HailDash Weather Model & Forecasting

2026 White Paper

Updated July 20, 2026 | Vantage Weather Technologies LLC

Overview

HailDash provides a continuously updating, high resolution multi-day hourly forecast of hail and thunderstorm risk, covering the continental United States, southern Canada, and northern Mexico. Our proprietary weather forecasting model generates a seamless, multi-class probability grid across the entire forecast domain, which is used by the HailDash web and mobile applications to derive point forecasts, weather alerts, and map layers.

HailDash's forecasts are continuously verified against independent observations. As of July 9, 2026, we have scored ~15 billion forecast-observation pairs from the 2026 season:

- 0.98 AUC discrimination at one-hour lead, still 0.89 at three days
- High-probability cells verify at 200–370× the background hail rate
- 97% of verified severe-hail events occurred where HailDash had flagged elevated probability

We provide an explanation of this data, the methodology behind the verification analysis, and additional performance metrics in the "Forecast Performance & Verification" section.

HailDash supplements, and does not replace, official National Weather Service watches and warnings. In a warned situation, follow official guidance.

Objective

The primary intent of HailDash is to serve the space between official storm warnings (e.g., the National Weather Service's near-real-time Severe Thunderstorm Warnings) and convective outlooks (e.g., the Storm Prediction Center's broad, categorical multi-day outlooks), though HailDash updates quickly enough to provide near-real-time guidance and alerting as well. The system was developed to answer questions like:

- “Do I call in additional staff to prepare our sites for this afternoon's hail threat, and when does that preparation need to be finished?”
- “It’s 1pm and I’m not seeing any storm clouds building. I don’t want to waste resources on hail mitigation. Has the chance of severe hail decreased, or is the forecast still on track?”
- “We have three days to complete work at an exposed site, and thunderstorms are possible on each of them. Which day gives us the longest workable window, and when does it end?”
- “Storms are expected overnight while our lot is unstaffed. What's the chance of damaging hail occurring during that time, and should we spend this evening preparing?”
- “Ten of our sites are inside today's outlook area. Which three actually need action, and in what order?”
- “A Severe Thunderstorm Warning just expired and the sky is still dark. Is the threat over here, or is the next round forming behind this one?”
- “Conditions look identical to last Tuesday, when nothing happened. What's different, and how confident is the model this time?”
- “Should I pay for covered parking at the airport, or is an open lot safe?”

Forecast Model Specifications

Outputs (per grid cell)

- Probability of thunderstorm (specifically, the probability of cloud-to-ground lightning)
- Probability of hail >0.25”
- Probability of hail >0.75” (severe hail)
- Probability of hail >2” (destructive hail)
- 3hr, 6hr, and 12hr aggregations of the above products, if applicable

Temporal

- **Domain:** 0 hours to 72+ hours
- **Resolution:** Hourly across the entire domain

Spatial

- **Domain:** CONUS, southern Canada, northern Mexico
- **Resolution:** ~8km x ~8km

Cycle Tempo / Update Frequency

- Up to every five minutes, enabled by a highly parallelized compute architecture.

Why Probabilistic Forecasts?

HailDash issues probabilistic forecasts to address two distinct kinds of uncertainty. The first concerns the storm itself: whether a particular storm becomes a hail producer depends on fine-scale processes that remain uncertain until close to impact. The second is placement and timing, and it grows to dominate as forecast range increases.

The environment that will produce hail-bearing storms is often predictable a day or more ahead. Exactly where and when storms fire within it is not. At range, forecast error is driven far more by a storm arriving two hours late or thirty miles east than by hail failing to materialize at all. A yes/no forecast collapses all of this into a single answer, and when the storm arrives a county away it ends up wrong twice: a false alarm where hail was predicted, a miss where it actually occurred. A probabilistic forecast carries the uncertainty forward instead, so each user can act on their own costs and risk tolerance.

Probabilities also convey structure that a binary forecast cannot. HailDash forecasts probabilities at multiple size thresholds simultaneously, so a 10% chance of severe hail under an 80% chance of thunderstorms means something different than the same 10% under a 15% chance of thunderstorms. The first is a busy day of mostly weak storms. The second is an otherwise quiet day that might produce one dangerous storm. And because the numbers update continuously, users can watch a threat sharpen or dissolve across forecast cycles rather than flip between yes and no.

HailDash Weather Model Technology

The HailDash model is a hybrid machine-learning (ML) and numerical weather prediction (NWP) forecasting system. Physics-based NWP outputs are ingested by a convolutional neural net (CNN) alongside weather observation and analysis data (e.g. data from radar, satellites, and weather stations).

Hybrid ML/NWP Architecture

The benefits of a ML/NWP hybrid model include:

- Forecasts that are grounded by actual physical parameters due to the use and augmentation of physics-based NWP models. Pure ML models have no concept of physics, which can result in weather predictions that would never be possible in the real world.
- The ability to quickly generate extended range forecasts in response to new data, even with several gigabytes of input data driving calculations across millions of parameters.
- Automatically bias correcting and enhancing NWP prediction data through the use of rapidly-updating observation and analysis data.

Competitor forecast products we have evaluated are typically built on a single upstream source, and in some cases simply repackage it:

- **Radar extrapolation** projects existing storms forward along their current motion. This can only forecast storms that already exist, and forecasts built upon this technique rarely extend past 60-90 minutes. NOAA freely provides storm-track products that are derived from radar observations, and are often repackaged directly into storm or hail forecasting solutions.
- **Single-model repackaging** presents one NWP run's hail-related and/or convective fields. The forecast updates only when that model runs, carries the model's biases uncorrected, and cannot incorporate what radar and satellites are showing between runs. Products that offer forecasts out to 18 hours are likely to

be a repackaged set of parameters that are otherwise freely available from one of NOAA's NWPs.

- **Regional outlook repackaging** covers the multi-day range but at coarse resolution, rarely narrower than a several-hour window over a broad area. It can say a region is at risk on a particular day, but it cannot say which hour or which site.

All of these techniques and products have their purpose, but each covers a slice of the problem. HailDash was built to cover these domains as one continuous forecast.

Multi-Member Ensemble

The final HailDash forecast is an ensemble, not the output of a single model run.

Several model cycles (or members) are run simultaneously with different inputs. Though each individual member produces a probability grid for each class (thunderstorm, hail exceedance), the resulting published product is derived by combining and weighting each of those outputs. Where they agree, probabilities sharpen, and where they disagree, the forecast widens.

The ensemble varies along two axes: each member runs with different inputs, and each member is itself built from specialized model components trained for different temporal ranges.

Model Inputs

Our forecasting system is supported by a robust data pipeline which ingests and transforms several terabytes of weather data per day. The pipeline is fault-tolerant, monitored, and distributed across multiple systems with high SLAs. The system is flexible enough to handle delayed or missing data sources, and has the ability to quickly fallback across multiple upstream origins in the event of upstream outages. The pipeline provides all of the inputs that the HailDash model uses, including:

- Operational deterministic weather model guidance
- Operational ensemble guidance

- Live weather radar
- Satellite imagery
- Lightning detection
- Surface observations
- Quality control and analysis weighting gates

Planned Upgrade / Next-Generation

The HailDash model is undergoing an upgrade intended to significantly improve its performance and latency in the real-time through short-term domain. This next-generation model has been trained to nowcast more quickly and confidently, producing updated guidance within five to ten minutes of new radar and satellite data being published

The upgrade also addresses two shortcomings as described in the Forecast Performance & Verification section:

- Better calibration of the sub-model handoff at the real-time to short-term boundary (~4-8 hours).
- Final calibration step using Platt remaps so that published probabilities more closely match observed event frequencies at every lead time and threshold.

This upgrade is planned for release in Summer 2026.

Forecast Performance & Verification

A forecast is not useful if its outputs are worse than random chance or a simple prediction of "no hail will ever occur anywhere, ever." The latter case is a higher bar than might be expected, due to the rarity of hail events – in our sample set, there are about 7 occurrences of severe hail at a given location per 100,000 hours. This means that a hail forecast that predicts 0% across the entire grid for every hour can objectively and provably claim an accuracy of over 99% while also being completely useless.

To measure HailDash's forecast performance, we have collected statistics beyond simple accuracy (which is in excess of 99% at all lead times).

Verification Methodology

Forecast performance statistics were generated against the following truth data:

- Cloud to ground lightning detection
- Ground-based hail observations
- Radar-derived hail size estimates with quality control, artifact removal, and masking/weighting

Areas of no data (no observation was made, or the radar beam was blocked) are treated as unknown rather than scored as a non-event. Scoring is otherwise strict: each grid cell is scored for each hour, and an event either occurred in that cell during that hour or it did not. A "near-miss" forecast that is an hour early or one cell away receives no credit. Every metric below is computed the way a user experiences the product: the probability a location's forecast page showed, against the weather that actually happened there.

We measured several thousand production forecasts issued between March 1 and July 9, 2026. This time period spans the full transition from early-spring storm systems to mid-summer pulse storms, and includes every hour of the day and night. In total we scored roughly fifteen billion forecast-observation pairs.

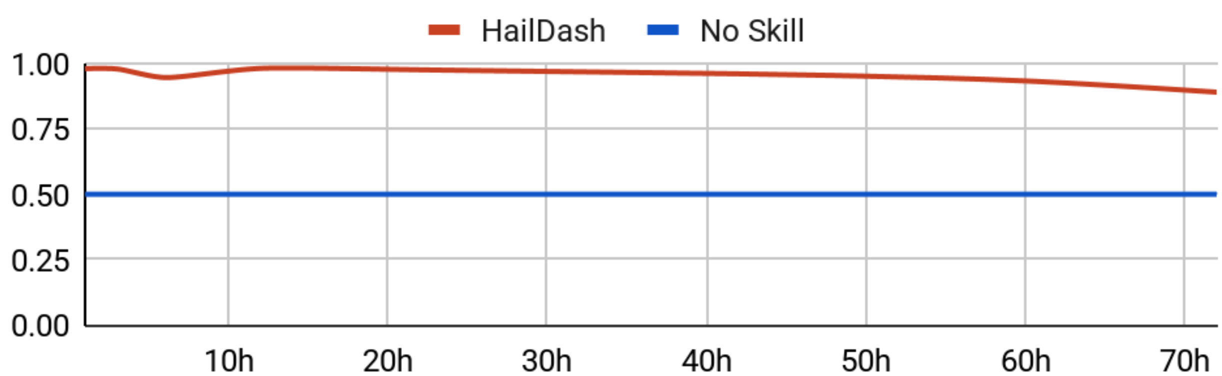
Verification Results

Discrimination: AUC

AUC (area under the [receiver-operating characteristic] curve) measures how reliably the forecast ranks locations that will experience an event above locations that won't (0.5 = coin flip, 1.0 = perfect).

For severe hail (≥0.75″):

Lead Time	1hr	3hr	6hr	12hr	24hr	36hr	48hr	60hr	72hr
HailDash	**0.981**	**0.980**	**0.947**	**0.982**	**0.975**	**0.966**	**0.955**	**0.934**	**0.891**
No Skill	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**



Thunderstorm discrimination follows the same shape (0.958 at one hour). Skill declines gradually with lead and remains far above chance at three days. Discrimination at three days is high enough to rank sites and plan work windows, a regime where per-hour, per-cell guidance is otherwise essentially unavailable.

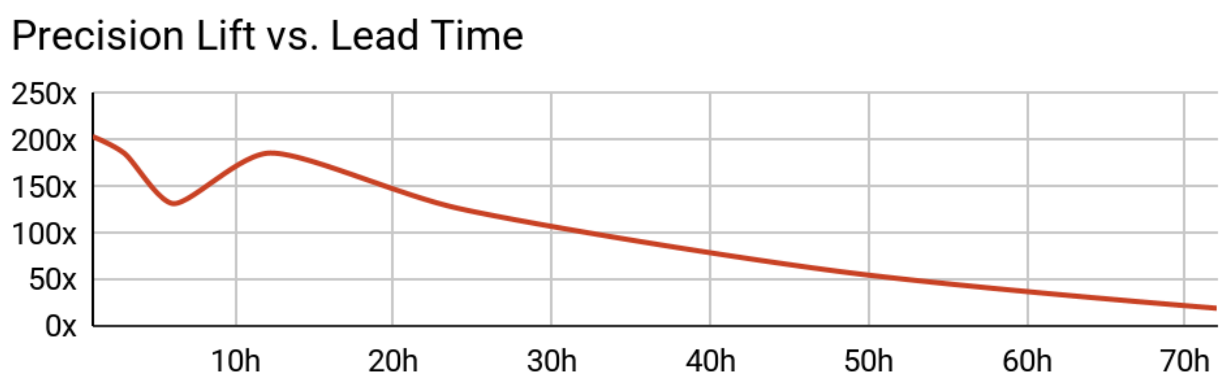
The localized dip at 6 hours is a known artifact of our ensemble blending and aggregation process: this is being addressed in the new version of the model (described earlier).

Precision

Severe hail occurred in roughly 7 of every 100,000 cell-hours in the sample. Against a base rate that low, an unskilled forecast achieves precision equal to chance. HailDash runs two orders of magnitude above that floor at short leads.

For severe hail (≥0.75″):

Lead Time	1hr	3hr	6hr	12hr	24hr	48hr	72hr
Precision vs. Chance	**203x**	**185x**	**131x**	**185x**	**126x**	**58x**	**18x**



Coverage and Relative Risk

Two statistics describe how forecast probability corresponded to observed severe hail at one-to-three-hour leads.

Coverage. 97% of verified severe-hail events occurred in cells where the forecast indicated a probability of at least 0.1%. Cells below that threshold produced the remaining 3% of events. This 3% value is consistent across all output classes, indicating that some or most of these events are attributable to erroneous observation data.

Relative risk. Cells with a forecast probability of 10% or greater experienced severe hail at 200 to 370 times the background frequency. Such cells constitute a small fraction of the grid at any given time.

Observed event frequency rises with forecast probability between these two thresholds. An operational action threshold can therefore be selected according to the cost of preparation relative to the expected loss from an unprepared event.

Limitations

1. The 2026 window contained too few destructive hail (≥2.00″) events to produce statistically meaningful scores. We will be able to publish these scores after additional data is collected over the course of the season.

2. Coverage is not uniform. Radar coverage, terrain, and observation density vary regionally. During the verification process, we mask and de-weight grid cells where observations and estimates are likely to be missing or unreliable. In practice, this means that the performance metrics are more applicable to populated areas than to remote mountainous terrain.

Alerting & Notifications System

HailDash converts its forecast probabilities into location-specific alerts through a three-tier system that follows National Weather Service conventions, so the vocabulary carries the meaning users already know:

- Hail Warnings: an imminent, high-confidence threat that requires immediate action.
- Hail Watches: a period of elevated risk spanning the next several hours. HailDash's watches are graded by severity (yellow, orange, red).
- Hail Outlooks: planning information covering the remainder of the forecast range. These are provided via optional scheduled email digests. Outlooks do not trigger notifications.

Because the forecast is continuous from three days down to the current hour, a single threat escalates naturally through the tiers: outlook, then watch, then warning. Users are notified only when an alert's state changes (issued, upgraded, downgraded, cancelled, or expired). Routine revisions (such as a change to the expected expiration time) update what is shown inside the application but do not issue new notifications. Since notifications are external to the app, users are instructed to treat an alert as active until it is explicitly cancelled, it expires, or they confirm its status within the app.

Alerts use a hysteresis mechanism to absorb HailDash's frequent forecast updates, tuned to respond immediately to new or increasing threats while minimizing desensitization from repeated cycles of cancellation and re-issuance.

Notifications are the delivery layer for alerts: changes to alert state reach users by email and/or push notification to all of their registered devices, independently for every monitored location, filtered by each user's severity and alert type preferences.