Available online at www.sciencedirect.com

ScienceDirect

Procedia Structural Integrity 75 (2025) 65–71

Structural Integrity Procedia

www.elsevier.com/locate/procedia

Fatigue Design 2025 (FatDes 2025)

Fatigue Assessment of Welded Joints using Extreme Value Statistics and Laser-Based Weld Toe Geometry Analysis

Marike Schwickardi[a], Mathis Harder[a], Finn Renken[b], Moritz Braun[a,b]*

[a]German Aerospace Center (DLR), Institute of Maritime Energy Systems, Düneberger Str. 108, D-21502 Geesthacht, Germany

[b]Hamburg University of Technology, Institute of Ship Structural Design and Analysis, Am Schwarzenberg-Campus 4C, D-21073 Hamburg, Germany

Abstract

This paper presents a novel approach for fatigue assessment of welded joints by employing extreme value statistics to predict the most critical weld toe geometry in welded structures. The methodology integrates advanced laser line sensor technology for precise weld geometry measurements, allowing the identification of geometrical features that are most likely to contribute to fatigue failure. By applying extreme value statistical techniques, the approach emphasizes the criticality of localized weld toe imperfections, providing a more accurate prediction of fatigue life compared to traditional methods. This technique offers improved reliability in fatigue life estimation for large welded structures, particularly in high-stress applications, and enhances the safety and durability of engineering designs.

Peer-review under the responsibility of Dr Fabien Lefebvre with at least 2 reviewers per paper

Keywords: Non-Destructive Testing; Weld Quality Assessment; Weld Imperfections; Laser Line Sensors; Geometrical Feature Extraction

1. Introduction

The challenge of ensuring fatigue resistance in welded joints—and the inherent variability of their fatigue behavior—is closely tied to the localized nature of fatigue. In welded structures, fatigue cracks often initiate at sites of local stress concentrations within the weld metal or heat-affected zone, such as weld transition notches, residual

* Corresponding author. Tel.: +49-4152-8488147.
E-mail address: moritz.braun@dlr.de

Peer-review under the responsibility of Dr Fabien Lefebvre with at least 2 reviewers per paper
10.1016/j.prostr.2025.11.008

stress concentrations, and regions with locally weak microstructural units. Often, geometric and metallurgical discontinuities introduced during the welding process—like undercuts, porosity, lack of fusion, and other imperfections—act as stress raisers that trigger the initiation of fatigue cracks. Thus, ensuring high fatigue strength of welded joints is inherently linked to ensuring a high weld quality with as few imperfections as possible. Thus, weld qualities are closely monitored during fabrication.

Despite the significant progress in optical measurement techniques, weld quality measurements are still manually performed. Traditional manual measurement approaches are time-consuming and susceptible to human error, which limits their reliability and consistency (Hammersberg and Olsson 2010). In comparison, automated surface digitization technologies provide a more efficient alternative for evaluating weld profiles (Jung et al. 2024; Renken et al. 2025). Modern non-destructive optical measurement systems, such as 3D scanners, combined with advances in computing power and data storage, have opened up new ways for assessing weld quality. These tools also enable data-driven analysis to link weld seam geometry with predicted fatigue performance, c.f., (Braun and Kellner 2022; Hultgren et al. 2023; Rohani Raftar et al. 2024).

On the other hand, extensive measurements would be required to detect severe imperfections reliable in large welded structures. Thus, methods for estimating the occurrence of rare or extreme imperfections is essential. For non-welded components—such as cast or additively manufactured components—*Extreme Value Analysis (EVA)* has been successful applied for more than two decades after its introduction by Murakami (1994).

By sampling polished cross-sections and applying block maxima techniques, he estimated the maximum inclusion or defect size. The key insight is that fatigue failure occurs not based on the average defect size, but when the largest flaw in the structure exceeds the critical size defined by the fatigue threshold at a given stress level. This reinforces that fatigue quality in welded joints cannot be judged by average geometrical parameters or imperfection metrics alone, but must account for the extreme cases that govern structural reliability.

The block maxima method has certain limitations in practical applications, thus the *Peak-Over-Threshold (POT)* method has become popular in recent years. Herein, the distribution of data points that exceed a predefined high threshold is modelled. The method focuses on the tail behaviour of the underlying distribution. By fitting these exceedances to distribution functions such as the Generalized Pareto Distribution, the POT method enables reliable estimation of the probability and magnitude of rare observations.

This study proposes an innovative approach to fatigue assessment of welded joints by employing the POT method to identify the most critical weld toe geometries within welded structures. High-resolution laser line sensor technology is used for weld geometry measurements, which enables us to detect fatigue-critical locations along weld seams. In addition, the method allows us to estimate worst-case scenarios in structures, where weld scanning cannot be performed for all welded joints. As a result, the proposed technique enhances the reliability of fatigue evaluations for large-scale welded structures.

2. Data Basics

The presented method was applied in an exploratory manner to 23 butt joint samples made of *A36 structural steel*, joined using the *Submerged Arc Welding (SAW)* process. Weld seam geometry was captured using a laser profile scan along each seam and subsequently processed through a python script—based on the curvature method (Renken et al. 2021; Renken et al. 2024)—to extract relevant features. The resulting dataset contains a wide range of geometrical and welding-related parameters.

The following parameters were extracted for each sample:

- Radius 1 / 2 [mm], Angle 1 / 2 [°], Undercut 1 / 2 [mm]
  (captured on both sides of the weld seam)
- Weld Width [mm], Maximum Weld Reinforcement [mm]
- Global Weld Angle [°], Angular Misalignment [°]
- Linear Misalignment [mm], Z Position [mm], Thickness [mm]

- Status and Error Codes, Iteration Number
- SCF values (Stress Concentration Factors) for in-plane bending and tension loading *on both sides of the weld seam*

The samples were cut from one large steel plate, which was welded in a continuous submerged arc welding process. Due to inherent process variability, the measured geometric weld parameters exhibit natural fluctuations, enabling a robust statistical evaluation.

To validate the applicability of extreme value theory methods, the independence of data points was checked by evaluating the spatial autocorrelation of the parameters. It was confirmed that, for the majority of parameters, the spatial autocorrelation decreases rapidly. In principle, ensuring a random spacing larger than the critical correlation length would be necessary to guarantee statistical independence.

However, fully resampling the data at a fixed spatial interval to enforce independence would result in the loss of a significant amount of data, particularly for parameters with slow spatial variation such as *Weld Width* or *Reinforcement Height*. For parameters characterized by more localized fluctuations—such as *Undercut Depth* or *Stress Concentration Factors*—the autocorrelation naturally decays more rapidly, making them better suited for classical POT analysis without aggressive down sampling.

3. Methodology

This study applies an *Extreme Value Analysis* based on the *Peak-Over-Threshold* approach to characterize rare but critical deviations in geometric weld parameters, which might act as fatigue crack starters. This is indicated in Figure 1. The analysis pipeline consists of four main steps: threshold selection, excess value extraction, normalization, and statistical evaluation. The POT method is well established in extreme value theory and allows for the statistical modelling of tail behaviour beyond a chosen high threshold (Davison and Smith 1990; Coles 2001).

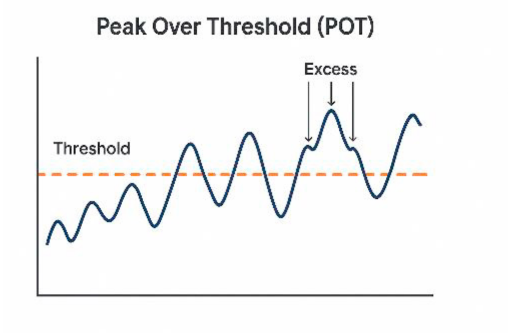

Fig 1: Illustration of the POT method. Data above the threshold (dashed line) are extracted as excess values for modelling the tail behaviour.

3.1 Threshold Selection and Excess Values

To isolate the extreme values, the 95th percentile threshold was calculated individually for each sample based on the selected weld parameter (e.g. weld width). All data points exceeding this threshold were classified as excess values, representing the upper tail of the distribution relevant for reliability and structural assessment:

$$\boldsymbol{x_{excess} = x - threshold} \quad (1)$$

Only excess values $x_{excess} > 0$ were used in the subsequent analysis.

In principle, the same methodology can be applied to detect lower-tail extremes by using a lower quantile threshold (e.g., the 5th percentile) and inverting the excess direction. This approach was also applied to parameters such as Radius 1 and Radius 2, where particularly small values are considered critical.

3.2 Normalization

To ensure comparability across samples with different absolute scales, all values were standardized using the classical *z*-transformation prior to thresholding. Each sample was normalized by subtracting the sample-specific mean μ and dividing by the standard deviation σ. The same transformation was applied to the threshold, resulting in dimensionless excess values that are independent of scale and location. This allows for consistent comparison of extreme value behavior across different weld seams.

$$x_{excess,norm} = \frac{(x - threshold)}{\sigma} \tag{2}$$

This form retains the original *exceedance above threshold*, but scales it relative to the sample's standard deviation. The result is a set of standardized excess values that are *independent of scale and location*, and can be meaningfully compared across different samples.

3.3 Distribution Fitting

The normalized excess values from all samples were pooled and fitted against a selection of theoretical distributions that are either *commonly used in extreme value statistics* or serve as *reference models* for comparison. The goal of this step is to identify which statistical models best describe the behaviour of the extracted extremes and to provide a robust basis for interpretation and further analysis.

The following statistical distributions were selected to model the excess values in the POT framework. A detailed overview of the properties and applications of these distributions can be found in Kotz and Nadarajah (2000).

- *Generalized Pareto distribution (GPD)*:
  The GPD is the canonical distribution for modelling exceedances in the POT framework. It is highly flexible and can represent heavy-tailed, light-tailed, and exponential-type behaviour depending on its shape parameter. It serves as a theoretical reference model in EVA.
- *Gumbel distribution*:
  A special case within the family of extreme value distributions, the Gumbel distribution is often used to model the distribution of maxima in block maxima methods but can also approximate upper tails in POT analysis when the underlying data exhibits light tails.
- *Weibull distribution*:
  This distribution is suitable for modelling *bounded* extreme values and is often used in reliability engineering and material fatigue. Its inclusion allows the detection of distributions where an upper bound might exist for weld seam features.
- *Exponential distribution*:
  The exponential distribution represents the simplest case of the GPD (with shape parameter ξ=0). Its inclusion allows assessment of whether a memoryless, purely scale-based model can sufficiently explain the extreme values.
- *Log-normal distribution*:
  While not part of classical EVA theory, the log-normal distribution is included due to its relevance in modelling skewed, multiplicative physical processes. It is often encountered in real-world measurement data, including geometry and material parameters.
- *Normal distribution*:
  Included as a benchmark, the normal distribution provides a reference for symmetric, non-heavy-tailed behaviour. While not expected to perform well for extreme values, its comparison highlights the departure of the data from Gaussian assumptions.

All distributions were fitted using *maximum likelihood estimation* as implemented in `scipy.stats`.

This comparative approach allows us to determine not only whether classical EVA models like the GPD are suitable, but also whether simpler or alternative distributions might better capture the observed behaviour of weld seam extremes.

3.4 Validation and Evaluation

To assess the goodness-of-fit of each candidate distribution, a combination of *statistical tests* and *information criteria* was used. The primary method applied was the *Kolmogorov–Smirnov (KS) test*, which compares the empirical cumulative distribution function (CDF) of the data to the theoretical CDF of the fitted model (Stephens 1974).

However, the *KS test* has well-known limitations, particularly when applied to distributions of *extreme values*. The test is uniformly sensitive across the entire data range, but it does *not place sufficient emphasis on the tails*, where the differences between extreme value models are most pronounced. Furthermore, since the distributions were fitted to the same data used in the test, the resulting p-values may be biased due to overfitting.

To address these limitations, additional methods were considered:

- The *Anderson–Darling test* was used as a tail-sensitive alternative. In contrast to the KS test, it applies greater weight to the tails of the distribution, making it more suitable for POT-based analyses (Stephens 1974).
- The *Akaike Information Criterion (AIC)* was calculated for each fitted distribution. As a likelihood-based model selection criterion, it provides a trade-off between model complexity and fit quality, independent of the specific shape of the tail (Akaike 1974).

The Kolmogorov–Smirnov (KS) test was chosen as a primary evaluation method due to its widespread use and ease of implementation, particularly in early-stage analyses. Future work will involve the generation and evaluation of additional data, which will allow for a broader assessment including tail-sensitive and information-theoretic approaches to improve the robustness of the analysis.

4. Results

Due to the large number of parameters and distributions considered in this study, the procedure is exemplified using the parameter *SCF Tension 1* to illustrate the core methodology and typical outcomes of the analysis.

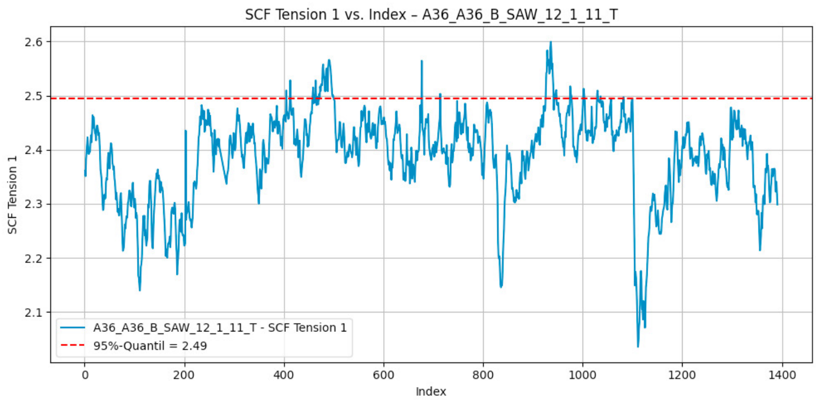

Fig 2: SCF Tension across a representative sample. The 95th percentile threshold is shown as a horizontal line. Excess values used for the Peak-Over-Threshold analysis are highlighted. This plot illustrates the identification and isolation of extreme values from the measurement data.

Figure 2 shows the behaviour of the *SCF Tension* across one representative sample of 40 mm length, which is the size of fatigue test specimens taken from a welded plate. The selected threshold (95th percentile) is marked as a horizontal line, and all values exceeding this threshold are identified as excess values. These form the basis for the POT analysis. This figure illustrates the initial filtering process and highlights the deviation of the extreme values from the bulk of the data. The visualization demonstrates how local irregularities in the weld geometry are systematically

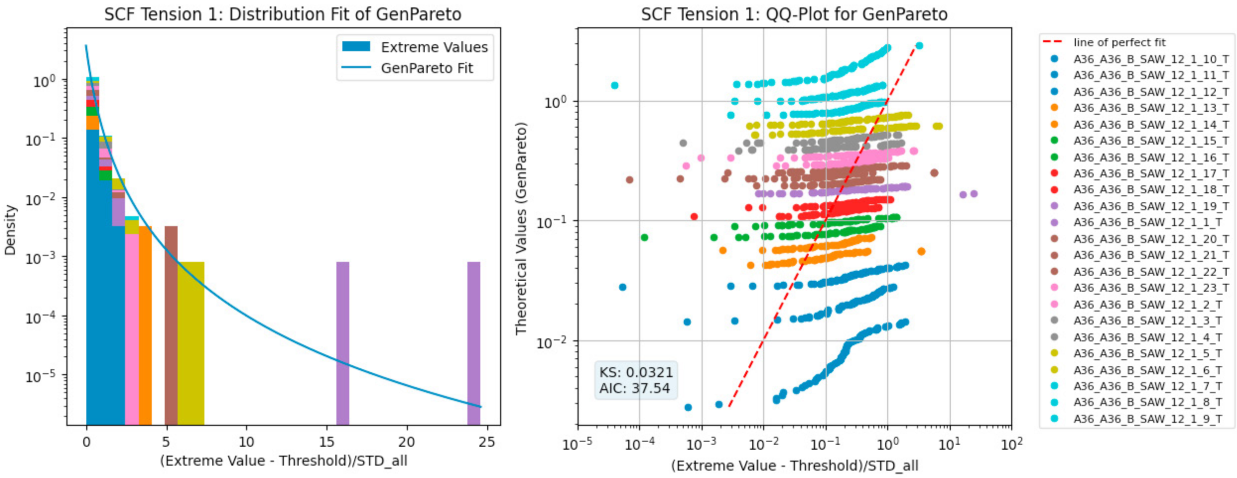

Fig 3 Distribution fit and Q–Q plot for the normalized excess values of SCF Tension. Left: Histogram overlaid with the fitted Gumbel distribution. Middle: Quantile–Quantile plot showing agreement between empirical and theoretical quantiles. The figure demonstrates the application of distribution modelling and visual goodness-of-fit assessment. Right: Legend with all used samples.

isolated for further statistical modeling.

Figure 3 presents the analysis of the parameter *SCF Tension* (Stress Concentration Factor) using the fitted Gumbel distribution as an example. The left panel shows a histogram of the normalized excess values overlaid with the probability density function (PDF) of the fitted distribution. The right panel displays a *Q–Q plot*, comparing empirical quantiles to the theoretical quantiles derived from the fit. The alignment along the line of perfect fit in the Q–Q plot suggests a reasonable agreement between the data and the model. This figure illustrates the application of distribution fitting and visual goodness-of-fit evaluation.

These plots exemplify the workflow applied to all analysed parameters: thresholding, extraction of excess values, normalization, distribution fitting, and evaluation through visual and statistical criteria. Having shown that EVA is a suitable tool to describe the distribution of critical weld geometry parameters, in a next step this theory can be applied to estimate worst-case geometrical locations at weld transitions of large-scale structures, where weld scanning cannot be performed for all welded joints.

5. Conclusions

The presented methodology demonstrates that Extreme Value Analysis using the Peak-Over-Threshold approach offers a powerful framework to detect and model rare but critical deviations in weld geometry. The standardized workflow allows for consistent evaluation across different weld samples and parameters.

However, to further improve the reliability of the analysis, the statistical evaluation of distribution fits should be refined. In particular, goodness-of-fit testing using the Kolmogorov–Smirnov and Anderson–Darling tests requires careful interpretation, especially in the tail regions.

In addition, the assumption of independence among data points must be explicitly validated. If dependencies—such as spatial or sequential correlations—are present in the data, the standard POT method may yield biased results. In such cases, more advanced approaches from time-dependent or clustered extreme value theory may be required.

Moreover, it remains to be evaluated whether the POT approach—designed to capture global extremes—is optimal for all parameters. An alternative strategy using the block maxima method could be explored in future work to better identify local extremes within individual weld segments.

6. Acknowledgements

The work was performed within the research project WeldScanPro – "Reduzierung der Fertigungs- und Wartungskosten von Offshore Windenergieanlangen mittels lnline-lnspektion und maschineller Lernverfahren" funded by the German Federal Ministry for Economic Affairs and Climate Action (project number 03SX595B).

Supported by:

Federal Ministry
for Economic Affairs
and Climate Action

on the basis of a decision
by the German Bundestag

References

Akaike H. 1974. A new look at the statistical model identification. *IEEE transactions on automatic control*. 19(6), 716-723

Braun M, Kellner L. 2022. Comparison of machine learning and stress concentration factors-based fatigue failure prediction in small-scale butt-welded joints. *Fatigue Fract Eng M*. 45(11), 3403-3417. https://doi.org/10.1111/ffe.13800

Coles S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Vol. 208. Springer. (Springer Series in Statistics). https://doi.org/10.1007/978-1-4471-3675-0

Davison AC, Smith RL. 1990. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 52(3), 393-425

Hammersberg P, Olsson H. 2010. Statistical evaluation of welding quality in production. Swedish Conference on Light Weight Optimized Welded Structures; March 24-25; Borlänge, Sweden.

Hultgren G, Mansour R, Barsoum Z. 2023. Fatigue strength assessment of welded joints incorporating the variability in local weld geometry using a probabilistic framework. *International Journal of Fatigue*. 167. https://doi.org/10.1016/j.ijfatigue.2022.107364

Jung M, Braun M, Schubnell J, Remes H. 2024. Round robin study on the determination of weld geometry parameters - Part A: Analysis of a reference specimen. *Welding in the World*. 69(1), 169-176. https://doi.org/10.1007/s40194-024-01829-y

Kotz S, Nadarajah S. 2000. *Extreme value distributions: theory and applications*. world scientific.

Murakami Y. 1994. Inclusion rating by statistics of extreme values and its application to fatigue strength prediction and quality control of materials. *Journal of Research of the National Institute of Standards and Technology*. 99(4), 345

Renken F, Jung M, Ehlers S, Braun M. 2024. A filter calibration method for laser-scanned weld toe geometries. *Applications in Engineering Science*. 20. https://doi.org/10.1016/j.apples.2024.100200

Renken F, Schubnell J, Jung M, Braun M, Remes H. 2025. Round-robin study on the determination of weld geometry parameters—part B: analysis of welded specimen. *Welding in the World*. https://doi.org/10.1007/s40194-025-02033-2

Renken F, von Bock und Polach RUF, Schubnell J, Jung M, Oswald M, Rother K, Ehlers S, Braun M. 2021. An algorithm for statistical evaluation of weld toe geometries using laser triangulation. *International Journal of Fatigue*. 149. https://doi.org/10.1016/j.ijfatigue.2021.106293

Rohani Raftar H, Ghanadi M, Hultgren G, Ahola A, Barsoum Z, Björk T. 2024. Assessing local stresses in scanned fillet weld geometry using bagged decision trees. *Journal of Constructional Steel Research*. 218. https://doi.org/10.1016/j.jcsr.2024.108745

Stephens MA. 1974. EDF statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*. 69(347), 730-737