

Forecast Accuracy For December 2011 For San Francisco, California, USA

This report describes the accuracy of the forecasts issued by several forecasting agencies at the San Francisco International Airport (San Francisco, California, United States) weather station. The error metrics are computed by comparing the forecasts issued during December 2011 to what actually happened, as recorded by the weather station.

Forecasting Agencies in This Report

Agency	Full Name	Number of Forecasts	Forecast Length
NOAA	National Weather Service	30	6.8 days
Met.no	Norwegian Meteorological Institute	43	9.4 days
WVO	World Weather Online	27	4.8 days
WXC	Weather Central	31	15.4 days

In the table above, the number of forecasts listed in the third column is the number of forecasts that we have in our archive. Each agency may have issued more forecasts during the month, depending on their forecasting schedule.

Three-Day Forecast Accuracy

The table below reports the three-day hourly root mean square error (RMSE) for each of the forecasting agencies. For reference, we include the baseline error rate resulting from forecasting each variable to be the average value of that variable from the same time in past years. For each variable, the agency with the smallest error (winner) is indicated in bold.

Three-Day Hourly Forecast Error

	NOAA	Met.no	WVO	WXC	Baseline	Winner
Temperature	3.7°F	3.1°F	4.6°F	3.5°F	4.3°F	Met.no
Cloud Cover	33%	41%	43%	37%	46%	NOAA
Wind Speed	4 mph	5 mph	4 mph	5 mph	6 mph	WVO
Wind Direction	82°	88°	83°	88°	82°	NOAA
Dew Point	5.8°F	9.6°F	6.0°F	7.5°F	8.1°F	NOAA
Humidity	14%	20%	15%	18%	19%	NOAA

Precipitation is a bit different from the other weather variables in that it consists of discrete events separated by periods of inactivity, and is not well represented by a smoothly varying number, as for example temperature. The hourly RMSE is consequently not a very good measure of forecast accuracy. Instead, we break the forecasts up into 24-hour chunks (days), and categorize each forecast as either predicting precipitation or not on each day. We then compare those yes-or-no predictions to what actually happened and count the cases where the forecasts were correct or incorrect in predicting precipitation or not.

The table below reports several precipitation forecast error metrics over the first three days of the forecasts.

Three-Day Precipitation Forecast Error

	NOAA	Met.no	WVO	WXC	Winner
Overall Error Rate	13%	14%	15%	18%	NOAA
Predicted but Not Observed*	31%	34%	52%	53%	NOAA
Observed but Not Predicted**	11%	11%	9%	9%	WVO

* as a fraction of all days where precipitation was predicted

** as a fraction of all days where precipitation was not predicted

Depending on your preferences, each of these error rates may be important.

Additional Three-Day Error Metrics

In addition to the RMSE reported above, we also present two additional error metrics on the three-day forecast. The first of these (shown below) is the maximum absolute error (MAE). This provides a measure of how far off each agency was at their worst.

Three-Day Hourly Mean Absolute Forecast Error

	NOAA	Met.no	WVO	WXC	Winner
Temperature	17.0°F	9.5°F	15.7°F	11.7°F	Met.no
Cloud Cover	92%	100%	100%	100%	NOAA
Wind Speed	21 mph	17 mph	14 mph	17 mph	WVO
Wind Direction	180°	180°	177°	180°	WVO
Dew Point	23.9°F	32.8°F	22.5°F	26.5°F	WVO
Humidity	62%	55%	56%	62%	Met.no

The second (shown below) is the mean error, a measure of how biased the forecasts were. A negative bias means the forecasting agency predicted a lower value on average than what was observed; a positive bias means they predicted higher values.

Three-Day Hourly Mean Forecast Error (Bias)

	NOAA	Met.no	WVO	WXC	Winner
Temperature	1.2°F	-0.2°F	0.6°F	0.0°F	WXC
Cloud Cover	-12%	-11%	-16%	-12%	Met.no
Wind Speed	0 mph	1 mph	1 mph	-1 mph	NOAA
Wind Direction	1°	-2°	14°	-0°	WXC
Dew Point	2.7°F	-5.1°F	0.1°F	-4.4°F	WVO
Humidity	3%	-9%	0%	-6%	WVO

While the RMSE is probably the most generally useful error metric for understanding the accuracy of the various forecasting agencies, these additional error metrics provide additional depth.

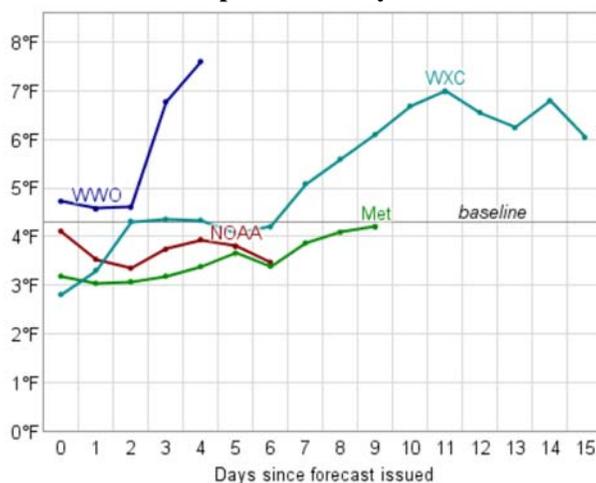
Daily Error Graphs

The graphs below show how the forecasting error evolves over time. They are useful for understanding how far out a forecast should be trusted, and provide a generally more nuanced picture than the three-day error rates reported above.

In the graphs below, we provide the baseline as a reference. The baseline is the error rate resulting from forecasting each variable to be the average from the same time in past years. If the forecasts consistently perform worse than the baseline a certain number of days out, we recommend looking at the averages more than the forecast to get a sense of what that aspect of the weather will be like.

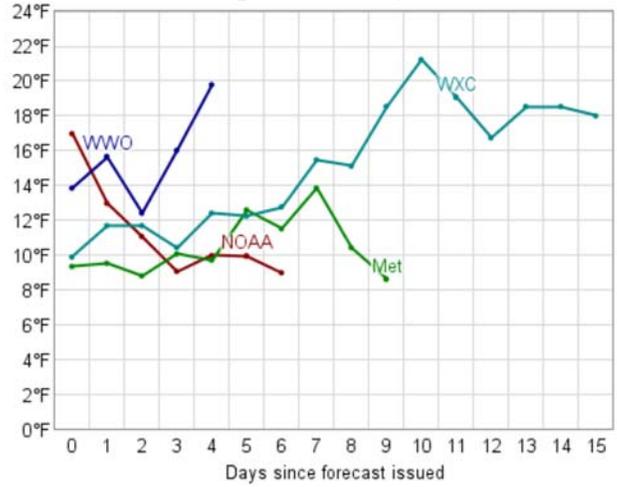
Temperature

Temperature: Daily RMSE



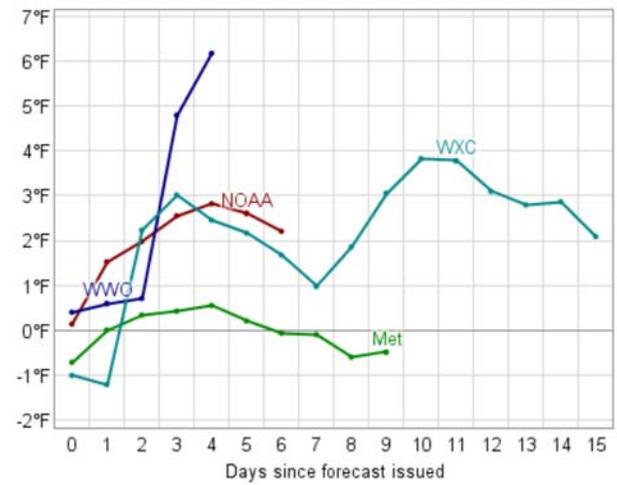
The daily root mean square error (RMSE) of the forecasted temperature. The baseline results from forecasting the temperature to be the average temperature from the same time in past years.

Temperature: Daily MAE



The daily maximum absolute error (MAE) of the forecasted temperature.

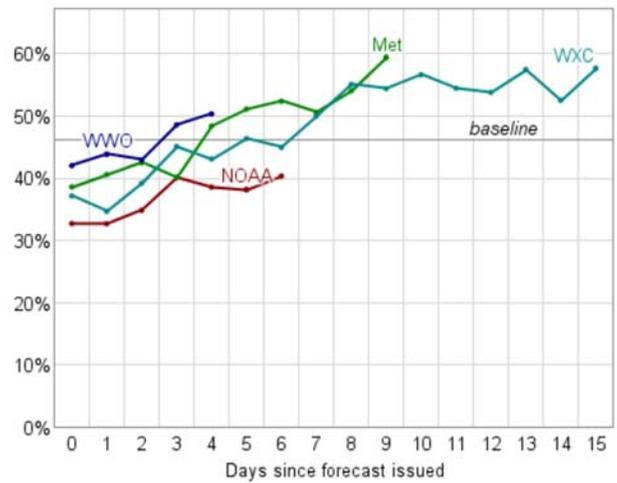
Temperature: Daily Forecast Bias



The mean error of the forecasted temperature for the various forecast sources. Positive bias means that on average the forecast predicted warmer temperatures than actually ended up happening; negative bias indicates colder temperatures were forecasted than happened on average.

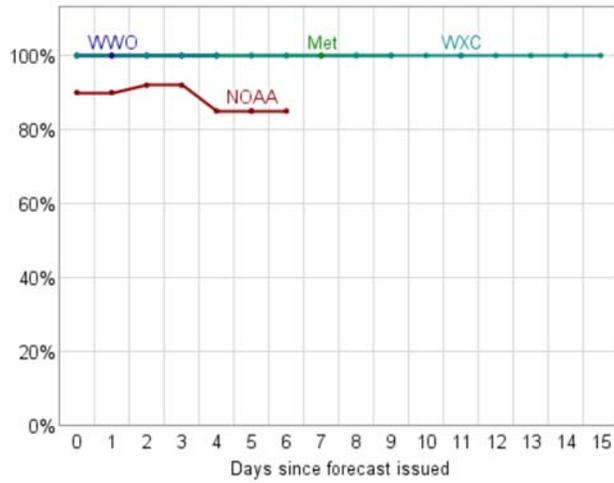
Cloud Cover

Cloud Cover: Daily RMSE



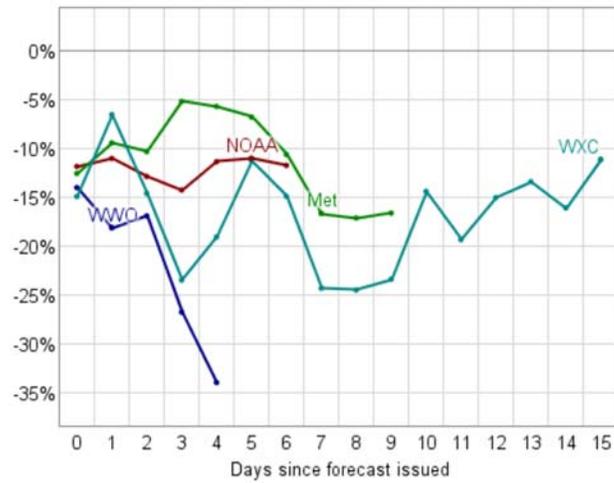
The daily root mean square error (RMSE) of the forecasted cloud cover. The baseline results from forecasting the cloud cover to be the average cloud cover from the same time in past years.

Cloud Cover: Daily MAE



The daily maximum absolute error (MAE) of the forecasted cloud cover.

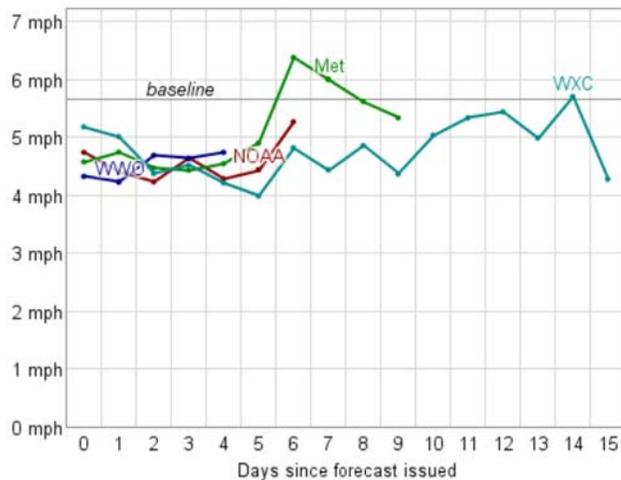
Cloud Cover: Daily Forecast Bias



The mean error of the forecasted cloud cover for the various forecast sources. Positive bias means that on average the forecast predicted cloudier skies than actually ended up happening; negative bias indicates clearer skies were forecasted than happened on average.

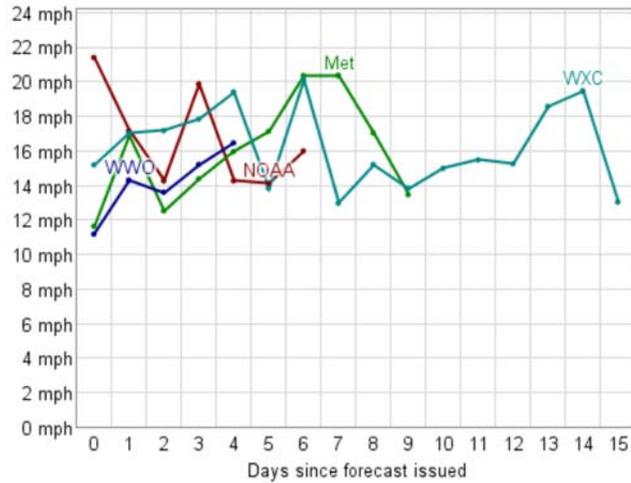
Wind Speed

Wind Speed: Daily RMSE



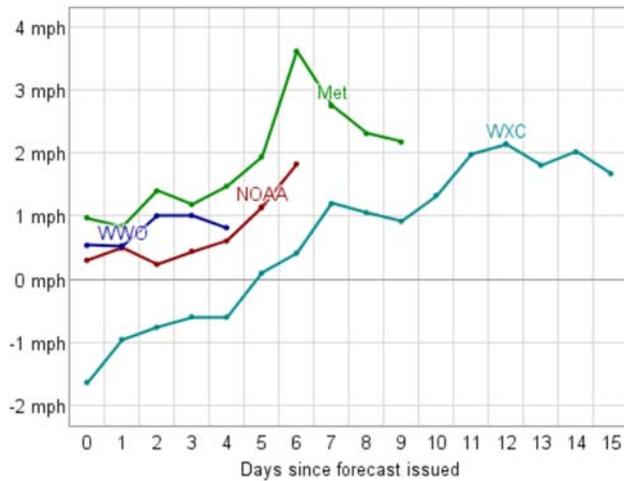
The daily root mean square error (RMSE) of the forecasted wind speed. The baseline results from forecasting the wind speed to be the average wind speed from the same time in past years.

Wind Speed: Daily MAE



The daily maximum absolute error (MAE) of the forecasted wind speed.

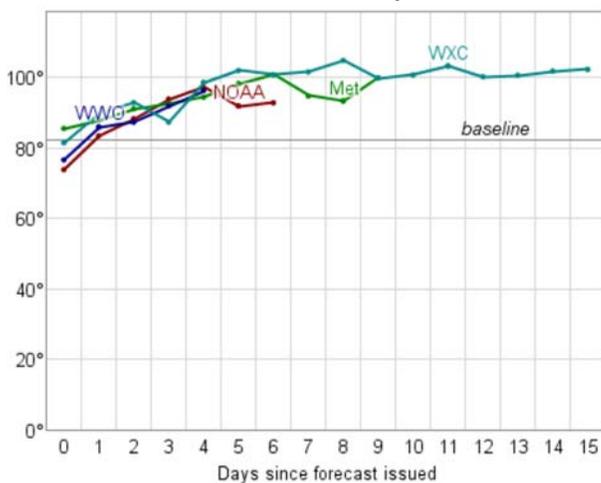
Wind Speed: Daily Forecast Bias



The mean error of the forecasted wind speed for the various forecast sources. Positive bias means that on average the forecast predicted higher winds than actually ended up happening; negative bias indicates calmer winds were forecasted than happened on average.

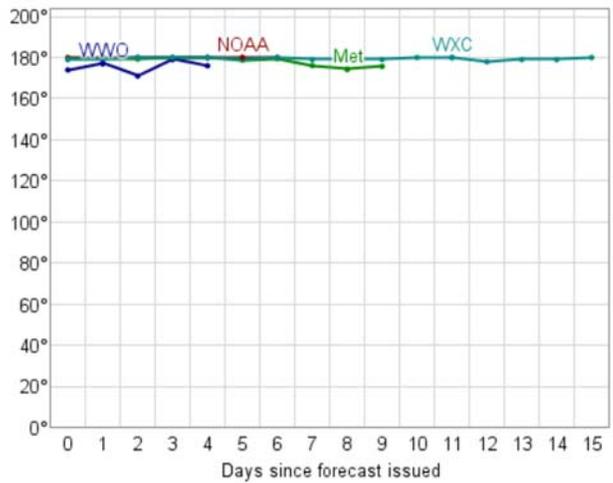
Wind Direction

Wind Direction: Daily RMSE



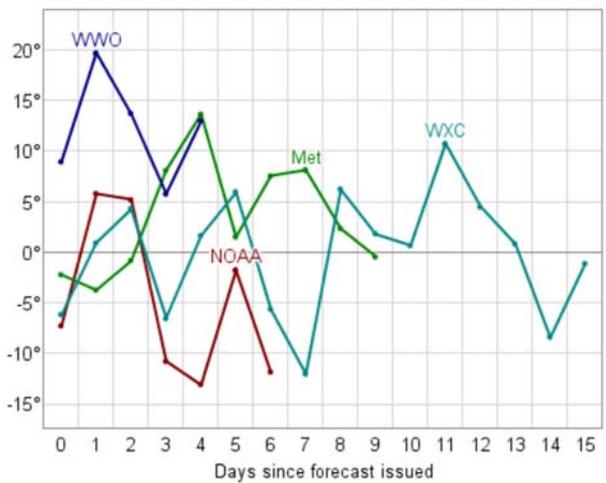
The daily root mean square error (RMSE) of the forecasted wind direction. The baseline results from forecasting the wind direction to be the average wind direction from the same time in past years.

Wind Direction: Daily MAE



The daily maximum absolute error (MAE) of the forecasted wind direction.

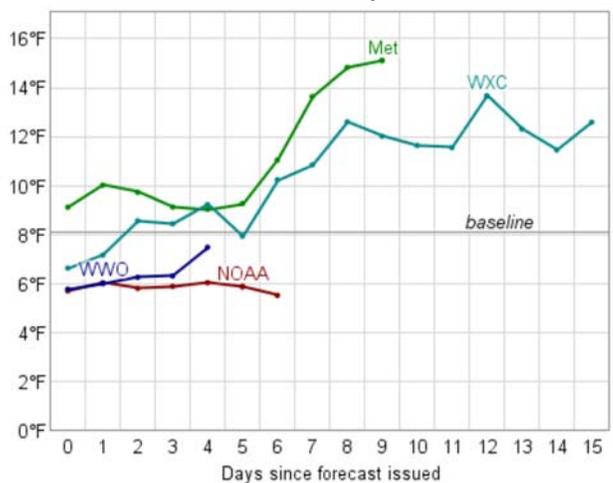
Wind Direction: Daily Forecast Bias



The mean error of the forecasted wind direction for the various forecast sources. Positive bias means that on average the forecast predicted winds to the right of what actually ended up happening; negative bias indicates winds to the left were forecasted than happened on average.

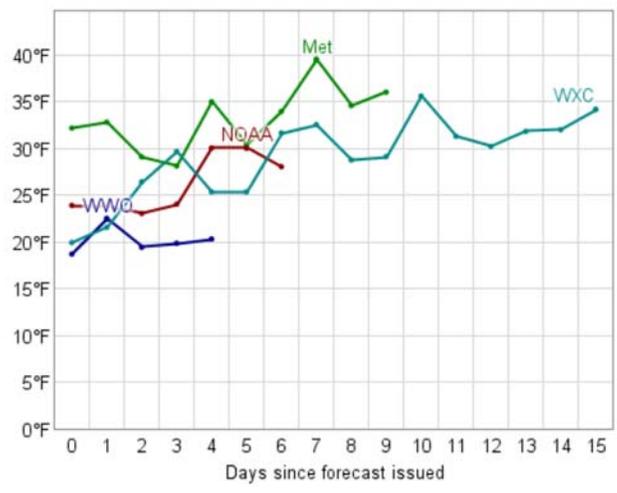
Dew Point

Dew Point: Daily RMSE



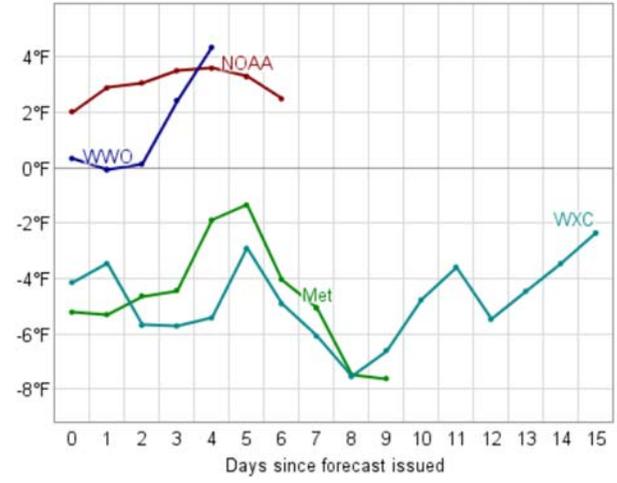
The daily root mean square error (RMSE) of the forecasted dew point. The baseline results from forecasting the dew point to be the average dew point from the same time in past years.

Dew Point: Daily MAE



The daily maximum absolute error (MAE) of the forecasted dew point.

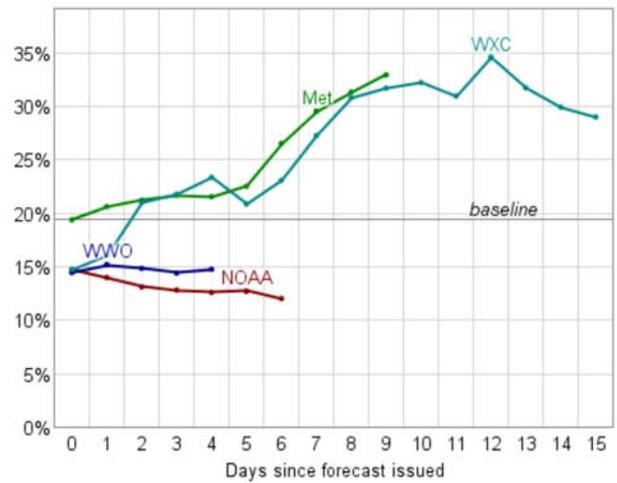
Dew Point: Daily Forecast Bias



The mean error of the forecasted dew point for the various forecast sources. Positive bias means that on average the forecast predicted higher dew points than actually ended up happening; negative bias indicates lower dew points were forecasted than happened on average.

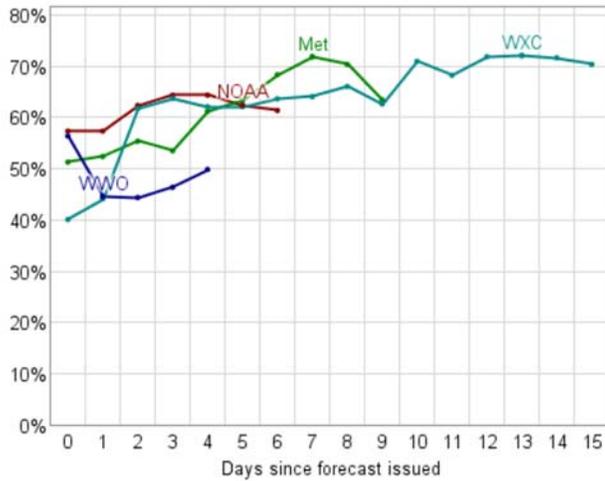
Humidity

Humidity: Daily RMSE



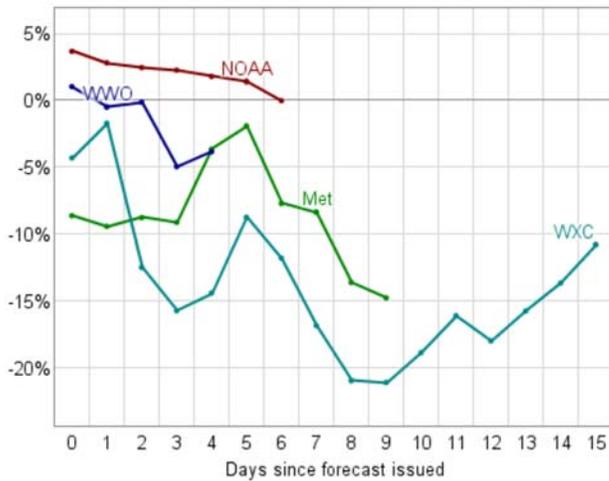
The daily root mean square error (RMSE) of the forecasted humidity. The baseline results from forecasting the humidity to be the average humidity from the same time in past years.

Humidity: Daily MAE



The daily maximum absolute error (MAE) of the forecasted humidity.

Humidity: Daily Forecast Bias

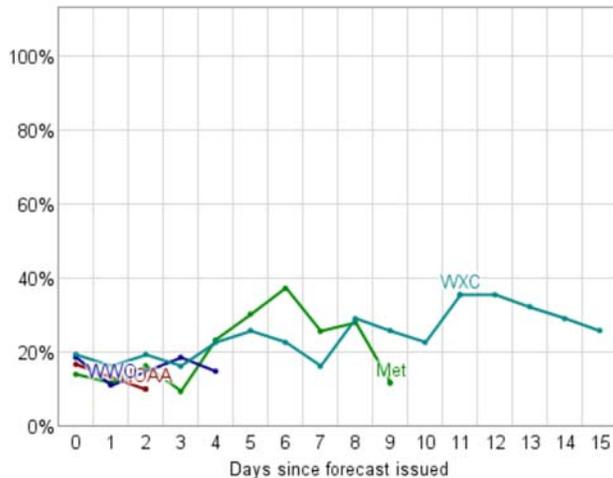


The mean error of the forecasted humidity for the various forecast sources. Positive bias means that on average the forecast predicted higher relative humidity than actually ended up happening; negative bias indicates lower humidity were forecasted than happened on average.

Precipitation

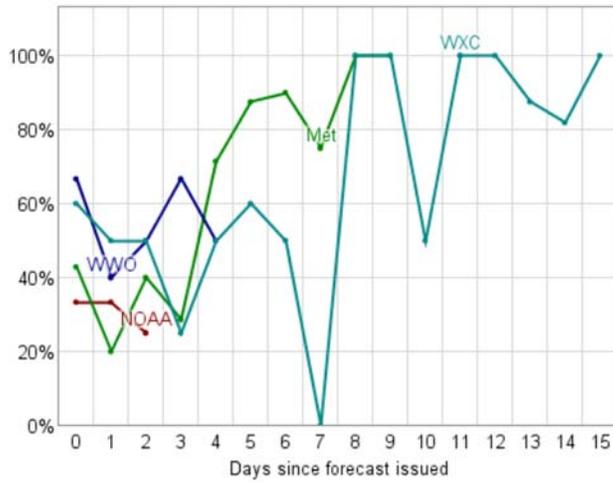
For the same reasons listed in the Three-Day Forecast Accuracy section above, the daily precipitation forecast error graphs differ from those of the other variables above.

Overall Precipitation Error Rate



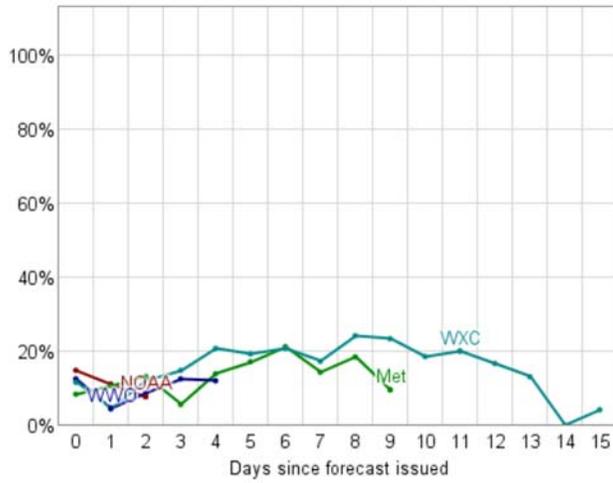
The percentage of cases in which either the forecast predicted precipitation that did not occur or that the forecast failed to predict precipitation that did occur.

Precipitation Predicted but Not Observed



The percentage of cases in which precipitation did not occur despite having been predicted (as a fraction of all cases where precipitation was predicted).

Precipitation Observed but Not Predicted



The percentage of cases in the forecast failed to predict precipitation that did occur (as a fraction of all cases where precipitation was not predicted).