Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P1838-cd

Measurement uncertainty and remaining useful life prediction: A case study using testing data from shape memory alloy wires

Alicia Auer

Chair of Reliability Engineering and Risk Analytics, University of Wuppertal, Germany.
E-mail: aauer@uni-wuppertal.de

Stefan Bracke

Chair of Reliability Engineering and Risk Analytics, University of Wuppertal, Germany.
E-mail: bracke@uni-wuppertal.de

In a globalized and interconnected world, the importance of reliability, maintenance, and quality continues to grow. At the same time, the prediction of the remaining useful life (RUL) is essential for maintenance planning, facing Prognostics and Health Management of technically complex products. Given the importance of maintenance measures for a reliable product life cycle, it is all the more important to achieve the most accurate prediction results possible. A relevant influencing factor here is the measurement uncertainty of the database used.
This paper presents a case study on the impact of measurement uncertainty on the RUL prediction. The study employs data from tests of cyclically stressed shape memory alloy wires. Real data from long-term life tests conducted in a test rig are given. Firstly, the RUL of the shape memory alloy wires is predicted using linear regression models. The training data is fitted with a Gaussian least-squared regression model, and the RUL is estimated using forecasts generated by this model and an adaptive y-target value for the failure time.
The second step is a comprehensive measurement uncertainty analysis, which determines and quantifies all relevant uncertainty components of the measurement process. The extended measurement uncertainty is determined according to ISO 22514-7 and VDA 5.
Thirdly, Monte Carlo simulations are conducted based on the original time series and the determined measurement uncertainty. Representative time series are generated, for which each the RUL is predicted. Subsequently, descriptive statistics are applied to the obtained set of simulated RULs, with the results compared to the original time series and the true RUL. The paper concludes with a discussion of the results and an outlook on future work.

Keywords: remaining useful life, measurement uncertainty, ISO 22514-7, VDA 5, linear regression, Monte Carlo simulation, shape memory alloys.

1. Introduction

In a globalized and interconnected world, the importance of reliability, maintenance, and quality continues to grow. At the same time, the prediction of the remaining useful life (RUL) is essential for maintenance planning, facing Prognostics and Health Management of technically complex products. Given the importance of maintenance measures for a reliable product life cycle, it is all the more important to achieve the most accurate prediction results possible. A relevant influencing factor here is the measurement uncertainty of the database used.

This paper presents a case study on the impact of measurement uncertainty on RUL prediction. The study employs data from tests of cyclically stressed shape memory alloy wires. Real data from long-term life tests conducted in a test rig are given. Three main steps are carried out. Firstly, the RUL is predicted using a linear regression approach and a determined adaptive failure threshold. As second, there is a comprehensive measurement uncertainty analysis according to ISO 22514-7 (International Organization for Standardization 2021) and VDA 5 (VDA 2021) of the test rig used as example of the case study. Lastly, the impact of measurement uncertainty on the predicted RUL is exemplarily determined on the basis of Monte Carlo simulations (Zio 2013).

2. Methodology

This chapter delineates the methodological approaches utilized in the present study. Section 2.1 elucidates the prediction of the remaining useful life (RUL) employing linear regression models. Section 2.2 addresses the measurement uncertainty analysis, and 2.3 describes the simulation concept used the present case study for determining the impact of measurement uncertainties on the RUL prediction.

2.1. *Remaining useful life prediction*

In this paper, the RUL is predicted based on a linear regression model. Regression analysis is a statistical method that can be used to establish a functional relationship between two or more variables. In this paper, the course of a variable y over the time variable x is modeled using a linear regression approach. The estimation of the model parameters is conducted using the Gaussian least-squares method. A linear regression model is fitted with gradient a and intercept b, cf. Eq. (1). For the use case, boundary conditions for a and b are set. (Bracke 2024)

$$y = ax + b;\;\; a < 0, b > 0 \qquad (1)$$

2.2. M*easurement uncertainty analysis*

The approach presented for determining measurement uncertainties is based on the ISO 22514-7 (International Organization for Standardization 2021) and VDA 5 (VDA 2021) standards. The objective of this approach is to systematically analyze and calculate combined and extended measurement uncertainties to ensure a precise and traceable measurement process.

1. Measurement task and measurement process description
Definition of the measurement task and specification of the measurement process, including relevant boundary conditions.

2. Analysis of the uncertainty components
Identification of all relevant uncertainty components of the measurement process. One part of these are the uncertainty components of the measurement system.

3. Definition of the determination method
Selection of suitable methods for determining the standard uncertainties:

- Error limits: calculation based on given information, e.g. data sheets, considering various distribution models,
- Standard deviation: Determination based on measurement series with at least 20 values,
- ANOVA: application of the analysis of variance for uncertainty estimation.

4. Calculation of the standard uncertainties
Quantification of the standard uncertainties for the measurement process (MP) with the selected determination method.

5. Formation of the combined uncertainty
Quadratic addition of the determined standard uncertainties to form the combined uncertainty of the measurement process u_{MP} (see Eq. (2)).

$$u_{MP} = \sqrt{\begin{array}{c} u_{CAL}^2 + max\{u_{RE}^2, u_{EVR}^2, u_{EVo}^2\} \\ +u_{LIN}^2 + u_{BI}^2 + u_{MS_{Other},i}^2 \\ +u_{OBJ}^2 + u_{GV}^2 + u_{AV}^2 + u_{STAB}^2 \\ + \sum_i u_{IA,i}^2 + u_T^2 + {u_{MP}}_{Other,i}^2 \end{array}} \qquad (2)$$

6. Calculation of the expanded measurement uncertainty
Derivation of the expanded uncertainty U_{MP} on the basis of the combined standard uncertainty u_{MP} using expansion factor k in dependence of the used scatter range, see Eq. (3). In this case study, k is selected as 2, corresponding to a scatter range with a coverage of 95.45%, assuming a normal distribution model.

$$U_{MP} = k \times u_{MP} \qquad (3)$$

2.3. *Simulation concept*

This chapter outlines the simulation concept utilized in the case study to determine the impact of measurement uncertainties on the RUL prediction.

1. Specification of the measurement result
The measurement result y is specified using Eq. (4) with the extended uncertainty U_{MP} and measured value x.

$$y = x \pm U_{MP} \qquad (4)$$

Given the relationship between the scattering range (here: 95.45%) and the quantile of the normal distribution (here: ± 2s), Eq. (5) is defined with standard deviation *s*.

$$x - 2s \leq y \leq x + 2s \tag{5}$$

2. Transformation
It can be deduced from Eq. (4) and (5) that U_{MP} is equivalent to 2s. By rearranging the equations, Eq. (6) is derived:

$$s = \frac{U_{MP}}{2} \tag{6}$$

3. Simulation of time series using Monte Carlo simulation (MCS)
A data pool with *M* data points per measured value is simulated for a time series with *n* measured values x_i ($i = 1,\ldots,n$). This Monte Carlo simulation is based on the assumption of a normal distribution model, whereby standard deviation *s* and mean $\bar{x}$ according to Eq. (6) resp. (7) are used. It is assumed that the measurement uncertainty is constant over the time.

$$\bar{x} = x_i \tag{7}$$

The selection of the number *M* of simulated data points per measured value *x* is based on the coverage *p*, as outlined in ISO/IEC GUIDE 98-3/Suppl.1 (International Organization for Standardization 2008) and Guimaraes Couto et al. (2013). The condition according to Eq. (8) applies in this context, ensuring a minimum number of data points to achieve the desired coverage probability.

$$M > \frac{10^4}{1 - p} \tag{8}$$

For a specified $p = 0.9$, a data point count of $M = 100{,}000$ is selected, thereby achieving an optimal balance between coverage, computational effort, and the available computing resources in the presented case study. This step completes in the generation of a set of *M* simulated time series of length *n*, which are expanded by the original time series.

4. Determination of the remaining useful life (RUL)
The RUL is determined for each simulated time series using linear regression with a adaptive y-target value for the failure.

5. Analysis of the RUL results
The *M* obtained RUL values are analyzed using descriptive statistics, and patterns are derived to evaluate the impact of the measurement uncertainty on the RUL prediction.

3. Use Case

This chapter outlines the experimental setup and the data sources used in the case study. Section 3.1 provides a description of the test rig used for life testing of cyclically stressed SMA wires. Section 3.2 presents the data sources, with Section 3.2.1 focusing on the dataset for the RUL prediction and Section 3.2.2 describing the data used for the measurement uncertainty analysis.

3.1. *Test rig*

This case study is conducted using real data from life testing of cyclically stressed linear actuators made of shape memory alloys (SMA). Multivariate cycle data from long-term tests carried out experimentally in a test rig are given. *Fig. 1* presents a photograph of the test rig. In the test rig, three SMA wires are parallel fatigue tested. The wires are subjected to cyclic current feed and cooling phases until failure. Further details regarding the test rig construction, functionality, and the collected data can be found in Theren et al. (2022).

3.2. *Data base*

This section provides an overview of the data sources utilized for the study. Section 3.2.1 describes the dataset derived from twelve tests for the RUL prediction, including measured variables and their processing. Section 3.2.2 details the additional data sources used for the measurement uncertainty analysis, including sensor specifications, and a custom experimental setup for repeatability and reproducibility studies.

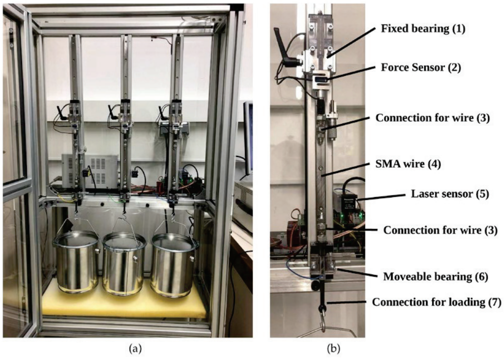

Fig. 1. Test rig for fatigue tests of three shape memory alloys (SMA) as used as application example in this study. (a) complete test rig, (b) detailed view of one test channel. (Theren et al. 2022)

3.2.1. *Remaining useful life prediction*

The RUL prediction is conducted using a series of twelve tests, each involving three SMA wires and executed under identical testing conditions. This results in a total of approximately 350,000 cycles with measured variables including temperature [°C], voltage [V], current [A], force [N], and distance [mm]. The present study utilizes a summarized dataset, which compiles the minimum, maximum, and delta of the aforementioned measured variables for each test cycle. The primary focus of the analyses is on the main output variable that exhibits a high correlation to runtime, namely the minimum distance per cycle. In order to ensure robust and comparable results, the dataset is divided into three distinct sets: training (70%), testing (20%), and validation (10%) wires. Each complete time series of a wire is randomly assigned to one group. The training wires are used to develop forecasting models, with each time series split into 70% training data (model fitting) and 30% testing data (verification). The testing wires evaluate model transferability using a similar split within their time series. The results presented herein focus on the test wires, while the validation wires are reserved for future analyses that extend beyond the scope of this study.

3.2.2. *Measurement uncertainty analysis*

The uncertainty analysis is based on three primary data sources, selected depending on the specific uncertainty component under consideration. Firstly, specifications obtained from the datasheet of the used laser sensor (Panasonic HG-C1050) is utilized. The standard uncertainty u can be determined from the available information according to the following formula, under the assumption of a rectangular distribution with a scatter range of $\pm a$:

$$u = \frac{2a}{\sqrt{12}} = \frac{a}{\sqrt{3}} \qquad (9)$$

Secondly, complete time series of training and test wires from the RUL prediction database, with selected data points (e.g., repeated measurements after the run-in phase), are employed.

Thirdly, a custom experimental series designed to assess repeatability and reproducibility precision is used. The experiment is grounded in the %GRR method (AIAG 2010). The test objects, SMA wires ($n = 3$), are evaluated across the three test channels ($k = 3$) with repeated measurements ($r = 10$). The configuration of the setup ensures the requirement set in Eq. (10).

$$n \times k \times r \geq 30 \qquad (10)$$

Each test channel is equipped with one wire, and the procedure is as follows:

1. Initiate the test, including a single run-in phase of about. 3,000 cycles.
2. Perform r measurement cycles.
3. Terminate the test and remove the wires.
4. Reinsert the wires, swapping test channels.

This sequence of steps is repeated until each wire has been tested in every channel.

4. Case Study

This chapter deals with the application of the methodology presented using the example of SMA wires. First, the approach and results of the RUL prediction are described in Section 4.1. This is followed in Section 4.2 by an analysis of the measurement uncertainty that arises when collecting the input data for the RUL prediction. Finally, Section 4.3 examines the impact of this measurement uncertainty on the prediction of the RUL.

4.1. *Remaining useful life prediction*

In order to make an accurate RUL prediction using linear regression as a statistical prognosis approach, it is necessary to determine a y-target value corresponding to the failure time. The database's primary output variable, the minimum distance per cycle, displays a distinct pattern: an exponential decrease during the initial running-in characteristic, followed by an almost linear decline, and a sharp drop at failure. A notable observation is the nearly constant ratio between the distance at the end of the running-in characteristic and the distance just before failure across all testing wires. *Fig. 2* illustrates this relationship, highlighting the two critical points in red, with their quotient displayed as horizontal lines.

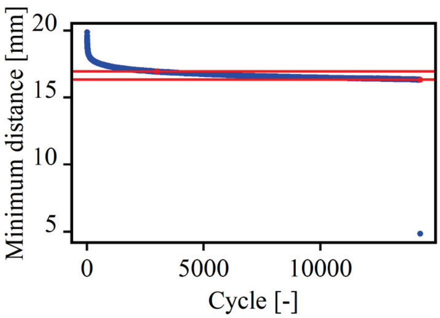

Fig. 2. Determination of the failure threshold: Visualization of the distance at the end of the running-in characteristic (first red point) and the distance before failure (second red point) with their quotient (horizontal red lines).

Fig. 3 presents a regression analysis of these distance values for the training wires, revealing a strong linear correlation with a Pearson correlation coefficient of r = 0.99713. Utilizing the average quotient of these distances, the y-target value for failure prediction is adaptively determined for each time series based on its minimum distance at the end of the running-in characteristic. The RUL is calculated as the difference between the predicted failure cycle and cycle of the end of the training data. Further details regarding this determination and application of the failure threshold as well as the following boundary conditions for the forecast can be found in Auer and Bracke (2025).

As RUL prediction approach, a linear regression model is selected as a straightforward yet effective method. As shown in Auer and Bracke (2025), in the specified use case, this approach outperforms more complex models. However, the fundamental concept can also be implemented with other RUL prediction methods.

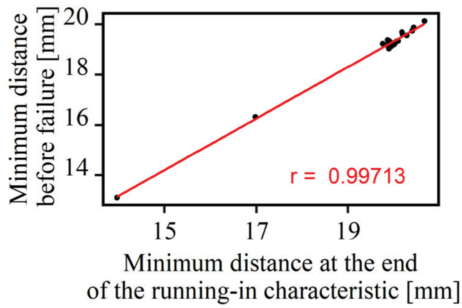

Fig. 3. Determination of the failure threshold: Regression analysis of distance at the end of the running-in characteristics with distance before failure with correlation coefficient according to Pearson's *r*.

The linear regression model for the RUL prediction is fitted using exclusively the training data from the end of the running-in characteristic. The failure time is predicted by equating the regression function to the calculated y-target value (last distance value) and implementing mathematical transformations.

Fig. 4 shows a representative RUL prediction result using this approach: The training data is represented in blue, with the subset utilized for fitting the model highlighted in light blue. The test data is displayed in green. The black smooth curve represents the prognosis model, while the dashed line indicates the forecast. The predicted failure time is marked by a red point.

The linear regression model accurately captures the linear range of the training data. However, as the time series progresses, the model underestimates the test data, resulting in an earlier predicted failure time, in this particular example, the discrepancy amounts to approximately 220 cycles.

With regard to the other testing wires, the mean absolute error between the real and predicted RUL is 1,701 cycles. In 71.43% of the cases, the RUL is underestimated. This underestimation of the RUL is advantageous for a practical application, as it enables the implementation of maintenance measures prior to the failure of the product.

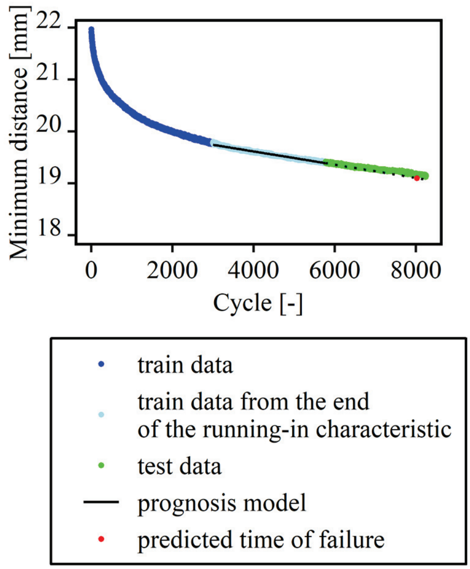

Fig. 4. Representative RUL prediction result using linear regression.

4.2. *Measurement uncertainty analysis*

Analyzing the application case, eight uncertainty factors were identified. Four regarding the measurement system:

- resolution u_{RE},
- repeatability (at reference) u_{EVR},
- non-linearity u_{LIN} and
- other influences regarding measurement system u_{MS_Other}

as well as four more regarding the measurement process:

- repeatability (at test object) u_{EVO},
- inhomogeneity of test object u_{OBJ},
- comparability of gages u_{GV} and
- comparability of times u_{STAB}.

In the following, the quantification of the standard uncertainties of all identified uncertainty factors is described before the combined standard uncertainties and the expanded measurement uncertainty are determined. The measurement system's uncertainties were determined using a conservative approach with the assumption of a rectangular distribution. The individual contributions were considered as follows: u_{RE} was derived from the scale division of the system, which is 0.001 mm. Half of the scale division was used as a measure of the uncertainty. u_{EVR} is based on the repeatability of the device, which is specified in the data sheet of the distance sensor as 30 µm. This information was included directly in the calculation of the uncertainty. For u_{LIN}, the value specified in the data sheet of ±0.1% of the measuring range (full scale) was used, with the measuring range set at ±15 mm to account for the linearity deviation. Additionally, the temperature dependence of the sensor was identified as an uncertainty factor (u_{MS_Other}). According to the data sheet, the temperature dependence of the sensor is 0.03% of the measuring range per degree Celsius. The measured temperature range from 20.28 °C to 21.75 °C was considered.

u_{OBJ} was determined using the test series for RUL prediction. To this end, 20 repeat measurements were used following the end of the running-in characteristic. For each channel, a single wire was designated as the representative of all wires within that channel, with the objective of mitigating the introduction of additional uncertainty factors. The maximum of the determined standard deviations was adopted as u_{OBJ}. u_{STAB} was also determined using the test series for RUL prediction. Due to the degradation of the wire, the first three measured values after the running-in characteristics were used for each wire channel. In this case, the highest standard deviation was used as a measure of u_{STAB}.

The determination of the uncertainty values, u_{GV} and u_{EVO}, was achieved through the implementation of an analysis of variance (ANOVA) of the measurement data obtained from the custom experimental series, cf. Sec. 3.2.2. A comprehensive analysis of the variances within the repetitions and between the test channels was conducted to quantify the repeatability and reproducibility of the system.

The squared summation of all determined standard uncertainties, when root extracted, results in a combined standard uncertainty of u_{MP} = 0.79716 mm. Assuming a normal distribution model with $P = 0.9545$, the expanded measurement uncertainty results in U_{MP} = 1.59432 mm.

4.3. *Impact of measurement uncertainty to the predicted remaining useful life*

In this case study, the simulation concept described in Chapter 2.3 is applied to one test wire as an example in order to obtain initial findings on the impact of measurement uncertainties on the RUL prediction. *Fig. 5* shows the results of the RUL prediction for all M = 100,000 simulated time series as a box plot. The real RUL and the RUL predicted from the original data set are also shown for comparison.

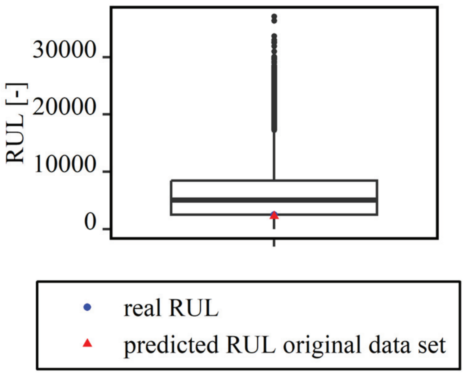

Fig. 5. Box plot of predicted RULs from simulated time series in comparison to real RUL and predicted RUL form original data set.

The majority of the simulated RUL values falls within the range of 0 and 10,000 cycles, as evidenced by the interquartile range in the box plot. However, there are a few outliers with notably higher values that exceed 30,000 cycles. It is noteworthy that both the real RUL (represented by the blue dot) and the RUL predicted using the original data set (illustrated by the red triangle) are significantly below the median of predicted RULs of the simulated time series. The distribution of the simulated RULs exhibits a right skew, which can be attributed to the inherent lower limit of the RUL set at zero.

To obtain a more nuanced understanding of the distribution of the simulated RULs, these are categorized in a histogram, as illustrated in *Fig. 6*. The real RUL and the RUL predicted from the original data set are once again presented for comparison.

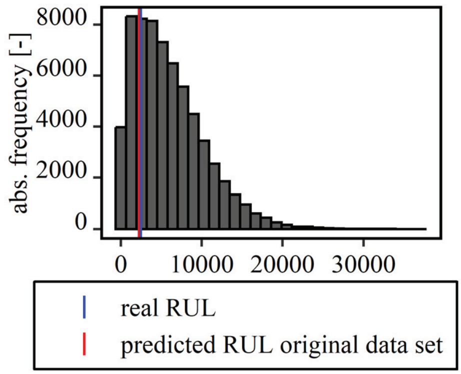

Fig. 6. Histogram of predicted RULs from simulated time series in comparison to real RUL and predicted RUL form original data set.

The hypothesis that there is a right-skewed distribution of the RULs is confirmed by the histogram. It is noteworthy that again both the real RUL and the RUL derived from the original data set are positioned to the left of the distribution curve.

Finally, the quantiles of the simulated results are calculated to depict a scatter range of the RUL based on the impact of the measurement uncertainty. These quantiles are calculated non-parametrically according to Hyndman and Fan (1996) and are documented in Table 1. All essential parameters of the simulation are also listed there.

Table 1. Overview of parameters of the simulation and key results.

Parameter	Value
U_{MP}	1.59432 mm
Number M of simulated data points	100,000
Real RUL	2,472
Predicted RUL original data set	2,252
0.01 quantile simulated RULs	110
0.05 quantile simulated RULs	522
0.1 quantile simulated RULs	1,019
0.9 quantile simulated RULs	11,896
0.95 quantile simulated RULs	14,222
0.99 quantile simulated RULs	18,953

The analysis of the simulated RULs reveals a substantial variation in the resulting values when compared to the two reference values, the real RUL (2,472) and the predicted RUL of the original data set (2,252). The underlying cause of this dispersion can be attributed to the significant expanded measurement uncertainty, particularly in relation to the value range of the degradation (see *Fig. 2*).

5. Summary and conclusion

In this paper, a simulation-based framework for analyzing the impact of measurement uncertainty on the prediction of remaining useful life (RUL) is presented and exemplarily applied in a case study with shape memory alloy (SMA) wires under cyclic stress. The study consists of three key steps:

1. *Measurement uncertainty analysis*

The measurement uncertainty of the a SMA test rig is assessed using ISO 22514-7 and VDA 5 standards, leading to the calculation of combined and expanded uncertainties, providing a replicable template for similar studies.

2. *RUL prediction*

Linear regression as a straightforward yet effective method is applied to predict the RUL based on the minimum distance per cycle, with an adaptive failure threshold.

3. *Simulation of impact of measurement uncertainty on the RUL prediction*

Monte Carlo simulations are performed to assess how measurement uncertainty influence the RUL prediction, showing a wide range of predicted RUL values of the simulated time series.

The results indicate that measurement uncertainty significantly impacts the RUL prediction. The study highlights the importance of minimizing the measurement uncertainty to improve the accuracy and reliability of RUL predictions, ultimately leading to the optimization of maintenance strategies.

Future research will focus on optimizing the experimental setup to achieve more practical and accurate results. Additional simulations with diverse test wires and load scenarios are planned to enhance robustness and transferability. Furthermore, development efforts will include adapting distribution models and utilizing functional relationships to better approximate RUL quantiles, enabling a more comprehensive assessment of the impact of measurement uncertainty on the RUL prediction.

Acknowledgement

The authors would like to thank Philipp Heß for providing the information regarding the test rig, for preprocessing the data and for giving important SMA wire specific information for the discussion of the results of this paper The authors thank both Philipp Heß and Marcel Schmidt for conducting the test series used in this paper.

References

AIAG. 2010. *Measurement Systems Analysis.* 4th ed.

Auer, Alicia, and Stefan Bracke. 2025. "Comparative Study on Remaining useful life Prediction Using Testing Data." *2025 Annual Reliability and Maintainability Symposium (RAMS), Miramar Beach, FL, USA, 2025*. paper in print.

Bracke, Stefan. 2024. *Reliability Engineering: Data Analytics, Modeling, Risk Prediction.* Berlin, Heidelberg: Springer.

Guimaraes Couto, Paulo Roberto, Jailton Carreteiro, and Sergio Pinheiro de Oliveir. 2013. "Monte Carlo Simulations Applied to Uncertainty in Measurement." In *Theory and Applications of Monte Carlo Simulations*, edited by Victor Chan.

Hyndman, Rob J., and Yanan Fan. 1996. "Sample Quantiles in Statistical Packages." *The American Statistician* 50 (4): 361.

International Organization for Standardization. 2008. *Uncertainty of Measurement: Guide to the Expression of Uncertainty in Measurement (GUM:1995)*, no. ISO/IEC GUIDE 98-3:2008.

International Organization for Standardization. 2021. *Statistical Methods in Process Management — Capability and Performance: Part 7: Capability of Measurement Processes*, no. ISO 22514-7:2021.

Theren, Benedict, Philipp Heß, Stefan Bracke, and Bernd Kuhlenkötter. 2022. "The Development and Verification of a Simulation Model of Shape-Memory Alloy Wires for Strain Prediction." *Crystals* 12 (8): 1121.

VDA. 2021. *Band 5 Mess- Und Prüfprozesse: Eignung, Planung Und Management.* 3., überarbeitete Auflage: Qualitäts Management Center im Verband der Automobilindustrie.

Zio, Enrico. 2013. *The Monte Carlo Simulation Method for System Reliability and Risk Analysis.* 1st ed. London: Springer.