408
doi:10.1017/S1431927621001987
Microsc. Microanal. 27 (Suppl 1), 2021

A Machine Learning Approach to Cluster Characterization for Atom Probe Tomography

Roland Bennett[1], Andrew Proudian[2] and Jeramy Zimmerman[3]

[1]Colorado School of Mines, GOLDEN, Colorado, United States, [2]Colorado School of Mines, Physics, Golden, Colorado, United States, [3]Colorado School of Mines, Physics, GOLDEN, Colorado, United States

The clustering properties of solute species drive performance in materials ranging from metal alloys to organic light emitting diodes [1][2][3]. In atom probe tomography (APT) people have typically used cluster detection algorithms, such as the maximum separation algorithm (MSA), to detect and then characterize individual clusters. MSA, in particular, has the drawbacks of requiring user input parameters and not being sensitive to low density clustering [4][5]. While other cluster detection techniques have been developed, they often share similar limitations requiring high contrast between clusters and background [6][7][8]. In this work, we advance a machine learning model implemented using rapt [9], which we have developed in the statistical computing language R. Our model is based on spatial statistics summary functions that characterize global clustering properties and behavior in APT data sets, providing an alternative characterization approach to cluster detection analysis. In previous work, we utilized a Bayesian regularized neural network (BRNN) machine learning model, trained on features derived from Ripley's K function to measure four metrics that characterize clusters: the cluster dopant density (ρ1), the background dopant density (ρ1), the mean cluster radius (r), and the radius blur (δr) (i.e. the standard deviation of the cluster radius divided by the mean cluster radius) [5]. Here, we improve upon our previous work by incorporating features derived from the first-order summary functions *G*, *G-cross*, and *F* into our models, resulting in more accurate cluster analysis.

These first order summary functions enable ρ1 and ρ2 to be predictions with very low error: in simulated training and testing data sets, 90% of predictions for both were within 3.5% of the actual value, as shown for ρ1 in Figure 1 (an improvement from 18% for ρ1 in our previous work). While the percent error of ρ2 was not measured in our previous work, the value of absolute error of the 90th error percentile was reduced by 94%. These first order summary functions enable decoupling of *r* from δr, enabling the prediction of the average cluster radius itself, with 90% of predictions falling within 15% of the actual value (in comparison to 18% for predictions based solely on the K-function). The predicted value vs true value for ρ_1 is shown in Figure 2.

The simulated data sets used in this work (clustered point patterns of random clustering metrics, 10,000 for training and 2,500 for testing) were created on a single 28-core high-performance computing node and required only 90 minutes to generate and 15 minutes to train and create predictions, meaning it is also feasible on common desktop computers. A larger model using ten times as much data was also examined, but only minor improvements were seen, therefore not justifying the greater computational costs.

In this talk, we discuss development and results of this algorithm, its applications to experimental data, and the implications of global clustering behavior. We have made example analyses available on our website, enabling other users of the APT community to easily adopt this method of cluster analysis.

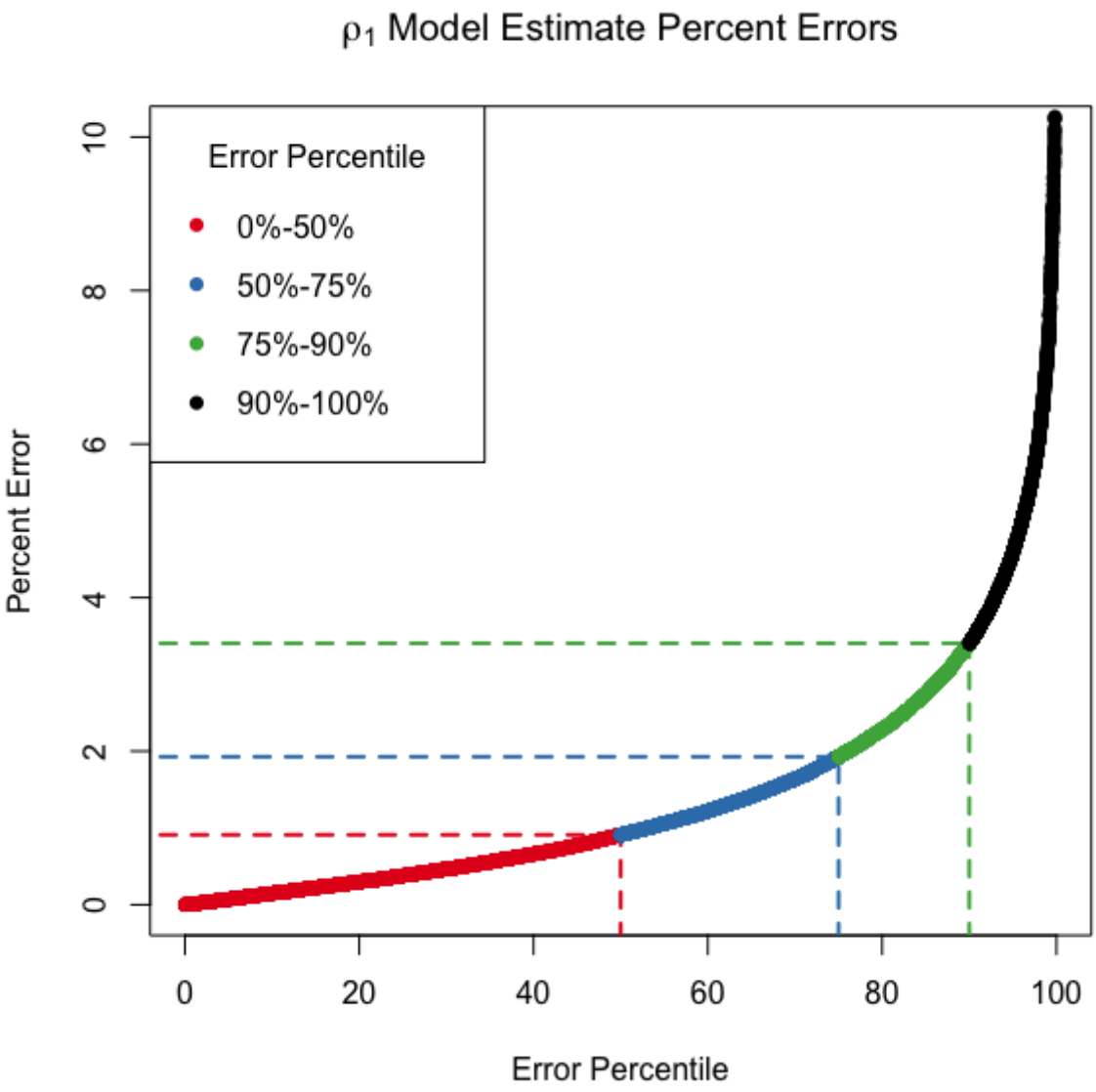

Figure 1. Percent error vs. error percentile for ρ1. 90% of predictions fall within 3.5% error. 75% of predictions fall within 2% error, and 50% of predictions fall within 1% error.

https://doi.org/10.1017/S1431927621001987 Published online by Cambridge University Press

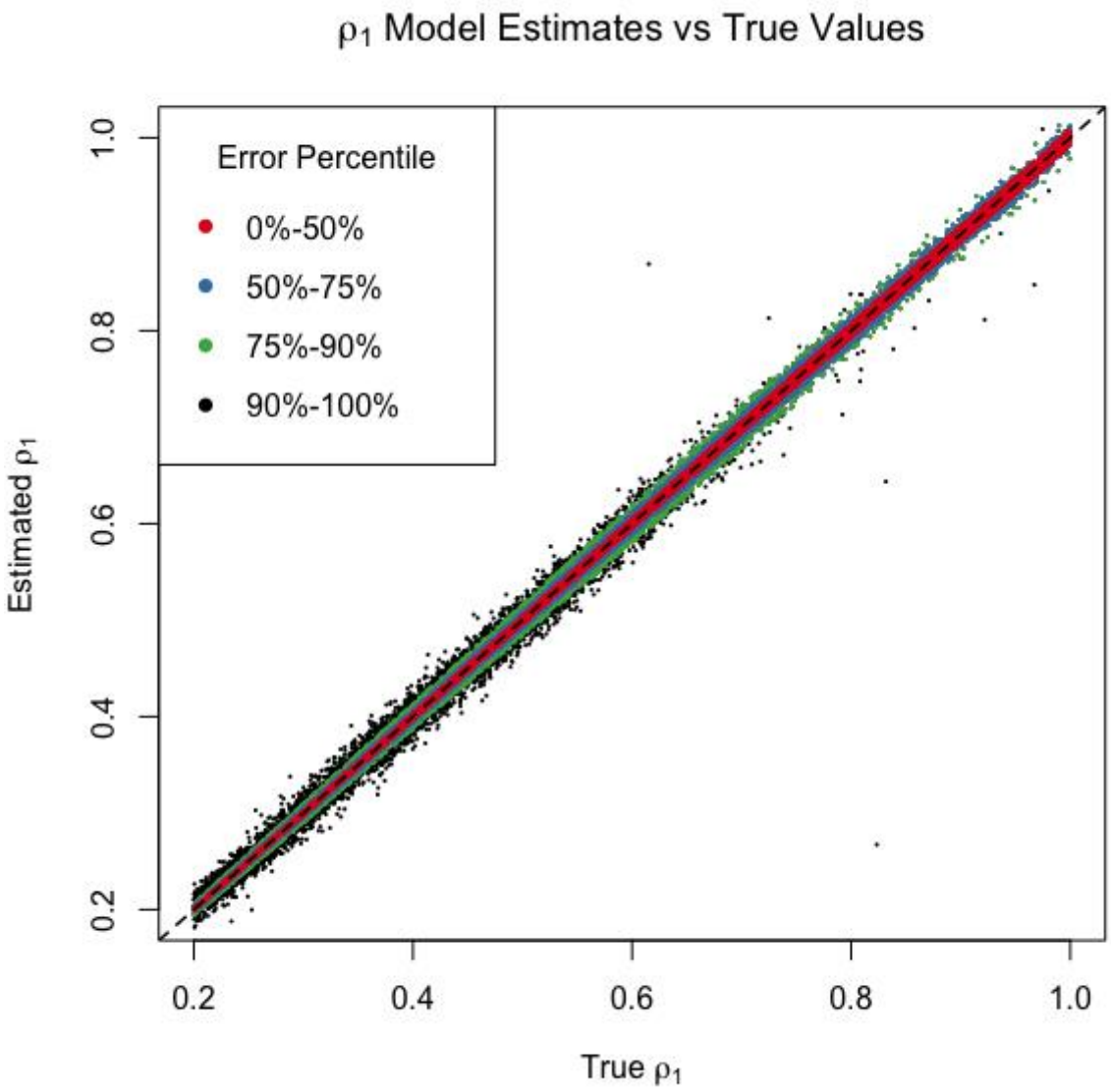

Figure 2. Predicted value vs true value for ρ1, with colors corresponding to error percentile. This contains data from 10 different simulations in order to increase the number of data points present.

References

[1] Jaskot, M.B. (2020), Linking Morphology to Electronic Properties in Small-Molecular Organic Semiconductors

[2] Deschamps, A, Pechiney-Centre de Recherches de Voreppe, & Brechet, Y. Influence of predeformation and ageing of an Al-Zn- Mg alloy. 2: Modeling of precipitation kinetics and yield stress. United States. doi:10.1016/S1359-6454(98)00296-1.

[3] Hirata, A., Fujita, T., Wen, Y. R., Schneibel, J. H., Liu, C. T., & Chen, M. W. (2011). Atomic structure of nanoclusters in oxide- dispersion-strengthened steels. Nature Materials, 10(12), 922-926. https://doi.org/10.1038/nmat3150

https://doi.org/10.1017/S1431927621001987 Published online by Cambridge University Press

[4] Jägle, E., Choi, P., & Raabe, D. (2014). The Maximum Separation Cluster Analysis Algorithm for Atom-Probe Tomography: Parameter Determination and Accuracy. Microscopy and Microanalysis, 20(6), 1662-1671. doi:10.1017/S1431927614013294
[5] Vincent, G., Proudian, A.P., & Zimmerman, J. (2020). Three Dimensional Cluster Analysis for Atom Probe Tomography Using Ripley's K-function and Machine Learning.
[6] Leigh T Stephenson, Michael P Moody, Peter V Liddicoat, and Simon P Ringer. New techniques for the analysis of fine-scaled clustering phenomena within atom probe tomography (apt) data. Microscopy and Microanalysis, 13(6):448–463, 2007.
[7] D Vaumousse, A Cerezo, and PJ Warren. A procedure for quantification of precipitate microstructures from three-dimensional atom probe data. Ultramicroscopy, 95:215–221, 2003.
[8] Anna V. Ceguerra , Michael P. Moody , Leigh T. Stephenson , Ross K.W. Marceau & Simon P. Ringer (2010) A three- dimensional Markov field approach for the analysis of atomic clustering in atom probe data, Philosophical Magazine, 90:12, 1657- 1683, DOI: 10.1080/14786430903441475
[9] A.P. Proudian. Rapt: R for atom probe tomography. https://github.com/ aproudian2/rapt, 2020.

https://doi.org/10.1017/S1431927621001987 Published online by Cambridge University Press