ForecastWatch
Accuracy Defined

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

January–December 2023

Intellovations, LLC

6724 Perimeter Loop Road, #275
Dublin, OH 43017
info@forecastwatch.com
http://www.forecastwatch.com
855-609-9609

Introduction

Everyone is affected by the weather. Most people are interested in knowing what the weather will be like in the coming days. Relying on weather forecasts helps people plan for severe and deadly weather events to stay safe and keep property protected. Weather forecasts also help everyone make normal daily plans, such as for commutes, picnics, hikes, day trips, selection of appropriate clothing, environmental allergies, or other health-related issues, and more. Substantial financial impacts can be at stake for businesses that depend highly on weather forecasts and need to plan for weather events, such as businesses that depend on energy, agriculture, and transportation. In the past four years (2020–2023), the United States alone saw a total of 88 individual billion-plus-dollar weather and climate disasters (22 in 2020, 20 in 2021, 18 in 2022, and a record-smashing 28 in 2023). Before 2020, the most for a single year was 19 in 2017.

This leads to people and businesses looking to weather forecasts—and their accuracy—more than ever to try to mitigate losses by preparing for what is coming. If weather forecasts are not accurate, damages and fatalities from severe and extreme weather increase, and the perception of forecasting value decreases. It is important for businesses that rely on weather forecasts to evaluate the accuracy of past forecasts to help evaluate risk and opportunities. It is also imperative for weather forecast providers themselves to evaluate the accuracy of past forecasts to identify areas where improvement is needed. Not only is this improvement important to sustain relationships and trust with customers and users of their forecasts, but it is also necessary to save lives and minimize damages from severe and extreme weather events.

This report provides a holistic analysis of weather forecast performance for 13 different metrics for global forecasts made one to five days in advance for the year 2023. The 13 different metrics include eight metrics for high and low temperatures, one for wind speed, two for precipitation, and two for opacity (sky cover). This report examines forecasts from 22 different weather forecast providers. Fifteen providers forecasted globally and nine forecasted for all 13 metrics. These same nine providers also forecasted globally.

Each provider was given one point for each month that provider finished in first place for a given metric at a given location. For ties, a point was given to each provider involved in the tie. The total number of points was counted and used to rank forecast providers. By evaluating forecast accuracy on a per-location granularity, we come up with a holistic approach that evaluates which forecast provider offers the most accurate forecasts at the most locations around the globe.

In this analysis, the overall most accurate weather forecast provider globally is Microsoft. Microsoft also had the highest metrics for temperatures (other than the mean absolute errors for MOS high

and 24-hour low temperatures) and opacity percentage. The Weather Channel was the most accurate for the previous two temperature metrics mentioned, as well as precipitation equitable threat score and wind speed mean absolute error, while Foreca/Vaisala was the most accurate for probability of precipitation Brier score. Weerplaza was the most accurate in the text sky cover category.

One- to Five-Day-Out Skill Summary

Microsoft scored the most first-place finishes (50,624) globally in 2023 for all 13 metrics and one through five days out. The Weather Channel had the second most with 39,145, and Foreca/Vaisala had the third most with 12,917. Weerplaza was fourth with 8,696, AccuWeather had 7.364, and CFAN had 5.482. The other 16 providers had fewer than 4,000 each.

Percentagewise, Microsoft had 34.73%, The Weather Channel claimed 26.86%, and Foreca/Vaisala took 8.86%. When weighting metrics such that each of the five weather variables has equal importance. these numbers are 37.01% for Microsoft, 27.37% for The Weather Channel, and 9.55% for Foreca/Vaisala.

In these tables, “w/in 3°F” means within 3 Fahrenheit degrees (an increment of 3 degrees), “MAE” stands for mean absolute error, and “ETS” stands for equitable threat score. The Brier score is for POP (probability of precipitation), while the equitable threat score is also for precipitation.

The most first-place finishes by Microsoft included the most in MOS and 24-hour high temperatures within 3 Fahrenheit degrees, MOS and 24-hour low temperature within 3 Fahrenheit degrees, 24-hour high mean absolute error, MOS low temperature mean absolute error, and opacity percent mean absolute error. The second greatest number of first-place finishes obtained by Microsoft were in MOS high temperature mean absolute error, 24-hour low temperature mean absolute error, precipitation (POP) Brier score, and wind speed mean absolute error. Microsoft had the third greatest number of first-place finishes in the two remaining metrics: precipitation equitable threat score and sky cover/opacity within one category.

The Weather Channel claimed the highest number of first-place finishes for MOS high temperature mean absolute error, 24-hour low temperature mean temperature error, precipitation equitable threat score, and wind speed mean absolute error, while Foreca/Vaisala had the most for precipitation (POP) Brier score, and Weerplaza had the most for sky cover/opacity within one category.

Table 1 shows the number of first-place finishes by forecast provider and metric, along with the total number of first-place finishes by forecast provider, for the nine providers that provided forecast data globally and for all 13 metrics. Table 2 shows the percentage of first-place finishes for these

providers. In both tables, the yellow highlighting indicates the top provider in each metric and total. Table 3 shows the overall rank of each of these 13 providers, by metric and overall number/percentage of first-place finishes.

Table 4 indicates the reasons why each of the other 12 providers were not eligible to receive all possible first-place points, either because they did not provide forecast data globally or did not provide forecast data for all 13 of the metrics, or a combination of those factors.

Provider	Total First-Place Finishes	MOS High w/in 3°F	MOS High MAE	MOS Low w/in 3°F	MOS Low MAE	24h High w/in 3°F	24h High MAE	24h Low w/in 3°F	24h Low MAE	Brier Score	ETS	Wind Speed MAE	Text Sky Cover	Opacity % MAE
AccuWeather	**7364**	832	606	807	629	654	565	788	496	30	72	1132	536	217
AerisWeather	**2442**	97	33	114	29	75	26	81	26	23	195	10	1017	716
Foreca/Vaisala	**12917**	818	352	582	251	798	370	1148	583	3961	3102	50	646	256
Microsoft	**50624**	4679	3794	5629	4660	5144	4268	3974	2605	3620	2708	3496	2051	3996
OpenWeather	**1119**	4	0	6	0	0	0	4	0	63	170	64	539	269
Pirate Weather	**1365**	6	5	7	1	0	0	117	37	94	345	48	562	143
The Weather Channel	**39145**	3550	4264	2748	3624	2524	3332	3326	4139	2698	3269	4310	591	770
Weatherbit	**2192**	65	21	150	95	47	14	241	174	49	92	128	526	590
World Weather Online	**3735**	1	0	2	0	0	0	4	0	6	63	16	1413	2230

Table 1: Number of First-Place Finishes by Provider and Metric for Providers That Provided Forecast Data Globally for All Thirteen Metrics and All Five Days Out

Provider	Total First-Place Finishes	MOS High w/in 3°F	MOS High MAE	MOS Low w/in 3°F	MOS Low MAE	24h High w/in 3°F	24h High MAE	24h Low w/in 3°F	24h Low MAE	Brier Score	ETS	Wind Speed MAE	Text Sky Cover	Opacity % MAE	Normalized
AccuWeather	**5.05%**	7.53%	5.64%	7.36%	5.86%	5.98%	5.26%	7.00%	4.55%	0.28%	0.67%	10.54%	3.15%	2.36%	5.21%
AerisWeather	**1.68%**	0.88%	0.31%	1.04%	0.27%	0.69%	0.24%	0.72%	0.24%	0.21%	1.81%	0.09%	5.99%	7.79%	1.82%
Foreca/Vaisala	**8.86%**	7.40%	3.28%	5.30%	2.34%	7.29%	3.44%	10.20%	5.35%	36.86%	28.82%	0.47%	3.80%	2.79%	9.55%
Microsoft	**34.73%**	42.34%	35.32%	51.31%	43.39%	47.02%	39.72%	35.32%	23.91%	33.69%	25.16%	32.54%	12.07%	43.50%	37.01%
OpenWeather	**0.77%**	0.04%	0.00%	0.05%	0.00%	0.00%	0.00%	0.04%	0.00%	0.59%	1.58%	0.60%	3.17%	2.93%	0.95%
Pirate Weather	**0.94%**	0.05%	0.05%	0.06%	0.01%	0.00%	0.00%	1.04%	0.34%	0.87%	3.21%	0.45%	3.31%	1.56%	1.06%
The Weather Channel	**26.86%**	32.13%	39.70%	25.05%	33.74%	23.07%	31.01%	29.56%	37.99%	25.11%	30.38%	40.12%	3.48%	8.38%	27.37%
Weatherbit	**1.50%**	0.59%	0.20%	1.37%	0.88%	0.43%	0.13%	2.14%	1.60%	0.46%	0.85%	1.19%	3.10%	6.42%	1.69%
World Weather Online	**2.56%**	0.01%	0.00%	0.02%	0.00%	0.00%	0.00%	0.04%	0.00%	0.06%	0.59%	0.15%	8.32%	24.27%	3.36%

Table 2: Percentage of First-Place Finishes by Provider and Metric for Providers That Provided Forecast Data Globally for All Thirteen Metrics and All Five Days Out

Provider	Rank	MOS High w/in 3°F	MOS High MAE	MOS Low w/in 3°F	MOS Low MAE	24h High w/in 3°F	24h High MAE	24h Low w/in 3°F	24h Low MAE	Brier Score	ETS	Wind Speed MAE	Text Sky Cover	Opacity % MAE	Normalized Rank
Microsoft	**1**	1	2	1	1	1	1	1	2	2	3	2	3	1	1
The Weather Channel	**2**	2	1	2	2	2	2	2	1	3	1	1	8	3	2
Foreca/Vaisala	**3**	4	5	4	6	4	6	3	5	1	2	11	7	7	3
AccuWeather	**5**	3	4	3	3	5	5	5	6	9	10	3	12	8	5
World Weather Online	**7**	20	21	19	19	17	19	16	19	14	11	13	4	2	8
AerisWeather	**10**	7	12	8	12	7	13	10	12	10	6	14	5	4	9
Weatherbit	**11**	9	14	7	11	10	14	6	8	7	9	7	13	5	10
Pirate Weather	**12**	16	18	14	16	17	19	9	11	5	5	12	10	9	11
OpenWeather	**13**	18	21	15	19	17	19	16	19	6	7	9	11	6	12

Table 3: Rank for First-Place Finishes by Provider and Metric for Providers That Provided Forecast Data Globally for All Thirteen Metrics and All Five Days Out

Provider	Regions with Forecast Data If Not Global	Metrics for Which Forecast Data Not Provided
Australian BOM	Australia Only	Wind Speed and Opacity
BBC	UK Only	POP, Wind Speed, and Opacity
CFAN	US Only	Wind Speed, Precipitation (POP and ETS), Sky Cover and Opacity
Environment Canada	Canada Only	Wind Speed and Opacity
Global News	Canada Only	Wind Speed and Opacity
MeteoFrance	France Only	Opacity
Moji		Opacity
National Weather Service	US Only	Wind Speed and Opacity
UK MetOffice		Opacity
Weather News		Wind Speed, Sky Cover, and Opacity
Weerplaza		Opacity
Wetter		Opacity
yr.no		Opacity

Table 4: Forecast Providers Not Eligible for All First-Place Points and Reason(s)

Table 5 and Table 6 indicate the total number and percentage, respectively, of first-place finishes for these other 13 providers. Table 7 indicates the overall rank of these other 13 providers that were inhibited in at least one way from receiving consideration for all the potential first-place finishes.

For MOS high temperatures within 3 Fahrenheit degrees, Microsoft (42.34%) and The Weather Channel (32.13%) together claimed about 75% of the first-place finishes. No other provider had more than 7.53% (AccuWeather), with 7.40% going to Foreca/Vaisala, for a total of 89.4% of all first-place finishes to these four providers of the 22 total providers. These same four providers together had 84% of the first-place finishes for MOS high temperature mean absolute error, 89% for MOS low temperature within 3 Fahrenheit degrees (over 51% of those for Microsoft), 85% for MOS low temperature mean absolute error, 83% for 24-hour high temperature within 3 Fahrenheit degrees, 79% for 24-hour high temperature mean absolute error, 82% for 24-hour low temperature within 3 Fahrenheit degrees, and 72% for 24-hour low temperature mean absolute error. In every one of those temperature metrics, Microsoft and The Weather Channel combined for between 62% and 77% of the total number of first-place finishes.

Provider	Total First-Place Finishes	MOS High w/in 3°F	MOS High MAE	MOS Low w/in 3°F	MOS Low MAE	24h High w/in 3°F	24h High MAE	24h Low w/in 3°F	24h Low MAE	Brier Score	ETS	Wind Speed MAE	Text Sky Cover	Opacity % MAE
Australian BOM	**55**	0	5	0	3	0	5	0	4	9	29		0	
BBC	**93**	24	8	3	0	9	4	1	0		9		35	
CFAN	**5482**	354	191	364	297	1159	1012	1017	1088	0	0			
Environment Canada	**130**	2	1	2	0	0	0	1	3	2	3		116	
Global News	**414**	0	163	0	156	0	35	0	3	0	5		52	
MeteoFrance	**623**	25	8	18	1	28	9	44	5	0	0	0	485	
Moji	**867**	69	37	53	12	35	29	30	6	0	143	104	349	
National Weather Service	**326**	26	25	22	6	53	28	27	11	19	31		78	
UK MetOffice	**3672**	398	213	377	219	332	186	194	99	46	25	770	813	
Weather News	**524**	6	102	3	119	2	91	18	177	6	0			
Weerplaza	**8696**	54	771	66	525	58	673	223	1355	105	458	393	4015	
Wetter	**3268**	18	134	15	112	9	94	12	85	14	31	167	2577	
yr.no	**708**	22	8	3	1	13	4	2	0	0	12	55	588	

Table 5: Number of First-Place Finishes by Provider and Metric for Providers That Did Not Provide Forecast Data Globally for All Thirteen Metrics and/or All Five Days Out

Provider	Total First-Place Finishes	MOS High w/in 3°F	MOS High MAE	MOS Low w/in 3°F	MOS Low MAE	24h High w/in 3°F	24h High MAE	24h Low w/in 3°F	24h Low MAE	Brier Score	ETS	Wind Speed MAE	Text Sky Cover	Opacity % MAE	Normalized %
Australian BOM	**0.04%**	0.00%	0.05%	0.00%	0.03%	0.00%	0.05%	0.00%	0.04%	0.08%	0.27%		0.00%		0.04%
BBC	**0.06%**	0.22%	0.07%	0.03%	0.00%	0.08%	0.04%	0.01%	0.00%		0.08%		0.21%		0.08%
CFAN	**3.76%**	3.20%	1.78%	3.32%	2.77%	10.59%	9.42%	9.04%	9.99%	0.00%	0.00%				N/A
Environment Canada	**0.09%**	0.02%	0.01%	0.02%	0.00%	0.00%	0.00%	0.01%	0.03%	0.02%	0.03%		0.68%		0.15%
Global News	**0.28%**	0.00%	1.52%	0.00%	1.45%	0.00%	0.33%	0.00%	0.03%	0.00%	0.05%		0.31%		0.23%
MeteoFrance	**0.43%**	0.23%	0.07%	0.16%	0.01%	0.26%	0.08%	0.39%	0.05%	0.00%	0.00%	0.00%	2.85%		0.63%
Moji	**0.59%**	0.62%	0.34%	0.48%	0.11%	0.32%	0.27%	0.27%	0.06%	0.00%	1.33%	0.97%	2.05%		0.86%
National Weather Service	**0.22%**	0.24%	0.23%	0.20%	0.06%	0.48%	0.26%	0.24%	0.10%	0.18%	0.29%		0.46%		0.23%
UK MetOffice	**2.52%**	3.60%	1.98%	3.44%	2.04%	3.03%	1.73%	1.72%	0.91%	0.43%	0.23%	7.17%	4.79%		3.38%
Weather News	**0.36%**	0.05%	0.95%	0.03%	1.11%	0.02%	0.85%	0.16%	1.62%	0.06%	0.00%				N/A
Weerplaza	**5.97%**	0.49%	7.18%	0.60%	4.89%	0.53%	6.26%	1.98%	12.44%	0.98%	4.26%	3.66%	23.63%		7.70%
Wetter	**2.24%**	0.16%	1.25%	0.14%	1.04%	0.08%	0.87%	0.11%	0.78%	0.13%	0.29%	1.55%	15.17%		3.61%
yr.no	**0.49%**	0.20%	0.07%	0.03%	0.01%	0.12%	0.04%	0.02%	0.00%	0.00%	0.11%	0.51%	3.46%		0.83%

Table 6: Percentage of First-Place Finishes by Provider and Metric for Providers That Did Not Provide Forecast Data Globally for All Thirteen Metrics and/or All Five Days Out

Provider	Rank	MOS High w/in 3°F	MOS High MAE	MOS Low w/in 3°F	MOS Low MAE	24h High w/in 3°F	24h High MAE	24h Low w/in 3°F	24h Low MAE	Brier Score	ETS	Wind Speed MAE	Text Sky Cover	Opacity % MAE	Normalized Rank
Weerplaza	**4**	10	3	9	4	8	4	7	3	4	4	5	1		4
CFAN	**6**	6	7	6	5	3	3	4	4	17	20				N/A
UK MetOffice	**8**	5	6	5	7	6	7	8	9	8	15	4	6		7
Wetter	**9**	15	9	13	10	14	8	15	10	12	12	6	2		6
Moji	**14**	8	11	10	13	11	11	12	14	17	8	8	15		13
yr.no	**15**	14	15	16	16	13	17	18	19	17	16	10	9		14
MeteoFrance	**16**	12	15	12	16	12	15	11	15	17	20	15	14		15
Weather News	**17**	16	10	16	9	16	9	14	7	14	20				N/A
Global News	**18**	21	8	21	8	17	10	21	17	17	18		18		16
National Weather Service	**19**	11	13	11	14	9	12	13	13	11	12		17		16
Environment Canada	**20**	19	20	19	19	17	19	19	17	16	19		16		18
BBC	**21**	13	15	16	19	14	17	19	19		17		19		19
Australian BOM	**22**	21	18	21	15	17	16	21	16	13	14		20		20

Table 7: Rank of Number/Percentage of First-Place Finishes by Provider and Metric for Providers That Did Not Provide Forecast Data Globally for All Thirteen Metrics and/or All Five Days Out

For precipitation (POP) Brier score, first-place finishes were basically distributed among Foreca/Vaisala (36.86%), Microsoft (33.69%), and The Weather Channel (25.11%), which, when combined, accounted for 95.66% of the total first-place finishes for this metric. All other providers recorded less than 1% of the total (BBC was the only provider in this analysis that did not forecast POP).

There was a little more variety for the precipitation equitable threat score, which was forecast by all providers in the analysis, but Foreca/Vaisala, Microsoft, and The Weather Channel still all combined for 84.4% of the total number of first-place finishes (30.38% for The Weather Channel, 28.82% for Foreca/Vaisala, and 25.16% for Microsoft). Weerplaza was in a distant fourth with 4.26%, while Pirate Weather had 3.21%.

Seven providers—the Australian BOM, BBC, CFAN, Environment Canada, Global News, National Weather Service, and Weather News—did not provide forecast data for wind speed and thus were not able to receive first-place finishes for wind speed mean absolute error. The Weather Channel (40.12%) and Microsoft (32.54%) combined for nearly 73% of the first-place finishes for wind speed mean absolute error, with AccuWeather claiming 10.54% of the share in third. The UK Met Office claimed another 7.17%, and Weerplaza claimed 3.66% in fifth. None of the other ten providers had more than 1.55% of the share of first-place finishes for wind speed.

Two providers, CFAN and Weather News, did not have provided forecasts for text sky cover, and thus were not evaluated for sky cover/opacity within one category. This metric, unsurprisingly, was more divided in terms of first-place finishes; the top three providers had just over half the total (Weerplaza with 23.63%, Wetter with 15.17%, and Microsoft with 12.07%). World Weather Online was fourth with 8.32%, AerisWeather was fifth with 5.99%, and the UK MetOffice was sixth with 4.79%.

Only nine providers had forecasts evaluated for opacity percentage mean absolute error, because most providers don't provide a specific sky cover percentage. These nine providers were AccuWeather, AerisWeather, Foreca/Vaisala, Microsoft, OpenWeather, Pirate Weather, The Weather Channel, Weatherbit, and World Weather Online, and are the nine providers found in Table 1, Table 2, and Table 3. Microsoft (43.50%) and World Weather Online (24.27%) combined garnered around two-thirds of the total first-place finishes for this metric. The Weather Channel was third with 8.38%, AerisWeather had 7.79%, and Weatherbit was fifth with 6.42%. The other four providers had between 1.5% and 3% of the share of first-place finishes.

For all metrics combined, the share of first-place finishes earned by Microsoft (34.73%) and The Weather Channel (26.86%) total comprised nearly 62% of all first-place finishes. The other 20 providers earned the other 38.4%. The nine providers that forecasted both globally and for all 13 metrics earned a total of 83% of all first-place finishes, and those that forecasted globally but not

necessarily for all metrics (15 of the 22 providers) earned a total of 95.1% of all first-place finishes. Providers that gave global forecast data but not for all 13 metrics include Moji, the UK Met Office, Weather News, Weerplaza, Wetter, and yr.no. There were no providers that forecasted for all metrics but not globally. Providers that did not provide global forecast data include the Australian BOM, BBC, CFAN, Environment Canada, Global News, MeteoFrance, and the National Weather Service.

Weather Metrics Analyzed

Temperature Metrics

ForecastWatch analyzed Model Output Statistics (MOS) and 24-hour high and low temperature forecasts. The metrics analyzed for these temperatures were the percentage of forecasts within three Fahrenheit degrees and mean absolute errors, for a total of eight temperature metrics:

- MOS high temperature within 3 Fahrenheit degrees
- MOS low temperature within 3 Fahrenheit degrees
- 24-hour high temperature within 3 Fahrenheit degrees
- 24-hour low temperature within 3 Fahrenheit degrees
- MOS high temperature mean absolute error
- MOS low temperature mean absolute error
- 24-hour high temperature mean absolute error
- 24-hour low temperature mean absolute error

Forecast Time Period Definitions

The day-part forecast period used for the MOS high metrics is defined as the time period typically defined as the "daytime" in local standard time of the observation location. This is defined as 7 a.m. to 7 p.m. local standard time. The night-part forecast period (used for the MOS low metrics) is similar, but it uses the period of 7 p.m. to 8 a.m. local standard time. The 24-hour temperature forecast period is midnight to midnight local standard time.

For all but the MOS low temperature forecasts, the forecasted days out is determined by subtracting the forecast day from the day the forecast was collected. For example, a MOS or 24-hour high temperature forecast for March 10, 2023, collected on March 3, 2023, would be a seven-day-out forecast. For the MOS low, the end day of the MOS low time period is used. Thus, if the MOS low period was 7 p.m. March 9, 2023, through 8 a.m. March 10, 2023, the March 10 date would be used for days-out calculations.

Observed High and Low Temperature

The maximum temperature from the 7 a.m. to 7 p.m. local standard time hourly observations was used to construct the high temperature observation. The minimum temperature from the 7 p.m. to 8 a.m. local standard time hourly observations was used to construct the low temperature observation. No attempt to determine high and low temperatures outside these time periods, curve fit, or otherwise determine an intra-hour temperature estimate was performed.

Percentage of Forecasts within Three Fahrenheit degrees

This metric refers to the percentage of forecasts that were within three Fahrenheit degrees, either too high or too low, of the observation. The winner in this metric would be the provider with the highest percentage of forecasts within three Fahrenheit degrees.

Mean Absolute Error

The mean absolute error (MAE) represents the average of the absolute value of the difference between forecast temperatures and observed temperatures. This means all errors are positive and represent the average magnitude of error. This measures how far off the set of forecasts is on average without regard for whether they are too high or too low. The provider with the smallest average absolute error would be the winner in this metric.

Wind Metric

ForecastWatch analyzed 24-hour wind forecasts using one metric: wind speed mean absolute error.

Wind Speed Mean Absolute Error

Average wind speed bias is the positive or negative difference between forecast wind speed and observed wind speed. The mean absolute error takes the absolute value of the error (bias) 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 whether they were too high or too low. The provider with the lowest mean absolute error wins this metric.

Precipitation Metrics

ForecastWatch analyzed 24-hour probability of precipitation using the metric of Brier score and the 24-hour precipitation forecast accuracy using the Equitable Threat Score.

Observed Precipitation

Precipitation measurements were taken from 24-hour local standard time precipitation observations. In the United States and Canada, if 0.01 inches or more of liquid-equivalent precipitation fell during any hour of that day, it was classified as a precipitation day. Internationally, if the observation report of weather noted precipitation, it was considered a precipitation day. Otherwise, it was considered a non-precipitation day.

Brier Score

Probability of precipitation (POP) can be assessed in two ways: reliability and resolution. A POP forecast is reliable when precipitation occurs the same percentage of time it is predicted. For example, if it rained 10% of the time, and the POP forecast called for a 10% chance of rain, the POP forecast would be considered reliable. But reliability is only half of the equation. Consider a scenario in which a forecaster always predicts a 30% chance of rain for a given location and, on average, there is precipitation in that location three out of every ten days. The POP forecast would be reliable, but not necessarily useful because it does not define with absolute certainty whether rain will take place; it is not resolved.

The other measure of a forecast 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 does not. There is no place for POP 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.

Therefore, to evaluate a POP forecast fully, both reliability and resolution must be considered. A Brier score, which considers both reliability and resolution, is one measure used to evaluate POP forecasts. A Brier score ranges from zero to one, with zero being perfectly reliable and resolved. The winner in this metric would be the provider with the lowest Brier score.

Precipitation Forecast Equitable Threat Score

Equitable threat score (ETS), also known as the critical success index, measures the skill of a forecast relative to chance. ETS helps answer the question of how well forecast precipitation events corresponded to the observed precipitation events (accounting for hits due to chance). Zero indicates no skill, and one indicates perfect skill. The provider with the highest ETS would win this metric.

Opacity Metrics

ForecastWatch analyzed daytime (7 a.m. to 7 p.m.) MOS opacity forecasts using two metrics for this report. These metrics were the percentage of text forecasts with the same sky forecast or one category off and the opacity mean absolute error. Observations for both metrics were derived from a total sky coverage code in the Integrated Surface Database (ISD).

Percentage of Text Forecasts with Same Sky or One Category Off

This is defined as the percentage of forecasts when the observed sky cover was the same as the text forecast provided, or one category off. Sky cover observations from the ISD were converted, and text forecasts were normalized, to canonical reported sky covers: clear (sunny), few (mostly sunny), scattered (partly sunny/partly cloudy), broken (mostly cloudy), and overcast (cloudy). For example, a forecast of "scattered clouds" would be categorized as scattered. If the observation was scattered, few, or broken, the forecast would be considered the same or one category off. The provider with the highest percentage of text forecasts with the same sky cover, or one category off, would win this metric.

Opacity Mean Absolute Error

Opacity mean absolute error is defined as the average absolute error of the opacity (sky cover percentage) forecasts, after the ISD observation is converted to a percentage opacity. This means that all errors are positive and represent the average magnitude of error. This measures how far off the set of forecasts is on average without regard for whether they are too high or too low. The lowest average absolute error would win this metric.

Weather Forecast Providers

- **AccuWeather**: Forecasts were collected using the AccuWeather API at http://api.accuweather.com using a specific location code.
- **AerisWeather**: Forecasts were collected using the AerisWeather API at https://www.aerisweather.com/develop/api/.
- **Australian BOM** (Australia only): Forecasts were collected from http://reg.bom.gov.au/.
- **BBC** (UK Only): Forecasts were collected from the BBC website at https://www.bbc.co.uk/weather/.
- **CFAN** (US only): Forecasts were collected from the feed provided by CFAN.

- **Environment Canada** (Canada only): Forecasts were collected from the Environment Canada weather forecast website at https://weather.gc.ca/.
- **Foreca/Vaisala**: Forecasts were collected from an API that populates the ten-day forecast page at http://www.foreca.com. The location parameter used was the city and state of the observation location for the website, and a location code (either ICAO or WMO) for the API.
- **Global News**: Forecasts were collected from the Global News website at https://globalnews.ca/.
- **MeteoFrance** (France only): Forecasts were collected using the unofficial MeteoFrance Python client library at https://github.com/hacf-fr/meteofrance-api.
- **Microsoft**: Forecasts were collected from an API provided by Microsoft.
- **Moji**: Forecasts were collected from an API provided by Moji.
- **National Weather Service** (US only): Forecasts were collected using the NWS Digital Forecast API at https://graphical.weather.gov/xml/.
- **OpenWeather**: Forecasts were collected using the Open Weather API at https://api.openweathermap.org/.
- **Pirate Weather**: Forecasts were collected via the API available at https://pirateweather.net/en/latest/API/.
- **UK MetOffice** (UK only): Forecasts were collected using the UK MetOffice Datapoint API at http://datapoint.metoffice.gov.uk/.
- **The Weather Channel**: Forecasts were provided by a private commercial API provided by The Weather Company. Latitude and longitude of the observation station were used to retrieve specific forecasts.
- **Weather News**: Forecasts were collected from the Weather News website http://weathernews.jp/onebox/.
- **Weatherbit**: Forecasts were collected using the Weatherbit API from https://www.weatherbit.io/api/weather-forecast-api.
- **Weerplaza**: Forecasts were collected from the Weerplaza website at https://www.weerplaza.nl/.
- **Wetter**: Forecasts were collected from the Wetter website https://www.wetter.com.

- **World Weather Online**: Forecasts were collected using the World Weather Online API from https://www.worldweatheronline.com/weather-api/.
- **yr.no**: Forecasts were collected from the yr.no website.

Collection Methodology

Observation Data

Observation data were collected from the primary Automated Surface Observing System (ASOS) network in the United States as well as international equivalents. United States and International data were collected from the Integrated Surface Database (ISD) product. Canadian data were collected from Environment Canada. All products consisted of hourly and daily observation parameters.

Observed High and Low Temperature

The maximum and minimum temperature observations are from the 24-hour local standard time temperature observations and were used to construct the high and low temperature observation. These observations are generally either FM-12 (SYNOP), FM-15 (METAR), or FM-16 (SPECI) records. United States 24-hour high and low temperature observations were collected from the Summary of the Day (SOD) records, which use five-minute sampling. All 24-hour high and low international observations were derived from hourly and special report observations. No attempt to curve fit or otherwise determine an intra-hour temperature estimate was performed.

Observed Wind

Wind conditions were taken from hourly observations over the course of a 24-hour period from local midnight to midnight standard time. These observations were then averaged to construct the daily wind observation.

Observed Precipitation

Precipitation measurements were taken from 24-hour local standard time precipitation observations. For United States and Canada locations, 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. Internationally,

precipitation reports were used to determine precipitation observations. The occurrence or non-occurrence of precipitation was then compared to the Probability of Precipitation (POP) forecast.

Observed Opacity

Sky cover conditions were taken from hourly observations over the course of a 24-hour period from local midnight to midnight standard time. These observations were then averaged to construct the daily opacity observation.

Forecast Collection

Data were collected from each of eight global regions at specific times during the day. Table 8 shows collection times for each region. For example, daily temperature forecasts were collected beginning at 22:00 UTC in the United States and continued until all forecasts were collected. For each location, forecasts from all providers were collected at the exact same time.

Validity

Forecasts were considered valid if they were complete (i.e., they contained a high and low temperature forecast, a POP/precipitation 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, forecast 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.

Region	Collection Time	Number of Stations
United States	22:00 UTC	900
Canada	21:40 UTC	125
Europe	16:00 UTC	644
Asia Pacific	08:00 UTC	241
Africa	15:30 UTC	38
Middle East	13:00 UTC	105
Central America	23:00 UTC	17
South America	21:00 UTC	78

Table 8: Regions and Collection Times

About ForecastWatch.com

ForecastWatch, a service of Intellovations, LLC, has been the world's premier weather forecast monitoring and analytics company since 2003. Our passion for data drives us every day. We collect weather forecast data from several thousand locations throughout the US and around the world. This information is added to an ever-growing and unparalleled historical database of more than two billion weather forecasts gathered from dozens of weather forecast providers and systems.

We use this vast collection of data to evaluate and compare weather forecast providers, improve decision-making by governments and business entities affected by weather, improve weather forecasting by meteorologists around the world, and educate customers with unbiased reporting. We strive to improve and expand our offerings to meet the needs of our current and future clients, finding ways to partner with them to help them evaluate their own deliverables, keep their customers safe, or help make business-critical decisions by analyzing weather forecasts to positively affect revenue, operating costs and risk mitigation costs.

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. Even consumers benefit from our ForecastAdvisor product. Our data meets the highest standard of scientific inquiry and has been used in several peer-reviewed studies.

Disclosure and Grant of License

The “Analysis of One- to Five-Day-Out Global Temperature, Wind Speed, Precipitation, and Opacity Forecasts” was commissioned and sponsored by Microsoft.

This report is made available to the public subject to certain restrictions. You may use this report for your personal use, for educational or scholarly use, or for research purposes only.

This report may not be copied, distributed, transmitted, duplicated, reduced, or altered in any way for commercial purposes, or for the purpose of redistribution, without a license from Intellovations, LLC. Requests for information regarding a license for commercial use or redistribution of the “Analysis of One- to Five-Day-Out Global Temperature, Wind Speed, Precipitation, and Opacity Forecasts” may be sent via email to reports@intellovations.com.