Model Analyses of Complex Systems Behavior using MADS

Velimir V. Vesselinov vvv@lanl.gov
Daniel O'Malley omalled@lanl.gov

Computational Earth Science, Los Alamos National Laboratory, USA

AGU Fall meeting, December, 2016
Unclassified: LA-UR-16-29120

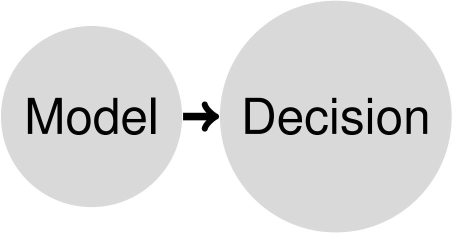
Our work inform important decisions

- **Climate Science**: Should we cap carbon emissions or not?
- **Meteorology**: Should we evacuate a city due to a hurricane?
- **Geology**: How much should we bid on a fossil fuel play?
- **Seismology**: Should we inject fluids in the underground (and how to do it without causing earthquakes and contamination)
- **Hydrogeology**: How to provide clean water supply?
- **Hydrogeology**: Which remediation option will clean up the groundwater?

- **Climate Science**: Should we cap carbon emissions or not?
- **Meteorology**: Should we evacuate a city due to a hurricane?
- **Geology**: How much should we bid on a fossil fuel play?
- **Seismology**: Should we inject fluids in the underground (and how to do it without causing earthquakes and contamination)
- **Hydrogeology**: How to provide clean water supply?
- **Hydrogeology**: Which remediation option will clean up the groundwater?

We rely on **data** & **models** to make scientifically defensible **decisions**

How should we support these decisions?

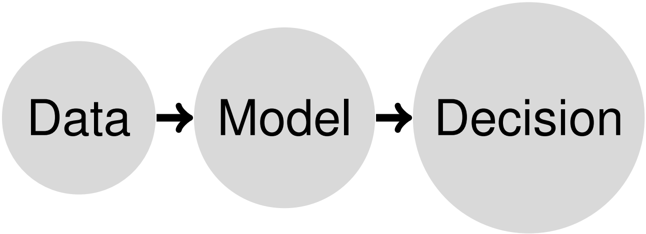


- Build a “representative” model
- Use the “representative” model to make a decision

How should we support these decisions?

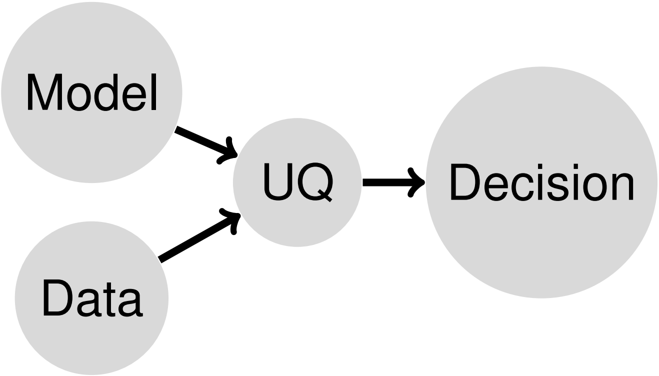
Model ➜ Decision

- Build a “representative” model
- Use the “representative” model to make a decision
- **However:**
  - many real-world models cannot be validated (especially in the earth sciences)
  - data can be highly uncertain
  - conceptualization can be highly uncertain
  - model predictions can be highly uncertain



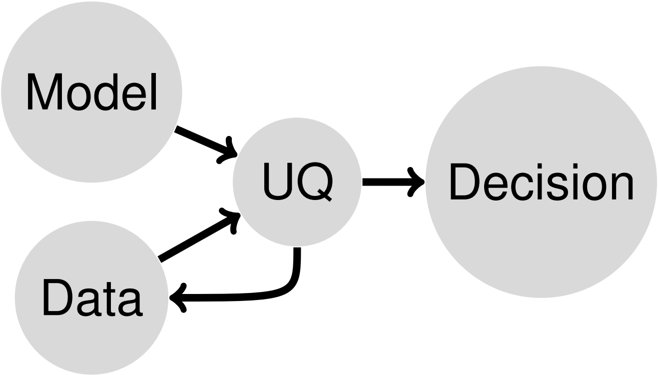
- Use data to calibrate the model
- Use the model to make a decision

How should we support these decisions?



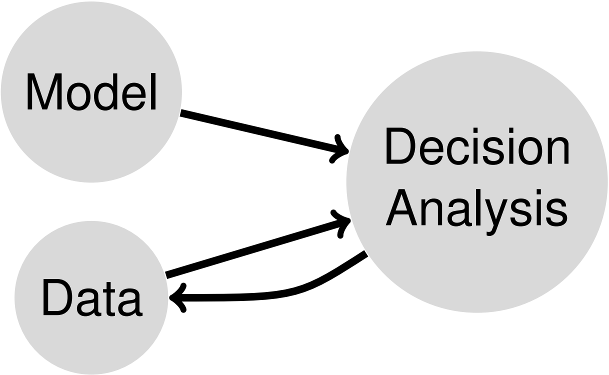
- Quantify uncertainty in the data and model
- Use estimated uncertainties in model predictions to make a decision

How should we support these decisions?



- Quantify uncertainty in the data and model
- Estimate uncertainties in model predictions
- Collect data that reduces the prediction uncertainties (optimal experimental design)
- Use the new data to quantify uncertainty again
- Use updated uncertainties to make decision (**are we done?**)

How should we support these decisions?



- Perform **decision analysis** coupling UQ with the decision process to evaluate uncertainty in decisions (**not** uncertainty in model parameters/predictions)
- Use the **decision analysis** to guide collection of data that can influence a better decision (it may not always be feasible)
- Use the new collected data to make a better decision

Decision Analysis Methodologies and Tools

- We need robust and versatile decision analysis methodologies and tools
- Recently, we have developed a series of novel methods and techniques for **data**- and **model**-based **decision** analyses
- Most of them are implemented in **MADS**

Model **A**nalysis & **D**ecision **S**upport

MADS: Model Analysis and Decision Support

- **MADS** is a **high-performance** computational framework
- **MADS** performs a wide range of **data**- & **model**-based analyses including
  - Sensitivity Analysis
  - Parameter Estimation, Model Inversion/Calibration
  - Uncertainty Quantification
  - Machine Learning Methods
  - Reduced Order Modeling (ROM)
  - Optimal Experimental Design (OED)
  - Decision Analysis
- **MADS** is open source code (GPL3) written in julia
- julia is a high-level, dynamic programming language for technical computing
- julia has C speed but with MatLab/Python flexibility
- julia provides access to a vast number of mathematical, statical, and visualization packages

MADS: Model Analysis and Decision Support

- **MADS** can be applied to perform analyses using any existing physics simulator
- **MADS** provides tools for model development, integration and couping
- **MADS** utilizes advanced code development tools for
  - version control (git)
    `https://github.com/madsjulia`
  - continuous integration (Travis-CI)
    `https://travis-ci.org/madsjulia/Mads.jl`
  - tracking code test coverage (Coveralls)
    `https://coveralls.io/github/madsjulia/Mads.jl`
- **MADS** contributors and developers are welcome
- **MADS** examples, manuals and publications are available at:
  `https://mads.lanl.gov`
  `https://madsjulia.github.io/Mads.jl`
  `https://mads.readthedocs.io`

Advanced and novel methods implemented in MADS

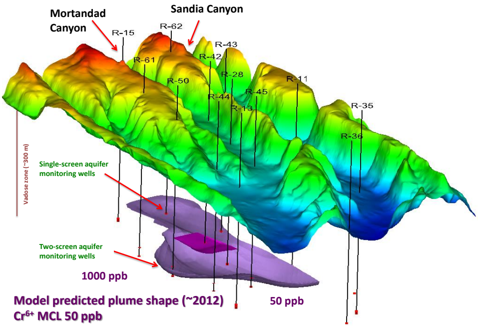
- Information-Gap Decision Theory (**IGDT**)
  - O'Malley, D., Vesselinov, V.V., Groundwater remediation using the information gap decision theory, Water Resources Research, doi: 10.1002/2013WR014718, 2014.
  - Harp, D.R., Vesselinov. V.V., Contaminant remediation decision analysis using information gap theory, Stochastic Environmental Research and Risk Assessment (SERRA), doi:10.1007/s00477-012-0573-1, 2012.
- Bayesian-Information-Gap Decision Theory (**BIG-DT**)
  - O'Malley, Vesselinov: Groundwater Remediation using Bayesian Information-Gap Decision Theory (**West 3024**, Thursday, 17:00 - 17:15, **H44E-05**)
  - Grasinger, M., O'Malley, D., Vesselinov, V.V., Karra, S., Decision Analysis for Robust CO2 Injection: Application of Bayesian-Information-Gap Decision Theory, International Journal of Greenhouse Gas Control, doi: 10.1016/j.ijggc.2016.02.017, 2016.
  - O'Malley, D., Vesselinov, V.V., Bayesian-Information-Gap decision theory with an application to CO2 sequestration, Water Resources Research, doi: 10.1002/2015WR017413, 2015.
  - O'Malley, D., Vesselinov, V.V., A combined probabilistic/non-probabilistic decision analysis for contaminant remediation, Journal on Uncertainty Quantification, SIAM/ASA, doi: 10.1137/140965132, 2014.
- Optimal Experimental Design (**OED**) driven by decision analysis
  - O'Malley, D., Vesselinov, V.V., (in preparation).
- Measure-theoretic Uncertainty Quantification (**UQ**)
  - Dawson, Butler, Mattis, Westerink, Vesselinov, Estep: Parameter Estimation for Geoscience Applications Using a Measure-Theoretic Approach (**West 3024**, Thursday, 17:30 - 17:45, **H44E-07**)
  - Mattis, S.A., Butler, T.D. Dawson, C.N., Estep, D., Vesselinov, V.V., Parameter estimation and prediction for groundwater contamination based on measure theory, Water Resources Research, doi: 10.1002/2015WR017295, 2015.
- Novel Levenberg-Marquardt (**LM**) optimization method using a dimensionality reduction based on Krylov subspace method
  - Lin, Y, O'Malley, D., Vesselinov, V.V., A computationally efficient parallel Levenberg-Marquardt algorithm for highly parameterized inverse model analyses, Water Resources Research, doi: 10.1002/2016WR019028, 2016.

- Model inversion using modified Total-Variation (**TV**) regularization
  - Lin, O'Malley, Vesselinov: Hydraulic Inverse Modeling with Modified Total-Variation Regularization with Relaxed Variable-Splitting (**poster**, Thursday, 8:00 - 12:00, **H41B-1301**)
- Model inversion using Principal Component Geostatistical Approach (**PCGA**) and Randomized Geostatistical Approach (**RGA**)
  - Lin, Y, Le, E.B, O'Malley, D., Vesselinov, V.V., Bui-Thanh, T., Large-Scale Inverse Model Analyses Employing Fast Randomized Data Reduction, 2016, (submitted).
- Blind Source Separation (**BSS**) using Non-negative Matrix Factorization (**NMF**)
  - Vesselinov, V.V., O'Malley, D., Alexandrov, B.S., Source identification of groundwater contamination sources and groundwater types using semi-supervised machine learning, (in preparation).
  - Iliev, F.L., Stanev, V.G., Vesselinov, V.V., Alexandrov, B.S., Sources identification using shifted non-negative matrix factorization combined with semi-supervised clustering, 2016, (submitted).
  - Stanev, V.G., Iliev, F.L., Vesselinov, V.V., Alexandrov, B.S., Machine learning approach for identification of release sources in advection-diffusion systems, 2016, (submitted).
  - Alexandrov, B., Vesselinov, V.V., Blind source separation for groundwater level analysis based on non-negative matrix factorization, Water Resources Research, doi: 10.1002/2013WR015037, 2014.
- Support Vector Regression (**SVR**) methods for surrogate modeling
  - Alexandrov, B.S., O'Malley, D., Vesselinov, V.V., (in preparation).
  - Vesselinov, O'Malley, Alexandrov, Moore: Reduced Order Models for Decision Analysis and Upscaling of Aquifer Heterogeneity (**South 302**, Monday, 8:45 - 9:00, **NG11A-04**)

- Advanced Monte Carlo Methods: Robust Adaptive Metropolis (**RAM**) and Affine Invariant Markov Chain Monte Carlo Ensemble Sampler (aka **Emcee**)
  - Vihola: Robust adaptive Metropolis algorithm with coerced acceptance rate, Statistics and Computing, 2012.
  - Goodman, Weare: Ensemble samplers with affine invariance. Communications in applied mathematics and computational science, 2010.
- Extended Fourier Amplitude Sensitivity Testing (**eFAST**) global sensitivity analysis
  - Saltelli, et al. Global sensitivity analysis, John Wiley & Sons, 2008.
- Multifidelity Global Sensitivity Analysis (**MFSA**) under given computational budget
  - Qian, Peherstorfer, O'Malley, Vesselinov, Wilcox: Multifidelity Global Sensitivity Analysis, SIAM, 2016, (submitted).

- Groundwater contaminant remediation (LANL Chromium & RDX)
  - Mattis, S.A., Butler, T.D. Dawson, C.N., Estep, D., Vesselinov, V.V., Parameter estimation and prediction for groundwater contamination based on measure theory, Water Resources Research, doi: 10.1002/2015WR017295, 2015.
  - Vesselinov, V.V., O'Malley, D., Katzman, D., Model-Assisted Decision Analyses Related to a Chromium Plume at Los Alamos National Laboratory, Waste Management, 2015.
  - O'Malley, D., Vesselinov, V.V., A combined probabilistic/non-probabilistic decision analysis for contaminant remediation, Journal on Uncertainty Quantification, SIAM/ASA, doi: 10.1137/140965132, 2014.
  - O'Malley, D., Vesselinov, V.V., Analytical solutions for anomalous dispersion transport, Advances in Water Resources, doi: 10.1016/j.advwatres.2014.02.006, 2014.
- Water/Energy/Food Nexus
  - Zhang, Vesselinov: Bi-Level Decision Making for Supporting Energy and Water Nexus (**West 3016**: Wednesday, 09:15 - 09:30, **H31J-06**)
  - Zhang, X., Vesselinov, V.V., Integrated Modeling Approach for Optimal Management of Water, Energy and Food Security Nexus Advances in Water Resources, Advances in Water Resources, 2016 (submitted).
  - Zhang, X., Vesselinov, V.V., Energy-Water Nexus: Balancing the Tradeoffs between Two-Level Decision Makers Applied Energy, Applied Energy, DOI: 10.1016/j.apenergy.2016.08.156, 2016.
- CO_2 injection
  - Grasinger, M., O'Malley, D., Vesselinov, V.V., Karra, S., Decision Analysis for Robust CO2 Injection: Application of Bayesian-Information-Gap Decision Theory, International Journal of Greenhouse Gas Control, doi: 10.1016/j.ijggc.2016.02.017, 2016.
  - O'Malley, D., Vesselinov, V.V., Bayesian-Information-Gap decision theory with an application to CO2 sequestration, Water Resources Research, doi: 10.1002/2015WR017413, 2015.

LANL Chromium site



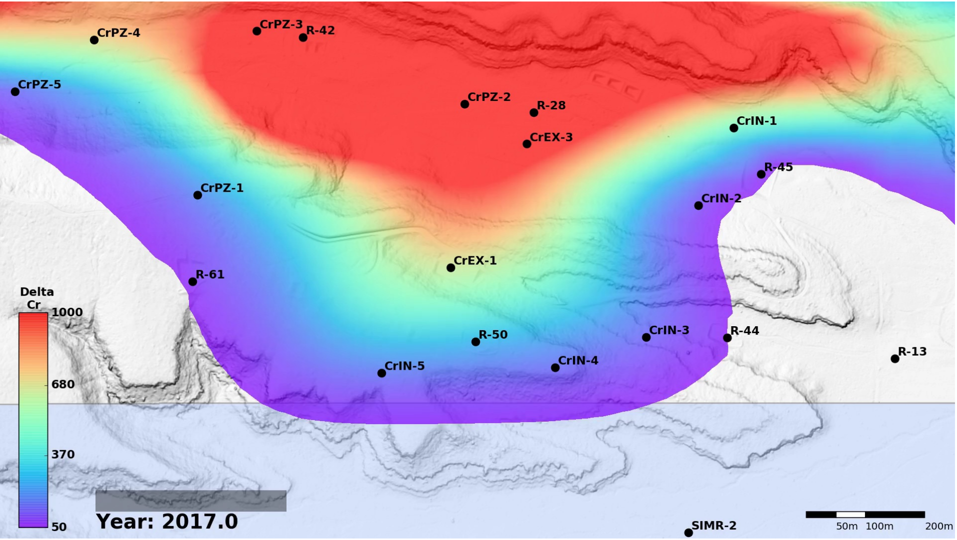
- Groundwater contamination site with high visibility (DOE)
- More than 20 wells drilled since 2007 (each wells costs $2-3M)
- **Limited remedial options**

- Complex **uncertainties/unknowns**
- Plume is located near **LANL boundary** and **water-supply wells**
- Modeling accounts for complex biogeochemical processes in highly heterogeneous media
- In the last 5 years, we have accumulated close to 2,000 years computational time on the LANL HPC clusters for various model analyses
- Used up to 4,096 processors simultaneously

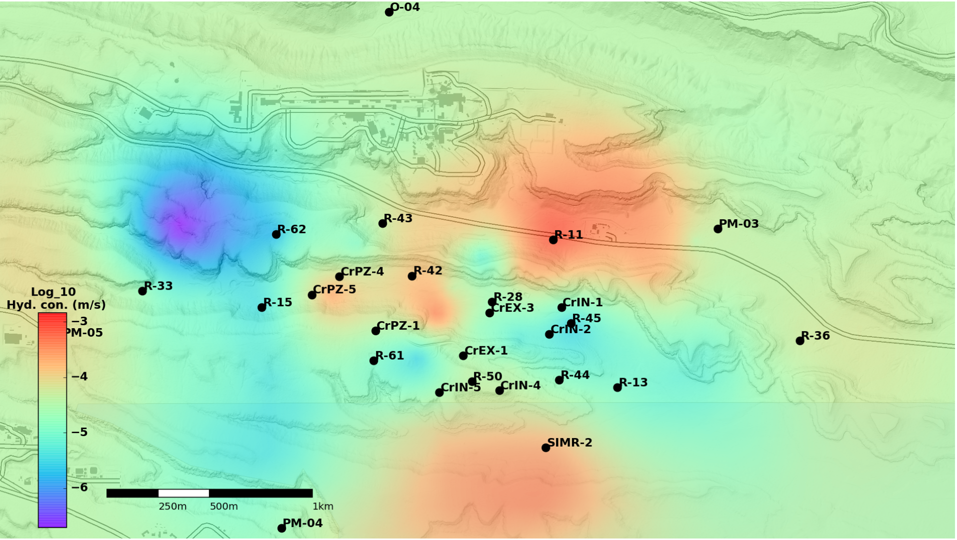
LANL regional aquifer model

- Model domain encompasses the regional aquifer beneath LANL ($\approx 8 \times 4 \times 0.3\ km$)
- 766,283 nodes / 4,659,062 cells
- 193 concentration calibration targets (representing annual transients for about 10 years)
- 182,090 water-level calibration targets (representing daily transients for about 4 years)
- water-level transients represent pumping effects caused by 9 wells (6 water-supply wells and 3 site wells where pumping tests are conducted)
- 230 unknown model parameters representing groundwater flow and transport (including aquifer heterogeneity and spatial location/strength of 3 unknown contaminant sources)
- Calibration required about 100 years computational time
- ... more data and physics/biogeochemistry needs to be incorporated in the model soon!

LANL regional aquifer model: Chromium transients



LANL regional aquifer model: Aquifer heterogeneity



LANL ADEM: Los Alamos National Laboratory Environmental Management Directorate

DiaMonD: An Integrated Multifaceted Approach to Mathematics at the Interfaces of **D**ata, **M**odels, and **D**ecisions https://mads.lanl.gov

https://github.com/madsjulia

https://madsjulia.github.io/Mads.jl

Related model and decision analyses presentations at AGU 2016

- Vesselinov, O'Malley, Alexandrov, Moore: Reduced Order Models for Decision Analysis and Upscaling of Aquifer Heterogeneity (**South 302**, Monday, 8:45 - 9:00, **NG11A-04**, **invited**)
- Lu, Vesselinov, Lei: Identifying Aquifer Heterogeneities using the Level Set Method (**poster**, Wednesday, 8:00 - 12:00, **H31F-1462**)
- Zhang, Vesselinov: Bi-Level Decision Making for Supporting Energy and Water Nexus (**West 3016**: Wednesday, 09:15 - 09:30, **H31J-06**)
- Vesselinov, O'Malley: Model Analysis of Complex Systems Behavior using MADS (**West 3024**: Wednesday, 15:06 - 15:18, **H33Q-08**)
- Hansen, Vesselinov: Analysis of hydrologic time series reconstruction uncertainty due to inverse model inadequacy using Laguerre expansion method (**West 3024**: Wednesday, 16:30 - 16:45, **H34E-03**)
- Lin, O'Malley, Vesselinov: Hydraulic Inverse Modeling with Modified Total-Variation Regularization with Relaxed Variable-Splitting (**poster**, Thursday, 8:00 - 12:00, **H41B-1301**)
- Hansen, Haslauer, Cirpka, Vesselinov: Prediction of Breakthrough Curves for Conservative and Reactive Transport from the Structural Parameters of Highly Heterogeneous Media (**West 3014**, Thursday, 14:25 - 14:40, **H43N-04**)
- O'Malley, Vesselinov: Groundwater Remediation using Bayesian Information-Gap Decision Theory (**West 3024**, Thursday, 17:00 - 17:15, **H44E-05**)
- Dawson, Butler, Mattis, Westerink, Vesselinov, Estep: Parameter Estimation for Geoscience Applications Using a Measure-Theoretic Approach (**West 3024**, Thursday, 17:30 - 17:45, **H44E-07**)