Probability of detection applied to X-ray inspection using numerical simulations

M. Yosifov\textsuperscript{a,b}, M. Reiter\textsuperscript{c}, S. Heupl\textsuperscript{a}, C. Gusenbauer\textsuperscript{c}, B. Fröhler\textsuperscript{b}, R. Fernández-Gutiérrez\textsuperscript{b}, J. De Beenhouwer\textsuperscript{b}, J. Sijbers\textsuperscript{b}, J. Kastner\textsuperscript{d} and C. Heinzl\textsuperscript{d}

\textsuperscript{a}Research Group X-ray Computed Tomography, University of Applied Sciences Upper Austria, Wels, Austria; \textsuperscript{b}Lab, Department of Physics, University of Antwerp, Antwerpen, Belgium; \textsuperscript{c}Department of Process Engineering, Nemak Linz GmbH, Linz, Austria

\textbf{ABSTRACT}

In this work, we apply and adapt established probability of detection (POD) methods on the in-line inspection of aluminium cylinder heads using X-ray computed tomography (XCT). We propose to use the XCT simulation tool SimCT to simulate virtual X-ray radiographs from the specimen including artificial defects, which avoids the manufacturing of specimens with calibrated defects of known type (e.g. pores, inclusions, cracks) and characteristics (e.g. size, shape, location). To quantify the POD, these virtual images are analysed using ZEISS automated defect detection (ZADD) to determine defects automatically. ZADD is a deep learning application for anomaly defect detection, classification and segmentation. To create respective POD curves, we apply a hit/miss approach. We demonstrate our method on artificial defects of different sizes, location and material types. Eight representative defects are discussed in detail together with the generated POD curves as well as their characteristics. We finally discuss the advantages of numerical simulations with respect to the probability of detection in order to quantify and improve detection limits.

\textbf{1. Introduction}

Visual analysis and investigations on three-dimensional structures play an important role in understanding the behaviour of materials and defects in terms of material characterisation and quality control. For Non-Destructive Testing (NDT), X-ray computed tomography (XCT) is considered as one of the most important true volumetric imaging methods [2]. Nevertheless, XCT requires tailored data processing and visualisation pipelines to extract the structures of interest [3]. In recent years, machine learning (ML) has demonstrated remarkable results in various domains and especially for defect detection in XCT data, e.g. in-line and at-line applications [4]. A new generation of high-speed XCT scanners such as the ZEISS VoluMax 1500 G2 allows users to detect and segment defects, as well as to determine and visualise defect properties in three dimensions. However, there is still a demand for improved measurement capabilities of detection algorithms and devices.
One of the methods to determine the defect detection capacity is the concept of probability of detection (POD) [5–7]. The POD concept is currently gaining interest in the field of X-ray-based digital radiography (XDR) and XCT, as there is a strong trend in industry towards fast in-line and at-line inspections for integrated quality control as well as for 100% inspections [8]. Typically, such in-line or at-line inspections require much shorter scanning and evaluation times than to conventional lab-based inspections, in order to acquire the X-ray data and the decision-making process between success or failure of detection within cycle time. However, short scanning times typically yield more noise in the acquired radiography or XCT images and therefore strongly demand the use of POD methods to characterise and optimise the detection performance of the inspection processes. For the computation of POD and respective curves, typically specimens with different features (i.e. artificial defects) are manufactured, and then scanned using a high number of XCT acquisitions [7]. The POD method provides statistical evidence regarding the detection capability of defects, by a high number of XCT scans. The artificial defects need to be characterised in terms of shape and size, while assuring that they are representative for real defects. As opposed to the conventional POD curve generation based on a high number of regular XCT acquisitions, this work focuses on the application of the POD method on virtual XCT data generated by numerical XCT simulation by modelling industrial specimens and respective XCT devices. The use of numerical XCT simulations vastly simplifies the process of applying the POD method to many slight modifications of the specimen. Overall, the integration of XCT simulation leads to a reduction of effort both in time and cost, as compared to standard POD methods.

2. Related work

For the related work, we consider the areas of XCT simulation (Subsection 2.1), probability of detection applications (Subsection 2.2), and defect detection and segmentation methods (Subsection 2.3).

2.1. XCT simulation

XCT simulation is often invoked as a preliminary step in biomedical applications [9] and material science [10], where optimised parameters are required for recurring scans of similar specimens, e.g. for in-line XCT inspections of industrial specimens. For this purpose, Fuchs et al. [11] use XCT simulation for generating a meaningful synthetic ground truth for defect detection. Generally, XCT simulation tools are rather diverse and can be classified into two types. The first type are numeric simulation tools, which include, e.g. Scorpius XLab [12] and ASTRA toolbox [13]. These numeric XCT simulation tools enable virtual radiographic testing (RT) and can assist in optimising system parameters and predicting the reliability and qualification of NDT or Nondestructive Evaluation (NDE) system. They are capable of creating radiographic images by efficient ray tracing algorithms similar to SimCT. They typically also provide a data visualisation modules. Monte Carlo XCT simulation is the second type of simulation focused on accurately simulating X-ray photon interactions. For example, the XCT simulation tools MC-GPU [14] and PEN ELOPE [15] are capable of generating
realistic radiography projection images and XCT scans. Another simulation tool in this area called GATE was introduced by Jan et al. [16] for modelling positron emission tomography (PET) and single-photon emission computed tomography (SPECT). An extended version of the GATE framework for accurate X-ray phase contrast imaging simulations using Monte Carlo methods was presented by Sanctorum et al. [17]. As Monte Carlo XCT simulation tools tend to have extensive computation times, these tools are considered as too time consuming for our work, which is also why our focus is on numerical simulation.

2.2. Probability of detection applications

The Probability of detection (POD) method has been introduced in the 1970s and was originally developed for the American military to be used on NDT and NDE data [5,7]. POD curves and their applications and limitations were presented by Georgiou in 2007 [18]. POD is a statistical measure for the performance of an inspection and its binary classification into the categories of ‘flaw found’ or ‘no flaw found’ [5]. POD represents the fraction of inspections that will detect an existing flaw. The POD largely depends on characteristics of the flaw (e.g. size, shape and location), the flaw detection method and the image quality provided by the XCT device [7]. In this work, we employ XCT simulations coupled to the computation of their POD as well as respective curves in terms of artificial defects.

2.3. Defect detection and segmentation methods

Traditional defect detection and segmentation methods, such as Otsu thresholding, region growing, watershed, K-means segmentation or template matching, are still popular with material science experts [19]. Improving X-ray images and respective methods is however essential to advance defect detection. For this reason, many ML algorithms have been developed for the segmentation and detection of objects for medicine [20] and material science [21] applications. In the current work, ZEISS automated defect detection (ZADD) [22] was used for segmentation and defect detection. ZADD uses deep neural networks for this purpose. The workflow of the used serial inspection process for ZADD was described by Reiter et al. [22] and consists of three steps. In the first step, a ground truth is created from several good parts scanned with XCT, which is referred to as the reference part. To obtain a usable reference part, all parts have to be scanned with the same XCT scan settings and matched with the simulation parameters. ZADD then automatically detects defects by comparing the reference part with the examined components. In the second step, the defect segmentation and classification parts is completed with determining the defect properties. In the final step, ZADD automatically decides between ‘OK’ and ‘NOT-OK’ parts.

3. Methods and investigations

The following paragraphs introduce the POD method (Subsection 3.1), the generation of virtual XCT data, our data analysis set-up (Subsection 3.2), the hit/miss model (Subsection 3.3), as well as the Akaike information criterion (Subsection 3.4).
3.1. Probability of detection curves

Figure 1 shows a one-dimensional schematic curve where the POD is a function of the flaw size $a$. If the flaw is shaped spherically, it is fully described by its radius and position. For a cylindrical shape, the orientation is required as additional parameter. In case the flaw is not representable by a simple geometric object, it will be described by volume and position. Generally, as shown in Figure 1, the estimate of the flaw size $a_{90}$ is defined as the size for which $POD(a_{90}) = 0.90$, and $a_{90/95}$ is the lower 95% confidence level of $a_{90}$. Blue dots display success (at probability 1) and failure (at probability 0). The size $a_{90/95}$ is an accepted measure for the minimum reliably detectable defect size in NDT. With other words, it characterises detection limits.

Typically, more than one model parameters are needed for POD, which further means that the POD curve should be theoretically an $N$-dimensional function, where $N$ corresponds to the number of model parameters, or the POD is only valid for specific boundary conditions. Defects such as cracks and pores often show too complex shapes to be accurately modelled and are therefore strongly simplified in NDE, e.g. to the length of a crack. This reduces the POD modelling to a regression process in one dimension, where various methods are available. Berens et al. [5] proposed the hit/miss analysis (Subsection 3.3) to determine POD curves. This work focuses mainly on the hit/miss analysis with the main application of fully automated XCT inspection and decisions between success and failure.
3.2. **XCT data generation and analysis set-up**

ZEISS VoluMax devices were used in-line and at-line with robot loading for the serial inspection of aluminium cylinder heads at Nemak Linz. The modelled device was a VoluMax 1500 G2 at-line, which allows fast data acquisition, ensured by continuous rotation of the rotary table. This XCT device with cone beam geometry uses a Comet 225 kV dual-focus X-ray source and a Perkin Elmer flat panel detector XRD 1620 AN14. The scanning parameters listed in Table 1 were selected to achieve a cycle time (part to part time) of 1 min including all robot movements to load and unload the part. The total part of specimen was bigger than the detector. Therefore, two scans were performed to cover the complete height of a part. The exposure time was 13.2 s per scan height. The reconstruction and joining of the reconstructed volumes to one volume as well as the subsequent evaluation were fully automated.

SimCT [1] as numerical XCT simulation was used for virtual XCT data generation. It is able to generate virtual XCT images based on a fine grained model in terms of scanning geometry (source object distance, source detector distance, focal spot size, etc.), XCT components (source, detector, rotary table, etc.) and their respective specifications, noise, or any kind of blur or artefacts due to beam-hardening and scattered radiation. SimCT was adapted in terms of X-ray source and detector to model the ZEISS VoluMax G2 XCT device. The specimen is defined by surface descriptions of aluminium cylinder heads using an STL model exported from CAD data. All defects were positioned in different locations (partially or full) inside the specimen. Those defect locations represent worst-cases for detecting pores, as highest penetration lengths occur during an XCT scan and thus the lowest CNR values of the whole volume can be expected at these positions. For reconstruction CERA [23] was used to compute the volumetric data sets of the simulated data. A single reconstruction took approximately 1 min 45 s. In total, the generation of all virtual data took approximately 17 days for 200 simulations and their reconstruction. The resulting reconstructed volumes have a size of $790 \times 630 \times 892$ voxels. All numerical simulations were done accounting for all relevant X-ray effects as realistically as possible, while considering uncertainties by repeated simulations with statistically selected input parameters. We have applied several physical effects, such as focal spot Gaussian blur, modulation transfer function, photon and Gaussian noise, image noise and scattering for the simulations to reach results as close as possible to the scanned specimen. In particular, specimen placement with respect to the reconstruction grid, is of importance for the detection of small flaws, since it influences the sampling of the flaw and thus the partial volume effect.

| Parameters                  | Unit  | Value   | Parameters                     | Unit | Value   |
|-----------------------------|-------|---------|--------------------------------|------|---------|
| Acceleration voltage        | kV    | 200     | Source–object-distance         | mm   | 1186.36 |
| Anode current               | µA    | 7200    | Source–detector-distance       | mm   | 1518.35 |
| Integration time            | ms    | 33      | Magnification                  |      | 1.2766  |
| Detector pixels             |       | 2048x2048| Voxel size                     | µm   | 313.32  |
| Binning                     |       | 2x2     | Focal spot size                | mm   | 1       |
| Number of projections       |       | 400     | Total cycle time               | min  | 1       |

Table 1. XCT scanning and device parameters.
3.3. Hit/miss model

The hit/miss approach [5] models the binary responses of detection, which means the output of an inspection is reduced to a binary signal (success or failure). For XCT, these binary responses can either be generated by a human inspector visually analysing XCT images, by an automatic image processing pipeline employing image segmentation algorithms, or by ML using a convolutional neural network.

Generalized Linear Models (GLM) [24] are continuous link functions with model parameters providing probabilities for the binary outcome of hit/miss scenarios. The most used models are the logit, probit, loglog and comloglog functions, which are applied for binary regression to create POD curves. The probit function is given by

$$\text{Probit}(\text{POD}(a)) = \Phi^{-1}(\text{POD}(a))$$

where $\Phi$ is the cumulative distribution function of the standard normal distribution and $\text{POD}(a)$ is the probability detection function linked with the size($a$) of defects. All functions have two horizontal asymptotes at zero and one. Typically, this is in accordance with data of XCT inspections, when the human inspector is trained or the image processing–based evaluation is designed to be robust with awareness of image noise. Nevertheless, the asymptotes need to be checked during data modelling. Generally, the use of a GLM is preferable compared to the grouping of binary data into defect size bins in order to estimate POD as the fraction found in that size range. This grouping could be done for repeated evaluations of the same defect or for defects with slightly varying size. A drawback of the grouping approach are rough size intervals and thus a lack of resolution in size and POD values.

3.4. Akaike information criterion (AIC) and linking function comparison

AIC is a criterion to balance precision and accuracy [25]. So, the selection of the linking function for each defect is made by applying the Akaike information criterion (AIC), which is a measure of model quality to find the most accurate model, that is, with smallest value of AIC. The comparison of the AIC results using the different linking function models is presented in Table 2. AIC is given by

$$\text{AIC} = 2k - 2\ln(L)$$

Suppose we have a statistical model of some data where $k$ is the number of model parameters to be estimated and $L$ is the maximum value of the likelihood function [26] for the model. The minimum value of AIC needs to be determined in order to select the proper linking function on binary data [27]. In this study, the probit linking function shows minimal values for AIC. For this reason, our POD curves were created with the probit linking function. Further linking functions, their formulae and a comparison of respective POD curves are presented in the supplemental material.
4. Processing and data characteristics

Our processing pipeline is shown in Figure 2. In the first block of the pipeline, the aluminium cylinder heads were scanned (1) by the ZEISS VoluMax 1500 G2. The specimen surface is defined by an STL exported from CAD data (3). To reduce scanning artefacts, the specimen was placed in the holder with a given tilt and shift. To compare corresponding simulations, the STL specimen and artificial defects have to be positioned exactly as in the real scan. Therefore, we determined specimen transformations by rigid registration between XCT scan and STL exported from CAD data (5). These transformations and physical effects (see Subsection 3.2) were applied to the virtual specimen in SimCT before the start of the simulation (8).

In the next step, artificial defect properties such as shape, location, radius, rotations (for cylindrical defect), density, and material type were defined for each defect (4). We planned to create 100 different sizes and locations for each defect. To do this, a rectangular distribution function was used to create 100 random sizes (e.g. the radius)

| Defect | Defect 1 | Defect 2 | Defect 3 | Defect 4 | Defect 5 | Defect 6 | Defect 7 | Defect 8 |
|--------|----------|----------|----------|----------|----------|----------|----------|----------|
| Logit AIC | 4        | 4        | 31.371   | 17.82    | 25.767   | 4        | 17.285   | 11.188   |
| Probit AIC | 4       | 4        | 30.813   | 17.732   | 25.389   | 4        | 16.039   | 11.137   |
| Loglog AIC | 4       | 4        | 30.858   | 18.185   | 25.847   | 4        | 17.475   | 11.351   |
| Comloglog AIC | 4    | 4        | 31.569   | 17.581   | 25.426   | 4        | 19.403   | 10.951   |

Figure 2. POD curve generation and data analysis pipeline.
in the range 31 µm to maximum range (see Table 3) for the artificial defects. To create 100 random positions for each defect, the artificial defects were moved randomly by ±0.5 voxels in all three dimension using the uniform distribution function. For the extracted defects, we have applied the same process to create again 100 random sizes (e.g. the volume size) in the range a and b (see Table 4). Additionally, 100 STL data sets were created for the extracted defects (see Subsection 4.2) with varying sizes. After that, 100 different geometry files were written which contain the STL data of the specimen, combined with artificial defects and extracted STL defects with their properties. We also created a geometry file, which only contains STL information of the specimen. Geometry files were used as input to simulate the data in SimCT. Before the simulation process was started, the computed transformations (tilt and shift) were applied in SimCT for all geometry files (8). Simulated scans were created of the specimen with 100 variations of differing defects (11), and 100 scans of the specimen without any defects were simulated with differences in the physical effects (10). Reconstructions were automatically computed by CERA per volume stack. Simulations were carried out in two parts as in the real scan. Two volume stacks were created, then combined as a final result for each sample (scan). Defect-free scans were created for the obtained reference part and the training part. For the second block, the anomaly detection software ZADD was trained with OK parts beforehand (12, 13). After that, the defect detection was performed using ZADD (14). In the final block, we implemented a MATLAB script that automatically reads all produced data and generates POD(a) curves (16) for probit, logit, loglog and comloglog.

4.1. Generated artificial defects

The artificial defect properties are listed in Table 3 for the defect detection and POD(a) curves. The table lists their shape, material type, class, minimum and maximum of the radius range. Six artificial defects are presented in this study: two differently shaped cylinders and spheres, and four types of material (Al, Fe, SiO₂ and pore). Defect classes were closure of the water jacket, inclusion, damage, porosity and residual sand. The minimum size was set to 31 µm on the defect radius size, the minimum defect radius size

| Shape     | Material type | Class          | Minimum range a | Maximum range b |
|-----------|---------------|----------------|-----------------|-----------------|
| Defect 1  | Cylinder      | Al             | 31 µm           | 2310 µm         |
| Defect 2  | Sphere        | Fe             | 31 µm           | 1320 µm         |
| Defect 3  | Sphere        | Air            | 31 µm           | 3140 µm         |
| Defect 4  | Sphere        | Air            | 31 µm           | 1320 µm         |
| Defect 5  | Sphere        | Air            | 31 µm           | 1320 µm         |
| Defect 6  | Sphere        | SiO₂           | 31 µm           | 4960 µm         |

Table 3. Artificial defect properties.
to 1320 µm for defect 2, 4, 5 and the maximum to 4960 µm for defect 6. The cylinder diameter was set to a constant value of four mm in our simulation, while the height was changed randomly (uniformly distributed).

4.2. *Extracted real defects*

For further POD studies with real defects, defect parts from an aluminium cylinder head were scanned with a voxel size of (25 µm)$^3$. The two biggest defects were then extracted as STL-files out of the aluminium cylinder heads. The observed defects are shown in Figure 3, and their properties in Table 4. Extracting real defects in STL format is mainly done to see the behaviour of the detection on the real type of defects and finally to compare them to artificially defined defects. In this paper, these defects are called extracted defects.

5. *Results and discussion*

In this section, we present the results of the defect detection in a comparison to the smallest detected defect, largest undetected defect, as well as $a_{90}$ and $a_{90/95}$ for each defect. Furthermore, two of the defects are presented visually in the figures below for 100 scans with their hit-and-miss points. In the final step, we show the POD curves of each defect.

5.1. *Visualisation of artificial defects*

In this part, one of the artificial defects with its variation in cylinder height (defect 1) and one of the extracted defects with its variation in volumes size (defect 7) is presented. Images of scans for defects 1 and 7 are presented while changing their size in a sorted range. Figures 4 and 5 also display the detected defects (in green), undetected defects (in black), smallest detected defect (in blue) and largest undetected defect (in orange) with their respective size. The material type of defect 1 was aluminium, while defect 7 was modelled as air in different sizes and located fully inside of the specimen. The smallest detected size was recorded at 208 µm for defect 1 and 1.75 mm$^3$ for defect 7. The largest missed defect was computed to be 188 µm for defect 1 and 2.74 mm$^3$ for defect 7. The

![Figure 3. Aluminium cylinder head with marked defects and extracted defects from an XCT scan.](image)
detection results from ZADD show that there is a perfect separation (P.S.) between hit and miss for defect detection of defect 1. For defect 7, we observed several differences (hit or miss) between minimum hit and maximum miss for the detection. This behaviour of detection plays an important role in creating the POD curves and in computing exact values for $a_{90}$ and $a_{90/95}$.

| 0.35 | 0.44 | 0.54 | 0.65 | 0.76 | 0.87 | 0.98 | 1.04 | 1.15 | 1.25 | 1.35 | 1.45 | 1.55 | 1.65 | 1.75 | 1.85 | 1.95 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 0.42 | 0.47 | 0.50 | 0.52 | 0.74 | 0.97 | 1.23 | 1.35 | 1.49 | 1.67 | 2.07 | 2.49 | 3.12 | 3.89 | 4.98 | 6.29 | 8.64 |
| 0.99 | 1.13 | 1.28 | 1.42 | 1.57 | 1.72 | 1.87 | 2.02 | 2.17 | 2.32 | 2.47 | 2.62 | 2.78 | 3.04 | 3.30 | 3.56 | 3.82 |
| 12.37 | 12.54 | 13.67 | 14.20 | 15.24 | 15.80 | 16.36 | 16.92 | 17.48 | 18.04 | 18.60 | 19.16 | 19.72 | 20.28 | 20.84 | 21.40 | 21.96 |
| 25.31 | 25.40 | 25.57 | 25.74 | 25.91 | 26.08 | 26.25 | 26.42 | 26.59 | 26.76 | 26.93 | 27.10 | 27.27 | 27.44 | 27.61 | 27.78 | 27.95 |
| 40.46 | 42.03 | 42.59 | 43.15 | 43.71 | 44.28 | 44.85 | 45.42 | 45.98 | 46.55 | 47.12 | 47.69 | 48.26 | 48.83 | 49.40 | 49.97 |
| 57.71 | 59.27 | 61.07 | 62.56 | 64.05 | 64.51 | 65.00 | 65.49 | 65.98 | 66.47 | 66.96 | 67.45 | 67.94 | 68.42 | 68.91 | 69.39 |

**Figure 4.** Slice images for artificial defect 1. Images are sorted according to cylinder height size. Detected defects are indicated in green, and undetected defects in black, while the defect size changes for the 100 virtual XCT data sets generated by SimCT. The smallest detected defect is indicated in blue and the largest undetected defect in orange.

**Figure 5.** Slice images for extracted defect 7. Images are sorted according to volume size (mm$^3$). Detected defects are indicated in green, and undetected defects in black, while the defect size changes for the 100 virtual XCT data sets generated by SimCT. The smallest detected defect is indicated in blue and the largest undetected defect in orange.
5.2. Virtual defect visualization and comparison

In this section, we show a visualisation of the eight defects using slice images (39 × 39 pixels each) extracted from the 3D volume. Statistics and estimated results for each defect are shown in Tables 5–7.

According to the results from the hit-and-miss process, defects 1, 2 and 6 have a perfect separation (P.S.) while defect 8 has an almost perfect separation (A.P.S) between success and failure of the detection (see Table 5 and Figure 6). Almost perfect separation means that there was only one miss between minimum hit and maximum miss. This means that it was not possible to create POD curves for these subjects (defect 1, 2, 6, 8). To obtain a POD curve on these subjects, one would need to repeat the process with different initial conditions, especially reducing the size range. Our results show that defect 1 (Al, cylinder) was the most detectable type of defect in this experiment (see Tables 5 and 6). This is probably because of the shape and volume size of defect 1. It features a cylindrical shape and the defined volume size range was bigger than that of other defects. POD curves for defects 3, 4, 5 and 7 were successfully created as well. We found out that the smallest detected defect (cylinder height) size was 208 µm for defect 1, while the largest undetected defect size was 745 µm for defect 4. In Figure 6, smallest detected defect size, the largest undetected defect size, and a_{90/95} are shown according to radius size for defect 3. However, as shown in Figure 6, a part of the sphere (damage) is positioned outside of the specimen. Therefore, the volume size was determined by approximate comparison of the scaled volume for the responsible defect size. Table 6 displays the volume size for each defect and their parameters. The minimum detected defect volume size was 0.135 mm³ for defect 2. This is most probably due to the material type. Defect 2 is an iron sphere. So, it might be easier to detect than other defects. According to Table 6, the detection limit for the volume size (a_{90/95}) was determined as 1.237, 4.458, 3.764 and 6.42 mm³ for defects 3, 4, 5 and 7, respectively. The detection limit for the volume size of defect 3 is in addition quite low when compared to defects 4 and 5.

Another important point becomes obvious when comparing defects 2 and 6. They have exactly same initial conditions (properties) except for the defect material type and location. The smallest detected defect radius is determined to be 319 µm for defect 2 and 580 µm for defect 5. Moreover, their detection numbers are 81 and 50, respectively. This also shows that it is easier to detect defects consisting of iron. Table 7 presents detection numbers for each defined defects out of 100 samples. This shows that defect 1 (Al), defect 2 (Fe) and defect 6 (SiO₂) were about 30% more likely to be detected than the other artificial defects. We also noticed, when the defect type is different than air or pore, the

| Table 5. Radius and cylinder height (defect 1) values. |
|----------------------------------|--------|--------|--------|--------|--------|--------|
| Defect 1                        | Defect 2 | Defect 3 | Defect 4 | Defect 5 | Defect 6 |
| Min hit size (µm)               | 208     | 319     | 1378    | 654     | 580     | 487     |
| Max miss size (µm)              | 188     | 318     | 1729    | 745     | 667     | 449     |
| a_{90} (µm)                     | P.S.    | P.S.    | 1788    | 758     | 690     | P.S.    |
| a_{90/95} (µm)                  | P.S.    | P.S.    | 2384    | 1021    | 965     | P.S.    |
Figure 6. While the first row shows 2D slices without any defects, the other rows show the minimum hit defect, the maximum missed defect and $a_{90/95}$, respectively, together with their radius (a) and volume (b).

Table 6. Volume size of defects (mm$^3$).

|            | Defect 1 | Defect 2 | Defect 3 | Defect 4 | Defect 5 | Defect 6 | Defect 7 | Defect 8 |
|------------|----------|----------|----------|----------|----------|----------|----------|----------|
| Min hit size | 2.613    | 0.135    | 0.239    | 1.171    | 0.817    | 0.483    | 1.75     | 2.48     |
| Max miss size | 2.362    | 0.134    | 0.429    | 1.732    | 1.242    | 0.379    | 2.74     | 2.68     |
| $a_{90}$    | P.S.     | P.S.     | 0.474    | 1.824    | 1.376    | P.S.     | 2.25     | A.P.S.   |
| $a_{90/95}$ | P.S.     | P.S.     | 1.237    | 4.458    | 3.764    | P.S.     | 6.42     | A.P.S.   |

Table 7. Detected defects number from 100 samples for each defined defect.

|            | Defect 1 | Defect 2 | Defect 3 | Defect 4 | Defect 5 | Defect 6 | Defect 7 | Defect 8 |
|------------|----------|----------|----------|----------|----------|----------|----------|----------|
| Total detected | 100/94   | 100/81   | 100/55   | 100/44   | 100/50   | 100/89   | 100/88   | 100/78   |

detected defect numbers were increasing in our experiment. Specifically defect 7 was detected 88 times; however, its volume size range was much larger than that of other defects.

Finally, the detection of the defect 1 and extracted defects from real specimens (7–8) seems harder than the detection of other defects (2–6). According to detection numbers, defects 1, 7 and 8 have quite high values (see Table 7); however, the smallest detected defects volume size are 2.61, 1.75 and 2.48 mm$^3$, respectively. These values are quite high compared to other defects (spherical defects).

5.3. POD curves presentation

In this section, we present the POD curves for each defect with their confidence levels at 95%. As mentioned in the previous section, there is a perfect separation between success and failure for three of the experiments (see Figures 7 and 8). Four POD curves were successfully created for defects 3, 4, 5 and 7. Furthermore, Figure 7 shows the resulting POD curve using the probit linking function for the hit-and-miss process with the detection limit
values ($a_{90/95}$) for the radius of the sphere or cylinder height. The POD($a$) model and its confidence bounds for 100 random sizes are shown in red. Figure 7 demonstrates POD curves for the six artificial defects, while Figure 8 shows the same for the two extracted defects. The resulting POD curves is presented with the use of the probit linking function for the hit-and-miss process with the detection limit values ($a_{90/95}$) for the volume size (see Figure 8). There are detected defects with defect sizes up to 80 mm$^3$ for defect 7 (see
supplemental material). However, the POD curve and its confidence level curves at x-values of between 1 and 10 mm³ already. Therefore, for a better visibility of the POD and confidence level curves, the x-axis was limited to a range of 0–20 mm³.

6. Conclusion

In this paper, POD methods were presented for analysing virtual XCT data as generated by the XCT simulation tool SimCT. We investigated uncertainty sources in terms of decision-making in success and failure. Different kinds of artificially created defect shapes, materials and locations were investigated in our experiment. Additionally, POD curves were created for two defects extracted from real specimen and their differently scaled versions.

Our analysis revealed that non-spherical defects are more difficult to detect as compared to others. We learned that even for the smallest detected non-spherical defects, respective values in terms of volume size are rather high for the extracted defects (defects 7 and 8) as well as for defect 1 compared to spherical defects (defects 2, 3, 4, 5, 6). As demonstrated in our paper, the POD method can be considered as an effective means to reliably determine the detection limit volume and radius in terms of defect detection. Furthermore, the detection capabilities of ZADD for specific types of defects have been investigated. In our findings, ZADD tends to be more sensitive to iron type defects. Figure 7 displays that especially the estimation of the confidence bounds for POD(a) should be improved for some of the analysed defects by more simulations or changing the defined range scale. Another highly important finding is, that POD curves can be used to optimise the acquisition parameters of XCT devices, to evaluate and compare the performance of different XCT devices, or to quantify capabilities of image processing algorithms by using the linking functions. Different linking functions were investigated and compared with AIC values to select the most suitable one on our data, rendering the probit function as the most suitable function for our application. The selection of the linking function also plays an important role in order to find out detection limits, especially on the confidence bounds curves. To sum up, we have demonstrated detection limits in terms of volume size depending on defect characteristic. With the knowledge of the final POD(a) shapes the upper limit for the random size picks for defects 1, 2, 6 and 8 could be lowered to maximise the efficiency of the analysis of POD. We further found out, that the efforts of a POD analysis can be significantly reduced by using numerical simulations as compared to real experiments, and the problem of manufacturing specimens with known and quantified defect sizes does not arise. Nevertheless, using numeric simulations for extracting POD curves is still a time-consuming process, especially, when low POD values are expected. Furthermore, the accuracy of PODs determined by simulations remains challenging.

For future work, extracted defects and non-spherical type defects will be the priority to discover detection limits for specific detection algorithms. Additionally, as only ZADD was used for defect detection in our current work, different detection devices or algorithms could be explored in order to compare their respective results.
We continuously improve our simulation quality to make it as realistic as possible. For example, we are planning to advance the object scattering model in SimCT using a more sophisticated scattering model. This might help to create even more realistic simulation results as well as better POD curves with their respective confidence bounds. Another potential improvement is speeding up the simulation process in SimCT. Performing more than 100 inspections would lead to better estimates of the confidence bounds, considering the POD analysis with repeated inspections as ‘ground truth’.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project has received funding from the K-Project “Photonic Sensing for Smarter Processes” financed by FFG and the governments of Upper Austria and Styria, as well as the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No. 956172 - xCTing

ORCID

M. Yosifov [http://orcid.org/0000-0002-6318-4869](http://orcid.org/0000-0002-6318-4869)
B. Fröhler [http://orcid.org/0000-0003-1271-0838](http://orcid.org/0000-0003-1271-0838)
J. De Beenhouwer [http://orcid.org/0000-0001-5253-1274](http://orcid.org/0000-0001-5253-1274)
J. Sijbers [http://orcid.org/0000-0003-4225-2487](http://orcid.org/0000-0003-4225-2487)
J. Kastner [http://orcid.org/0000-0002-3163-0916](http://orcid.org/0000-0002-3163-0916)
C. Heinzl [http://orcid.org/0000-0003-8871](http://orcid.org/0000-0003-8871)

References

[1] Reiter M, Erler M, Kuhn C, et al. SimCT: a simulation tool for X-ray imaging. In: 6th Conference on Industrial Computed Tomography (iCT 2016); 2016; Wels, Austria. [https://www.ndt.net/article/ict2016/papers/ICT2016_paper_id20.pdf](https://www.ndt.net/article/ict2016/papers/ICT2016_paper_id20.pdf)

[2] Heinzl C, Amirkhanov A, Kastner J. Processing, analysis and visualization of CT data. In: Industrial X-ray computed tomography. Vol. 52. Springer International Publishing; 2018, pp. 99–142. DOI:10.1007/978-3-319-59573-3_4

[3] Heinzl C, Stappen S. STAR: visual computing in materials science. Comput Graph Forum. 2017;36(3):647–666. DOI:10.1111/cgf.13214

[4] Hirsch B, Holom R, Gusenbauer C, et al. Automatic classification of XCT images in manufacturing. Case Stud Nondestr Test Eval. 2021;AICT-627:220–231. DOI:10.1007/978-3-030-79150-6_18

[5] Berens AP. NDE reliability data analysis. In: Aquil Ahmad, Eaton Corporation, (Retired) Leonard J. Bond and Iowa State University. Asm handbook, volume 17 - nondestructive evaluation and quality control. Vol. 17. Materials Park, Ohio 44073-0002: AMS International; 1989. pp. 689–701.

[6] Amrhein S, Rauer M, Kaloudis M. Characterization of computer tomography scanners using the probability of detection method. J Nondestr Evalu. 2014;33(4):643–650. DOI:10.1007/s10921-014-0258-4
[7] US Department Of Defense. Handbook Non-destructive evaluation system reliability assessment. Vol. MIL-HDBK-. Washington DC, United States: U.S. Department Of Defense; 1999.

[8] Kim F, Pintar A, Moylan S, et al. The influence of X-ray computed tomography acquisition parameters on image quality and probability of detection of additive manufacturing defects. J Manuf Sci Eng. 2019;11:141. DOI:10.1115/1.4044515

[9] Thompson A, McNally D, Maskery I, et al. X-Ray computed tomography and additive manufacturing in medicine: a review. Int J Metrol Qual Eng. 2017;8:17. DOI:10.1051/ijmpe/2017015

[10] Brock J, Sutton M. Materials science and X-ray techniques. Mater Today. 2008;11 (11):52–55. DOI:10.1016/S1369-7021(08)70239-6

[11] Fuchs P, Kröger T, Garbe CS. Defect detection in CT scans of cast aluminum parts: a machine vision perspective. Neurocomputing. 2021;453:85–96.

[12] Reisinger S, Kasperl S, Franz M, et al. Simulation-Based planning of optimal conditions for industrial computed tomography. In: International Symposium on Digital Industrial Radiology and Computed Tomography; 2011; Berlin, Germany. https://www.ndt.net/search/docs.php?id=11125&file=article/dir2011/papers/mo31.pdf

[13] Aarle WV, Palenstijn WJ, Cant J, et al. Fast and flexible X-ray tomography using the ASTRA toolbox. Opt Express. 2016 ;24(22):25129–25147. DOI:10.1364/OE.24.025129

[14] Badal A, Badano A. Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit. Med Phys. 2009;36 (11):4878–4880. DOI:10.1118/1.3231824

[15] Nuclear Energy Agency. PENEOPE 2018: a code system for Monte Carlo simulation of electron and photon transport. Oecd. 2019. DOI:10.1787/32da5043-en

[16] Jan S, Santin G, Strul D, et al. GATE: a simulation toolkit for PET and SPECT. Phys Med Biol. 2004 ;49(19):4543–4561. DOI:10.1088/0031-9155/49/19/007

[17] Sanctorum J, Beenhouwer JD, Sijbers J. X-Ray phase contrast simulation for grating-based interferometry using GATE. Opt Express. 2020 ;28(22):33390–33412. DOI:10.1364/ OE.392337

[18] Georgiou G. PoD curves, their derivation, applications and limitations. Insight. 2007;49 (7):409–414. DOI:10.1784/insi.2007.49.7.409

[19] Sarma R, Gupta YK. A comparative study of new and existing segmentation techniques. IOP Conf Ser Mater Sci Eng. 2021;1022(1):12–27. DOI:10.1088/1757-899x/1022/1/012027

[20] Çiçek Ö, Abdulkadir A, Lienkamp SS, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation. Med Image Comput Comput Assist Interv – MICCAI Lecture Notes Comput Sci. 2016;9901:424–432. DOI:10.1007/978-3-319-46723-8_49

[21] Yosifov M. Extraction and quantification of features in XCT datasets of fibre reinforced polymers using machine learning techniques [master’s thesis]. Umeå University, Department of Computing Science; 2020.

[22] Reiter M, Gusenbauer C, Huemer R, et al. At-Line X-ray computed tomography of serial parts optimized by numerical simulations. International Symposium on Digital Industrial Radiology and Computed Tomography (DIR2019); 2019; Fürth, Germany. https://www.dir2019.com/

[23] CERA user manual version 6.0.0; 2019. Available from: https://www.cera-systeem.nl/en/downloads.

[24] McCullagh P, Nelder J. Generalized linear models. 2nd ed. London: Chapman and Hall/CRC; 1989.

[25] Bozdogan H. Model selection and Akaike’s information criterion (AIC): the general theory and its analytical extensions. Psychometrika. 1987;52:345–370.

[26] Demidenko E. Computational aspects of probit model. Math Commun. 1997;6:233–247.

[27] Li J. Choosing the proper link function for binary data [master’s thesis]; 2014.