Geostatistical LTM Optimization Using GTS

Kirk Cameron, Ph.D.
MacStat Consulting, Ltd.
Colorado Springs, CO

What Kind of Optimization?

- Highly statistical, data-driven
  - Not linked to fate/transport models
  - What do observed sampling measurements tell us?
- Minimize waste; maximize usefulness of data collected
  - Look for redundant sampling information
- Geostatistical Temporal/Spatial (GTS) algorithm looks at two areas:
  - Monitoring network locations
  - Sampling frequencies in network

Other Applications

- Optimization of treatment systems
  - Example: sampling frequencies of influent/effluent for pump & treat operations
- Characterization & mapping of sites
  - Change in contamination patterns over time
  - Hydrogeologic parameters needed for flow-based geophysical or fate/transport models
  - Determining optimal locations for new sampling or drilling

GTS Algorithm

- Designed with decision-logic framework
  - Version 2.0 available by end of July
- Separate identification of temporal & spatial redundancy
- Uses geostatistical and trend optimization methods
  - Variogram = spatial correlation measure
  - Kriging = spatial interpolation = spatial regression
  - Locally-Weighted Quadratic Regression (LWQR)

Note on Redundancy

- Practical definition: What happens when data removed from current system?
- Temporal
  - Can trends be re-constructed?
  - Do consecutive sampling events become uncorrelated?
- Spatial
  - Can surface map be re-constructed?
    - Plume extent and intensity

Optimality vs. Redundancy

- Redundancy a misnomer
  - All unique data points valuable
  - Always have loss of information if removed
- Must balance tradeoff between cost savings and loss of accuracy
  - Optimal system = minor information loss but large gain in resource savings

Optimality (cont.)

- Common strategy
  - Use existing data to estimate baseline
  - Remove some data (wells, sampling events)
  - Re-estimate baseline with reduced data set
  - Measure relative error incurred
  - Examine cost-accuracy tradeoff

Key Steps to Optimality

- Exploration/Screening
- Temporal Optimization
  - Too many sampling events at individual wells?
- Spatial Optimization
  - Too many wells in network?

Exploration & Screening

- COC analysis
  - Statistical summaries of detections, hits above regulatory limits
  - Spatial spread & intensity
    - Post-plots of wells by detection, regulatory status
- Horizon analysis
  - 2-D or 3-D?
  - Compute spatial spread & intensity for each horizon
  - Overlay horizon-specific correlograms

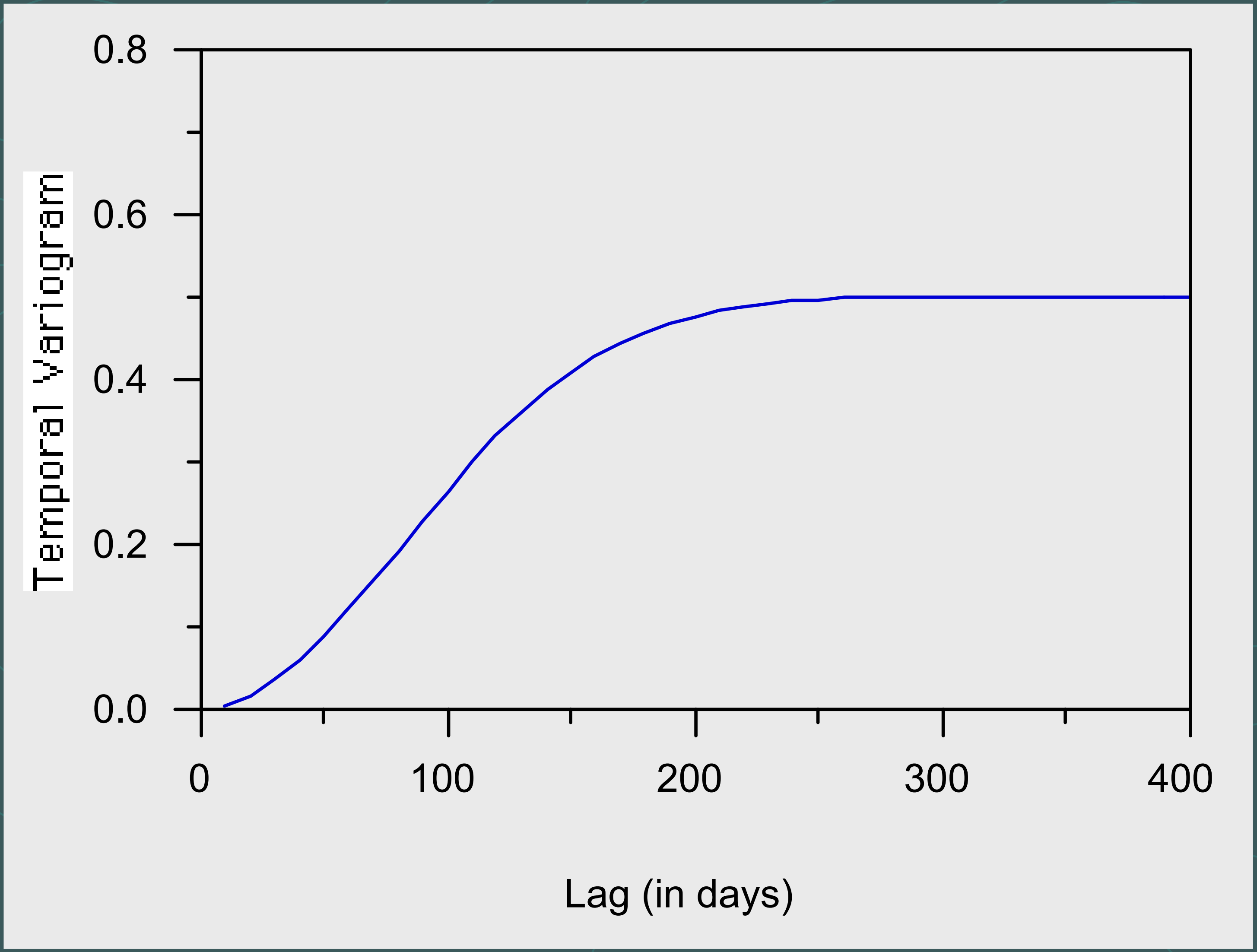
Temporal Optimization

- Two approaches
  - Temporal variogram to estimate average correlation between sampling events
  - Iterative “thinning” of individual wells to adjust well-specific sampling frequencies

Temporal Variogram

- Measures correlation between sampling events over time
- Variogram maps inverted correlation against sampling interval (Δt)
  - High correlation leads to low variogram, & vice-versa

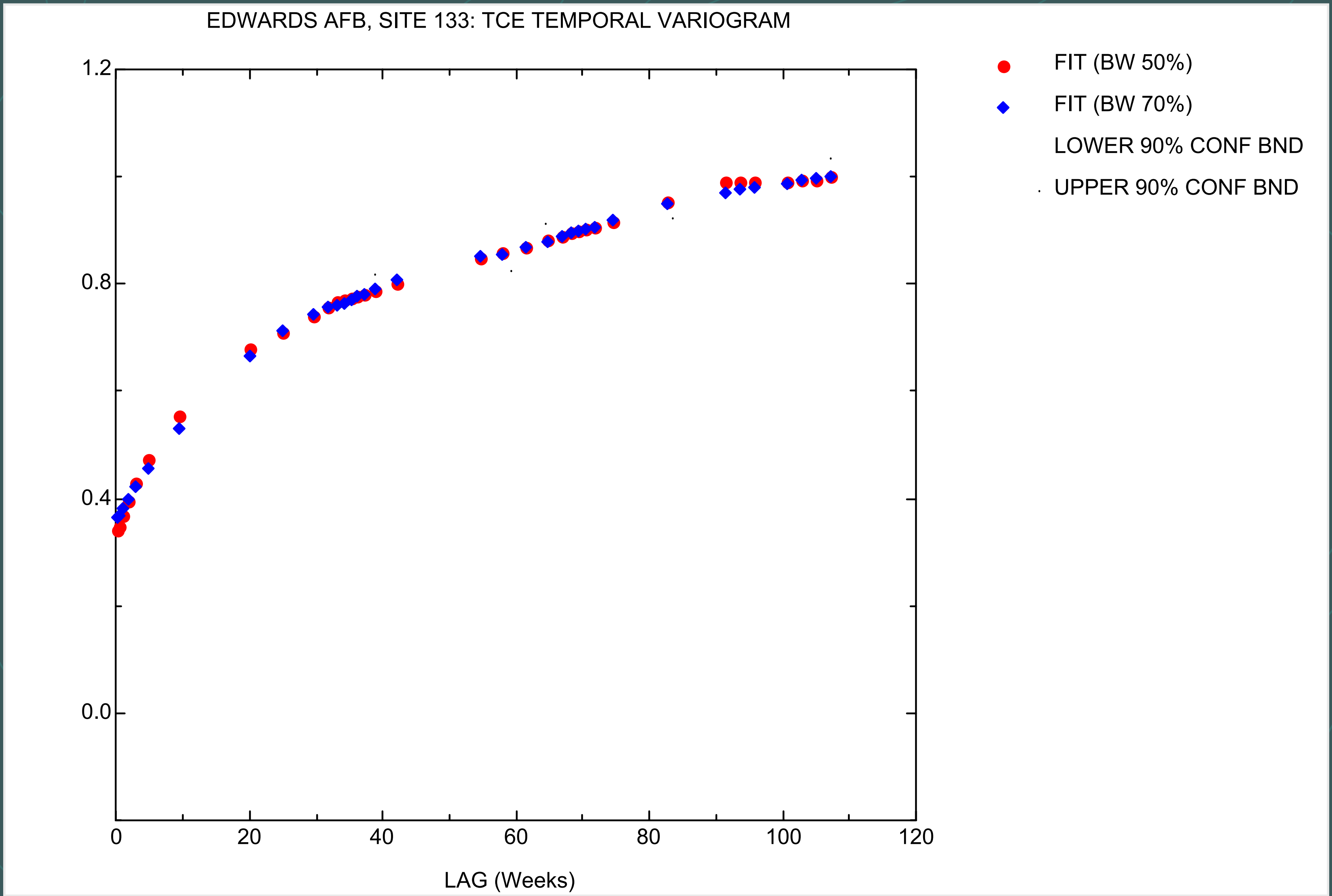
Ideal Variogram



Temporal Variogram (cont.)

- Advantages of temporal variogram
  - Useful with irregularly spaced data
  - Data from multiple wells can be included
  - Single graph shows optimal global sampling interval
- Just need to determine sill and where it begins (“range”)
  - Range can be taken as minimum global sampling interval

TCE Temporal Variogram



Temporal Variogram Requirements

- Sampling events from ≥3-4 yr period for some wells; others can have few as two events
- All wells contribute to overall variogram
  - Except if dominated by NDs
  - Data re-scaled to put all wells on equal statistical footing
- LWQR used to estimate overall pattern & range

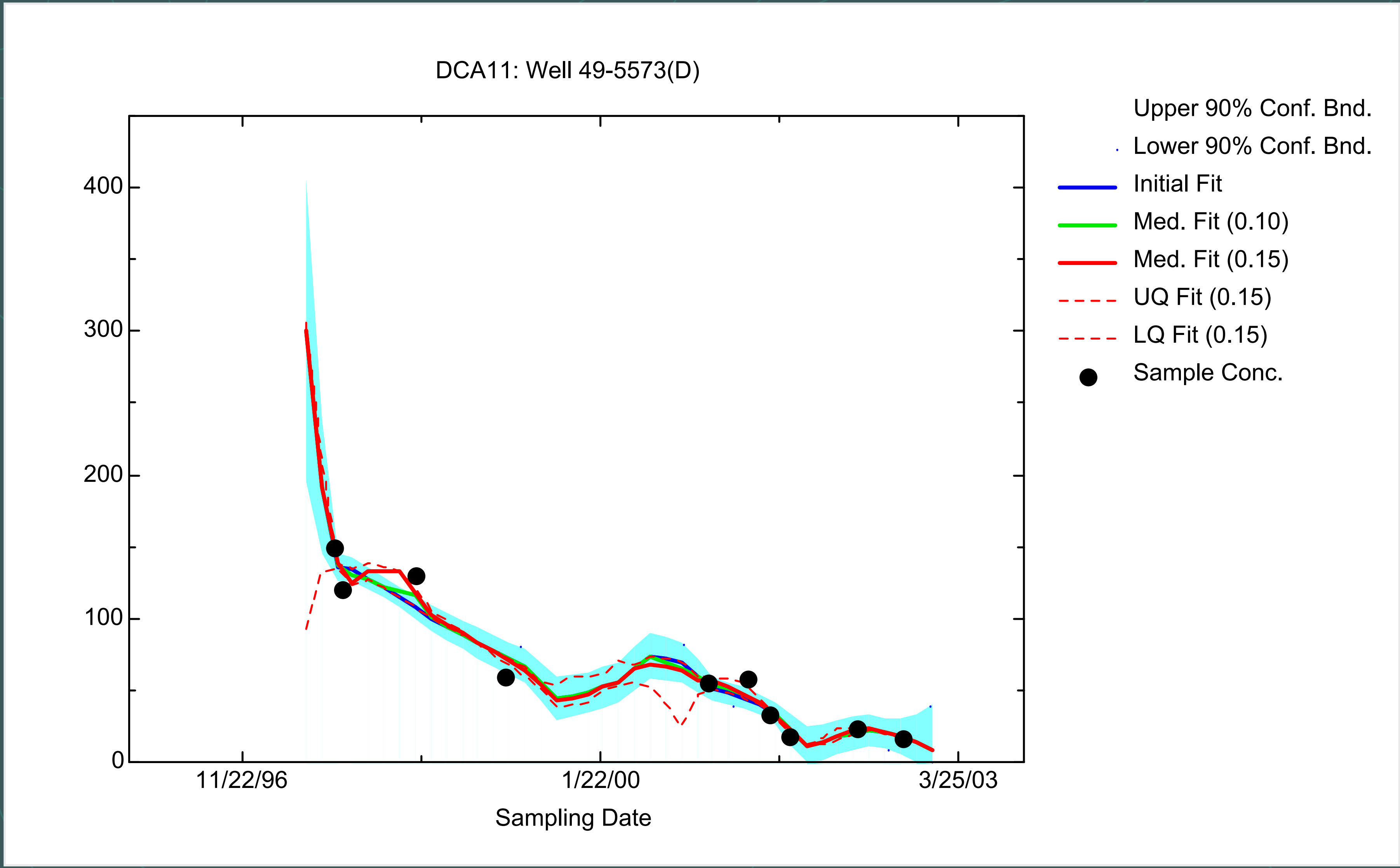
Iterative Thinning

- Adjust individual well sampling frequencies
  - Global sill might not be evident
  - Key wells in network might behave differently
  - Operational target frequency = median of individual wells
- Iterative thinning approach: overview
  - Estimate baseline trend
  - Randomly “weed out” data points
  - Re-estimate trend

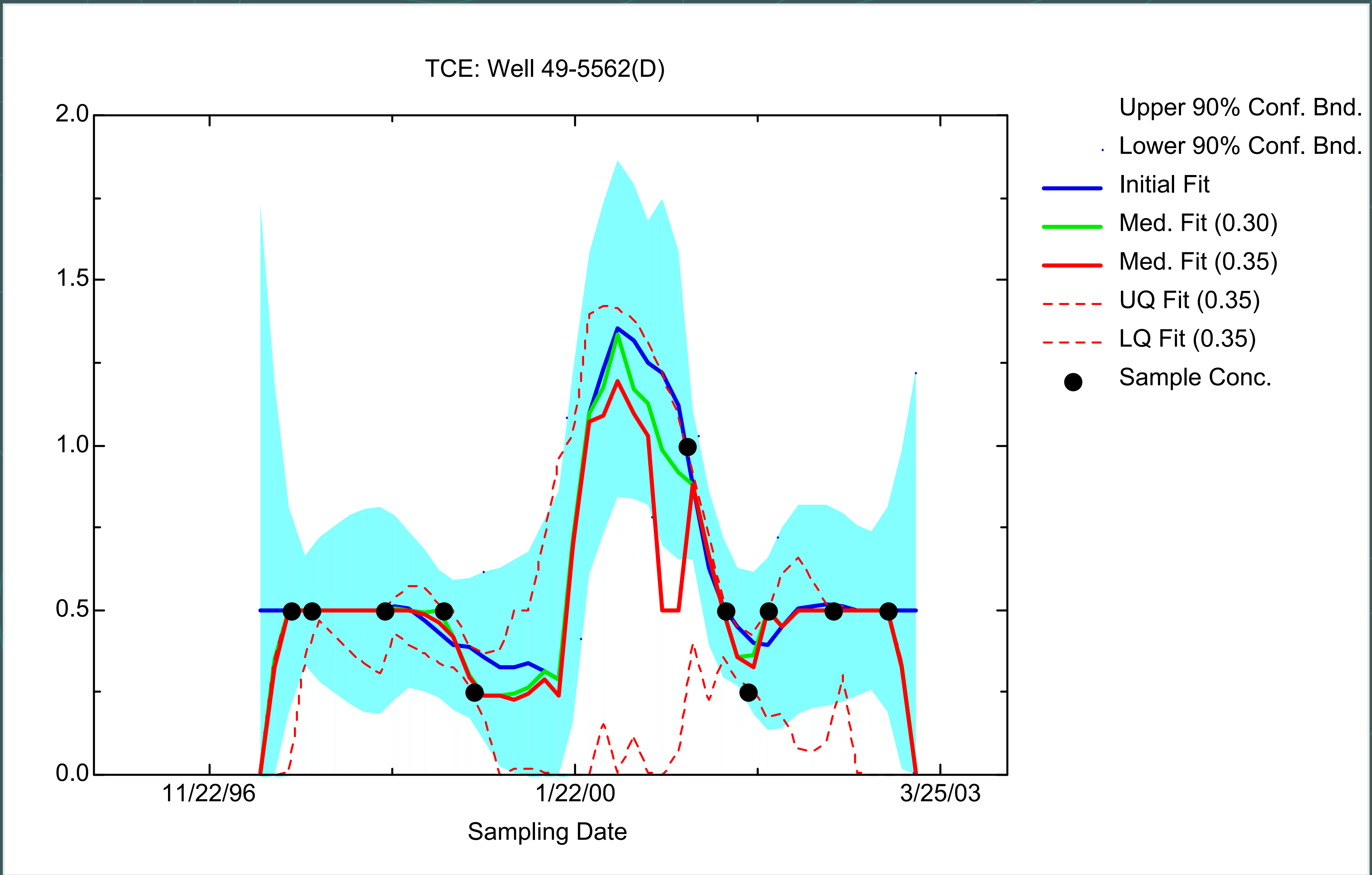
Iterative Thinning Details

- At least 8 sampling events per well
- NDs set to common imputed value
  - Guards against ‘apparent’ trend from varying DLs
- Complex trends, seasonal patterns OK
  - LWQR fits non-linear trends

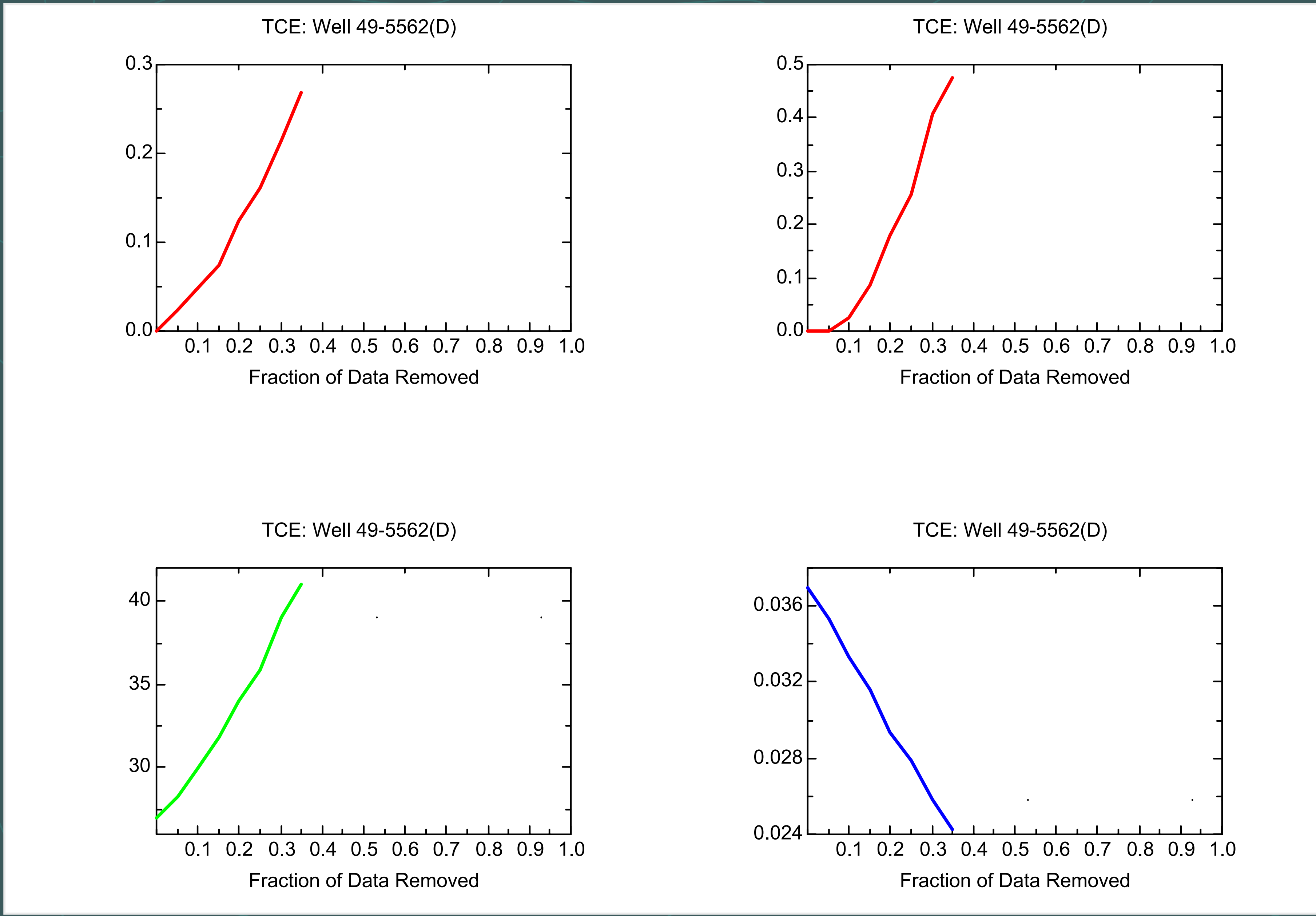
Iterative Thinning Example 1



Iterative Thinning Example 2



Iterative Fitting Diagnostics



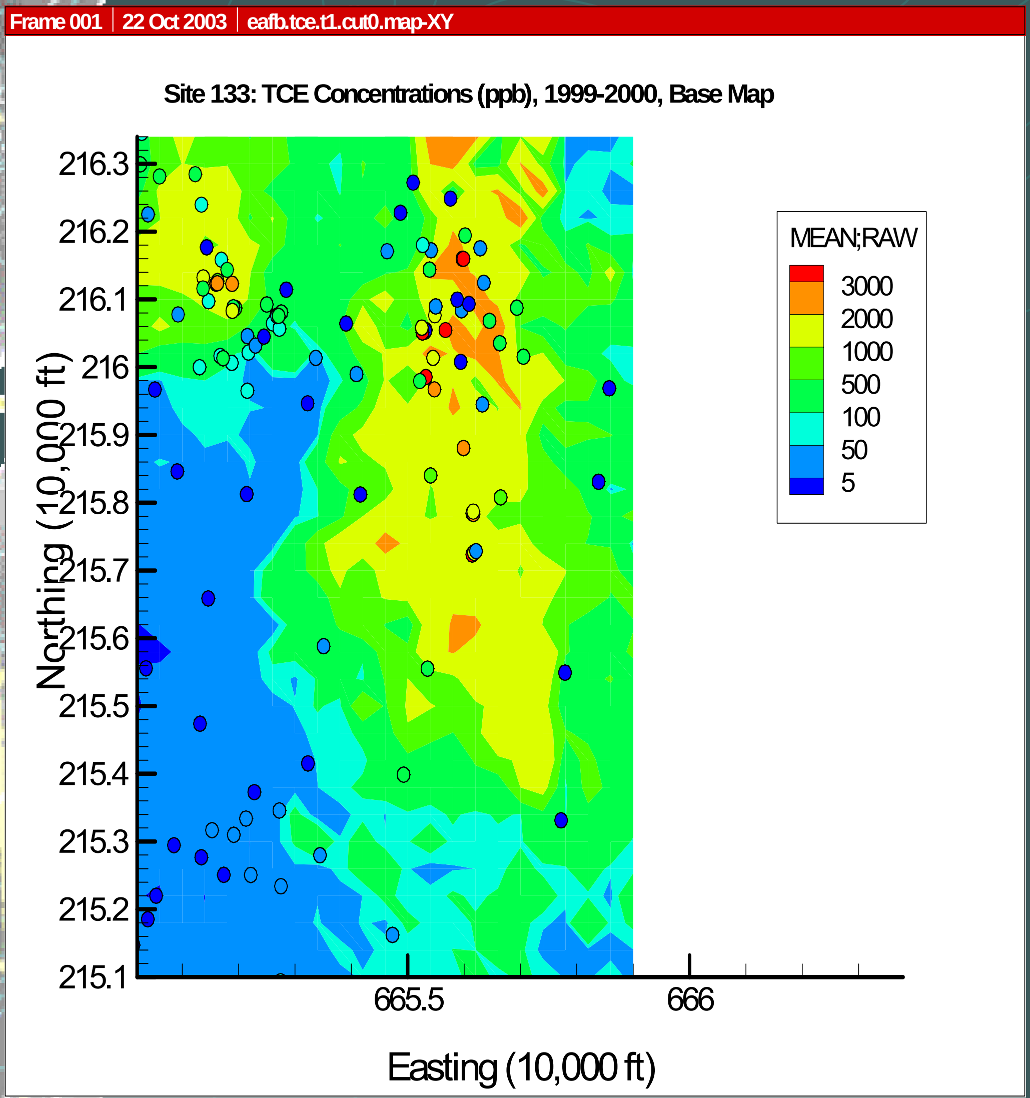
Spatial Optimization

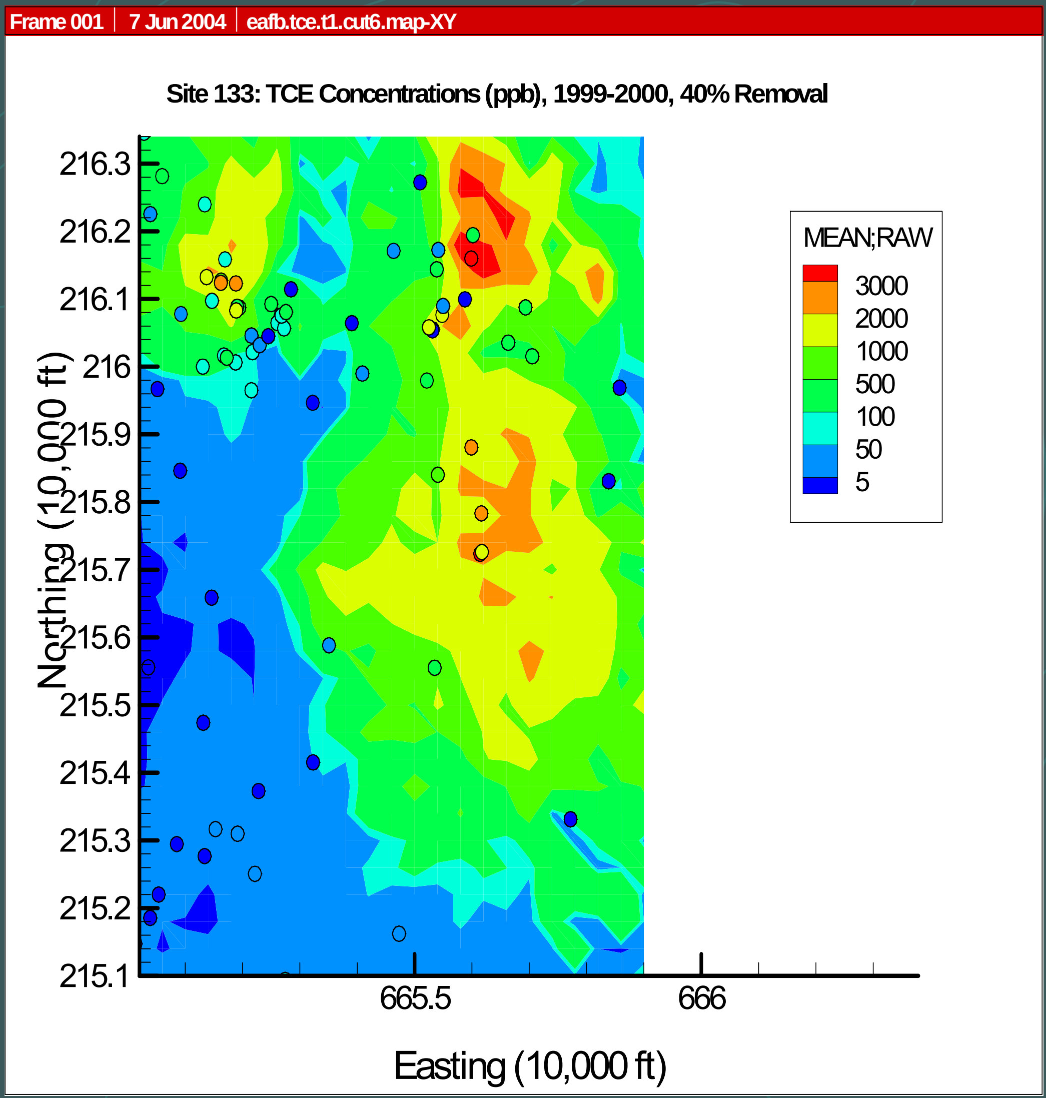
- Create base map, then series of "reduced-data" maps
  - LWQR used instead of kriging
- Spatial analysis
  - Estimate typical contribution from each well to site/plume maps (global regression wgts)
- Wells tagged for removal if their contributions are essentially duplicated by nearby wells
  - Redundant wells have low regression wgts

Basic Approach

- Baseline map uses all available data
- Iteratively remove lowest contributing wells; re-estimate map
- Measure loss of map quality/accuracy compared to baseline
  - Stop when maps deteriorate too much

Spatial Comparison





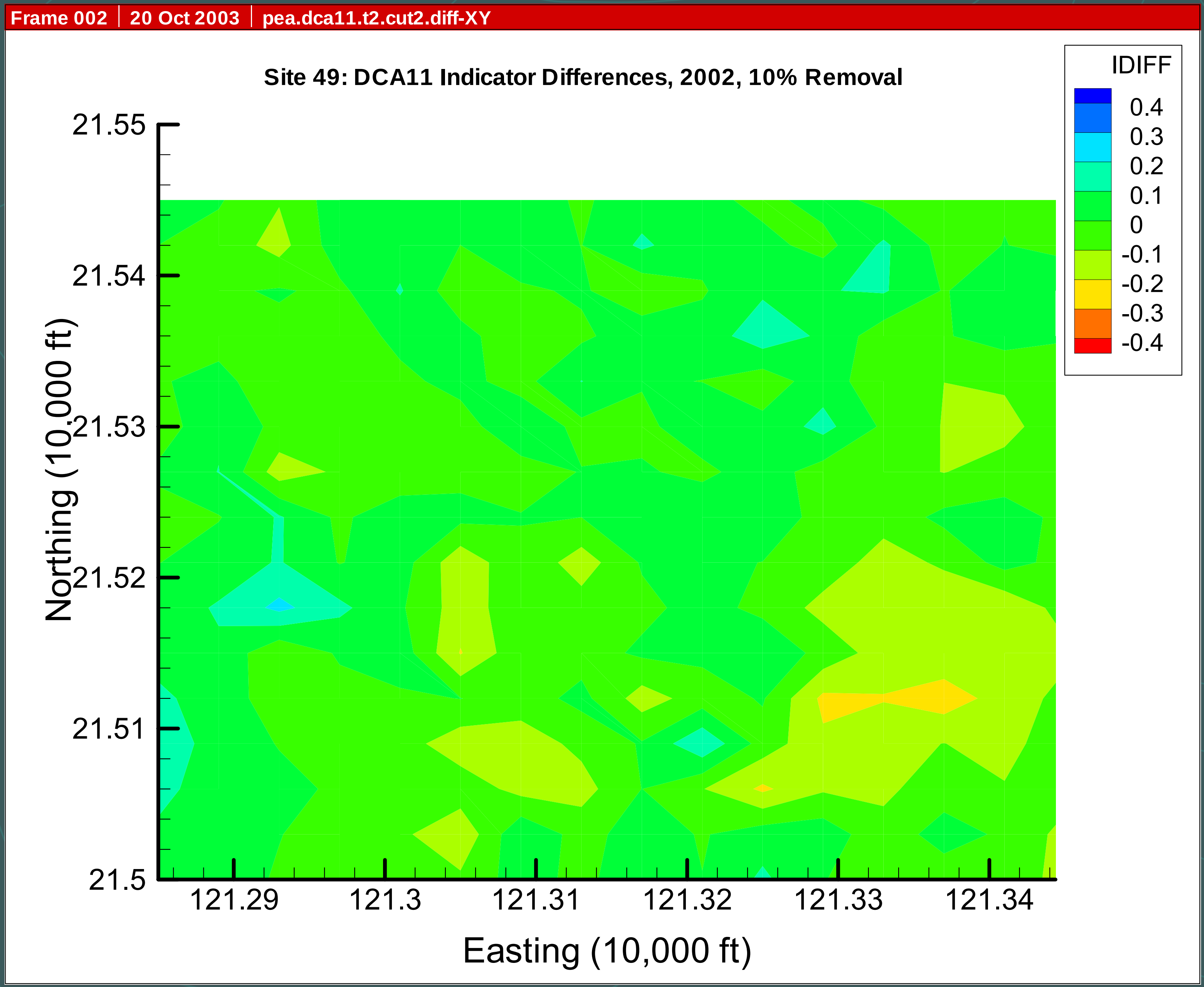
Features of Spatial Algorithm

- Advantages to LWQR approach
  - A priori spatial model not required
  - Smoother, not an interpolator
- Can build site maps either in:
  - 3-D space
  - Separately by depth horizon or geologic unit
  - Separately by regulatory or geographic unit
    - As long as enough data available per unit

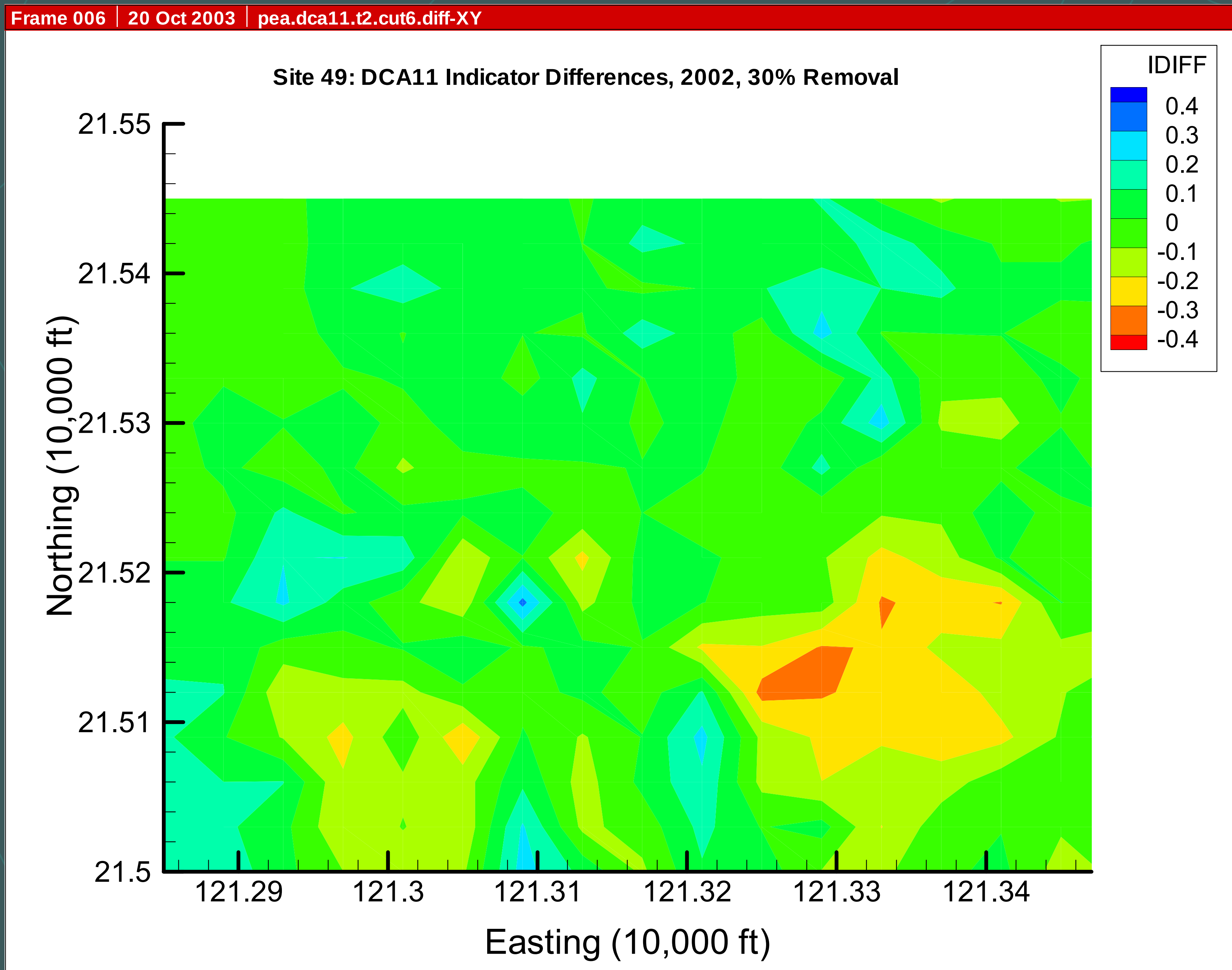
Features (cont.)

- Semi-objective spatial optimization
  - Iterative "removal" of lowest contributing wells/sampling locations
  - At each stage, measure:
    - Differences in site maps from baseline
    - Increases in global uncertainty and average bias
    - Prevalence of areas of high local uncertainty
    - Misclassification bias

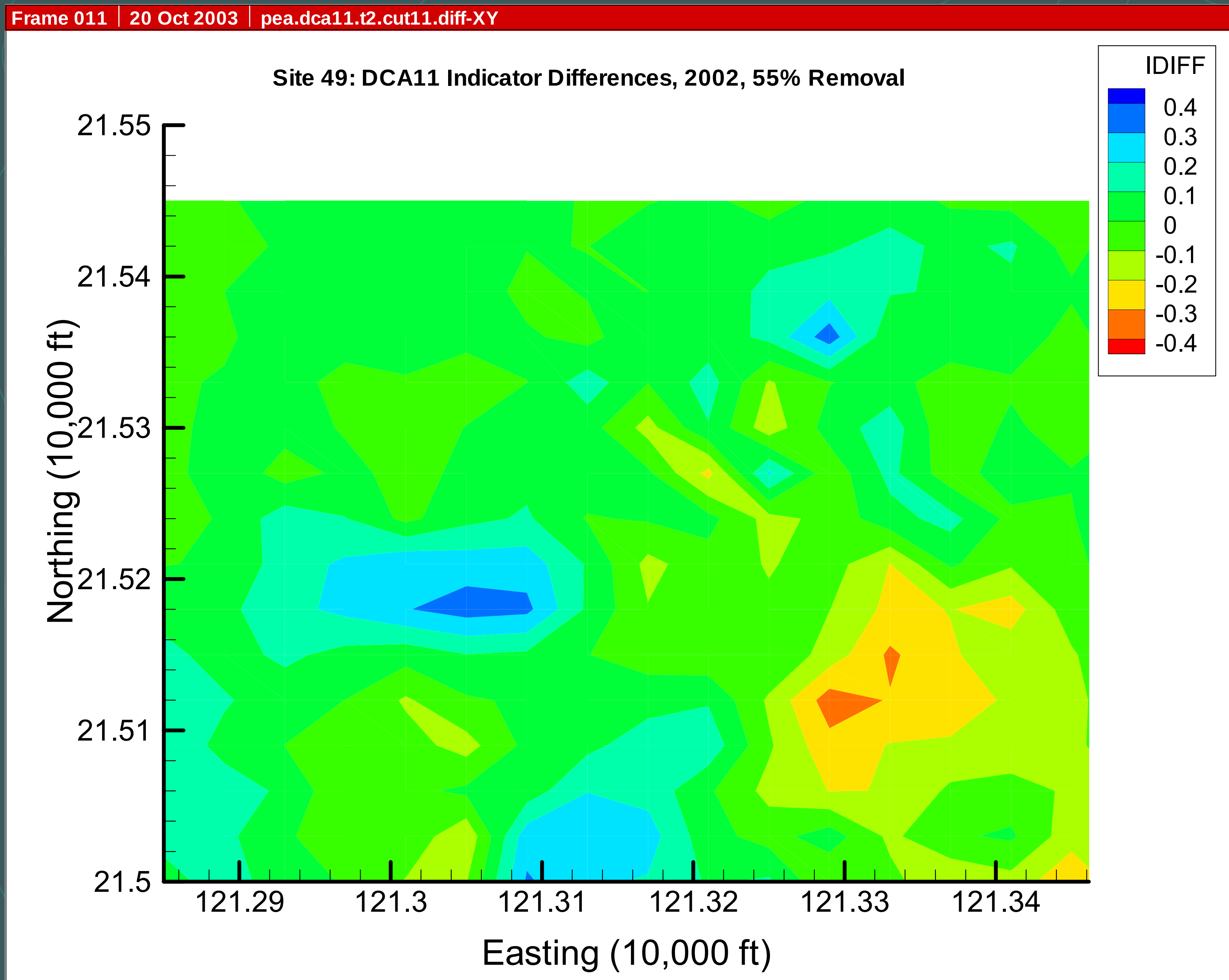
DCA11, 10% Removal



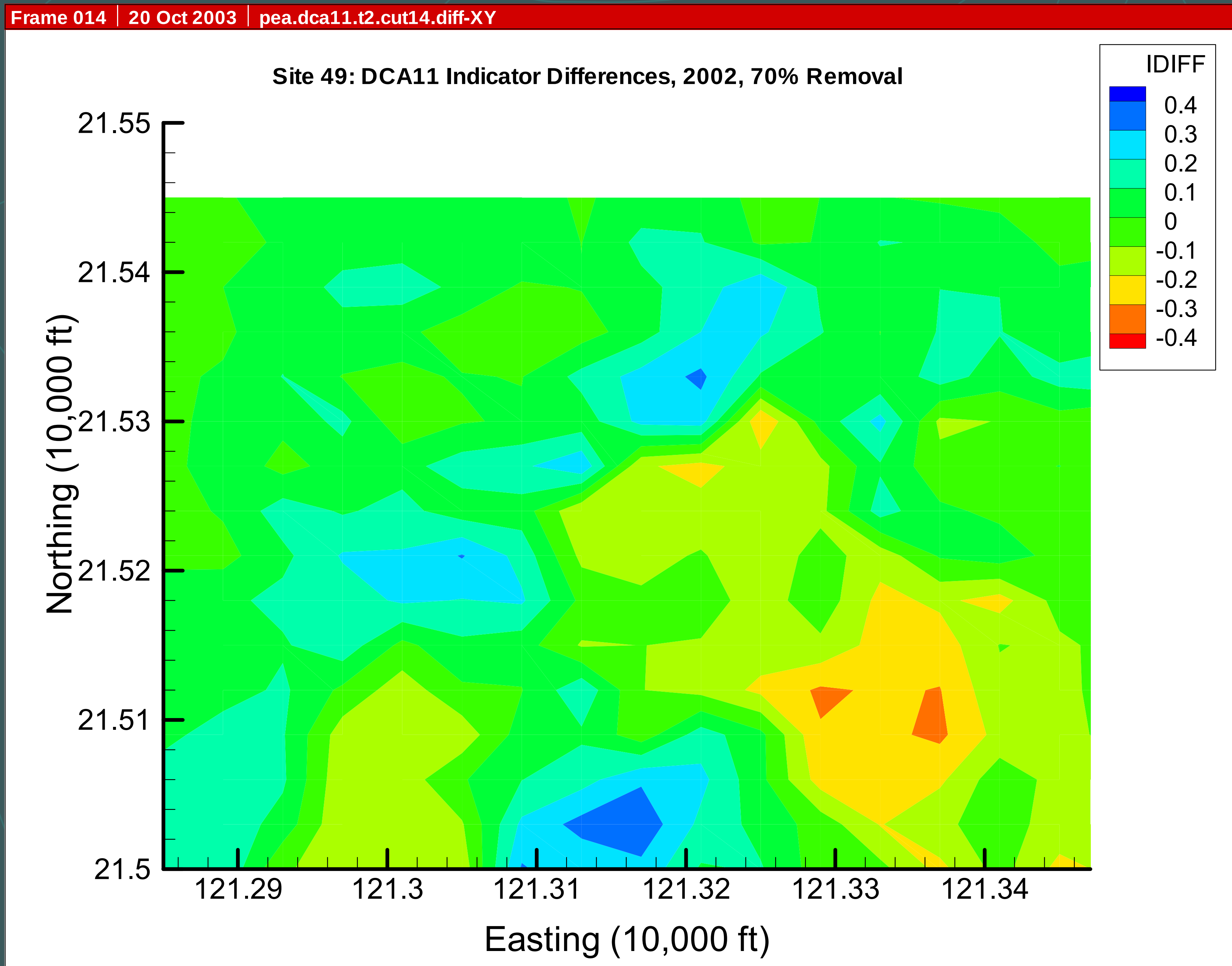
DCA11, 30% Removal



DCA11, 55% Removal



DCA11, 70% Removal



Spatial Requirements

- At least 20-30 regularly-monitored wells
  - Irregular sampling schedules OK
- Best COCs have:
  - Higher detection frequencies
  - Greater spatial spread & intensity
- Good to have 2-3 years of most recent monitoring data at each well
  - Data broken into ‘time slices’ (sampling date ranges); each time slice separately estimated
  - Multiple “snapshots” account for changing plumes

Robust Estimation Method

- Data never ideal, but method guards against skewness, transformation bias, impact of NDs
- Observed concentration range chopped into deciles; cutoff levels chosen
- Raw data converted to indicators at each cutoff
  - 0 or 1 according to whether cutoff exceeded
  - Logit link function
- Updated probability distribution (CCDF) created at each unknown location
  - Combines estimated probabilities of non-exceedance at each cutoff
  - CCDF used to construct site maps

Some Recent Results

	Edwards	Loring	Pease
Original Interval	Annual	Qtrly	Annual
Optimized Interval	Every 7 Qtrs	Every 2-3 Qtrs	Biennial
Redundant Wells	20-34%	20-30%	10-36%
Cost Reduction	54-62%	33-39%	49-52%
Annual Cost Savings	$230 K-$266 K	$306 K-$358 K	$85 K-$89 K

Summary: GTS Advantages

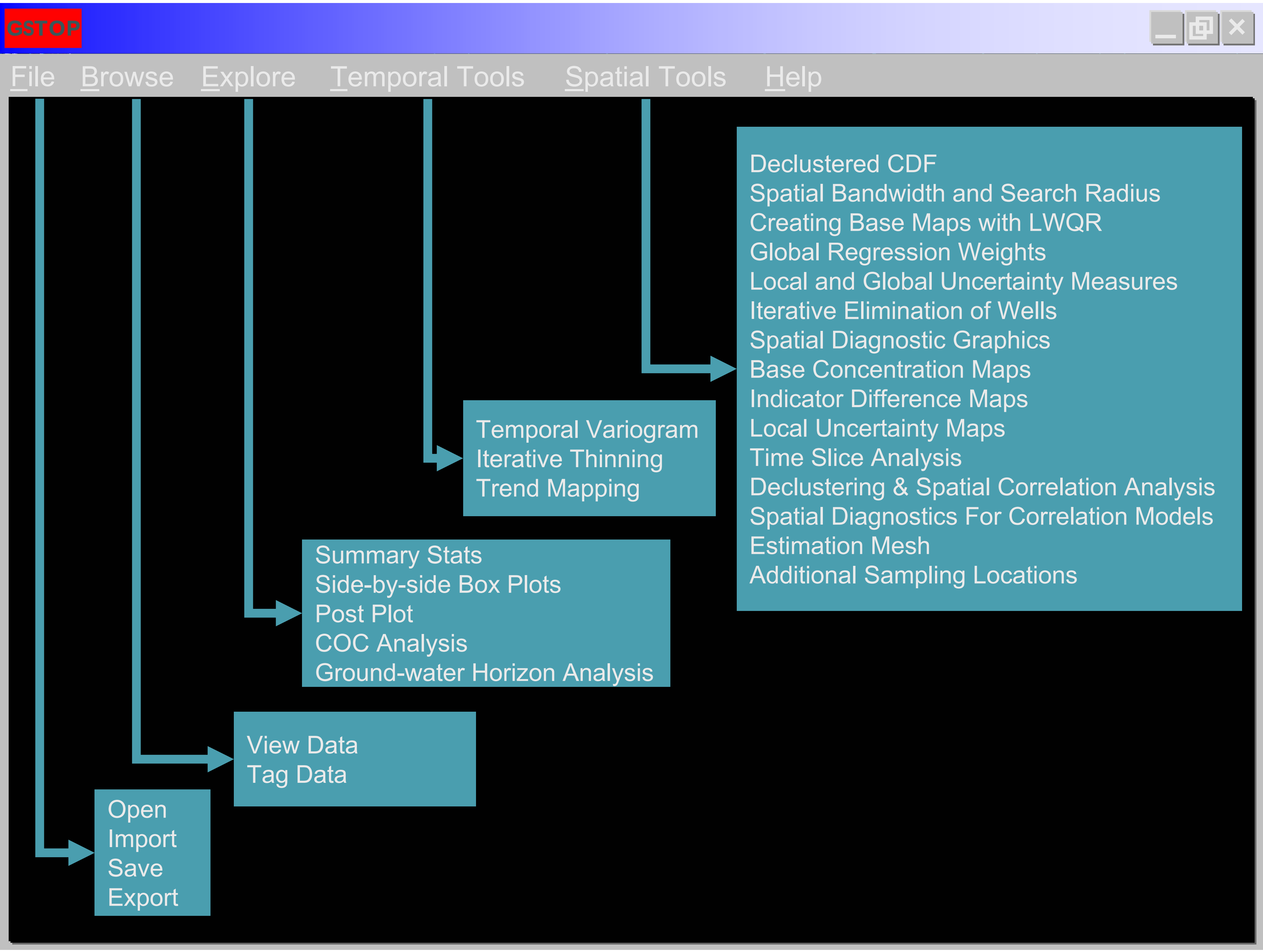
- Flexible strategies for optimizing sampling frequencies
  - Iterative thinning for individual wells
  - Temporal variogram for broad selection of sampling locations
- Objective criterion for ranking well locations according to redundancy
  - Global regression weights remove most subjectivity from process

Advantages (cont.)

- Emphasis on visual/graphical output
  - Graphs of temporal variograms
  - Site maps of concentration levels
  - Maps of local uncertainty
  - Plots of redundant and essential sampling locations

Toward the Future

- GTS uses geostatistical & spatial tools in a novel manner
- "Plug-in" architecture, flexibility
  - Temporal, spatial, or both
- Cost savings on order of at least 20-40%
- Stand-alone software coming


GSTOP
File
Browse
Explore
Temporal Tools
Spatial Tools
Help
Declustered CDF
Spatial Bandwidth and Search Radius
Creating Base Maps with LWQR
Global Regression Weights
Local and Global Uncertainty Measures
Iterative Elimination of Wells
Spatial Diagnostic Graphics
Base Concentration Maps
Indicator Difference Maps
Local Uncertainty Maps
Time Slice Analysis
Declustering & Spatial Correlation Analysis
Spatial Diagnostics For Correlation Models
Estimation Mesh
Additional Sampling Locations
Temporal Variogram
Iterative Thinning
Trend Mapping
Summary Stats
Side-by-side Box Plots
Post Plot
COC Analysis
Ground-water Horizon Analysis
View Data
Tag Data
Open
Import
Save
Export