Alternative Optimization Techniques

Kirk Cameron, Ph.D.
MacStat Consulting, Ltd.
kcmacstat@qwest.net

LTMO

Ways to Skin a Cat

- Optimization techniques for LTM have advanced over last several years
  - Improved mathematics
  - Increased complexity
  - Combination of statistical & physical data
- Highlight some alternate methods
  - Genetic algorithms
  - Kalman-type filtering, updating

A Question of Tradeoffs

- Computational horsepower
  - More analysis time, computer power often required
  - More complex models, greater reliance on advanced numerical methods
- More complex data needs
  - “Garbage in, garbage out” principle
  - More kinds and quantities of data typically needed
  - Needed information not always available

LTMO

It’s an Empirical Thing

- No optimization possible without data
  - Moving from simpler to more complex techniques exposes a dilemma
  - Advanced approaches may look ‘slick,’ but without extra data, results may be no better
- Determine which data inputs are crucial
  - Forward thinking: how can this data be collected in the future?

Multiobjective Groundwater Monitoring Design

- Patrick Reed and Venkat Devireddy
  - Department of Civil and Environmental Engineering
  - The Pennsylvania State University
- Contact info: preed@engr.psu.edu

LTMO

Genetic Algorithm Basics

- Genetic algorithms search among all possible monitoring plans
  - Create an initial "population" of sampling plans
  - Sampling plans are represented as a string binary digits
    - 1 – i^{th} well sampled
    - 0 – i^{th} well not sampled

0	1	0	0

  - Judge each sampling plan based on how well it satisfies monitoring objectives
    - Plans that satisfy design objectives are "highly fit"
    - Have higher likelihood of mating and passing traits
  - Sampling plans undergo Darwinian "natural selection" until the best set of plans are evolved

LTMO

How do GAs work?

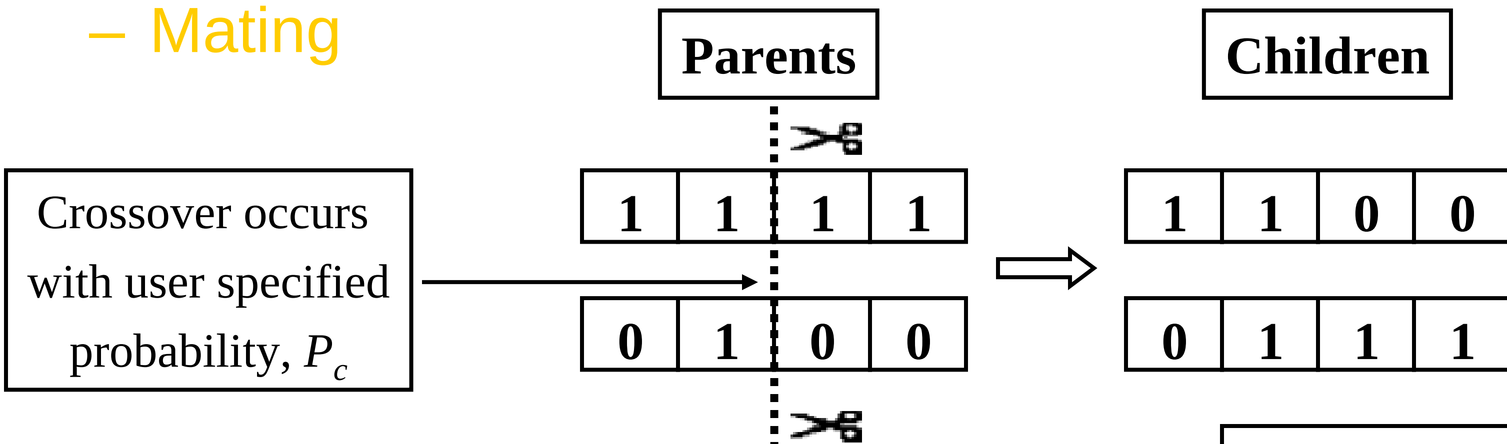
– Selection

– Measure of the fitness of string

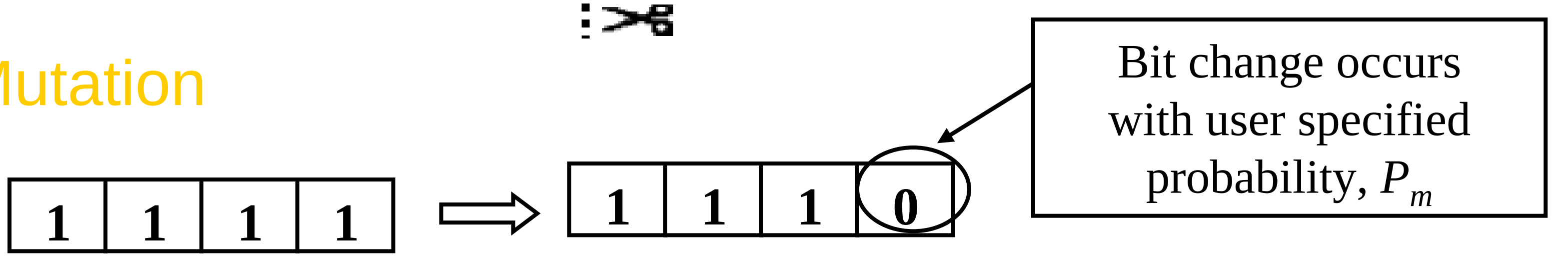
– Fitness is rated in terms of the objective functions

– Population members compete to mate & pass traits

– Mating

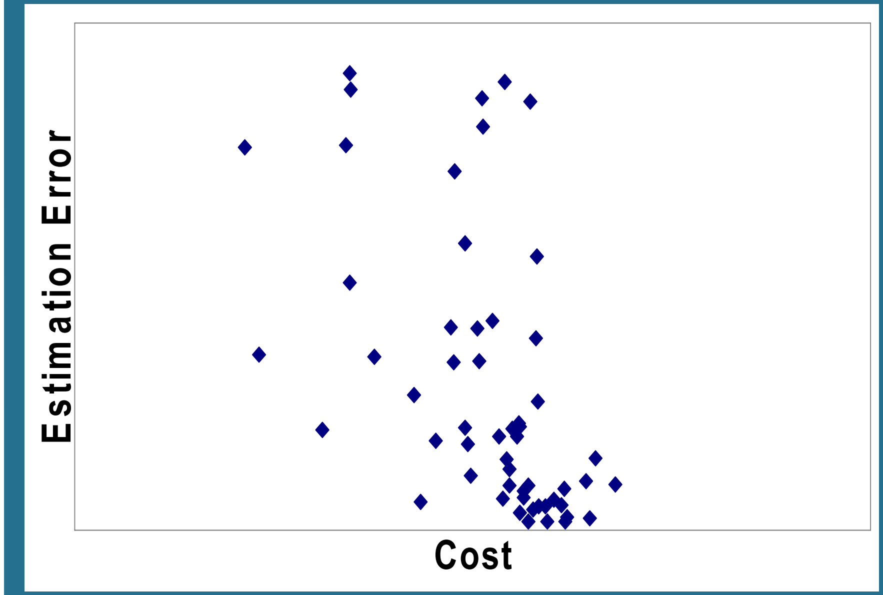


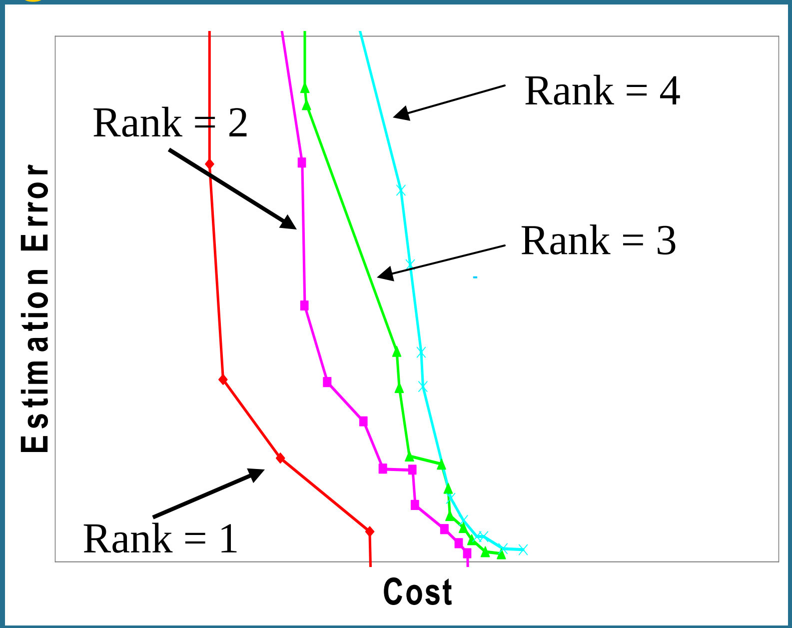
– Mutation



Multiobjective Evolution

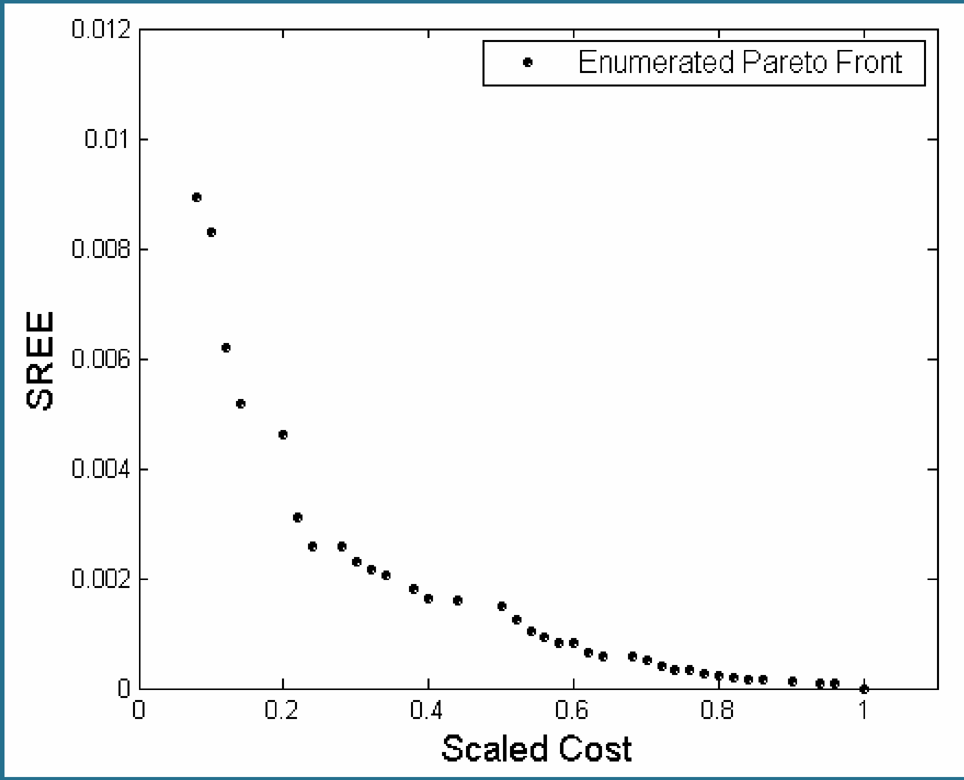
- The Nondominated Sorted Genetic Algorithm-II:
  - Population classified into fronts using non-dominated sorting
  - Better ranking = Higher likelihood of passing traits
  - Entire tradeoffs evolved in single run





Case Study: Enumerated Tradeoff

- Two-Objectives:
  - Minimize Cost
  - Minimize Relative Mapping Error (termed SREE)
- 2^{20} possible designs
- Evaluated using a nonlinear least squares interpolation
- Best published result
  - Reed et al. (2004) WRR
  - 38000 evaluations



LTMO

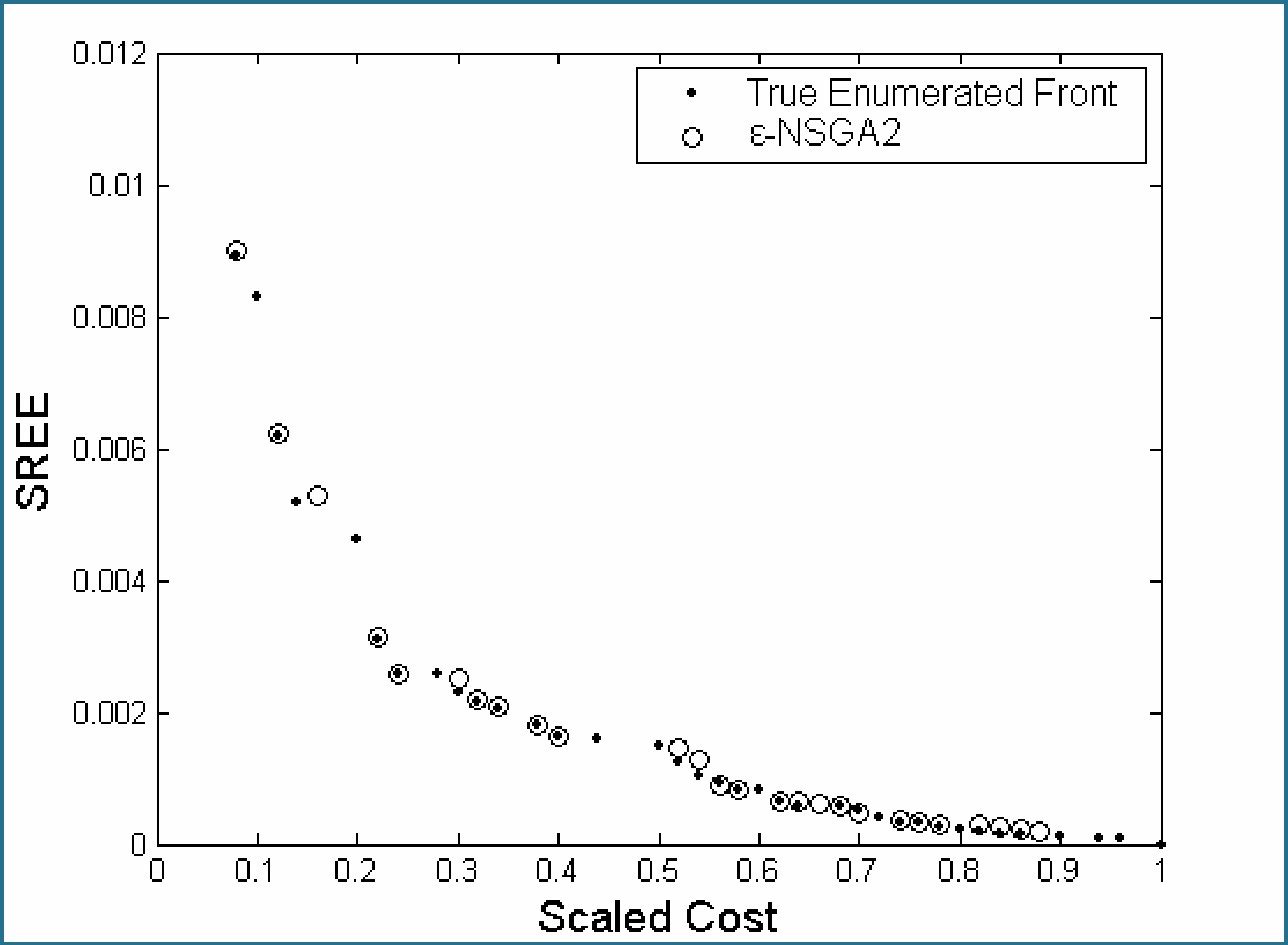
Case Study: Efficiency of New Tools

- Results for 50 trials (using interactive archiving)
  - Reduced solution time from hours to seconds
  - New tool is up to 90% more efficient

	Reed et al. 2003	ε-NSGA2 (New Tool)
Min. no. of design evaluations	16800	1400
Avg. no. of design evaluations	39889	7992
Max. no. of design evaluations	74400	25400

LTMO

Case Study: Typical Result



Adaptive Environmental Monitoring System (AEMS)

- Under development at RiverGlass Inc., Champaign, IL
  - Software development company launched by the University of Illinois
  - Project lead: Barbara Minsker, PhD minskerconsulting@insightbb.com
- Beta testing of AEMS expected to begin in late Summer 2005
  - Completed functionality available now on a project basis – contact Dr. Minsker for more information

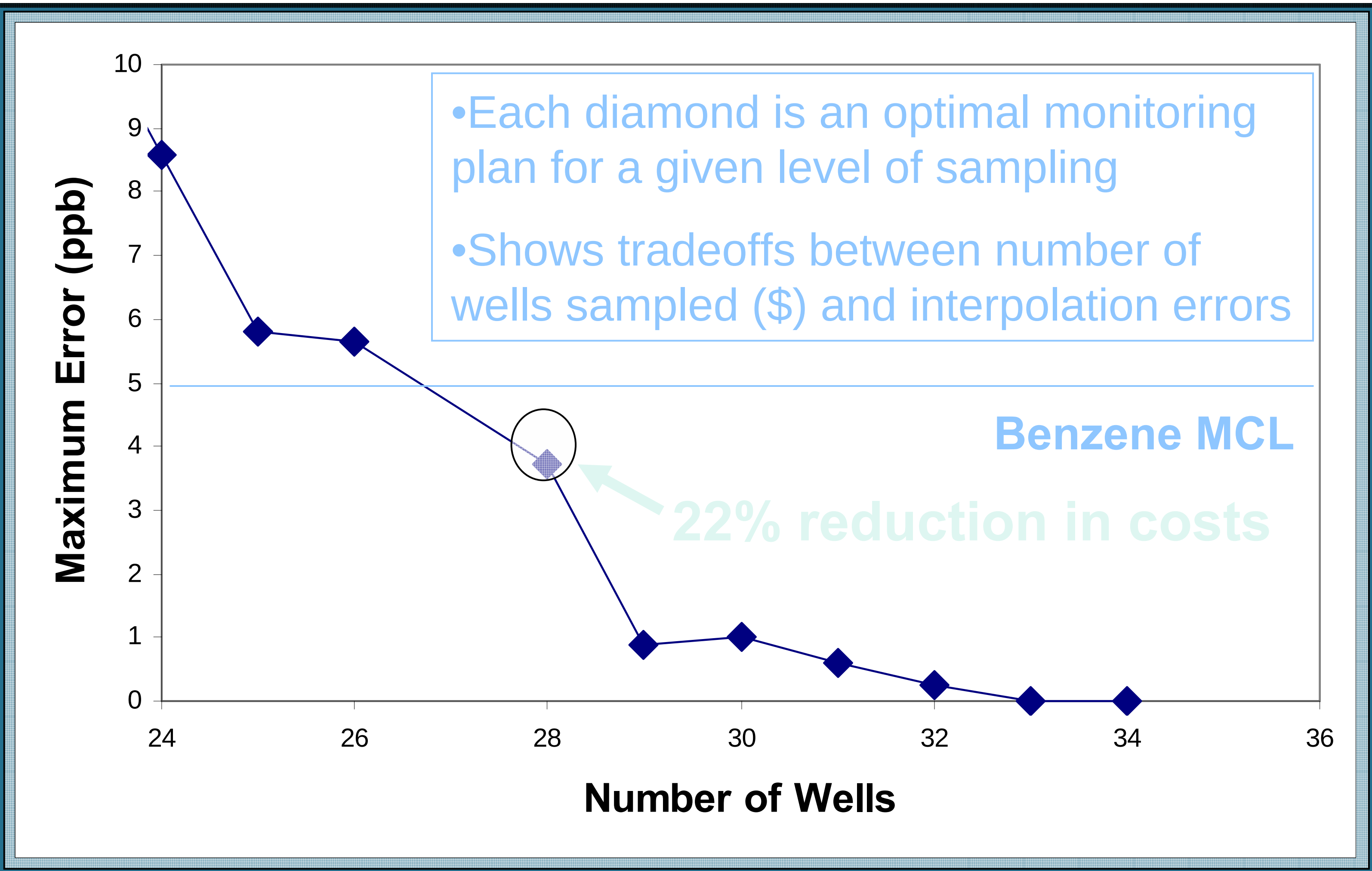
AEMS Capabilities

- Build data-driven trend models based on historical data (geostatistical or analytical models)
- Assess new data in real time to
  - Identify significant deviations from previous trends (spatial or temporal), providing automated alerts
  - Identify locations/times where additional data would be most beneficial to reducing risks
- Identify temporal and spatial redundancies in existing monitoring regimes

Redundancy Analyses with AEMS

- Method considers
  - Tradeoffs among multiple objectives (e.g., cost and error) with mathematical optimization
  - Uncertainty in identifying robust designs
  - Temporal and spatial redundancy simultaneously
- Applied at 2 BP sites and 1 DOE site
  - 22-36% redundancy found in well sampling plans

AEMS Identifies Monitoring Tradeoffs



LTMO

Optimization of Large Scale Subsurface Environmental Impacts

- Larry M. Deschaine, PE
  - Engineering Physicist
  - SAIC & Chalmers University of Technology
- Contact info: larry.m.deschaine@alum.mit.edu

LTMO

Information Content Fused Approach

- Integrated algorithm(s) consist of:
  - Simulation models based on physics
  - Data models based on sampling
  - Uncertainty handled through (geo)-statistics
- Information content fusion (Data & Physics):
  - Signal processing (i.e. Kalman Filters, etc.)
  - Genetic Programming
- Optimal System Estimate
  - Optimal estimate of “system” for locating plume at given time, or time-space correlated estimates of long term monitoring programs

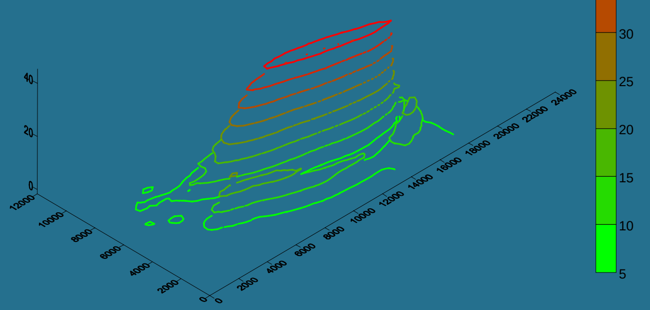
Fundamentally Different From Estimation

- Current methods typically gather data, calibrate model and use model for predictions
  - Models break down as physics becomes complex, data sparse or input parameters not well known
- This method fuses the information content via signal processing / machine learning algorithms:
  - Integrated data/physics model provide optimal estimates based on knowledge gained from both the physical simulator *and* the data, updated as new information is obtained

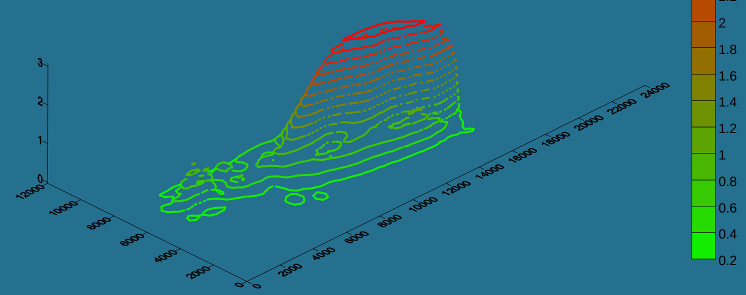
LTMO

Determining optimal well locations & their benefit to understanding plume location – when to stop adding wells, how to justify

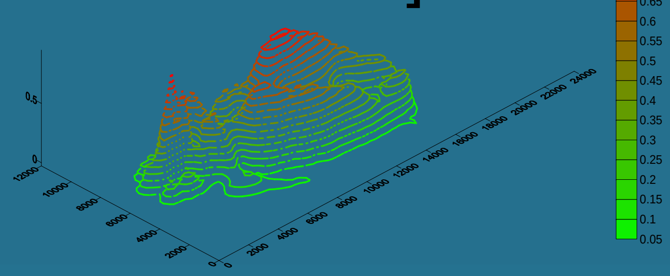
No wells
[Max uncertainty]



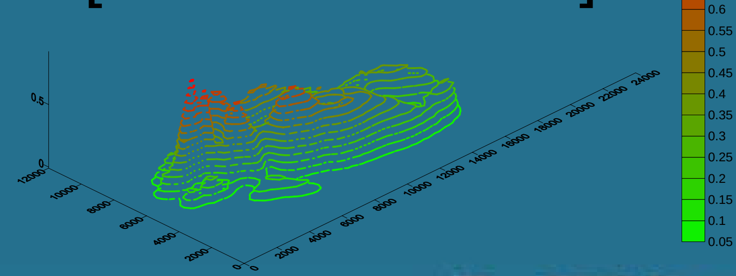
Existing MW system
[93.4% reduction]



Add One Well
[98% reduction]

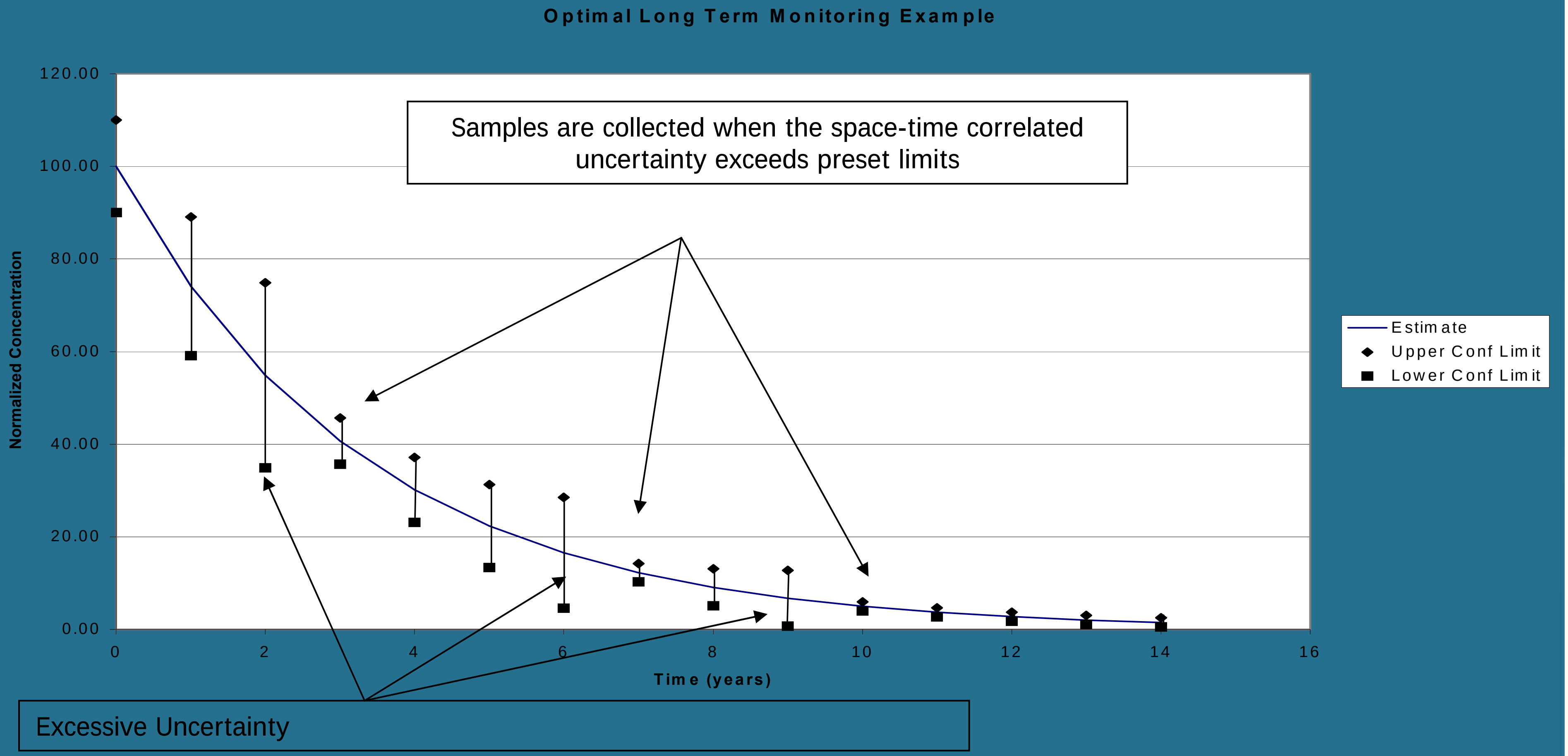


Add 2nd Well
[98.5% reduction]



LTMO

LTM: Extend the reduction of uncertainty to include value of historic samples in space and time -> Results in less samples needed to understand long term system behavior



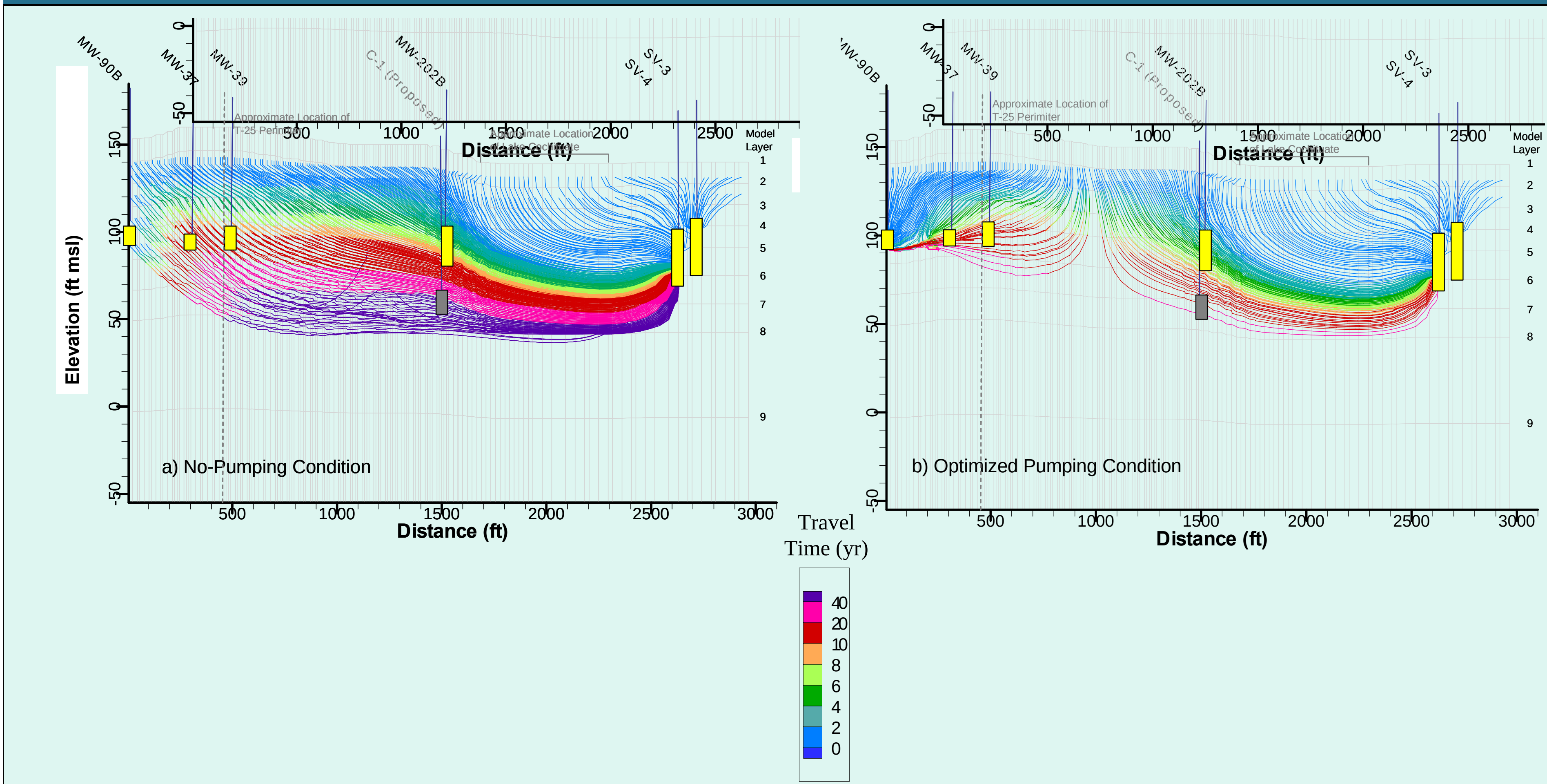
OPDATE

- David Dougherty, PE
  – Subterranean Research, Inc.
- Contact info: ddougher@subterra.com

Fusion of Data, Simulation

- Site-specific, consulting approach
  – Strong emphasis on decision-trees
- Simulation models first used to both:
  – Help plan remedy phases during RI, FS and design LTM program
    - One goal of LTM design should be to improve simulation model for anticipated future remedy/risk assessments
  – Make predictions about future subsurface conditions

Use of Model in LTM Planning: Placement of MWs Accounts for Different Phases of Remedy



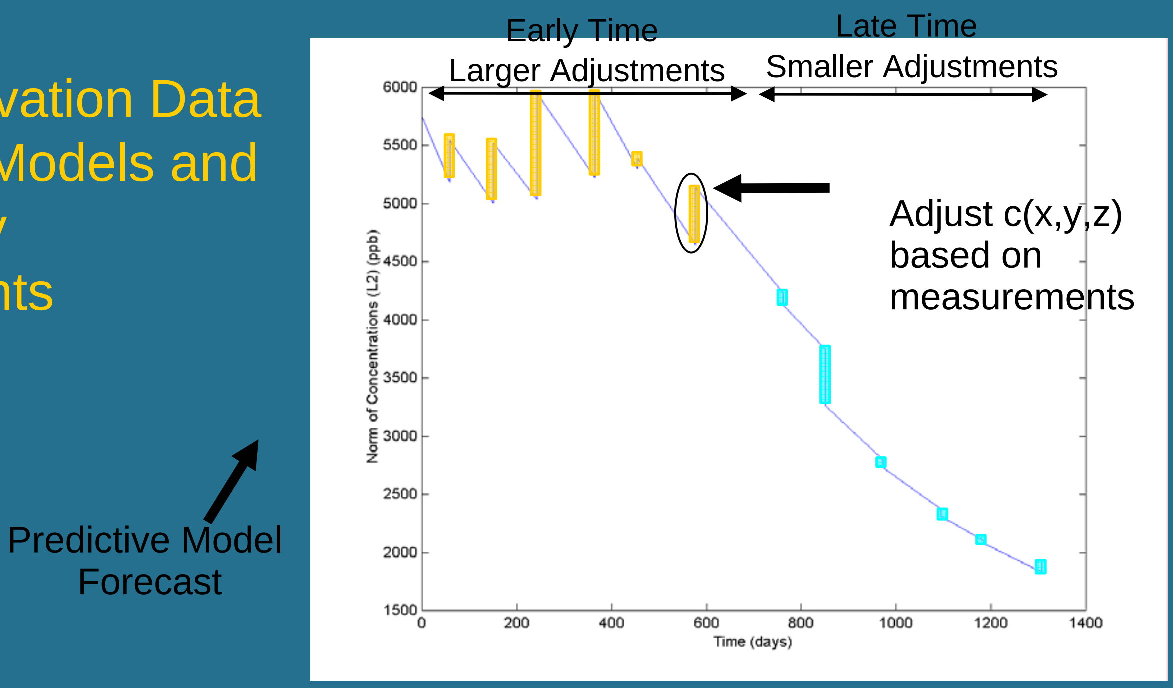
Retrospective Analysis

- Retrospective hydrologic data assimilation
  - Applied to observed data to:
    - 1) add branches to LTM decision trees, and
    - 2) improve quality of forecasts
- Measured data used to update and optimize (“opdate”) simulation models

LTMO

Using Data Assimilation

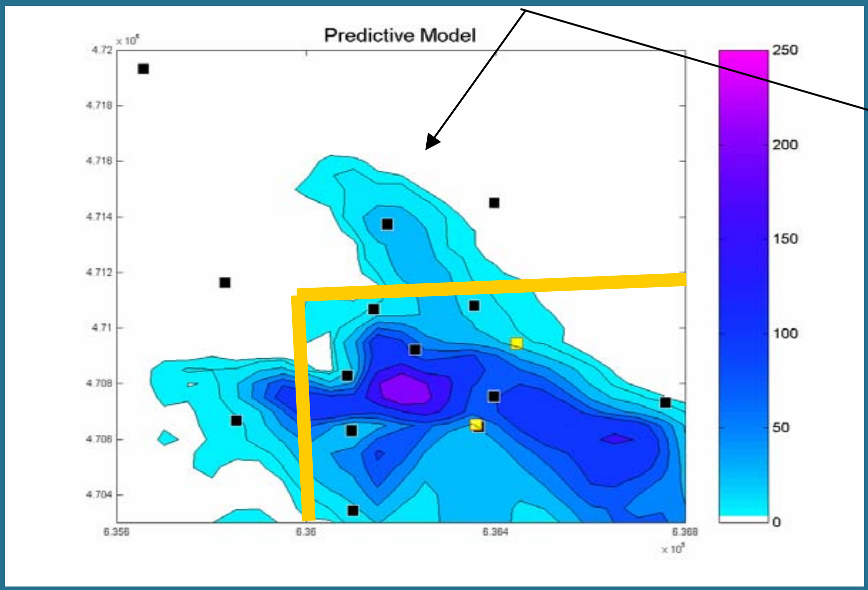
Use Observation Data to Update Models and Uncertainty Assessments

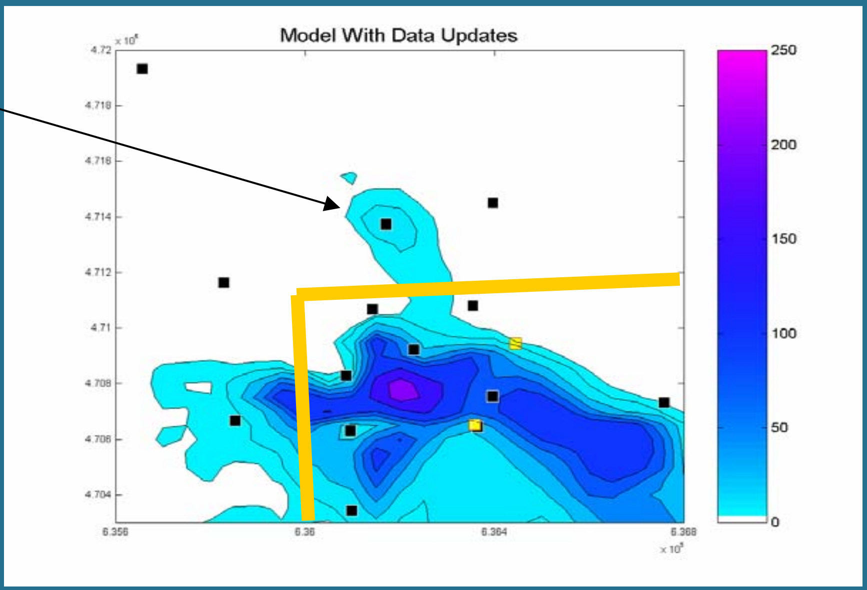


LTMO

“Opdate”-ing

“Snapshot in time” (July, 1999), prior to and after updating the predicted concentration plumes with monitoring data.





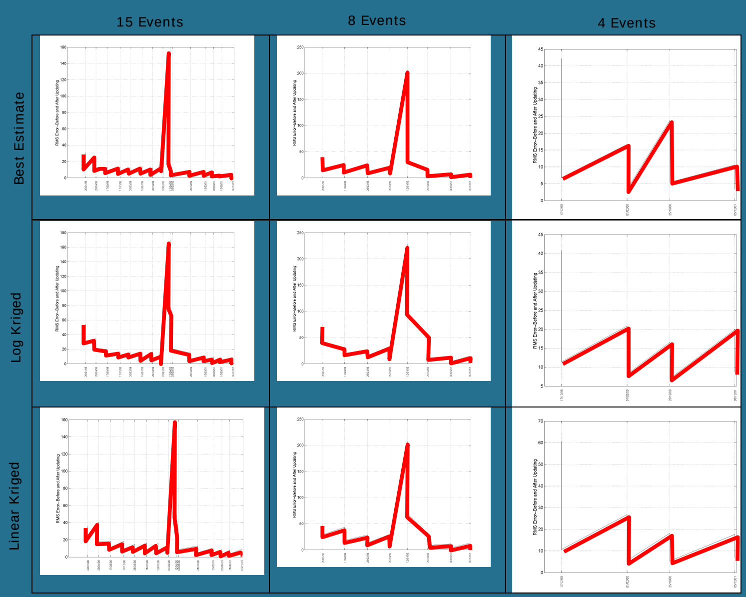
LTMO

Retrospective Data Assimilation—OPDATE™

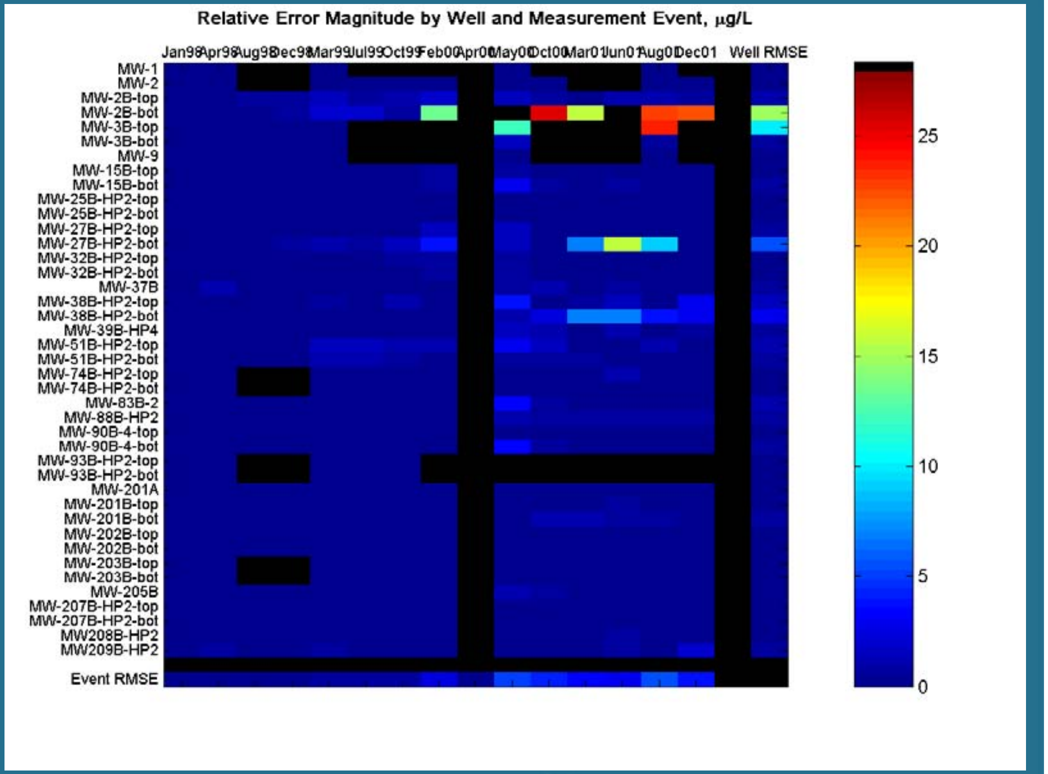
Sensitivity to
Historical Sampling Frequency

When and Where Do
Errors Occur?

Scenarios



Quarterly Semi-Annual Annual
Updating Frequency



LTMO

Summary

- ‘Next Generation’ methods offer great promise… however
  - Many sites currently lack sufficiently detailed data or right kind of data
  - More expertise required to correctly set-up, implement
- Power, adaptability of genetic algorithms, Kalman filter techniques cannot be denied
  - Especially true as computer horsepower gets faster & cheaper

LTMO

Summary (cont.)

- Each method being developed for real-world applications
  - Savannah River, BP, DOE, AF facilities
  - Some “scalability” possible, especially with genetic algorithm approaches
- Key benefits
  - Formal incorporation of additional kinds of data
  - Enumeration of entire cost tradeoff curve, sort of