Analysis of One- to Five-Day-Out Global 24-Hour Temperature, Probability of Precipitation, and Wind Speed Forecasts

2015 – 2018

By ForecastWatch.com,
a Service of Intellovations, LLC

Eric Floehr
Owner
Intellovations, LLC
6724 Perimeter Loop Road, #275
Dublin, OH 43017
eric@forecastwatch.com
http://www.forecastwatch.com
855-609-9609

Executive Summary

The disruptive impact of weather on lives and businesses is significant. Severe weather causes billions of dollars in damage each year while weather-caused automobile accidents cause 250,000 fatalities annually. In 2018, the total economic cost of weather disasters totaled $215 billion dollars. Costliest back to back years, 2017 and 2018, for weather disasters on record was $653 billion dollars.

People rely on weather forecast providers to help them keep their families safe and their property protected. Businesses in many industries — including recreation, construction, sports, energy and utility, safety and insurance — evaluate past performance of forecast accuracy so they may more effectively assess future risk and opportunity.

Forecast accuracy is also critical for companies that specialize in weather prognostication and for website and app users who rely on the forecast to plan their day. By demonstrating the accuracy of their weather predictions, forecast providers can build long-term success with current—and future—clients and establish trust with the people who rely on them daily to plan their lives and protect their families.

This report provides an analysis of three important aspects of weather forecasts—temperature, precipitation, and wind—for the one- to five-day-out forecast period. Data for these analyses was gathered from 1,108 locations around the world for the four-year period ending December 31, 2018. More than 120 million forecasts were obtained and analyzed from six providers: AccuWeather, Dark Sky, Foreca, Intellicast, The Weather Channel, and Weather Underground. (Note: Precipitation forecasts from Foreca were not included in the PoP analysis because collection of probability of precipitation forecasts from Foreca didn't start until January 1, 2018.) Although both owned by IBM, The Weather Company (which owns The Weather Channel web properties and forecasts) and Weather Underground forecasts are obtained separately from each provider.

In the overall analysis, AccuWeather was the most accurate provider for temperature and precipitation forecasts. AccuWeather's wind speed forecasts were the best among the six providers while its wind bias ranked second behind Foreca.

Accuracy in the three major forecast areas is summarized below:

Temperature Forecasts. AccuWeather was the most accurate provider for both 24-hour high temperature and 24-hour low temperature forecasts for one- to five-day-out. This was reflected in both lowest average absolute error and greatest percentage of forecasts coming within 3°F of actual temperature observations.

Precipitation Forecasts. AccuWeather was the most accurate provider among the five providers analyzed in the evaluation of 24-hour POP forecasts for one- to five-day-out compared to observed precipitation events.

Wind Speed Forecasts. AccuWeather had the most accurate forecasts for one- to five-day-out 24-hour wind speed during the 48-month period. Foreca had the least amount of bias in 24-hour wind speed forecasts, with AccuWeather second.

Analysis of Temperature Forecasts

Forecasts were collected from six top global providers of consumer weather forecasts. Results are expressed as **mean absolute error**—an average of the absolute temperature errors—and the **percentage of forecasts within 3°F**.

High Temperature Forecasts

The mean absolute error for one- to five-day-out 24-hour high temperature forecasts for 2015 – 2018 is shown in **Table 1**.

Findings: AccuWeather had the lowest mean absolute error among the six providers for 24-hour high temperature forecasts at 2.77°F. This was slightly better than The Weather Channel (at 2.81°F) and Weather Underground (2.82°F). Foreca had an error that was 6.1% higher than AccuWeather. Dark Sky's mean absolute error was 3.53°F, considerably worse than all other providers and 27.4% higher than AccuWeather's.

Rank	Provider	Means Abs Error
1	AccuWeather	2.77
2	The Weather Channel	2.81
3	Weather Underground	2.82
4	Foreca	2.94
5	Intellicast	3.02
6	Dark Sky	3.53

Table 1 - One- to five-day-out 24-hour high temperature forecast mean absolute error for 2015 – 2018

Table 2 shows the one- to five-day-out 24-hour high temperature forecasts within 3°F of the actual observed temperature.

Findings: Provider performance for the percentage of one- to five-day-out 24-hour high temperature forecasts falling within 3°F of actual observations followed in the same order as the mean absolute error. For AccuWeather, 71.83% of high temperature forecasts were within 3°F. The Weather Channel and Weather Underground were close behind at 71.35% and 71.23%, respectively. Dark Sky had only 60.70% of forecasts fall within the 3°F range.

Rank	Provider	% within 3°F
1	AccuWeather	71.83%
2	The Weather Channel	71.35%
3	Weather Underground	71.23%
4	Foreca	69.50%
5	Intellicast	69.37%
6	Dark Sky	60.70%

Table 2 - One- to five-day-out 24-hour high temperature forecasts within three degrees for 2015 - 2018

Low Temperature Forecasts

The error in low temperature forecasts tends to be higher than the error in high temperature forecasts. The reasons for this include both definition and collection methodology. Forecasts are collected in the mid-afternoon, so, a zero-day-out high is defined as the current day high, and the zero-day-out low is defined as the next day's low. Therefore, a same day-out low always follows its high.

Temperature forecast error, whether high or low, increases as the forecast time moves further out, and the low temperature observations occur approximately twelve hours after the corresponding high temperatures. However, this doesn't account for the entire difference in accuracy between high and low temperature forecast. In general, low temperatures tend to be slightly less predictable than high temperatures.

Table 3 below shows the mean absolute error for global one- to five-day-out 24-hour low temperature forecasts.

Rank	Provider	Means Abs Error
1	AccuWeather	3.21
2	The Weather Channel	3.36
3	Weather Underground	3.36
4	Foreca	3.44
5	Intellicast	3.54
6	Dark Sky	3.79

Table 3 - One- to five-day-out 24-hour low temperature forecast mean absolute error for 2015 – 2018

Findings: AccuWeather had the lowest mean absolute error among the six providers for one-to five-day-out 24-hour low temperature forecasts for the 48-month period ending December 2018. The Weather Channel and Weather Underground were tied for second in accuracy, with an error that was 4.7% higher. Intellicast's error was 10.3% higher while Dark Sky's was 18.1% higher.

Table 4 shows the percentage of one- to five-day-out 24-hour low temperature forecasts that fell within 3°F of the observed low temperature.

Findings: Nearly two-thirds (65.90%) of AccuWeather's 24-hour low temperature forecasts from one- to five-day-out fell within 3°F of actual observations. This was notably higher than the next four providers (The Weather Channel, Weather Underground, Foreca, and Intellicast), who ranged from 62.49% to 64.04%, and significantly higher than Dark Sky, which came in at 57.70%.

Rank	Provider	% within 3°F
1	AccuWeather	65.90%
2	The Weather Channel	64.04%
3	Weather Underground	63.92%
4	Foreca	62.59%
5	Intellicast	62.49%
6	Dark Sky	57.70%

Table 4 - One- to five-day-out 24-hour low temperature forecasts within 3°F for 2015 –2018

Analysis of Probability of Precipitation Forecasts

How Precipitation Forecasts Are Evaluated

Precipitation forecasts are expressed as probability of precipitation (POP). Probability forecasts cannot be evaluated individually — a precipitation observation either happened or it didn't. However, in aggregate, the most accurate POP forecasts would describe the percentage of time precipitation happened. For example, of all the times 10% probability of precipitation was forecast, ideally there should have been precipitation on 10% of those days. To analyze accuracy, the difference between the percentage of precipitation days for a given POP forecast was used, and then averaged over each POP value to find the average absolute error for POP forecasts. For example, if precipitation was observed on 13.05% of days where 10% POP was forecast, the error would be 13.05-10.00 or 3.05.

POP Analysis

Table 5 reflects the percentage of time that measurable precipitation occurred in a full 24-hour day for the full range of POP forecasts. AccuWeather performed best in the 10%, 20%, 40%, and 50% bins, while The Weather Channel was closest to the actual percentage at 0%, 30%, 80%, and 100% POP. Dark Sky was best at 60% and 70%. Overall, when averaging absolute error over the entire POP range, AccuWeather had the lowest error at 8.58 percentage points, followed by The Weather Channel at 8.72.

Provider	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
AccuWeather	6%	14%	34%	48%	42%	53%	68%	82%	86%	86%	84%
Dark Sky	9%	21%	35%	44%	52%	58%	67%	74%	78%	81%	83%
Intellicast	6%	16%	38%	48%	55%	63%	71%	82%	78%	89%	95%
The Weather Channel	5%	15%	37%	44%	53%	63%	71%	81%	79%	89%	95%
Weather Underground	5%	16%	38%	48%	55%	63%	71%	83%	79%	89%	95%

Table 5 - Percentage of time measurable precipitation occurred given different POP forecasts, 2015 – 2018

Provider	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Average Abs Error
AccuWeather	5.79	4.04	14.08	18.40	2.35	3.23	8.01	12.18	6.20	4.42	15.63	8.58
The Weather Channel	5.05	5.22	17.10	14.35	13.30	12.71	10.75	10.93	0.66	1.07	4.74	8.72
Weather Underground	5.30	6.00	17.94	18.31	15.43	13.36	11.37	12.94	0.90	0.80	4.77	9.74
Dark Sky	8.62	10.74	15.37	14.43	11.65	7.92	6.99	4.45	2.47	8.84	16.60	9.83
Intellicast	5.53	6.31	18.03	18.25	15.20	13.11	10.98	12.47	1.70	1.39	5.20	9.83

Table 6 – Average absolute error of POP vs. actual precipitation percentage 2015 - 2018

Table 6 shows the average absolute error of POP versus actual precipitation percentage. The data in this table is derived from the data in Table 5.

Findings: AccuWeather was the most accurate provider for POP forecasts with an average absolute error of 8.58, followed by The Weather Channel at 8.72. Intellicast and Dark Sky were tied as the least accurate providers for POP.

Analysis of Wind Speed Forecasts

Accurate wind forecasts are critical for businesses that rely on wind for the efficient use of resources. In particular, wind farm operators and utility operators make crucial decisions based on anticipated wind conditions. Accurate forecasting allows operators to achieve favorable trading performances on the electricity markets. The further in advance an operator can make a reliable estimate about how much electricity that can be produced, the more profit they can make.

Wind forecast results are presented in two ways: 1) average absolute error – the difference between the average daily wind speed and the forecast wind speed, and 2) average bias – the positive or negative difference between forecast wind speed and actual wind speed.

How Wind Accuracy Is Measured

There are several ways that wind accuracy can be assessed. This analysis calculated the absolute error between the observed daily wind speeds (an average of 24-hourly observations) and the provider's wind forecast. This analysis does not take wind direction (wind vector) into account and thus strictly measures the difference in wind speed.

The wind forecast accuracy is also assessed by examining bias in wind speed forecasts. Bias measures the tendency for a wind forecast to over- or underestimate actual wind conditions. Providers that have a positive bias are more apt to predict wind speeds that are higher than those observed. Conversely, providers whose forecasts have a negative bias tend to predict wind speeds that are lower than actual observed wind speeds.

Average Absolute Error

Table 7 shows the average absolute error for global one- to five-day-out 24-hour wind speed forecasts for 2015 - 2018.

Rank	Provider	Abs Error (kph) (lower is better)
1	AccuWeather	3.37
2	Foreca	4.31
3	Dark Sky	4.77
4	The Weather Channel	4.77
5	Weather Underground	4.79
6	Intellicast	4.94

Table 7 - One- to five-day-out average absolute error for 24-hour average wind speed forecasts 2015 – 2018

Findings: AccuWeather stood out as the most accurate wind speed forecaster during the period of analysis with an absolute error of 3.37 kph. This was 28% better than Foreca, the second most accurate provider with an error of 4.31 kph. Dark Sky, The Weather Channel, and Weather

Underground were bunched closely together with error rates ranging from 4.77 kph to 4.79 kph. Intellicast was the least accurate wind speed forecaster with an absolute error of 4.94 kph.

Average Bias

Table 8 shows the average bias of one- to five-day-out 24-hour average wind speed forecasts.

Findings: The most accurate provider with respect to one- to five-day-out 24-hour wind speed bias was Foreca, which had a bias of 0.30 kph for the period. AccuWeather was the second most accurate at 0.74 kph. The Weather Channel (3.68 kph) and Intellicast (3.69 kph) had the highest wind speed bias. All providers exhibited a positive bias (forecasting wind speed higher than actual) except for Dark Sky.

Rank	Provider	Bias (kph)
1	Foreca	0.30
2	AccuWeather	0.74
3	Dark Sky	-2.77
4	Weather Underground	3.63
5	The Weather Channel	3.68
6	Intellicast	3.69

Table 8 - One- to five-day-out 24-hour average bias in wind speed forecasts, 2015 - 2018

Methodology

Temperature

Error is determined by subtracting the actual temperature from the forecast temperature. Both measures are stored in whole degrees Fahrenheit. A forecast that predicts too low a temperature will have a **negative error**, while a forecast that is too high will have a **positive error**.

After the error is established, the average absolute error can be determined. This calculation takes the absolute value of the error of each forecast, so that all errors are positive, and then averages all errors.

This measures how far off the set of forecasts is on average without regard to whether they are too high or too low. Finally, if the mean absolute error was three degrees or less, the forecast was considered within 3°F.

Probability of Precipitation

POP forecasts were compared against precipitation measured or observed at the observation locations analyzed. The forecasts were grouped into eleven bins, 0%, 10%, through 100%. A POP forecast between 0% and 5% (inclusive) was binned into the 0% bin, 5% (exclusive) to 15% (inclusive) into the 10% bin, and so on until the 100% bin, which consisted of forecasts between 95% and 100% (inclusive). The percentage of precipitation events for the forecasts in each bin was then calculated.

Wind Speed

Error is determined by subtracting the daily average wind speed from the forecast wind speed. A forecast that predicts too low a wind speed will have a **negative error**, while a forecast that predicts too high a wind speed will have a **positive error**.

After the error is established, the average absolute error can be determined. This measure takes the absolute value of the error of each forecast so that all errors are positive, and then averages all errors. This measures how far off the set of forecasts is on average without regard for if they are too high or too low.

ForecastWatch employed the commonly used method of confidence intervals for a normal distribution of error to determine if providers should be considered statistically tied. This is based on the total number of samples, the mean absolute error of the samples, and the standard deviation of absolute error. A confidence interval is a set of values that are all reasonable estimates for a population (true) parameter, based on a particular sample. Not all intervals will contain the true value of the statistic, and the accuracy of the interval is dependent on the assumptions of independence and the underlying distribution of the sample. Because of such assumptions, other statistical means of assessing ties may occasionally lead to different results.

Providers

- **AccuWeather** Forecasts were collected using the AccuWeather API at http://api.accuweather.com using a specific location code.
- **Dark Sky** Forecasts were collected using the Dark Sky API at http://api.forecast.io. Latitude and longitude of the observation station were used to retrieve specific forecasts.
- **Foreca** Forecasts were collected from the 10-day forecast page at http://www.foreca.com. During the analysis period, this transitioned from scraping the website to using the API that populates the page. The location parameter used was the city and state of the observation location for the website, and a location code (either ICAO or WMO) for the API.
- **Intellicast** http://intellicast.com. Extended forecast page. Location parameter was a site-specific code for the location.
- **The Weather Channel** Forecasts were collected from the 10-day forecast page at http://www.weather.com. During the analysis period, this transitioned from scraping the website to using the API that populates the page. Latitude and longitude of the observation stat were used to retrieve specific forecasts.
- **Weather Underground** Forecasts were collected using the Weather Underground API located at http://www.wunderground.com/api. The location parameter used to retrieve specific forecasts was the International Civil Aviation Organization (ICAO) code or surface synoptic observations (SYNOP) of the observation station.

Observation Collection

Data was collected from eight regions at specific times during the day. In **Table 9**, for example, daily temperature forecasts were collected at 22:00 UTC (6 p.m. Eastern Standard Time) in the United States and continued until all forecasts were collected. For each location, forecasts from all providers were collected at the exact same time.

Region	Collection Time	Number of Stations
United States	22:00 UTC	759
Canada	21:40 UTC	41
Europe	16:00 UTC	187
Asia Pacific	08:00 UTC	63
Africa	15:30 UTC	13
Middle East	13:00 UTC	21
Central America	23:00 UTC	10
South America	21:00 UTC	14

Table 9 - Forecast collection times and regions

Validity

Forecasts were considered **valid** if they were complete (i.e., they contained a high and low temperature forecast, a POP forecast, and a wind forecast), and if they passed both manual and automated audits. These audits checked for out-of-bounds values and other indicators that suggested the forecast should be marked as invalid. Forecasts that were simply **bad** (inaccurate or wrong) were not considered invalid. However, forecasts issues caused by system errors or delivery problems (such as a -32768 degree high temperature, a 120% chance of rain or a 270 kph wind speed) were declared invalid.

Observation Data

Observation data was collected from the primary Automated Surface Observing System (ASOS) network in the United States as well as international equivalents. United States and International data collected from the Integrated Surface Database (ISD) product. Canadian data was collected from Environment Canada. All products consisted of hourly and daily observation parameters.

Observed High and Low Temperature

The maximum and minimum temperature observations are from the 24-hour local time temperature observations and were used to construct the high and low temperature observation. United States 24-hour high and low temperature observations were collected from the Summary of The Day (SOD)

records which use 5-minute sampling. All 24-hour high and low international observations were derived from hourly and special report observations. No attempt to curve fit or otherwise determine an intra-hour temperature estimate was performed.

Observed Precipitation

Precipitation measurements were taken from 24-hour local time precipitation observations. For United States and Canada locations if 0.01 inches or more of liquid-equivalent precipitation fell during any hour of that day, it was considered to be a day with precipitation. Internationally, precipitation reports were used to determine precipitation observations. The occurrence or non-occurrence of precipitation was then compared to the POP forecast.

Observed Wind

Wind conditions were taken from hourly observations over the course of a 24-hour period from local midnight to midnight. These observations were then averaged to construct the daily wind observation.

Calculation Methodology

Tables 10, 11, and 12 show the number of high/low temperature, POP, and wind forecasts collected and compared for each provider for the one- to five-day-out forecasts. The percent of possible forecasts collected and compared is less than 100% because of invalid forecasts, problems in collecting forecasts successfully, including the unavailability of a provider's website or feed due to network or other issues, and days in which observations were not available for a particular site. Overall, across all providers, the percentages of possible forecasts and observations available for comparison were 92.45% for temperature, 92.71% for POP, and 79.38% for wind.

Provider	Number of Temperature Forecasts	Percent of Possible Forecasts
AccuWeather	7,493,516	92.58%
Dark Sky	7,504,977	92.72%
Foreca	7,473,857	92.34%
Intellicast	7,513,606	92.83%
The Weather Channel	7,514,118	92.84%
Weather Underground	7,397,938	91.40%

Table 10 - One- to five-day-out high and low temperature forecasts analyzed and percent of possible forecasts, 2015 –2018

Provider	Number of Precipitation Forecasts	Percent of Possible Forecasts
AccuWeather	7,493,516	92.58%
Dark Sky	7,504,977	92.72%
Intellicast	7,513,606	92.83%
The Weather Channel	7,514,118	92.84%
Weather Underground	7,397,938	91.40%

Table 11 - One- to five-day-out POP forecasts analyzed and percent of possible forecasts, 2015 – 2018

Provider	Number of Wind Forecasts	Percent of Possible Forecasts
AccuWeather	6,418,458	79.51%
Dark Sky	6,428,210	79.64%
Foreca	6,401,229	79.30%
Intellicast	6,435,404	79.72%
The Weather Channel	6,423,338	79.58%
Weather Underground	6,336,885	78.50%

Table 12 - One- to five-day-out wind speed forecasts analyzed and percent of possible forecasts, 2015 – 2018

About ForecastWatch

ForecastWatch, a service of Intellovations, LLC, has been the world's premier weather forecast monitoring and assessment company since 2003, when it released the largest public weather forecast accuracy study at the time. ForecastWatch compiles weather forecasts and observations from more than 1,200 locations around the world, including the United States, Canada, Europe, South America, Central America, Africa, and the Asian Pacific. ForecastWatch maintains a historical database of more than 950 million weather forecasts from a number of providers and provides unbiased reporting.

Meteorologists, utilities, and energy companies depend on ForecastWatch's accurate data and analysis. Agriculture, futures traders, and other companies whose business depends on being right about the weather put their trust in ForecastWatch to help them achieve success. The data meets the highest standard of scientific inquiry and has been used in several peer-reviewed studies.