Salient

Probabilistic Decision Modeling using S2S Forecasts

The Importance of Calibration and Reliability to Decision Support System Design

2024 CDPW–CPASW Joint Workshop
Brian Zimmerman - Climate Scientist - bz@salientpredictions.com

Executive Summary

- Humans have a hard time thinking probabilistically
- Good probabilistic forecasts ***need great probabilistic forecast users***
- Effective decision making on S2S timescales requires:
  a. Reliable probabilistic forecasts
  b. Decision support systems
  c. ***Education***
- Synthetic models → intuitive insight → ***constructive dialogue*** with end users

Probabilistic thinking is hard

“The human brain expects more regularity and patterns than randomness actually exhibits” - Bill Notz, OSU

Probabilistic thinking is hard

“The reason people find probability unintuitive and difficult is because it is unintuitive and difficult.” - David Spiegelhalter, Cambridge

Probabilistic thinking is hard

- Utilizing **expected value** to make decisions **is insufficient for comprehensive risk management**
  - This includes classic cost/loss calculations (though these are useful)
- There is a need for a ***more nuanced and educational approach***
  - Education for complex concepts → ***intuitive primers***
  - Synthetic model → ***ask illuminating questions***
  - Engage user in conversation → ***get people thinking***

S2S forecasts - who (do we think) needs them?

Energy

- Load forecasting
- Demand Forecasting
- Renewable production
- Trading and hedging

Insurance

- Better price & underwrite private products
- Provide value-add tools to reduce claims

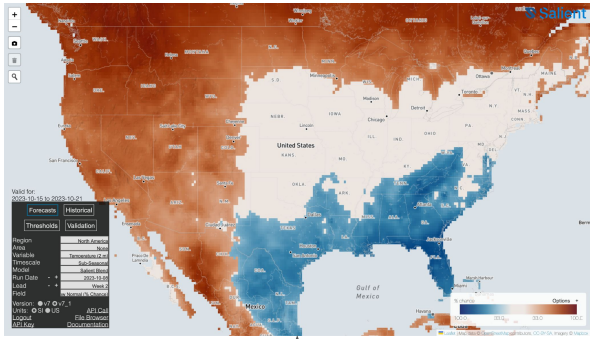
Agriculture

- Optimized crop planning
- Harvest timing
- Pest and disease risk
- Fertilizer applications
- Labor management

The solution Salient provides

Salient Blend forecast - global, reliable, probabilistic

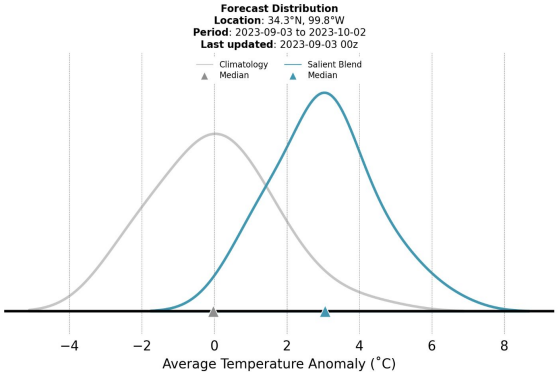
Temperature - Precipitation - Wind Speed - Solar Insolation		
Weekly to 5 weeks	**Monthly** to 3 months	**Quarterly** to 1 year
Updated weekly	Updated Weekly	Updated Monthly
Global (1/4° grid / 25km) and point-based debiasing		



API ACCESS

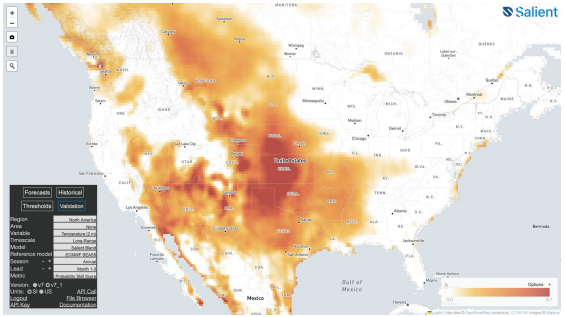
S2S forecasts require attention to detail

Probabilistic



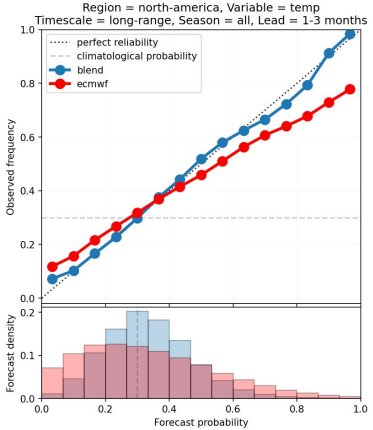
- Salient ML/AI/Clim models operate in quantile space
- Quantify uncertainty
- Can make specific forecasts

Proper Scores



- Proper Scoring metrics (CRPS) are critical for fair evaluation
- Skill scores allow easy comparison among models
- Broken out by timescale, season, and lead time

Reliability



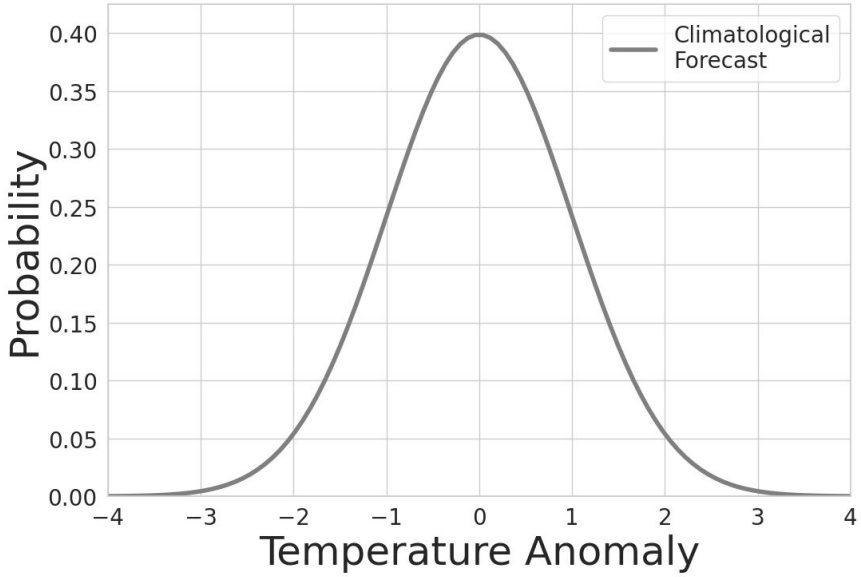
- All models are calibrated
- Reliability diagrams available for above/below tercile categorical forecasts

Intuitive Primers for Probabilistic S2S Concepts

- Tercile Forecasts
- Continuous Ranked Probability Score
- Categorical Metrics
- Reliability Diagrams
  - Effects of forecast calibration
- Other ideas?

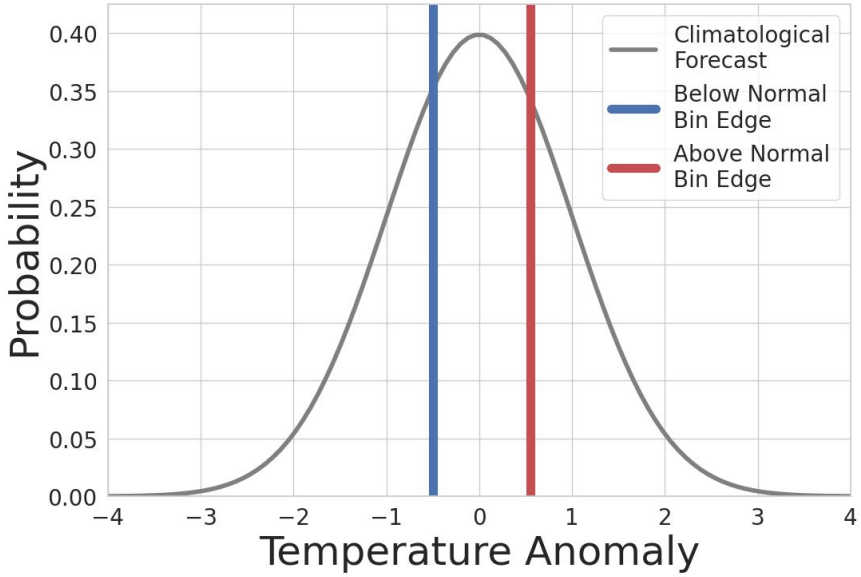
Probabilistic Tercile Forecasts

Typical climate forecasting formulation. Three categories - Above, Near, and Below Normal. By definition, the climatological ‘random chance’ forecast is a 33% of being in any one of the three categories



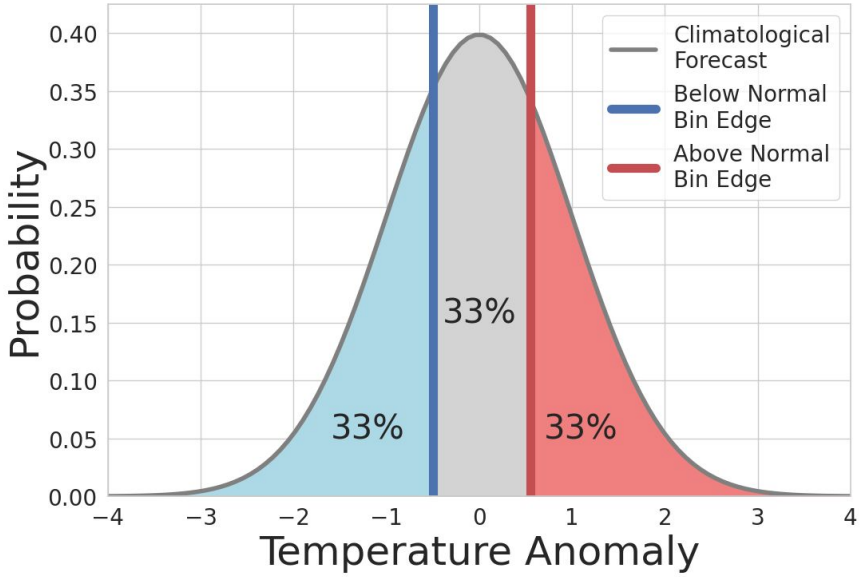
Probabilistic Tercile Forecasts

Typical climate forecasting formulation. Three categories - Above, Near, and Below Normal. By definition, the climatological ‘random chance’ forecast is a 33% of being in any one of the three categories



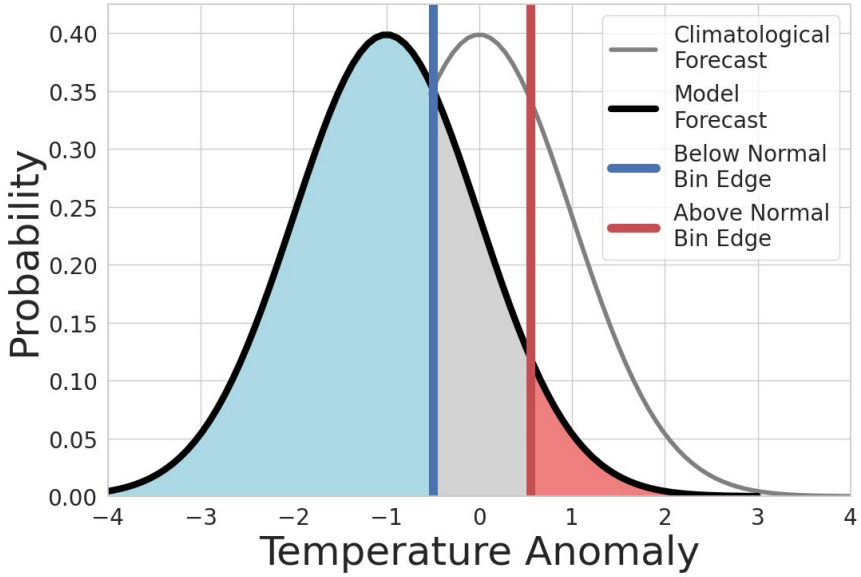
Probabilistic Tercile Forecasts

Typical climate forecasting formulation. Three categories - Above, Near, and Below Normal. By definition, the climatological ‘random chance’ forecast is a 33% of being in any one of the three categories



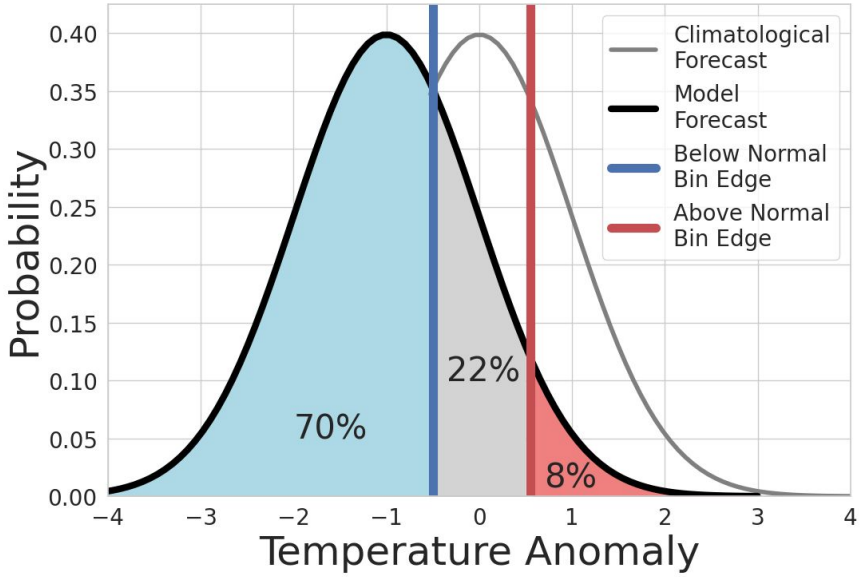
Probabilistic Tercile Forecasts

Typical climate forecasting formulation. Three categories - Above, Near, and Below Normal. By definition, the climatological ‘random chance’ forecast is a 33% of being in any one of the three categories



Probabilistic Tercile Forecasts

Typical climate forecasting formulation. Three categories - Above, Near, and Below Normal. By definition, the climatological ‘random chance’ forecast is a 33% of being in any one of the three categories



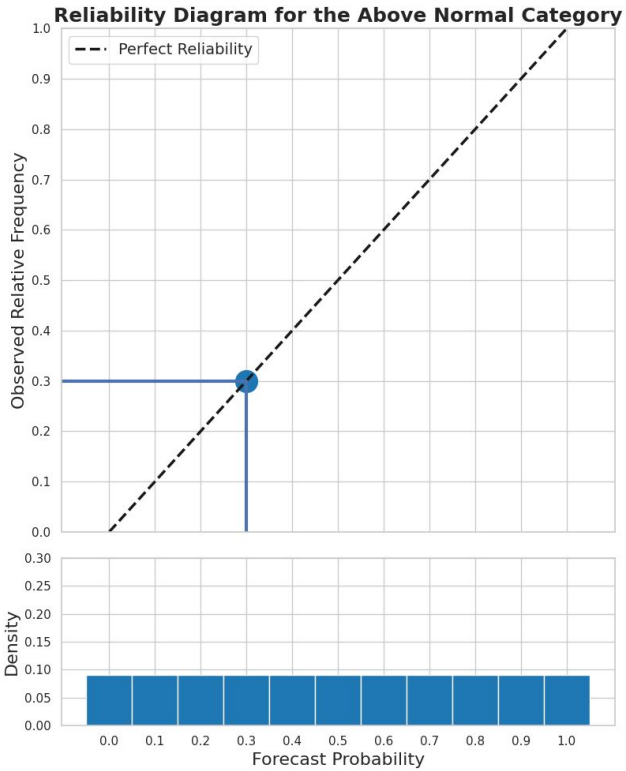
Reliability Diagram

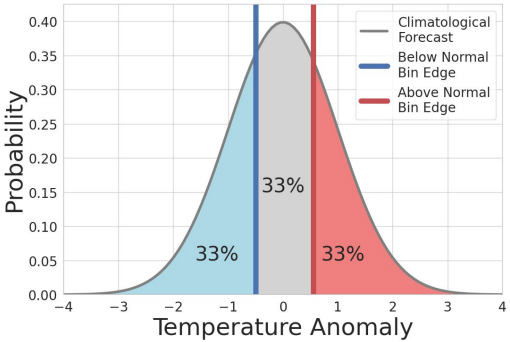
Histograms - Forecast Count

Shows the relative number of times a forecast at a certain probability levels occur - typically aggregated over region and season

Typical Three Category Forecast

33% forecast probability is “random chance” or “climatological” forecast - there is an equal chance of being in any one of the three terciles.





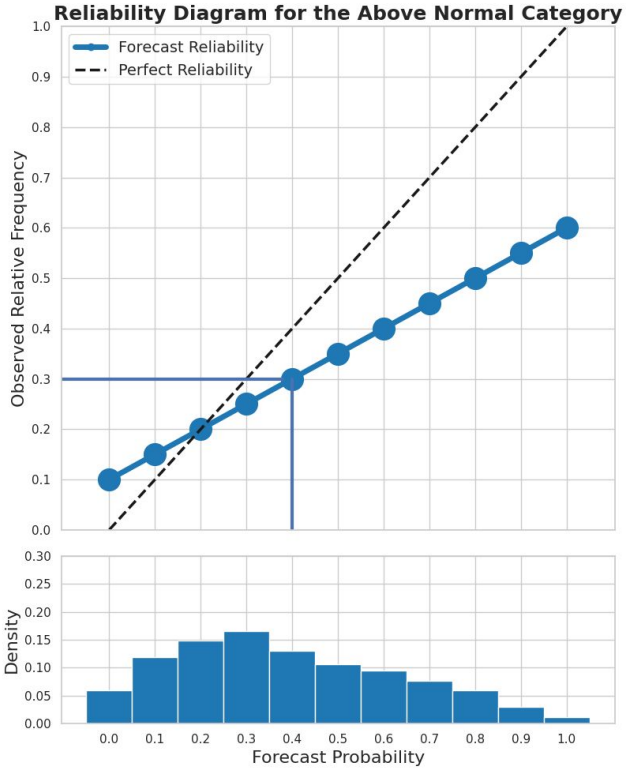
Reliability Diagram

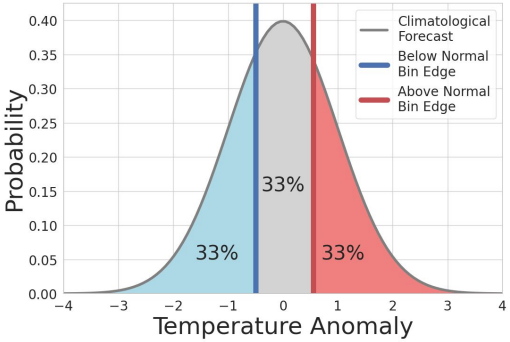
Histograms - Forecast Count

Shows the relative number of times a forecast at a certain probability levels occur - typically aggregated over region and season

Typical Three Category Forecast

33% forecast probability is “random chance” or “climatological” forecast - there is an equal chance of being in any one of the three terciles.





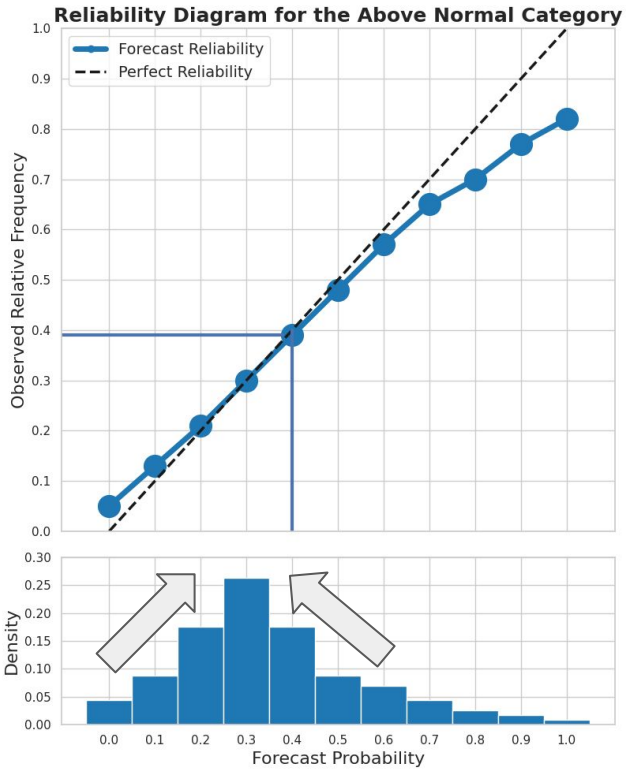
Reliability Diagram

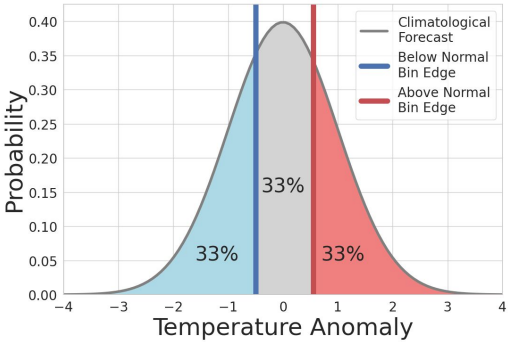
Histograms - Forecast Count

Shows the relative number of times a forecast at a certain probability levels occur - typically aggregated over region and season

Typical Three Category Forecast

33% forecast probability is “random chance” or “climatological” forecast - there is an equal chance of being in any one of the three terciles.





Synthetic Models - An Exploratory Approach

- What does using a DSS look like?
- How does forecast reliability affect user outcomes?
- What level of risk am I prepared to take?
- Questions one can answer with synthetic models:
  - What are the odds of being wrong **N** times in a row?
    - How does the reliability of a forecast affect this?
    - How does choice of categories affect this?
- This is not separate or antithetical to the cost/loss framework, but complementary

From forecasts to decisions

Understanding probabilistic forecasts facilitates better decision making

Probabilistic decision modelling using S2S forecasts

Decision support systems: the importance of calibration and reliability

Brian Zimmerman, Climate Scientist | bzimmerman@salientpredictions.com | www.salientpredictions.com

Calibration and Reliability

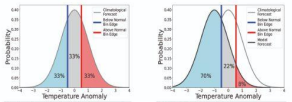


Fig 1: Climatological Tercile Forecast (i.e. equal chances)

Fig 2: Salient Tercile Forecast

What is reliability?

A reliable forecast is one where the forecast probabilities correctly correspond to the frequency of verification. A forecast with 90% confidence should be correct 9 out of 10 times.

What does calibration do?

For ***dynamical models***, reforecast data and validation data are used to calculate parameters used to correct the forecast.

For ***Salient's AI model***, loss penalties to encourage calibration are applied during model training.

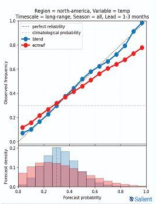


Fig 3: Reliability Diagram comparing Salient Blend to ECMWF

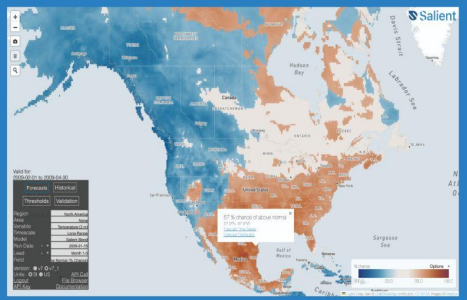
Non-Technical Summary

An uncalibrated model forecast is like a friend who tells you that the next coin flip they make has a 70% chance of being heads.

Calibration is you knowing (from experience) that it's only 50% and betting accordingly.

For the **best odds** at success in an **uncertain world**, two things are necessary:

reliable **probabilistic forecasts** and a **decision support system**



Salient

Let's chat about how reliable probabilistic forecasts could help inform your decision making process

Idealized Agriculture Case Study

A farmer uses a decision support system (DSS) and seasonal probabilistic tercile temperature forecasts to make decisions. Reliability can be turned on and off.

- Farmer has 1000 acres
- Can plant a mix of wheat / barley
- Crop yield affected by which tercile verifies for the season (Table 1)
- Planted acreage depends on the strength of the forecast (Table 2)

Table 1: Harvest Yield Multiplier

	Above Normal	Near Normal	Below Normal
Wheat	1.5x	1.0x	0.5x
Barley	0.5x	1.0x	1.5x

Table 2: Planting DSS

P(Above Normal)	Wheat Acreage
> 90%	900
90% - 80%	800
80% - 70%	700
70% - 60%	600
60% - 40%	500
40% - 30%	400
30% - 20%	300
20% - 10%	200
< 10%	100

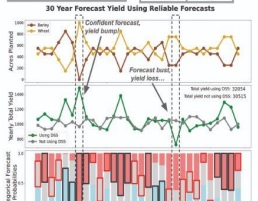


Fig 4: One Realization of a Case Study Simulation - Reliability ON

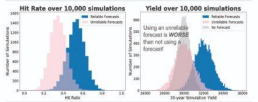


Fig 5: Histograms, Forecast Hit Rate (correct forecasts / total forecasts)

Fig 6: Histograms, Total Yield

Idealized agriculture case study

A farmer uses a decision support system (DSS) and seasonal probabilistic tercile temperature forecasts to make decisions. Reliability can be turned on and off.

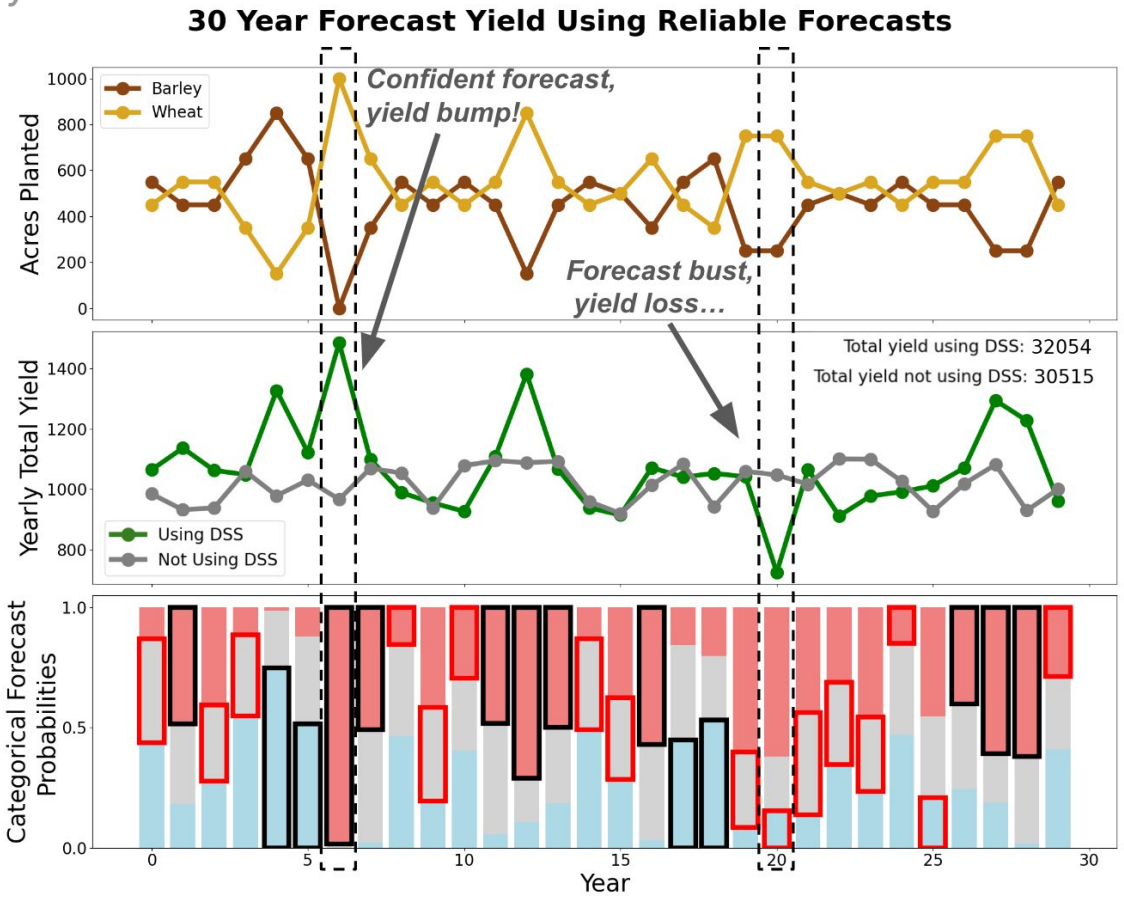
- Farmer has 1,000 acres
- Can plant a mix of wheat / barley
- Crop yield affected by which tercile verifies for the season (Table 1)
- Planted acreage depends on the strength of the forecast (Table 2)

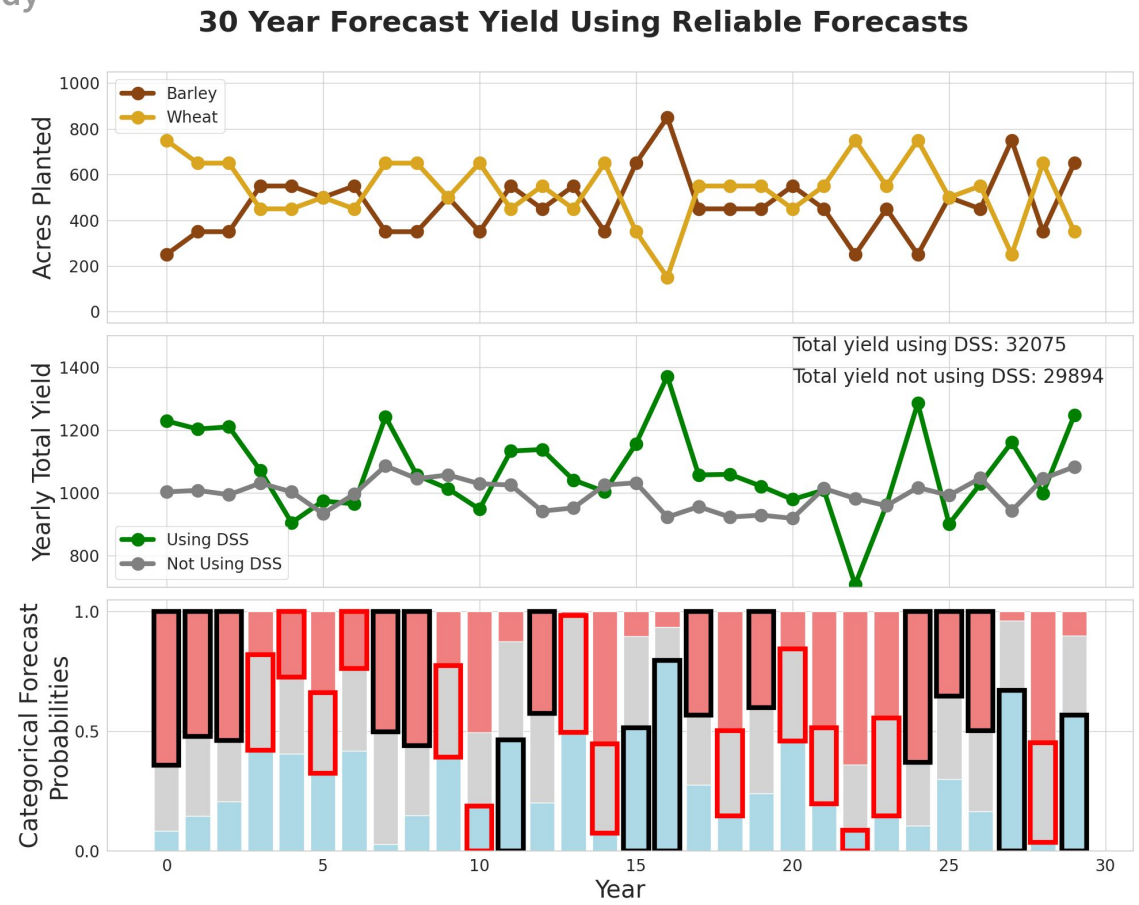
Table 1: Harvest yield multiplier

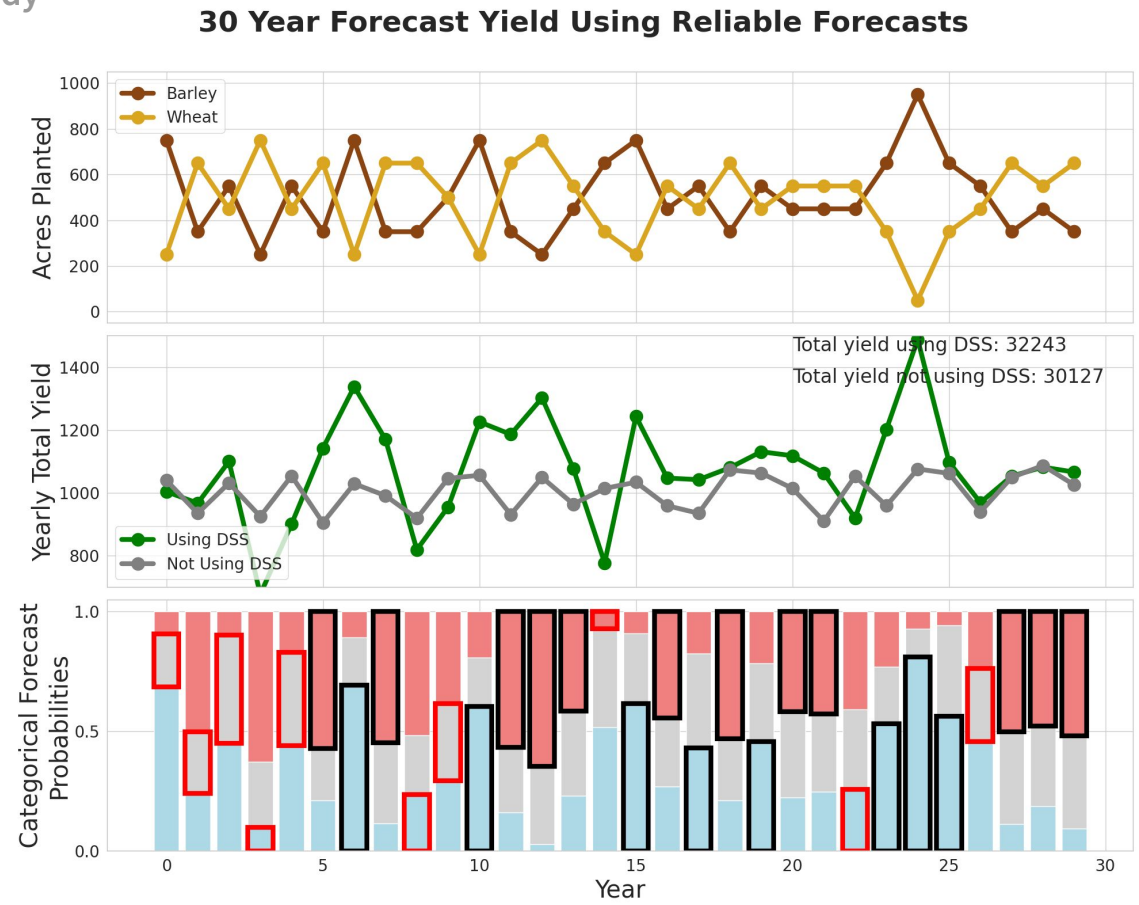
	Above normal	Near normal	Below normal
Wheat	1.5x	1.0x	0.5x
Barley	0.5x	1.0x	1.5x

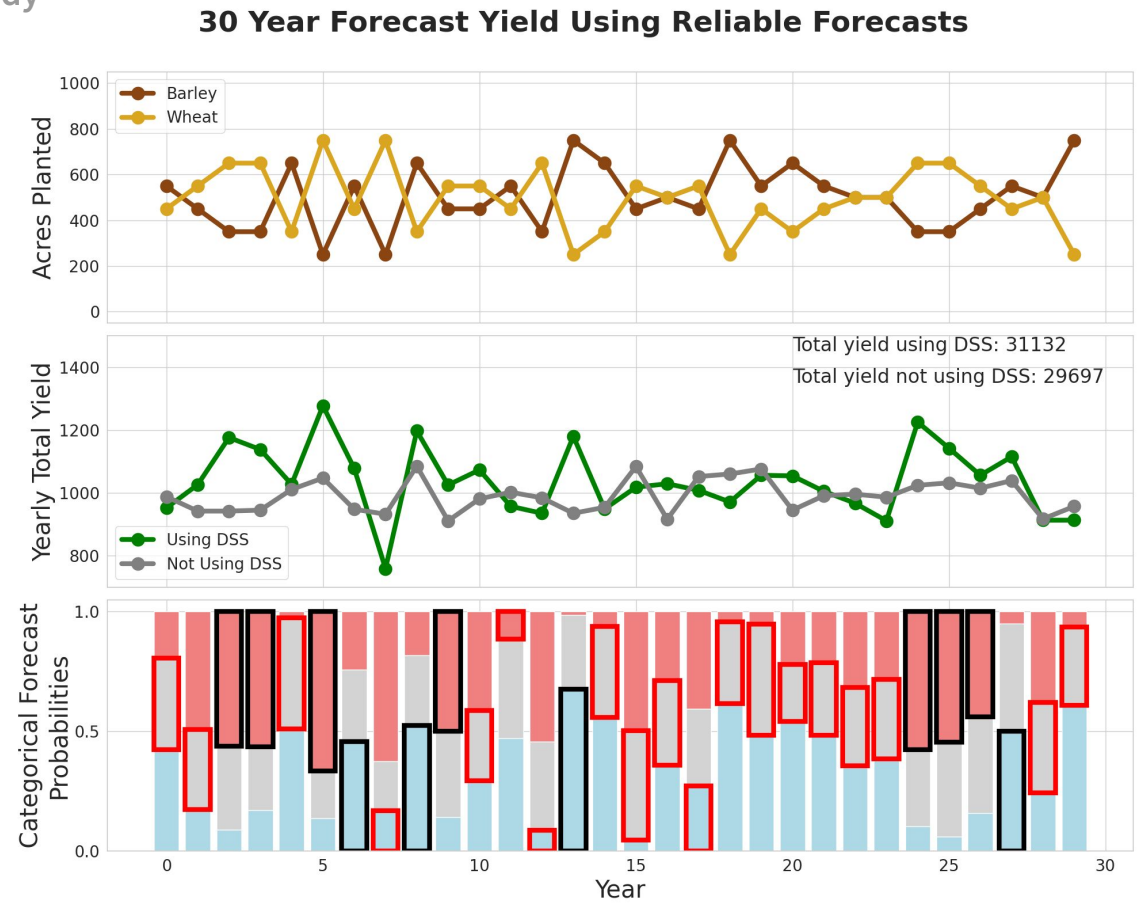
Table 2: Planting DSS

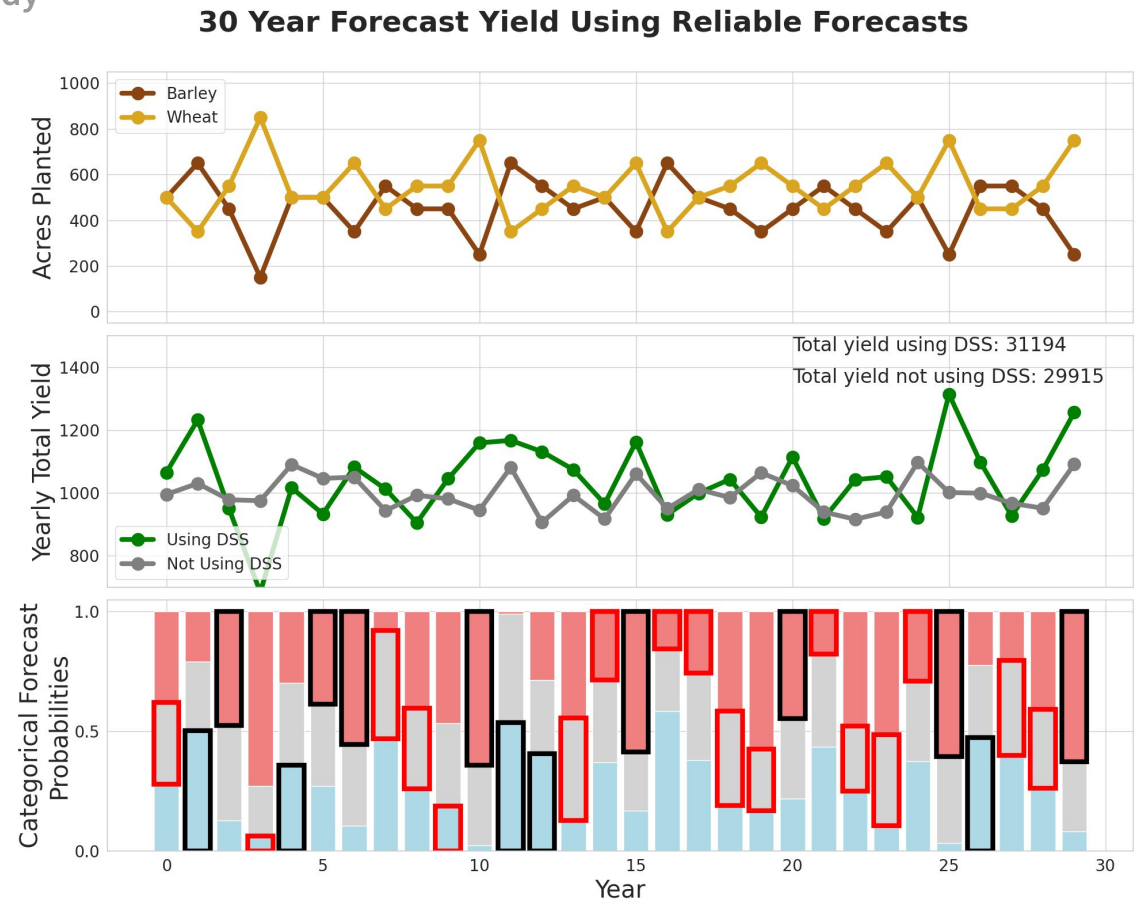
P(Above Normal)	Wheat Acreage
> 90%	900
90% - 80%	800
80% - 70%	700
70% - 60%	600
60% - 40%	500
40% - 30%	400
30% - 20%	300
20% - 10%	200
< 10%	100

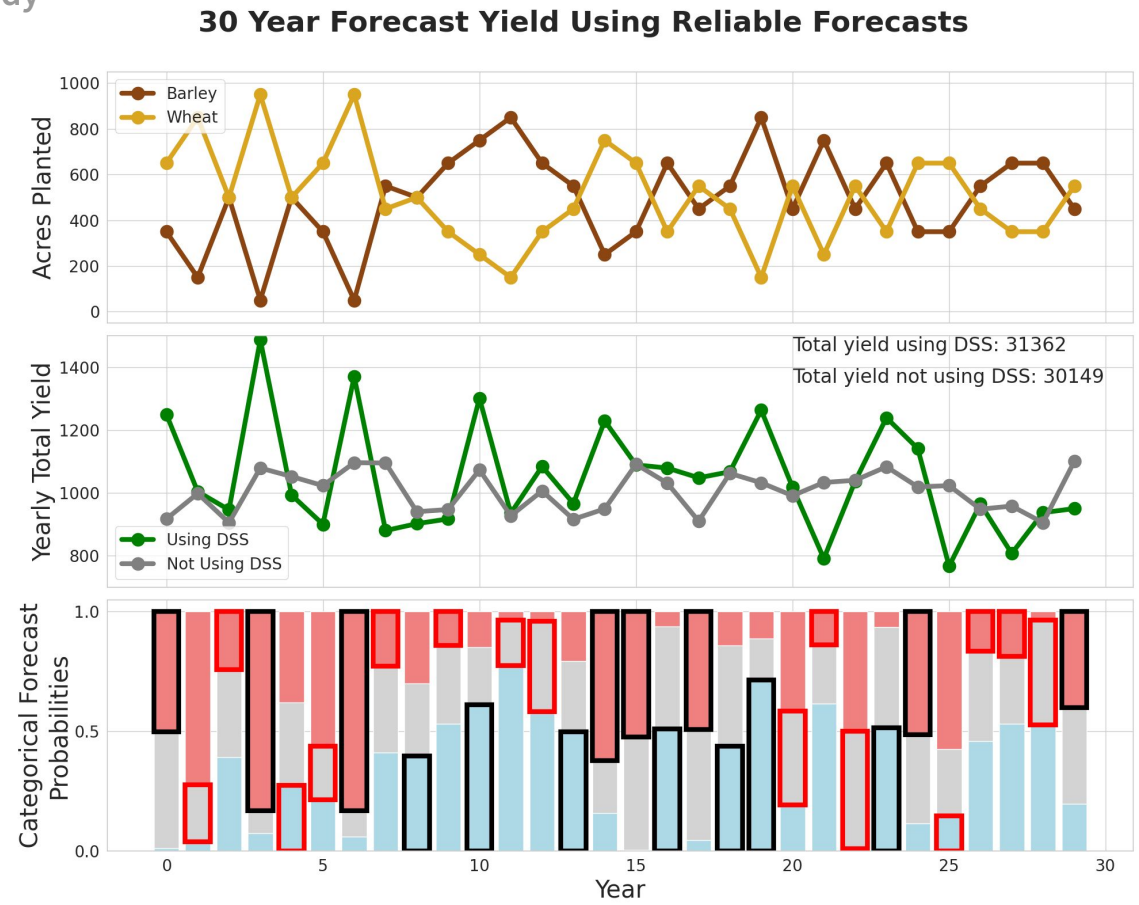


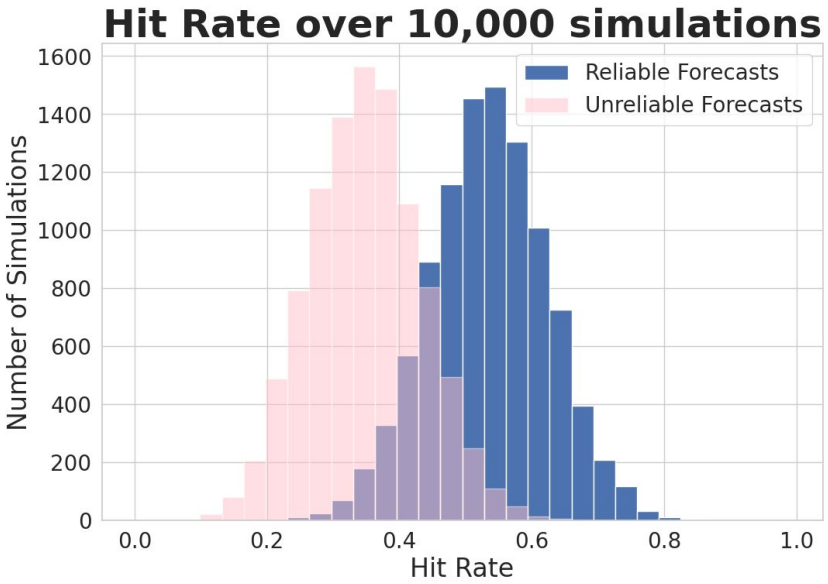


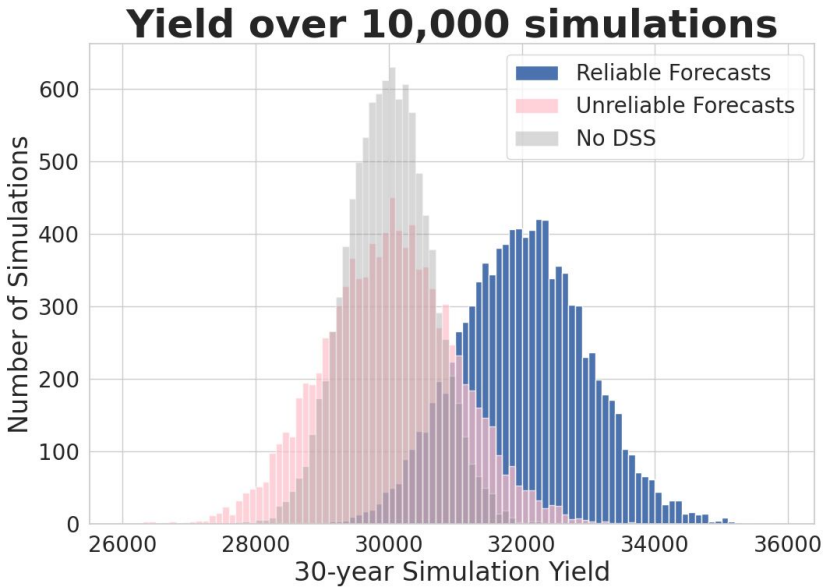


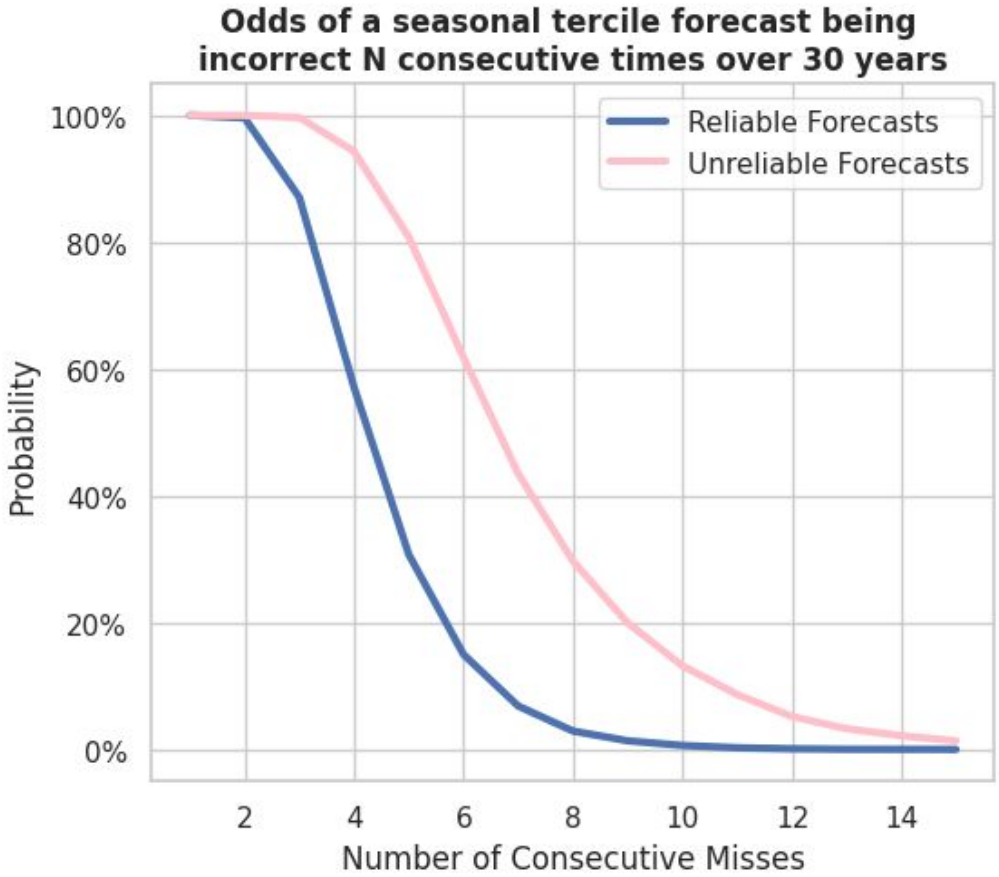


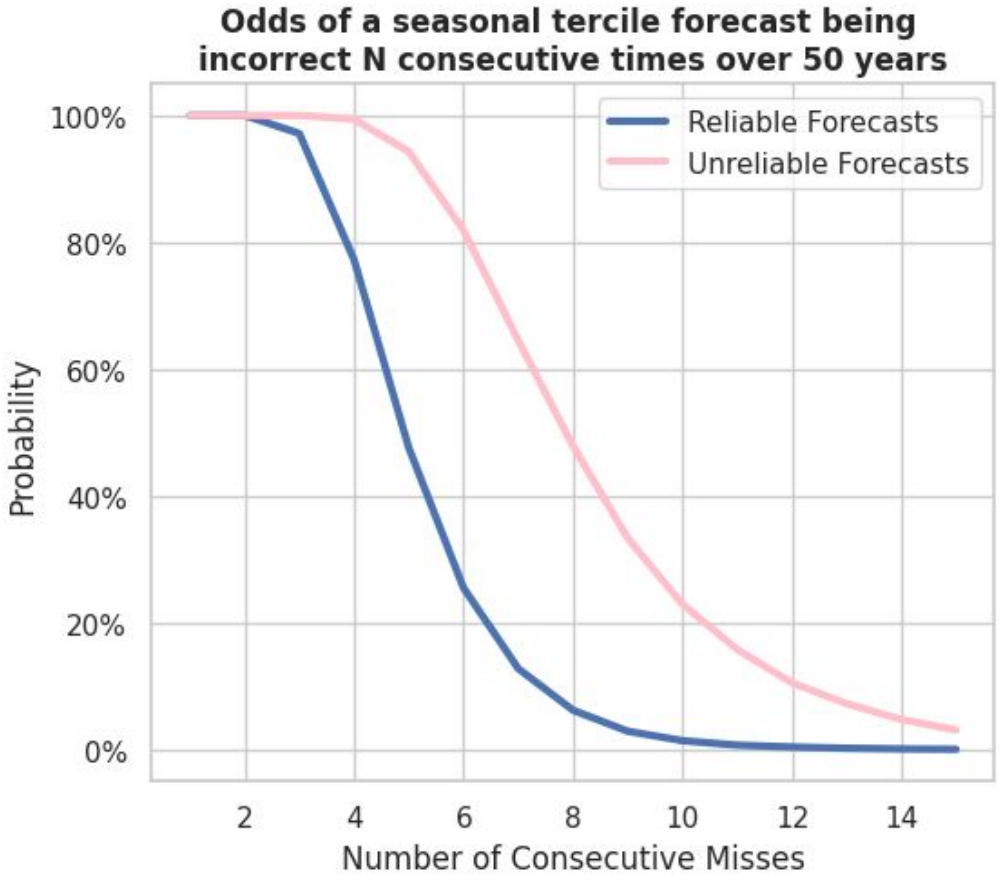


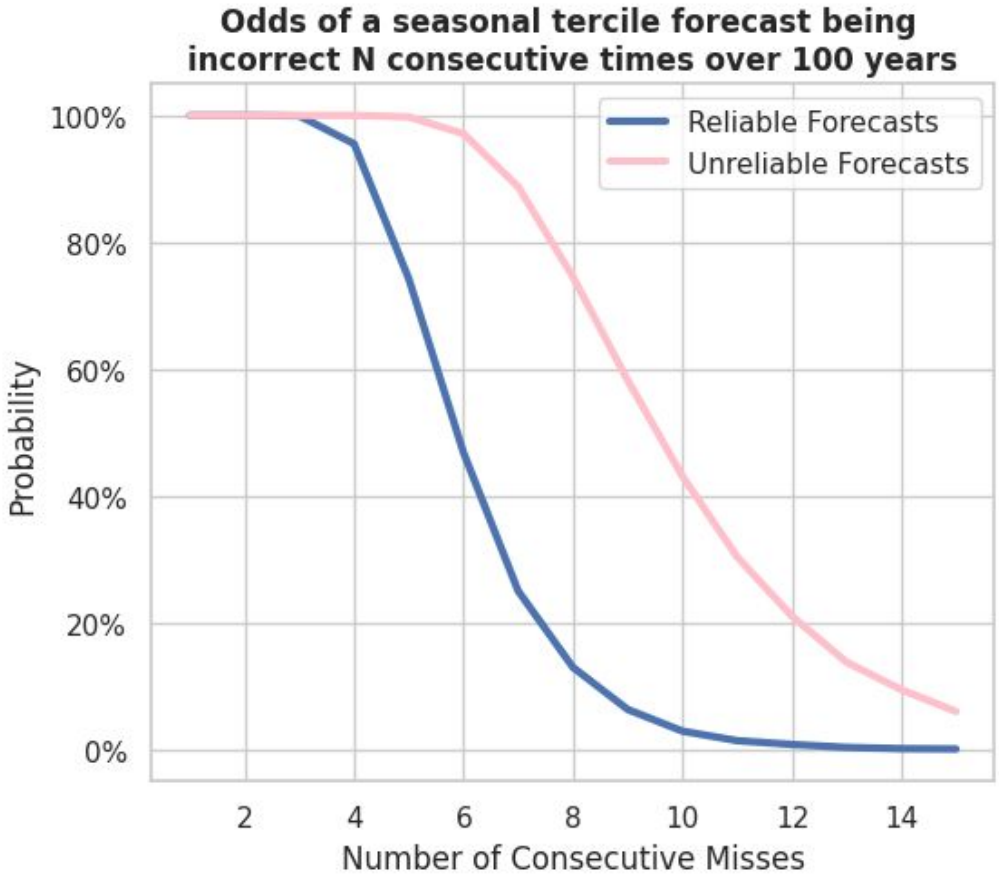








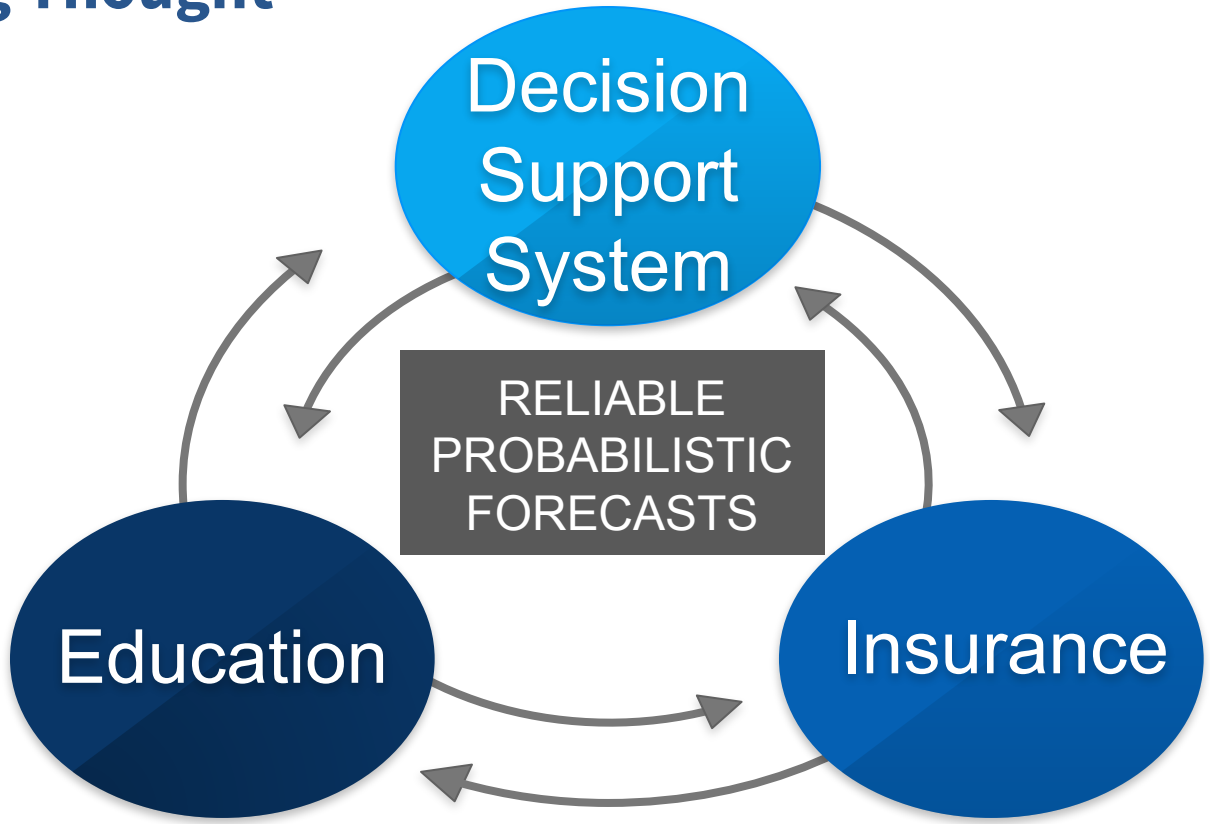




Wrapping up

- Intuitive primers help end users understand probabilistic concepts
- Simple synthetic models can answer interesting questions
- Goal is to ***get the end user thinking about*** how the concepts could be applied to ***their particular situation***

A Parting Thought



Questions?

Thanks! Let’s discuss!

bz@salientpredictions.com