LIFE2018 — Thoughts on Data Science from Case Western Reserve University

By: Roger H. French, YeongAe Heo, Jennifer L. W. Carter, Laura S. Bruckman and Rigoberto Advincula, Case Western Reserve University

In August, three of the authors (French, Heo and Carter) had the opportunity to attend LIFE2018, the Halliburton Landmark Innovation Forum and Expo. This event highlighted the importance of maintaining an innovative spirit to drive exploration in a field ripe with legacy data and infrastructure. It underlined the need for the oil and gas industry to embrace new methods of data science and analytics to solve the fast-moving technological challenges.

Immediately preceding LIFE2018 was the OpenEarth® Community Developers summit, which generated much excitement and discussion throughout the week. The OpenEarth Community, of which Halliburton is a founding member, is a free and open community of scientists, engineers and software developers working to speed up and lower the cost of digital innovation for the entire industry. It provides unique value for commercial solutions and production optimization to drive productivity.

For Case Western Reserve, the OpenEarth Community gave an insight into the challenges that the industry faces today, and the state-of-the-art solutions that are being engineered. This has helped our university to focus its research in order to best support the scientific endeavors of the industry.

Case Western Reserve is committed to translating innovations into reality. Departments across the university perform integrated and team research that enables breakthroughs to be made in a variety of fields, from high-performance materials to infrastructure. We have incorporated the recent rise of the *Digital Transformation,* centered on data science and statistical and machine learning, to push the boundaries of our fields — an approach that we feel will benefit the oil and gas industry.

DEVELOPING DATA-DRIVEN DECISION TOOLS

A risk-based approach will improve the structural integrity management of any high-risk and high-value facilities outfitted with structural health monitoring systems, which could include petrochemical facilities. Some of our recent research has focused on: developing data-driven decision tools; validating and updating models of the risks and performance of systems in real time; and assessing long-term changes to

systems during their service life — all of which have clear overlaps with the oil and gas industry.

SIMULATING DOWNHOLE CONDITIONS

We conduct research to test high pressure-high temperature and flow conditions. This synthesis and testing using flow chemistry methods allows us to better simulate downhole conditions (Figure 1).

At the same time, flow chemistry is used for the synthesis of new polymer architectures. Surface modification, anti-corrosion and anti-scaling coatings are of high interest for asset integrity. 3D printing for the oil and gas industry, using high-performance polymers, is one of our top research priorities.

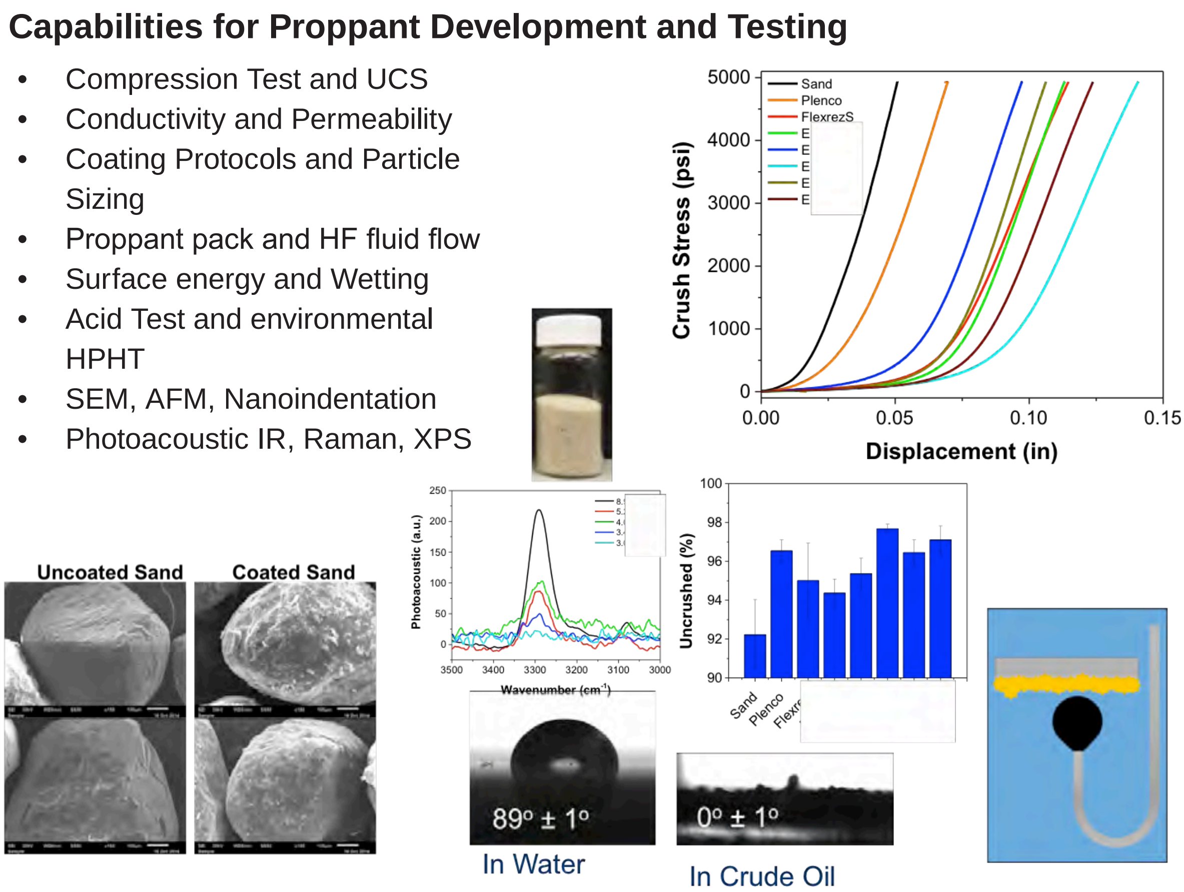

Figure 1 > Proppants testing and evaluation of resin-coated proppants, including the use of thermosetting polybenzoxazine chemistry and hydrogels. The goal is to enhance well productivity.

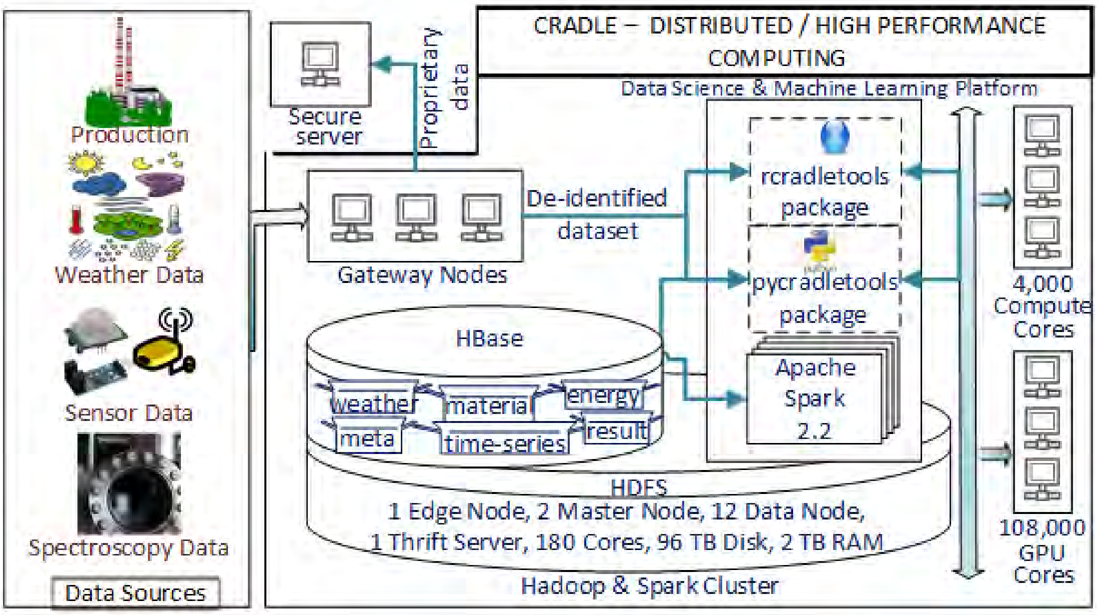

Figure 2 > CRADLE is a 96 Tb, 180 core, distributed computing environment based on Hadoop/HBase/Spark, embedded in Case Western Reserve's high-performance computing cluster.

MACHINE LEARNING AND BIG DATA IN REAL-WORLD APPLICATIONS

Systems deployed outdoors demand high levels of performance over long expected lifetimes and are affected by geographic location, climate and weather. Accelerated qualification testing can be used in the lab to design materials and systems for use in the real world.

We approach these problems with spatial-temporal, graph-based network modeling. Using distributed computing (based on Hadoop/HBase and Spark), datasets are ingested into our high-performance computing cluster (referred to as CRADLE, and shown in Figure 2) to study thousands of systems and their properties and behavior over decade-long time frames. Applying machine learning methods to images demonstrates the broad applicability of this research.

Lifetime prediction models relate the stress to a material performance response, like developing cracks or undergoing a color change, to predict lifetime under various conditions (French et al., 2015). These stress-response models are used to cross-correlate in-use responses to accelerated exposed results. Data science approaches have been used to integrate real-world studies that include spectroscopic, image analysis, and other analytical techniques to produce lifetime prediction models (French et al., 2015). For example, an optical profilometry technique (Figure 3) was developed to quantify automatically the number and depths of cracks using a machine learning algorithm (Huang et al., 2018).

Data science has also been used to integrate novel experimental approaches and explore the physics of microstructure evolution in interface-rich materials (Bruckman et al., 2013; Klinke et al., 2018). Data-enabled models require the collection and quantification of metrics of processing-structure-performance; as well as developing data-driven decision tools, methodologies for real-time validation and updating of models of system performance and risks, and frameworks for assessing long-term changes to a system during its service life (or possible service life extension) due to evolving operational and environmental conditions; and for developing in-service inspection and maintenance policies that enable the system to

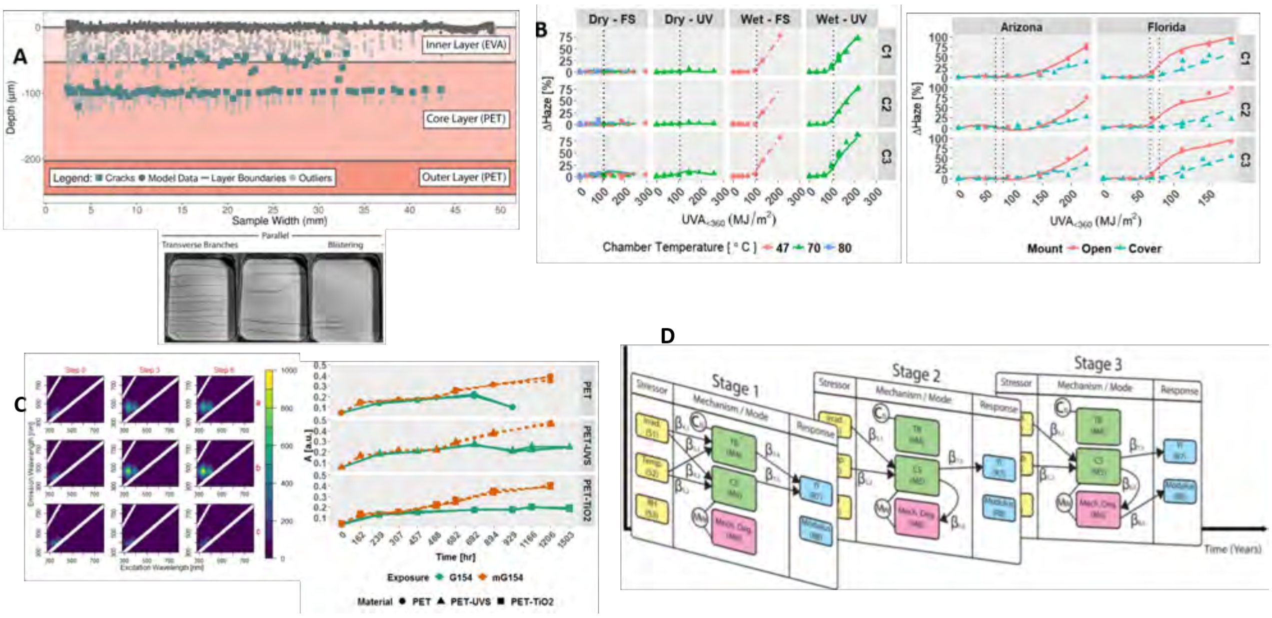

Figure 3 > A. Data science informed optical profilometry technique to determine crack formation and depth in photovoltaic back sheet materials. B. Multi-variate, multiple regression models of haze formation in poly (ethylene-terephthalate) (PET) under various exposure conditions. C. Parallel factor analysis of excitation-emission fluorescence of PET degradation to determine contribution degradation products. D. Network degradation pathway models (netSEM) for acrylic degradation.

adapt to those changes (Hufnagel et al., 2014; Ellis et al., 2015; Verma et al., 2017).

COLLABORATION IS VITAL

The LIFE2018 event illustrated unique opportunities to integrate novel technical solutions in the oil and gas E&P industry. Case Western Reserve can contribute significant value to driving solutions forward in this demanding field. In particular, the techniques we have developed for the disparate technologies presented here can be applied in new and exciting methods by collaborating with the engineers and scientists at Halliburton Landmark and beyond, contributing to the Digital Transformation.

The work carried out at Case Western Reserve demonstrates the university's ability to deliver impactful solutions. While these endeavors encourage and spark further research, they also provide opportunities for commercialization of products capable of directly impacting society. To find out more about the research being conducted at Case Western Reserve, visit our website.

To relive LIFE2018 and the OpenEarth Community Developers summit, watch the keynote speeches on iEnergy®.

REFERENCES

Bruckman, L.S., Wheeler, N.R., Ma, J., Wang, E., Wang, C.K., Chou, I., Sun, J. and French, R.H., 2013. Statistical and domain analytics applied to PV module lifetime and degradation science. IEEE Access, 1, pp.384-403.

Ellis, D. L., Carter, J. L. W., & Ferry, M. H. (2015). A statistical study of the effects of processing upon the creep properties of GRCop-84. Materials Science and Engineering: A, 640, 1–15. https://doi.org/10.1016/j.msea.2015.05.028

French, R. H., Podgornik, R., Peshek, T. J., Bruckman, L. S., Xu, Y., Wheeler, N. R., Gok, A., Hu, Y., Hossain, M.A., Gordon, D.A., Zhao, P., Sun, J., Zhang, G. Q. (2015). Degradation science: Mesoscopic evolution and temporal analytics of photovoltaic energy materials. Current Opinion in Solid State and Materials Science, 19(4), 212–226.

Huang, W. H., N. Wheeler, A. Klinke, Y. Xu, W. Du, A. Gok, D. Gordon, Y. Wang, J. Liu, A. Curran, J. Fada, X. Ma, J. Braid, L. Bruckman, R. French, netSEM: Network Structural Equation Modeling, 2018. https://CRAN.R-project.org/package=netSEM.

Hufnagel, A. C., Heo, Y. A., & Kang, T. H. K. (2014). Modeling Parameters for Reinforced Concrete Slab-Column Connections. Special Publication, 297(1-16).

Klinke, A.G., Gok, A., Ifeanyi, S.I. and Bruckman, L.S., 2018. A non-destructive method for crack quantification in photovoltaic backsheets under accelerated and real-world exposures. Polymer Degradation and Stability, 153, pp.244-254.

Verma, A.K., French, R.H. and Carter, J.L.W. (2017). Physics-Informed Network Models: a Data Science Approach to Metal Design. Integrating Materials and Manufacturing Innovation, 6(4), 279–287. https://doi.org/10.1007/s40192-017-0104-5

AUTHORS

Prof. Roger H. French is the director of the SDLE Research Center, faculty director of the Applied Data Science program, and a professor within the Department of Materials Science and Engineering. His work focuses on degradation science: integrating real world performance datasets with lab-based experimental datasets to better characterize the durability of various materials and systems used in demanding applications and large-scale global systems, in a variety of energy and materials applications. He has been developing spatio-temporal GIS data analytics on the long term performance of globally distributed systems.

Prof. Carter is an assistant professor within the Department of Materials Science and Engineering and has aligned her work towards the applications of data analytics and machine learning to structural materials design. In 2016, Prof. Carter was awarded an NSF-CAREER grant for her work in this interdisciplinary field, spanning network modeling and computer vision to quantify microstructural evolution.

Prof. Heo, an assistant professor in the Department of Civil Engineering, supervises her SAFE C.I.T.I. (Community and InTegrated Infrastructure) Research group aimed to advance knowledge of damage propagation mechanisms and risk-based structural integrity management of complex civil infrastructure systems subjected to multiple extreme natural and man-made hazards. Prof. Heo's early research focused on nonlinear dynamic performance evaluation for structural materials and systems, damage modeling, and probabilistic risk and loss modeling of various hazards (earthquakes, winds, waves, explosions, fires, and impact loads) for multistory buildings and offshore structures. Prof. Heo is the recipient of the 2017 Early-Career Research Fellowship of the NASEM Gulf Research Program which sponsors her research on the resilience of oil and gas systems and coastal communities.

Prof. Laura Bruckman is an associate research professor in the SDLE Research Center. Her research interests incorporate data science and analytics to describe material degradation under in-use conditions and accelerated exposures. The material systems she investigates include photovoltaics, exterior coatings, and building materials. Her work focuses on employing a data science approach to integrate laboratory and real-world studies that include spectroscopic, image analysis, and other analytical techniques to develop lifetime prediction models. Prof. Bruckman developed an optical profilometry technique used to automatically quantify the number of cracks and depth of cracks using a machine learning algorithm.

Prof. Rigoberto Advincula, Professor of Macromolecular Science and Engineering, is investigating high-performance polymers and additives for upstream and especially challenging environments. He is also the Director of PETRO Case, a laboratory dedicated to advanced polymer materials development in the industry.

DISCLAIMER