

# Towards proper sampling and statistical analysis of defects

Ali Cetin<sup>1,a</sup>, Andrew Roiko<sup>2</sup> and Lind Mina<sup>3</sup>

<sup>1</sup>Subsea AS, Norway

<sup>2</sup>VTT, Finland

<sup>3</sup>Aalto University, Finland

**Abstract.** Advancements in applied statistics with great relevance to defect sampling and analysis are presented. Three main issues are considered; (i) proper handling of multiple defect types, (ii) relating sample data originating from polished inspection surfaces (2D) to finite material volumes (3D), and (iii) application of advanced extreme value theory in statistical analysis of block maximum data. Original and rigorous, but practical mathematical solutions are presented. Finally, these methods are applied to make prediction regarding defect sizes in a steel alloy containing multiple defect types.

## 1. Introduction

Statistical methods for defect characterization are widely studied, and numerous papers are published on the topic including some standards [1]. Due to practical limitations in sampling capabilities, current efforts are mainly aimed at block maximum methods with sample data originating from polished inspection areas. Many comprehensive and complicated mathematical models are available for the analysis of these sample data. Unfortunately, the simplest of these methods fall short of transparently modelling the actual physical phenomena, often leading to confusion when the model predictions are incorrect.

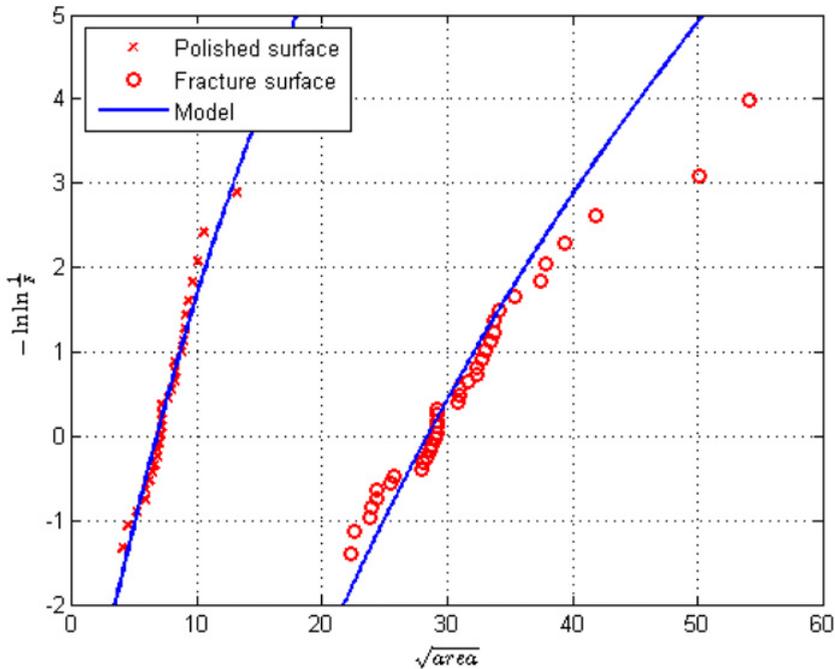
The present work aims at presenting rigorous mathematical and statistical models that are practical and simple to apply. Three main issues are considered; (i) proper handling of multiple defect types, (ii) relating sample data originating from polished inspection surfaces (2D) to finite material volumes (3D), and (iii) application of advanced extreme value theory in statistical analysis of block maximum data.

## 2. Method and results

A steel alloy (bearing steel) with multiple defect types has been utilized as an initial verification of the model performances. Defects have been sampled from polished surfaces by means of automated optical microscope inspection. Finally, fatigue tests were performed to sample defects at the fracture surface. A key observation was that the largest defect in polished surfaces was often MnS, while the defects at the fracture surface after fatigue testing was mostly  $\text{Al}_2\text{O}_3$ .

---

<sup>a</sup> Corresponding author: [ali.cetin@subsea.com](mailto:ali.cetin@subsea.com)



**Figure 1.** Probability plots of the  $\text{Al}_2\text{O}_3$  defects observed on the polished surfaces and the fracture surface.

The applied methods may be summarized as:

- Identify each defect type and handle them separately.
- Determine the representative inspection volume for the inspection area. A generalization to the Murakami method is proposed.
- Convert the 2D data to equivalent 3D data. A novel approach is proposed.
- Fit the converted 3D data to a new non-classical extreme value distribution.

The above method was applied to the  $\text{Al}_2\text{O}_3$  defects sampled from the polished surfaces. The defect sizes expected to occur on the fracture surface after the fatigue tests were predicted and compared with the actual observations. The result is presented in Fig. 1.

The presented method is capable of characterizing the defects expected to occur in a finite volume based on sample data obtained from polished surfaces. Note that the data from the polished surfaces and the fracture surfaces do not align on parallel lines despite coming from the same population. However, the presented model is able to capture this non-linear feature.

## Reference

- [1] ASTM E 2283 – 03: *Standard Practice for Extreme Value Analysis of Nonmetallic Inclusions in Steel and Other Microstructural Features*, ASTM, 2003