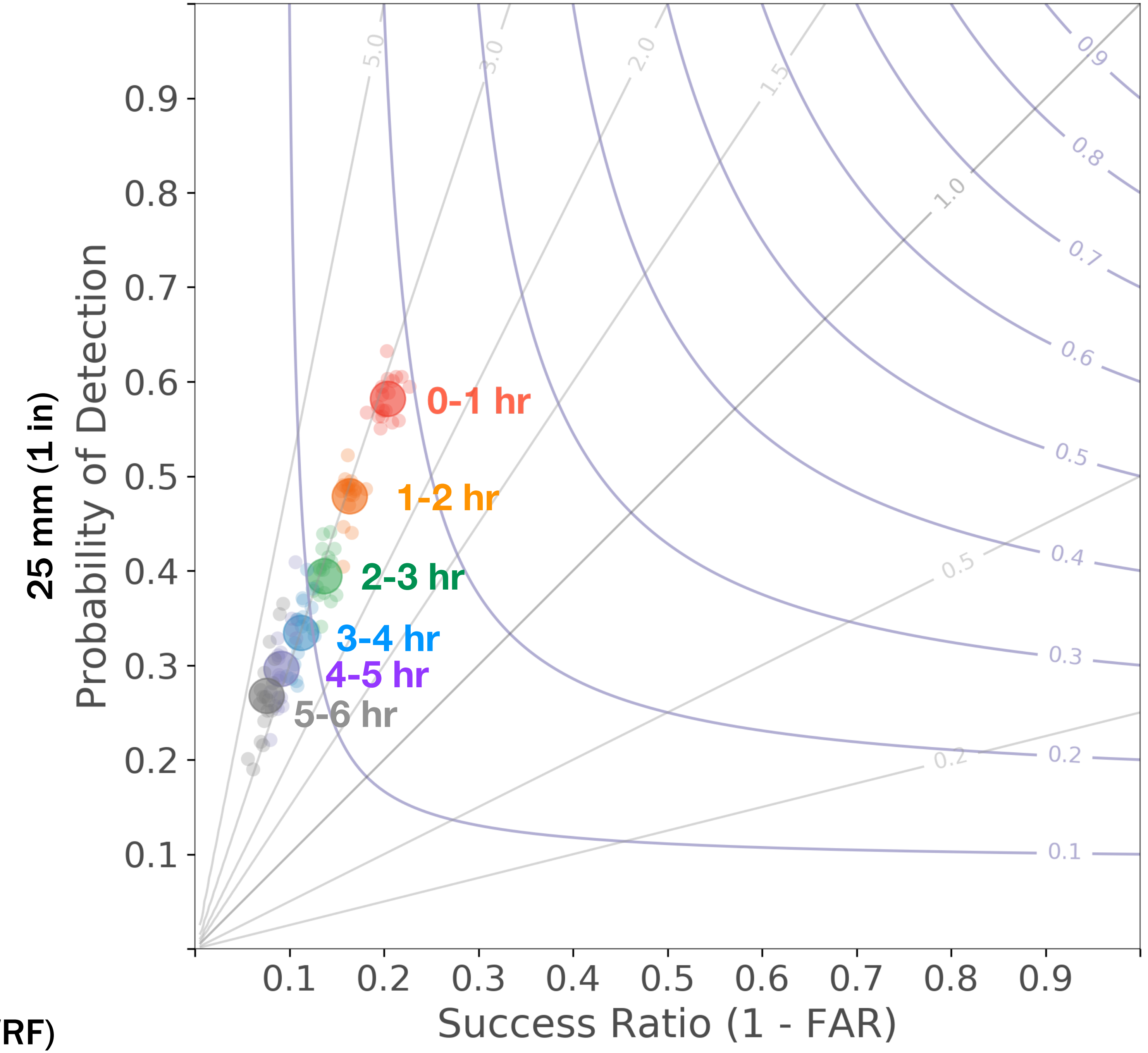
Skinner configuration

Warn-on-Forecast System ensemble (WRF)



POD: 0.56 FAR: 0.55 BIAS: 1.41 CSI: 0.32

Skinner configuration



Warn-on-Forecast System ensemble (WRF)

Conclusions

- Depending on your end user of your convective hazard forecast, you should vary the verification method. What do you want to know?
- Can the model provide skillful forecasts of hail's existence, on a meso-spatial scale? (e.g., an SPC outlook?)
  - Grid-based neighborhood verification with upscaling for a given hail size threshold
  - CAPS-FV3 SAR HAILCAST: wide variations in skill among FV3 members. Member pert_sfcl1 appears to offer the most reliable 25-mm hail forecast.

Conclusions

- Depending on your end user of your convective hazard forecast, you should vary the verification method. What do you want to know?
- If the model correctly forecasts convection, are HAILCAST's forecasts of hail size skillful? (E.g., info for a model developer)
  - Object-based verification with matching object requirement
  - CAPS-FV3 SAR HAILCAST:
    - Pert_sfcl1 strongly underforecasts peak hail magnitudes within objects
    - Core_mp1 offers higher CSI values for forecasting 38-mm hail objects, but with larger error (forecast-MESH) among matched objects
    - Wide variation in skill among members
  - Warn-on-Forecast System Ensemble
    - Much tighter distribution of skill among members
    - Slight overforecasting by HAILCAST when WRF correctly forecasts convection

Conclusions

- Depending on your end user of your convective hazard forecast, you should vary the verification method. What do you want to know?
- Does the model produce skillful forecasts of hail's existence and location on a storm scale? (E.g., can it be used for warnings?)
  - Object-based verification **without** matching object requirement (Skinner configuration)
  - Warn-on-Forecast System Ensemble
    - Tight distribution of skill among members
    - Overforecasting of convection and 25-mm hail by WRF-HAILCAST
- How are using this information to improve HAILCAST?
  - Poster #165: Adams-Selin, R: *Comparison of One-Dimensional Pseudo-Lagrangian and Three-Dimensional Fully Lagrangian Trajectories when Forecasting Hail Size*
  - Poster #166: Dahl, Nathan, R. Adams-Selin, R. E. D. Jewell, and I. L. Jirak: *Updating HAILCAST Hail Size Predictions in NSHARP*