



# ForecastWatch

*Accuracy Defined*

## **Analysis of Combined One- to Five-Day-Out Global Temperature Forecasts** (Calendar Year 2015)

*By ForecastWatch.com, a Service of Intellovations, LLC  
May 12, 2016*

**Contact:**

Eric Floehr  
*Intellovations, LLC*

P.O. Box 164442  
Columbus, Ohio 43216

[eric@forecastwatch.com](mailto:eric@forecastwatch.com)  
<http://www.forecastwatch.com>  
(614) 440-0130

## Executive Summary

ForecastWatch collected one- to five-day-out high and low temperature forecasts from six top global providers of consumer weather forecasts in 2015. Almost 11.7 million high and low temperature forecasts were collected for 1,148 locations throughout the world and compared against the 7am to 7pm local-time high temperature and 7pm to 8am local-time low temperature recorded by the observation station at each location. Mean absolute error and the number of forecasts within three degrees Fahrenheit were calculated for high and low temperature forecasts, as well as for overall temperature (taken as the average of the high and low temperature absolute errors and percentage within three degrees). AccuWeather's high temperature and overall (high and low combined) temperature forecasts had the highest accuracy in both measurements, as well as the percentage of low temperature forecasts within three degrees. The Weather Channel, Weather Underground, and AccuWeather statistically tied at the 99% confidence level for mean absolute error for low temperature forecasts.

## How Temperature Forecasts Are Evaluated

Temperature forecast accuracy is measured a number of ways. All accuracy calculations begin with taking the forecast high or low temperature and subtracting the actual observed high or low temperature. This number is called the error. 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.

Average absolute error is a measure of the accuracy of temperature forecasts. This measure takes the absolute value of the error of each forecast, so that all errors are positive, and then averages all errors. This is a measure of how far off, on average, the set of forecasts is, regardless if they are too high or too low. High and low temperature forecasts and observations are integer numbers in degrees Fahrenheit. If the mean absolute error is three degrees or less, the forecast is within three degrees Fahrenheit.

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 three degrees was calculated as the average of the percentage of high temperature forecasts within three degrees and low temperature forecasts within three degrees.

Most providers forecast the high temperature that will occur between 7am and 7pm local time, and forecast the low temperature that will occur between 7pm and 8am. This is because most consumers of weather forecasts expect to see that the high will occur during the day and the low will occur overnight. These forecasts are then compared with the high and low temperature observations that occurred during those respective time periods.

## Results of High Temperature Forecast Comparison

The mean absolute error for global one- to five-day-out high temperature forecasts for calendar year 2015 are shown in Table 1. There is a difference of nearly one degree Fahrenheit between the most accurate provider, AccuWeather, and the least accurate provider, Dark Sky. At the 99% confidence level, AccuWeather was the most accurate provider of one- to five-day-out high temperature forecasts for the study period. The Weather Channel and Weather Underground were tied for second place at the 99% confidence level. Intellicast and Foreca were tied for fourth at that confidence level.

Table 2 shows the percentage of one- to five-day-out high temperature forecasts within three degrees of the observation. This means the high temperature forecast was within the range of three degrees lower to three degrees higher than the observed high temperature. Three degrees of error is often used by meteorologists as a standard measure of forecast performance, as differences in forecasts greater than

three degrees can often be noticed by consumers, hence the often-heard “three degree guarantee”. AccuWeather had the highest percentage of forecasts within three degrees, at 71.50%, whereas Dark Sky had the fewest at 58.63%. This is nearly 22% more high temperatures within three degrees Fahrenheit than the last place provider, Dark Sky.

Rank	Provider	Mean Abs Error
1	<i>AccuWeather</i>	2.799
2*	<i>The Weather Channel</i>	2.869
2*	<i>Weather Underground</i>	2.872
4*	<i>Intellicast</i>	3.091
4*	<i>Foreca</i>	3.097
6	<i>Dark Sky (forecast.io)</i>	3.715

*Table 1: One- to five-day-out high temperature forecast mean absolute error for calendar year 2015*

*\* – difference not statistically significant at the 99% confidence level*

Rank	Provider	% within 3°F
1	<i>AccuWeather</i>	71.50%
2	<i>The Weather Channel</i>	70.64%
3	<i>Weather Underground</i>	70.63%
4	<i>Intellicast</i>	68.60%
5	<i>Foreca</i>	67.46%
6	<i>Dark Sky (forecast.io)</i>	58.63%

*Table 2: One- to five-day-out high temperature forecasts within 3 degrees for calendar year 2015*

## Results of Low Temperature Forecast Comparison

Table 3 shows the mean absolute error for global one- to five-day-out low temperature forecasts. Mean absolute error tends to be higher than for high temperatures. Part of the difference is that a one-day-out low temperature forecast is defined as occurring overnight after the one-day-out high temperature, so on average occurs twelve hours after. Temperature forecast error, whether high or low temperature, increases as the forecast time is further out. However, that doesn't account for the entire difference, so in general low temperatures tend to be slightly less predictable than high temperatures. The number of low temperature forecasts within three degrees is also lower than the comparable high temperature forecasts within three degrees.

The Weather Channel, Weather Underground, and AccuWeather were statistically tied for first for one- to five-day-out global low temperature forecast mean absolute error, at about 3.06 degrees Fahrenheit. AccuWeather had the most low temperature forecasts within three degrees Fahrenheit at 66.59%. This is 14% more low temperature forecasts within three degrees than the last place provider, Dark Sky.

Rank	Provider	Mean Abs Error
1*	<i>The Weather Channel</i>	3.059
1*	<i>Weather Underground</i>	3.060
1*	<i>AccuWeather</i>	3.064
4	<i>Intellicast</i>	3.240
5	<i>Foreca</i>	3.351
6	<i>Dark Sky (forecast.io)</i>	3.742

Table 3: One- to five-day-out low temperature forecast mean absolute error for calendar year 2015

\* – difference not statistically significant at the 99% confidence level

Rank	Provider	% within 3°F
1	<i>AccuWeather</i>	66.59%
2	<i>The Weather Channel</i>	66.58%
3	<i>Weather Underground</i>	66.57%
4	<i>Intellicast</i>	65.08%
5	<i>Foreca</i>	62.80%
6	<i>Dark Sky (forecast.io)</i>	58.23%

Table 4: One- to five-day-out low temperature forecasts within 3 degrees for calendar year 2015

## Results of Overall Temperature Forecast Comparison

Combining the high and low temperature mean absolute error and averaging shows us the overall temperature forecast performance for each provider. This is shown in Table 5. AccuWeather had the lowest mean absolute error for overall temperature one- to five-day-out forecasts. The difference between first and last place is 0.796 degrees Fahrenheit, which is significant in many weather-sensitive industries.

Rank	Provider	Mean Abs Error
1	<i>AccuWeather</i>	2.932
2	<i>The Weather Channel</i>	2.964
3	<i>Weather Underground</i>	2.966
4	<i>Intellicast</i>	3.166
5	<i>Foreca</i>	3.224
6	<i>Dark Sky (forecast.io)</i>	3.728

Table 5: One- to five-day-out overall temperature forecast mean absolute error for calendar year 2015

Table 6 shows the combined average of high and low one- to five-day-out temperature forecasts within three degrees Fahrenheit. AccuWeather again has the highest average at 69.04% forecasts, while Dark Sky has the least at 58.43%.

Rank	Provider	% within 3°F
1	<i>AccuWeather</i>	69.04%
2	<i>The Weather Channel</i>	68.61%
3	<i>Weather Underground</i>	68.60%
4	<i>Intellicast</i>	66.84%
5	<i>Foreca</i>	65.13%
6	<i>Dark Sky (forecast.io)</i>	58.43%

Table 6: One- to five-day-out overall temperature forecasts within 3 degrees for calendar year 2015

## Forecast Collection Methodology

ForecastWatch collects forecasts from eight regions. The regions are the United States, Canada, Europe, Asia Pacific, Africa, Middle East, Central America, and South America. Collection in each region starts approximately late afternoon local time for the region. Collection times for each region and the number of stations in each region are shown in Table 7.

Region	Collection Time	Number of Stations
<i>United States</i>	22:00 UTC	792
<i>Canada</i>	21:40 UTC	41
<i>Europe</i>	16:00 UTC	193
<i>Asia Pacific</i>	08:00 UTC	65
<i>Africa</i>	15:30 UTC	13
<i>Middle East</i>	13:00 UTC	20
<i>Central America</i>	23:00 UTC	10
<i>South America</i>	21:00 UTC	14

Table 7: When forecasts were collected for each region and how many locations

For example, daily temperature forecasts were collected from each provider starting at 22:00 UTC (6 pm 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.

Forecasts were considered valid if they were complete (i.e. they contained a high and low temperature 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 were not considered invalid. However, forecast issues caused by system bugs or delivery problems (such as a -32768 degree high temperature) were declared invalid.

### ***Collecting AccuWeather Forecasts***

Forecasts from AccuWeather were collected from <http://www.accuweather.com> using the Premium, ad-free 15-day forecast page in January. The location parameters used to retrieve the specific forecasts were of the form {number}\_PC. However, beginning in February, forecasts were collected using the AccuWeather API at <http://api.accuweather.com>.

### ***Collecting Foreca Forecasts***

Forecasts from Foreca were collected from <http://www.foreca.com> using the ten-day forecast page. The location parameter used was the city and state of the observation location.

### ***Collecting Intellicast Forecasts***

Forecasts from Intellicast were collected from <http://www.intellicast.com> using the extended forecast page. The location parameter used was a site-specific code for the location.

### ***Collecting The Weather Channel Forecasts***

Forecasts from The Weather Channel were collected from <http://www.weather.com> using the ten-day forecast page. The latitude and longitude of the observation station were used to retrieve specific forecasts.

### ***Collecting Weather Underground Forecasts***

Forecasts from Weather Underground were collected from the Weather Underground API at <http://www.wunderground.com/api>. The location parameter used to retrieve specific forecasts was the ICAO code or SYNOP of the observation station.

### ***Collecting Dark Sky Forecasts***

Forecasts from Dark Sky were collected from the forecast.io API at <http://api.forecast.io>. The latitude and longitude of the observation station were used to retrieve specific forecasts.

## **Observation Collection Methodology**

Observational data was procured from the primary ASOS weather observation network in the United States, and comparable international equivalents. United States data were quality controlled by National Climatic Data Center (NCDC) systems and personnel prior to delivery to ForecastWatch via the Quality-Controlled Local Climatic Data (QCLCD) product data set. International data came from the Integrated Surface Database product. Both products consisted of both hourly and daily observation parameters.

The maximum temperature from the 7am to 7pm local time hourly observations was used to construct the 7am to 7pm (MOS) high temperature. The minimum temperature from the 7pm to 8am local time hourly observations was used to construct the 7pm to 8am (MOS) low temperature. No attempt to curve fit or otherwise determine an intra-hour temperature estimate was performed.

## Calculation Methodology

Table 8 shows the number of high temperature forecasts collected and compared for each provider for the one- to five-day-out forecasts. Table 9 shows the number of low temperature forecasts collected and compared for each provider for the one- to five-day-out forecasts. For both high and low temperature forecasts, there were around 1,950,000 forecasts used for each provider for the period. The percent of possible forecasts collected and compared is less than 100% because of invalid forecasts, problems in collecting forecasts successfully, and days in which observations were not available for a particular site.

Differences in the forecast count are due to a number of factors. First, invalid forecasts are removed. Second, occasionally a provider's website or feed would be off-line or not be complete due to network issues, production issues in the provider's forecast creation, or other issues. Finally, observations stations are down and don't provide data for a day or two every few months due to maintenance. Overall, around 93% of the possible forecasts and observations that could be compared for each provider were able to be compared.

Provider	Number of Forecasts	Percent of Possible Forecasts
<b>AccuWeather</b>	1,942,617	92.72%
<b>Dark Sky (forecast.io)</b>	1,952,921	93.21%
<b>Foreca</b>	1,944,641	92.82%
<b>Intellicast</b>	1,953,933	93.26%
<b>The Weather Channel</b>	1,954,117	93.27%
<b>Weather Underground</b>	1,946,694	92.92%

Table 8: Number of one- to five-day-out high temperature forecasts analyzed and percent of possible

Provider	Number of Forecasts	Percent of Possible Forecasts
<b>AccuWeather</b>	1,942,185	92.70%
<b>Dark Sky (forecast.io)</b>	1,952,921	93.21%
<b>Foreca</b>	1,944,236	92.80%
<b>Intellicast</b>	1,953,526	93.24%
<b>The Weather Channel</b>	1,953,711	93.25%
<b>Weather Underground</b>	1,946,320	92.90%

Table 9: Number of one- to five-day-out low temperature forecasts analyzed and percent of possible

## About ForecastWatch.com

ForecastWatch is the nation's premier weather forecast monitoring and assessment company. A full-service, technology consulting firm, ForecastWatch compiles weather forecasts and observations at more than 1,200 locations around the world, including the U.S., Canada, Europe, South America, Central America, Africa and Asia Pacific. ForecastWatch also maintains a historical database of over 600 million weather forecasts from a number of providers.

ForecastWatch data and analysis has been used by meteorologists, utilities and energy companies, the agriculture industry, futures traders, and others whose business success depends on being right about the weather. Our data meets the highest standard of scientific inquiry, and has been used in several peer-reviewed studies, including a paper published in the Monthly Weather Review.

ForecastWatch services have been used to evaluate weather forecast providers, improve decision-making where weather forecasts are used as input, improve weather forecasts by providing useful feedback, compare weather forecast performance between providers, educate customers with unbiased reporting, and improve the quality of weather forecast websites.