Long-Term Analysis of Short-Term Probability of Precipitation Forecasts
(July 2006 through June 2015)

By ForecastWatch.com, a Service of Intellovations, LLC
February 11, 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 October 2006, ForecastWatch has been collecting and verifying probability of precipitation forecasts from CustomWeather, The National Weather Service, The Weather Channel, Weather Underground, and a non-public feed from Schneider Electric (collected at the same time as the public forecasts) against actual observations from the ASOS national weather observation network. This year, one-day-out probability of precipitation (POP) forecasts were evaluated over approximately 800 locations in the United States between July 1, 2014 and June 30, 2015.

A total of almost 1.1 million one-day-out probability of precipitation forecasts were collected over the year-long period, and approximately 12 million forecasts have been collected over the nine years of the study. However, on March 12, 2015, CustomWeather blocked ForecastWatch from accessing it's freely available forecasts at www.myforecast.com. Prior to blocking CustomWeather tended to be near the bottom of overall accuracy. For nine straight years, Schneider Electric had the best scores for the year-long POP accuracy period, as well as for the winter months only.

Importance of POP Forecasts

Many organizations rely on good precipitation forecasts. Concrete pouring and asphalting decisions depend on reliable rain forecasts. Missing the rain can result in costly re-dos, while forecasting rain when there ends up being none results in lost revenue opportunities. Public works departments and state DOTs rely on accurate snow and ice forecasts to know when to call out crews and pre-treat roads. For them reliable forecasts are critical for public safety, and for avoiding unnecessary and costly crew call outs.

Accurate precipitation forecasts are similarly important to electrical utilities, airports, golf courses, outdoor sports and recreation, and police/emergency management. Accurate precipitation forecasts add to the bottom line of weather-dependent businesses. And they help cities, counties and other organizations better meet their mission.

How POP Forecasts Are Evaluated

There are two components to measuring the accuracy of a probability-of-precipitation forecast. The first is accuracy. If, over the forecasts being measured, there was precipitation the same percentage of time as forecast, the forecast is said to be *accurate*. For example, if it rained 10% of the time the POP forecast called for a 10% chance of rain, the POP forecasts would be accurate. If, on average, there is precipitation three out of ten days at a given location, and the forecaster always predicted a 30% chance of precipitation every day, the forecaster would be accurate. While accurate, the forecast isn't useful.

The second measure of a POP forecast is *resolution*. A perfectly resolved POP forecaster would always predict no chance of precipitation for dry days, and 100% precipitation for days on which there was rain or snow. The forecaster above who always forecast a 30% chance of precipitation would be said to be fully *unresolved*. However, a forecaster who predicted 100% chance of precipitation on dry days, and zero percent on wet ones is still perfectly resolved, but completely inaccurate. While resolved, the forecast isn't useful.

Evaluating a POP forecast fully, therefore, must take both the accuracy and the resolution of the forecast into account. The calculation used to evaluate POP forecasts is called the Brier score. The Brier score takes both accuracy and resolution into account. A Brier score ranges from zero to one, with zero being perfectly accurate and resolved (0% POP forecast on dry days, 100% POP forecast on days with precipitation).

Results of Short-Term POP Forecast Comparison

The following tables detail the Brier scores for each weather forecast provider for one-day-out probability of precipitation forecasts. Table 1 shows Brier scores for the one-year period of July 1, 2014 through June 30, 2015. This year we made a change to how we interpret National Weather Service POP forecasts (note b). Prior to this year, we interpreted NWS forecasts without a POP forecast as a POP forecast of 0%. However, that is not true, as the NWS will in certain regions and under certain conditions, not show a POP if their forecasted POP is 20% or lower. Therefore, interpreting a missing POP as 0% chance of precipitation is incorrect. This year, we only evaluated NWS forecasts that contained a POP.

This year, Schneider Electric again had a lower Brier score than all public forecasts collected by ForecastWatch in the United States, and has maintained that rank for all nine years of the study.

Rank	Provider	Brier Score (lower is better)
1	*Schneider Electric*	0.1078
2	*Weather Underground*	0.1223
3	*The Weather Channel*	0.1230
4	*National Weather Service*	0.2111[b]
5	*CustomWeather*	*

Table 1: Results of one year (July 2014–June 2015) short-term POP forecast analysis (lower is better)
* CustomWeather blocked ForecastWatch on March 12, 2015

Many businesses, governments, and individuals are especially interested in winter forecasts. Preparations for snow or ice, such as changing business processes, salting roads in advance, or keeping employees on standby are real costs. Better prediction of winter precipitation results in money saved, better service, and improved planning for those businesses, governments, and individuals.

POP scores for the winter months of December 2014 through February 2015 have been broken out in Table 2. Schneider Electric forecasts led all collected public forecasts for the winter months, and it has led for all nine years of the study.

Rank	Provider	Brier Score (lower is better)
1	*Schneider Electric*	0.0958
2	*The Weather Channel*	0.1152
3	*Weather Underground*	0.1154
4	*CustomWeather*	0.1155
5	*National Weather Service*	0.2035[b]

Table 2: Results of short-term POP forecast analysis for winter 2014-2015 (lower is better)

Nine Year Comparison

This is the ninth year that this POP study has been conducted. Table 3 shows full-year Brier scores for each of the nine years. This year, Schneider Electric and The Weather Channel had their lowest (and best) Brier scores for the entire period. This is the fifth year in a row Schneider Electric has improved their Brier score for POP forecasts, and is unmatched by any other provider in the study.

Year	Schneider Electric		WX Underground		The Weather Channel		NWS		Custom-Weather	
Jul 2014 – Jun 2015	0.1078	1	0.1223	2	0.1230	3	0.2111[b]	4		
Jul 2013 – Jun 2014	0.1090	1	0.1247	3	0.1415	5	0.1244	2	0.1252	4
Jul 2012 – Jun 2013	0.1115	1	0.1209	2	0.1315	5	0.1242	4	0.1217	3
Jul 2011 – Jun 2012	0.1163	1	0.1778	5	0.1288	4	0.1288	3	0.1258	2
Jul 2010 – Jun 2011	0.1260	1			0.1390	3	0.1393	4	0.1351	2
Jul 2009 – Jun 2010	0.1330	1	0.1534	5	0.1468	3	0.1486	4	0.1386	2
Jul 2008 – Jun 2009	0.1290	1	0.1458	5	0.1399	3	0.1428	4	0.1329	2
Jul 2007 – Jun 2008	0.1300	1			0.1459	3	0.2182[a]	4	0.1379	2
Oct 2006 – Jun 2007	0.1219	1			0.1382	3	0.1903[a]	4	0.1271	2

Table 3: Nine-year comparison of one year short-term POP forecast analysis (lower is better)

Note that the Weather Underground redesigned their site in February 2011 which temporarily removed POP forecasts for the June 2010 – June 2011 period. Because we could not collect a full year of forecasts for them, they were not included in the study that period. CustomWeather blocked ForecastWatch collection on March 12, 2015, so a full year could not be compared.

Winter Season	Schneider Electric		The Weather Channel		WX Underground		Custom-Weather		NWS	
Dec 2014 – Feb 2015	0.0958	1	0.1152	2	0.1154	3	0.1155	4	0.2035[b]	5
Dec 2013 – Feb 2014	0.0971	1	0.1569	5	0.1174	4	0.1164	3	0.1128	2
Dec 2012 – Feb 2013	0.1052	1	0.1307	5	0.1206	2	0.1216	3	0.1237	4
Dec 2011 – Feb 2012	0.1008	1	0.1198	5	0.1167	2	0.1180	4	0.1177	3
Dec 2010 – Feb 2011	0.1145	1	0.1272	4			0.1252	3	0.1249	2
Dec 2009 – Feb 2010	0.1183	1	0.1277	3	0.1345	5	0.1223	2	0.1303	4
Dec 2008 – Feb 2009	0.1194	1	0.1275	3	0.1458	5	0.1204	2	0.1300	4
Dec 2007 – Feb 2008	0.1224	1	0.1445	3			0.1363	2	0.2017[a]	4
Dec 2006 – Feb 2007	0.1104	1	0.1846	4			0.1228	2	0.1351[a]	3

Table 4: Nine-year comparison of winter season short-term POP forecast analysis (lower is better)

Please also note (a) that The National Weather Service scores for 2006-2007 and 2007-2008 were calculated on a small percentage of POP forecasts relative to the other providers, due to NWS website forecast issues which were corrected in 2008 when the NWS began migrating to point forecasts on its website. For 2014-2015 only National Weather Service scores forecasts that provided a POP forecast were evaluated, no attempt to assign a POP forecast to forecasts without a POP were made.

Table 4 shows the nine year winter month data. The winter-only rankings are similar to the full-year rankings. Schneider Electric had the lowest Brier score this year, and has improved its score six out of the last seven years. This year, Schneider Electric also achieved the lowest Brier score measured in the nine years of the study.

Methodology of the Comparison

Brier scores are more useful the larger number of forecasts and observations there are to calculate. This study evaluated POP forecasts for 792 locations within the United States over the period of July 1, 2014 through June 30, 2015. Forecasts were collected 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. A forecast without a probability of precipitation value was not evaluated and no attempt at assigning a POP value was attempted. Forecasts with invalid POP percentages (less than 0% or greater than 100%) were declared invalid and not included. Forecasts declared invalid through ForecastWatch's extensive audit process (i.e. bad high or low temperatures) were also not included.

Provider	Number of Forecasts	Percent of Possible Forecasts
CustomWeather	*	*
NWS	262,171	90.7%
Schneider Electric	272,161	94.1%
The Weather Channel	270,571	93.6%
Weather Underground	270,810	93.7%

Table 5: Number of forecasts collected during the period and percent of possible, by provider

Provider	Number of Forecasts	Percent of Possible Forecasts
CustomWeather	64,594	90.6%
NWS	63,048	88.5%
Schneider Electric	64,946	91.1%
The Weather Channel	64,932	91.1%
Weather Underground	64,806	90.9%

Table 6: Number of forecasts collected during the winter period and percent of possible, by provider

The forecasts were compared against daily 24-hour local time precipitation measured by the ASOS/AWOS observation network maintained by the National Weather Service and the Federal Aviation Administration. If 0.01 inches of liquid-equivalent precipitation or more fell that in that day, it was considered to have been a precipitation event. 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.

Despite the occasional collection problem or invalid forecasts, collection rates for the full year were above 93% for all providers except The National Weather Service. The National Weather Service website does not publish POP values for about 20 locations which don't produce a point forecast, which keeps their rates a little lower than the other providers. Collection rates for the 90 day winter period from December 1, 2014 through February 28, 2015 are also above 90% except for NWS.

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.