Advancing IDSS Through Probabilities

FY22 Western Region Goals & Guidance

Photo by CDC on Unsplash

NATIONAL OCEANIC & ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE

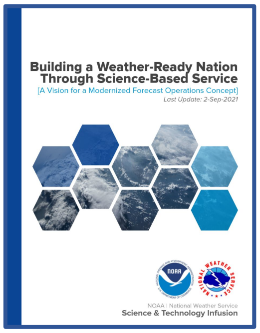
WRH WESTERN REGION HEADQUARTERS

BACKGROUND

BLUF

The science is clear that we need to use probabilistic information to enable targeted messaging for our partners and maximize the value of our information. Probabilistic data can make every kind of deterministic forecast but deterministic data can't make any kind of probabilistic forecast. Effective probabilistic communication avoids words like "possible" and "expected", instead explicitly stating the probability along with or in lieu of an adjective. Proficiency in ensemble thinking, probabilistic understanding, and the harnessing the power of probabilistic information to enhance partner decisions are at the heart of our goals this year.

In 2021 NWS Western Region adopted the Modernized Forecast Operations Concept as the guiding vision for the future of forecast operations. This vision can be summarized as follows: "**If we fully leverage probabilistic information, focus on the things that humans do well, embrace change, and learn what questions our partners really need answers to, we can revolutionize our ability to protect life and property, enhance the economy of the United States and build a Weather-Ready Nation.**" There are three joint (STID/ISD/WFO) goals that we will be pursuing this year to pursue that vision.



GOAL 1

Photo by CDC on Unsplash

WR offices will develop and provide targeted probabilistic messaging with 1-2 partners as a regionally supported activity

WHY?

Modernized Forecast Operations are all about maximizing the decision-making value of our IDSS. Social science research shows that people make better decisions and have more trust in the forecast when they are shown probabilistic information.

WHAT?

The focus will be on developing short term wins with receptive partners who are ready to make decisions based on a richer understanding of scenarios, probabilities, and uncertainty using targeted products.

MORE INFO

1.1 - Am I supposed to make every piece of DSS probabilistic now?

1.2 - Next-level partner engagement with Goal 1

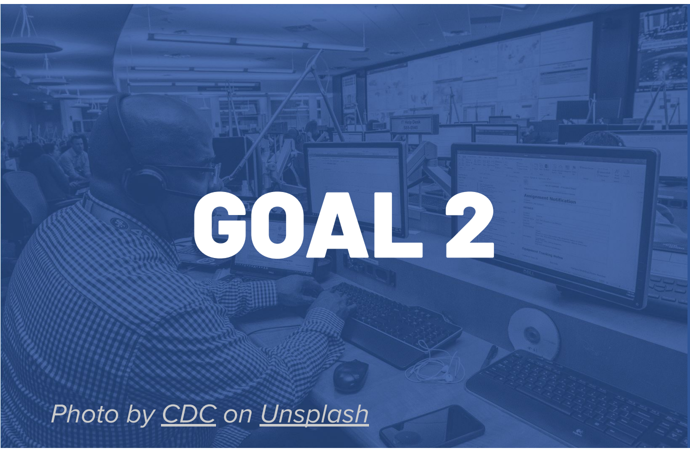


Photo by CDC on Unsplash

GOAL 2

WR offices will become fluent using probabilistic data and tools through routine use

WHY?

Probabilistic data can power every kind of deterministic forecast. Deterministic data can only power one forecast. The foundation of our IDSS must be probabilistic to give all our partners the specific info they need.

WHAT?

The focus will be on the Weather Risk Decision Matrix, which requires probabilistic thinking to populate the likelihood criteria, and providing routine partner emails.

MORE INFO

2.1 - What tools should I use for assessing likelihood in IDSS Consistency Phase 2?

Photo by CDC on Unsplash

GOAL 3

WR offices will message using numbers or numbers & adjectives, versus just qualitative terms, in our IDSS product suite

WHY?

The science is clear that unqualified words of estimative probability (e.g. “possible”) are dramatically reducing the quality of the information we provide to our partners and the public.

WHAT?

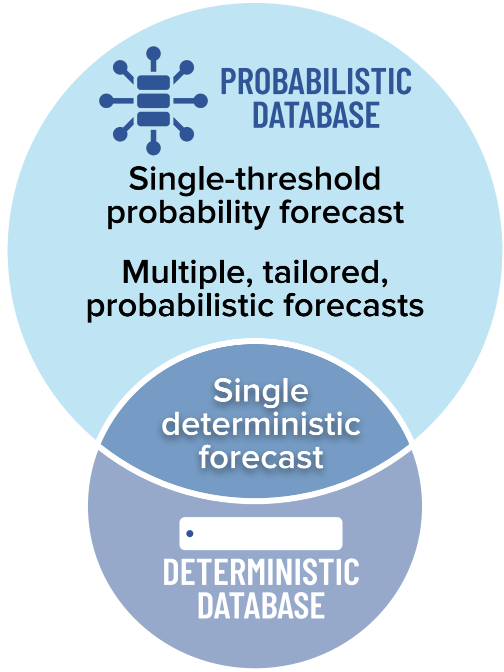
The focus will be on introducing more quantitative language into the products we control.

MORE INFO

3.1 - What is the big deal with numbers instead of just words?

1.1 - Am I supposed to make every piece of IDSS probabilistic now?

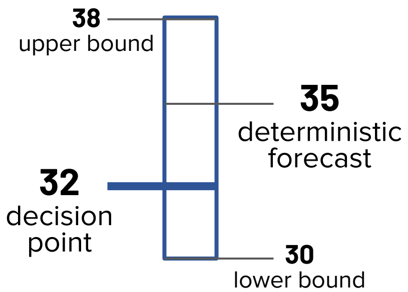
The goal of shifting our operational process to be probabilistic is not to make every piece of information we put out a predictive interval or probability of exceedance. **But in a world where we target our information to specific partners' needs, we always need our tools and our mental model of the atmosphere to be probabilistic regardless of the format that eventually comes out at the end.** When the range of outcomes is very narrow, that single number can do a good job. When the range of outcomes is wide or the partner need is very sensitive, a single number is often woefully inadequate.



EXAMPLE | **ROAD ICING** (Joslyn & LeClerc, 2011)

Using a 32°F decision point for salting roads...

Non-experts make **better decisions** and have **higher trust** in forecast information when provided probabilistic forecasts (probability of freeze) vs a deterministic forecast.



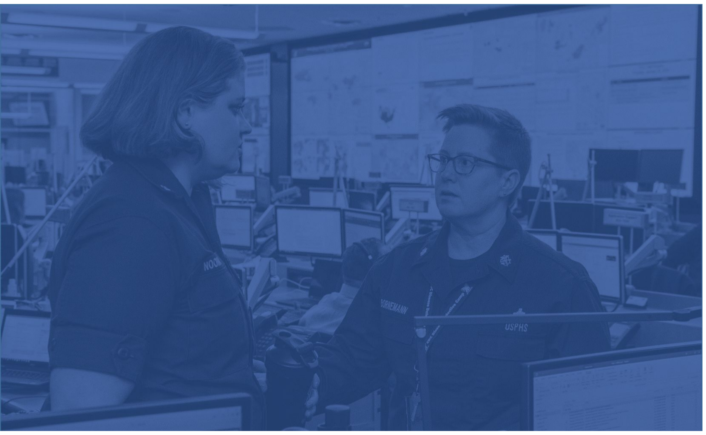
In a series of experiments, non experts were tasked with deciding to salt roads or not using low temperature forecast information. There was a cost to salting, and a penalty for not salting when needed. Groups given probabilistic information made better decisions than the groups given the deterministic low forecast alone.

Does the range of solutions in the forecast cause different outcomes for the user?

When we go deep with our partners this year to assess their needs, this is one of the questions we should be thinking about.

- Does the range of solutions in the forecast cause different outcomes for the user? Probabilistic info is probably best. It is these situations and partners that we will be focusing on with Goal 1.
- Is the partner insensitive to the range of solutions in the forecast? A deterministic number might be just fine. Members of the public who are casual viewers of weather information will often fall into the latter category.

1.2 - Next-level partner engagement with Goal 1



A huge part of executing Goal 1 will be developing or deepening partner relationships such that we fully understand their operational needs. We call this "going deep". You might think "That is what we have been doing for 10 years!". Here we are going to try to illustrate what it looks like to accomplish this goal and how it is different from what we have generally been doing up to this point. Undoubtedly there are examples from the field where all of what we are about to describe has been done. In those cases, great job and keep it up!

What? Temperature?
How much? **Where?**
Wind? Rain? Snow?
What is your threshold?

The questions we've been asking

+

Why? How?

What has happened in the past?

The questions we need to ask

We have long asked our partners about the "What?" and "Where?" of their needs. But it doesn't allow our partners to fully tap into our knowledge as professional meteorologists. We need to ask "**How?**" and "**Why?** as well"

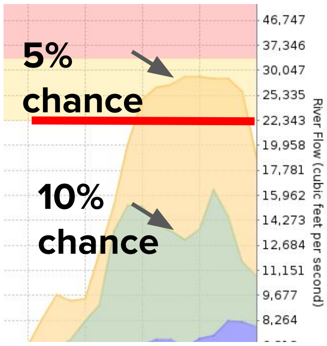
- *Why* is that QPF amount important to you?
- *How* do you use QPF to make operational decisions?
- "Tell me what impacts you saw in *X* event."
- "Can I shadow you on a day of active weather?"
- "How much does it cost you to prepare for this event? How much do you lose if it happens?"

Asking "Why?" and "How?' will allow us to understand what the need truly is and how it can best be met.

Cost/loss assessment is one new way we can improve our service and leverage probabilistic information. At a basic level, cost/loss is simply a ratio of costs associated with taking action ahead of a weather event relative to the loss that will occur if the event happens and action was not taken.

Every user has a set of costs, benefits and risk tolerances that they apply to every decision they make with regard to upcoming weather. Some advanced users might have the actual dollars and cents worked out. Others know generally what their costs and benefits are but haven't put it on paper. Some just follow their gut.

EXAMPLE | **FLOODING**



$5,000
cost to act

$100,000
cost to recover

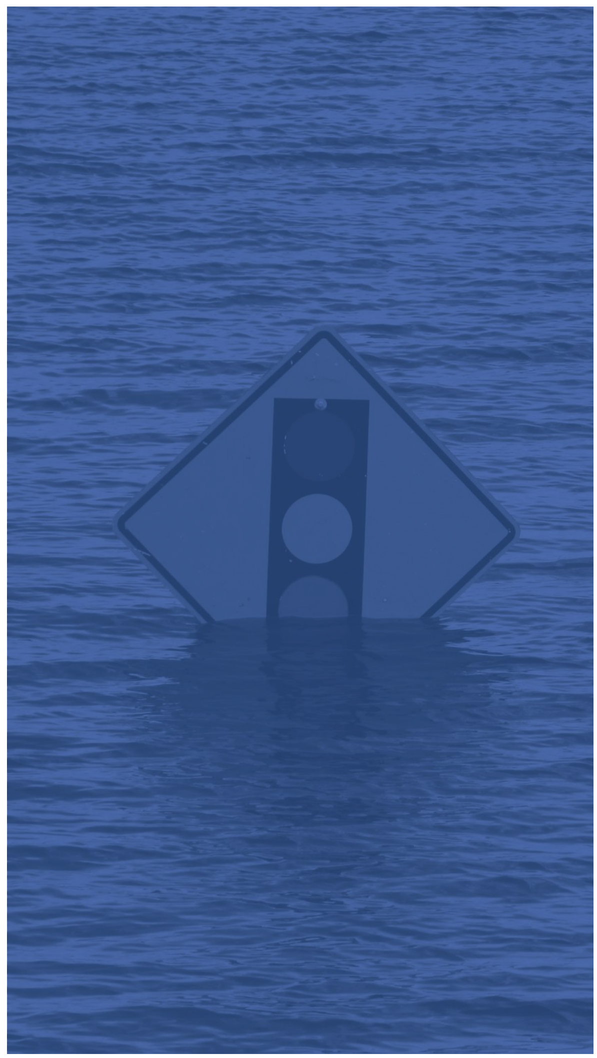
5% probability of occurrence threshold to act

It might cost a city manager $5,000 to sandbag a low spot in a city ahead of a possible flood. But if the river reaches 32 feet, flooding would cause $100,000 in damage. The cost/loss ratio here is 5k/100k=5%. So if the probability of 32 feet being reached is anything higher than 5%, it would be well worth sandbagging! A deterministic river level forecast of 25 feet in this case might actively prevent sound decision-making!

If we put into practice the next-level partner interaction described above, it is possible for us to get a reasonable assessment of the cost/loss ratio of our partners. This cost/loss ratio can be paired with probabilistic forecast information to let the partner makes the go/no-go decision!

Not every situation will be as clear-cut as the example above. **But it is a great example of how, even when the ultimate decision is yes or no, go or no-go, probabilistic information is key to helping the decision-maker decide.** Getting this kind of information to our partners and the public is critical for the NWS achieving our vision of a Weather-Ready Nation.

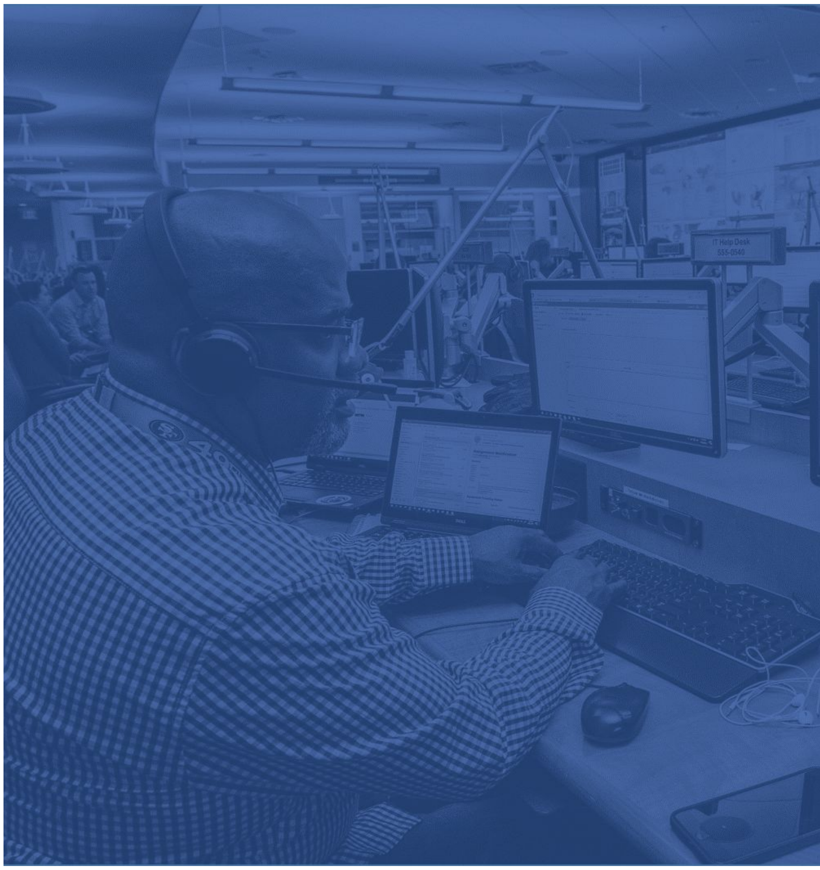
Sometimes it can be tough to get time with your partners. They might not be interested in changing what you give them or maybe they are just busy. In these cases, it may be worth taking advantage of a high-impact event (maybe even a busted forecast that cost them money!) to enter into a conversation where you help them evaluate their cost/loss ratios and therefore actionable probabilities.



GOAL 2

2.1 - What tools should I use for assessing likelihood in IDSS Consistency Phase 2?

It's important to acknowledge that the objective probabilistic output we currently have is imperfect. And some variables have no probabilistic version at all. For these reasons meteorologists will often have to make subjective assessments of likelihood.



It is ok if an estimate of likelihood is subjective. In fact, we have done this with much less and much worse data for decades. But any subjective estimate should be derived from the ensemble information we have.

The complex terrain of WR reduces the range of time and locations that those data can be used without modification. But you can apply the **90-90 rule** that we apply to grid editing for likelihood assessment. Over 90% of the area, and for 90% of the forecasts you make, the NBM is good enough. The other 10% is when you leverage the full ensemble toolset to modify the numbers from the NBM.

Research has shown that even without ensemble guidance, humans can make reliable probabilistic forecasts. That means that if you leverage ensemble tools, you are able to create useful probabilities. Just remember to quantify these probabilities as much as possible. This means avoiding solely words like "chance". While we may associate that word with a 50% probability, our users may think 15%.

> *Nearly all of the studies ... indicate that people make **better decisions**, have **higher trust** in information, and/or display a **greater understanding** of forecast information when shown a probabilistic forecast instead of a deterministic one*

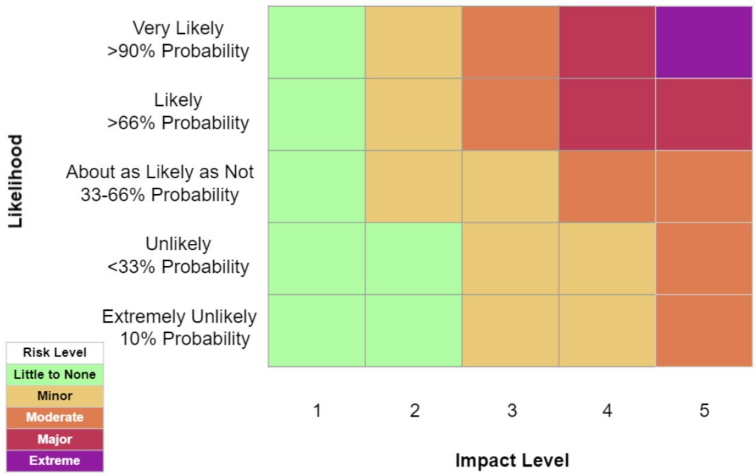
We also know that the social science literature is nearly unanimous that people make better decisions, have higher trust in information, and/or display a greater understanding of forecast information when shown a probabilistic forecast instead of a deterministic one. So let's do our best to communicate effectively the full range of outcomes our partners face, even when it's not fully objective.

Western Region meteorologists will need to assess the likelihood of a range of weather events as part of WR IDSS Consistency Phase 2. What database is the *best* to assess likelihood for each variable?

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8+
QPF	1,2,4,11	1,2,4,11	1,4,5,6,11	1,3,5,6	1,3,5,6	1,3,5,6	1,3,5,6	1,3,5,6
SNOW	1,2,4,9	1,2,4,9	1,4,6,9	1,3,6	1,3,6	1,3,6	1,3,6	1,3,6
WIND	1,2,4	1,2,3,4	1,3,4	1,3,6	1,3,6	1,3,6	1,3,6	1,3,6
TEMP	1,2,4	1,4,6	1,4,5,6	1,3,5,6	1,3,5,6	1,3,5,6	1,3,5,6	1,3,5,6
RH	4	4	4	*none yet*				
FIRE	2,4,8	2,4,8	4,8	8	8	8	8	8
LTNG	1,2	1,2	1	1	1	1	1	1
SKY	1,2,3	1,2,3	1,3	1,3	1,3	1,3	1,3	1,3
SVR	2,10,11	2,10,11	10,11	10	10	10	10	
WAVES	*none yet*							
SYNOP	3,5,6,7							

1	**NBM** (1-D or WSUP)
2	**HREF** (SPC or EMC)
3	**3rd Party** (Weatherbell/WeatherModels)
4	**UW Ens** (UW)
5	**WPC Clusters** (WPC)
6	**EPS EFI** (NCEP/WPC)
7/8	**GEFS/HDW** (WSUP, HDW)
9	**WPC PWPF**
10	**SPC Outlooks**
11	**CSU MLP** (CSU)

One major question that might come to your mind is what you do when there are multiple scenarios possible with a given storm? For instance, the storm has a <33% probability of of impact level 5 and a >66% chance of an impact level 2. Do you put equal emphasis on both outcomes? For the purpose of selecting a single color to populate colors in the partner email, we recommend that you use the most severe indicated color. Additional background and information on the rest of the Weather Risk Decision Matrix can be found [here](). ISD and STID will be collecting best practices and cases to share with the field through additional calls and information.



Weather Risk Decision Matrix from [IDSS Consistency Phase II Implementation Plan]()

GOAL 3

3.1 - What is the big deal with numbers instead of just words?

There is a large and expanding body of research demonstrating the drawbacks of words like "possible" and "expected" in the communication of weather information.

For instance, [Lenhardt et al. (2020)]() showed that a term like "chance" is interpreted to mean anything from 5-60% with a tail out to 80% in a survey of ~1500 people. This is a similarly wide range to what WR-STID has found in polling of Western Region meteorologists, so even we can't consistently interpret these words among ourselves!

Do you care if your PoP grid says 60% when you think it should be 5%? If yes, then you should also care about removing words like "possible" and "chance" from your communication! There are times when we want to use squishy terms because they will let our forecast look right no matter what happens. But that unfortunately means that the information is largely useless for decision-making as well.

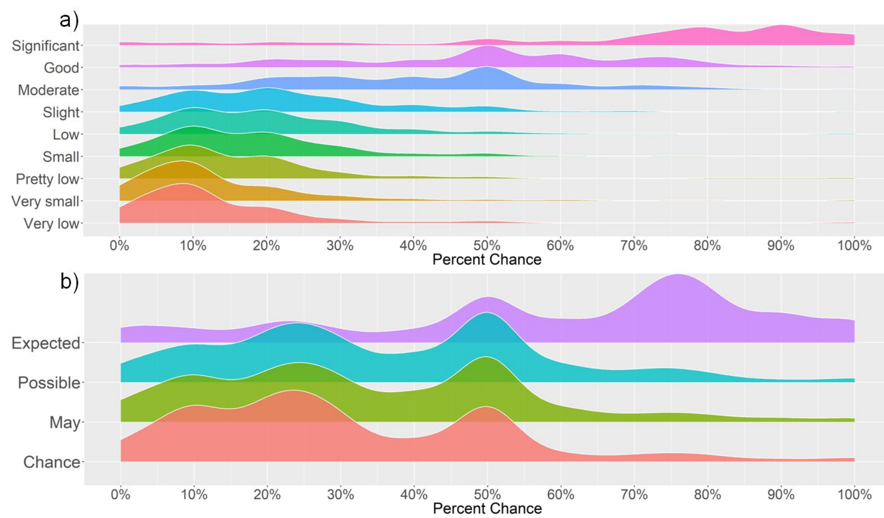


figure from [Lenhardt et al (2020)]()

(continued on next page)

A study of 300+ papers showed that solving this problem doesn't mean that we have to use *only* numbers and no words to express uncertainty or probability.

We can of course just replace words with numbers (e.g. **"60%"**).

If we want to keep the word we should put a numeric translation next to it (e.g. **"chance (60%)"**).

At the very least we should use words like high, medium, or low to modify (e.g. **"low chance"**)

The literature review even found that this rule of using numbers is especially true when comprehension is critical, like with high-impact events. If you need to make sure people understand you, use a number.

Since this is a significant change that feels weird even to us at STID who have been doing it (or trying to at least!) for years, we understand if you find it hard to put into practice. Even if people don't know the definition of PoP, a numeric expression is more likely to elicit the right response. Since the science is clear and the stakes are high, this is something that we are going to pursue together region-wide this year.

Modernized Forecast Operations Concept
https://docs.google.com/document/d/1fqfJ--dycK8fWbCZwrfn4pR-JWhmo5PgV4ehQlxLJpo/edit?usp=sharing

Ripberger et al. ProbCom
https://crcm.shinyapps.io/probcom/#section-executive-summary

Joslyn & LeClerc (2012) Uncertainty Forecasts Improve Weather-related Decisions and Attenuate the Effects of Forecast Error
https://www.apa.org/pubs/journals/features/xap-18-1-126.pdf

Murphy & Winkler (1977) Reliability of Subjective Probability Forecasts of Precipitation and Temperature
https://www.jstor.org/stable/2346866?origin=crossref

Western Region IDSS Consistency Phase 2 Implementation Plan
https://docs.google.com/document/d/15vM0IJniX4EI9yi3LICScSksnZKjepbo4I-8LJwRrMY/edit#bookmark=id.bukf10yh3zjuf

Lenhardt et al. (2020) How Likely is That Chance of Thunderstorms? A Study of How National Weather Service Forecast Offices Use Words of Estimative Probability and What They Mean to the Public
https://objects-us-east-1.dream.io/nwafiles/jom/articles/2020/2020-JOM5/2020_JOM5.pdf