ForecastWatch
Accuracy Defined

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

January-December 2016

By ForecastWatch.com,
a Service of Intellovations, LLC

Eric Floehr
Owner
Intellovations, LLC

P.O. Box 164442
Columbus, Ohio 43216
eric@forecastwatch.com
www.forecastwatch.com
(614) 440-0130

Executive Summary

Forecast accuracy is extremely important for the vast number of businesses that make important business decisions based on weather and weather predictions. It's critical for many industries, both public and private, including recreation, construction, sports, energy and utility, safety and insurance. Understanding and evaluating the past is key to assessing future risk and opportunity.

Accuracy is also vitally important to companies that specialize in weather predictions. These organizations have a great investment in producing accurate forecasts and in establishing their capability to provide that accuracy. The ability to make dependable predictions and to meet the needs of their clients is critical to the health of their business.

This report contains three separate analyses of 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 1,145 locations around the world for the 12-month period ending December 31, 2016. In total, more than 30 million forecasts were obtained and analyzed from six providers: AccuWeather, Dark Sky, Foreca, Intellicast, The Weather Channel, and Weather Underground. Foreca was not included in the analysis of probability of precipitation.

AccuWeather had the highest forecast accuracy in each of these three forecast areas, as summarized below:

Temperature Forecasts. AccuWeather's high temperature and overall (high and low combined) temperature forecasts had the lowest average absolute error. AccuWeather also had the highest percentage of high temperature forecasts and overall temperatures within 3°F. The Weather Channel and Weather Underground tied for least average absolute error and most temperatures within 3°F for low temperature forecasts.

Precipitation Forecasts. AccuWeather's performance for probability of precipitation (POP) forecast accuracy was the best among the five providers analyzed. It achieved both the best Brier score as well as the best average absolute error when POP forecasts were compared to observed precipitation events.

Wind Speed Forecasts. AccuWeather's wind speed forecast accuracy was the best among the six providers analyzed, ranking the highest for both average absolute error and average bias.

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**.

AccuWeather's high temperature and overall (high and low combined) temperature forecasts had the lowest average absolute error. AccuWeather had the highest percentage of high temperature forecasts and overall temperature forecasts within 3°F.

High Temperature Forecasts

The mean absolute error for one- to five-day-out high temperature forecasts for January-December 2016 is shown in **Table 1**.

Findings: AccuWeather had the lowest mean absolute error. The Weather Channel and Weather Underground were essentially tied for second. Notably, there is a significant difference of nearly one full degree between AccuWeather and the sixth-place provider, Dark Sky.

Rank	Provider	Mean Abs Error
1	AccuWeather	2.651
2*	The Weather Channel	2.709
2*	Weather Underground	2.716
4	Foreca	2.831
5	Intellicast	2.929
6	Dark Sky (forecast.io)	3.580

Table 1: One- to five-day-out high temperature forecast mean absolute error for January-December 2016
** Difference not statistically significant at the 99% confidence level*

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

Findings: AccuWeather had the highest percentage of forecasts (73.50%) within 3°F of the observation whereas Dark Sky had the lowest percentage (60.59%).

The Weather Channel (72.74%) and Weather Underground (72.63%) followed AccuWeather in second and third place, respectively.

Rank	Provider	% within 3°F
1	AccuWeather	73.50%
2	The Weather Channel	72.74%
3	Weather Underground	72.63%
4	Foreca	70.90%
5	Intellicast	70.72%
6	Dark Sky (forecast.io)	60.59%

Table 2: One- to five-day-out high temperature forecasts within three degrees for January-December 2016

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. Low temperatures are defined (and collected) as the low temperature from 7:00 p.m. to 8:00 a.m. while high temperatures are defined (and collected) from 7:00 a.m. to 7:00 p.m. Therefore, the one-day-out low temperature forecast occurs overnight ***after*** the one-day-out high temperature.

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 low temperature forecasts.

Findings: The Weather Channel and Weather Underground were statistically tied for first. AccuWeather came in third with a mean absolute error of 2.978, which was .7% higher than The Weather Channel.

Rank	Provider	Mean Abs Error
1*	The Weather Channel	2.956
1*	Weather Underground	2.962
3	AccuWeather	2.978
4	Intellicast	3.143
5	Foreca	3.164
6	Dark Sky (forecast.io)	3.663

Table 3: One- to five-day-out low temperature forecast mean absolute error for January-December 2016
**Difference not statistically significant at the 99% confidence level*

Rank	Provider	% within 3°F
1	The Weather Channel	67.96%
2	Weather Underground	67.87%
3	AccuWeather	67.71%
4	Intellicast	66.41%
5	Foreca	65.06%
6	Dark Sky (forecast.io)	58.51%

Table 4: One- to five-day-out low temperature forecasts within 3°F for January-December 2016

Table 4 shows the accuracy rate for low temperature forecasts within 3°F.

Findings: Nearly 68% of The Weather Channel forecasts were within 3°F for low temperature, closely followed by Weather Underground at 67.87%. The top three providers were fairly closely packed, while Dark Sky, at 58.51%, was well behind.

Overall Temperature Forecast Comparison

The overall temperature forecast comparison was calculated by combining the high and low temperature mean absolute error and averaging. Each provider's overall temperature forecast performance is shown in **Table 5** below.

Findings: AccuWeather had the lowest mean absolute error for overall temperature one- to five-day-out forecasts. It was 0.12% better than second-place Weather Underground. AccuWeather's mean absolute error was 22.3% lower than Dark Sky's.

Rank	Provider	Mean Abs Error
1	AccuWeather	2.815
2	The Weather Channel	2.832
3	Weather Underground	2.839
4	Foreca	2.998
5	Intellicast	3.036
6	Dark Sky (forecast.io)	3.621

Table 5: One- to five-day-out overall temperature forecast mean absolute error for January-December 2016

Rank	Provider	% within 3°F
1	AccuWeather	70.61%
2	The Weather Channel	70.35%
3	Weather Underground	70.25%
4	Intellicast	68.56%
5	Foreca	67.98%
6	Dark Sky (forecast.io)	59.55%

Table 6: One- to five-day-out overall temperature forecasts within 3°F for January-December 2016

Table 6 shows the combined average of high and low one- to five-day-out temperature forecasts that came within 3°F.

Findings: At 70.61%, AccuWeather had the highest percentage of one- to five-day-out overall temperature forecasts within 3°F. AccuWeather's performance was 11.06 percentage points better than Dark Sky's.

Analysis of Probability of Precipitation Forecasts

Accurate precipitation forecasts are vital to organizations that must count on dry weather to perform certain tasks or make alternative plans when wet weather is forecast. For example, concrete pouring and asphalt paving are more successful on days without precipitation. When these operations are performed in the rain, the integrity of the material used can be compromised, necessitating costly re-installation. At the same time, however, it's also costly to reschedule these services when precipitation that was forecast does not materialize.

AccuWeather's performance for probability of precipitation (POP) forecast accuracy was the best among the five providers analyzed. It achieved both the best Brier score as well as the best average absolute error when POP forecasts were compared to observed precipitation events.

How Precipitation Forecasts Are Evaluated

Precipitation forecasts are expressed as probability of precipitation (POP). POP can be assessed in two ways: reliability and resolution. To fully evaluate a POP forecast, both must be considered. A Brier score, which takes into account both reliability and resolution, is one measure and is used to evaluate POP forecasts in this report. The range of the Brier score is zero to one, with zero being perfectly reliable and resolved.

Reliability

A *reliable* forecast is one where it rains 10% of the time the POP forecast called for a 10% chance of rain. However, consider a scenario where a 30% chance of precipitation is always predicted at a location and, on average, it precipitates three out of every ten days. This forecast would be reliable but not necessarily useful because it does not define with absolute certainty whether rain will take place or not.

Resolution

To paint a fuller picture, the other measure is resolution. A *resolved* POP forecast would always predict either no chance of precipitation or a 100% chance of precipitation, since precipitation either occurs or it doesn't. There is no place for probability of precipitation in a resolved forecast. Now consider a forecast that predicts a 100% chance of precipitation on dry days and a 0% chance of precipitation on rainy or snowy days. It is resolved because it predicted either precipitation or completely dry conditions, but this example shows that a resolved forecast may not necessarily be reliable.

Brier Score

As shown in **Table 7**, AccuWeather had the lowest Brier score for global one- to five-day-out POP forecasts for January through December 2016 among the five providers analyzed. Weather Underground and The Weather Channel followed closely, while Dark Sky finished last among the five providers in precipitation forecast accuracy.

Rank	Provider	Brier Score (lower is better)
1	AccuWeather	0.1529
2	Weather Underground	0.1533
3	The Weather Channel	0.1534
4	Intellicast	0.1545
5	Dark Sky (forecast.io)	0.1584

Table 7: One- to five-day-out Brier score for POP forecasts, January-December 2016

POP Analysis

Provider	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
AccuWeather	5%	12%	32%	47%	41%	51%	66%	81%	86%	85%	83%
Dark Sky (forecast.io)	8%	31%	39%	43%	47%	55%	65%	76%	86%	93%	n/a*
Intellicast	5%	16%	38%	48%	54%	61%	68%	82%	77%	87%	94%
The Weather Channel	5%	16%	38%	48%	55%	62%	69%	82%	78%	88%	94%
Weather Underground	5%	16%	38%	48%	55%	62%	69%	82%	78%	88%	94%

Table 8: Percentage of time measurable precipitation occurred given different POP forecasts, January-December 2016.
**Dark Sky never forecasted a 100% POP so there is no data to report.*

Table 8 above shows the percentage of time that measurable precipitation occurred for the full range of POP forecasts. AccuWeather's forecasts were particularly accurate at the 40% POP level, where precipitation occurred 41% of the time, and at the 50% POP level, where precipitation occurred 51% of the time. However, at AccuWeather's 100% POP forecast, precipitation occurred a comparatively low 83% of the time. For the 100% POP forecasts for Intellicast, The Weather Channel and Weather Underground, precipitation occurred 94% of the time. Dark Sky never forecasted a 100% POP, so their performance is not ranked.

Table 9 shows the average error of POP versus actual precipitation. The data in this table is derived from the data in Table 2. For example, AccuWeather's 2.05 average error for the 10% POP that appears in Table 3 is calculated as 12.05% (shown in Table 2 as 12%) less the 10% POP.

AccuWeather POP forecasts were the most accurate among providers for several POP forecasts, including 10%, 20%, 40% and 50% probabilities. The Weather Channel and Weather Underground had the most accurate forecasts at the 80% and 90% POP levels, with a variance between POP forecast and actual precipitation occurrence of approximately 2%. AccuWeather also had the lowest average absolute error for the period of analysis, 21.3% better than The Weather Channel, which had the second best average absolute error.

Provider	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Average Abs Error
AccuWeather	5.31	2.05	11.83	17.17	0.88	1.39	6.46	10.94	5.81	5.40	17.40	7.70
The Weather Channel	5.16	5.62	17.76	17.86	14.53	11.69	8.73	11.80	2.00	1.95	5.62	9.34
Weather Underground	5.15	5.64	17.73	17.85	14.52	11.69	8.66	12.08	1.97	1.92	5.62	9.35
Intellicast	5.34	5.77	17.76	17.93	14.13	11.20	8.29	11.52	2.88	2.53	6.15	9.41
Dark Sky (forecast.io)	8.16	21.21	19.41	13.37	7.15	4.60	5.48	6.26	5.99	2.73	n/a*	9.44

Table 9: Average error of POP vs. actual precipitation percentage, averaged over each cohort, January-December 2016.
**Dark Sky never forecasted a 100% POP so there is no data to report.*

Analysis of Wind Speed Forecasts

Accurate wind forecasts are critical for businesses that rely on wind for the proper 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.

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.

AccuWeather's performance with regard to wind speed forecast accuracy was the best among the six providers analyzed. It was ranked the highest for both average absolute error as well as average bias for the 12-month period ending December 31, 2016.

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 actually 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

The average absolute error for global one- to five-day-out wind speed forecasts for January through December 2016 is shown in **Table 10**. AccuWeather had the lowest mean absolute error at 3.28 kph. This was 1.15 kph (or approximately 25%) better than the next best provider, Foreca. The difference between AccuWeather and the last provider was 1.67 kph (34%).

Rank	Provider	Abs Error (kph) (lower is better)
1	AccuWeather	3.28
2	Foreca	4.43
3	The Weather Channel	4.76
4	Weather Underground	4.78
5	Dark Sky (forecast.io)	4.86
6	Intellicast	4.95

Table 10: One- to five-day-out average absolute error for 24-hour average wind speed forecasts, January-December 2016

Average Bias

Table 11 shows the average bias of one- to five-day-out 24-hour average wind speed forecasts. At 0.51 kph, AccuWeather had the lowest average wind speed bias of any of the providers analyzed. This means that on average — for one year's worth of one- to five-day-out wind forecasts — AccuWeather over forecast wind speeds by 0.51 kph. Foreca, with a bias of 1.08 kph, was second. All other providers had a wind forecast bias greater than 3 kph. All providers had a positive bias except for Dark Sky.

Rank	Provider	Bias (kph)
1	AccuWeather	.51
2	Foreca	1.08
3	Dark Sky (forecast.io)	-3.01
4	Weather Underground	3.54
5	The Weather Channel	3.60
6	Intellicast	3.67

Table 11: One- to five-day-out 24-hour average bias in wind speed forecasts, January-December 2016

Methodology

Temperature

Error is determined by subtracting the actual temperature from the forecast temperature. 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 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.

Overall temperature accuracy was calculated by taking the average of the mean absolute error for the high and low temperature forecasts. Overall temperature forecasts within 3°F were calculated as the average of the percentage of high temperature forecasts within 3°F and low temperature forecasts within 3°F.

High and low temperature forecasts and observations were collected and stored as whole degrees **Fahrenheit**. Therefore, 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 at the various locations analyzed. If 0.01 inches or more of liquid-equivalent precipitation fell during that day, it was considered to have been a precipitation event.

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

Observation Collection

Data was collected from eight regions at specific times during the day. In **Table 12**, for example, daily temperature forecasts were collected at 22:00 UTC (6 p.m. Eastern Standard Time) in the United States region, 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	792
Canada	21:40 UTC	39
Europe	16:00 UTC	193
Asia Pacific	08:00 UTC	64
Africa	15:30 UTC	13
Middle East	13:00 UTC	21
Central America	23:00 UTC	10
South America	21:00 UTC	13

Table 12: 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 data was quality controlled by the National Climatic Data Center (NCDC) prior to delivery to ForecastWatch via the Quality-Controlled Local Climatic Data (QCLCD) product data set. Canadian data was collected from Environment Canada. Other international data came from the Integrated Surface Database (ISD) product. All products consisted of hourly and daily observation parameters.

Observed High and Low Temperature

The maximum temperature from the 7 a.m. to 7 p.m. local time hourly observations was used to construct the high temperature observation. The minimum temperature from the 7 p.m. to 8 a.m. local

time hourly observations were used to construct the low temperature observation. 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. 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. 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 13, 14, and 15 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 the percentages of possible forecasts and observations available for comparison were 92% for temperature, 93% for POP, and 83% for wind.

Provider	Number of Forecasts	Percent of Possible Forecasts
AccuWeather	1,910,453	91.68%
Dark Sky (forecast.io)	1,913,592	91.83%
Foreca	1,909,872	91.65%
Intellicast	1,914,207	91.86%
The Weather Channel	1,914,396	91.87%
Weather Underground	1,899,214	91.14%

Table 13: One- to five-day-out high and low temperature forecasts analysed and percent of possible forecasts

Provider	Number of Forecasts	Percent of Possible Forecasts
AccuWeather	1,943,977	93.11%
The Weather Channel	1,942,044	93.02%
Weather Underground	1,947,168	93.26%
Intellicast	1,947,364	93.27%
Dark Sky (forecast.io)	1,928,981	92.39%
AccuWeather	1,943,977	93.11%

Table 14: One- to five-day-out POP forecasts analysed and percent of possible forecasts

Provider	Number of Forecasts	Percent of Possible Forecasts
AccuWeather	1,730,792	83.12%
Dark Sky (forecast.io)	1,729,038	83.03%
Foreca	1,729,091	83.04%
Intellicast	1,733,604	83.25%
The Weather Channel	1,733,786	83.26%
Weather Underground	1,717,951	82.50%

Table 15: One- to five-day-out wind speed forecasts analysed and percent of possible forecasts

About ForecastWatch.com

ForecastWatch, a service of Intellovations, LLC, has been the nation'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 600 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 us to help them achieve success. The data meets the highest standard of scientific inquiry and has been used in several peer-reviewed studies.