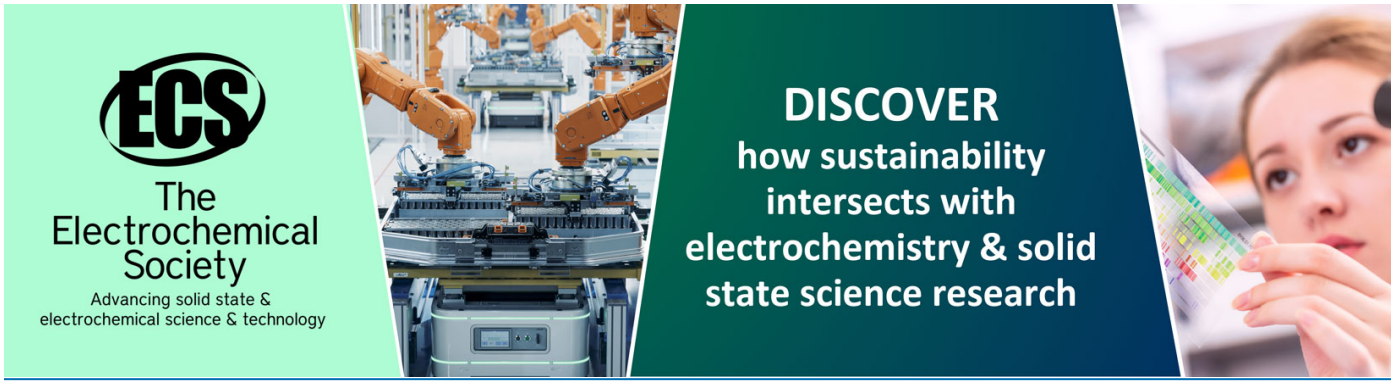
Analysis of grain size distribution evolution of steel during recrystallization and grain growth



Analysis of grain size distribution evolution of steel during recrystallization and grain growth

O Seppälä[1,*], A Pohjonen[1], M Ali[1,2], J Larkiola[1] and T Fabritius[3]
[1] Materials and Mechanical Engineering, Oulu University, Oulu, 90014, Finland
[2] Steel Technology Department, Central Metallurgical Research and Development Institute, Helwan 11421, Cairo, Egypt
[3] Process Metallurgy Research Unit, Oulu University, 90014 Oulu, Finland
[*] Corresponding author, E-mail: oskari.seppala@oulu.fi

Abstract. Controlling the hot rolling process requires a deep understanding of the underlying metallurgical phenomena. Quantitative methods are of paramount importance for achieving the capability of controlling microstructural evolution. Since the final mechanical properties of steel result from microstructural evolution in the whole process, analysis of the microstructure provides an important input for numerical simulations that can be used for tailoring the mechanical properties of steel. The evolution of grain size distribution of a low-carbon CrNiMnB ultrahigh-steel in austenitic state is studied in hot forming and annealing using experimental data obtained with the Gleeble 3800 thermo-mechanical simulator. A general method is described that can be utilized to systematically compare the grain size distributions obtained from the experimental studies. The experimental data has been obtained from laser scanning confocal microscopy images using the mean linear intercept method. A custom-made semi-automatic software has been utilized to process the data rapidly and reliably.

1. Introduction

The microstructure of steel plays an essential role when considering its mechanical properties. Understanding the evolution of the microstructure of steel becomes more important in order to apply the material in more demanding conditions. In [1], microstructure and its effect on the mechanical properties of steel was studied. Refining austenite grain size and austenite pancaking were found to improve mechanical properties for low-carbon low-alloyed TMCP-DQ martensitic-bainitic steels. In [2], the Hall-Petch relation between flow stress and grain size was discussed. Conventional and ultra-fine grain structures of Cu were studied, and dislocation structure evolution was modelled.

Many different models have been developed to simulate microstructural evolution. In [3], a finite element method (FEM) model for the thermomechanical simulator Gleeble 3800 was developed and a recrystallization model was implemented to simulate recrystallization in different parts around the test piece cross-section. In addition, various mean field models simulating the average grain size after recrystallization were investigated. In [4], a physically based static recrystallization model was implemented and fitted successfully into experimental stress relaxation test data. The results are promising, but further studies are required to improve model prediction reliability.

Validating models that simulate microstructural evolution requires high-quality experimental studies that produce diverse data depicting the microstructure as accurately as possible. One way to study microstructure is with a single mean value, like the mean linear intercept grain size [5]. In the current study, experimental data is used to construct a probability density function for linear intercept interval

length, which is capable of depicting grain size distribution and allows for comparison of statistical parameters to better understand test parameter dependencies.

2. Theory

In the current study a technique is proposed for finding a suitable continuous and differentiable function for describing the cumulative probability distribution (*CPD*) of the linear intercepts x. The probability density (*PD*) can then be obtained by differentiating the obtained function with respect to the linear intercept interval length.

2.1. Polynomial fitting

The cumulative probability density can be described by equation (1)

$$CPD(x) = 1 - \exp\big(-f(x)\big) \tag{1}$$

Since $\mathrm{d}CPD(x)/\mathrm{d}x = \mathrm{d}f(x)/\mathrm{d}x \cdot \exp\big(-f(x)\big)$, *CPD* is increasing if and only if $f(x)$ is increasing. Since CPD must be increasing or zero for every x, the derivative of $f(x)$ is always positive or zero and for this reason there exists $g(x)$ such that $\mathrm{d}f/\mathrm{d}x = g(x)^2 \geq 0$. Since $f(x) = 0$ when $x = 0$, the function is described by equation (2)

$$f(x) = \int_0^x g(x')^2 \, dx' \tag{2}$$

where x' is an integration variable. Assuming that the function $g(x)$ is sufficiently smooth, so that n:th order derivative exists for every x, it can be expressed as n:th order Taylor polynomial $g(x) = \Sigma_0^n(a_n x^n)$. Solving equation (1) for $f(x)$ and inserting it to equation (2) yields then equation (3).

$$-\ln\big(1 - CPD(x)\big) = \int_0^x (\Sigma_0^n a_n x^n)^2 \, dx' \tag{3}$$

In equation (3) the left-hand side can be calculated directly from the obtained data. An analytic expression for the right-hand side of equation (3) was obtained by raising the polynomial to the second power and integrating the terms by using computer algebra system, such as Maxima [6] or Sympy [7]. The resulting function has the coefficients a_n that were used as fitting parameters. The chosen approach has the benefit that the derivative of the function describing *CPD*(x) is guaranteed to be always non-negative.

2.2. Log-normal fitting

Another way to describe distributions is the log-normal method [8], where *CPD* is obtained with equation (4).

$$F_X(x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\ln x - \mu}{\sigma\sqrt{2}}\right)\right] \tag{4}$$

Where μ and σ are fitting parameters. A convenient way to obtain the PD function is using the form in equation (5).

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \tag{5}$$

3. Experimental work

An ultrahigh strength steel with the chemical composition given in table 1 was investigated. The material was received in the form of homogenized hot-rolled 12 mm-thick plates. Cylindrical specimen of dimensions Ø10x12 mm were machined. Interrupted stress relaxation tests were performed using Gleeble 3800 thermomechanical simulator. The samples were heated at 10 °C/s to 1250 °C and held for 120 s for homogenization, then cooled at 2 °C/s to the deformation temperature where samples were held for 15 s prior compression up to the prescribed strain. Using the stroke mode, the strain was held constant after the deformation and the compression force relaxed as a function of

holding time. The interrupted stress relaxation tests were performed using a single strain rate 10 s^{-1} in the temperature range 950-1250 °C, strain range 0.2-0.6 and holding time range 0-60 s to have different deformation parameters for evaluating the grain growth. To reveal the prior austenite grains, the polished samples were etched using picric acid solution.

Table 1. Chemical composition of the investigated steel (in wt.%).

C	Si	Mn	Cr	Ni	Al	B	N
0.16	0.2	1.0	0.5	0.5	0.03	0.0015	0.0050

4. Methods

4.1. Grain size calculator

In this study, all the grain size distributions are based on experimental data obtained from laser scanning confocal microscopy images. The images are processed with a novel self-made GUI tool for grain size calculations, called *grain size calculator* [9], see figure 1. The tool is based on the mean linear intercept method, where multiple lines are drawn on the image and each intercept with a grain boundary is marked and counted together to calculate a mean linear intercept grain size [5]. With the tool, user draws lines and marks each boundary. Then it calculates various values, most importantly for the current study the length between each consecutive linear intercept. Although the linear intercept intervals do not directly represent the grain diameter, they are closely related to the grain size. The relation between linear intercept interval and grain size is discussed in more detail in [5].

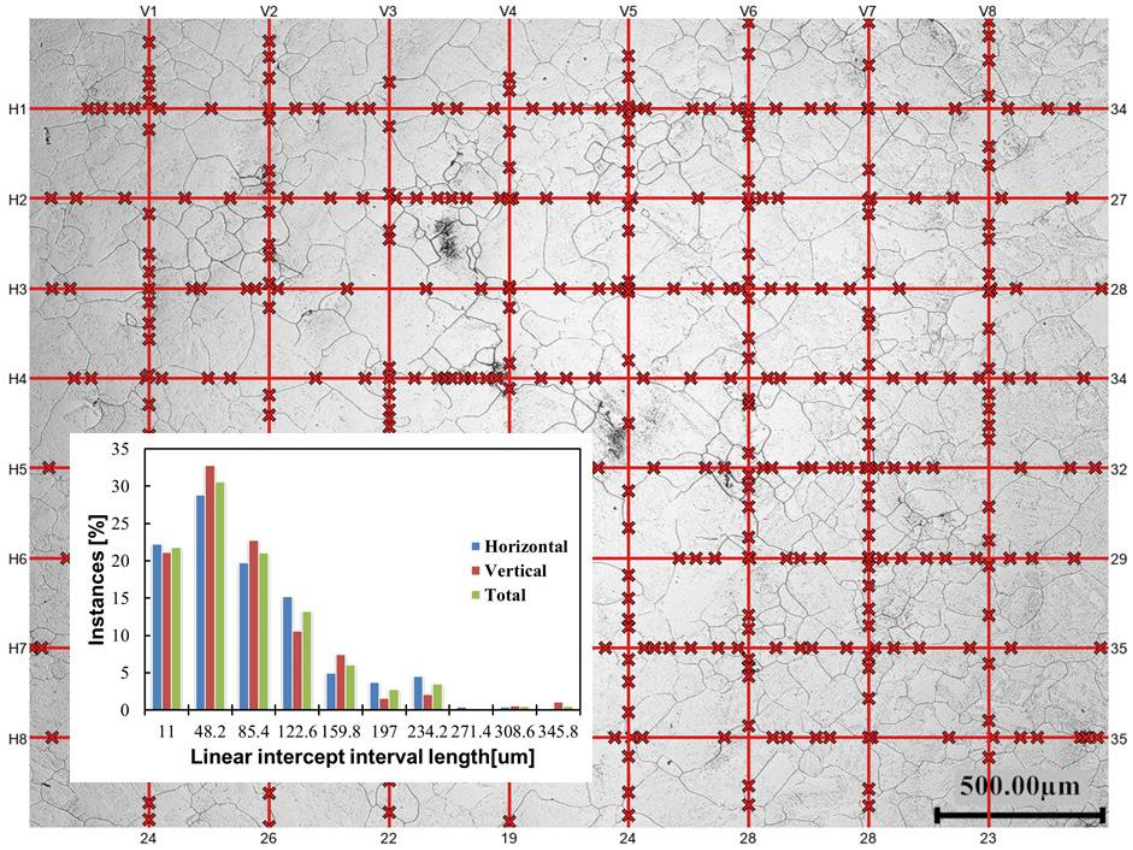


Figure 1. example grain structure image with linear intercepts marked, including the resulting histogram with horizontal, vertical and total instance percentages.

The linear intercept interval lengths are used to form a random variable data set. This can be used to create a histogram to study linear intercept interval distribution. The problem with histograms is choosing bin length, which always affects the data visualization and is arbitrarily chosen by user. A new approach was chosen in this study to avoid the problem of choosing the bin length. In the current approach the data set is converted into a continuous probability density function by applying the method described below.

4.2. Fitting technique – polynomial

First, the *CPD* is obtained from the data set by ordering the data by size, and then normalizing their cumulative count. *CPD* function is then calculated in the way explained by equation (3) and an n^{th} Taylor polynomial is fitted to the data, see figure 2-a). Next, the analytical expression of *CPD* is calculated by equation (1), see figure 2-b). Finally, the analytical equation is derivated by linear intercept interval length, which produces the probability density function in figure 2-c).

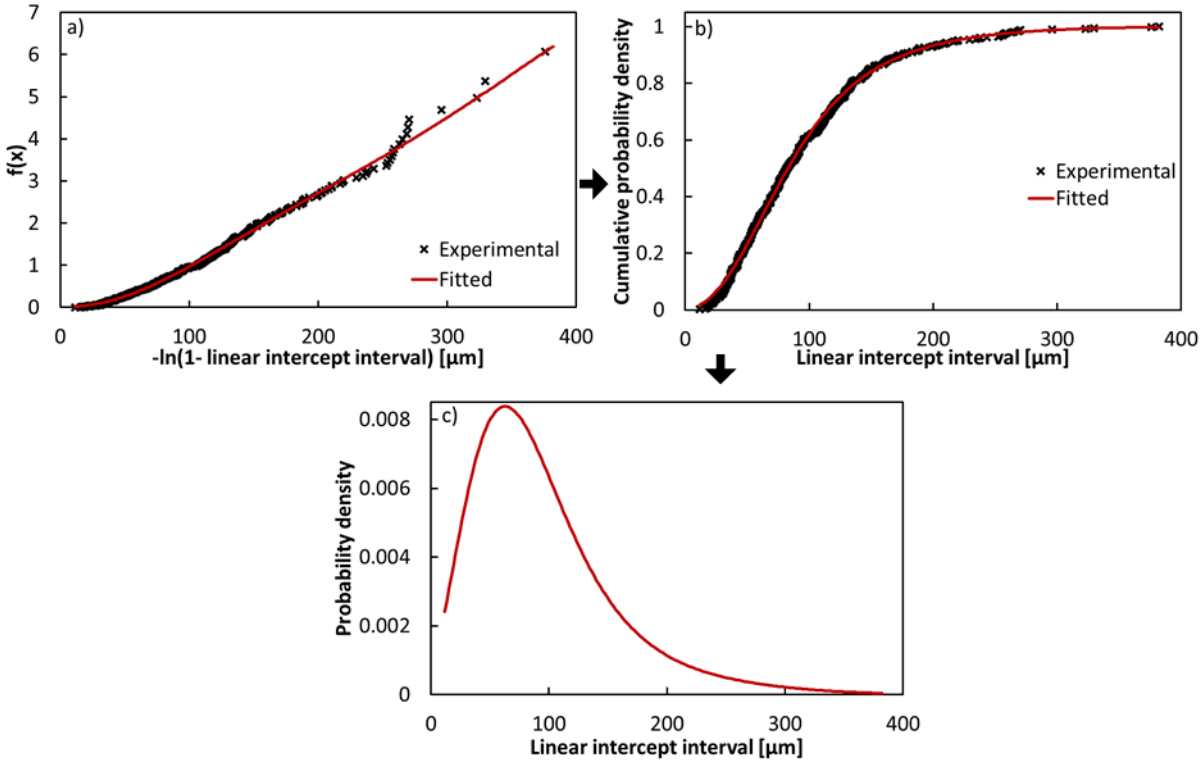


Figure 2. Method to create a probability density function for the linear intercept interval lengths with the polynomial fitting technique.

During the fitting many techniques were tested, and the following methods produced satisfactory results. Sum of difference minimization algorithm was used, using the *Nelder-Mead* -algorithm in the *minimize*-function of the Python3 *scipy*-module. Taylor polynomials of different order were tested for the function $g(x)$. They were generated with Python3 module *sympy* and individually fitted, the order of the polynomial was varied from 1 to 10. For the fitting, the initial values for parameters a_i were randomized values in closed interval [0,1]. For each a_i, the order of the initial guess was 10^{-i}, where $i \in [1,10] \subset \mathbb{N}$, to prevent too large emphasis for higher order terms. The linear intercept intervals were normalized for the same reason. The fitting was executed twice, the second fit having the first fitted values as starting values, to help the fitting algorithm in finding the minimum. Each polynomial was tested with 10 different random starting values, and the best fit was chosen according to the smallest sum of difference. Finally, the best polynomial was chosen based on this principle.

4.3. Fitting technique – Log-normal

The log-normal fitting is similar, but more straight-forward than the polynomial fitting described in section 4.2. Equation (4) is fitted with the sum of difference minimization algorithm directly into the data set CPD, see figure 3-a, and the fitting parameters μ and σ are used in equation (5) to obtain the PD function, see figure 3(b).

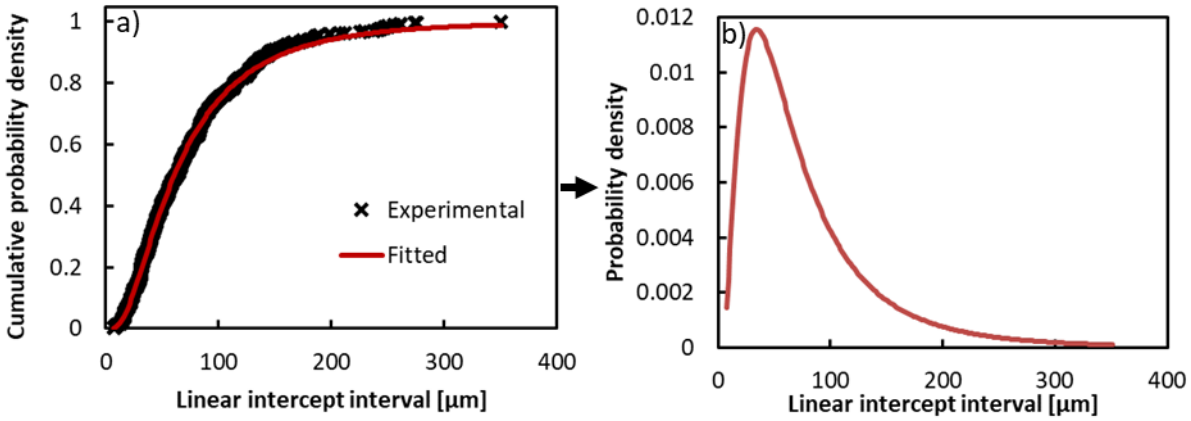


Figure 3. Method to create a probability density function for the linear intercept interval lengths with the log-normal fitting technique.

5. Results

5.1. Comparison of methods

An example case of temperature 1050 °C, strain 0.2 and holding time 0 s was chosen in figure 4 for comparing histograms and probability density functions. As can be seen, changing the bin length has a major impact on the data set -based histogram data, whereas the probability density function does not require this definition.

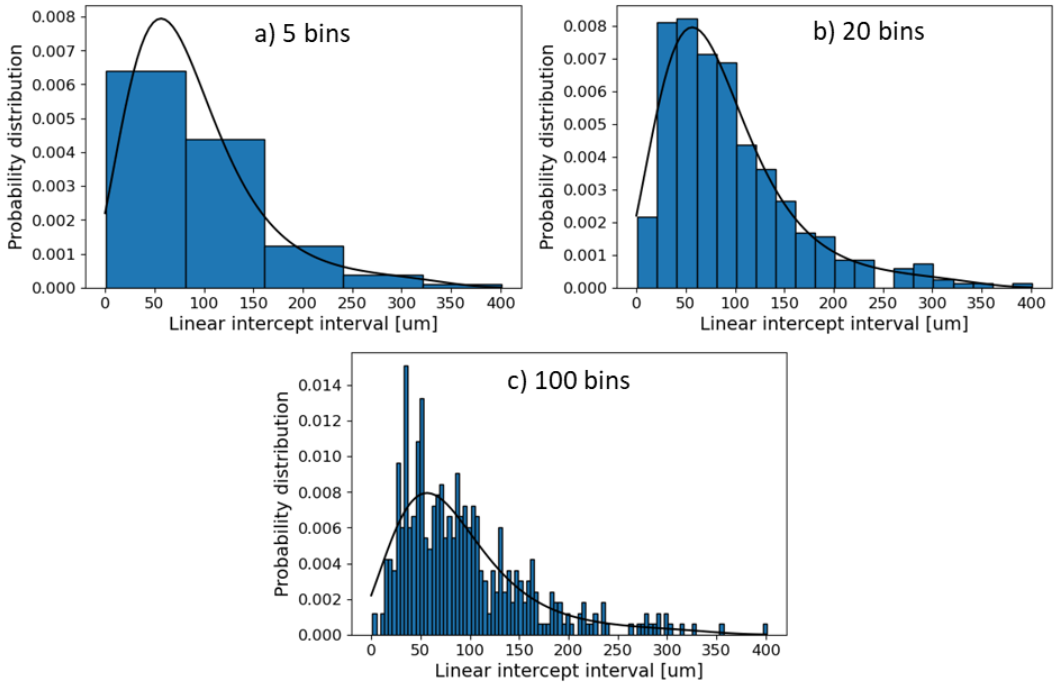


Figure 4. Comparison between a polynomial PD function and histograms with (a) 5, (b) 20 and (c) 100 bins.

The same test case (1050 °C – 0.2 – 0 s) was used to compare the polynomial and log-normal fitting methods, see figure 5. From the fitted cumulative probability densities, it can be seen that both methods are able to describe the experimental data. Log-normal fit matches the start of experimental data slightly better and polynomial fit matches the end better.

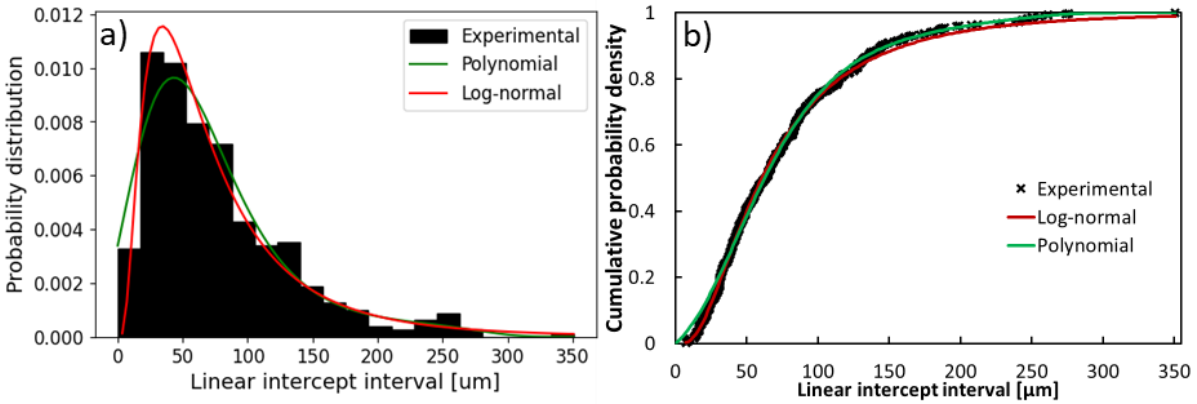


Figure 5. Comparison between 20 bin histogram data set, polynomial fitting and log-normal fitting.

The two fitting methods are not compared further in this study, but they will be thoroughly studied in the future. The rest of this study will focus on the polynomial fitting method. Multiple temperatures, strains and holding times are compared with the goal of learning more about how each variable affects grain size distribution. Another goal is to use the data as a test case for the proposed probability density polynomial function generation method.

5.2. Holding time

Figure 6 includes microstructural images for holding times 0, 30 and 60 s with temperature 1050 °C and strain 0.4. Their probability density functions, and data set histograms are collected to a single figure for comparison. Note the slightly different sized scale bar in figure 6(c). Figure 7 shows the effect of holding time for each temperature, with strain set to 0.2.

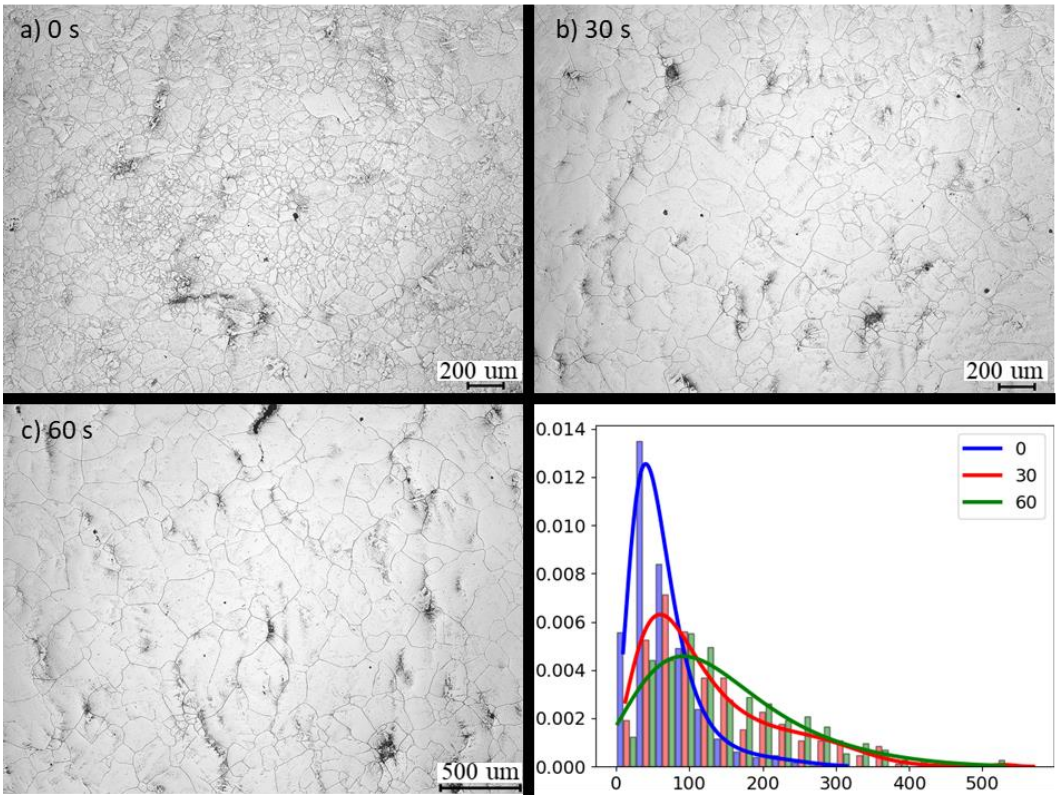


Figure 6. Microstructural images and corresponding probability density functions and data sets for temperature 1050 °C, strain 0.4 and holding times (a) 0, (b) 30 and (c) 60 s.

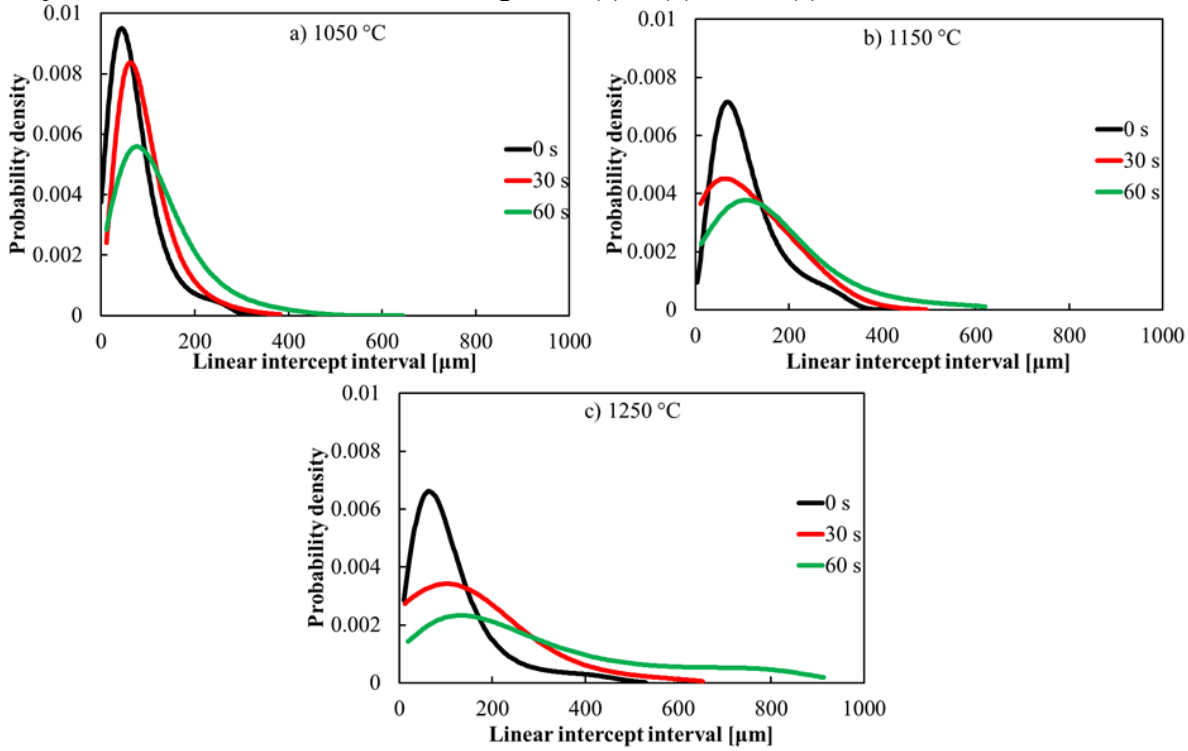


Figure 7. Distribution density functions for holding times 0, 30 and 60 s with strain 0.2 and temperatures (a) 1050, (b) 1150 and (c) 1250 °C.

As shown in figure 7, holding time has a major impact on the grain size. At 0 s there are many small grains, most likely because new nucleating grains occupy most of the test piece. 30 s and 60 s cases have a smaller likelihood for small grains, but the difference is not as distinct between the two. This is expected to be due to the fact that recrystallization has completed, and the test sample has reached the grain growth stage, where new nucleated grains have collided into each other and continue growing with a much slower pace.

5.3. Strain

Figure 8 includes microstructural images for strains 0.2, 0.4 and 0.6 with temperature 1050 °C and holding time 0 s. Their probability density functions, and data set histograms are collected to a single figure for comparison. Figure 9 shows the effect of strain for each holding time, with temperature set to 1050 °C.

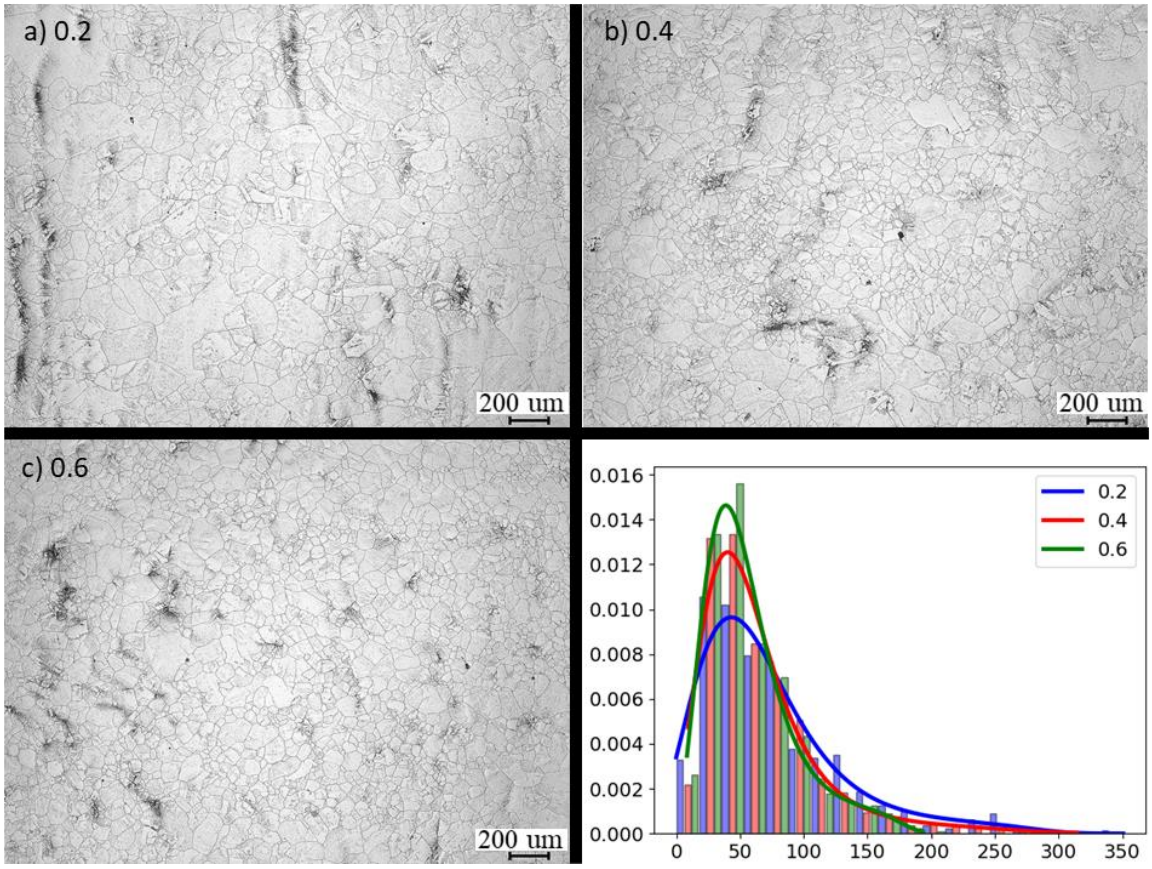


Figure 8. Microstructural images and corresponding probability density functions and data sets for temperature 1050 °C, holding time 0 s and strains (a) 0.2, (b) 0.4 and (c) 0.6.

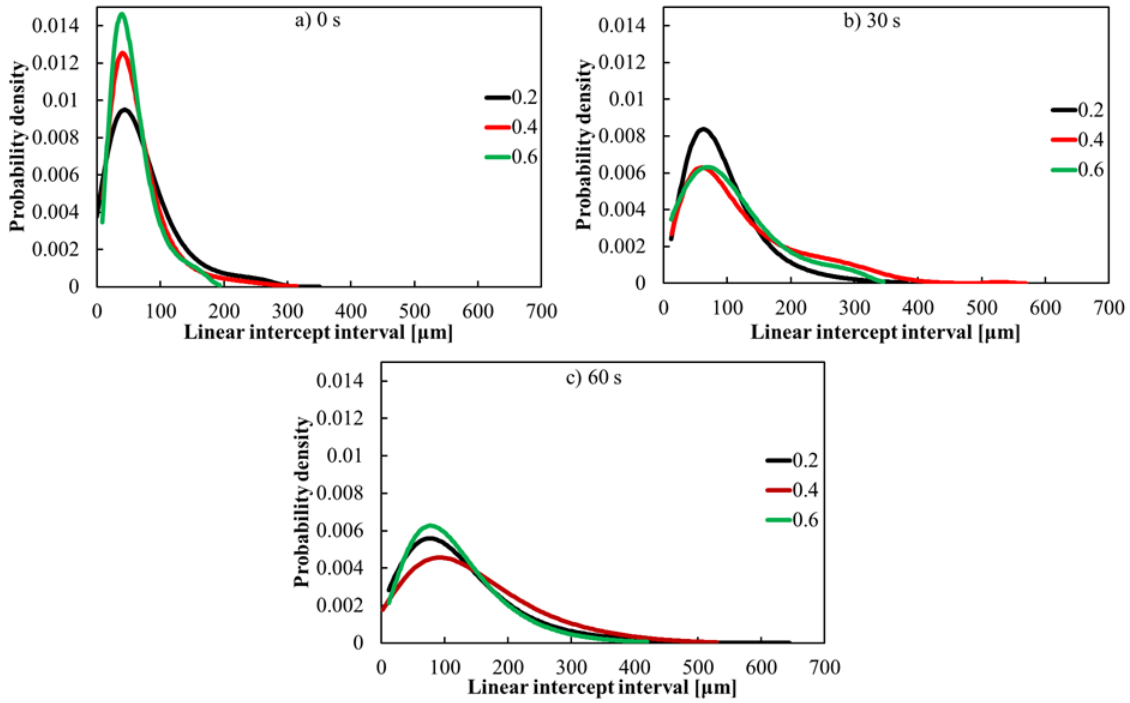


Figure 9. Probability density functions for strains 0.2, 0.4 and 0.6 with temperature 1050 °C and holding times (a) 0, (b) 30 and (c) 60 s.

At 0 s, higher strain causes a higher number of small grains. The small grains are likely freshly nucleated recrystallized grains, and their increasing amount suggests faster recrystallization start with higher strains. The effect of strain is difficult to observe with the presented method for 30 and 60 s. The effect of strain could be further investigated by inspecting the aspect ratio, but in the current study aspect ratio analysis is omitted.

5.4. Temperature

Figure 10 includes microstructural images for temperatures 1050, 1150 and 1250 °C with strain 0.4 and holding time 30 s. Their probability density functions, and data set histograms are collected to a single figure for comparison. Figure 11 shows the effect of temperature for each holding time, with strain set to 0.2.

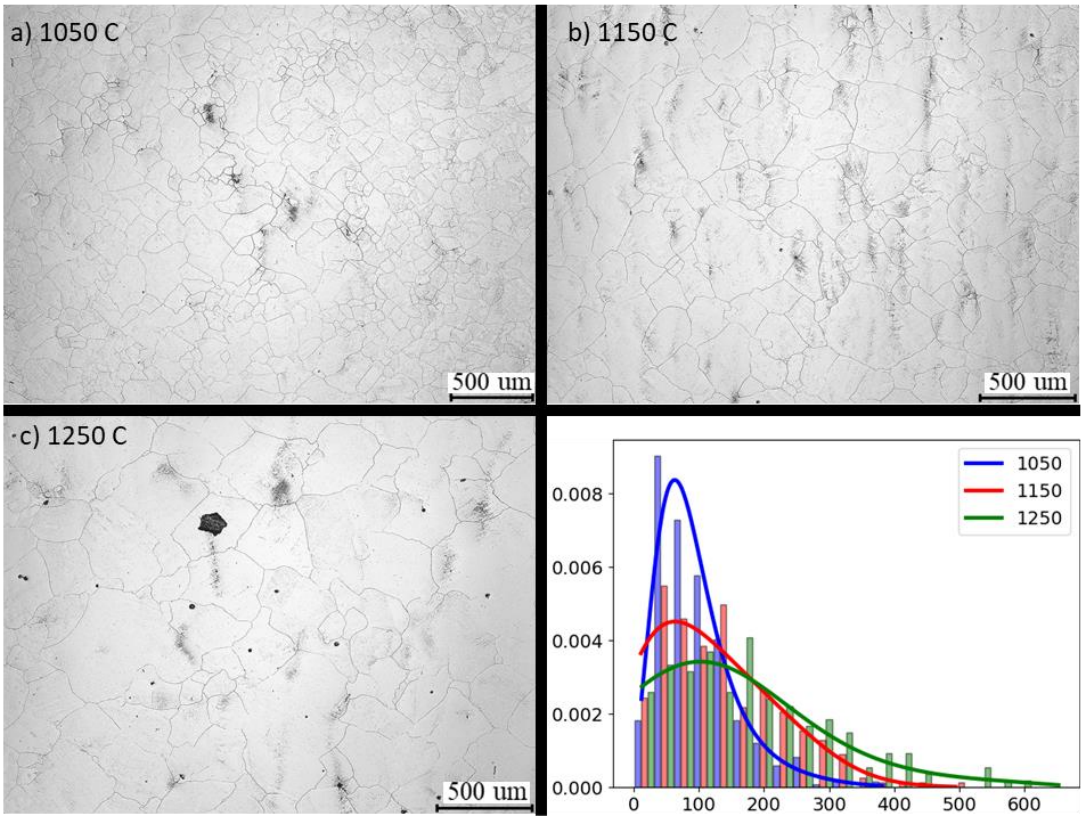


Figure 10. Microstructural images and corresponding probability density functions and data sets for strain 0.2, holding time 30 s and temperatures (a) 1050, (b) 1150 and (c) 1250 °C.

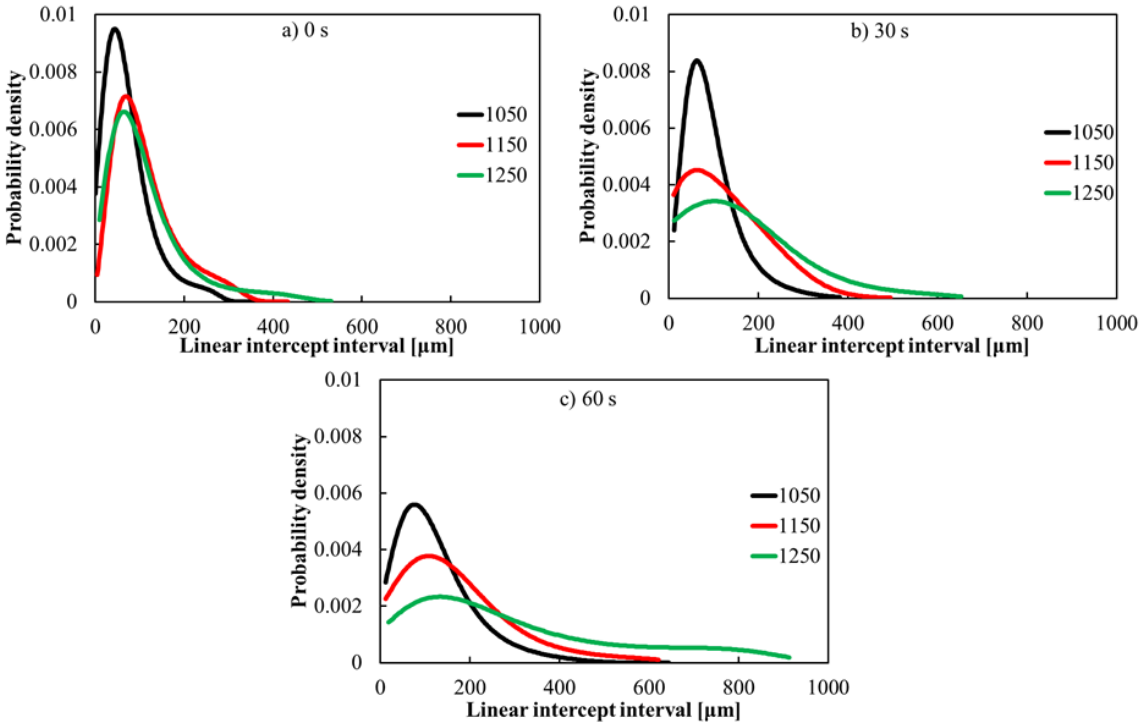


Figure 11. Probability density functions for temperatures 1050, 1150 and 1250 °C with strain 0.2 and holding times (a) 0, (b) 30 and (c) 60 s.

From the probability density functions it can be seen that temperature has a clear effect on linear intercept interval length: increasing temperature decreases the likelihood of small grains and increases the maximum interval length. Especially at 1250 °C the maximum grain size increases greatly.

5.5. Statistical parameters

Probability density functions are easy to compare against each other, as they have clearly defined function parameters. Central tendency parameters of the probability density functions and data set histograms for each case will be studied below. The number of bins in the histograms is set to 20 for easier comparison.

5.5.1. Mode. Mode is obtained for histogram data sets by selecting the highest bar in the histogram and for probability density functions by selecting the maximum value of the curve. In figure 12-a) they are compared against each other with varying temperature, strain and holding time. Data set (figure 12(b)) and probability density function (figure 12(c)) modes are plotted for different temperatures with strain and holding time.

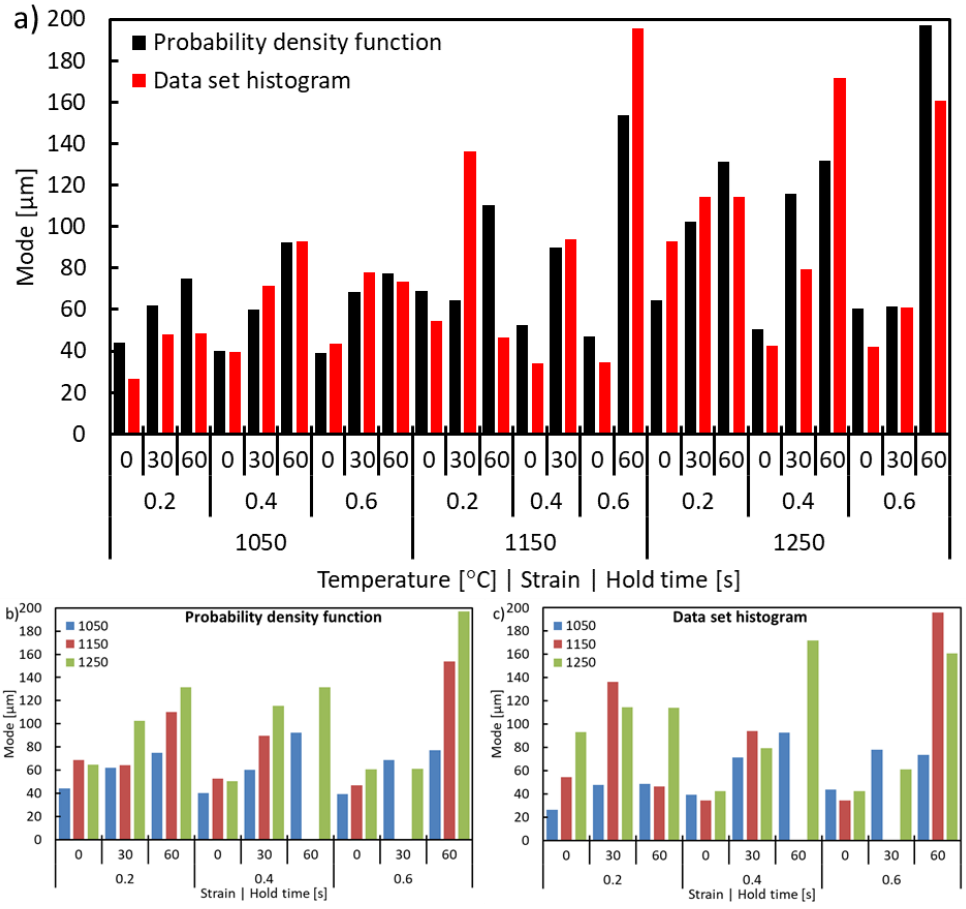


Figure 12. Modes with different temperatures, strains and holding times obtained from the probability density functions and histograms.

Modes show major changes between the probability density functions and data set histograms. Looking at both cases individually, function values show a clear increase for both temperature and holding time. Histogram data set values, on the other hand, have great differences and it takes a lot of imagination to find the trending effect of temperature and holding time. In this respect the probability density function seems to work better with modes.

5.5.2. Median. Median is obtained for data sets by choosing the midpoint of the ordered data set. Probability density function median is obtained by choosing the 50% cumulative distribution point. In figure 13 medians are presented similarly to modes.

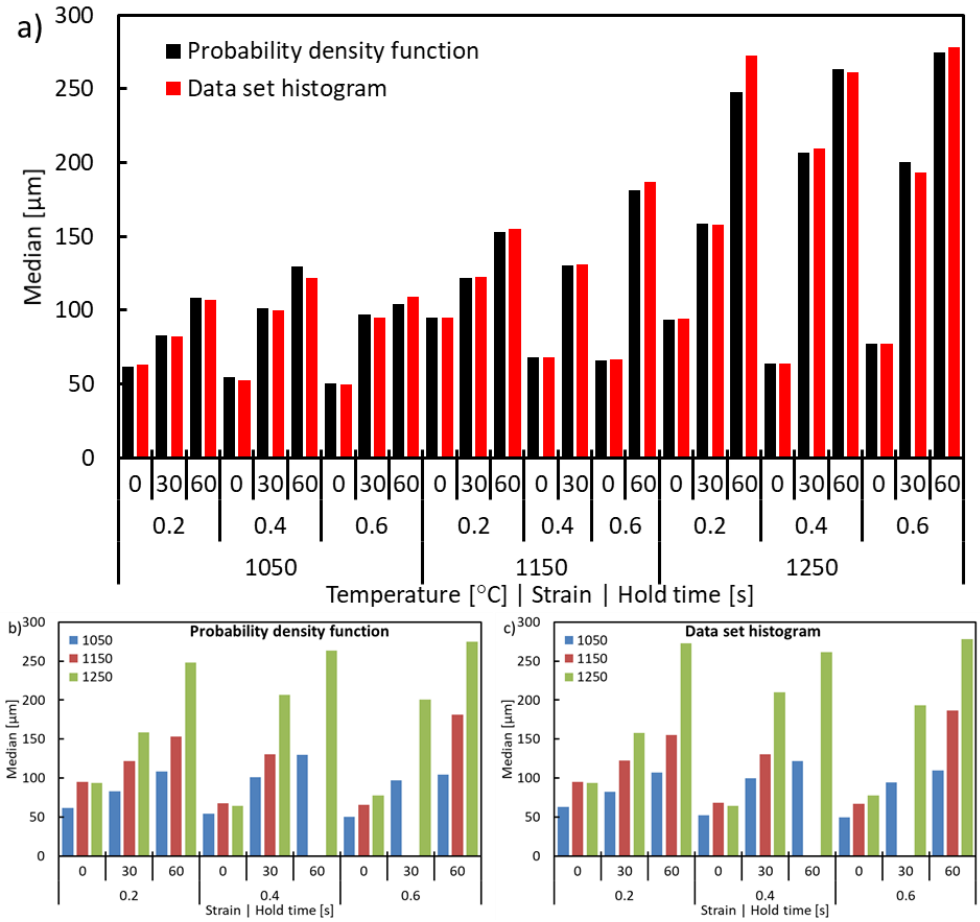


Figure 13. Medians with different temperatures, strains and holding times obtained from the probability density functions and histograms.

Medians have quite a bit of similarity between histogram data sets and their probability density functions. Both show the effect of temperature and holding time.

5.5.3. Average. Average value is obtained for the data sets by summing each datapoint together, then dividing that sum with the amount of datapoints. For probability density function the value is obtained by finding the linear intercept interval length value corresponding to the inverse of the division between maximum value and minimum value, $y_{avg} = 1/(x_{max} - x_{min})$. In figure 14 averages are presented similarly to modes.

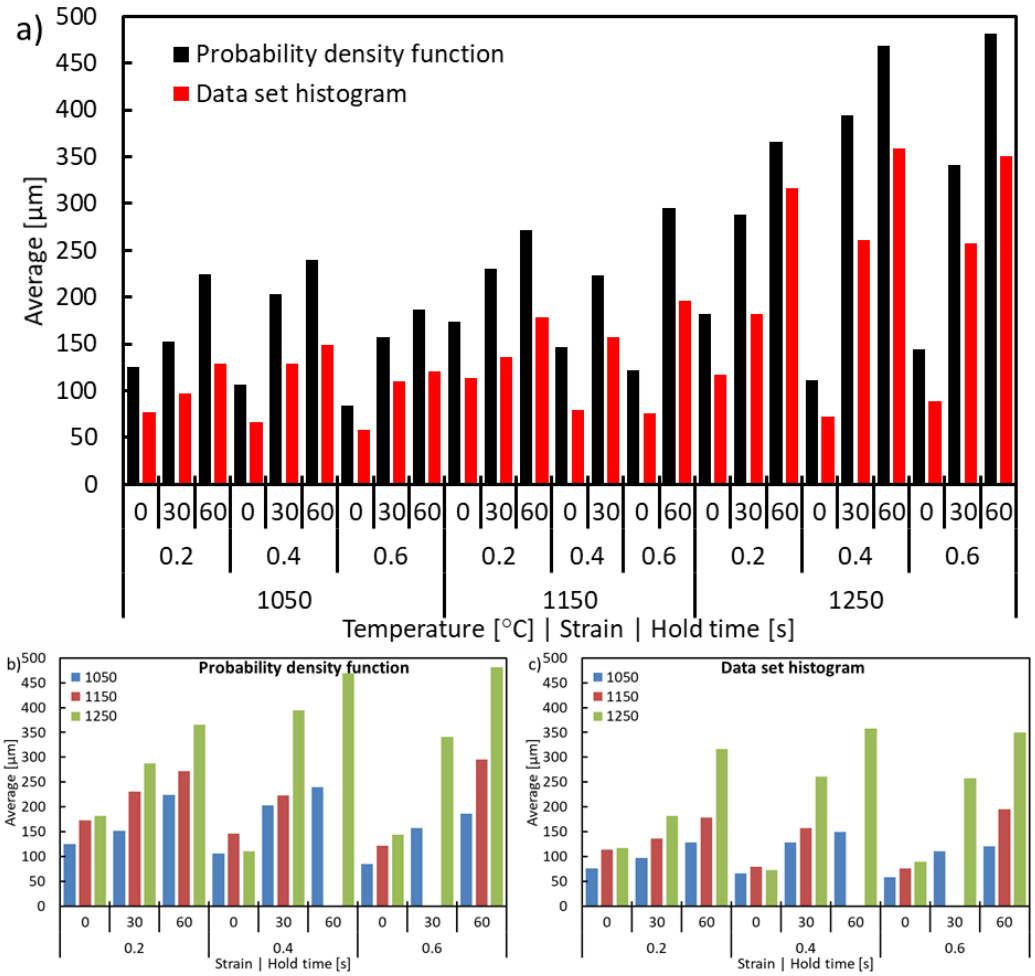


Figure 14. Average values with different temperatures, strains and holding times obtained from the probability density functions and histograms.

Averages show function values to be quite a bit larger than the data set. This is likely caused by the majority of the counted intercept intervals being small, which naturally emphasizes smaller grain sizes.

6. Conclusion

In summary, it was shown that the proposed polynomial probability density function generation method is viable and provides a probability density with an analytical function that is easy to use for further data analyzation. Data sets, here depicted as histograms, have some challenges in data visualization that can be avoided by using probability density functions.

Preliminary study between polynomial and log-normal fitting methods was conducted, and both seem to work quite well. The polynomial method is more complex and includes more fitting parameters than log-normal but is able to reproduce a wider range of data behaviors.

Chosen central tendencies were compared for both the data set and its analytical function and the expected, well-known, effect of temperature and holding time were found. Effect of strain could not be seen so easily with the chosen method. Grain aspect ratio might be a useful parameter to study the effect of strain further.

In the future, more in-depth probability density function comparison studies are planned, mainly focusing on the parameters defining curve shape. In addition, the possibility of finding a regression model capable of predicting grain size distribution evolution during static recrystallization and grain growth will be rigorously examined.

References

[1] Kaijalainen A J, Suikkanen P P, Limnell T J, Karjalainen L P, Kömi J I and Porter D A, 2013 *J. Alloys Compd* **577** 642–8
[2] Blum W, Li Y J, Chen J, Zeng X H and Lu K 2006 *Int. J. Mat. Res.* **97** 1661–6
[3] Seppälä O, Pohjonen A, Ilmola J, Jokiranta A, Kaijalainen A, Somani M and Larkiola J 2019 *J. Phys.: Conf. Ser.* **1270** 012027
[4] Pohjonen A, Seppälä O, Jokiranta A, Kaijalainen A, Somani M, Porter D, Larkiola J and Kömi J 2019 *J. Phys.: Conf. Ser.* **1270** 012013
[5] Higginson R L and Sellars C M 2003 *Worked Examples in Quantitative Metallography* (London: Maney)
[6] Maxima A *Computer Algebra System Version 5.46.0* (2022), https://maxima.sourceforge.io/
[7] Meurer A, et al 2017 SymPy: symbolic computing in Python *PeerJ Computer Science* 3: e103 https://doi.org/10.7717/peerj-cs.103
[8] Weisstein, Eric W "Log Normal Distribution." From MathWorld--A Wolfram Web Resource. https://mathworld.wolfram.com/LogNormalDistribution.html
[9] Koskenniska S, Seppälä O and Kömi J 2020 *Procedia Manuf.* **50** 684-8

Acknowledgements

The work of Mohammed Ali was conducted within the framework of the AMET project funded by Business Finland.
Financial assistance of Business Finland, project FOSSA- Fossil-Free Steel Applications, is acknowledged. Likewise, the authors also acknowledge the valuable scientific and technical input and knowhow to the research by M.Sc. Juha Uusitalo from the laboratory of Materials and Mechanical Engineering at the University of Oulu.