U.S. Department of Energy
Office of Environmental Management

Technical Solutions Report No. 031102-035

Technical Solutions Study – Hanford Operations

Optimization of Groundwater Monitoring at the Hanford Site

Prepared by

The Office of Cleanup Technologies (EM-21)

February 8, 2006

EXECUTIVE SUMMARY

At the request of the U.S. Department of Energy Richland Field Office (DOE/RL), the DOE Headquarters Core Technical Group (EM-23) assembled a Review Team to explore possibilities for and issues related to Groundwater Monitoring Optimization (GWMO) at the Hanford Site. Review Team members were selected for their expertise in a variety of disciplines including hydrogeology, environmental engineering, mathematical optimization, database management, environmental chemistry, and statistics. The Review Team was provided with numerous documents for preliminary reading during July 2004. During its site visit from August 2-5, 2004, the Team received briefings from PNNL, Fluor Hanford, and DOE/RL personnel, and also met with key U.S. Environmental Protection Agency (USEPA) and State of Washington Department of Ecology (Ecology) personnel. Based on its discussions during that week, the Review Team prepared an out-briefing that was presented Thursday afternoon, August 5. This report provides a more detailed analyses and demonstrations concerning specific areas of interest to DOE/RL.

Groundwater monitoring at the Hanford Site is a complex and costly undertaking. There are several regulatory drivers for that monitoring. Monitoring conducted under the provisions of the Resource Conservation and Recovery Act (RCRA) is tied to regulated units that are sources or potential sources of contamination; this is administered by Ecology. Monitoring conducted under the provisions of the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), on the other hand, is tied to contaminated Operable Units (OUs); jurisdiction is shared between USEPA and Ecology. Monitoring is also conducted under provisions of the Atomic Energy Act (AEA).

In addition to conducting a holistic review of groundwater monitoring at Hanford, the Review Team discussed applicable GWMO techniques for specific programs at the Hanford Site. DOE/RL initially suggested three candidate areas for demonstrations. The Review Team ultimately selected the uranium plume in the 300 Area CERCLA Groundwater Operating Unit (300-FF-5 OU). This report presents demonstrations of two methodologies for identifying temporal redundancies, and hence reducing monitoring frequencies in an appropriate, scientifically supportable manner. The results of these demonstrations suggest that long-term monitoring of the 300-FF-5 uranium plume should occur, at most, on an annual basis. Current monitoring frequencies for those wells range from every month-and-a-half to every year-and-a-half.

The report also presents two methodologies for identifying spatial redundancies and identifying an appropriate, scientifically supportable subset of monitoring wells. The results of these demonstrations suggest that long-term monitoring of that plume could be reduced by 12 to 35 percent in terms of the number of wells with minimal loss of information, beyond which there are quantifiable trade-offs between information and cost. In addition, the report discusses optimization demonstrations previously conducted by PNNL using data related to contaminant plumes in the 200 West Area. These results suggest that from 15 to 42 percent of the wells can be removed, depending on the specific plume and constraints placed on the monitoring program.

The Review Team demonstration results are not presented as final solutions for monitoring the uranium plume in the 300-FF-5 OU, but rather as potentially applicable methodologies. This CERCLA OU has been proceeding under an interim remedy and is not yet in a long-term monitoring mode. Prerequisites to being in a long-term monitoring mode include having an accepted, demonstrably viable final remedy in place (including a reliable conceptual model for the uranium transfer from vadose zone to groundwater to river water) and having at least quasi-steady-state groundwater chemistry in the OU. A forthcoming Focused Feasibility Study (FFS) will be reviewing these issues at this OU. Hence, the Review Team demonstrations are represented as hypothetical examples, even though they are based on historical data from the 300-FF-5 OU.

In the course of studying the 300-FF-5 OU, the Review Team identified a number of supplementary issues related to the FFS and presented several recommendations. In particular, "bounding" engineering studies of uranium fate and transport should be performed in parallel with the more detailed geochemical and modeling studies currently underway. These could be based on both available data and a quick medium-scale experiment of leaching, desorption, and adsorption using 300 Area soils. In addition, improved groundwater flux estimates should be obtained; the report briefly discusses several innovative technologies that might be useful.

Another of the areas of interest to DOE/RL and proposed for demonstration is the LLWMA-3 Waste Management Area, a RCRA unit. This is a poor candidate for GWMO demonstrations, primarily because the unit is in a detection monitoring program. Therefore it presumably has not yet impacted the groundwater, so GWMO methods appropriate for long-term monitoring of plumes and trends of known constituents of concern are not relevant.

Nonetheless, the Review Team did review the LLWMA-3 and identified a number of monitoring inefficiencies, primarily associated with choices of statistical methods and monitoring frequencies under various RCRA programs, as well as with well siting. The Review Team concurs with PNNL and DOE/RL that, so long as LLWMA-3 and similar RCRA units must be monitored individually rather than as part of a site-wide integrated monitoring program, monitoring based on intrawell comparisons, clearly allowed under the RCRA 40 CFR Part 264 Final Status rules, is preferable to monitoring based on upgradient-downgradient comparisons, apparently the only allowable plan under the 40 CFR Part 265 Interim Status rules.

The specific intrawell statistical comparisons proposed by PNNL and DOE/RL and reviewed by Ecology are those using Combined Shewhart-CUSUM Control Charts (CCCs). Unfortunately, the statistical performances of the CCC plans recommended by Ecology and those recommended by PNNL differ considerably. Some of the differences can be attributed to choices of numerical parameters, which obviously affect false positive and false negative rates. More importantly, some of the differences are due to the use or non-use of resamples and/or automatic updating of background datasets. This report presents a detailed evaluation of various plans and concludes that the CCC plan recommended by Ecology should not be considered because of a very high false positive rate, due to not allowing confirmatory resamples. In contrast, some of the plans recommended by PNNL have insufficient sensitivity to slow releases due to their use of automatic updating of background. The report discusses procedures for comparing the statistical performances of CCC plans and intrawell prediction limit plans.

With regard to monitoring frequency, the Review Team concurs that the regulatory requirements for both Interim Status and Final Status are excessive. DOE/RL should review the historical evolution of this regulation, however, in preparing its proposals with regard to monitoring frequency. The next set of rules promulgated by USEPA (i.e., 40 CFR Part 258 rules for groundwater monitoring at solid waste management facilities promulgated in 1991) are much more reasonable in this regard. Furthermore, it is clear from the preamble in the *Federal Register* to the 1988 revision of the Final Status rules that the monitoring frequency required (i.e., four independent samples per period, often per quarter) was (1) intended to support the use of the ANOVA statistical tests in a “one-point-in-time” fashion that is not appropriate for use with indicator parameters with inherent spatial variability; (2) is no longer encouraged in RCRA guidance; and (3) is clearly excessive when using intrawell comparisons.

In its proposals for alternative statistical tests, PNNL also suggests detrending data for certain indicator parameters with clear trends in incoming groundwater and adjusting uranium concentration measurements made in wells in the 300 Area near the river for the apparent dilution caused by mixing with river water during high river stages. The concern in the detrending proposal is that following decades of discharge of river water used for process cooling, the groundwater is returning to ambient conditions with higher conductivity related to higher concentrations of naturally occurring constituents. The Review Team suggests that two alternate approaches be considered. One is updating background data in a non-automatic fashion whenever confirmed exceedences are found. The other is using intrawell prediction limits combined with rudimentary modeling of the apparent return of the groundwater chemistry to ambient conditions. With regard to the mixing adjustments, it does appear that the procedure proposed by PNNL may reduce the variability in the data, which in turn should produce a more sensitive monitoring program. If, in addition to using the resulting adjusted data for intrawell comparisons, they are to be used for comparisons with a concentration limit, a simple experiment should be conducted to verify the mixing model, if it has not been done already.

Well siting is the other key issue at LLWMA-3. Wells are going dry and need to be replaced. Current plans are to emplace eight new wells, all in the current downgradient direction, at locations determined using the Monitoring Efficiency Model (MEMO). The Review Team recommends the following: (1) the proposed location strategy should be revised to ensure upgradient wells regardless of changes in groundwater flow direction; (2) the use of the MEMO be reviewed and made more appropriate in view of the known and unknown discontinuities and inhomogeneities in the vadose zone between LLWMA-3 and the groundwater; and (3) the current practice of locating monitoring wells at the waste boundary be reconsidered in view of (2), the likely future capping of LLWMA-3, and enhancing the usefulness of all new wells to all programs, not just RCRA monitoring at LLWMA-3.

The Review Team encourages continued efforts to coordinate RCRA monitoring of (potential) source units with CERCLA monitoring of remedies for known contamination, in the interests of efficiencies in both, not only at LLWMA-3 but also wherever RCRA units are embedded in CERCLA OUs. The Review Team suggests that DOE/RL, USEPA, and Ecology explore the utility of pre-negotiated decision frameworks for streamlining the decision process by which decisions regarding the evolution of groundwater monitoring at the Hanford Site may be made. Such pre-negotiated decision frameworks are currently being used at a variety of large facilities. Finally, DOE/RL, USEPA, and Ecology should explore the benefits and efficiencies of groundwater monitoring using innovative technologies, such as sensor systems and networks and non-intrusive devices, in both the saturated and vadose zones.

In summary, this report provides demonstrations of GWMO methods that will one day be applicable at the Hanford Site on a large scale. It also provides focused recommendations regarding information needs with regard to the forthcoming FFS for the 300 Area uranium plume and RCRA detection monitoring and well siting at the LLWMA-3 and similar facilities. The final set of recommendations addresses shifting both the regulatory oversight and the monitoring scope from the current localized concerns to integrated, sight-wide programs with pre-negotiated decision frameworks, eventually enhanced by the incorporation of innovative monitoring technologies.

TABLE OF CONTENTS

TABLE OF EXHIBITS

1.0 OVERVIEW

1.1 Purpose of this Technical Assistance Activity

At the request of the U.S. Department of Energy Richland Operations Office (DOE/RL), the DOE Headquarters Core Technical Group (EM-23) organized a Review Team to conduct an optimization review of the groundwater monitoring program at the Hanford Site. Appendix A contains a copy of the original technical assistance request from DOE/RL. The Review Team was asked to analyze and make recommendations on improving the effectiveness and achieving cost savings in groundwater monitoring. Specific areas of focus included the 200 West low-level mixed waste burial grounds (LLMWA-3) regulated under the Resource Conservation and Recovery Act (RCRA) and the 300 Area uranium groundwater plume operating unit (300-FF-5 OU) regulated under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA). The Review Team consisted of the following individuals:

- Beth Moore, US DOE/HQ, Hydrogeologist and Project Manager;
- Charles Davis, Ph.D., Environmetrics & Statistics Limited, Statistician and Review Team Leader;
- Kirk Cameron, Ph.D., MacStat Consulting, Ltd., Statistician;
- David Dougherty, Ph.D., Subterranean Research, Inc., Hydrogeologist;
- Rob Greenwald, Ph.D., GeoTrans, Inc., Hydrogeologist;
- Barbara Minsker, Ph.D., University of Illinois, Environmental Engineer; and
- Maureen Ridley, Lawrence Livermore National Laboratory (LLNL), Chemist.

Kathy Yager of the US Environmental Protection Agency (USEPA) Office of Superfund Remediation and Technology Innovation also participated in the site visit and discussions. Ms. Yager's role was to offer assistance in technical transfer of optimization methods and techniques through USEPA Region 10, should DOE/RL, the USEPA regulatory staff, and the State of Washington Department of Ecology (Ecology) regulatory staff adopt the Review Team recommendations.

A series of background and reference documents was made available to the Review Team in advance of the site visit. The site visit occurred during August 2-5, 2004, and included a tour of the Hanford Site with a special emphasis on the 200 West and 300 Areas. The site visit also included presentations by and discussions with PNNL, Fluor Hanford, DOE/RL, USEPA, and Ecology personnel. Additional reference materials were provided during and following the site visit to fill data gaps identified by the Review Team. All documents provided are listed in Appendix B.

The charge to the Review Team was to suggest potential strategies and new evaluation techniques that can improve the Hanford groundwater monitoring programs. The issues, optimization approaches, and analytical techniques evaluated by the Review Team ranged from specifics (e.g., the benefit of statistical analyses using intrawell versus interwell comparisons for RCRA detection monitoring) to broad issues (e.g., using a decision framework to guide program improvement in the out years and further enhancing regulatory coordination at the site). This report includes demonstrations of several groundwater monitoring optimization methods to remove data redundancy prepared by Review Team members using data from the 300-FF-5 OU. The Review Team's recommendations address both issues specifically relevant to 300-FF-5 OU and LLMWA-3 and broader site-wide and long-term monitoring issues as well.

1.2 Description of the Hanford Site

The Hanford Site is a DOE facility located along the Columbia River in Eastern Washington State, near the town of Richland (Exhibit 1). It is approximately 586 square miles in area. Plutonium was produced at Hanford from the 1940's through the late 1980's. Operations included the discharge of wastewater into trenches, pits, and ponds. Discharge of untreated wastewater to the ground ceased in 1995. There are also numerous landfills on the Hanford Site containing hazardous and radioactive wastes. Some landfills, including low-level and mixed low-level waste management areas, remain in service. Historical site activities have resulted in releases of contaminants to the environment including radionuclides, volatile organic compounds, and metals.

Tritium, nitrate, and iodine-129 are the most widespread contaminants in groundwater. The most prominent portions of these plumes originated at waste sites in the 200 Areas and spread toward the southeast. Nitrate and tritium also had significant sources in the 100 Areas. Other contaminant plumes include:

- Carbon tetrachloride and trichloroethene in the 200 West Area;
- Chromium in the 100 Areas;
- Chromium in the 600 Area south of the 200 Areas;
- Strontium-90 in the 100 Areas;
- Technetium-99 and uranium that extend eastward from the 200 West Area;
- Technetium-99 and uranium with minor amounts of cyanide and cobalt-60 in the northwest 200 East Area; and
- Uranium in the 300 Area.

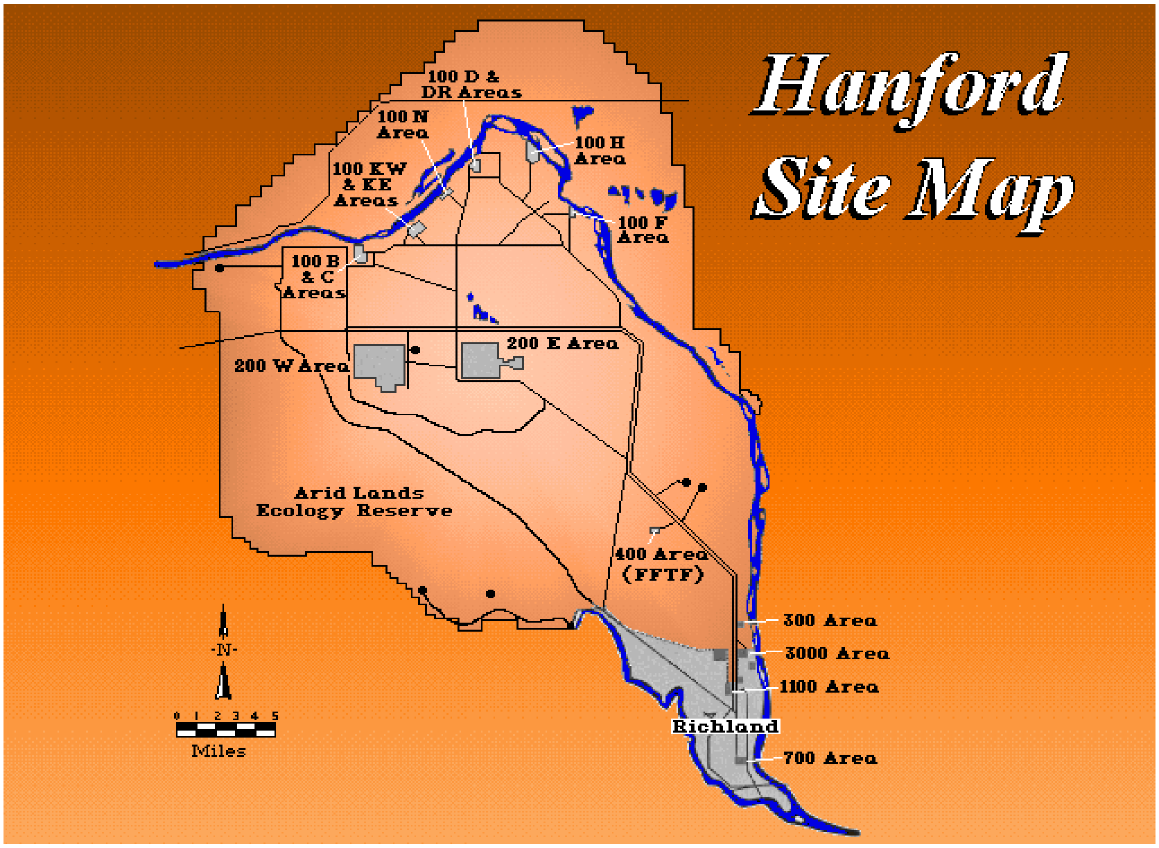
A series of interim groundwater remedies has been implemented at the Hanford Site, including several pump-and-treat operations in 100 Area and 200 Area Operable Units (OUs) and an interim Monitored Natural Attenuation (MNA) remedy in the 300 Area Groundwater Operable Unit (300-FF-5 OU).

1.3 Overview of Current Groundwater Monitoring Programs

Groundwater monitoring is conducted at the Hanford Site for a variety of reasons, including:

- Maintaining compliance with applicable regulations;
- Characterizing and defining trends in constituent concentrations and water levels;
- Assessing and monitoring the progress of groundwater remediation activities; and
- Identifying new contamination.

Exhibit 1. The Hanford Site



The Hanford Site is regulated under a variety of monitoring programs. In July 1989, the Hanford Site was listed on the National Priorities List (NPL) as four separate NPL sites (100 Areas, 200 Areas, 300 Area, and 1100 Area), each of which is further divided into OUs based on geographic area and waste sources. Additionally, the site operates under a single site-wide RCRA permit; RCRA groundwater monitoring was conducted during fiscal year 2003 at 24 waste management areas. For instance, the low level burial grounds, including LLWMA-3, are regulated as RCRA treatment, storage, and disposal facilities. RCRA regulations are administered by Ecology. Jurisdiction of CERCLA activities is shared between Ecology and USEPA Region 10, pursuant to the 1989 Tri-Party Agreement. Some units are also regulated under the Atomic Energy Act (AEA).

There is potential for redundant sampling at Hanford as a result of the multiple regulatory drivers; in particular, RCRA-regulated units are located within the footprints of plumes that are regulated under CERCLA. A further complication is that contaminants originating from some areas have migrated over large distances into other geographic areas of the site, causing overlapping plumes.

1.4 Summary of Current Conceptual Modeling Status

Conceptual modeling involves establishing a fundamental understanding of the following:

- Site hydrogeology, including depth to ground water, ground water flow direction (both horizontal and vertical), approximate ground water flow magnitude, existence and prevalence of flow channels and aquatards, interaction of ground water with surface water;
- Historic and continuing sources of groundwater contamination;
- Potential human and ecological receptors;
- Historic and current extent of contamination with respect to sources, potential receptors, and other landmarks such as property boundaries; and
- Historic site remedies, such as source removal or control activities, that impact the site conceptual model.

The Hanford Site hydrogeology has been extensively characterized, and the general stratigraphy is well understood. The site is underlain, from top to bottom, by Pleistocene-age catastrophic flood deposits (primarily sands and gravels) of the Hanford formation; semi-consolidated sands, silts, and clays of the Ringold Formation; and Tertiary-age flood basalts that typify the surrounding Columbia Plateau. However, localized heterogeneities are not always well defined, which creates uncertainty in the understanding of the fate and transport of contaminants. Water levels have declined, and continue to decline, since significant discharge of water to the land surface ceased in the mid 1990's. This has caused some monitoring wells in key areas to go dry, and has also caused regional groundwater flow directions to change over time. The stage of the Columbia River is highly variable, due to varying demands for hydroelectric power among other influences; this impacts both water levels and groundwater quality in the immediate vicinity of the river.

Many historic and continuing sources of groundwater contamination have been identified. Some potential future sources of groundwater impact exist. Receptors, contaminant extent, and impacts of previous remedial actions are somewhat specific to individual management areas and/or plumes. The Review Team was asked to focus on two areas:

- 300-FF-5 OU: a groundwater operating unit regulated under CERCLA, with an interim MNA remedy associated primarily with uranium, located along the banks of the Columbia River just north of Richland; and

- LLWMA-3: a low-level waste management area, with a RCRA interim status detection monitoring program, located in the 200 West Area of the Central Plateau.

The conceptual model for the 300-FF-5 OU is still evolving. An interim remedy for groundwater based on MNA was emplaced in 1994 to address groundwater contamination (primarily uranium). The selection of MNA for the interim remedy was based on a conceptual model that incorporated simplifications such as:

- No continuing source in the unsaturated zone was assumed, and no net recharge through ground surface was assumed;

- A constant K_d was assumed independent of water quality and geochemistry; and

- Interaction between the Columbia River and the aquifer (including bank storage and springs) was highly simplified.

These are addressed in greater detail in Section 3.2 of this report.

Since the interim MNA remedy was implemented, it has become apparent that there is likely a continuing source of uranium contamination to the groundwater in the vadose zone; that K_d varies as a function of water quality; and that river water mixing with aquifer water is an important process affecting the fate and transport of uranium. Therefore, the conceptual model is currently being updated in conjunction with remedy evaluations in preparation for the forthcoming Focused Feasibility Study (FFS) associated with selection of a final remedy.

For LLWMA-3, the issue addressed by the Review Team is detection monitoring of the waste facility. LLWMA-3 is located in the northern portion of the 200-West Area, on the Central Plateau. Conceptual modeling issues at LLMWA-3 include the following (also discussed in more detail later in this report):

- The direction of groundwater flow underneath LLWMA-3 has changed significantly during the past two decades and is expected to change further in the future, complicating the definition of "upgradient wells" and "compliance wells;"

- Many of the existing monitoring wells have gone dry due to declining water levels; and

- Many other facilities are located near LLWMA-3 and some have caused groundwater contamination that extends to and under LLWMA-3, presenting complications with respect to the RCRA detection monitoring being performed.

In addition, a modeling tool previously used to site wells is based on a very simplified conceptual model of vertical transport thru the vadose zone, which may not fully account for lateral spreading of contaminants in the vadose zone (see Section 4.4.2 below). Given these conceptual modeling issues, it is expected that the conceptual model of LLWMA-3 will continue to evolve over time.

1.5 Organization of this Report

Section 1 outlines the purpose of this Technical Assistance Activity and reviews the background and current situation at the Hanford Site with regard to groundwater monitoring and groundwater conceptual modeling.

Section 2 presents the broad topic of monitoring optimization. This presentation begins with a discussion of the requirements that an optimizable long-term monitoring program must possess and evaluates the current situation at Hanford with respect to those requirements. It reviews a variety of optimization techniques briefly; Appendix C contains a more extended introduction. It then presents a number of examples and demonstrations of optimization using Hanford Site data, including several performed previously by PNNL personnel as well as several prepared by Review Team members using methods different from those in the PNNL demonstrations.

Although its demonstrations are based on the 300-FF-5 OU data, the Review Team does not believe that the 300-FF-5 OU is ready for the establishment of a long-term monitoring program at this time. Therefore these demonstrations should be regarded as hypothetical, showing the efficiencies that could be obtained in monitoring a situation like that of the 300-F-5 OU if it were in long-term monitoring.

Section 3 discusses additional issues identified by the Review Team with regard to the 300-FF-5 OU, particularly with regard to the forthcoming FFS and the advances needed before long-term monitoring would become appropriate at that OU.

Section 4 reviews the situation with regard to RCRA groundwater detection monitoring at LLWMA-3, which is very similar to that of numerous other RCRA regulated units at Hanford. RCRA detection monitoring is different from plume monitoring or long-term monitoring of a unit at which there is a known history of contaminant data. The issues at this and similar units are two: the choice of appropriate statistical methods and monitoring frequencies; and the choice of appropriate monitoring well numbers and locations. The optimization techniques discussed and demonstrated in Section 2 are only tangentially related to these issues.

A common theme throughout the discussions of the specific areas is the benefit to be derived from encouraging the continued evolution of regulatory coordination between RCRA and CERCLA programs and between DOE/RL, USEPA, and Ecology. **Section 5** revisits this topic in the general context of integrated site-wide long-term monitoring optimization. **Section 6** contains a brief discussion of pre-established decision frameworks, which have proven to be quite valuable in providing efficiencies in making monitoring program decisions at major facilities comparable in size and complexity to the Hanford Site.

In the future, groundwater monitoring may become much less dependent on data obtained via traditional sampling and analysis, and instead rely on data from innovative technologies based on *in situ* methods and sensor networks. Potential technologies in this area are presented in **Section 7**. Finally, **Section 8** summarizes the Review Team's specific recommendations contained in previous sections.

2.0 GROUNDWATER MONITORING OPTIMIZATION

2.1 Prerequisites to Optimization

An analysis of groundwater monitoring optimization (see Appendix C) is typically conducted within the context of long-term monitoring in a quasi-steady-state setting, such as monitoring the progress of selected remedies in attenuating known contaminant plumes at a facility. Before attempting to optimize such a groundwater monitoring program, the underlying situation must be well understood and the selected remedies must be agreed upon by the facility and relevant stakeholders. Understanding the situation requires a demonstrably reliable conceptual model for the evolution of the contaminant plumes under the selected remedies. It also requires a consensus among the relevant parties about the desired end state and the likelihood that the selected remedies will achieve that end state. Nearly always there is a data history for the constituents of concern (COCs) and other appropriate monitoring parameters. There must also be a consensus about the performance objectives for monitoring.

These concerns do not arise in detection monitoring because the regulated unit is assumed not to be impacting the groundwater. In many cases COCs and other monitoring parameters can be and should be identified based on the prevalence, mobility, and detectability of waste constituents. There is seldom any data history for these on which to base any characterization of plume behavior or validate conceptual hydrogeological or fate and transport models. Moreover, at the present time, detection monitoring of a unit regulated under RCRA is subject to regulations that have historically been quite inflexible. Hence, optimization of groundwater monitoring for RCRA-regulated operating units is more a matter of establishing an efficient program under the applicable regulation as well as determining an appropriate regulatory framework for the regulated unit. Appendix G contains an extended discussion of these matters, with particular reference to one of the RCRA-regulated units in detection monitoring status that appears prominently in PNNL and DOE/RL discussions of the issue. That unit (at the 300 Area Process Trenches) is not LLWMA-3, which the Review Team was asked to investigate; however, the issues are virtually identical.

Even if the RCRA-regulated units at Hanford are not considered, there are a number of CERCLA operating units, related to multiple contaminant plumes and their remedies. Each of these has its regulatory drivers, with some overlap in jurisdiction. The remedies for these units are in different states of finality. Moreover, as stated previously and discussed in detail in Section 3.2, the conceptual models needed to construct long-term monitoring programs are likewise in a variety of states of validity.

The prerequisites for entering into long-term monitoring scenarios at the Hanford Site are therefore not well satisfied at this time. The Review Team suggests that, in addition to satisfying these prerequisites with respect to individual monitoring efforts, Hanford should coordinate its various monitoring efforts in pursuit of an overall integrated groundwater end-state solution.

2.2 Review of Optimization

A monograph recently published by ASCE (EWRI [2003]) discusses optimization of long-term monitoring programs. Appendix C presents a brief summary of major points, and

Section 7 discusses some of the more visionary concepts contained in that monograph. The analyses alluded to in Section 2.1 pertain to optimizing monitoring in existing networks with data histories involving the COCs. The goals of the optimization exercises are to identify opportunities for improving efficiencies through one or more of the following:

- Removing wells from the monitoring network if their data are too highly correlated with the data from neighboring wells;

- Adding wells in areas where the model underlying the optimization suggests that data would not be highly correlated with that from neighboring wells and/or information gaps may exist; and

- Adjusting the frequency of monitoring with the objective of minimizing the redundancy of information provided during each monitoring event.

Of the approaches discussed in Appendix C, the demonstrations to follow are based on statistical and geostatistical models for the measurements for a single COC or other monitoring parameter. In two of the four Review Team demonstrations as well as the PNNL demonstrations, the statistical model is based on spatial or temporal autocorrelations fit to the data as a whole; in another an alternate method is used for evaluating the local predictability of measurements; and in the fourth the redundancy is evaluated through trend reconstruction using subsets of the data, comparing actual observations with predictions made using reduced data sets. Each case establishes a quantitative measure of the loss of information suffered when wells or events are removed. This measure is optimized over the specific wells to be removed, in the spatial case. The loss of information is then evaluated as a function of the reduction in cost of the program, so that one can evaluate the trade-offs involved. The number of choices to be made in the optimization process is typically quite large, to the extent that mathematical or computational decision support is typically called upon to assist in the search for desirable solutions.

2.3 Optimization Demonstrations and Examples

Demonstrations of specific groundwater monitoring optimization approaches are included in this section to illustrate certain specific techniques as well as the value of optimization. Given the time and budget limitations associated with this GWMO review, the Review Team's intent is to demonstrate principles, recognizing that comprehensive analyses to solve particular problems at the Hanford Site will be likely to require additional effort beyond this technical assistance activity, and may very well be premature given the realities discussed in Section 2.1 and elsewhere. The following types of demonstrations were considered initially:

- Identification of redundant well locations (spatial redundancy) in a plume- mapping monitoring network (i.e., monitoring wells are used to generate maps of plume concentration values and contours);

- Identification of spatial redundancy in a natural attenuation monitoring network (i.e., monitoring wells are used to assess the stability of plume shape and the dissipation of contaminant); and

- Identification of unnecessarily frequent monitoring (temporal redundancy) in a groundwater monitoring network.

Discussions and presentations during the site visit identified already-completed and on-going plume-mapping monitoring optimization at PNNL that made the first candidate demonstration unnecessary. Pre-visit communications between DOE/RL and DOE/HQ identified two locations at which groundwater monitoring optimization was of interest and within the scope possible by the Review Team: the Area 300-FF-5 OU uranium plume; and the low-level waste burial ground LLWMA-3. The issues at LLWMA-3 are well siting and statistical method selection rather than optimization of an existing network. These issues will be discussed in Section 4. The Review Team has therefore opted to perform four in-principle optimization demonstrations for the 300-FF-5 OU uranium plume: spatial optimization; temporal optimization using methods tailored to two somewhat different situations; and spatial geostatistical optimization.

The remainder of this subsection discusses the PNNL applications and the four Review Team demonstrations. Section 2.3.6 summarizes and compares the results of these demonstrations.

2.3.1 PNNL Geostatistical Optimization of Plume-Mapping Networks

Perhaps the most frequently used method to optimize a monitoring well network used for on-going mapping of a contaminant plume comprises a combination of geostatistical interpolation of concentration data and minimization of variance (or another measure of variability). PNNL has applied such a method to identify redundant monitoring wells for three plumes at the Central Plateau of the Hanford Site: the carbon tetrachloride (CCl_4) plume (200-ZP-1 Operable Unit); the technetium-99 (^{99}Tc) plume (200-BP-5 Operable Unit); and the tritium (^{3}H) plume (200-PO-1 Operable Unit). A more detailed description of the methodology appears in Appendix D.

The results PNNL achieved by applying this procedure to the CCl_4, ^{99}Tc, and ^{3}H plumes are summarized in Exhibit 2. Although presented together, it should be noted that the three plumes have different monitoring networks, different measures of network quality, different locations around the site, and different "off-limits" monitoring wells. The resulting percentages of wells removable are similar to the range of results reported elsewhere (EWRI [2003] and USEPA [2004]). If all monitoring wells were considered to be possibly redundant (i.e., there were no "off-limits" wells), about 25 to 40 percent of the total number of monitoring wells (including the "off-limits" wells) could be removed for each plume. In the cases where the optimization was constrained such that some monitoring wells were "off-limits" then 15 to 20 percent of the total number of monitoring wells were found redundant. (Additional discussions are given in Michael *et al.* [2000] and Murray *et al.* [personal communication, 2004]).

Exhibit 2. Summary of Geostatistical Network Optimizations by PNNL[1]

Summary Statistic	*^{3}H Plume*	*^{99}Tc Plume*	*CCl_4 Plume*
# of MWs, total, considered	293	59	117
# of MWs on "constrained" list	[2]0	[2]0	58
# of MWs optimization says removable (unconstrained/constrained)	[3]72 / [4]60	[5]25 / --	[6]30 / 17
Percent of MWs removable (unconstrained/constrained) as percent of total # of MWs	[3]25% / [4]20%	[5]42% / --	[6]26% / 15%

[1] Data from Michael *et al*. [2001], DOE [2001], Murray *et al*. [personal communication, 2004], and Murray [personal communication, 2004].
[2] No constrained set was used.
[3] Based on mean absolute error between concentration maps and change in variogram
[4] Based on post-review of unconstrained case against need for regulatory and geographic coverage
[5] Based on ad hoc decision and lack of plateau in RMSE vs # removed plot
[6] Based on plateau in RMSE vs # removed plot and change in variogram

The optimization results were combined with other considerations and folded into proposals for revised monitoring programs at the respective facilities. Regulatory approvals have been received for revised monitoring plans [E. Dresel, personal communication, 2004] and are being implemented. Because the optimization results were combined with other considerations, it is not possible to assign a specific savings to optimization. PNNL has also used a modification of this scheme at the ^{99}Tc plume to identify and rank locations that variability estimates indicate would most benefit from new well installations. One of the recommended wells has been added in the monitoring well installation plan under the Milestone 24 process [E. Dresel, personal communication, 2004].

The success of these applications notwithstanding, PNNL has plans for several improvements in this methodology. Among these are (1) replacing the "greedy" or "local" optimization technique by a global optimization approach, (2) improving geostatistical representation of concentration data to account for asymmetrical shapes induced by groundwater flow, and (3) for mobile plumes, coupling the optimization approach to simulation models so that monitoring well optimization accounts for anticipated plume migration over the useful life of the wells. Discussion of these plans is deferred until Section 4.5 on RCRA-CERCLA coordination for monitoring at LLWMA-3 and Section 5 on regulatory coordination in general.

2.3.2 Demonstrations at the 300-FF-5 OU

The Review Team opted to base its GWMO demonstrations on the data from the uranium (U) plume in the 300-FF-5 OU. Four approaches are shown, two for identifying temporal redundancies and two for spatial redundancies. As emphasized previously, these should be considered to be demonstrations of potentially useful techniques, rather than final solutions to the monitoring problem at the 300-FF-5 OU, particularly in view of the forthcoming FFS. In view of previous comments about the evolutionary state of conceptual modeling of the U plume in this area, only statistical and geostatistical methods were considered for these demonstrations.

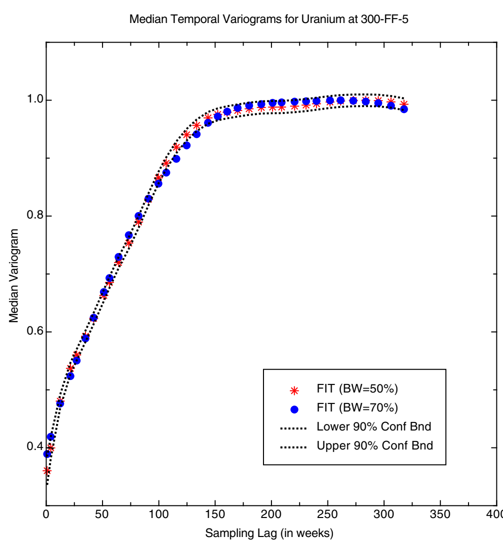
The two temporal redundancy analyses and one of the spatial redundancy analyses were performed using the Geostatistical Temporal/Spatial (GTS) methodology (Cameron and Hunter [2002]); the other spatial redundancy analysis was performed using the Multi-Objective Long Term Monitoring Optimizer (M-LTMO) software (see Appendix F). Appendix D of *Hanford Site Groundwater Monitoring for Fiscal Year 2003* [PNNL, 2004] provided a very thorough and detailed account of the groundwater data quality and precision. This document was very helpful in assessing the current state of the quality assurance and quality control for groundwater data at the Hanford Site. The web site https://www4.hanford.gov/groundwater/ and the stand-alone version of the Hanford Site Groundwater Monitoring Project Data Viewer and Evaluator (DaVE) were also very helpful in accessing the groundwater data needed for the Review Team's evaluations and demonstrations.

2.3.3 Demonstrations Using GTS: Temporal Redundancy

These demonstrations of the GTS methodology were conducted on total uranium data from the 300-FF-5 OU. GTS has two optimization components: a temporal module for optimizing sampling frequency and a spatial module for optimizing well network locations. Both are designed to reduce information redundancy in LTM by lengthening recommended sampling intervals or by removing sampling locations.

To provide flexibility, the temporal module includes two techniques: temporal variograms and iterative thinning. The first of these creates a one-dimensional temporal variogram from historical data at the combined set of wells to determine the minimal average interval at which consecutive measurements become serially uncorrelated. The idea is that the fluctuations in serially correlated data are more-or-less redundant (i.e., predictable from prior data in the series), whereas as the serial correlation decreases the new data points add new information more efficiently. The GTS temporal variogram approach is most useful where historical data may be limited at some or all wells, but an efficient common sampling schedule is desired.

Exhibit 3. Temporal Variogram per GTS for 300-FF-5 OU U Data



The temporal variogram constructed from recent U data from 38 wells at the 300-FF-5 OU (see Appendix F) is given in Exhibit 3. This shows that consecutive sampling events become uncorrelated in approximately three years, on average. This contrasts sharply with the current sampling schedule, which varies from every month-and-a-half at selected RCRA wells to semi-annually or annually at other wells. (The frequency requirements for the RCRA wells are discussed in Section 4.3 and Appendix G).

In iterative thinning a temporal trend, with confidence limits, is fitted at each location. Data are removed at random, and the trend re-fitted, proceeding iteratively until the reduced-data trend no longer falls within the confidence limits. A well-specific optimized sampling interval is then estimated. Iterative thinning requires more data than the temporal variogram. Because the goal of iterative thinning is reconstruction of trends, the results will not always correspond to those of the temporal variogram. Also, each well may receive a different recommended sampling frequency. For U monitoring at the 300-FF-5 OU, the wells amenable to iterative thinning are listed in Table E.2 of Appendix E; the median recommended sampling interval is one year. A side benefit of the trend fitting in iterative thinning is that trend estimates can be mapped across the site. For example, Figure E.4 of Appendix E illustrates that recent trends (since 2000) near the river have been generally downward.

2.3.4 Demonstration Using GTS: Spatial Redundancy

The spatial component of GTS is based on an idea similar to iterative thinning, as follows:

- Construct a base plume map using all available wells;
- Remove selected wells (e.g., a small percentage) and reconstruct the map, determining how much it has changed;
- Stop when the map deteriorates significantly; and
- Propose the reduced well set as the optimal network for the site.

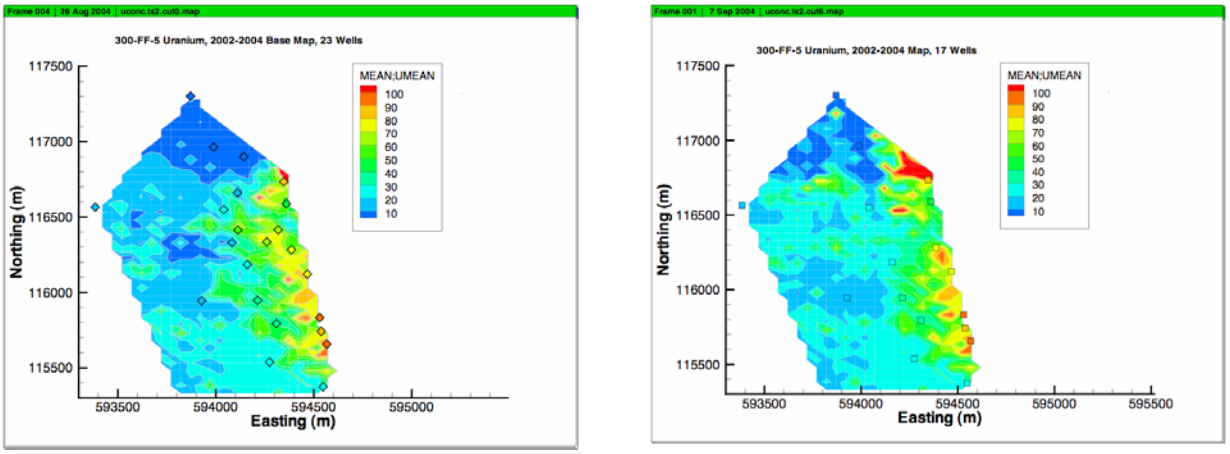
GTS tracks the tradeoff between the cost of the reduced network and the loss of accuracy in subsequent maps. Selected measures of this tradeoff are presented in Appendix E; the key statistical indicators are bias and uncertainty.

One key benefit of the GTS approach is its use of multiple indicator local regression (MILR) to estimate site maps, sidestepping the effort and art involved in developing a spatial variogram for kriging. Often, as at the 300-FF-5 OU, wells are not sampled at the same times. Hence, data from an interval (a 'time slice') are included to ensure enough spatial coverage for estimating the spatial redundancy.

At the 300-FF-5 OU there is the added complexity of the effect on U measurements of seasonal water level fluctuations near the river. There are several practical approaches that might be taken to deal with this factor. In one approach redundancy could be estimated at both low flow and high flow periods. Doing so would require enough data from both flow regimes to achieve appropriate spatial coverage. A second approach was taken in this GTS demonstration, which is to broaden the time slices to provide not only better spatial coverage, but also to include multiple periods of low and high flow in a single time slice, akin to estimating a longer-term average.

For demonstration purposes, two time slices were analyzed, one from 1999-2001 and the other from 2002-2004. Base maps computed for the 2002-2004 time slice using 23 and 17 wells are shown in Exhibit 4. Preliminary analyses suggested that the uranium plume was more intense and widespread during the first time slice than the second; see Figures E.9 and E.10 of Appendix E. To decide on an optimized network, stakeholders must balance cost savings and loss of accuracy. GTS aids in this decision process by providing not only global assessments of the cost-accuracy tradeoff, but also maps of the local bias and uncertainty from which one can determine when specific areas of the site are being over- or under-sampled using the reduced-well network or at what point too many areas of substantial uncertainty have appeared.

Exhibit 4. 300-FF-5 OU Base Maps for U (2002-2004) per GTS, Using 23 Wells (left) and 17 Wells (right)

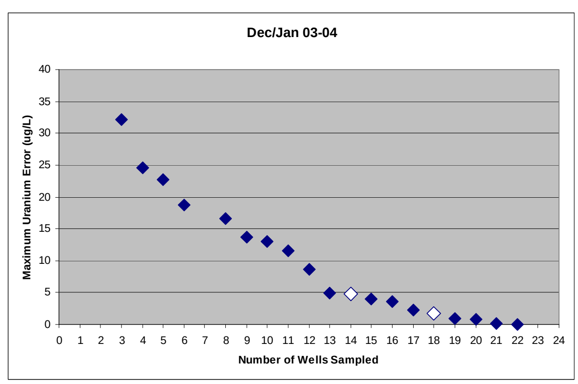


2.3.5 Demonstration Using M-LTMO: Spatial Redundancy

The use of multi-objective optimization to identify spatial redundancy at the 300-FF-5 OU is demonstrated in this section. Software called M-LTMO (Multi-objective Long Term Monitoring Optimizer), developed at the University of Illinois and Moiré Inc., was used. The approach and findings are summarized here; more details can be found in Appendix F.

Multi-objective optimization is a type of mathematical optimization in which solutions are sought to a problem that has conflicting objectives. For the 300-FF-5 OU, the optimization seeks to identify the optimal set of wells in the 300-FF-5 OU that should be sampled in each monitoring period, given the desired monitoring objectives. For the purposes of this demonstration, the goal is to identify redundant sampling within an existing monitoring network. The objectives are assumed to be to minimize the number of wells sampled and to minimize the maximum error in interpolated uranium concentrations caused by eliminating monitoring wells from the network.

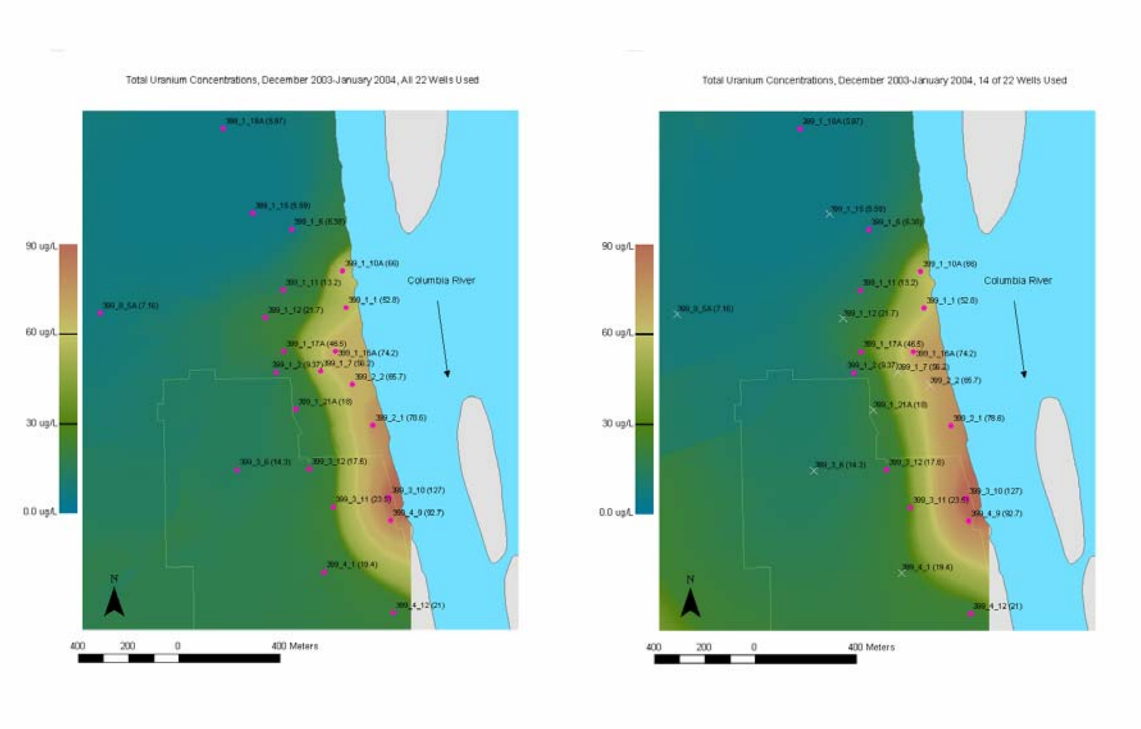
Exhibit 5. Optimal Tradeoffs For 300-FF-5 OU U Monitoring During Low-Flow Periods, per M-LTMO



A spatial interpolation model was created that estimates contaminant concentrations at all locations in the 300-FF-5 OU using quantile kriging with data from two recent monitoring periods. The periods, June-July 2003 and December 2003-January 2004, were selected because they have significant amounts of recent data available and they appear to be representative of high- and low-flow periods, respectively, in the Columbia River. The optimization algorithm automatically identifies the minimum number of wells that can be sampled for a range of interpolation errors (i.e., the optimal tradeoffs between the objectives). The interpolation model is used with a reduced dataset to estimate the errors caused by deleting wells from the monitored network.

An example of the optimal tradeoffs at the 300-FF-5 OU is given in Exhibit 5 for the low-flow period. Each diamond in the figures represents the optimal monitoring design that minimizes errors for a given level of sampling. Compared with the drinking water standard (DWS) for uranium of 30 µg/L, it is clear that relatively low errors are possible with fewer wells. Exhibit 6 compares the interpolated maps from the optimal 14-well sampling plan, one of two plans highlighted as white diamonds in Exhibit 5, with the map created using data from all 22 wells sampled in the low-flow period. This Exhibit shows that removal of eight redundant wells has little significant effect on the plume map. Similar results were found for the high-flow period; see Appendix F. Only three wells were found to be redundant in both periods.

Exhibit 6. Interpolated U Maps for 300-FF-5 OU Using 22 Wells (left) and 14 Wells (right), per M-LTMO. Crosses represent redundant wells.



These results suggest that the 300-FF-5 OU has significant spatial redundancy with regard to these objectives. The wells identified as redundant in the low- and high-flow periods differ significantly, indicating that the temporal variability in the 300-FF-5 OU is an important factor that should be considered in future spatial redundancy optimization efforts in this area. Spatial redundancy could be optimized for multiple sampling periods simultaneously, identifying a configuration of wells that would be nearly optimal in both high- and low-flow periods. Alternatively, both optimal spatial locations and temporal frequencies could be identified simultaneously.

Another factor that should be considered is the uncertainty inherent in the uranium plume maps. This analysis has assumed that the map obtained by interpolating with all wells is the "ground truth" and that any sampling plan that can reproduce essentially the same map is sufficiently accurate. However, any interpolated map, particularly one based on 22 samples, can have significant error. M-LTMO can be configured to consider uncertainty, as well as temporal sampling, but such analyses are beyond the scope of this demonstration.

2.3.6 Summary of Demonstration Approaches and Results

The approaches and results of the various demonstrations are summarized in Exhibit 7. Again, the Review Team results are for a hypothetical situation assuming that the 300-FF-5 OU data are from a quasi-steady-state long-term monitoring program, which is not that case in fact. Therefore these results are suggestions of the kinds of information obtainable through optimization, not specific recommendations for Hanford.

Exhibit 7. Summary of Demonstration Approaches and Results

By	Data	Spatial or Temporal	Method	Result and comments
PNNL	^{3}H, ^{99}Tc, CCl_4 plumes in Central Plateau	Spatial	Geostatistical with evaluation of prediction variance	15-26%, 20-25%, 42% of wells can be removed for the different plumes
Review Team (GTS)	300-FF-5 U plume	Temporal	Geostatistical with global temporal variogram	On average, serial correlation vanishes in three years; useful for setting common frequency for all wells
Review Team (GTS)	300-FF-5 U plume	Temporal	Iterative thinning with trend reconstruction	Well-specific optimal frequency ranges from quarterly to every year-and-a-half; median recommendation is annual
Review Team (GTS)	300-FF-5 U plume	Spatial	Spatial using local quadratic regression and evaluation of plume map reconstruction	Little information loss with up to 12% reduction in number of wells; thereafter must evaluate tradeoff
Review Team (M-LTMO)	300-FF-5 U plume	Spatial	Geostatistical with evaluation of prediction variance	Little change in plume maps with 35% reduction in wells, considering low- and high-flow conditions separately, but specific wells retained are different. Must evaluate trade-off; no clear jumps in information vs cost curve

3.0 ADDITIONAL REVIEW TEAM COMMENTS ON 300-FF-5 OU ISSUES

3.1 Current Remedy Status of the 300-FF-5 OU

The 300-FF-5 OU comprises groundwater that has been impacted by activities at the 300 Area, including buildings, disposal trenches and ponds, and other facilities, and the so-called 300 North Area, including the 618-10 and 618-11 burial grounds and the 316-4 cribs. It is regulated under CERCLA with an embedded RCRA permitted facility at the 316-5 process trenches.

Among several COCs, uranium is the primary driver, with current concentrations above the 30 µg/L DWS. Other radionuclides, such as strontium-90, tritium, iodine-129, and technetium-99, either have been observed as limited occurrences or migrate into the 300 Area from 200 East Area sources. Tritium observed in the 300-FF-5 OU is part of a site-wide tritium plume originating in the 200 East Area. Tritium concentrations are about 20,000 pCi/L in the 300-FF-5 OU North Area and are lower to the south. A localized tritium plume emanating from the 618-11 burial ground exceeded the DOE derived concentration guide (DCG) of 2,000,000 pCi/L. Levels of the other radionuclides are below their DWS's.

Chlorinated hydrocarbons have been detected at the 300-FF-5 OU, but during fiscal year 2003 only TCE and cis-1,2-DCE were found above their DWS's. TCE above its DWS was found in two deeper wells near the 316-5 process trenches and the 300 North Area facilities, and also in the upper part of the unconfined aquifer due to migration into the 300 Area from offsite sources. Cis-1,2-DCE was found above its DWS only at two deeper wells in the unconfined aquifer downstream of the 316-5 process trenches. Nitrate concentrations exceeding background levels have been found across the 300-FF-5 OU. The sources of the nitrate appear to be offsite industry and agriculture for the southwest part of the 300-FF-5 OU and the 200 East Area for nitrate in the 300-FF-5 North Area. The remainder of this section focuses on the portion of the 300-FF-5 OU located near the southern 300 Area buildings, ponds, and trenches.

The current interim remedy for the 300-FF-5 OU is MNA. Implemented in 1994, the MNA remedy has not performed as expected, and the final remediation remedy is yet to be determined. The initial selection of MNA may have been misguided by the fact that observed decreases in the uranium concentrations, which were observed after the surface sediments were removed from disposal ponds, trenches, and other appurtenances and which were used to estimate attenuation rates, were likely to have been related to dilution of the plume by uranium-free water that continued to be discharged into the trenches. It was noted during the 5-year review in 2001 that the U concentrations rebounded once the discharges ended.

USEPA expects to receive a draft feasibility study for the 300-FF-5 OU, at which point a final remedy may or may not be selected. Until the selection process is completed and the final remedy is operational, the 300-FF-5 OU is, by definition, not in the long-term monitoring phase. Once the final remedy is in place, then optimization of the long-term monitoring would be recommended.

3.2 Conceptual Modeling of the 300-FF-5 OU

The conceptual model for the fate and transport of uranium in the 300-FF-5 OU is undergoing re-evaluation. The Phase I Remedial Investigation [USDOE, 1994] included a hypothetical conceptual model of dissolved and sorbed (linear K_d) uranium; also, total uranium was modeled as a conceptualized dissolving flocculant. The mass of this flocculant was estimated to be large enough to provide a constant source concentration for about ten years. Modeling of total uranium proceeded based upon complete dissolution of the flocculant. Total uranium was found to move completely from the unconfined aquifer and into the Columbia River in an uncertain period of time, anywhere from a couple of years to decades following flocculant depletion, depending on the actual K_d. The Phase I RI also noted that at that time "insufficient data currently exists to more accurately predict future uranium concentrations in the year 2018 and beyond" [pg 345 of 532 in the electronic image of Part 1]. The soil concentration of 7.4 pC/g was assumed to be the concentration of flocculant that would dissolve. Also [pg 333 of 532 in the electronic image of Part 1], a source height of three meters "was chosen because it corresponds roughly to the *assumed* thickness of contaminated groundwater." An analytical model and a numerical model (PORFLO3) were applied in the Phase I RI; see the description in Appendix D of [USDOE, 1994].

A contemporaneous study [PNNL, 1994] indicated that for a soil sample used "to represent uncontaminated sediments below the 300 Area North Process Pond, ...[a]fter the evaporite salts [of uranium] flush from the columns, it appears that Cr and U present in 300 Area North Process Pond sediment leachates does not adsorb significantly" (page 6.2). This may be a consequence of large amounts of evaporites interfering with the test procedures. It goes on "we recommend that any future groundwater impact analyses be performed using the simulated rainwater leachate data to estimate solution concentrations that would percolate through vadose zone sediments."

The subsequent RI/FS [USDOE, 1995] rejected the solubility-controlled (or U flocculant) hypothesis (see Section 4.3.2, pages 4-17 and 4-18, of the RI/FS), arguing that U concentrations were decreasing at monitored wells while they should remain constant if there were a solubility-limited concentration. (This interpretation appears to be incorrect, because the monitoring wells are in an open-flow system and concentrations were diluted by continued discharge of uranium-free water). The three- to ten- year estimate given in the RI/FS for MNA to accomplish cleanup concentrations used a simplified spreadsheet model from the Phase I RI, but with "refined" coefficients. Total uranium was assumed to adsorb/desorb using a linear sorption model with a single K_d value. The main features of the conceptual model used in this study were (1) steady, essentially one-dimensional groundwater flow due to regional flow, (2) linear desorption of U from the saturated zone, (3) no additional sources of U, and (4) well-characterized initial mass of U in the saturated zone. Several problems have been identified in this conceptual model: (1) the initial mass estimates, (2) the assumption that desorption is the only continuing source of U, (3) the absence of interactions between the aquifer and river (other than the river receiving groundwater discharge), (4) the assumed uranium contamination and plume in upper three meters of the saturated zone, and (5) the usual complications of heterogeneity.

The conceptual model that was discussed during the Review Team's site visit adds (1) time-varying river-water migration into the aquifer, (2) the need for a more complicated sorption model, per S. Yabusaki's presentation, (3) vadose zone impacts, at least in the portion of the vadose zone immediately above the water table that is sometimes saturated through river-water incursions, (4) a more sophisticated dissolution model for the vadose zone based on both the fluctuating water table and infiltration, and (5) clarification of the mass and flux in Hanford versus Ringold materials.

PNNL is constructing a detailed three-dimensional, transient numerical model of uranium transport in the 300-FF-5 OU that will only be partially completed by March 2005, when the draft FFS is due. Features are anticipated to include non-equilibrium mass transfer, more detailed simulation of river water coming into the system, and STOMP simulations of the vadose zone including spatially variable recharge. It is expected that a one-dimensional preliminary model will be completed in time for the FFS.

Due to the time required to complete the numerical modeling, the Review Team recommends concurrently calculating a uranium mass inventory for the 300 Area using existing data, such as production concentrations, soil data, groundwater data, and K_d values, while the numerical modeling continues. Depending on the results of the estimates and their uncertainties, it may be possible to make decisions as to whether more data are necessary to produce a more accurate estimate. A mass inventory result used in conjunction with a mass flux value (mass leaving the site via the Columbia River), would allow for a quick, rough estimate of the number of years required for the complete removal of uranium from the site. This estimate may provide input allowing the remedy selection process to proceed. This estimate may even show that MNA is possible and the appropriate remedy for the site. Additional short-term suggestions are given in the following section.

3.3 Review Team Recommendations

3.3.1 Parallel Engineering Analysis to Meet FFS Needs

Modeling performed in 1995 in support of the interim MNA remedy had many simplifications that, with the benefit of hindsight, appear not to be justified, as discussed previously. The Review Ream is very supportive of the geochemical analysis and modeling work currently being performed by PNNL to better understand and simulate the complex environmental fate and transport of uranium in the 300-FF-5 OU. The modeling program presented to the Review Team is very ambitious given the time frame of the FFS, however. During the site visit, USEPA indicated its expectation that by March 2005 they would be presented with strong arguments to support a specific alternative for the future (e.g., MNA, justification for technical infeasibility (TI), or other viable active remedial alternatives). Based on information provided during the site visit, the modeling being performed by PNNL will be only partially finished, in the form of 1-D simulations, by March 2005. The Review Team is concerned that this modeling, while ultimately beneficial, may not be completed early enough to address the engineering issues associated with the FFS remedial evaluations.

Furthermore, the Review Team notes that significant uncertainties will be associated with the resulting predictions. Given the complex nature of the model being developed, there will be many sources of uncertainty, including aquifer parameters, source terms, reactive transport parameters, and surface water interactions. During the site visit, PNNL suggested that a primary use of the model will be to determine if concentrations of uranium in groundwater will decline through natural attenuation mechanisms over a "reasonable" period of time, so an evaluation can be made of whether MNA is feasible relative to TI or other active remedial actions. The Review Team is concerned that the model-estimated time frame for adequate natural attenuation may ultimately be reported with such a large range, due to uncertainties, that the underlying question about the efficacy of MNA may not be adequately answered.

To address these concerns, the Review Team recommends that a separate, simpler, directed engineering analysis be performed in parallel to the more detailed geochemical analysis and modeling work being performed by PNNL. This parallel analysis would address a number of questions, including the following:

- What is the amount of (potentially) mobile U in the vadose and saturated zones? This involves calculations of residual uranium mass in the vadose zone, based on previously collected soil boring data (best estimate, high estimate, low estimate). For example, PNNL [1994] found uranium concentrations of 26 to 46.2 ppm (µg/g) in soils samples from outside, yet near, the North Process Trench. Preliminary results from the ongoing geochemical study are on the same order of magnitude.

- What is the vertical distribution of U, as well as the horizontal distribution, and how do they translate into dissolved U mass flux? This involves calculations of mass flux rate of uranium from the aquifer to the river (best estimate, high estimate, low estimate). The 1994 Phase I RI [USDOE, 1994] assumed that uranium contaminants were present in a source of height of 3 m, which "was chosen because it corresponds roughly to the ***assumed*** thickness of contaminated groundwater" [emphasis added].

- What are the rate and the rate-limiting process for transferring U from the vadose zone to the groundwater? This involves calculations of mass flux rate of uranium from the vadose zone to the aquifer (best estimate, high estimate, low estimate). The RI modeling assumed that U in flocculants was present at 7.4 pCi/g in the soils (approximately 10.5 µg/g) and would be completely removed by dissolution in about 10 years based on a solubility limit of 270 µg/L. If the original concentration were larger, compared with that identified in the PNNL [1994] report, or the solubility limit were smaller, then the time to dissolution would be larger. On the other hand, if the solubility were larger (for example, Table I-10 in Appendix I of the Phase I RI indicates the solubility for total uranium is 32,500 pCi/L, i.e., about 46,400 µg/L), then the time to dissolution would be smaller. Will the solubility change as flocculants are weathered? How is solubility affected by carbonate concentration in the leaching water?

- What is the groundwater cleanup time under various scenarios? This involves mass balance estimates for uranium over time, based on the above calculations and estimates.

These calculations would be based on simplifications of the actual system, and would be based on previously collected data, such as soil borings conducted during the RI, and/or previously conducted scientific studies, such as work being performed on sorption K_d values for radionuclides reported by Ken Krupka of PNNL [Beth Moore, personal communication, 2004].

Data gaps should be identified, and those gaps that are important and that can be quickly resolved should be placed on a prioritized list. As a hypothetical example, if the vertical distribution of U is not known adequately, a number of GeoProbe pushes may be advanced at a small number of locations parallel to the riverbank though the vadose zone, Hanford formation, and the conductive Ringold formation. Vertical profiling at a small number of sites can test the hypothesis that U is essentially limited to the Hanford formation, and can also form the basis for mass flux estimates to the Columbia River.

The Review Team and others present during the site visit agreed that there are significant difficulties associated with up-scaling results from laboratory studies to field scale applications. The Review Team therefore discussed the possibility that a medium-scale leaching experiment could potentially be conducted to better estimate mass flux rate from the vadose zone to the aquifer, to improve estimates the amount of potentially mobile U in the vadose zone, and to identify key mass transfer processes. Materials would be obtained and used at a scale that accounts for the variation of aquifer materials (such as a 2x2x2-meter cube of vadose zone material collected near the water table, Columbia River water, and undiluted groundwater). A previous soil leaching study [PNNL, 1994] focused on soil samples obtained from within former disposal ponds, which were subsequently removed from the ponds. That study also included one "nearby" sample that appears more representative of what would have remained after removal. Initial U concentrations of 26 to 46 ppm were reported for this soil sample; these concentrations are consistent with the carbonate extractable concentrations in the current work by PNNL as reported by Yabusaki during the Review Team's site visit. The laboratory-scale leaching study suggested that there were significant solubility issues, and eluted concentrations barely decrease after flushing with 800 pore volumes of simulated acid rain (see Figure 4.1 of PNNL [2004]). Unknown thousands of leach cycles would be required to obtain effluent concentrations of 21 pCi/L (which, using the conversion factor of 0.7 pCi/μg given in Table 4-4 of the Phase I RI, corresponds to 30 μg/L, the water quality standard for U). The tested soils may not, however, be representative of currently existing conditions. A fast-tracked, updated, and larger-scale test may have significant value for the engineering studies.

The calculations would include variants to the simplifications, to provide the "best", "high", and "low" estimates suggested above. By performing each of the above calculations as a range (best estimate, high estimate, low estimate), a matrix of uranium mass balance possibilities can be evaluated, and this will presumably represent the range of possibilities to ultimately be assessed with the more detailed modeling by PNNL (i.e., these will be "bounding calculations"). An advantage of this parallel approach is that it can be performed within months, so that results will be available for the draft FFS analysis due in March 2005. The results should indicate one of following:

- If the results indicate that uranium impacts will meaningfully dissipate due to natural flushing in a time frame that is acceptable to the regulators for most/all scenarios, then the FFS analysis can include an emphasis on MNA.

- If the results indicate that uranium impacts will not meaningfully dissipate due to natural flushing in a time frame that is acceptable to the regulators for most/all scenarios, the FFS analysis should emphasize alternatives other than MNA.

- If the results indicate that uranium impacts may meaningfully dissipate due to natural flushing, but that large uncertainties in the time frame result from the bounding calculations, then the FFS analysis should provide information as to the specific uncertainties that cause an unacceptably large uncertainty in the time estimates. This discussion should provide a basis for identifying specific characterization activities that would reduce the range of uncertainty for those specific parameters and a preliminary implementation plan.

If either of the first two cases listed above occurs, it is possible that FFS decisions could be made on the basis of this simplified analysis based on existing data, without needing to wait for the results of the more detailed modeling analysis.

The Review Team believes that process of performing these bounding calculations, and deciding on the types of simplifications and associated bounding values, will also be of benefit to the team conducting the more detailed geochemical analysis and modeling. Therefore, if the parallel engineering analysis is conducted, it is also recommended that communication between the engineering and geochemistry teams be open and frequent to ensure that best available site-specific science is being used and that the engineering needs motivate the research directions selected.

3.3.2 Current Data Needs Related to Possible FFS Options

Since the upcoming FFS-related work is taking on larger proportions, the Review Team considered whether any additional monitoring needs should be considered. Currently 300-FF-5 OU samples are being analyzed for the constituents necessary to support the interim MNA remedy as well as those required for RCRA compliance. Draft USEPA guidance on MNA needs for radionuclides is reported to require five-year data histories for geochemical parameters of site-specific importance. A preliminary list of parameters addressing those guidelines should be included with any consideration of MNA in the upcoming FFS. The parameters should include, but not be limited to, those identified as significant in the geochemical analysis and modeling work. Moreover, synoptic groundwater head and river stage observations should be collected regularly.

Uranium isotopic composition of the vadose sediments and groundwater may be a fruitful area of investigation, particularly examining uranium-236, a by-product of the reprocessing of uranium ore and fuel rods. This type of isotopic information would support the differentiation of the waste sources and provide information on contributions from natural upgradient water and mineral weathering. In addition, the ratios among ^{234}U, ^{235}U, and ^{238}U would provide information on the presence of enriched and depleted uranium, both of which were processed in the 300 Area.

Because an interim remedy appears to be an attractive result from the March 2005 FFS, sufficient monitoring should be maintained to support contingencies that might be required. For example, if a permeable reactive barrier (PRB) were selected for a treatability study, the monitoring schedule should not be allowed to lose focus on behavior of the plume even as attention would be placed on the PRB; plume extent and in-plume concentrations should continue to be monitored.

Contaminant flux monitoring would be a useful addition to the current sampling program at the 300-FF-5 OU. *In situ* flux meters provide location-specific mass transport and groundwater flow data. Field flux data is extremely useful regardless of the chosen cleanup strategy. As a tool for long-term monitoring, rows of flux meters installed downgradient from a contaminant source and perpendicular to groundwater flow can be used to characterize source attenuation. The spatial distribution of measured fluxes can be integrated across the transect of wells to produce estimates of monthly and annual contaminant mass loadings through groundwater.

Although flux monitoring is less mature than head or concentration monitoring, several technologies are currently available or in development. Seepage meters and temperature anomalies are well known approaches. The following are some newer technologies:

- Benthic Flux Sampling Device**.** The Benthic Flux Sampling Device (BFSD) has been tested and certified by the USEPA to measure the contaminant flux in coastal marine sediments affected by tides. The changing water levels in the Columbia River present a similar issue for measuring fluxes across the groundwater-surface water interface. The BFSD is an automated, *in situ* water sampling device designed to collect data for quantifying the flux of trace metals, including arsenic, cadmium, copper, nickel, lead, and zinc. It collects and filters discrete water samples periodically over a deployment of up to four days, which are then preserved and delivered to an analytical laboratory for analysis. The technology has been used at several DoD sites. Additional information can be found at http://www.calepa.ca.gov/calcert/CertifiedTech/BenthFlu.

- UFL Passive Flux Meter**.** The University of Florida passive flux meter (PFM) is a self-contained permeable device inserted into a well or boring that acts as an integrating sampler. Groundwater flows through it and is not retained. The interior of the device is a matrix of hydrophobic and hydrophilic permeable sorbents that retain dissolved organic and inorganic contaminants present in fluid intercepted by the unit. The sorbent matrix is also impregnated with known amounts of water-soluble resident tracers (typically benzoate), which are leached from the sorbent at rates proportional to the fluid flux. Following exposure to groundwater flow for a period ranging from days to months, the PFM is removed from the monitoring well and the sorbent extracted to quantify the masses of contaminants intercepted and the residual masses of resident tracers. The contaminant masses are used to calculate time-averaged contaminant mass fluxes, and residual resident tracer masses are used to calculate time-averaged groundwater flux. Suitability to the alternating groundwater flow directions encountered near the Columbia River has not been established. This device is being validated through ESTCP demonstrations at three DoD sites; see http://www.estcp.org/projects/cleanup/200114o.cfm.

- Immission Pumping Test. The University of Tübingen has developed an integral method for estimating mass fluxes using a transect of wells. The immission pumping test (IPT) has two stages. In the first stage, each monitoring well in the transect is pumped at a constant rate and the time-varying concentration of the contaminant of interest is recorded. In the second stage, the concentration versus time data for all of the wells are combined using a numerical inversion technique to estimate the mass flux of contaminant across the transect. Unlike the PFM, it also allows the maximum and average concentrations over the transect to be estimated. Being an active method, the IPT can overcome localized issues (e.g., borehole damage) that interfere with passive methods and can obtain measurements over a length scale greater than the well diameter. On the other hand, one must dispose of the water pumped from the wells and multiple chemical analyses are required. This method has been tested at several European sites. Results have been published since 1998; see, e.g., the article by Ptak in http://www.image-train.net/products/proceedings_sec/Session2_Integrated_Solutions.pdf.

Finally, a tremendous amount of groundwater data is available for general access and review, as mentioned in Section 2.5.2. By contrast, a large amount of soil data has been gathered during environmental investigations of the Hanford Site, but most of these data are not available electronically. For the entire site, approximately 10,250 soil results are not electronically available. These data could be extremely useful in assessing the initial conditions of a site and allow for the calculation of contaminant inventories, for example. Making the data available electronically is a virtual prerequisite to assessing their usefulness.

4.0 REVIEW TEAM COMMENTS ON GROUNDWATER MONITORING AT LLWMA-3

The other unit that the Review Team was explicitly asked to review is the Low-Level Waste Management Area 3 (LLWMA-3) in the 200 West Area. Groundwater monitoring at this unit is regulated under RCRA; the unit is in detection monitoring status. It is currently being regulated under the provisions of RCRA interim status (40 CFR Part 265) rather than permitted (final) status (40 CFR Part 264), although PNNL and DOE/RL have proposed modifications that are much more nearly in line with the Part 264 monitoring requirements.

LLWMA-3 is located in the northern portion of the 200-West Area, and is approximately 79 hectares in area. Wastes have been received within LLWMA-3 since 1970, and there are several active trenches still receiving wastes. Wastes placed in LLWMA-3 include ion exchange resins, failed equipment, tanks, pumps, ovens, agitators, heaters, hoods, jumpers, vehicles, rags, paper, rubber gloves, disposable supplies, and broken tools. Many other facilities are located near LLWMA-3; some have caused groundwater contamination extending to and under LLWMA-3. Contaminants from other sources include nitrate and carbon tetrachloride. These other impacts present a complication with respect to the RCRA detection monitoring being performed.

LLWMA-3 sits in the Hanford formation, which is underlain by the Plio-Pleistocene, the Ringold Formation, and basalts, in order of increasing depth. The water table of the uppermost aquifer under LLWMA-3 is located within Hydrogeologic Unit 5 (Ringold Unit E). The vadose zone is approximately 75 meters thick and the base of the uppermost aquifer is generally the Ringold lower mud unit; at LLWMA-3 the lower Ringold mud is present under the southern portion of the facility, but is absent to the north.

4.1 Primary Issues

There are three major issues that lead to inefficiencies in groundwater monitoring at and around LLWMA-3. These include:

- The statistical requirements of 40 CFR Part 265, which date from 1980, are primitive by comparison with those of the revised Part 264, promulgated in 1988. Those, in turn, have been improved substantially in the subsequent Part 258 Solid Waste rules; in addition, subsequent guidance, particularly USEPA [1992], shows further evolution of statistical concepts and implementation over the discussions in the *Federal Register* accompanying the various regulatory pronouncements.

- Wells have been sited in this area using the simplistic MEMO model, which dictates a high density of downgradient wells located at the compliance boundary, in spite of the regulatory requirement for upgradient wells, the changing groundwater flow conditions, and the complexities of flow through the vadose zone in this area.

- There are overlapping monitoring requirements in this part of the Hanford Site resulting from the presence of both RCRA-regulated units, each with its own requirements, in the midst of the overall CERCLA monitoring of contaminant plumes and remedial activities. Although there have been efforts to achieve coordination among the programs, inefficiencies remain.

4.2 Exacerbating Factors

These regulatory realities combine with several other factors to create costly inefficiencies in groundwater monitoring. The direction of groundwater flow under LLWMA-3 has changed significantly during the past two decades and is expected to change further in the future:

- Groundwater flow direction was historically west-to-east.
- Groundwater flow direction was east-to-west at the time RCRA monitoring began in 1991 (when the B-1 water disposal ponds were still in operation).
- Groundwater flow direction was south-to-north until water disposal in the U-1 ponds was terminated.
- Groundwater flow direction has been shifting toward the east as the mound from previous operations has dissipated.
- Current groundwater flow direction is estimated to be about 70° east of north.

More than flow directions has changed, however. Groundwater water levels have dropped significantly (~20 meters) since the B-1 and U-1 ponds stopped receiving water, and an additional 20 to 30 feet of decline are expected. The dropping water levels have caused most of the monitoring wells at LLWMA-3 (twelve out of eighteen) to go dry; the remaining four wells at the water table are expected to go dry in the next several years.

Future groundwater-related activities, such as possibly more intense pumping at the nearby 200-ZP-1 CCl_4 plume, may cause further changes in flow directions and heads, but these future activities are uncertain. Some analyses have reportedly been performed for modest increases in pumping at the CCl_4 plume pump-and-treat system, but the impacts of the full range of possible scenarios being discussed by stakeholders have not been evaluated.

The high cost of installing new wells is a significant issue with respect to siting, due to the depth to groundwater, drilling difficulties, and health and safety issues. DOE/RL and PNNL provided estimates that the cost of installing a new monitoring well is approximately $1,000 per foot, which translates to approximately $250,000 to $300,000 per water table monitoring well.

4.3 RCRA Groundwater Monitoring Statistical Methods

Groundwater monitoring is mandated at facilities regulated under RCRA. At the Hanford Site there are 15 units in RCRA detection monitoring status, seven in assessment monitoring, and two in corrective action [PNNL, 2004]. The monitoring requirements differ considerably, depending on which RCRA program (Part 265 interim status or Part 264 final status) governs the particular regulated unit, as discussed in Appendix G.

PNNL and DOE/RL have proposed the use of intrawell comparisons using Combined Shewhart-CUSUM Control Charts (CCCs) for RCRA detection monitoring at both permitted (final) status units, clearly within the regulation and guidance, and interim status units, clearly deviating from the regulation. As discussed in Appendix G, the Review Team concurs that the use of intrawell tests is warranted for these regulated units; there will be challenges in selecting appropriate indicator parameters for detecting releases from these units effectively, undistracted by the other influences on the groundwater chemistry. One must of be convinced that the regulated unit belongs in detection monitoring, of course; otherwise, at least under Federal regulations for assessment monitoring or corrective action, the appropriate mode of monitoring should rely solely on comparing monitoring data with groundwater protection standards.

One problem with the evolving detection monitoring statistical regulation and guidance was the failure to take into account the (sometimes quite large) number of statistical comparisons made during a single monitoring event, with its resulting multiple exposure to the statistical false positives. USEPA [1992] explicitly allows for Facility-Wide False Positive Rate (FWFPR) control in detection monitoring statistical programs; in doing so, however, it establishes statistical power requirements for alternative tests. Carrying out these statistical power comparisons can be subtle, particularly in the case of the CCCs, as shown by Davis [1999].

Appendix G discusses these subtleties and presents a case study based on the 300 Area Process Trenches, one of the Hanford RCRA units discussed in the PNNL and DOE/RL proposals. The Appendix G case study concludes that the statistical performance of CCC procedures recommended by Jandhyala and Zhang [1999] are far different from those recommended by PNNL [Chou, 2004], and that neither should be considered acceptable. The version recommended by Jandhyala and Zhang (with no resamples) provides inadequate false positive protection, whereas that recommended by PNNL (with both resamples and automatic updating of background data) provides inadequate sensitivity to releases. The acceptable CCC procedures of those evaluated in the case study in Appendix G do allow for resamples before declaring a statistically significant increase, but do not allow for automatic updating of background. It turns out that the statistical performance of the CCC procedures in the case study is relatively less sensitive to the choices of control limits than to the use or non-use of resamples and automatic updating. Appendix G includes further discussion of the updating issue.

The CCCs proposed by PNNL include two somewhat non-standard features: detrending the data; and adjusting for apparent dilution of groundwater during high river level periods in wells near the river. These are discussed in Appendix G as well. An alternate approach to detrending, if such is really appropriate, is suggested for consideration. This alternate approach is based on intrawell prediction limits that incorporate both historical data and a rudimentary conceptual model for the anticipated return of groundwater chemistry (for the selected indicator parameters) to pre-operational conditions. Concerning the dilution adjustment, it seems that using such an adjustment should reduce the “noise” in the data, which should in turn make the resulting monitoring program more sensitive to real changes. If, in addition to comparing the adjusted data to adjusted background data, one wishes to compare them with a groundwater protection standard, a little more work to verify the mechanics of the adjustment might be done, if it hasn't been done already.

One additional issue comes up in considering the RCRA detection monitoring statistical rules: the issue of the required frequency of observations. In this matter USEPA regulation and guidance has shown a continuing evolution. This evolution is reviewed in Appendix G, and suggestions are provided for approaching this problem. The fundamental problem is that the specific monitoring frequency requirements in both the Part 265 and Part 264 regulations were written with statistical tests different from the highly desirable intrawell comparisons in mind, and one should argue for the flexibility to use sample sizes and sampling frequencies that are suited to the intrawell statistical tests and the groundwater flow conditions actually encountered. The advances in this regard seen in the most recent RCRA regulation (the 1991 Part 258 rules for Solid Waste Management Facilities) can be called upon in discussions and negotiations.

4.4 Well Siting for the LLWMA-3 RCRA Unit

4.4.1 Upgradient Background Monitoring

RCRA interim status monitoring currently requires a minimum of one background (upgradient) and three sentinel (downgradient) wells. Even if this unit will eventually be regulated under the final status rules and detection monitoring will use intrawell comparisons, there is good reason to retain an adequate number of upgradient background wells, if for no other reason than to be able to evaluate (a) the chemistry of groundwater coming into the monitoring network and (b) any possible systematic changes in the performance of the measurement system. The loss of monitoring wells that have gone dry and changes in the water table have resulted in insufficient background monitoring even for the RCRA interim status rules. This situation could be further compromised in the future, as it is anticipated that forthcoming RCRA guidance will recommend at least two background wells [Cameron, personal communication, 2004].

In its plans for new wells at LLWMA-3, DOE/RL has proposed thirteen new monitoring wells [Dresel, personal communication, 2004], eight of which have been vetted and appear on the three-year well installation schedule (signed August 3, 2004) established under the Milestone-24-57 process. These eight prioritized monitoring wells are generally located on the eastern side of LLWMA-3, which is currently the downgradient side of the facility; these wells would be placed just outside the waste management line. The status of the remaining five proposed wells is unresolved. Moreover, Ecology has been encouraging the installation of many more monitoring wells for this unit.

The Review Team recommends that the locations of eight already-prioritized new wells be revised to include at least two wells that would be upgradient based on current flow conditions, so that minimal compliance with the RCRA interim status rules will be accomplished in a timely manner. As new wells are added, high priority should be given to maintaining at least one or two wells on each side of the facility, so that at least one upgradient well will be present regardless of any changes in groundwater flow direction over the anticipated 25-year useful lifespan of the new wells.

4.4.2 Compliance Well Locations

As noted in the previous subsection, the eight proposed well locations reflect a considerable emphasis on installing wells on one side of the facility, based on the current groundwater flow direction. The Review Team recommends that this be avoided. Significant and once-unanticipated changes in groundwater head and flow direction, approaching 180°, have occurred during the past two decades. As a result, history argues against excessive bias in well locations. Given this and given that there may be significant changes in groundwater levels and flows resulting from increases in nearby pump-and-treat operations within the useful life of new wells, the Review Team suggests that new well locations be distributed around the facility and that placing them only on one side is not appropriate.

One of the tools used in siting new wells for RCRA interim status monitoring has been the Monitoring Efficiency Model (MEMO), developed by Golder Associates in the late 1980's. MEMO, whose use has been accepted by Ecology, is computer software that uses a highly simplified model of subsurface conditions to determine "monitoring efficiencies" of alternative monitoring networks. Among the assumptions made by MEMO are the following:

- Fixed, uniform, and steady groundwater flow direction; and
- Homogeneous aquifer materials.

Input parameters to MEMO include:

- Longitudinal and transverse dispersivities; and
- The size of the "patch" representing the size of the groundwater source for the MEMO analysis, which is the result of vertical transport through the vadose zone to the water table of wastes released from a 55-gallon drum.

MEMO is not required by the RCRA interim status monitoring regulation itself, but the use of a geohydrologically-based tool, such as MEMO, can be helpful in siting new monitoring wells. It is important, however, that the conceptual model used by the tool and the site release and vadose zone conceptual models be compatible and that coefficients be appropriately selected. Reportedly there is some disagreement between PNNL and Ecology over key parameter values.

The Review Team believes that the results of the MEMO model as currently applied are likely to be excessively conservative, resulting in more compliance wells than are likely necessary to achieve an acceptable "monitoring efficiency." This conclusion results from the following observations:

- <u>Oversimplified representation of transport through the vadose zone</u>. The vadose zone below the burial ground contains carbonate-rich layers, such as caliche, which may support perched water, but these layers are expected to be of limited horizontal extent due to pinch-outs and fractures. Current interpretations of the vadose zone at and near LLWMA-3 suggest highly variable materials in laminae, some of which are fairly extensive horizontally and others of which have limited lateral extent. These conditions argue against the idealized conceptual model that a RCRA facility discharge travels vertically downward before impacting the groundwater. The vadose zone conditions at and near LLWMA-3 suggest a more complicated scenario in which a discharge travels downward, impinges on a heterogeneity and migrates horizontally, before again traveling downward toward the water table. One implication is that the "patch" may be wider than implemented in the MEMO model input, due to horizontal spreading within the vadose zone. Another implication is that groundwater monitoring wells located immediately adjacent to the edge of the waste management unit may not be as protective as a well located further away from the facility, due to potential for horizontal spreading in the vadose zone.
- <u>Locating compliance wells immediately outside the waste area</u>. Locating wells in the immediate vicinity of the waste, as opposed to some distance from the waste, will result in a need for closer well spacing. By placing compliance wells further from the

boundary of the waste, fewer wells will be needed for a fixed "monitoring efficiency" due to transverse dispersion. It is also noted that placing wells further from the waste boundary is sensible with respect to future capping of the waste; emplacing covering materials may cause damage to monitoring wells located too close to construction operations.

- Oversimplified groundwater flow direction assumptions. The Review Team notes that changing groundwater flow directions over time will spread impacts over a wider area than simulated by the MEMO model, which is based on one fixed flow direction. Thus, if some variation in flow directions occurs, fewer wells would likely be required to detect impacts. Representing and implementing this in the MEMO approach may represent a somewhat more difficult modification than suggested by the other items discussed above. The Review Team notes, however, that the usual definition of dispersivity is based on spatial averaging of flow variations over unobserved scales and that an analogous temporal averaging of flow can be interpreted as an enhanced lateral dispersion.

The Review Team suggests that, if the use of the MEMO continues, these issues should be considered more rigorously.

4.5 Coordination between RCRA and CERCLA

Another issue discussed by the Review Team relates to the coordination of RCRA monitoring with CERCLA monitoring of contaminants that have migrated from other areas. Given that LLWMA-3 and other RCRA units in detection monitoring status are embedded in an area with numerous waste facilities, along with the expense of adding new wells in this area, installing a high concentration of new monitoring wells around the boundary of LLWMA-3 would draw resources away from installing monitoring devices at locations that might be more useful for monitoring other contaminants and facilities.

As discussed elsewhere in this document, the Review Team recommends a holistic approach to groundwater monitoring across programs. Achieving this might involve an aggressive multi-lateral evaluation of possible flexibilities in the RCRA monitoring, predicated on the objective of more robust monitoring of the overall area, subject to the limited resources that exist. A detailed site-wide groundwater model or a 200-W Area sub-model could potentially be used with well location optimization tools and/or a decision framework (as discussed in Section 6) to help establish prioritized monitoring well locations for an integrated RCRA/CERCLA approach.

One possibility discussed during the site visit is using vadose zone monitoring (VZM) for groundwater protection that could potentially be used to detect impacts from LLWMA-3 before they reach groundwater. An evaluation of VZM alternatives is beyond the scope of this review; none-the-less, the Review Team notes the comprehensive review of this topic presented by Fluor Hanford [2003]. The Fluor Hanford document includes detailed information regarding VZM, including a discussion of different potential monitoring approaches for the vadose zone in Hanford's Central Plateau, some of which involve the use of innovative *in situ* and non-invasive monitoring technologies.

5.0 REGULATORY COORDINATION

5.1 Progress to Date

The Review Team observed during the site visit that all parties are aware of the potential for overlapping regulatory programs to cause redundancy in monitoring, create uncertainty in the decision-making process, and hamper attempts to optimize monitoring programs. It was reported to the Review Team that reaching agreement on regulatory issues has been slow. It was also reported that only limited funding is available for new well construction over the entire Hanford complex, making it extremely important to make sure that decisions on how to use those funds are made on the basis of holistic goals and constraints that account for the costs and benefits of localized actions.

The Review Team observed that some progress has been made in overcoming these obstacles. In 1989, federal and state regulators and the DOE signed the “Tri-Party Agreement”, or TPA, which coordinates the relevant regulations, sets the agenda for cleanup, and establishes a schedule of deadlines, known as milestones. The Review Team was also made aware of the *Hanford Site Groundwater Strategy* (DOE/RL-2002-59) document, dated July 2003, prepared as a collaborative effort by Ecology, USEPA, and DOE (the Tri-Parties). That document presents a strategy for coordinatin among multiple regulatory authorities and government agencies, but it does not specifically change regulatory requirements or procedures. The stated goals of the strategy document include:

- Identifying regulatory requirements and environmental objectives to protect, monitor, and remediate groundwater;
- Providing a clear mechanism to achieve the mission of the Hanford Groundwater Program by integrating RCRA, CERCLA, and AEA requirements and minimizing overlap, duplication, and inconsistencies; and
- Providing a framework that relates data needs to the decisions needed for remedial activities and monitoring.

With respect to monitoring, the strategy document indicates that USEPA’s data quality objectives (DQO) process was successfully used to integrate the RCRA, CERCLA, and AEA groundwater monitoring requirements in the 200 West Area, and that those efforts can be used as a model for the remaining groundwater regions.

Another item described to the Review Team was Milestone M-24-57 associated with the Tri-Party Agreement. This milestone specifies a minimum number of new wells to be installed per year, and includes a mechanism for prioritizing those wells based on overall benefits given the limited resources.

The Review Team also observed that some documents are now being prepared in a manner that integrates the requirements of CERCLA, RCRA, and AEA monitoring. For example, the 300-Area RCRA TSD was integrated with the 300-FF-5 CERCLA OU in a recent RCRA permit modification submittal. Also, recently approved CERCLA sampling and analysis plans (SAPs) have integrated CERCLA, RCRA, and AEA elements.

5.2 Further Evolution Needed

Although progress is evident regarding integration of the various regulatory programs, the Review Team recommends additional effort in this regard. The primary objective in this recommendation is cost-effectiveness, given the limited resources available for monitoring, in particular for siting new wells in the Central Plateau. The Review Team suggests that combined RCRA/CERCLA monitoring for areas that might contain multiple RCRA units may be more cost-effective, and ultimately equally protective of human health and the environment, compared to the approach where each individual RCRA is monitored separately. Embedding of RCRA requirements within larger CERCLA units would likely be a cost-effective approach. Of course, a consensus regarding protectiveness will need to be established between the regulatory agencies. A detailed decision framework (see Section 6) would help in this regard.

The Review Team also recommends periodic evaluation of new monitoring technologies that might provide direct or indirect information to augment traditional monitoring. Such methods (e.g., *in situ* sensors) could provide opportunities to further integrate the monitoring needs of the various regulatory programs, by providing more cost-effective approaches to monitoring. These technologies are innovative in both technical and regulatory terms, but there is growing momentum toward considering and allowing their use in closely related areas, as discussed in Section 7.

6.0 PRE-ESTABLISHED DECISION FRAMEWORKS

As an extension of the issue of regulatory coordination, and in view of the forthcoming FFS-related efforts, the Review Team suggests that the use of pre-established (i.e., pre-negotiated) decision frameworks may be beneficial. The use of a decision support system provides a generalized framework for consistent, technically defensible, and traceable environmental decisions. These decisions, depending on the type of framework, can evaluate the adequacy of system performance, resource management, prioritization of data collection activities, and remediation selection and design. Frameworks provide a consistent, structured foundation for negotiation between site operators, regulators, and stakeholders, and facilitate coming to closure on decisions in a timely and cost-effective manner.

Hanford is repeatedly making many of the decisions listed above. Having an agreed-upon decision framework would greatly improve speed and efficiency of the decision-making process. Many decision framework systems/programs are currently in use at other sites, and could provide a useful starting point from which Hanford could create its own decision framework system; alternatively, Hanford might use commercially available decision support software (DSS).

Typical DSS packages integrate environmental data and simulation models into a framework for making site characterization, monitoring, and cleanup decisions. Applications for DSS packages include optimization of sampling locations and cost-benefit analysis of additional or reduced sampling, as well as human health and ecological risk analysis as influenced by contaminant cleanup endpoints. An effective DSS package integrates, analyzes, and presents environmental information to assist a project manager in developing a cost-effective and defensible cleanup or monitoring strategy. Some examples of systems currently in use are the following:

- The Vandenberg Air Force Base Basewide Groundwater Monitoring Program consists of a site-specific optimization decision tree to recommend wells or analyses for reduced sampling frequency and to recommend elimination of redundant wells and analytes. It includes tools for visualization, temporal trend analysis, and geostatistical analysis of plume mass, size, and movement. A presentation on the system may be found at www.clu-in.org/siteopt/ataglance.htm.

- Cost-Effective Sampling (CES) is a methodology for reviewing and assessing ground water data. The CES program produces a data information sheet and a lowest-frequency sampling schedule for a given groundwater monitoring location that provides the needed information for regulatory and remedial decision-making. The determination of sampling frequency for a given location is based on trend, variability, and magnitude statistics describing the contaminants at that location. For more information on CES go to www.erd.llnl.gov/library/JC-118909.pdf

- CTech's EVS-PRO unites interpolation, geostatistical analysis, and fully 3-D visualization tools into a software system developed to address, among other things, sample optimization and cost-benefit analysis. EVS-PRO's capabilities can be used to provide 3-D maps of geologic structure, subsurface contamination, and regions containing contamination above specified threshold concentrations at a fixed probability level. EVS-PRO can also perform geostatistical analyses that optimize sample locations for site characterization and can estimate volumes and mass of contaminated media for use in cost-benefit analysis. EVS-PRO can quantify the statistical variation in the contaminant volume and mass estimates resulting from the current level of characterization. Further information may be found at www.ctech.com/products/products.htm.

- Sampling*FX* is a geostatistics-based software program intended to provide decision makers and analysts a means of evaluating environmental information relative to the nature and extent of contamination in surface and subsurface soils. Key attributes of the product include the ability to delineate, provide visual feedback on, and quantify uncertainties in the nature and extent of soil contamination (e.g., concentration distribution, probability of exceeding a soil cleanup guideline); to provide objective recommendations on the number and location of sample locations; and to provide statistical information about the contamination (e.g., average volume of contamination, standard deviation, etc.). See www.decisionfx.com/SampFX.html for more information.

- Spatial Analysis and Decision Assistance (SADA) is an environmental software product that incorporates tools from various fields, including visualization, geospatial analysis, statistical analysis, human health risk assessment, cost-benefit analysis, sampling design, and decision analysis, into a dynamic and interactive environment. The modules can be used independently or in an integrated fashion to address site-specific concerns in the characterization and remedial action design. SADA was designed to simplify and streamline several of the processes in environmental characterization, risk assessment, and cost-benefit analysis to bring the information together in a way that can help users make decisions about their particular site in a quick and cost-effective manner. Further information may be found at www.tiem.utk.edu/~sada/decision_analysis.html.

- MAROS is a decision support tool based on statistical methods applied to site-specific data that account for relevant current and historical site data as well as hydrogeologic factors (e.g. seepage velocity) and the location of potential receptors (e.g., wells, discharge points, or property boundaries). Based on this site-specific information, the software suggests an optimization plan for the current monitoring system in order to efficiently achieve the termination of the monitoring program. See www.gsi-net.com/software/Maros.htm for further information.

- The Sandia Environmental Decision Support System (SEDSS) is a methodology and tool for risk assessment, site characterization, and comparing remedial alternatives for environmental restoration. For more information on SEDSS go to www.sandia.gov/Subsurface/factshts/ert/sedss.pdf.

- Visual Sample Plan (VSP) provides simple, defensible tools for defining an optimal, technically defensible sampling scheme for characterization. VSP is applicable for any two-dimensional sampling plan including surface soil, building surfaces, water bodies, or other similar applications. Further information may be found at http://dqo.pnl.gov/vsp/.

7.0 Future Visions for Groundwater Monitoring

All current RCRA groundwater monitoring regulation and guidance are framed in terms of the traditional paradigm:

- Install wells;
- Purge the wells;
- Obtain samples periodically;
- Preserve the samples;
- Transport the samples to a laboratory with necessary chain of custody documents;
- Analyze the samples for selected waste constituents and indicator parameters; and
- Report the analytical results to the facility and (if and when necessary) to the regulatory agency.

This paradigm is labor-intensive and costly, particularly in a long-term setting such as post-closure monitoring.

A great deal of effort has been expended by DOE, USEPA, and other groups on developing field analytical techniques with the potential to greatly enhance the efficacy of characterization and monitoring efforts. These techniques range from minor modifications of standard procedures, such as the use of field-portable analytical instruments with samples collected as usual, to *in situ* instruments and sensors, and ultimately to autonomous sensor networks for monitoring the groundwater itself. In addition, Fluor Hanford's [2003] comprehensive review of vadose zone monitoring techniques for monitoring for groundwater protection that would be applicable at the Hanford Site, particularly in the elevated Central Plateau, has already been noted.

DOE has sponsored a good deal of this research. Basic and applied research have been funded by the Environmental Management Science Program (http://emsp.em.doe.gov), including work by PNNL and others at Hanford. Through Fiscal Year 2002 the DOE also maintained the Characterization, Monitoring, and Sensor Technology program in the Office of Science and Technology; see its archived web site (www.external.ameslab.gov/cmst) as well as the archived Technology Management System (http://tms.em.doe.gov). A successor to that program has been the Advanced Monitoring Systems Initiative (AMSI), which is currently sponsoring research and development at Hanford of *in situ* technologies for monitoring hexavalent chromium in the Columbia River and technetium-99 and strontium-90 in the groundwater. A presentation on AMSI can be found at www.clu-in.org/siteopt.ataglance.htm.

The USEPA Technology Innovation Program (Office of Superfund Remediation and Technology Innovation) has been at the forefront of this development in recent years as well (www.clu-in.org). In particular, that program champions the use of the Triad approach to sampling and characterization during site remediation. The Triad approach emphasizes combining field measurement technologies with more conventional ones, emphasizing the fact that gains in spatial coverage and quick turn-around can often more than compensate for lower analytical precision or sensitivity. The Triad approach currently addresses characterization and active remediation, but the same concepts can apply to long-term monitoring as well.

Another group championing the use of innovative technologies is the Interstate Technology Regulatory Council, which sponsors product evaluations, short courses, and documents. ITRC is a joint effort of several states along with USEPA, the Department of Defense, and DOE; relevant materials may be found under the SCM (Sampling, Characterization, and Monitoring) team tab on its web site www.itrcweb.org. The Department of Defense as well has contributed to this effort; see the web site www.denix.osd.mil (Defense Environmental Network and Information Exchange), for example.

Chapter 6, "Future Needs", of the ASCE's [EWRI, 2003] *Long-Term Groundwater Monitoring: The State of the Art* contains a substantive discussion of possibilities for the future of long-term monitoring. That chapter was written with the Hanford Site and other large DOE and Department of Defense facilities specifically in mind. It calls for continuing research and development of technology and concepts (such as appropriate Data Quality Objectives and Quality Assurance concepts for sensor networks), technology verification and transfer for current R&D efforts, and the regulatory evolution needed to implement these innovations. The efforts of the USEPA Technology Innovation Program and the ITRC to implement the latter are encouraging.

8.0 SUMMARY OF RECOMMENDATIONS

At the request of the DOE/RL, DOE/HQ assembled a Review Team to address issues related to groundwater monitoring optimization at the Hanford Site. The Review Team included technical expertise across a broad range of disciplines, in order to address numerous issues that exist at Hanford. The charge to the Review Team was derived through interactions between DOE/RL and DOE/HQ, and included a request to provide demonstrations and/or advice specific to two areas: the 300-FF-5 Operable Unit and LLWMA-3. The Review Team was provided with numerous documents and databases of field observations and also participated in a four-day site visit that included presentations and discussions with DOE/RL, PNNL, Fluor Hanford, Ecology, and USEPA personnel with site responsibilities. The Review Team provided an out-briefing at the conclusion of the site visit and has subsequently collaborated on the analyses and demonstrations provided in this report. The Review Team offers the following recommendations and suggestions.

8.1 Perform Quick Studies Related to Modeling at the 300-FF-5 OU

A separate, simpler, directed engineering analysis in parallel with the more detailed geochemical analysis and modeling being performed by PNNL ("science task") should be initiated immediately. This effort will support the FFS, for which the first submittal is due in March 2005. This activity will be of greatest benefit if there is regular, bi-directional communication with the science task, but must be a mission-oriented activity. Its work must support and flesh out revisions of the conceptual model for the 300-FF-5 OU, particularly for uranium, and must ensure that those baseline and bounding calculations necessary to examine alternatives are completed using the best science available within the next several months. Key questions to be addressed include the following: What amount of potentially mobile U is in the vadose and saturated zones? What is the vertical distribution of dissolved U in the aquifer? What is the rate and rate-limiting process for transferring U from the vadose zone to the saturated zone; and what is the anticipated cleanup time (and its uncertainty) under various scenarios.

A medium-scale leaching experiment, using for example a 2x2x2-meter cube of vadose zone material collected near the water table, may be valuable to this effort. It would provide for controlled experiments on the effects of fluctuating water table and carbonate concentration (due to river water incursion), mixing rules to support alternative statistics, and up-scaling of laboratory results.

8.2 Improve Flux Estimates for the 300-FF-5 OU

Although sufficient monitoring of concentrations of various chemical constituents is occurring, improved estimates of dissolved mass flux out of the saturated groundwater zone would be beneficial. The previously mentioned vertical distribution of U in the saturated zone is one important element of mass flux estimates. Another element, of course, is estimating the groundwater flow magnitude and direction. Some newer flux samplers use existing solutes to estimate dissolved mass flux; one approach uses a passive instrument and another uses an active sampling approach (see Section 3.3.2).

8.3 Implement Alternate Statistical Tests for RCRA Detection Monitoring

Regarding not only the 300-FF-5 OU and LLWMA-3, but also all of the RCRA regulated units in detection monitoring status, alternative statistical methods using intrawell tests such as CCCs are warranted. Neither the Jandhyala-Zhang [1999] nor the PNNL [Chou, 2004] proposals for using CCCs appear adequate, however. An alternative approach to evaluating the statistical performance of CCCs that addresses the sensitivity to the use or non-use of resamples and the related issue of automatic updating are detailed in Appendix G.

Another significant factor is the use of CCCs when there is an underlying trend (e.g., the rise in conductance from very low to natural levels as a result of the cessation of discharges to the surface in some areas). Ignoring the presence of trends can lead to false positives or false negatives. If the trend were constant, adjustments could readily be made. However, if the trending concentration approaches an asymptote, a more sophisticated adjustment may be warranted. Suggestions are provided on approaches to this problem. Also, because of the difference in water chemistry between the Columbia River and the groundwater, the interpretation of near-river groundwater sample results is affected by the dynamics of Columbia River flows. The simple mixing rule proposed by PNNL [2004] appears to offer a useful approach to remove noise from the sample results; a short-term bench-scale experiment is suggested to verify the mixing model.

8.4 Reconsider Monitoring Well Siting at LLWMA-3

A new set of monitoring wells for LLWMA-3 is being planned, because changing groundwater flow conditions have caused flow directions to shift and many of the 18 existing monitoring wells to go dry. DOE/RL has proposed 13 new wells, eight of which have been scheduled and tentatively located (mostly on the eastern side of LLWMA-3) through the Milestone 24-57 protocol. The status of the other five proposed wells is unclear, and Ecology is requesting a significant increase in the number of monitoring wells. With respect to the eight scheduled wells, the Review Team recommends their locations be changed to ensure the presence of two background (or upgradient) wells and to monitor all sides of LLWMA-3. For example, two monitoring wells might be located on each side of the (roughly) quadrilateral facility. The Review Team also recommends that the conceptual models for potential vertical transport of waste through the vadose zone, with its heterogeneities and laterally discontinuous carbonate layers, and for direction of groundwater flow be compared to the conceptual models embodied in current planning methods. Specifically, if the MEMO approach that has been used for over a decade continues to be used for planning purposes, a comparison of conceptual models will lead, the team believes, to locating new monitoring wells a greater distance from the edge of the burial ground and to increasing the spacing between new monitoring wells.

8.5 Continue Efforts toward Regulatory Coordination

The Review Team observes that coordination among DOE, USEPA, and Ecology over the past several years has led to improvements in planning, reporting, and costs. The Tri-Party Agreement, Milestone 24-57 process, and unified annual monitoring reports are concrete examples resulting from these efforts. Additional opportunities exist, and the Review Team recommends that the tri-parties pursue them vigorously.

For example, LLWMA-3 is unit regulated under interim status RCRA rules that is embedded within an area that is regulated under CERCLA; disjoint monitoring requirements promote conflict among the parties regarding the scope (and costs) of monitoring programs without a rational, holistic assessment of protectiveness in the context of both CERCLA and RCRA programs. The Review Team encourages establishing formal decision frameworks, such as have been developed for Vandenberg AFB, that encapsulate pre-negotiated decision points and accommodate resource requirements and limitations.

Another area that should be pursued relates to innovative monitoring technologies. New mass flux assessment systems and *in situ* concentration sensors offer potential to improve the quality of assessments and reduce the costs of obtaining them. Because innovative monitoring methods are novel in regulatory terms, as well as technical terms, their adoption will require all three parties to commit to on-going communication and flexibility. The Review Team is optimistic in this regard, as there are many examples of innovative environmental technologies that have garnered regulatory support with concerted efforts, such as direct push techniques.

8.6 Undertake Groundwater Monitoring Optimization as Areas Transition to Final Remedies

Despite extensive groundwater monitoring activities across the Hanford Site, most areas are not yet in long-term monitoring status because they have interim remedies. The Review Team anticipates, therefore, that optimization will become increasingly important as they make the transition to final remedy. The demonstrations performed for this report and the applications that have already been performed by PNNL provide useful bounds on the savings that may accrue from the use of mathematical optimization. The probability of realizing these potential savings increases with the amount of stakeholder input to setting the optimization objectives and constraints, the ability to quantitatively incorporate the disparate preferences that different parties will give to different objectives, the flexibility to simultaneously address multiple regulatory frameworks, and the degree to which the mathematical optimization is included within the recommended decision frameworks.

9.0 ACKNOWLEDGEMENTS

The members of the Review Team express their appreciation to all of the Hanford Site personnel for their efforts at providing documents, presentations, and other information and their willingness to participate in candid exchanges of information and viewpoints. These include, from DOE/RL, Doug Hildebrand and Mike Foster; from PNNL, Evan Dresel, Bob Peterson, Stuaart Luttrell, Paul Thorne, Charissa Chou, Chris Murray, and John Fruchter; and from Fluor Hanford, Tom Fogwell. The discussions with Dib Goswami of Ecology and Mike Goldstein, Dennis Falk, and Marcia Knadle of USEPA are likewise acknowledged and appreciated.

10.0 References

Cameron, K.C. and P. Hunter [2002]. "Using Spatial Models and Kriging Techniques to Optimize Long-Term Groundwater Monitoring Networks: A Case Study", *Environmetrics* v. 13, 629-656.

Chou, C.J. [2004]. *Evaluation of an Alternative Statistical Method for Analysis of RCRA Groundwater Monitoring Data at the Hanford Site*, PNNL-14521, Richland, WA.

Davis, C.B. [1999]. *Comparisons of Control Chart and Prediction Limit Procedures Recommended for Groundwater Detection Monitoring*, EnviroStat TR 99-1, EnviroStat, Henderson, NV.

USDOE [2001]. *Groundwater Sampling and Analysis Plan for the 200-BP-5 Operable Unit*, DOE/RL-2001-49, Richland, WA.

USDOE, 1994. Phase I Remedial Investigation Report for the 300-FF-5 Operable Unit, DOE/RL-93-21, Rev. 0, Richland, WA.

USDOE, 1995. Remedial Investigation/Feasibilitity Study Report for the 300-FF-5 Operable Unit, DOE/RL-94-85, Richland, WA.

EWRI Task Committee on the State of the Art in Long-Term Groundwater Monitoring Design [EWRI, 2003]. *Long-Term Groundwater Monitoring: The State of the Art*, ed. by B.S. Minsker, American Society of Civil Engineers, Reston, VA, 2003.

Fluor Hanford [2003]. *Study of Vadose Zone Monitoring at the Hanford Site: Task II: Potential Applications at the Central Plateau Remediation Project*, prepared for DOE/RL by Fluor Hanford, Inc., Richland, WA.

Jandhyala, V.K. and H. Zhang [1999]. *Report on Statistical Analysis for B-Pond Groundwater Monitoring Program*, prepared for State of Washington Department of Ecology.

Michael, D, E. Dresel, C. Chou, C. Murray, D. Gilbert, and B. Pulsipher [2003]. "Rethinking Groundwater Monitoring at the Hanford Site", *Remediation*, 19-44.

PNNL, 1994. Leaching Tendencies of Uranium and Regulated Trace Metals from the Hanford Site 300 Area North Process Pond Sediments, PNL-10109, Richland, WA.

PNNL [2004]. *Hanford Site Groundwater Monitoring for Fiscal Year 2003*, M.J. Hartman, L.F. Morasch, and W.D. Webber, eds., PNNL,-14548, Richland, WA.

USEPA [1992].]. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities: Addendum to Interim Final Guidance*, USEPA, Washington, DC.

USEPA [2004]. *Demonstraiton of Two Long-Term Groundwater Monitoring Optimization Methods*, EPA/542/R-04/001a and EPA/542/R-04/001b, DRAFT, USEPA Office of Solid Waste and Emergency Response, Washington, DC.

APPENDIX A: TECHNICAL ASSISTANCE REQUEST

TECHNICAL ASSISTANCE REQUEST

Tracking Number:	
Request Title:	Groundwater Monitoring Optimization at Hanford
Contact Individual:	MS McCormick-RL, JA Frey-RL, RD Hildebrand - RL
Requesting Organization:	Richland Operations Office
E-Mail Address:	Matthew_S_McCormick@rl.gov Jeffrey_A_Frey@rl.gov r_d_doug_hildebrand@rl.gov
Phone Number:	509-373-9971 (McCormick) 509-376-7727 (Frey) 509-373-9626 (Hildebrand)
Fax Number:	509-372-3548 (McCormick) 509-376-0306 (Frey) 509-372-1926 (Hildebrand)

Scope of Work:

Groundwater monitoring for the Hanford Site currently requires expenditure of $9.9 million each year, an expense that is expected to continue for decades. Technical assistance is requested to evaluate the efficacy of the current approach and determine how the efficiency can be increased. Considering redundancies of CERCLA/RCRA drivers and other considerations, Hanford groundwater monitoring needs for the coming decades need to be evaluated and recommendations presented for future monitoring activities.

A team of experts from several organizations working together can provide this analysis and identify how cost savings can be achieved. The work can be performed in two phases: phase I (included in this Technical Assistance requested) will be provided by internal and outside experts and include analysis of current drivers as well as expected changes in the next several decades.

Gaining efficiency in groundwater monitoring will be a significant benefit to the upcoming 5-year CERCLA review of groundwater treatment systems by showing cleanup progress, and it can also translate to major cost savings for the RCRA monitoring program (in terms of reducing both the frequency and number of analytes).

Technical Assistance

Description of Problem

Groundwater monitoring consumes a significant portion of the Hanford budget and will continue to do so for decades. Realizing efficiencies in these systems would release valuable resources that could be used to implement final solutions for the groundwater Operable Units.

Ongoing modification of the groundwater-monitoring network at Hanford is needed to accommodate the changing groundwater characteristics and to ensure focus on future needs. Past efforts to evaluate groundwater monitoring have focused on identifying wells that provide redundant information and on data gaps in terms of defining contaminant plumes. Modifications that will be needed to address changes in contaminant distributions over time have only been addressed qualitatively and little has been done to predict how monitoring needs will change in the future or to design a system that can be effective over the next several decades. With $9.9 million spent in FY2004 operating and maintaining the groundwater program, a 20% increase in efficiency would provide approximately $20 million in cost savings over the next decade at Hanford.

Technical Assistance Approach

The Hanford Site is seeking to implement new and innovative approaches to more efficiently monitor groundwater. A two-phase approach is needed that will evaluate the drivers and opportunities for optimization of the monitoring system in phase 1 followed by development of the actual technology required in Phase II. Phase I will include experts from across the DOE complex and will bring some of the best minds to Hanford for discussions leading to recommendations on how to make groundwater monitoring at Hanford more efficient. The Phase I team will provide timely input to Hanford and provide technical information as well as credibility with regulators and the public for the chosen path forward.

Support:

What resource(s) have been selected?

Tentative Schedule

Select Phase I Technical Assistance Team – May 1, 2004
Provide information to Technical Assistance Team – May 15, 2004
Meetings in Richland – June 15, 2004
Final report – July 31, 2004

What resources were offered, but not selected?

N/A

Requested Start Date: 5/1/2004 **Requested Phase I Completion Date:** 7/31/2004

Estimated Phase I Cost:

Labor:	$43K
Mat'ls:	$ 1K
Supplies:	$ 1K
Travel costs:	$15K
Total:	$60K

Submitted By: Jeffery A. Frey [signature] 4/15/04

Approved By: Matt McCormick, AMCP [signature] 4/29/04

Appendix B: Documents Provided to the Review Team

Site-wide

Hanford Site Groundwater Monitoring for Fiscal Year 2003. M.J. Hartmann, L.F. Morasch, and W.D. Webber, editors. PNNL, March 2004.

Hanford Site Groundwater Monitoring Project Data Viewer And Evaluator along with a July 2004 snapshot of the groundwater data in the *Hanford Environmental Information System*.

Michael, D., E. Dresel, C. Chou, C. Murray, D. Gilbert, and B. Pulsipher, "Rethinking Groundwater Monitoring at the Hanford Site", *Remediation*, Spring 2000.

Murray, C.J., Y.-J. Chien, and P.E. Dresel, "Ranking of Monitoring Wells Using Geostatistics", Preprint.

300 Area

Operation and Maintenance Plan for the 300-FF-5 Operable Unit, DOE/RL-95-73 Rev. 1. Prepared by CH2M Hill Hanford, Inc. for USDOE, May 2002.

Remedial Investigation/Feasibility Study Report for the 300-FF-5 Operable Unit, DOE/RL-94-85 Rev. 0. USDOE, 1994.

300-FF-5 Operable Unit Sampling and Analysis Plan, DOE/RL-2002-11 Rev. 0. Prepared by CH2M Hill Hanford, Inc. for USDOE, May 2002.

Patton, G.W., B.L. Tiller, E.J. Antonio, T.M. Poston, and S.P. VenVerst, *Survey of Radiological and Chemical Contaminants in the Near-Shore Environment at the Hanford Site 300 Area*, PNNL-13692 Rev. 1. PNNL, March 2003.

Lerch, J.A., *Data Quality Objectives for the 300 Area Leach/Kd Study*, BHI-01441 Rev. 0. Prepared by Bechtel Hanford, Inc. for USDOE, November 2000.

Serne, R.J., V.L. LeGore, and S.V. Mattigod, *Leaching Tendencies of Uranium and Regulated Trace Metals for the Hanford Site 300 Area North Process Pond Sediments*, PNL-10109/UC-510. Prepared by PNL for USDOE, September 1994.

Kaplan, D.I. and R.J. Serne, *Distribution Coefficient Values Describing Iodine, Neptunium, Selenium, Technetium, and Uranium Sorption to Hanford Sediments*, PNL-10379 Sup. 1/UC-721. Prepared by PNL for USDOE, March 1995.

Gamerdinger, A.J., D.I. Kaplan, and C.T. Resch, *Uranium (VI) Sorption and Transport in Unsaturated, Subsurface Hanford Site Sediments – Effect of Moisture Content and Sediment Texture: Final Report for Subtask 2b*, PNNL-11975. Prepared by PNNL for USDOE, October 1998.

Lindburg, J.W. and C.J. Chou, *300 Area Process Trenches Groundwater Monitoring Plan*, PNNL-13645. Prepared by PNNL for USDOE, August 2001.

Poston, T.M., *A Survey for Elevated Levels of Uranium North of the 300 Area on the Hanford Site*, PNL-7271/UC-602. Prepared by PNL for USDOE, April 1990.

Price, K.R. and R.R. Kinnison, *Uranium and Other Heavy Metals in Soil and Vegetation From the Hanford Environs*, PNL-4466/UC-41. Prepared by PNL for USDOE, November 1982.

LLWMA-3 and Similar Areas

Performance Assessment Monitoring Plan for the Hanford Site Low-Level Burial Grounds, DOE/RL-2000-72 Rev. 0. USDOE, November 2000.

Performance Assessment for the Disposal of Low-Level Waste in the 200 West Area Burial Grounds, WHC-EP-0645. Prepared by Westinghouse Hanford Company for USDOE, June 1995.

Dresel, P.E., *Interim Status Groundwater Monitoring Plan for the Low-Level Burial Ground Waste Management Areas 1 to 4 Resource Conservation and Recovery Act Facilities, Hanford, Washington*. INTERNAL DRAFT, PNNL, July 2004.

Hanford Site Risk-Based End State Vision, DOE/RL-2003-59. USDOE, DRAFT, April 2004.

Study of Vadose Zone Monitoring at the Hanford Site: Task II: Potential Applications at the Central Plateau Remediation Project, Fluor Hanford, October 2003.

Groundwater Sampling and Analysis Plan for the 200-BP-5 Operable Unit, DOE/RL-2000-49. Prepared by PNNL for USDOE, March 2003.

Appendix 5D: Monitoring Efficiency Model Output. Revised May 2003.

RCRA Statistical Methods

Chou, C.J., O'Brien, R.F., and D. B. Barnett, "Application of Intrawell Testing of RCRA Groundwater Monitoring Data When No Upgradient Well Exists", *Environmental Monitoring and Assessment* v. 71, 91-106, 2001.

Chou, C.J., "Groundwater Monitoring: Statistical Methods for Testing Special Background Conditions", Chapter 8 of *Environmental Monitoring*, G. B. Wiersma, ed., CRC Press, 2004.

Chou, C.J., *Evaluation of an Alternative Statistical Method for Analysis of RCRA Groundwater Monitoring Data at the Hanford Site*, PNNL-14521, PNNL, June 2004.

V.K. Jandhyala and J. Zhang, *Report on Statistical Analysis for B-Pond Groundwater Monitoring Program*, prepared for Ecology, February 2005.

APPENDIX C: A PRIMER ON OPTIMIZATION

C.1 Why Optimize?

Operating and managing long-term monitoring (LTM) systems for groundwater can be expensive. Because these systems will, by definition, operate over a long period of time and will have recurring costs, small improvements in efficiency or reductions in unit costs can accumulate to a significant amount over the life of the LTM program. Optimization of groundwater monitoring networks attempts to minimize resource requirements (such as capital, labor, or time) while satisfying the goals, objectives, and constraints of the LTM program.

Optimization of LTM systems can also formalize procedures by which a LTM system design and operation can be revised as new information and/or methods become available. A properly optimized LTM system, which accounts for the uncertainties and variabilities that always are present in groundwater environments and monitoring systems, continually generates new information that is evaluated to determine if further efficiencies can be realized. Such evaluations can be used to determine whether the correct data are being collected in light of changed conditions; whether the LTM system design and operation continue to be appropriate and meet stated objectives; and whether some of the long-term objectives of the LTM program have been accomplished or, in fact, whether these objectives remain appropriate in view of changing site conditions, technology, regulations, policies, and stakeholder expectations.

C.2 The Optimization Process

Optimization is a problem-solving process like design. It is iterative in nature; a single pass through the process is usually not sufficient. The optimization approach to problem-solving is outlined in Figure C.1.

The starting point is an understanding what needs optimizing, how the system being optimized came into being, and all of the impinging factors that limit or provide flexibility in decision-making. Next is defining the benefits sought from the particular optimization exercise, such as reducing cost. One also describes the resources (e.g., budget) that can be brought to bear; the solution of an optimization problem accordingly indicates how much of the limited resource is actually needed in the optimal solution. Once specified, a solution procedure is needed. There are many solution methods, which need to be matched to the problem definition. In some cases, mathematical optimization should be used; in others, the basis of decision may be the collected opinion of a set of experts. Once an optimal solution is obtained, it is important that it be subjected to review. Often in this stage, the optimization results provide an indication that the problem was not properly specified. For example, it may be determined that minimizing the cost of well sampling may be less important than providing validation data needed for a treatability study. It is quite common to revise the optimization problem and cycle through the process.

Figure C.1. Outline of the Optimization Process.

1. Understand the problem. Groundwater optimization problems have many facets, including these specifics:
   - understand the technical and regulatory drivers;
   - understand objectives, goals, and constraints of problem owner and of stakeholders;
   - understand the physical setting and processes of significance.;
   - understand institutional and historical background, such as Data Quality Objectives (DQOs);
   - understand the conceptual site model (CSM); and
   - understand cost constraints and drivers.

2. Specify the problem as an optimization problem. Optimization problems are usually discussed in terms of objectives and constraints, benefits and costs, or some similar set of conflicting, yet complementary, features:
   - set or determine objectives (objective functions);
   - set resource limitations (constraints);
   - identify critical non-technical or non-quantifiable objectives and constraints.

3. Select and implement an appropriate method. Two of the general categories are:
   - mathematical optimization;
   - expert-guided optimization.

4. Review results with the problem owner; reassess the problem specification or optimization approach as needed.

5. Iterate as necessary.

C.3 Benefits of Optimization

The optimization approach to problem solving can yield many benefits. Cost savings that result from optimization are perhaps the most obvious: optimization can identify redundant sampling that can be eliminated with an acceptable (or even inconsequential) reduction of information, enabling a reduction in labor, analysis, and data management costs. Such redundancies may be caused by too frequent sampling or too closely spaced samples. A slightly different formulation of the optimization problem can lead to the identification of information gaps, providing guidance on where additional sampling will be most beneficial. Weaknesses in conceptual models and in their numerical implementations may also be identified by finding locations where errors committed by simulation models are consistently larger than elsewhere.

Specific benefits that can be obtained for a particular problem are actually elicited in the problem specification step (Step 2) of Figure C.1. Characterizing the problem is the step of converting problem understanding into a specific set of tasks, goals, objectives, constraints, and resource limitations that the optimization method will address. Some elements of the characterization step may be obvious, e.g., a particular site may have monitoring objectives that include estimating both the mass of a contaminant remaining in an aquifer and the flux of a solute across a boundary region and the resulting discharge rate into a receiving water body. Others may be more challenging, such as assigning priorities or preferences to competing objectives when multiple stakeholders are involved.

The physical design of LTM networks generally involves the locations of monitoring points, what is to be sampled at those points, and how frequently these points should be sampled. Optimizing the design of a monitoring network often has the goal of reducing one or more of these parameters, while still meeting the objectives of the LTM system. In other cases, the goal is to optimize the augmentation of a network (either adding a sample location or revising the sample schedule) so as to maximize the marginal information provided about a needed plume characteristic. Doing this is not a simple task. Sampling location decisions can be influenced by remedy selection, geologic heterogeneity, contaminant types, hydrogeologic conditions, and potential future data needs. Sampling frequency can be influenced by sampling methods, sample cost and accuracy (screening or indicator samples), contaminant types, rates of groundwater flow, potential risks, governing regulatory framework, and various other factors. The impacts of these issues can be modeled to determine the sensitivity of optimal configurations to them. Correctly establishing the optimization problem can be quite challenging, particularly in the context of geohydrologic and source uncertainties, multiple objectives, and multiple stakeholders.

C.4 Optimization Approaches

Optimization of long-term groundwater monitoring systems can be accomplished using a variety of approaches. The relative merits of the various available approaches and techniques depend on several factors, including (1) the scale of the monitoring program, (2) the objectives of the monitoring program, (3) the type of data being used in the analysis (subsurface stratigraphy, water levels, and groundwater chemistry), (4) the main processes affecting the contaminant (for example, transport and fate of chemicals in the vadose and saturated zones), (5) whether the long-term or transient nature of groundwater quality properties is pertinent, (6) the time-dependent goals of a long-term monitoring program, and (7) the stakeholders involved in the site and their legal and regulatory standing. A recent monograph published by ASCE (EWRI [2003]) provides a more detailed review of methods than can be presented here. There are four general approaches to optimization, relying on the following bases:

- hydrogeological expert judgment;
- statistical and geostatistical;
- physical models; and
- a combination of statistics with physical models.

They can be employed in an *ad hoc* way or in a mathematical optimization approach.

The hydrogeologic expert judgment approach is the case where the long-term monitoring program is optimized on the basis of the calculations and judgment of the hydrogeologist without resorting to advanced physics-based, statistical, or probabilistic techniques. In this approach sampling locations and frequencies are determined by the hydrogeologic conditions at and near the monitoring network of interest.

Statistical methods, including both time series and spatial (i.e., geostatistical) statistics, are widely used to determine the information content of a set of measurements and the value of a given measurement. For example, statistical analysis can be used to identify the temporal autocorrelation in a series of measurements, which might better be sampled less frequently to reduce redundancy.

Methods based on physical modeling use mechanistic (phenomenological) relationships expressed as numerical simulation models to express how contaminant plumes are affected by changes in the operation of the groundwater system. These combine physical principles (such as mass balance and charge balance) with flux relations (e.g., Darcy's law) to develop input-output relationships. Such models can be used to identify where concentrations are expected to be at a specified value (e.g., the plume boundary), which can be compared to monitoring well locations to identify data gaps and redundancies.

Recognizing that both physical-based models and observations have errors, another approach uses a combined statistical treatment of both of these spatially and temporally varying errors, in which a data-driven scheme is used to decide how to best combine sample and model data. This approach allows one to identify where the physical understanding (as encapsulated in the physical-based model) allows confident interpolation between samples, as well as locations where a purely statistical interpolation method is used because of large and persistent errors in the physical-based model.

These four approaches are often used in an *ad hoc* manner by consultants and called optimization. When a sequence of decisions needs to be taken optimally and each decision can be broken down into yes-or-no decisions, then the optimization strategy can be expressed as a decision tree. Section 6 presents a further discussion of decision frameworks.

Mathematical optimization provides a systematic (although not necessarily deterministic) way of searching for the optimal set of decisions for a given optimization problem. These methods are most appropriate for use when the number or complexity of possible designs makes manual search cumbersome. There are many mathematical optimization methods, beginning with the traditional linear programming, nonlinear programming, and integer methods such as branch-and-bound. In the past decade or two, a large number of heuristic optimization methods have been devised to address large-scale and difficult optimization problems. These include simulated annealing, tabu search, and genetic algorithms, all of which may be called evolutionary algorithms. These methods are exceedingly flexible, and if an optimization problem can be expressed quantitatively these methods can be brought to bear; the price is that these methods can sometimes require very large computational resources. Other mathematical optimization methods have been devised for incorporating the disparate objectives of multiple stakeholders, evaluating optimal solutions under uncertainty, and other complex situations. Detailed discussion of these methods is far beyond the scope of this report; see EWRI [2003] and references cited therein for further information.

C.5 Case Study of Long-Term Monitoring Optimization at DOE/Pantex

In 2001, G. Rice published "Evaluation of the Groundwater Characterization and Modeling at the Pantex Plant," on behalf of a local (to Pantex in Texas) stakeholder group. The report concluded that the onsite groundwater modeling performed to date was inadequate to simulate significant physical processes, the assumptions made were questionable, and impacts to environmental receptors such as the Ogalalla Aquifer from major contaminants were not accounted for. DOE sought a path forward at the Pantex Site to improve site-specific model predictions, monitoring, and surveillance through process optimization, incorporating a scientific advisory, stakeholder, and public review and acceptance program.

DOE convened a Technical Advisory Group with a directed scope of work to review state-of-the-art modeling, simulation, and evaluation methods, ultimately to develop a site-wide transport model to represent the 500-foot-thick vadose zone at the facility. DOE/Pantex undertook optimization of the groundwater modeling and monitoring using Mathematical Optimization Design Methods (MODM), a human expert/systems approach by supplying advanced formal optimization algorithms for key aspects of a project:

- Plume finding;
- Long-term plume monitoring;
- Optimal remedial design; and
- Optimal site-wide resource apportionment under uncertainty.

This methodology was endorsed by the Technical Advisory Group, a peer review scientific group, regulators, and community stakeholders. The plume-finding technology was used to evaluate the effectiveness of the existing groundwater monitoring well network, identifying optimal locations for additional wells. Monitoring optimization resulted in cost savings of approximately $1 million in new monitoring wells not needed. The deployed system is believed to be one of the largest peer-reviewed, fully three-dimensional variable saturated flow and transport models of multiple perched aquifers. The model is used extensively for simulation, risk assessment, and other advanced engineering evaluations.

More specifically, the methodology provides for optimal management of subsurface remediation activities using linear and non-linear elliptic and parabolic equations to model multiphase flow (e.g., of water, soil gas, non-aqueous phase liquid [NAPL]) and multi-component transport (e.g., of radionuclides, heavy metals, volatile organics, and explosives). These transport equations are resolved using numerical methods such as finite elements. Genetic programming is used to generate simulators from data when simulation models do not exist, are inaccurate, or are inefficient.

Geostatistic numerical models, Kalman filtering, and optimization tools are integrated to define subsurface conditions. Optimal contaminant plume finding is obtained for the plume fringe at a specified time in order to identify the least possible and most cost-effective number of monitoring wells. The model minimizes the uncertainty of the target boundary location. When applied to long-term monitoring, cost is minimized by integrating spatial-time correlations in order to increase the certainty of what is known about plume behavior. MODM uses prior sampling data (if available), knowledge of subsurface physics, and model simulations to indicate where and when to sample during the project life cycle.

The expert system handles estimation of life-cycle costs, maximum annual costs, or maximum allowable annual discharge (i.e., for MNA), and determines the best location of remedial system components. It also has the capacity to force certain solutions or preclude others. A suite of optimization techniques is available, including the outer approximation method, Lipschitz global optimization, and evolutionary algorithms.

C.6 References

EWRI Task Committee on the State of the Art in Long-Term Groundwater Monitoring Design [EWRI, 2003]. *Long-Term Groundwater Monitoring: The State of the Art*, ed. by B.S. Minsker, American Society of Civil Engineers, Reston, VA, 2003.

Rice, G. 2001 Evaluation of the Groundwater Characterization and Modeling at the Pantex Plant. Prepared for Serious Texans Against Nuclear Dumping.

APPENDIX D: PNNL GEOSTATISTICAL OPTIMIZATION ALGORITHM

The PNNL geostatistical optimization methodology described in Section 2.5.1 is summarized in the following flowchart (Figure D.1). This figure is based on descriptions made available to the Review Team. The highlights of the methodology follow.

D.1 Geostatistical Highlights

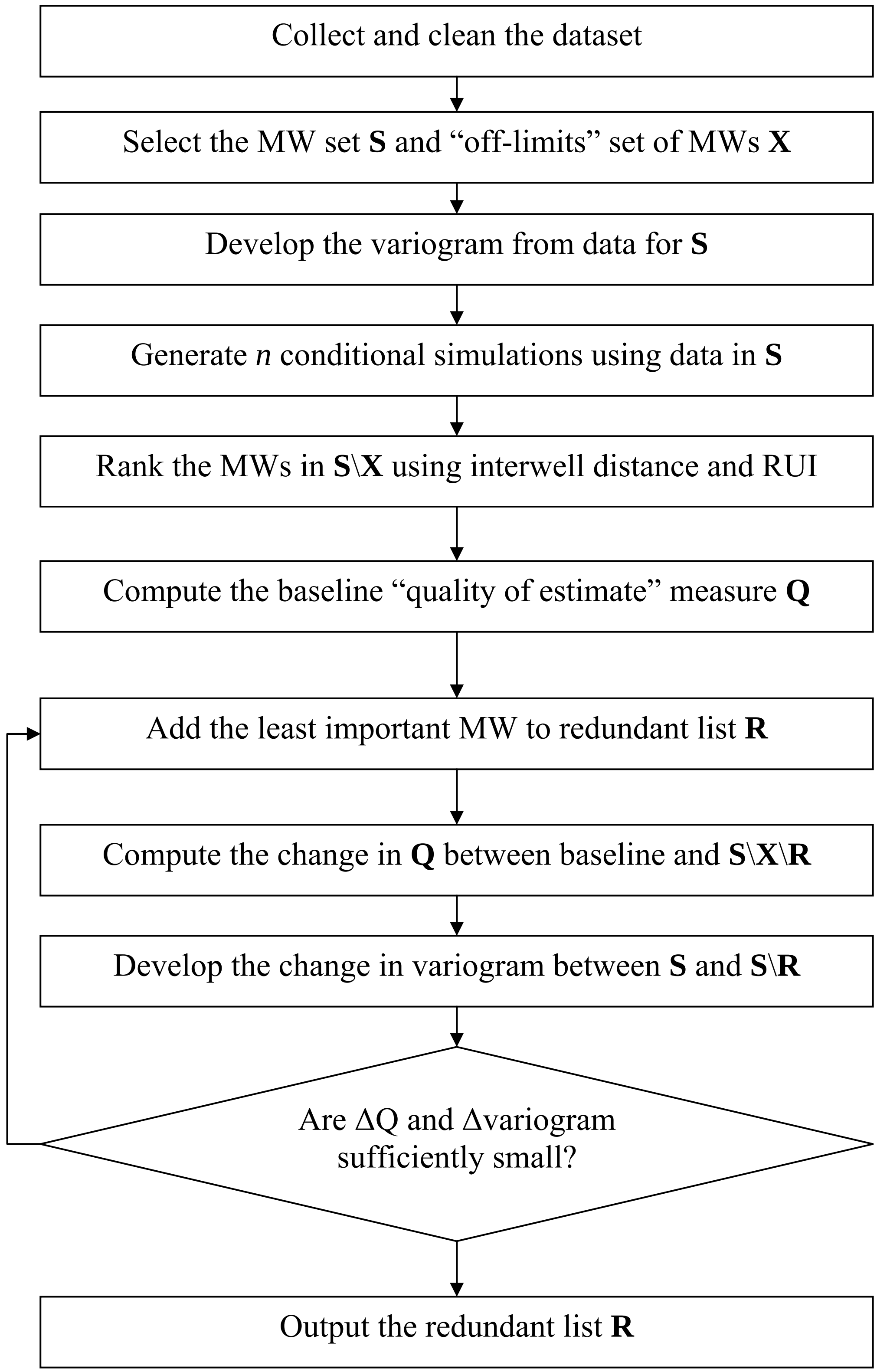
Once the plume (including the analyte of concern) is selected, the monitoring wells comprising its monitoring network are identified. A snapshot of concentration data (concentration values for all monitoring wells obtained from samples taken in a small range of dates) is collected and subjected to variography (the process of fitting a model variogram to the data). A grid of points covering the plume is selected. Using the observed concentrations at all of the monitoring wells, conditional simulation is used to generate a number of different plumes that are consistent with the observed concentrations and the model variogram, are equally likely to occur, and are defined by concentration values on the selected grid.

For every grid point, a measure of the variability among the generated plumes is calculated. Since there is particular interest in the location of the plume boundary (where the concentration equals a "target" value), the variability is weighted by how close the concentration is to the target boundary concentration. The resulting weighted variability, called the RUI, is quite small in areas near observation data and is large near the target concentration boundary and where there is large variability among the simulated plumes. In addition, for each observation well the distance to the nearest neighboring observation well is calculated.

D.2 Optimization Highlights

The significance of each monitoring well in the network that is not "off-limits" is determined by (1) excluding that well from the observed data set, (2) repeating the variography, conditional simulation, and variability assessments using the reduced data set, and (3) calculating the change in the plume (actually the change in the RUI across the plume) caused by excluding that monitoring well. The wells are then ranked (1) by the RUI from low to high, so that a monitoring well with low rank is less informative than a well with high rank, and (2) by the increasing distance to the nearest neighboring observation well. The average of the two rankings are used to reduce the network by removing the well with rank 1, determining the change in the concentration map that results, removing the well with rank 2, determining the change in the concentration map caused by deleting the 2 wells, and so on. The optimization procedure ends when either a specified number of redundant wells has been identified or the root mean square error between the observed and median estimates of concentrations at observation wells is excessive.

Figure D.1. Flowchart of the Basic PNNL Geostatistical Groundwater Monitoring Network Optimization Application.



Appendix E: GTS Demonstration Details

E.1 Introduction

This appendix provides an expanded description of the GTS methodology, along with selected technical results and descriptions from the demonstration of the GTS methodology at the 300-FF-5 OU. Each table or graphic is explained in sufficient detail to allow the reader an understanding of the basic results. A description of the GTS methodology can be found in Cameron and Hunter [2002] and on the AFCEE web site at http://www.afcee.brooks.af.mil/products/techtrans/models.asp.

Table E.1. List of Wells Used in GTS Demonstration

WELL_ID	Wellname	EASTING	NORTHING	Screen Top Elev	Screen Bottom Elev
A5018	399-1-1	594359.985	116588.842	102.88	92.21
A5020	399-1-11	594109.837	116660.156	107.75	101.35
A5021	399-1-12	594040.258	116548.511	103.75	99.18
A5024	399-1-15	593988.332	116964.243	107.34	102.76
A5025	399-1-16A	594318.118	116414.139	106.85	102.28
A5026	399-1-16B	594324.688	116411.619	84.85	81.8
A5028	399-1-17A	594112.87	116413.787	107.74	103.17
A5029	399-1-17B	594104.816	116417.723	84.99	81.94
A5031	399-1-18A	593870.644	117301.572	107.32	102.75
A5032	399-1-18B	593866.062	117297.232	85.83	82.78
A5035	399-1-2	594082.376	116329.576	110.58	95.34
A5036	399-1-3	594254.205	116334.906	110.65	96.93
A5038	399-1-5	594111.658	116552.122	109	102.29
A5039	399-1-6	594142.511	116900.079	107.53	100.82
A5040	399-1-7	594260.1	116335.071	110.11	94.87
A5041	399-1-8	594257.76	116329.57	91.92	85.83
A5043	399-2-1	594467.248	116121.223	109.18	91.81
A5044	399-2-2	594385.728	116282.618	104.66	98.56
A5046	399-3-1	594481.297	116008.028	111.1	97.38
A5047	399-3-10	594530.122	115832.268	107.12	102.55
A5048	399-3-12	594213.754	115946.786	107.87	103.6
A5049	399-3-6	593927.103	115944.403	106.96	102.99
A5052	399-4-1	594274.1	115537.349	113.22	96.45
A5053	399-4-10	594566.442	115655.802	106.53	100.43
A5054	399-4-11	594087.926	115709.187	107.03	102.45
A5055	399-4-7	594603.236	115492.6	109.25	69.94
A5056	399-4-9	594537.847	115741.45	105.16	99.07
A5057	399-5-1	593750.741	115525.32	113.95	90.48
A5059	399-8-1	593632.233	116331.956	110.34	95.71
A5411	399-1-10A	594346.573	116733.99	106.76	102.19
A5413	399-1-14A	593901.125	116778.217	107.21	102.64
A5414	399-1-21A	594160.776	116183.875	107.29	100.95
A5415	399-1-21B	594157.212	116176.798	86.06	83.01
A5416	399-8-5A	593384.245	116565.517	107.09	100.99
A8064	399-1-10B	594350.847	116728.788	82.93	79.89
A8072	399-3-3	594049.574	115804.753	106.31	97.47
A8077	399-3-11	594309.088	115793.674	107.43	101.33
A8089	399-4-12	594548	115373	104.76	98.67

E.2 Temporal Redundancy Analysis

To provide recommendations concerning long-term groundwater monitoring (LTM) at Hanford, a demonstration of the Geostatistical Temporal/Spatial (GTS) methodology was conducted on total (unfiltered) uranium data from the 300-FF-5 OU. Clearly, since a final remedy at this site has not been implemented, the results described below merely offer a proof-of-concept of the technology. In general, GTS assumes that the site being analyzed is roughly in the 'steady-state' condition associated with LTM. It is also particularly valuable in cases where an abundance of monitoring wells provides 'redundant' statistical information about the underlying plume(s).

GTS has two optimization components: a temporal module for optimizing sampling frequency and a spatial module for optimizing well network locations. Both modules are designed to reduce information redundancy in the LTM network (if it exists) by either lengthening the recommended sampling intervals and/or by decreasing the number sampling locations in the well network.

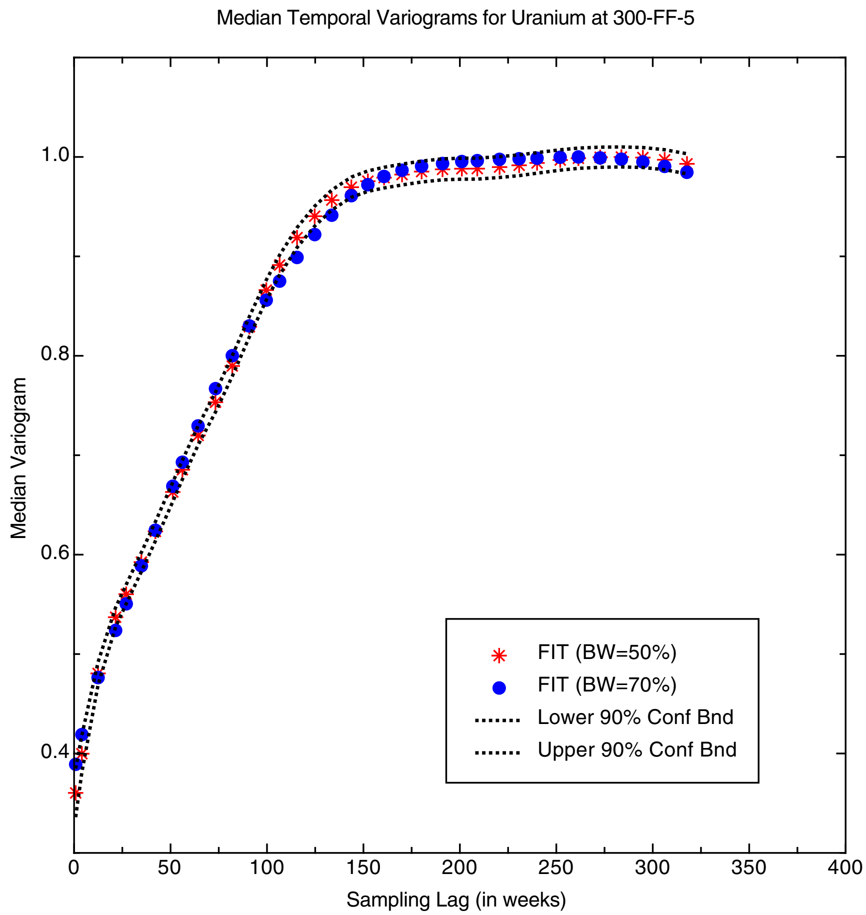
To provide maximal flexibility, the temporal module includes two techniques: temporal variograms and iterative thinning. The first of these creates a one-dimensional variogram from the historical data at the combined set of wells in order to determine the minimal average sampling interval at which consecutive sampling events become uncorrelated. This strategy thus eliminates temporal redundancy by ensuring a lack of statistical correlation between events. It is most useful in cases where the historical data may be limited at some or all the wells, yet a common sampling schedule is desired.

At the 300-FF-5 OU, data from 38 wells with total uranium concentrations were downloaded from the chemical summaries at https://www4.hanford.gov/groundwater/. Data collected prior to 1990 were eliminated from the analysis. Also, any well with a last reported sampling date prior to 1999 was not used. In addition, a subset of these 38 wells is apparently no longer being sampled, and so would not be considered part of the current sampling network. A few wells from the active list (as documented to the Review Team), however, are not currently listed on the Hanford web site. So for purposes of the demonstration, these 38 wells constituted the well network (see Table E.1).

The temporal variogram constructed from these data is shown in Figure E.1. Of interest, the interval at which consecutive sampling events become uncorrelated (on average) is approximately 3 years. This stands in contrast to the current sampling schedule for the 300-FF-5 OU monitoring network, which varies from almost monthly at selected RCRA wells to semi-annually or annually at other wells.

Note that the temporal variogram in Figure E.1 is a smoothed fit, using local regression, of the squared lag pair differences of the measurements after conversion by a uniform transformation. The confidence bands are also estimated using local regression. Two different local bandwidth parameters (i.e., smoothing windows) were used in creating these variograms, as shown in the legend. In both cases, the range of the variogram is approximately 150 weeks or nearly 3 years to the point of non-correlation between sampling events.

Figure E.1. Temporal Variograms for U at the 300-FF-5 OU



The second technique for temporal analysis in GTS is iterative thinning. This method requires at least 8-10 historical measurements per well, since a temporal trend is fitted at each location. The basic idea is to first construct at each well both a baseline trend and a confidence interval around that trend. Sampling points are then removed at random in percentage increments, and the trend re-fitted to the reduced data. Removal proceeds iteratively until the reduced-data trend no longer falls within the baseline confidence limits. A well-specific optimized sampling interval is then estimated based on the difference between the average baseline interval between events and the average interval for the reduced-data set.

Iterative thinning is more ambitious than the temporal variogram from the standpoints that (1) more historical data is required, and (2) reconstruction of trends is the goal, rather than an estimate of autocorrelation. Because of this, the results of iterative thinning will not always correspond to those of the temporal variogram. Also, each well receives a potentially different recommended sampling frequency. If a common schedule is desired, the average or median recommended sampling interval can be adopted. For the 300-FF-5 OU well network, the wells amenable to iterative thinning are listed in Table E.2. From this list, it can be seen that the median recommended sampling interval is approximately 1 year, corresponding to annual sampling.

Table E.2. Iterative Thinning Results for U at the 300-FF-5 OU

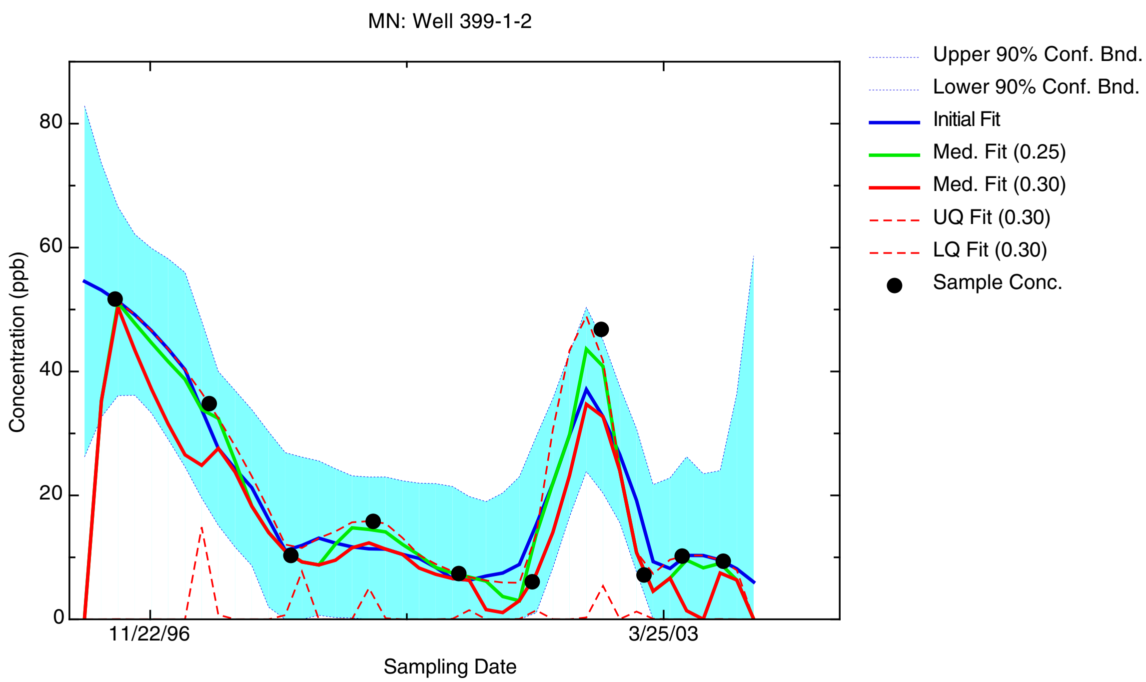
WELLID	EASTING	NORTHING	INTERVAL	FREQ	QINT
399-1-10A	594346.57	116733.99	19.58543	0.0473	2
399-1-10B	594350.85	116728.79	19.15698	0.0485	1
399-1-11	594109.84	116660.16	21.97117	0.0467	2
399-1-12	594040.26	116548.51	78.73023	0.0122	6
399-1-14A	593901.13	116778.22	65.02378	0.0149	5
399-1-16A	594318.12	116414.14	18.31668	0.0506	1
399-1-16B	594324.69	116411.62	25.29443	0.0366	2
399-1-17A	594112.87	116413.79	9.48298	0.0965	1
399-1-17B	594104.82	116417.72	38.46325	0.0239	3
399-1-18A	593870.64	117301.57	28.75132	0.0324	2
399-1-18B	593866.06	117297.23	17.58449	0.0529	1
399-1-2	594082.38	116329.58	57.3907	0.0176	4
399-1-21A	594160.78	116183.88	63.29311	0.0155	5
399-1-6	594142.51	116900.08	79.45867	0.0127	6
399-1-7	594260.1	116335.07	40.28041	0.0241	3
399-2-1	594467.25	116121.22	79.53136	0.0121	6
399-2-2	594385.73	116282.62	81.24461	0.0121	6
399-3-10	594530.12	115832.27	69.6725	0.0138	5
399-3-11	594309.09	115793.67	41.88171	0.0233	3
399-3-12	594213.75	115946.79	76.98588	0.0131	6
399-3-6	593927.1	115944.4	56.5556	0.0177	4
399-4-1	594274.1	115537.35	45.63344	0.0213	4
399-4-12	594548	115373	47.42766	0.0206	4
399-4-9	594537.85	115741.45	56.92144	0.0178	4
399-8-5A	593384.25	116565.52	56.52067	0.0177	4

Only 25 wells had sufficient data in the period since 1990 to be eligible for iterative thinning. The column INTERVAL refers to the estimated optimal sampling interval in weeks. Column QINT translates this optimized interval into the nearest number of quarters between events. The median result of this list is 4 quarters or annual sampling. The column FREQ is another expression of the optimal result, this time inverting the sampling interval into a recommended sampling frequency *per week*.

The key benefits of iterative thinning are (1) trends of almost any sort can be fit, including trends with complex or seasonal components, and (2) the methodology is data-driven and highly empirical. Selected examples of fitted trends and re-fitted trends at levels of optimal data removal are provided in Figures E.2 and E.3.

As shown in the legend, the initial local regression fit is shown in deep blue in Figure E.2. The green and red traces are denoted as median fits, since at each percentage level of random data removal, the fits are repeated 500 times to protect against bias from any particular fit. The green trace represents the median fit at the optimal level of data removal of 25 percent. The red trace is the median fit when too much data has been removed. Not only does it depart more clearly from both the initial fit and the 90 percent confidence band shown in aqua, but two related statistics, the red dashed traces, denoting the upper and lower quartiles of the 500 fits at 30 percent removal, also show significant departures from the initial confidence band, especially on the lower end.

Figure E.2. Fitted Trends for Uranium at Well 399-1-2



A key criterion used in judging optimality, in fact, is the percentage of repeated fits that fall outside the initial confidence band. This is particularly important in cases such as Figure E.3, where even with 50 percent removal the median fit (in red) is still rather similar to the baseline fit (in deep blue).

Another side benefit of the trend fitting done in iterative thinning is the fact that estimates of recent or historical trends can be mapped across the site, as in, for instance, Figure E.4. This representation illustrates that recent trends (since 2000) near the river have been generally downward. This trend map is constructed by estimating the median slope at each well across a dense set of fitting points spanning the time interval of interest (2000 and later in this case). Using the same set of slope estimates, a non-parametric confidence interval (Sen's method) is constructed to determine the statistical confidence associated with each trend estimate. Confidence intervals not including zero are deemed surely upward or downward, and mapped in red or deep blue. Other trends are shown in pink or aqua. The relative magnitude of each trend (relative to those estimated) is indicated by the symbol size.

Figure E.3. Fitted Trends for Uranium at Well 399-1-10A

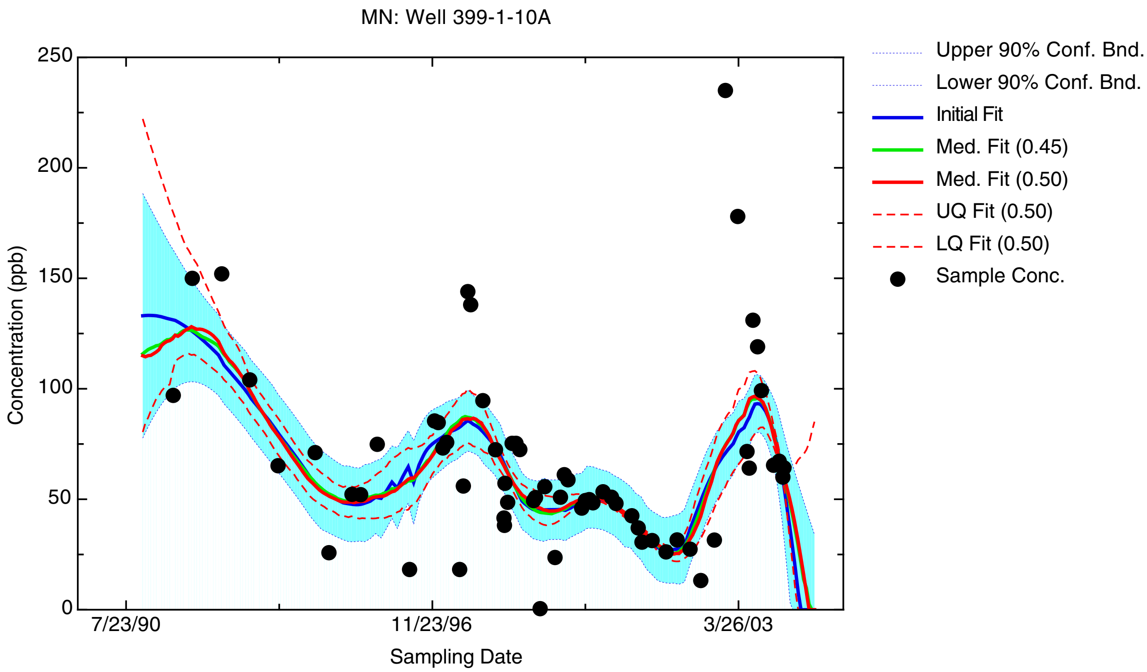
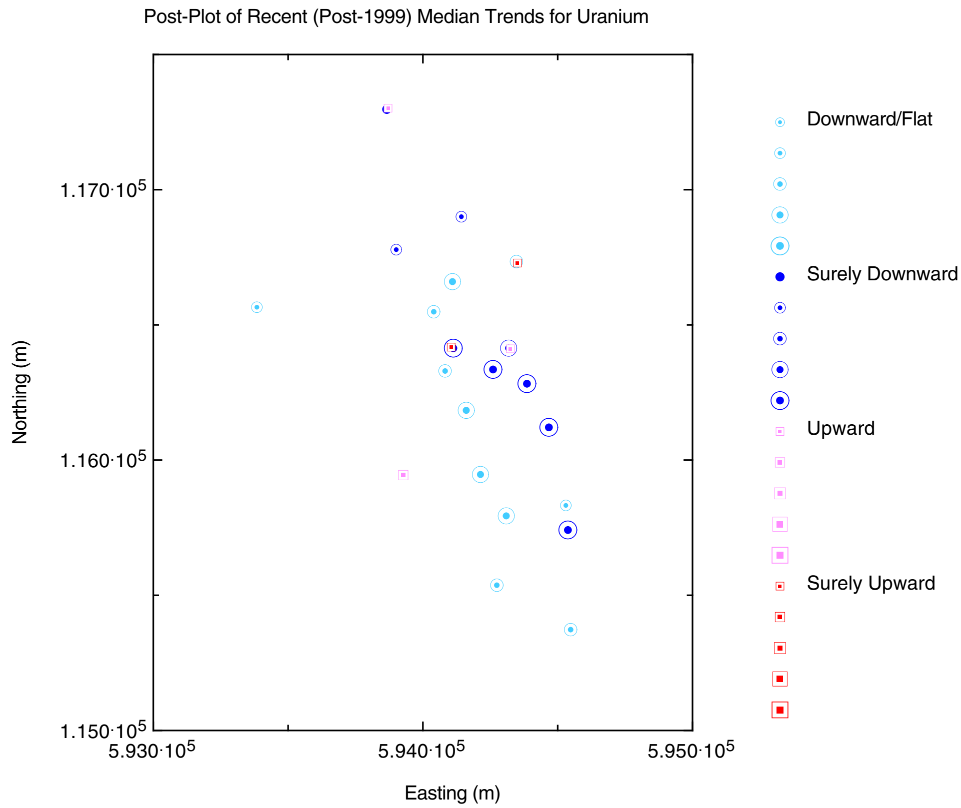


Figure E.4. Recent (post-1999) Trend Map for Uranium at the 300-FF-5 OU

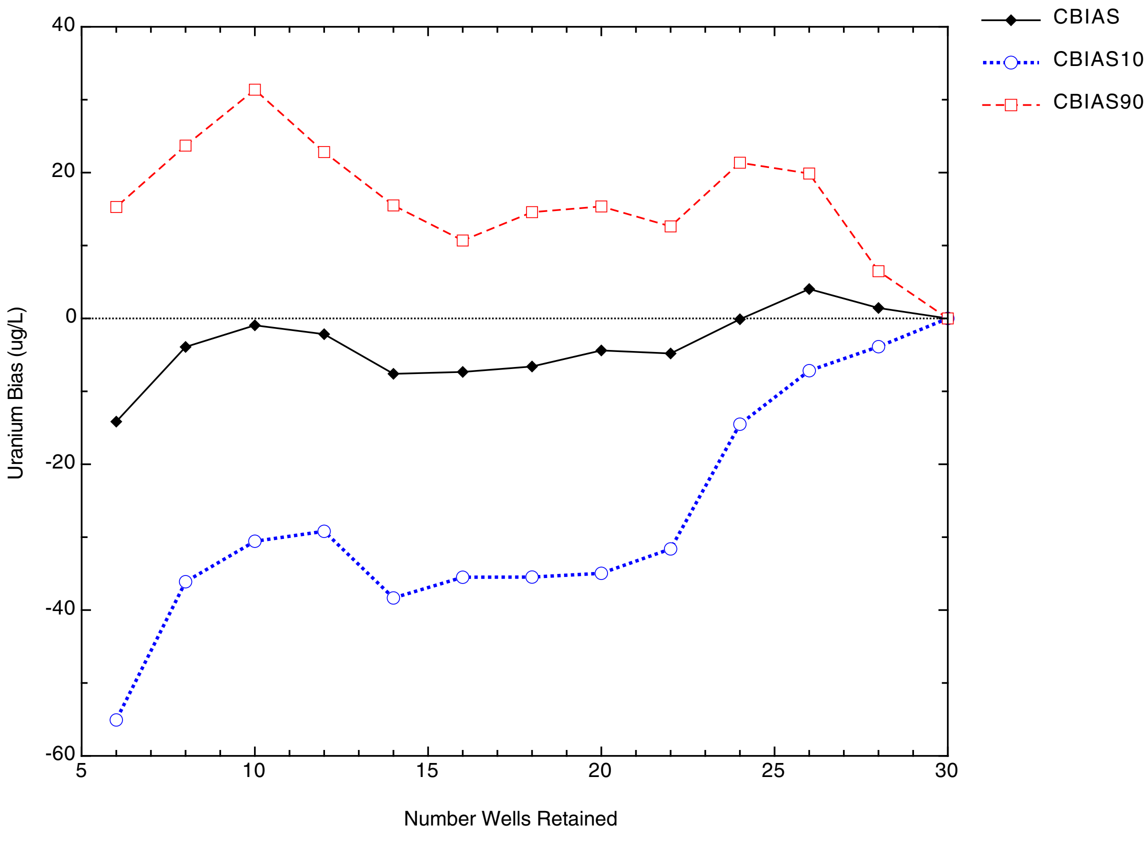


E.3 Spatial Redundancy Analysis

The spatial component of GTS is based on an idea similar to iterative thinning. First, construct a base map of the site using all available wells. Then remove selected wells (e.g., a small percentage) and reconstruct the map to determine how much it has changed. Stop removing wells when the map deteriorates significantly compared to the base map, and propose the reduced well set as the optimal monitoring network for that site.

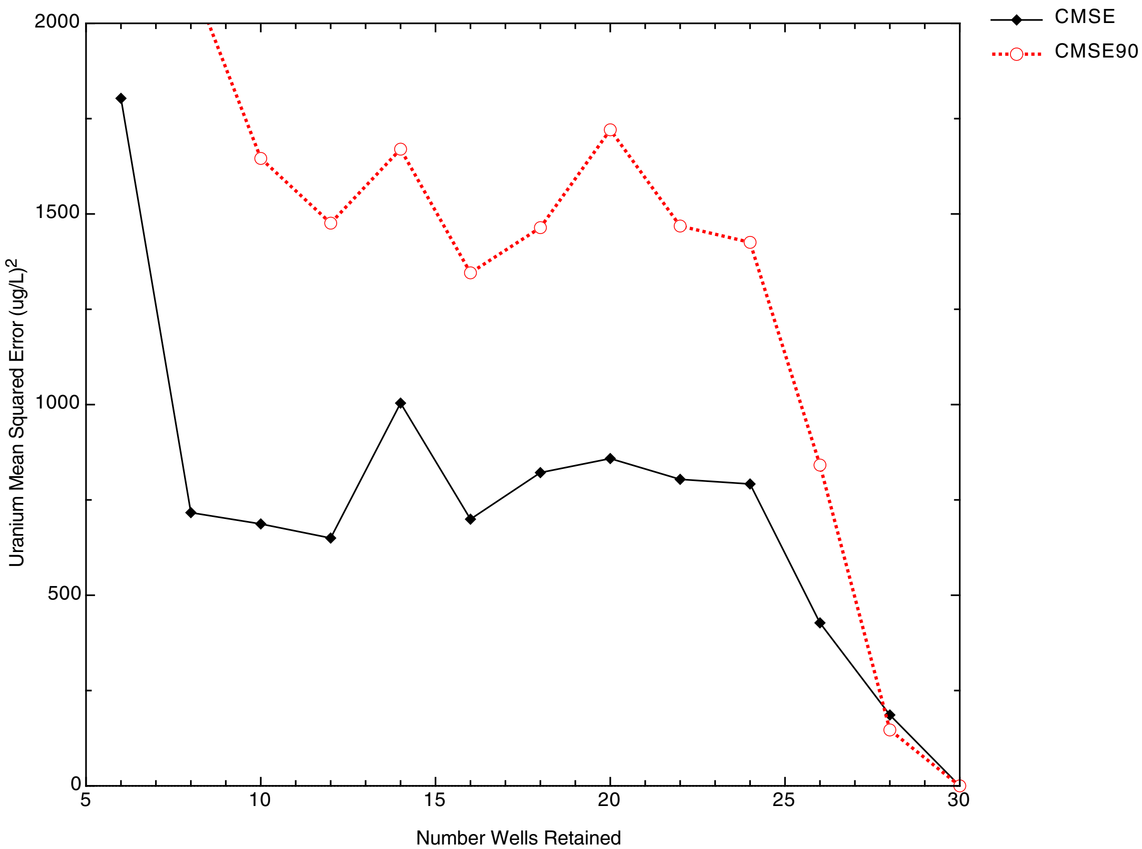
Since a loss of statistical information is inevitable when wells are removed from the network, GTS tracks the tradeoff between the cost of the reduced network and the loss of accuracy in subsequent maps. Selected measures of this cost-accuracy tradeoff are presented in Figures E.5 through E.8. The key statistical indicators are bias and uncertainty, bias reflected through the average measurement differences displayed between the reduced-data and base maps, and uncertainty reflected through the relative mean squared error.

Figure E.5. Uranium Bias Tradeoff, 1999-2001 Data



The bias trace in black (CBIAS) in Figure E.5 denotes the average pixel-by-pixel difference in concentrations between the reduced-data and base maps. CBIAS10 and CBIAS90 denote the 10th and 90th percentiles respectively of the distribution of local bias values. It can be seen that the bias hovers near zero until more than six wells are removed (reading from right to left). After this point, the level of extreme underestimation (denoted by the lower 10^{th} percentile of the bias distribution) drops significantly.

Figure E.6. Uranium Mean Squared Error Tradeoff, 1999-2001 Data



In Figure E.6 CMSE denotes the mean squared relative error between the pixel-by-pixel estimates of the reduced-data and base maps. CMSE90 denotes the upper 90th percentile of the distribution of local relative squared differences. Of note, the mean squared error rises rapidly with any data reduction, but then 'levels out' across a broad range of data reductions.

Figure E.7 illustrates the uranium bias tradeoff for data from 2002-2004. The most significant difference from Figure 5 is that instead of average bias in concentration, this measure indicates average bias in *cumulative probability* between the reduced-data and base maps, the cumulative probability associated with the univariate distribution of observed uranium concentration values in the 300-FF-5 OU over the period analyzed. The advantage of also considering probability level bias is that the unit-less scaling of the measure stays constant across time slices, chemical parameters, and/or sites. The term 'indicator' refers to the fact that multiple indicator local regression (MILR) computes cumulative probability estimates at each of a series of increasing indicator level variables prior to re-assembling the estimates into a conditional cumulative probability distribution (CCDF). This CCDF is used in turn to compute a final concentration estimate.

Figure E.7. Uranium Indicator Bias Tradeoff, 2002-2004 Data

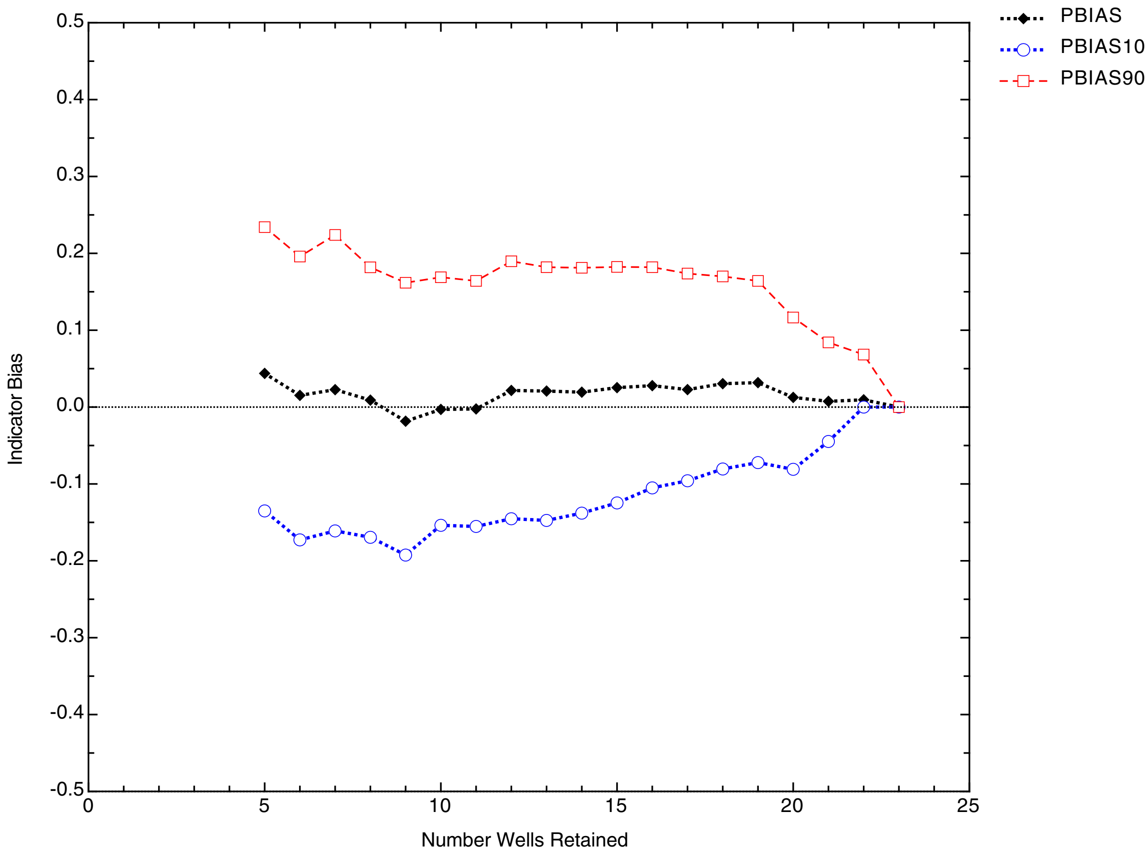
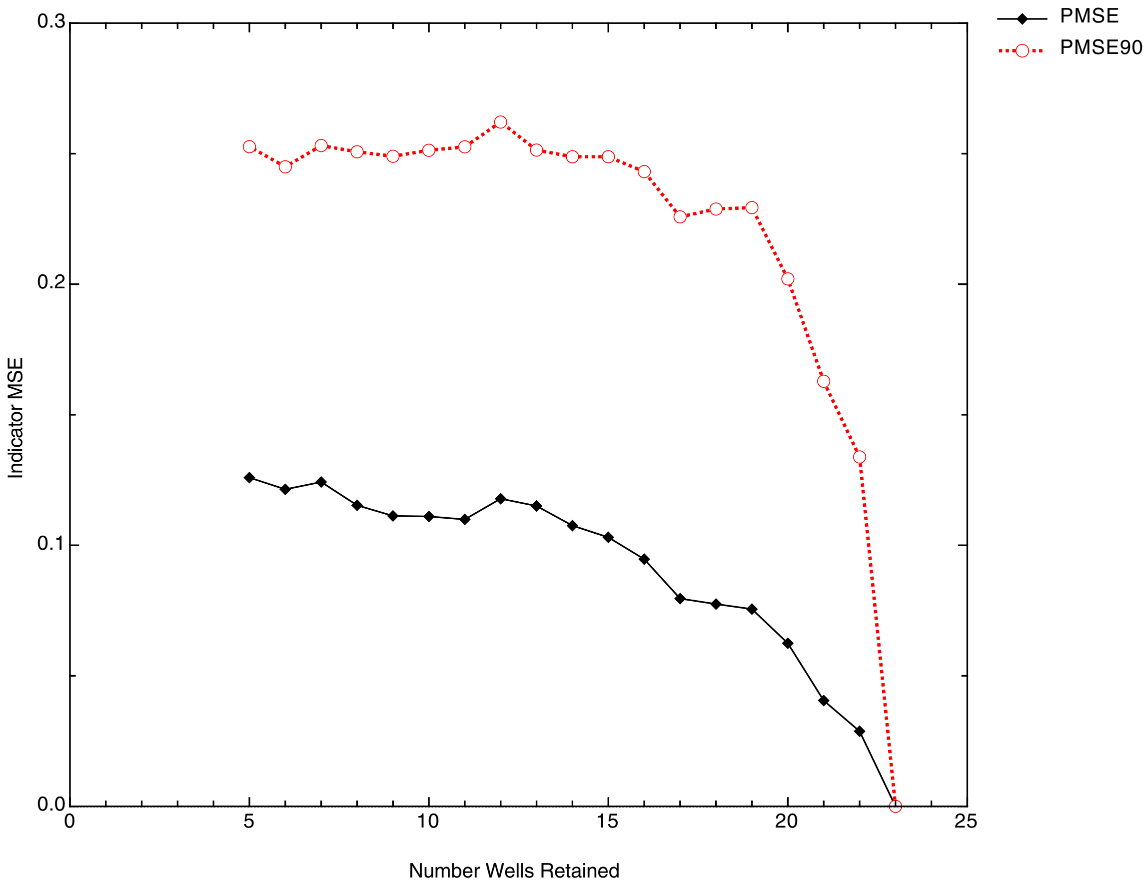


Figure E.8 is similar to Figure E.6, only recorded on the cumulative probability scale. Like the results for the first time slice, the mean squared error rises quickly with any level of data removal, with a lessening slope as more wells are removed (reading right to left).

The key benefits to the GTS approach to spatial redundancy include: 1) a highly data-driven methodology, and 2) the ability to estimate site maps using multiple indicator local regression (MILR) instead of kriging. This second benefit means one can essentially sidestep the effort and art involved in developing a spatial correlation model. The use of local regression also allows for multiple data values at a single location, for instance by including all of the sampling results from a well over a specified range of time, rather than having to perhaps average them when performing kriging. Often, as with the 300-FF-5 OU, the wells in the network are not sampled at precisely the same times. To ensure enough spatial coverage in the data for an estimate of spatial redundancy, sampling events from an interval of time (called a 'time slice' in GTS) must be included.

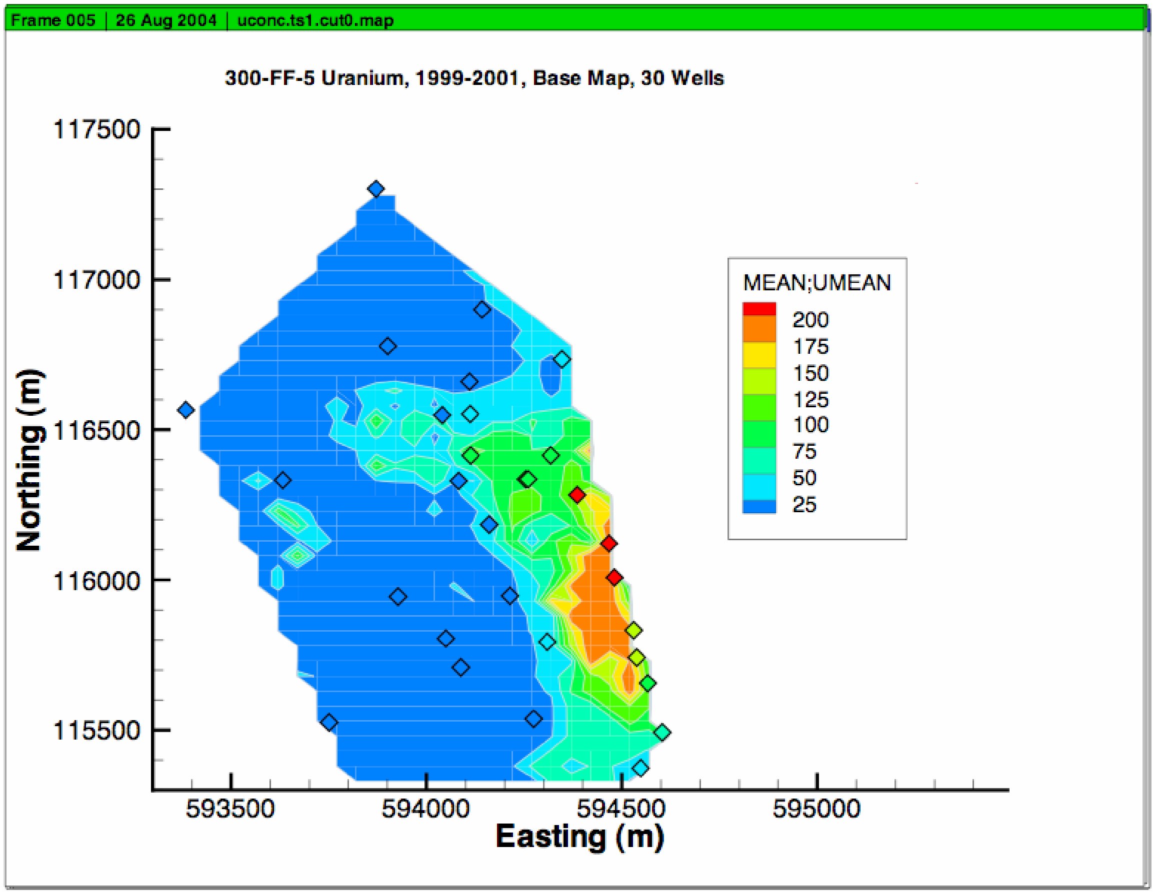
Figure E.8. Uranium Indicator Mean Squared Error Tradeoff, 2002-2004 Data



At the 300-FF-5 OU, there is also the added complexity of seasonal tidal fluctuations near the river and the impact of these tidal fluctuations on measured uranium levels. One of two practical approaches might be taken to deal with this factor. First, redundancy could be estimated at perhaps a 'low flow' period and then again at a 'high flow' period. A second approach was taken in the GTS demonstration, and that was to broaden the length of the time slices to include not only better spatial coverage, but also multiple periods of low and high flow in a single slice. This is akin to estimating a longer-term average within the uranium data. The major potential drawback is if the uranium plume changes substantially over the course of the time slice.

For purposes of the demonstration, two time slices were analyzed, one covering data from 1999-2001 and the other covering data from 2002-2004. Preliminary analysis suggested that the uranium plume was more intense and widespread during the first time slice than the second, an observation corroborated by base maps created for each time slice and presented in Figures E.9 and E.10.

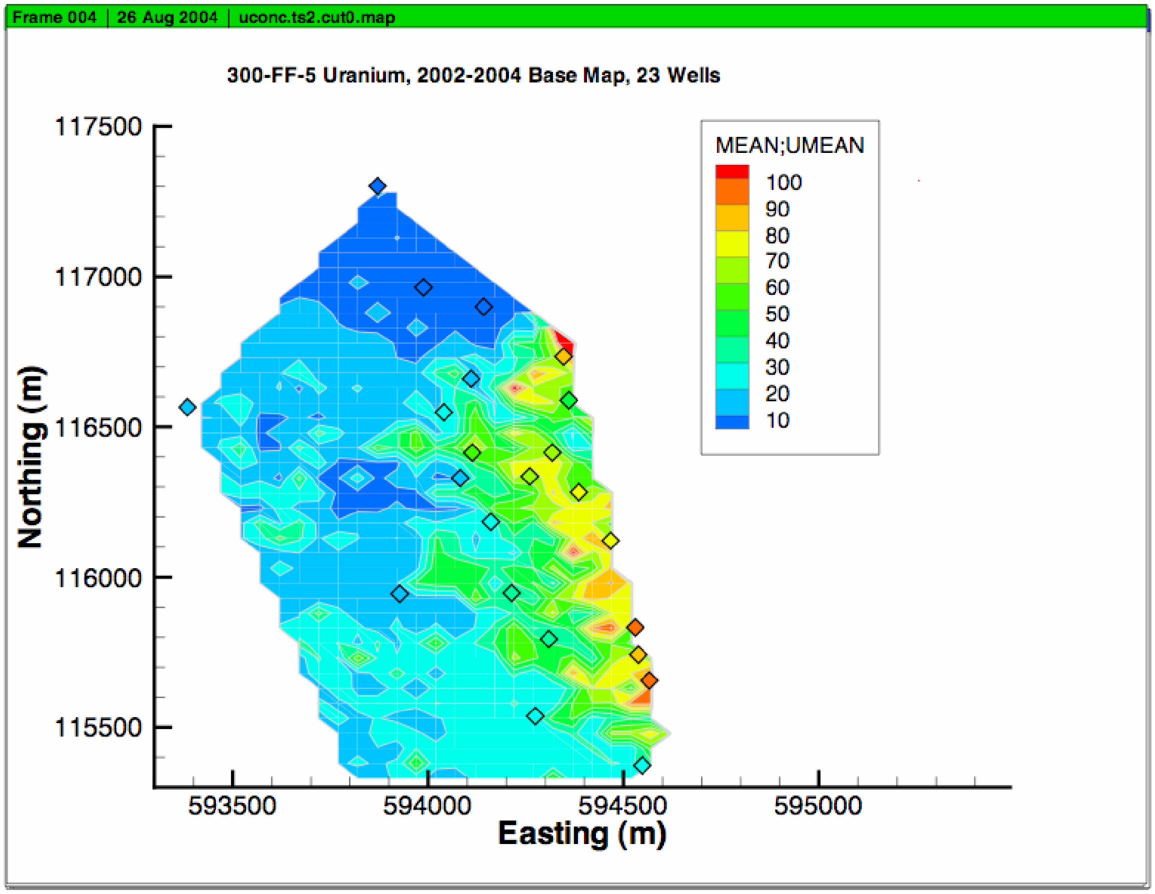
Figure E.9. Uranium Base Concentration Map, 1999-2001



The base maps show in Figures E.9 and E.10 are derived by calculating pixel-by-pixel estimates using MILR. An overlay of the average concentration (over the time slice) at each known well in the base set (filled with the same colors as the contour flood levels of the base map) is shown to visually assess how well the base map matches the observed concentrations. Good agreement between the known data and estimated values is seen in both time slices (see Figure E.10 below). Clearly, the uranium plume (averaged over the 3-year period) is seen to be concentrated near the river, but is more intense in the first time slice than the second.

To actually decide on an optimized monitoring network, the stakeholders must decide on an appropriate balance of cost savings and loss of accuracy. GTS aids in this decision process not only by providing 'global' assessments of the cost-accuracy tradeoff, but also by offering maps of the 'local' bias and uncertainty, from which it can be determined when specific, critical areas of the site are being over- or under-estimated using the reduced-well network, or at what point too many areas of substantial uncertainty have appeared. Selected examples of these maps are provided in Appendix Figures E.11 through E.14.

Figure E.10. Uranium Base Concentration Map, 2002-2004

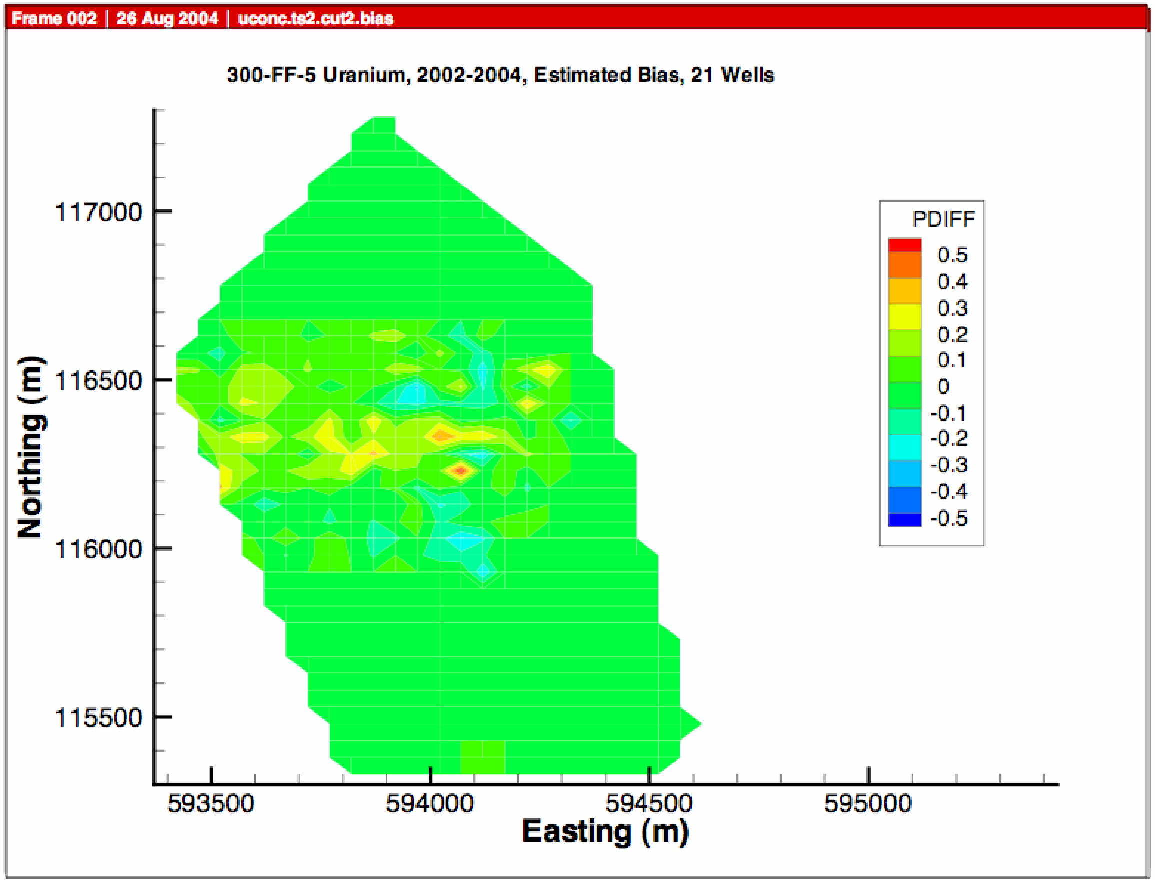


Comparing the local bias maps in Figures E.11 and E.12, it is seen that many more areas of substantial local bias are indicated when only 13 wells are retained from the base set of 23. In particular, a large area of significant relative overestimation is evident toward the northeastern portion of the site, and strong bands of underestimation further to the south. Comparisons of this sort are valuable in determining where areas of bias tend to occur, how frequently, and at what intensity.

A similar comparison of the indicator probability-level local squared relative errors in Figures E.13 and E.14 indicates that substantially more areas of local uncertainty exist when 13 wells are retained compared to when 21 wells are kept. Again, the stakeholders must decide when the level and intensity of the uncertainty rises past a tolerable level. Such a decision may also depend to some degree on *where* the local uncertainty is greatest.

The most significant potential drawbacks to the GTS spatial analysis approach are probably the following: (1) a moderate to larger number of wells is helpful, in addition to decent spatial coverage of the area of interest; (2) wells are selected for removal based on a kind of 'steepest descent' approach using the lowest 'global regression weights' rather than through, for instance, a genetic algorithm selection strategy; and (3) the GTS algorithm is entirely statistical, and does not incorporate any geophysical modeling or make predictions based on a combined statistical and geophysical framework.

Figure E.11. Map of Uranium Local Indicator Bias Values, 2002-2004, 21 Wells Retained



Not all of these are always drawbacks, of course. At sites lacking a detailed geophysical conceptual model, or where the model is highly uncertain, a more highly empirical approach makes sense. If only a small number of wells is available, discussions of spatial redundancy are probably best replaced by questions of data adequacy. GTS is designed for a specific purpose. At Hanford, it ought to be considered as one alternative methodology for assessing various groundwater monitoring networks, provided those networks are truly engaged in long-term monitoring.

E.4 Reference

Cameron, K.C. and P. Hunter [2002]. "Using Spatial Models and Kriging Techniques to Optimize Long-Term Groundwater Monitoring Networks: A Case Study", *Environmetrics* v. 13, 629-656.

Figure E.12. Map of Uranium Local Indicator Bias Values, 2002-2004, 13 Wells Retained.

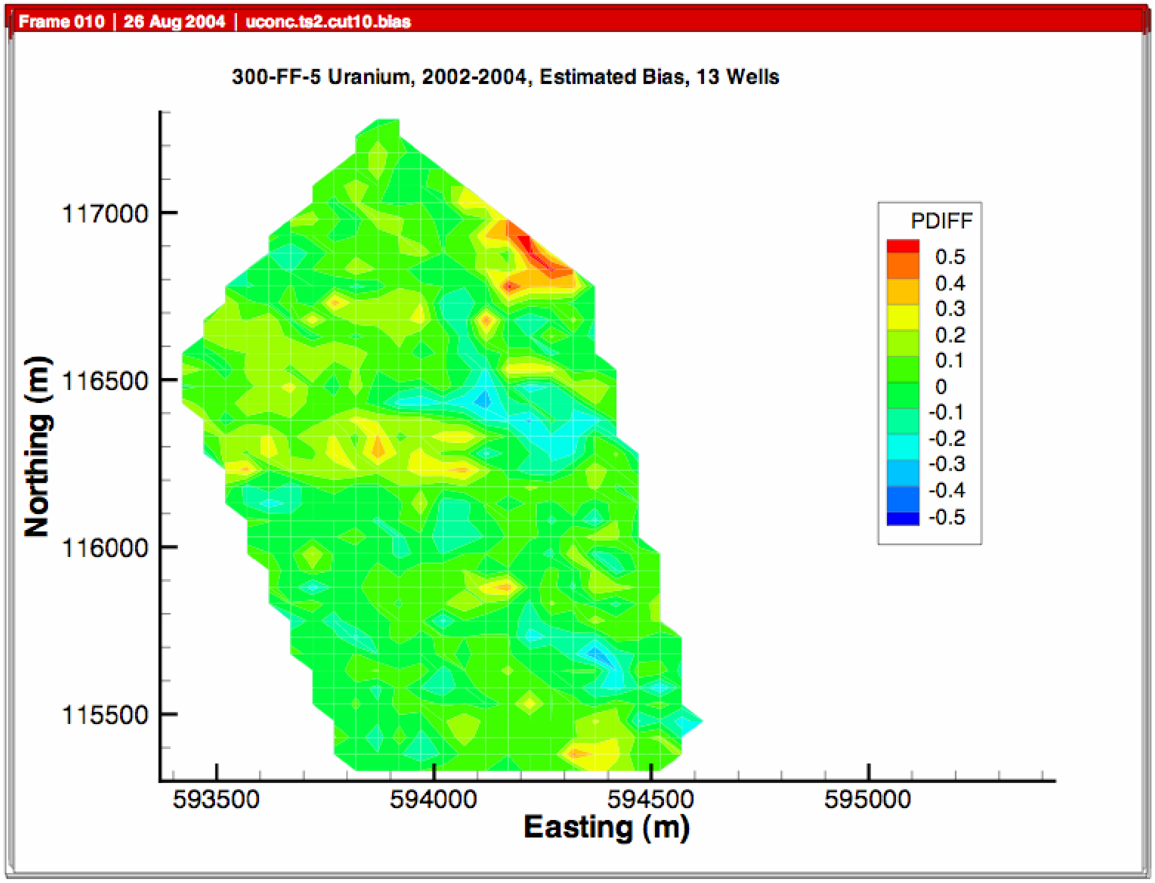


Figure E.13. Map of Uranium Local Indicator Squared Error Values, 2002-2004, 21 Wells Retained

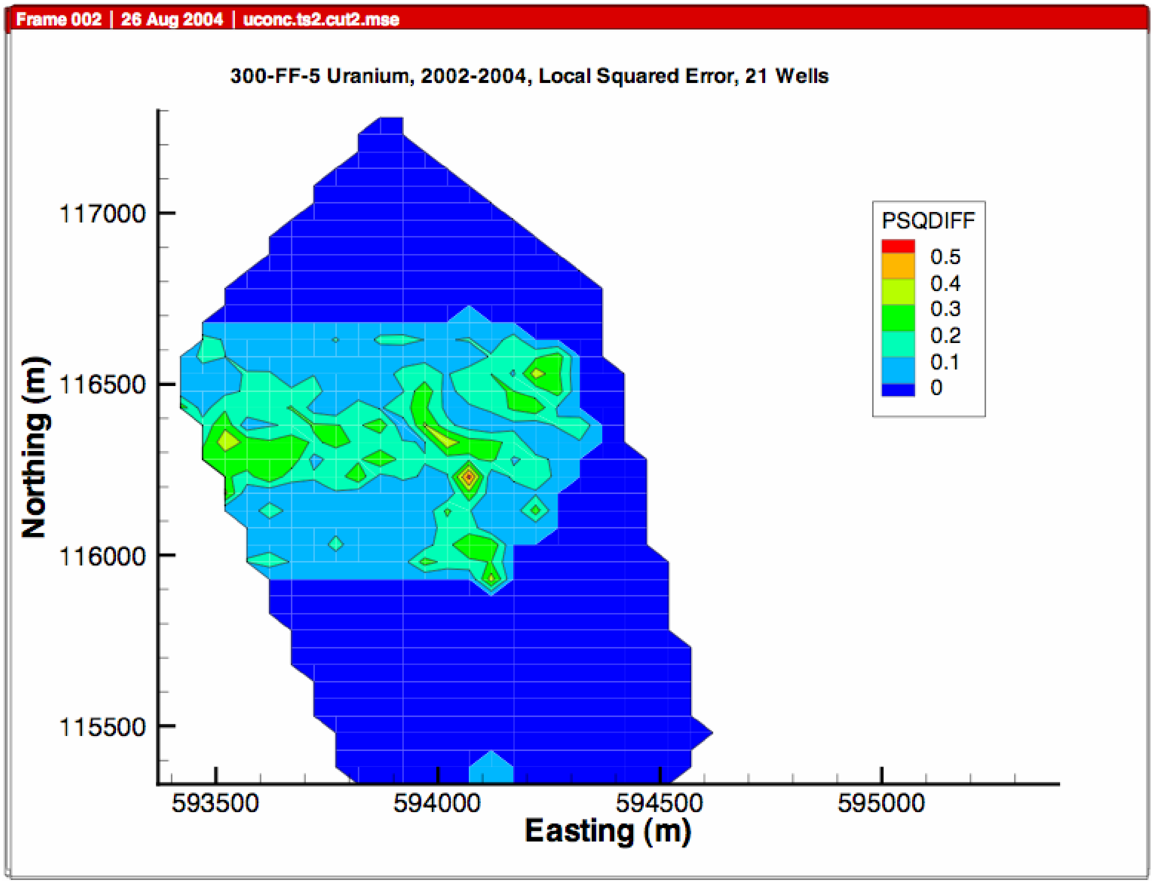
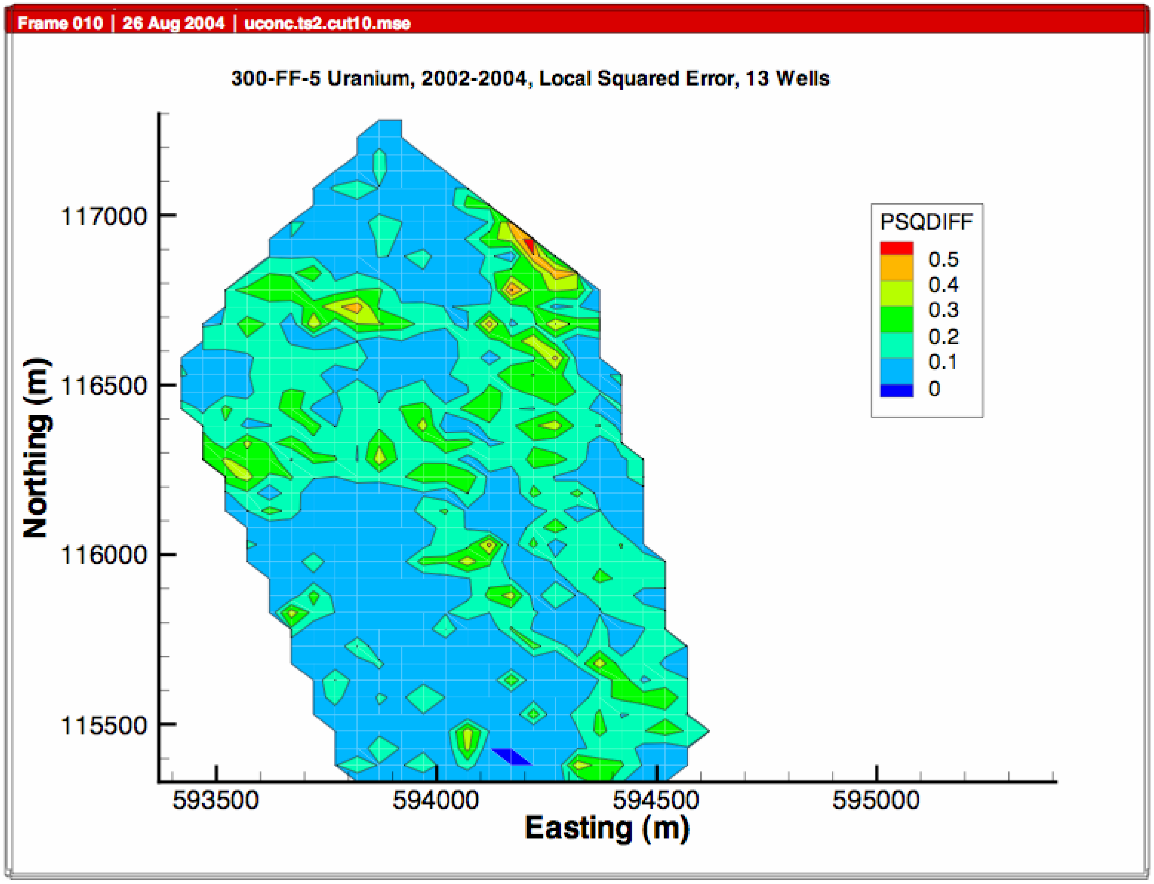


Figure E.14. Map of Uranium Local Indicator Squared Error Values, 2002-2004, 13 Wells Retained



APPENDIX F: DEMONSTRATION OF MULTI-OBJECTIVE OPTIMIZATION AT THE 300-FF-5 OPERABLE UNIT

F.1 Introduction

This appendix gives details on the multi-objective optimization demonstration performed at the 300-FF-5 OU and summarized in Section 2.5.5. This demonstration applied multi-objective optimization to identify spatial redundancy using an existing software package called M-LTMO (Multi-objective Long Term Monitoring Optimizer), developed at the University of Illinois and Moiré Inc.

F.2 Methodology

Figure F.1 illustrates the spatial redundancy optimization approach used in this demonstration. The first step of the approach is to identify a mathematical definition of the long-term monitoring objectives and constraints for the site. In this demonstration, monitoring objectives are assumed to be: (1) minimizing the number of wells sampled and (2) minimizing the maximum error between concentrations estimated using all sampled wells and using a reduced number of wells. The second objective, minimizing the maximum error associated with any particular monitoring design *K* and a particular contaminant of concern (COC) *c*, is represented using the maximum absolute error at any of the removed wells, i.e.,

$$\underset{K}{Minimize}\ Error = \max_{i}\left|c_i^s - c_i^{est}(K)\right|$$

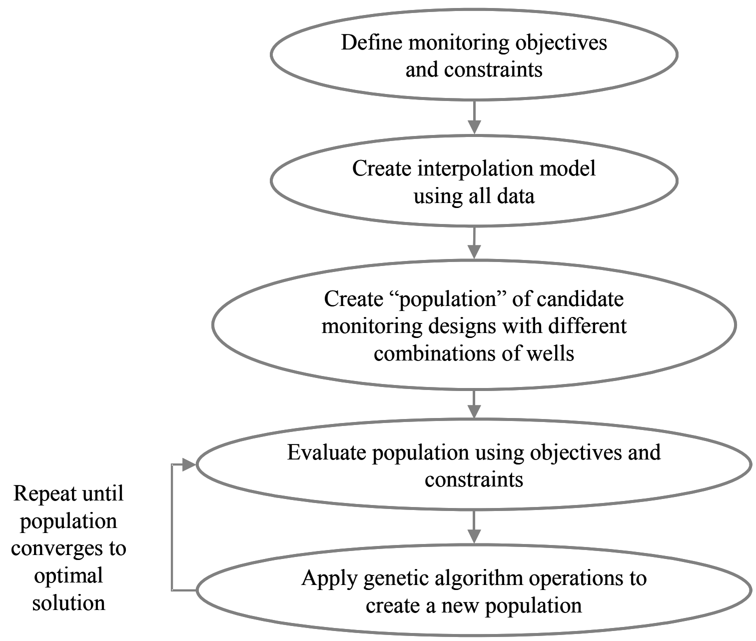
where c_i^s is the sample result concentration at point *i*, and $c_i^{est}(K)$ is the concentration at point *i* estimated using only the monitoring wells included in design *K*. This error metric accounts only for the locations at which measurements are available and yet are not being used in the interpolation model, since those are the only locations where errors can actually be calculated.

The next step of the optimization process, when one of the objectives involves spatial interpolation error, involves creating a suitable spatial interpolation model using data from all wells that were sampled in a recent monitoring period. A spatial interpolation model estimates contaminant concentrations at all locations in the 300-FF-5 OU where concentrations are of interest, using a number of measured data points. Once this interpolation model is created, a redundancy analysis is undertaken in which data from one or more wells are removed from the data set, concentrations are re-interpolated, and the interpolated concentrations at the removed well locations are compared to their true values. If the interpolated concentrations are insignificantly different from the sample results, then the data are redundant and further sampling at those well locations may not be necessary as long as the concentrations are relatively stable in time (i.e., the plume is not migrating substantially).

This type of problem has been tackled at a number of field sites; see EWRI [2003]. In most cases, however, identifying which data could be removed was done by trial-and-error elimination of one or a few wells. Eliminating single wells sequentially, or even several at a time, is not likely to find the best combination of wells that minimizes both sampling error and cost, unless all of the possible combinations of monitoring wells can be enumerated and evaluated. Mathematical optimization can be used to efficiently sort through the many combinations of wells that could be considered and to identify the optimal set of wells to meet the user's objectives and constraints. Moreover, the mathematical optimization approach can identify optimal tradeoffs among multiple objectives, a task that cannot be readily accomplished using trial-and-error approaches to optimization.

The optimization model, a multi-objective genetic algorithm , in the case of M-LTMO, becomes a wrapper around the interpolation model, as shown in Figure D; see Reed *et al*. [2003] and Reed and Minsker [2004] for more details on the genetic algorithm implemented in M-LTMO. When minimizing error is one of the monitoring objectives, the optimization method generates candidate sets of monitoring wells and uses the spatial interpolation model to determine how much error each set creates. Using this information, the mathematical optimization method will ultimately converge on the best set of wells to meet the user's objectives and constraints.

Figure F.1. Overview of Optimization Process



F.3 Demonstration at the 300-FF-5 OU

To demonstrate the usefulness of this approach using data from the 300-FF-5 OU, the method was implemented using data from two recent monitoring periods: June and July 2003 (a high-flow period) and December 2003 and January 2004 (a low-flow period). Multiple sample results within a monitoring period at the same monitoring well were averaged, and then a spatial interpolation model was fit to each period.

Quantile kriging was selected as the spatial interpolation model, as it has been shown in previous work to be highly accurate for plume concentration data compared with many other approaches [Reed *et al*., 2004]. Quantile kriging uses the same procedure as ordinary kriging, except that the concentration data are transformed to quantiles prior to application of kriging, and afterward transformed back to concentrations [Reed *et al*., 2004].

Applying kriging requires that variograms, which model the spatial correlation structure of the data, be fit to the data. Figures F.2a and F.2b show the variograms determined for each period. These variograms show reasonable spatial correlation structure, particularly considering the relatively low number of monitoring wells (22) available in each monitoring period. Note that the variogram values for gamma(h) in the figures are equal to one-half of the difference in uranium concentration *ranking* squared.

Using these variograms, Figures F.3a and F.3b show the resulting interpolated plumes found with quantile kriging for each monitoring period. The plumes in the low- and high-flow periods have similar shapes, but the river water infiltration in the high-flow period causes somewhat lower concentrations. These plumes are used as the basis for evaluating potential monitoring designs in the next section.

Figure F.2a. Spatial variogram for the high-flow period.

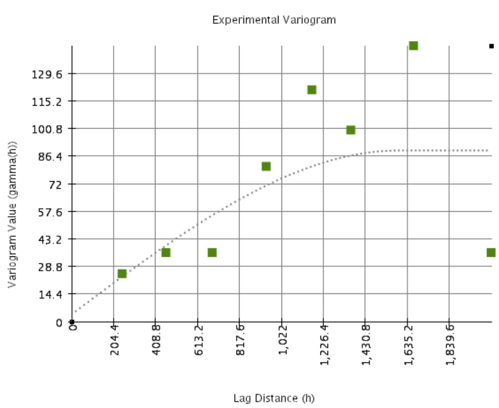


Figure F.2b. Spatial variogram for the low-flow period.

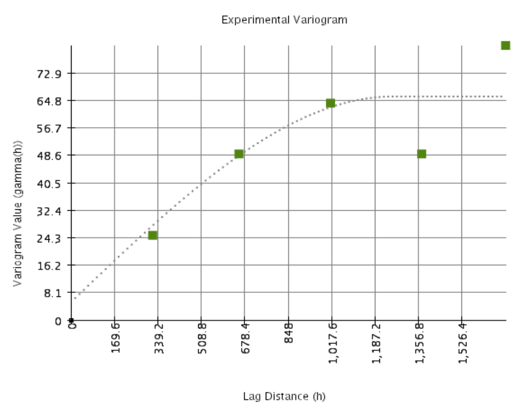


Figure F.3a. Interpolated uranium plume using all 22 sampled wells in the high-flow period.

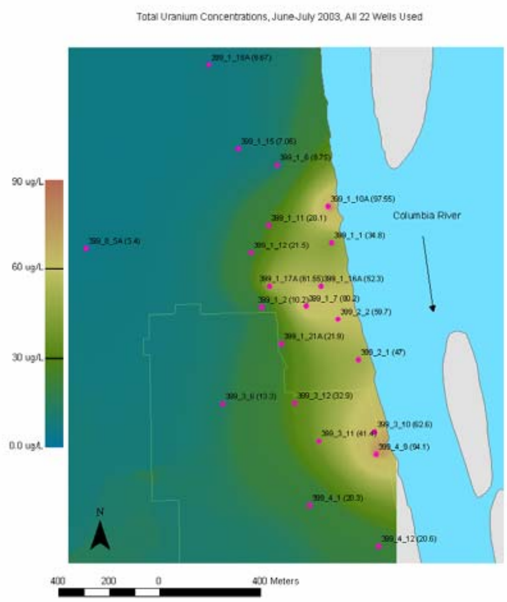
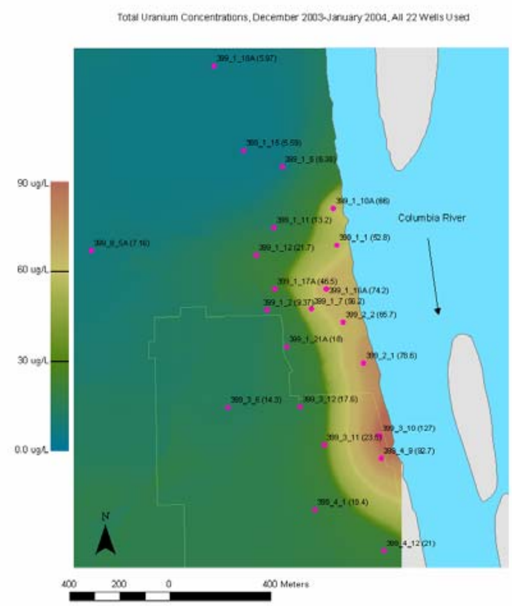


Figure F.3b. Interpolated uranium plume using all 22 sampled wells in the low-flow period.



F.4 Results

Figures F.4a and F.4b show the optimal tradeoffs between number of wells sampled and maximum interpolation error in the high-flow and low-flow monitoring periods, respectively. Each diamond-shaped marker in the figures represents an optimal monitoring design that minimizes errors for a given level of sampling; if increasing the number of wells did not lead to an improved error measure, the marker is omitted. The figures show that the concentration estimation errors rise somewhat more quickly as sampling is reduced during the high-flow period (June/July) than during the low-flow period (Dec/Jan). Compared with the drinking water standard (DWS) for uranium of 30 μg/L and the variability in sample results, errors below 10 μg/L are relatively small and would likely be considered acceptable.

The next step of the analysis is to compare the spatially interpolated plumes for several designs with errors at acceptable levels. In this case, the 14-well and 18-well designs highlighted with white markers in Figures F.4a and F.4b are compared with sampling all wells in each period. Figures F.5a and F.5b compare the interpolated map from the optimal 18-well sampling plan with the map from sampling all wells in the high-flow period. Figures F.5a and F.5b confirm that removing four redundant wells (marked as crosses in Figure F.5a) has little significant effect on the interpolated plume map in the high-flow period. Similarly, Figures F.6a and F.6b confirm that removing four redundant wells during the low-flow period also has little real effect on the interpolated plume map. However, comparing Figures F.5a and F.6a, only two wells are selected as redundant for both periods.

Figure F.4a. Optimal tradeoffs between number of wells sampled and interpolation errors in the high-flow period.

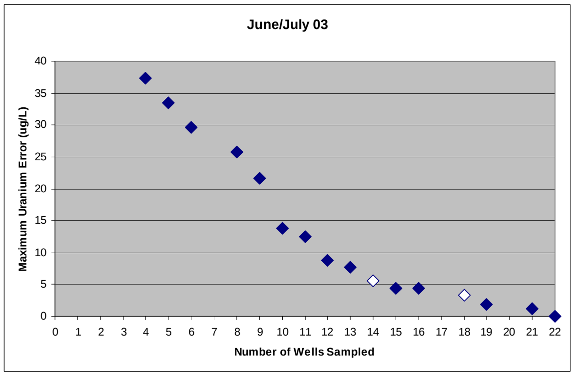


Figure F.4b. Optimal tradeoffs between number of wells sampled and interpolation errors in the low-flow period.

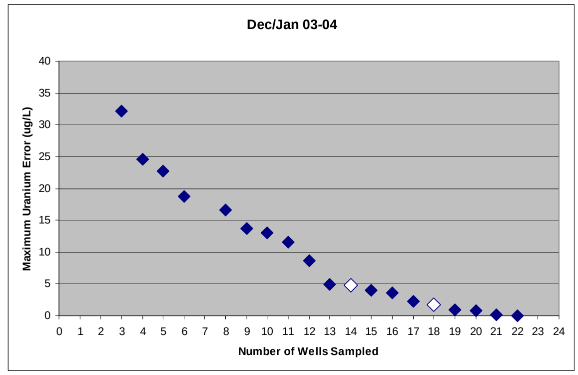


Figure F.5a. Interpolated uranium plume using optimal 18-well sampling design in the high-flow period. Crosses represent redundant wells.

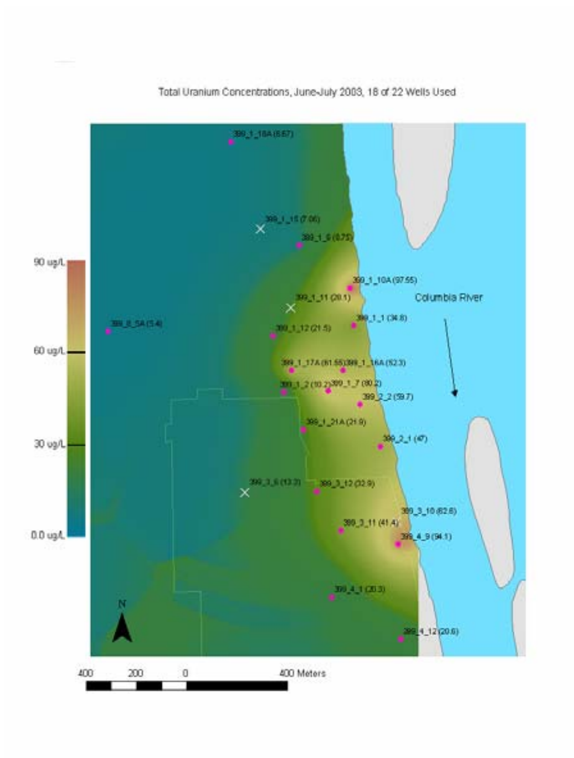


Figure F.5b. Interpolated uranium plume using all 22 sampled wells in the high-flow period.

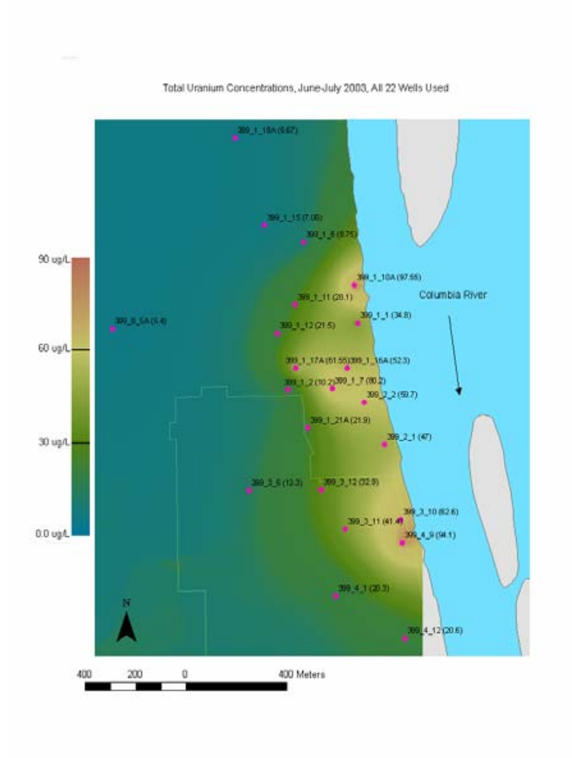
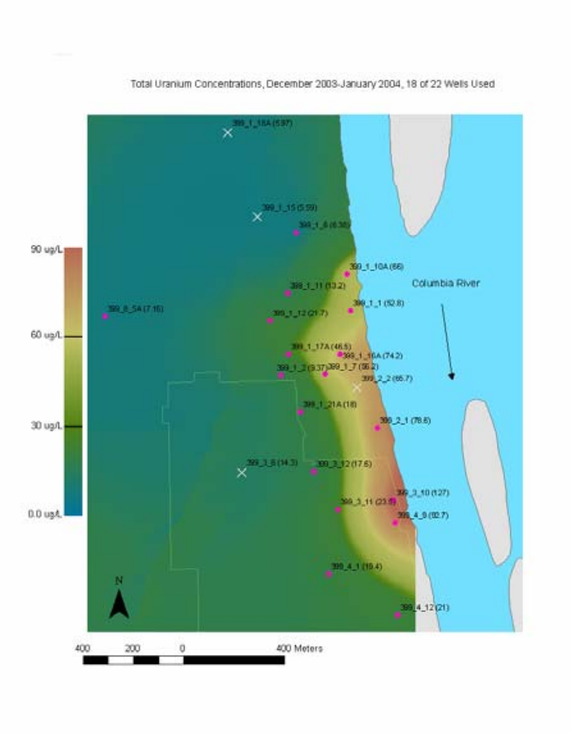
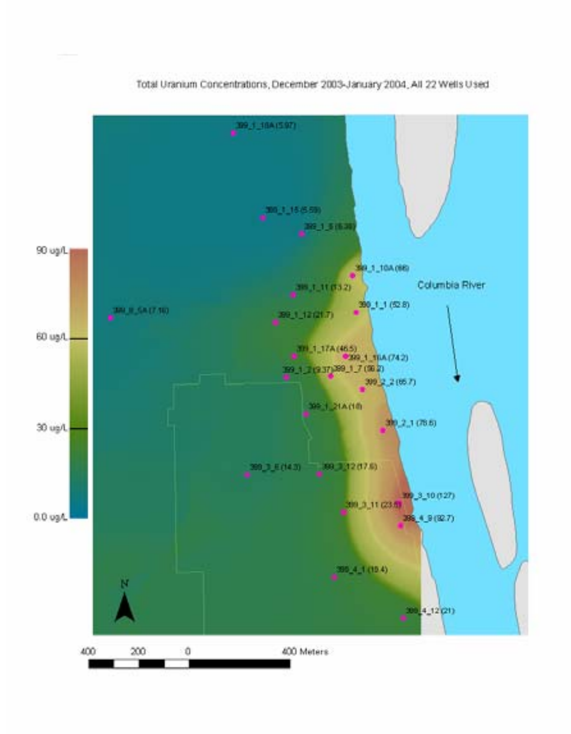


Figure F.6a. Interpolated uranium plume using optimal 18-well sampling design in the low-flow period. Crosses represent redundant wells.

Figure F.6b. Interpolated uranium plume using all 22 sampled wells in the low-flow period.





Figures F.7a and F.7b show the same comparisons for the 14-well solution during the high-flow period. These results indicate that removal of eight redundant wells has somewhat more effect on the interpolated plume map, but the differences are still not substantial compared to the likely variability in sample results. Figures F.8a and F.8b show similar findings for the 14-well solution during the low-flow period. Comparing Figures F.7a and F.8a, note that only three out of eight redundant wells are common to both periods.

F.5 Discussion

This demonstration clearly shows that there is redundancy in the 300- FF-5 OU sampling network, based on the objectives considered here (minimizing the number of samples and minimizing maximum spatial interpolation error). However, the redundant wells identified in the low- and high-flow periods differ significantly, indicating that temporal variability is an important factor in the 300-FF-5 OU and should be considered in any future redundancy optimization efforts in this area. Although M-LTMO could easily be configured to identify sets of monitoring wells that are redundant for both periods, such an analysis was beyond the scope of this demonstration. Similarly, an optimization analysis that simultaneously identifies both optimal sampling frequency and locations would also be possible within M-LTMO.

Figure F.7a. Interpolated uranium plume using optimal 14-well sampling design in the high-flow period. Crosses represent redundant wells.

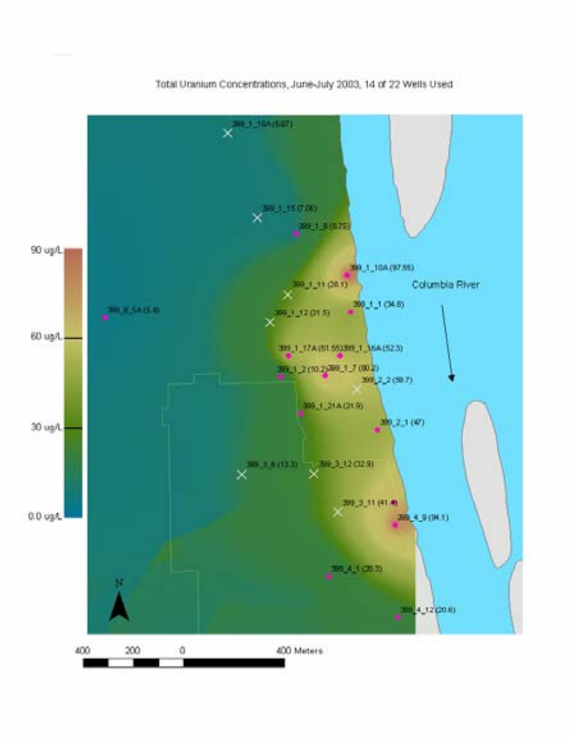
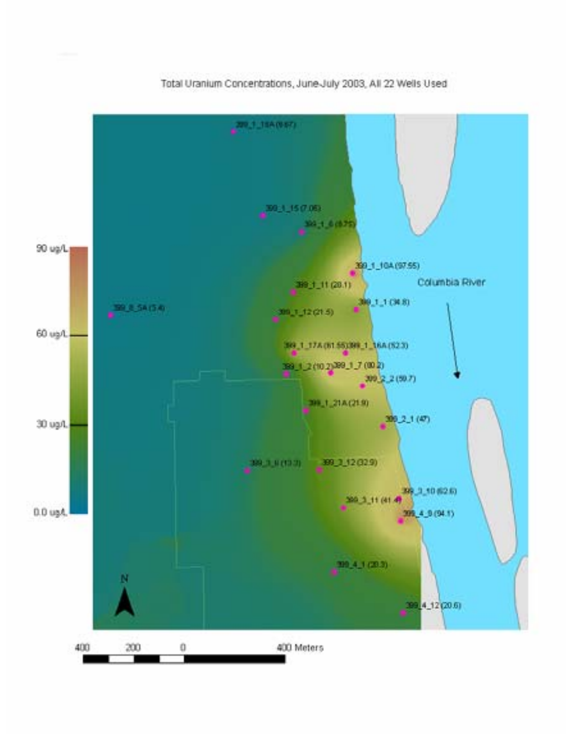


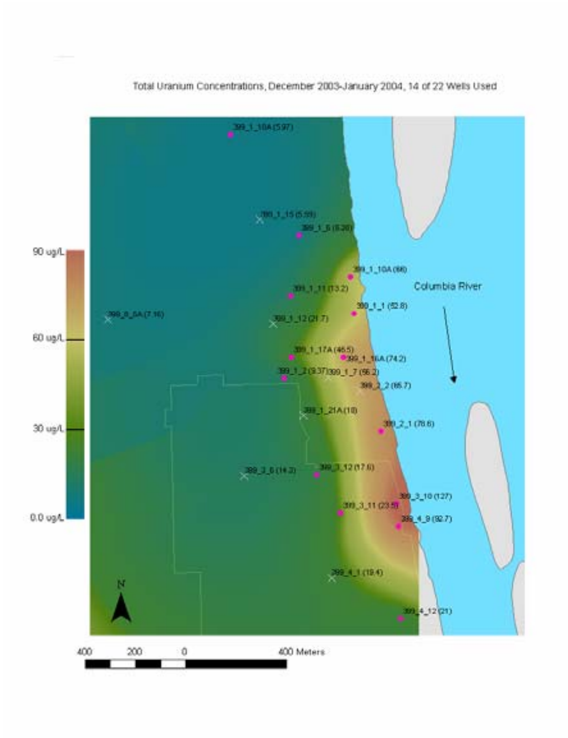
Figure F.7b. Interpolated uranium plume using all 22 sampled wells in the high-flow period.

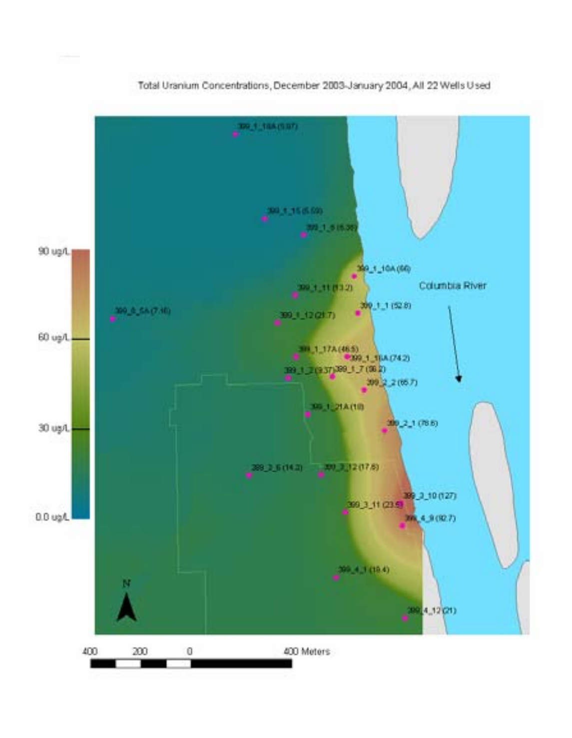


Another factor that should be considered is the uncertainty inherent in the uranium plume maps. The analysis above assumes that the interpolated map obtained using sample results from all of the 22 monitoring wells is the "ground truth" and that any sampling plan that can reproduce the same map is sufficiently accurate. However, any baseline map obtained by interpolation, particularly one based on 22 samples, can have significant error. This error can result from both sampling error in the data used for the interpolation, spatial variability that is not captured by the available data, and spatial structure that may not be fully captured by the fitted variogram model. M-LTMO can also be configured to identify sampling plans that will be robust for multiple baseline plume maps ("realizations") that would be plausible given the observation data and the site-specific uncertainty. Of course, such an analysis can only consider uncertainty that can be quantified using stochastic approaches (such as conditional simulation using the kriged model), expert judgment, or a combination of both approaches.

Figure F.8a. Interpolated uranium plume using optimal 14-well sampling design in the low-flow period. Crosses represent redundant wells.

Figure F.8b. Interpolated uranium plume using all 22 sampled wells in the low-flow period.





F.6 References

EWRI Task Committee on the State of the Art in Long-Term Groundwater Monitoring Design [EWRI, 2003]. *Long-Term Groundwater Monitoring: The State of the Art*, ed. by B.S. Minsker, American Society of Civil Engineers, Reston, VA, 2003.

Reed, P., T. R. Ellsworth, and B. S. Minsker [2004]. "Spatial Interpolation Methods for Nonstationary Plume Data", *Ground Water* v. 42, 190-202.

Reed, P. and B. S. Minsker [2004]. "Striking the Balance: Long Term Groundwater Monitoring Design for Multiple, Conflicting Objectives", *Journal of Water Resources and Planning Management* v. 130, 140-149.

Reed, P., B. S. Minsker, and D. E. Goldberg [2003]. "Simplifying Multiobjective Optimization: An Automated Design Methodology for the Nondominated Sorted Genetic Algorithm-II", *Water Resources Research* v. 39, 1196, doi:10.1029/2002WR001483.

APPENDIX G: RCRA DETECTION MONITORING STATISTICAL TESTS

G.1 Summary

PNNL and DOE/RL have proposed the use of intrawell statistical comparisons based on USEPA's Combined Shewhart-CUSUM Control Charts (CCCs) for use in RCRA detection monitoring programs (Chou, O'Brien, and Barnett [2001]; Chou [2004a]; Chou [2004b]). Statisticians at Washington State University reviewed a version of such comparisons on behalf of Ecology (Jandhyala and Zhang [1999]). Unfortunately, these proposals and reviews fail to take into account certain important aspects of the history and evolution of the relevant regulation and guidance, and also fail to take into account the significant impact that seemingly subtle modifications of the implementation of the CCCs can have on their statistical performance. The net result is that the statistical performance of the procedures recommended by Jandhyala and Zhang [1999] is considerably different from the performance of those recommended by PNNL [Chou, 2004b], particularly those recommended for the 300 Area Process Trenches. The purpose of this Appendix is to discuss the regulatory background of intrawell comparisons and the statistical evaluation of CCC tests. In this context we will also discuss an alternate intrawell statistical test that has met with widespread regulatory approval.

Copies of a relevant technical report [Davis, 1999] and short-course text [Davis, 1998] have been provided to PNNL personnel, along with software used in the investigation reported in the former. A case study related to the monitoring program proposed for the 300 Area Process Trenches follows in this Appendix.

G.2 A Brief History of RCRA Statistical Rules

First, there is the question of whether the unit is regulated under 40 CFR Part 265 (Interim Status) or 40 CFR Part 264 (Permitted or Final Status). Interim status (Part 265) was intended to apply to facilities newly coming under RCRA jurisdiction, but only until they were granted an operating permit. In interim status the rules govern the facility directly. There is essentially only one type of monitoring prescribed, which is a form of detection monitoring. If a facility finds evidence of impact to the groundwater during interim status detection monitoring, it must go into corrective action adequate to meet final closure requirements, or else shift over to the permitted status program. Although the monitoring requirements during closure are very flexible, being negotiated between the facility and the agency, the detection monitoring requirements are extremely rigid and misguided, as explained succinctly by Chou, O'Brien, and Barnett [2001]. The indicator parameters mandated (pH, specific conductance, total organic carbon, and total organic halogen) are not selected based on constituents characteristic of the wastes, and are (except possibly for specific conductance) frequently only the grossest of potential indicators of releases.

Moreover, the interim status statistical rules seem to be designed to guarantee that "statistically significant" differences between upgradient and downgradient wells will be found nearly regardless of the true state of affairs, by ignoring the variance components structure of the measurements of these constituents in real groundwater (involving systematic spatial and temporal variation in addition to sampling and analytical error) as well as by insisting that a 1 percent significance level (false positive rate) be used for each of the 4 (indicator parameters) * (# of downgradient wells, minimum 3) $\geq$ 12 comparisons conducted during each monitoring event. The resulting Facility-Wide False Positive Rate (FWFPR) would therefore be expected to be around 11 percent per event, IF the extremely simplistic model ignoring the variance components were true. In actuality it must be expected to be higher, unless the facility happens to have an upgradient well whose pH is about the same as that of every other well and whose other indicator parameters have upgradient values uniformly higher than downgradient. In that happy or unhappy circumstance (depending on one's perspective), the monitoring program would have little chance of finding real releases, even those that might affect the latter three indicator parameters substantially.

USEPA recognized the problems with the statistical requirements in its *Technical Enforcement Guidance Document* (*TEGD*) [USEPA, 1986]. Although the *TEGD* did provide some relief along the FWFPR lines, by allowing a Bonferroni *t*-test with a rather lower significance level per comparison, and recognized some of the variance components structure through allowing the Averaged Replicate *t*-test, the result was still extremely insensitive to any potential releases, and still handled inherent spatial variation inadequately.

Meanwhile, in 1982 USEPA promulgated statistical rules in 40 CFR Part 264 for Final Status facilities. These rules provided for three stages of monitoring, presented in this order: compliance monitoring (when hazardous constituents have been found in the groundwater); corrective action (when those constituents exceed a groundwater protection standard); and detection monitoring (otherwise). The facility could propose detection monitoring indicator parameters that would provide a "reliable indication" of incipient contamination, a considerable improvement over the Part 265 situation. The statistical test required, so long as a rudimentary test of distributional normality was passed, was the Cochran's Approximation to the Behrens-Fisher (CABF) *t*-test. In mandating this test, USEAP recognized that upgradient (or background) and downgradient data might routinely have different variabilities, which is an implicit acknowledgement of the variance-components structure of groundwater monitoring data for naturally occurring indicator parameters. One wrinkle was that, if a statistical test failed the comparison, one should resample and repeat the test. A "statistically significant" exceedence was not declared unless both comparisons failed. The test was to be performed using a 5 percent significance level for each comparison. The actual significance level of this process explicitly incorporating (correlated) resample comparisons was unclear at the time; Davis and McNichols [1987] developed algorithms for computing significance levels and powers for statistical tests of this nature taking resampling plans into account. The Davis-McNichols algorithms, however, assume that initial samples and resamples are statistically independent, which is not the case where upgradient-downgradient comparisons are used and inherent spatial variation is present.

The weaknesses of that state of affairs were soon recognized, and USEPA provided revised Part 264 rules in 1988, followed shortly by the *Interim Final Guidance* [USEPA, 1989]. The Combined Shewhart-CUSUM Control Chart approach to monitoring made its first appearance in the 1988 rules and 1989 guidance. The *Addendum to Interim Final Guidance* followed [USEPA, 1992]. One of the advances in the *Addendum* is the explicit recognition of the legitimacy of controlling the FWFPR.

In the revised Part 264 permitted (final) status rules there are still three programs possible. If contaminants have been detected in groundwater (statistically significantly) above a groundwater protection standard, the regulated unit should be in corrective action. If they have been detected, but are not (statistically significantly) above a groundwater protection standard (including possibly a negotiated alternate concentration limit), the regulated unit should be in compliance monitoring. The monitoring requirements specified for compliance monitoring are statistical comparisons with the groundwater protection standard.

Otherwise, the regulated unit should be in detection monitoring. In detection monitoring under Part 264 one collects background and monitoring data and performs statistical comparisons. Several statistical tests are specifically mentioned in the regulation. Those mentioned first are parametric and nonparametric Analyses of Variance, followed by downgradient-upgradient comparisons if the ANOVA result is statistically significant, which essentially equates statistically significant spatial variation (with downgradient higher than upgradient) with *prima facie* evidence of releases. Also specifically mentioned are prediction or tolerance limits, which come in both an upgradient-downgradient version (which also equates spatial differences with evidence of releases) and an intrawell version (which compares data from a given well with its own historical background). Finally control charts are mentioned; these are also inherently intrawell.

The idea of equating statistically significant evidence of spatial variation with *prima facie* evidence of releases has lost favor in USEPA [1992] for good reason (see [Davis, 1998]). When used with constituents or indicator parameters whose values vary inherently among wells, one of the following will happen: (a) the test will repeatedly "find" the inherent variation, resulting in cycling through detection monitoring to assessment to detection monitoring; (b) the test will be insensitive to real impacts due to releases; or possibly (c) both with different wells in the same regulated unit. Accordingly, recent USEPA guidance [1992] as well as the forthcoming *Unified Guidance* (*UG*, currently in preparation) favor statistical methods based on intrawell comparisons in a detection monitoring program, so long as one is convinced that the spatial variation is not due to prior releases from the regulated unit. The intrawell comparisons then are geared toward finding deviations from the current, acceptable state of affairs.

The preceding remarks apply for constituents that exhibit statistically significant inherent spatial variation, which is nearly always present with naturally occurring constituents. The exceptions are trace constituents, for which the data distributions tend to be dominated by nondetects. For these nonparametric statistical tests are often appropriate. Nonparametric tests generally require larger datasets; on the other hand, the spatial variation demanding intrawell comparisons is generally lacking, so it may make sense to use the nonparametric analyses of variance or nonparametric prediction limits comparing monitoring measurements with pooled upgradient (or all historical) background data.

G.3 When Intrawell Statistical Comparisons Are Appropriate

Intrawell comparisons are allowed for a RCRA final (permitted) status facility in a detection monitoring program. The operative Federal regulations are contained in 40 CFR Part 264. §264.97(h) allows specific statistical tests that are inherently intrawell, and §264.97(i)(6) requires that "if necessary" the statistical method should control for spatial variation. As outlined in §264.91(a), detection monitoring status is appropriate when releases of hazardous constituents from the particular regulated unit have not been detected at the compliance boundary. Accordingly, if a facility is in detection monitoring status, one assumes that any spatial variation that exists is due to sources other than that regulated unit. Those sources may include natural variation or other influences.

As a practical matter, when instituting a RCRA detection monitoring program at other than a "green-field" site, one should perform a preliminary evaluation of the existing groundwater conditions. Unless that preliminary evaluation concludes that either (a) corrective action is needed, per §264.97(a)(2) or §264.97(a)(3), or (b) hazardous constituents are present and therefore a compliance monitoring program should be instituted, per §264.97(a)(1), the regulated unit should be deemed to be in detection monitoring status, and therefore any existing spatial variation should be deemed to be inherent, or at least an acceptable state of affairs, as discussed by Davis [1997]. In such situations the goal of a "detection" monitoring program should be to detect changes from that acceptable state of affairs.

Two conditions are laid out in §264.97(a)(1)(i) for deciding that wells other than hydraulically upgradient wells may be used to provide background data for statistical tests. One is the case where one cannot adequately determine the hydraulic gradient. The other is where wells other than upgradient wells provide an indication of background groundwater quality that is at least as representative of what future monitoring data should be like as that which would be obtained from upgradient wells. In the presence of statistically significant spatial variation, and in view of the discussion of the previous paragraph, one can argue that the most appropriate background data for any well is previous data from the same well, once the decision has been made that a detection monitoring program is appropriate.

Attempts to control for inherent spatial variation other than by using intrawell comparisons, if they are statistically valid, tend to be more complicated and less certain of success. Moreover, if they are successful in controlling for false positives due simply to inherent spatial variation, they can be much less sensitive to real changes than intrawell statistical tests; see Davis [1998].

G.4 Regulatory Subtleties

Two intrawell statistical tests are specifically approved in RCRA regulation and widely accepted in guidance and practice: intrawell prediction limits and control charts. Intrawell tests had come to the attention of USEPA between the time of the advance notice of the proposed revision of 40 CFR Part 264 and the promulgation in 1988 of that revision, as had the concept of FWFPR control. It is instructive to compare the discussion of the relevant issues in Davis and McNichols [1988], written during that interval, with §264.97(i)(2) of the revised rule. For example, the latter requires that in general individual comparisons (i.e., statistical tests for a single constituent at a single well) should have a 1 percent significance level, but specifically excludes control charts and prediction limits from that requirement.

The legitimacy of FWFPR control was firmly established in USEPA [1992]. That document allows for the FWFPR of the entire monitoring program for a unit in RCRA detection monitoring to be held to 5 percent, so long as certain power comparison criteria are met for the sensitivity of the tests to increases or changes in any individual constituent at any individual well. Following this principle, one evaluates false positive rates on a program-wide basis, but evaluates statistical power on a one-constituent, one-well basis.

USEPA [1992] proposed a set of reference power curves, which are the power curves for ordinary 1 percent significance level *t*-tests using $n = 8, 16, 24,$ or 32. These curves give the probability of obtaining a "statistically significant" measurement as a function of the increase in mean, measured in standard deviation units (sigmas) as is customary. The idea is that an alternate statistical test should be acceptable so long as it is as sensitive to releases as the (presumably automatically acceptable) traditional *t*-test. Of course, in order to attain decent FWFPR control, the power at zero shift in mean will have to be far lower than 1 percent, so this requirement means that the power curve for the alternate test will have to cross the reference power curve at some point. The comparison is successful if the crossing occurs by the point that the reference *t*-test itself has attained some reasonable power, such as (say) 50 percent.

It was pointed out to USEPA that keying the reference power curve to the anticipated background sample size (*n* = 8, 16, 24, or 32, e.g.), thereby making a more stringent standard for larger sample sizes, would effectively discourage facilities planning their background sampling programs from obtaining larger amounts of background data that might provide for better background characterization. USEPA agreed, and the 1 percent significance level *t*-test with *n* = 8 has been widely used as the reference standard since then (see Davis and McNichols [1994b]). More recently, USEPA has been preparing a *Unified Guidance* (*UG*) that subsumes and improves upon previous guidance (USEPA [1989] and USEPA [1992]). The *UG* is still in preparation; along the way there has been some discussion of changing the reference standard to the 1 percent significance level *t*-test with *n* = 10. Some implications of this possible change are discussed in Davis [1999]. There has also been some discussion of changing the 5 percent FWFPR target per monitoring event to 10 percent ÷ (# events per year), to achieve greater parity between facilities with quarterly monitoring, for example, and those with only annual monitoring.

Finally, one should point out that the requirement for numbers of samples in §264.97(g) does not distinguish between background samples and future monitoring samples in an intrawell setting. The default of four samples per monitoring event stated in §264.97(g)(1) applies most clearly when the selected statistical test is the parametric or nonparametric ANOVA of §264.97(h)(1) or §264.97(h)(2). When USEPA promulgated this regulation it believed that the use of "one point in time" ANOVA procedures using data from only the current monitoring period would be an effective way of controlling for systematic temporal variation. It turns out, however, that inherent spatial variation is nearly always a much greater concern than systematic temporal variation, and using such "one point in time" ANOVAs will serve only to repeatedly confirm that such spatial variation is present, providing little or no sensitivity to changes in the patterns. USEPA [1992] explicitly discourages the use of ANOVA procedures in detection monitoring; see also Davis [1998].

When considering intrawell tests, then, the operative parts of regulation regarding sample sizes should be the language in §264.97(g) about "generally accepted statistical principles", for the background dataset, and §264.97(g)(2) "an alternate sampling procedure as approved" for monitoring post the background period. Specifically, in determining post-background sampling frequency one should take into account both groundwater flow rates as well as other information. The language in the detection monitoring section §264.98(d) is more prescriptive about requiring four "independent" observations per monitoring event. To circumvent this, one should remember that this regulation is directed toward the agency's permit writers, and that the Regional Administrator has considerable flexibility in its implementation. In addition, one should recall that between the promulgation of this regulation and that of the Solid Waste rules two years later in 1991 (40 CFR 258), nearly all of the statistical rules were unchanged, except notably for that of the sampling frequency requirement during detection monitoring. §258.54(b) requires that during detection monitoring at least four observations per well should be obtained during the first semiannual sampling event, and at least one sample per well per event thereafter.

In that 1991 regulation there is also the explicit allowance for variance from this prescriptive requirement. Cameron's [1996] article "RCRA Leapfrog: How Statistics Shape and In Turn Are Shaped by Regulatory Mandates" provides further insight into the continuing evolution of RCRA groundwater monitoring statistical regulation and guidance; see also Davis ([1994a] and [1998]).

G.5 Available Intrawell Tests

The most straightforward intrawell statistical tests are based on intrawell prediction limits. Using background data one constructs a prediction limit PL = (mean) + K (standard deviation) for each constituent for each well. If a future observation exceeds PL (and the verification resampling plan is exhausted, if one is used), a statistically significant exceedence is declared and the provisions of §264.98(g) kick in. The PLs are computed on the original or transformed scale, as appropriate, but expressed on the original scale.

The statistical properties of a PL procedure depend on the multiplier K, the background sample size, the background degrees of freedom, and the resampling plan allowed. Davis and McNichols [1994b] discuss various resampling plans, of which only the "1 of 2" plan is of practical interest for the Hanford situation because of the limited numbers of constituents and wells per regulated unit. The "1 of 2" designation means that the well-constituent combination "passes" the statistical test so long as either the initial measurement or the resample measurement does not exceed the PL; this allows two attempts to obtain one "inbounds" measurement.

Davis and McNichols [1987] presented algorithms for computing exceedence probabilities for a generalization of "1 of *m*" plans when sampling from normal distributions; their algorithms along with those for other types of resampling plans are included in the program RESAMPLE which is being provided to PNNL personnel. The program GETK inverts these computations for the null distribution case, to provide the K value for given resampling plan, sample size, and degrees of freedom. Having obtained K, one then uses RESAMPLE to obtain power curves.

With large monitoring networks and small amounts of background data, it can be difficult to find a feasible resampling plan for which the power curves compare adequately with the reference curve using strictly intrawell background data. If it is reasonable to assume that the intrawell standard deviation remains approximately constant across wells, one can achieve better power by using a pooled standard deviation estimate. When using pooled standard deviation estimates, one should evaluate possible data transformations for both their normalizing and variance-stabilizing efficacy.

Another situation in which one may wish to use a pooled standard deviation estimate is the case where the data from some (but not all) wells are dominated by nondetects. In this situation one could select a transformation based on variance stabilization considering the remaining wells (not dominated by nondetects), compute the pooled standard deviation estimate based on those wells, and use it to compute "well-specific" PLs for all wells. Nondetects are typically simply assigned a nominal value, usually half their reporting limits (prior to transformation).

In using a pooled standard deviation estimate one assumes that data from all wells are statistically independent. If there is a temporal correlation among wells, Davis (1994) provides a simple adjustment to the pooled degrees of freedom to correct for that correlation.

The significance level used for each comparison should be 0.05 ÷ (# comparisons); for example, if there are four constituents and five wells, each individual test should use a significance level of 0.0025. This is a conservative worst-case computation, due to Bonferroni. It is very mildly conservative if different measurements are statistically independent. It is less conservative if measurements are positively correlated. Positive correlation in measurements can occur when (a) there are systematic time effects or (b) different constituents are correlated. If there are systematic time effects, such as variable analytical response from week to week, one can avoid sending resamples to the lab immediately as one precaution. One can also insist on obtaining data from at least one upgradient well to help demonstrate the existence of such effects. If two constituents are highly correlated, one may question the need for including both in the monitoring program.

CCC's have been described in detail in several documents, such as Chou [2004b]. The Shewhart Control Limit portion is essentially a prediction limit for individual measurements; the critical value for standardized measurements is denoted SCL. The CUSUM (CUmulative SUM) portion accumulates the exceedences of individual standardized measurements over a value (k) (usually 0.75 or 1.0). If the accumulated sum exceeds a critical value (h) an exceedence is declared. If the accumulated sum becomes negative it is reset to zero.

The statistical properties of CCCs depend obviously on the critical values and parameters SCL, k, and h and the background sample size n. They also depend strongly, but not so obviously, on the resampling plan allowed (if any), and any automatic background updating incorporated into the plan. (PL plans typically do not allow for automatic updating of background, in our experience). The latter two factors can have quite substantial impacts on the statistical properties, as will be seen in the case study to follow.

Since the CUSUM portion of the CCC depends on data from multiple sampling events, comparing the performance of CCCs with that of reference *t*-tests has rarely been attempted. This may be one reason why the impacts of resampling and background updating are not well understood. In the course of providing peer review input to the forthcoming *UG*, Davis [1999] developed a suite of simulation programs that allow for comparison of various CCC plans with various PL plans, using Average Run Length (ARL) and run length distributions as the figures of merit rather than false positive rate and power. The programs are CCNUL and PLNUL, for simulating the Facility-Wide Average Run Length (FWARL), and CCPOW and PLPOW, for simulating the sensitivity of the various plans to releases. These programs have been provided to PNNL personnel.

Typical guidance (USEPA [1989], USEPA [1992], and ASTM [1998], e.g.) sets SCL ≈ h at a rather high value (around 4.5), with the expectation that the Shewhart portion of the plan will find large releases immediately and the CUSUM portion will find moderate releases reasonably soon. The conventional wisdom is that the CUSUM portion will also be sensitive to small, gradual increases in measurement levels. Accordingly, CCPOW and PLPOW simulate both sudden and gradual increases. One of the findings of Davis [1999] is that allowing automatic updating of background can substantially compromise the sensitivity of these CCC plans to moderate and gradual releases, contrary to the expectations of the conventional wisdom.

G.6 Case Study Based on the 300 Area Process Trench RCRA Program

For a quantitative illustration of these concepts, we consider RCRA detection monitoring for the 300 Area Process Trenches, at which there are six wells and three constituents. With 18 comparisons per monitoring event, this is intermediate between the 12 or 15 comparisons anticipated in the LLWMA-3 RCRA detection monitoring program and the 32 comparisons evaluated by Jandhyala and Zhang [1999]. There are approximately 12 background observations available for each constituent for each well.

Ten intrawell statistical procedures are compared:

- ▸ reference 1% significance level *t*-tests with n = 8;
- ▸ reference 1% significance level *t*-tests with n = 10;
- A> "1 of 2" PL test with background sample size 12 and 11 degrees of freedom;
- B> "1 of 2" PL test with background sample size 12 and 33 degrees of freedom (see the previous discussion of pooled standard deviation estimates);
- C> CCC plan with SCL = h = 3.0 and k = 0.75, with no resamples or background updating, one of the plans favored by Jandhyala and Zhang [1999];
- D> CCC plan with SCL = h = 3.0 and k = 0.75, with one resample and no background updating;
- E> CCC plan with SCL = h = 3.0 and k = 0.75, with one resample and automatic background updating every four events;
- F> CCC plan with SCL = h = 4.5 and k = 0.75, with no resamples or background updating;
- G> CCC plan with SCL = h = 4.5 and k = 0.75, with one resample and no background updating; and
- H> CCC plan with SCL = h = 4.5 and k = 0.75, with one resample and automatic background updating every four events, as recommended by PNNL [Chou, 2004b].

Figure G.1. Power Comparisons of PL and Reference Tests

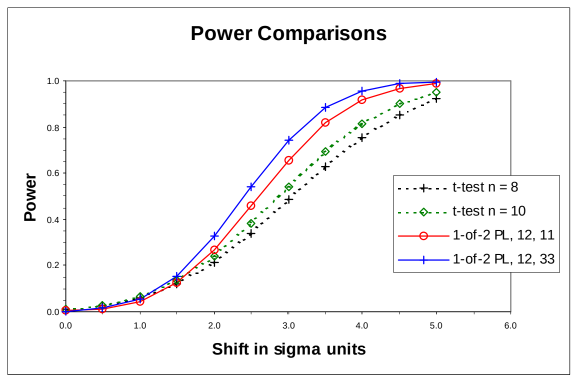
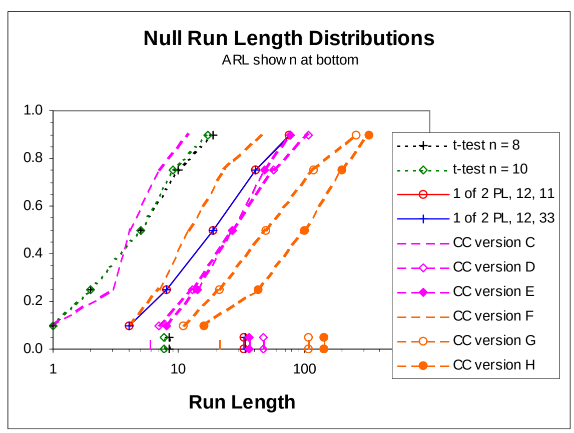


Figure G.2. Null Run Length Distributions and FWARLs



The PL procedures have K selected to achieve an FWFPR of 5 percent; the K values are 2.10 and 1.87 respectively. The power curves compare very favorably with either of the reference *t*-test power curves, as the seen in Figure G.1. For this regulated unit one would not need to bother with pooled standard deviations, unless the data for some constituent for some, but not all, wells were dominated by nondetects.

Figure G.2 shows the empirical FWARL and null run length distributions for the various procedures. Note that the distributions and FWARLs are very similar for the two reference *t*-tests. Those for the two PL tests are also very similar, with considerably greater run lengths. The run length distributions and FWARLs vary considerably among the CCC plans, however, with the run lengths of the worst case (recommended by Jandhyala and Zhang [1999]) being even lower than those of the *t*-tests, and the CCC procedures involving resamples with or without automatic updating having higher run lengths than the PL tests. All resamples are “delayed”, meaning that the measurement from the next sampling event is used as the resample. The highest run lengths belong to the CCC plans with resamples and larger values of SCL and h, with or without automatic updating. Table G.1 gives the estimated FWARLs.

Table G.1. FWARLs With No Release

Test	FWARL
t n = 8	8.43
t n = 10	7.57
A	32.96
B	33.92
C	5.95
D	47.62
E	36.42
F	21.23
G	109.23
H	142.39

Figure G.3 shows the ARLs for sudden shifts for a single well-constituent combination for each procedure. Observe that both PL procedures and four of the six CCC procedures compare favorably with the results for the reference *t*-tests. The reason these curves exceed the *t*-test curves at the far right is simply the use of the delayed resample. The two CCC procedures that do not compare favorably are those that allow for an automatic updating of the background dataset.

Figure G.4 shows the <u>ratios</u> of ARLs for procedures A, C, and H with the ARL of the 1 percent *t*-test with n = 8; procedure H can take up to six times as long, on average, to find a release as the reference test, whereas the “1 of 2” PL takes longer than the reference test only when the effect of the delayed resample is seen. CCC version “C”, by contrast, is rather more sensitive to sudden releases than the reference *t*-test; this is expected, since it is very quick to “find” a release even when none is present, as shown by its null FWARL.

Figure G.3. ARLs for Sudden Shifts

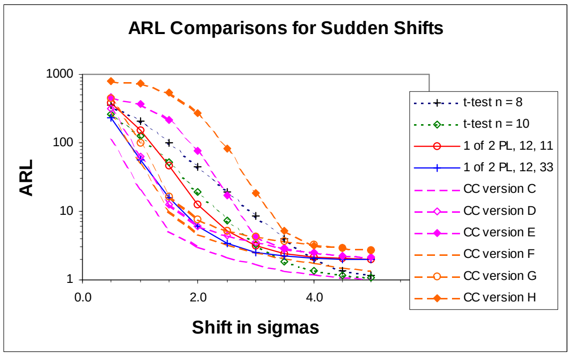
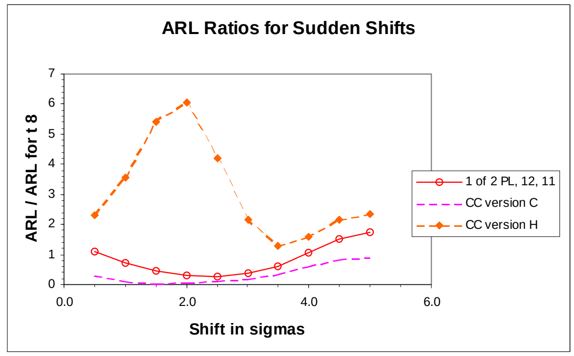


Figure G.4. ARL Ratios for Sudden Shifts



Figures G.5 and G.6 provide the same information, but for gradual (linear) increases. Again, the performances of the *t*-tests and PLs are quite similar, and those are similar to the performances of the CCC plans with resamples but no automatic updating. CCC plans with no resamples or automatic updating are more sensitive to releases (but also have lower ARL distributions when no releases are present). CCC plans with automatic updating miss small gradual releases more easily, contradicting one of the claimed advantages of the CUSUM portion of the CCC procedures.

Figure G.5. ARLs for Gradual Increases

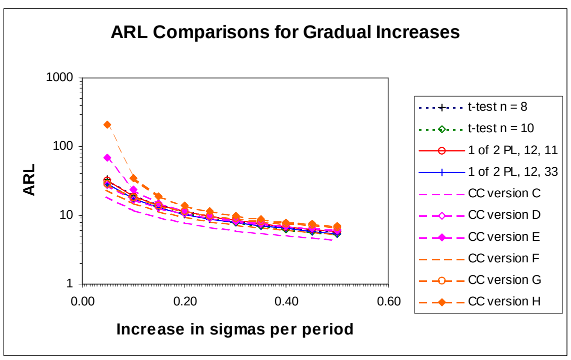
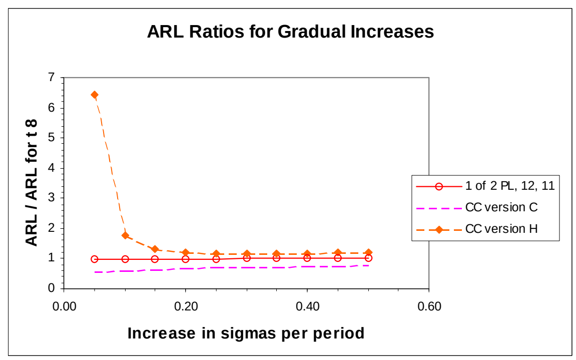


Figure G.6. ARL Ratios for Gradual Increases



The bottom line in these comparisons is that CCC plan "C", recommended by Jandahyala and Zhang [1999], is no better than the reference *t*-tests in terms of false positives, which means that it offers no Facility-Wide False Positive Rate control, whereas all of the other procedures are better. CCC plan "F", having no resamples or updating but higher control limits, provides nearly as good protection against false positives as do the PL tests. The other CCC plans, with resamples and with or without automatic updating, do better than the PL tests in this regard. With regard to sensitivity to sudden shifts, the ranking of the *t*-tests and PLs is the same using ARLs as in using conventional power computations, although the "1 of 2" PL with n = 12 and 11 degrees of freedom and the *t*-test with n = 10 are relatively closer together in the ARL analysis than in the conventional power analysis. The CCC plans with automatic background updating fail the comparison with the ARL curves for the reference *t*-tests. For gradual increases, the PL plans, *t*-tests, and CC plans with resamples but no updating are all about equivalent. The CCC plans with no resampling are more powerful, and the CCC plans with resampling and automatic updating less powerful. So the desirable plans will be CCC plans “D” and “G”, and then the PL plans.

A technical note: In these simulations the maximum run length was capped at 1000 periods (500 years of semi-annual monitoring). That maximum was achieved with some regularity for the smaller increases. Hence the ARLs shown in the plots, particularly the higher values, are generally underestimates of the actual ARL that would have been achieved in the absence of the maximum run length cap.

G.7 Updating Background Datasets

The updating of background discussed previously is an automatic updating every *x* events. Another form of updating is an updating of background, or re-establishing background, following a (confirmed) statistical exceedence that is determined not to be caused by the operation of the regulated unit. In such a situation, it may make little sense to return to detection monitoring using the previous background. Rather, one should update the background data. This can be the situation, for example, if there is a systematic change in measurement levels due to a change in instrument or laboratory method (or even laboratory) or a change in the underlying groundwater regime.

It should be emphasized that intrawell statistical tests can be quite sensitive to even small perturbations of the groundwater measurement process. The power and ARLs presented here are in terms of the intrawell standard deviation, which can be relatively much smaller than the overall variability of a constituent’s measurements at a regulated unit. Accordingly, the need to deal with the consequences of true positives not related to releases should be anticipated in establishing the monitoring program. The requirement in §264.98(g)(2) for an immediate sampling of all wells with analysis for all Part 264 Appendix IX constituents is not appropriate. The 1991 Part 258 Solid Waste rules are a considerable improvement: following the discovery of a statistically significant exceedence, one first has 90 days to evaluate the situation and determine if the change is due to benign causes. If not, and the Part 258 Appendix II sampling is required, it may be limited to a subset of wells and a subset of constituents rather than all of each as is required under the 1988 Part 264 rules. Again, although the Part 264 Hazardous Waste rules are nominally quite stringent, in practice the Regional Administrator has considerable latitude in their enforcement.

G.8 Pre-Existing Conditions and Statistical Assumptions

All of these statistical tests implicitly assume that the measurement data are independent and identically distributed. Chou [2004b] accurately describes the implications of the independence assumption with respect to sampling frequency. In addition to those arguments, one should add the possibility that the laboratory (or measurement system as a whole) may have temporal patterns that can affect data values in a manner resembling an autocorrelation. That is a further argument for avoiding sampling too frequently, regardless of what might be suggested by Darcy's equation computations. These arguments are, of course, at odds with the nominal regulatory requirement for four independent observations per (semi-annual) monitoring event.

The possibility of pre-existing trends is another matter, and a difficult one to deal with. With respect to the B Pond system, the claim is made that apparent upward trends in the measurements of all three indicator parameters (specific conductance, gross alpha, and gross beta) are due to the cessation in the mid-1990s of the discharge of relatively dilute cooling water to the pond, with a consequent return to ambient groundwater conditions. If this is the current conceptual model for groundwater in the vicinity of the B Pond, perhaps one should quantify this conceptual model in a fashion that explicitly incorporates an asymptotic value into the projected values, and then construct prediction limits about that projection, rather than simply extrapolating a continued linear increase. At a minimum, any extrapolated linear trend should be shut off at some point in the future when the projected mean value equals the anticipated ambient mean.

In examining plots of data for the B Pond wells and constituents extracted from the Hanford Environmental Information System (HEIS), it appears that the depressed values of conductance in several wells, including the deep upgradient well 699-44-39B but not the deep downgradient well 699-42-42B, are of fairly recent origin (mid-1993) and ceased in mid-1998. One well is dry now; two seem to show mild increases in conductance in recent years, one is new, and the deep upgradient well seems to show a decrease. There seems to be a general upward <u>shift</u> in gross alpha and gross beta measurements in mid-1998, but no particular <u>trends</u> since then. The HEIS data go back in time only to 1988; perhaps a longer perspective would allow one to see the target ambient levels better. On the other hand, it is possible that a systematic change in sampling, sample preparation, or measurement methodology could have caused the mid-July changes in all three constituents.

Nonetheless, the concept that failing to take real trends into account can mask real departures from those trends [Chou, 2004b] is valid. Perhaps an alternate approach would be to fit the projection described above to recent data (beginning, say, in mid-1998) and incorporate an anticipated ambient level as an asymptote. A liberal criterion could be used in determining whether or not to fit a trend. Residuals from the fitted curve could be used to estimate the standard deviation around the fitted line, and the standard error of the fit could be used to obtain an approximate "pseudo-sample size", which could then be used with GETK to obtain appropriate "1 of 2" PL coefficients. In this application one may wish to use two-sided PLs, to determine whether there are positive <u>or</u> negative departures from the anticipated trend.

Another form of deviation from the "identically distributed" assumption is found in wells near the river in the 300 Area Process Trench monitoring network. In these wells high levels of the river may dilute the groundwater, leading to lower levels of dissolved contaminants in the groundwater. In order to adjust for this phenomenon, PNNL proposes using specific conductance as a surrogate for the mixing ratio, assume that the constituents in question are essentially absent from the river water, and compute what the concentrations in undiluted groundwater would have been. Such an approach seems quite reasonable. Our only question is whether specific conductance behaves enough like a concentration under mixing of waters of different characters for this approach to be reliable. A simple off-line experiment should be enough to establish that fact.

G.9 References

ASTM [1998]. *Standard Guide for Developing Appropriate Statistical Approaches for Ground-Water Detection Monitoring Programs*, D-6312-98, ASTM, West Conshohocken, PA.

Cameron, K.C. [1996]. "RCRA Leapfrog: How Statistics Shape and In Turn Are Shaped by Regulatory Mandates", *Remediation* v. 7, 15-25.

Chou, C.J. [2004a]. "Groundwater Monitoring: Statistical Methods for Testing Special Background Conditions", chapter 8 of *Environmental Monitoring*, G.B. Wiersma, ed., CRC Press, Boca Raton, FL.

Chou, C.J. [2004b]. *Evaluation of an Alternative Statistical Method for Analysis of RCRA Groundwater Monitoring Data at the Hanford Site*, PNNL-14521, Richland, WA.

Chou,C.J, R.F. O'Brien, and D.B. Barnett [2001]. "Application of Intrawell Testing of RCRA Groundwater Monitoring Data When No Upgradient Well Exists", *Environmental Monitoring and Assessment* v. 71, 91-106.

Davis, C.B. [1998]. *Groundwater Statistics & Regulations: Principles, Progress and Problems*, 2nd Edition, EnviroStat, Henderson,NV.

Davis, C.B. [1999]. *Comparisons of Control Chart and Prediction Limit Procedures Recommended for Groundwater Detection Monitoring*, EnviroStat TR 99-1, EnviroStat, Henderson, NV.

Davis, C.B. and R.J. McNichols [1987]. "One-Sided Intervals for at Least *p* of *m* Observations From a Normal Population on Each of *r* Future Occasions", *Technometrics* v. 29, 359-370.

Davis, C.B. [1997]. "Challenges in Regulatory Environmetrics", *Chemometrics and Intelligent Laboratory Systems* v. 37, 43-53.

Davis, C.B. and R.J. McNichols [1987]. "One-Sided Intervals for at Least *p* of *m* Observations From a Normal Population on Each of *r* Future Occasions", *Technometrics* v. 29, 359-370.

Davis, C.B. and R.J. McNichols [1988]. "Statistical Issues and Problems in Ground Water Detection Monitoring at Hazardous Waste Facilities", *Ground Water Monitoring Review*, 135-160.

Davis, C.B. and R.J. McNichols [1994]. "Ground-Water Monitoring Statistics Update: Part I: Progress Since 1988", *Ground Water Monitoring and Remediation*, 148-158.

Jandhyala, V.K. and H. Zhang [1999]. *Report on Statistical Analysis for B-Pond Groundwater Monitoring Program*, prepared for State of Washington Department of Ecology.

McNichols, R.J. and C.B. Davis [1994]. "Ground-Water Monitoring Statistics Update: Part II: Nonparametric Prediction Limits", *Ground Water Monitoring and Remediation*, 159-175.

USEPA [1986]. *RCRA Ground-Water Monitoring Technical Enforcement Guidance Document*, Office of Solid Waste and Emergency Response, USEPA, Washington, DC.

USEPA [1989]. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities: Interim Final Guidance*, EPA/530-F-93-003, USEPA, Washington, DC.

USEPA [1992]. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities: Addendum to Interim Final Guidance*, USEPA, Washington, DC.