AI and Applications to Hazardous Weather Forecasting

Dr. Aaron J. Hill

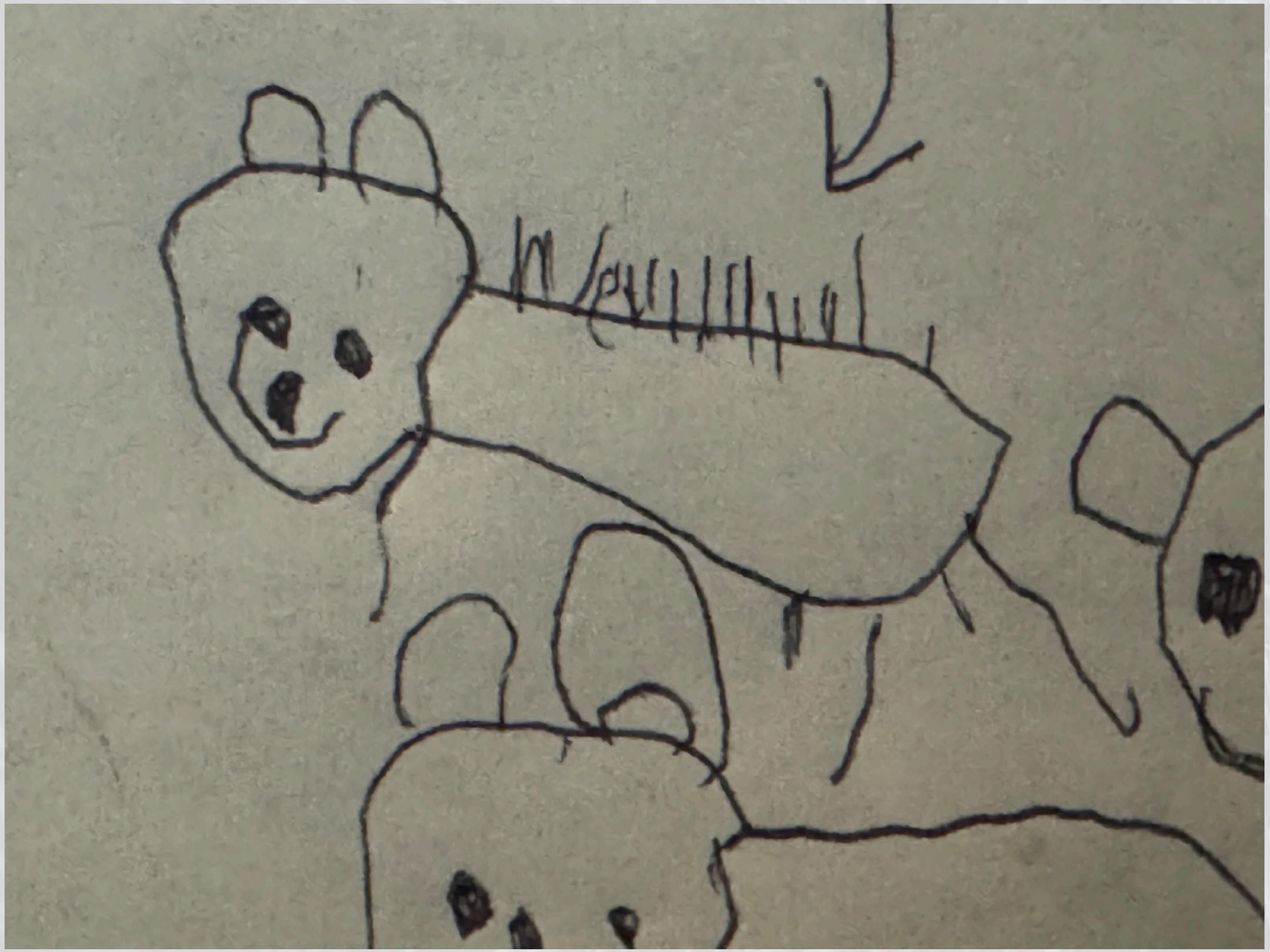
School of Meteorology, University of Oklahoma

NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES)

ESIG 2024 Forecasting and Markets Workshop, Salt Lake City, UT

11 June 2024

Question: What kind of animal is this?



Options

A: Bear
B: Horse
C: Dog
D: Cat

More specifically: A “husky puppy”

Source: My 5-year-old

How did you arrive at your answer?

You likely considered any number of the following:

- The image looked like something you had seen before
- The shape or proportions resembled a dog
- The shape had some legs, even if not the amount (or position) you might have expected…
- It had a tail…maybe
- It had ears that looked like a dog's ears
- You have small children that have made similar-looking animal figures

And so on…

We have been trained through experience to key in on features and patterns to help us identify animals.

How did you arrive at your answer?

You likely considered any number of the following:

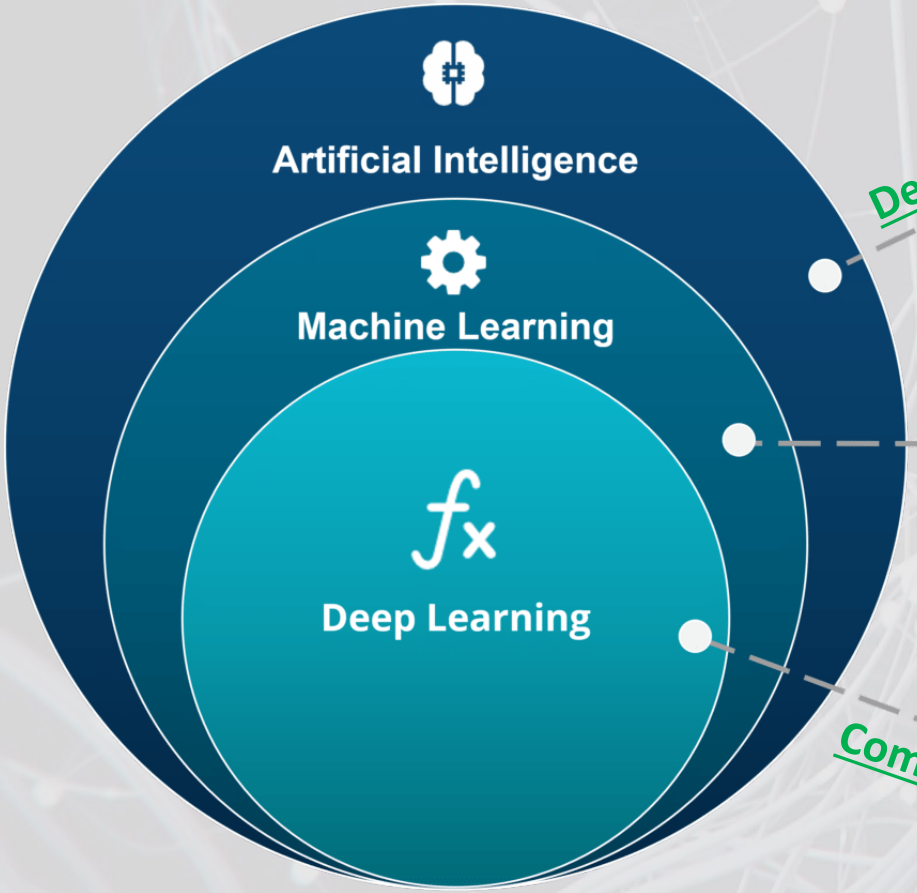
- The image looked like something you had seen before
- You have small children that have made similar looking animal figures

And so on…

We have been trained through experience to key in on features and patterns to help us identify animals.

Artificial Intelligence replaces our finite human processing capability with a computer's power and resources

So what is AI anyways?



Decision making

ARTIFICIAL INTELLIGENCE
A technique which enables machines to mimic human behaviour

Pattern recognition

MACHINE LEARNING
Subset of AI technique which use statistical methods to enable machines to improve with experience

Complex Patterning

DEEP LEARNING
Subset of ML which make the computation of multi-layer neural network feasible

More complex algorithms that enable richer insights into complex data

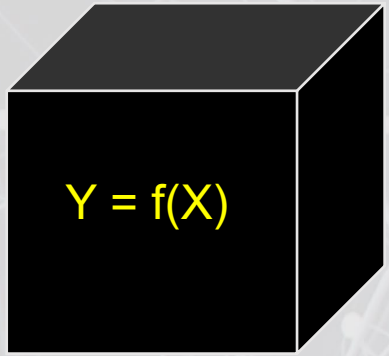
Source: https://www.edureka.co/blog/wp-content/uploads/2018/03/AI-vs-ML-vs-Deep-Learning.png

Domain-specific questions control the problem setup

What controls the prediction?

What do you want to predict?

Data
X(s)

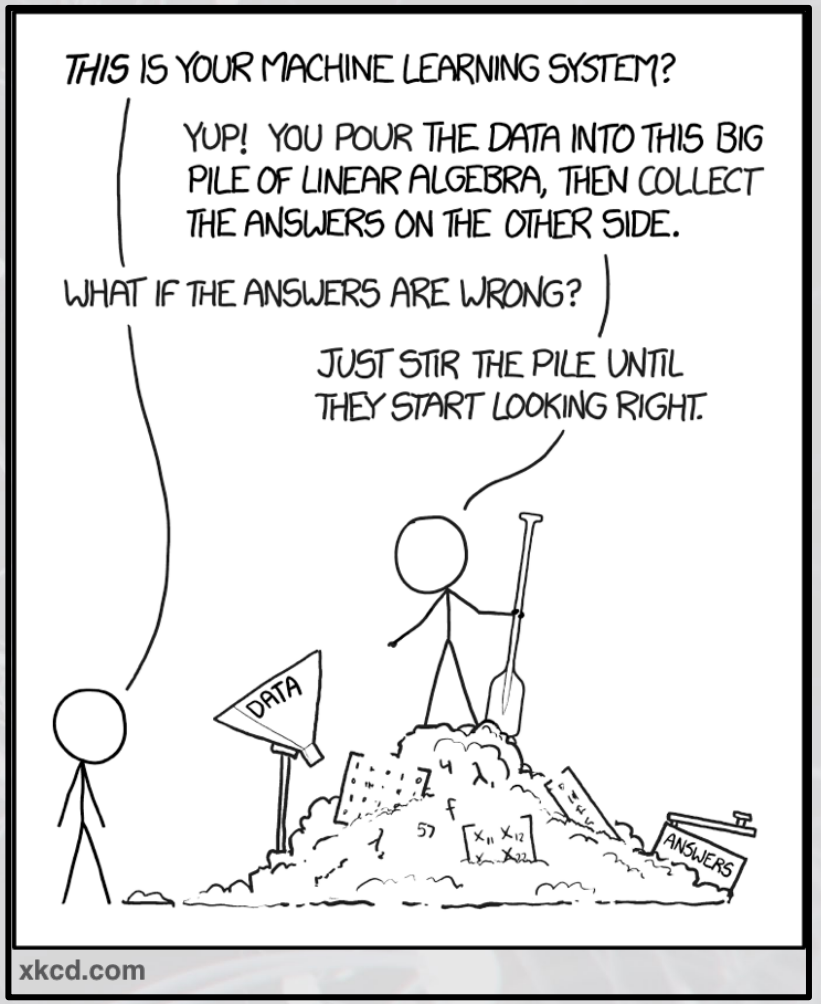


Prediction
Y

Choose f(X)

Linear/logistic regression
Decision Trees/Random Forests
Neural Networks
Support Vector Machines

Convolutional Neural Networks, Generative Adversarial Networks, Diffusion Models, Variational Autoencoders…



Domain-specific questions control the problem setup

What controls the prediction?

What do you want to predict?

Heat, humidity, mountains, snow…

Y = f(X)

Aaron is miserable

Decision Trees

Is it snowing?

no

Is it humid?

no

yes

Not miserable!

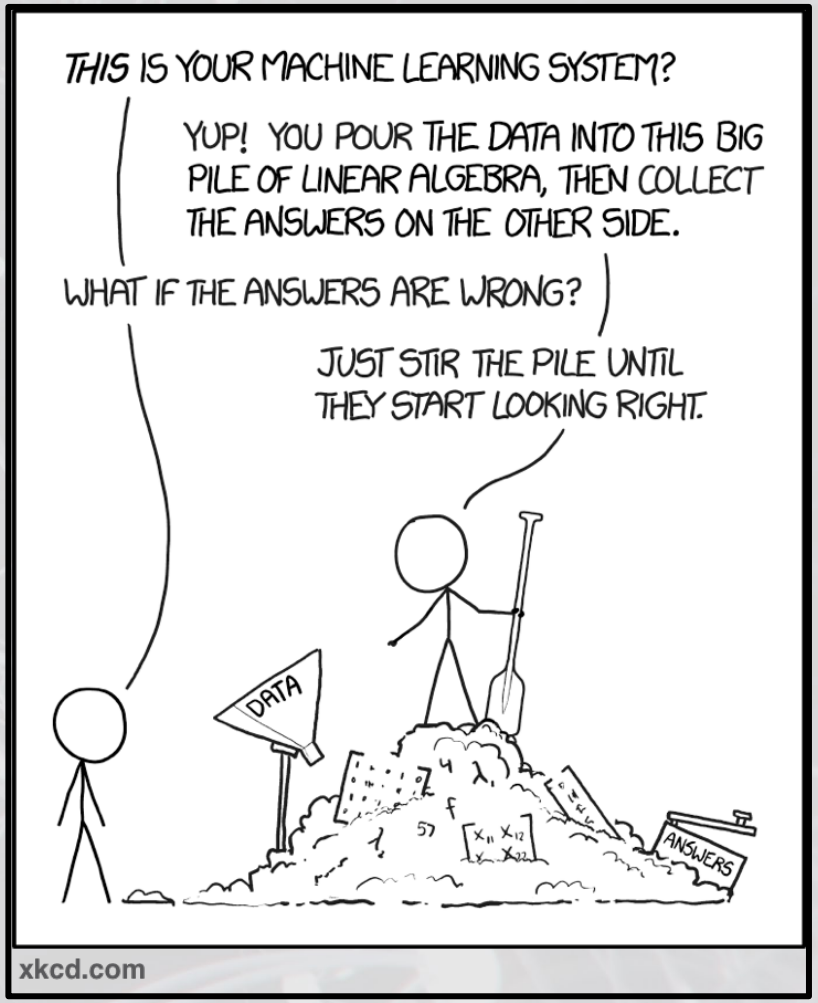
Is it hot?

yes

no

Miserable…

Not miserable!



Weather is a great domain for pattern recognition and classification

Weather is a crucial variable across a variety of industries (emergency management, agricultural commodities, energy markets, air travel)

We have petabytes of data being generated daily in weather domains

Forecasting is often a matter of pattern recognition – humans are actually quite good at this when given enough time!

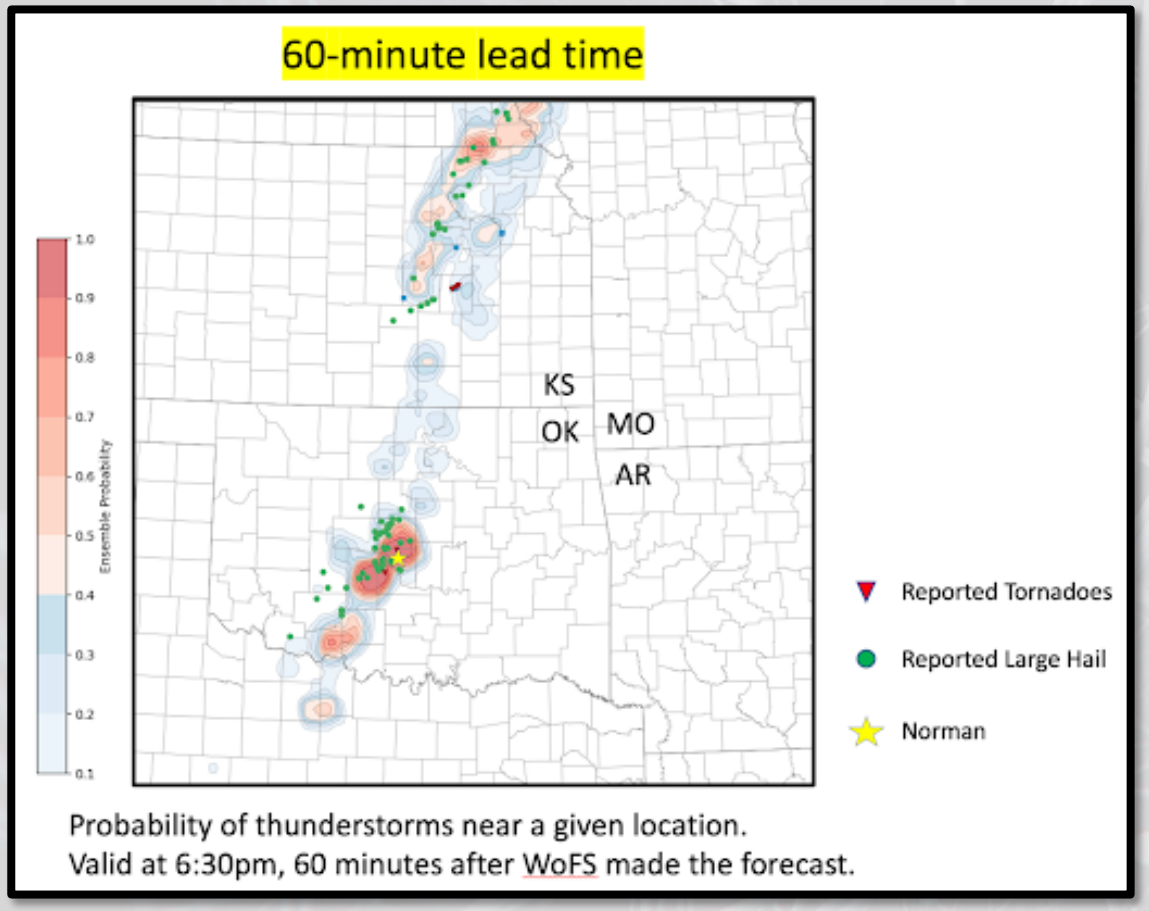
Weather is a great domain for pattern recognition and classification

With improved computing we can now train the **computer** to identify patterns (e.g, tropical cyclone structure) and objects (e.g., hook echo in radar) in data – traditionally, this was done by humans alone (or other statistical algorithms). Can have complex patterns in data!

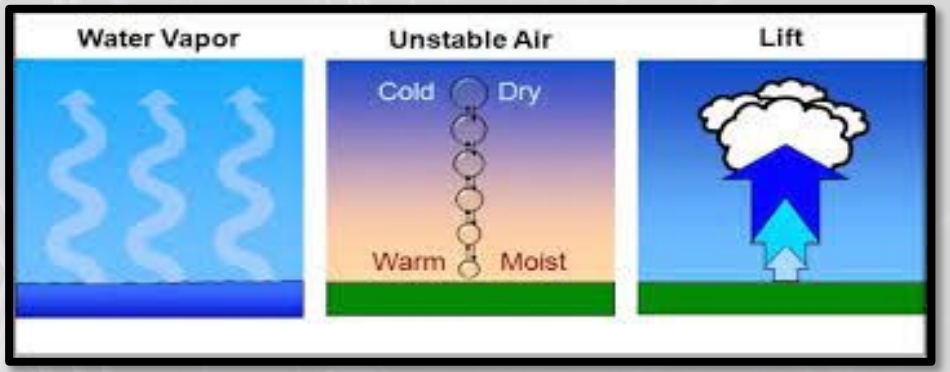
Can postprocess observations, numerical weather prediction model output to bias-correct forecasts, generate new forecasts, and meet client/end-user needs

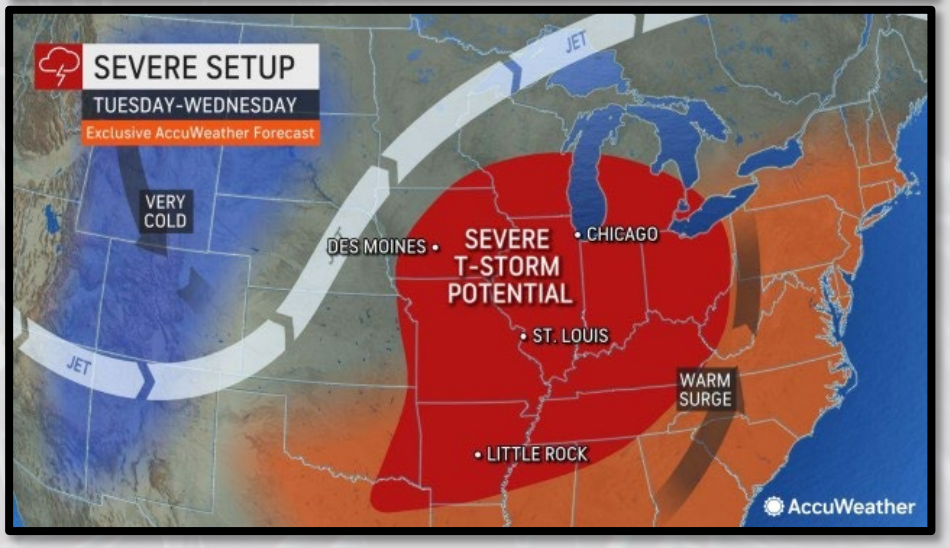
Current approaches for predictions of high-impact weather hazards

High-resolution models;
Logistical/predictability limits > 3 days



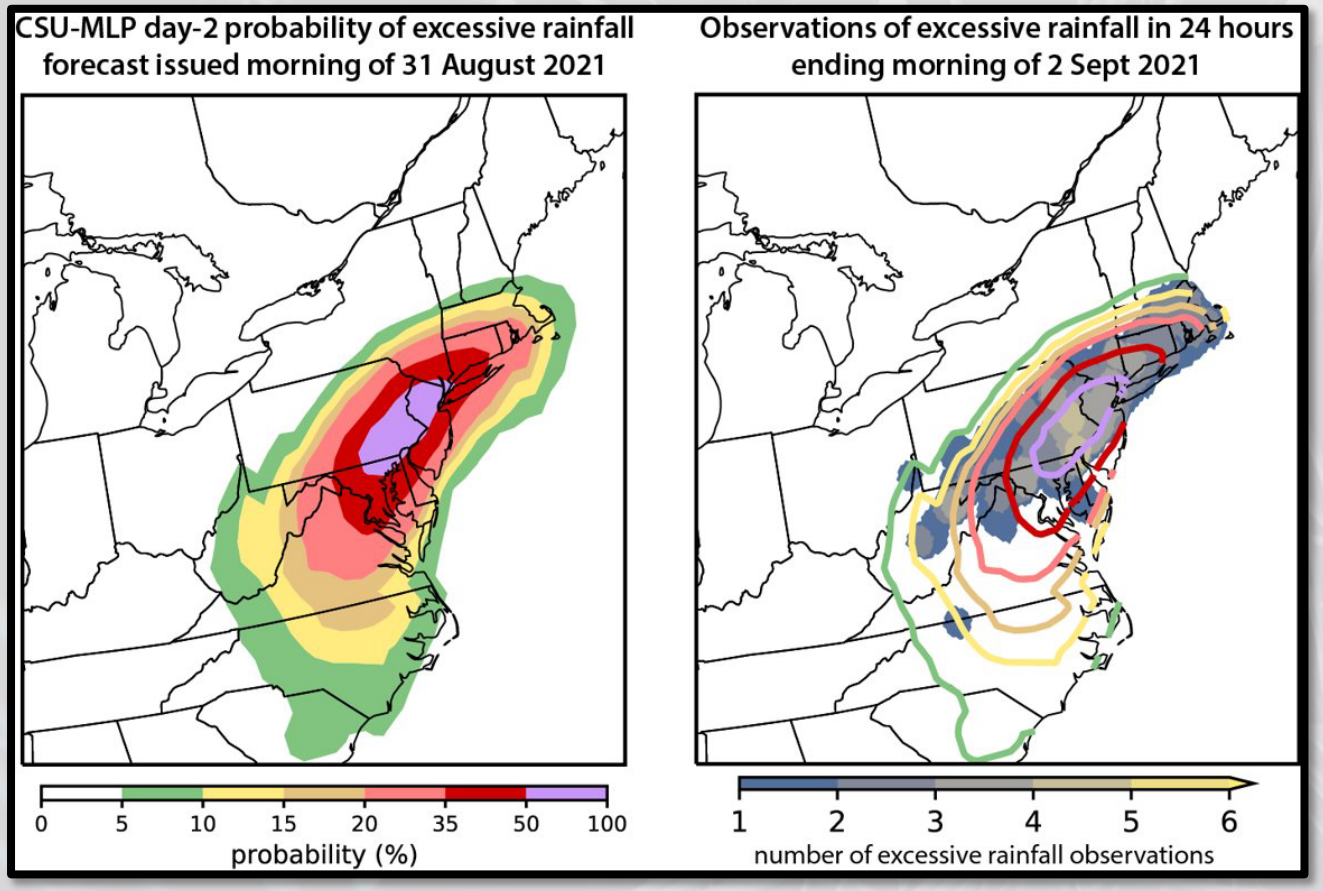
Ingredients-based forecasting at large scales
for day(s) ahead prediction (Johns and Doswell 1992;
Doswell et al. 1996)





Sources: https://inside.nssl.noaa.gov/nsslnews/wp-content/uploads/sites/21/2023/06/unnamed.png, https://www.weather.gov/media/zhu/ZHU_Training_Page/thunderstorm_stuff/Thunderstorms/thunderstorms.pdf, https://cms.accuweather.com/wp-content/uploads/2023/03/SevSetUpTuesWedChan31Mar11a.jpg?w=632

ML can extend (and improve) the prediction



The Conversation (2023): "AI and machine learning are improving weather forecasts, but they won't replace human experts"

High-resolution models still don't offer the necessary resolution to ***resolve*** weather hazards

Proxies for hazards (e.g., updraft helicity) are used to forecast hazard location and storm evolution – an ***implicit*** prediction

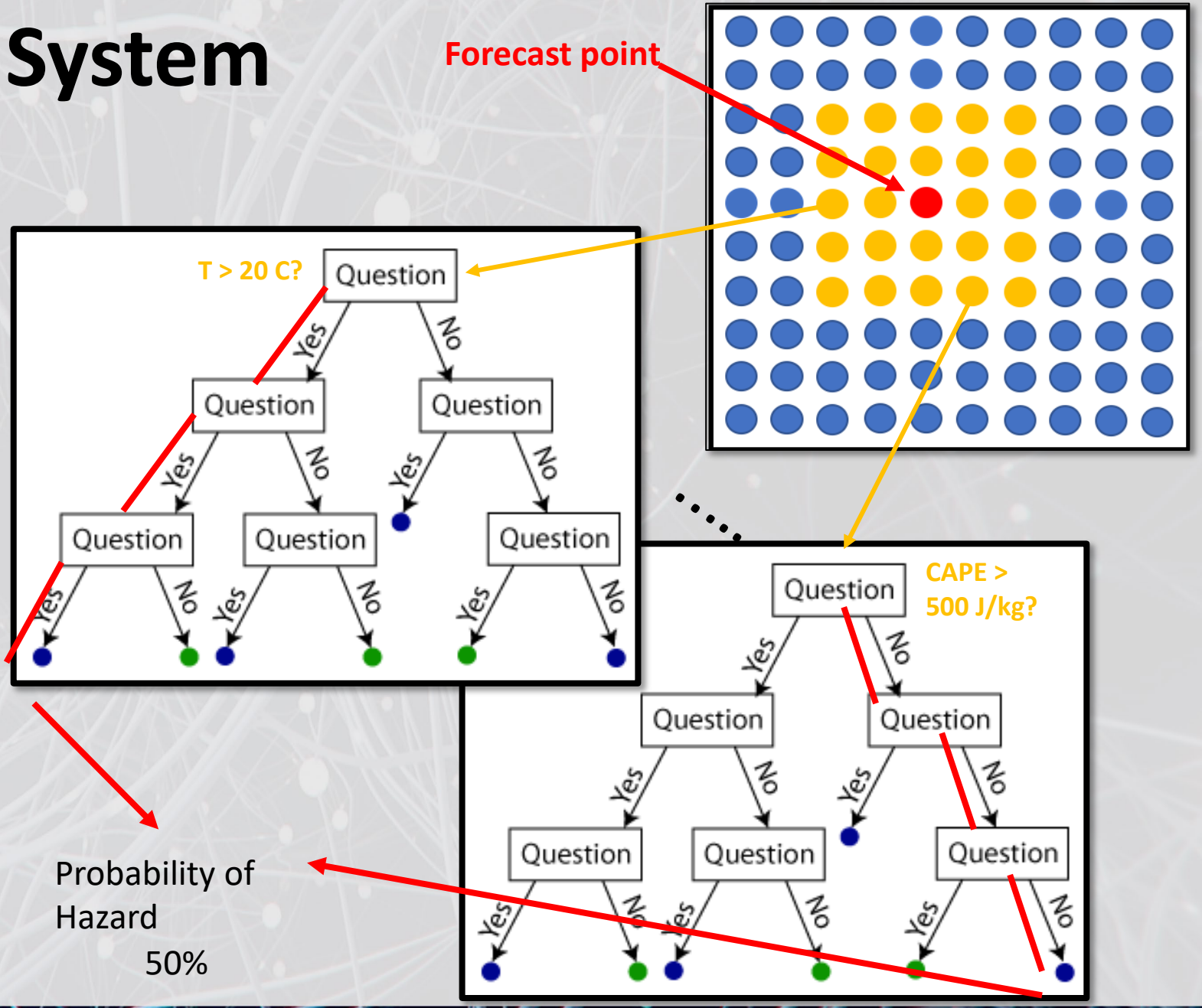
Using historical observations of events, we can train machine learning (ML) models to predict events ***explicitly***

MLP Prediction System

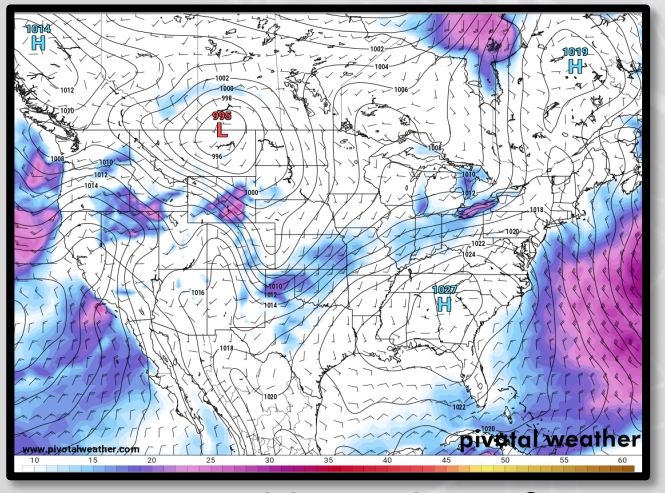
Random Forests (RFs) trained on environments from an ensemble forecast system and historical observations of excessive rainfall, tornadoes, severe hail, and severe wind.

RFs: A series of questions that determine environmental conditions favorable for hazards over a large historical record

Forecasts constructed to mimic operational outlook products – "first-guess" forecasts, 24-hour probabilistic predictions of a hazard



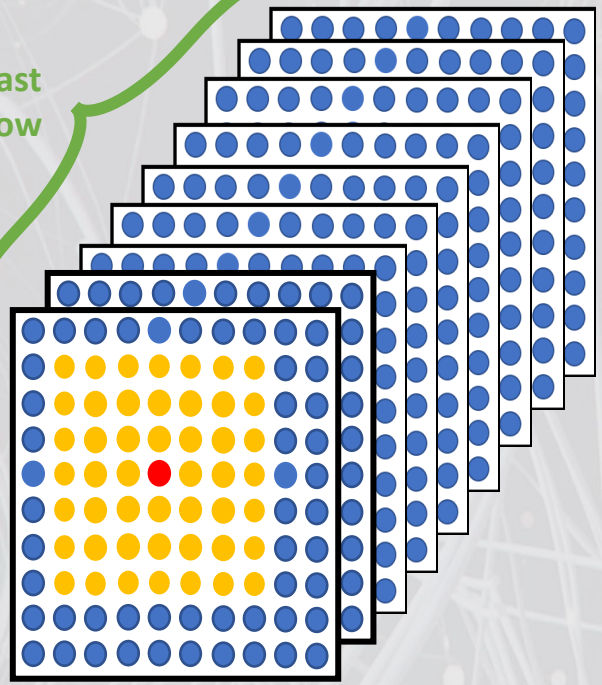
MLP Prediction System



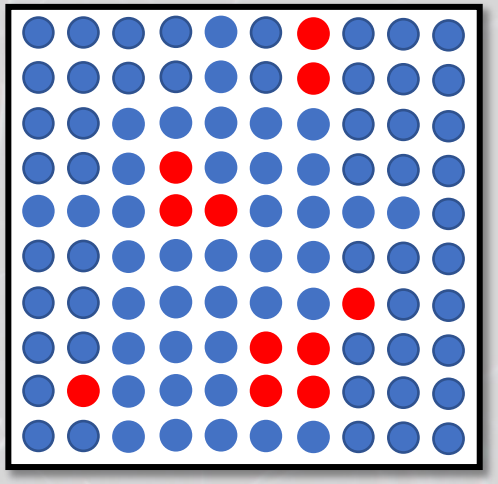
Ensemble median of environmental parameters

Simulated rainfall
Instability
Pressure
Winds
etc.

24 h forecast window



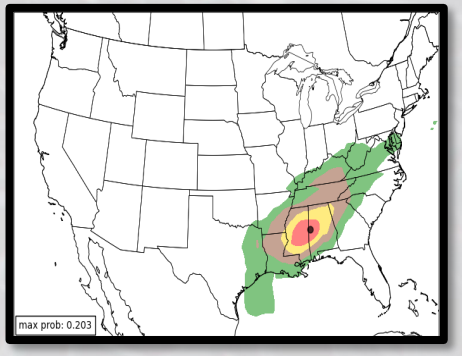
Historical Daily Events
(red event; blue no event)



Real-time predictors

MODEL

Real-time Predictions



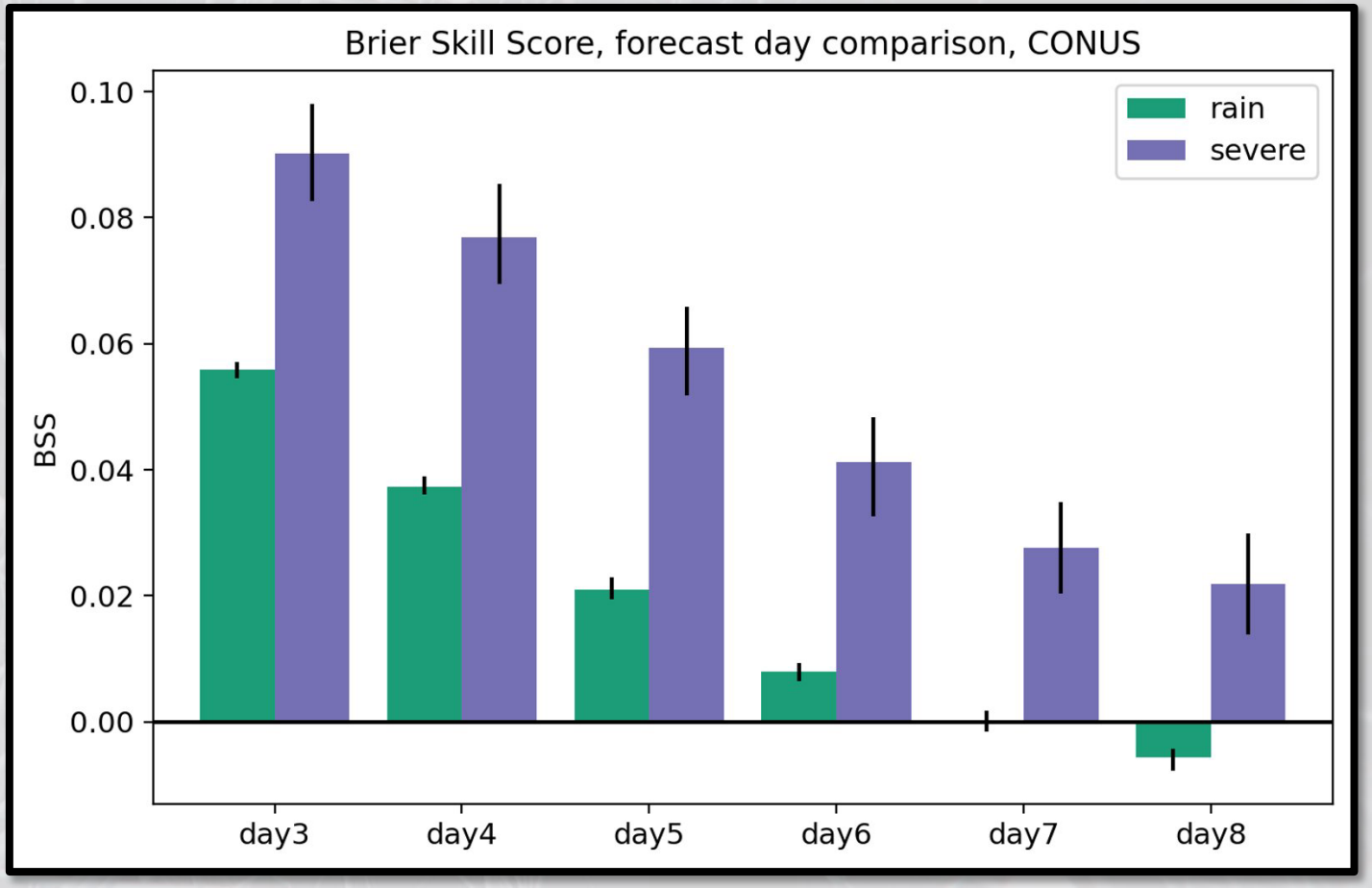
Extending skillful predictions out 6-7 days

Skill as compared to a baseline, temporally varying daily climatology

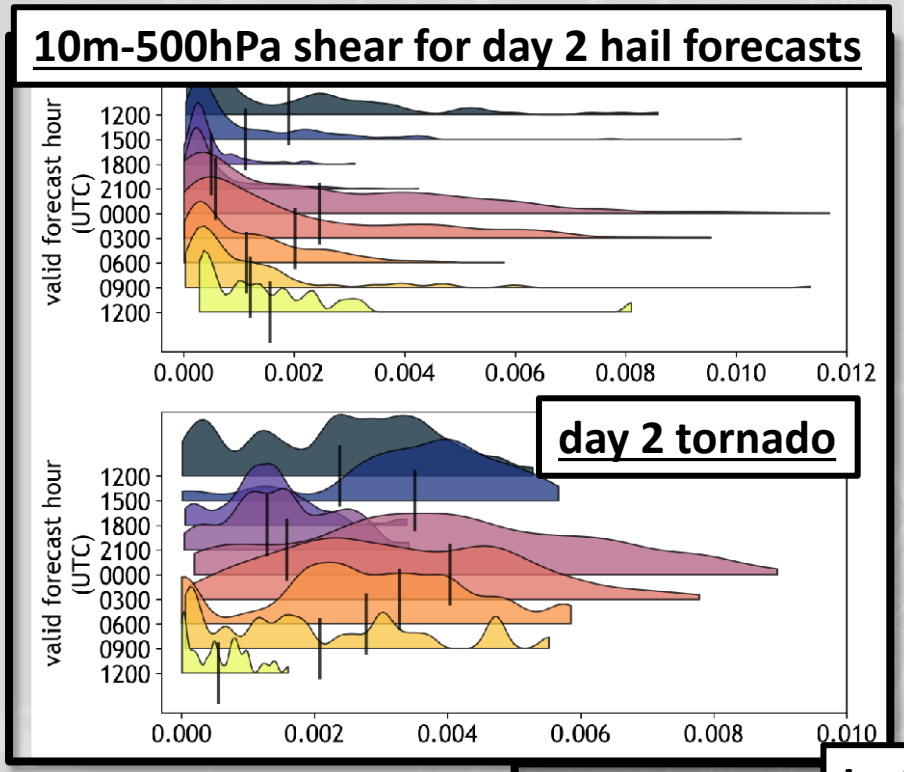
Both prediction systems outperform their human-based outlook counterparts, particularly at day 3 and beyond (Hill et al. 2023, 2024)

All systems now “operational”

To make useful predictions in this framework you have to have clear definitions of the “event”: this is not always the case with rainfall! (Hill and Schumacher 2021, Hill et al. 2024).

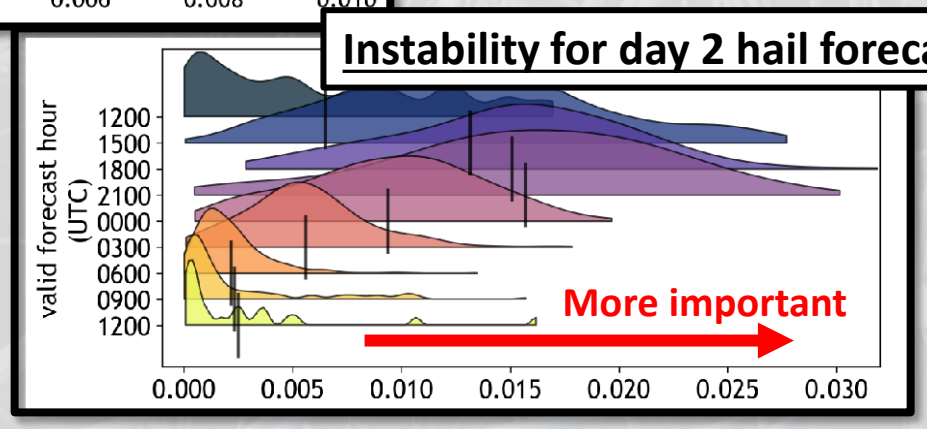


We can explain and interpret the ML models



Despite not trained to do so, the RFs are capable of learning long-standing relationships between environmental parameters and severe weather occurrence – instability and wind shear important for hail production; diurnal patterns of instability

Interpreting and explaining the model forecasts builds ***trust*** with forecasters, especially when the models are using traditional ingredients that forecasters would use



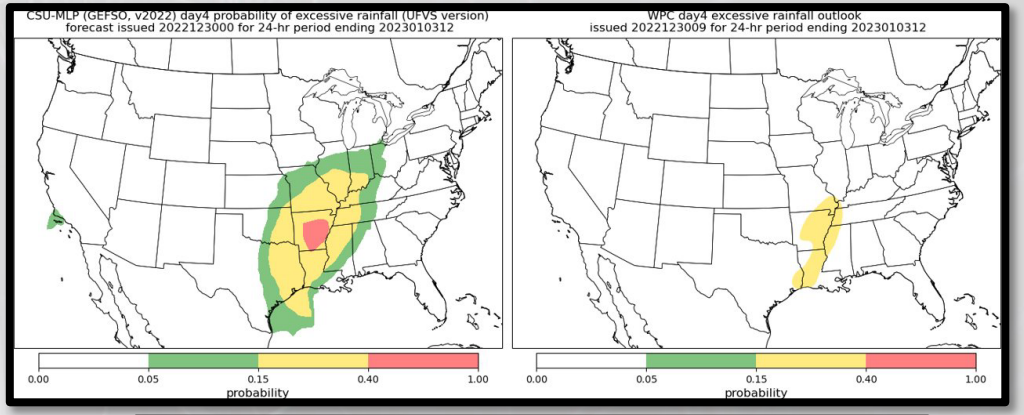
Diurnal cycle

Mazurek et al., 2024: *"Can Ingredients-Based Forecasting be Learned? Disentangling a Random Forest's Severe Weather Predictions"* (in review)

Summary and Discussion

AI and ML are extending our capability to predict hazardous weather much farther than current weather prediction models are capable, out 6-7 days

In practical applications, like weather, AI is simply recognizing patterns in data (those already known or perhaps unknown) to relate environments, shapes of storms, or other special characteristics to hazard occurrences through a rigorous training process



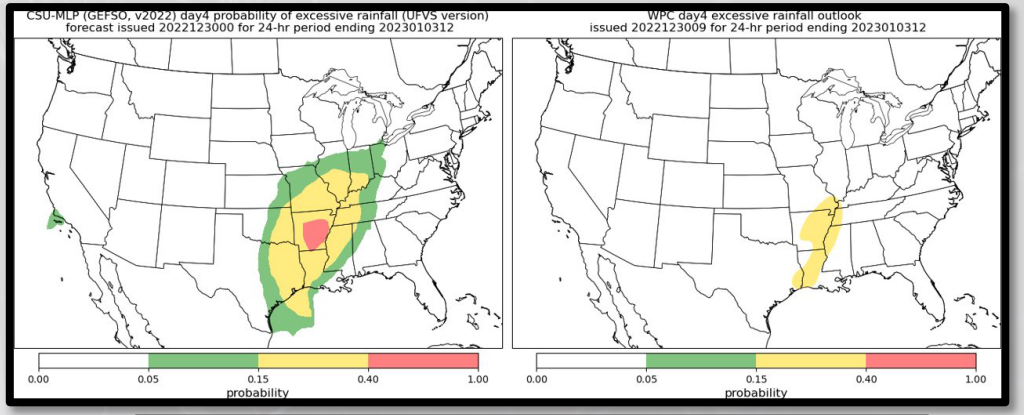
MLP real-time forecast graphics: https://schumacher.atmos.colostate.edu/hilla/csu_mlp

Contact with questions/comments: ahill@ou.edu

Summary and Discussion

One of the most important aspects of AI-based predictions is building trust with end-users (e.g., forecasters) – models and forecasts should be <u>interpretable, explainable, ethical, and trustworthy</u>

Outside of post-processing methodologies like the MLP, other data-driven approaches present ensembles (e.g., 1000 members) as the next step for AI models in weather applications



MLP real-time forecast graphics: https://schumacher.atmos.colostate.edu/hilla/csu_mlp

Contact with questions/comments: ahill@ou.edu