Achieving Gage-Capable Process Control for High-Speed, Precision Manufacturing

Solid, gage-proven metrology is an essential tool for process control across all industrial settings. A good measurement system enables manufacturers – ranging from factories that operate 24/7 and incorporate high-volume, high-end processes to medical instrument production performing critical FDA inspections – to keep parts within tight specification limits. However, if the measurement system is unsuitable and the amount of measurement error exceeds the allowable tolerance, it may incorrectly overlook and accept bad parts and/or reject good parts. Consequently, a company's ability to comply with evolving standards, participate in product and process innovation, and remain competitive in modern markets are all contingent upon highly accurate measurement system performance. It is no surprise, then, that gage studies for metrology measurement system performance testing and validation are now commonplace. Such gage studies are a method for determining and demonstrating a measurement system's capability to meet given manufacturing specification tolerances, as well as its suitability for its intended purpose. This application note discusses gage studies and Bruker's latest metrology advancements that help manufacturers and engineers meet the increasingly stringent demands for high-end applications.

Measurement System Error

Ultimately a measurement system is designed and purchased with the needed accuracy necessary to measure a part to a known calibration standard that is ultimately traceable to a national governing body, such as NIST (National Institute of Standards and Technology) in the United States, NPL (National Physical Laboratory) in UK, JISC (Japan Industrial Standards Committee, or PTB (Physikalisch-Technische Bundesanstalt) in Germany. A suitable measurement system requires the repeatability and accuracy (precision) to achieve the necessary standard for a particular measurement. Given the rapid evolution of modern industrial standards, it is not uncommon – either during measurement system development or deployment – to identify system errors that can compromise product quality and other manufacturing process outcomes. Most measurement system errors fall into one of the following categories: stability, accuracy, linearity, repeatability, reproducibility, and bias. Figure 1 provides a visual representation of various accuracy and repeatability errors, including both singular and concurrent errors, with the center bull's-eye representing the part manufacturing target ideal specification.
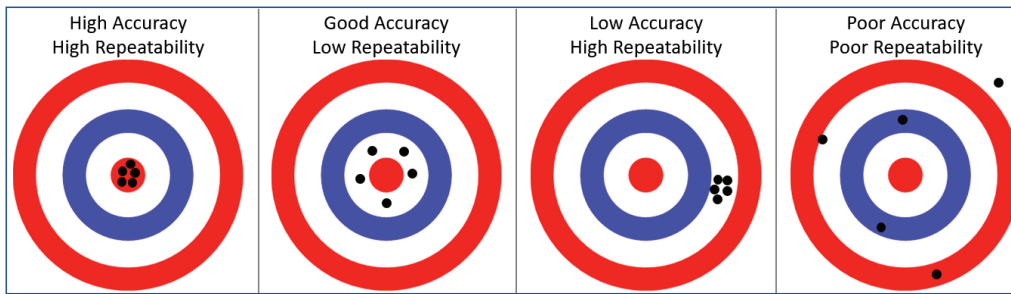
Gage Studies

To identify and understand potential measurement system errors, manufacturers perform gage studies. The scope of a true gage study – commonly referred to as a Gage Repeatability and Reproducibility (GR&R) study – not only considers the capabilities and performance of the measurement tool itself, but also those of the manufacturing instrument(s), operators, environment, and parts. In a GR&R study, the amount of measurement error expressed by the repeatability and reproducibility of a system enables one to determine the performance impact of these different system variables.

The quantitative "repeatability" of a measurement system describes the variation in measurement observed when the same operator uses the same instrument to measure the same parts multiple times under unchanged operating conditions. Conversely, the "reproducibility" of a system describes the amount of variation observed when different operators use the same instrument to measure the same parts multiple times. Reproducibility can also reflect the longer term stability of a measurement system using data collected via a longitudinal study, which also brings tool stability and environment into consideration. Put simply, the first "R" in GR&R represents instrument/fixturing/algorithm variation, while the second "R" represents the measurement system stability across various operators. Both gage study aspects are reflected in a part's specifications; they are applied as a percentage and contribute to its overall part stack up tolerance.

One key strength of a GR&R study is that it ensures the quality of a measurement system's results or initially qualifies it via a systematic approach to gaging – and then improving – the system. That is, if the GR&R results are not acceptable on the first attempt, reasonable and well-documented adjustments can be made and the GR&R repeated to confirm that the results have improved.

GR&R Project Setup and Analysis

Due to the potential large number of repetitions, determining the size of the project is a crucial first step for any GR&R study. This includes quantifying the number of parts (including the number of features to be measured on each part used for process control), the number of operators, and the number of measurements (replicates/repeats) for each operator and each part, as well as the associated resources and costs. Doing so enables manufacturers to weigh the time and financial resources allocated to the study against the practical usability of its results. For example, executing many measurement trials and capturing large sets of gage data may yield more accurate results, yet doing so also requires more time and resources – the consequences of which depend on the measurement time and part cost. Conversely, limiting the number of repetitions in order to minimize resource usage may diminish the reliability of the results and expose the firm to increased risk of process failure.

Most commonly, standard operating procedures within a company dictate how a GR&R study is performed. These should include best practices for the number of samples, for the analysis of the resulting data, and should by and large be followed across various companies. For example, data collection cannot begin without first developing a practical template for recording data points. Typical short-term GR&R spreadsheets, as shown in Figure 2, might consist of ten parts, run nine times, using three operators. This same GR&R study can then be repeated across

Short Term Gauge Potential

| Parameter | AAB ABS Quad Crown | | Gauge | Slider-nm | | Date | TBD |
| Tolerance | 20 | | Part Name | Conventional | | Location | Factory D |

| | Operator1 | | | | Operator2 | | | | Operator3 | | | |
Sample	Trial 1	Trial 2	Trial 3	Range	Trial 1	Trial 2	Trial 3	Range	Trial 1	Trial 2	Trial 3	Range
1	45.93	45.94	45.93	0.006	45.18	45.42	45.75	0.568	45.26	45.75	45.69	0.484
2	48.29	48.3	48.29	0.010	48.3	48.18	48.42	0.240	48.75	49.26	48.75	0.516
3	47.9	48.29	48.3	0.393	48.29	48.3	48.18	0.126	48.42	48.46	48.26	0.201
4	48.07	47.9	48.03	0.165	48.09	48.29	48.09	0.201	48.18	48.42	48.07	0.343
5	48.75	48.69	48.9	0.211	48.94	48.93	48.94	0.006	48.93	48.78	48.94	0.164
6	48.26	48.75	48.69	0.484	47.9	48.29	48.3	0.393	48.29	48.3	48.18	0.126
7	48.75	49.26	48.75	0.516	48.69	47.05	48.29	1.648	48.93	48.29	48.93	0.637
8	49.42	48.46	49.26	0.955	48.75	48.69	47.9	0.841	48.29	48.93	48.29	0.637
9	53.18	53.42	53.05	0.371	53.26	53.05	53.09	0.216	53.05	53.29	52.93	0.363
10	45.93	45.78	45.75	0.185	45.75	45.26	45.75	0.485	45.69	45.9	45.29	0.611
Sum Total	484.476	484.796	484.954	3.302	483.156	481.462	482.712	4.730	483.790	485.389	483.342	4.082

| | | | R-Bar | 0.330 | | | R-Bar | 0.472962 | | | R-Bar | 0.408165 |

		Sum A	1454.225			Sum B	1447.330			Sum C	1452.521
		Average A	48.474			Average B	48.244			Average C	48.41737
									Average R-Bar	0.404

Upper Control Limit = D4*R-Bar

| D4 = 3.267 | For 2 Trials |
| D4 = 2.574 | For 3 Trials |

UCL - Range 1.039

*If any individual Range exceeds this limit, the measurement or reading should be reviewed, repeated, corrected or discarded.

Repeatability
For 2 Trials, REPT = UCL/1.128
For 3 Trials, REPT = UCL/1.693

| Repeatability Standard Deviation | 0.238 | | Total Standard Deviation | 0.267 |

Reproducibiltiy
For 2 Operators, 0.709 * (Max Average - Min Average)
For 3 Operators, 0.524 * (Max Average - Min Average)

| Reproducibility Standard Deviation | 0.120 | | Total Variability Spread @6s | 1.603 |

| | | | | P / T RATIO | 8.015 |

multiple days and all results combined to provide the long-term reproducibility.

Several typical analysis variations can be used to calculate the GR&R results. A simple X-bar and R-chart, which used to be done by hand and now more likely uses a spreadsheet or control software, graphs the data captured with the tolerance bars being the part tolerance. However, this method is generally considered to be less accurate and is often used more for process control. The previously discussed short-term GR&R spreadsheet analysis or an even more advanced ANOVA analysis, which requires specialized computer software, are considered the most accurate methods. These analyses of the GR&R data can provide information on part variation, repeatability, reproducibility, standard deviation, and percent of total variation, as compared to solely the percent of part tolerance.

GR&R studies results can be judged by the resulting Standard Deviation in the units the data was captured in or as a percentage of the part process tolerance (P/T ratio). In Figure 2, the part tolerance is entered into the header and the P/T ratio is calculated in the lower right corner. If the GR&R results are reported in percent of process tolerance, many high-precision manufacturing facilities want the measurement system or gage tolerance contribution to be less than 10% of that part's process tolerance. A P/T ratio <20% means the measurement tool can meet the process tolerance, where a greater P/T ratio risks a false positive or false negative measurement.

Another example consists of on-fly statistical analysis, as shown in Figure 3. Reproducibility measurements for areal roughness (Sq, ISO 25178) are performed over a couple of days through Bruker's Vision Map software. The flow of data is graphically represented within the upper and lower control limits (UCL/LCL), which determines the safe boundaries for production process together with expected 2-sigma dispersion (95.45%) from measurement mean. Any outliers are easily spotted, helping quality engineers to further investigate the root causes of discrepancy. It also illustrates how widely the results spread within the specified limits. Quantitative values like process capability (Cp) and process capability index (Cpk), which must be considered as respectively machine capability (Cm) and machine capability index
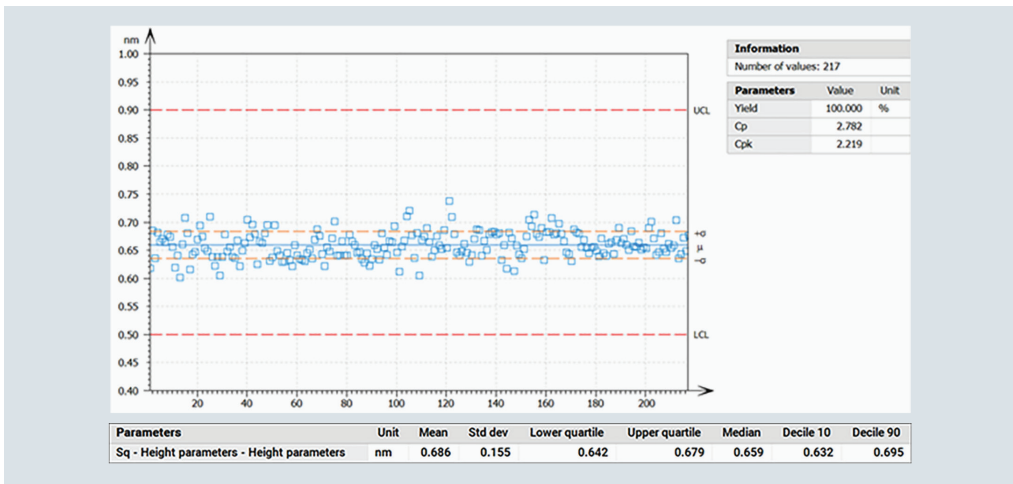
Parameters	Unit	Mean	Std dev	Lower quartile	Upper quartile	Median	Decile 10	Decile 90
Sq - Height parameters - Height parameters	nm	0.686	0.155	0.642	0.679	0.659	0.632	0.695

(Cmk) during a GR&R study, allow quick understanding of whether the measurement mean is metrologically capable to assess the process.

Maintaining Good Gage Performance

Before a GR&R study is initially performed, the measurement system must be set up and calibrated to manufacturer specifications. Long-term standard measurements must be performed at periodic calibration cycles to a known traceable standard to maintain the measurement system traceability and correlation. When performing a periodic calibration on a measurement system, the standard must be measured first, before any adjustments are made. This initial measurement not only indicates if the measurement system has drifted, but ensures that adjustments only happen when the results approach the calibration specification limit since continued adjustment adds additional measurement variation to that system. The ideal way to monitor a measurement system's stability is to periodically run "golden" part standards graphed with specifications, which should help to determine the system reproducibility over time.

Unfortunately, the initial GR&R is often the last time a GR&R is performed on the measurement system until there is a measurement problem, which is often flagged by seemingly ambiguous measurements or false part fails/passes found during a customer part audit. To avoid these "surprise" incidents, an ideal practice is to run a periodic GR&R of a measurement system with the golden parts. Periodic degrading of a GR&R can indicate issues with the measurement system itself or bring to light new environmental effects. Good practices are to maintain those initial GR&R parts (or a subset of them) to periodically gage or requalify the measurement system after repair or maintenance.

The Advantages of White Light Interferometry and Self-Calibration

Metrology instruments utilizing white light Interferometery (WLI) are non-contact, non-destructive and are among the most repeatable measurement systems on the market today. They provide true three-dimensional images quickly over a large area that fully characterize the surface with sub-angstrom repeatability. Many factors can contribute to the previously discussed measurement errors, and Bruker's WLI profilers are specifically designed to minimize these random and systematic errors with a core designed measurement component and superior environmental isolation.

All Bruker WLI tools have some form of standard or optional built-in active air isolation. Air isolation mounts allow for precise, vibration-free measurements in all environments. Low-noise digital cameras with ultralow-noise electronics assist in providing true characterization of the surfaces under test. The measurement head, tip-tilt cradles, and covers are all designed specifically for improved vibration and acoustic isolation in extreme production environments. The tool's base castings are CAD designed and modeled for maximum part isolation, as well as for minimum deformity from the environment (see Figure 4).
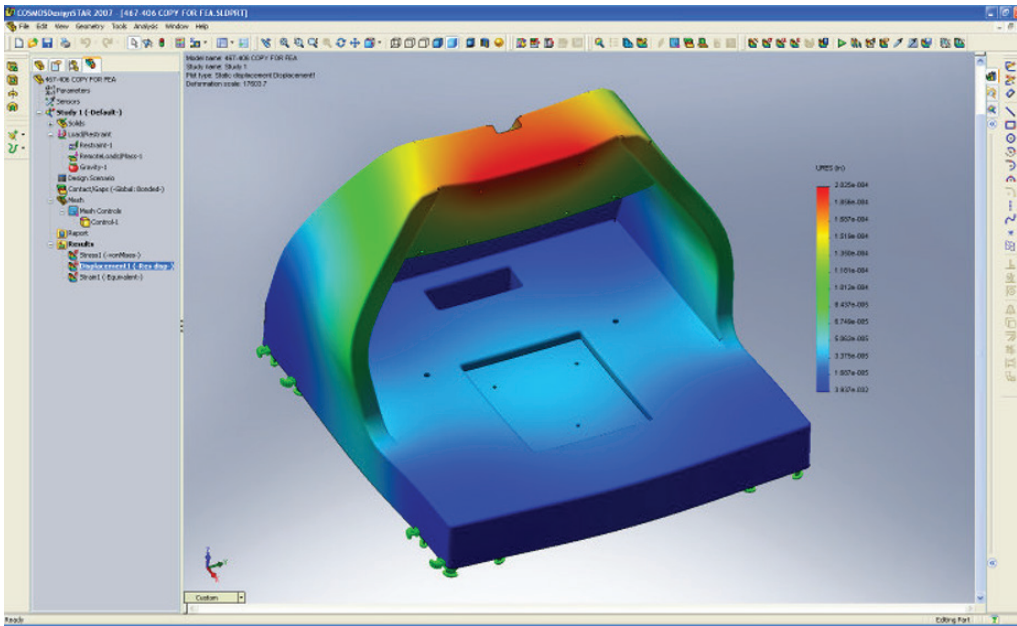
Bruker's high end WLI measurement systems incorporate an industry-leading laser interferometery reference signal module that monitors each measurement (see Figure 5). This type of technology is known as a secondary level traceable internal standard, which is close to being an absolute standard as the accuracy is traceable to the known wavelength of its stabilized HeNe laser (632.82±0.01nm). Every measurement is monitored and adjusted continuously over the entire measurement scan to minimize any short-term mechanical irregularities and any long-term thermal environmental drift. This self-calibration HeNe laser was incorporated into the measurement system to remove any minor measurement design errors from Abbe error (lateral offset between the reference signal and optical measurement axis), cosine error (angular offset between the reference signal and scan axis), and dead path error (differences between the two laser reference signals).
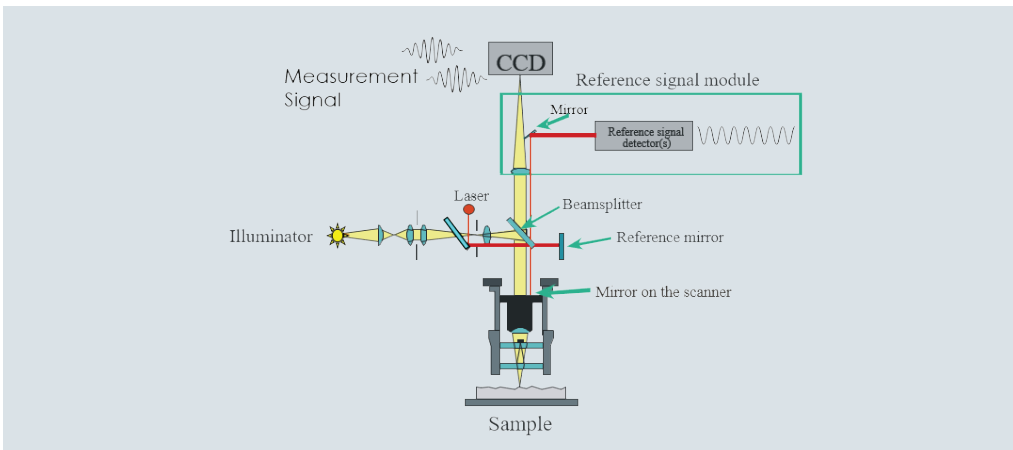
Not all industries require the precision of a WLI-based system with this self-calibrating reference signal technology, but it is a necessity for many sophisticated applications in semiconductor, data storage, MEMS, automotive, aerospace, optics, precision machining, medical or precision films. Self-calibration greatly improves system accuracy over time while improving tool-to-tool correlation in manufacturing plants around the world. Figure 6 shows a 23-µm height measurement with an extreme ambient temperature shift where the self-calibration reference signal removes the tens-of-nanometers thermal expansion error introduced to the entire volume of the tool.
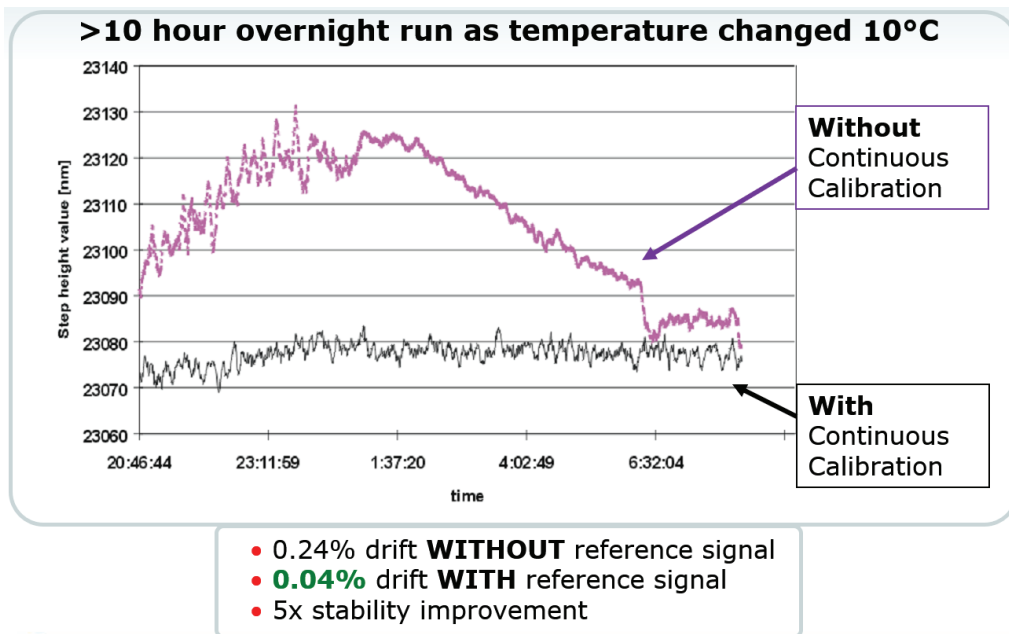
>10 hour overnight run as temperature changed 10°C

Without Continuous Calibration

With Continuous Calibration

- 0.24% drift **WITHOUT** reference signal
- **0.04%** drift **WITH** reference signal
- 5x stability improvement

All Bruker WLI profilers are designed with stabilized, uniform LED illumination that delivers superior image quality at any magnification. Bruker also has worked with major objective manufacturers on the design of interferometer objectives, making them thermally stable (see Figure 7). Stabilized optics and lighting are standard on all Bruker WLI measurement systems, leading directly to measurement stability for improved GR&R capability.

Bruker has also developed specialized software, and has incorporated it into the multi-core 64-bit Vision processing software to improve repeatability and gaging of the measurement systems. Phase-shifting algorithms for capturing mirror-like surface measurements that are insensitive to noise and vibration are standard on all systems. Variations of Vision's automatic pattern-matching software can find fiducials for repeatable part alignment, centering of the part, or centering the current capture image for repeatable analysis. This type of measurement part centering eliminates user intervention, thereby maximizing efficiency while achieving the best measurement gage capability by measuring the part in a repeatable location. Even the automated detection and analysis software has pattern finding built in to achieve optimum GR&R (see Figure 8).
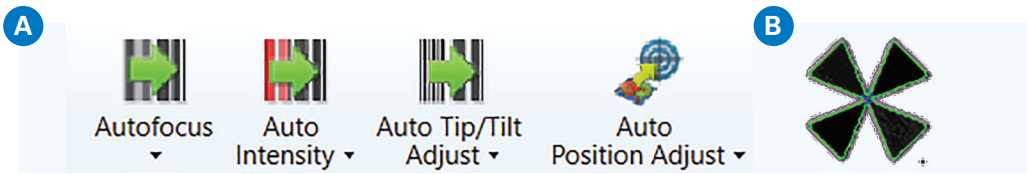
WLI has for many years played a critical role in the data storage industry for various demanding in-process measurements. The most demanding applications require data acquisition and analysis in a matter of seconds per part while achieving GR&R sub-angstrom measurement results over multiple days. Figure 9 shows the final GR&R analysis of a three-feature part study captured over three days to create a final sub-angstrom 1-sigma reproducibility result utilizing the aforementioned measurement system designs. This study used specialized Anova software to calculate the total 1-sigma GR&R for each measurement part region, and is compared directly to the desired specification.

nm	Day 1	Day 2	Day 3	Reproducibility
	0.042	0.045	0.041	0.052
	0.035	0.034	0.032	0.038
	0.066	0.076	0.068	0.086

Conclusion

Following a solid GR&R measurement analysis will give you, your company, and your customer confidence in the process measurement system and minimize waste in production due to unacceptable measurement variation. WLI non-contact optical profiling continues to be the technology of choice for high-volume throughput, highly accurate gaging, and long-term reproducibility for process control measurement. The advantage of WLI over other optical measurement technology is its ability to achieve the same sub-angstrom vertical resolution at any magnification. By working closely with customers at early stages of their measurement road maps, Bruker continues to evolve its measurement systems to face the ever increasing challenges of demanding production applications and environments. The inclusion of a self-calibration HeNe laser module has eliminated the need for regular calibration, reduced maintenance down time and improved the cost of ownership for manufacturers. Bruker's combination of other system improvements and isolation provide superior imaging while minimizing measurement error with advanced data analysis. A variety of Bruker WLI-based profiler configurations make it easy to ideally meet nearly every customer's individual measurement GR&R needs.

Authors

Roger Posusta, Senior Marketing Application Specialist, Bruker Nano Surfaces and Metrology Division (roger.posusta@bruker.com)

Samuel Lesko, Application Development Director, Bruker Nano Surfaces and Metrology Division (samuel.lesko@bruker.com)

www.bruker.com/MetrologySolutions