Effect of Assembly Errors in Back-to-Back Gear Efficiency Testing

M. Andersson, M. Sosa, S. Sjöberg and U. Olofsson

As gear efficiency is improved in small steps, it is important to be able to distinguish actual improvements from scatter that can occur while testing. An FZG back-to-back gear test rig was used to investigate how the assembly and re-assembly of the same test setup affects the measurements. A spread in torque loss between one assembly and another of the same test setup were observed. Rig conditions also affected the spread in input torque. With knowledge of how the spread in torque loss varies due to assembly, test results could be distinguished between changes due to assembly and actual differences between tests.

Introduction

Testing of gears can be performed in various types of test rigs with different degrees of complexity, from clean model test rigs via functional back-to-back test rigs to full scale system test rigs. Because of the high efficiency of precision gear drives, and because progress in gear drive development are measured in tenth of percentages, for instance (Ref. 1) and (Ref. 2), makes it necessary to have low scatter between test setups, and to minimize errors that can occur while testing. This will make it possible to distinguish natural errors from actual improvements. While investigating noise and vibration, Åkerblom (Ref. 3) discussed that variations in measured noise and vibration with the same gear pair could be due to assembly. Oswald et al. (Ref. 4) also discussed the differences in noise level with the same gear system could be due to assembly. Sound and vibration are system parameters, just like efficiency. The sound generated from the contact between a gear pair is affected by bearings and bolted joints. Gearbox efficiency is affected by bearings, seals, oil level, as well as the contact friction between two gear teeth in contact. No sensitivity or uncertainty analysis has been found in literature on the effect of re-assembly on the efficiency of a back-to-back test rig.

In general, efficiency tests do not quantify the effect of assembly errors. With knowledge on how measured results might spread, future tests can then be compared to determine whether the test results should be ascribed to assembly, or to some other external factor, such as surface roughness and gear geometry.

The goal of this study is to increase the understanding on how and whether gearbox efficiency can vary due to assembly and re-assembly. To investigate this, both theoretical and experimental sensitivity studies were performed on an FZG back-to-back gear test rig.

Method and Materials

Test equipment. Theoretical and experimental sensitivity studies are defined as follows. The theoretical sensitivity study focuses on the uncertainty of the different parameters measured during an efficiency test. The experimental sensitivity study analyses the effect on the efficiency results of assembly, as well as other parameters such as oil level, preheating of the test rig, and unloading of the inside torque. Both studies were performed by analyzing an FZG back-to-back gear test rig with an efficiency test setup (Ref. 5). Efficiency was measured as input torque, torque loss, from the motor to the power loop. The power loop consists of the two gearboxes connected by a load clutch. A sketch of the rig can be seen in Figure 1.

The gear test rig was taken apart and reassembled between assembly tests. The order in which the test rig was disassembled is as follows. To be able to remove the gears in the slave gearbox (#3), the motor (#5), followed by the torque sensor (#4) were removed. The gears in the test gearbox (#1) were then removed. The opposite procedure was made to put the test rig back together, tightening all bolts to a specified torque. To minimize human error the same operators were used in all tests. The same standard FZG C-PT spur gears were used in the slave and test gearbox in all tests; their dimensions can be seen in Table 1. A running-in procedure (Ref. 5) was followed.

For dip lubrication, a commercially available polyalphaolefin with a viscosity of 64.1 cSt @ 40 °C and 11.8 cSt @ 100 °C, and a density of 837 kg/m^3 was then added to both gearboxes. An oil level to the center of the shaft was used in both gearboxes in all tests.

The procedure of taking the rig apart, putting it back together and adding lubricant is defined as one assembly.

FZG efficiency testing. To test the efficiency variation due to different assemblies the following test procedure was devised. The FZG gear test rig was assembled and then loaded to

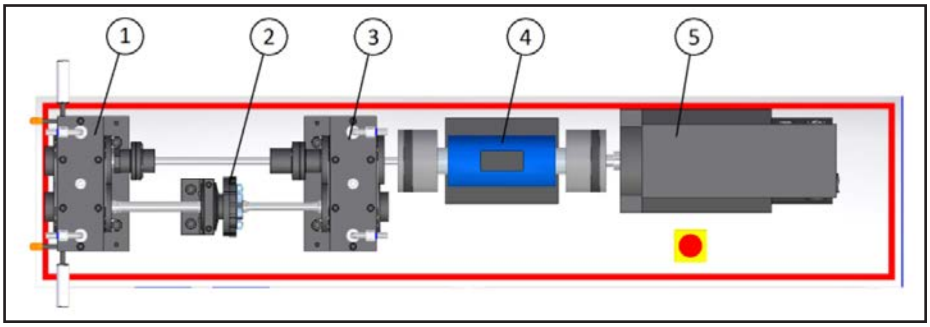

Figure 1 Schematic of the FZG back-to-back gear test rig with its most important parts: #1 test gearbox; #2 load clutch; #3 slave gearbox; #4 torque and speed sensor and #5 motor.

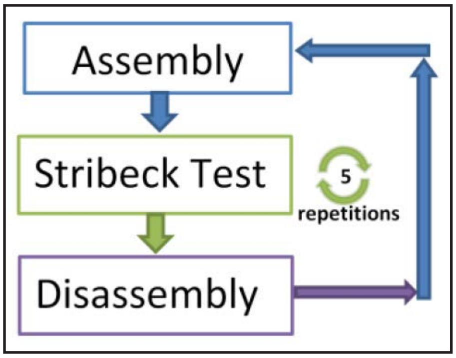

Figure 2 Flow diagram showing the order of the Stribeck test method.

94 Nm (FZG KS 5). Pitch speeds from 0.5–20 m/s were tested for five minutes at a lubricant temperature of 90 °C in both gearboxes. This test procedure is known as a Stribeck test (Table 2). These tests were repeated five times sequentially for each assembly. A schematic diagram showing this procedure is shown in Figure 2. In all assemblies, an oil level to the center of the shaft was used. The same assembly and disassembly procedure was used in all assembly tests.

The first four assemblies were used as a benchmark for the subsequent assemblies; these consisted of the assembly procedure described above and running the test combination shown in Table 2. In the first four benchmark tests the oil level was controlled by visual inspection.

Three further test conditions were chosen to investigate their effect on the spread in measured torque loss. They are as follows:

Loading and unloading, assembly 4UN Loads between 0–535 Nm can be applied to the gear rig; it is most likely that the same load will be tested several times — not necessarily in the same test plan. To study the influence of deviations in inside power loop loading, the torque was unloaded and loaded to the same value at the beginning of each speed repetition. This test was denoted as 4UN because it was performed using the same assembly setup as in test four.

Oil level, assemblies 5 and 6. Dip lubrication is the most common way to lubricate a gearbox, where the gears splash in the lubricant. The oil level is important when studying efficiency, since churning losses can be significant. Following the five initial tests, two assembly tests (5 and 6) were performed to determine the influence of oil level on efficiency. The initial four assemblies and the unloading test (4UN) were performed by observing that the oil level was at the center of the shafts, without measuring the oil level itself. In assemblies 5 and 6 the oil level was set to 103 mm, with a precision of ±1 mm from the bottom of the gearbox, corresponding to the center of the shafts.

Pre-heating of the gear test rig-assemblies 7 and 8. Components expand due to the substantial increase in temperature that occurs when performing standard efficiency tests, which in turn might affect results. In order to determine the influence of pre-heating the test rig before a Stribeck test, two assemblies were pre-heated for twelve hours to 90 °C prior to testing, with testing oil at standstill. The oil level was measured to be 103 mm in these assemblies as well.

Table 1 Basic geometry of the test gears used

Parameter	Unit	Gear	Pinion
Number of teeth	–	24	16
Module	mm	4.5	
Centre distance	mm	91.5	
Face width	mm	14	
Tip diameter	mm	118.4	82.5
Pitch diameter	mm	109.8	73.2
Pressure angle	°	20	
Working pressure angle	°	22.4	

Table 2 Test schedule of the Stribeck tests

Test number	Load [Nm]	Seed [m/s]	Duration [min]	Lubricant temperature [°C]
1	94	0.5	5	90
2	94	1	5	90
3	94	2	5	90
4	94	8.3	5	90
5	94	15	5	90
6	94	20	5	90

Theoretical sensitivity study, uncertainty in measured data from Stribeck tests. Test results can also be affected by the uncertainty of results of the measured variables. In each Stribeck test the temperatures in the test and slave gearbox, the torque inside the power loop, and the speed and input torque from the motor were measured. In this test rig eight outputs with a range of 0–10 V and a 12-bit resolution are provided by the manufacturers. These voltages are multiplied by pre-set scaling factors to achieve the appropriate sensor reading. The voltages are sampled using a DAQ NI-6009 12-bit resolution analogue to digital converter, over the range of 0–10 V. A sampling rate of 1 Hz was used. The sensors are described below.

Torque meter inside power loop. Torque is loaded onto the shaft to the right of the load clutch (Fig. 1; # 2). However, the inside torque is measured on the shaft left of the load clutch by a full bridge torsional strain gauge connected to a telemetry system. The left shaft is calibrated to have a linear relationship between applied torque and angular deformation.

The torsional strain gauge is a full bridge configuration, with the four equal strain gauges to be connected around the perimeter of the shaft at 60 mm from the gearbox sidewall. Being a full bridge configuration, strain gauge measurements are insensitive to temperature. The telemetry system as a whole has a signal bandwidth from 0–10 kHz; sensitivity drift of 0.015% /°C; a resolution of ±0.030 Nm over the full-scale output; and a full-scale output error of 0.3% and nonlinearity of 0.2% — from 60 to 368 Nm.

Loading is performed by applying a torque onto the clutch (Fig. 1; #2). The torque can be decomposed into two components, the force and the lever arm. Force is applied by dead weights onto the lever arm (the weights have a tolerance of ±5 g). The lever arm is 500 mm from the center of rotation. But due to large shear strains when testing, the action line of the force may not be perpendicular to the lever arm if the lever arm at the start of the test is perpendicular to the real torque loaded onto the rig. The position of the load on the lever arm is controlled by a wedge on the loading bar, as well as a wedge slot. A quantifiable error could arise from the overall deviation in load torque due to weight, positioning, and angle of the loading; however, the inside power loop torque is corrected during the test by the operator who can adjust the torque to match the torque reading shown by the torsional strain gauge. The operator always ensures that the inside torque is ±2 % of the nominal load.

Torque meter outside power loop and speed. In each test the input torque was measured by an accurate torque sensor (Fig. 1; #4). This torque sensor can measure torques up to 200 Nm and speeds up to 8,000 RPM. It measures torque with a sensitivity of 0.05%/10° C and with a measurement uncertainty of ±0.08 Nm over the full scale output, including hysteresis.

A flexible coupling, tightened to a specific torque, connects the slave gearbox to the outside torque sensor, and this device in turn is connected to the input motor by an identical coupling. Guiding pins are used to align all three components — thus ensuring minimal misalignment between the shafts. These backlash-free flexible couplings help ensure that the system can self-align if misalignments do occur.

The input power is supplied by a three-phase servo-controlled induction motor (Fig. 1; #5). To determine its speed uncertainty during operation, its speed is measured during testing at all testing speeds. Results show a deviation of ±2 RPM over all speeds when compared to the nominal speed. It is, however, important to note that in this study speed is of minor importance because all measurements are compared using torque loss, effectively decoupling speed from efficiency.

Temperature sensors and control in the gearboxes. A PT-100 sensor is mounted in the oil sump between the gearbox casing and the gear, both in the test gearbox and slave gearbox. Typically, this sensor has an uncertainty of 0.03 °C – 0.15 °C (Ref. 6). Independently, the temperature sensor was shown to have an uncertainty of 0.2 °C when compared to a known source. Note that all tests were performed at a controlled temperature of 90 °C. No further inves-

Table 3 DAQ resolution versus sensor uncertainty

Parameter	Unit	Testing range	DAQ resolution	Sensor measurement uncertainty
Temperature TGB	°C	30-120	0.024	0.2
Temperature SGB	°C	30-120	0.024	0.2
Inside power loop torque	Nm	0-372	0.049	0.3% FSO Noninearity 0.2% (60 to 368 Nm)
Outside power loop torque	Nm	0-20	0.012	0.5 % FSO (0.08 Nm)

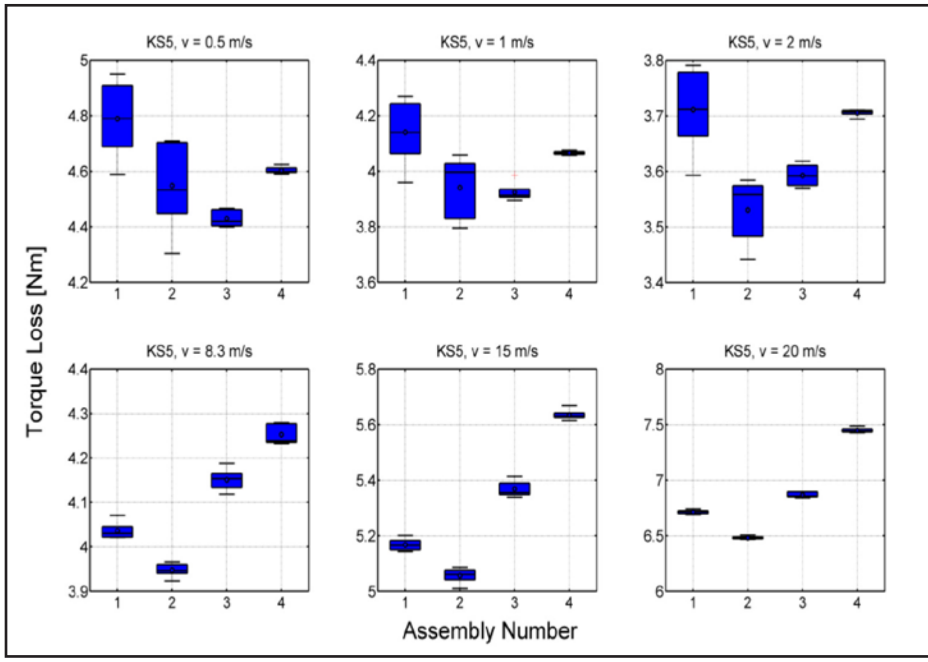

Figure 3 Reference error plot at 94 Nm, at 6 different speeds.

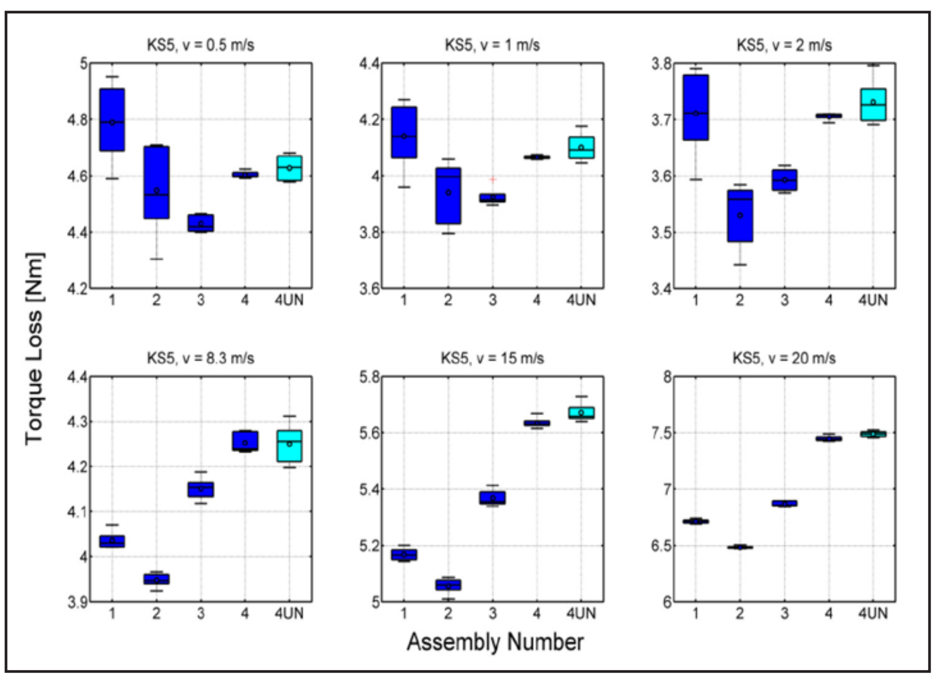

Figure 4 Comparing unloading of assembly 4 with reference error plot at 94 Nm, at 6 different speeds.

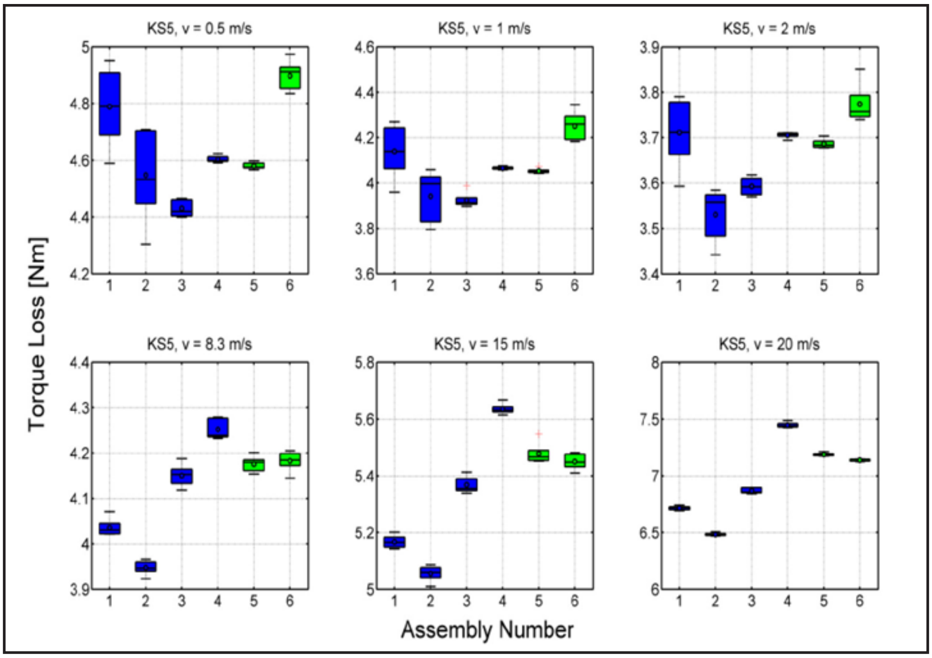

Figure 5 Comparing effect of oil level with reference error plot at 94 Nm, at 6 different speeds.

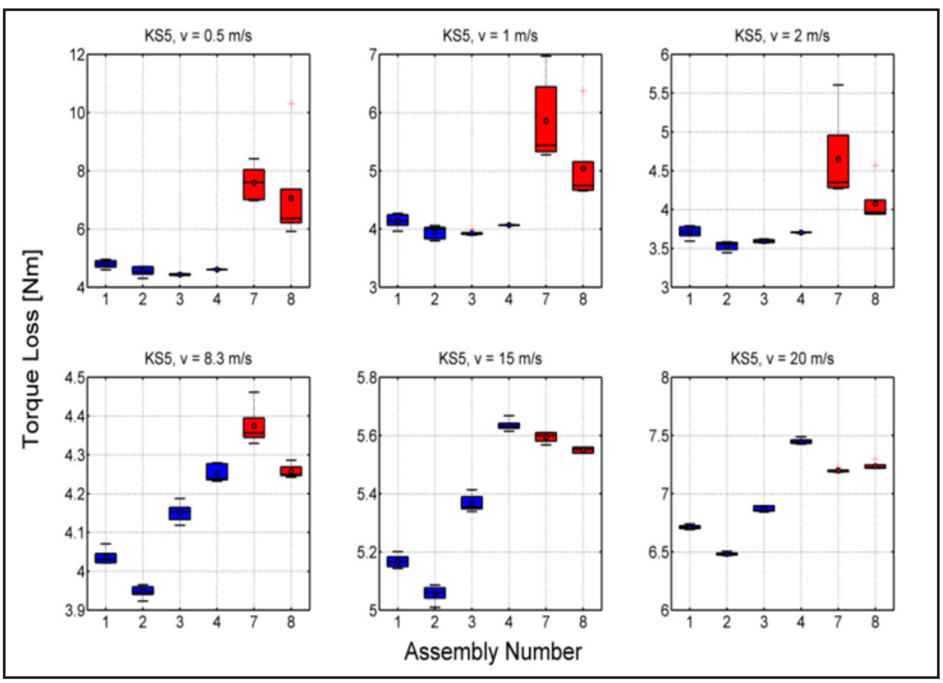

Figure 6 Comparing effect of preheating with reference error plot at 94 Nm, at 6 different speeds.

tigations were made to determine the uncertainty of the temperature sensors, as during testing the temperature is controlled to ±3 °C, making the sensor at least 3.3% accurate over this range of six degrees.

Signal processing. Table 3 shows efficiency test parameters as well as the DAQ resolution. At a 1 Hz rate over five minutes, each test has three hundred samples of all the parameters shown in the table below, which allows the means to be calculated. In the specific case of the outside torque, a tare value is calculated from the load at the start and end times of each test for each assembly. A statistically accurate tare is thus calculated, thus eliminating the zero drift over time.

Results for Stribeck Test and Assembly

The first four tests were devised as a benchmark measurement and to analyze the spread between them. Results are presented in the form of box plots. Box plots are defined as a graphical way to represent the median, upper and lower quartiles. In this paper a circle was added in each box plot showing the mean value, and the whiskers represent the maximum and minimum value in one dataset. Furthermore, if the boxes are separated from each other they represent a statistically significant result and can be interpreted as a graphical ANOVA with 95% confidence that the medians do differ. It can be observed by analyzing the different assemblies at different speeds that the spread is speed-dependent. Additionally, the scatter in assemblies 1 and 2 at low speed covers the same spread as the two subsequent tests. Furthermore, from 8.3 m/s onward, each assembly is statistically different.

Figure 4 shows the influence of unloading the inside torque. This variation in the test procedure compares the effect of assembly versus the effect of unloading and loading the inside torque. It can be observed that unloading does not change the level of torque loss, but does change the scatter for a specific test.

In Figure 5 the influence of setting a precise oil level (assemblies 5 and 6) is compared to the reference test (assemblies 1 – 4). The figure shows oil level does not influence the scatter of the torque loss at any speed, but influences the torque level in speeds from 8.3 m/s onwards.

Lastly, Figure 6 compares the effect of a long pre-heating period, 12 hours, to the first four assembly tests. A large significant difference can be seen below 8.3 m/s, in which both the scatter and level increase considerably. Mean torque loss at 0.5 and 1 m/s for the pre-heated test increased by almost 30–40%, while extreme values differ by about 200% at 0.5 m/s and 100% at 1 m/s. At high speeds the scatter is similar to tests 1 to 4.

In order to achieve one of the aims in this work, to determine the influence of the assembly methodology when measuring efficiency, tests 1 – 6, a pooled standard deviation was calculated to determine the spread in torque loss at each tested speed. Figure 7a shows the mean torque loss (continuous line), as well as dashed lines showing the expected spread for ±1 standard deviation. As previously discussed, each speed has a different scatter in the torque loss due to the assembly methodology. From Figure 7a the maximum scatter over the entire speed range is ±0.45 Nm and a minimum scatter of ±0.10 Nm.

In order to estimate the reproducibility (signal-to-noise ratio) of assembly 1–6, standard deviation of the torque loss at each speed was divided with the mean value of the torque loss at each speed (Fig. 7b). The reproducibility varied between 2.42% and 5.04%; the best reproducibility was yielded at 2 m/s and the worst at 20 m/s.

Results for measurement uncertainties. The measurement uncertainties calculated from the root mean square of the DAQ bit resolution and the sensor measurement uncertainty are presented in Table 4. The measurement uncertainties in temperatures, inside power loop torque and outside power loop torque measurements are dominated by the sensor uncertainty.

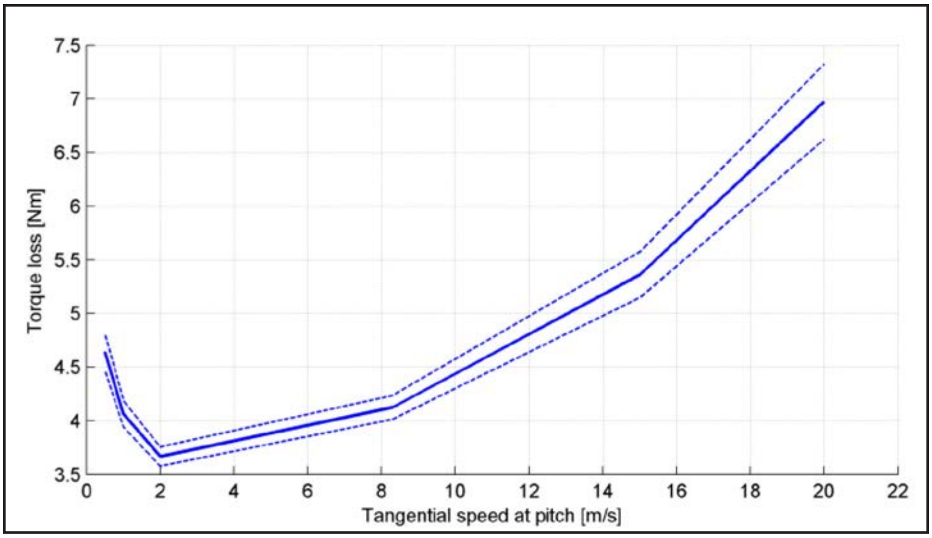

Figure 7a Mean and ±1σ of test 1 to 6. Continuous line mean, dashed line ±1σ.

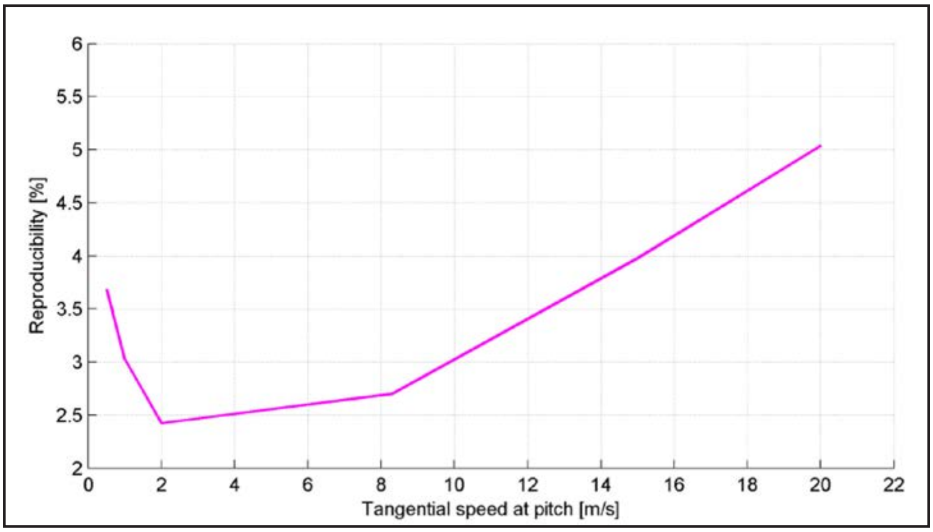

Figure 7b Reproducibility of test 1 to 6. Best reproducibility at 2 m/s

Discussion

The four initial benchmark tests (Fig. 3) show a spread in torque loss between the same test setup for different assemblies. This is similar to the sound transmission results from Åkerblom (Ref. 3) and Oswald et al. (Ref. 4) regarding gearboxes where assembly influenced the system parameter sound. For system-level parameters, including gearbox efficiency, one must take into account the way the gearbox is assembled. The spread in torque loss for the three lowest speeds decreases for each assembly. It is not known whether this is due to wear after each assembly, or whether the operators had increased their assembly skill. The spread is lower for the three higher speeds between each assembly than what is shown in Figure 7a. This is because at higher speeds the mean torque loss is dominated by the amount of lubricant in the gearboxes (Fig. 5). To achieve a more realistic torque loss spread between assemblies (Fig. 7a), more tests are needed with the same oil height measurement method as in assembly 5 and 6.

Figure 4 shows a slightly higher spread in assembly 4UN, when unloading and loading to the same load in the same assembly. However, the torque loss is at the same level as the initial assembly 4. With a measurement uncertainty of 0.081 Nm in torque loss, all tests at each speed are within that range and thus can be said to come from the same assembly. It seems that unloading and loading with the same load does not have as large an effect on the measured torque loss as does a new assembly.

In assemblies 5 and 6—in which the oil level was set to 103 mm from the bottom of the gearboxes—a difference in torque loss can be seen between the two assemblies at lower speeds (Fig. 5). In fact, assemblies 4 and 5 are very similar, but as the speed increases, the effect of an accurate oil level is shown. It seems that the differences in torque loss at slower speeds disappear as the speed increases, and speed-dependent losses dominate as full film lubrication prevails. A precise oil level minimizes unwanted differences in torque loss at higher speeds.

A comparison between the four initial assemblies and preheating for twelve hours can be seen in Figure 6. The torque loss between assemblies 1–4, versus 7–8, is significantly larger at slow speeds, but decreases as the

Table 4 Measurement uncertainties in the Stribeck tests

Parameter	Measurement uncertainty
Temperature TGB	0.2 °C
Temperature SGB	0.2 °C
Inside power loop torque	0.3% FSO Nonlinearity 0.2% (60 to 368 Nm)
Outside power loop torque	0.5 % FSO (0.08 Nm)

speed increases. The unloading (4UN) and lubricant level tests (5 and 6) do not differ as much in torque loss as when the gearboxes were pre-heated. The reason for this could be that the lubricant chemically reacts on the gear surfaces as the gearboxes are heated for an extended time (Ref. 7). The effect at low speeds may be explained by unwanted stresses in the assembly caused by expansion of components due to the increase in temperature. It is not known why the behavior changes at high speeds.

The sensitivity study investigated the uncertainty from each sensor in the Stribeck tests; except for the outside torque, the uncertainty with which test parameters can be measured in each Stribeck test is limited by the DAQ. An uncertainty of 0.081 Nm in outside torque loss, 0.3% FSO inside the power loop, and 0.2 °C in lubricant temperature should be sufficient resolution in this type of testing. Rather than utilizing a data acquisition device with better resolution, minimizing assembly spread (maximum value of ±0.45 Nm and minimum value of ±0.10 Nm) is more important to distinguish differences between tests. If maximum measured test results can be estimated, the scaling factors could then be lowered for better resolution.

Future efficiency tests can be compared using Figure 7a. It describes the mean and the standard deviation by pooling the four initial tests as well as assembly tests 5 and 6 (assembly tests in which the oil level was precisely controlled). Results of future efficiency tests can be compared to this spread to determine whether the results lie within the variation due to assembly or not. Furthermore, Figure 7b shows the reproducibility of assemblies 1–6. The lowest chance of a repeated test is at 20 m/s, where at that speed the oil level has a significant effect on reproducibility. Higher reproducibility will be achieved when the oil level is strictly controlled.

The test parameters in this study are commonly used with regard to the tested load and speeds; in order to achieve a more statistically accurate comparison, more tests at a controlled oil level should be performed. In standard efficiency tests efficiency is characterized by torque loss; however, if the true efficiency is measured, the speed spread should also be carefully considered. Also, since testing is performed at other loads, the effect of variations in load should also be determined. A test procedure that quantifies how much each test parameter affects torque spread is also required.

Conclusions

The spread in torque loss due to assembly methodology was quantified, as well as the overall uncertainty of measurements of temperature, torque and speed. From this study the following conclusions can be drawn:

- In the performed tests different assemblies having the same test setup give different measured torque loss. The spread in torque loss due to assembly methodology has been quantified. In these tests the smallest difference in torque spread is ±0.10 Nm and the largest difference is ±0.35 Nm (within a torque loss of 3.6 Nm – 7.5 Nm) when running at a load torque of 94 Nm between 87-3479 RPM in an FZG gear test rig.
- The overall uncertainty of measurements in temperature, torque and speed has been quantified in the gear test rig used. The measured uncertainty for the torque loss is smaller than the scatter from the different assemblies.
- Unloading and loading does not affect the torque loss level, but the spread is slightly increased.
- Variations in oil level are detrimental to torque loss level at higher speeds, and should be controlled for accurate results at those speeds.
- Pre-heating of the gear test rig increases the spread and level in torque loss at low speeds. **PTE**

References

1. Petry-Johnson, T. T., N.E. Anderson, D.R. Chase, and A. Kahraman. "An Experimental Investigation of Spur Gear Efficiency," *J Mech Des., American Society of Mechanical Engineers*. 130 (6):62601, 2008.
2. Martins, R. C., N.F.R. Cardoso, H. Bock, A. Igartua and J.H.O. Seabra. Power Loss Performance of High-Pressure Nitrided Steel Gears," Tribol. Int. Elsevier, **42** (11):1807–15, 2009.
3. Åkerblom, M. "Gearbox Noise: Correlation with Transmission Error and Influence of Bearing Preload," KTH, Royal Institute of Technology, 2008.
4. Oswald, F.B., D.P. Townsend, M. J. Valco, R.H. Spencer and R.J. Drago. "Influence of Gear Design Parameters on Gearbox-Radiated Noise," 1994, NASA TM, 106511.
5. Doleschel, A., K. Michaelis and B.R. Höhn. "Method to determine the frictional behavior of lubricants using a FZG gear test rig," Technical Report Research Project No.345, FVA, Mar 2002.
6. ELFA. *www.elfa.se.*, 2014.
7. Dizdar, S., Andersson, S. "Influence of Pre-Formed Boundary Layers on Wear Transition in Sliding Lubricated Contacts, *Wear*, Elsevier, 213 (1):117–22, 1997.

Martin Andersson is a PhD student at the department of machine design at the KTH Royal Institute of Technology in Stockholm, with an M.Sc. in materials engineering earned in 2011 at Uppsala University, Sweden. Andersson's current research topic is gearbox efficiency, with particular focus on load-independent energy losses.

Mario Sosa is a PhD student at the department of machine design at the KTH Royal Institute of Technology in Stockholm, with an M. Sc. in 2012 in machine design from the KTH Royal Institute of Technology, Sweden. Sosa's current research topic is running-in of gears.

Sören Sjöberg received his PhD in machine design in 2014 at Stockholm's KTH Royal Institute of Technology before joining the University of Gävle as a lecturer. His area of interest is gear tribology, with the focus on gear efficiency. Sjöberg has more than twenty years of experience in both mechanical design and production engineering, in that time working for ABB, Sandvik and Volvo.

Ulf Olofsson, a professor in tribology at the department of machine design at the KTH Royal Institute of Technology (Stockholm), received his PhD in 1996 based upon on a thesis on wear as failure mechanism in boundary lubricated rolling bearings. His main research interests include interfaces and especially simulation and prediction of friction and wear — mainly applied to problems in mechanical, automotive and railway engineering. Olofsson has published more than 100 articles in scientific journals.