Machine Learning and Data Analytics Overview

DOE Upstream Oil and Gas Workshop

Grant Bromhal
Senior Fellow, Geological and Environmental Sciences
National Energy Technology Laboratory

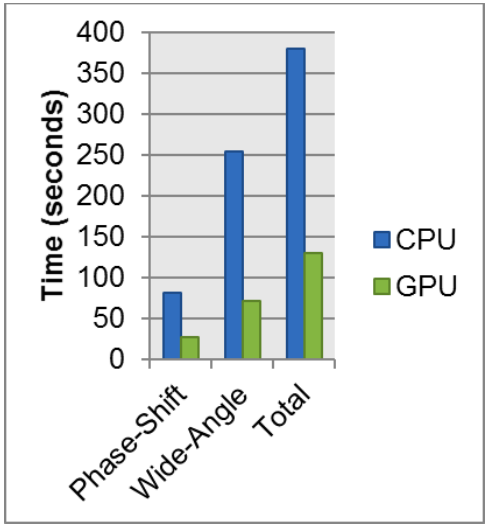
February 14, 2018

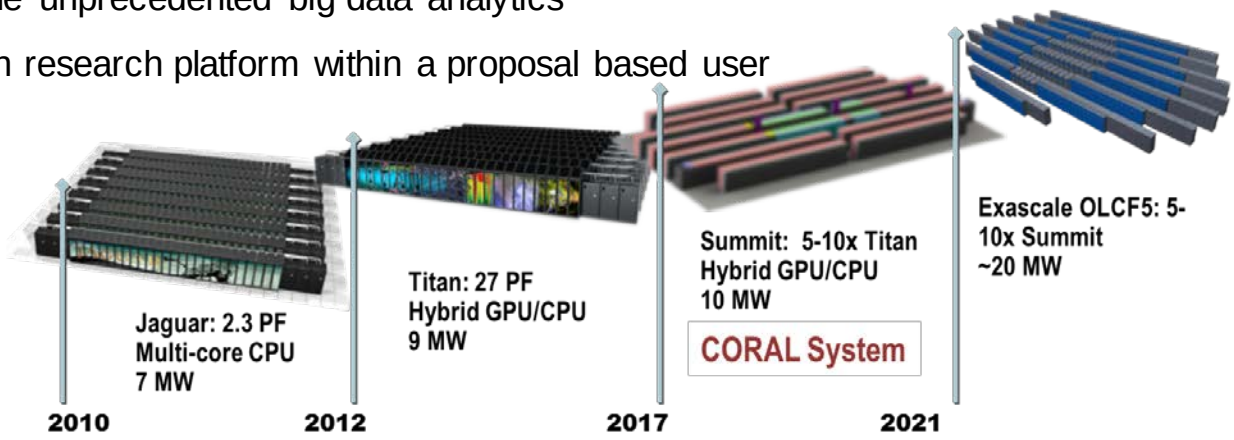


U.S. DEPARTMENT OF ENERGY | Fossil Energy
OFFICE OF OIL & NATURAL GAS

HIGH PERFORMANCE COMPUTING

- High performance computing systems and applications are key DOE lab capabilities for big data analytics and machine learning
  - GPU accelerators generate very rapid calculations for many problems, including subsurface flow predictions
  - HPC resources within Oak Ridge Leadership Computing Facility (OLCF) and at other labs have been evolving into “big data” and AI supercomputers
    - Titan is a hybrid GPU/CPU architectures that allow for enhanced AI capabilities
    - Summit and next-gen, exascale HPC will deliver very powerful, “fat” nodes that enable unprecedented big data analytics
    - OLCF is an open research platform within a proposal based user program



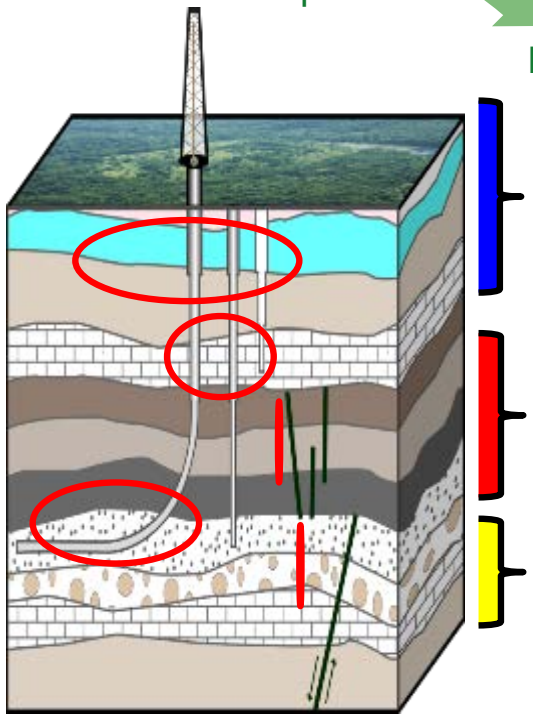


DOE R&D IN MACHINE LEARNING

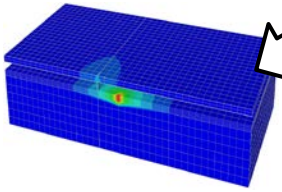
- DOE's Office of Science recently held a workshop on Machine Learning
- Core R&D ongoing in multiscale modeling and image processing
  - Identifying correlations in data
  - Reducing dimensionality
  - Improving sampling techniques
  - Automating characterization methods
- Improving robustness and interpretability of ML approaches
- Uncertainty quantification
- Quantum computing

NATIONAL RISK ASSESSMENT PARTNERSHIP USES ROMS TO PROBE UNCERTAINTY IN THE SYSTEM

A. Divide system into discrete components

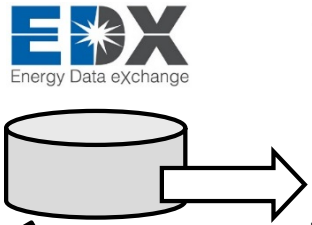


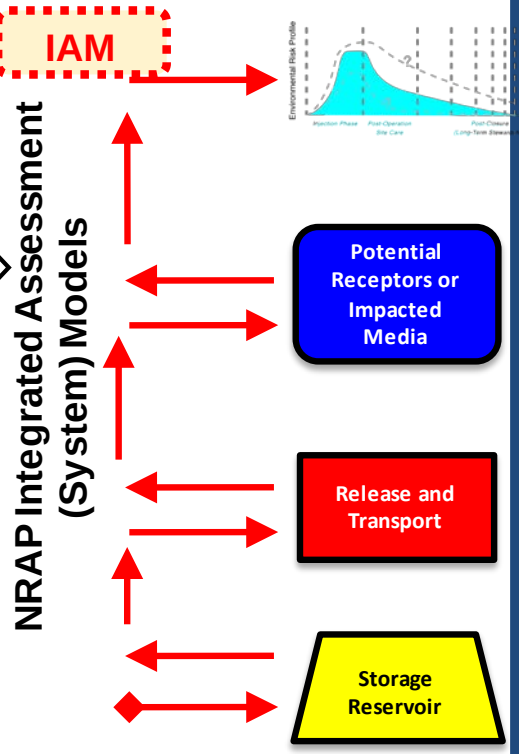
B. Develop detailed component models that are validated against lab/field data



C. Develop reduced-order models (ROMs) that rapidly reproduce component model predictions

D. Link ROMs via integrated assessment models (IAMs) to predict system performance

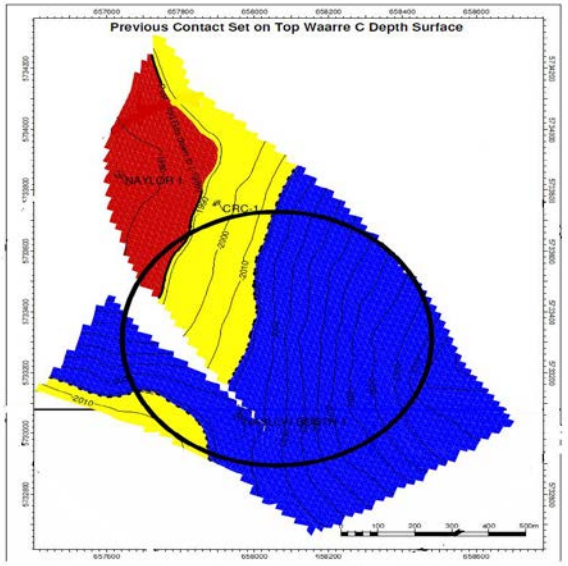


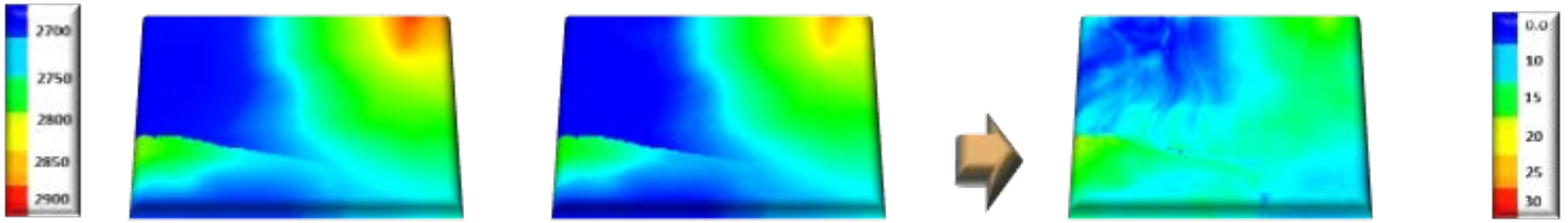


E. Exercise whole system model to explore risk performance

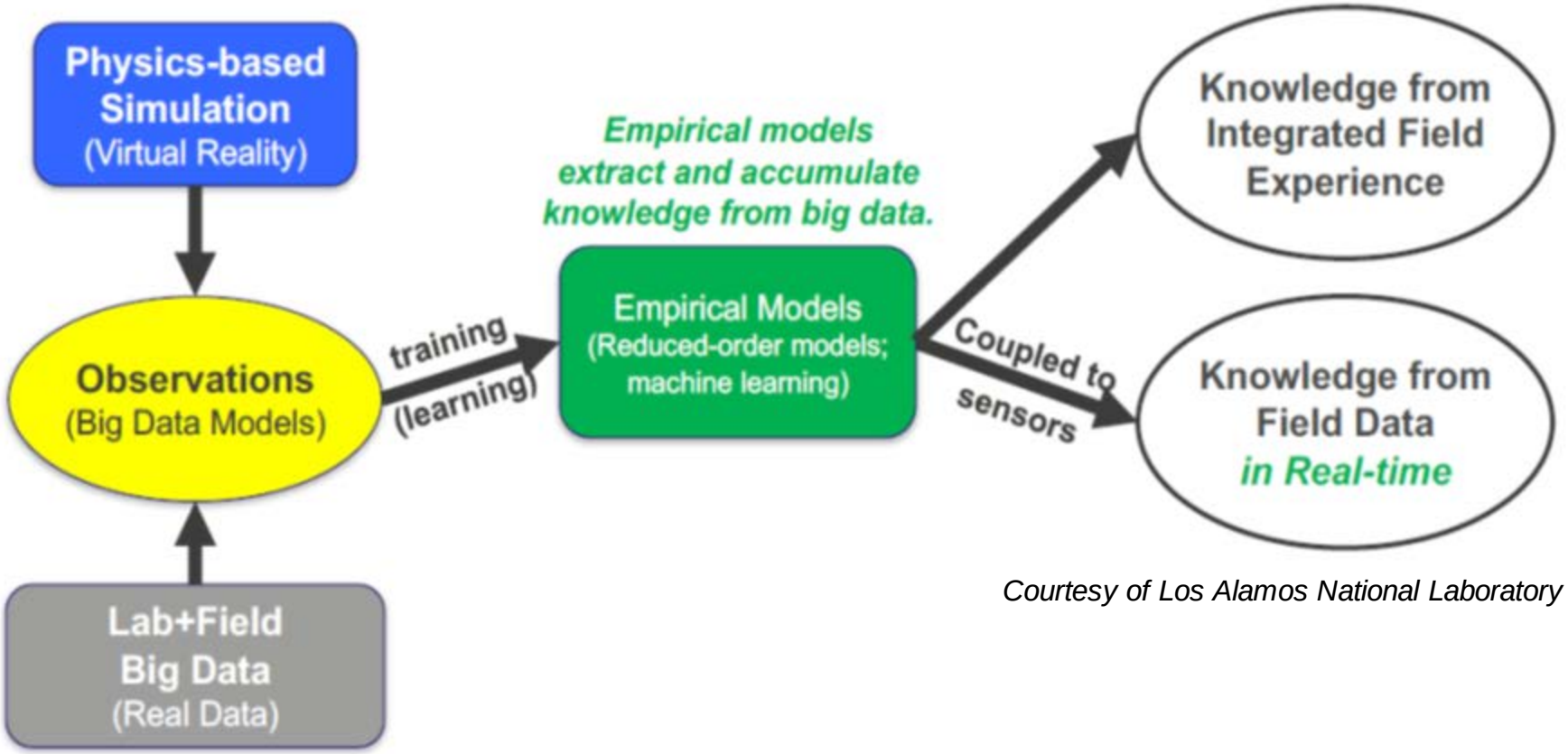
SURROGATE RESERVOIR MODELING

- **Utilizes data mining and neural networks to predict reservoir behavior**
  - Pressures
  - Saturations
- **Can be trained with a relatively small number of reservoir simulations**
- **Allows for spatial heterogeneity and operational variability**
- **Shrinks time for reservoir simulation from days and hours to minutes and seconds**





ROMS/BIG DATA ANALYTICS WILL ENABLE REAL-TIME UNDERSTANDING OF RESERVOIRS



Courtesy of Los Alamos National Laboratory

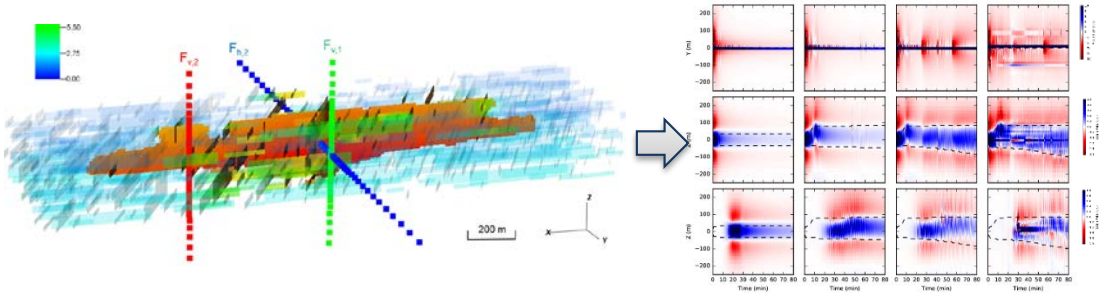
MACHINE LEARNING AND DISTRIBUTED SENSING

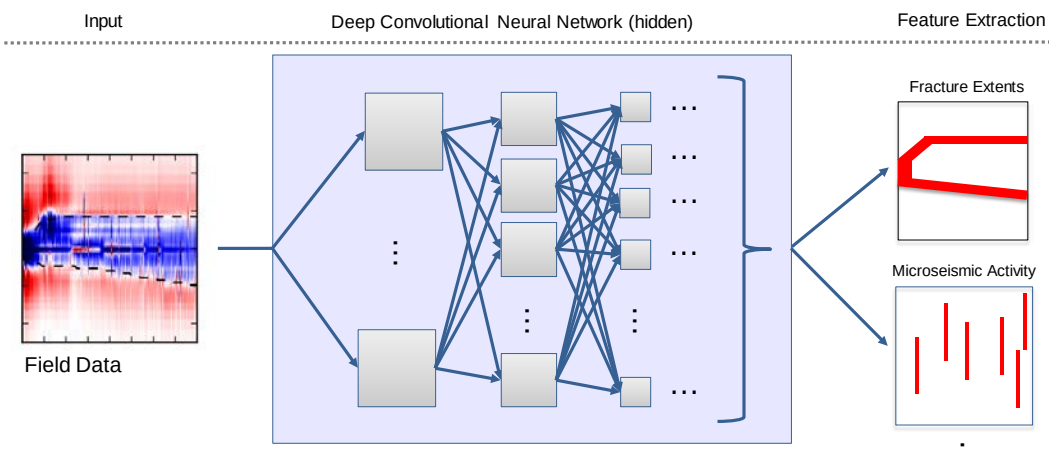
FIBER-OPTIC SYSTEMS WILL CREATE MASSIVE DATA STREAMS

- Simulate subsurface processes using the GEOS code and record synthetic DAS measurements
- Design and train a deep convolutional neural network (DCNN) to identify features
- Optimize the DCNN and test it on field data

Capability could be used to make pumping schedule changes in real time during stimulation to target pay zone more effectively

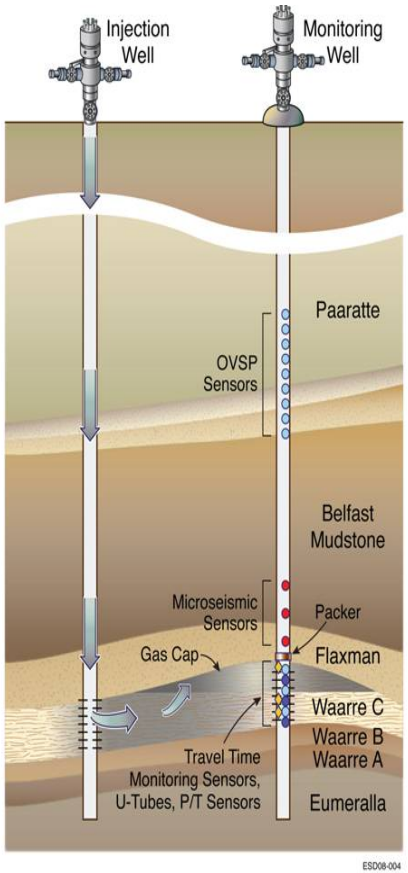
- Dark fiber also contains a potential wealth of data that is being mined (LBNL)





Courtesy of Lawrence Livermore National Laboratory

INTELLIGENT WELL TECHNOLOGY



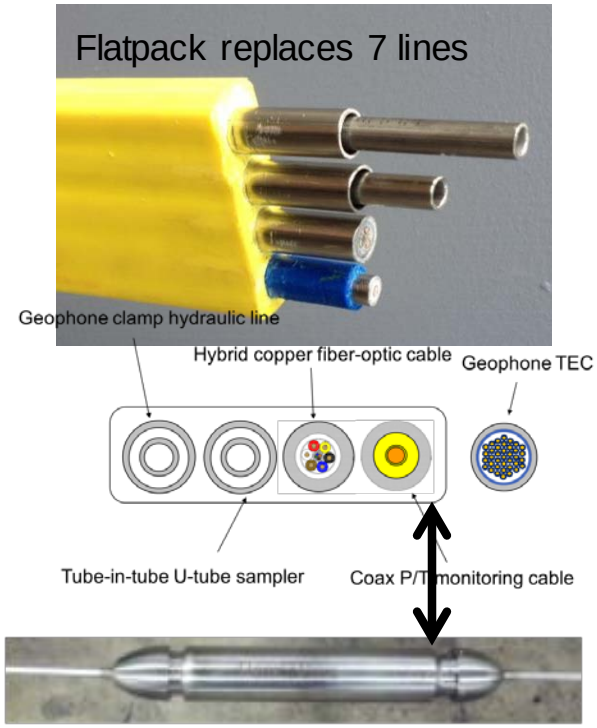
Monitoring CO2 Storage at 2 km (Otw ay)

Decades of experience developing/deploying intelligent well technology

Example: Modular Borehole Monitoring

- Integrated, robust suite of well-based tools
- Configurable for application, deployment method, geologic environment
- Combine deployments from casing, tubing, wireline, behind casing
- Continue development and deployment of inexpensive fiber optic distributed temperature sensing (DTS), distributed acoustic sensing (DAS), and distributed strain sensing (DSS) to identify local anomalies.

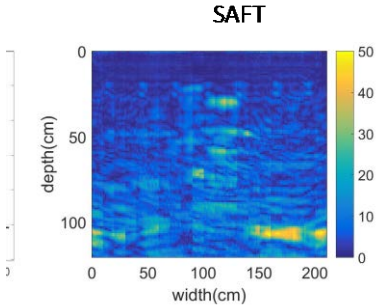
Courtesy of Lawrence Berkeley National Lab



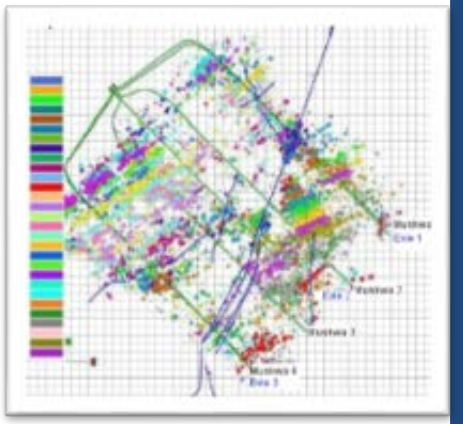
Monitor seismicity with DAS and clamp-on geophone with welded geophone line

OTHER EXAMPLES OF O&G USE

- Improved ultrasound reconstruction algorithm development for Oil, Gas, and Geothermal wells using model based iterative methods (ORNL)
- Advance novel deep and machine learning techniques to develop methods for predicting unconventional oil and gas production characteristics (SNL)
- Detecting anomalies in petabytes of geophysical and geochemical data (SNL)
- Data-driven approaches to addressing induced seismicity (INL, LANL, LBNL, LLNL, NETL, PNNL, SNL)
- Advanced approaches to understand shale reservoir performance by analyzing production data (INL, NETL, LANL)
- Fusing pore and reservoir scale information in models (LBNL)
- Toolset Development for Massive Datasets in Immersive Environments (INL, PNNL)
- Real-time subsurface imaging, combining forward coupled simulation with process feedback and control (PNNL)
- Deriving empirical models for complex fractured flow systems (LANL)



Courtesy of Oak Ridge National Lab



Courtesy of Sandia National Lab



Computer Assisted Virtual Environment (CAVE) system displaying and earth model

Courtesy of Idaho National Lab

KEY POINTS

- DOE and labs have significant capabilities in high performance computing, data analytics, and machine learning
  - Application to geoscience applications is in varying stages
- National labs have the tools needed to handle large oil and gas data sets
- There is a rich set up opportunities to apply machine learning to oil and gas applications
  - Shale behavior/performance
  - Reservoir uncertainty quantification, enhanced production
  - Intelligent well technology
  - Seismicity
  - Others
- Machine learning and distributed sensing together have significant potential to improve recovery from both unconventional and conventional resources
- A focus moving forward will be on machine learning approaches can enable real-time reservoir management/control

THANK YOU!

- Questions?