Weather Forecasts:
Just how reliable are they?

Do you regularly visit your favorite weather Web site, or check the forecast every morning on your cell phone? Some of us really obsess about the weather. In the United States, we have an entire television channel dedicated to it, and I have friends who pay more attention to this channel than the news, their favorite sports team…perhaps even their spouse!

When you check the weather, you can just get the next-day forecast, or you can look ahead at a 5-day or even a 10-day forecast. But have you ever wondered how truly trustworthy these predictions are? Many people count on these long-range forecasts to accurately predict future weather, but are they really reliable, or are they more like the meteorological equivalent of gazing into a magic crystal ball?

In this article, we will look at temperature data and use various statistical tools to see just how reliable the forecast is.

Data Collection

We visited a leading weather Web site and collected the predicted next-day, 5-day and 10-day high temperatures (in ºF) for State College, Pennsylvania—the location of Minitab's World Headquarters. We did this every day for 30 days, and recorded the forecasted temperatures, in addition to the actual high temperature for each day, in Minitab Statistical Software (Fig. 1).

Then, using **Calc > Calculator**, we computed the differences between the 3 forecasted high temperatures and the actual temperature reading for each day.

	C1-D	C2	C3	C4	C5	C6
	Date	10-day high	5-day high	next day high	actual high	10-day diff
1	May 12	69	77	77	77	-8
2	May 13	65	70	67	66	-1
3	May 14	64	64	63	63	1
4	May 15	65	66	66	68	-3
5	May 16	68	67	61	64	4
6	May 17	69	59	62	61	8
7	May 18	67	66	66	64	3
8	May 19	67	66	66	66	1

Figure 1. We recorded the 10-day, 5-day and next-day forecasts in Minitab, then calculated the difference between the forecasted and actual high temperatures.

Forecasts: Predictable or Patchy?

With any analysis, graphing your data is always a good place to start. Since these data were collected over time, we can use **Graph > Time Series Plot** to graph the actual and forecasted temperatures for each day.

The graph (Fig. 2) shows that the forecast was more reliable on some days compared

to others. It also shows the rises and falls in temperature that occurred in State College during the late spring.

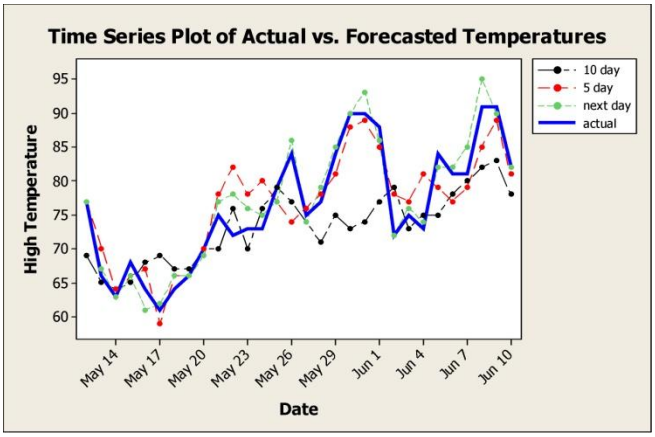

Figure 2. We can use a Time Series Plot to compare the forecast to the actual temperature for each day of the study.

To easily compare the center and spread for the difference between the actual temperature and the three forecasts, we can use **Graph > Individual Value Plot**.

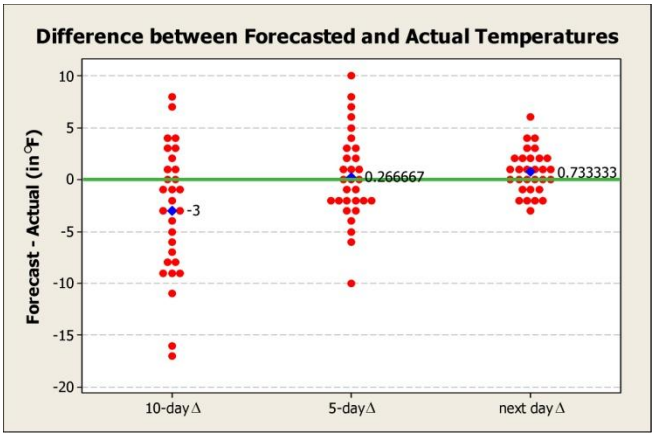

Figure 3. The individual value plot shows that the 10-day forecast exhibits more variation than the other two forecasts.

With a standard deviation of 6.2 degrees, we can see that the 10-day forecast overestimated the high temperature by as much as 8 degrees and underestimated it up to 17 degrees (Fig. 3). The 5-day and next day forecasts were less variable with standard deviations of 4.3 and 2.1, respectively.

Now, is this difference in variability between the 3 forecasts statistically significant, or are these observed samples likely given that the variances are truly equal? Let's use an equal variances test to find out.

There are multiple ways to conduct a variances test in Minitab. In this case, we will use **Assistant > Hypothesis Tests** to obtain a comprehensive report of our results. With a p-value of 0.000 (not shown), we can conclude that there is a statistically significant difference between the variances. Specifically, the variance for the next-day forecast difference is significantly smaller than the variances for the 5-day and 10-day forecast differences.

Conclusion: The next day forecast is significantly more precise than the other two forecasts.

Next, let's use **Assistant > Hypothesis Tests** to conduct a Welch's ANOVA for unequal variances and evaluate the forecasts' accuracy. Using an α=0.05, the p-value of 0.011 is significant. Therefore, the three averages are not all equal (Fig. 4). Specifically, the 0.7º average overestimate for the next day forecast is significantly greater than the -3º average underestimate for the 10-day forecast.

Conclusion: The next- and 5-day average differences appear to be equally accurate, with both confidence intervals including 0.

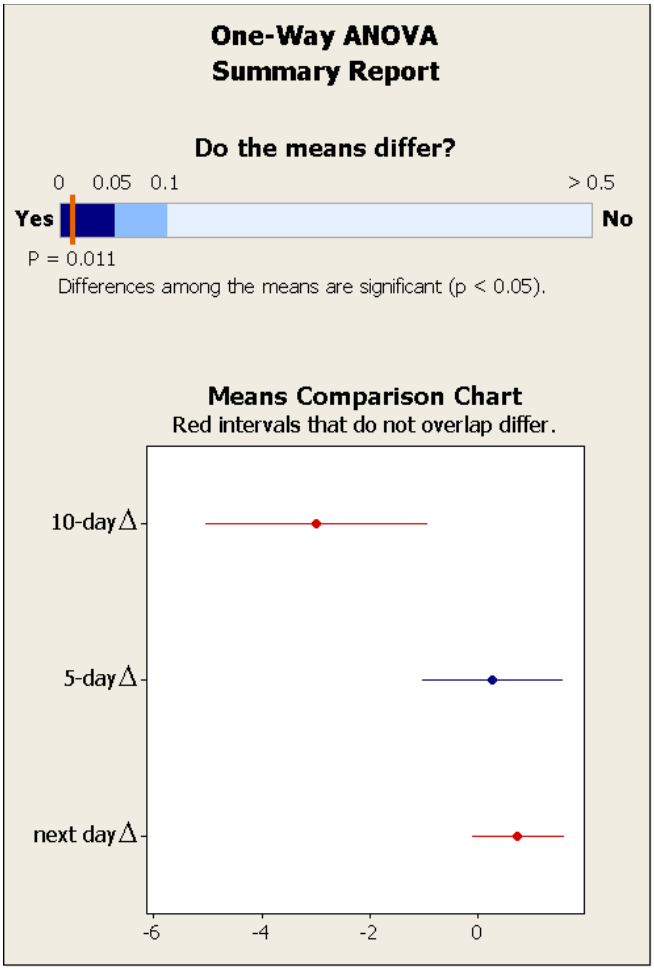

Figure 4. The average disparity for the next-day forecast is significantly less than the 10-day forecast.

Swimsuit or Sweater?

Now that we've done our comparisons and determined that the 10-day forecast is significantly less accurate, let's evaluate just how well the 5-day and next-day forecasts can be used to predict whether you should plan to wear a swimsuit or a sweater.

5-Day Forecast

Although the 5-day forecast is more variable than the next-day forecast, we still want a crystal ball: we want to look into the future, and perhaps by more than 24 hours. To visually and statistically assess how well we can predict the actual temperature using the 5-day forecast, we can conduct a regression analysis using **Stat > Regression > Fitted Line Plot**.

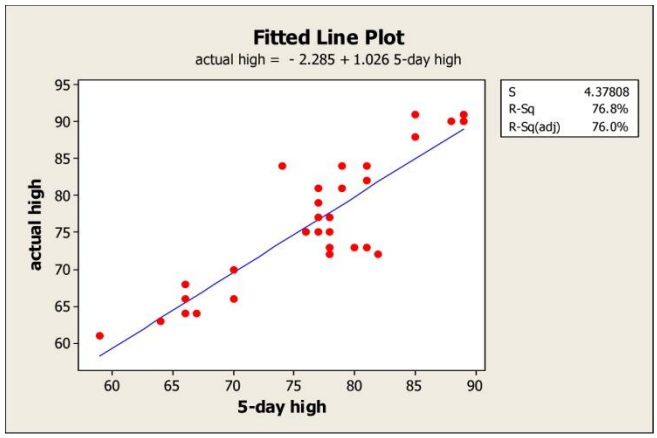

Figure 5. We can use a fitted line plot to explore the relationship between the actual temperature and the 5-day forecast.

The p-value of 0.000, found in the Session window, indicates that a significant linear relationship exists between the actual temperature and the 5-day forecast.

The R-squared value tells us that this model accounts for 77% of the variability seen in the actual high temperature, which is likely better than your average crystal ball.

Also, using residual plots (not shown) we can verify the model assumptions and conclude that the analysis is valid.

Next-Day Forecast

Now, let's run a similar analysis for the next-day forecast. The graph (Fig. 6) and high R-squared value indicate that the next day forecast is a better and very reliable predictor of the actual high temperature.

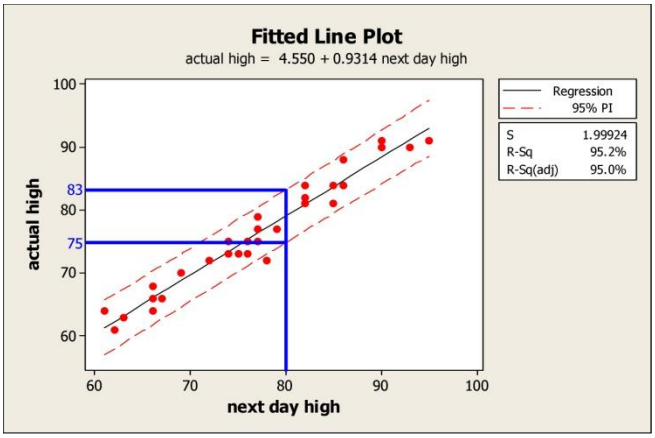

Figure 6. Because the points fall close to the line, the regression model appears to be a good fit.

Using prediction intervals, we can calculate a likely range of values for a given next day forecast. For example, we can be 95% certain that a next-day forecast of 80ºF will likely correspond to an actual temperature between 75 and 83ºF.

It's important to note that although regression tells us if a linear relationship exists, it does not tell us if this is a 1-to-1 relationship. In other words, our p-value could be significant because a forecast of 80ºF indicates an actual high of 80ºF, or it could be significant because a forecast of 80ºF does a good job at predicting an actual high of, say, 40ºF (e.g., if the slope coefficient is 0.5).

To see if there is a 1-to-1 relationship, we can calculate the confidence intervals for the coefficients using **Stat > Regression > General Regression > Results**.

Coefficients

Term	Coef	P	95% CI
Constant	4.54983	0.150	(-1.75304, 10.8527)
next day high	0.93142	0.000	(0.85017, 1.0127)

Figure 7. We can use a confidence interval to assess the coefficients for the y-intercept and slope.

Because the confidence interval (Fig. 7) for the constant (-1.75, 10.85) includes 0 and the confidence interval for the slope coefficient (0.85, 1.01) includes 1, we can conclude that the relationship between the next-day forecast and the actual temperature is in fact a 1-to-1 relationship.

Conclusion

Given all of the factors that influence it, the weather is an undeniably complex process—and like any process, it can exhibit a lot of variation. However, if you're going to make any big plans based on weather and you want to minimize the variation, the data we collected suggest it's best to rely on the next-day forecast.

There is not much we can accurately predict 5 days into the future, so relatively speaking, the 5-day forecast comes a lot closer to doing that than most aspects of life. As for the 10-day forecast, it's likely that meteorologists know exactly how unpredictable the weather conditions 10 days in the future can be. And they provide it to us weather-watchers nonetheless because we still want some sense of what the future holds, despite the unreliability of the predictions. But it's good to know which forecasts we can really count on, and which come closer to fortune-telling!

Michelle Paret
Product Marketing Manager, Minitab Inc.

Eston Martz
Senior Creative Services Specialist, Minitab Inc.

*Visit **www.minitab.com** for more articles.*