Poster #21-55

Cloud Based Cyber Infrastructure to Enable Process Understanding

Roelof Versteeg[1*], Baptiste Dafflon[2], Haruko Wainwright[2], Nicola Falco[2], Doug Johnson[1], Erek Alper[1], Satish Karra[3], Ye Zhang[4], Peter Lichtner[5], Reza Soltanian[6], Kelly Wrighton[7], Chris Henry[8], Haiyan Zhou[1], Rebecca Rubinstein[1]

[1] Subsurface Insights, Hanover, NH
[2] Lawrence Berkeley National Laboratory, Berkeley, CA
[3] Los Alamos National Laboratory, Los Alamos, NM
[4] University of Wyoming, Laramie, WY
[5] OFM Research, Santa Fe, NM
[6] University of Cincinnati, Cincinnati, OH
[7] Colorado State University, Fort Collins, CO
[8] Argonne National Laboratory, Lemont, IL

Contact: roelof.versteeg@subsurfaceinsights.com

BER Program: SBR
Project: SBIR
Project Website: www.subsurfaceinsights.com

It is increasingly recognized that for both scientific, operational and regulatory reasons an in depth, near real time understanding of subsurface processes at thousands of sites will be required.

Obtaining such an understanding will require an enabling cyberinfrastructure which can support data collection, ingestion, management and analysis as well as collaborative research and result delivery and information use at scale. Over the last several years, through both in house funding and under multiple DOE SBIR awards and funding Subsurface Insights has developed a cloud based cyber infrastructure and associated autonomous hardware for process understanding which increasingly enables this large scale understanding.

This cyber infrastructure (PAF – Predictive Assimilation Framework) leverages open source scientific software components developed by DOE scientists (PFLOTRAN, E4D, PyFlotran), and university scientists (ODM2, PEST, Landlab). It also takes full advantage of open source frontend libraries for graphing (Plotly, D3) and user interfaces (JQuery, React) and backend tools for data analysis (e.g. Scikitlearn and R). As part of this development Subsurface Insights has been contributing to the development of several of these open source codes.

PAF has been developed and demonstrated using data from both DOE funded projects and data from other research institutions and the private sector. PAF currently can ingest geochemical, geophysical, hydrological and remote sensed data and uses the PFLOTRAN reactive transport model for modeling and analysis. Both data and capabilities within PAF can be accessed through browser and mobile interfaces and software APIs, allowing for easy integration with third party computational capabilities. An example of this is our ability to parameterize and run PFLOTRAN models both through a browser, a mobile phone and through an API.

Subsurface Insights was recently funded under a new DOE Phase I SBIR to integrate PAF with the DOE developed KBASE platform to couple microbiological and field data for enhanced subsurface process understanding. In our poster we will give an overview of the overall architecture of PAF and of the new comprehensive datamodel developed by Subsurface Insights (ODMX) which can accommodate multiple key geoscience datasets. Meeting attendees will be able to get guest accounts allowing them to run models and perform analyses.