



ForecastWatch

Accuracy Defined

Analysis of Combined One- to Five- Day-Out Global Temperature Forecasts, January-June 2016

By ForecastWatch.com,
a Service of Intellovations, LLC

Eric Floehr
Owner
Intellovations, LLC

P.O. Box 164442
Columbus, Ohio 43216
eric@forecastwatch.com
<http://www.forecastwatch.com>



Executive Summary

Accurate weather forecasting is critical for many industries, both public and private, including recreation, construction, sports, energy and utility companies, safety and insurance. Understanding and evaluating the past is key to assessing future risk and opportunity.

Companies that specialize in weather prediction have the greatest investment of all in accuracy 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 in the health of their business.

This report analyses one- to five-day-out high, low, and overall temperature forecasts across the globe for the period of January through June of 2016. Over 11.4 million high and low temperature forecasts were collected from 1,145 United States and international locations. 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 degrees Fahrenheit**.

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, low, and overall temperature forecasts within three degrees. The Weather Channel, AccuWeather, and Weather Underground were statistically tied at the 99% confidence interval for lowest average absolute error in low temperature forecasts.

Results

High temperature forecasts

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

Findings: AccuWeather had the lowest mean absolute error. The difference between AccuWeather and the sixth place provider, Dark Sky, is nearly one full degree, which is significant.



Rank	Provider	Mean Abs Error
1	AccuWeather	2.844
2*	The Weather Channel	2.946
2*	Weather Underground	2.956
4	Foreca	3.033
5	Intellicast	3.153
6	Dark Sky (forecast.io)	3.864

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

Table 2 shows the percentage of one- to five-day-out high temperature forecasts within three degrees, lower or higher, of the observation.

Rank	Provider	% within 3°F
1	AccuWeather	70.71%
2	The Weather Channel	69.16%
3	Weather Underground	68.99%
4	Foreca	67.90%
5	Intellicast	67.39%
6	Dark Sky (forecast.io)	56.98%

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

Findings: AccuWeather had the highest percentage of forecasts within three degrees Fahrenheit of the observation, at 70.71%, whereas Dark Sky had the fewest, at 56.98%. This means that AccuWeather forecasted nearly 25% more high temperature forecasts within three degrees than Dark Sky, and 2.2% more than the second place provider Weather Channel.



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. 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 shows the mean absolute error for global one- to five-day-out low temperature forecasts.

Findings: The Weather Channel, AccuWeather, and Weather Underground were statistically tied for first for one- to five-day-out global low temperature forecast mean absolute error, at about 3.08 degrees Fahrenheit. Each had about a two-tenths of a degree lower error than Intellicast and Foreca, and a seven-tenths of a degree Fahrenheit lower error than Dark Sky.

Rank	Provider	Mean Abs Error
1*	The Weather Channel	3.081
1*	AccuWeather	3.082
1*	Weather Underground	3.088
4	Intellicast	3.261
5	Foreca	3.286
6	Dark Sky (forecast.io)	3.787

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

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

Findings: AccuWeather had the highest percentage of low temperature forecasts within three degrees, at 66.18%, whereas Dark Sky had the fewest, at 56.98%.

Rank	Provider	% within 3°F
1	AccuWeather	66.18%
2	The Weather Channel	65.99%
3	Weather Underground	65.86%
4	Intellicast	64.63%
5	Foreca	63.50%
6	Dark Sky (forecast.io)	56.98%

Table 4: One- to five-day-out low temperature forecasts within 3 degrees January-June 2016.

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 in **Table 5**.

Findings: AccuWeather had the lowest mean absolute error for overall temperature one- to five-day out forecasts. The difference between first and last place is 0.863 degrees Fahrenheit, which is significant.

Table 6 shows the combined average of high and low one- to five-day-out temperature forecasts within three degrees Fahrenheit.

Findings: AccuWeather had the highest percentage of one- to five-day-out overall temperature forecasts within three degrees, at 69.04%.



Rank	Provider	Mean Abs Error
1	AccuWeather	2.963
2	The Weather Channel	3.013
3	Weather Underground	3.022
4	Foreca	3.159
5	Intellicast	3.207
6	Dark Sky (forecast.io)	3.826

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

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

Table 6: One- to five-day-out overall temperature forecasts within 3 degrees for January-June 2016.

Methodology

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

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 within three degrees Fahrenheit.

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 7**, 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 7: Forecast collection times and regions.

Validity

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** (inaccurate or wrong) were not considered invalid. However forecasts issues caused by system errors or delivery problems (such as a -32768 degree high temperature) 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 were 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.

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.

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, around 91% of the possible forecasts and observations were available for comparison.



Provider	Number of Forecasts	Percent of Possible Forecasts
AccuWeather	954,957	91.65%
Dark Sky (forecast.io)	956,796	91.83%
Foreca	954,662	91.62%
Intellicast	956,827	91.83%
The Weather Channel	956,921	91.84%
Weather Underground	949,301	91.11%

Table 8: One- to five-day-out high temperature forecasts analysed and percent of possible forecasts.

Provider	Number of Forecasts	Percent of Possible Forecasts
AccuWeather	955,496	91.70%
Dark Sky (forecast.io)	956,796	91.83%
Foreca	955,210	91.68%
Intellicast	957,380	91.88%
The Weather Channel	957,475	91.89%
Weather Underground	949,913	91.17%

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



About ForecastWatch.com

ForecastWatch 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 maintain a historical database of over 600 million weather forecasts from a number of providers and provide 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.