**46th Annual Review of Progress
in Quantitative Nondestructive Evaluation
QNDE2019
July 14-19, 2019, Portland, OR, USA**

QNDE2019-####

ON ADOPTION OF CLOUD COMPUTING AND MACHINE LEARNING FOR THE INTEPRETATION OF SUBSURFACE MEASUREMENTS DATA

Smaine Zeroug
Schlumberger-Doll Research, Cambridge, MA

ABSTRACT

The significant progress being made in HPC cloud infrastructure and machine learning has brought about plenty of opportunities to industrial Research & Development teams to enhance their delivery prowess with the aim of supporting business growth around bringing efficiency through automation and effectiveness through data insight and leverage. But the challenges to take advantage of the technological progress also abound and encompass both organizational and technical aspects. This is especially true for R&D teams in the oil & gas services industry confronted with extracting value from subsurface measurements data that are sparse, rich, and often non- or ill-labeled. In this talk, we will share our experience in 'navigating the digital transformation' with focus on the technical aspects of adopting cloud compute infrastructure as well as in developing data-driven approaches for the interpretation of complex signals acquired by devices lowered in boreholes or at the surface of an oilfield.

On cloud-based compute, we conduct modeling and inversion tasks of downhole measurements data on the cloud that have reduced the compute time from tens of days to a few hours or from several hours to a few minutes. This has enabled us to tackle computationally-expensive tasks in configurations that address realistic 3D geometries with material heterogeneities and at scales large enough to appropriately explore the parametric sensitivities inherent in the measurements modeling and inversions; tasks that we couldn't tackle hitherto with our local powerful (1000+ core) cluster.

On leveraging machine learning (ML), our objectives encompass attaining efficiencies through automating algorithmic interpretations to support real-time and on-site semi-autonomous decision making, improving their accuracy and reducing their uncertainty through integration of multi-physics data and prior knowledge, and gaining insight into the physics of the problem where deterministic approaches have proven to be limited. Beyond the critical tasks of data curation, the key data challenges encountered include their sparsity (a sampling along a rectilinear well of a 3D subsurface parametric volume), their richness (heterogeneity in character or multi-parameter sensitivity), and lack of labeling (proper interpretation). We are finding that these challenges can be mitigated, and for certain problems adequately addressed, with the injection of physical modeling and domain knowledge in supervised and unsupervised ML and deep learning (DL) tasks.

Domain knowledge and associated processing tools can be used in the data feature engineering task (which often is a first step in dimensionality reduction) as well as in feature labeling. Physical modeling can generate synthetic data that can be used to enrich the field (real) data to conduct supervised or semi-supervised learning and can also be employed via transfer learning with first sequential training on synthetic data followed by training on real data. In unsupervised learning, where the aim is to gain insight on the data structure, domain knowledge is critical to interpret the resulting clusters. Synthetic data can also be used in the latter task to tune and learn how best to optimize the ML/DL neural network characteristics and go about interpreting the resulting clusters. We will illustrate these aspects using various ML and DL schemes applied to problems related to the interpretation of complex borehole measurements (acoustic, electromagnetic, etc) acquired in open and cased holes as well as, at a larger scale, for subsurface geological modeling.

Keywords: Cloud computing, machine learning, artificial intelligence, deep learning, automation, physics modeling, domain knowledge, oil and gas geophysical applications.

ACKNOWLEDGEMENTS

In this overview, the author will feature results from research conducted by many of his colleagues within the applied math and data analytics research department he manages at Schlumberger-Doll Research: S. Bose, P. Tilke, V. Polyakov, L. Liang, L. Venkataramanan, D. Omeragic, L. Zhu, Z. Bayraktar, V.-M. Gkortsas, T. Zhang, Y. Liu, S. Omar, and M. Thiel.

