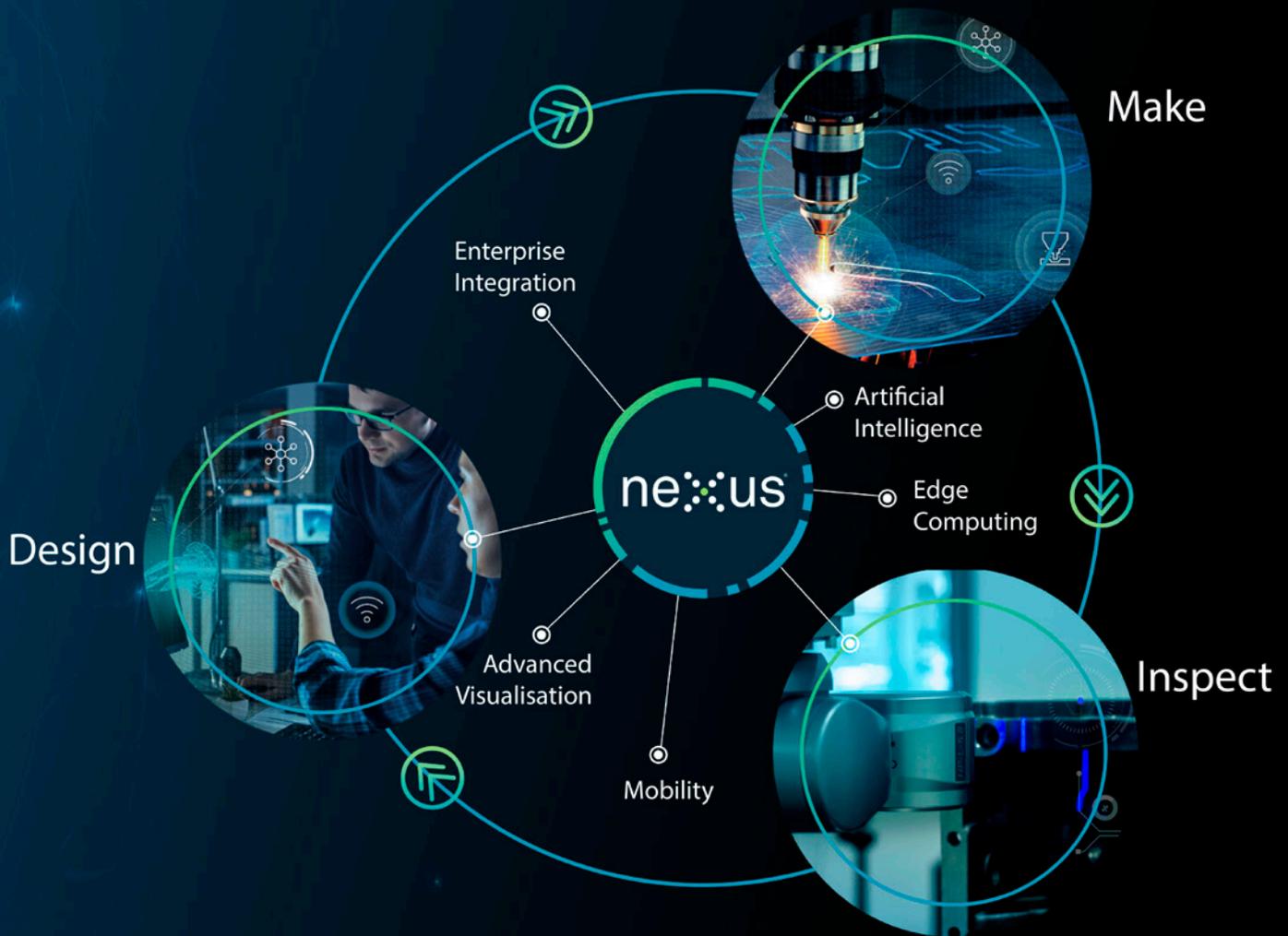


Delivering on the Promises of Digital Twins and Big Data

Balancing accuracy and precision in ISO 6336

Barry James, Senior Technical Leader, Hexagon Manufacturing Intelligence



As a community, we gear engineers collaborate and share ideas to progress our collective capability. Technology progresses based on our efforts, and we have seen solid advances in the performance of our products as they become quieter, cheaper, more efficient, and more power dense. The pages of this magazine (past and present editions) are filled with examples where talented engineers have dug deeper into a subject using a more precise approach to a particular area concerning gear performance. The implied belief is always that greater precision (complexity) in the calculations brings greater accuracy (alignment with reality).

Of all the performance characteristics of gears, durability/reliability is the most important. No matter how quiet, cheap, or efficient a machine is, it counts for nothing if operation fails. ISO 6336 delivers the standard for how to design gears, so one might assume that every gear engineer universally has a firm grasp of how to target a specific level of reliability—how to achieve the specific trade-off between over-design (excess size, weight, cost) and under-design (excess failures)—along with a statement on what is the anticipated failure rate for a given design. However, this turns out not to be the case.

A safety factor of 1.0 indicates 1 percent of the population will fail. However, if a different safety factor (say 1.2) is achieved, no information is given by ISO 6336 about the failure rate. A range of different research papers exist that explore the variability of gear failures, converting this data into numbers suitable for prediction through simulation, and these were summarized in a Gear Solutions article a few years ago (Ref. 1). Surely this provides the answers.

Unfortunately, the different sources suggest data that give wildly varying outcomes in predicted reliability (Ref. 2). The spread of results is not trivial. Take the safety factor of 1.2. Depending on which reference values for reliability you take, the predicted failure rate may be either 0.38 percent or 3×10^{-14} percent, (Ref. 3) i.e., from “reasonably frequent” to “vanishingly improbable even if applied to all the machines mankind has ever made”! It is almost as though you can “decide what result you want, and you will find a paper to give you that result.”

This is not to say that gear designers are clueless or negligent. In practice, each company has values in target safety factor/stress that have been developed and refined over the years based on experience and ‘not getting into trouble’, to be handed down to the next generation. However, this is still a long way from really being able to carry out a quantifiable trade-off between gear center distance (or any other design parameter) and failure rate. Hands up—which gear engineer wants to admit to their client or their boss that they do not really have any idea how many failures will occur for the gears that have just been designed? This is not how it is supposed to be.

So, despite all the efforts of gear researchers over the decades, there is surprisingly little agreement on how to predict what is the most important performance characteristic during design.

What is more is that this is reflected in another group of engineers with an interest in gear reliability—the maintenance profession. For any expensive or safety-critical asset, knowing how long to run a machine, when to maintain it, and when to shut it down, is vitally important. Maintain too often and large

Standard Deviation as % of Mean	Safety Factor			
	1.0	1.1	1.2	1.3
3	1	1.3e-5	3e-14	-
8	1	0.057	2.8e-3	1.47e-4
13	1	0.24	0.064	0.018
18	1	0.43	0.21	0.1
23	1	0.6	0.38	0.26

Table 1—Percentage failure rates at different safety factors depending on which reference data you take for the variability of gear material strength.




www.involutegearmachine.com

A Perfect Mesh

American ingenuity, service and support teamed with Japanese efficiency, quality and technology.

Kashifuji

CNC Gear Hobbing Machines



100+ Years of Manufacturing Gear Hobbing Machines

8 Models of Machines from 50 to 1000mm



CNC Gear Inspection and Double Flank Rolling Machines

65+ Years of Manufacturing Gear Inspection Machines

Machine Models to 850mm OD Capacity



CLP-35DDSF Gear Tester



GTR25 Double Flank Gear Roller

46449 Continental Drive
 Chesterfield, MI 48047
 Phone: 1-586-329-3755
 Fax: 1-586-329-3965
 rodney.soenen@involutegearmachine.com

costs are guaranteed, too infrequently and failure may occur, risking even higher costs and safety problems.

Practices such as Reliability Centered Maintenance (RCM) were developed to balance risks and optimize maintenance schedules, using a data-driven approach. So, what do RCM practitioners say about gears? Yet again, we see a variation in the recommended data for gear reliability (Ref. 4) that extends even wider than that from the design analysis research papers. Again, you can pretty much make up whatever result you like, and you can find data to give you that result. The situation is no better for our close cousins in the world of rolling element bearings. Here, the judgment is clear—major organizations such as NASA (Ref. 5) and the principal reference book for RCM (Ref. 6) written by one of its founders (Ref. 7) have decreed that bearings fail at random.

So, it is a salutary thought that, despite all our efforts to characterize the behavior of gears, designers do not really know what their failure rate will be, and those who maintain the machines that contain gears think they fail at random.

Some have suggested that this will change. “Big Data” and “machine learning” are on their way to rescue us! Everything will be monitored, observed, and correlated. “In the future, there will be no need for physics-based simulations, as everything will be a regression algorithm,” was how one leading engineer put it at a conference (Ref. 8).

This approach has profound problems. Firstly, without a physics-based framework, all data, on every machine, needs to be recorded and stored forever—you can never know which snippet of data is irrelevant and which will hold the key to insight. A few short minutes of rough calculations indicate that this would lead to absurd quantities of data being retained that would incur vast costs and, incidentally, lead to huge energy bills and environmental damage.

The second problem has been highlighted by the RCM community for over 45 years, which is that catastrophic failures tend to be very rare, meaning that there is little or no data from which failure models can be derived—take helicopter gearbox failures as an example. This is what is known as Resnikoff’s conundrum (Ref. 9).

Hexagon takes a different approach. Essentially, it is not physics *or* data, but physics *and* data—combining the established physics-based framework with the opportunities that Big Data and big processing capability bring. Big Data has its role to play in providing hitherto unprecedented quantities of data, but it needs to be interpreted within the framework of existing methods such as ISO 6336.

It is all very good to describe grand ideas with a broad brush, but the devil is in the detail and the proof that it works relies on demonstrating an implementation that delivers value. This has been the focus of Hexagon’s work over recent years with a number of globally renowned vehicle manufacturers.

The starting point is the vehicle usage data. Every vehicle gearbox is designed according to a duty cycle, which is intended to represent the usage that the gearbox will experience in operation. It may be accelerated/condensed (for the purposes of rig testing), but it is supposed to represent in-service usage regarding fatigue damage and reliability.

The problem is that, out of a fleet of nominally identical vehicles, each one will be driven differently. Small-scale studies have taken place to try to quantify this over the decades (Refs. 10, 11), but this has done nothing more than to scratch the surface of the issue and it has remained far too expensive to install telemetry equipment such as strain-gauged driveshafts on more than a handful of vehicles.

The situation regarding vehicle data is now changing. The development of connected and autonomous vehicles (CAVs) means that vehicle connectivity has reached levels not previously seen and large quantities of valuable data is available. For each company and each vehicle this varies, however, in general, it means that signals from the CAN (Controlled Area Network)

bus are available, including the torque and speed of the prime mover (the internal combustion engine or electric machine, as appropriate) and the selected ratio (if appropriate). Depending on the company, this can be collected on the vehicle and uploaded to the Cloud, whether this is periodically, daily, or more frequently (Ref. 12). Essentially, the driving profile of every individual vehicle can be known.

This provides the opportunity for a digital twin of the gearbox. Alongside Big Data and the Internet of Things (IoT), a *digital twin* is another phrase that is widely used, often misused, promises much, and usually under-delivers. What does it mean within this context?

Hexagon’s recent work (Refs. 2, 3) goes into the details of the origins of the term *digital twin* (Ref. 13) and the basis on which Hexagon uses the term in relation to others. In summary, the Hexagon digital twin is where the digital asset shadows the physical asset (the machine) during its in-service operation, extracting operational data and, potentially, feeding this back to the machine for performance optimization.

This is separate from a *design twin*, which is essentially a computer-aided engineering (CAE) model. However, there is a close relationship between a CAE model and an in-service digital twin. In the case of gearboxes, the digital twin was based on Hexagon’s Romax software, which has been used worldwide for gearbox design since its release in 1994. Meanwhile, data handling, processing, queuing, results plotting, etc. was handled by the various capabilities of Nexus, Hexagon’s open digital reality platform for manufacturing that is developed to provide connectivity and interoperability across all aspects of design, manufacturing, metrology, and in-service operation for all Hexagon’s client industries.

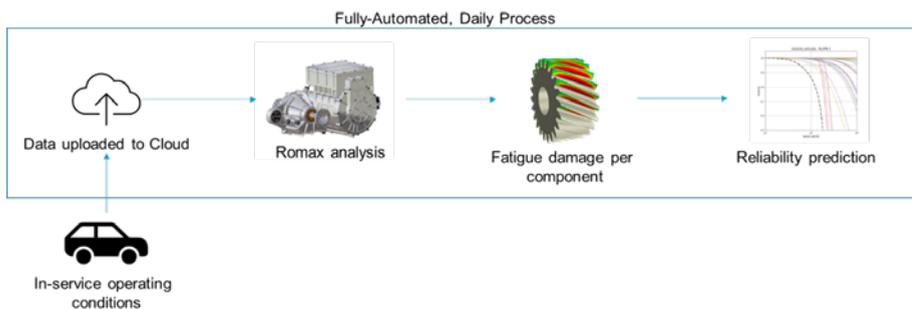


Figure 1—Overall workflow for Hexagon’s gearbox digital twin.

The work saw vehicle operational data uploaded to the Cloud where the Romax Digital Twin calculates fatigue damage and then predicts reliability for the gears and bearings. The reliability methods used were based on recommendations from the standards in the case of bearings and Hexagon's best judgment from the literature search. Detailed explanation and justification of this is covered in Hexagon's papers at 2023's VDI International Gear Conference (Ref. 3) and AGMA Fall Technical Meeting (Ref. 2).

One insight has been apparent from the start, which is the extent to which vehicle usage varies across the fleet and therefore affects gearbox fatigue. Of course, this makes sense qualitatively—everyone knows that some vehicles are driven hard, some gently, and this will influence gear failure rates—but acquiring the data meant that this impact could be quantified.

What this showed (Ref. 14) is that the variation in vehicle usage is so great that carrying out failure analysis purely based on vehicle mileage/hours of usage is so grossly inaccurate as to be meaningless. Even if the simulation model of the gearbox is perfect and the components behave with great consistency, omitting the variation of vehicle usage from the input data injects such enormous variability into fatigue damage that the results are nonsense. It is simply a case of "garbage in, garbage out."

In many respects, this explains the broad range of reliability data proposed by the research papers and used by RCM professionals, and the reliance by gear designers on experience-based values for safety factor. Since it has not previously been possible to know the usage data for each vehicle, failure analyses have been based on mileage/hours and components have appeared to fail (approximately) randomly, when this may well not have been the case had suitable data been available.

Hexagon's work is a process of identifying these key sources of "garbage" that cause this randomization and filtering them out, converting "unknowns" into "knowns," which can be quantified in the analysis. Rome was not built in a day and not all of these have been covered to date—the process is slow and steady and each time a new factor is incorporated into the analysis checks must be made to ensure that the results make sense. Nonetheless, sufficient work has been carried out to

Our High Torque Gears Fill a Key Design Gap



For over six decades, we've helped top engineers tackle tough gearing problems.

Learn how our gearing systems can be leveraged in your sector. Contact our engineers today for more information.

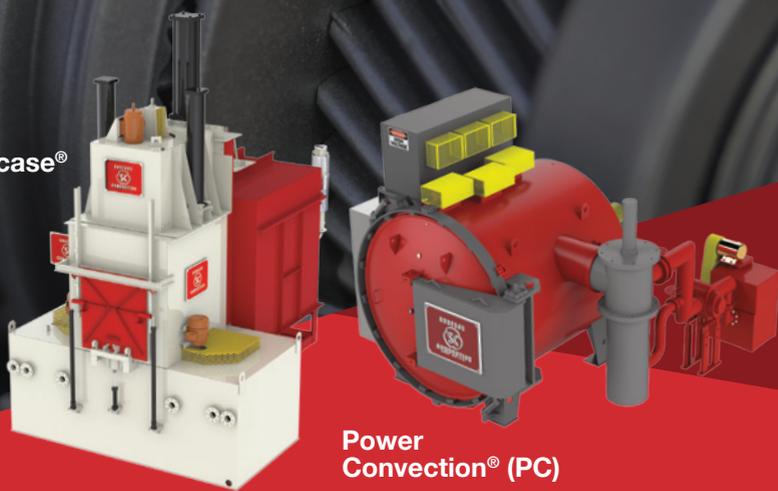


(320) 762-7133 • SpiroidGearing.com

ROBOTICS • AEROSPACE • DEFENSE • TRANSPORTATION • MEDICAL

Furnace Solutions for Atmosphere and Low Pressure Carburizing (LPC)

Allcase®



Power Convection® (PC)

Surface is Your Source.



info@surfacecombustion.com | 1-800-537-8980 | surfacecombustion.com

show that, within Hexagon's wider portfolio, it can account for variations in gear manufacturing (Ref. 15) and road-induced shock loads in its further work.

Thus, there is the potential to deliver unprecedented accuracy in the prediction of gear (and bearing) failures. The intention is to use this to recommend service intervals or maintenance inspections based on vehicle usage.

The description to this point assumes that the prediction of component reliability is perfect and all that is needed is accurate input data on vehicle usage, gear manufacturing, and shock loads. However, this may well not be sufficient. Most companies only have a vague estimation as to the fatigue strength of their gear materials, and it is highly likely, for example, that the data selected by the gear designer does not match the material's real performance. Frequently it is observed that Romax users automatically plump for "Medium Grade Case Carburized Steel" regardless of the actual composition of their material, its cleanliness, the capability of their heat treatment process, or their grit blasting/shot blasting/shot peening processes.

This is where the next phase of the digital twin work comes in. For each vehicle in the fleet, the usage data will be known. Not many of them will see their gearboxes fail but out of a fleet of many thousands, some will. Understanding these failures on a case-by-case basis will be assisted by being able to inspect the usage data to see if the vehicle was driven aggressively/unusually compared to the design assumptions. However, the key value comes from aggregating the failure data across the fleet.

Meecker and Escobar (Ref. 16) have illustrated "sample size analysis," whereby the confidence in a result relates to the number of samples being tested plus how far along the test regime they have lasted. Since we have a very large sample size, we can derive insight with great confidence even when there are just a few failures.

Adoption of the digital twin across a fleet of vehicles will provide sample sizes and confidence that will dwarf the approaches used to date to understand reliability. Often companies sign off on a gearbox design by rig testing a sample of prototype gearboxes prior to the start of production. This sample is

small—somewhere around 5, perhaps a few more. Analysis shows (Ref. 17) that this actually provides very little confidence in the result. Even university research programs, on which the S-N curves of gear materials are based, extend to a couple of hundred samples, although FZG can point to a dataset of around 1,000 pitting and 1,700 bending test results for the common case hardening steels 16/20MnCr5 or 18CrNiMo7-6 (Ref. 18). This is all very good if you use one of these steels, but Hexagon has still seen large variations in results for 20MnCr5, for example, owing to variability of heat treatment and shot peening, factors which are defined by each manufacturer. By comparison, vehicle manufacturing companies could build up sample sizes running into the hundreds of thousands or even millions of gears.

This is not to say that rig testing before the start of production should be stopped, but rather that the processing of vehicle usage data and component failure data, combined across a sizeable fleet of vehicles/machines, could provide unprecedented accuracy in the ability to predict component fatigue and reliability.

How would this all actually work? Recently (Ref. 17), Hexagon carefully stepped through the process for how a digital twin could automatically adjust the calculation such that the predicted failure rate matched the observed failure rate. This would not involve throwing away ISO 6336, or appending additional complexity to its calculation, but simply adjusting parameters such as application factor, allowable stress, and Weibull shape parameter. Note the change in emphasis from precision (adding complexity) to accuracy (making sure the predictions match the observation).

Basing the method on ISO 6336 has another key advantage, which is that it uses the language that existing gear engineers understand. The leading experts in

each manufacturer will have learned their trade using this approach and have had at least 2 decades of experience in its application. This approach builds on that experience and tweaks the input values, providing evidence to justify such changes.

The current limitations of ISO 6336 described apply to all implementations of the standards in all codes. ISO 6336 permits companies to adjust values for gear material fatigue strength and application factor. What we show here is that it is now possible to get a clear indication of these values. The answer is not to throw away ISO 6336 (as some advocates of Big Data suggest (Ref. 8) or add further layers of complexity (precision) to the calculation, but to develop an approach that uses each company's data to derive the correct input values to make their ISO calculation accurate. Prioritize accuracy over precision.

In summary, there exists the possibility to deliver on the grand promises of Big Data, digital twins, IoT, etc., promises that are too often spoken about in revolutionary, utopian terms, but which usually fail to deliver. By taking a pragmatic approach and taking advantage of recent advances to convert important input data (vehicle loading, manufacturing accuracy, shock loads) from assumed to confirmed values, Hexagon is working with its clients to deliver these promises.

Behind all the gear engineering and maths is *Nexus*, the framework that coordinates all the data and makes the implementation of the digital twin possible. Hexagon's roadmap sees *Nexus* taking a significant role in the development of digital twins, extending from vehicle gearboxes to other fleets of geared machines such as wind turbines, then other fatiguing components and other physics that Hexagon covers (noise, heat, etc.).

Despite failing to live up to expectations so far, digital twins can and will deliver, as current projects are demonstrating.

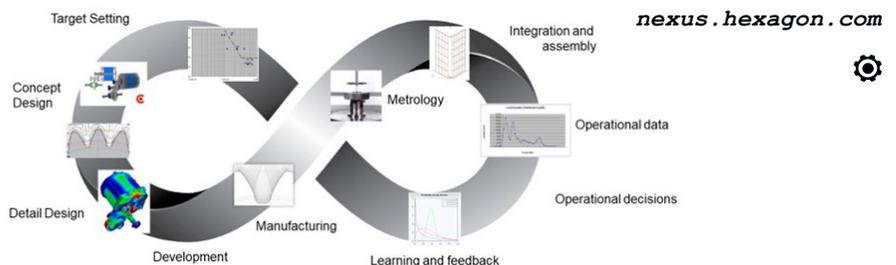


Figure 2—Completed digital thread covering the life cycle of gears.

References

1. Heim, M. et al. "Reliability of Gears: Determination of Statistically Validated Material Strength Numbers," *Gear Solutions*, February 2020.
2. James, B., et al. "Use of Gear Reliability Data in a Cloud-based Gearbox Digital Twin," AGMA Fall Technical Meeting, Detroit, October 16–18 2023.
3. James, B., Runkel, D., "Use of Gear Reliability Data in a Cloud-based Gearbox Digital Twin," VDI International Conference on Gears, Munich, September 13–15, 2023.
4. <https://www.ge.com/digital/documentation/predix-apm/latest/reliability-analytics-distribution-types.html>
5. Figure 2.1, page 2-2, "NASA Reliability-Centered Maintenance Guide for Facilities and Collateral Equipment," September 2008
6. Page 239 and page 159 of "Reliability Centred Maintenance—John Moubray, The Aladon Network, 1999; <https://www.elsevier.com/books/reliability-centered-maintenance/moubray/978-0-7506-3358-1>
7. Paragraph 3, https://en.wikipedia.org/wiki/Reliability-centered_maintenance
8. Dr. Victoria van Camp, CTO, SKF, Public discussion, eDSIM conference, Hotel Maritim in Darmstadt, Germany, 25–26 September 2018.
9. Resnikoff, H. L. *Mathematical Aspects of Reliability-Centered Maintenance*, Defense Technical Information Center, San Francisco, 1977.
10. "Manual Gearbox Endurance Rig Test Schedule for Gear and Main Bearing Life," BLS.57.02.505—British Leyland Motor Corporation Standard, December 1973.
11. MacDonald, C., et al. "Manual Gearbox Fatigue Test Schedule Predictions from Car Road Load Data," C462/6/170, Autotech '93, The Institution of Mechanical Engineers.
12. Mingfei Li, Dr. Fabian Noering. "Customer oriented electric drivetrain durability development," VW AG, CTI Berlin, December 2023.
13. Grieves, M., 2016, "Origins of the Digital Twin Concept," DOI: 10.13140/RG.2.2.26367.61609.
14. James, B., and Long, L. "Demonstrating a Cloud-based Gearbox Digital Twin using telematics data for reliability predictions," CTI USA, Novi, MI, 15–16 May 2023.
15. James, B., Runkel, R., "Developing a digital thread linking gear design to manufacturing simulation and metrology," VDI International Conference on Gears, Munich, 12 September 2022.
16. Meeke, W. Q. and L. A. Escobar. *Statistical Methods for Reliability Data*, New York: Wiley, 1998.
17. James, B., Runkel, R., "Establishing a Digital Twin with Self-Correcting Reliability Predictions Based on Large Fleet Data," CTI Berlin, December 2023.
18. Email correspondence with Michael Geitner, Gear Research Center (FZG), Technical University of Munich, February 2024.

For Related Articles Search

software

at geartechnology.com

Footnote: For those perhaps not so familiar with the English language and the nuances of different definitions, it is perhaps necessary to emphasize the difference between *accuracy* and *precision*. If two shots are close to the bullseye but far apart from each other they are *accurate* but not *precise*. If two shots are tightly grouped but away from the bullseye, they are *precise* but not *accurate*. If both shots are close to the bullseye and tightly grouped, they are *accurate* and *precise*.

RINGS 4" - 144" OD • DISCS UP TO 60"

CARBON • ALLOY • STAINLESS

FORGING
AHEAD
OF THE PACK

**Fast.
No Fine Print.
No Premium.**

At McInnes Rolled Rings, we provide quality products, shipped in as few as 4 days. And we partner that with exceptional customer service to forge the perfect partnership with our customers.

MADE IN THE
USA

**McINNES
ROLLED RINGS**

1.877.695.0280 • www.McInnesRolledRings.com