Model-Assisted Decision Analyses Related to a Chromium Plume at Los Alamos National Laboratory

Velimir V. Vesselinov, Dan O'Malley, Danny Katzman

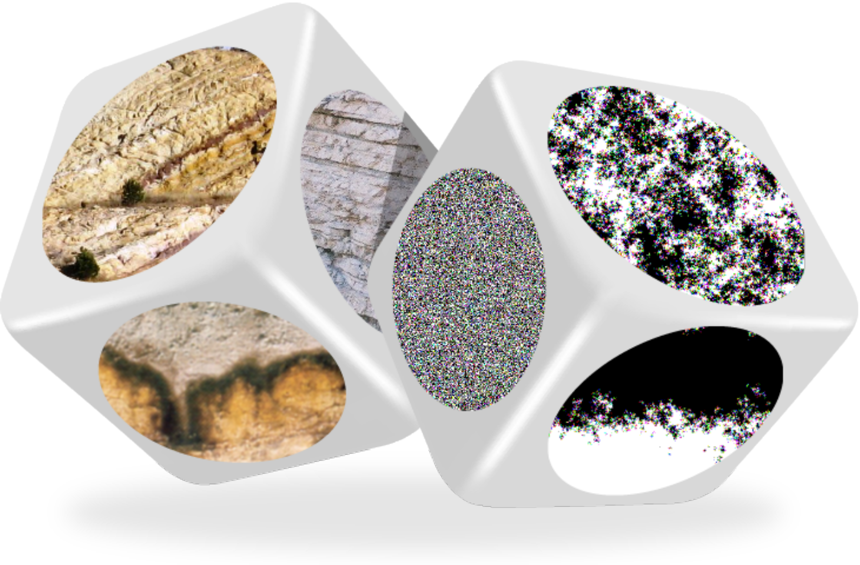
Computational Earth Science, Los Alamos National Laboratory

Waste Management Symposium, March 19, 2015
Unclassified: LA-UR-15-21965

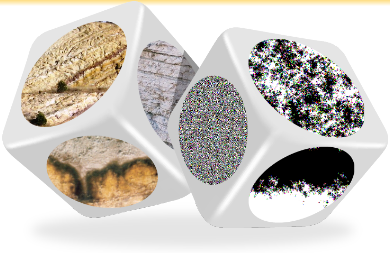
- **Uncertainties in Environmental Management Problems**
- **BIG-DT: Bayesian-Information-Gap Decision Theory for Uncertainty Quantification & Decision Analysis**
- **Los Alamos National Laboratory (LANL) Chromium Problem**
- **ZEM workflow: Data $\Leftrightarrow$ Models $\Leftrightarrow$ Decisions**

Probabilistic Uncertainty

Non-probabilistic Uncertainty



Uncertainties in Environmental Management Problems



- Probabilistic methods work very well for dice-rolling predictions
- However, many environmental management uncertainties cannot be represented probabilistically
- For example, geologic heterogeneity is typically unknown (**left die**)
- We also do not know which model of heterogeneity is representative (**right die**), but we must choose a single representative model conditioned on the available data
- We also do not know what all the sides of the dice look like, and how many sides there are
- Therefore, we cannot **enumerate all possible outcomes**
- All these issues make purely probabilistic (Bayesian) analyses **flawed** for many environmental-management problems

- Model uncertainties (conceptualization and model implementation; typically **non-probabilistic**)
- Parameter uncertainties (model parameters)
- Data uncertainties (measurement errors)
- Uncertainties in the performance of the engineered environmental management system (remediation, waste management, etc.; typically **non-probabilistic**)
- All of these uncertainties can have both:
  - **probabilistic** components, and
  - **non-probabilistic** components
- How to address these uncertainties?
  - Recently we have developed a novel methodology and advanced computational tools that can address **probabilistic** and **non-probabilistic** uncertainties
  - **BIG-DT**: **B**ayesian-**I**nformation **G**ap **D**ecision **T**heory
  - O'Malley & Vesselinov 2014 SIAM UQ.

Bayesian problem

Bayes' theorem is mathematically rigorous, but its application in science and engineering is not always rigorous.

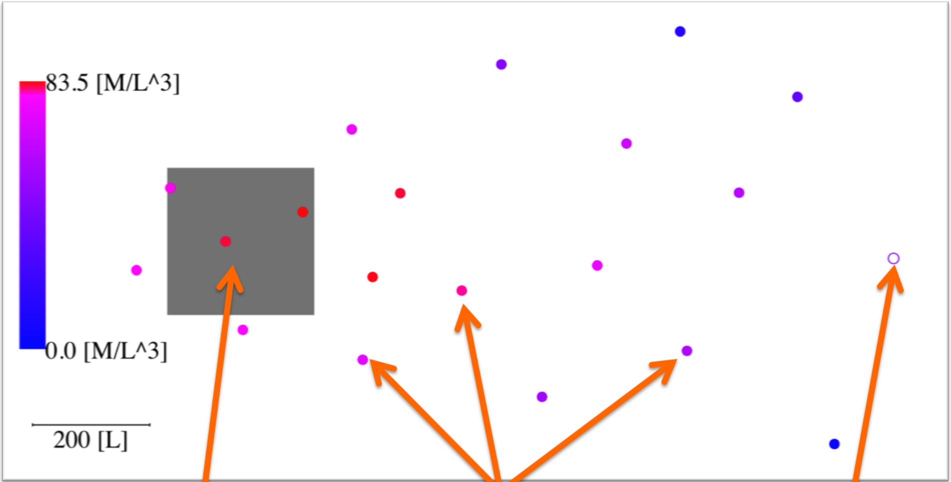
There are two reasons for this:

- We can enumerate the possible outcomes of **dice-rolling**, but not all the possible outcomes for **real-world engineering problems**.
- We can precisely determine conditional probabilities for **coin-tossing**, but substantial uncertainty surrounds the conditional probabilities for **real-world engineering problems**.

Note from R.A. Fisher to A.C. Aitken

From this point of view it is almost unfortunate that a group of cases has been found in which inductive inference ***may properly be expressed in terms of probability****, using the fiducial mode of argument; for this has tempted some mathematicians, and will, I fear, tempt more, to imagine that this type of argument is more widely applicable than is really the case, and to avoid enlarging their imaginations sufficiently to grasp* ***the cases where no probability statement is adequate****.*

BIG-DT Contaminant remediation problem: Scenario 1

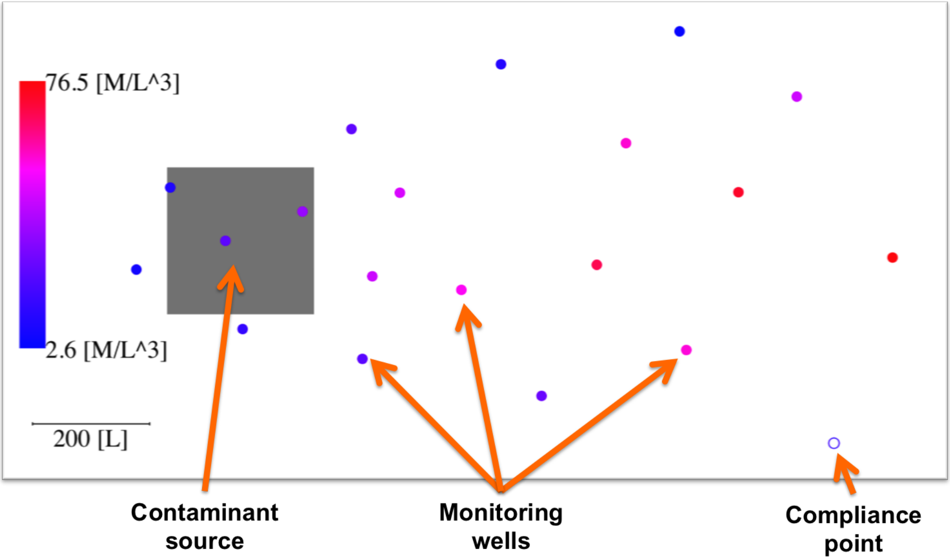


Contaminant source

Monitoring wells

Compliance point

BIG-DT Contaminant remediation problem: Scenario 2

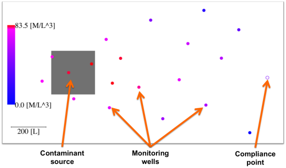


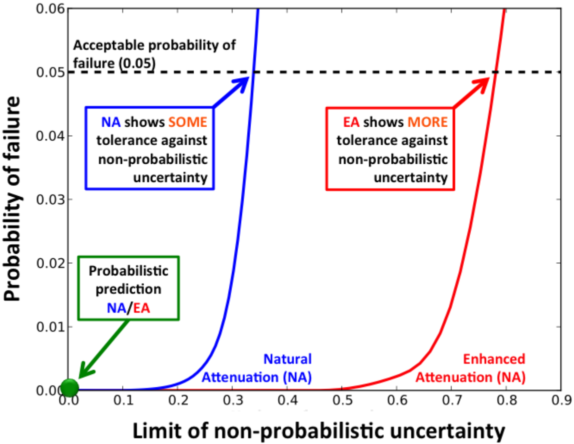
BIG-DT Contaminant remediation problem:

- **Known**:
  - 10 annual concentration observations at 19 wells (190 data items)
  - Location of compliance point
- **Estimated (probabilistic uncertainties)**:
  - location, size, contaminant mass flux at the source source
  - aquifer flow properties (groundwater flow direction, magnitude, etc.)
  - aquifer transport properties (porosity, dispersivity, etc.)
- **Unknown (non-probabilistic uncertainties)**:
  - geochemical reaction rate (natural/enhanced)
  - contaminant dispersion mechanism (Fickian or non-Fickian)
- Limit of non-probabilistic uncertainty (as defined in the Information Gap Decision Theory) applied to represent the **unknowns (non-probabilistic uncertainties)**

BIG-DT Contaminant remediation problem: Scenario 1

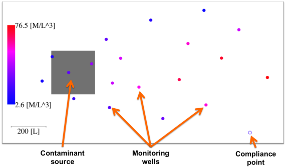
- To Act or Not to Act? That is the Question.
  - Act = Perform Enhanced Attenuation (EA)
  - Not to Act = Natural Attenuation (NA)
- **To Act is the Answer**

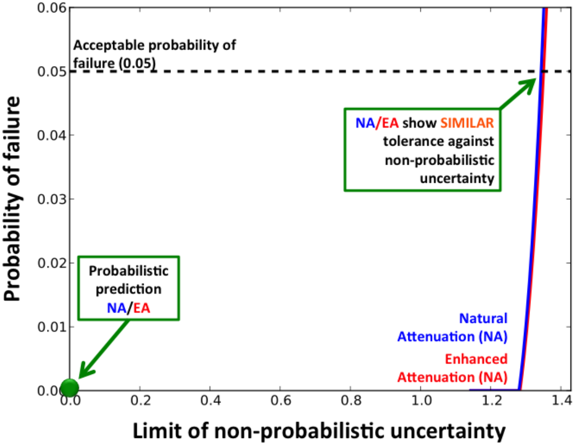




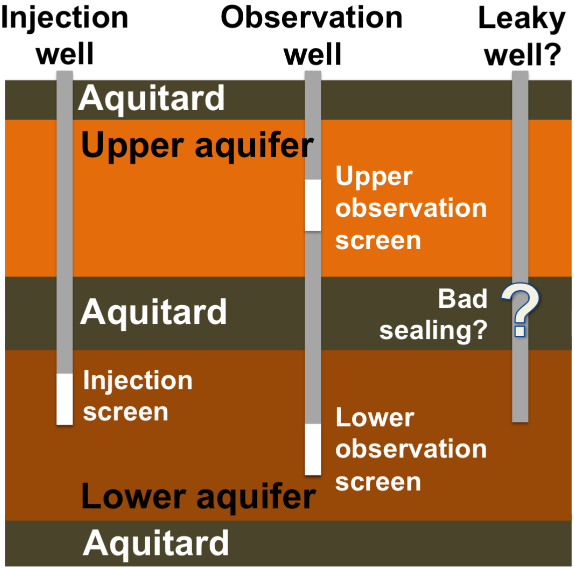
BIG-DT Contaminant remediation problem: Scenario 2

- To Act or Not to Act? That is the Question.
  - Act = Perform Enhanced Attenuation (EA)
  - Not to Act = Natural Attenuation (NA)
- **Not To Act is the Answer**



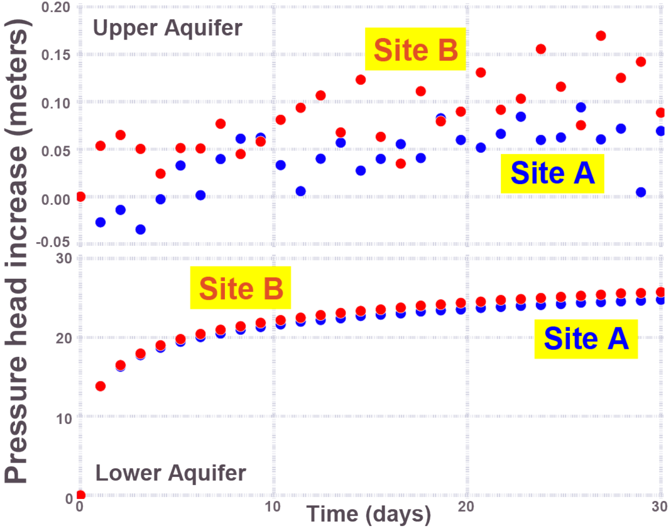


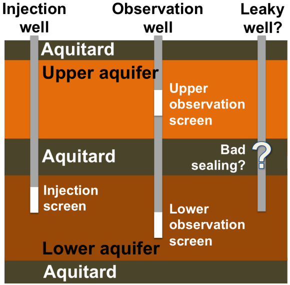
BIG-DT Geologic Carbon Sequestration problem: Problem Setup



BIG-DT Geologic Carbon Sequestration (GCS) problem: Data

Pressure data collected during injection tests at two potential sites

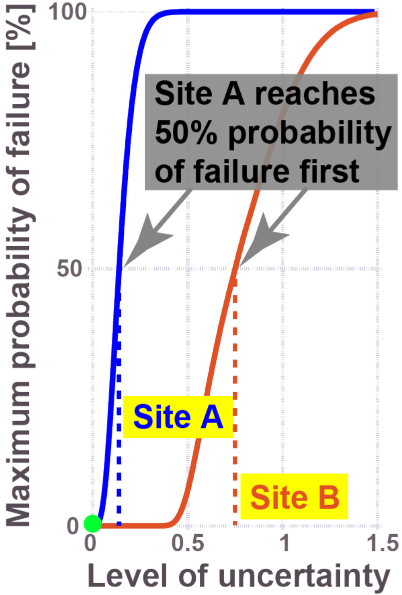


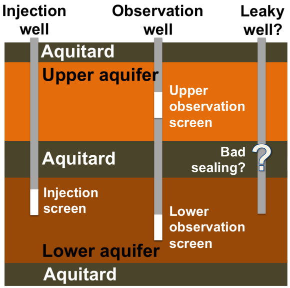


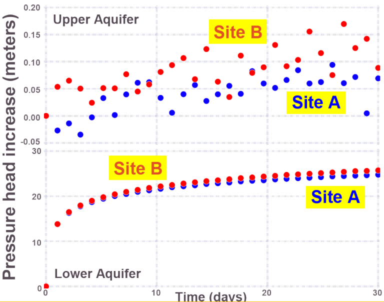
BIG-DT Geologic Carbon Sequestration (GCS) problem: Analysis

- Choose between two sites (A and B) with different geologic properties
- Pressure data collected during injection tests at two sites are applied to estimate the location and properties of the unknown leaky well
- **Residuals** are defined to represent the mismatch (discrepancy) between model predictions and field observations of the pressure in the upper and lower aquifer during the injection tests
- Nominally, **the residuals are assumed to be representative of a Gaussian white noise**
- This assumes that the selected model is representative of the actual site conditions
- Actually, we do not know if this is the correct model; as a result, **the residuals can be correlated**
- Information Gap analysis deals with **uncertainty in the distribution of the residuals**
- In this way, **non-probabilistic uncertainties** (unknowns) in the physics model representing the site conditions are captured

BIG-DT Geologic Carbon Sequestration (GCS) problem: Decision

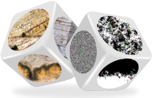




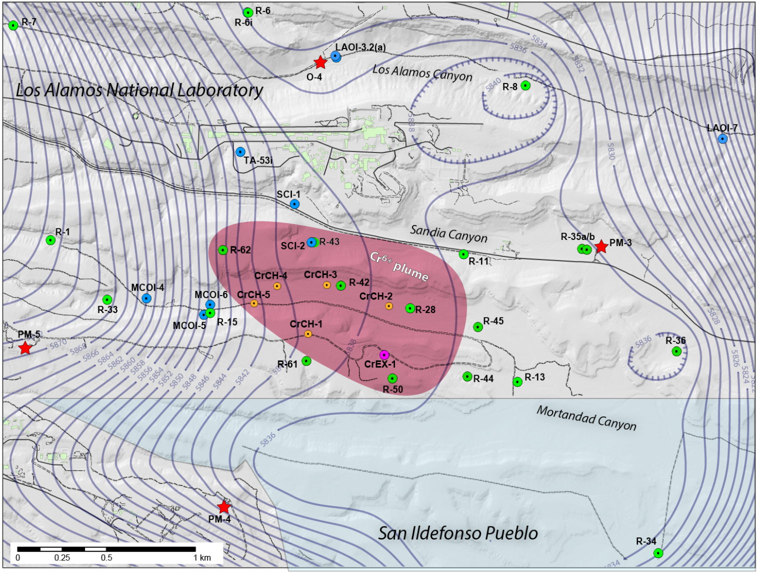


BIG-DT: Highlights

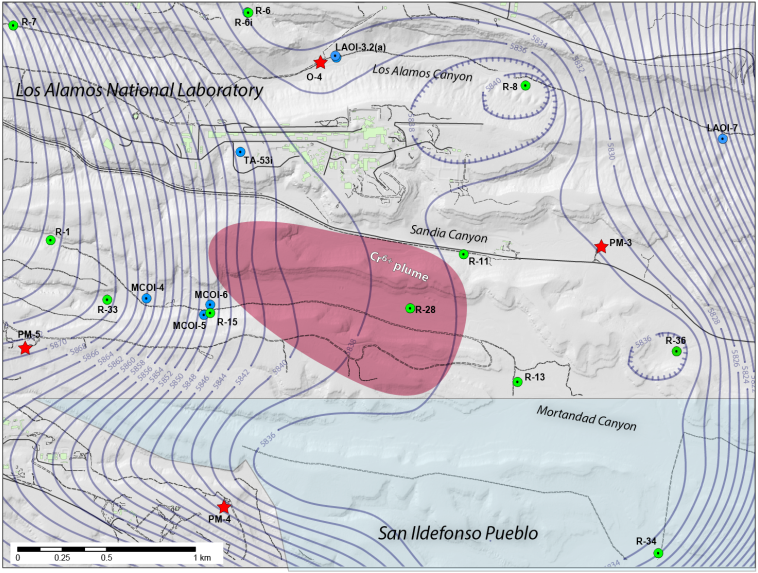
- By combining Bayesian analysis with Information-Gap analysis, **BIG-DT**
  - Circumvents the shortcomings of Bayesian analysis
  - Accounts for unknowns (non-probabilistic uncertainties)
  - Provides scientifically-defensible decisions
- Theory is already published (O'Malley & Vesselinov 2014 SIAM UQ)
- High-Performance Computational (HPC) framework for **BIG-DT** analyses has been already developed and tested (http://mads.lanl.gov)
- A series of synthetic problems have been solved
- Currently, we are applying **BIG-DT** for the LANL Chromium site



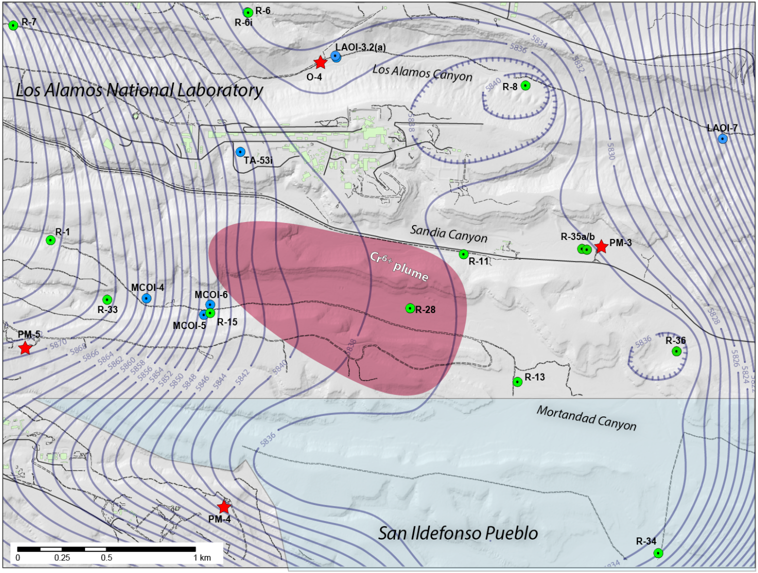
LANL Chromium site (2015)



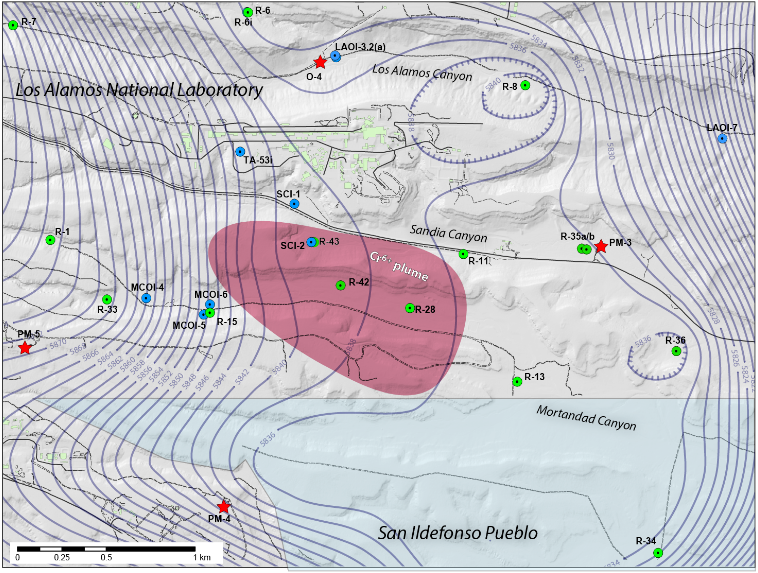
LANL Chromium site (2005)



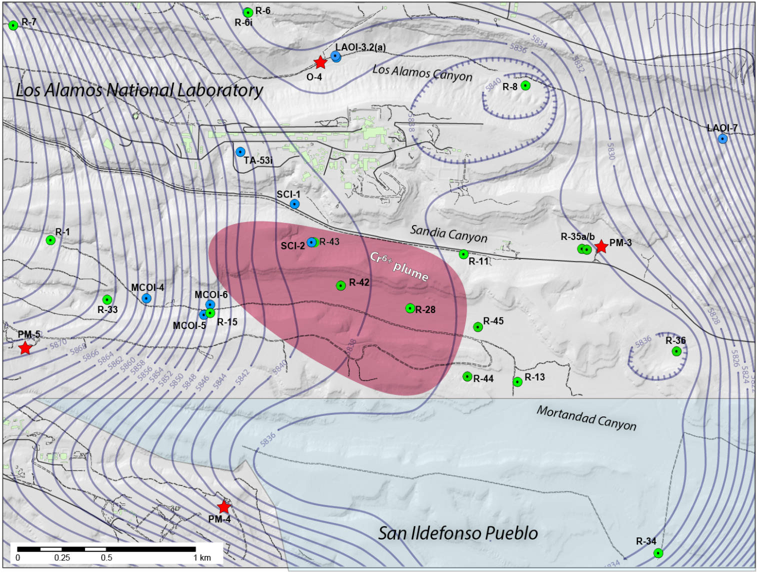
LANL Chromium site (~2006)

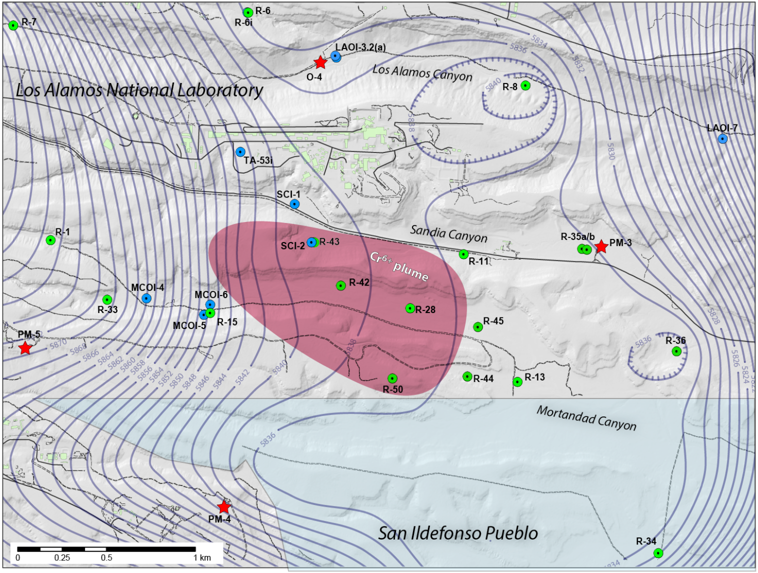


LANL Chromium site (~2007)



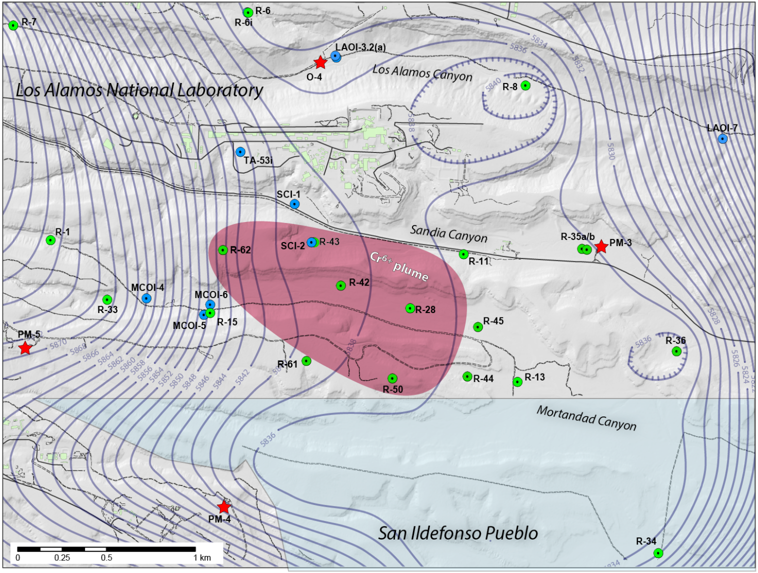
LANL Chromium site (~2008)



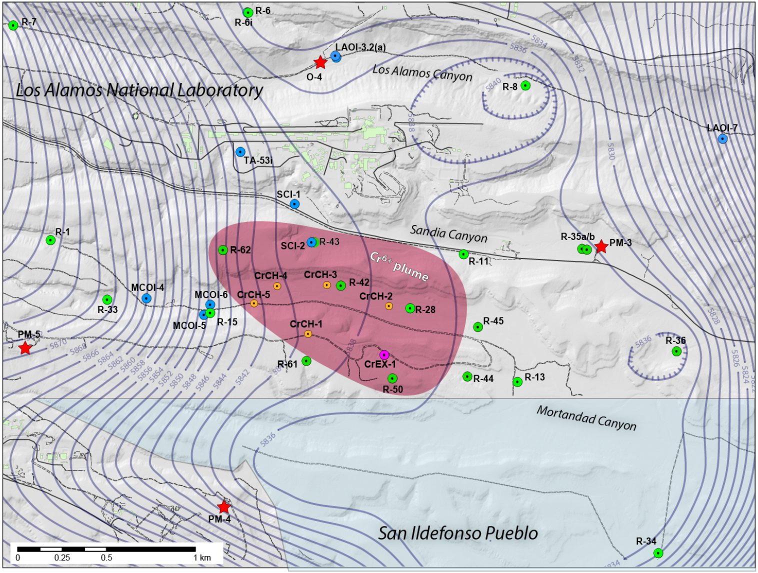


Los Alamos National Laboratory
Los Alamos Canyon
Sandia Canyon
Mortandad Canyon
San Ildefonso Pueblo
Cr6+ plume
R-7
R-6
R-6i
LAOI-3.2(a)
O-4
R-8
LAOI-7
TA-53i
SCI-1
SCI-2
R-43
R-11
R-35a/b
PM-3
R-1
MCOI-4
MCOI-6
MCOI-5
R-15
R-33
R-42
R-28
R-45
PM-5
R-36
R-50
R-44
R-13
PM-4
R-34
0
0.25
0.5
1 km

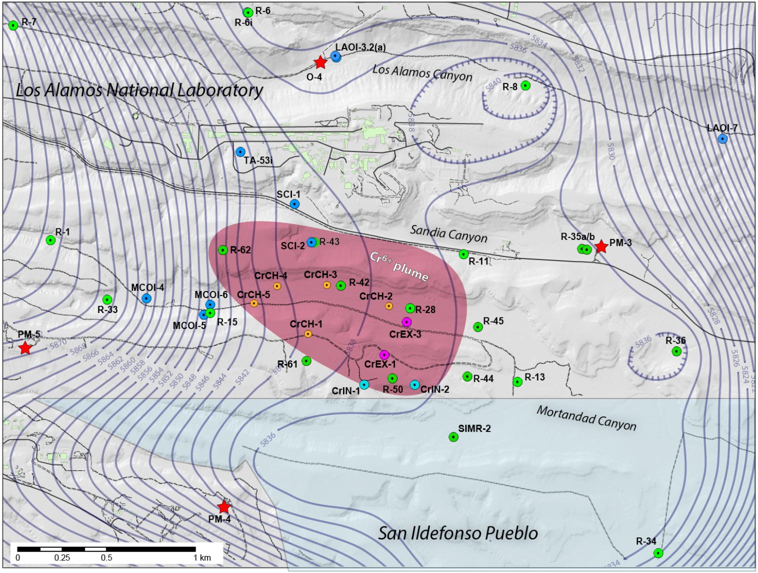
LANL Chromium site (~2011)



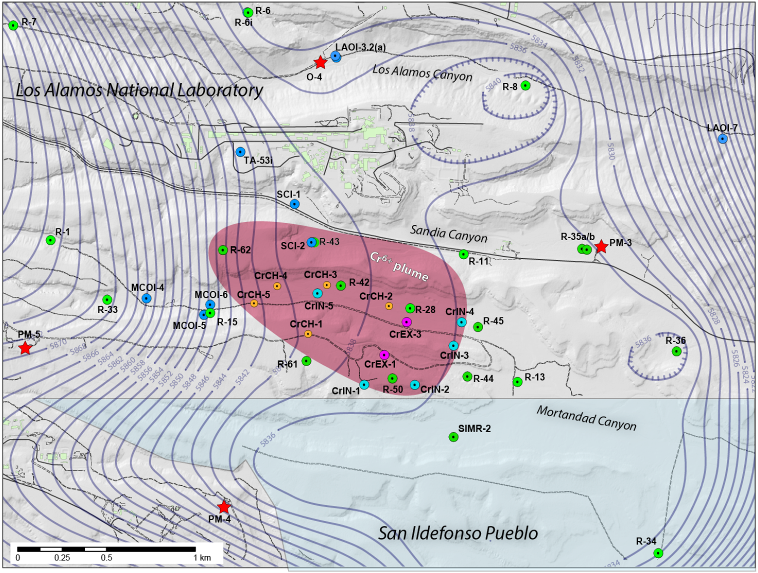
LANL Chromium site (2015)



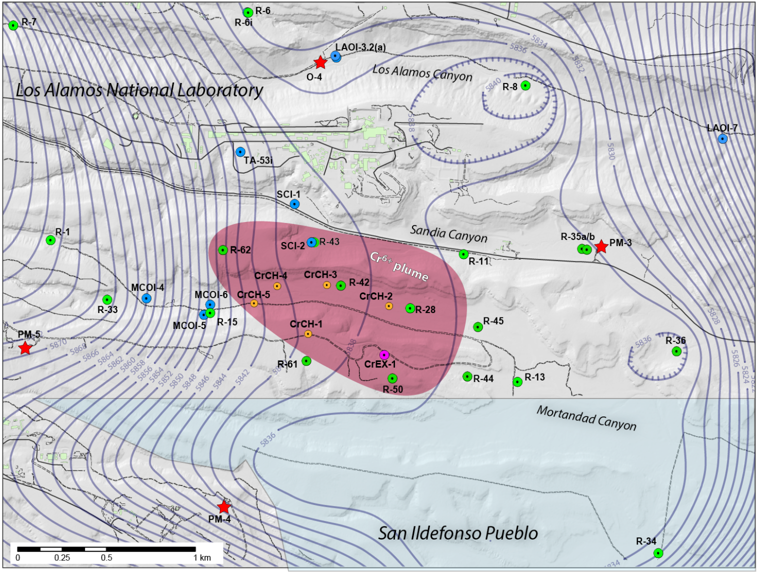
LANL Chromium site (~2015)



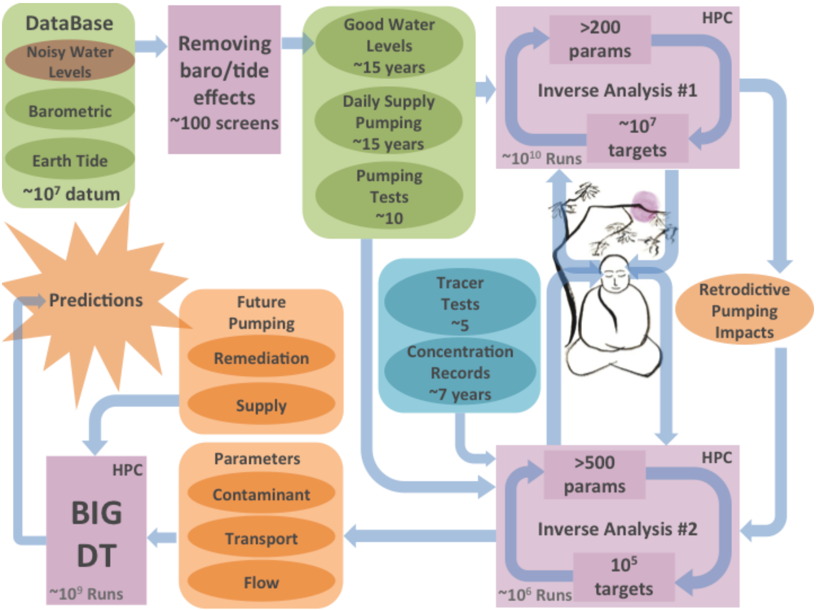
LANL Chromium site (~2016)



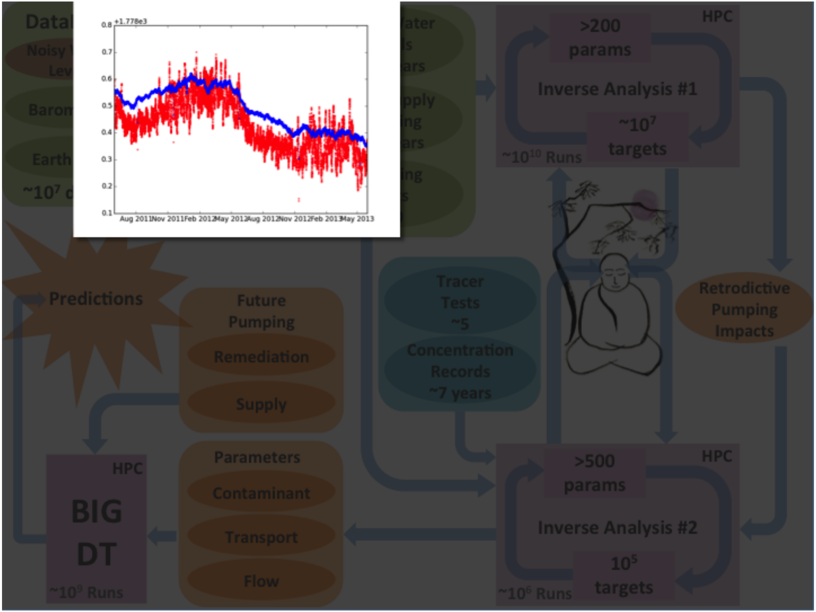
LANL Chromium site (2015)



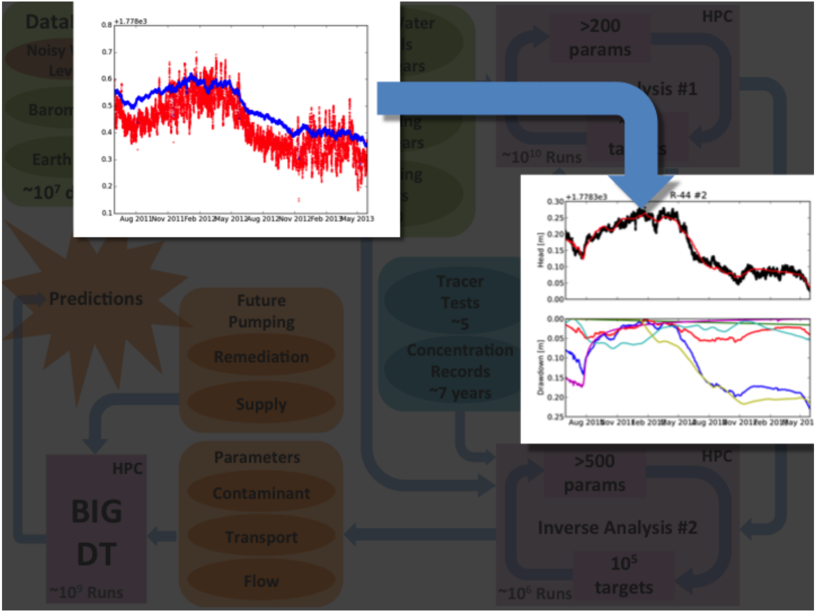
ZEM workflow: Data ⇔ Models ⇔ Decisions



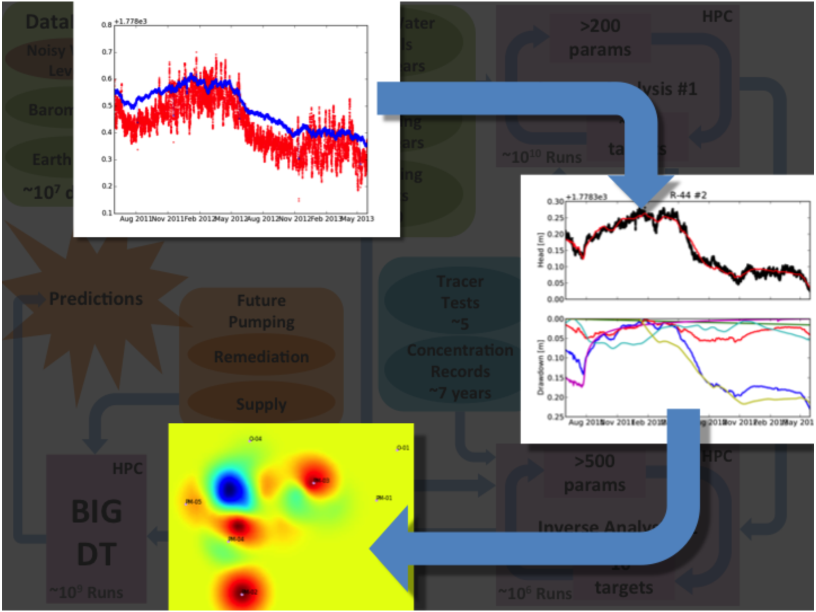
ZEM workflow: Data ⇔ Models ⇔ Decisions



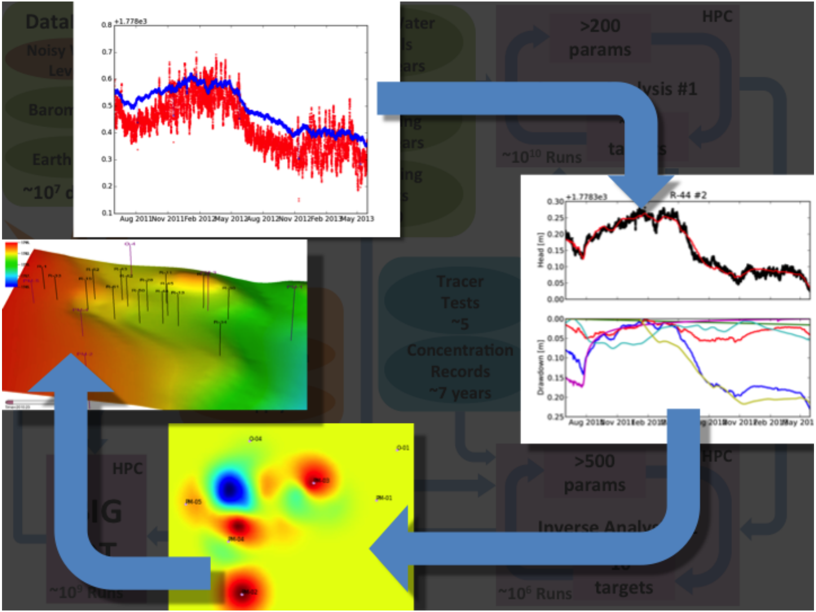
ZEM workflow: Data ⇔ Models ⇔ Decisions



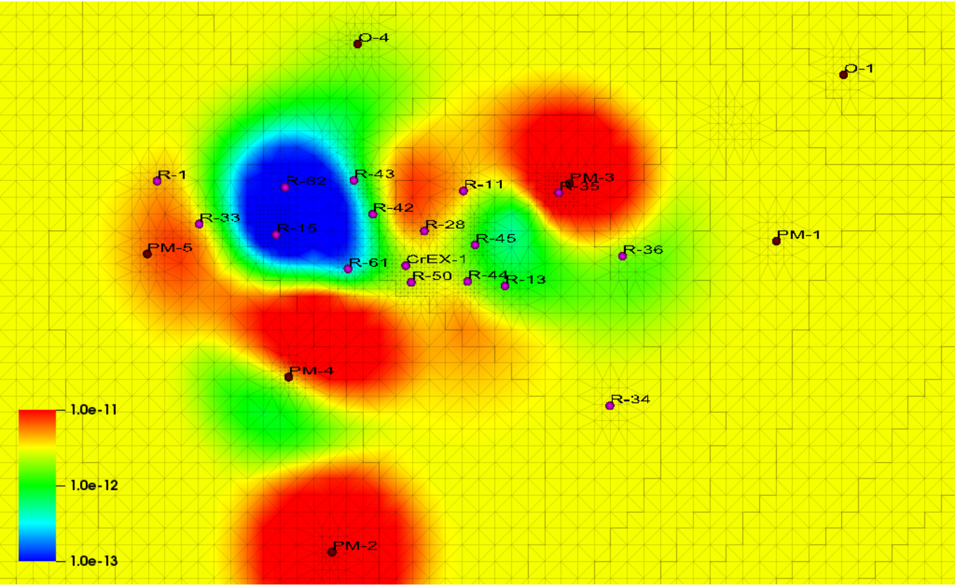
ZEM workflow: Data ⇔ Models ⇔ Decisions



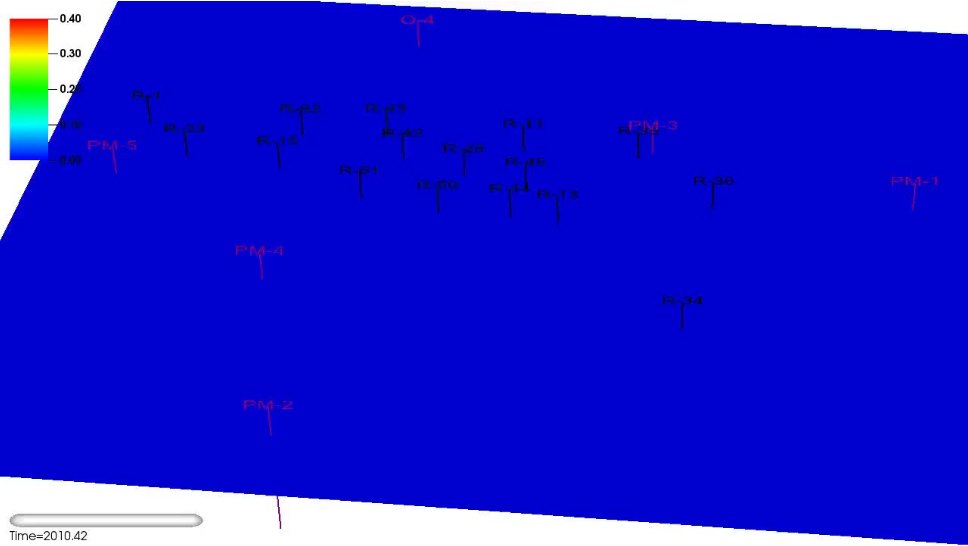
ZEM workflow: Data ⇔ Models ⇔ Decisions



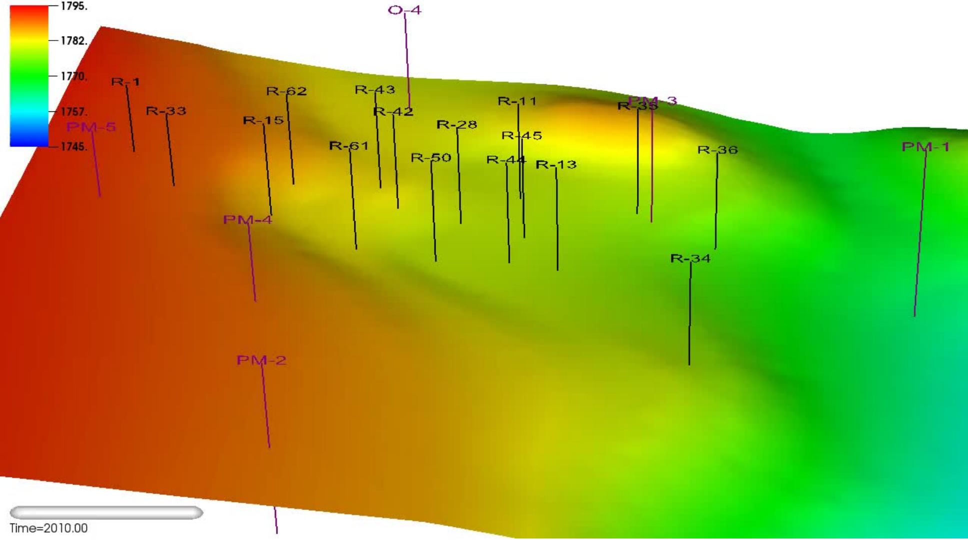
Potential Aquifer Heterogeneity (work in progress)



What are the drawdowns from the existing supply wells?



What are the water-level impacts?



Highlights: Uncertainties

- Many uncertainties in the environmental management problems **cannot** be represented probabilistically
- Newly developed methodology **BIG-DT** (**B**ayesian-**I**nformation **G**ap **D**ecision **T**heory) is developed to address this issue
- **BIG-DT** can quantify and address probabilistic and non-probabilistic uncertainties
- **BIG-DT** circumvents the shortcomings of Bayesian analysis
- **BIG-DT** accounts for unknowns and surprises
- **BIG-DT** provides scientifically-defensible decisions
- **BIG-DT** is currently applied for the LANL Chromium site
- Presented **BIG-DT** analyses are motivated by applications to decision support for geologic environmental management problems
- However, **BIG-DT** is applicable to any real-world engineering (environmental management) problem

- **ZEM** provides automated and reproducible workflow interconnecting Data ⇔ Models ⇔ Decisions
- **ZEM** provides quality assurance of the performance assessment process
- **ZEM** uses GIT to allow **version control** and **team collaboration** in the model development process based on cloud (web) repositories (gitlab.com/git.lanl.gov)
- **ZEM** is written predominantly using julia scripts
- julia: High-performance language for technical computing (MIT)
- For example, a single script can:
  - perform automated data query from the database
  - place the data in the model input files
  - initiate the simulations on HPC clusters
  - generate plots and movies with the final results

Highlights: LANL Chromium site

- Monitoring network at the site was augmented over the years using model-assisted decision analyses
- So far the model predictions have been consistent with the new observations
- Model-assisted decision analyses are currently performed to design site remediation activities

Highlights: BIG-DT & ZEM Implementation

- **BIG-DT** & **ZEM** implementation is based on **MADS**
- **MADS**: **M**odel **A**nalysis & **D**ecision **S**upport
  Open source C/C++ code (GPL v.3) `http://mads.lanl.gov`
- ASCEM: Advanced Simulation Capability for EM (DOE-EM)
- DiaMonD: An Integrated Multifaceted Approach to Mathematics at the Interfaces of Data, Models, and Decisions (DOE-SC; MIT, UT-Austin, U of Colorado, LANL)