



By Tom Klemens  
Senior Editor/Engineering  
tklemens@hanleywood.com

## Surrounded by Data

A large part of Colin Milberg's project that began in 2007 is simply collecting design and as-built data on concrete construction projects—and then figuring out exactly what and how to measure. After more than a year, he and his team have accumulated data from their firsthand use of laser scanners on nearby jobsites and from contractors submitting data from additional projects in other locations.

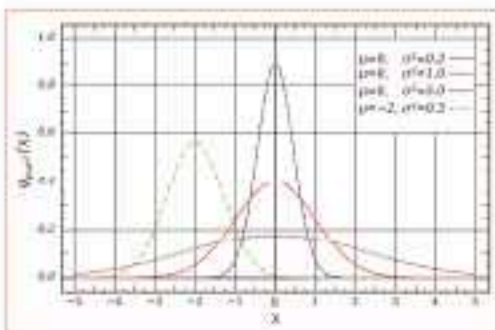
Their overall goal is to compare as-built data with design data to establish baselines for how accurately the industry is constructing various concrete elements. From that, Milberg expects to establish process capabilities for the different elements that also take into account key variables, such as the type of formwork and placement methods used.

Laser scanner data is of a very high density compared with more traditional methods, such as using a total station, and that adds another variable to the data analysis. Both approaches collect 3-D coordinates for as-built points and compare them with the design coordinates for those points. Scanning provides data for each point on a rough grid. As a simplistic example, a 3-inch grid on a 4-foot-wide 10-foot-high wall results in nearly 700 points. Using a 1-inch grid gives you 5900 or more. However, participants using total stations to collect the data on such a wall are provided with the vertical and horizontal coordinates for 30 randomly generated points, for which they measure the third dimension (distance from the instrument).

The data collection procedure for using standard equipment is well explained on the project's "Construction Process Capability (CPC) Database Web site for

Concrete Construction" ([www.itmcc.sdu.edu/index.php](http://www.itmcc.sdu.edu/index.php)). Besides an overview of the project and its goals, the site now offers a Powerpoint download that walks through the entire procedure, from signing up and logging in, to data-entry options.

The question is how many points one must measure to provide enough information to evaluate the finished product. The answer, in this case, comes from a standard application of statistical analysis.



A bell curve represents data with a normal distribution, which means you can be 95% confident that your measurements reflect reality.

Among other things, Milberg's team is determining whether the data fit a normal distribution curve. In statistical terms, the word *normal* is just a name that refers to a bell-shaped curve where the data points are distributed evenly above and below the mean value, with most data points near the mean and fewer further away from it.

The bell's width for a normal distribution is described by the standard deviation, simply a value reflecting the data spread. A small standard deviation is like a small slump value for concrete—the shape is tall and relatively thin. For a larger standard deviation, the bell is relatively shorter and wider.

Making sure the data is normally distributed defines the confidence level. Without identifying that natural distribution of concrete surface variation, you won't know that your measurements are giving a good picture of reality.

"The standard deviation is an example of an estimator that is the best we can do if the underlying distribution is normal. However ... confidence intervals based on the standard deviation tend to lack precision if the underlying distribution is in fact not normal," according to section 1.3.5.6 of the NIST/SEMATECH e-Handbook of Statistical Methods, (<http://www.itl.nist.gov/div898/handbook/>, December 2008).

You might spend some time with the Web book based on *Handbook 91, Experimental Statistics*, written by Mary Natrella of the National Bureau of Standards' Statistical Engineering Laboratory and published in 1963. In presenting its updated online form free of charge, the National Institutes of Science and Technology (NIST) says its goal "is to help scientists and engineers incorporate statistical methods into their work as efficiently as possible. Ideally it will serve as a reference that will help scientists and engineers design their own experiments and carry out the appropriate analyses when a statistician is not available to help."

Then take another look at the reference materials on the CPC Database Web site. Between a better understanding of statistics concepts and Milberg's team approach, you'll be able to make a lot more sense of the results they'll be posting next year. **CC**