SAND2011-1954C

Model-Based Decision Making: A New Perspective

John Doherty[1], Scott James[2], and Paul Reimus[3]

April13th, 2011

[1]National Centre for Groundwater Research and Training, Flinders University, Australia
[2]Thermal/Fluid Science & Engineering, Sandia National Laboratories, USA
[3]Earth and Environmental Sciences, Los Alamos National Laboratory, USA

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Not So New Perspective?

"All models are wrong, but some are useful."

George Edward Pelham Box, *Robustness in the strategy of scientific model building*, in "Robustness in Statistics," R.L. Launer and G.N. Wilkinson, Editors, Academic Press: New York, 1979.

"…We have a large reservoir of engineers (and scientists) with a vast background of engineering know how. They need to learn statistical methods that can tap into the knowledge. Statistics used as a catalyst to engineering creation will, I believe, always result in the fastest and most economical progress…."

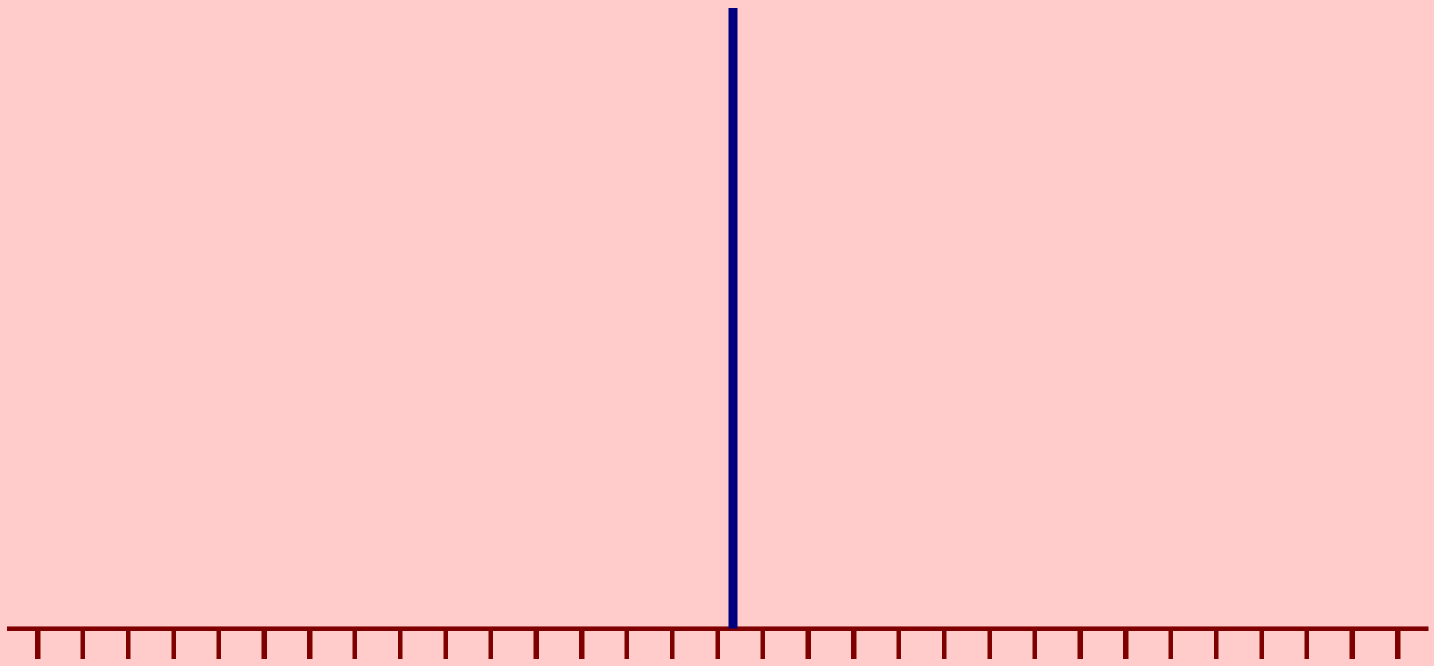
Sandia National Laboratories

All a Model Can Promise

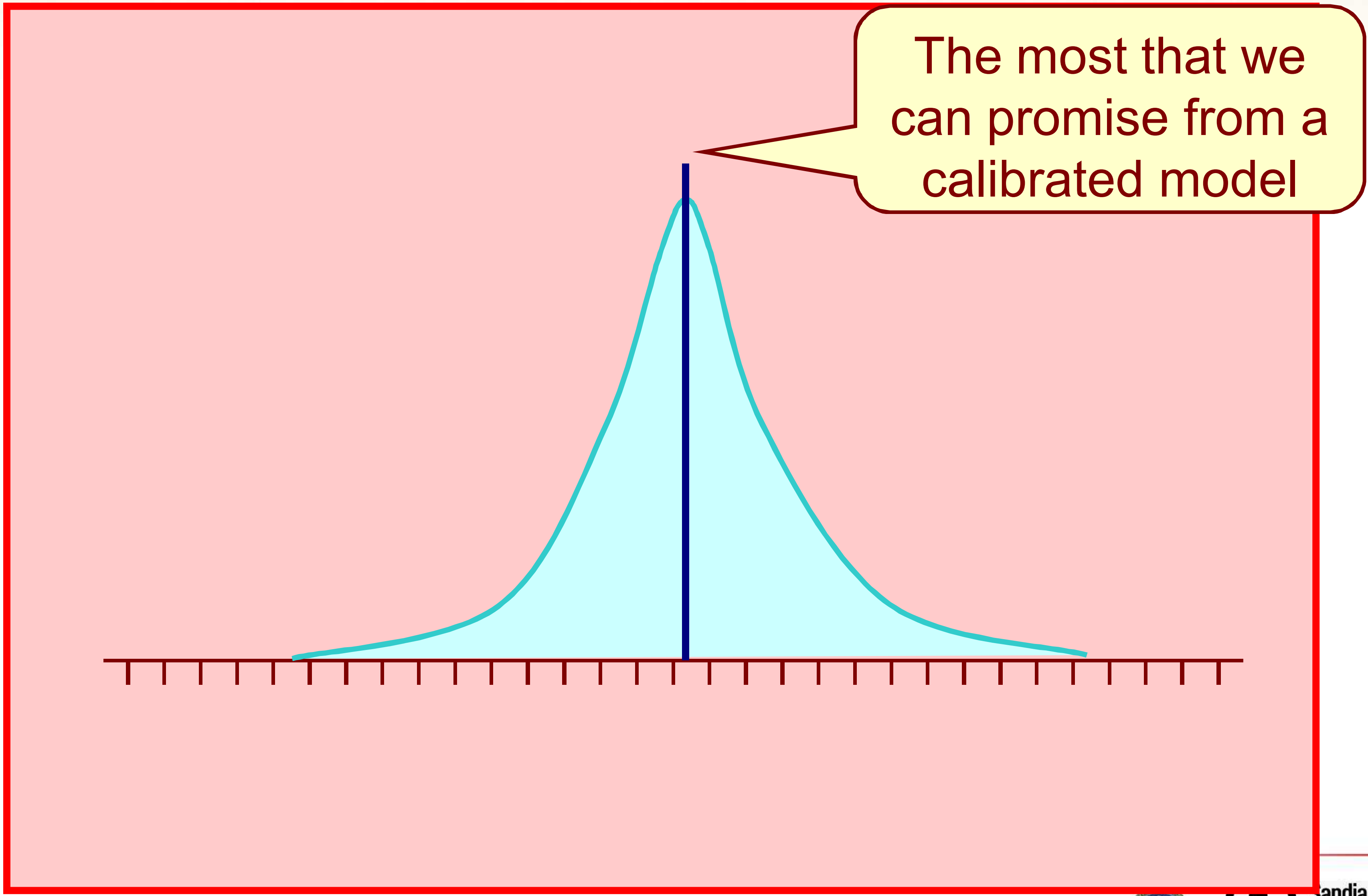
- The correct prediction lies within model-computed uncertainty intervals (quantify predictive uncertainty)
- These intervals are reduced to as close to their theoretical minimum as possible (strictly a function of information content of data and expert knowledge).

A model cannot promise the correct prediction!

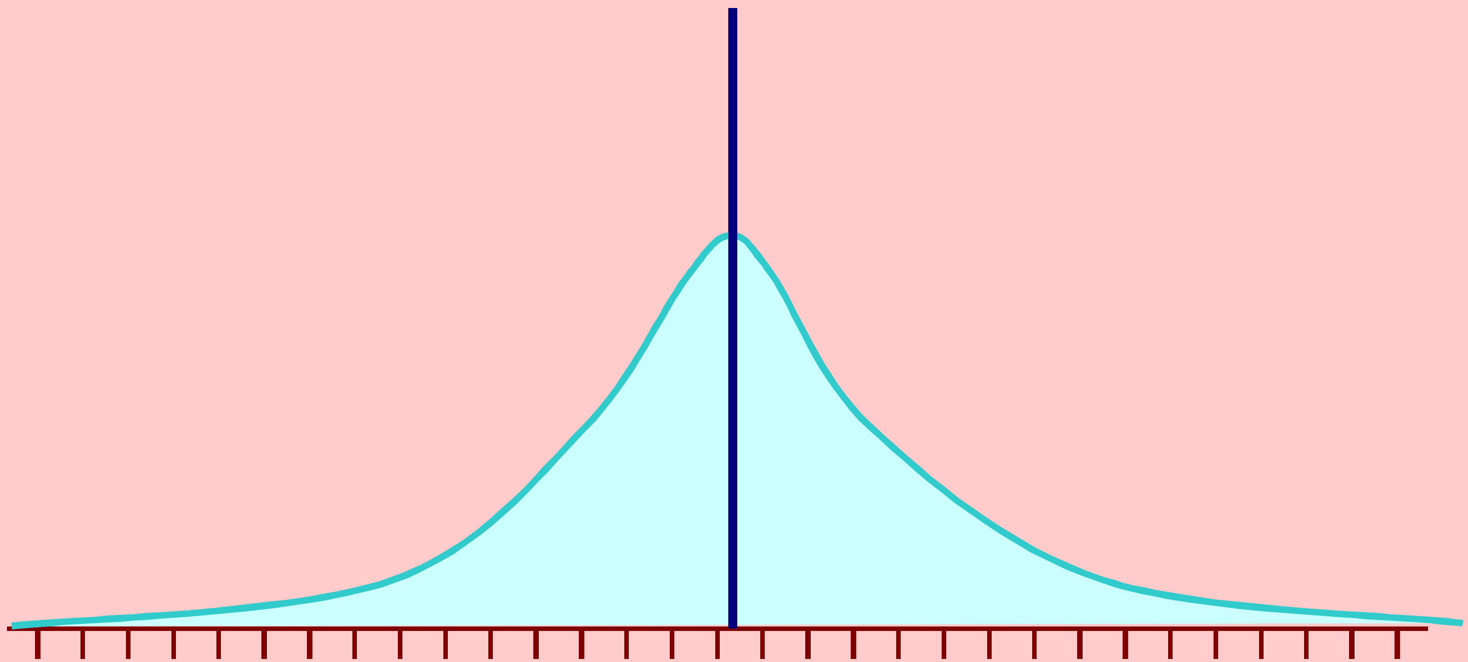
The Prediction of Minimum Error Variance



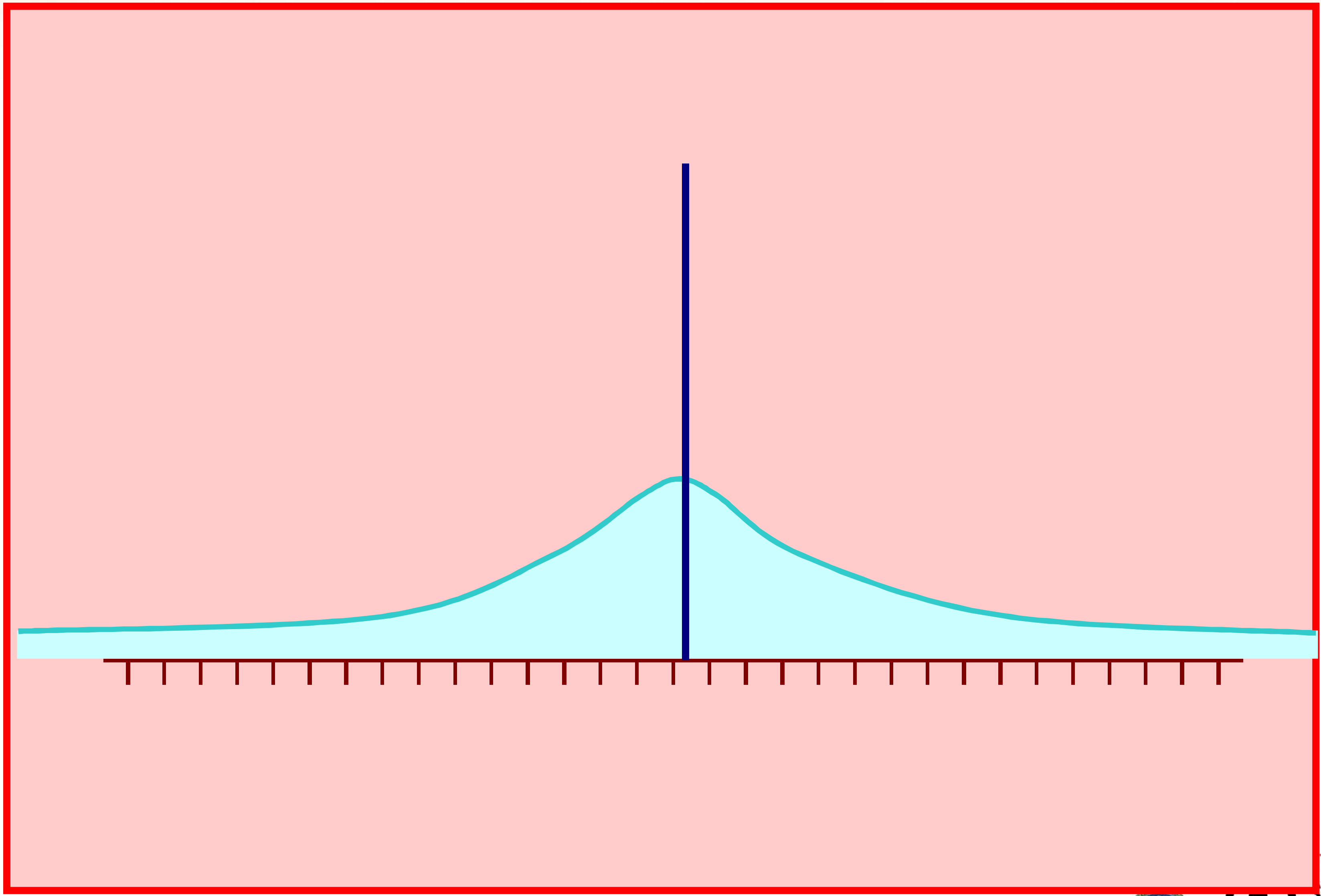
The Prediction of Minimum Error Variance



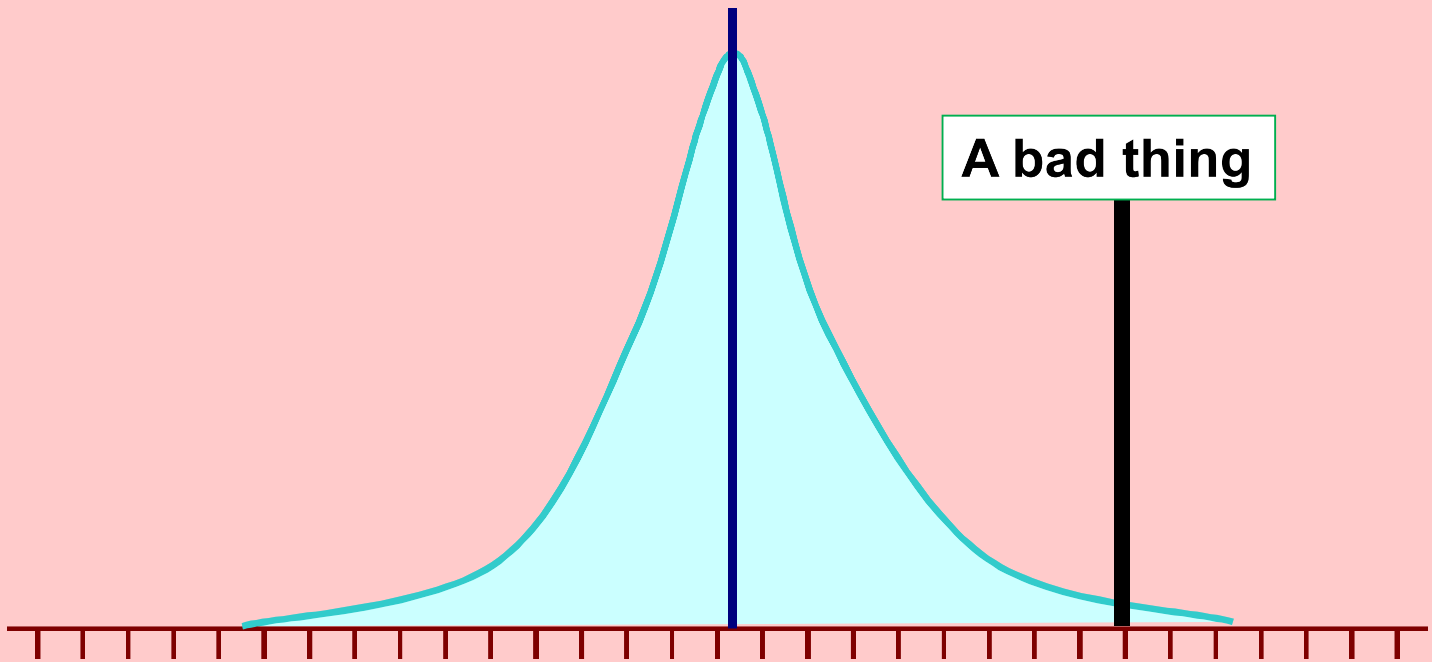
The Prediction of Minimum Error Variance



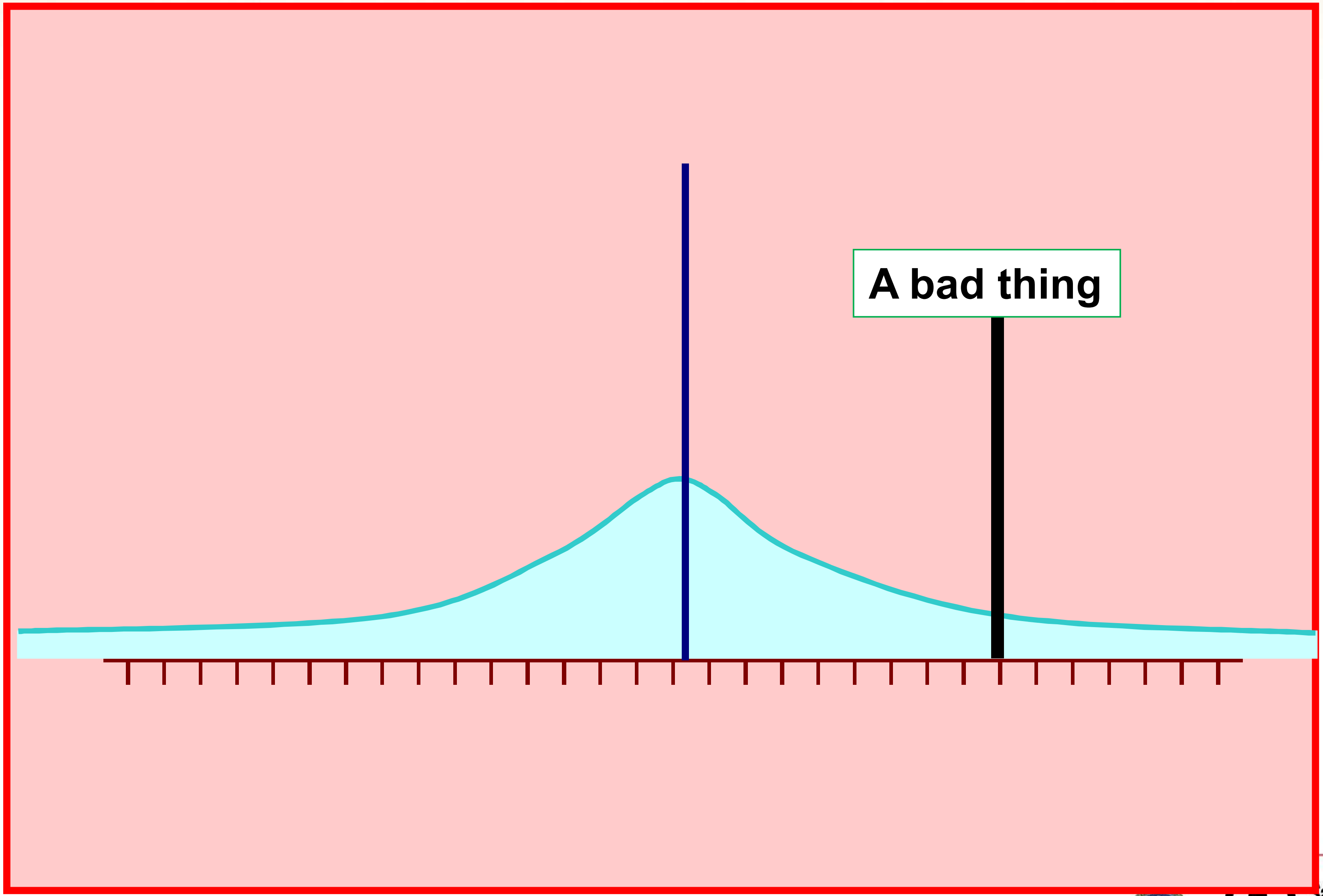
The Prediction of Minimum Error Variance



Confidence of Adverse-Occurrence Avoidance



Confidence of Adverse-Occurrence Avoidance

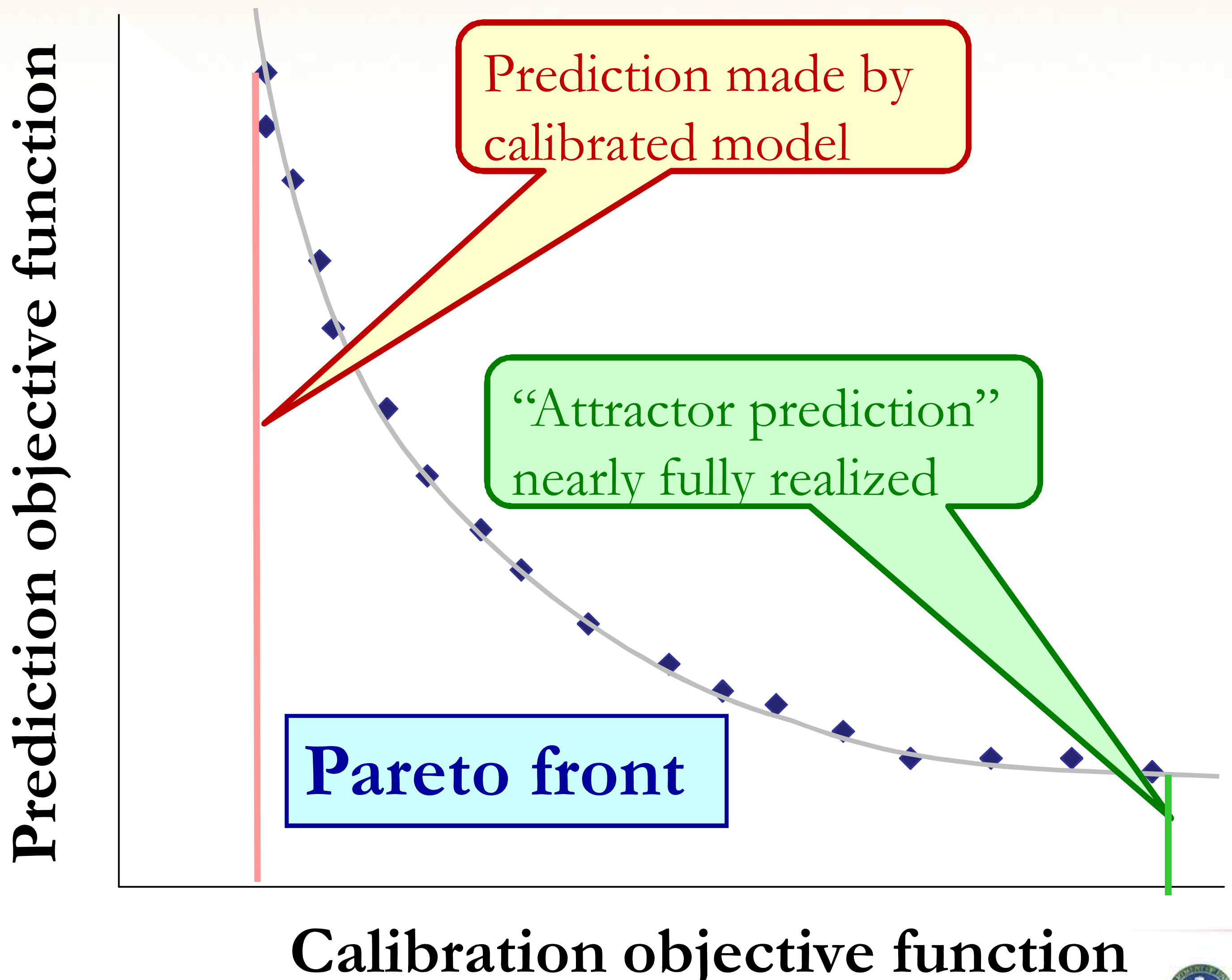


Pareto Optimization

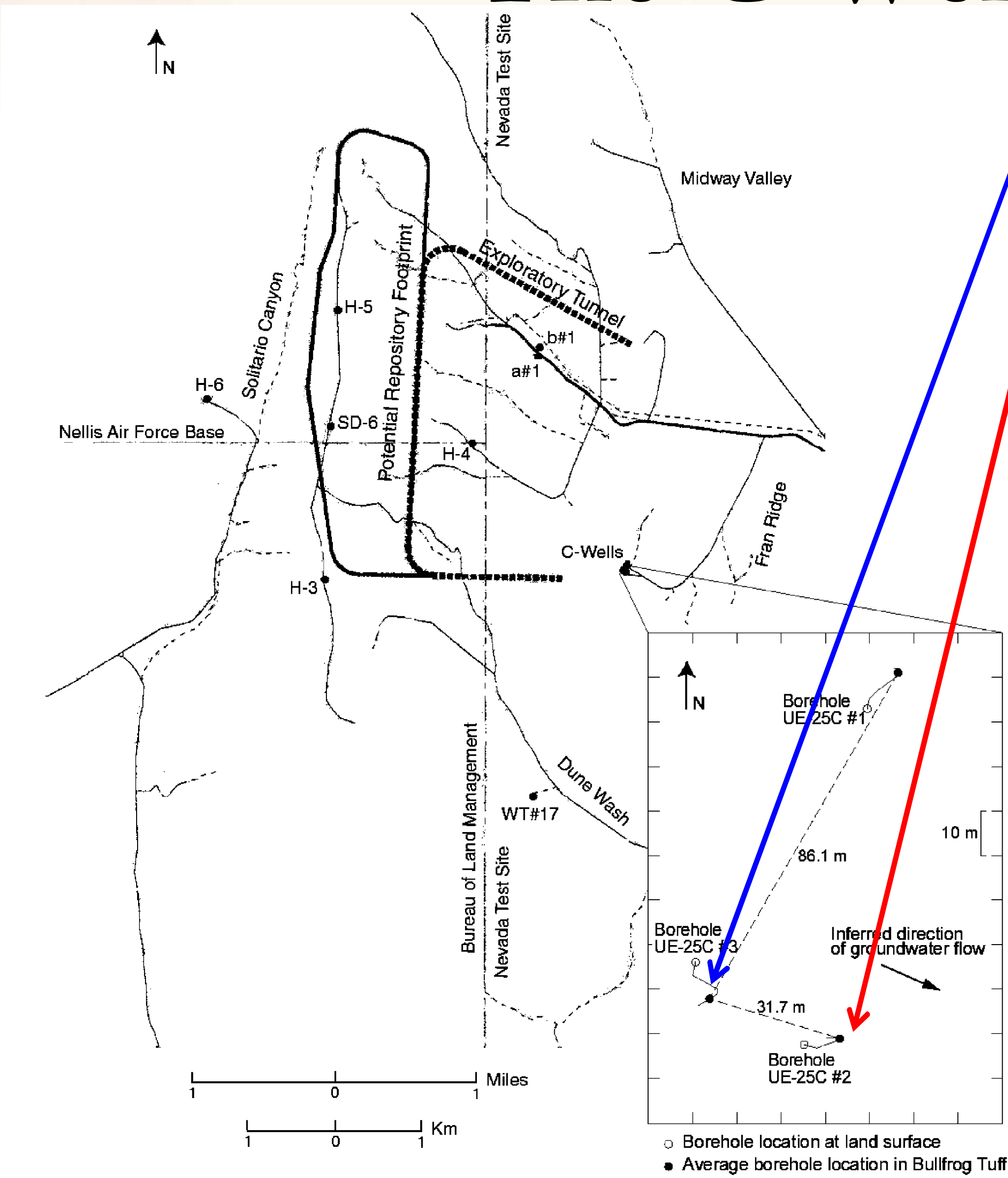
Given an initial allocation of goods among a set of individuals, a change to a different allocation that makes at least one individual better off without making any other individual worse off is called a Pareto improvement. An allocation is defined as "Pareto efficient" or "Pareto optimal" when no further Pareto improvements can be made.

− Italian economist Vilfredo Pareto (1848−1923)

Pareto Optimization



The C-Wells Test



Injection at well UE25C#3

Withdrawal at well UE25C#2

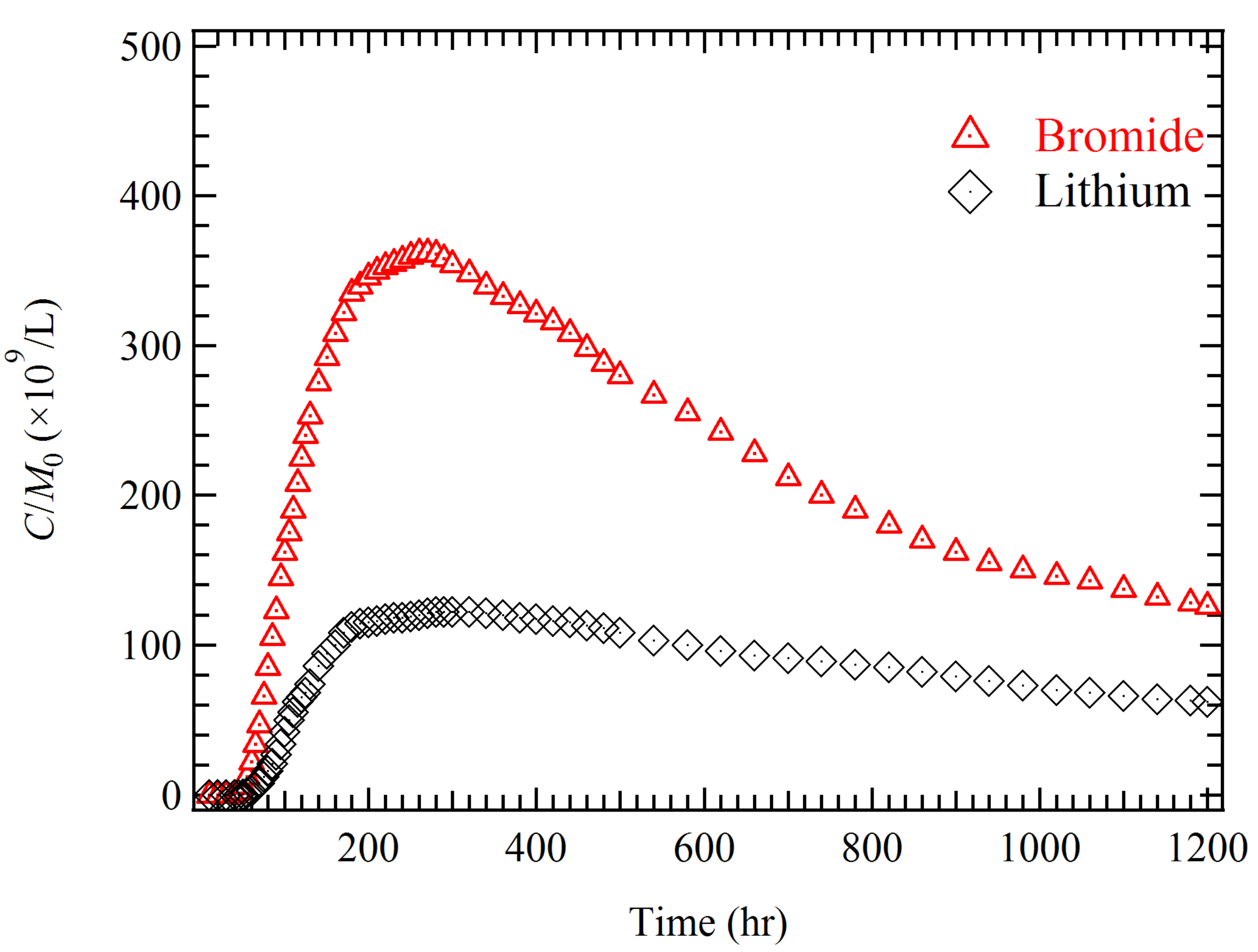
- Dipole established along the inferred groundwater flow direction.
- Wells separated by 31.7 m.
- Three tracers injected:
  - PFBA
  - Bromide
  - Lithium

Tracers

- 30.6 kg Bromide (Br)
- 16.0 kg lithium (Li)
- PFBA not considered in this analysis
- Br non-sorbing
- Li weakly sorbing
- Br: $D_{\mathrm{f}} = 2.1\times10^{-5}$ cm^2/s
- Li: $D_{\mathrm{f}} = 1.0\times10^{-5}$ cm^2/s

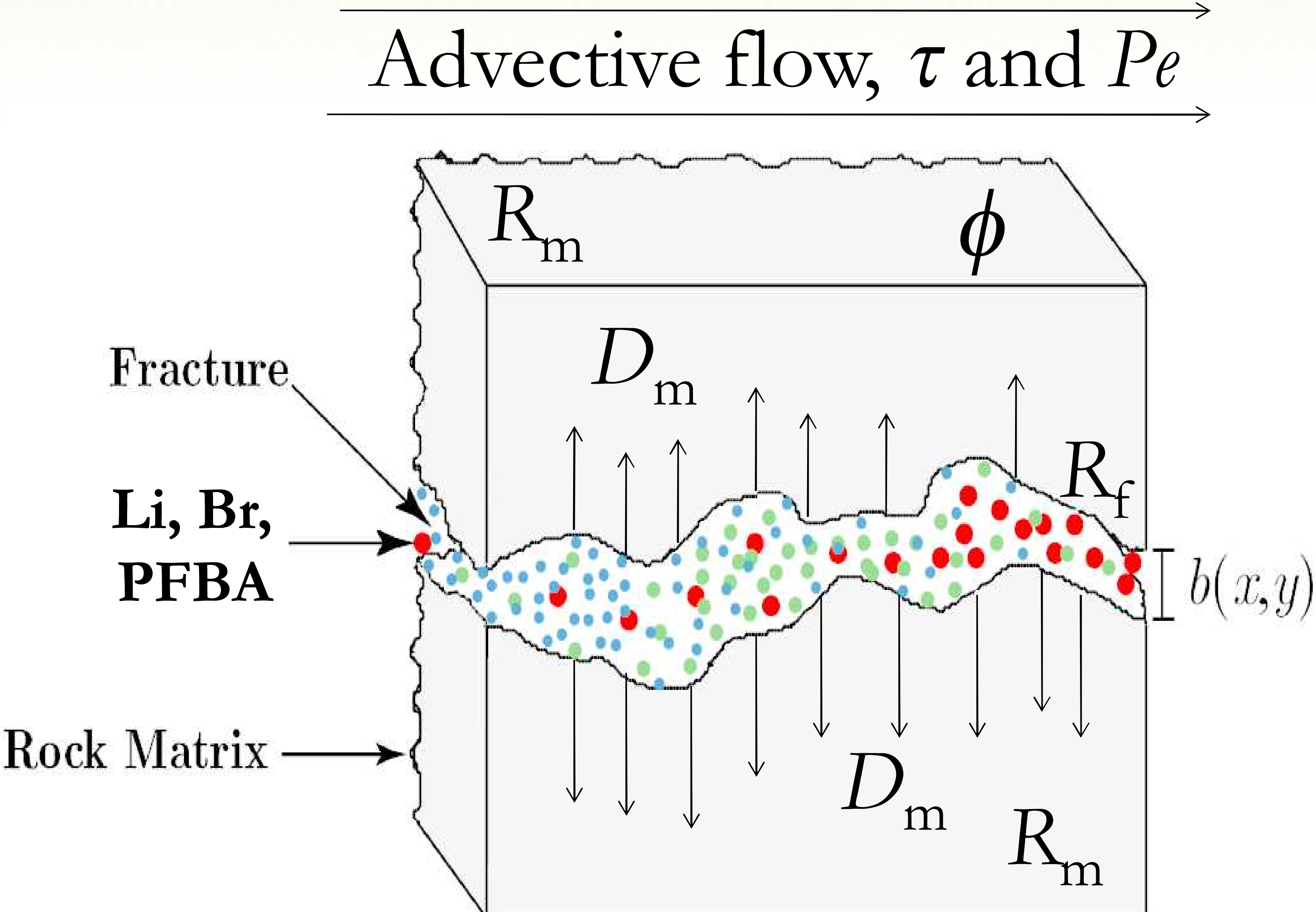
Sandia National Laboratories

Breakthrough Data



PFBA is not considered in this example uncertainty assessment.

Conceptual Model



Model: RELAP

- Solves the dual-porosity (fracture/matrix) transport equations undergoing linear sorption subject to a finite pulse injection, wellbore mixing, and recirculation.
- There are 8 loosely constrained input parameters that control the characteristics of the simulated breakthrough curve.
- These are estimated with PEST.

Adjustable Parameter Ranges

- Mass fraction of tracers in the test: $f = 0.2, 0.6, 1.0$ (−)
- Li fracture retardation factor: $R_f = 1, 5, 10$ (−)
- Li matrix retardation factor: $R_m = 1, 25, 50$ (−)
- Fracture aperture: $b = 0.01, 0.15, 1.0$ (cm)
- Porosity: $\phi = 0.1, 0.2, 0.4$ (−)
- Matrix diffusion coefficient: $D_m = 0.03, 3.1, 6\times10^{-6}$ (cm^2/s)
- Mean fluid residence time in fractures: $\tau = 200, 320, 8000$ (hr)
- Péclét number: $Pe = 0.05, 8.4, 10$ (−)

Initial values (lots of uncertainty)

Sandia National Laboratories

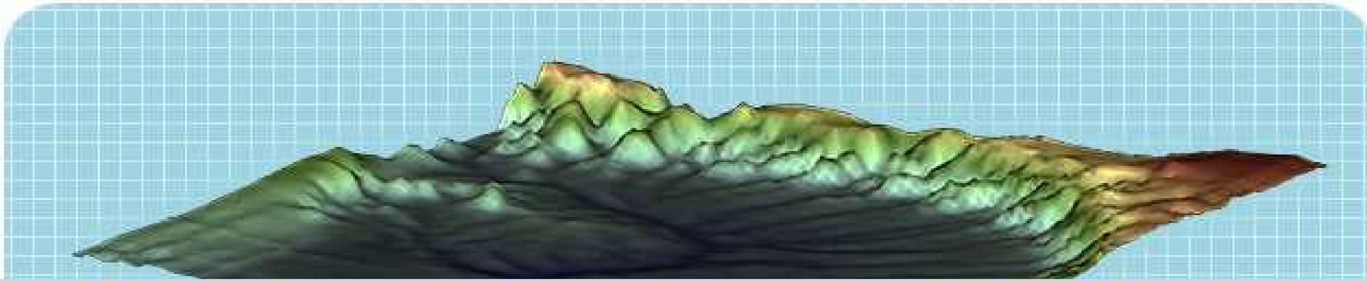
Home

PEST

Utility Support Software

Downloads

PEST FAQ

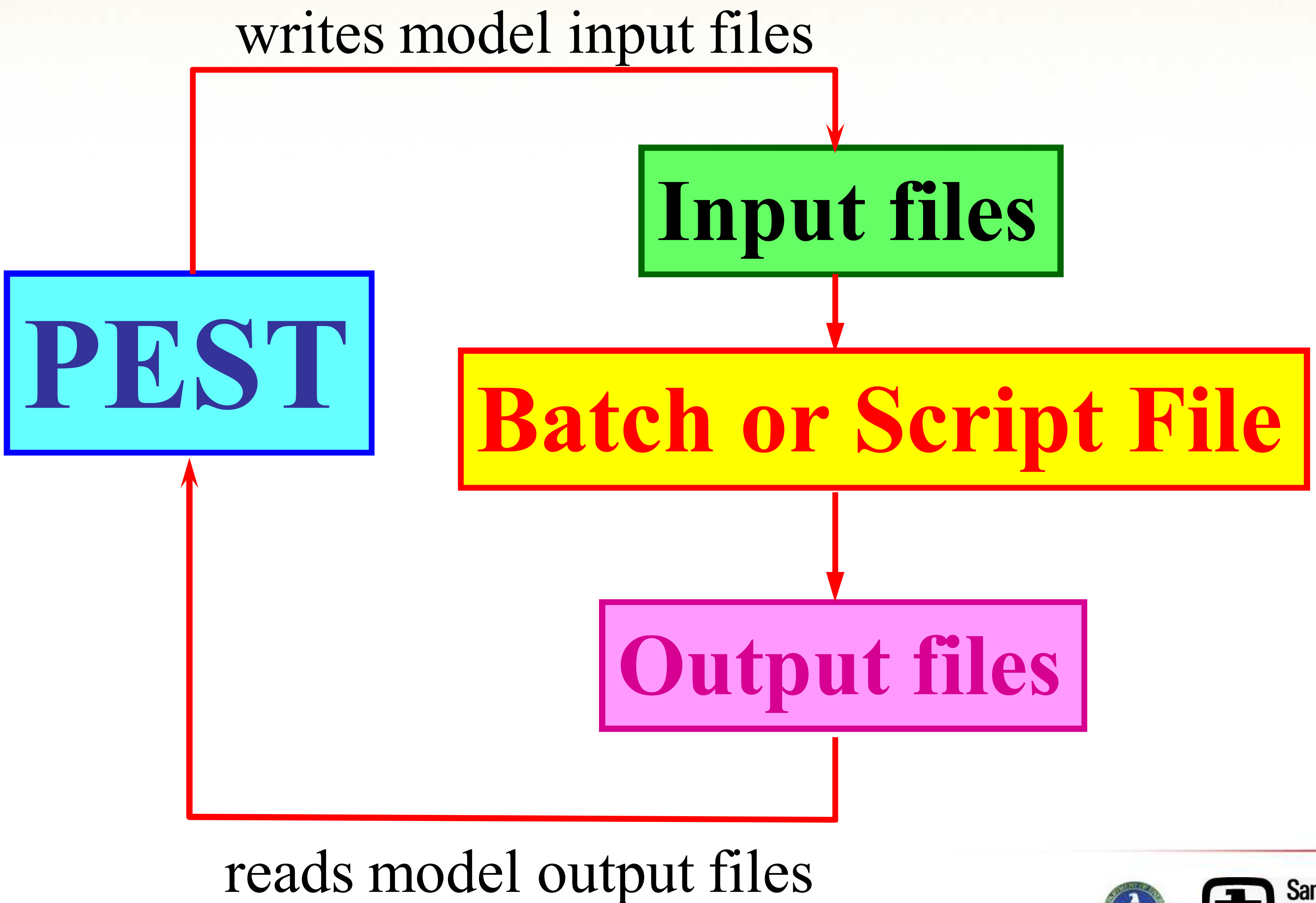


- Model-independent Parameter ESTimation
- Powerful gradient-based optimization code
- Parameter "identifiability"
- Linear and nonlinear uncertainty analyses
- Optimization of data acquisition
- Parameter contribution to predictive uncertainty
- Null-space Monte Carlo (calibration-constrained predictive uncertainty)
- New Pareto mode for predictive uncertainty assessment

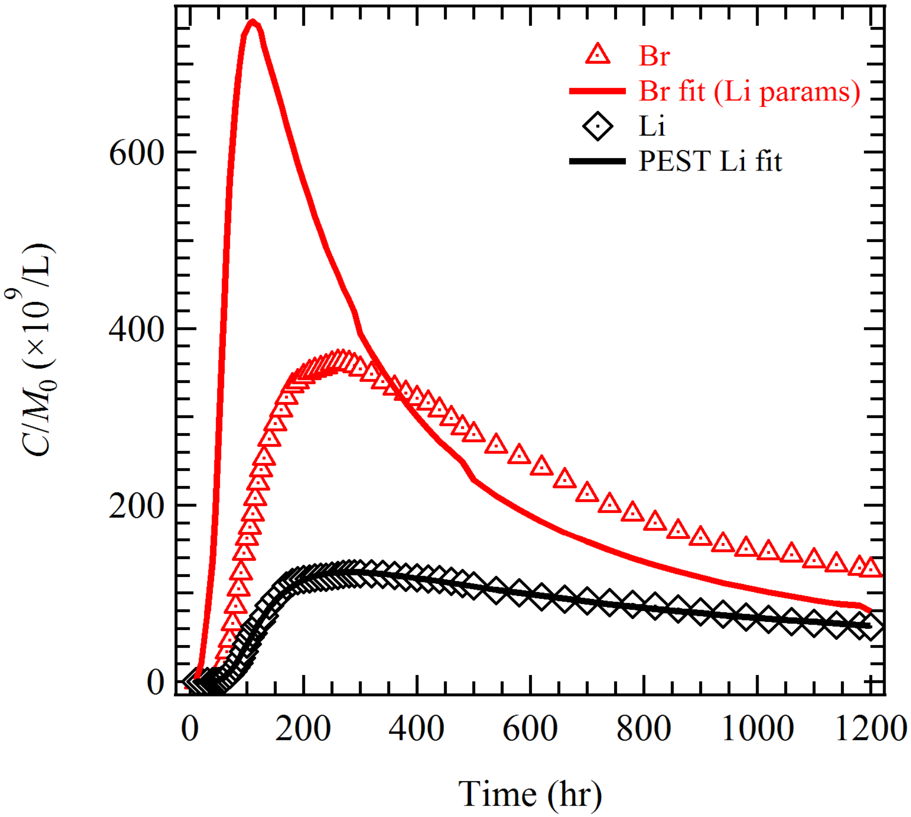
Model-Independent Parameter Estimation and Uncertainty Analysis

Welcome to the PEST web pages.

Model Independent



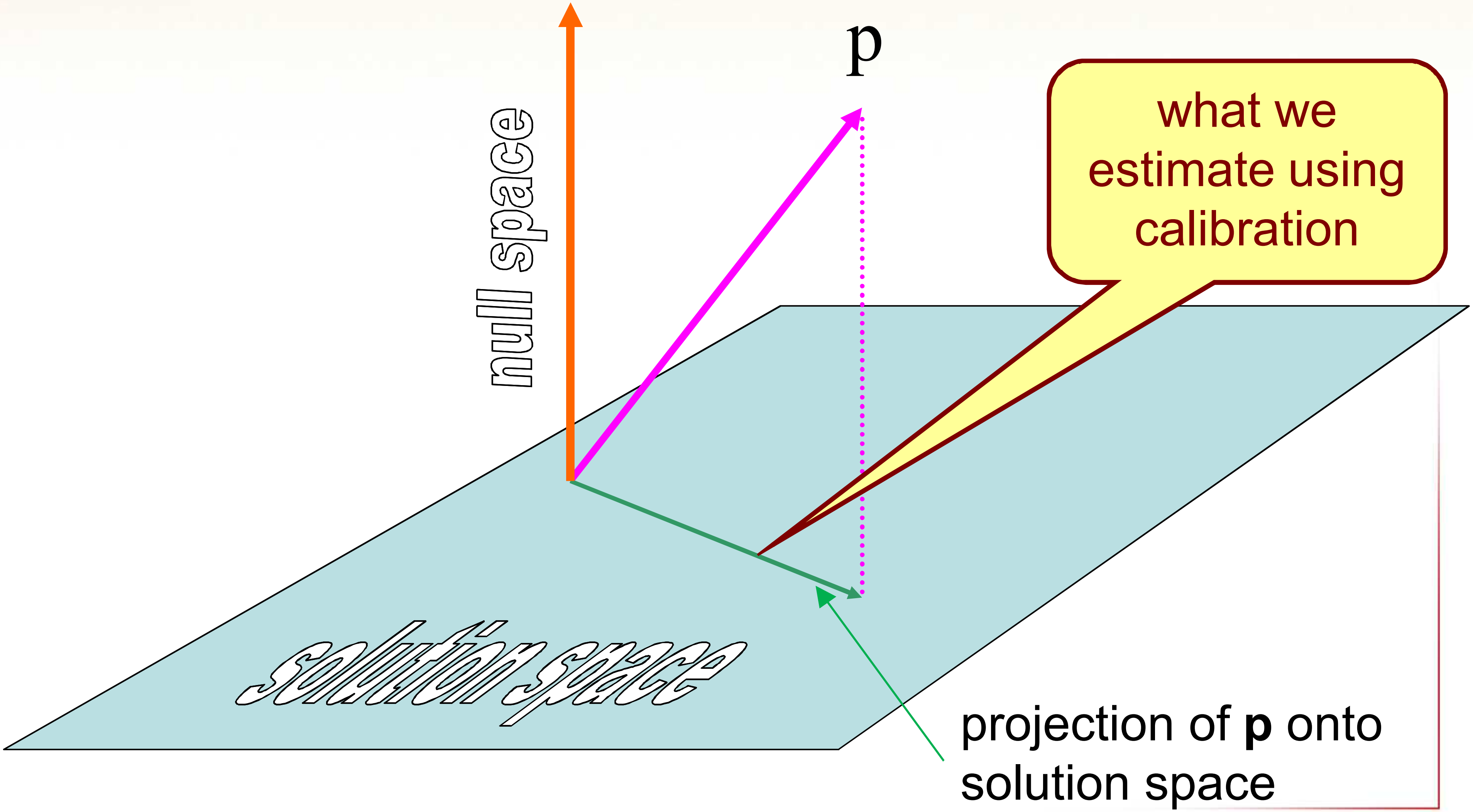
Fit to Conservative Li Only

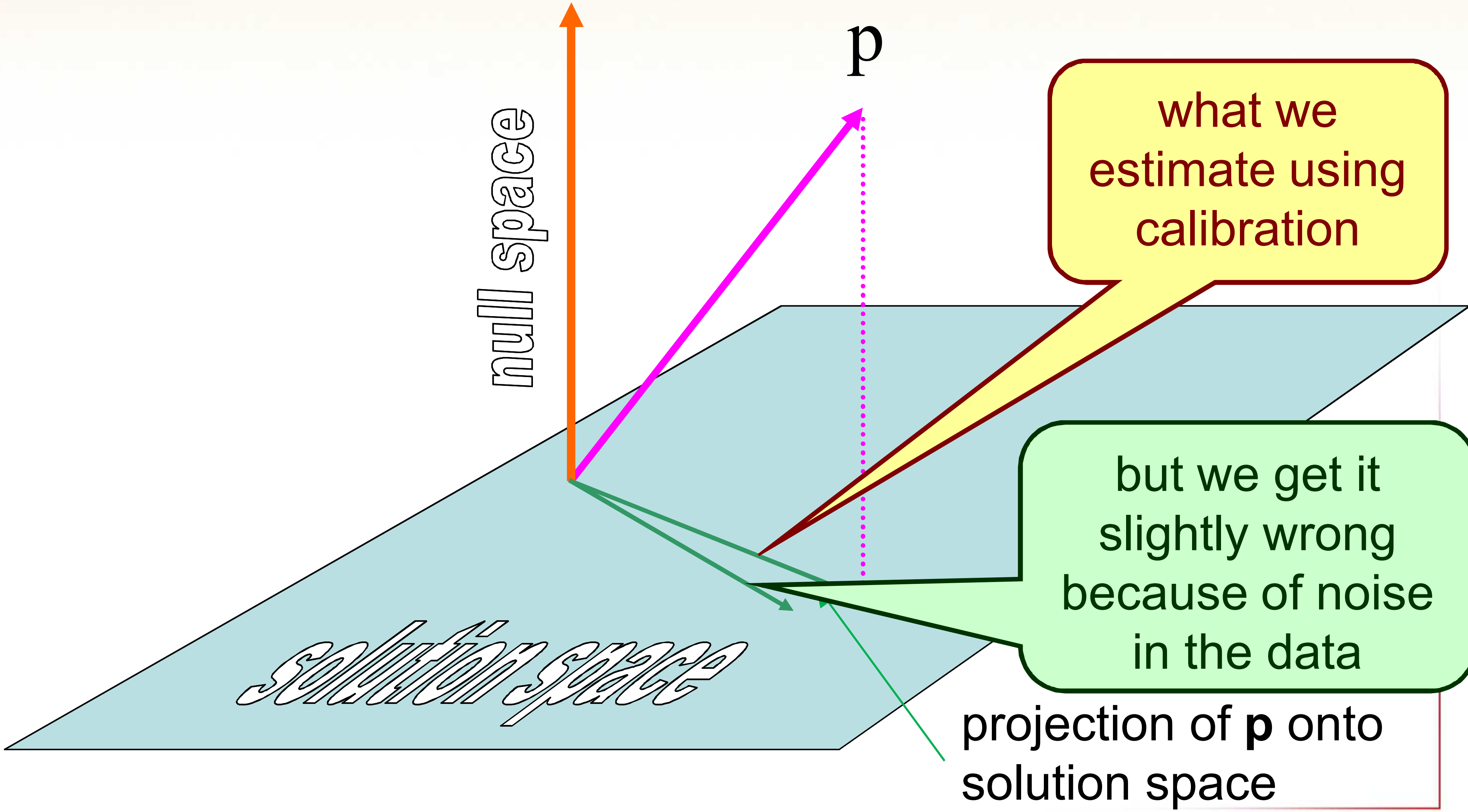


- $f = 0.6 \rightarrow 0.49$ (−)
- $R_f = 5 \rightarrow 2.3$ (−)
- $R_m = 25 \rightarrow 22.8$ (−)
- $b = 0.15 \rightarrow 0.18$ (cm)
- $\phi = 0.2 \rightarrow 0.17$ (−)
- $D_m = 3.1 \rightarrow 2.83 \times 10^{-6}$ (cm^2/s)
- $\tau = 320 \rightarrow 256$ (hr)
- $Pe = 8.4 \rightarrow 1.9$ (−)

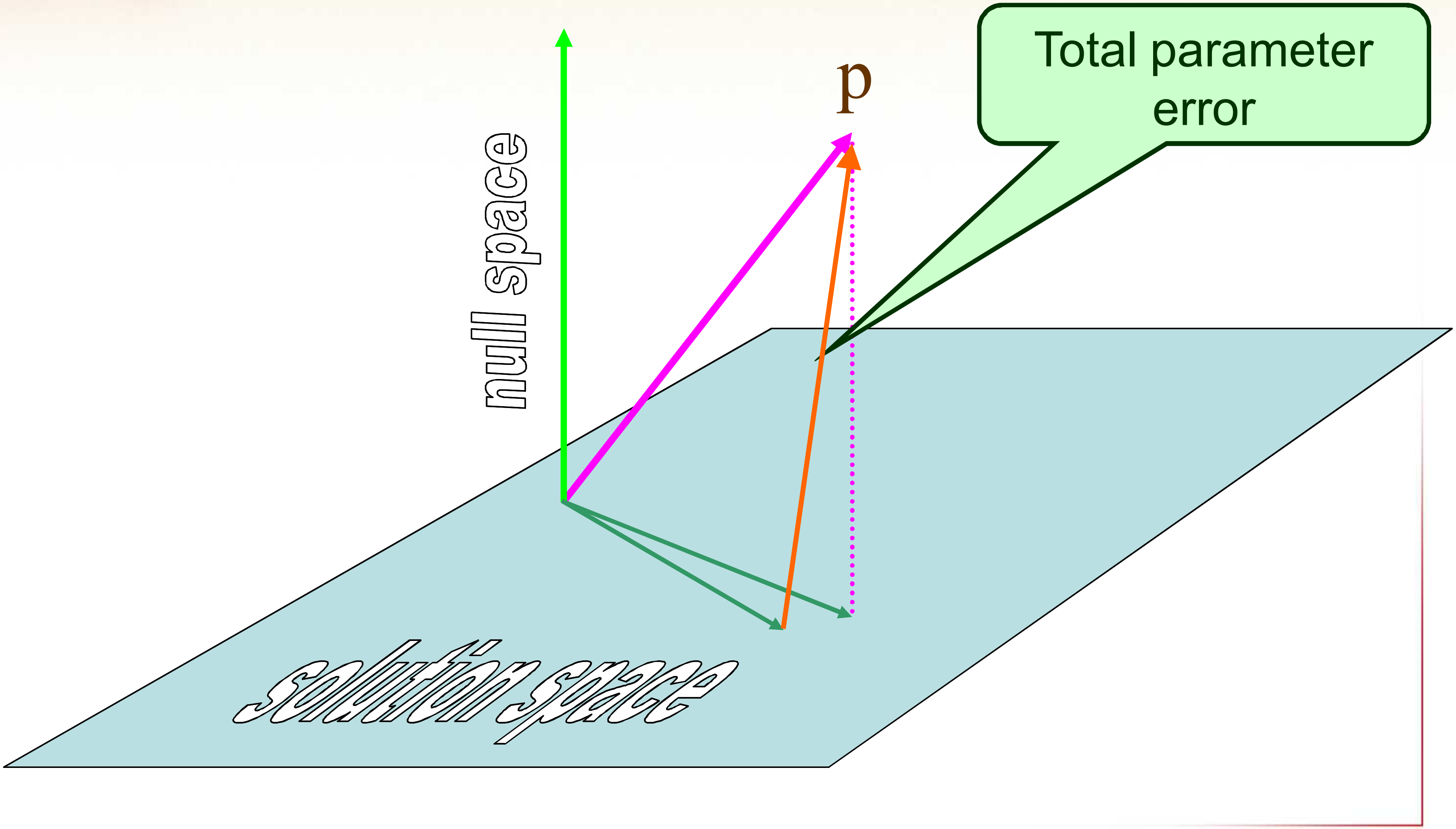
Poor Bromide fit (its data were not used to constrain the parameters)

Sandia National Laboratories


p
null space
what we estimate using calibration
solution space
projection of **p** onto solution space

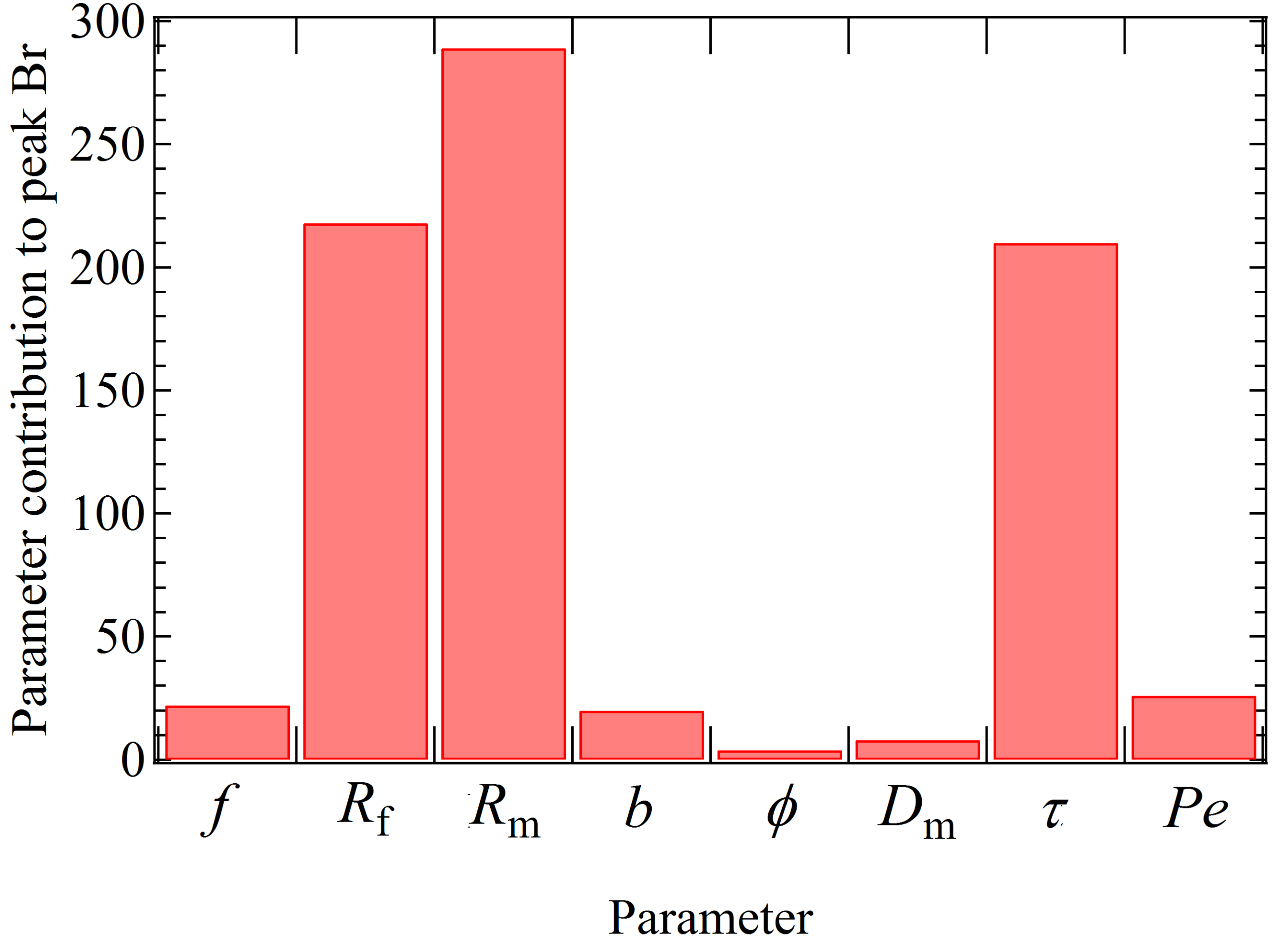


p
null space
what we estimate using calibration
but we get it slightly wrong because of noise in the data
solution space
projection of **p** onto solution space
Sandia National Laboratories



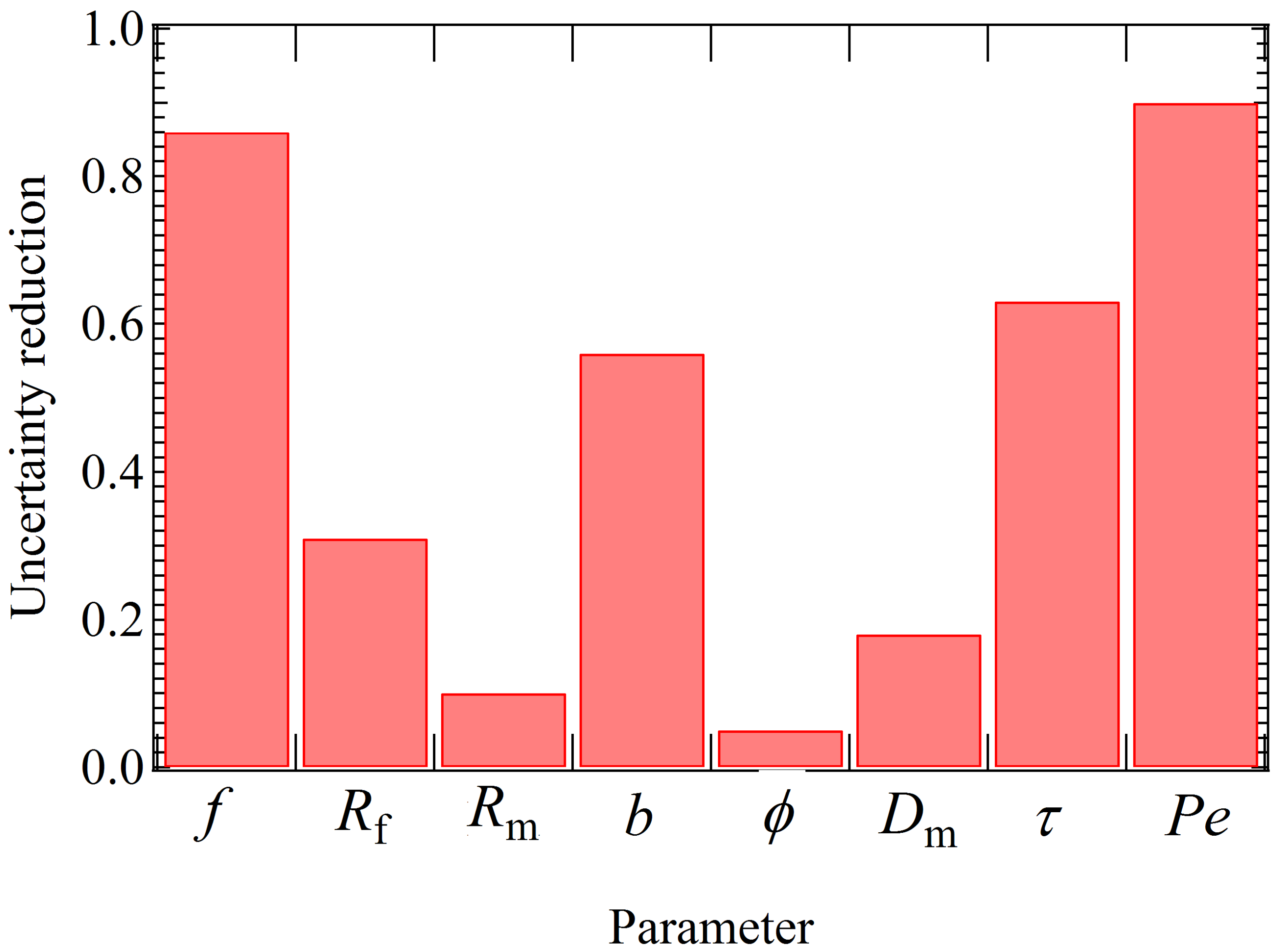
Total parameter error
p
null space
solution space

Parameter Contribution to Peak Br



Peak is 2,100

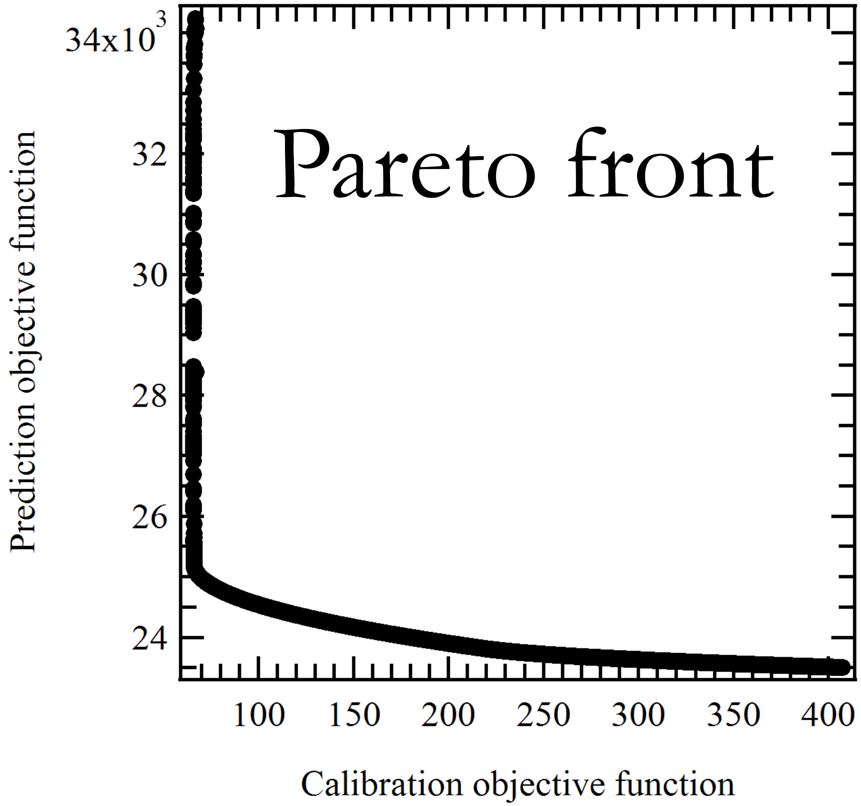
Reduction in Parameter Uncertainty



Pareto Analysis

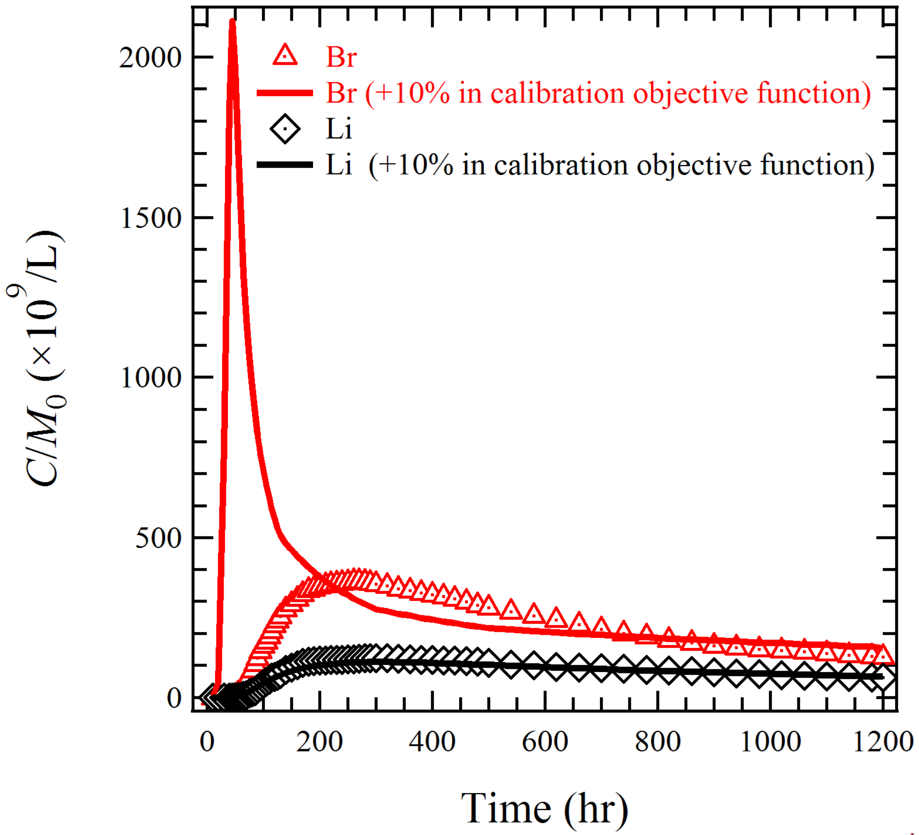
- The Br fit was poor, but how poor could it be while still honoring the Li data?
- Li (calibration) objective function: Minimize the sum of squared differences between Li data and RELAP simulation.
- New Br (prediction) objective function: Maximize peak Br.
- Allow an arbitrary 10% increase in the Li objective function.

Pareto Results

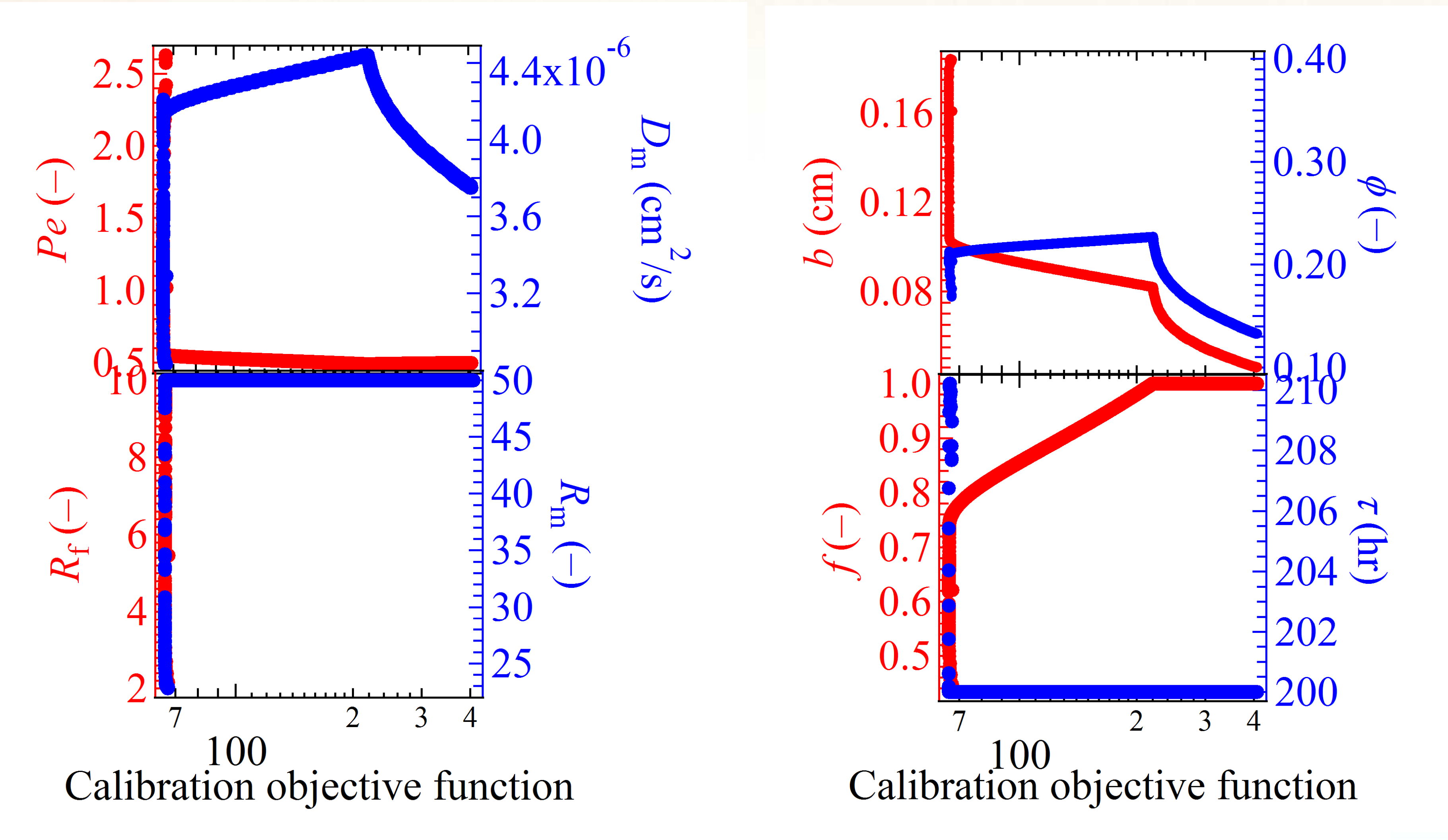


Sharp decrease in prediction objective function indicates that there are many values of parameters that can maximize peak Br and still fit Li well.

Peak increased from 740 to 2,100!



Parameters



Note the log axis and confluence of points on left of the *x*-axis.

Uncertainties

- Parameter uncertainty.
- Measurement noise in the data set.
- Typically large null space (the data do not inform the parameter values).

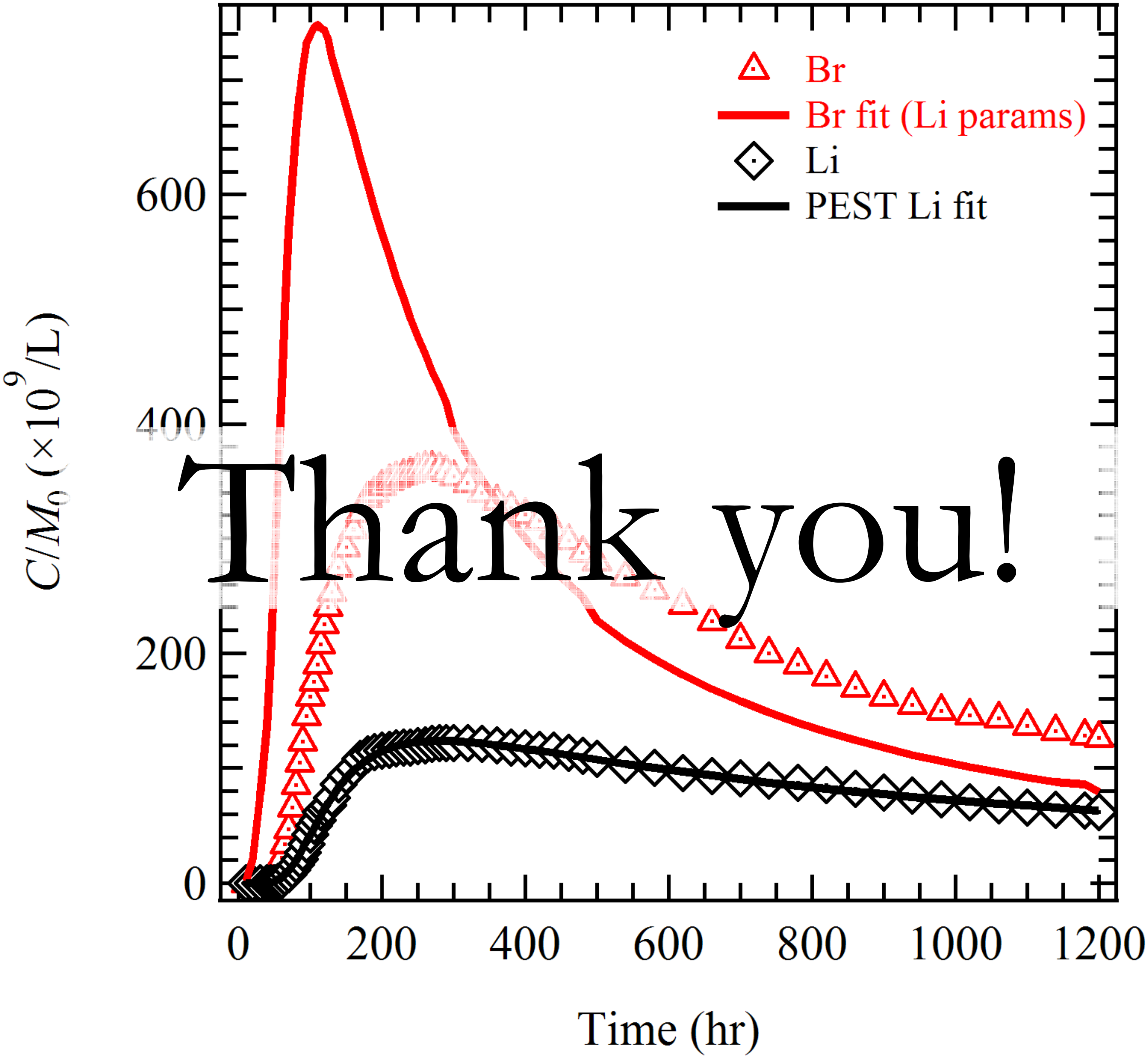
These are not the fault of the model, but of data and expert knowledge inadequacy!

If Not a Correct Prediction Then?

- The correct prediction lies within model-computed uncertainty intervals.
- These intervals are reduced to as close to their theoretical minimum as possible (strictly a function of information content of data and expert knowledge).

Take Home Message

This example should instill decision makers with the concept that all models should be thoroughly interrogated as to their predictive uncertainties. Even as seemingly excellent fits are made to data, non-uniqueness can allow for significantly different predictions for a calibration objective function that is not too far from its minimum.



Thank you!

Sandia National Laboratories