



ForecastWatch

Accuracy Defined

Long-Term Analysis of Short-Term High Temperature Forecasts

(September 2006 through September 2015)

*By ForecastWatch.com, a Service of Intellovations, LLC
February 4, 2016*

Contact:

Eric Floehr
Owner
Intellovations, LLC

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

Executive Summary

Since September 2006, ForecastWatch.com has been collecting and verifying high temperature forecasts from the public websites of AccuWeather, CustomWeather, Intellicast, the National Weather Service, The Weather Channel, a forecast feed from the NWS National Digital Forecast Database, Weather Underground, and a forecast feed from Schneider Electric. This year, one- and two-day-out high temperature forecasts were evaluated over approximately 800 locations in the United States between September 1, 2014 and August 31, 2015.

A total of nearly 3.8 million one- and two-day-out high temperature forecasts were collected from seven weather forecast providers and compared with daily high temperature observations from the Automated Surface Observation Network. However, on March 12, 2015, CustomWeather blocked ForecastWatch from accessing its freely available forecasts at www.myforecast.com. Prior to blocking CustomWeather tended to be near the bottom of overall accuracy.

Over the nine years of this study, almost 40 million forecasts have been collected and compared with observations. Over that nine year period, the weather forecast providers' rankings have remained remarkably consistent. For every year of the study, Schneider Electric's forecasts have had the highest accuracy of all providers.

Why Accurate Temperature Forecasts Are Important

Accurate temperature forecasts are particularly valuable for electric and gas utilities. These weather forecasts are the main driver of their load forecasting, which is used for generation and purchasing decisions. Accurate temperatures pay many benefits including avoiding excessive base loads, reducing spinning reserves and reducing costs overall. Trading and power marketers also greatly benefit from accurate temperature forecasts resulting in smarter market transactions and higher revenues.

Public works departments and state DOTs also benefit from accurate temperature forecasts which are one of the drivers of accurate pavement temperature and frost forecasts. This allows better decision making in crew call-outs, pre-treating roads and overall public safety.

How Temperature Forecasts Are Evaluated

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

Root-mean-squared (RMS) error takes the square of each absolute error, averages all errors for the set of forecasts, and then takes the square root of the average. Samples are weighted linearly in estimating absolute error but as squares in estimating RMS error. Thus, a set of forecasts that show greater variance are penalized more than forecasts with consistent error. Consistent forecast accuracy is very important to industries and government agencies who make decisions based on the weather where lives and livelihoods are at stake.

Results of Temperature Forecast Comparison

The one-day-out high temperature forecast RMS errors in degrees Fahrenheit for the period September 1, 2014 through August 31, 2015 are shown in Table 1. For consumers of forecasts who rely on the most accurate forecasts and where temperature differences can significantly change decision-making, RMS error is usually the best accuracy measurement. With RMS, larger forecast error variability is penalized much more than small forecast error variability. The RMS Error column is the calculated RMS error in degrees Fahrenheit for the period. Rank is the ordered rank of providers for this period (a lower RMS is better).

One-Day-Out High Temperature Forecast Error for period 9/1/2014-8/31/2015

Rank	Provider	RMS Error (lower is better)
1	<i>Schneider Electric</i>	3.16
2	<i>AccuWeather</i>	3.25
3	<i>The Weather Channel</i>	3.40
4	<i>WX Underground</i>	3.41
5	<i>Intellicast</i>	3.54
6	<i>NDFD</i>	3.59
7	<i>NWS Web</i>	3.96
8	<i>CustomWeather</i>	*

Table 1: Results of one year one-day-out high temperature forecast analysis (lower is better)
* – CustomWeather blocked ForecastWatch March 12, 2015

Two-Day-Out High Temperature Forecast Error for period 9/1/2014-8/31/2015

Rank	Provider	RMS Error (lower is better)
1	<i>Schneider Electric</i>	3.60
2	<i>AccuWeather</i>	3.68
3	<i>The Weather Channel</i>	3.78
4	<i>WX Underground</i>	3.79
5	<i>Intellicast</i>	3.91
6	<i>NDFD</i>	4.04
7	<i>NWS Web</i>	4.36
8	<i>CustomWeather</i>	*

Table 2: Results of one year two-day-out high temperature forecast analysis (lower is better)
* – CustomWeather blocked ForecastWatch March 12, 2015

There was a 0.80 degree Fahrenheit (0.44 degree Celsius) difference between the least accurate provider (the NWS website) and the most accurate provider (Schneider Electric), which was slightly less than last year. The two-day-out high temperature forecast RMS errors are shown in Table 2. The two-day-out results are similar to the one-day-out results, with all providers having the same relative rank as the one-day-out results. Overall, error for two-day-out forecasts are greater than one-day-out forecasts, as expected. On average, there is an approximately 0.40 degree Fahrenheit (0.22 degree Celsius) increase in error between the one- and two-day-out high temperature forecasts.

The spread between first and last place was 0.76 degree Fahrenheit (0.42 degree Celsius) for two-day-out forecasts. For both one- and two-day-out high temperature forecasts, the difference in accuracy between first and last place providers is nearly a degree Fahrenheit. For companies where a degree of difference in a forecast can have life-and-death or monetary implications, it is clear that selecting the right forecast provider and continually monitoring their forecasts is important.

Nine Year Comparison

For the ninth year, Schneider Electric had the lowest RMS error. While the gap between first and last place remains large, all providers except Schneider Electric lowered their RMS error from the previous year. AccuWeather had the largest improvement in accuracy, moving from seventh place to second place in both one- and two-day-out accuracy. AccuWeather's one-day-out RMS error improved 0.59 degree Fahrenheit (0.33 degree Celsius) and two-day-out RMS error improved 0.54 degree Fahrenheit (0.30 degree Celsius). These were the largest year-over-year improvements in forecast accuracy for any provider over the nine year study period.

One-Day-Out High Temperature Forecast Error By Year

Year	Schneider Electric		Accu-Weather		TWC		WXU		Intellicast		NDFD		NWS Web		Custom-Weather	
2014-2015	3.16	1	3.25	2	3.40	3	3.41	4	3.54	5	3.59	6	3.96	7		
2013-2014	3.15	1	3.84	7	3.52	2	3.60	3	3.67	4	3.71	5	4.06	8	3.72	6
2012-2013	3.25	1	3.73	6	3.28	2	3.75	7	3.42	3	3.50	4	3.90	8	3.67	5
2011-2012	3.34	1	4.06	7	3.40	2	3.93	5	3.48	3	3.60	4	3.95	6	4.08	8
2010-2011	3.44	1	4.27	8	3.70	2	4.07	6	3.74	4	3.72	3	4.02	5	4.21	7
2009-2010	3.25	1	4.06	8	3.52	2	3.75	5	3.57	4	3.53	3	3.86	7	3.84	6
2008-2009	3.45	1	4.31	8	3.91	3	4.12	5	3.97	4	3.84	2	4.14	6	4.23	7
2007-2008	3.68	1	4.29	7	3.83	2			3.89	4	3.87	3	3.99	5	4.09	6
2006-2007	3.60	1	4.18	7	3.71	2			3.77	4	3.75	3	3.92	5	4.02	6

Table 3: Nine year comparison of one-day-out high temperature forecast error (lower is better)

Table 3 lists the one-day-out rankings and RMS error for all nine year-long periods. After accounting for AccuWeather's move from seventh to second place, other providers' relative rankings remained the same again this year. The Weather Channel, Weather Underground, Intellicast, the National Digital Forecast Database, and the NWS website forecasts all lowered their RMS error from the previous year. While the overall trend for all providers is positive, only AccuWeather and Weather Underground obtained their lowest forecast error over the nine year study period this year.

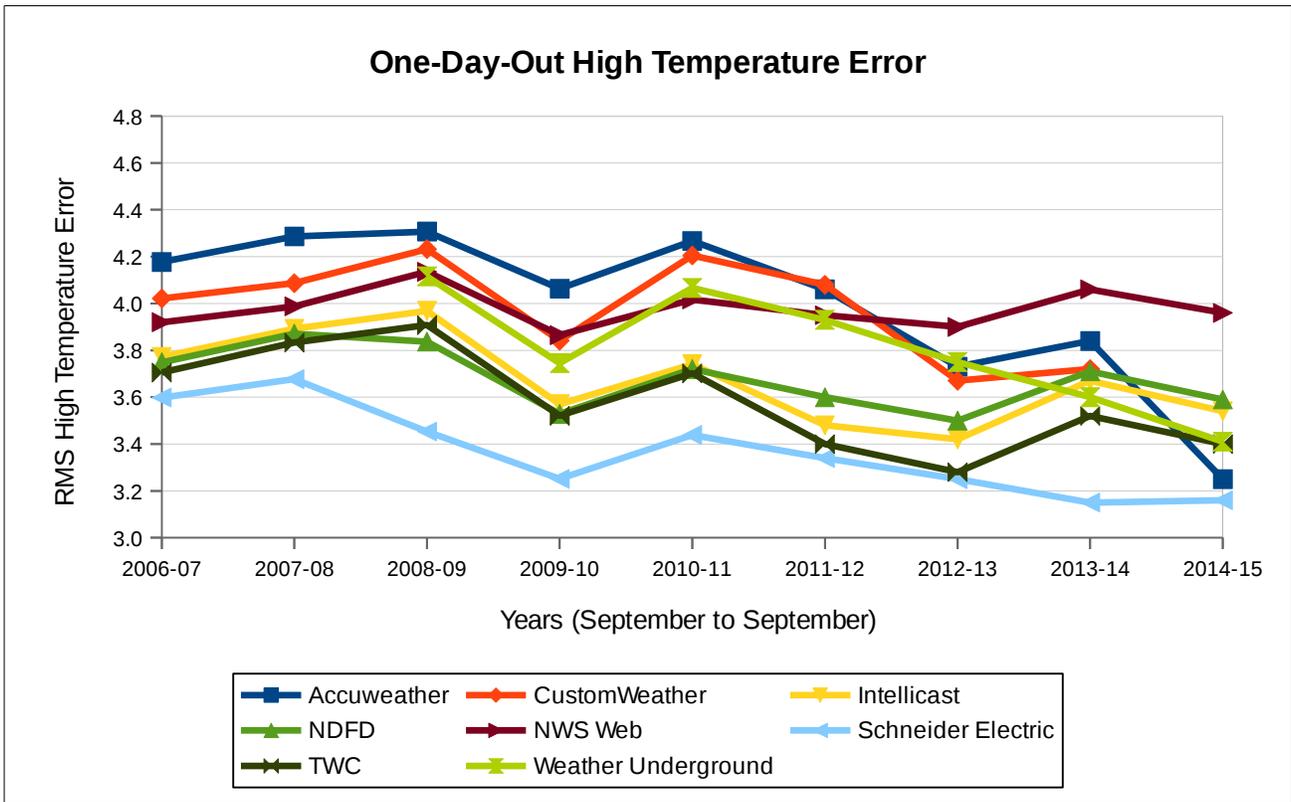
Two-Day-Out High Temperature Forecast Error By Year

Year	Schneider Electric		Accu-Weather		TWC		WXU		Intellicast		NDFD		NWS Web		Custom-Weather	
2014-2015	3.60	1	3.68	2	3.78	3	3.79	4	3.91	5	4.04	6	4.36	7		
2013-2014	3.59	1	4.22	7	3.92	2	4.00	3	4.05	4	4.08	5	4.38	8	4.11	6
2012-2013	3.70	1	4.13	6	3.72	2	4.17	7	3.83	3	3.89	4	4.23	8	4.01	5
2011-2012	3.85	1	4.42	7	3.88	2	4.37	6	3.94	3	4.03	4	4.35	5	4.43	8
2010-2011	3.92	1	4.63	8	3.70	3	4.45	6	4.20	4	4.12	2	4.39	5	4.57	7
2009-2010	3.70	1	4.45	8	3.97	2	4.17	5	4.00	4	3.97	3	4.25	6	4.27	7
2008-2009	3.97	1	4.76	8	4.34	3	4.59	6	4.43	4	4.33	2	4.57	5	4.67	7
2007-2008	4.20	1	4.67	7	4.32	2			4.39	4	4.37	3	4.47	5	4.50	6
2006-2007	4.13	1	4.64	7	4.20	2			4.29	3	4.29	4	4.44	5	4.54	6

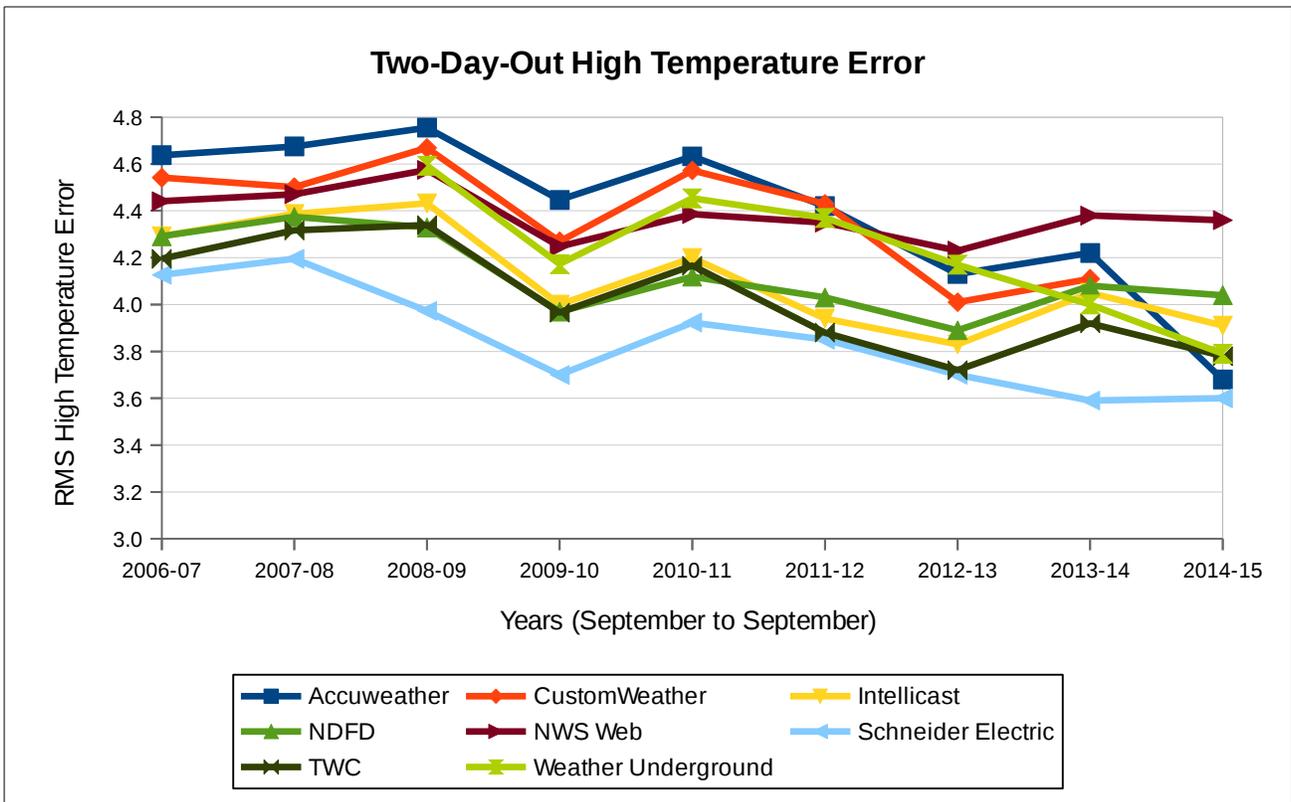
Table 4: Nine year comparison of two-day-out high temperature forecast error (lower is better)

Table 4 lists the two-day-out rankings and RMS error for all nine years. Like the one-day-out year-over-year results, AccuWeather moved from seventh to second, moving the other providers down one rank. For the two-day-out results, only AccuWeather and Weather Underground achieved their lowest RMS error for the entire nine years this year, while The Weather Channel tied their best year. Schneider Electric, for both one-day-out and two-day-out results, missed their best year this year by just 0.01 degree Fahrenheit (0.006 degree Celsius).

The two line graphs, Graph 1 and Graph 2, illustrate the changes in high temperature forecast error over the nine year study period and the accuracy differences between providers. Graph 1 shows the nine year history of high temperature forecast RMS error for one-day-out forecasts, while Graph 2 shows the nine year history for two-day-out forecasts. The y-axis of both graphs do not begin at zero, in order to highlight differences and trends in accuracy over the nine-year period, however the vertical scale of both graphs is identical.



Graph 1: Nine year graph of one-day-out high temperature forecast error (lower is better)



Graph 2: Nine year graph of two-day-out high temperature forecast error (lower is better)

Methodology of the Comparison

Daily high temperature forecasts were collected from each provider starting at 22:00 UTC (6pm Eastern Standard Time) and continuing until all forecasts are collected. For each location, forecasts from all providers were collected at the exact same time. Forecasts from AccuWeather were collected from the AccuWeather API at <http://apidev.accuweather.com/developers/>. Intellicast was collected from <http://www.intellicast.com>, The Weather Channel (TWC) forecasts were from <http://www.weather.com>, and the National Weather Service forecasts from <http://www.weather.gov>. The NDFD forecasts were collected using a SOAP interface at <http://www.weather.gov/ndfd/>. Weather Underground forecasts were collected from <http://www.wunderground.com>.

A one-day-out high temperature forecast is the forecast for the next day, whereas the two-day-out forecast is for the day after that. For example, for a forecast collected on January 1, 2015, the one-day-out high temperature forecast would be the forecast for January 2, 2015, and the two-day-out forecast would be the forecast for January 3, 2015.

Provider	Number of Forecasts	Percent of Possible Forecasts
AccuWeather	271,122	93.8%
CustomWeather	*	*
Intellicast	272,488	94.3%
NDFD	270,238	93.5%
NWS Web	262,413	90.8%
Schneider Electric	272,493	94.3%
The Weather Channel	270,912	93.7%
WX Underground	271,899	94.1%

Table 5: Number of one-day-out forecasts analyzed and percent of possible, by provider
* – CustomWeather blocked ForecastWatch March 12, 2015

Provider	Number of Forecasts	Percent of Possible Forecasts
AccuWeather	271,084	93.8%
CustomWeather	*	*
Intellicast	272,446	94.2%
NDFD	270,215	93.5%
NWS Web	262,568	90.8%
Schneider Electric	272,449	94.2%
The Weather Channel	270,876	93.7%
WX Underground	271,881	94.1%

Table 6: Number of two-day-out forecasts analyzed and percent of possible, by provider
* – CustomWeather blocked ForecastWatch March 12, 2015

For this study, the 24-hour high temperatures as reported by the observation stations in the ASOS/AWOS observation network maintained by the National Weather Service and the Federal Aviation Administration were used. These observations were collected from the Quality-Controlled Local Climatic Data (QCLCD) product from the National Climatic Data Center (NCDC).

The Root-Mean-Squared (RMS) error was calculated by subtracting the high temperature observation from the high temperature forecast, and that value was squared. The square root of the average of all the squared errors for each year was calculated, and that is the value shown in the report. Each yearly period includes forecasts collected September 1 through August 31 the following year.

There were around 270,000 forecasts used for each provider and year. 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.

Table 5 shows the number of forecasts collected and compared for each provider for one-day-out forecasts, and Table 6 shows the same for two-day-out forecasts. 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. For example, the relatively lower number of NWS web forecasts is primarily due to quality issues (invalid forecasts) and website availability issues (the weather.gov site was unavailable or could not provide a forecast). Finally, ASOS/AWOS stations are down and don't provide data for a day or two every few months due to maintenance. Overall, around 93.5% of the possible forecasts and observations that could be compared for each provider were able to be compared.

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 500 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. In 2003, ForecastWatch.com released the largest public weather forecast accuracy study undertaken to that point.

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.