



Quantification of Aleatoric and Epistemic Uncertainty of Microstructures using Experiments and Markov Random Fields

Matthew Long^{*}, Arulmurugan Senthilnathan[†], and Pinar Acar[‡]
Virginia Tech, Blacksburg, Virginia, 24060

The present work addresses uncertainty quantification within the application of Markov-Random Fields (MRF) on the multi-scale modeling of microstructures. The aleatoric uncertainty of experimental measurements, as well as the epistemic uncertainty arising from computational microstructure reconstruction, is explored. The study is performed on the experimental data of Titanium-7wt%Aluminum (Ti-7Al) alloy, which is a candidate material for many aerospace systems owing to its desirable mechanical performance under large thermo-mechanical stresses. The aleatoric uncertainty of the experimental electron backscatter diffraction (EBSD) data is quantified by identifying the noises in data clusters for crystallographic orientations. After the identification and correction of such uncertainty, microstructure reconstruction is performed on the experimental samples using the MRF algorithm to create synthetic data. This reconstructed data provides a statistically similar representation of the test samples while introducing epistemic uncertainty on the microstructural features which is captured in this study by computing the distance metrics between experimental and synthetic data.

I. Introduction

Component-scale properties and performance of metallic materials are fundamentally driven by the underlying microstructural features. Therefore, multi-scale modeling requiring the exploration of the micro-scale features of materials, such as crystallographic orientations and grain topology, has become an important and growing research area. When it comes to visualizing the microstructural features of certain materials, instruments such as an optical microscope, Scanning Electron Microscope (SEM), Transmission Electron Microscope (TEM), and Atomic Force Microscopy (AFM) are used [1, 2]. To identify the grain orientation information for metallic microstructures, techniques such as Electron Backscatter Diffraction (EBSD) [3–8] or X-ray diffraction [9–11] are used. While these instruments and techniques provide us with real-world data, the process of getting this data consumes a large amount of time and materials for even small-scale samples. In order to minimize the costs associated with gathering this data, computational microstructure reconstruction methods are utilized to predict the evolution of the microstructure in larger domains by utilizing the small-scale information already gathered. One of the methods that is implemented for computational microstructure reconstruction is the Markov Random Field (MRF) approach, and while it has been proven to be efficient in predicting large-scale microstructures in 2D [12–15], this algorithm, similar to other computational methods, is a source of epistemic uncertainty that needs to be quantified in order to fully understand its capabilities and limitations.

Moreover, microstructures are inherently stochastic due to the aleatoric uncertainty associated with their processing and experimental measurements. When it comes to uncertainty quantification of synthetic metallic microstructure data, previous work has solely focused on the analysis of either aleatoric uncertainty [16–20] or epistemic uncertainty [21]. This work will focus on a more in-depth analysis of the uncertainty quantification of microstructures and will do so by isolating two main sources of uncertainty present in the microstructure reconstruction process used.

Microstructure reconstruction is designed to generate statistically similar synthetic microstructures from test samples. One of the methods that is used for microstructure reconstruction is MRF, which is based on a high-order Ising model. In this application, the MRF-generated synthetic samples are expected to be statistically similar to the input test data by utilizing conditional probability information captured from the test sample to estimate the location of certain color pixels during reconstruction. Due to this, there is some expected algorithmic randomness in each generated synthetic sample which contributes to epistemic uncertainty. In addition, a source of aleatoric uncertainty can be found in the EBSD data that is measured after thermo-mechanical tests. This aleatoric uncertainty can be visualized through the grains of an EBSD sample that do not have the exact same color values (due to perturbations in pixel values) even though they

^{*}Graduate Research Assistant, Department of Mechanical Engineering

[†]Graduate Research Assistant, Department of Mechanical Engineering

[‡]Assistant Professor, Department of Mechanical Engineering

demonstrate identical crystallographic orientations, which can cause additional uncertainty for the MRF reconstruction. In the present study, we identify the aleatoric uncertainty first by color-correcting the sample for consistent grain colors to limit the uncertainty present in the synthetic microstructure. Ti-7Al alloy is selected for this study as the application material since it has a wide range of use cases in the aerospace industry owing to its superior mechanical performance under large thermal and mechanical stresses [22].

The paper is structured as follows. Section II discusses the methods that are used to identify and remove the aleatoric uncertainty in the experimental data. Section III explains the concept behind MRF and how it is used to reconstruct a colored image. The summary of all of the work, as well as the results from the above methods and future plans, are provided in Section IV.

II. Identifying Aleatoric Uncertainty in Experimental Images

A. Identifying Grains using K-Means Clustering

In order to identify the aleatoric uncertainty present in our given EBSD experimental data, shown in Figure 1, the data must first be broken down into usable sizes. The EBSD data consists of a 1029×1024 matrix of pixel values, so in order to generate accurate images within a reasonable amount of time, samples with a resolution of 150×150 are used. Next, to identify the grains in those samples, an adjusted k-means clustering algorithm is used to follow the methodology developed by our group for grain topology quantification [21]. This algorithm is designed to segment an image into a given number of clusters based on the Red-Green-Blue (RGB) channel values of each pixel. However, due to the noise from the experimental data, the algorithm would frequently mark grain boundaries as separate new grains or as a part of existing grains, which can be seen in Figure 2. These grain definitions would cause inaccuracies in the evaluation of epistemic uncertainty by both introducing new grains and blurring grains together, an example of which can be seen in Figure 3. Due to this issue, an algorithm is developed to both refine grain separation and add grain boundaries back to their original grains.

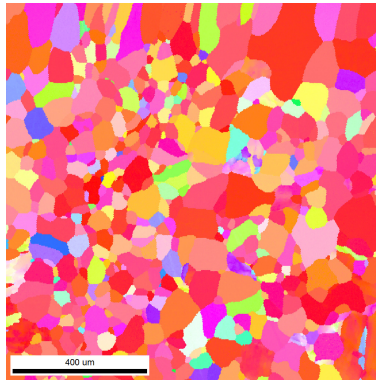


Fig. 1 Experimental EBSD image of Ti-7Al

B. Post Processing of Initial Grain Separation

The post-processing algorithm is designed to utilize two main sources of information. The first is the output of the k-means algorithm, which is an $N \times N$ array of integer labels for the original image. This data is used as an initial springboard for the algorithm, both for the individual label assignments and for the analysis of the neighboring pixels. The second source of information utilized is the colors of the original experimental sample, translated into the CIELAB color space.

The algorithm starts by checking the pixel labels around the pixel located at (i, j) . It checks a square area with sides equal to $(2r + 1)$, where r is a user-defined radius. This data is then tabulated, and the label value that has the most occurrences in the specified area is compared to the pixel label that was assigned at (i, j) . If the two values are different, then the pixel at (i, j) becomes unassigned unless the majority label occupies more than 80% of the searched area. Once each pixel has been checked, the algorithm is left with a set of unassigned pixels that can be seen in Figure 3.

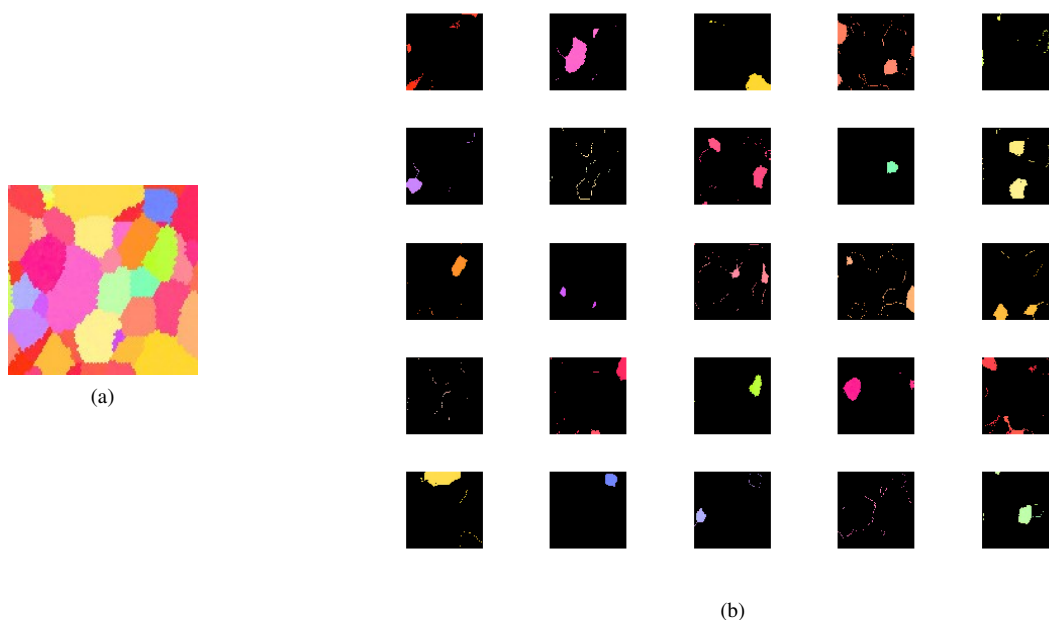


Fig. 2 (a) A 150×150 cropped section of the original experimental image, (b) Separation of 25 clusters via the adjusted k-means clustering algorithm

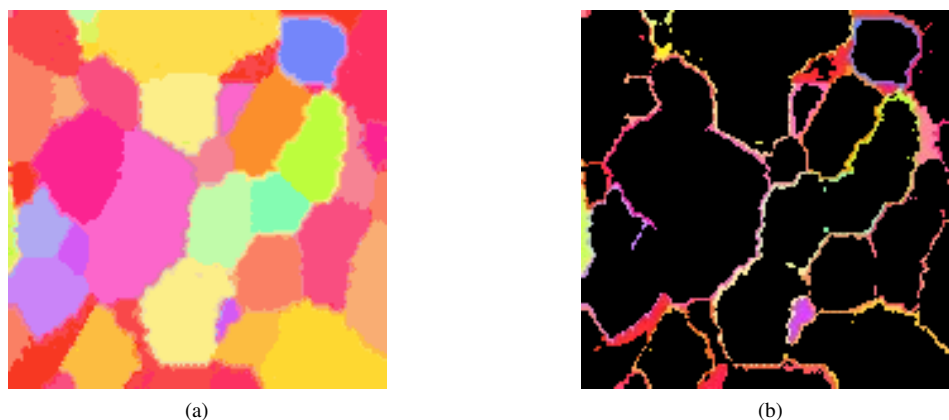


Fig. 3 (a) Example of aleatoric uncertainty analysis without post-processing, (b) Visualizations of pixels that are misassigned with k-means clustering

The unassigned pixels can be classified into four main groups, which are ‘loose pixels’, ‘broken grains’, ‘small grains’, and ‘grain boundaries’.

Each of those groups can be added to their correct label once the CIELAB color space data is added to the post-processing algorithm. The CIELAB color space is designed to modify the RGB values of a pixel into numbers that can be used to identify similar colors in a way that more accurately matches human perception of those colors, which also helps identify accurate grain boundaries. This is done by checking the color of the selected pixel against the average color of each of the labels around it, including pixels that have been unassigned from labels. Then if there is a suitably similar color around the selected pixel, it will be added to that label. This also allows for grains that have become unassigned to be assigned to an unused label. This process ends with grains that are shown in Figure 4.

Once grains have been assigned through the above process, their RGB values are then averaged together to get a

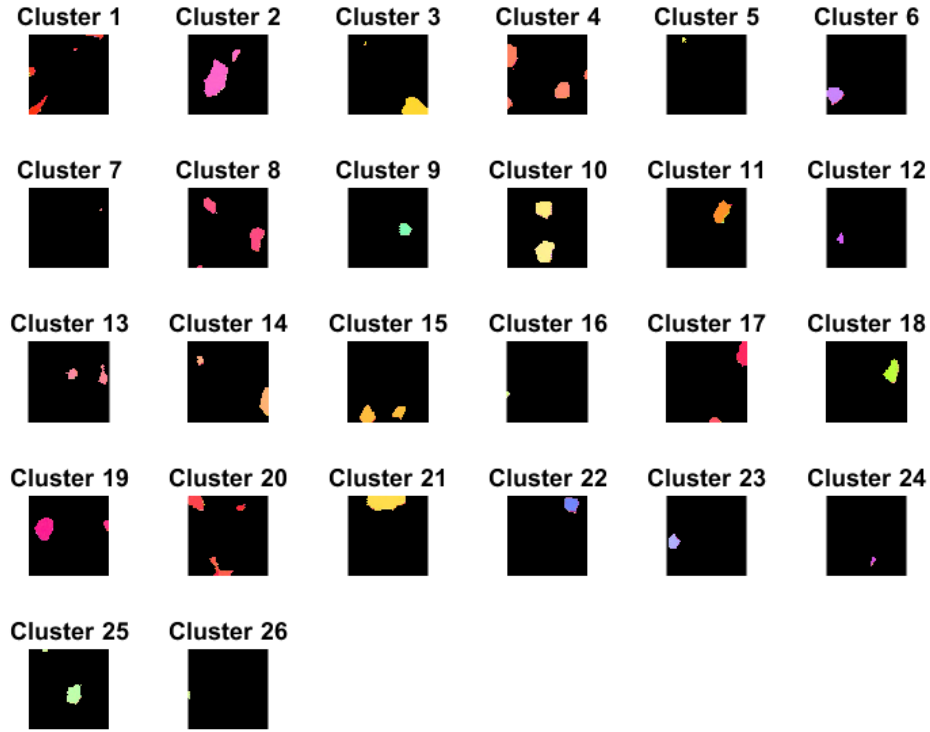


Fig. 4 The final clusters developed by the post-processing algorithm

product that has a reduced aleatoric uncertainty. The entire process is repeated until the percent difference between the initial EBSD image and the color-corrected image converges to 0.1%, which can be seen in Figure 5 and is calculated using Equation 1. This process is completed for each of the five samples that are used for this study, the data for which can be seen in Figure 6 and resulted in an average percent difference of 8.41%.

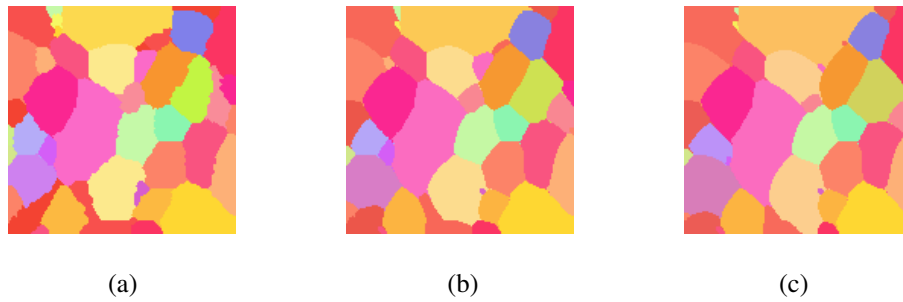


Fig. 5 (a-c) Evolution of the clusters developed by the post-processing algorithm during convergence

$$\sum_{i=1}^N \sum_{j=1}^N |Real_{i,j} - Corrected_{i,j}| \quad (1)$$

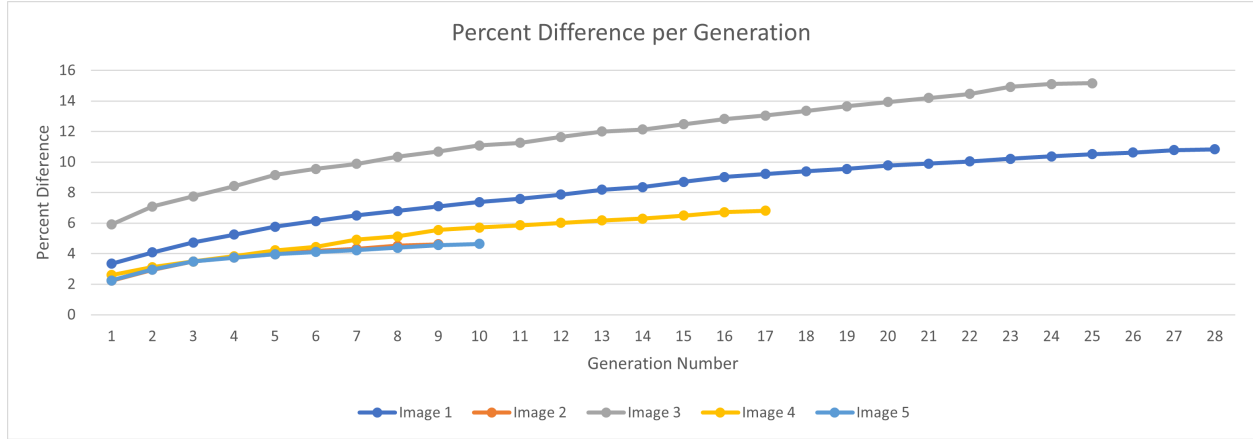


Fig. 6 Convergence of each of the cropped sections of the initial experimental data

III. Microstructure Reconstruction using Markov Random Fields

The MRF is a graphical model that relies on probability analysis. It is used in this study to generate synthetic microstructures using small-scale EBSD test data to predict the evolution of microstructures in large domains. The algorithm determines the probabilities based on the color values of each individual pixel in the EBSD data and then assembles the synthetic microstructures by assuming an Ising model.

An Ising model requires a square image of dimensions $N \times N$, which then is represented as an $N \times N$ grid of pixel values that contain any value X_i among one of the color levels G of the original image in the range $0, 1, \dots, G - 1$. While a basic Ising model only captures the nearest-neighbor pixel information, shown in Figure 7(a), a higher-order representation can be achieved with the MRF method by modeling every pixel in a certain area, defined with an external parameter that is called the window size. This higher-order model illustration can be seen in Figure 7(b), and the pixel of interest for this model is found in the center and connected to all other pixels in the defined window. Inside the MRF approach, the probability of any value (X) is assumed to be conditionally independent of all values outside of its window due to the probability density of all particles being represented by the local Markovian property [21, 23–25].

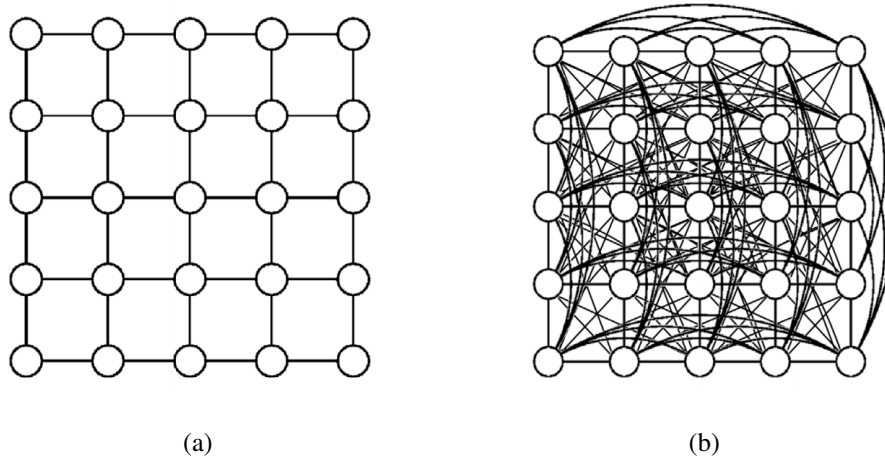


Fig. 7 Visualization of an MRF model, where the circles shown are the pixels of the chosen image and the lines connect each neighbor of the model. (a) Nearest-neighbor Ising model. (b) The model used by the MRF method to perform microstructure reconstruction is a higher-order Ising model.

The cropped image in Figure 8 (a) shows a microstructure sample with multiple large grains spread out over the entire sample. When dealing with EBSD data that has patterns similar to this, a large window size is typically used to capture the necessary data to recreate statistically similar samples using MRFs, a sample of which can be seen in



Fig. 8 (a) Cropped section of the color corrected image used in MRF reconstruction, (b) Reconstruction of the image with a window size of 29 and a resolution of 150×150 pixels.

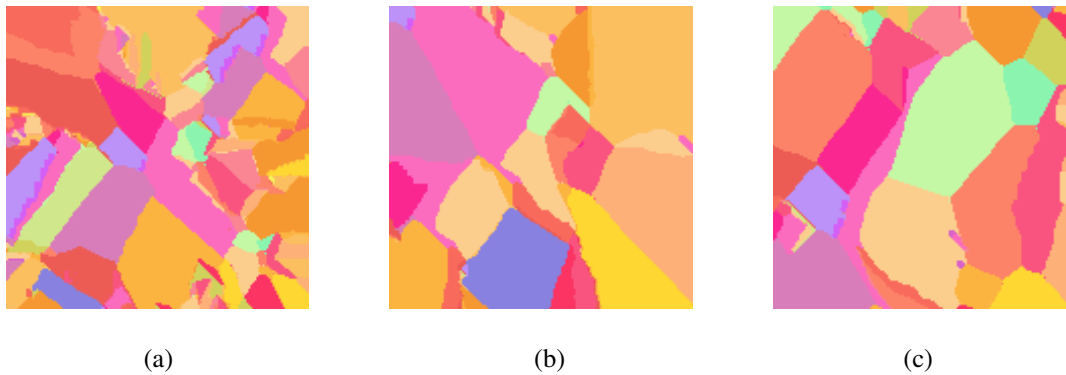


Fig. 9 Reconstruction of the cropped image with window sizes of 5, 7, and 9, respectively, and with a resolution of 150×150 pixels.

Figure 8 (b) with a large window size of 29. The difficulty in using large window size values, especially for uncertainty quantification purposes, is associated with the exponential growth of computing times required to generate a synthetic image, with the sample image (shown in Figure 8 (b)) taking 45 minutes in a desktop computer to create. Due to computing time expenses and the resulting numerically intractable nature of the uncertainty quantification problem with large window size values, the uncertainty analysis is performed for window sizes of 5, 7, and 9 as shown in Figure 9 so as to identify both potential sources of error and limitations of the MRF algorithm as a result of the epistemic uncertainty.

The epistemic uncertainty created from the MRF is calculated once the average percent difference between the reconstructed image and the cropped image has converged to 0.1%, causing each data point to have a varying number of samples used in the calculations. Equation 2 shows the calculation of these values which can be seen in Figure 10 and Table 1. These compare the converged average percent difference of the five samples that are used in the study for each window size that was used in the reconstruction. The variation in the convergence values is most substantial when using a window size of five, which can be directly attributed to the fact that this window size value cannot capture the grain information of the samples precisely. Window size seven has both the smallest variation out of each size that was tested and the least number of data points necessary for convergence. Window size nine has the smallest average epistemic uncertainty compared to the other two, but the epistemic uncertainty in these reconstructed images is larger than ideal. The randomness that can be seen in the reconstructed samples can be directly attributed to the MRF algorithm due to the mismatching color density and larger grain sizes. Due to these variables, a smaller window size is not able to accurately retrieve the neighbor relationships needed for a statistically more equivalent synthetic representation of the test sample.

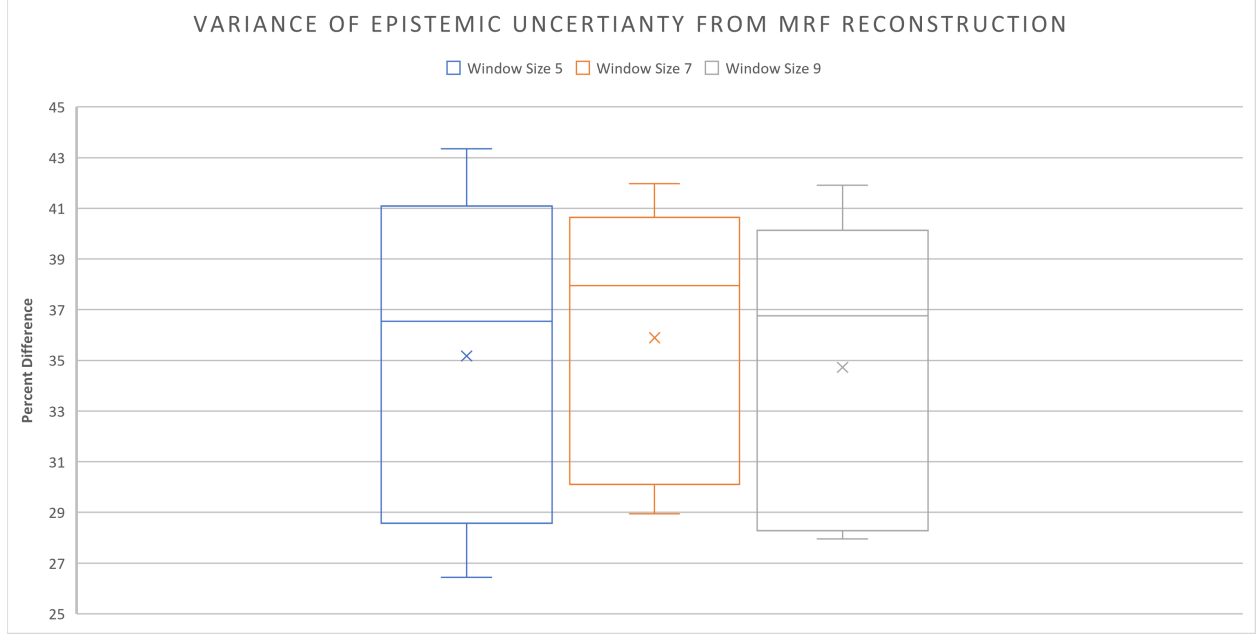


Fig. 10 Variation in epistemic uncertainty related to the window size.

$$\sum_{i=1}^N \sum_{j=1}^N |MRF_{i,j} - Corrected_{i,j}| \quad (2)$$

Table 1 Epistemic uncertainty values for reconstructed images

	Window Size 5	Window Size 7	Window Size 9
Mean (%)	35.169	35.89	34.72
Variation (%)	16.912	13.023	13.963
Average Number of Data Points	13.0	9.6	12.0
Average Time (sec)	255.523	420.01	614.988

IV. Conclusion

The present work focuses on uncertainty quantification of both aleatoric and epistemic uncertainty present in synthetic microstructure samples. The formulation of this analysis is based on the separation of uncertainties to better understand the effects of the epistemic uncertainty stemming from the Markov Random Field approach. While aleatoric uncertainty is identified in the EBSD sample and minimized with the application of a post-processing algorithm, the epistemic uncertainty created during the microstructure reconstruction process is found to be large due to the use of small window size values to create necessary statistics as the use of large window size values is limited because of required computing times. The framework developed here is tested on forged Ti-7Al microstructure samples, but it can be applied to other polycrystalline materials in future work. In addition, future work can also be focused on applications of larger window sizes using more efficient uncertainty quantification approaches.

Acknowledgments

The authors acknowledge the financial support received from the Air Force Office of Scientific Research (AFOSR) through the Young Investigator Program (YIP) under Grant FA9550-21-1-0120 (PO: Dr. Ali Sayir) and from the

References

- [1] Zambaldi, C., Yang, Y., Bieler, T. R. and Raabe, D., "Orientation informed nanoindentation of α -titanium: Indentation pileup in hexagonal metals deforming by prismatic slip," *Journal of Materials Research*, Vol. 27, No. 1, 2012, pp: 356-367.
- [2] Wright, Stuart I., and Matthew M. Nowell, "EBSD image quality mapping," *Microscopy and microanalysis*, Vol. 12, No. 1, 2006, pp: 72-84.
- [3] Wright, Stuart I, "Fundamentals of automated EBSD", *Electron backscatter diffraction in materials science*, 1st ed., Vol. 2, Springer, Boston, MA, 2000, pp. 51-64.
- [4] Schwartz, A.J., Kumar, M., Adams, B.L. and Field, D.P. eds., *Electron backscatter diffraction in materials science*, Vol. 2, Springer, New York, 2009, p. 403.
- [5] Adams, Brent L., Stuart I. Wright, and Karsten Kunze, "Orientation imaging: the emergence of a new microscopy," *Metallurgical Transactions A*, Vol. 24, No. 4, 1993, pp: 819-831.
- [6] Dingley, Do J., and V. Randle, "Microtexture determination by electron back-scatter diffraction," *Journal of materials science*, Vol. 27, No. 17, 1992, pp: 4545-4566.
- [7] Humphreys, F. J., "Characterisation of fine-scale microstructures by electron backscatter diffraction (EBSD)," *Scripta materialia*, Vol. 51, No. 8, 2004, pp: 771-776.
- [8] Kulkarni, A. K., Kirk H. Schulz, T. S. Lim, and M. Khan. "Dependence of the sheet resistance of indium-tin-oxide thin films on grain size and grain orientation determined from X-ray diffraction techniques," *Thin solid films*, Vol. 345, No. 2, 1999, pp: 273-277.
- [9] Tamura, N., MacDowell, A.A., Celestre, R.S., Padmore, H.A., Valek, B., Bravman, J.C., Spolenak, R., Brown, W.L., Marieb, T., Fujimoto, H. and Batterman, B.W., "High spatial resolution grain orientation and strain mapping in thin films using polychromatic submicron x-ray diffraction." *Applied Physics Letters*, Vol. 80, No. 20, 2002, pp: 3724-3726.
- [10] Ludwig, W., Reischig, P., King, A., Herbig, M., Lauridsen, E.M., Johnson, G., Marrow, T.J. and Buffiere, J.Y., "Three-dimensional grain mapping by x-ray diffraction contrast tomography and the use of Friedel pairs in diffraction data analysis." *Review of scientific instruments*, Vol. 80, No.3, 2009, 033905.
- [11] Ludwig, W., Schmidt, S., Lauridsen, E.M. and Poulsen, H.F., 2008. "X-ray diffraction contrast tomography: a novel technique for three-dimensional grain mapping of polycrystals. I. Direct beam case," *Journal of Applied Crystallography*, Vol. 41, No. 2, 2008, pp: 302-309.
- [12] Acar, Pinar, and Veera Sundararaghavan, A Markov random field approach for modeling spatio-temporal evolution of microstructures, *Modelling and Simulation in Materials Science and Engineering*, Vol. 24, No. 7, 2016, 075005.
- [13] Kumar, A., Nguyen, L., DeGraef, M. and Sundararaghavan, V., "A Markov random field approach for microstructure synthesis", *Modelling and Simulation in Materials Science and Engineering*, Vol. 24, No. 3, 2016, 035015.
- [14] Liu, Y., Greene, M.S., Chen, W., Dikin, D.A. and Liu, W.K., "Computational microstructure characterization and reconstruction for stochastic multiscale material design," *Computer-Aided Design*, Vol. 45, No. 1, 2013, pp: 65-76.
- [15] Bostanabad, R., Bui, A.T., Xie, W., Apley, D.W. and Chen, W., "Stochastic microstructure characterization and reconstruction via supervised learning," *Acta Materialia*, Vol. 103, January 2016, pp: 89-102.
- [16] Acar P., "Uncertainty quantification for Ti-7Al alloy microstructure with an inverse analytical model (AUQLin)," *Materials* 12 (11) (2019) 1773.
- [17] Acar P., Sundararaghavan V., "Uncertainty quantification of microstructural properties due to variability in measured pole figures," *Acta Materialia*, Volume 124, 2017, pp. 100–108. <https://doi.org/10.1016/j.actamat.2016.10.070>.
- [18] Acar P., Sundararaghavan V., "Uncertainty quantification of microstructural properties due to experimental variations," *AIAA Journal*, Vol. 55, No. 8, 2017, pp. 2824–2832.
- [19] Acar P., "Machine learning reinforced crystal plasticity modeling under experimental uncertainty," *AIAA Journal*, Vol. 58, No. 8, 2020, pp. 3569-3576.

- [20] Senthilnathan A., Acar P., "Shape moment invariants as a new methodology for uncertainty quantification in microstructures", AIAA Paper 2021-1697, Jan. 2021. <https://arc.aiaa.org/doi/abs/10.2514/6.2021-1697>
- [21] Senthilnathan, A., and Acar, P., "Multi-Scale Modeling for Texture and Grain Topology of Polycrystalline Microstructures under Uncertainty," AIAA Paper 2022-2106, Dec. 2021. <https://arc.aiaa.org/doi/abs/10.2514/6.2022-2106>
- [22] Dutta Bhaskar, Froes Francis, "Additive Manufacturing of Titanium Alloys: State of the Art, Challenges and Opportunities," Netherlands: Elsevier Science, 2016.
- [23] Acar. P. and Sundararaghavan, V., "A Markov Random Field Approach for Modeling Spatio-Temporal Evolution of Microstructures", Modelling and Simulation in Materials Science and Engineering, Vol. 24, No. 7, 075005, 2016.
- [24] Acar, P. and Sundararaghavan, V., "Do Epistemic Uncertainties Allow for Replacing Microstructural Experiments with Reconstruction Algorithms?", AIAA Journal, Vol. 57, No. 3, pp: 1078-1091, 2019.
- [25] Senthilnathan, A., Acar, P. and De Graef, M., "Markov Random Field based microstructure reconstruction using the principal image moments", Materials Characterization, Vol. 178, 111281, August 2021.