NASA Earth Science DISASTERS PROGRAM

Deriving Severe Hail Likelihood from Satellite Observations and Model Reanalysis Parameters using a Deep Neural Network

Benjamin Scarino[1*], Kristopher Bedka[2], Kyle Itterly[1], Cameron Homeyer[3], John Allen[4]

[1]Science Systems and Applications, Inc., Hampton, VA, USA
[2]NASA Langley Research Center, Hampton, VA, USA
[3]School of Meteorology, University of Oklahoma, Norman, OK, USA
[4]Central Michigan University Institute for Great Lakes Research, Mount Pleasant, MI, USA

Motivation: Deep Neural Net for Hailstorm Detection

- **Geostationary satellites such as the GOES series have been observing severe convection at 15-60 minute intervals (year/region dependent) for over 40 years**
  - A valuable data record for estimating severe storm risk throughout the diurnal cycle
  - GOES-8 through -15 provides GEO climate data record for W. Hemisphere between 1994-2018
    - Assess detection skill sensitivity to imager resolution and sampling (GOES-R series)

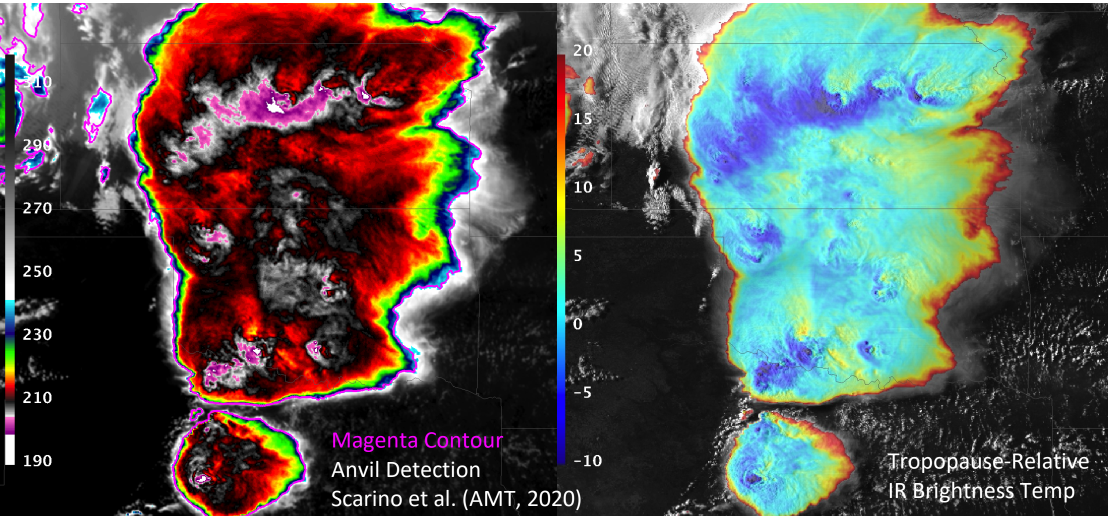
- **Environmental parameters favorable for hailstorm formation are well-known and are captured fairly well by reanalyses such as MERRA-2 and ERA-5**
  - e.g., $(2*\text{MUCAPE})^{0.5}*\text{SHEAR}_{0\text{-}6\text{ km}}$ (WMAXSHEAR, Taszarek et al. 2017, 2020)

- **Wealth of data over United States to train and test a deep neural network (DNN)**
  - NEXRAD radar-estimated Maximum Expected Size of Hail (**MESH**) as truth
    - (Murillo and Homeyer 2019, Murillo et al. 2021)
  - <u>Apply to other regions of the world</u>

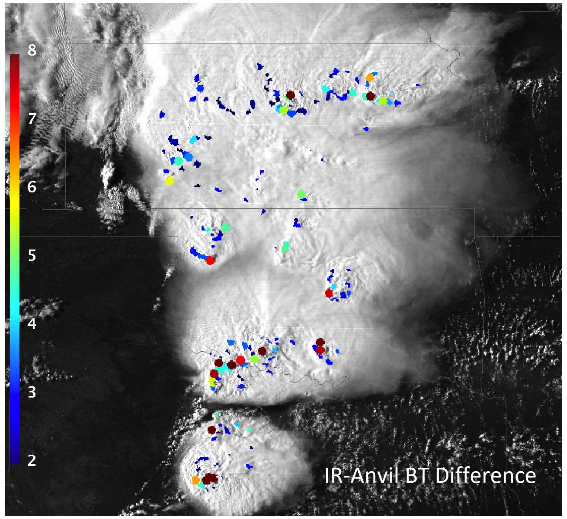
Automated Overshooting Cloud Top Detection and Validation

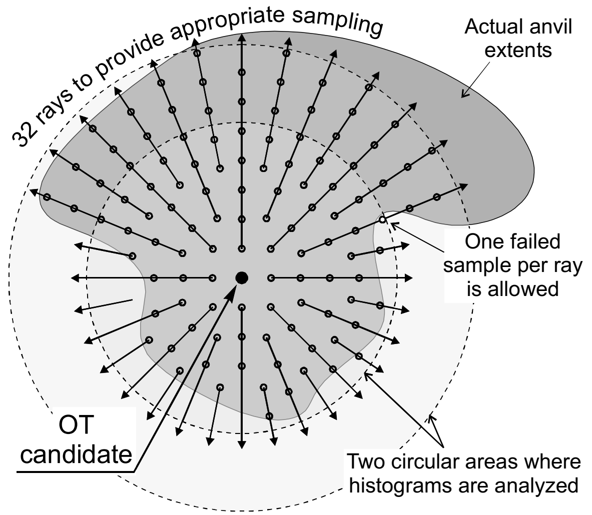
Khlopenkov et al., Cooney et al. (JGR, 2021)

1) Normalize IR temperature using reanalysis tropopause temperature and identify convective anvil clouds

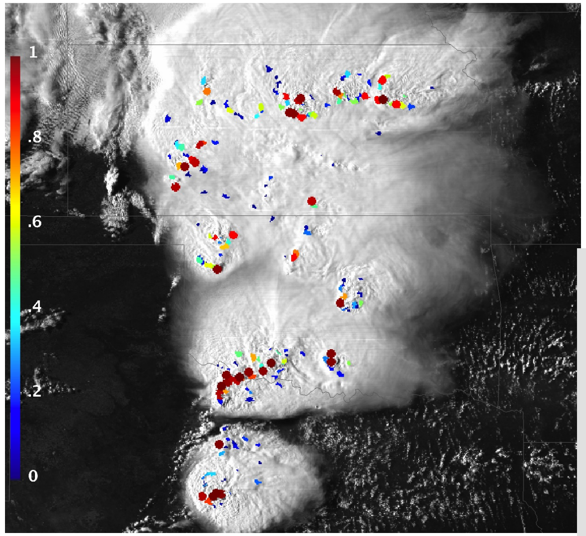


2) Identify cold spots embedded within anvils, and compute temperature difference relative to surrounding anvil

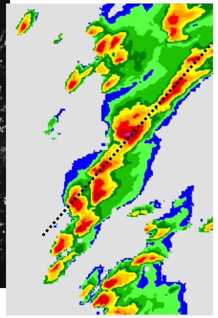


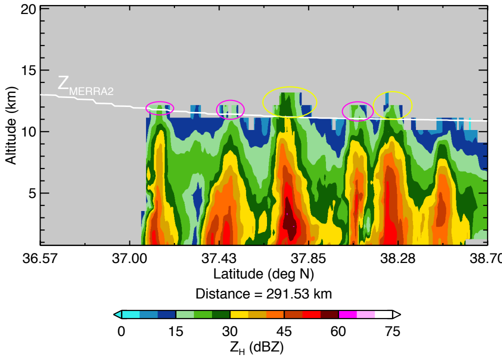


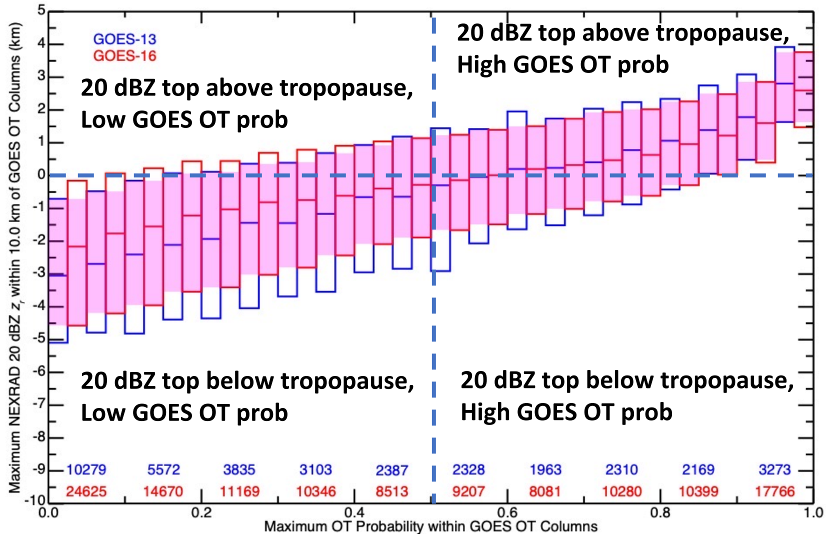
3) Use statistical functions based on human expert and NEXRAD OT identifications to combine IR-anvil, IR-tropopause, anvil area, and anvil spatial uniformity to derive OT Probability



4) Perform quantitative validation of geostationary OT detection using NEXRAD precipitation echo tops



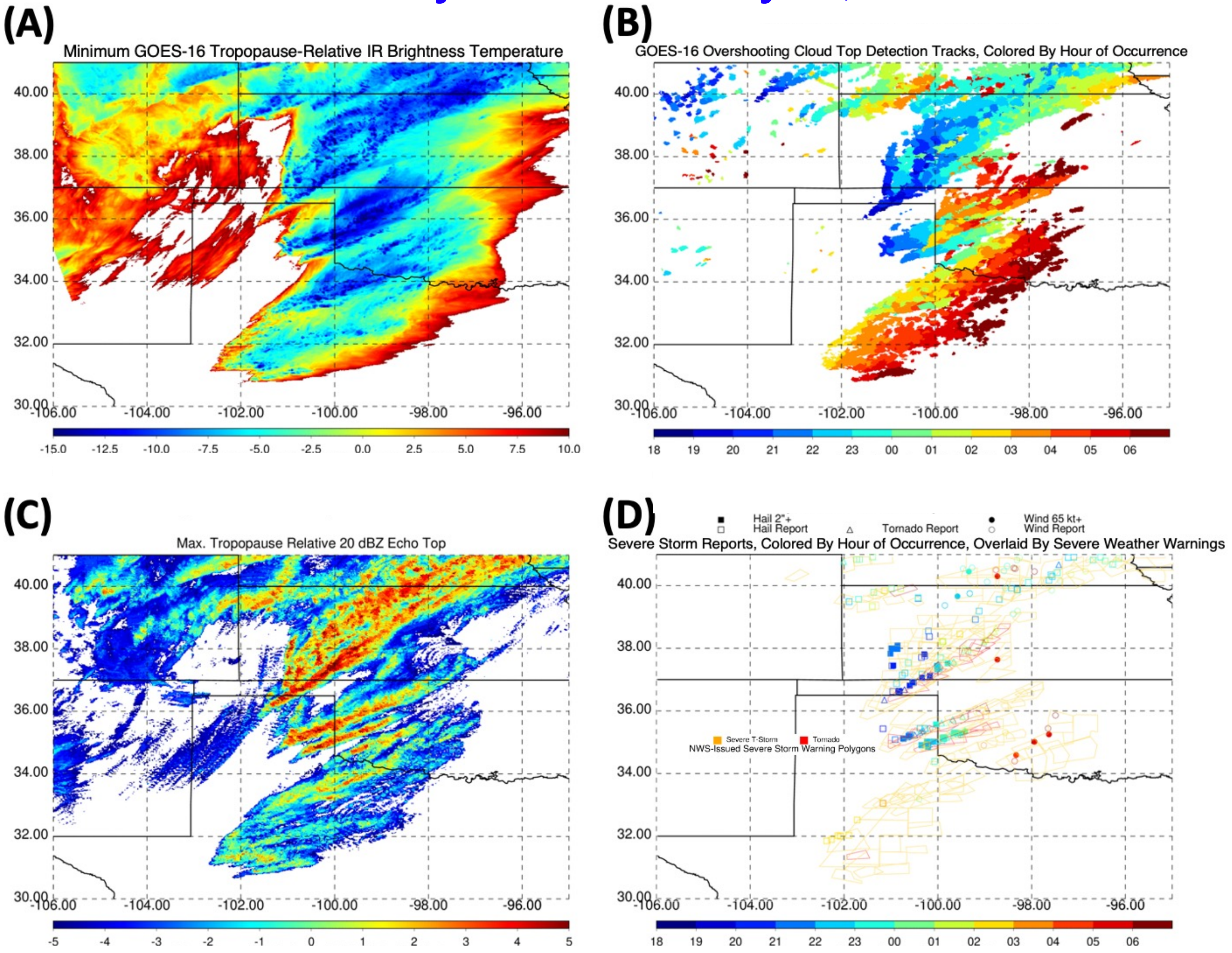




Storm Event Lifetime Overshooting Top Detection Map

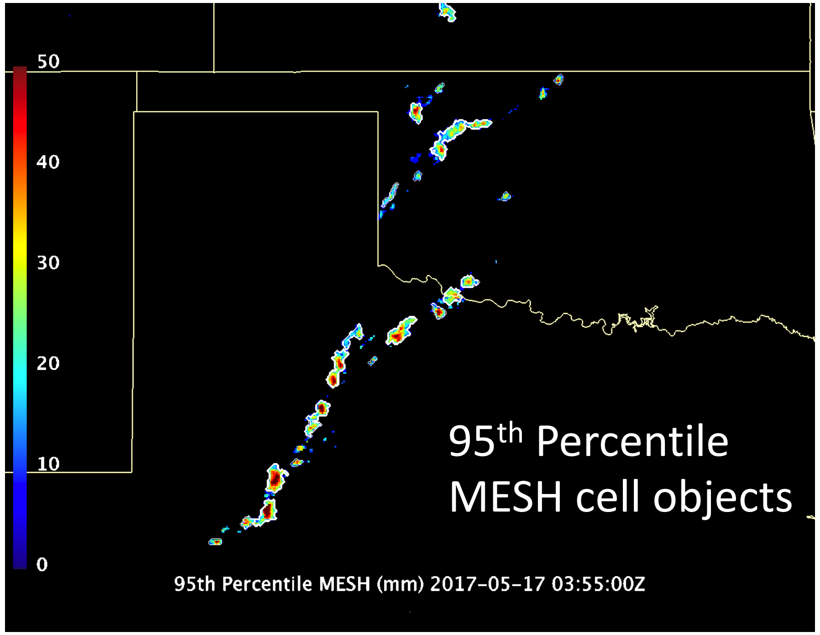
- **Good qualitative agreement between GOES and NEXRAD overshooting top regions**
- **Properties derived at the 2-4 km pixel scale for each detection**
  - OT Probability
  - OT Cloud Top Height
  - OT IR - Tropopause BTD
  - OT IR – Anvil BTD
  - Tropopause Height
  - Area of Cloud < 225 K
  - Mean Anvil Height
  - Cold Spot Area
- **How are these properties distributed for severe hail-producing storms compared to non-severe storms?**

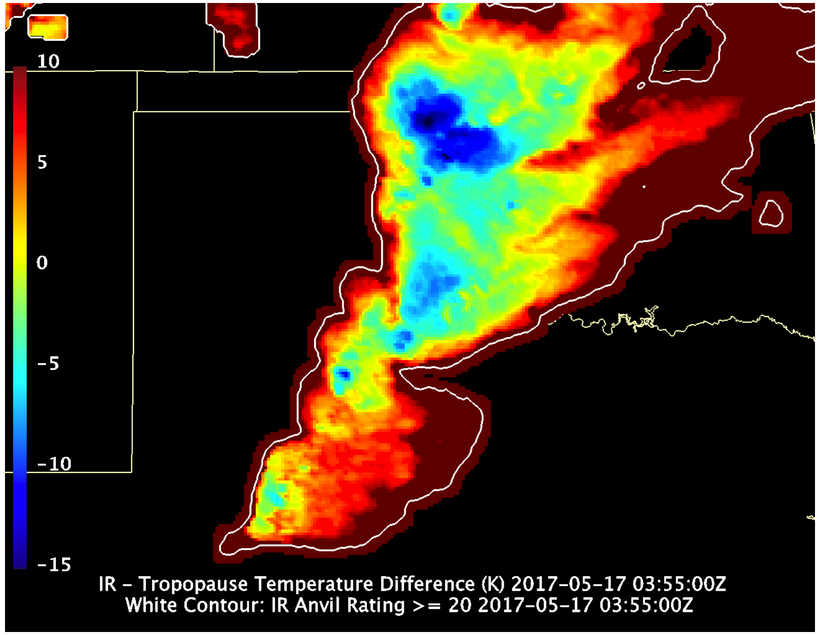
18 UTC May 16 – 7 UTC May 17, 2017

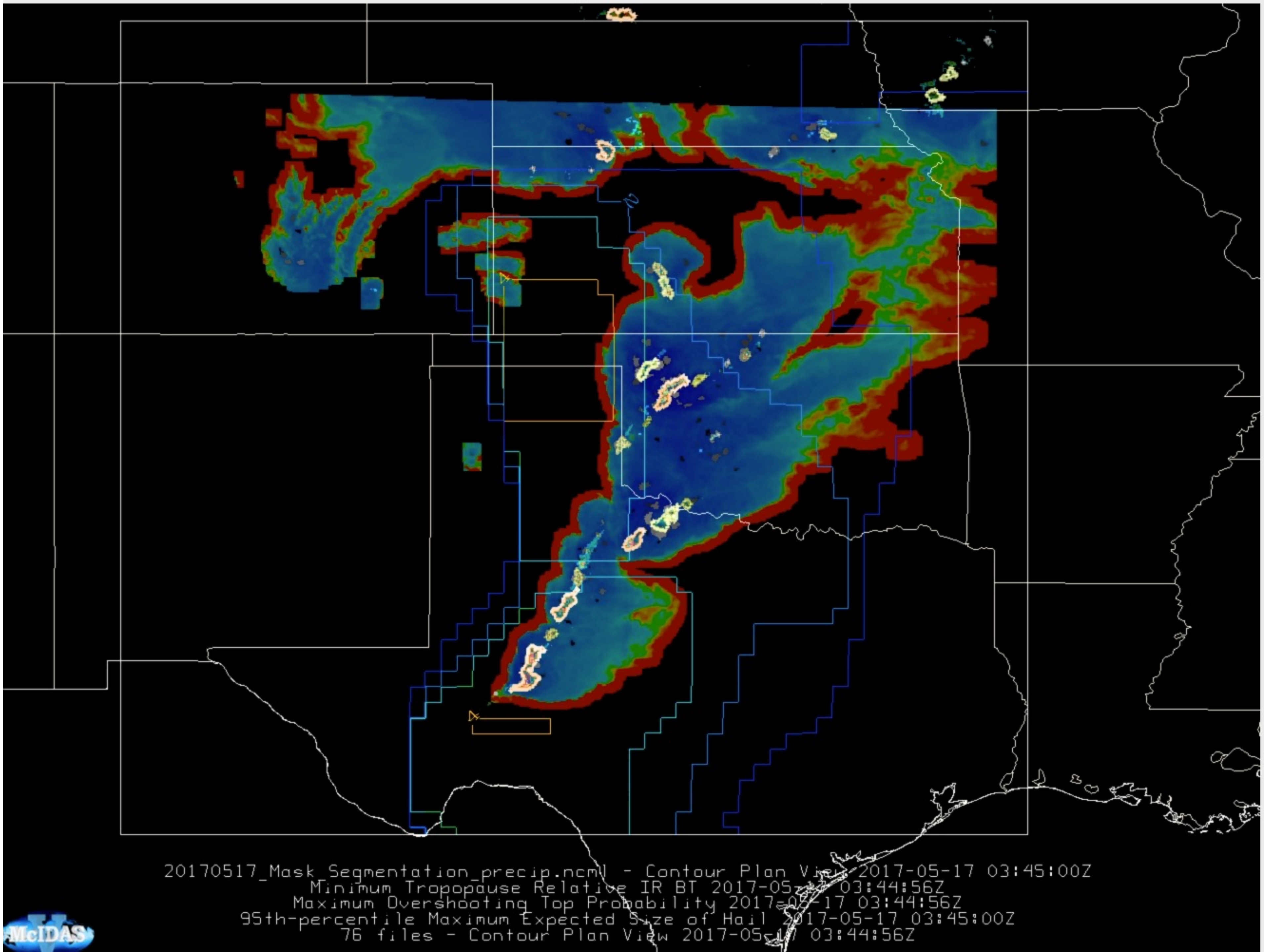


Analysis Process

- **Aggregated over 2013-2017 CONUS warm seasons:**
  - GOES-13 IR properties over anvils
    - Embedded cold spots
      - Local IR BT minima within intense anvils
  - 95th Percentile MESH local maxima cell objects
    - Hourly NEXRAD GridRad (5-minute for validation)
    - Spotter reports matched (within 28 km^2 and +/- 15 min.)
  - Microwave Hail Probability (Bang and Cecil, 2021)
  - ERA5/MERRA-2 environmental parameters (John Allen)
    - Max value within ~1.5° x1.8° lat/lon box
- **Combine GOES-13 and MERRA-2 to quantify severe hail discrimination skill**
  - Let deep neural net objectively decide which parameter combination has best skill
  - GOES-13 proxy for historical GEO imagers



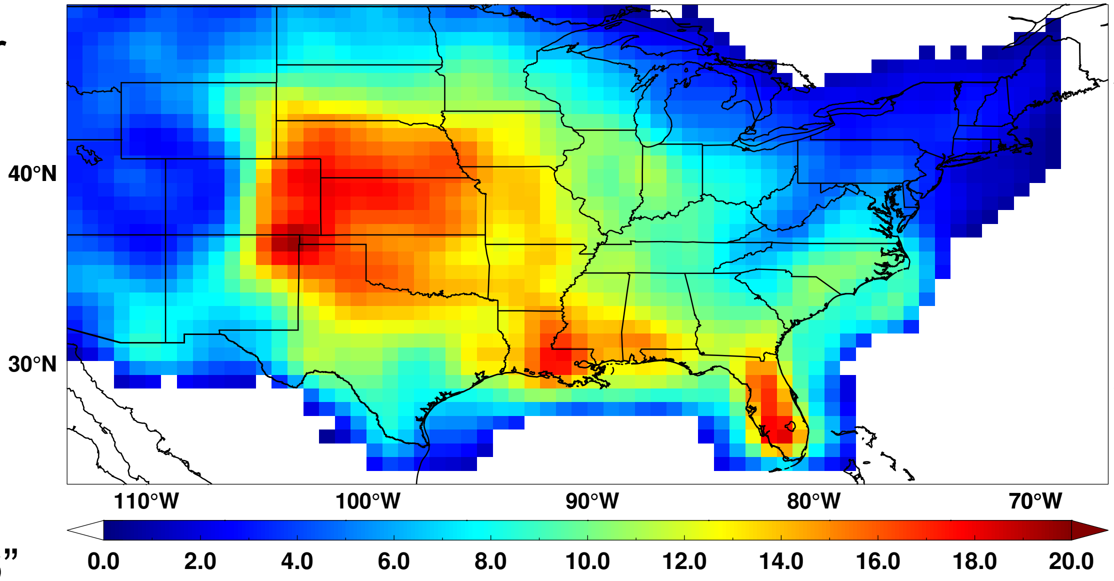



20170517_Mask_Segmentation_precip.ncml - Contour Plan View 2017-05-17 03:45:00Z
Minimum Tropopause Relative IR BT 2017-05-17 03:44:56Z
Maximum Overshooting Top Probability 2017-05-17 03:44:56Z
95th-percentile Maximum Expected Size of Hail 2017-05-17 03:45:00Z
76 files - Contour Plan View 2017-05-17 03:44:56Z
McIDAS

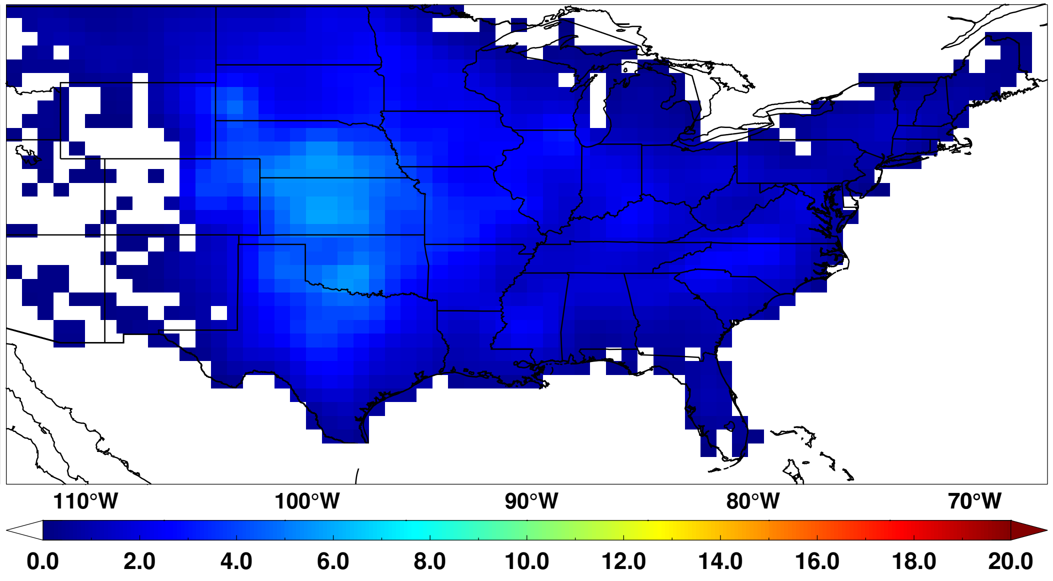
Observed Hail Climatologies

1 detection per day (00:00-23:00 UTC) is required to be considered an **event**

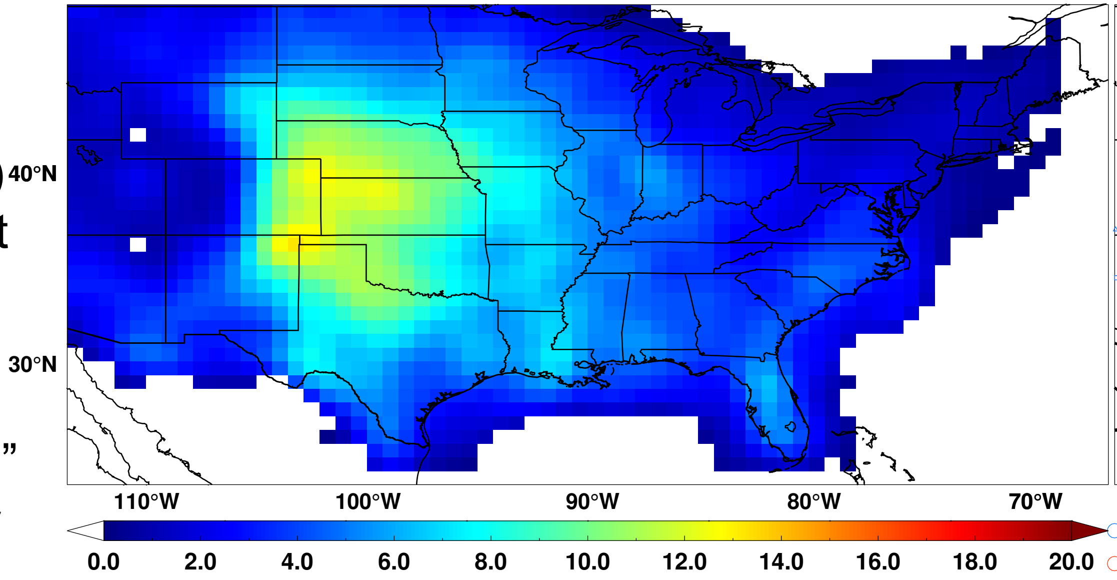
MESH95>=1.6” (40 mm, lower left) more representative of 1”+ (25 mm) hail reported at the ground (upper right) than MESH95>1.18” (30 mm, upper left)



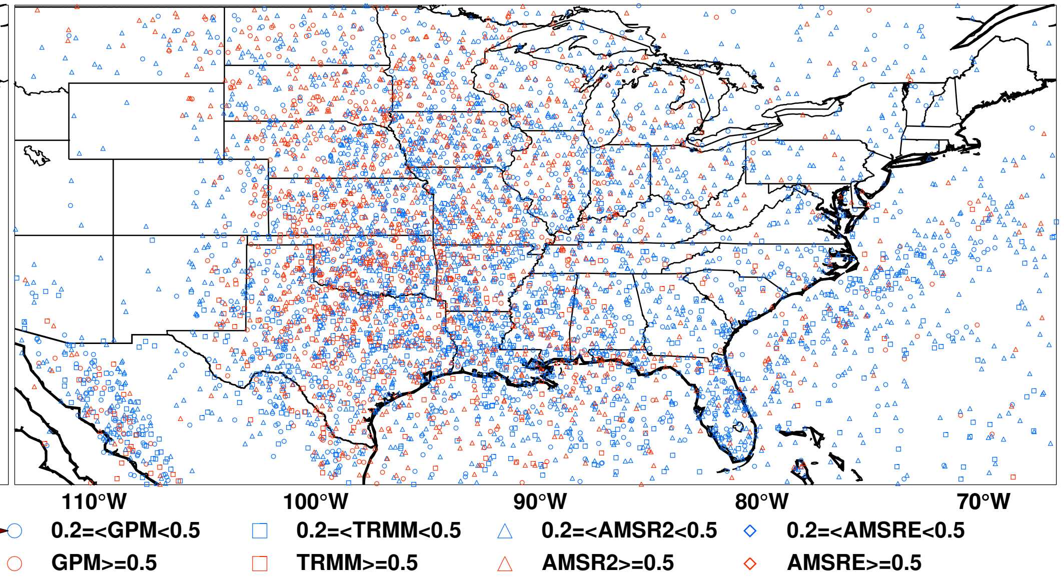
MESH95>= 1.18" Local Maxima, Jan–Dec, 2013–2017 Yearly Average Events



SPC Hail Reports (>=1") Jan–Dec, 2013–2017 Yearly Average Events



MESH95>= 1.60" Local Maxima, Jan–Dec, 2013–2017 Yearly Average Events



Hail Probability (>=0.2), Jan–Dec, 2013–2017

ERA5 and MERRA-2 derived parameters

Parameter short name	Description	Units
surfCape	Surface CAPE	J kg^{-1}
ml1000Cape	1000-m Mixed-Layer CAPE	J kg^{-1}
ml500Cape	500-m Mixed-Layer CAPE	J kg^{-1}
mucape	Most Unstable CAPE	J kg^{-1}
surfCin	Surface CIN	J kg^{-1}
ml1000Cin	1000-m Mixed-Layer CIN	J kg^{-1}
ml500Cin*	500-m Mixed-Layer CIN	J kg^{-1}
muCin*	Most Unstable CIN	J kg^{-1}
fzlv	Freezing Level	m
shear01*	0-1-km Vertical Wind Shear	m s^{-1}
shear06*	0-6-km Vertical Wind Shear	m s^{-1}
t500*	500 hPa Temperature	K
midLapse*	700-500-hPa Lapse Rate	ºC km^{-1}
ship*	Significant Hail Parameter	
stpLM	Significant Tornado Parameter Left-mover	m^2 s^{-2}
stpRM	Significant Tornado Parameter Right-mover	m^2 s^{-2}
scpLM	Supercell Composite Parameter Left-mover	m^2 s^{-2}
scpRM*	Supercell Composite Parameter Right-mover	m^2 s^{-2}
ehi01LM*	0-1-km Energy Helicity Index Left-mover	m^2 s^{-2}
ehi01RM*	0-1-km Energy Helicity Index Right-mover	m^2 s^{-2}
ehi03LM	0-3-km Energy Helicity Index Left-mover	m^2 s^{-2}
ehi03RM*	0-3-km Energy Helicity Index Right-mover	m^2 s^{-2}
srh01LM	0-1-km Storm Relative Helicity Left-mover	m^2 s^{-2}
srh01RM	0-1-km Storm Relative Helicity Right-mover	m^2 s^{-2}
srh03LM	0-3-km Storm Relative Helicity Left-mover	m^2 s^{-2}
srh03RM	0-3-km Storm Relative Helicity Right-mover	m^2 s^{-2}
lapse3*	0-3-km Lapse Rate	ºC km^{-1}
lapse24*	2-4-km Lapse Rate	ºC km^{-1}
thgz	Thickness of Hail Growth Zone	m
sblcl	Surface-based Lifted Condensation Level	m
lclml500	500-m Mixed-Layer Lifted Condensation Level	m

Satellite IR derived cloud top parameters

Parameter short name	Description	Units
OT Probability*	Overshooting Top Probability (Khlopenkov et al. 2021)	
ECS–Anvil BTD*	Brightness temperature difference between coldest pixel and mean anvil background	K
ECS Area	Area of pixels for each unique ECS region	km^2
IR–Tropopause*	GOES IR brightness temperature – MERRA-2 Tropopause temperature	K
Anvil Height*	Average cloud top height of pixels with IR anvil rating>=20 within ~30 km^2	km
Area of Cold Cloud	Average area of pixels with IR BT<225 K within ~30 km^2	km^2
Anvil Frequency*	Percentage of pixels with IR anvil rating>=20 within ~30 km^2	%
Cloud Top Height*	Derived from IR BT match with MERRA-2 temperature profile. OT regions are height-assigned using a constant lapse rate assumption from Griffin et al. (2016)	km
Tropopause Height*	Smoothed tropopause height from MERRA-2 (Khlopenkov et al. 2021)	km

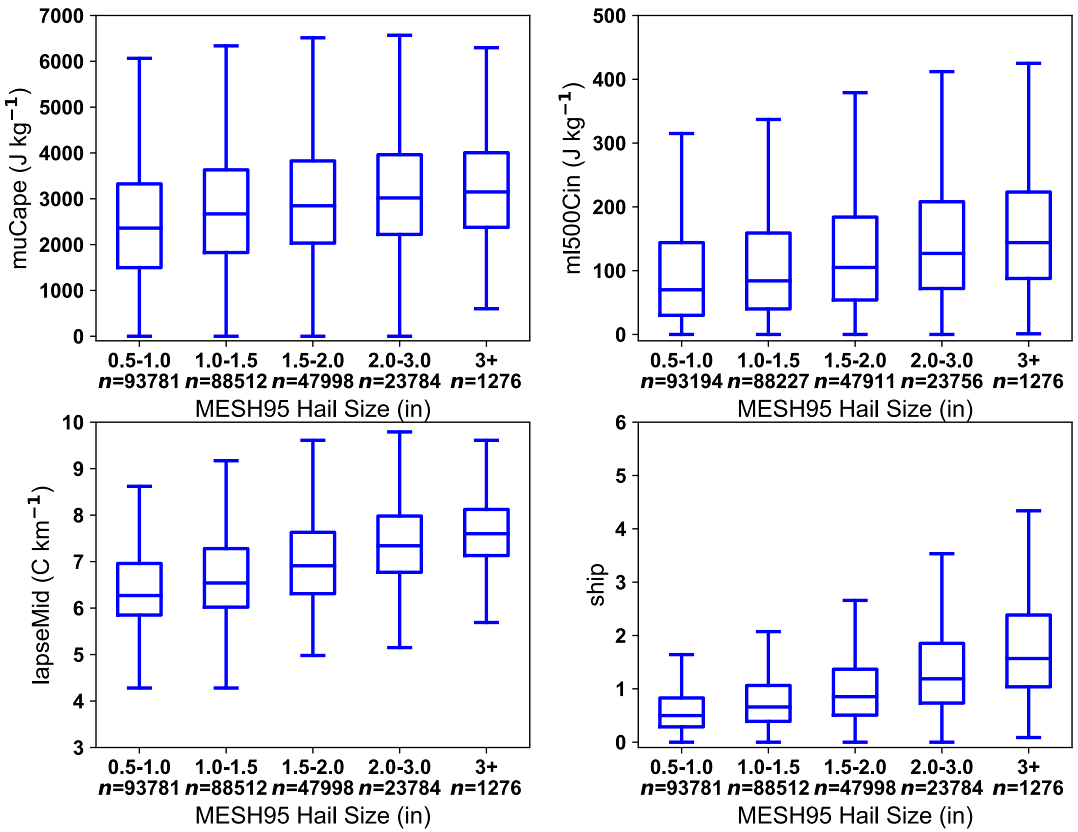
DNN experiments

Experiment name	Description
GOES-13 + ERA5	GOES-13 and ERA5
ERA5 Only	Only ERA5
GOES-13 + MERRA-2	GOES-13 and MERRA2
MERRA-2 Only	Only MERRA-2
GOES-13 Only	Only GOES-13
Direction Independent	GOES-13 and ERA5 excluding those with Left/Right-mover dependency
GOES-16 + ERA5	GOES-16 and ERA5, i.e., GOES-13-based training applied to GOES-16
Warm Season	GOES-13 and ERA5 for days of year 91–273
Cold Season	GOES-13 and ERA5 for days of year 1–90 and 274–366.

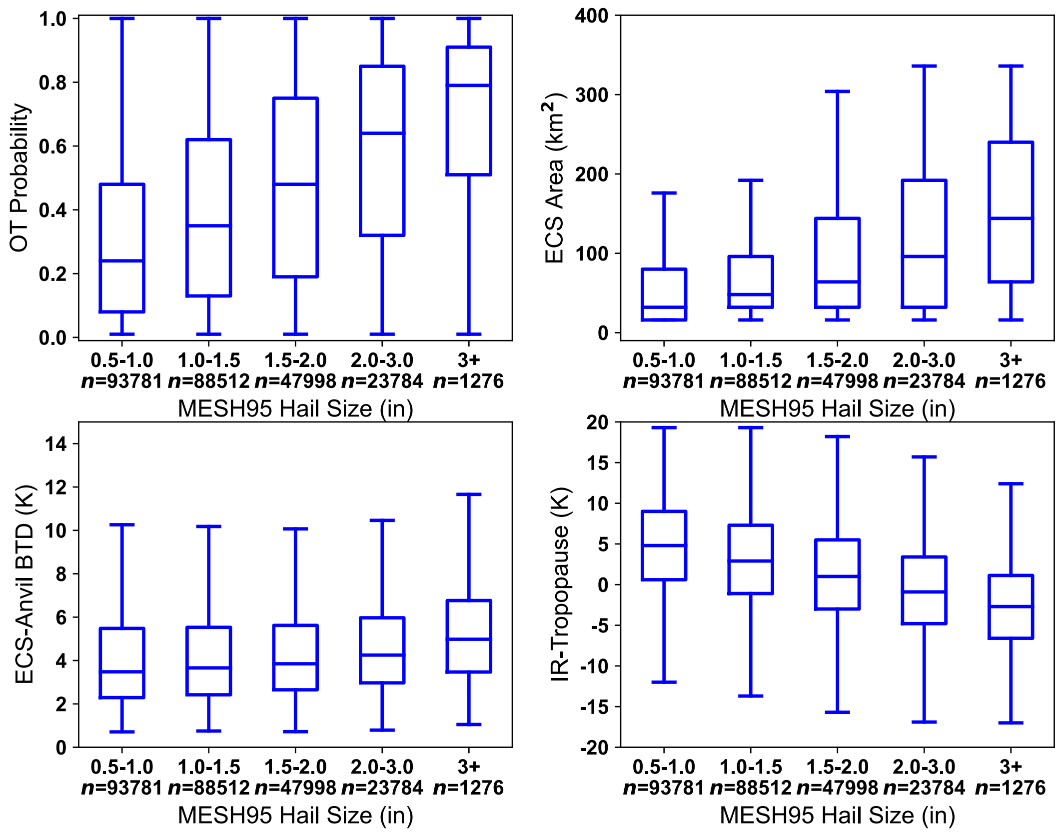
optimal predictors are marked with an asterisk (*)

Parameter Space of Severe Hail

2013-2017 ERA5 Parameters Binned by MESH95 Hail Size (in)



2013-2017 GOES-13 Parameters Binned by MESH95 Hail Size (in)



Marion et al. (2019) found significant correlation between OT Area + tornado intensity

median lapseMid sensitivity to observed hail size agrees closely with those reported by Johnson and Sugden (2014)

Modern reanalysis shows promising ability to discriminate large hail from small hail (SHIP)

Large (2"+) hail producing ECS are:

~5 K colder than tropopause

~3x wider

~60% more likely to be an overshooting top

compared to <1" hail-producing ECS

2013-2017 GOES-13 Matched with ERA5 and Hail Observations

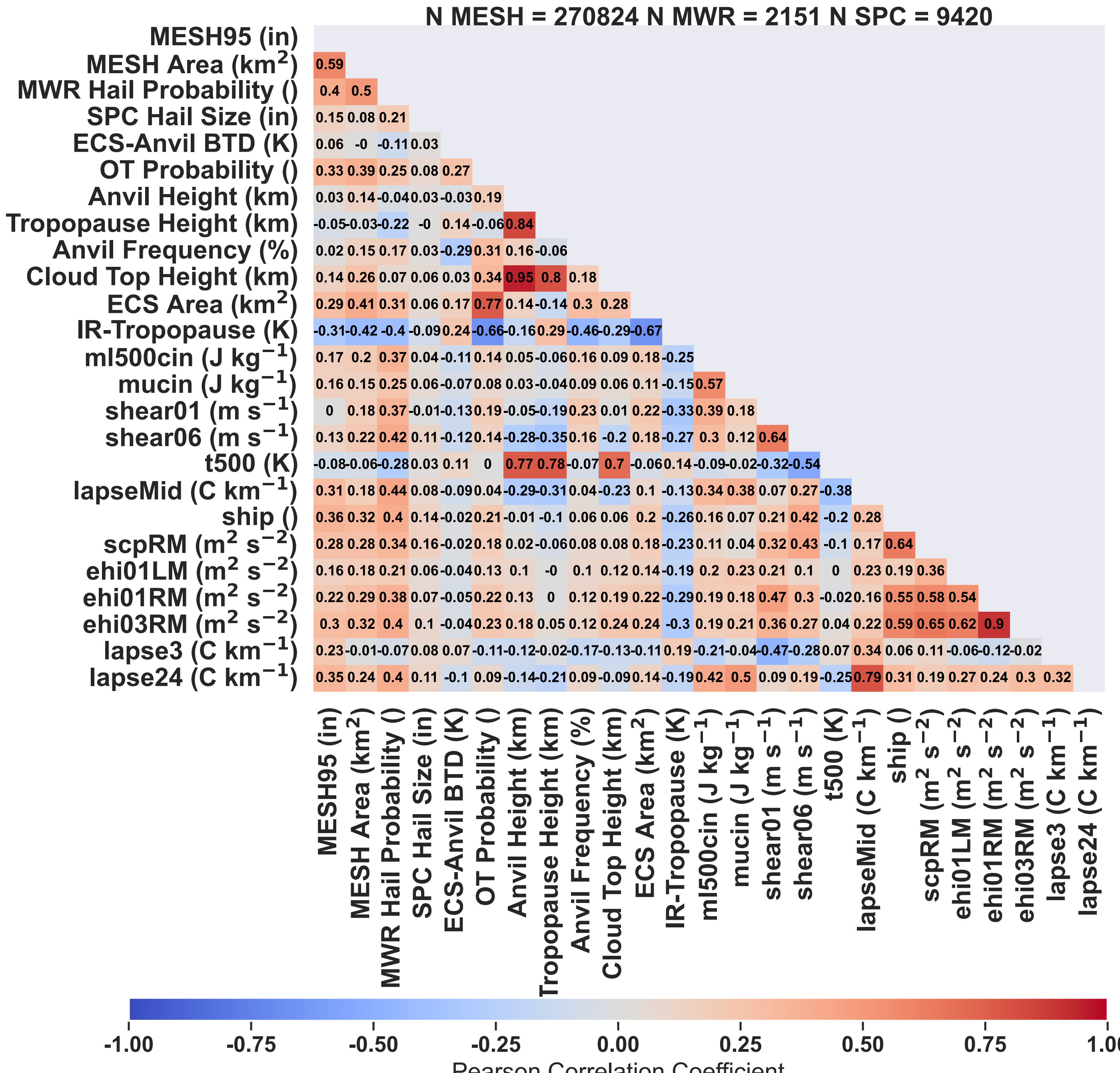
- **preliminary indicators of predictor importance for each observation type**

- **strong collinearity among certain predictors, e.g.,**
- **OT Probability vs. ECS Area (*r*=0.78), ECS Area vs IR-Tropopause (*r*=-0.67), and lapseMid vs. 2-4-km lapse rates (*r*=0.79)**

- **MWR hail probability is best correlated with most ERA5 parameters**

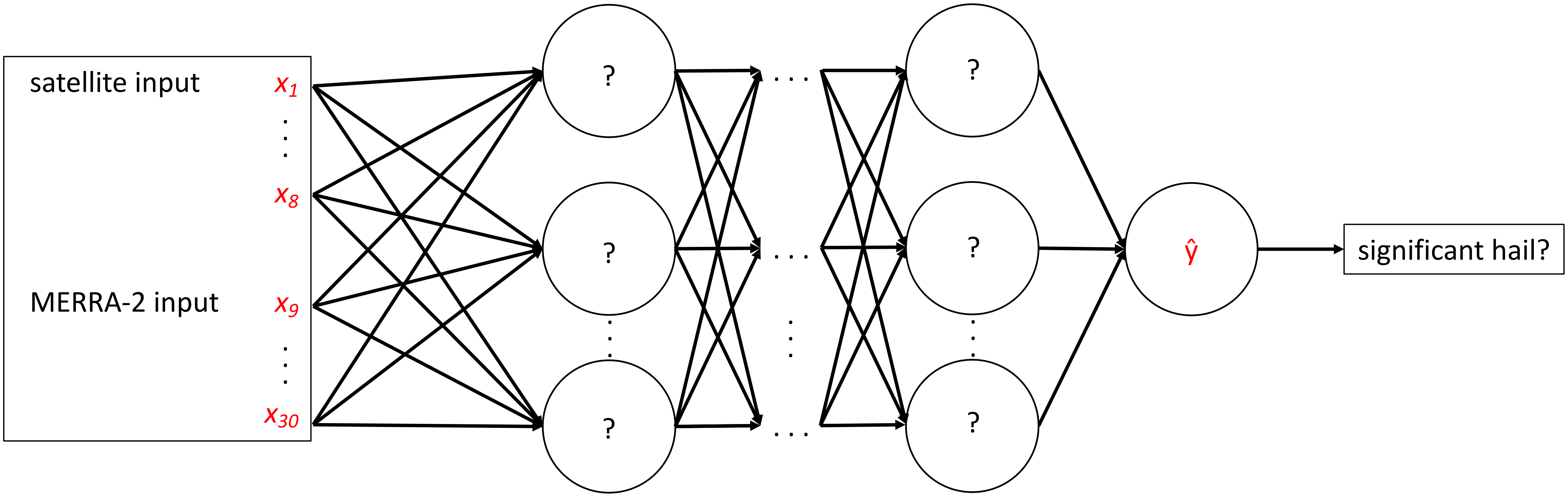
- **Significantly weaker relationships are found between reported hail size and reanalysis/satellite parameters**

- **GOES-16 warm season parameters, revealing 0.05-0.10 improved correlation strength for many**



A Deep Neural Network (DNN) for Predicting Significant Hail

Model computes likelihood of each ECS producing 1.18”+ MESH95 (severe class)



- How well can satellite observations + NWP distinguish significant/insignificant MESH hail events?
- Evaluate skill across multiple DNN model permutations (2013-2017 GOES-13 CONUS)

k-fold Cross-validation Results

- Randomly divide input set into *k* equal divisions (e.g., *k*=10)
- Train on, e.g., 9 divisions and Test on 1 (repeat 9 times)
- Evaluate overall skill across *k* permutations
  - **Recall**: What fraction of true classifications are predicted?
  - **Precision**: What fraction of true predictions correct? (**False Alarms = 1 – Precision**)

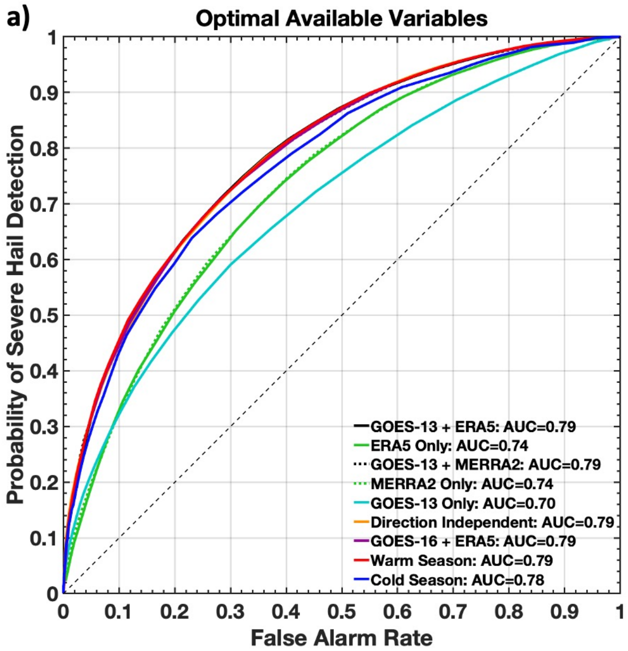
skill metrics evaluated with *k*-fold cross-validation*
for **optimal available DNN inputs**

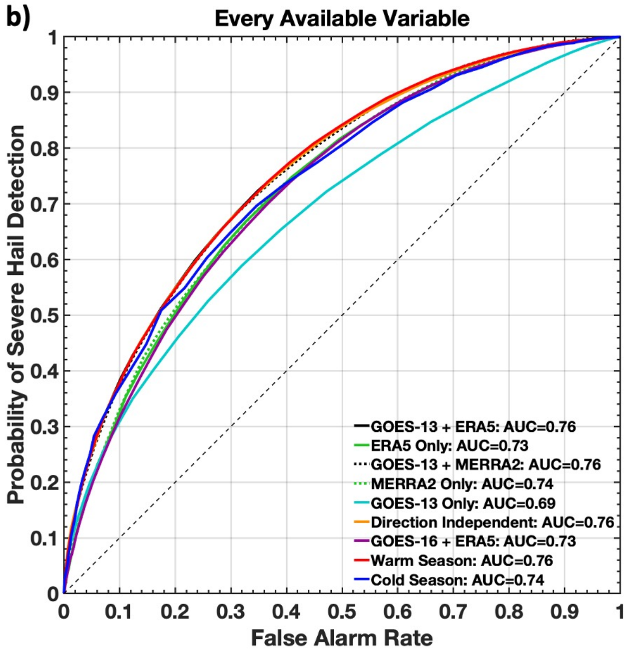
	POD	FA Ratio	FA Rate	CSI	HSS	AUC*	PD Prob*	ROC Prob*
GOES-13 + ERA5	**0.718**	**0.303**	**0.287**	**0.547**	**0.361**	**0.79**	**0.34**	**0.38**
ERA5 Only	0.714	0.356	0.363	0.512	0.300	0.74	0.33	0.38
GOES-13 + MERRA-2	0.711	0.304	0.285	0.543	0.358	0.79	0.33	0.37
MERRA-2 Only	0.697	0.351	0.346	0.506	0.302	0.74	0.34	0.39
GOES-13 Only	0.600	0.364	0.317	0.445	0.255	0.70	0.32	0.37
Direction Independent	0.704	0.300	0.278	0.541	0.359	0.79	0.32	0.37
GOES-16 + ERA5*	0.711	0.281	0.284	0.556	0.350	0.79	0.32	0.37
Warm Season	0.722	0.296	0.295	0.553	0.353	0.79	0.34	0.38
Cold Season	0.699	0.399	0.265	0.477	0.388	0.78	0.33	0.35

Receiver operating characteristic (ROC) curve to assess impact of **probability thresholding** on prediction skill

Probability of Detection (POD) = Recall

False Alarm Rate (FAR) = 1 – Precision





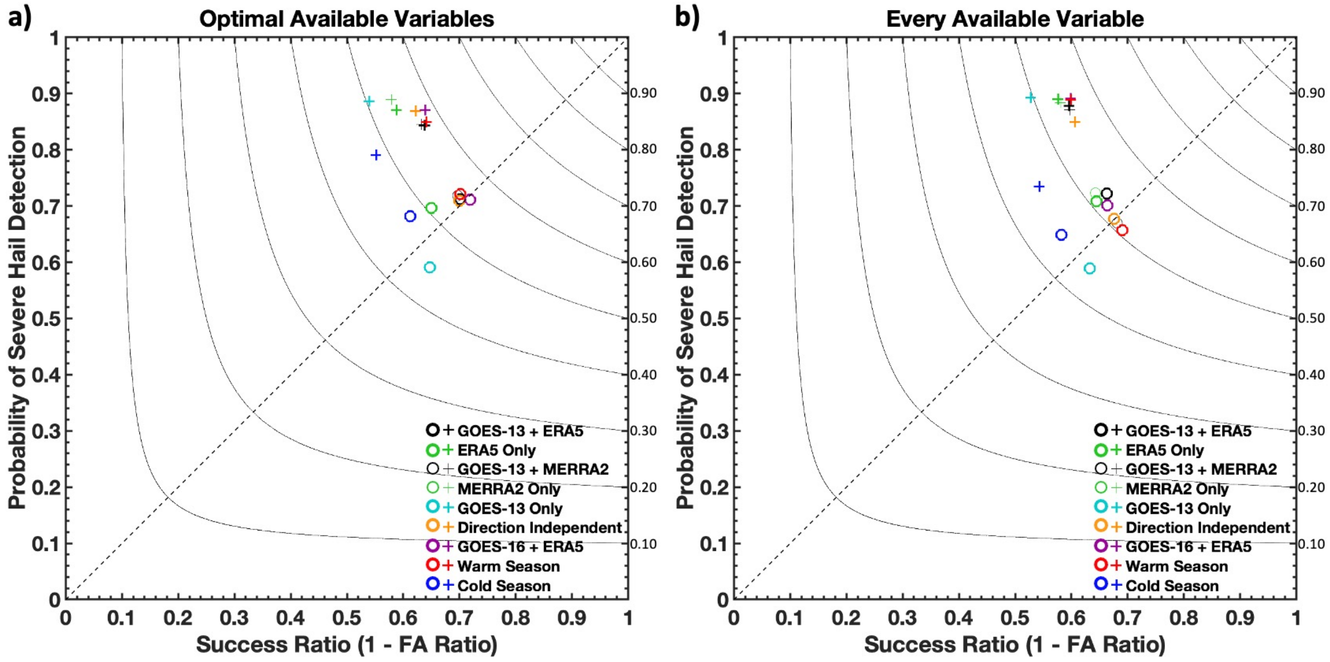
At near 50% confidence in threat of 1.18” hail, model identifies 72% of threats with 28% false alarms

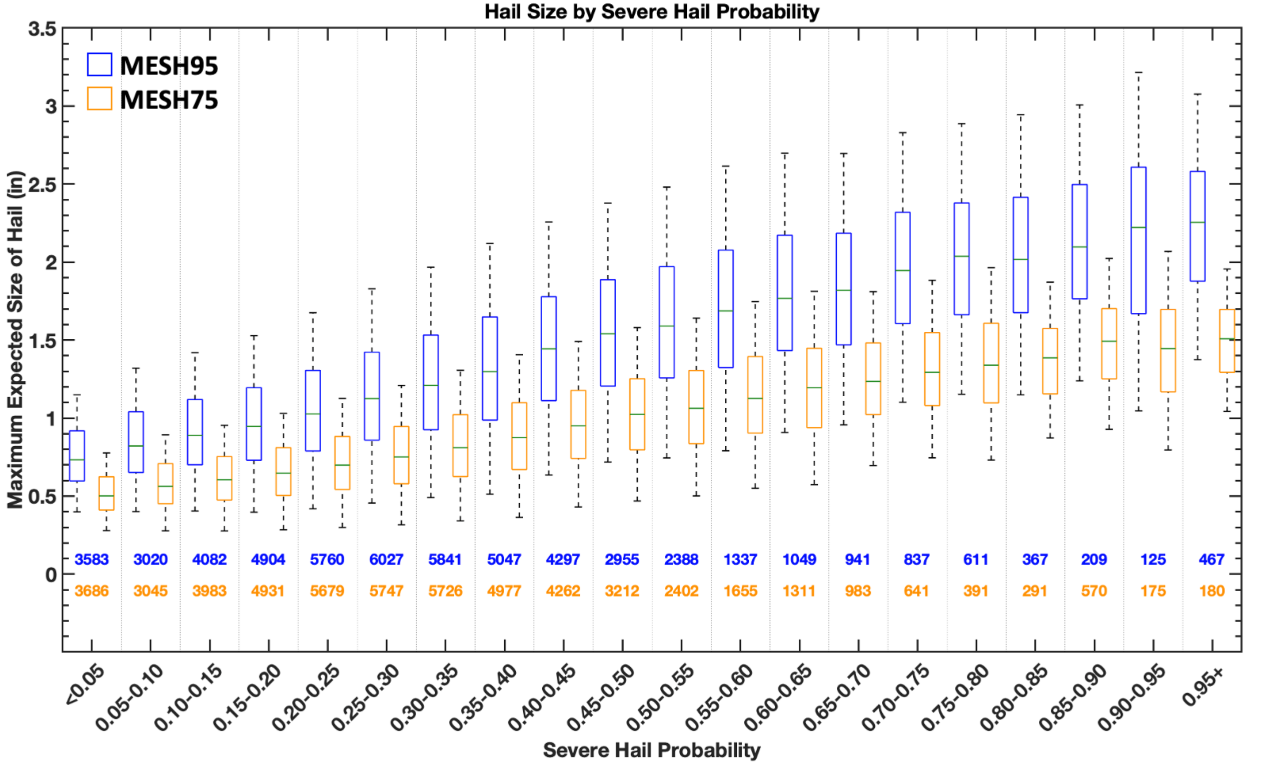
Satellite resolution impacts false alarms

Model data are imperfect

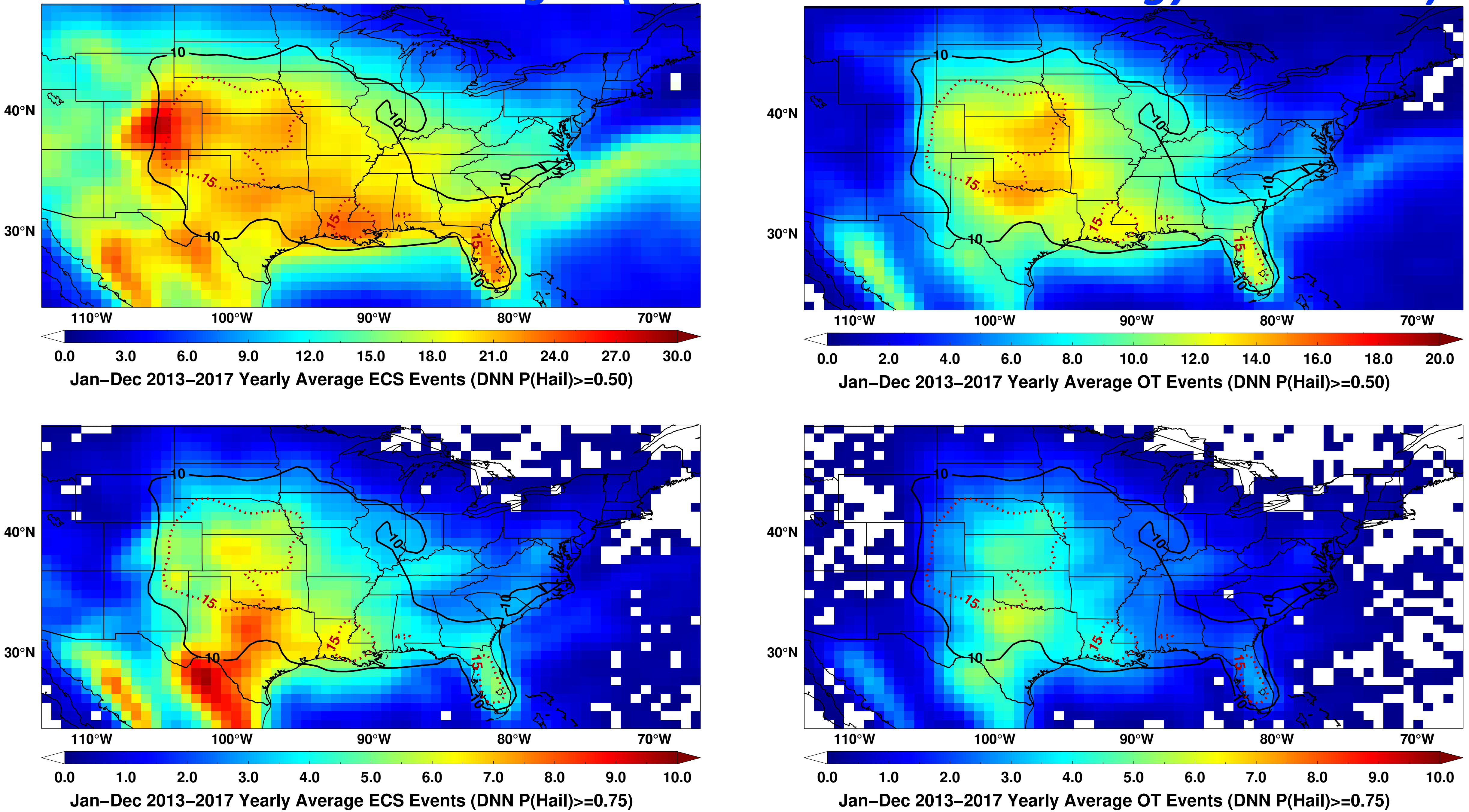
Hail Model Validation

- black curved lines are intervals of CSI from 0.10 through 0.90 (right axis)
- Circles mark the severe classification probability positioned at the peak of the ROC curve
- crosses mark the classification probability where CSI is largest
- Warm season CSI over 0.1 better than cold season
- Large overlap among neighboring bins highlights satellite's limit in discerning marginal features
- DNN is assigning the highest probabilities to the storms which are more likely to generate the largest hail

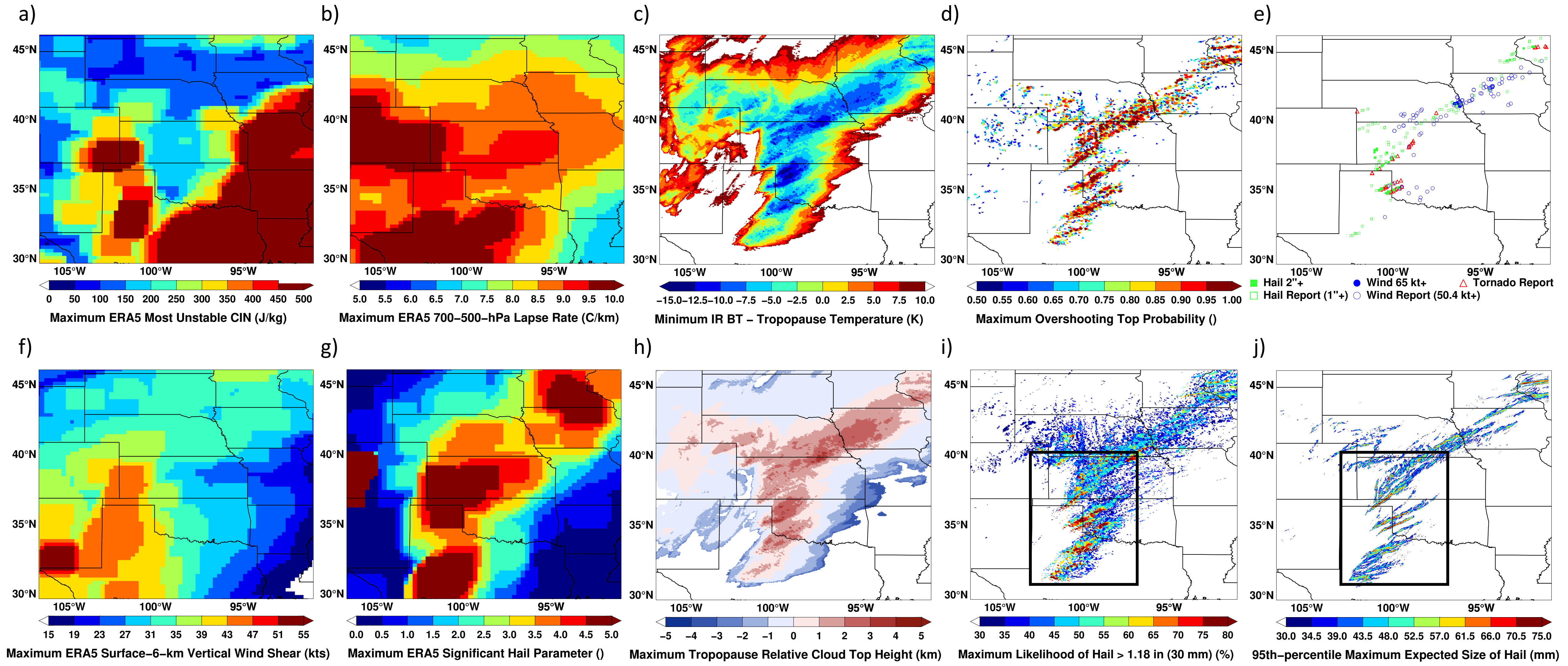




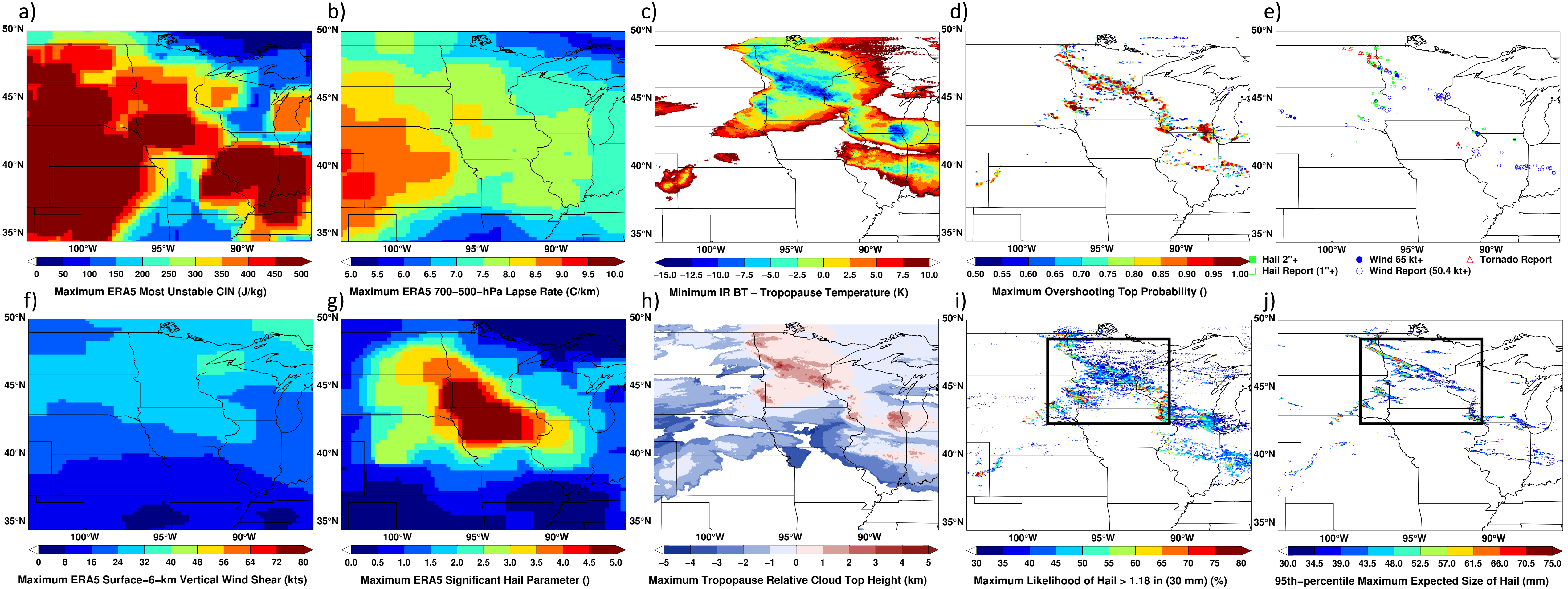
DNN Filtered Climatologies (MESH>=1.18” Climatology Contoured)



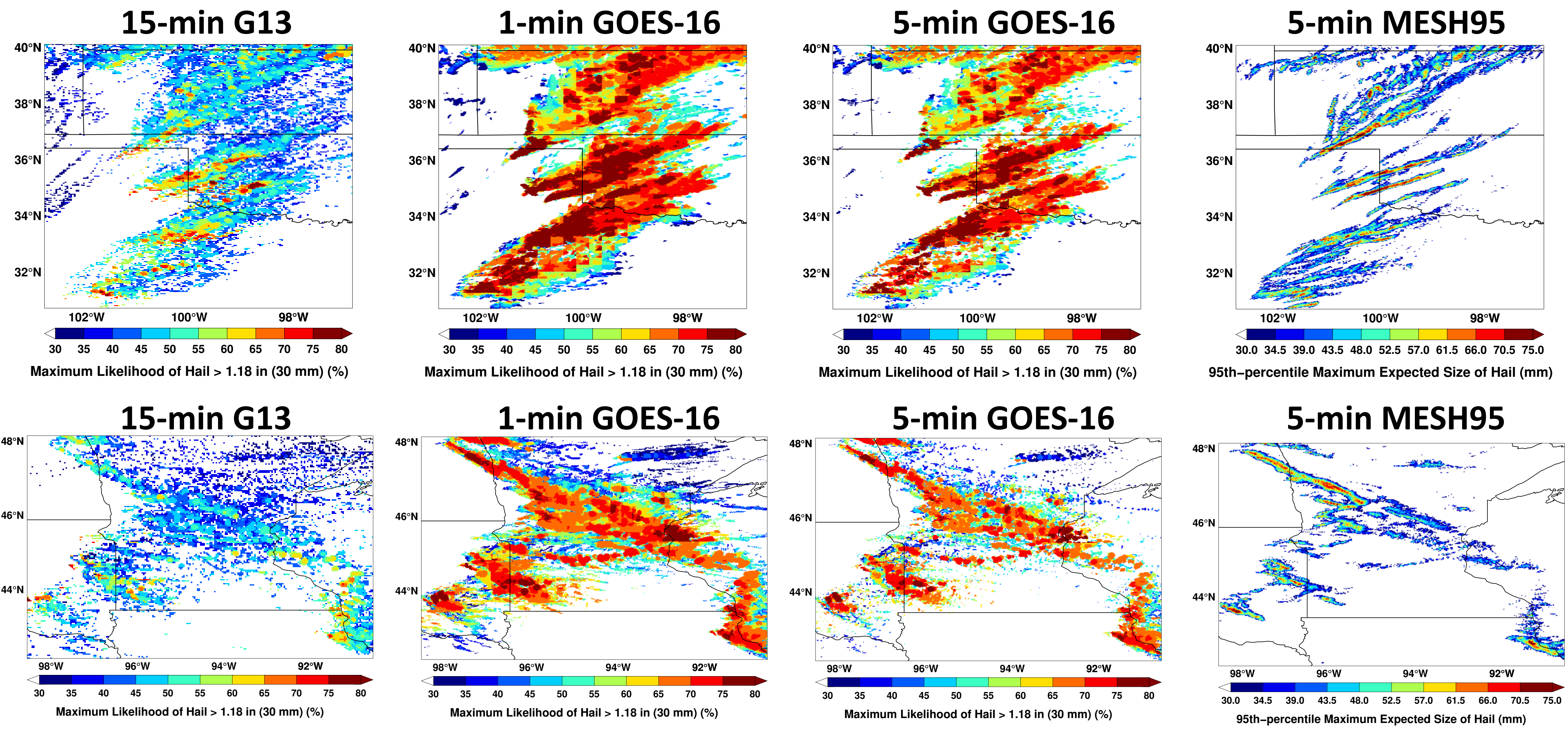
18Z 16 May – 6Z 17 May, 2017 Event Summary



12Z 11 July – 8Z 12 July, 2017 Event Summary



Enhanced Case Studies: Sensitivity to Spatio-temporal Resolution



Summary and Future Plans

We seek to:

- Develop the highest possible resolution and longest duration satellite and reanalysis-based severe storm climatologies
    - Other regions of the world with less reliable warning systems
- Support CatModel development within the reinsurance industry
- Map damage to the land surface damage that hail and other severe weather events generate

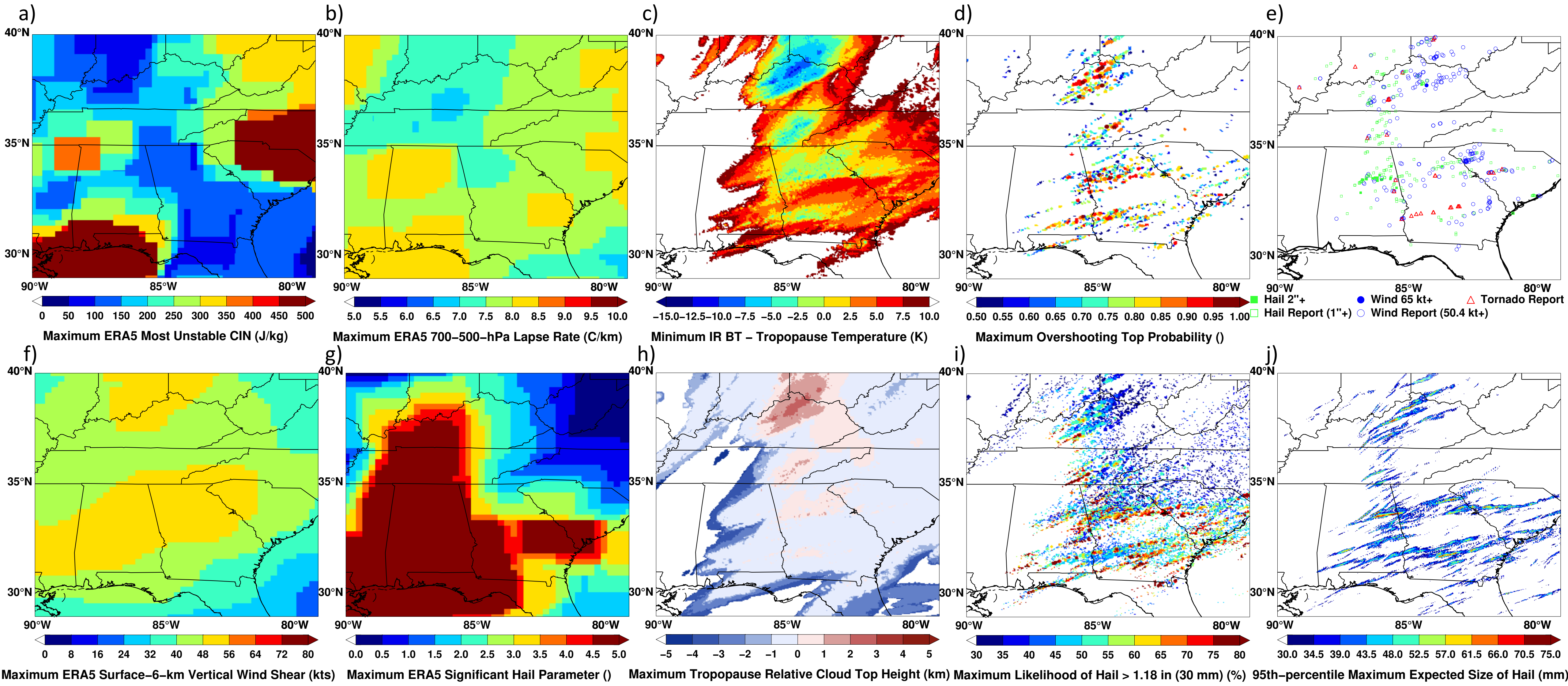
Upcoming Activities Include:

- Generate a 25-year GOES-8 to -16 convection climatology over South America
- Use machine learning to develop an optimal combination of GEO IR + reanalysis parameters to discriminate hail from non-hailstorms observed by passive microwave imagers, human spotters, and U.S. NEXRAD observations
- Better understand the strengths and limitations of using Geostationary Lightning Mapper data for hailstorm detection
- Assess the potential for developing passive microwave hailstorm climatologies, and determine the best way to combine data from various satellites to analyze the diurnal cycle
- Automate a workflow for damage mapping using SAR and optical imager data

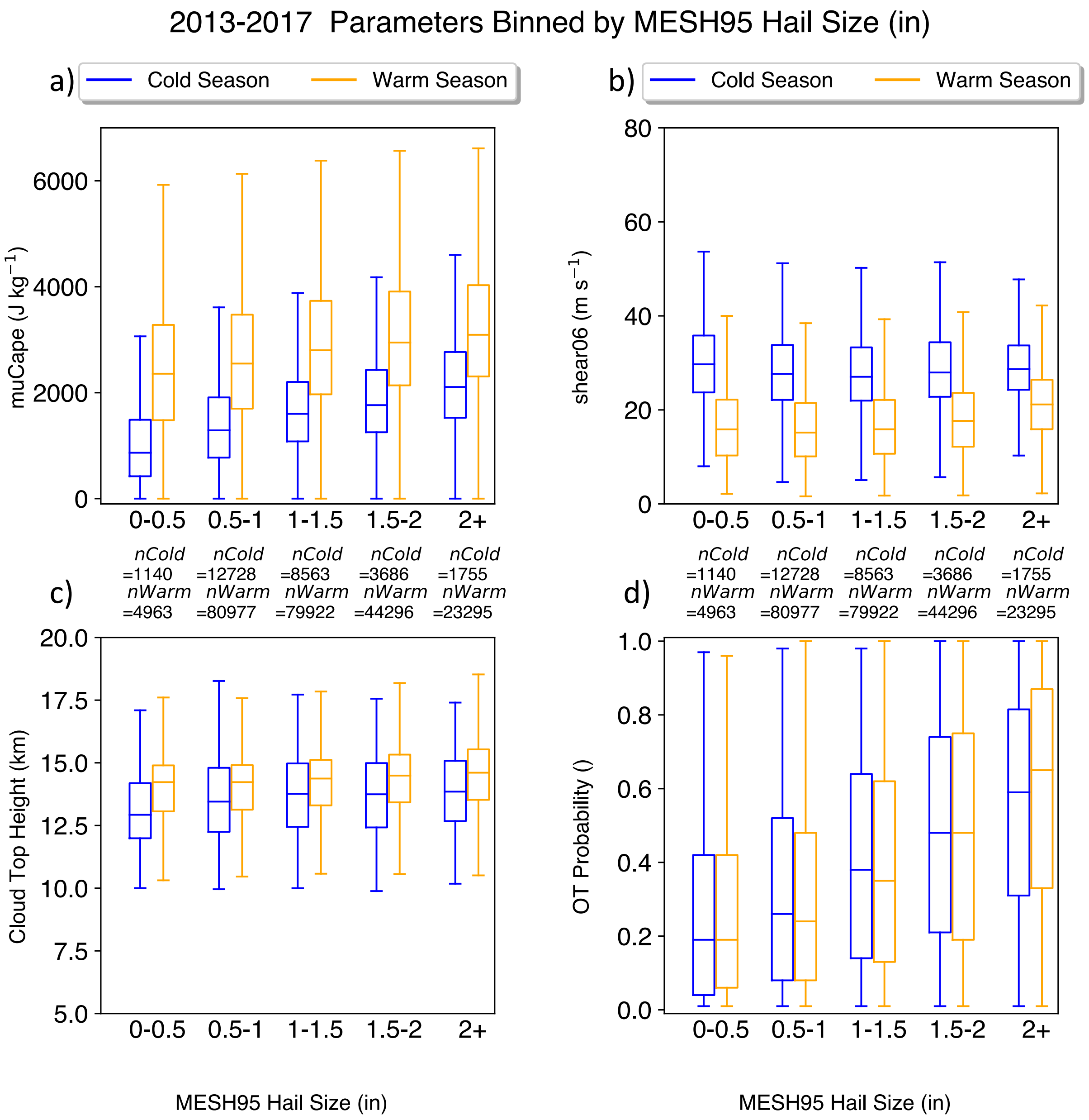
QUESTIONS/COLLABORATIONS?

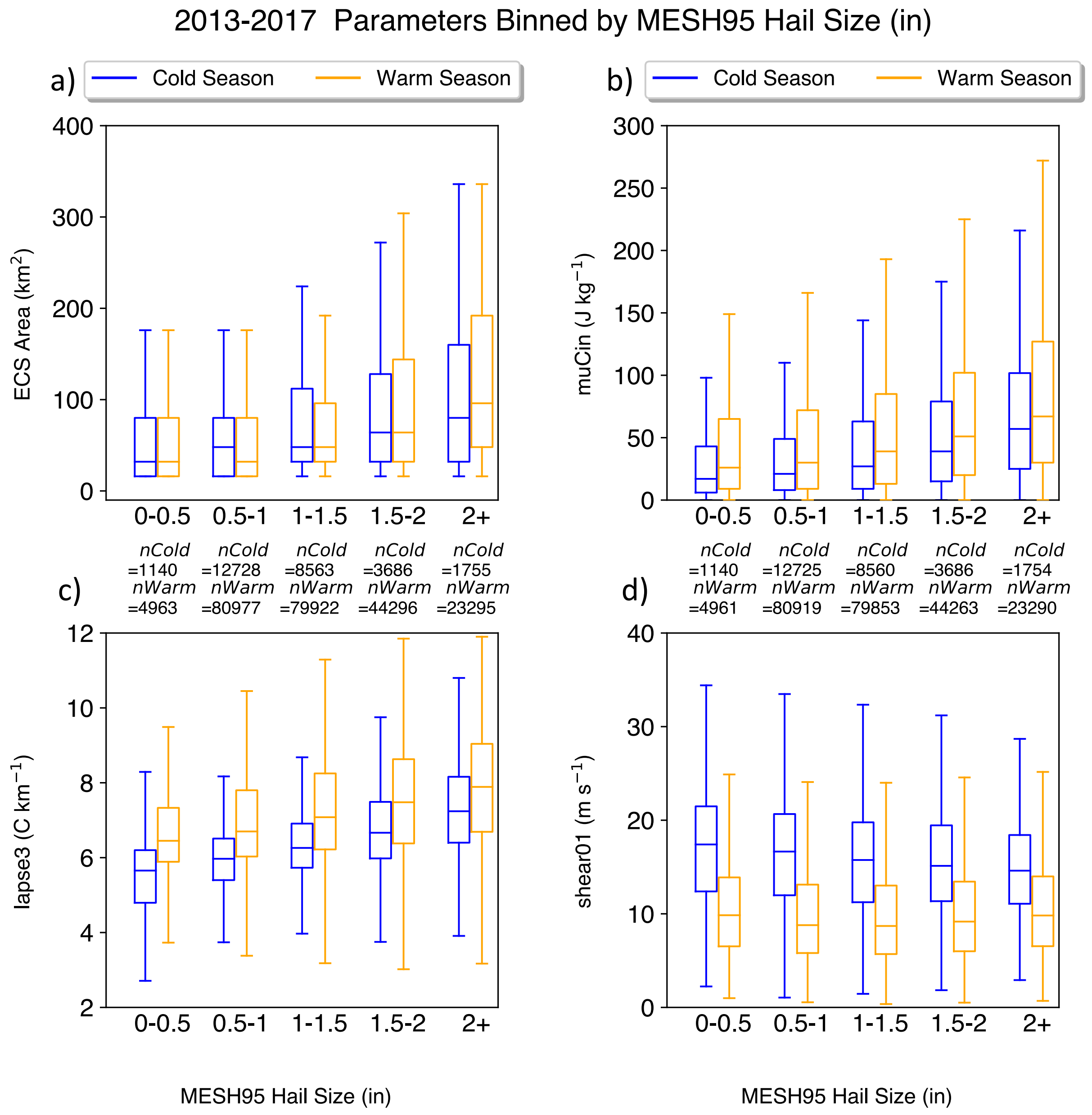
benjamin.r.scarino@nasa.gov
kristopher.m.bedka@nasa.gov
kyle.f.Itterly@nasa.gov (presenting)

13Z 5 April– 6Z 6 April, 2017 Event Summary

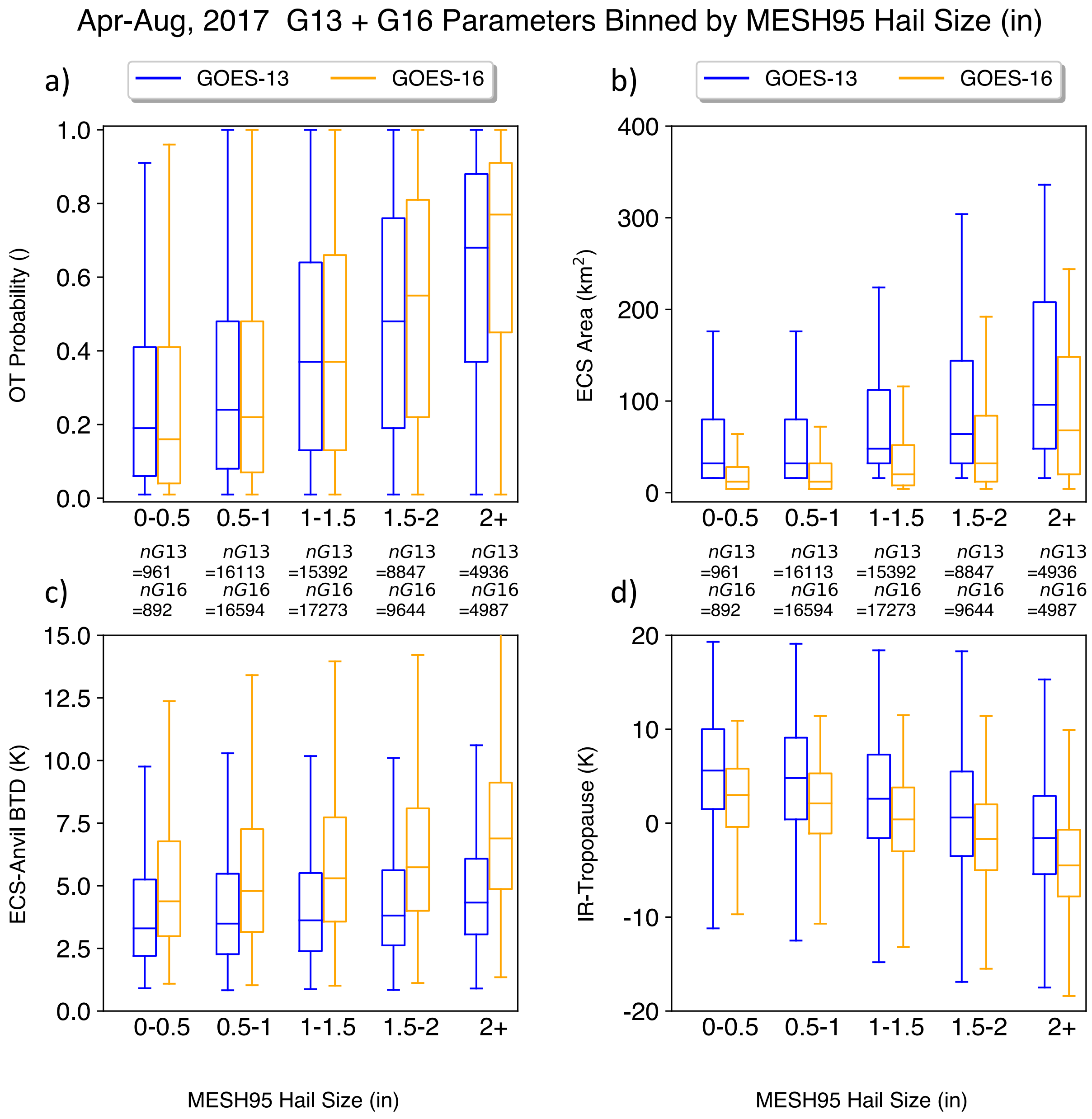


Seasonal Characteristics of Satellite and Reanalysis Hail Predictors





GOES-13 vs. GOES-16 Parameter Sensitivity to Observed Hail



Simplest Example: Logistic Regression

- Think of Logistic Regression as a 1-layer Neural Network, with 1 node
- The node must predict ($\hat{y}$) the outcome (y) given input (x)

- I.e., "Linear Unit"
- Defined by parameters ***w*** and ***b***

1) Linear Regression

$$z(x) = w^T \cdot x + b$$

2) Non-linear Activation

$$\sigma(z) = \frac{1}{1 + e^{-z}} = \hat{y}$$

"sigmoid" function:
$0 \leq \sigma(z) \leq 1$

- I.E., non-linear "Activation Function"
- Restricts output to finite range
- Defines output of the node, i.e., whether it "activates" or "fires"

Logistic Regression as a Neural Network

- Make initial prediction
- Compute cost (error)
- Apply "backward propagation"
  - "Learning" step of "deep learning"
  - Find "downhill direction" toward minimal error
  - Update *w* and *b* at learning rate $\propto$
- Repeat process until convergence
- A larger Neural Networks looks like a Logistic Regression stacked and connected many more times