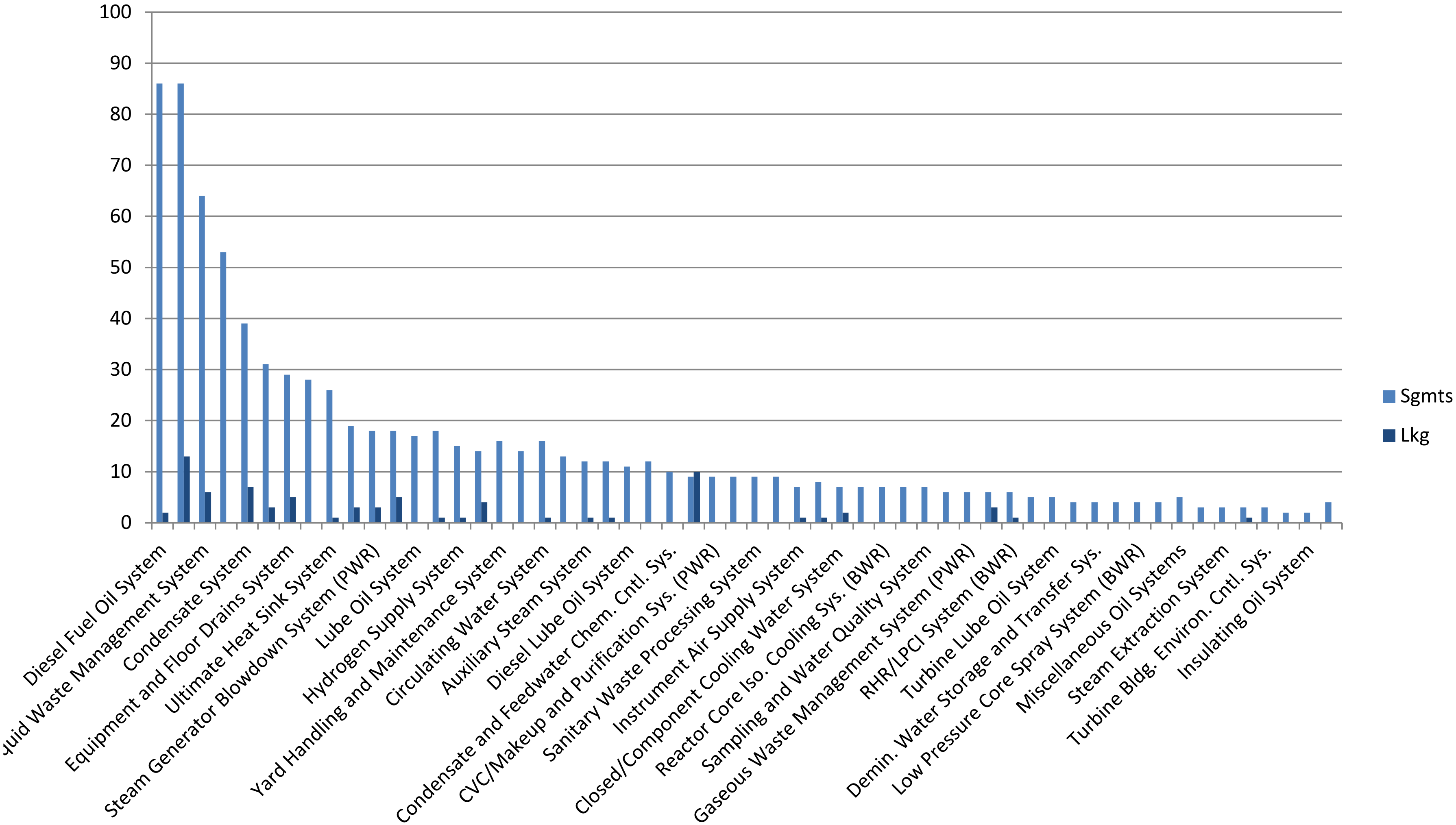
Targeting Data Collection for Decision Support

Ty Ferré
Hydrology and Water Resources
University of Arizona

- NSIAC approved the Underground Piping and Tank Integrity Initiative in November 2010
  - Added scope
    - Underground piping and tanks whether or not they are in direct contact with the soil if they are outside of buildings and
      - Contain licensed radioactive material or
      - Are safety related
- Goal is reasonable assurance of structural and leakage integrity of in-scope piping and tanks with special emphasis on components containing licensed material

Riley – problem definition

Leakage Events



Riley – problem definition

Equipment

- LDM acquisition systems
  - Installed at Hanford since 200
  - >99.99% reliable
  - Accurate
  - UL rated
  - NQA-1
- Geotection
  - 180 channels
  - UL Rated
  - Undergoing V&V

Rucker – imaging

Proposed In-situ Real-time Leak Detection System

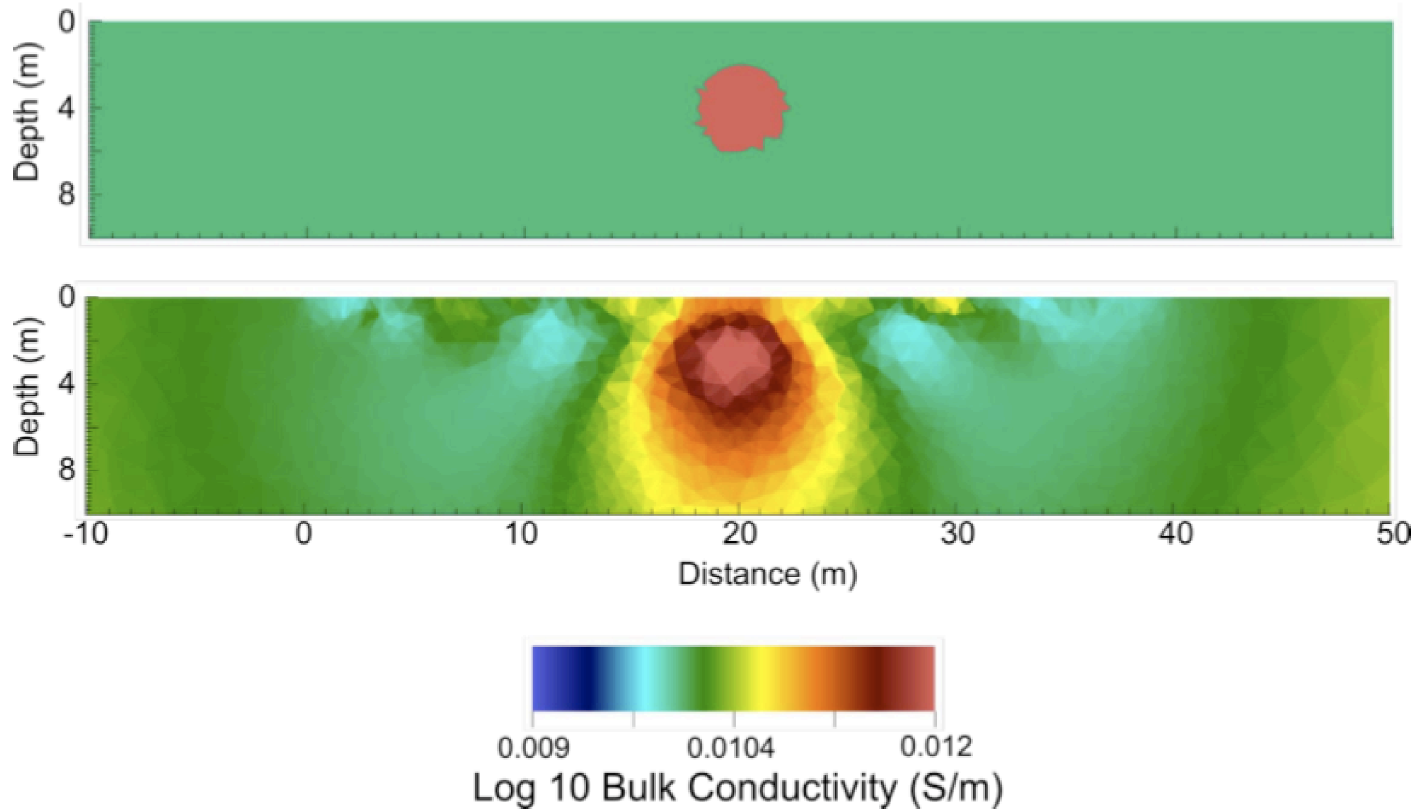
- A multiple sensor system including
  - Acoustic sensor
  - Moisture sensor
  - Temperature sensor
  - Radiation sensors (tritium sensor)
    - Plastic scintillators to cover a large area
    - Monolithic active pixel sensors for tritium autoradiography (Nuclear Instruments and Methods in Physics research A 543, 537-548, 2005)
    - Liquid scintillation counting coupled with fiber-optics
    - Laser absorption spectroscopy coupled with fiber optics
- Leak detection algorithms
  - Sensor fusion and data processing
  - Subsurface modeling for leak location

This work was sponsored by the U.S. Nuclear Regulatory Commission under Office of Research

Inversion Challenges

True bulk conductivity

ERT image



Johnson – inversion with filtering

Joint Hydrological-Geophysical Inversion

- *Geophysical* data provide *high-resolution* information on soil *structure* and state
- *Hydrological* data provide *process-specific* information on soil *properties* and state
- The hydrological forward model provides *physical regularization* to geophysical imaging
- Joint hydrogeophysical data analysis *reduces ill-posedness* of inverse problem
- Joint estimation of geometry and properties *reduces estimation and prediction bias*

Some basic considerations

IMAGING MIGRATION OF LEAK CONSTITUENTS WITHIN THE SOIL

Is there a strong

DETECTING

BOTH

Is there a geophysical technique with the spatial

THESE QUESTIONS SHOULD BE ADDRESSED VIA SYNTHETIC STUDIES **PRIOR TO EXPENDITURE IN THE FIELD**

source to measurable perturb the electrical potentials?

Slater - prescreening

If “no” then do not pursue a geophysical survey!!!

DIRECT

Discrimination/Inference to Reduce Expected Cost Technique

Motivation

Our data are sparse;

our models are uncertain;

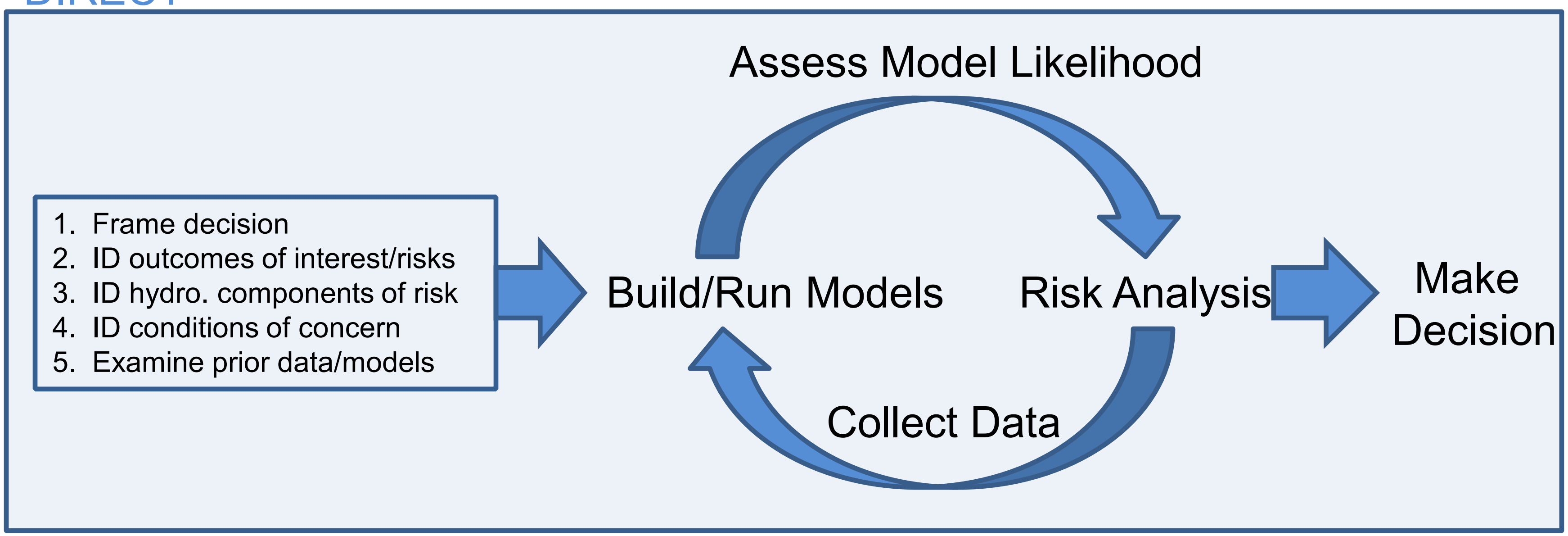
but, we must decide.

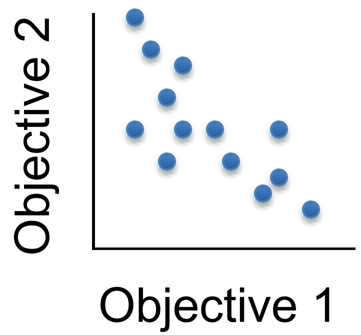
Traditional Approach

Frame Hydro. Problem → Collect Data → Build/Run Model → Make Decision

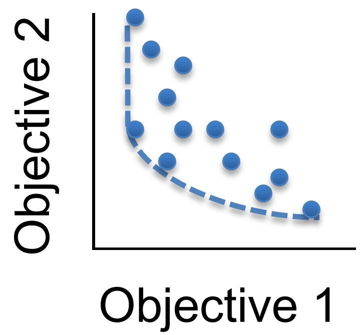
Traditional Approach

DIRECT



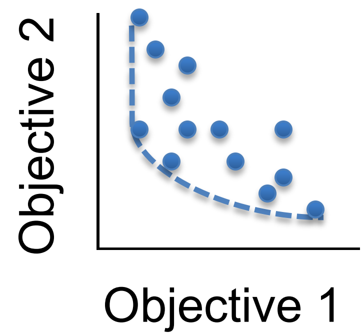


Most decisions involve trade-off solutions: risk/cost; development/environmental protection; treatment A/B.



Most decisions involve trade-off solutions: risk/cost; development/environmental protection; treatment A/B.

Tools have been developed to find the Pareto solutions.



Most decisions involve trade-off solutions: risk/cost; development/environmental protection; treatment A/B.

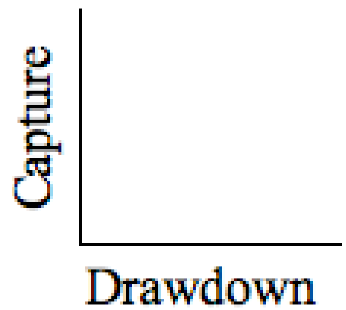
Tools have been developed to find the Pareto solutions.

Few of these methods address the influence of uncertainty. Fewer still offer guidance on which data to collect to reduce relevant uncertainties and improve decisions.

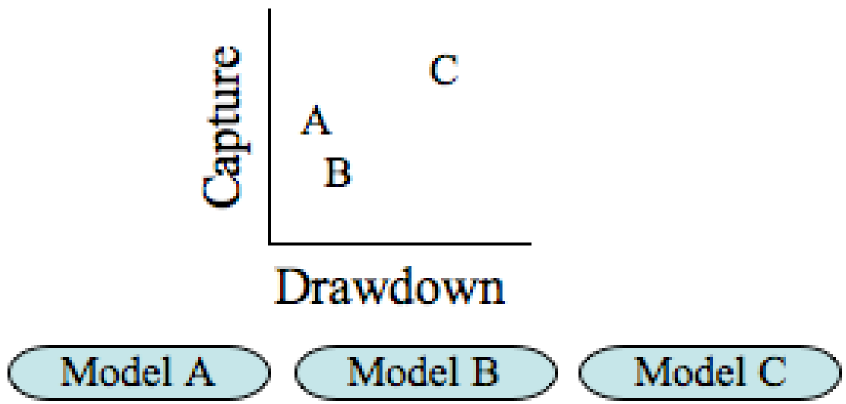
DIRECT is a two part process based on multimodel analysis that uses the model ensemble to:

1. identify likely valuable measurements, and

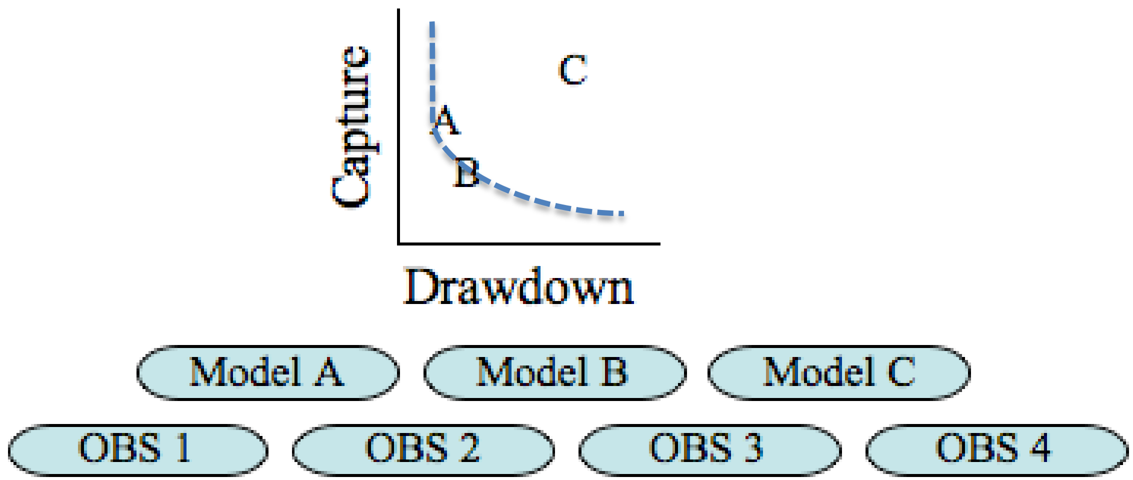
2. optimize decisions under uncertainty.



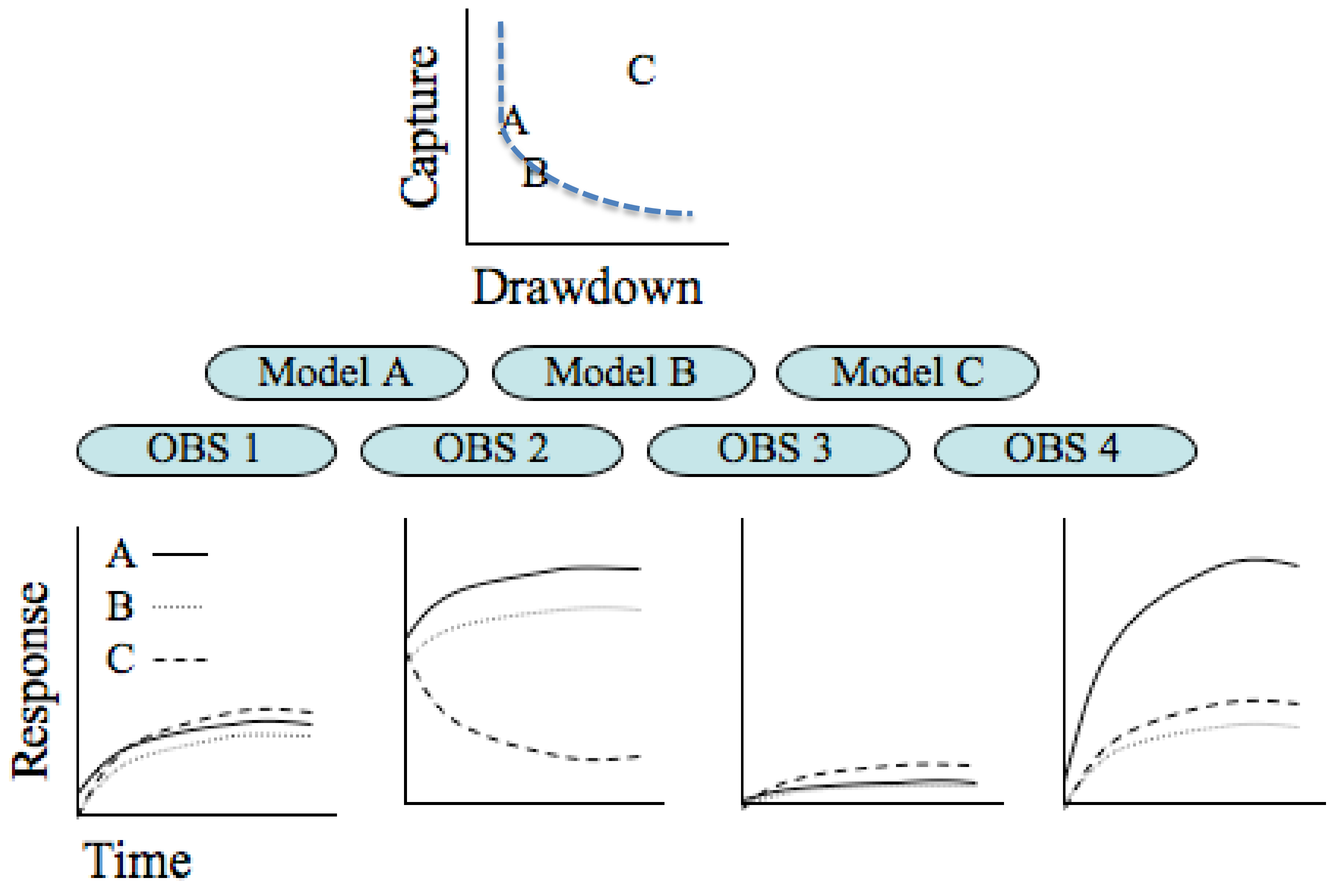
Consider a water resource problem that trades off capture with drawdown.



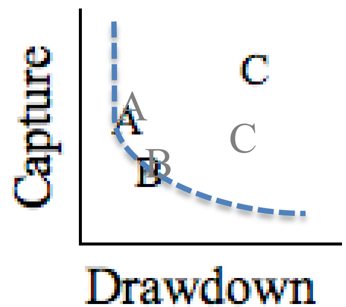
Given existing data, we have a high degree of uncertainty regarding the system response, as shown by the range of predictions of our model ensemble.



We are considering multiple future measurements to help resolve this uncertainty, hopefully leading toward definition of a Pareto solution.



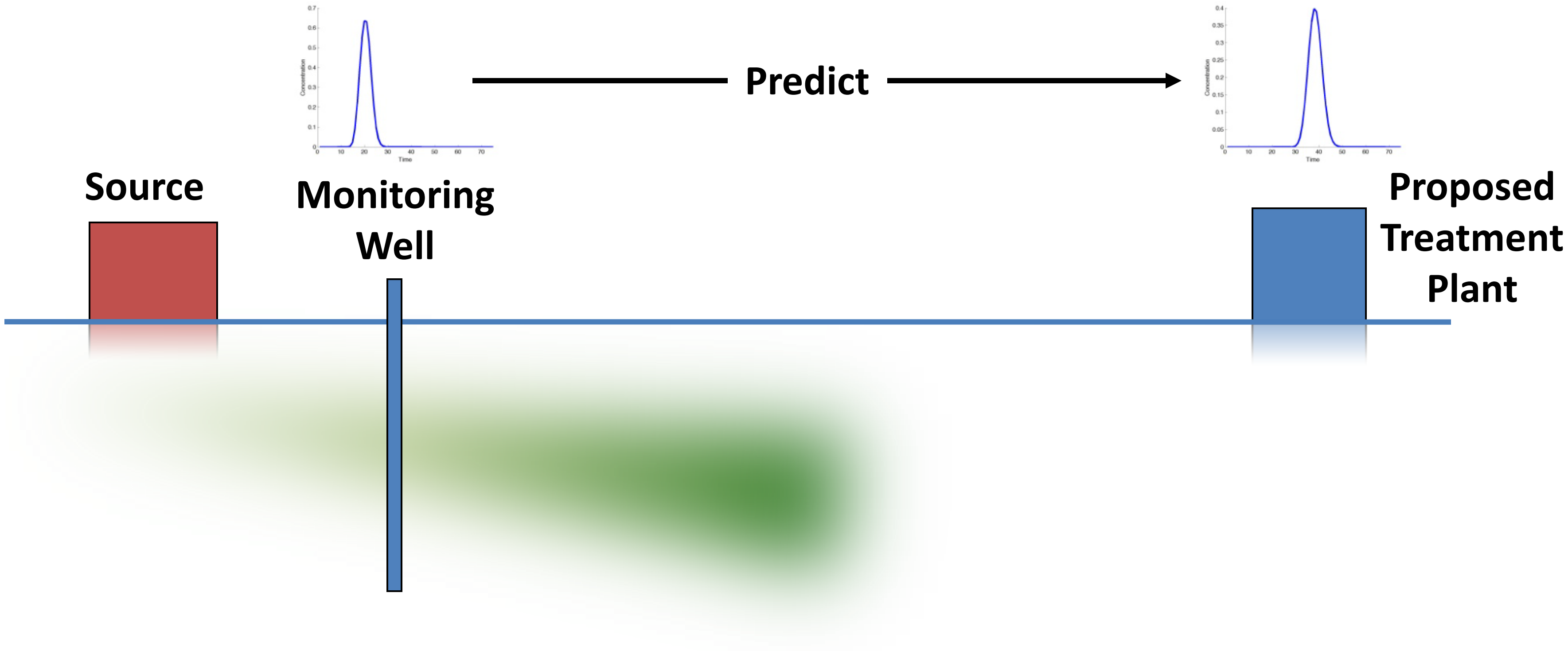
DIRECT is based on choose observations that are most likely to *discriminate* among important, high-likelihood models and that preferentially inform predictions of interest.



After updating with new data, it is still likely that we will have prediction uncertainty. The second part of DIRECT deals with using the model ensemble to make optimal decisions.

DIRECT

Simple Contaminant Transport Example

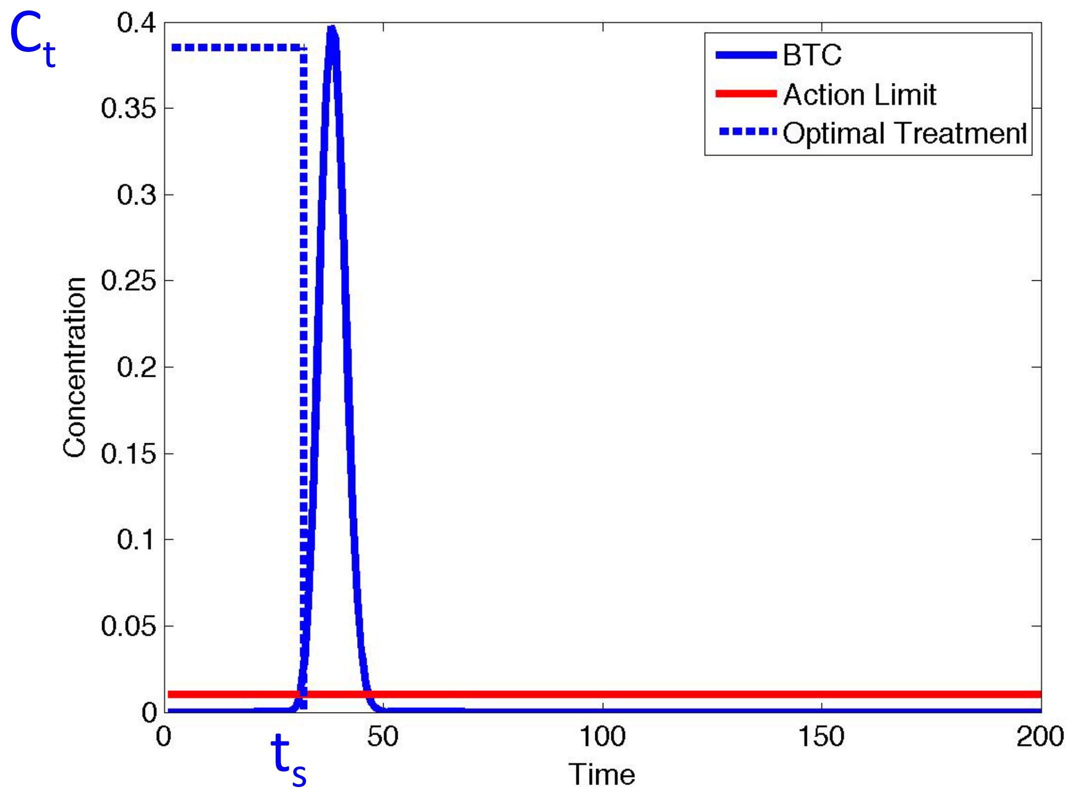


Decision:

When should treatment start?

How aggressively to treat (design concentration)?

Lowest-Cost Treatment Based on BTC



Treatment decision variables

Maximum concentration (C_t) and start time (t_s)

Objective: Minimize Total Cost

Total Cost = *Capital* + *Operations* + *Penalty*

$$\text{Capital Cost} = K_1 * C_t$$

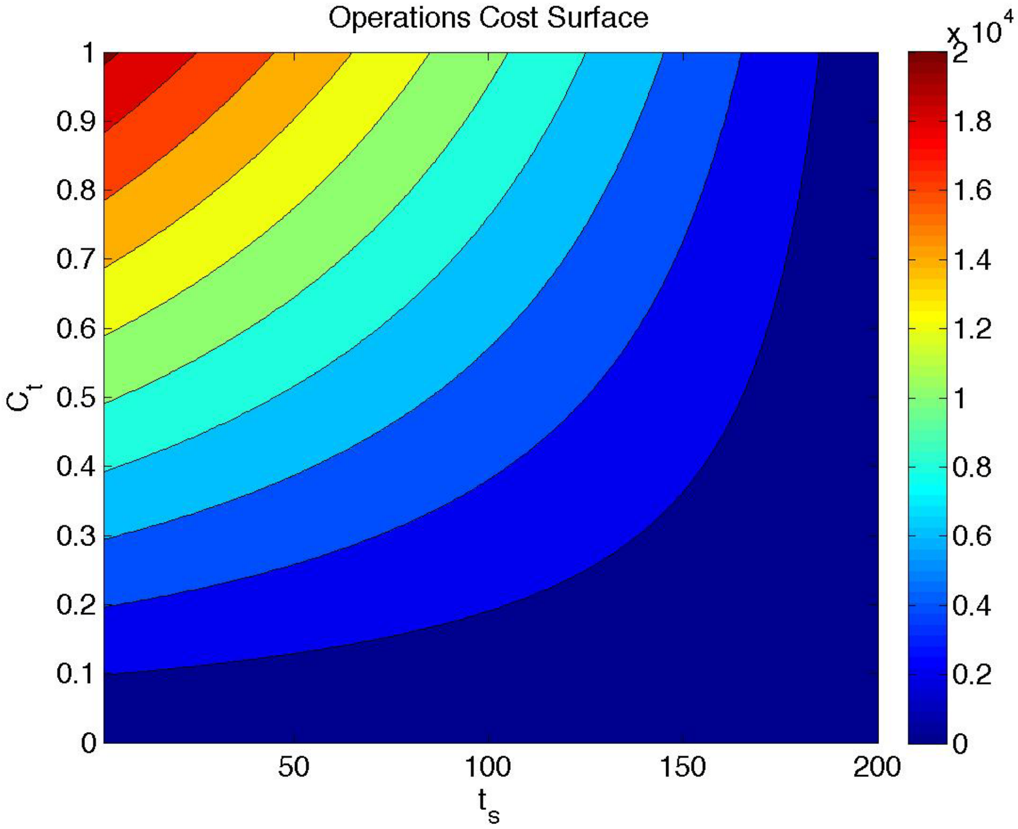
$$\text{Operations Cost} = K_2 * C_t * (t_{final} - t_s)$$

$$\text{Penalty Cost} = K_3 * n_{exceedence}$$

$$Exceedence = C(t) - C_t > C_{action}$$

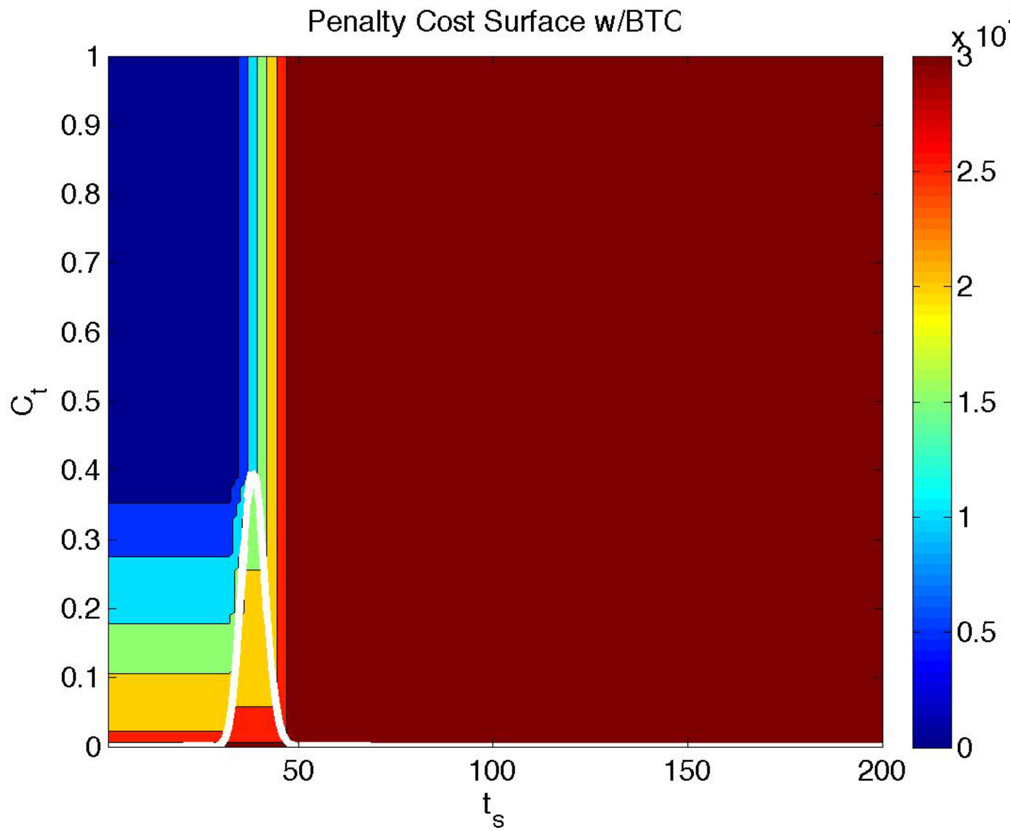
$\mathbf{K_\#}$ = problem-specific constants
C = contaminant concentration
$\mathbf{C_{action}}$ = maximum allowable concentration
$\mathbf{C_t}$ = treatment design concentration
$\mathbf{t_s}$ = treatment start time
n = number of monitored periods

Capital & Operations Cost Surface



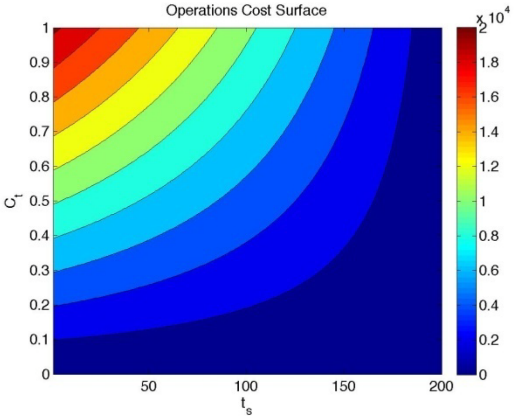
The capital and operations costs don't depend on the true BTC, so they can be defined exactly *a priori*.

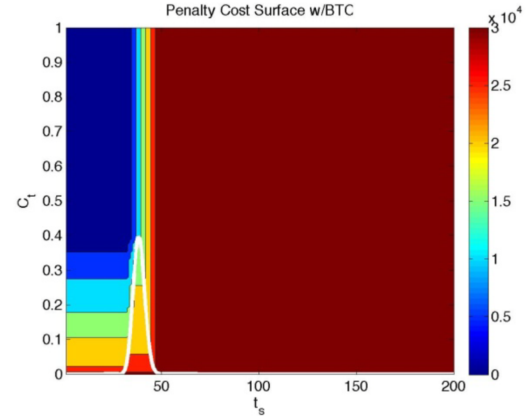
Penalty Cost Surface

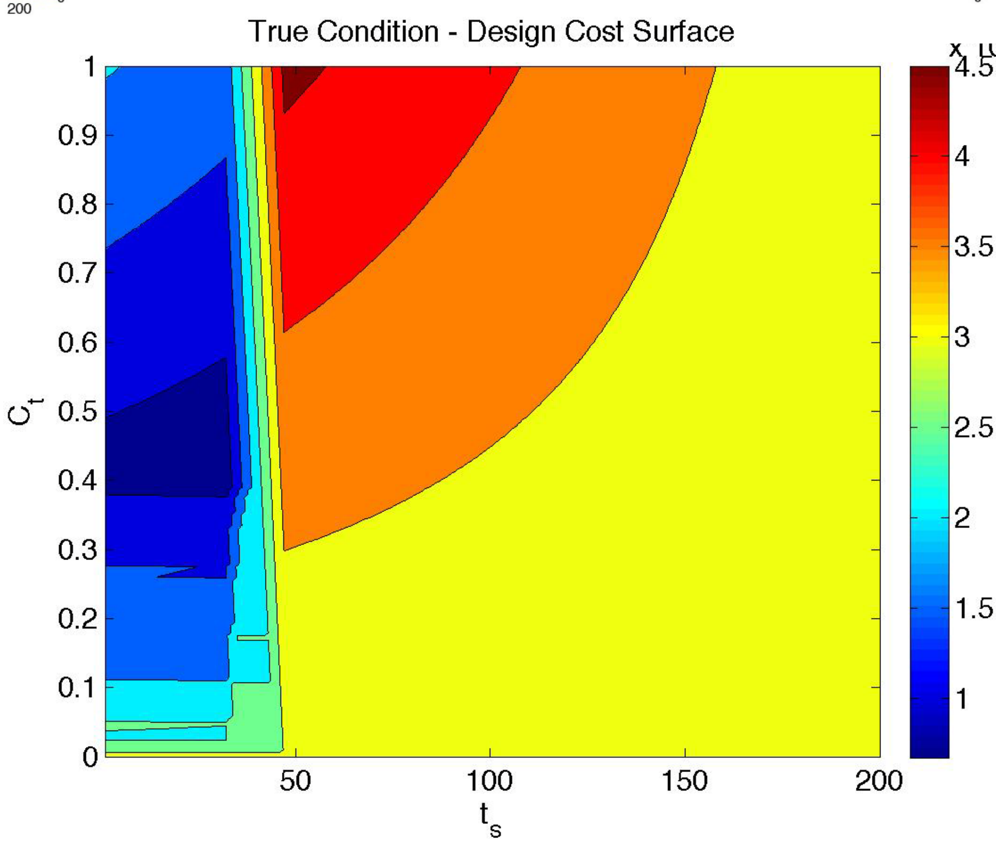


The penalty costs depend on the selected C_t and t_s relative to the *true* BTC. They cannot be defined without knowledge of the system.

Total Cost

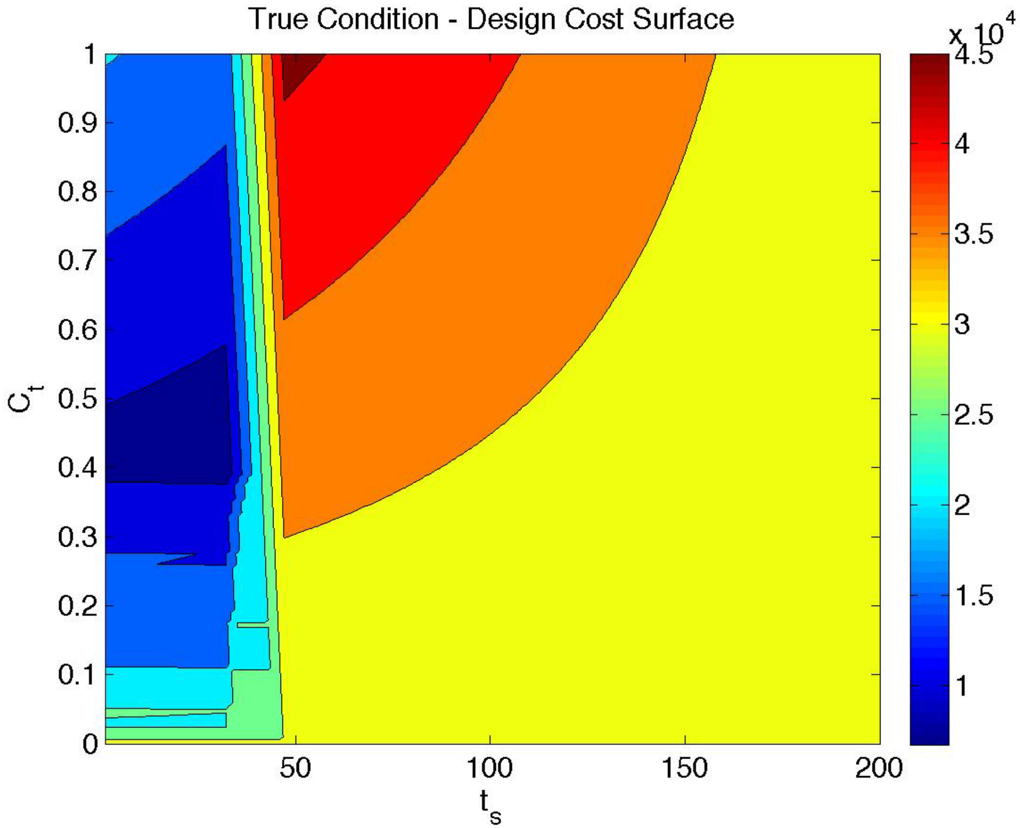






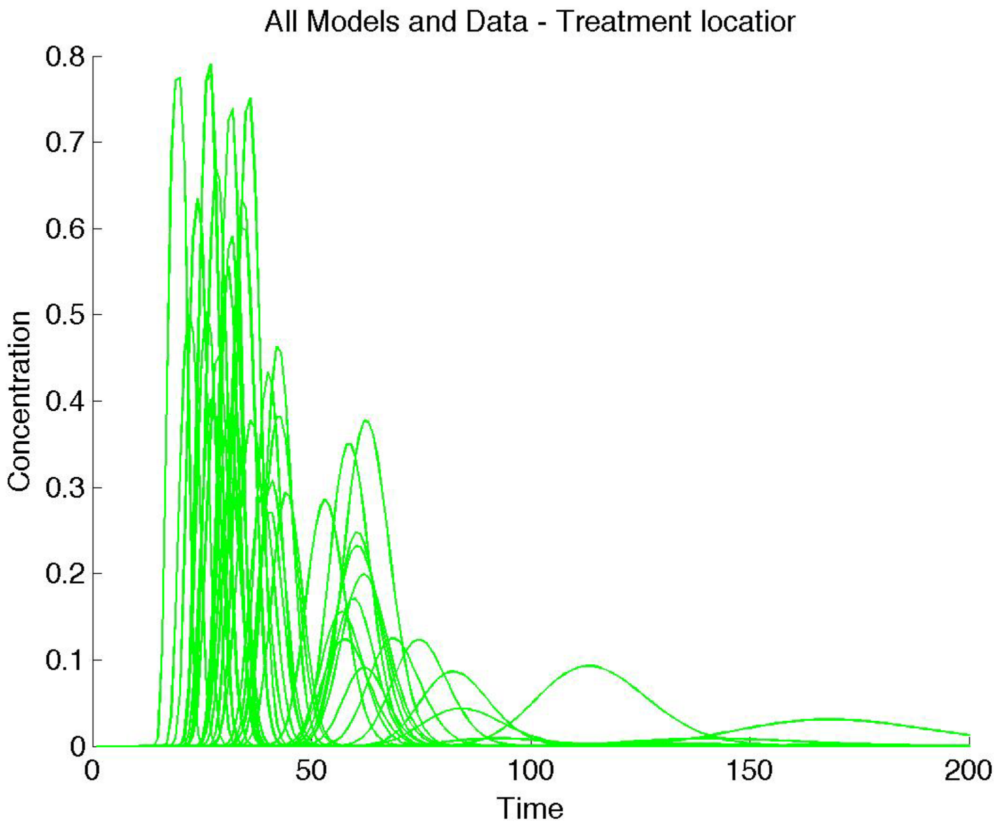
The total cost surface is the sum of the operations and penalty cost surfaces. It depends on the true BTC.

Design Surface



The project objective is to minimize total cost. The hydrologic objective is to provide information to identify the optimal design.

Sparse Data → Many Possible BTCs

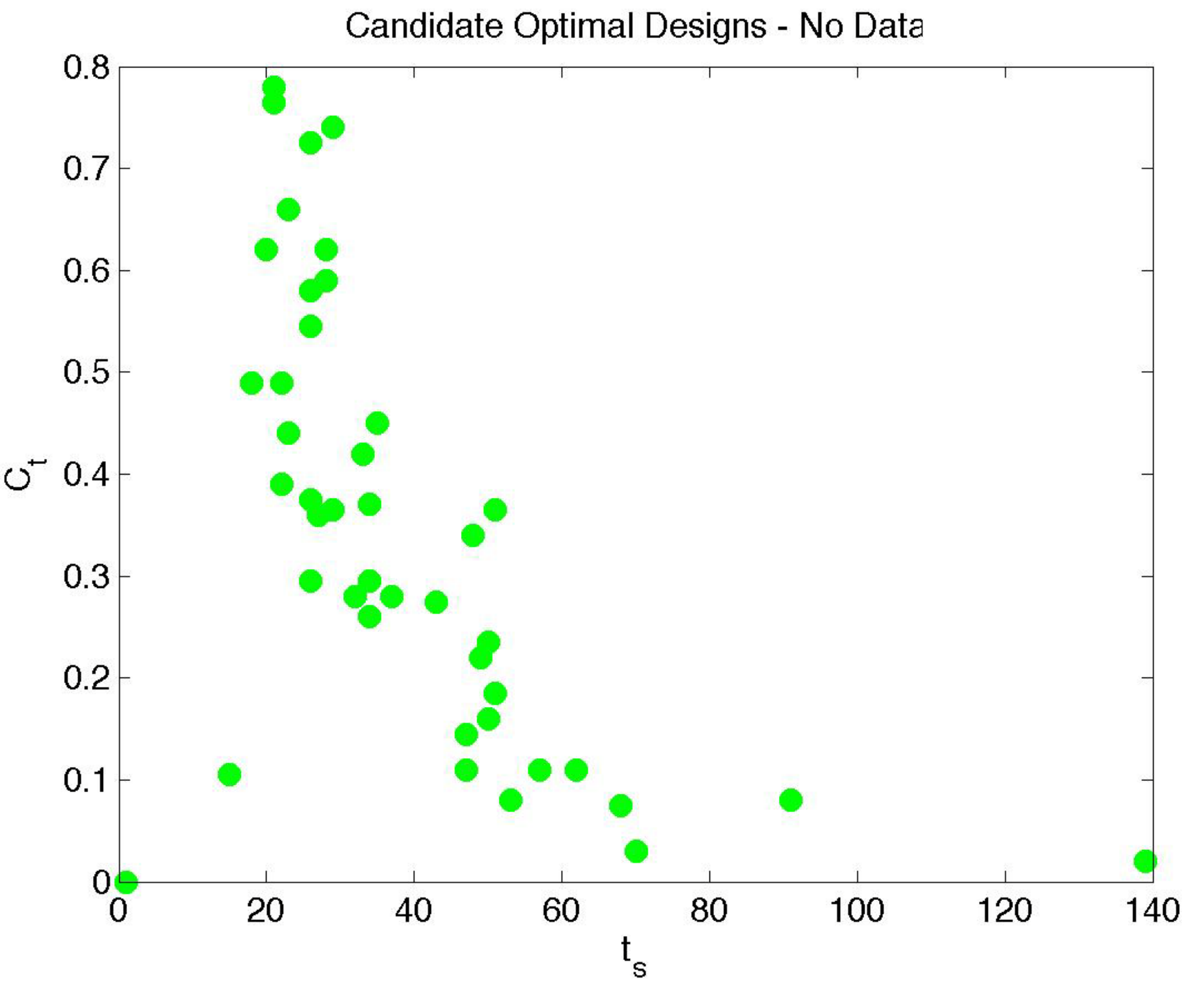


Varied Parameters

- Hydraulic Conductivity
- Retardation
- Decay

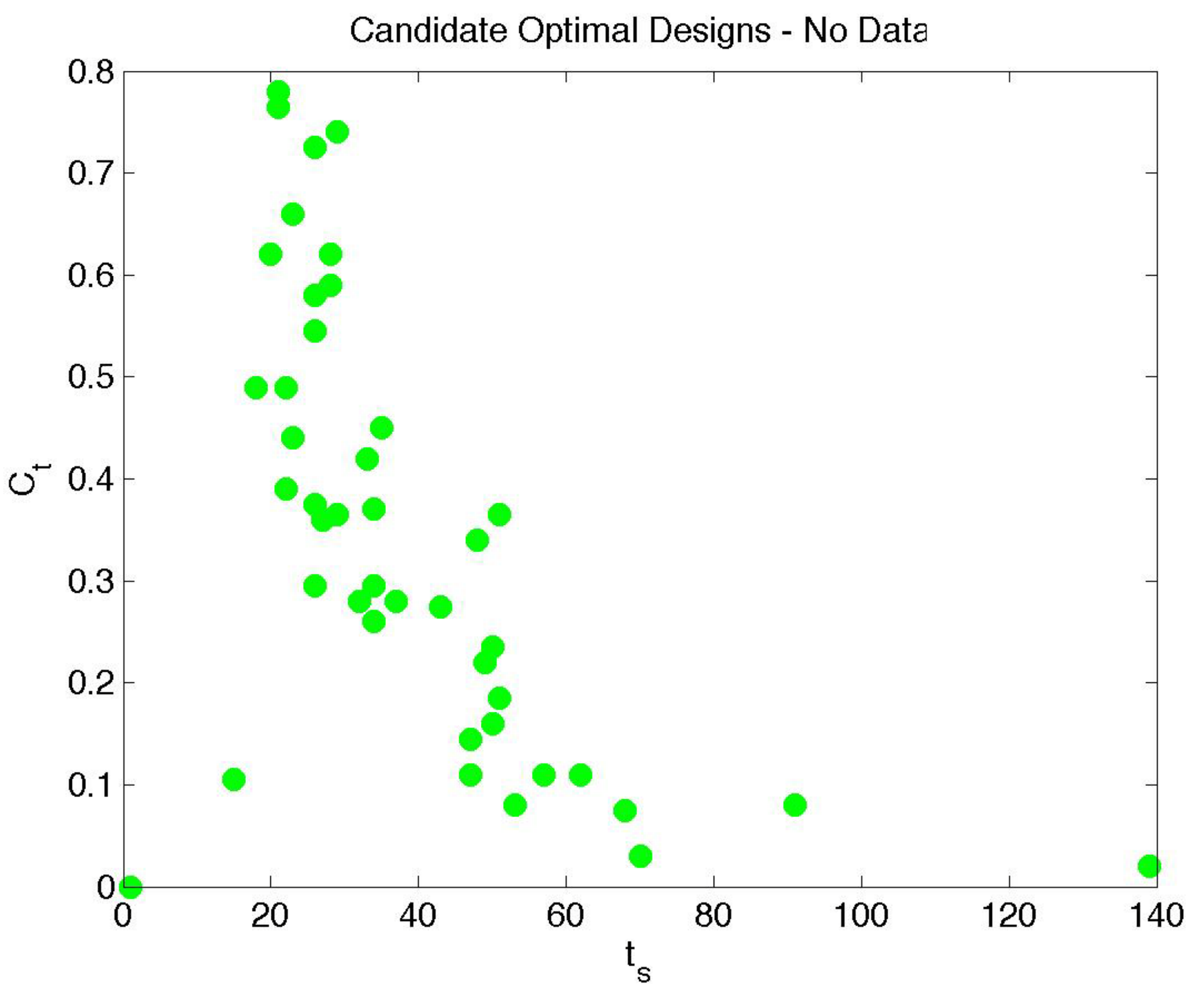
Geologic and limited hydrologic information can define parameter ranges. But, this can lead to a wide range of *predictions of interest*.

Deterministic Design



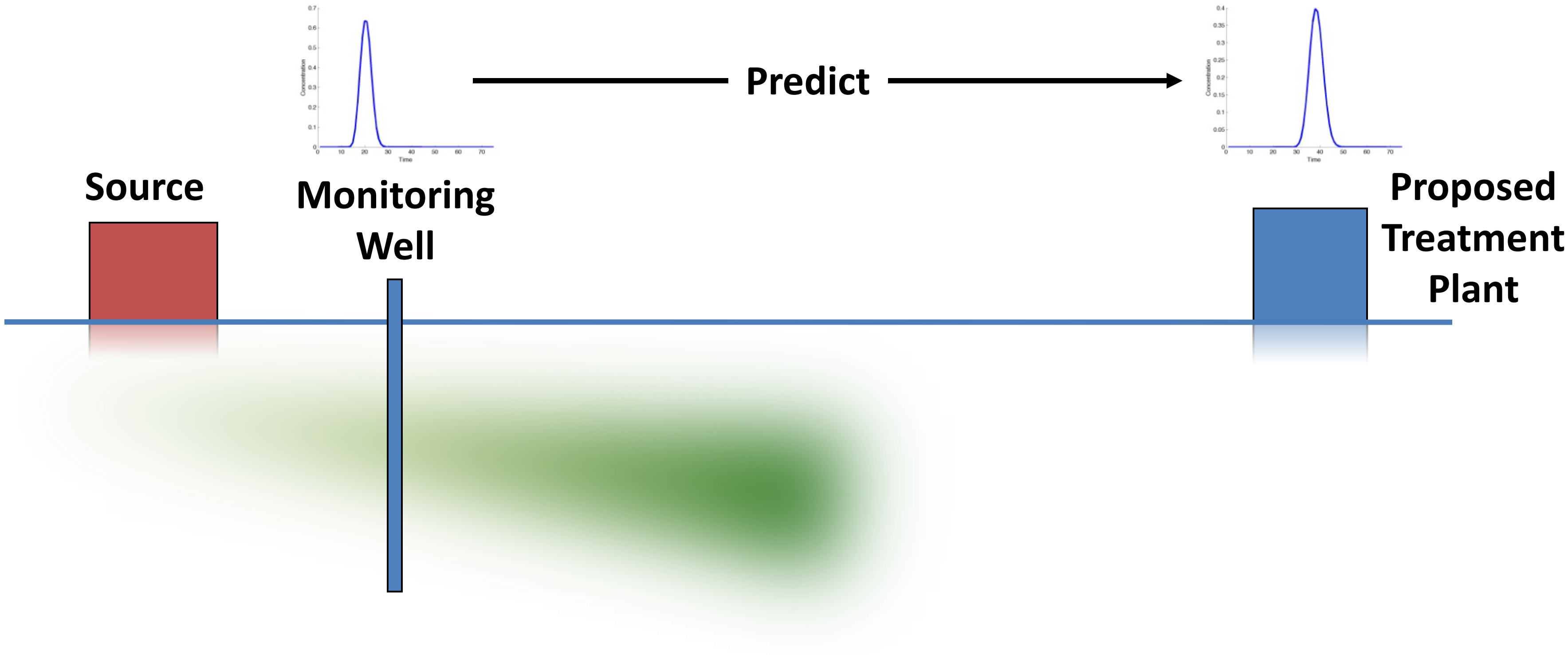
Each model-based prediction has a unique total cost surface, leading to a unique lowest-cost design.

Deterministic Design



Traditionally, the role of hydrologists has been to find the *true* conditions, to select the appropriate optimal design.

Simple Case Study

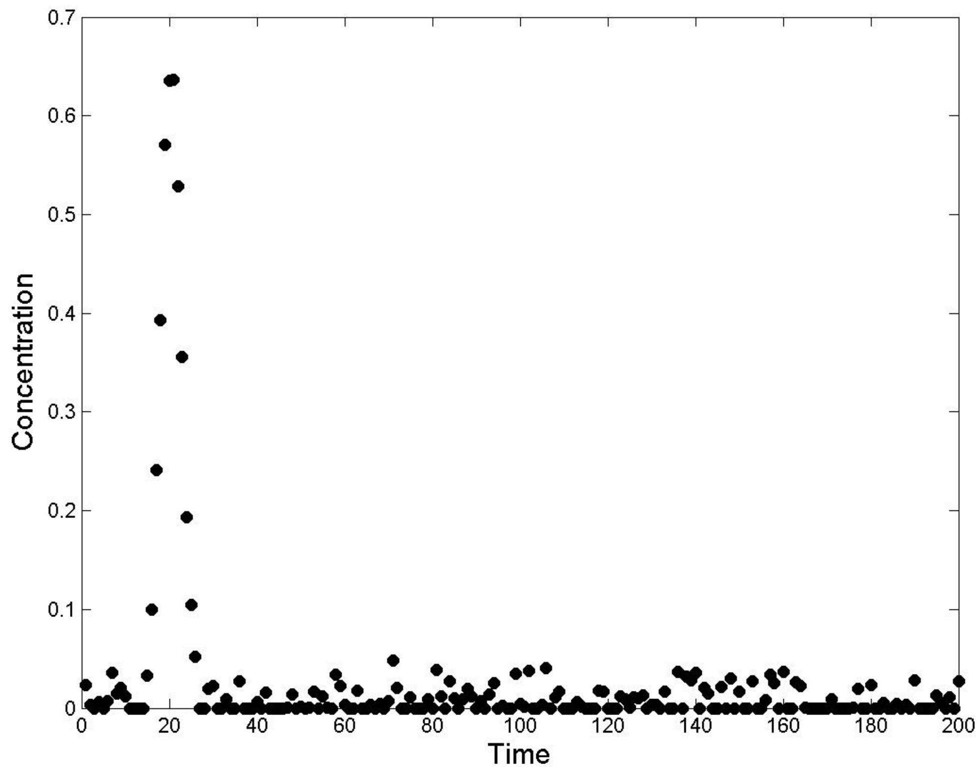


Decision:

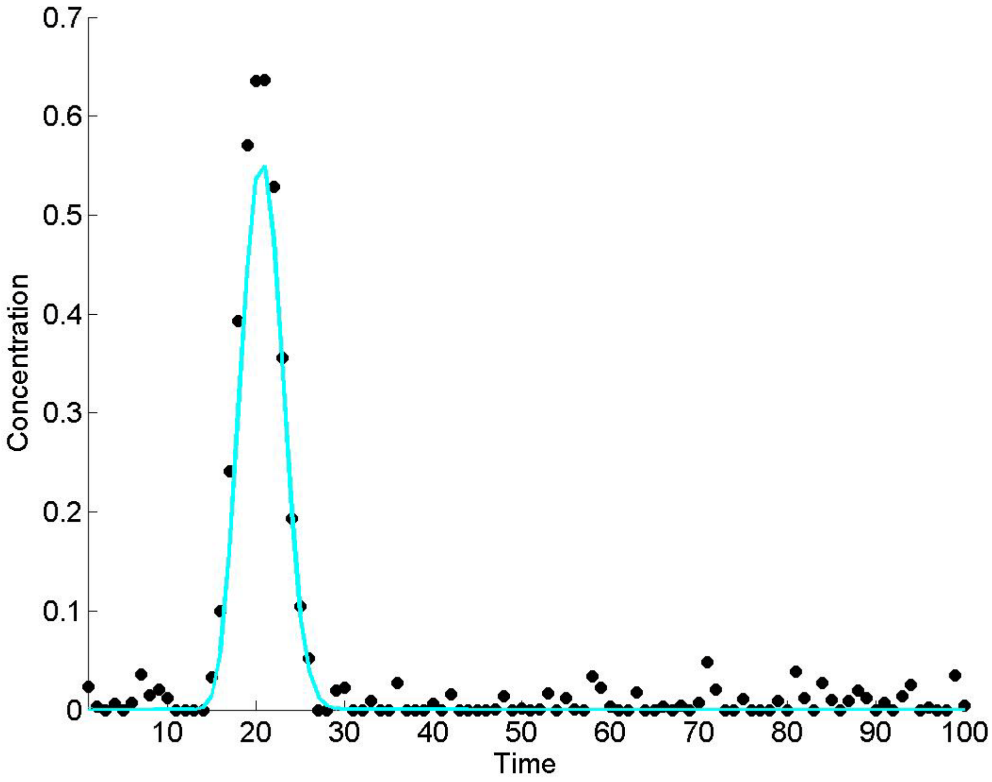
When should treatment start?

How aggressively to treat (design concentration)?

Collect Data at Monitoring Well



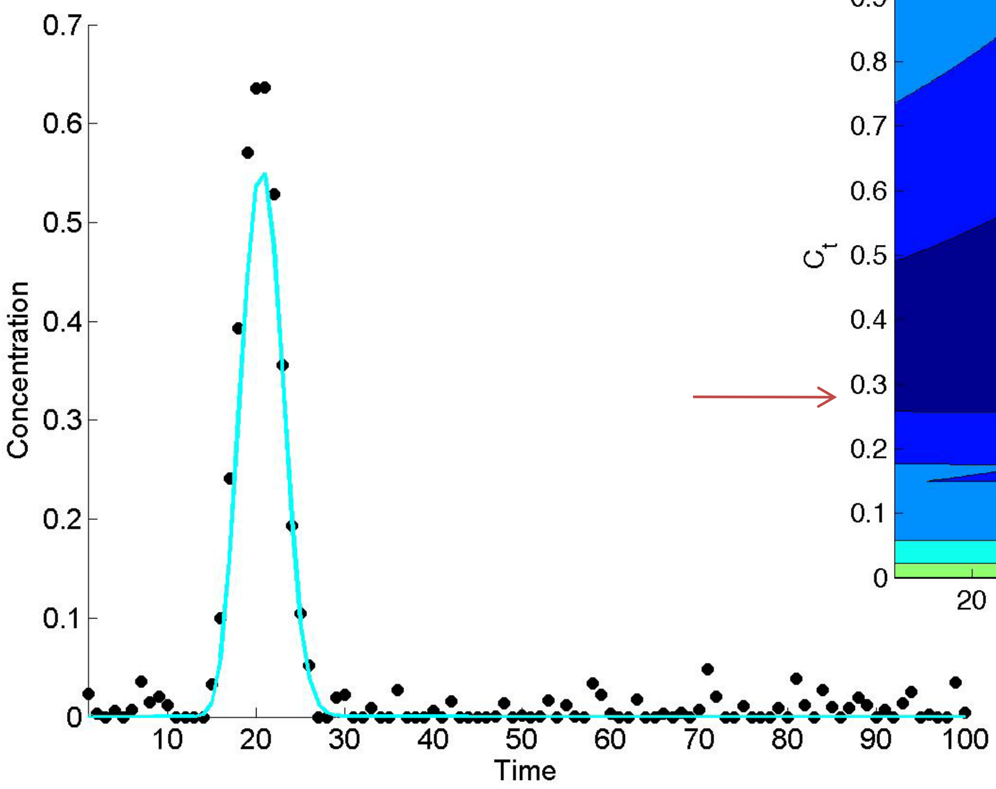
Deterministic Approach:
Select a Best-Fit Model

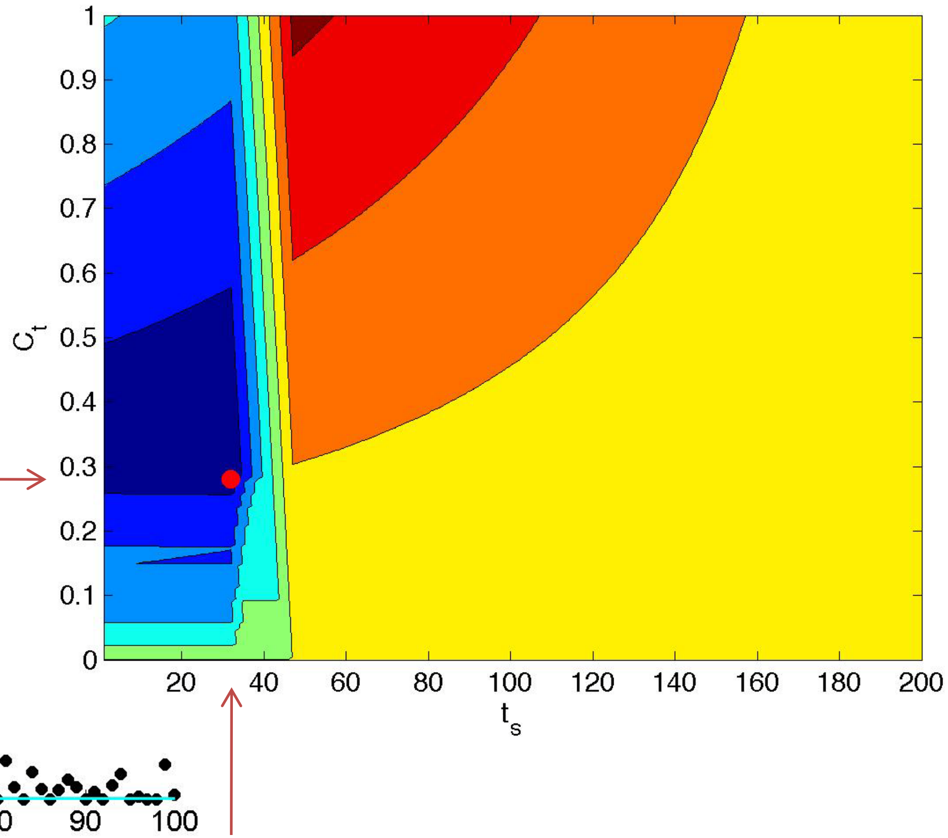


Select best-fit model based on goodness of fit to the data.

Deterministic Approach:

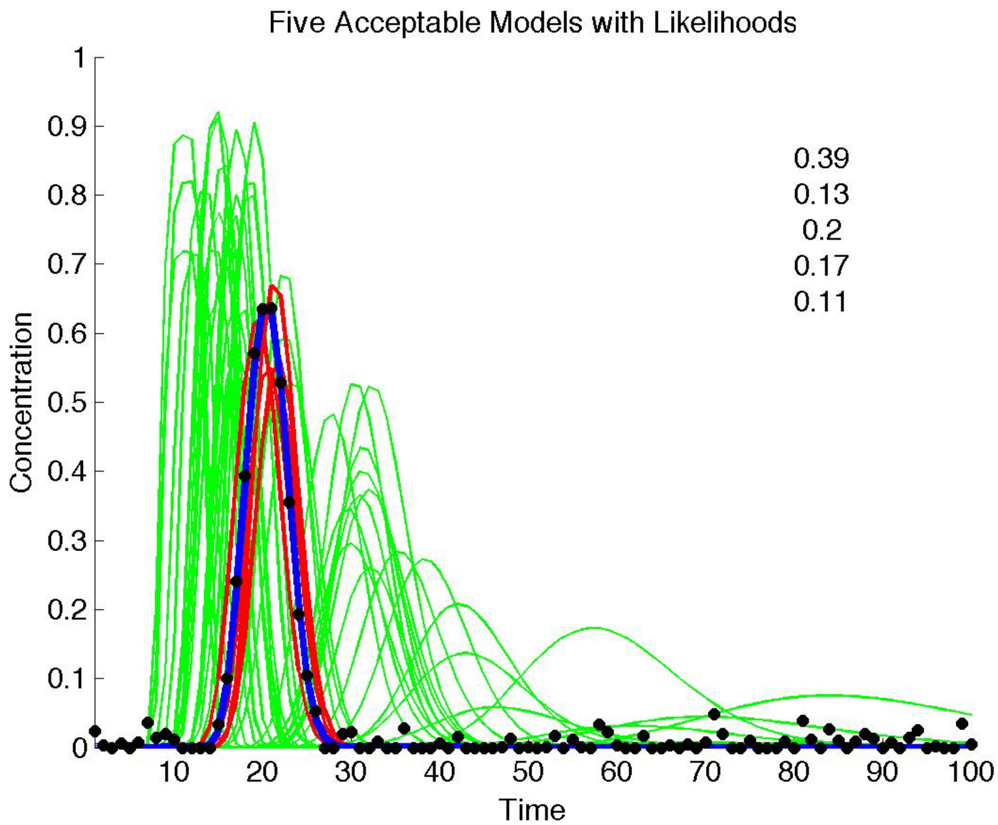
Construct Cost Surface Based on Model





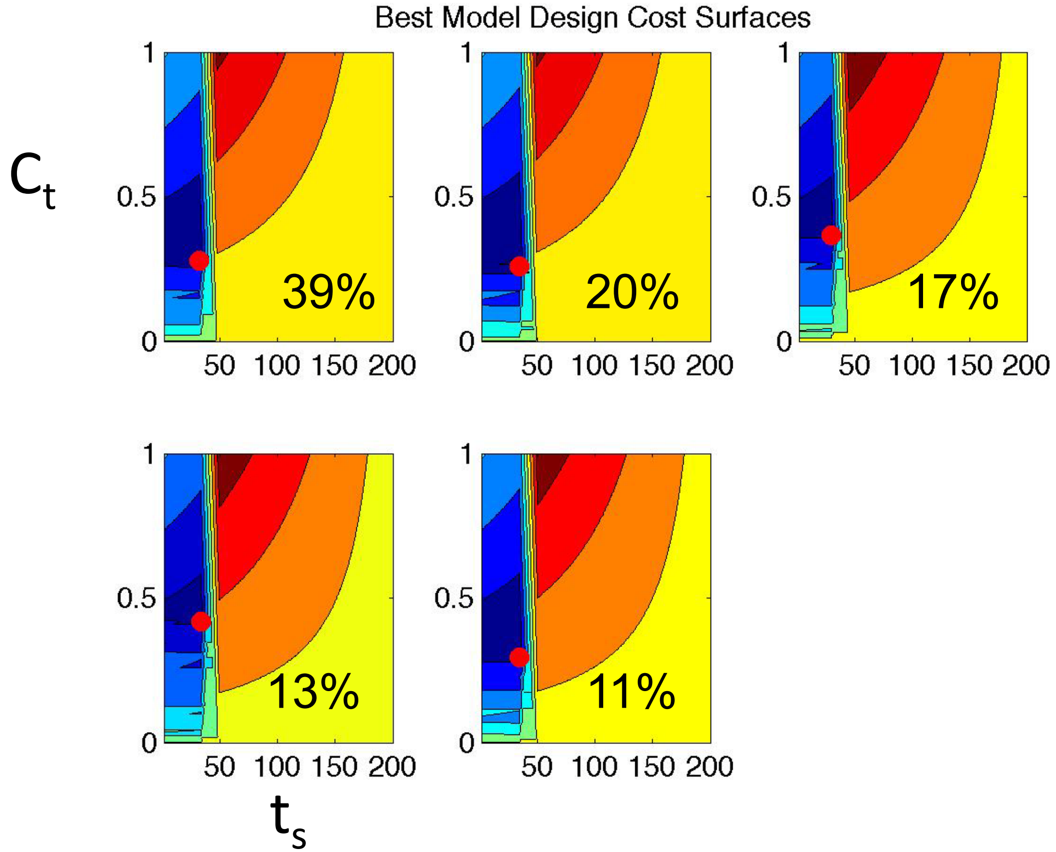
Construct the true design surface and choose the design with the lowest total cost.

Probabilistic Approach:
Select an Ensemble of Best-Fit Models



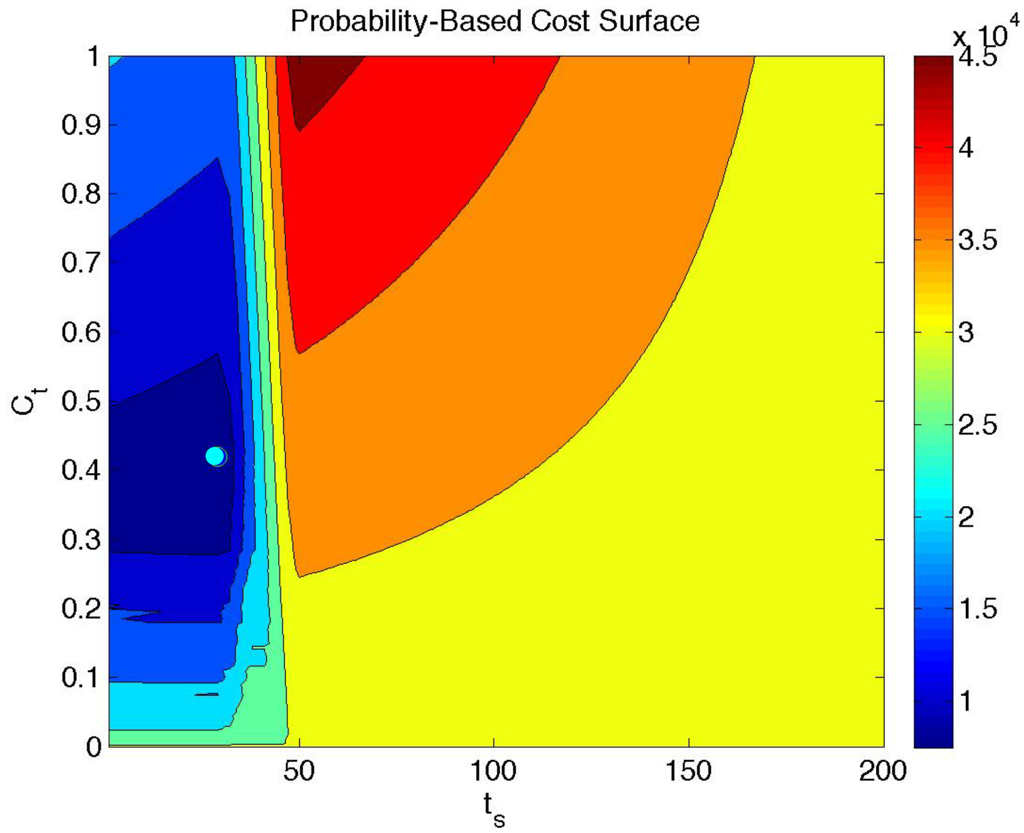
Select several models based on goodness of fit. BMA approaches then produce a likelihood-weighted prediction, which can be used to define a design surface.

Each Total Cost Surface Has a Relative Probability of Being Correct



We take a different approach. We produce a likelihood weighted cost surface based on the individual cost surfaces.

Probability-Weighted Cost Surface

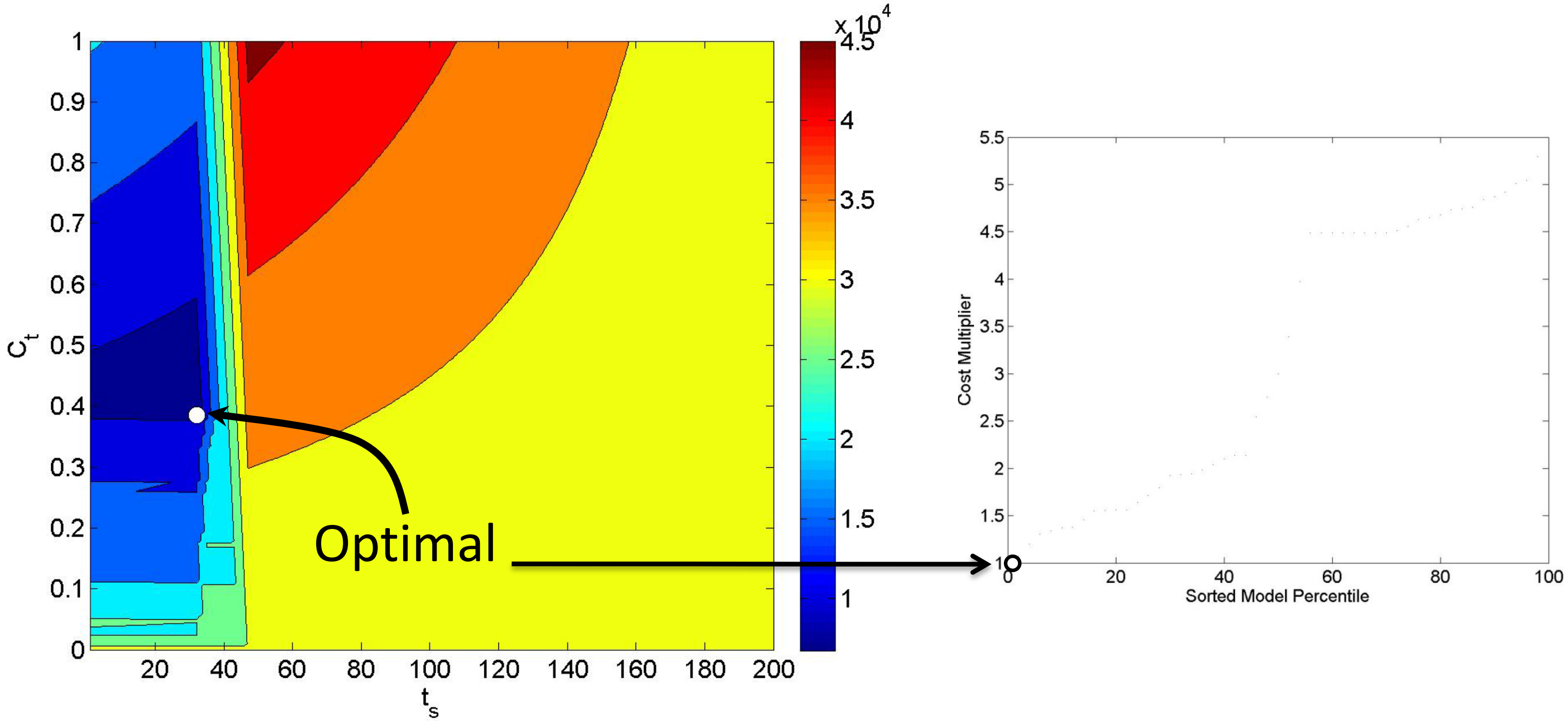


The optimal design on this surface is equivalent to considering the expected cost of each model, given the likelihoods that every other model is true.

Post-Audit

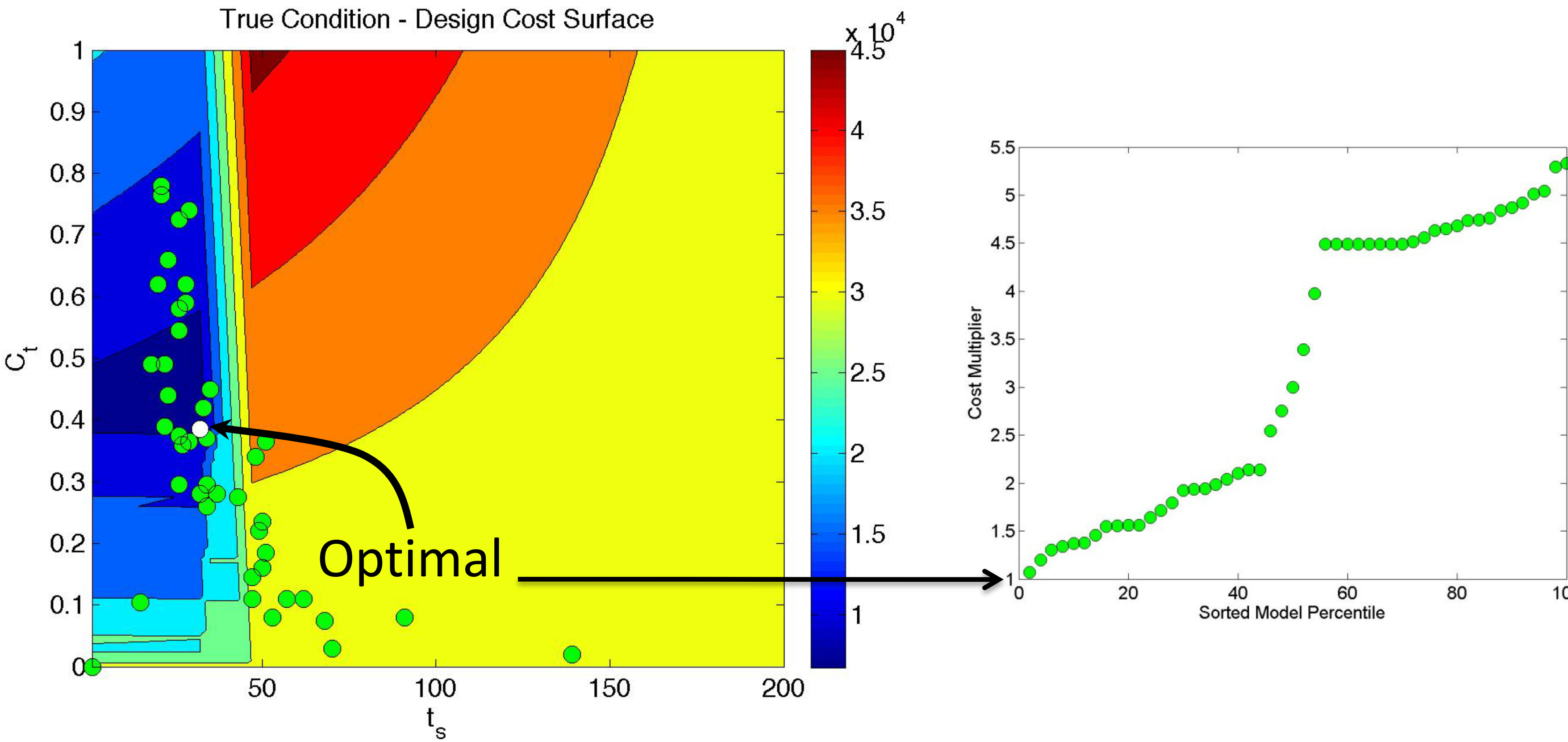
Now we know the true condition, leading to the true total cost curve.

True Design Cost Surface



There is a true optimal design (C_t, t_s). We will compare all other designs costs to this minimum.

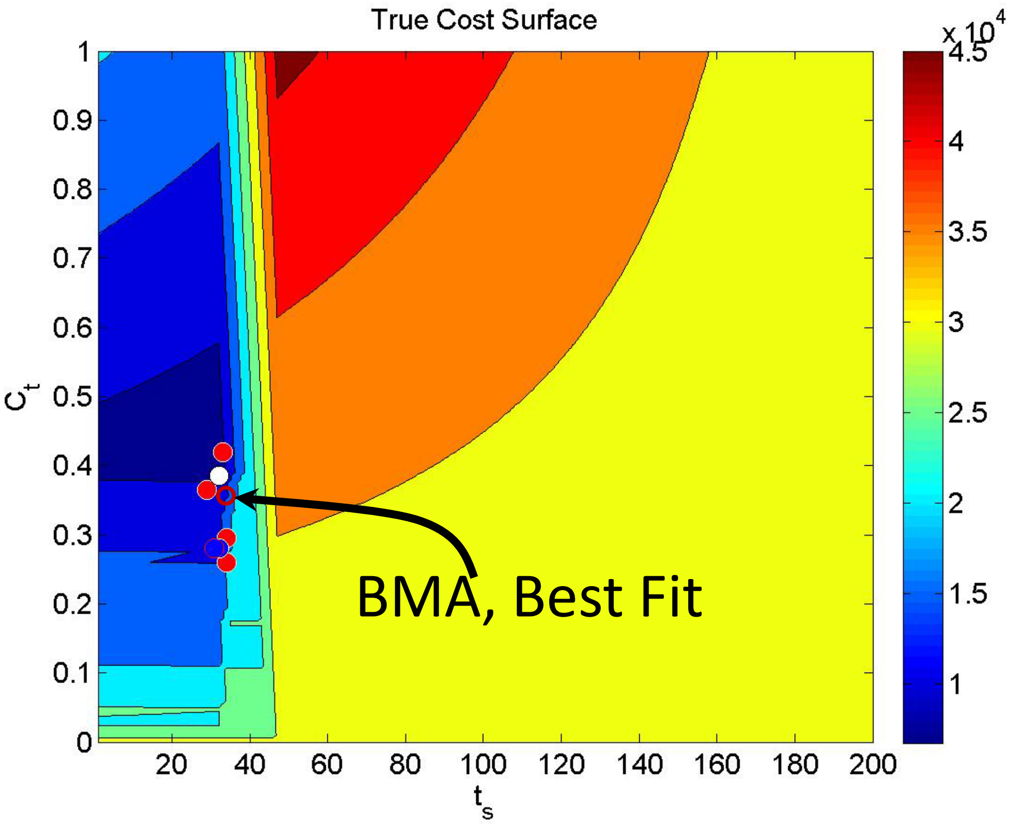
Designs on True Cost Surface

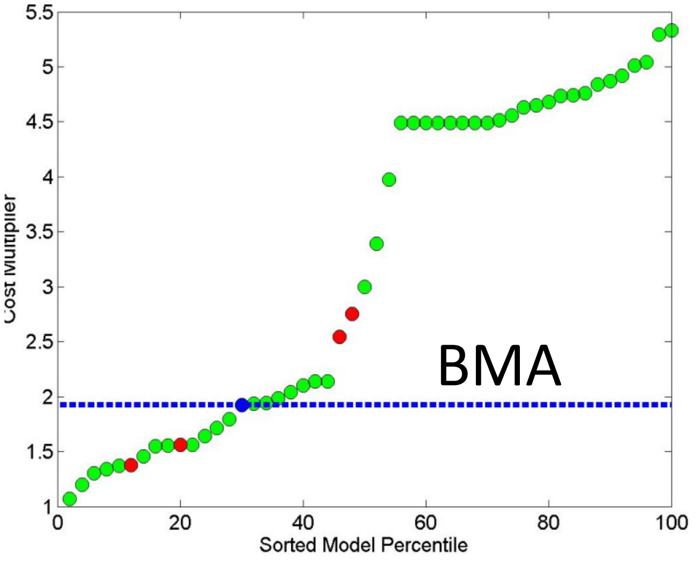


Uncertainty in the model can translate to added cost in the design.

But, the added cost depends on the (unknown) true BTC.

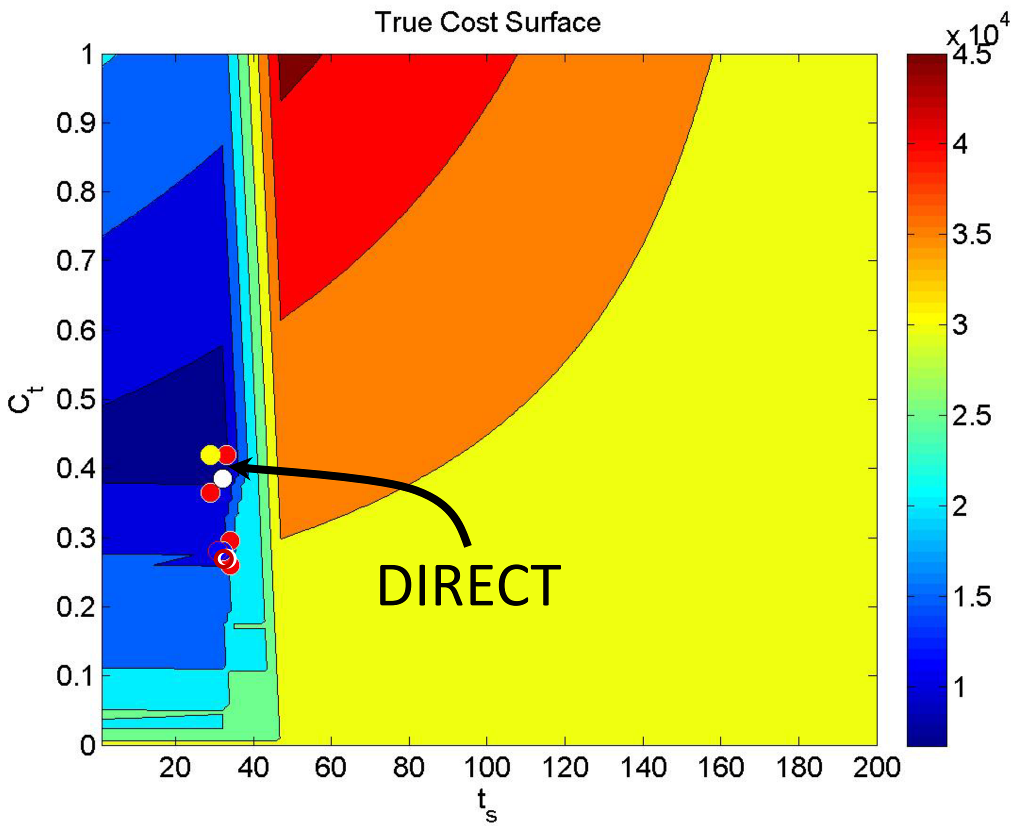
Best Fit & BMA Designs with True Cost

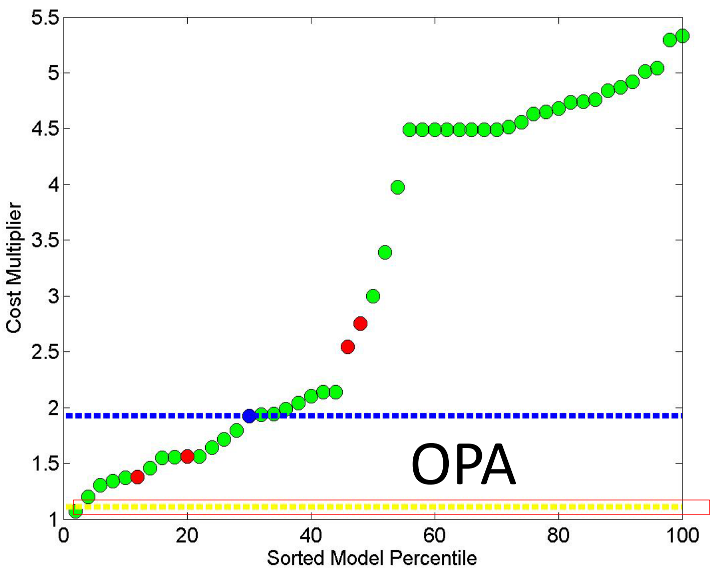




Even the best fit and Bayesian model average approaches would cost more than the optimal.

DIRECT Design with True Cost





OPA approach performed better than the best fit and BMA approaches

Targeting Data Collection for Decision Support

- Identify predictions of concern and related risks/cost (Riley);

Targeting Data Collection for Decision Support

- Identify predictions of concern and related risks/cost (Riley);
- Identify potentially valuable measurement methods and implementations (Rucker, Sheen);

Targeting Data Collection for Decision Support

- Identify predictions of concern and related risks/cost (Riley);
- Identify potentially valuable measurement methods and implementations (Rucker, Sheen);
- Maximize information extraction through data analysis (Johnson) and context (Finsterle);

Targeting Data Collection for Decision Support

- Identify predictions of concern and related risks/cost (Riley);
- Identify potentially valuable measurement methods and implementations (Rucker, Sheen);
- Maximize information extraction through data analysis (Johnson) and context (Finsterle);
- Conduct prescreening to eliminate non-informative measurements (Slater); and

Targeting Data Collection for Decision Support

- Identify predictions of concern and related risks/cost (Riley);
- Identify potentially valuable measurement methods and implementations (Rucker, Sheen);
- Maximize information extraction through data analysis (Johnson) and context (Finsterle);
- Conduct prescreening to eliminate non-informative measurements (Slater); and
- Explicitly consider uncertainty when selecting measurements and making decisions.