SAND2022-8162J

JOM manuscript No.
(will be inserted by the editor)

Applications of Autonomous Data Collection and Active Learning

Andrew T. Polonsky[1*] · **Patrick G. Callahan**[2]

Received: date / Accepted: date

Advances in sensors and robotics have dramatically improved the diversity of experimental approaches available to the materials community. Autonomous data collection platforms, either custom-made or commercially available, provide researchers with novel tools with which to probe materials behavior and perform advanced materials characterization. The application of novel controls algorithms and active learning approaches can create much more robust experimental data, or can be used to improve the performance of existing characterization tools. Five papers within this special topic focus on experimental and computational methodologies for use in automatic data collection routines for materials characterization. From novel platforms for materials discovery to new statistical frameworks for assessing the autonomous experimentation process, the papers highlight the diverse range of applications of automation for advancing materials science.

Reyes, Liu and Vargas investigate the role of screening operations for use in autonomous data collection platforms in their manuscript entitled "Decision-Making Under Uncertainty for Multi-stage Pipeline: Simulation Studies to Benchmark Screening Strategies". Such pipelines are critical for evaluating candidate samples or materials in autonomous experimental platforms, but are generally loosely-defined, motivating the need for more statistically rigorous assessments of screening procedures. The authors present a generalizable framework for simulation and testing of screening pipelines, and demonstrate its efficacy in assessing various screening approaches for a given experimental problem definition.

Pandey et al., in their paper entitled "Machine learning interatomic potential for high-throughput screening of high-entropy alloys" have used machine learning to quickly predict interatomic potentials, enabling rapid prediction the mechanical properties in 4- and 5- element refractory high entropy alloys. There is a significant

[1] Materials Mechanics & Tribology Department, Sandia National Laboratories, Albuquerque, New Mexico, USA 87185

[2] Materials Science and Technology Division, US Naval Research Laboratory, Washington, DC, USA 20375

*Corresponding author: E-mail: apolon@sandia.gov

reduction in computation time using these ML techniques over more traditional DFT methods, and using these techniques is enabling researchers to determine candidate alloys with necessary properties such as ductility for further study.

In their manuscript, "Automatic detection of slag eye area based on a hue-saturation-value image segmentation algorithm", Wang et al. use image processing methods along with a change of color space from RGB to HSV in order to measure the size and shape of the slag eye that develops during inert gas bubbling in ladle refining of molten steel. Their method was robust to changes in image quality and camera position, so could be used in various configurations, and could be used as a critical step in the development of automated industrial metallurgy.

In "A Framework for Closed-Loop Optimization of An Automated Mechanical Serial-Sectioning System via Run-to-Run Control as Applied to a Robo-Met.3D", Gallegos-Patterson et al. have implemented a closed-loop controls framework to target material removal rates in a robotic serial-sectioning tool. Leveraging over a decade of historical data, the authors demonstrate the validity of their approach via both simulated experiments as well as several real-world tests. Such a closed-loop control approach is expected to improve the efficiency of the data collection process, as well as make the mechanical serial-sectioning tool more robust to the characterization of novel materials and complex sample geometries.

DeCost and collaborators describe applications of an autonomous experimental platform in "Towards automated design of corrosion resistant alloy coatings with an autonomous scanning droplet cell". This flexible tool enables the fabrication of new materials via electrodeposition and performance of corrosion assays, all incorporated into a fully-automated station operating within a synchrotron facility, thereby providing potential for x-ray flourescence and diffraction measurement in additional to optical and laser surface imaging. Using active learning approaches, the corrosion properties of complex multi-component alloy systems can be optimized much more rapidly than via more conventional approaches, even when multiple optimization targets are given to the system. The authors further demonstrate the challenges and opportunities in incorporating physics-based models into these types of autonomous platforms, leveraging theoretical understanding of materials to gain a better understanding of the results of so called "black-box" machine-learning optimizations.

To read or download any of the papers in this special topic, follow the URL: [insert link] to the table of contents for the August 2022 issue of JOM.

Acknowledgements Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. PGC would like to acknowledge support provided by the Naval Research Laboratory under the auspices of the Office of Naval Research.